EE 512 — Week 7 [05-Oct-2018]

This week:

I. DTFS Markov Chains
- review + examples
- Evolution in time
  - Chapman-Kolmogorov Eqns
- Classification of States
- Stationary and Limiting Distributions
  Global Balance
- Ergodicity
- Reversible Chains and "Detailed Balance"

II. Markov Chain Monte Carlo (MCMC)
- Extending DTFS with kernels
- Metropolis Methods
- Gibbs Sampling
- Markov Chain LLN + CLT

Note:
- HW5 due next week

I. DTFS Markov Chains:

Recall: A Markov Chain $\{X_k\}_{k\geq 1}$ is a random process such that

- given an index set I
- $X_k : \Omega \to \Lambda \quad \forall\, k \in I$
- where Λ is a state (or phase) space.

and

$$\boxed{P(X_{k+1} \mid \{X_j\}_{j=1}^{k}) = P(X_{k+1} \mid X_k)}$$

[Markov or memoryless Property]

$\Rightarrow$ All finite-order distributions of $\{X_k\}_k$ can be expressed in terms of 2^{nd}-order distributions. i.e $\forall\, n \in \mathbb{Z}^+$

$$P\left(\bigcap_{k=0}^{n}\{X_k = x_k\}\right) = \overbrace{P(X_0 = x_0)}^{\text{initial distribution}} \cdot \prod_{k=1}^{n} P(X_k = x_k \mid X_{k-1} = x_{k-1})$$

When $\left.\begin{aligned} I &= \mathbb{N} \\ \Lambda &= \{1, \cdots, m\} \end{aligned}\right\} \Rightarrow$ Discrete Time Finite State (DTFS) MC

$\mathbb{Z}$

The chain is <u>homogeneous</u> if

$$P(X_{k+1} = \lambda_j \mid X_k = \lambda_i) = P(X_1 = \lambda_j \mid X_0 = \lambda_i)$$

$\forall k, i, j$

[i.e. $i \Rightarrow j$ transition probability is constant $\forall$ time]

<u>Canonical Representation</u>

Many homogeneous Markov Chains (HMCs) can be written as a transformation function or recurrence relation driven by "white noise".

$$X_{k+1} = f(X_k, Z_{k+1})$$

where $\{Z_k\}_{k \geq 1}$ is iid ("white")

Why? Observe that

$$P(X_{k+1} = \lambda_{k+1} \mid \{X_i = \lambda_i\}_i^k) = P(f(\lambda_k, Z_{k+1}) \mid \{X_i = \lambda_i\}_i^k)$$
$$= P(f(\lambda_k, Z_{k+1}))$$

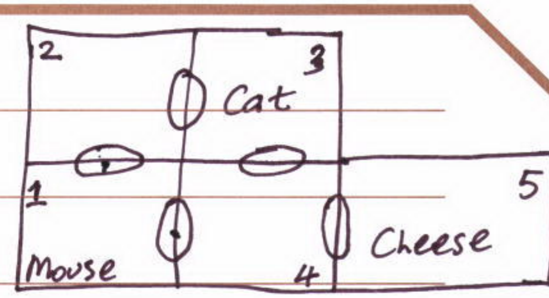
(since Z_{k+1} indep of $\{X_i\}_i^k$)

$\Rightarrow \forall k.$

$$P(X_{k+1} = \lambda_{k+1} \mid X_k = \lambda_k) = P(f(\lambda_k, Z_{k+1}))$$

[not all (H)MCs admit such a representation]

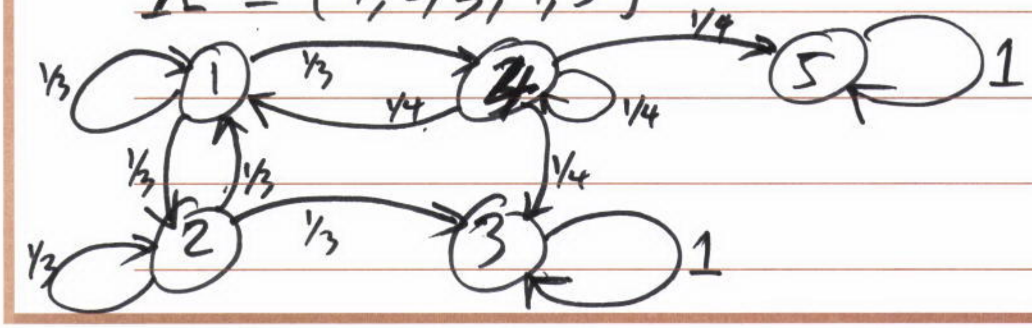
Ex [Cat-Mouse-Cheese]



Mouse seeking cheese in indicated floorplan. Cat in room #3. Mouse doing complete random walk in graph with 2 absorbing states (#3, #5).

X_n = room location of the mouse @ time n

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



Ex. Wright-Fisher model of Genotype Drift.

