Probability and Stochastic Processes II - Lecture 4a

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

2024

IV Markov Chain Convergence

IV 1 Stationary Distributions

- suppose $\{X_n : n \in \mathbb{N}_0\}$ is a MC and assume $\lim_{n\to\infty} P(X_n = j) = \pi_j$ for every j so

$$\begin{aligned}
\pi_j &= \lim_{n\to\infty} P(X_{n+1} = j) \\
&= \lim_{n\to\infty} \sum_{i\in\mathcal{S}} P(X_n = i) P(X_{n+1} = j \,|\, X_n = i) \text{ by TTP} \\
&= \lim_{n\to\infty} \sum_{i\in\mathcal{S}} P(X_n = i) p_{ij} \text{ by TH} \\
&= \sum_{i\in\mathcal{S}} \lim_{n\to\infty} P(X_n = i) p_{ij} \text{ (assume limit and sum can be interchanged)} \\
&= \sum_{i\in\mathcal{S}} \pi_i p_{ij}
\end{aligned}$$

so these limiting probabilities, when they exist, satisfy the system of linear equations $\pi P = \pi$

- in particular π is a left eigenvector of P associated with the eigenvalue 1

Example IV.1

- recall Example III.2

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

- then the probability distributions of X_1, X_{10} and X_{100} are given by

$$
\begin{aligned}
vP^1 &= (0.2291667, 0.1458333, 0.4166667, 0.2083333) \\
vP^{10} &= (0.2121419, 0.2121462, 0.3030114, 0.2727005) \\
vP^{100} &= (0.2121212, 0.2121212, 0.3030303, 0.2727273)
\end{aligned}
$$

which definitely indicates convergence ■

Definition IV.1 A probability distribution π on S is *stationary* for the MC with transition matrix P if $\pi P = \pi$, $(\sum_{i \in \mathcal{S}} \pi_i p_{ij} = \pi_j$ for every j.)■

- suppose π is the initial distribution then

$$(\ldots, P(X_n = j-1), P(X_n = j), P(X_n = j+1), \ldots) = \pi P^n = \pi$$

and so the distribution of X_n is the same (stationary) for all n

Example IV.2 *(double stochasticity)*

- suppose S is finite and P is doubly stochastic (each column also sums to 1)

- then if $\pi_i = 1/\#(S)$ for all i, then

$$\sum_{k \in \mathcal{S}} \pi_k p_{kj} = \frac{1}{\#(S)} \sum_{k \in \mathcal{S}} p_{kj} = \frac{1}{\#(S)} = \pi_j$$

and the uniform distribution on S is stationary
■

- how to obtain π when it exists?

- calculate vP^n for large n or compute left eigenvalues of P

Example IV.1 (continued)

- $P^t \pi^t = Q\pi^t$ so π^t is a right eigenvector of $Q = P^t$ associated with eigenvalue 1

- in R (recall any constant times an eigenvector is an eigenvector associated with the same eigenvalue so need to normalize) when P is finite dimensional

```
#computing stationary distribution
r1=c(0,1/3,1/2,1/6)
r2=c(1/3,0,1/2,1/6)
r3=c(1/6,1/6,0,2/3)
r4=c(1/3,1/3,1/3,0)
P=rbind(r1,r2,r3,r4)
Q=t(P)
e=eigen(Q)
evals=e$values
evals
E=e$vectors
E
norm=Re(sum(E[,1]))
pi=Re(t(E[,1]/norm))
pi
pi%*%P
```

```
- output
> evals
[1] 1.0000000+0.0000000i -0.3333333+0.2357023i -0.3333333-0.2357023i -0.3333333+0.0000000i
> E=e$vectors
> E
     [,1]          [,2]                    [,3]                    [,4]
[1,] 0.4190790+0i 0.2581989+0.1825742i 0.2581989-0.1825742i -7.071068e-01+0i
[2,] 0.4190790+0i 0.2581989+0.1825742i 0.2581989-0.1825742i 7.071068e-01+0i
[3,] 0.5986843+0i 0.2581989-0.3651484i 0.2581989+0.3651484i 3.238467e-16+0i
[4,] 0.5388159+0i -0.7745967+0.0000000i -0.7745967+0.0000000i -5.867608e-16+0i
> norm=Re(sum(E[,1]))
> pi=Re(t(E[,1]/norm))
> pi
     [,1]      [,2]      [,3]      [,4]
[1,] 0.2121212 0.2121212 0.3030303 0.2727273
> pi%*%P
     [,1]      [,2]      [,3]      [,4]
[1,] 0.2121212 0.2121212 0.3030303 0.2727273
```

■

Definition IV.2 A MC is *time reversible* (or satisfies *detailed balance*) wrt probability distribution π if $\pi_i p_{ij} = \pi_j p_{ji}$ for every $i, j \in \mathcal{S}$. ■

