Feb 01 2013

마르코프 체인, 이론과 그 응용

Kyomin Jung
KAIST
Applied Algorithm Lab

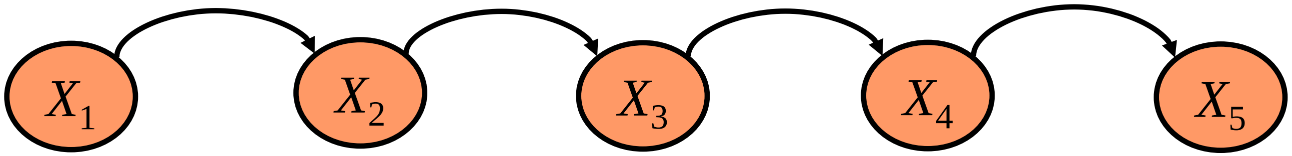
Overview

- Markov Chain
- Example
- Stationary Distribution
- Banach Fixed Point Theorem
- Application: Page Rank

Markov Process

- Markov Property: The state of the system at time t+1 depends only on the state of the system at time t

$$\Pr\left[X_{t+1} = x_{t+1} \mid X_1 \cdots X_t = x_1 \cdots x_t\right] = \Pr\left[X_{t+1} = x_{t+1} \mid X_t = x_t\right]$$



- Transition probabilities are independent of time

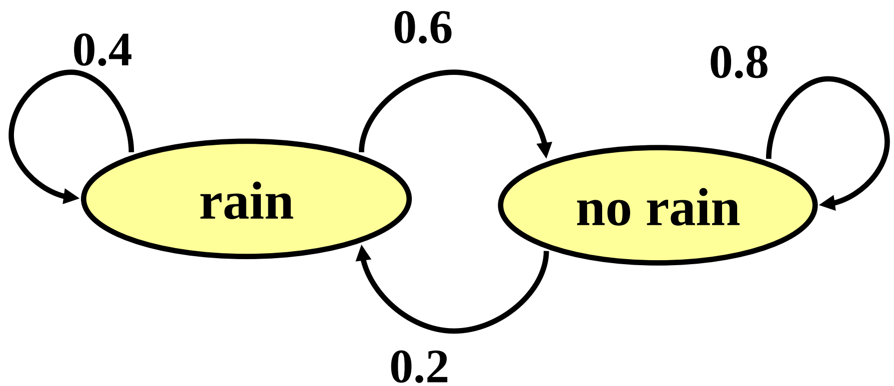
$$\Pr\left[X_{t+1} = b \mid X_t = a\right] = p_{ab}$$

Markov Chain : A Simple Example

Weather:

- raining today → 40% rain tomorrow
  → 60% no rain tomorrow

- not raining today → 20% rain tomorrow
  → 80% no rain tomorrow



Markov Chain : A Simple Example

Weather:

- **raining today** ⇒ **40% rain tomorrow**
  ⇒ **60% no rain tomorrow**

- **not raining today** ⇒ **20% rain tomorrow**
  ⇒ **80% no rain tomorrow**

The transition matrix:

$$P = \begin{pmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{pmatrix}$$

Rows sum up to 1

Markov Property

- "**Future**" is independent of "**Past**" and depend only on "**Present**"

- In other words: Memoryless

- Useful for modeling and analyzing many real systems

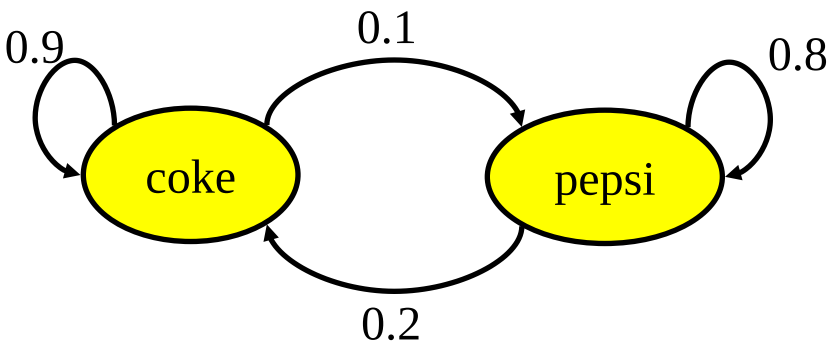
Markov Chain

Coke vs Pepsi example

- Given that a person's last cola purchase was Coke, there is a 90% chance that his next cola purchase will also be Coke.
- If a person's last cola purchase was Pepsi, there is an 80% chance that his next cola purchase will also be Pepsi.

transition matrix:

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$



Markov Chain

Coke vs Pepsi example

Given that a person is currently a Pepsi purchaser, what is the probability that he will purchase Coke two purchases from now?

