STA 447/2006 Lecture 4

Midterm exam #1

- During class 6:15pm – 8:45pm
- 4 pages of cheat sheet (8 sides)
- No electronics, no calculators.
- Find previous years' midterm on course website
- Office hours: 2-hr for TA's, prof. announced on quercus.
- Coverage: first 4 weeks.

Recall from last time: stationary distribution, convergence / non-convergence to stationary.

(convergence is false for reducible, periodic, non-existing stat. distr.)

Thm. If P is irreducible, aperiodic, $\exists$ stationary distribution π.

then $\forall\, i,j \in S$, $\lim_{n\to\infty} P_{ij}^{(n)} = \pi_j$.

(implies uniqueness of π under these assumptions)

Proof. Main lemma (Markov forgetting).

$\forall i, j, k \in S$ $\quad \lim_{n \to +\infty} \left| P_{ik}^{(n)} - P_{jk}^{(n)} \right| = 0.$

Proof of the forgetting lemma.

Idea: coupling

Want to compare $(X_n^{(1)})_{n \geq 0}$ (starting from $i \in S$)

with $(X_n^{(2)})_{n \geq 0}$ (starting from $j \in S$).

Construct a new MC on $\bar{S} := S \times S$.

starting from $(i, j) \in \bar{S}$

where $\underline{(X_n^{(1)})_{n \geq 0}}$ and $(X_n^{(2)})_{n \geq 0}$ evolve independently.

i.e. $\quad \bar{P}_{(i_1, i_2), (i_3, i_4)} = P_{i_1, i_3} \cdot P_{i_2, i_4}$

Facts about the joint chain.

- P has stat. distr. π

$\Rightarrow$ joint chain has stat. distr.

$$\bar{\pi}_{(i,j)} = \pi_i \cdot \pi_j$$

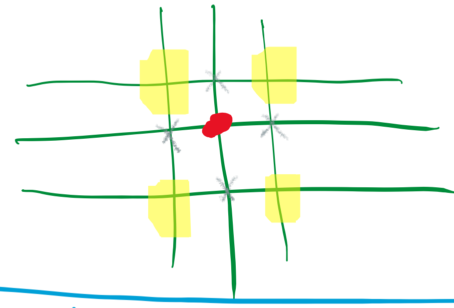
- irreducible & aperiodic P $\Rightarrow$ irreducible joint chain.

Avoid pathology. e.g.

1-D SRW is irreducible, period $=2$.

2-D SRW (running 2 1-D SRWs independently)

reducible



Lemma: If state $i \in S$ is aperiodic and $f_{ii} > 0$.

then $\exists\, n_0(i) \in \mathbb{N}_+$ s.t. $\forall n \geq n_0(i)$, $P_{ii}^{(n)} > 0$.

(Proof idea: $A = \{n : P_{ii}^{(n)} > 0\}$.

$m \in A$, $n \in A$ then $m+n \in A$ $\left(P_{ii}^{(m+n)} \geq P_{ii}^{(m)} \cdot P_{ii}^{(n)}\right)$

$am+bn \in A$ $(a, b \in \mathbb{N}^+)$

then invoke Bezout identity in number theory)

Using this lemma, we can conclude

Corollary: under irreducible & aperiodic, $\forall i, j \in S$ $\exists\, n_0(i,j) > 0$,

s.t. $P_{ij}^{(n)} > 0$ $\forall n > n_0(i,j)$.

(Proof. $P_{ij}^{(r)} > 0$ for some $r > 0$, $P_{jj}^{(n)} > 0$ $\forall n \geq n_0(j)$,

So $P_{ij}^{(r+n)} > 0$ $\forall n \geq n_0(j)$. Let $n_0(i,j) = r + n_0(j)$).

For the joint chain, we want to show

$P_{(i,j),(k,l)}^{(n)} > 0$ for sufficiently large n.

$$P^{(n)}_{(i,j)(k,l)} = P^{(n)}_{ik} \cdot P^{(n)}_{jl} > 0 \qquad n \geq \max\{n_0(i,k), n_0(j,l)\}$$

therefore, irreducible & aperiodic $P \Rightarrow$ irreducible joint chain (aperiodic).

Joint chain has stationary $\bar{\pi} \Rightarrow$ recurrence.

