Probability and Stochastic Processes II - Lecture 4d

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

2024

IV.5 Random walks on graphs

Definition IV.4 A *weighted graph* is given by (V, w) where V is the *vertex set* and the *weight function* $w : V \times V \to [0, \infty)$ satisfies $w(u, v) = w(v, u)$.

- note - if $w(u, v) \in \{0, 1\}$ for all u, v, then this is an *unweighted graph* and (u, v) is an *edge* of the graph when $w(u, v) = 1$ and since this implies $w(v, u) = 1$ it is an *undirected graph* which will be assumed hereafter

Definition IV.5 For graph (V, w) and $u \in V$ the *degree* of u is

$$d(u) = \sum_{v \in V} w(u, v).$$

- assume hereafter that $d(u) > 0$ for all $u \in V$

Definition IV.6 For graph (V, w) define a *simple random walk* by the MC with state space $\mathcal{S} = V$ and transition probabilities

$$p_{uv} = \frac{w(u, v)}{d(u)}.$$

- for the unweighted graph the walk moves from u to one of the vertices connected to u with uniform probability $1/d(u)$

- put $Z = \sum_{u,v \in V} w(u,v)$ which in the unweighted case is 2 times the number of edges with different vertices plus the number of self-edges

Example IV.11 *Simple symmetric random walk on* $\mathbb{Z}$

- $V = \mathbb{Z}$ and $w(u,v) = 1$ when $|u - v| = 1$ and is 0 otherwise

- $\#(V) = \infty$, $d(u) = 2$ for all u, $p_{uv} = 1/2$ when $|u - v| = 1$ and is 0 otherwise and $Z = \infty$ ■

Example IV.11 *Random walk on* $\mathbb{Z}_m = \{0, \ldots, m-1\}$ *(Ring graph)*

- $V = \mathbb{Z}_m$ and

$$w(u,v) = \begin{cases} 1 & (u \pm 1) \equiv v \bmod m \\ 0 & \text{otherwise} \end{cases}$$

- then $d(u) = 2$ for all u, $p_{uv} = 1/2$ when $(u \pm 1) \equiv v \bmod m$ and is 0 otherwise and $Z = 2m$ ■

Proposition IV.4.14 (*Graph stationary distribution*) For graph (W, w) with *finite* Z, then $\pi_u = \frac{d(u)}{Z}$ is a stationary distribution.

Proof: We have

$$\pi_u p_{uv} = \frac{d(u)}{Z}\frac{w(u,v)}{d(u)} = \frac{w(u,v)}{Z} = \frac{w(v,u)}{Z} = \pi_v p_{vu}$$

and so the MC is time reversible wrt to π which implies π is a stationary distribution. ■

Definition IV.6 A graph (V, w) is *connected* if for any $u, v \in V$ there exist $u_0, u_1, \ldots, u_n \in V$ with $u_0 = u, u_n = v$ and $w(u_i, u_{i+1}) > 0$.

- for a connected graph the simple random walk is irreducible

- the period of the simple random walk on a connected graph is 1 or 2 and is 1 whenever there is a self-edge $w(u, u) > 0$

- simple random walk on $\mathbb{Z}_m$ has period 2

Definition IV.7 A graph (V, w) is a *bipartite graph whenever* $V = V_1 \cup V_2$ with $V_1 \cap V_2 = \phi$ and $w(u, v) > 0$ iff $u \in V_1, v \in V_2$ or conversely.

- a srw on a connected bipartite graph has period 2 and if the graph is not bipartite it has period 1 since $\{n : p_{uu}^{(n)} > 0\}$ contains 2 and an odd number for some $u \in V$ (Kőnig's Theorem: a graph is bipartitite iff all its cycles are even)

Proposition IV.4.15 (*Graph convergence theorem*) For a srw on a connected, nonbipartite graph (i) $\lim_{n\to\infty} p_{uv}^{(n)} = \pi_v$ and (ii) $\lim_{n\to\infty} P(X_n = v) = \pi_v$.

