Probability and Stochastic Processes II - Lecture 4b

Michael Evans
University of Toronto
https://utstat.utoronto.ca/mikevans/stac62/staC632024.html

2024

- when a stationary distribution π exists, do we have $\lim_{n\to\infty} P(X_n = j) = \pi_j$?

Example IV.5

- suppose $\mathcal{S} = \{1, 2\}$ and $P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, then $\pi = (1/2, 1/2)$ is stationary but

$$P(X_n = 1) = \begin{cases} 1 & \text{if } X_0 = 1 \\ 0 & \text{if } X_0 = 2 \end{cases}$$

so $\lim_{n\to\infty} P(X_n = j) \neq \pi_j$

- note that the chain is reducible ■

- another concern is periodicity

IV.2 Periodicity

Definition IV.3 The *period* of state i is gcd $T(i)$ where

$$T(i) = \{n \geq 1 : p_{ii}^{(n)} > 0\}.$$

A chain where all states have period 1 is called *aperiodic*.

- so determine the set $\{n \geq 1 : p_{ii}^{(n)} > 0\}$ and then find the largest $m \in \mathbb{N}$ s.t. m divides evenly into every n in this set

Example IV.6 - suppose

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

- then $p_{11}^{(2)} > 0, p_{11}^{(3)} > 0$ so period of state 1 must be 1 because that is the largest integer dividing 1 and 2

- note that we didn't need to determine all the elements of $T(1) = \{n \geq 1 : p_{11}^{(n)} > 0\}$ to determine the period

- by the same argument the states 2 and 3 also have period 1 so the chain is aperiodic ■

- in general if $p_{ii}^{(n)} > 0, p_{ii}^{(n+1)} > 0$ for some n, then the period of i is 1

Proposition IV.5 If $i \leftrightarrow j$, then i and j have the same period.

Proof: There exist r, s s.t. $p_{ij}^{(r)} > 0$ and $p_{ji}^{(s)} > 0$. Now $p_{ii}^{(r+n+s)} \geq p_{ij}^{(r)} p_{jj}^{(n)} p_{ji}^{(s)}$, so if $p_{jj}^{(n)} > 0$, then $p_{ii}^{(r+n+s)} > 0$ which implies that the period(i) must divide $r + n + s$. Since $p_{ii}^{(r+s)} \geq p_{ij}^{(r)} p_{ji}^{(s)} > 0$, then period$(i)$ must divide $r + s$ and so it must divide n as well. Therefore period(i) divides every element of $T(j)$ which implies period$(i) \leq$ period(j). Reversing the argument establishes that period$(i) \geq$ period(j) and so they are equal. ■

Corollary IV.6 If the MC is irreducible, then all states have the same period and if $p_{ii} > 0$ for some i, then the chain is aperiodic.

IV.3 The Convergence Theorem

Proposition IV.6 If an irreducible chain has a stationary distribution then it is recurrent.

Proof: Recall Corollary IV.4 that a transient chain cannot have a stationary distribution. ■

Proposition IV.7 If state i has $f_{ii} > 0$ and is aperiodic, then there exists $n_0(i)$ s.t. $p_{ii}^{(n)} > 0$ for all $n \geq n_0(i)$.

Proof: Suppose m, n satisfy $p_{ii}^{(m)} > 0, p_{ii}^{(n)} > 0$. Therefore $p_{ii}^{(m+n)} \geq p_{ii}^{(m)} p_{ii}^{(n)} > 0$ (Chapman-Kolmogorov) which implies $T(i) = \{n \geq 1 : p_{ii}^{(n)} > 0\}$ is additive, namely if $m, n \in T(i)$, then $m + n \in T(i)$.
Note that we need only show that there exists k s.t. $k, k+1 \in T(i)$ since

$$p_{ii}^{(2k)} \geq p_{ii}^{(k)} p_{ii}^{(k)} > 0, p_{ii}^{(2k+1)} \geq p_{ii}^{(k)} p_{ii}^{(k+1)} > 0, p_{ii}^{(2k+2)} \geq p_{ii}^{(k+1)} p_{ii}^{(k+1)} > 0$$

implies $2k, 2k+1, 2k+2 \in T(i)$ and similarly $jk, jk+1, \ldots, jk+j \in T(i)$ for any j. As soon as $j \geq k-1$ the blocks overlap and this implies the result.

