This Week:

- Markov Chains
  - Markov Property
  - Definition: Markov Chains (MCs)
  - Examples
  - Visualizing Markov Chains
- Long-run behavior of MCs
  - Fixed-point vs steady-state distributions
  - Convergence for Ergodic MCs
  - Transience vs Recurrence

H/W:

L-G: 6.71, 6.72

G: 12.5, 12.6, 12.7

I// Markov Chains:

Recall: for events $\{A_k\}_{k\geq 1}^{n}$

$\{A_k\}_k$ independent

(a) $\iff$ $P(\bigcap_{k=1}^{n} A_k) = \prod_{k=1}^{n} P(A_k)$

In the general dependent case, multiplication theorem applies:

(b) $P(\bigcap_{k=1}^{n} A_k) = P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1 \cap A_2)\cdots P(A_n|\{A_k\}_{k=1}^{n-1})$

We can identify a case between complete independence (a) and complete dependence (b):

$$\boxed{P(\bigcap_{k}^{n} A_k) = P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_2)\cdots P(A_n|A_{n-1})}$$

This is the Markov Property on the events $\{A_k\}_k^n$

This property only allows for dependence between events that are local in the index k. i.e. the immediate future A_{k+1} depends only on the present A_k. Or the present A_k depends only on the immediate past A_{k-1}. This is also called the memoryless property.

We can extend this to random variables $\{X_k\}_{k\geq 1}$

Definition : Markov Chain

Given :

i) an index set : I

ii) a state space (or phase space) Λ

iii) a sequence of random variables

$$X_n : \Omega \longrightarrow \Lambda$$

indexed by $n \in I$

$\{X_n\}_{n \in I}$ is a Markov Chain (MC)

$$\Longleftrightarrow \boxed{P\left(X_{n+1} \,\middle|\, \{X_k\}_{k=1}^{n}\right) = P(X_{n+1} \mid X_n)}$$

If Λ is finite and I is countable then the MC is a Discrete time Finite State MC. (DTFS)

Ex : $S_n = \sum_{k=1}^{n} X_k$, $\{X_k\}$ iid

Claim : The random walk S_n is a MC

Proof :
$$P(S_{n+1} = s_{n+1} \mid S_1 = s_1, \ldots, S_n = s_n) = P\left(S_n + X_{n+1} = s_{n+1} \,\middle|\, \{S_k = s_k\}_{k=1}^{n}\right)$$
$$\overset{\text{sub}}{=} P\left(X_{n+1} + s_n = s_{n+1} \,\middle|\, \{S_k = s_k\}_{k=1}^{n}\right)$$
$$= P\left(X_{n+1} = s_{n+1} - s_n \,\middle|\, \{S_k = s_k\}_{k=1}^{n}\right)$$
$$\overset{\text{ind}}{=} P(X_{n+1} = s_{n+1} - s_n)$$
$$= P(X_{n+1} + s_n = s_{n+1})$$
$$\overset{\text{sub}}{=} P(X_{n+1} + S_n = s_{n+1} \mid S_n = s_n)$$
$$= P(S_{n+1} = s_{n+1} \mid S_n = s_n)$$

$\therefore \{S_n\}$ is a MC // QED

$\Lambda = \mathbb{Z}$

$I = \mathbb{N}$

Ex: LA weather MC: (DTFS)

$\Lambda = \{ r = \text{"rain"},\ s = \text{"sunshine"} \}$

$I = \mathbb{N}$; X_n = LA weather on n^{th} day.

