Probabilistic Method and Random Graphs

Lecture 12. Excursions, Stationary Distributions, and Applications of Markov Chains [1]

Xingwu Liu

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

January 6, 2020

[1] The slides are mainly based on *Introductory Lecture Notes on Markov Chains And Random Walks* by Takis Konstantopoulos and *Lecture Notes of Stochastic Processes* by Glen Takahara.

Questions, comments, or suggestions?

Outline

1. A recap of Lecture 11
2. Excursions
3. Stationary Distribution
4. Calculation of Stationary Distribution
5. Applications of Stationary Distribution

Basic concepts of Markov chains

Definition

- A stochastic process which has the Markov perperty
  - Time homogeneous

Representations

- Transition diagram: weighted directed graph
- Transition matrix: $P = (p_{ij})_{i,j \in S}$

Reachability

- Period
- Hitting time T_{ij}
  - $f_{ij}^{(t)} \triangleq \Pr(T_{ij} = t)$, $f_{ij} \triangleq \sum_t f_{ij}^{(t)}$

Classification of states

Definition

- Transient if $f_{ii} < 1$, otherwise recurrent
- Positive recurrent if $\mathbb{E}[T_{ii}] < \infty$, otherwise null recurrent

Equivalient definitions of recurrent states

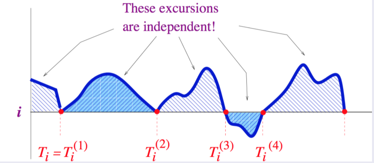
- $\sum_n p_{ii}^{(n)} = \infty$
- $\mathbb{E}[J_i | X_0 = i] = \infty$, J_i is the number of times i is reached
- Pr$(J_i = \infty | X_0 = i)$=1

Corollary

- If $i \leftrightsquigarrow j$ and i is recurrent, then so is j
- Cool! Counterpart of positive recurrent? See excursions...

Excursions: independent structure in Markov chains

If a Markov chain $\{X_t\}$ visits state i again and again



Hitting times

- $T_i = T_i^{(1)} = \min\{t > 0 : X_t = i\}$
- $T_i^{(r)} = \min\left\{t > T_i^{(r-1)} : X_t = i\right\}$

Excursion: trajectory between two successive visits to state i

- $\chi_i^{(r)} = \left\{X_t : T_i^{(r)} \leq t < T_i^{(r+1)}\right\}, r \geq 1$
- $\chi_i^{(0)} = \left\{X_t : 0 \leq t < T_i^{(1)}\right\}$

Excursions are i.i.d. random variables

Theorem

- $\chi_i^{(r)}, r \geq 0$, are independent
- $\chi_i^{(r)}, r \geq 1$, have the same distribution

Proof

It follows from the strong Markov property

Remark

- Dependence is annoying, but excursions decouple the chain into independent blocks
- The independent structure means so much ...

Revisiting positive recurrent states

Theorem

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

Proof

- i is positive recurrent, so there are infinitely many excursions
- The length of $\chi_i^{(r)}$: $T_i^{(r+1)} - T_i^{(r)} = T_{ii}$ with finite expectation
- Since $i \rightsquigarrow j$, starting from i,
  $p =$Pr(reach j before returning to i) > 0
- For each $r > 0$, j appears in $\chi_i^{(r)}$ with probability $p > 0$
- Pr(j is reached)=1. Wlog., j is first reached in $\chi_i^{(0)}$
- R s.t. j is reached next in $\chi_i^{(R)}$? Geometric distribution
- $T_{jj} \leq RT_{ii} \Rightarrow j$ is positive recurrent (by Wald's equation)

Wow! One more example

Law of large number in Markov chains

Law of large number

For i.i.d. r.v. $\{X_n\}$, $\Pr(\lim_{n\to\infty} \frac{X_1+\cdots X_n}{n} = \mathbb{E}[X_1]) = 1$

What if X_n's are states of a Markov chain?

