EE 512 – Week 6 [Post-Midterm I] [28-Sept-2018]

Today:

0/ EM Recap

I/ Markov Chains: Basics

- The Markov Property
- Markov Chains
  - Definitions
  - Examples and Visualizations
- Characterizing Markov Chains

Next week:

- Time-evolution of Markov Chains
- Stationary Behaviors

No Homework this week.

I. Markov Chains

$\{A_k\}_{k\geq 1}^{n}$ $\quad I = \mathbb{N}$

(a) Dependent:

$$\iff P\left(\bigcap_{k=1}^{n} A_k\right) = P(A_1)\cdot P(A_2|A_1)\cdots P\left(A_n \mid \{A_k\}_{k=1}^{n-1}\right)$$

(b) Independent:

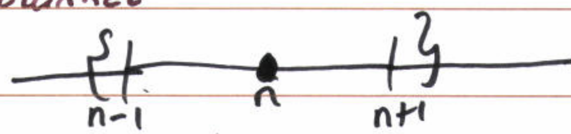
$$\iff P\left(\bigcap_{k=1}^{n} A_k\right) = P(A_1)\cdot P(A_2)\cdots P(A_n)$$

In between (a) and (b), we have:

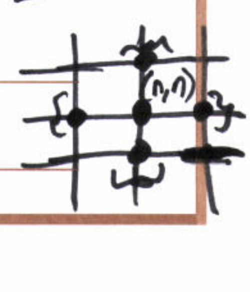
(c) Markov Property:

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

- Models "local" dependence between events with locality defined in terms of the index set.
- General term for local neighborhood of dependence: the "Markov blanket"



$I = \mathbb{N}\times\mathbb{N}$



Extend this to random variables.

Definition: Markov Chain

Given

i) an index set : I

ii) a state space (or phase space) : Λ

iii) a sequence of random variables (indexed)

$$X_k : \Omega \longrightarrow \Lambda$$

$\{X_k\}_{k \in I}$ is a Markov Chain (MC)

iv) $\Longleftrightarrow$ $\boxed{P(X_{k+1} \mid \{X_i\}_{i=1}^{k}) = P(X_{k+1} \mid X_k)}$

Typical types of MCs : $(|\underset{\text{state}}{\Lambda}|, |\underset{\text{time}}{I}|)$

- DTFS : Discrete Time Finite State
- MCMC $\rightarrow$ DTIS : Discrete Time Infinite State
- CTIS : Continuous Time Infinite State

Ex: LA (MC Assumption) weather: $I = \mathbb{N}$, $\Lambda = \{\underset{\text{rain}}{r}, \underset{\text{sun}}{s}\}$, X_k

$$P(X_{k+1} = r \mid X_k = s) = P_{s \to r}$$

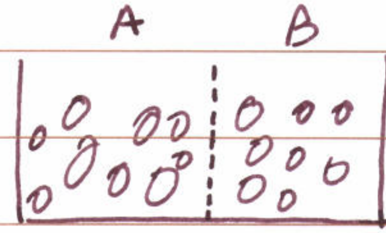
$$P(X_{k+1} = s \mid X_k = r) = P_{r \to s}$$

$$P(X_{k+1} = s \mid X_k = s) = P_{s \to s}$$

$$P(X_{k+1} = r \mid X_k = r) = P_{r \to r}$$

Ex. Ehrenfest's Model

$N = |A| + |B|$



- @ step k, pick 1 out of N balls in $A \cup B$
  ↳ transfer it to other bin @ step $k+1$

X_k = # of balls in urn A @ step k

$$\begin{cases} P(X_{k+1} = m+1 \mid X_k = m) = \dfrac{N-m}{N} & \text{(picked from B urn)} \\ P(X_{k+1} = m-1 \mid X_k = m) = \dfrac{m}{N} & \text{(picked from A urn)} \end{cases}$$

- Observe that @ step t, the probability of transitioning between states a and b ($b = a \pm 1$) changes with the time t.
- Ehrenfest's Urn = small-scale atomistic model of diffusion across a barrier

The relative transition probabilities $\left(\frac{N-m}{N}, \frac{m}{N}\right)$ reflect "osmotic" pressure.