Starting from $(i,j) \in \bar{S}$, choose any $i_0 \in S$ and fix it.

$$\tau = \inf\{n \geq 1, (X^{(1)}_n, X^{(2)}_n) = (i_0, i_0)\} < +\infty \quad (w.p. 1).$$

$(\forall k \in S)$

$$p^{(n)}_{ik} = \mathbb{P}_{(i,j)}(X^{(1)}_n = k)$$

$$= \sum_{m=1}^{+\infty} \mathbb{P}_{(i,j)}(X^{(1)}_n = k, \tau = m).$$

$$= \underbrace{\sum_{m=1}^{n} \mathbb{P}_{(i,j)}(X^{(1)}_n = k, \tau = m)}_{\text{meet at } i_0 \text{ before } n} + \underbrace{\sum_{m=n+1}^{+\infty} \mathbb{P}_{i,j}(X^{(1)}_n = k, \tau = m)}_{\text{---- after } n.}.$$

For $m \leq n$,

$$\mathbb{P}_{(i,j)}(X^{(1)}_n = k, \tau = m)$$

$$= \mathbb{P}_{(i,j)}(\tau = m) \cdot \underbrace{\mathbb{P}_{(i,j)}(X^{(1)}_n = k \mid \tau = m)}_{\text{By strong Markov property, } = P^{(n-m)}_{i_0 k}}$$

$$\sum_{m=1}^{n} \mathbb{P}_{(i,j)}\left(X_n^{(1)}=k,\ \tau=m\right)$$

$$= \sum_{m=1}^{n} \mathbb{P}_{(i,j)}(\tau=m)\cdot p_{l_0 k}^{(n-m)}$$

(by symmetry)

$$= \sum_{m=1}^{n} \mathbb{P}_{(j,i)}(\tau=m)\cdot p_{l_0 k}^{(n-m)} = \sum_{m=1}^{n} \mathbb{P}_{(i,j)}\left(X_n^{(2)}=k,\ \tau=m\right)$$

Doing the same decomposition for $p_{jk}^{(n)} = \mathbb{P}_{(i,j)}\left(X_n^{(2)}=k\right)$

First sum term cancels

So we have

$$\left|p_{ik}^{(n)} - p_{jk}^{(n)}\right| \le \sum_{m=n+1}^{+\infty} \left|\mathbb{P}_{(i,j)}\left(X_n^{(1)}=k,\ \tau=m\right) - \mathbb{P}_{(i,j)}\left(X_n^{(2)}=k,\ \tau=m\right)\right|$$

$$\le 2\sum_{m=n+1}^{+\infty} \mathbb{P}_{(i,j)}(\tau=m)$$

$$= 2\cdot \mathbb{P}_{(i,j)}(\tau \ge n+1) \longrightarrow 0 \quad (\text{as } n\to+\infty)$$

since $\tau<+\infty$ w.p. 1.

This proves the forgetting lemma.

In general, coupling construction is used to prove speed of convergence. (e.g. card shuffling)

Make the two particles meet as early as possible.

From Markov forgetting to convergence.

Counter-example: 1-D Lazy SRW, $X_{n+1} = \begin{cases} X_n + 1 & \text{w.p. } \frac{1}{4} \\ X_n - 1 & \text{w.p. } \frac{1}{4} \\ X_n & \text{w.p. } \frac{1}{2} \end{cases}$

Recurrence of joint chain is true

So we have Markov forgetting.

But we know that

$$\lim_{n\to+\infty} P_{ik}^{(n)} = 0, \quad \lim_{n\to+\infty} P_{jk}^{(n)} = 0.$$

Need to use "$\exists$ stationary".

$$\left| P_{ij}^{(n)} - \pi_j \right| = \left| P_{ij}^{(n)} - \underbrace{\sum_{k\in S} \pi_k P_{kj}^{(n)}}_{\text{marginal of } X_n \text{ for } X_0 \sim \pi} \right|$$

(Idea: compare a chain from i to a chain from π)

$$\leq \sum_{k\in S} \pi_k \cdot \underbrace{\left| P_{ij}^{(n)} - P_{kj}^{(n)} \right|}_{\to 0}$$

Justify "interchanging Σ and lim":

M-test.

$$\sum_{k\in S} \pi_k \cdot \sup_{n\geq 0} \left| P_{ij}^{(n)} - P_{kj}^{(n)} \right| \leq \sum_{k\in S} \pi_k = 1 < +\infty.$$

So we have $\lim_{n\to+\infty} \left| P_{ij}^{(n)} - \pi_j \right| = 0$. QED.