Law of large number in Markov Chains

Assume Markov chain $\{X_n\}$ has a positive recurrent state a, $\Pr(a \text{ is reached} | X_0)$=1, and $f : S \to \mathbb{R}$ is bounded. Then

$$\Pr\left(\lim_{t\to\infty} \frac{f(X_0) + \cdots f(X_t)}{t} = \bar{f}\right) = 1$$

where

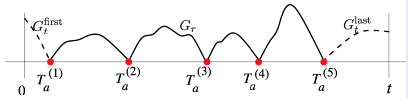
$$\bar{f} = \frac{\mathbb{E}\left[\sum_{n\in\chi_a^{(1)}} f(X_n)\right]}{\mathbb{E}[T_{aa}]} = \mathbb{E}_{\pi}[f]$$

Proof

Basic idea

Break the sum into subsum over excursions, reducing to law of large number of i.i.d. r.v.

$N_t = \max\{r \geq 1 : T_a^{(r)} \leq t\}$: $\#$ a-excursions occurring in [0, t]



$N_t = 5$ in this example

Irregular parts vanish

Full excursions are i.i.d. with expectation $\bar{f} = \dfrac{\mathbb{E}\left[\sum_{n \in \chi_a^{(1)}} f(X_n)\right]}{\mathbb{E}[T_{aa}]}$

1. A recap of Lecture 11
2. Excursions
3. **Stationary Distribution**
4. Calculation of Stationary Distribution
5. Applications of Stationary Distribution

Stationary Distribution

Motivation

1. Given a positive recurrent state i, how to calculate $\mathbb{E}[T_{ii}]$?
2. If t is sufficiently large, what is the distribution of X_t?

Definition

A distribution π over S satisfying $\pi P = \pi$ is a stationary distribution of the Markov chain.

Fundamental Problems

1. Given a Markov chain, does it have a stationary distribution?
2. Is the stationary distribution unique?
3. How to calculate it?

An example

Waiting for a bus

- When a bus arrives, $\Pr$(the next arrives in i minutes)$= p(i)$
- X_t: the time till the arrival of the next bus
- Transition probability: $p_{i,i-1} = 1, p_{0,i} = p(i), i \geq 1$

If a stationary distribution π exists

- It holds that $\pi(i) = \sum_{j\geq 0} \pi(j)p_{j,i} = \pi(0)p(i) + \pi(i+1)$
- This implies $\pi(i) = \pi(0)\sum_{j\geq i} p(j)$
- Since $\sum_{i\geq 0} \pi(i) = 1$, we have $\pi(0) = (\sum_{i\geq 0}\sum_{j\geq i} p(j))^{-1}$
- The stationary distribution exists iff $\sum_{i\geq 0}\sum_{j\geq i} p(j) < +\infty$
- $\sum_{i\geq 0}\sum_{j\geq i} p(j) = \sum_{j\geq 0}\sum_{0\leq i\leq j} p(j)$
  $= \sum_{j\geq 0}(j+1)p(j) = \mathbb{E}[T_{00}]$
- The stationary distribution exists iff 0 is positive recurrent

Is this correct in general?

Yes!

Existence Theorem of Stationary Distribution

Assume that a is a recurrent state. For any state x, define $\nu^{[a]}(x) = \mathbb{E}\left[\sum_{n=0}^{T_{aa}-1} \mathbf{1}(X_n = x) | X_0 = a\right]$.

Lemma

$\nu^{[a]} = \nu^{[a]} P$.

For any state x s.t. $a \rightsquigarrow x$, we have $0 < \nu^{[a]}(x) < +\infty$.

Theorem of existence

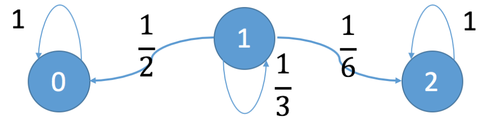
If the Markov chain has a positive recurrent state a, $\pi^{[a]} \triangleq \frac{\nu^{[a]}}{\mathbb{E}[T_{aa}]}$ is a stationary distribution.

