Probabilistic Method and Random Graphs

Lecture 11. A Brief Introduction to Markov Chains [1]

Xingwu Liu

Institute of Computing Technology
Chinese Academy of Sciences, Beijing, China

[1] The slides are mainly based on *Introductory Lecture Notes on Markov Chains And Random Walks* by Takis Konstantopoulos.

Preface

Questions, comments, or suggestions?

A recap of Lovász local lemma

Mission

- Do events $A_1, ...A_n$ satisfy $\Pr(\bigcup_{i=1}^n A_i) < 1$?

Symmetric version: $\Pr(\bigcup_{i=1}^n A_i) < 1$ when

- $edp \leq 1$ for all i, with $p = \max_i \Pr(A_i), d = \max_i |\Gamma(A_i)|$

Asymmetric version: $\Pr(\bigcup_{i=1}^n A_i) < 1$ when

- $\forall i, \sum_{A_j \in \Gamma(A_i)} \Pr(A_j) \leq \frac{1}{4}$, or
- $\exists x_1, ...x_n \in (0,1)$ s.t. $\forall i, \Pr(A_i) \leq x_i \prod_{A_j \in \Gamma(A_i)} (1 - x_j)$
- Shearer's bound is tight
- Moser-Tardos algorithm is efficient up to Shearer's bound

An overall review of probabilistic method

Handling dependence, exploiting independence

- Counting (union bound): mutually exclusive
- First moment: linearity doesnt care dependence
- Second moment: pairwise dependence
- LLL: global dependence

Continue this trend in stochastic process

Markov Chains

Informal definition

A mathematical model of a random phenomenon evolving with time such that the past affects the future only through the present

Time can be discrete or continuous (Markov process)

Debut of the concept of Markov chains

Andrey Markov. Extension of the law of large numbers to dependent quantities, Izvestiia Fiz.-Matem. Obsch. Kazan Univ., (2nd Ser.), 15(1906), pp. 135-156

From an individual to a sequence of random variables

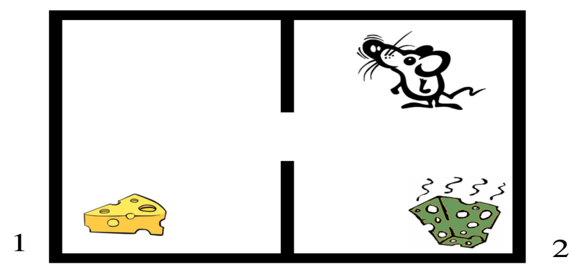
- Asymptotical behavior matters

Andrey Andreyevich Markov

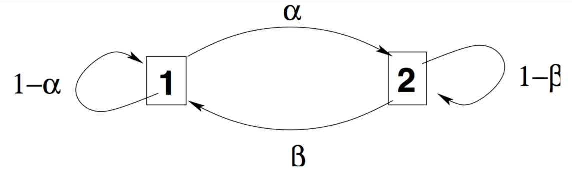
А. А. Марков (1886).

Russian Mathematician
1856-1922

Example: a mouse in cage



Behavior of the mouse (transition diagram): $\alpha = 0.05, \beta = 0.99$



Example: a mouse in cage

Behavior of the mouse (transition matrix)

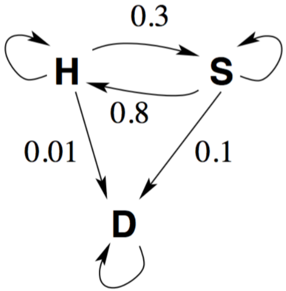
$$\mathsf{P} = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} = \begin{pmatrix} 0.95 & 0.05 \\ 0.99 & 0.01 \end{pmatrix}$$

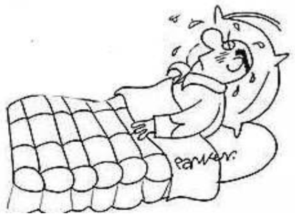
Interesting questions

- How long does it take for the mouse, on the average, to move from cell 1 to cell 2?
  - Easy to solve due to the geometric distribution
- How often is the mouse in room 1?
  - Hard to answer it in one minute

Example: insurance company's puzzle

Human health on a monthly basis





Transition matrix

$$P = \begin{pmatrix} 0.69 & 0.3 & 0.01 \\ 0.8 & 0.1 & 0.1 \\ 0 & 0 & 1 \end{pmatrix}$$

What is the distribution of the lifetime of a currently healthy one?