A result from number theory (Niven and Zuckerman (1972) An Introduction to the Theory of Numbers, Thm 1.5) : if the gcd of a set T is g, then there exist distinct $n_1, \ldots, n_k \in T$ and integers x_i s.t. $g = \sum_{i=1}^{k} x_i n_i$. In this case $g = 1$. Now write $x_i = x_i^+ - x_i^-$ where

$$x_i^+ = \begin{cases} x_i & \text{if } x_i \geq 0 \\ 0 & \text{otherwise} \end{cases} \qquad x_i^- = \begin{cases} -x_i & \text{if } x_i < 0 \\ 0 & \text{otherwise.} \end{cases}$$

Therefore

$$1 = \sum_{i=1}^{k} x_i n_i = \sum_{x_i^+ > 0} x_i^+ n_i - \sum_{x_i^- > 0} x_i^- n_i$$

where $\sum_{x_i^+ > 0} x_i^+ n_i > 0, \sum_{x_i^- > 0} x_i^- n_i > 0$. Therefore

$$\sum_{x_i^+ > 0} x_i^+ n_i = \left(\sum_{x_i^- > 0} x_i^- n_i \right) + 1$$

and since $n_1, \ldots, n_k \in T(i)$, and $T(i)$ is additive then $\sum_{x_i^+ > 0} x_i^+ n_i \in T(i)$, $\sum_{x_i^- > 0} x_i^- n_i \in T(i)$ and so it has been shown that there are consecutive elements of $T(i)$. ■

Corollary IV.8 If a chain is irreducible and aperiodic then for any $i, j \in \mathcal{S}$ there is $n_0(i,j) \in \mathbb{N}$ s.t. $p_{ij}^{(n)} > 0$ for all $n \geq n_0(i,j)$.

Proof: Let $n_0(i)$ be as in the proposition, m be s.t. $p_{ij}^{(m)} > 0$ and put $n_0(i,j) = n_0(i) + m$. When $n \geq n_0(i,j)$ then $n - m \geq n_0(i)$ so $p_{ii}^{(n-m)} > 0$ which implies $p_{ij}^{(n)} \geq p_{ii}^{(n-m)} p_{ij}^{(m)} > 0$. ■

- the following result shows that, under conditions, the initial distribution of the chain has no long term effect, also the proof introduces the important technique of *coupling*

Proposition IV.8 If a MC $\{X_n : n \in \mathbb{N}_0\}$ is irreducible and aperiodic with stationary distribution π then for all $i, j, k \in \mathcal{S}$

$$\lim_{n\to\infty} |p_{ik}^{(n)} - p_{jk}^{(n)}| = 0.$$

Proof: Define a new chain $\{(X_n, Y_n) : n \in \mathbb{N}_0\}$ where $\{Y_n : n \in \mathbb{N}_0\}$ is an independent copy of $\{X_n : n \in \mathbb{N}_0\}$. So the state space of the new chain is $\mathcal{S} \times \mathcal{S}$ with transition probabilities $p_{(i,j)(k,l)} = p_{ik}p_{jl}$. This new chain has stationary distribution $\pi^*_{(i,j)} = \pi_i \pi_j$ since

$$\sum_{(i,j)\in\mathcal{S}\times\mathcal{S}} \pi^*_{(i,j)} p_{(i,j)(k,l)} = \sum_{i\in\mathcal{S}} \pi_i p_{ik} \sum_{k\in\mathcal{S}} \pi_j p_{jl} = \pi_k \pi_l = \pi^*_{(k,l)}.$$

The new chain is irreducible, since $p^{(n)}_{(i,j)(k,l)} = p^{(n)}_{ik} p^{(n)}_{jl} > 0$ whenever $n \geq \max\{n_0(i,j), n_0(k,l)\}$. Also the new chain is aperiodic since $p^{(n)}_{(i,j)(i,j)} = p^{(n)}_{ii} p^{(n)}_{jj} > 0$ iff $p^{(n)}_{ii} > 0$ and $p^{(n)}_{jj} > 0$ which is true whenever $n \geq \max\{n_0(i), n_0(j)\}$. By Prop. IV.6 the new chain is recurrent.