Assuming: Tomorrow's weather depends only on today's weather. i.e. $P(X_{n+1} = t \mid \{X_k\}_{k=1}^{n}) = P(X_{n+1} = t \mid X_n)$

$$P(X_{n+1} = r \mid X_n = s) = p_{s \to r} = p_{sr}$$

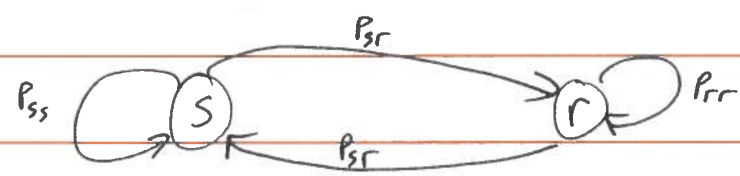
$$P(X_{n+1} = s \mid X_n = r) = p_{r \to s} = p_{rs}$$

$$P(X_{n+1} = s \mid X_n = s) = p_{s \to s} = p_{ss}$$

$$P(X_{n+1} = r \mid X_n = r) = p_{r \to r} = p_{rr}$$

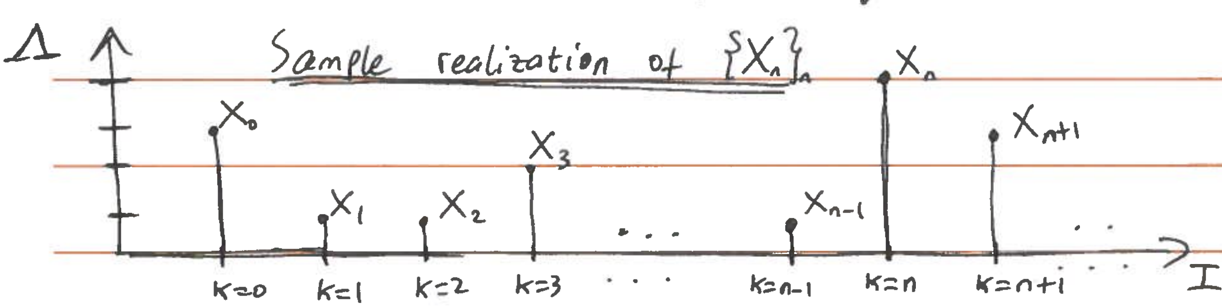
Visualizing Markov Chains:

Transition graph for previous example



Note that MCs are random sequences too:

Sample realization of $\{X_n\}_n$



Transition probability matrix ($\underline{P}$)

e.g. $\underline{P} = \begin{array}{c|cc} & s & r \\ \hline s & p_{ss} & p_{sr} \\ r & p_{rs} & p_{rr} \end{array}$ $\leftarrow$ each row sums to 1 [i.e. Stochastic Matrix]

$\underline{P} = ((p_{ij}))_{i,j}$ where $p_{ij} = P(X_{n+1} = j \mid X_n = i)$ //

If p_{ij} is independent of $n \in I$ $\forall n \in I$

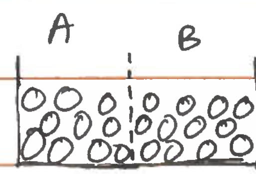
$\Leftrightarrow$ the MC is homogeneous

i.e. $P(X_{n+1} = j \mid X_n = i) \equiv p_{ij}$ $\forall n \in I$ constant

If $p_{ij} = P(X_{n+1} = j \mid X_n = i)$ changes with n

$\Leftrightarrow$ the MC is inhomogeneous.

EX: [Ehrenfest's Urn model]:



$N = |A| + |B|$

@ step n, pick 1 out of N balls in A and B

and transfer it to the other urn @ step $n+1$.

X_n = # of balls in urn A @ step n

$$\Rightarrow \begin{cases} P(X_{n+1} = m+1 \mid X_n = m) = (N-m)/N = 1-p \\ P(X_{n+1} = m-1 \mid X_n = m) = m/N = p \end{cases}$$

i.e. $\underline{P}$ changes at each step n. $\therefore$ Inhomogeneous MC.

- This is a small-scale model of diffusion across a barrier.
- The distribution of balls tends to the Max entropy distribution.