Similarly, if MC starts from $X_0 \sim \nu$

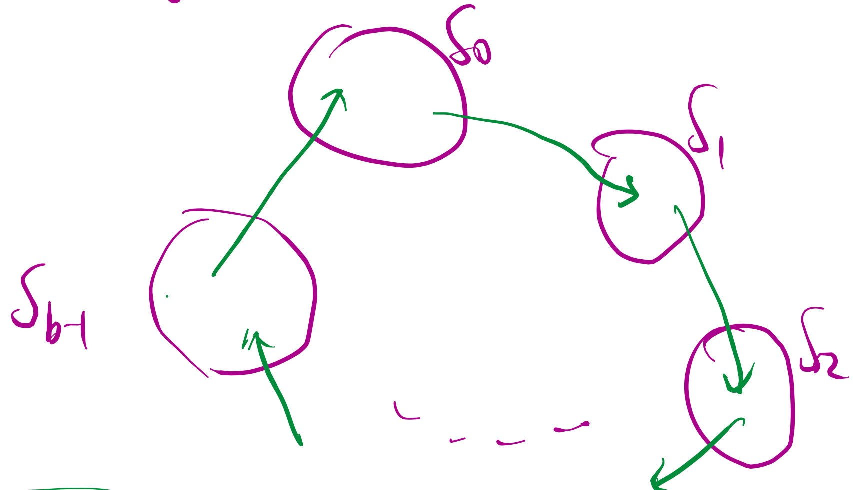
we can show $\mathbb{P}(X_n = j) \to \pi_j$

Periodic MCs.

Thm. If MC is irreducible, period $b \geq 2$, and has stationary distr. π. then

$$\lim_{n \to +\infty} \frac{1}{b}\left(P_{ij}^{(n)} + P_{ij}^{(n+1)} + \cdots + P_{ij}^{(n+b-1)}\right) = \pi_j.$$

Proof idea: "cyclic decomposition". (structures of period-b MC)



$$\pi(S_0) = \pi(S_1) = \cdots = \pi(S_{b-1}) = \frac{1}{b}.$$

Key obs: P^b is a MC on S_i for each i

irreducible, aperiodic, $\exists$ stationary.

Corollary. If P is irreducible, $\exists$ stat. distr. π

then $\displaystyle\lim_{n \to +\infty} \frac{1}{n}\sum_{m=1}^{n} P_{ij}^{(m)} = \pi_j.$

(this also implies uniqueness of π).

Question: $\exists$ stationary msr / distr. ?

When does a stat. msr. become a stat. distr.?

Answer "positive recurrence".

Def. State $i \in S$ is positive recurrent if

$\mathbb{E}_i[T_i] < +\infty$, where T_i is hitting time of i.

(cf. recurrence $\Longleftrightarrow$ $\mathbb{P}_i(T_i < +\infty) = 1$).

For "$\exists$ stat measure".

We'll show that recurrence is sufficient.

Step I: Construction of stat. msr. under recurrence.

Thm. If P irreducible & recurrent, then for fixed $i_0 \in S$ $(\mu_{i_0}(i_0) = 1)$.

$$\forall j \in S, \quad \mu_{i_0}(j) = \sum_{n=0}^{+\infty} \mathbb{P}_{i_0}(X_n = j, T_{i_0} > n)$$

is positive & finite. and μ_{i_0} is a stat. msr.

i.e. $\mu_{i_0} = \mu_{i_0} P$.

Proof. "stationarity".

$$\mu_{i_0}(j) = \mathbb{E}_{i_0}[\#\text{visits to } j \text{ in } \{0, 1, \cdots, T_{i_0} - 1\}]$$

$$(\mu_{i_0} P)(j) = \mathbb{E}_{i_0}\left[\#\text{visits to } j \text{ in } \{1,2,\cdots,T_{i_0}\}\right].$$

Can be verified through compute.

Note that $X_0 = X_{T_{i_0}} = i_0$

So shift does not change # visits.

"finiteness": we know $\mu_{i_0}(i_0) = 1 < +\infty$.

For $j \in S$ $j \neq i_0$, by stationarity

$$1 = \mu_{i_0}(i_0) = \sum_{j \in S} \mu_{i_0}(j) \cdot P^{(n)}_{j,i_0}$$

By irreducible, $\exists n$, s.t. $P^{(n)}_{j,i_0} > 0$ for this specific j.

$$\text{So} \quad \mu_{i_0}(j) \leq \frac{1}{P^{(n)}_{j,i_0}} < +\infty$$

"Positivity". by hit lemma.

$$\mu_{i_0}(j) \geq \mathbb{P}_{i_0}(\text{visit } j \text{ before return to } i_0) > 0.$$

Remark: $\exists$ transient MC that has stat. msr. (e.g. biased RW). not through this construction.

When does it become a stat distn?

$$\sum_{j \in S} \mu_i(j) = \sum_{j \in S} \sum_{n=0}^{+\infty} \mathbb{P}_i(X_n = j, T_i > n)$$

$$= \sum_{n=0}^{+\infty} \sum_{j \in S} \mathbb{P}_i(X_n = j, T_i > n).$$

$$= \sum_{n=0}^{+\infty} \mathbb{P}_i(T_i > n)$$

$$= \mathbb{E}_i[T_i].$$

Immediately, when $\mathbb{E}_i[T_i] < +\infty$ (positive recurrence).

$$\tilde{\mu}(j) = \frac{\mu_i(j)}{\mathbb{E}_i[T_i]} \quad \forall j \in S$$

is a stationary distribution

Structural properties of pos. rec.