Proof: MCCT. ■

- there is a corresponding version for the periodic case

Exercises

IV.8 Text 2.7.10

IV.9 Text 2.7.11

IV.6 Mean Recurrence Times

Definition IV.8 For MC $\{X_n : n \in \mathbb{N}_0\}$ define the *mean return time* of state i by $m_i = E_i(\inf\{n \geq 1 : X_n = i\})$. A recurrent state is called *null recurrent* if $m_i = \infty$ and *positive recurrent* if $m_i < \infty$.

- so $m_i =$ is the average amount of time it takes the chain to return to state i having started in state i

- if state i is transient then $f_{ii} < 1$ so there is a positive probability the chain never returns which implies $m_i = \infty$

- note even with a recurrent state where $f_{ii} = 1$, so we are certain to return, but still could have $m_i = \infty$ which indicates that the distribution of the return time has a long tail

- let T_j^r be the time of the r-th visit to state j then

$$
\begin{aligned}
& P_j(T_j^1 = k) = P_j(X_1 \neq j, \ldots, X_{k-1} \neq j, X_k = j) \\
= & \sum_{i_1,\ldots,i_{k-1} \in \mathcal{S} \setminus \{j\}} P_j(X_1 = i_1, \ldots, X_{k-1} = i_{k-1}, X_k = j) \\
= & \sum_{i_1,\ldots,i_{k-1} \in \mathcal{S} \setminus \{j\}} p_{ji_1} \cdots p_{i_{k-1}j} \\
& P_j(T_j^1 = k, T_j^2 - T_j^1 = l) \\
= & P_j \left(\begin{array}{l} X_1 \neq j, \ldots, X_{k-1} \neq j, X_k = j, \\ X_{k+1} \neq j, \ldots, X_{k+l-1} \neq j, X_{k+l} = j \end{array} \right) \\
= & \sum_{i_1,\ldots,i_{k-1}, i_{k+1},\ldots,i_{k+l-1} \in \mathcal{S} \setminus \{j\}} P_j \left(\begin{array}{l} X_1 = i_1, \ldots, X_{k-1} = i_{k-1}, X_k = j, \\ X_{k+1} = i_{k+1}, \ldots, X_{k+l-1} = i_{k+l-1}, \\ X_{k+l} = j \end{array} \right) \\
= & \sum_{i_1,\ldots,i_{k-1}, i_{k+1},\ldots,i_{k+l-1} \in \mathcal{S} \setminus \{j\}} p_{ji_1} \cdots p_{i_{k-1}j} p_{ji_{k+1}} \cdots p_{i_{k+l-1}j} \\
= & \sum_{i_1,\ldots,i_{k-1} \in \mathcal{S} \setminus \{j\}} p_{ji_1} \cdots p_{i_{k-1}j} \sum_{i_{k+1},\ldots,i_{k+l-1} \in \mathcal{S} \setminus \{j\}} p_{ji_{k+1}} \cdots p_{i_{k+l-1}j} \\
= & P_j(T_j^1 = k) P_j(T_j^2 - T_j^1 = l)
\end{aligned}
$$

which implies that from distribution P_j then $T_j^1, T_j^2 - T_j^1$ are statistically independent with the same distribution and similarly $T_j^1, T_j^2 - T_j^1, \ldots, T_j^r - T_j^{r-1}$ are i.i.d. with mean $E_j(T_j^1) = m_j$

- also if we start from state i, then $T_j^2 - T_j^1, \ldots, T_j^r - T_j^{r-1}$ are i.i.d. with mean m_j

- define $r_j(n) = \#\{l : X_l = j, 1 \leq l \leq n\} =$ number of visits to j in first n steps and put

$$G_n(i,j) = E_i(r_j(n)) = E_i\left(\sum_{l=1}^{n} I_{\{j\}}(X_l)\right) = \sum_{l=1}^{n} p_{ij}^{(l)}$$

Lemma IV.16 If MC is irreducible, recurrent, $\lim_{n\to\infty} G_n(i,j)/n = 1/m_j$.