Can also write

$$X_{n+1} = X_n + Z_{n+1}$$

$$Z_{n+1} = \begin{cases} +1 & p = (N-m)/N \\ -1 & p = (m/N \end{cases}$$

Theorem:

Suppose $X_{n+1} = f(X_n, Z_{n+1})$

$\{Z_n\}$ iid $\Rightarrow$ $\{X_n\}_n$ is a homogeneous MC

$\{Z_n\}$ independent $\Rightarrow$ $\{X_n\}_n$ is a MC (maybe inhomogeneous)

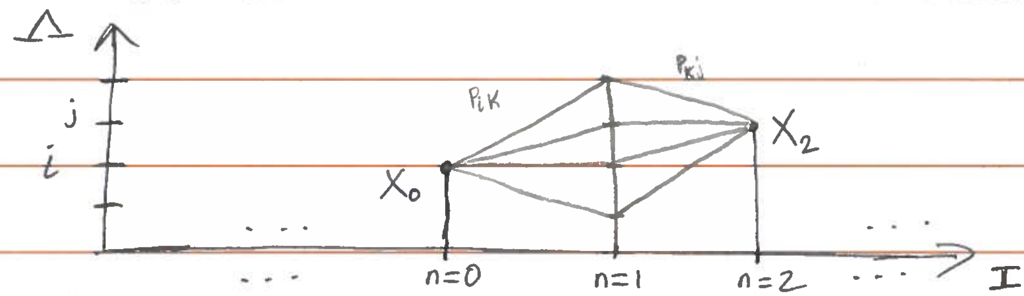
Chapman – Kolmogorov Equations:

Assume Homogeneity and a MC with transition matrix $\underline{P}$.

The probability

$$p_{ij}(n=2) = P(X_2 = j \mid X_0 = i)$$

should be expressible in terms of $\underline{P} = ((p_{ij}(n=1)))_{ij}$ since $\{X_n\}$ is a MC. The idea is to sum up the product probabilities along all paths/evolutions that lead from $X_0 = i$ to $X_2 = j$ in 2 steps. i.e



$$\Rightarrow P(X_2=j|X_0=i) \overset{\text{Marg.}}{=} \sum_{k\in\Lambda} P(X_2=j, X_1=k \mid X_0=i)$$

$$\overset{\text{Bayes}}{=} \sum_{k\in\Lambda} P(X_2=j \mid X_1=k, X_0=i)\cdot P(X_1=k \mid X_0=i)$$

$$\overset{\text{Markov}}{=} \sum_{k\in\Lambda} P(X_2=j \mid X_1=k)\cdot P(X_1=k \mid X_0=i)$$

i.e. $P(X_2=j \mid X_0=i) = \sum_{k\in\Lambda} P_{ik}\cdot P_{kj} = P_{ij}(2)$

Recall that

$$((\underline{A}\cdot\underline{B}))_{i,j} = \sum_k a_{ik}\cdot b_{kj}$$

$$\Rightarrow \quad \underline{P}(2) = \underline{P}(1)\cdot\underline{P}(1) = \underline{P}^2$$

In general:

$$\boxed{\underline{P}(n+m) = \underline{P}^n\cdot\underline{P}^m}$$

[Chapman–Kolmogorov]

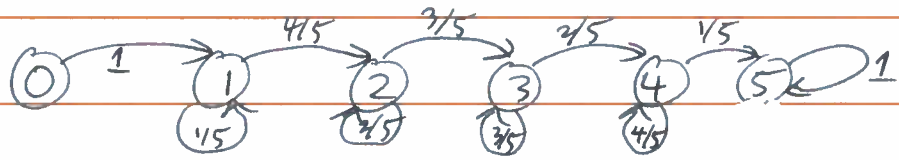
Ex: ∞ # of Balls distributed among 5 urns at random in discrete time.

$X_n \in \{0, 1, \ldots, 5\}$

X_n = # of non-empty urns at step n

Qu: find $P_{03}(3)$?

Ans:



$$P(X=3 \mid X_0=0) = P_{03}(3) = \underset{1}{P_{01}(1)}\cdot P_{13}(2)$$

$$= \sum_k P_{1k}\cdot P_{k3} = \left(\tfrac{1}{5}\right)(0) + \left(\tfrac{4}{5}\right)\left(\tfrac{3}{5}\right) = \tfrac{12}{25}$$

Classification of states and Chains:

State j is *accessible* from state i $[i \to j]$

$\iff \exists n : P_{ij}(n) > 0$

States i and j *communicate* $[i \leftrightarrow j]$

$\iff (i \to j)$ and $(j \to i)$

The communication relation "$\leftrightarrow$" is an *equivalence relation* on the state space Λ. So we can partition Λ into classes of communicating states.

