11/29

Markov Chain

Markov Chain — J. R. Norris

Probability Space $(\Omega, \mathcal{F}, P)$

$\Omega \rightarrow$ Sample space

$\mathcal{F} \rightarrow$ Events

$P \rightarrow$ Probability.

Discrete time Markov chains

* I : a countable set

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

Each $i \in I$ is called a state

where I is called a state-space

* ~~$\lambda_i = P($~~

Define a random variable X with values in I.

* Define: $\lambda_i = P(X = i)$

THEN: λ defines a distribution of X.

X	Prob.
1	λ_1
2	λ_2
3	λ_3
$\vdots$	$\vdots$

* We think of X as <u>modelling</u> a random state (i) with probability λ_i.

* Consider a matrix

$$P = (p_{ij} : i, j \in I)$$

$I = \{1, 2, 3, \dots, n\}$

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & & & \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

It is called a <u>stochastic</u> (random) if every row $\boxed{(p_{ij} : j \in I)}$ is a distribution (p.m.f.)

i.e. for our example:

$$\begin{aligned} p_{11} + p_{12} + \cdots + p_{1n} &= 1 \\ p_{21} + p_{22} + \cdots + p_{2n} &= 1 \\ &\vdots \\ p_{n1} + p_{n2} + \cdots + p_{nn} &= 1 \end{aligned}$$

* <u>THEN</u> there is a one-to-one correspondence between the <u>stochastic matric</u> and the "Markov diagrams".

Markov diagram

fig 1

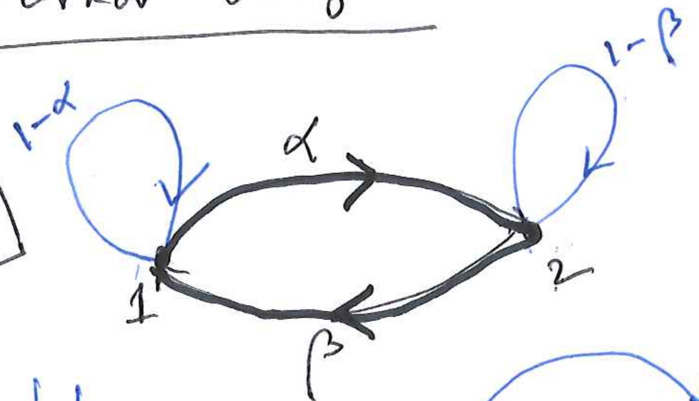
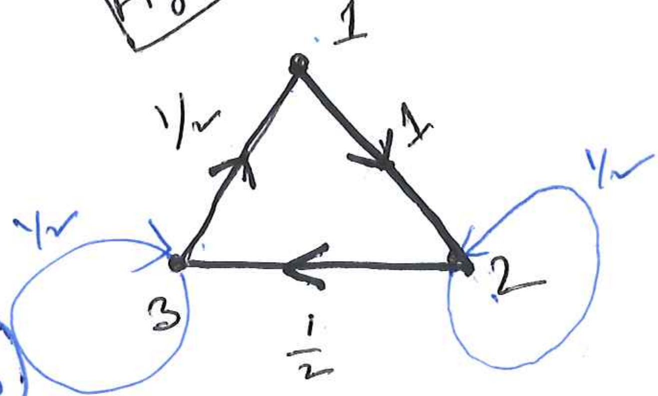


fig 2



$I = \{1, 2\}$

We assign this way

fig 1

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

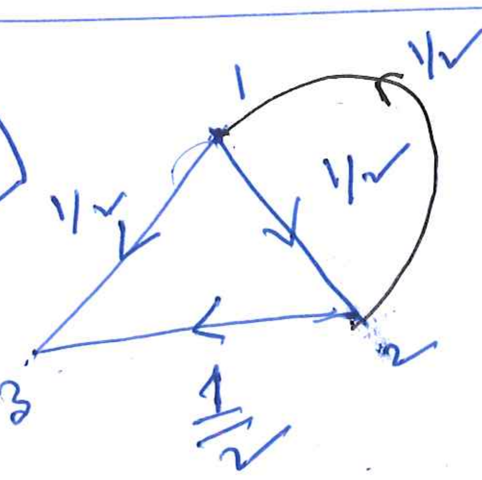
fig 2

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

$$P = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix}.$$

---

fig 3



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

Markov chain

We say that $(X_n)_{n \geq 0}$ is a Markov chain with initial distribution λ and transition matrix P if