Assumptions:

- Constant population in each generation (N)
- Diploid genotype $\Rightarrow$ 2 alleles, 1 from each parent
- Random mating + non-overlapping generations

$\{a, A\}$

X_n = # of 'A' allele in the pop$^{\underline{n}}$ @ time n

$n=0$ $\boxed{a \mid a}_1$ $\boxed{A \mid A}_2$, $\boxed{a \mid A}$, ... $\boxed{A \mid A}_N$

$$X_{n+1} \sim b\left(n = 2N,\ p = (X_{n-1})/2N\right)$$

$$|\underline{\underline{P}}| = |\Delta| \times |\Delta|$$

$$\underline{\underline{P}} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 1/3 & 1/3 & 0 & 1/3 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1/4 & 0 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

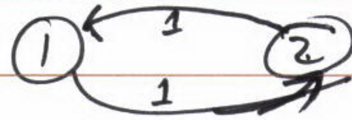
$$\vec{\pi}_0 = (1, 0, 0, 0, 0)$$

$$\vec{\pi}_1 = \vec{\pi}_0 \cdot \underline{\underline{P}}$$

$$\vec{\pi}_n = \left(P(X_n = 1), P(X_n = 2), \ldots, P(X_n = |\Delta|)\right)$$

Examples of HMC Types :

i) [PI]



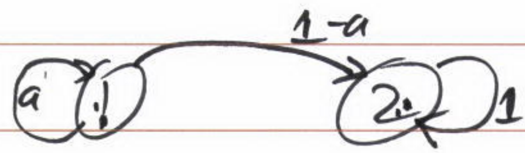
$i \in \{1, 2\}$ $\quad d(i) = 2$

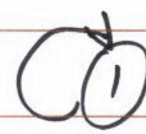
$$\underline{\underline{P}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$1 \longleftrightarrow 2$

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

ii) [AR]



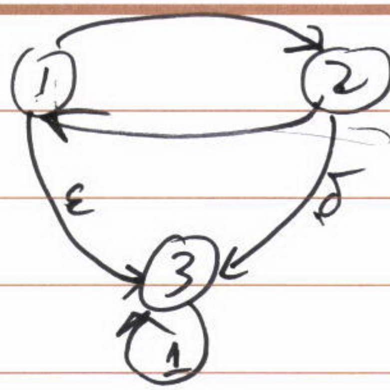


$i \in \{1, 2\}$

$1 \longrightarrow 2$

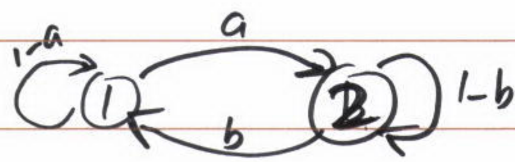
$$\underline{\underline{P}} = \begin{pmatrix} a & 1-a \\ 0 & 1 \end{pmatrix}$$

iii) [PR]



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

iv) [AI]



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

Time-Evolution of HMCs:

The transition matrix

$$\underline{\underline{P}} = ((p_{i,j}))_{i,j}$$

where

$$p_{ij} = P(X_{k+1} = j \mid X_k = i)$$

completely determines the evolution of the HMC in one time step $(k \to k+1)$.

$\underline{\underline{P}}$ also characterizes any $n \geq 1$ step evolution via the Chapman-Kolmogorov equations

i) 1-step Transition:

Assume an initial state distribution:

$$\vec{v}(0) = (P(X_0=1), P(X_0=2), \dots, P(X_0=m))$$
$$[\Delta = \{1, \dots, m\}]$$

QU: What is the distribution after evolving the MC 1-step forward in time? i.e

$$\vec{v}(1) = (P(X_1=1), \dots, P(X_1=m)) = ?$$

Ans: Multiplication Theorem / Conditional Probability / Total Probab

$$P(X_1 = j) = \sum_{i \in \Delta} P(X_0 = i) \cdot \underbrace{P(X_1 = j \mid X_0 = i)}_{P_{i \to j}} \quad \vec{\pi} \cdot \underline{\underline{P}}$$

$$\Rightarrow \boxed{\vec{v}(1) = \vec{v}(0) \cdot \underline{\underline{P}}}$$

Homogeneity $\Rightarrow \boxed{\vec{v}(k+1) = \vec{v}(k) \cdot \underline{\underline{P}}} \quad \forall k$

ii) 2-Step Transition:

Qu: $\vec{v}(2)$ given $\underline{\underline{P}}$ and $\vec{v}(0)$?

Ans: rely on recurrence

$$\vec{v}(2) = \vec{v}(1) \cdot \underline{\underline{P}}$$

$$\vec{v}(1) = \vec{v}(0) \cdot \underline{\underline{P}}$$

$$\Rightarrow \vec{v}(2) = \vec{v}(0) \cdot \underline{\underline{P}} \cdot \underline{\underline{P}}$$

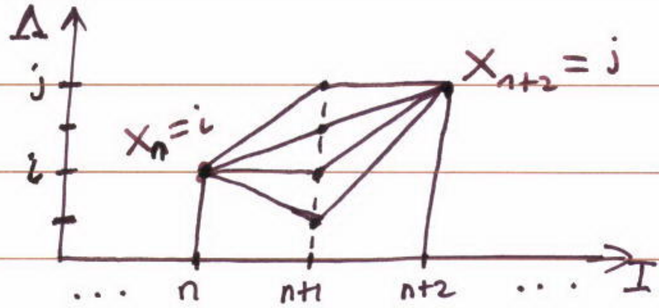
$$\boxed{\vec{v}(2) = \vec{v}(0)\, \underline{\underline{P}}^2}$$

By homogeneity again:

$$\boxed{\vec{v}(k+2) = \vec{v}(k) \cdot \underline{\underline{P}}^2} \quad \forall k$$