We can also talk about the *period* of state i, $d(i)$:

$$d(i) = \gcd\{n : P_{ii}(n) > 0\}$$

This is, intuitively, the period of all the # of steps it takes to return to state i.

If $d(i) = 1 \iff$ state i is *aperiodic*.

If $d(i) = 1 \ \forall i \in \Lambda \iff$ the whole MC is *aperiodic* [A]

If all states $i, j \in \Lambda$ communicate [i.e $i \leftrightarrow j \ \forall i \neq j$]

$\iff$ the whole MC is *Irreducible* [I].

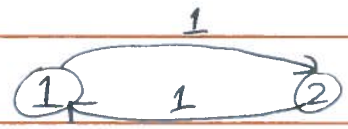
If a finite state MC is both aperiodic and irreducible (AI) $\iff$ MC is *Ergodic*

Example Cases for 2-State Markov Chains:

i/ **[PI]**: <u>P</u>eriodic, <u>I</u>rreducible

$\exists i,\ d(i) > 1$ no "absorbing" states.

e.g

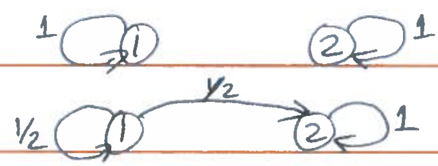


$$\underline{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\quad d(i) = 2$$

ii/ **[AR]**: <u>A</u>periodic, <u>R</u>educible

$\forall i,\ d(i) = 1$ $\exists$ absorbing states

e.g



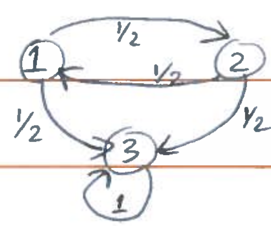
$$\underline{P} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

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

iii/ **[PR]**: <u>P</u>eriodic, <u>R</u>educible

need ≥ 3 states . . .

e.g

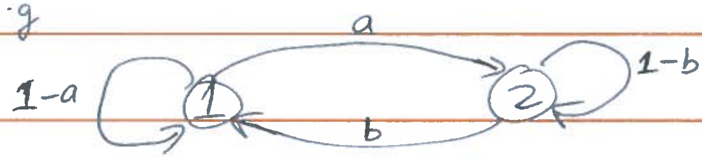


$$\underline{P} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 1 \end{pmatrix}$$

iv **[AI]**: <u>A</u>periodic, <u>I</u>rreducible (aka. <u>Ergodic</u>)

↑ for DTFS MC

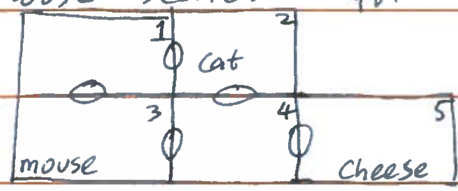
e.g



$$\underline{P} = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

Ex [Cat-mouse-cheese]

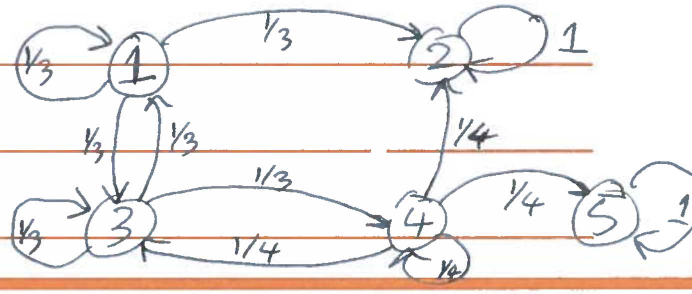
A mouse searches for cheese in a 5-room house laid out as:



The mouse moves at integer time n. It chooses to stay in the same room or use any doors available, all with equal probability. The cat kills the mouse if the mouse enters room 2. In room 5, the mouse eats the cheese and sleeps there.