Proof: Now $T_j^2 - T_j^1, \ldots, T_j^r - T_j^{r-1}$ are i.i.d. with mean m_j and $T_j^r = T_j^1 + (T_j^2 - T_j^1) + \cdots + (T_j^r - T_j^{r-1})$.Also $P_i(T_j^1 < \infty) = 1$ so $T_j^1/r \overset{wp1}{\to} 0$.This together with the SLLN implies, as $r \to \infty$

$$\frac{T_j^r}{r} = \frac{T_j^1}{r} + \frac{r-1}{r}\frac{1}{r-1}\sum_{k=2}^{r}(T_j^k - T_j^{k-1}) \overset{wp1}{\to} m_j.$$

Since j is recurrent then $r_j(n) \overset{wp1}{\rightarrow} \infty$ and since $n \leq$ the time of the $(r_j(n)+1)$-st visit,

$$\frac{T_j^{r_j(n)}}{r_j(n)} \leq \frac{n}{r_j(n)} \leq \frac{T_j^{r_j(n)+1}}{r_j(n)}$$

which implies

$$m_j \leq \liminf \frac{n}{r_j(n)} \leq \limsup \frac{n}{r_j(n)} \leq m_j$$

and so $\lim_{n\to\infty} n/r_j(n) = m_j$. Now $G_n(i,j) = E_i(r_j(n))$ which implies

$$\begin{aligned} \lim_{n\to\infty} \frac{G_n(i,j)}{n} &= \lim_{n\to\infty} E_i\left(\frac{r_j(n)}{n}\right) \\ &= E_i\left(\lim_{n\to\infty}\frac{r_j(n)}{n}\right) \text{ by DCT since } 0 \leq r_j(n)/n \leq 1 \\ &= 1/m_j. \end{aligned}$$

■

- note that $0 \leq 1/m_j \leq 1$

note - if $x_n \to x$ then for $\epsilon > 0$ there exists n_ϵ s.t. $|x_n - x| < \epsilon$ for $n \geq n_\epsilon$ and

$$\left| \frac{1}{n} \sum_{k=1}^{n} x_k - x \right| = \left| \frac{1}{n} \sum_{k=1}^{n} (x_k - x) \right| \leq \frac{1}{n} \sum_{k=1}^{n} |x_k - x|$$

$$= \frac{1}{n} \sum_{k=1}^{n_\epsilon - 1} |x_k - x| + \frac{1}{n} \sum_{k=n_\epsilon}^{n} |x_k - x| \leq \frac{1}{n} \sum_{k=1}^{n_\epsilon - 1} |x_k - x| + (1 - \frac{n_\epsilon}{n})\epsilon \to \epsilon$$

which implies $\frac{1}{n} \sum_{k=1}^{n} x_k \to x$ (Cesaro summation)

- so by MCCT since $p_{ij}^{(n)} \to \pi_j$, then $\frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)} \to \pi_j$ and similarly for the periodic case (average blocks)

Proposition IV.17 If MC is irreducible and each state i is positive recurrent, then the MC has a unique stationary distribution π with $\pi_i = 1/m_i$.

Proof: Suppose $\sum_i \alpha_i p_{ij} = \alpha_j$ for all i, j for some distribution α. As before this implies $\sum_i \alpha_i p_{ij}^{(n)} = \alpha_j$ for all i, j for all n. Therefore, applied to the periodic and aperiodic cases,

$$\begin{aligned} \alpha_j &= \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \sum_i \alpha_i p_{ij}^{(k)} = \sum_i \alpha_i \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)} \text{ by DCT and Cesaro} \\ &= \sum_i \alpha_i \frac{1}{m_j} \text{ (by Lemma IV.16)} = \frac{1}{m_j} \end{aligned}$$

and so, if a stationary distribution exists, it is unique.
Now

$$c = \sum_j \frac{1}{m_j} = \sum_j \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)} \leq \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \sum_j p_{ij}^{(k)} = 1 \qquad (1)$$

where the inequality follows since, if

$$1 = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \sum_{j} p_{ij}^{(k)} < \sum_{j} \frac{1}{m_j}$$

then there exist l states $j_1, \ldots, j_l$, s.t.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \sum_{j} p_{ij}^{(k)} < \sum_{j\in\{j_1,\ldots,j_l\}} \frac{1}{m_j} \text{ but}$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \sum_{j} p_{ij}^{(k)} \geq \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \sum_{j\in\{j_1,\ldots,j_l\}} p_{ij}^{(k)}$$

$$= \sum_{j\in\{j_1,\ldots,j_l\}} \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)} = \sum_{j\in\{j_1,\ldots,j_l\}} \frac{1}{m_j} \text{ by Lemma IV.16}$$

which is a contradiction.