This also gives us a way to write the 2-step transition matrix $\underline{\underline{P}}(2)$

$$\underline{\underline{P}}(2) = \left(\left(P(X_{n+2} = j \mid X_n = i)\right)\right)_{i,j}$$



To go from $X_n = i$ to $X_{n+2} = j$ we need to sum over all paths thru $n+1$

$$\text{"} P(X_{n+2} = j) = \sum_k P(X_{n+2} = j, X_{n+1} = k) \text{"}$$

$$P(X_{n+2} = j \mid X_n = i) \overset{\text{marg.}}{=} \sum_{k \in \Delta} P(X_{n+2} = j, X_{n+1} = k \mid X_n = i)$$

$$\overset{\text{Bayes}}{=} \sum_{k \in \Delta} P(X_{n+2} = j \mid X_{n+1} = k, X_n = i) \cdot P(X_{n+1} = k \mid X_n = i)$$

$$= \sum_{k \in \Delta} \overbrace{P(X_{n+2} = j \mid X_{n+1} = k)}^{\underline{\underline{P}}} \cdot \overbrace{P(X_{n+1} = k \mid X_n = i)}^{\underline{\underline{P}}}$$

$$P(X_{n+2} = j \mid X_n = i) = \sum_{k \in \Delta} p_{ik} \cdot p_{kj}$$

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

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

In general: [Chapman - Kolmogorov]

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

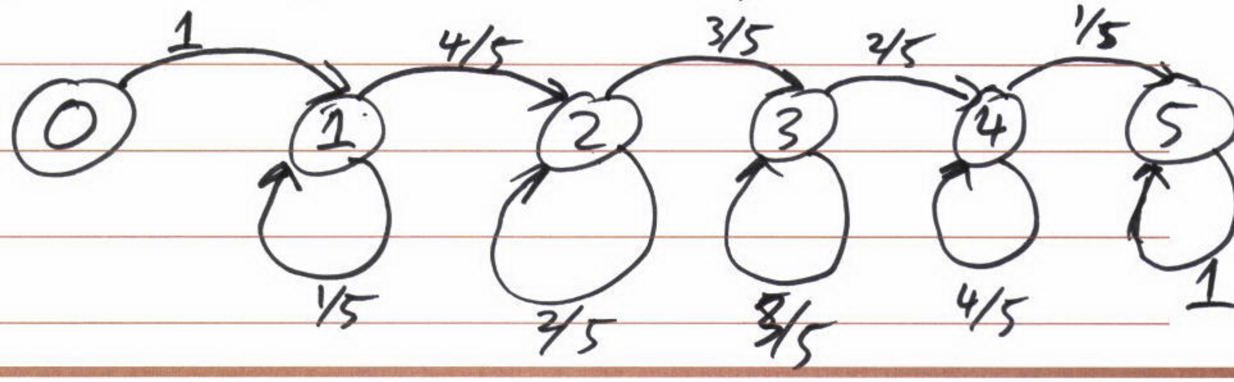
Ex | 1 | 2 | 3 | 4 | 5 |

Unlimited # of balls to distribute

X_n = # of non-empty buckets @ time n

$\Omega = \{0, 1, 2, 3, 4, 5\}$

Qu : $P(X_3 = 3 \mid X_0 = 0)$



$$P(X_3 = 3 \mid X_0 = 0) = P(X_0 = 0) \cdot P(X_1 = 1 \mid X_0 = 0) \cdot P(X_3 = 3 \mid X_1 = 1)$$

$P(X_0 = 0) = 1$

$P(X_1 = 1 \mid X_0 = 0) = 1$

$$P(X_3 = 3 \mid X_1 = 1) = \sum_{X_2 \in \{\cancel{1}, 2\}} P(X_3 = 3 \mid X_2 = k) \cdot P(X_2 = k \mid X_1 = 1)$$

$$= P(X_3 = 3 \mid X_2 = 2) \cdot P(X_2 = 2 \mid X_1 = 1)$$

$$= (3/5)(4/5)$$

$\underline{P}(3) = (\underline{P})^3 \leftarrow P_{03}$

Classification of HMC States:

Two controlling concepts:

Communication + Periodicity (+ Transience/Recurrence)

i) **Communicating States**: $(\Longleftrightarrow)$

$p_{ij}(n)$ is the $(i,j)^{th}$ entry of the n-step transition matrix $\underline{P}^n$

State j is accessible from state i

$\Longleftrightarrow \quad \exists n \in \mathbb{N}^+ : p_{ij}(n) > 0$

$[i \longrightarrow j]$

States (i,j) communicate

$\Longleftrightarrow \quad (i \to j)$ and $(j \to i)$

$[i \longleftrightarrow j]$

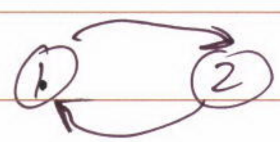
"$\leftrightarrow$" is an equivalence relation on Λ

What is this?