X_n = mouse location at n^{th} step (stationary)

$\Rightarrow \Lambda = \{1, 2, 3, 4, 5\}$



Ex Wright-Fisher Model for "genotype drift"

Assumptions: - Population (N) constant in each generation

- Diploid genotype i.e. individual genotypes made up of 2 alleles, one from each parent. alleles $\in \{a, A\}$

- Random mating + non-overlapping generations.

Population:

$\underset{1}{\boxed{a|A}}$ $\underset{2}{\boxed{A|A}}$ . . . . $\underset{N}{\boxed{a|a}}$

X_n = # of "a" alleles in population at the n^{th} generation.

$X_n \sim \text{bin}\left(n = 2N,\ p = \left(X_{n-1}/2N\right)\right)$ [approximately]

$\Rightarrow \{X_n\}_n$ is a Markov chain.

II/ Long-run Behavior of Markov Chains

$\underline{P}$ completely specifies the dynamics of a homogeneous MC. But we need an initial (unconditional) pdf vector

$\vec{\pi}_0 = \left(P(X_0=1),\ P(X_0=2),\ \dots,\ P(X_0=N) \right)$

on $\Delta = \{1, \cdots, N\}$ to talk about unconditional probabilities at step n e.g. $P(X_n = i) = ?$

We can get the pdf for X_n, $\vec{\pi}_n$, via:

$$\boxed{\begin{aligned} \vec{\pi}_n &= \vec{\pi}_0 \cdot \underline{P}(n) \\ \vec{\pi}_n &= \vec{\pi}_0 \cdot \underline{P}^n \end{aligned}}$$

We can make an analogy to differential equations (1st order, const coefficient):

$\underline{P}$ is like the diff. eqn., $\vec{\pi}_0$ is the initial condition.

Qu:

How does the distribution of X_n, $\vec{\pi}_n$, behave as $n \to \infty$?

Ans: Any long-term distribution vector $\vec{\pi}_\infty$ will need to have the property:

$$\vec{\pi}_\infty \underline{P} = \vec{\pi}_\infty$$

i/ This is the eigenvector with eigenvalue 1 for $\underline{P}$.

ii/ This is also a fixed-point of the vector transformation

$$P(\vec{\pi}) = \vec{\pi} \cdot \underline{P}$$

$P(\vec{\pi}) = \vec{\pi} \cdot \underline{\underline{P}}$ is a linear ($\therefore$ continuous) function.

$P : S \to S$ where (if $m = |\Delta|$)

$S = \{ \vec{\pi} \in \mathbb{R}^m : \pi(i) \in [0,1], \text{ and } \sum_i \pi(i) = 1 \}$

(also called the simplex on $\mathbb{R}^m$).

S is convex and compact (i.e closed + bounded).

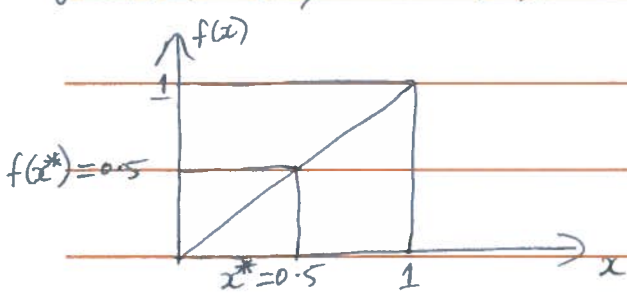
Brouwer's Fixed Point Theorem

If $f : S \to S$ is continuous and $S \in \mathbb{R}^m$ is compact and convex

$\Rightarrow$ $\exists$ (possibly non-unique) $\boxed{x^* \in S \; : \; f(x^*) = x^*}$

[x^* is called a fixed point of f.]

e.g. (a) $f(x) = x$, $S = [0,1]$



(b) $f(x) = x^2 - 2$, $S = [-1, 2]$

$f(x^*) = x^* \Rightarrow x^{*2} - 2 = x^*$

$\Rightarrow x^{*2} - x^* - 2 = 0$

$(x^* - 2)(x^* + 1) = 0$

$x^* \in \{-1, 2\}$ $\Leftarrow$ non-unique fixed points.