Proof of the theorem

It immediately follows from the lemma

The Stationary Distribution is not Necessarily Unique

- Consider the Markov chain



- States 0 and 2 are positive recurrent, so stationary distributions exist
- For any $0 \leq \alpha \leq 1$, $\pi = (\alpha, 0, 1-\alpha)$ is a stationary distribution. **Not Unique!**

- Note that the chain is reducible
- Does this cause the non-uniqueness?
- Yes!

Uniqueness Theorem

Theorem of uniqueness

For an irreducible Markov chain, its stationary distribution is unique if existent

Actually, when irreducible, if a stationary distribution π exists

- $\pi_i \mathbb{E}[T_{ii}] = 1$ for every state
- All states must be positive recurrent

Any Markov chain with stationary distribution π

- $\pi_j = 0$ if j is transient
- j is positive recurrent if $\pi_j > 0$

Calculating Expected Return Time

When irreducible, $\pi(i) = \frac{1}{\mathbb{E}[T_{ii}]}$ for any i.

Calculating $\mathbb{E}[T_{ii}]$ is reduced to calculating the stationary distribution.

But how to calculate the stationary distribution?

Stability Theorem

Theorem of stability

Let π be the stationary distribution of an irreducible and ergodic (positive recurrent, aperiodic) Markov chain. Then

1. $\lim_{n\to\infty} \Pr(X_n = x) = \pi(x)$ for any initial distribution and any $x \in S$;
2. $\lim_{n\to\infty} p_{yx}^{(n)} = \pi(x)$ for any $x, y \in S$.

Remarks

- Approximating by iteratively computing
- Each row of $P^{(n)}$ converges to π
- Though $\sum_n p_{yx}^{(n)} = +\infty$ when x is recurrent
  - $\lim_{n\to\infty} p_{yx}^{(n)} > 0$ if x is positive recurrent
  - $\lim_{n\to\infty} p_{yx}^{(n)} = 0$ if x is null recurrent

Sub-summary: fundamental theorems of Markov chains

Existence	Uniqueness	Stability
Positive recurrency	Irreducibility	Aperiodicity $\cdots$
$\pi^{[a]}(i) = \frac{\mathbb{E}\left[\sum_{n=1}^{T_{aa}} \mathbf{1}(X_n{=}i) \mid X_0{=}a\right]}{\mathbb{E}[T_{aa}]}$	$\pi = \pi^{[a]}$	$\lim_{n\to\infty} p_{ji}^{(n)} = \pi(i)$

Are the conditions necessary?

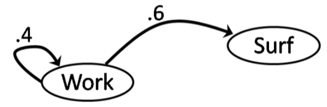
Positive Recurrency is Necessary for Existence

Theorem

If a Markov Chain has a stationary distribution π, then any state i with $\pi(i) > 0$ is positive recurrent.

Does uniqueness imply irreducibility?

No!



However

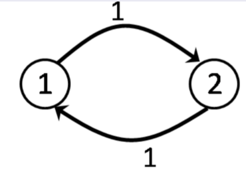
It is weakly irreducible:
only one communicating class of positive recurrent states

Theorem

If a Markov chain has a unique stationary distribution, it has a unique communicating class of positive recurrent states

No Aperiodicity, No Stability

Consider the Markov chain with period 2.



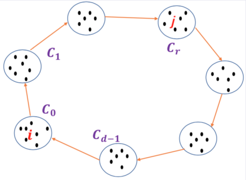
$p_{11}^{(2k)} = 1$, but $p_{11}^{(2k-1)} = 0$. So, $\lim_{n\to\infty} p_{11}^{(n)}$ does not exist.

- Generally, in case of period d, does $\lim_{n\to\infty} p_{jj}^{(nd)}$ exist?
  - If existent, what's it?
- Yes!