Pr[Pepsi→?→Coke] =

Pr[Pepsi→Coke→Coke] + Pr[Pepsi→ Pepsi →Coke] =

0.2 * 0.9 + 0.8 * 0.2 = 0.34

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}$$

Pepsi → ?   ? → Coke

Markov Chain

Coke vs Pepsi example

Given that a person is currently a Coke purchaser, what is the probability that he will purchase Pepsi **three** purchases from now?

$$P^3 = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix} = \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix}$$

Markov Chain

Coke vs Pepsi example

- Assume each person makes one cola purchase per week
- Suppose 60% of all people now drink Coke, and 40% drink Pepsi
- What fraction of people will be drinking Coke three weeks from now?

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \qquad P^3 = \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix}$$

$$\Pr[X_3=\text{Coke}] = 0.6 * 0.781 + 0.4 * 0.438 = 0.6438$$

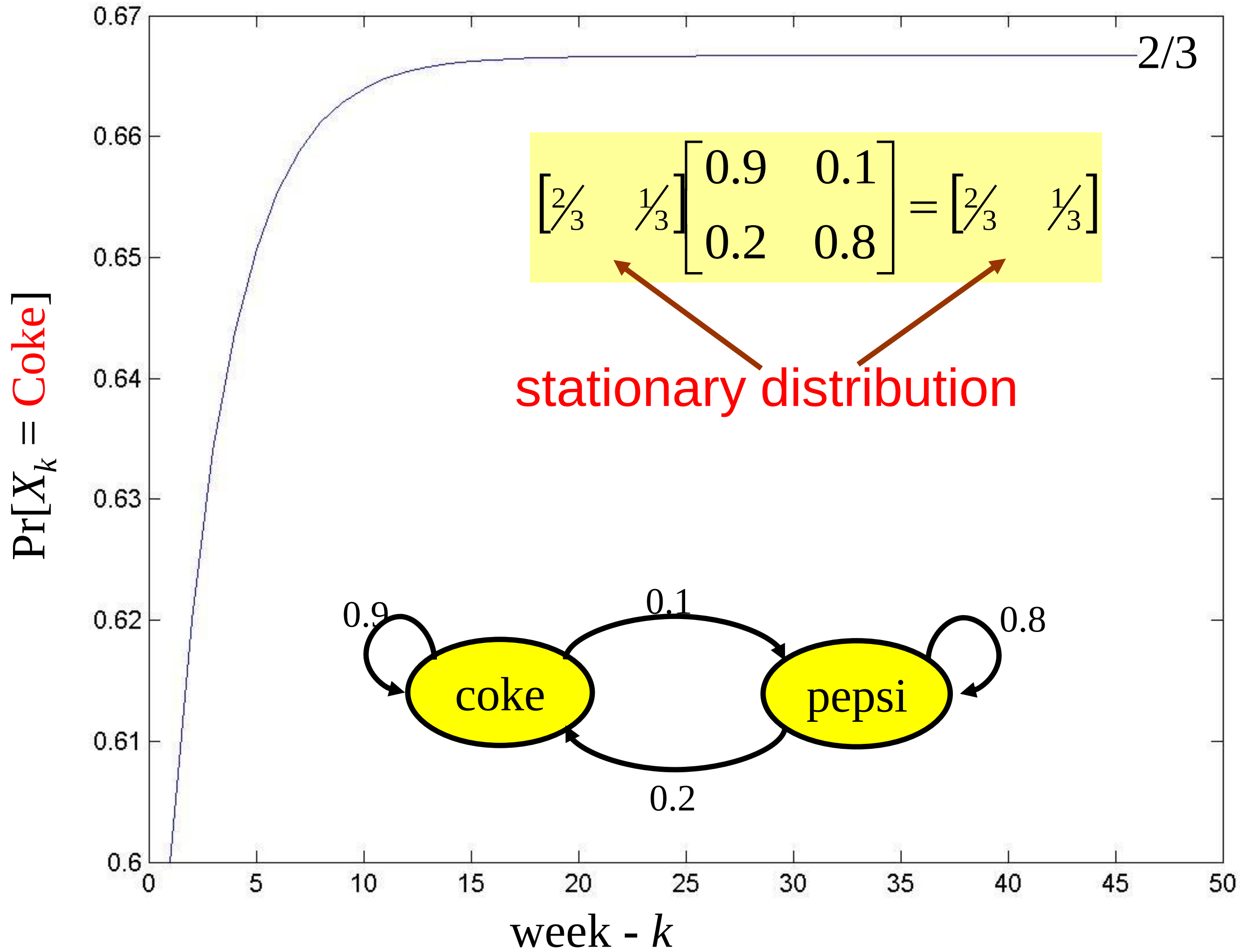
π_k - the distribution in week k

$\pi_0=(0.6,0.4)$ - initial distribution

$\pi_3 = \pi_0 * P^3 = (0.6438, 0.3562)$

Markov Chain

Coke vs Pepsi example



Stationary Distribution

Note that

$$\pi_{k+1} = \pi_k P$$

Stationary Distribution: $\pi = \lim_{k\to\infty} \pi_k$

for any π_0 .