$P(\vec{\pi}) = \vec{\pi} \cdot \underline{P}$ is continuous, S convex + compact

$\Rightarrow P(\vec{\pi})$ has a fixed point $\vec{\pi}^*$.

i.e $\boxed{\text{Every}}$ Markov Chain has a stationary (or equilibrium) distribution $\vec{\pi}^*$ such that:

$$\vec{\pi}^* \cdot \underline{P} = \vec{\pi}^*$$

eg : [PI]

$$\begin{tikzcd} 1 \arrow[r, bend left, "1"] & 2 \arrow[l, bend left, "1"] \end{tikzcd} \Rightarrow \underline{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$\vec{\pi}^* = (1/2, 1/2)$

$\vec{\pi}^* \cdot \underline{P} = (1/2, 1/2) \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

$= (1/2, 1/2) = \vec{\pi}^*$

(unique)

[AR]

$$\begin{tikzcd} 1 \arrow[loop left, "1"] & 2 \arrow[loop right, "1"] \end{tikzcd} \Rightarrow \underline{P} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \underline{I}$$

$\Rightarrow \forall \vec{\pi} \in S, \; \vec{\pi} \cdot \underline{P} = \vec{\pi}$

since $(a, b) \cdot \underline{I} = (a, b)$

(non-unique)

But (i) there is no guarantee that $\lim_{n \to \infty} \vec{\pi}_n = \lim_{n \to \infty} (\vec{\pi}_0 \underline{P}^n) = \vec{\pi}^*$ for all MC $\underline{P}$.

e.g. if $\underline{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $\vec{\pi}_0 = (a, b)$ [say $a, b \neq 1/2$]

then $\vec{\pi}_n$ alternates between $\{\vec{\pi}_{ev} = (a, b), \; \vec{\pi}_{odd} = (b, a)\}$ indefinitely. $\therefore \vec{\pi}_n \not\to \pi^* = (1/2, 1/2)$

(ii) If $\vec{\pi}^*$ is not unique then $\vec{\pi}_n$ may converge to different $\vec{\pi}_t^* \in \{\vec{\pi}_t\}$ [set of stationary $\vec{\pi}$] depending on initial condition $\vec{\pi}_0$.

These difficulties disappear in the case of Ergodic MCs...

Theorem:

If a Discrete Time Finite State MC is Ergodic (DTFS)

(i.e the DTFS MC is Aperiodic + Irreducible [AI])

$$\Rightarrow \begin{cases} i/ \ \vec{\pi}^* \text{ is unique} \\ ii/ \ \vec{\pi}_n \longrightarrow \pi^* \quad \forall \vec{\pi}_0 \in S \end{cases}$$

Condition i/ can fail for reducible MCs

" ii/ can fail for periodic MCs

ii/ $\Rightarrow X_n \xrightarrow{d} X^*$

where X^* has pdf $\vec{\pi}^*$ //

Qu: How fast does $\vec{\pi}_n \longrightarrow \pi^*$ in the Ergodic case?

Ans: MC convergence speed is related to the Second Largest Eigenvalue Modulus (SLEM). Recall that the $\vec{\pi}^*$ is the eigenvector with $\lambda^* = 1$. Perron-Frobenius's Theorem $\Rightarrow |\lambda_i| \leq |\lambda^*| = 1$. Denote λ_2 as the eigenvalue with the 2^{nd} largest modulus $|\lambda_2|$ after $|\lambda^*|$.

$\Rightarrow$ MC convergence $\propto 1/|\lambda_2| = 1/\text{SLEM}$.

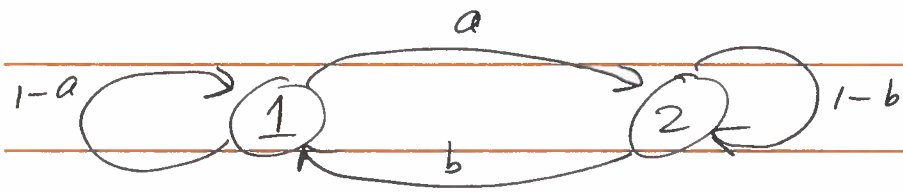
Smaller SLEM $\Rightarrow$ faster convergence.

