Invariant Disns for Markov Chains contd

Thomas Mountford

EPFL

March 29, 2021

Irreducible periodic case

This argument used explicitly aperiodicity. Indeed its conclusion is not true for periodic chains. Nonetheless it has consequences for periodic irreducible chain. We know that an irreducible chain has a single period d. If the chain is periodic, then $d > 1$. We suppose that to be the case. In that case for $j \in I$ we can define D_0 to be the set of sites so that for some $n, p_{jk}^{nd} > 0$. So equivalenetly, $k \in D_0$ if and only if $p_{jk}^{nd} > 0$ for all n large. We can then define for $0 \leq r < d$, $D_r = \{k : p_{jk}^{nd+r} > 0\}$ for some $n = \{k : p_{jk}^{nd+r} > 0\}$ for all n large.

Irreducible periodic case

It is easily seen that if $i \in D_s, k \in D_t$, then $p_{ik}^n > 0$ only if $s + n$ is congruent to t mod(d).

Thus starting at,say, j the chain moves from D_0 to D_1 to D_2 and so on mod(d). The transition matric P^d is irreducible on D_0, on D_1 . . but not on I.

Theorem

If π is invariant distribution for periodic irreducible Markov chain then for D_r as above, for fixed r

$$\lim p_{jk}^{nd+r} = d\pi(k)$$

for every $k \in D_r$

Remark: Obviously if k is not in D_r, then for every n, $p_{jk}^{nd+r} = 0$.

Reversible Proceses

An important class of Markov chains are socalled *Reversible* processes: A Markov chain is reversible with respect to a measure π if $\forall x, y \in I, \pi(x)p(x,y) = \pi(y)p(y,x)$. Most chains are not reversible with respect to a distribution or measure. However

Theorem

If distribution π is such that P is reversible with respect to it, then π is invariant for P.

Proof: Fix x.

$$\pi(x) = \sum_y \pi(x)p(x,y) = \sum_y \pi(y)p(y,x)$$

Examples

We first consider irreducible birth and death chains. IF the chain has an invariant law π, then this must be a reversible measure for the process: For $(X_n)_{n\geq 0}$ starting in equilibrium π, we have that for every $A \subset \mathbb{N}, P(X_n \in A$ does not change. Thus

$$P(X_0 \in A, X_1 \notin A) = P(X_0 \notin A, X_1 \in A)$$

Apply this to $A = [0, n]$ to get

$$\pi(n) p_{nn+1} = \pi(n+1) p_{n+1n}$$

Given our process only makes nearest neighbour moves, this is enough.

Examples

Thus to see if a B and D chain is positive recurrent we have the following recipe. We use p_i for p_{ii+1} and q_i for p_{ii-1}.

- Put $\nu(0) = 1$
- Recursively choose $\nu(n) = \nu(n-1)p_{n-1}/q_n = \prod_{k=1}^{n} p_{k-1}/q_k$
- If $\sum_k \nu(n) = M < \infty$, then the chian is positive recurrent with invariant $\pi(n) = \nu(n)/M$; otherwise it is not positive recurrent.

Examples contd

Consider a "random walk" on a finite graph $G = (V, E), : P_{x,y} = \frac{1}{d_x}$ (where d_x = the degree of x) if x and y are neighbours; otherwise $P_{x,y} = 0$. Then the measure $\nu(x) = d_x$ satisfies the detained balence and so (for a connected graph or equivalently an irreducible chain), the invariant distribution is ν/M where $M = \sum_x d_x$

Reversibility

If a Markov chain is reversible with respect to probability distribution π, then if $(X_n)_{n\geq 0}$ starts in equilibrium (i.e. X_n has law $\pi \forall n$), then if we fix $N < \infty$, then $(Y_n)_{0\leq n\leq N}$ given by $Y_n = X_{N-n}$ satisfies for any $i_0, i_1, \cdots i_N$

$$P(Y_0 = i_0, Y_1 = i_1 \cdots Y_N = i_N) = P(X_0 = i_N, X_1 = i_{N-1} \cdots X_N = i_0)$$

$$\pi(i_N) p_{i_N, i_{N-1}} \cdots p_{i_1 i_0}$$

But using now reversibility we see

$$p_{i_{N-1}, i_N,} \pi(i_{N-1}) p_{i_{N-1}, i_{N-2}} \cdots p_{i_1 i_0} = p_{i_{N-1}, i_N,} p_{i_{N-2}, i_{N-1}} \pi(i_{N-2}) \cdots p_{i_1 i_0} \cdots$$

$$= \pi(i_0) p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{N-1}, i_N},$$

Reversibility

Theorem

If $(X_n)_{n\geq 0}$ *is a Markov chain in starting with an invariant probability with respect to which it is reversible, then for each* N

$$(X_n)_{0\leq n\leq N} \stackrel{D}{=} (X_{N-n})_{0\leq n\leq N}$$

An invariant measure.

Up until now we have really been interested in invariant distributions rather than invariant measures. Recall a measure on I, ν is simply a collection of positive values $\nu(i) i \in I$. If $\sum_i \nu(i) < \infty$, then we may divide ν by its total mass to obtain a probability distribution. In later studies of continuous time process we will need some results on invariant measures, even if they are not finite.

An invariant measure.