- Converges to the max entropy distribution (uniform)

In General, if we have a process $\{X_k\}_{k\geq 1}$ such that

$$X_{k+1} = g(X_k, Z_k)$$

where Z_k is an independent innovation

$\Rightarrow \{X_k\}$ is an MC.

$\{Z_k\}_k \rightarrow$ iid $\Rightarrow$ homogeneous MC

$\rightarrow$ just indep $\Rightarrow$ inhomogeneous MC.

Markov Chain Representations

The Markov assumption is fundamentally a dimension reduction assumption

$$P\left(\bigcap_k A_k\right)$$

Dependence	Markov
$P(A_1)\cdot P(A_2 \mid A_1)\cdots P(A_n \mid \{A_j\}^{n-1})$	$P(A_1)\cdot P(A_2 \mid A_1)\cdots P(A_n \mid A_{n-1})$
$\uparrow$	$\uparrow$
keeping track of $\lvert\Lambda\rvert^n$ transition probabilities	keeping track of $\lvert\Lambda\rvert^2$ transition probabilities

---

For DTFS MCs, representation:

i) <u>The Transition Matrix $\underline{P}$</u>:

$$\underline{P} \in M_{|\Lambda|\times|\Lambda|}([0,1]), \text{ a } \underbrace{\text{"stochastic matrix"}}_{\text{i.e. every row is a pmf vector}} \qquad |\Lambda| \times |\Lambda|$$

eg $\Lambda = \{r, s\}$

$$\underline{P} = \begin{array}{c c} & \begin{array}{cc} r & s \end{array} \\ \begin{array}{c} r \\ s \end{array} & \begin{pmatrix} P_{r\to r} & P_{r\to s} \\ P_{s\to r} & P_{s\to s} \end{pmatrix} \end{array} ; \quad \underline{P} = ((p_{ij}))_{i,j\in\Lambda}$$

$$P_{r\to s} = P(X_{k+1} = s \mid X_k = r)$$

[note assumption of "homogeneity"]

Homogeneous MC

$\Leftrightarrow P(X_{k+1}=r \mid X_k=s) = P(X_{t+1}=r \mid X_t=s)$

$\forall\, k,t \in I, \quad \forall\, r,s \in \Lambda$

[i.e. transition probability independent of time]

$\Leftrightarrow \underline{\underline{P}}$ is constant in time

Inhomogeneous MC

$\Leftrightarrow$ the transitions are time-dependent

$\underline{\underline{P}}(t) = ((p_{ij}(t)))_{i,j \in \Lambda}$

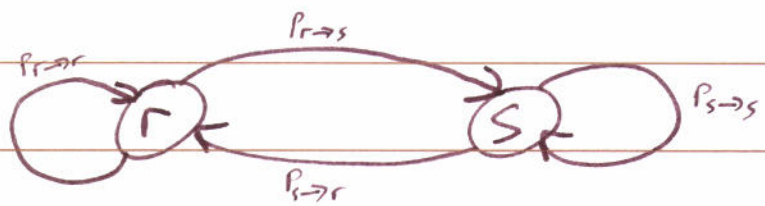
$\underline{\underline{P}}$ is non-constant over time

e.g. Ehrenfest's Urn

ii) Transition Graph :