In the Periodic Irreducible (PI) case, $\{\lambda_i\}_{i=1}^{d}$ are the d^{th} roots of unity

$\Rightarrow |\lambda_2| = 1 \Rightarrow$ no MC convergence.

[d = period of the MC]

Steady State distribution for 2-State MC (AI):



$$\underline{\underline{P}} = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} \qquad a, b \in (0,1)$$

find $\vec{\pi} = (x, y)$: $\vec{\pi}\,\underline{\underline{P}} = \vec{\pi}$

Soln:

$\Rightarrow \quad \vec{\pi}(\underline{\underline{P}} - \underline{\underline{I}}) = 0$

$$\underline{\underline{P}} - \underline{\underline{I}} = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -a & a \\ b & -b \end{pmatrix}$$

$$\Rightarrow \vec{\pi}(\underline{\underline{P}} - \underline{\underline{I}}) = (x \quad y)\begin{pmatrix} -a & a \\ b & -b \end{pmatrix}$$

$$= (-ax + by, \; ax - by) = (0, 0)$$

$\Rightarrow \quad ax = by \quad \ldots \text{i)}$

$\quad x + y = 1 \quad \ldots \text{ii)}$

$\Rightarrow ax = b(1-x) \Rightarrow ax + bx = b \Rightarrow x = \frac{b}{a+b}$

$y = 1 - x = \frac{a}{a+b}$

$$\Rightarrow \boxed{\vec{\pi}^* = \frac{1}{a+b}(b, a)}$$

Transience vs Recurrence:

We can formalize the idea of a transient state vs a recurrent state. A state is transient if there is a non-zero probability of never returning to that state. It is recurrent if we must revisit that state infinitely often.

More formally:

$$i \text{ is transient} \iff \sum_{n=1}^{\infty} P_{ii}^{n} < \infty$$

$$i \text{ is recurrent} \iff \sum_{n=1}^{\infty} P_{ii}^{n} = \infty$$

We can also use first passage time $T_1(i)$ to define this.

$$T_1(i) \triangleq \min\{k \geq 1 : X_k = i\} \quad \begin{bmatrix} \text{first time } k \\ \text{when } X_k = i \end{bmatrix}$$

Now define the probability of returning to i in finite time:

$$f_i = P(T_1(i) < \infty \mid X_0 = i)$$

Then

$$i \text{ is transient} \iff f_i < 1 \quad \begin{bmatrix} \text{non-zero probability} \\ \text{of no return} \end{bmatrix}$$

$$i \text{ is recurrent} \iff f_i = 1 \quad \begin{bmatrix} \text{zero probability} \\ \text{of no return} \end{bmatrix}$$

Theorem:

$$[i \leftrightarrow j] \iff [(i,j \text{ recurrent}) \text{ OR } (i,j \text{ transient})]$$

EX 1-D Random Walk:

$\Lambda = \mathbb{Z}$; $S_n = \sum_{i=1}^{n} X_i$; $X_i \sim \text{Bernoulli}(p=1/2)$ iid

$\{S_n\}_{n\geq 0}$ is a MC [proven earlier]

Qu: $\forall i \in \Lambda$, is i transient or recurrent?