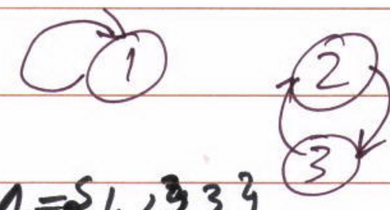
- reflexive
- transitive
- symmetric

$\Rightarrow$ "$\Longleftrightarrow$" partitions Λ into equivalence classes called communicating classes

e.g



vs



$\Lambda = \{1, 2\}$ $\qquad$ $\Lambda = \{1, 2, 3\}$

$\{\{1\}, \{2,3\}\}$

ii/ **State Periodicity**: $(d(i))$

State $i \in \Lambda$ has period $d(i)$:

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

$\rightarrow$ min # of steps required to return to a state / period of self-return path lengths.

$$d(i) = 1 \iff \text{aperiodic state}$$

$$d(i) = 1 \ \forall i \in \Lambda \iff \text{aperiodic MC}.$$

$C \subseteq \Lambda$, a communicating class $\Rightarrow \forall i,j \in C$, $d(i) = d(j)$

Definition:

An **Ergodic** Markov Chain is a MC such that:

i/ All states communicate [Irreducible]

i.e $\forall i,j \in \Lambda$, $i \leftrightarrow j$

ii/ $\forall i \in \Lambda$, $d(i) = 1$ [Aperiodic]

[Ergodic MC = AI MC]

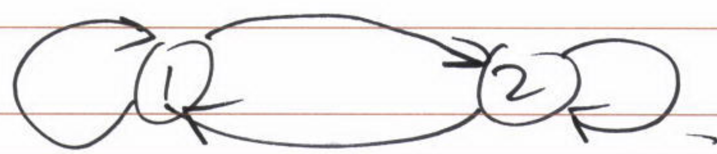
Ergodicity for an RP

Time average $=$ Ensemble Average

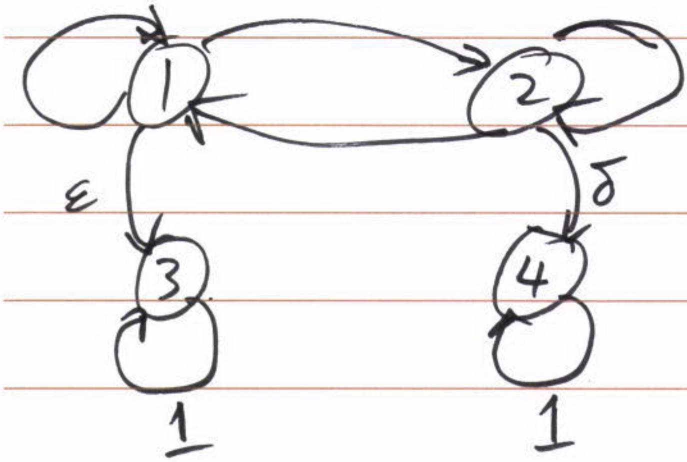
ensemble; fix n get ensemble avg

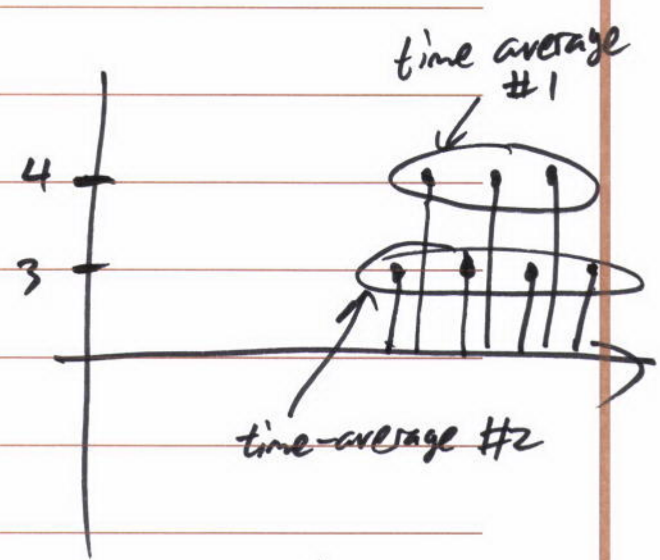
$$X(\omega, n)$$

time; fix ω get time avgs.



↳ $\vec{\pi}^*(n) \quad n \rightarrow \infty$





#1, #2 $\neq$ ensemble average (3, 4)

Stationary and Limiting Distributions

$\underline{P}$ specifies the dynamics of a homogeneous Markov Chain (HMC). An initial distribution vector specifies start point:

$$\vec{\pi}_0 = \left(P(X_0 = 1), \ \ldots, \ P(X_0 = |\Delta|) \right)$$

$(\vec{\pi}_0, \underline{\underline{P}})$ are complete for determining $\vec{\pi}_t$

$$\boxed{\begin{aligned} \vec{\pi}_t &= \vec{\pi}_0 \cdot \underline{\underline{P}}(t) \\ \vec{\pi}_t &= \vec{\pi}_0 \cdot \underline{\underline{P}}^t \end{aligned}} \qquad t \in \mathbb{N}^+$$

Think of: $\left\{ \begin{array}{l} \underline{\underline{P}} \text{ as LCC Diff. Eq} \\ \vec{\pi}_0 \text{ as initial condition} \end{array} \right\}$

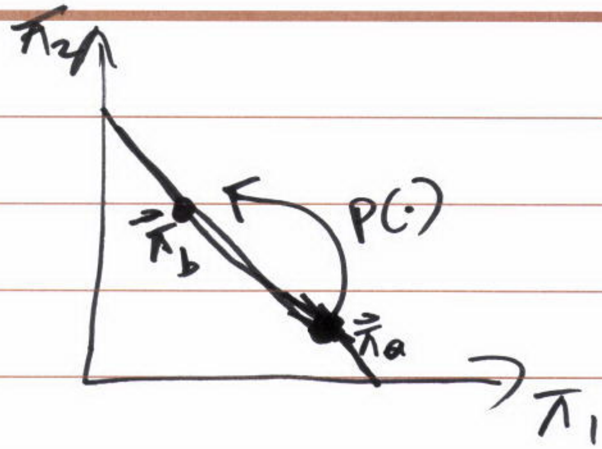
<u>Qu</u>: i) What does $\lim_{t \to \infty} \vec{\pi}_t$ look like?