Next for any state i, (and using the same argument for the inequality (1)) by Lemma IV.16

$$\begin{aligned} \frac{1}{m_j} &= \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} p_{ij}^{(t+1)} = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \sum_k p_{ik}^{(t)} p_{kj} \\ &\geq \sum_k \left(\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} p_{ik}^{(t)} \right) p_{kj} = \sum_k \frac{p_{kj}}{m_k} \end{aligned}$$

and so

$$\sum_j \frac{1}{m_j} \geq \sum_j \sum_k \frac{p_{kj}}{m_k} \overset{DCT}{=} \sum_k \sum_j \frac{p_{kj}}{m_k} = \sum_k \frac{1}{m_k}$$

which implies

$$\frac{1}{m_j} = \sum_k \frac{p_{kj}}{m_k}$$

which in turn implies $\pi_j = 1/cm_j$ is the unique stationary distribution. Since always $\frac{1}{n} \sum_{l=1}^{n} p_{ij}^{(l)} \to \pi_j$, then Lemma IV.16 implies $c = 1$. ■

Proposition IV.18 If a MC has a stationary distribution π and state j is not positive recurrent then $\pi_j = 0$.
Proof: We have

$$\begin{aligned} \pi_j &= \sum_i \pi_i p_{ij}^{(n)} \text{ for every } n \\ &= \frac{1}{n}\sum_{t=1}^{n}\sum_i \pi_i p_{ij}^{(t)} = \sum_i \pi_i \frac{1}{n}\sum_{t=1}^{n} p_{ij}^{(t)} \\ &= \lim_{n\to\infty}\sum_i \pi_i \frac{1}{n}\sum_{t=1}^{n} p_{ij}^{(t)} = \sum_i \pi_i \lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n} p_{ij}^{(t)} \text{ by DCT} \\ &= \sum_i \pi_i \frac{1}{m_j} = \sum_i \pi_i 0 = 0 \text{ since } m_j = \infty. \end{aligned}$$

■

Corollary IV.19 A MC with no positive recurrent states does not have a stationary distribution.

Proposition IV.20 If states i and j mutually communicate then if one is positive recurrent so is the other.

Proof: There exist $r, s > 0$ s.t. $p_{ij}^{(r)} > 0, p_{ji}^{(s)} > 0$. Therefore, using Lemma IV.16 for recurrent states and the result for transient states

$$\begin{aligned}
\frac{1}{m_j} &= \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} p_{jj}^{(t)} \geq \lim_{n\to\infty} \frac{1}{n}\sum_{t=s+r}^{n} p_{ji}^{(s)} p_{ii}^{(t-s-r)} p_{ij}^{(r)} \text{ by Chapman-Kolmog} \\
&= p_{ji}^{(s)} \left(\lim_{n\to\infty} \frac{n-s-r}{n} \frac{1}{n-s-r} \sum_{t=s+r}^{n} p_{ii}^{(t-s-r)} \right) p_{ij}^{(r)} \\
&= p_{ji}^{(s)} \frac{1}{m_i} p_{ij}^{(r)}
\end{aligned}$$

and so $1/m_j > 0$ when $1/m_i > 0$ so m_j is finite whenever m_i is finite and conversely. ■

Corollary IV.21 An irreducible MC either has all states positive recurrent or no states positive recurrent.

- so an irreducible chain has a stationary distribution given by π_i iff it is a positive recurrent chain

Proposition IV.22 An irreducible MC with a finite state space is positive recurrent and so has a stationary distribution.

Proof: We have

$$\lim_{n\to\infty} P_i(X_l = j \text{ for some } l, 1 \le l \le n) = f_{ij} > 0$$

for every i and j. Since $\mathcal{S}$ is finite, then there exists $\delta \in (0,1)$ and m s.t.

$$P_i(X_l = j \text{ for some } l, 1 \le l \le m) > \delta$$

for every i and j. Therefore,

$$P_i(X_l \neq j \text{ for all } l, 1 \le l \le m) < 1 - \delta$$

for every i and j.