Now choose $i_0 \in \mathcal{S}$ and let $\tau = \inf\{n : (X_n, Y_n) = (i_0, i_0)\}$. By the recurrence of the new chain we have $f_{(i,j)(i_0,i_0)} = 1$ for all (i,j) which implies $P_{(i,j)}(\tau < \infty) = 1$ for all (i,j). We have

$$\begin{aligned}
p_{ik}^{(n)} &= P_i(X_n = k) = \frac{P(X_0 = i, X_n = k)P(Y_0 = j)}{P(X_0 = i)P(Y_0 = j)} \\
&= \frac{P(X_0 = i, X_n = k, Y_0 = j)}{P(X_0 = i, Y_0 = j)} \text{ by independence} \\
&= P_{(i,j)}(X_n = k) \\
&= \sum_{m=1}^{\infty} P_{(i,j)}(X_n = k, \tau = m) \\
&= \left\{ \sum_{m=1}^{n} P_{(i,j)}(X_n = k, \tau = m) + P_{(i,j)}(X_n = k, \tau > n) \right\},
\end{aligned}$$

and by the same argument

$$p_{jk}^{(n)} = \left\{ \sum_{m=1}^{n} P_{(i,j)}(Y_n = k, \tau = m) + P_{(i,j)}(Y_n = k, \tau > n) \right\}.$$

For $n \geq m$

$$\begin{aligned}
& P_{(i,j)}(X_n = k, \tau = m) = P_{(i,j)}(\tau = m) P_{(i,j)}(X_n = k \,|\, \tau = m) \\
= \;& P_{(i,j)}(\tau = m) P_{(i,j)}(X_n = k \,|\, X_m = Y_m = i_0 \text{ and coupling} \\
& \qquad \text{doesn't happen before time m}) \\
= \;& P_{(i,j)}(\tau = m) P_{(i,j)}(X_n = k \,|\, X_m = Y_m = i_0) \text{ by MP} \\
= \;& P_{(i,j)}(\tau = m) P_i(X_n = k \,|\, X_m = i_0) \text{ by independence of the chains} \\
= \;& P_{(i,j)}(\tau = m) p_{i_0 k}^{(n-m)} \text{ by TH and} \\
& P_{(i,j)}(Y_n = k, \tau = m) = P_{(i,j)}(\tau = m) p_{i_0 k}^{(n-m)} \text{ by the same argument.}
\end{aligned}$$

This has proved that

$$P_{(i,j)}(X_n = k, \tau = m) = P_{(i,j)}(Y_n = k, \tau = m).$$

Now we have

$$
\begin{aligned}
& |p_{ik}^{(n)} - p_{jk}^{(n)}| \\
= \ & |\sum_{m=1}^{n} P_{(i,j)}(X_n = k, \tau = m) - \sum_{m=1}^{n} P_{(i,j)}(Y_n = k, \tau = m) + \\
& \qquad P_{(i,j)}(X_n = k, \tau > n) - P_{(i,j)}(Y_n = k, \tau > n)| \\
= \ & |P_{(i,j)}(X_n = k, \tau > n) - P_{(i,j)}(Y_n = k, \tau > n)| \\
\leq \ & (P_{(i,j)}(X_n = k, \tau > n) + P_{(i,j)}(Y_n = k, \tau > n)) \\
\leq \ & (P_{(i,j)}(\tau > n) + P_{(i,j)}(\tau > n)) = 2P_{(i,j)}(\tau > n).
\end{aligned}
$$

Therefore

$$
\lim_{n\to\infty} |p_{ik}^{(n)} - p_{jk}^{(n)}| \leq 2 \lim_{n\to\infty} P_{(i,j)}(\tau > n) = 2P_{(i,j)}(\tau = \infty) = 0.
$$

■

Proposition IV.9 (*Markov Chain Convergence Theorem - MCCT*) If a MC is irreducible, aperiodic and has stationary distribution π, then

(i) $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$ for every $i, j \in \mathcal{S}$,

(ii) $\lim_{n\to\infty} P(X_n = j) = \pi_j$ for any initial distribution ν.