[limiting distribution]

ii) Does $\exists\ \vec{\pi}^*$ such that

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

[stationary distribution(s)]



$$f : [0,1] \longrightarrow [0,1]$$

$$f(x) = x$$

$$x^* = 0.5 \quad ; \quad f(0.5) = 0.5$$

$$f : [-1,2] \longrightarrow \overset{=[0,3]}{D} \subseteq \mathbb{R}$$

$$f : \ x^2 - 2$$

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

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

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

$$x^* = \{+2, -1\}$$

Function POV:

an HMC's evolution can be viewed as an iterated function composition:

$$P(\cdot) : S^{|\Delta|} \longrightarrow S^{|\Delta|}$$

where:

$$S^{|\Delta|} = \{ \vec{\pi} \in \mathbb{R}_+^{|\Delta|} : \pi(i) \in [0,1], \sum \pi(i) = 1 \}$$

Simplex on $\mathbb{R}^{|\Delta|}$

Convex / compact / closed / Bounded

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

[a linear transformation]

Brouwer's Fixed Point Theorem

If

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

Then

$$\exists \; x^* \in S : f(x^*) = x^*$$

[i.e. x^* is a possibly non-unique fixed pt.]

HMC-specific Corollary:

- $P(\vec{\pi}) = \vec{\pi} \cdot \underline{\underline{P}}$ is a continuous fxn on a convex, compact space.
  - $\Rightarrow$ every HMC has a fixed point $\vec{\pi}^*$ such that

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

[eigenvector of $\lambda = \underline{1}$]

- However, $\vec{\pi}^*$ may not be unique. And... $\lim_{t \to \infty} \vec{\pi}_t$ and $\vec{\pi}^*$ may <u>not</u> be the same.

Perron-Frobenius Theorem:

If $\underline{\underline{P}}$ is a stochastic matrix then

- $\underline{\underline{P}}$ has an eigenvalue $\lambda_1 = 1$
- λ_1 has algebraic and geometric multiplicity equal to the # of communication classes.
- $\Rightarrow$ if $\underline{\underline{P}}$ is aperiodic + irreducible λ_1 is a unique stationary point.
- $\left| \underline{\underline{P}}(t) - \underline{\underline{P}}^{\infty} \right| \propto |\lambda_2|^t$

$\underline{\underline{P}}^{\infty} = \vec{1}\, \vec{\pi}_1$

2nd Largest Eigenvalue Modulus

Key Points/Theorem for DTFS MCs

If a DTFS HMC is Ergodic
[i.e Aperiodic and Irreducible]

$\Rightarrow$ i) $\vec{\pi}^*$ is unique

ii) $\vec{\pi}_t \longrightarrow \vec{\pi}^*$

Failure Modes:

(i) fails for reducible MCs

(ii) fails for periodic MCs

Also:

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

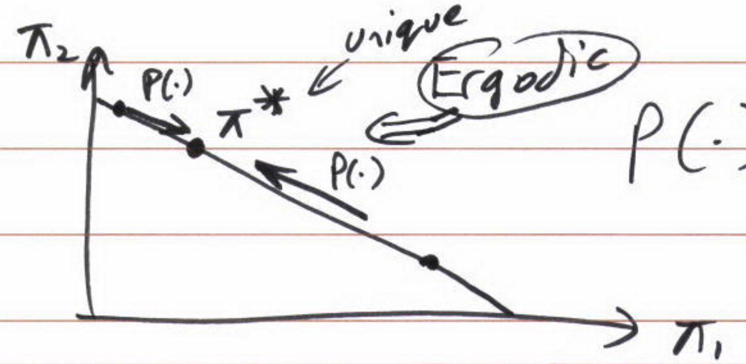
where $X^* \sim \vec{\pi}^*$

So when n is sufficiently large we can begin to explore limit laws like LLN/CLT for Markov Chains.