States as nodes

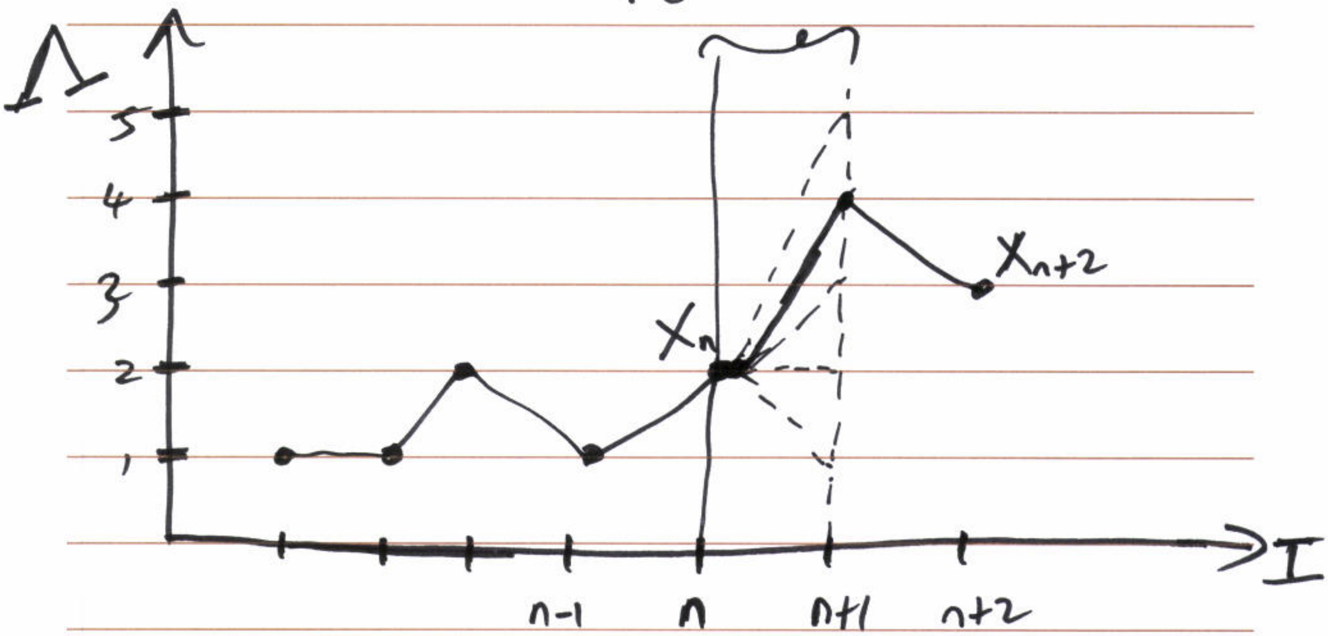
Transition probabilities as edges.



$\underline{\underline{P}}$ is the (directed) adjacency matrix of the the transition graph.

Sample Paths:

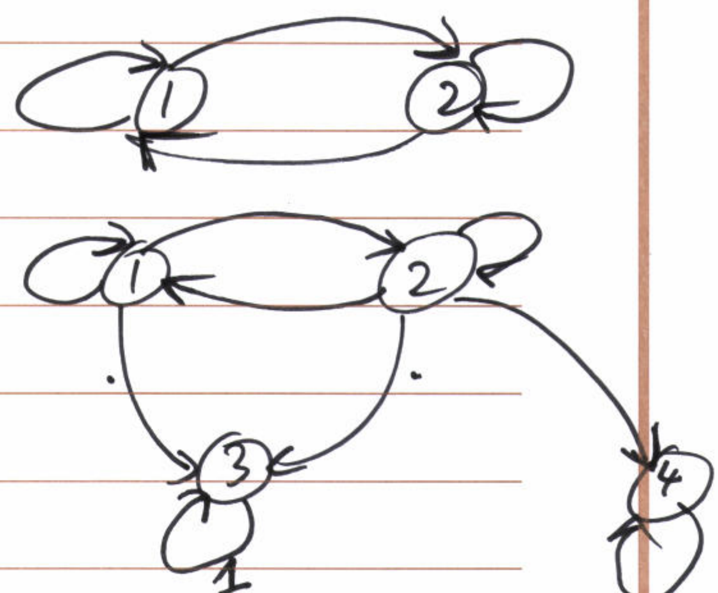
$$P(X_{n+1}=j \mid X_n = i) = P_{ij}$$



$$\{X_k\}_{k \geq n_0}^{\infty} \sim ???$$

initial Distribution

$$\vec{V}(0) = \begin{pmatrix} P(X_0 = \lambda_1) \\ P(X_0 = \lambda_2) \\ \vdots \\ P(X_0 = \lambda_3) \end{pmatrix}$$



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

$$\vec{v}_0 \longrightarrow \vec{v}_1 \longrightarrow \vec{v}_2 \rightarrow \cdots \longrightarrow \vec{v}_\infty$$