Proof: We have

$$|p_{ij}^{(n)} - \pi_j| = \left| \sum_{k\in\mathcal{S}} \pi_k \left(p_{ij}^{(n)} - p_{kj}^{(n)} \right) \right| \le \sum_{k\in\mathcal{S}} \pi_k \left| p_{ij}^{(n)} - p_{kj}^{(n)} \right|$$

so

$$\begin{aligned} & \lim_{n\to\infty} |p_{ij}^{(n)} - \pi_j| \le \lim_{n\to\infty} \sum_{k\in\mathcal{S}} \pi_k \left| p_{ij}^{(n)} - p_{kj}^{(n)} \right| \\ = & \sum_{k\in\mathcal{S}} \pi_k \lim_{n\to\infty} \left| p_{ij}^{(n)} - p_{kj}^{(n)} \right| \text{ by DCT since } \left| p_{ij}^{(n)} - p_{kj}^{(n)} \right| \le p_{ij}^{(n)} + p_{kj}^{(n)} \le 2 \\ = & \ 0 \text{ by Prop. IV.8} \end{aligned}$$

and this proves (i).

Now

$$\lim_{n\to\infty} P(X_n = j) = \lim_{n\to\infty} \sum_{i\in\mathcal{S}} \nu_i P_i(X_n = j) = \lim_{n\to\infty} \sum_{i\in\mathcal{S}} \nu_i p_{ij}^{(n)}$$

$$\begin{aligned} &= \sum_{i\in\mathcal{S}} \nu_i \lim_{n\to\infty} p_{ij}^{(n)} \text{ by DCT since } p_{ij}^{(n)} \le 1 \\ &= \sum_{i\in\mathcal{S}} \nu_i \pi_j \text{ by part (i)} \\ &= \pi_j \end{aligned}$$

and this proves (ii). ■

Example IV.7 *(Example III.2 continued)*

- recall

$$\begin{aligned} \mathcal{S} &= \{1,2,3,4\}, \upsilon = (1/4, 1/2, 1/8, 1/8), \\ P &= \begin{pmatrix} 0 & 1/3 & 1/2 & 1/6 \\ 1/3 & 0 & 1/2 & 1/6 \\ 1/6 & 1/6 & 0 & 2/3 \\ 1/3 & 1/3 & 1/3 & 0 \end{pmatrix} \end{aligned}$$

- chain is irreducible, recurrent and aperiodic $\{n : p_{11}^{(n)} > 0\} = \{2, 3 \ldots\}$
- as we saw, and now by the Markov Chain Convergence Thm,

$$P^n = (p_{ij}^{(n)}) \to \begin{pmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 \\ \pi_1 & \pi_2 & \pi_3 & \pi_4 \\ \pi_1 & \pi_2 & \pi_3 & \pi_4 \\ \pi_1 & \pi_2 & \pi_3 & \pi_4 \end{pmatrix}$$

and $P(X_n = j) \to \pi_j$ where

$$(\pi_1, \pi_2, \pi_3, \pi_4) = (0.2121212, 0.2121212, 0.3030303, 0.2727273)$$

■

Corollary IV.10 If a MC is irreducible and aperiodic then it has at most one stationary distribution.

Proof: If the chain has a stationary distribution then by the MCCT $\lim_{n\to\infty} P(X_n = j) = \pi_j$.

Example IV.8 *(Simple Random Walk)*

- the chain is irreducible but has period 2 so MCCT doesn't apply and no stationary distribution exists anyway

- recall $P(X_n = j) = P(\sum_{i=1}^n Z_i = j)$ where the Z_i are $i.i.d.$ with $P(Z_i = 1) = p, P(Z_i = -1) = 1 - p$ so

$$\begin{aligned} E(Z_i) &= p - (1-p) = 2p - 1, \\ Var(Z_i) &= 1 - (2p-1)^2 = 4p(1-p) \end{aligned}$$