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

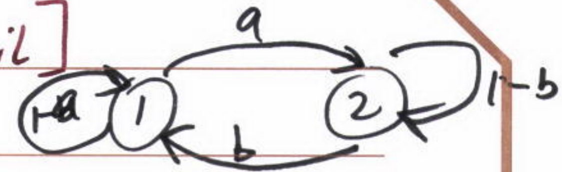
$\cancel{\forall}\ \forall\ \vec{\pi} \in S \qquad \vec{\pi} \cdot \underline{\underline{P}} = \vec{\pi}$

$\Rightarrow \forall\ \vec{\pi} \in S$ is $\vec{\pi}^*$



$P(\cdot)$

Ex [2-state MC in full detail]



Given HMC $\{X_n\}_n$ with

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

[$a,b \in \{0,1\} \Rightarrow$ not Ergodic]

$\underline{\underline{P}}$ AI $\Rightarrow \vec{\pi}_t \to \vec{\pi}^*$

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

Eigenvalues are roots of the characteristic eq, $C(\lambda)$

$$0 = C(\lambda) = \operatorname{Det}(\lambda \underline{\underline{I}} - \underline{\underline{P}}) = \begin{vmatrix} \lambda-(1-a) & -a \\ -b & \lambda-(1-b) \end{vmatrix}$$

$$= (\lambda - 1 + a)(\lambda - 1 + b) - ab$$

$$= (\lambda-1)^2 + a(\lambda-1) + b(\lambda-1)$$

$$\Rightarrow \quad 0 = (\lambda-1)(\lambda - 1 + a + b)$$

$$\Rightarrow \quad \lambda = \begin{cases} 1 & \Leftarrow \text{ determines } \vec{\pi}^* \\ (1-a-b) & \Leftarrow \text{ determines convergence speed} \end{cases}$$

Global balance $\boxed{\vec{\pi}^* \underline{\underline{P}} = \vec{\pi}^*} = (\pi^*(1), \pi^*(2))$

$$\Rightarrow \quad \pi^*(1)(1-a) + \pi^*(2)\, b = \pi^*(1)$$

$$-\left(\pi^*(1)\, a + \pi^*(2)(1-b) = \pi^*(2)\right)$$

$$\pi^*(1)(1-2a) + \pi^*(2)(2b-1) = \pi_1^* - \pi_2^*$$

$$\Rightarrow \qquad 2b\,\pi_2^* = 2a\,\pi_1^*$$

Giving $\begin{cases} \pi_1^* + \pi_2^* = 1 \\ a\pi_1^* = b\pi_2^* \end{cases}$ ← PMF constraint

$$\pi_2^* = \frac{a\pi_1^*}{b}$$

$$\pi_1^* = 1 - \frac{a\pi_1^*}{b}$$

$$\Rightarrow \pi_1^* = \frac{b}{a+b}$$

$$\Rightarrow \pi_2^* = \frac{a}{a+b}$$

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

More generally:

- $$\underline{\underline{P}}^n = \frac{1}{a+b}\left(\begin{pmatrix} b & a \\ b & a \end{pmatrix} + (1-a-b)^n \begin{pmatrix} a & -a \\ -b & b \end{pmatrix}\right)$$

- $$\left|\underline{\underline{P}}^n - P^\infty\right| \propto |1-a-b|^n$$

[geometrically fast convergence]

Reversible Chains and "Detailed Balance":

Suppose $\{X_n\}_n$ is a HMC. Assume X_n is stationary

Define the time-reversed HMC Y_n

$$Y_n = X_{-n}$$

We are interested in conditions under which

$$\{Y_n\}_n \stackrel{d}{=} \{X_n\}_n$$

This requires that forward transitions and backward transitions on $\{X_n\}_n$ are balanced.

i.e probability mass flowing $i \to j$

= mass flowing $j \to i$

$$\Rightarrow \boxed{\pi_i p_{ij} = \pi_j p_{ji}} \quad \forall i, j \in \Delta$$

[Detailed Balance]

A HMC (in) with stationary distribution $\vec{\pi}$ is called <u>reversible</u> if it satisfies detailed balance.

Compare Detailed Balance to

Global Balance:

$$\boxed{\vec{\pi} \cdot \underline{\underline{P}} = \vec{\pi}}$$

Claim: [Detailed Balance Test]

If $\vec{\pi}$ satisfies Detailed Balance for $\underline{\underline{P}}$

$\Rightarrow$ $\vec{\pi}$ satisfies Global Balance for $\underline{\underline{P}}$

i.e $\left[\pi_i P_{ij} = \pi_j P_{ji} \quad \forall i,j \in \Lambda\right] \Rightarrow \left[\vec{\pi} \cdot \underline{\underline{P}} = \vec{\pi}\right]$

Proof: $\sum_{i \in \Lambda} P_{ji} = 1$

$$\Rightarrow \forall j \in \Lambda, \quad \pi_j = \pi_j \cdot \left(\sum_{i \in \Lambda} P_{ji}\right)$$