Now consider

$$
\begin{aligned}
& P_i(X_l \neq i \text{ for } l, 1 \leq l \leq 2m) \\
= & \sum_{j \in \mathcal{S} \setminus \{i\}} P_i(X_l \neq i \text{ for } l \neq m, 1 \leq l \leq 2m, X_m = j) \\
= & \sum_{j \in \mathcal{S} \setminus \{i\}} \left(\begin{array}{l} P_i(X_l \neq i \text{ for } 1 \leq l \leq m-1, X_m = j) \times \\ P_j(X_l \neq i \text{ for}, 1 \leq l \leq m) \end{array} \right) \text{ using MP and T} \\
\leq & \left(\sum_{j \in \mathcal{S} \setminus \{i\}} P_i(X_l \neq i \text{ for } 1 \leq l \leq m-1, X_m = j) \right) (1-\delta) \\
= & P_i(X_l \neq i \text{ for } l, 1 \leq l \leq m)(1-\delta) \leq (1-\delta)^2
\end{aligned}
$$

and similarly

$$P_i(X_l \neq i \text{ for } 1 \leq l \leq km) \leq (1-\delta)^k.$$

Then for n satisfying $km \le n < (k+1)m$, we have ...

$$\begin{aligned} & P_i(X_l \neq i \text{ for } l, 1 \le l \le n) \\ \le \;& P_i(X_l \neq i \text{ for } l, 1 \le l \le km) \le (1-\delta)^k = (1-\delta)^{\lfloor n/m \rfloor}. \end{aligned}$$

This implies

$$\begin{aligned} m_i &= E_i(T_i^1) = \sum_{n=0}^{\infty} P(T_i^1 > n) = \sum_{n=0}^{\infty} P_i(X_l \neq i \text{ for } l, 1 \le l \le n) \\ &\le \sum_{n=0}^{\infty} (1-\delta)^{\lfloor n/m \rfloor} = \sum_{k=0}^{\infty} m(1-\delta)^k = m/\delta < \infty \end{aligned}$$

which proves the result. ■

IV.7 Markov chains of higher order

- suppose stochastic process $\{X_n : n \in \mathbb{N}_0\}$ satisfies

$$P(X_n \,|\, X_0, \ldots, X_{n-1}) = P(X_n \,|\, X_{n-2}, X_{n-1}),$$

a MC of order 2

- note a full description of the s.p. requires an intial distribution on (X_0, X_1) then, for example,

$$\begin{aligned} & P(X_0, X_1, \ldots, X_n) = P(X_2, X_3, \ldots, X_n \,|\, X_0, X_1) P(X_0, X_1) \\ = \; & P(X_3, \ldots, X_n \,|\, X_0, X_1, X_2) P(X_2 \,|\, X_0, X_1) P(X_0, X_1) \\ = \; & P(X_3, \ldots, X_n \,|\, X_1, X_2) P(X_2 \,|\, X_0, X_1) P(X_0, X_1) \text{ etc.} \end{aligned}$$

- define a s.p. by $\{Y_n : n \in \mathbb{N}_0\}$ by $Y_n = (X_n, X_{n+1})$, then

$$\begin{aligned} & P(Y_n \mid Y_0, \ldots, Y_{n-1}) \\ = \; & P((X_n, X_{n+1}) \mid (X_0, X_1), \ldots, (X_{n-1}, X_n)) \\ = \; & P((X_n, X_{n+1}) \mid X_0, X_1, \ldots, X_{n-1}, X_n) \\ = \; & P((X_n, X_{n+1}) \mid X_{n-1}, X_n) \text{ by MP of order 2} \\ = \; & P(Y_n \mid Y_{n-1}) \end{aligned}$$

and so $\{Y_n : n \in \mathbb{N}_0\}$ is a 2-dimensional MC with these transition probabilities and initial distribution specified by $Y_0 = (X_0, X_1)$

- so all the results for MC's apply here and clearly this can be generalized to MC's of order r where

$$P(X_n \mid X_0, \ldots, X_{n-1}) = P(X_n \mid X_{n-r}, \ldots, X_{n-1})$$