1. X_0 has distribution λ.
2. For $n \geq 0$, conditioned on $X_n = i$, X_{n+1} has distribution $(p_{ij} : j \in I)$ and is independent of $X_0, \ldots, X_{n-1}$

X_0	Prob
1	$\lambda_1 = P(X_0 = 1)$
2	$\lambda_2 = P(X_0 = 2)$
3	$\lambda_3 = P(X_0 = 3)$
$\vdots$	$\vdots$

X_1	Prob given X_0
1	P_{1j}, $j = 1, 2, \ldots$
2	P_{2j}, $j = 1, 2, \ldots$
3	P_{3j}, $j = 1, 2, \ldots$
$\vdots$	

$\lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots)$

↑ initial distribution.

$$P = \begin{pmatrix} P_{11} & P_{12} & \cdots \\ P_{21} & P_{22} & \cdots \\ \vdots & & \end{pmatrix}$$

Transition prob. matrix.

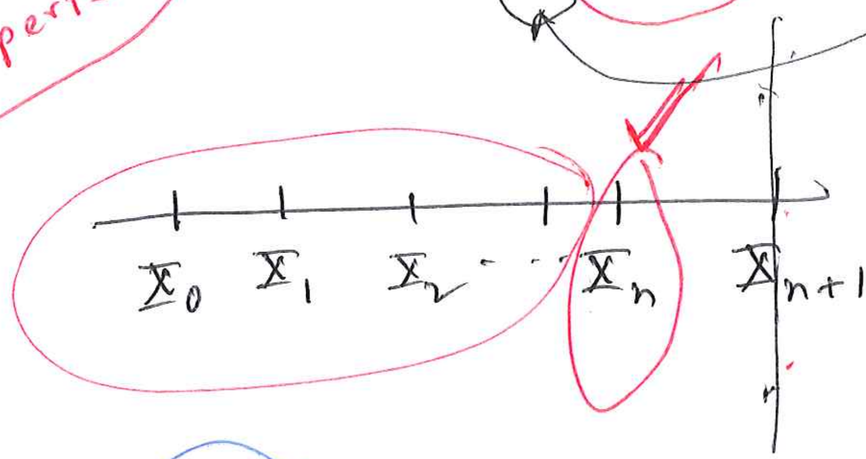
~~X_2 | Prob. given X_1~~
~~1 | P_{1j}, $j = 1, 2, \dots$~~
~~2~~
~~3~~
~~⋮~~

More explicitly, ~~th~~ ~~co~~ these conditions (① and ②) claim:

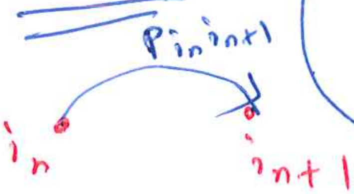
① $P(X_0 = i_0) = \lambda_{i_0}$

② $P(X_{n+1} = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) = P_{i_n i_{n+1}}$

Memory-less property



Note:



$P_{i_n i_{n+1}}$ probability ~~from~~ for going from state $= i_n$ to state $= i_{n+1}$

Theorem: A discrete-time random process

X_n, $0 \le n \le N$

is Markov (λ, P)

if and only if

for all states $i \in I$

$I = \{ i_0, i_1, \ldots, i_N \}$

$$P(X_0 = i_0, X_1 = i_1, \ldots, X_N = i_N)$$
$$= \lambda_{i_0} P_{i_0 i_1} P_{i_1 i_2} P_{i_2 i_3} \cdots P_{i_{N-1} i_N}$$

Aside

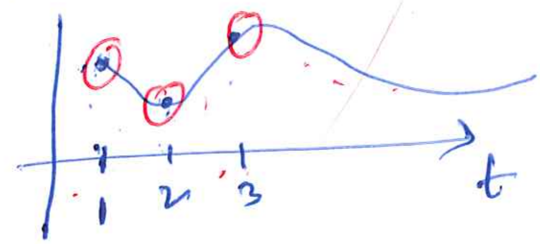
Random process.