Formal definition of Markov Chains

General setting

- A sequence of random variables $\{X_n : n \in \mathbb{N}\}$
- For all n, X_n is defined on the same state space S
  - Any $s \in S$ is called a state

Markov property

$\Pr(X_{n+1} = x_{n+1} | X_n = x_n, ... X_0 = x_0) = \Pr(X_{n+1} = x_{n+1} | X_n = x_n)$, for any $n \in \mathbb{N}$ and $x_0, ... x_n \in S$.

The future is independent of the past, given the present state

Homogeneous

$\Pr(X_{n+1} = y | X_n = x)$ is independent of n, denoted by p_{xy}

Focus on homogeneous Markov chains

Representation of a Markov chain

Transition diagram

Weighted directed graph $G = (V, E, W)$

- $V = S$, the state space
- $e_{ij} \in E$ if and only if $p_{ij} \triangleq \Pr(X_t = j | X_{t-1} = i) > 0$
- $W : e_{ij} \mapsto p_{ij}$

This provides **intuition**

- Example: state reachability is reachability over the graph

Transition matrix

$P = (p_{ij})_{i,j \in S}$, all entries are nonnegative, $\sum_j p_{ij} = 1$
This enables **calculation**

- Example: $P^{(n)} = P^n$, where $P^{(n)} = (p_{ij}^{(n)})_{i,j \in S}$,
  $p_{ij}^{(n)} \triangleq \Pr(X_n = j | X_0 = i)$

Multistep transition matrix

$P^{(n)} = P^n$

Proof by induction on n.

Remark: a summand of $p_{ij}^{(n)}$ corresponds to a path from i to j whose length is n

State distribution at time t

Given initial distribution π, $\pi^{(t)} = \pi P^{(t)} = \pi P^t$

Interesting questions

- Can a state j be reached from i?
- If yes, when?
- What's the state distribution at any t?
- What's the distribution in the long run (average frequency)?

Reachability

Equivalent conditions of reaching j from i

- There is a directed path in G from i to j
- $p_{ij}^{(n)} > 0$ for some n

Denoted by $i \rightsquigarrow j$

Communicating states

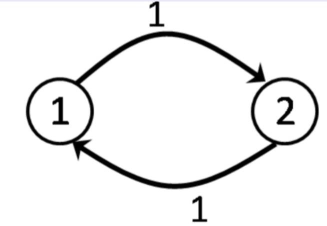
$i \leftrightsquigarrow j$ if $i \rightsquigarrow j$ and $j \rightsquigarrow i$

Communicating classes: equivalence classes of $\leftrightsquigarrow$

- Strongly connected components of G

Period

The life style of a pig



$p_{ii}^{(n)} > 0$ only if n is even. It is *periodic*

The period of state i of a Markov chain

d_i is the GCD of $D_i \triangleq \{n \geq 1 : p_{ii}^{(n)} > 0\}$.
If $d_i = 1$, i is said to be aperiodic

Communicating states have the same period

Theorem

If $i \leftrightsquigarrow j$, then $d_i = d_j$

Proof

- Since $i \leftrightsquigarrow j$, $p_{ij}^{(s)} > 0$ and $p_{ji}^{(t)} > 0$ for some $s, t > 0$
- $p_{ii}^{(s+t)} \geq p_{ij}^{(s)} p_{ji}^{(t)} > 0$, so d_i divides $s + t$
- For any $n \in D_j$, $p_{ii}^{(s+n+t)} \geq p_{ij}^{(s)} p_{jj}^{(n)} p_{ji}^{(t)} > 0$, so d_i divides $s + n + t$
- Since d_i divides $s + t$, d_i divides n
- d_i divides d_j
- Symmetrically, d_j divides d_i
- $d_j = d_i$

Nonzero multi-step transition probability of aperiodic states

Theorem

If i is aperiodic, $p_{ii}^{(n)} > 0$ for all large enough n

Proof

- Choose $n_1, n_2 \in D_i$ s.t. $n_2 - n_1 = 1$
- For any n, there are integers q and $r < n_1$ s.t. $n = qn_1 + r$
- $n = qn_1 + r(n_2 - n_1) = (q - r)n_1 + rn_2$
- When n is large enough, $q - r > 0$
- $p_{ii}^{(n)} \geq \left(p_{ii}^{(n_1)}\right)^{q-r} \left(p_{ii}^{(n_2)}\right)^{r} > 0$

Hitting time

Definition

T_{ij}: the first time that j is reached when the initial state is i

- $f_{ij}^{(n)} \triangleq \Pr(T_{ij} = n) = \Pr(X_n = j, X_k \neq j, 1 \leq k < n | X_0 = i)$
- $f_{ij} \triangleq \sum_n f_{ij}^{(n)}$

Recurrency

If $f_{ii} = 1$, the state i is recurrent (otherwise, transient)