- so suppose at time n the chain is at state j but you don't know what the previous states were, then when the chain is time reversible and the initial distribution was π

$$P(X_{n-1} = i \,|\, X_n = j) = \frac{P(X_{n-1} = i, X_n = j)}{P(X_n = j)}$$

$$\begin{aligned}
&= \frac{\sum_{i_0,\ldots,i_{n-2}} \pi_{i_0} p_{i_0 i_1} \cdots p_{i_{n-2} i} p_{ij}}{\sum_{i_0,\ldots,i_{n-1}} \pi_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} j}} \\
&= \frac{\sum_{i_0,\ldots,i_{n-2}} \pi_{i_1} p_{i_1 i_0} p_{i_1 i_2} \cdots p_{i_{n-2} i} p_{ij}}{\sum_{i_0,\ldots,i_{n-1}} \pi_{i_1} p_{i_1 i_0} \cdots p_{i_{n-1} j}} \\
&= \frac{\sum_{i_1,\ldots,i_{n-2}} \pi_{i_2} p_{i_2 i_1} p_{i_2 i_3} \cdots p_{i_{n-2} i} p_{ij}}{\sum_{i_1,\ldots,i_{n-1}} \pi_{i_2} p_{i_1 i_2} \cdots p_{i_{n-1} j}} \\
&= \cdots = \frac{\sum_{i_{n-2}} \pi_{i_{n-2}} p_{i_{n-2} i} p_{ij}}{\sum_{i_{n-1}} \pi_{i_{n-1}} p_{i_{n-1} j}} \\
&= \frac{\pi_j p_{ji}}{\pi_j} = p_{ji}
\end{aligned}$$

- also

$$
\begin{aligned}
& P(X_{n-1}=i \mid X_n=j, X_{n+1}=k) \\
= \; & \frac{P(X_{n-1}=i, X_n=j, X_{n+1}=k)}{P(\, X_n=j, X_{n+1}=k)} \\
= \; & \frac{P(X_{n+1}=k \mid X_{n-1}=i, X_n=j) P(X_{n-1}=i, X_n=j)}{P(\, X_{n+1}=k \mid X_n=j) P(X_n=j)} \\
= \; & \frac{P(X_{n+1}=k \mid X_n=j) P(X_{n-1}=i, X_n=j)}{P(\, X_{n+1}=k \mid X_n=j) P(X_n=j)}=P(X_{n-1}=i \mid X_n=j)
\end{aligned}
$$

- so running a chain backwards is always a MC (conditioning on the present and the future only depends on the present) and time reversibility means this has the same transition probabilities

Proposition IV.1 If a MC is time reversible wrt π, then π is stationary.

Proof: We have
$$\sum_{k\in\mathcal{S}} \pi_k p_{kj} = \sum_{k\in\mathcal{S}} \pi_j p_{jk} = \pi_j.$$

■

Example IV.3 *(Ehrenfest's Urn)*

- two urns Urn I and Urn II and d labelled balls (labelled $1, \ldots, d$) in total distributed into the two urns

- at each time one ball is chosen at "random" and moved to the other urn

- let $X_n = \#$ of balls in Urn I at time n so for $i, j \in S = \{0, 1, \ldots, d\}$

$$P(X_n = j \,|\, X_0 = i_0, \ldots, X_{n-1} = i) = P(X_n = j \,|\, X_{n-1} = i)$$

$$= \begin{cases} \frac{i}{d} & j = i - 1 \\ \frac{d-i}{d} & j = i + 1 \\ 0 & \text{otherwise} \end{cases} = p_{ij}$$

- so this is a time homogeneous MC with $P \in \mathbb{R}^{(d+1)\times(d+1)}$ given by

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

- it is clear that $\{X_n : n \in \mathbb{N}_0\}$ is an irreducible MC and since S is finite it is a recurrent chain

- if there is a distribution π s.t. $\pi_i p_{ij} = \pi_j p_{ji}$ it must satisfy the equations below (this is expressed pairwise in terms of the off-diagonal elements) and multiplying all elements of P by d

$$\begin{aligned}
\pi_1 &= d\pi_0 = \binom{d}{1}\pi_0 \\
2\pi_2 &= (d-1)\pi_1 = d(d-1)\pi_0 \text{ or } \pi_2 = \frac{d(d-1)}{2}\pi_0 = \binom{d}{2}\pi_0 \\
3\pi_3 &= (d-2)\pi_2 = \frac{d(d-1)(d-2)}{2}\pi_0 \text{ or } \pi_3 = \binom{d}{3}\pi_0 \\
&\vdots \\
d\pi_d &= \pi_{d-1} \text{ or } \pi_d = \binom{d}{d}\pi_0
\end{aligned}$$

- also π must satisfy

$$1 - \pi_0 = \sum_{i=1}^{d} \pi_i = \pi_0 \sum_{i=1}^{d} \binom{d}{i} = \pi_0(2^d - 1) \text{ so } \pi_0 = 1/2^d$$

which implies that detailed balance is satisfied by

$$\pi_i = \binom{d}{i} 2^{-d} \text{ for } i \in S$$

- therefore the MC is time reversible wrt π and π (the binomial$(d, 1/2)$ distribution) is a stationary distribution for the chain

■

- not all chains with stationary distributions are time reversible

Proposition IV.2 For a MC, if $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ for all i, j, then the chain doesn't have a stationary distribution.