- therefore by SLLN

$$\frac{1}{n}\sum_{i=1}^n Z_i \overset{wp1}{\rightarrow} 2p - 1 \qquad (1)$$

- when $p \neq 1/2$ the chain is transient

- when $p > 1/2$ there is $\epsilon > 0$ s.t. $2p - 1 - \epsilon > 0$ so

$$1 = P\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} Z_i > 2p - 1 - \epsilon\right)$$

so for $c > 0$ and ω s.t. (1) holds there is an $n(\omega, c)$ s.t. $\sum_{i=1}^{n} Z_i(\omega) > c$ for all $n \geq n(\omega, c)$ which implies $\lim_{n\to\infty} \sum_{i=1}^{n} Z_i(\omega) = \infty$ which implies

$$1 = P\left(\lim_{n\to\infty} \sum_{i=1}^{n} Z_i = \infty\right) = P\left(\lim_{n\to\infty} X_n = \infty\right)$$

- similarly when $p < 1/2$

$$1 = P\left(\lim_{n\to\infty} \sum_{i=1}^{n} Z_i = -\infty\right) = P\left(\lim_{n\to\infty} X_n = -\infty\right)$$

- what about when $p = 1/2$? since it is a recurrent chain in this case we have $f_{ij} = 1$ for all i, j which implies

$$f_{i_1 i_2} f_{i_2 i_3} \cdots f_{i_{m-1} i_m} = 1$$

for *any* sequence of distinct states $i_1, i_2, \ldots, i_m$ (and there are infinitely many such sequences) so the chain never settles down

- but note in all cases the CLT gives

$$\frac{X_n/n - (2p-1)}{\sqrt{4p(1-p)/n}} \xrightarrow{d} Z \sim N(0,1)$$

so

$$\begin{aligned} & P(X_n = j) = P(j - 1/2 \le X_n \le j + 1/2) \\ \approx \;& P\left(\frac{(j-1/2)/n - (2p-1)}{\sqrt{4p(1-p)/n}} \le Z \le \frac{(j+1/2)/n - (2p-1)}{\sqrt{4p(1-p)/n}}\right) \\ = \;& \Phi\left(\frac{(j+1/2)/n - (2p-1)}{\sqrt{4p(1-p)/n}}\right) - \Phi\left(\frac{(j-1/2)/n - (2p-1)}{\sqrt{4p(1-p)/n}}\right) \end{aligned}$$

- e.g. $p = 3/4, n = 100, j = 40$, then

$$P(X_n = j) \approx \Phi(-1.0970) - \Phi(-1.2124) = 0.0236 \;\blacksquare$$

Proposition IV.11 (*Periodic Convergence Theorem*) If a MC is irreducible with period $b \geq 2$ and stationary distribution π, then for all $i, j \in \mathcal{S}$

$$\lim_{n\to\infty} \frac{1}{b} \sum_{k=0}^{b-1} p_{ij}^{(n+k)} = \pi_j$$

$$\lim_{n\to\infty} \frac{1}{b} \sum_{k=0}^{b-1} P(X_{n+k} = j) = \pi_j$$

$$\lim_{n\to\infty} \frac{1}{b} P\left(\cup_{k=0}^{b-1}\{X_{n+k} = j\}\right) = \pi_j$$

Proof: See book.

- again this implies that, if a stationary distribution exists for an irreducible MC, then it is unique

Example IV.9 *Ehrenfest's Urn*

- recall $\mathcal{S} = \{0, 1, 2, \ldots, d\}$ and

$$P = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 1/d & 0 & (d-1)/d & 0 & & \\ 0 & 2/d & 0 & (d-2)/d & & \\ & & & & & \\ & & & (d-1)/d & 0 & 1/d \\ 0 & & & & 1 & 0 \end{pmatrix}$$

- the chain is irreducible, recurrent with stationary distribution $\pi_i = \binom{d}{i} 2^{-d}$

- so $T(1) = \{n : p_{11}^{(n)} > 0\} = \{2, 4, \ldots\}$ and period$(1) = 2$ which is the period of all the states but we can't apply the MCCT

- but

$$\frac{1}{2} \sum_{k=0}^{1} p_{ij}^{(n+k)} = \frac{1}{2}(p_{ij}^{(n)} + p_{ij}^{(n+1)}) \longrightarrow \pi_j = \binom{d}{j} 2^{-d} \text{ as } n \longrightarrow \infty$$

■

Exercises

IV.1 Text 2.1.3

IV.2 Text 2.2.4

IV.3 Text 2.3.5

IV.4 Text 2.4.18

IV.5 Text 2.5.4