If such π exists, it satisfies that

for all j, $\sum_i p_{ij}\pi(i) = \pi(j)$ and $\sum_i \pi(i) = 1$

Special case : Contraction Map

- Let (X, d) be a metric space, $f : X \to X$ is called a contraction if there is a $\rho \in [0,1)$ such that

$$d\big(f(x), f(y)\big) \le \rho d(x, y) \quad \forall x, y \in X$$

A Markov Chain is a contraction if for any two initial distributions π_1, π_2 ,

$$\mathrm{d}(\pi_1 \mathrm{P}, \pi_2 \mathrm{P}) \le \rho \mathrm{d}(\pi_1, \pi_2)$$

Banach Fixed Point Theorem

- Let (X,d) be a complete metric space, and $f:X \to X$ be a contraction. Then *f* has a fixed point. Further, this fixed point is unique.

- So if $\mathrm{d}(\pi_1 \mathrm{P}, \pi_2 \mathrm{P}) \le \rho \mathrm{d}(\pi_1, \pi_2)$ for some $\rho \in [0,1)$, then the Markov Chain P has a unique stationary distribution (fixed point).

PageRank: Markov Chain over the Web

- If a user starts at a random web page and surfs by clicking links and randomly entering the new web page, what is the probability that a random user is found at a web page?

- The PageRank of a page captures this notion
  - “popular” web page get a higher rank
  - This gives a rule for a random walk on The Web graph (a directed graph).