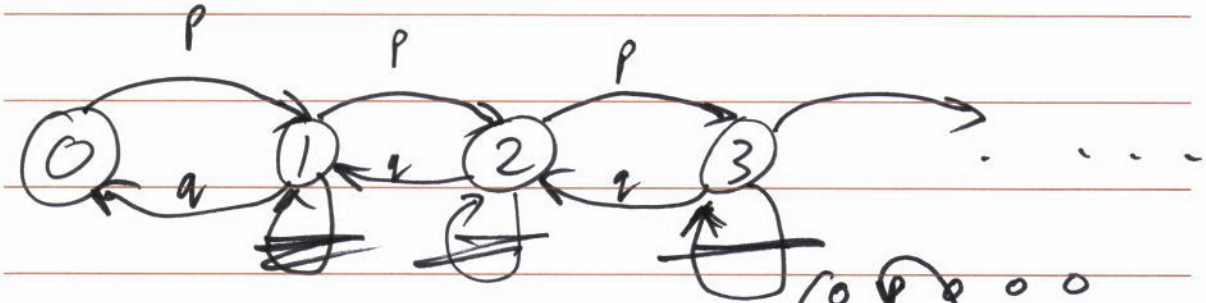
$$= \sum_i \pi_j P_{ji}$$

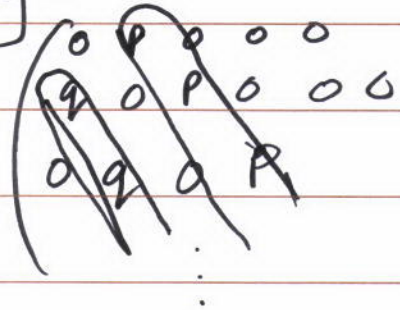
$$\overset{D.B.}{=} \sum_i \pi_i P_{ij} \qquad [\because \text{reversible}]$$

i.e $\forall j \in \Lambda, \ \pi_j = \sum_i \pi_i P_{ij} = \left(\vec{\pi} \cdot \underline{\underline{P}}\right)_j$

$$\Rightarrow \vec{\pi} = \vec{\pi} \cdot \underline{\underline{P}}$$

Birth - Death Process

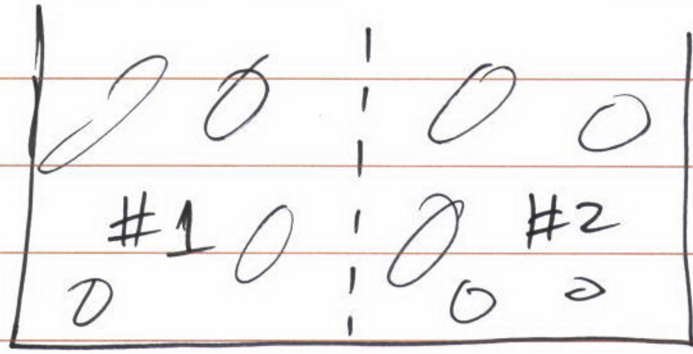




$$\vec{\pi}\,\underline{\underline{P}} = \vec{\pi}$$

$$\vec{\pi}(i) = \phi(i)$$

$$\phi(i) \cdot P_{ij} = \phi(j) \cdot P_{ji} \qquad \forall\, i,j \in 1$$



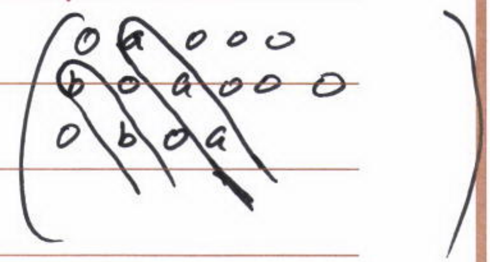
N = # Particles

X_n = # of Particle in #1

E.g : Stationary distribution for Ehrenfest Urn

Recall

$$X_{n+1} = \#\text{ of particles in bin 1}$$



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

$$Z_{n+1} \in \{-1, +1\}$$

$$P(Z_{n+1} = j \mid X_n = i) = \begin{cases} i/N & j = -1 \leftarrow \\ (N-i)/N & j = +1 \leftarrow \end{cases}$$

Test $\pi(i) = \frac{1}{2^N}\binom{N}{i} \leftarrow$ for detailed Balance

$$\pi(i)\, P_{i,i+1} \stackrel{?}{=} \pi(i+1)\, P_{i+1,i} \qquad \boxed{\pi(i)\, P_{i,i-1} \stackrel{?}{=} \pi(i-1)\, P_{i-1,i}}$$

$$\frac{1}{\cancel{2^N}}\binom{N}{i}\frac{N-i}{\cancel{N}} \stackrel{?}{=} \frac{1}{\cancel{2^N}}\binom{N}{i+1}\cdot\frac{i+1}{\cancel{N}}$$

$$\frac{N!}{i!\,(N-i)!}\cdot(N-i) \stackrel{?}{=} \frac{N!}{(i+1)!\,(N-i-1)!}\cdot(i+1)$$

$$\Rightarrow \frac{N!}{\underbrace{(i+1)\cdot i!}\,(N-i)!} \stackrel{?}{=} \frac{N!}{(i+1)\cdot\underbrace{(N-i)\cdot(N-i-1)!}}$$

$$\frac{N!}{(i+1)!\,(N-i)!} \stackrel{\checkmark}{=} \frac{N!}{(i+1)!\,(N-i)!}$$

$\underline{\underline{DTFS}} \leftarrow \text{Ergodic} \Rightarrow [A\ I]$

$DT\underline{CS} \leftarrow \text{Ergodic} = ?$

$\uparrow$ Countably Infinite space

Ergodic $\Leftrightarrow$ States $\underline{\text{Mix}}$

T_j = 1^{st} revisiting Time for state $j \in \Omega$

$= \min\{k \in \mathbb{N} : X_k = j \mid X_0 = j\}$

$P(T_j < \infty \mid X_0 = j) = 1 \ \forall j \in \Omega$

$E[T_j \mid X_0 = j] < \infty \leftarrow$ +ve recurrent

$= \infty \leftarrow$ null recurrent

$[A, I, \text{+ve recurrent}] \Leftrightarrow \text{Ergodic}$

Transience vs Recurrence:

For a state $j \in \Lambda$ of a HMC $\{X_k\}_{k \in I}$

We can examine (1^{st}) return times:

$$T_j = \min\{k \in I : X_k = j\}$$

$[1^{st}$ time $X_k = j]$

$\Rightarrow$ Probability of returning to $j \in \Lambda$

$$\boxed{f_j = P(T_j < \infty \mid X_0 = j)}$$

$\{T_j = \infty\} \Rightarrow$ no return to j

For $j \in \Lambda$

j is a **recurrent** state

$\Longleftrightarrow f_j = 1$

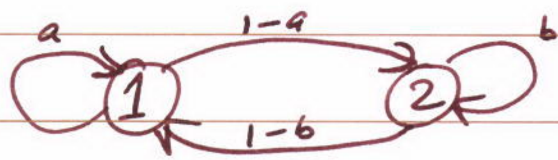
j is a **transient** state

$\Longleftrightarrow f_j < 1$

$$P(\{T_j = \infty\} \mid X_0 = j) = 1 - f_j$$

Also: $\underline{H}MC \Rightarrow P(\cdot \mid X_0 = j) \longrightarrow P(\cdot \mid X_n = j) \; \forall n$

Ex:



$a, b \in (0,1)$

Claim: State 2 is recurrent

proof:

(2) recurrent $\Longleftrightarrow f_2 = 1$

$f_2 \neq 1 \Rightarrow P(T_2 = \infty \mid X_0 = 2) \neq 0$

$\{T_2 = \infty\} \Rightarrow \{X_k = 1, \ \forall k > 0\}$

$$\Rightarrow 1 - f_2 = P\left(\bigcap_{k=1}^{\infty} \{X_k = 1\} \,\middle|\, X_0 = 2\right)$$

$$1 - f_2 = P_{21} \cdot \prod_{k=1}^{\infty} P_{11}^{k-1}$$

$$P_{11} = a \in (0,1) \Rightarrow \lim_{k \to \infty} a^k = 0$$

$$\Rightarrow 1 - f_2 = 0$$

$$\Rightarrow f_2 = 1 \Rightarrow (2) \text{ recurrent}$$

Theorem:

$[i \leftrightarrow j] \Longleftrightarrow [\{i, j\}$ both transient or recurrent$]$

Further recurrence characterizations:

$$j \text{ is transient} \iff \sum_{n=1}^{\infty} p_{jj}^{n} < \infty$$

$$j \text{ is recurrent} \iff \sum_{n=1}^{\infty} p_{jj}^{n} = \infty$$

Recall: $p_{jj}^{n} = P(X_n = j \mid X_0 = j)$
$= E\left[\mathbb{1}_{\{X_n = j\}} \mid X_0 = j \right]$

i.e p_{jj}^{n} is the average occurrence of $\{X_n = j\}$ for a HMC starting @ j

j is Positive Recurrent

$$\iff E[T_j \mid X_0 = j] < \infty$$

j is Null Recurrent

$$\iff E[T_j \mid X_0 = j] = \infty$$

DT MC with possibly Countably Infinite States

Ergodic $\iff$ $\left\{ \begin{array}{l} \text{Aperiodic, Irreducible} \\ \text{Positive Recurrence} \end{array} \right\}$

Limit Theorems for Markov Chains:

Suppose:

- $\{X_n\}_{n\geq 0}$ is an <u>ergodic</u> Markov chain with stationary pdf $\pi(x)$.
- $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ is a measurable function.
- $E_\pi[|f|] = \int |f(x)| \cdot \pi(x)\, dx < \infty$.

Then:

(a) [MC-LLN]

$$\overline{f_n(X)} = \frac{1}{n}\sum_{k=1}^{n} f(X_k) \xrightarrow{a.s} E_\pi[f(X)]$$

"$X_k \sim \pi^*$"

(b) [MC-CLT]

$$\sqrt{n}\left(\overline{f_n(X)} - E_\pi[f(X)]\right) \xrightarrow{d} W \sim N(0, \sigma_f^2)$$

where:

$$\sigma_f^2 = V_\pi[f(X_1)] + \underline{Cov_\pi}[f(X_1), f(X_k)]$$

- ident (in the limit)
- Dependent sample since MC

$\{S_n\}_{n\geq 1}$ is a Markov Chain

$$S_n = \sum_{i=1}^{n} X_n$$

$$S_n = S_{n-1} + X_n$$

$$S_n = f(S_{n-1}, \cancel{X}_n)$$

X_n iid Bernoulli

$X_n \in \{\pm 1\}$

Canonical representation of HMCs:

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

Z_{n+1} iid

$$\left.\begin{array}{l} I = \mathbb{N}\,? \\ \Lambda = \mathbb{Z} \end{array}\right\}.\quad S_n: \text{i/ Ergodic ?}$$

$: \text{ii/ } \underline{P} = ?$

iii/ Long-run distribution

$$\underline{P} = \begin{pmatrix} 0 & p & 0 & & \cdots & \\ 1-p & 0 & p & 0 & 0 & \\ 0 & 1-p & 0 & p & & \\ 0 & 0 & 1-p & 0 & p & \\ & \vdots & & & & \\ & \vdots & & & & \end{pmatrix}$$

Ergodicity 2, 4

Aperiodic ✓

Irreducible ✓

+ve Recurrent ← ?