Random variable indexed by time

X_t, $t = 1, 2, \ldots$

$X_1, X_2, X_3, \ldots$

r.v. r.v. r.v. ...



Random (Stochastic) Process

Discrete (X_n, n countable)

($X_1, X_2, \ldots$)

Random Walk

Continuous (X_t, $t \ge 0$)

Brownian motion

$X_t \sim N(0, t)$

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

$S_1 = X_1$

$S_2 = X_1 + X_2$

...

Theorem (Markov property)

Let $(X_n)_{n \geq 0}$ be Markov(λ, P).

Then, conditional on $X_m = i$ (i.e., when X_m is "known")

$(X_{m+n})_{n \geq 0}$ is ~~Markov~~ Markov(δ_i, P) and is independent of the random variables $X_0, \ldots, X_m$.

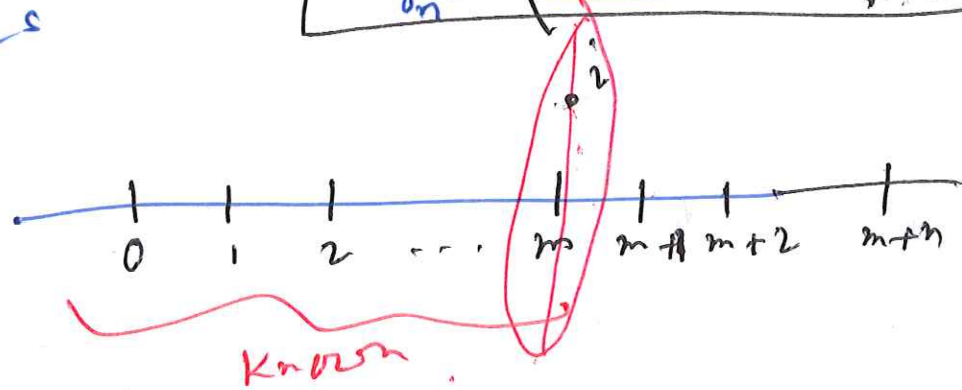
Notation:

$$\delta_i = (\delta_{ij} : j \in I)$$

where

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$$

$$\delta_{ij} = \begin{array}{c c} \delta_1 & \\ \delta_2 & \\ \delta_3 & \\ \delta_n & \end{array} \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & \vdots & & \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$



Conclusion: Markov chains have no memory!

~~Theorem 7~~

Notation:

* λ will be considered to be a row vector whose components are indexed by I.

$$\lambda = \Big(\underset{P(X_0 = i_0)}{\underset{\downarrow}{\lambda_{i_0}}} \quad \underset{P(X_0 = i_1)}{\underset{\downarrow}{\lambda_{i_1}}} \quad \lambda_{i_2} \cdots \quad \underset{P(X_0 = i_N)}{\underset{\downarrow}{\lambda_{i_N}}} \Big)$$

$$\underset{1\times N}{\underset{\downarrow}{\lambda}} \, \underset{N\times N}{\underset{\downarrow}{P}} = \Big(\cdot \; \cdots \; \underset{j^{th}\text{ entry}}{\underset{\uparrow}{-}} \; - \; - \; - \Big)_{1\times N}$$

$$(\lambda P)_j = \sum_{i\in I} \lambda_i \, P_{ij}$$

$$P^2 \longrightarrow N\times N \text{ matrix}$$

$$(P^2)_{ik} = \sum_{j\in I} P_{ij} \, P_{jk}$$

We set $P^0 = I$ $\leftarrow$ identity matrix //

Notation:

$$P_{ij}^{(n)} = (P^n)_{ij}$$

(i,j) entry of P^n.

Notation:

$$P_i(A) = P(A \mid X_0 = i)$$

(Note $P(X_0 = i) = \lambda_i$)

Theorem: Let $(X_n)_{n \geq 0}$ be Markov(λ, P).

THEN for all $n, m \geq 0$.

① $P(X_n = j) = \underbrace{(\lambda P^n)_j}$

(j^{th} component of the vector λP^n)

② $$\boxed{P_i(X_n = j) = P(X_{n+m} = j \mid X_m = i) = P_{ij}^{(n)}}$$