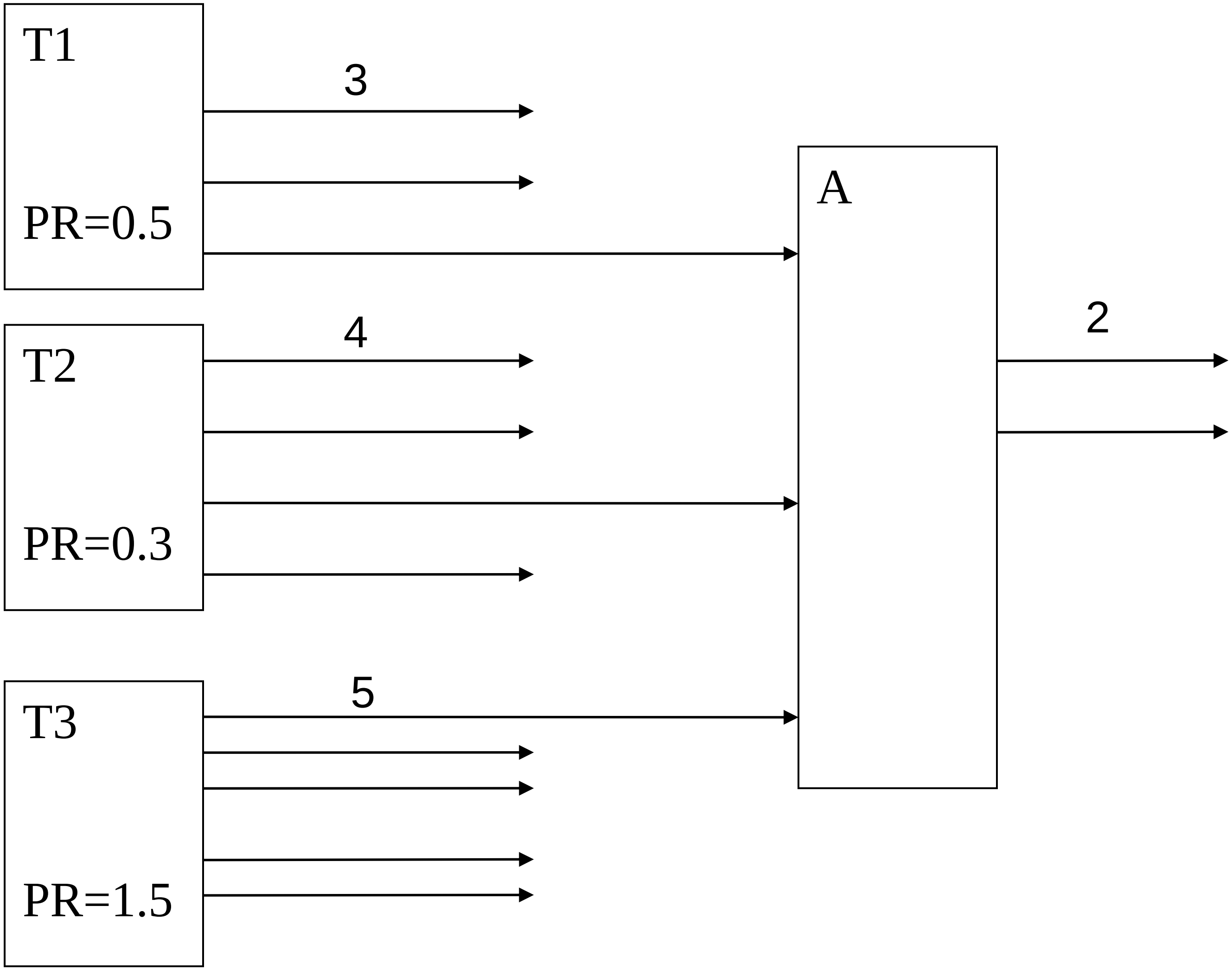
PageRank: Formula

Given page A, and pages T_1 ,… T_j linking to A, PageRank PR(A) of A is defined as the solution of :

$$PR(A) = (1-\rho) + \rho(PR(T_1)/D(T_1) + \ldots + PR(T_j)/C(T_j))$$

- D(T) is the out-degree of page T
- ρ is the "random URL" factor (=0.85)

- Then, PR/n is the stationary distribution of the corresponding Markov chain, where n is the total number of web pages.



$$PR(A)=(1-\rho) + \rho*(PR(T1)/D(T1) + PR(T2)/D(T2) + PR(T3)/D(T3))$$
$$=0.15+0.85*(0.5/3 + 0.3/4+ 1.5/5)$$

PageRank: Computation

- This Markov Chain is a **contraction** with $\rho=0.85$, and PR is its stationary distribution.
  - Banach fixed point theorem holds
- Hence PR can be calculated iteratively :

$$PR_{i+1}(A) = (1-\rho) + \rho(PR_i(T_1)/C(T_1) + \ldots + PR_i(T_n)/C(T_n))$$

- Each page distributes its PR_i to all pages it links to. Linkees add up their awarded ranks to compute their PR_{i+1}.
- PR_i converges exponentially fast over i.