Lemma: If $i \leftrightarrow j$, i is pos. rec $\Rightarrow$ j is pos. rec

Corollary: P irreducible, $i \in S$ pos. rec $\Rightarrow$ all state are pos. rec

(we call it pos. rec. MC). $\exists$ stat. distr.

Def. $i \in S$ is null recurrent if $\mathbb{E}_i[T_i] = +\infty$.

We can show null recurrent chain does not have stat. distr.

$$\begin{tikzcd}[column sep=large]
 & \text{Recurrent} \arrow[r, "\text{Yes}", "{\mathbb{E}_i[T_i] < +\infty?}"'] \arrow[dr, "\text{No}"'] & \text{Positive Rec.} \arrow[r, leftrightarrow] & \exists \text{ stat. distr.} \\
\text{Irreducible MCs} \arrow[ur, "\text{Yes}", "{\mathbb{P}_i(T_i < +\infty) = 1?}"'] \arrow[dr, "\text{No}"'] & & \text{Null Rec.} & \exists \text{ stat. msr.} \\
 & \text{Transient} & &
\end{tikzcd}$$

We have shown $\mathbb{E}_i[T_i] < +\infty \Rightarrow \exists$ stat distr.

From stationary distribution to $\mathbb{E}_i[T_i]$?

Thm. Suppose P is irreducible & Recurrent.

$N_n(i) = \sum_{t=1}^{n} 1\{X_t = i\}$. then $\frac{N_n(i)}{n} \xrightarrow{a.s.} \frac{1}{\mathbb{E}_i[T_i]}$

(for any starting state).

Remarks.

$$\begin{cases} \text{Null rec.} & \mathbb{E}_i[T_i] = +\infty \iff \frac{N_n(i)}{n} \xrightarrow{a.s.} 0 \\ \text{Pos. rec.} & \mathbb{E}_i[T_i] < +\infty \iff \frac{N_n(i)}{n} \xrightarrow{a.s.} \frac{1}{\mathbb{E}_i[T_i]}. \end{cases}$$

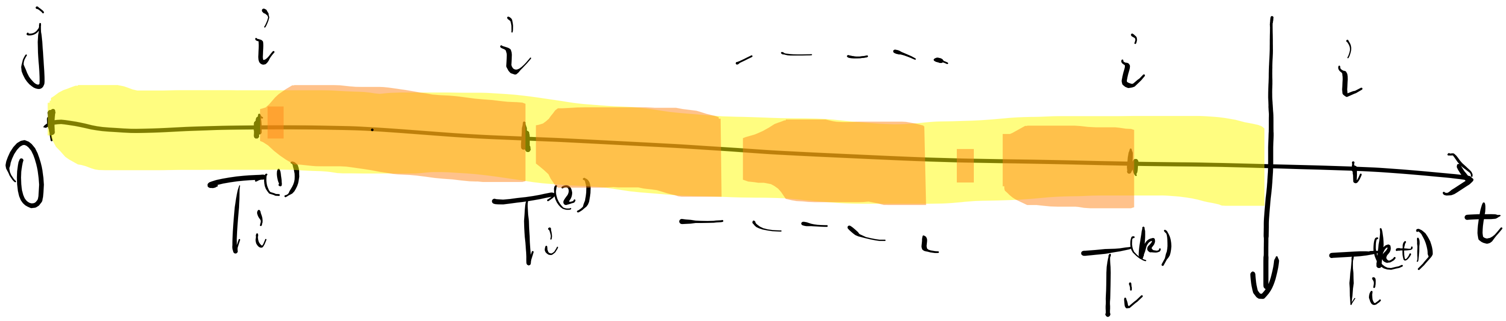
What does $\frac{N_n(i)}{n}$ converge to?

By MC convergence thm, when $\exists$ stat distr. π, $\mathbb{E}\left[\frac{N_n(i)}{n}\right] = \frac{1}{n}\sum_{t=1}^{n} P_{ji}^{(t)} \to \pi_i$

So this implies

- SLLN version of MC convergence
- $\pi_i = \frac{1}{\mathbb{E}_i[T_i]}$.
- $\exists$ stat msr. $\Rightarrow$ pos. rec.

Proof of the thm. $T_i^{(k)} :=$ k-th hitting time of i



$\{X_t : T_i^{(k)} < t \leq T_i^{(k+1)}\}$ are iid blocks.

use SLLN.

$$\frac{T_i^{(N_n(i)+1)}}{N_n(i)} \longrightarrow \mathbb{E}_i[T_i].$$