Ans: $P_{ii}^n = ?$

n odd $\Rightarrow P_{ii}^n = 0$ [$\because$ need even # of steps to return to i]

$\Rightarrow \sum_{n=1}^{\infty} P_{ii}^n = \sum_{n=1}^{\infty} P_{ii}^{2n}$

$P_{ii}^{2n} = 0 \Rightarrow$ # successes = # ~~fail~~ ures

$\Rightarrow P_{ii}^{2n} = P(n \text{ successes in } 2n \text{ iid trials})$

$P_{ii}^{2n} = \binom{2n}{n}\cdot\left(\tfrac{1}{2}\right)^n\cdot\left(\tfrac{1}{2}\right)^n = \frac{(2n)!}{n!\,n!}\ \frac{1}{2^{2n}}$

Use Stirling's Approximation: $\boxed{n! \approx n^{n+1/2}\cdot e^{-n}\cdot\sqrt{2\pi}}$

$$\Rightarrow \binom{2n}{n} \approxeq \frac{(2n)^{2n+1/2}\cdot e^{-2n}\cdot\sqrt{2\pi}}{n^{n+1/2}\cdot e^{-n}\cdot 2\pi\cdot n^{n+1/2}\cdot e^{-n}} = \frac{2^{2n+1/2}}{\sqrt{2\pi n}}$$

$$\Rightarrow P_{ii}^{2n} \approxeq \frac{2^{2n+1/2}}{\sqrt{2\pi n}}\cdot\frac{1}{2^{2n}} = \frac{\sqrt{2}}{\sqrt{2\pi n}} = \frac{1}{\sqrt{\pi n}}$$

p-series with $p = 0.5 \leq 1$ $\therefore$ diverges

$$\Rightarrow \sum_{n=1}^{\infty} P_{ii}^{2n} = 1/\sqrt{\pi}\cdot\sum_{n=1}^{\infty} 1/\sqrt{n} = \infty$$

$\Rightarrow \forall i \in \Lambda$, i is recurrent.

[transient if $p \neq 1/2$].

S_n in 2-D $\Rightarrow P_{ii}^{2n} \approxeq 1/\pi n \Rightarrow$ recurrent | S_n in d-Dim ($d \geq 3$) $\Rightarrow \sum_{n=1}^{\infty} P_{ii}^n < \infty \Rightarrow$ Transient

Ex : [Fitting Markov framework]

- Today's weather depends on previous 2-days :

X_n = today's weather

$Y_n = (X_n, X_{n-1})$

$P(X_{n+1} = 1 \mid Y_n = (00)) = 0.2$

$P(X_{n+1} = 1 \mid Y_n = (01)) = 0.4$

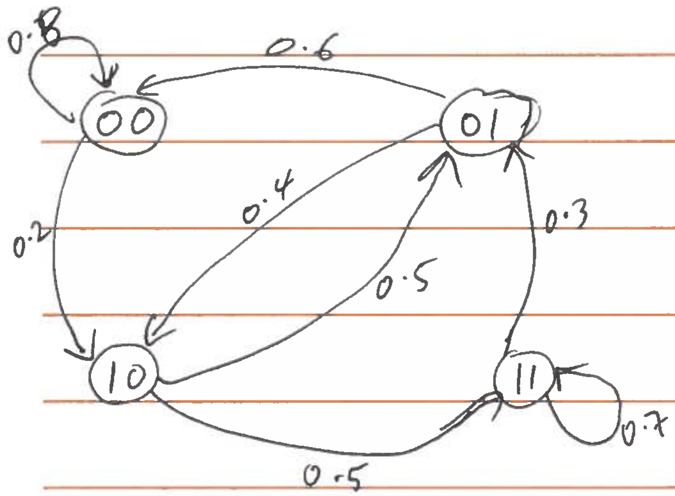
$P(X_{n+1} = 1 \mid Y_n = (10)) = 0.5$

$P(X_{n+1} = 1 \mid Y_n = (11)) = 0.7$

$P(X_{n+1} = 1 \mid X_n = 0) = 0.4 \quad Y_n = (0,1)$

X_n not MC $\because$ $P(X_{n+1} = 1 \mid X_n = 0) = 0.2 \quad Y_n = (0,0)$

Y_n is an MC $\in \Lambda = \{00, 01, 10, 11\}$



$$\underline{P} = \begin{array}{c} \begin{array}{cccc} 00 & 01 & 10 & 11 \end{array} \\ \begin{pmatrix} 0.8 & 0 & 0.2 & 0 \\ 0.6 & 0 & 0.4 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0 & 0.3 & 0 & 0.7 \end{pmatrix} \end{array}$$