Cpts 515.   11-2-2020

Today: Markov Chain

1. One of the most useful math. tool in Computer Science

2. For us, we only need focus on finite-state & Discrete time Markov Chain.

3. Always, for us, Markov Chain is a graph.

Markovian Property:

Tomorrow <u>only</u> depends on today, <u>NOT</u> yesterday or any past days.

1

Markov chain
(finite state &
Discrete Time)

$\sim$

Finite Automata
(finite state, no
probability).

||

Finite Automata – input symbols
+ probability

---

2 In entire mathematics, we don't have many areas where we can program. However, we do have some discrete models/math we can code for:

linear algebra
markov chain.

time is discrete

$t = 0, 1, 2, 3, \ldots$

start day — no end day.

Each day i, we associate a random variable X_i.

Now, we have a random process (stochastic process)

$$X_0, X_1, X_2, \ldots$$

In theory, dependency between the X_i's can be very complex. However, herein, we only look at a very special class of random processes that are Markovian.

Convention: we assume each X_n is a r.v. over k things. For instance, we have two things

$\mathcal{S}$, $\mathcal{D}$.

We use $i, j \in \{\mathcal{S}, \mathcal{D}\}$.

"Markovian" says:

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = \dots, X_{n-2} = \dots\dots)$$

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

$$= P_{ij}.$$

not dep. on n. Such a process is called homogeneous.

So, the markov chain

$$X_0, X_1, \ldots \ldots$$

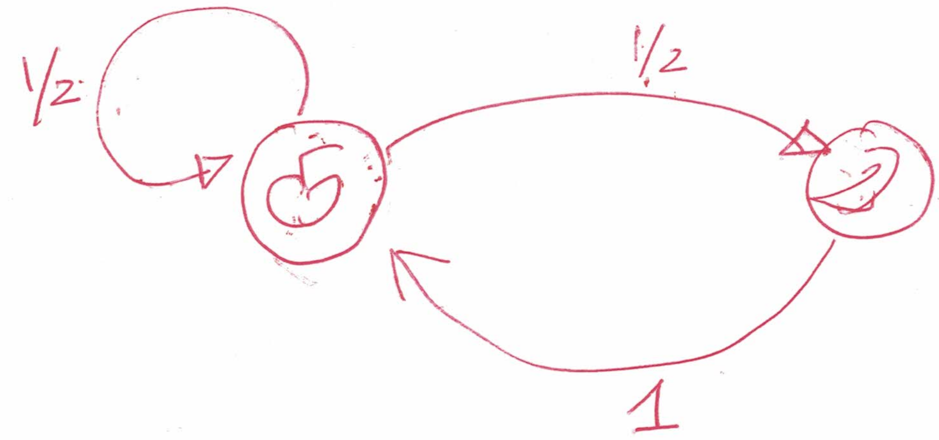
is given by a matrix $[P_{ij}] = P$

is of $k \times k$ matrix
and each X_n takes
values in k things.

Each P_{ij} means the transition probability from "state" i to "state" j.

a conditional
probability.

For us, we draw a Markov chain as a graph:



We translate the graph into a matrix: (☉ = 1, ☾ = 2)

$$P = \begin{bmatrix} P_{11} = \frac{1}{2} & P_{12} = \frac{1}{2} \\ P_{21} = 1 & P_{22} = 0 \end{bmatrix}$$

sum of each row = 1

Suppose that the Markov Chain

$$X_0, X_1, X_2, \ldots, X_n, \ldots$$

has the initial distribution (i.e., the distribution of X_0) as π_0. For instance, π_0 can be

$$\text{Prob}(X_0 = ☁) = 0.3$$

$$\text{Prob}(X_0 = ☾) = 0.7.$$

Then, what is the distribution π_1 of X_1? You can check:

$$\pi_1 = \pi_0 \cdot P.$$

$☁ = 1$, $☾ = 2$

Checking :

$\text{prob}(X_1 = ☁) =$

$\text{prob}(X_1 = ☁ \mid X_0 = ☁) \cdot \text{prob}(X_0 = ☁) +$

$\text{prob}(X_1 = ☁ \mid X_0 = ☾) \cdot \text{prob}(X_0 = ☾)$

$= p_{11} \cdot \pi_0[1] + p_{21} \cdot \pi_0[2]$

$= \pi_0 \cdot \begin{pmatrix} p_{11} \\ p_{21} \end{pmatrix}.$

Similarly $\text{Prob}(X_1 = ☾) = \pi_0 \cdot \begin{pmatrix} p_{12} \\ p_{22} \end{pmatrix}.$

Then, $\pi_1 = \pi_0 \cdot P.$

Continue with this process, we have

$$\begin{aligned}
\pi_2 &= \pi_1 \cdot P \\
&= \pi_0 \cdot P \cdot P \\
&= \pi_0 \cdot P^2 \\
&\vdots \\
\pi_n &= \pi_0 \cdot P^n .
\end{aligned}$$

Two <u>most</u> important questions on Markov Chain

(1). Do we have have a π^* s.t.

when $\pi_0 = \pi^*$ and $\pi_n = \pi^*$ for all n?

(Invariant).

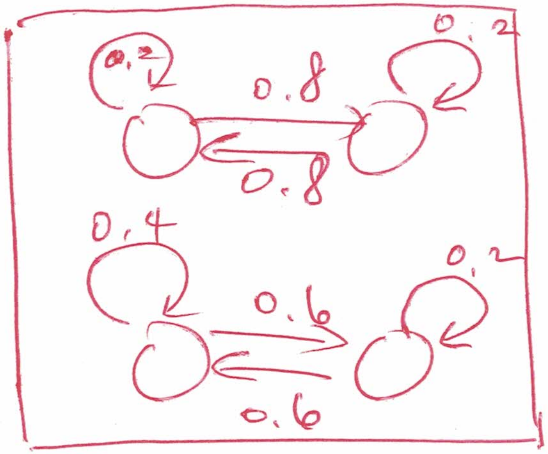
(2). Do we have

$$\lim_{n \to \infty} P^n = P^* \text{ for some } P^* ?$$

The π^* is called stationary distribution.

π^* always exists <u>but</u> may not be unique.

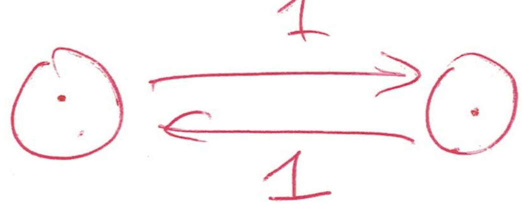
Exaple:



has many stationary distribution.

P^* may not exist.

Exaple:



or, $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Big mistake for many CS students:

when we learn Markov Chain, we forget that

it talks about graphs !!!!!!

Key Theorem.

intuition: P defines a graph G where we only keep edges with positive transition probability.

Example: $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \Rightarrow$

$$\begin{tikzcd} 1 \arrow[r, bend left, "1"] & 2 \arrow[l, bend left, "1"] \end{tikzcd}$$

$$\begin{tikzcd} 1 \arrow[loop above, "0"] \arrow[r, bend left, "1"] & 2 \arrow[loop right, "0"] \arrow[l, bend left, "1"] \end{tikzcd}$$

when G has only one "leaf" SCC

(It is a SCC that can not "walk" to other SCC)

// Recall: if we treat any SCC as a Big
// node, then all the Big nodes form
// a DAG !!! Above says this DAG
// has only <u>one</u> leaf node (which has
// outdegree 0.

then $\underline{P}$ has unique stationary distribution!