- Furthermore, if $\mathbb{E}[T_{ii}] < \infty$, i is positive recurrent
- Otherwise, it is null recurrent

Example

Human health chain, pig life style chain, and more

Decision theorem of recurrency

The following conditions are equivalent

1. i is recurrent
2. $\sum_n p_{ii}^{(n)} = \infty$
3. $\mathbb{E}[J_i|X_0 = i] = \infty$, J_i is the number of times i is reached
4. $\Pr(J_i = \infty|X_0 = i)$=1

Proof: 2⇔3

$$\begin{aligned} J_i &= \sum_n \mathbf{1}(X_n = i) \\ \mathbb{E}[J_i|X_0 = i] &= \mathbb{E}[\sum_n \mathbf{1}(X_n = i)|X_0 = i] \\ &= \sum_n \Pr(X_n = i|X_0 = i) \\ &= \sum_n p_{ii}^{(n)} \end{aligned}$$

Proof (continued)

1⇒ 4

- Let $J_i^{(l)}$ be the times of reaching i no earlier than step l
- Property: $J_i = J_i^{(1)}$
- $g_{ii} \triangleq \Pr(J_i = \infty | X_0 = i) = \lim_k \Pr(J_i^{(1)} \geq k | X_0 = i)$
- $(J_i^{(1)} \geq k + 1 | X_0 = i) = \cup_l (T_{ii} = l, J_i^{(l+1)} \geq k | X_0 = i)$
- $\Pr(J_i^{(1)} \geq k + 1 | X_0 = i) = f_{ii} \Pr(J_i^{(1)} \geq k | X_0 = i) = f_{ii}^{k+1}$
- $g_{ii} = \lim_k f_{ii}^k = 1$ since i is recurrent

4⇒ 3

Trivial

Proof: 2⇒ 1

- Chapman-Kolmogorov equation:
  $p_{ij}^{(n)} = \sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)}$, $p_{ii}^{(0)} = 1$
- For any N,

$$
\begin{aligned}
\sum_{n=1}^{N} p_{ii}^{(n)} &= \sum_{n=1}^{N} \sum_{k=1}^{n} f_{ii}^{(k)} p_{ii}^{(n-k)} \\
&= \sum_{k=1}^{N} f_{ii}^{(k)} \sum_{n=k}^{N} p_{ii}^{(n-k)} \\
&= \sum_{k=1}^{N} f_{ii}^{(k)} \sum_{n=0}^{N-k} p_{ii}^{(n)} \\
&\le \sum_{k=1}^{N} f_{ii}^{(k)} \sum_{n=0}^{N} p_{ii}^{(n)}
\end{aligned}
$$

- $\frac{\sum_{n=1}^{N} p_{ii}^{(n)}}{1+\sum_{n=1}^{N} p_{ii}^{(n)}} = \frac{\sum_{n=1}^{N} p_{ii}^{(n)}}{\sum_{n=0}^{N} p_{ii}^{(n)}} \le \sum_{k=1}^{N} f_{ii}^{(k)} \le f_{ii} \le 1$
- Since $\sum_{n=1}^{N} p_{ii}^{(n)} = \infty$, the lefthand side $\to 1$ as $N \to \infty$
- $f_{ii} = 1$, so i is recurrent

Recurrency is preserved by communicating relation

Theorem

If $i \leftrightsquigarrow j$ and i is recurrent, then so is j

Prove

It immediately follows from the above theorem

A necessary condition of transient states

Theorem

If j is a transient, $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$ for any i

Proof

- $p_{ij}^{(n)} = \sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)}$, $p_{ii}^{(0)} = 1$
- For any N,

$$\begin{aligned} \sum_{n=1}^{N} p_{ij}^{n} &= \sum_{n=1}^{N} \sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)} \\ &= \sum_{k=1}^{N} \sum_{n=k}^{N} f_{ij}^{(k)} p_{jj}^{(n-k)} \\ &= \sum_{k=1}^{N} f_{ij}^{(k)} \sum_{n=0}^{N-k} p_{jj}^{(n)} \\ &\leq \sum_{k=1}^{N} f_{ij}^{(k)} \sum_{n=0}^{N} p_{jj}^{(n)} \end{aligned}$$

- $\sum_{n=1}^{N} p_{ij}^{(n)} \leq \sum_{n=0}^{N} p_{jj}^{(n)} \leq 1 + \sum_{n=1}^{N} p_{jj}^{(n)} < \infty$

Positive recurrency

Any rule for deciding if a state is positive recurrent?

How to compute the expected hitting time of a positive recurrent state?

Reference

- Introductory Lecture Notes on Markov Chains And Random Walks by Takis Konstantopoulos
- Baidu Wenku