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.*

Proof (i) is immediate. Secondly, since the chain is irreducible for any $j, \mathbb{P}_k(T_j < T_k) > 0$ and $\mathbb{P}_j(T_k < T_j) > 0$. The first inequality implies that $\gamma(j) > 0$ while the second one implies that under $\mathbb{P}_k$, the number of visits to j before T_k is stochastically dominated by a Geometric random variable of parameter $\mathbb{P}_j(T_k < T_j) > 0$ and so $\gamma(j) < \infty$

An invariant measure.

For (iii), just as in the proof of the existence of the invariant probability, the key element is a temporal shift. Since the chain is recurrent we have T_k is a.s. finite. Thus

$$\gamma(j) = \mathbb{E}_k\left(\sum_{0<r\leq T_k} I_{X_r=j}\right)$$

Now as before, we can write

$$\gamma(j) = \mathbb{E}_k\left(\sum_{0\leq r<T_k} I_{X_r=i}\sum_i p_{ij}\right)$$

$$= \sum_i \nu(i)p_{ij}$$

Theorem

If λ is another invariant measure satisfying (i)-(iii) above, then $\gamma = \lambda$.

Proof of Theorem.

We first prove $\gamma \leq \lambda$. Fix $j \neq k$, then

$$\lambda(j) = \sum_{i_1} \lambda(i_1) p_{i_1 j} = \lambda(k) p_{kj} + \sum_{i_1 \neq k} \lambda(i_1) p_{i_1 j}$$

$$= p_{kj} + \sum_{i_1 \neq k} \lambda(i_1) p_{i_1 j}$$

$$\sum_{i_2} \sum_{i_1 \neq k} \lambda(i_2) p_{i_2 i_1} p_{i_1 j} + p_{kj} = \sum_{i_2 \neq k} \sum_{i_1 \neq k} \lambda(i_2) p_{i_2 i_1} p_{i_1 j} + \sum_{i_1 \neq k} p_{k i_1} p_{i_1 j} + p_{kj}$$

Continuing

$$= \sum_{i_3} \sum_{i_2 \neq k} \sum_{i_1 \neq k} \lambda(i_3) p_{i_3 i_2} p_{i_2 i_1} p_{i_1 j} + \sum_{i_1 \neq k} p_{k i_1} p_{i_1 j} + p_{kj}$$

$$= \sum_{i_3 \neq k} \sum_{i_2 \neq k} \sum_{i_1 \neq k} \lambda(i_3) p_{i_3 i_2} p_{i_2 i_1} p_{i_1 j} + \sum_{i_2 \neq k} \sum_{i_1 \neq k} p_{k i_2} p_{i_2 i_1} p_{i_1 j} + \sum_{i_1 \neq k} p_{k i_1} p_{i_1 j} + p_{kj}$$

Continuing to arbitrary many iterations and noting that all terms are positive we obtain the inequality $\lambda(j) \geq \gamma(j)$.

To complete the arguement we consider invariant measure $\lambda - \gamma$. This is positive everywhere and zero at k. But by iteration of the invariance, we must have the difference is zero for every j such that $p_{jk} > 0$, iterating again we see $\lambda(j) - \gamma(j)$ must be zero for every j with $p_{jk}^2 > 0$ and so on. By ireducibility we have every j must give value zero.

Ergodic Theorem

Theorem

For a positive recurrent irreducible Markov chain $(X_n)_{n\geq 0}$ and a bounded f on I

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k<n} f(X_k) \to \sum_{i\in I} \pi_i f(i)$$

a.s. over all initial distributions.

The proof is easy. we suppose that f is bounded by 1 i abolute value. We already know that for any k, the proportion of time spent in state k converges a.s, to π_k, so we get this law of large numbers for all k in any fixed finite subset of I. Let us fix $A \subset I$ finite. Then $\frac{1}{n}\sum_{k<n} f(X_k) =$

$$\sum_i N_n(i)f(i) = \sum_{i\in A} N_n(i)f(i) + \sum_{i\in A^c} N_n(i)f(i)$$

where $N_n(i)$ is the proportion of time spent in state i by time n

proof contd

But $\sum_{i \in A} N_n(i) f(i)$ converges to $\sum_{i \in A} \pi_n f(i)$. Recalling that f is bounded by one we get

$$\limsup_n \left| \frac{1}{n} \sum_{k<n} f(X_k) - \sum_{i \in I} \pi_i f(i) \right| \leq 2 \sum_{i \in A^c} \pi_i$$

Generalization

We simply observe that for any integer $r \geq 1$ (under the previous conditions)$(Y_n)_{n\geq 0}$ is an irreducible Markov chain on the set of "possible" $r$tuples $(i_1, i_2, \cdots i_r) \in I^r$ for $Y_n = (X_n, X_{n+1}, \cdots X_{n+r-1})$. It has unique invariant distribution

$$\pi^Y((i_1, i_2, \cdots i_r)) = \pi(i_1) p_{i_1,i_2} \cdots p_{i_{r-1},i_r}.$$

Theorem

For a positive recurrent irreducible Markov chain $(X_n)_{n\geq 0}$, positive integer r and and a bounded f on I^r

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k<n} f(X_k, X_{k+1}, \cdots, X_{k+r-1}) \to \sum_{\bar{i}\in I^r} f(\bar{i}) \pi(i_1) p_{i_1,i_2} \cdots p_{i_{r-1},i_r}$$

a.s. over all initial distributions.