Proof: Suppose a stationary distribution π exists. Then we have $\pi P^n = \pi$ for all n so $\lim_{n\to\infty} \pi P^n = \pi$ or

$$\begin{aligned} \pi_j &= \lim_{n\to\infty} \sum_{i\in\mathcal{S}} \pi_i p_{ij}^{(n)} = \lim_{n\to\infty} \int_{\mathcal{S}} p_{ij}^{(n)} \Pi(di) \text{ and by DCT} \\ &= \int_{\mathcal{S}} \lim_{n\to\infty} p_{ij}^{(n)} \Pi(di) \text{ since } |p_{ij}^{(n)}| \le 1 \text{ and } \int_{\mathcal{S}} 1\Pi(di) = \sum_{i\in\mathcal{S}} \pi_i = 1 \\ &= 0 \end{aligned}$$

so $\pi_j = 0$ for all j and so $\sum_{i\in\mathcal{S}} \pi_i = 0$ which is a contradiction. ■

Lemma IV.3 For a MC, if $\lim_{n\to\infty} p_{kl}^{(n)} = 0$ for some k, l and $k \to i, j \to l$, then $\lim_{n\to\infty} p_{ij}^{(n)} = 0$.

Proof: There eixist $r, s > 0$ s.t. $p_{ki}^{(r)} > 0, p_{jl}^{(s)} > 0$ so $p_{kl}^{(r+n+s)} \ge p_{ki}^{(r)} p_{ij}^{(n)} p_{jl}^{(s)}$ which implies $0 \le p_{ij}^{(n)} \le p_{kl}^{(r+n+s)} / p_{ki}^{(r)} p_{jl}^{(s)} \to 0$ as $n \to \infty$. ■

Corollary IV.4 For an irreducible MC (i) either $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ for all i, j or $\lim_{n\to\infty} p_{ij}^{(n)} \neq 0$ for all i, j, (ii) so if $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ for some i, j the MC does not have a stationary distribution and in particular (iii) a transient chain does not have a stationary distribution.

Proof: (i) is immediate from Lemma IV.3 and (ii) follows from Proposition IV.2. For (iii), the Cases Theorem implies $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$ for all i, j which implies $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ for all i, j and the result follows from Proposition IV.2. ■

Example IV.4 *Simple random walk*

- the chain is irreducible so the previous result applies

- when $p \neq 1/2$ we showed the chain is transient and so no stationary distribution exists

- when $p = 1/2$ we showed the chain is recurrent so we need only show $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ for some i, j to show that a stationary distribution doesn't exist

- we have

$$p_{00}^{(n)} = \begin{cases} 0 & \text{if } n \text{ not even} \\ \binom{n}{n/2}\left(\frac{1}{2}\right)^n & \text{if } n \text{ even} \end{cases}$$

and for $n = 2m$, using results from Example III.2.1, namely, $t_m = 2^{2m}/\sqrt{\pi m}$

$$\begin{aligned} p_{00}^{(2m)} &= \binom{2m}{m}\left(\frac{1}{2}\right)^{2m} = \frac{(2m)!}{m!m!}\left(\frac{1}{2}\right)^{2m} = \left(\frac{(2m)!}{m!m!t_m}\right) t_m \left(\frac{1}{2}\right)^{2m} \\ &= \left(\frac{(2m)!}{m!m!t_{2m}}\right)\frac{1}{\sqrt{\pi m}} \to 0 \text{ as } m \to 0 \end{aligned}$$

- therefore a stationary distribution doesn't exist for any srw

- note - when $p = 1/2$, by the Cases Theorem

$$\sum_n p_{ij}^{(n)} = \infty \text{ even though } p_{ij}^{(n)} \to 0$$

also

$$\sum_j p_{ij}^{(n)} = 1 \text{ for every } n \text{ so } \lim_{n\to\infty} \sum_j p_{ij}^{(n)} = 1 \neq 0 = \sum_j \lim_{n\to\infty} p_{ij}^{(n)}$$

so this is a situation where we can't interchange summation and limit (can't use MCT or DCT) ■

- see the book for a MC with $\#(S) = \infty$ and a stationary distribution exists