The Normal Form of Periodic Markov Chain

Normal form theorem

Given an irreducible Markov chain with period d, the state space S can be uniquely partitioned into disjoint sets $C_0, C_1, ...C_{d-1}$ such that $\sum_{j \in C_{r+1 \bmod d}} p_{ij} = 1$ for $i \in C_r$, $r = 0, 1, ...d - 1$.



$$\lim_{n\to\infty} p_{ij}^{(r+nd)} = d\pi(j) = \frac{d}{\mathbb{E}[T_{jj}]}$$

Sub-summary: fundamental theorems of Markov chains

Existence	Uniqueness	Stability
Positive recurrency	Irreducibility	Aperiodicity $\cdots$
$\pi^{[a]}(i) = \dfrac{\mathbb{E}\left[\sum_{n=1}^{T_{aa}} \mathbf{1}(X_n{=}i) \mid X_0{=}a\right]}{\mathbb{E}[T_{aa}]}$	$\pi = \pi^{[a]}$	$\lim_{n\to\infty} p_{ji}^{(n)} = \pi(i)$

All the conditions are (weakly) necessary!

1. A recap of Lecture 11
2. Excursions
3. Stationary Distribution
4. Calculation of Stationary Distribution
5. Applications of Stationary Distribution

By stability theorem

Compute iteratively or approximate by limits

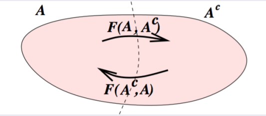
By definition

- Solve the linear equation system $\pi = \pi P$, $\sum_{i \in S} \pi(i) = 1$
- Flow balance theorem

Flow balance theorem

Flow balance

Let $A \subseteq S$ be a set of the states of a Markov chain, and π be a distribution over S. Define $F(A, A^c) = \sum_{i \in A, j \in A^c} \pi(i) p_{ij}$.



Theorem

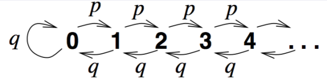
π is a stationary distribution if and only if $F(A, A^c) = F(A^c, A)$ for all $A \subseteq S$.

Proof

($\Leftarrow$) Prove by considering singletons A.
($\Rightarrow$) Observe that $\pi_i \sum_j p_{ij} = \pi_j \sum_j p_{ji}$.

Example

Walk with a barrier



Find the stationary distribution by definition

$\pi(i) = p\pi(i-1) + q\pi(i+1)$ for all $i > 0$.

By flow balance theorem

- For any $i > 0$, let $A = \{0, 1, ...i-1\}$
- $\pi(i-1)p = F(A, A^c) = F(A^c, A) = \pi(i)q$
- $\pi(i) = (p/q)^i \pi(0)$

Application

Natural language processing

A fundamental problem: Computing the probability that a sentence appears. The computation is made possible by Markov hypothesis.

PageRank

Task: Assign importance to web pages.
Model: Web graph consists of linked pages. A typical process of surfing the Web is to follow links and randomly jump in case of dangling. So we get a Markov chain with transition probability

$$\widehat{p}_{ij} = \begin{cases} 1/|L(i)| & \text{if } j \in L(i) \\ 1/|V| & \text{if } L(i) = \emptyset \\ 0 & \text{otherwise} \end{cases}$$

To guarantee irreducibility and aperiodicity, use **bored surfer** style.
Namely $p_{ij} \triangleq (1-\alpha)\widehat{p}_{ij} + \frac{\alpha}{|V|}$.
The stationary distribution is the rank. Compute iteratively.

Random Walks on Undirected Graphs

Random walks

Let $G = (V, E)$ be a finite, undirected, and connected graph. A random walk on G is a Markov chain with $p_{uv} = \frac{1}{d_u}$ for $(u, v) \in E$

Period

A random walk on G is aperiodic iff G is not k-partite

Stationary distribution

Stationary distribution of a random walk on G: $\pi(v) = \frac{d(v)}{2|E|}$.

So, expected return time $h_{uu} \triangleq \mathbb{E}[T_{uu}] = \frac{2|E|}{d(u)}$.

Expected Hitting Time and Cover Time

Claim: If $(u, v) \in E$, then $h_{vu} < 2|E|$.
Proof: Use the fact that $\frac{2|E|}{d(u)} = h_{uu} = \frac{1}{d(u)} \sum_{v \in N(u)} (1 + h_{vu})$.

Cover time

Claim: The cover time of $G = (V, E)$ is no more than $4|V| \cdot |E|$.
Proof: Explore the Eulerian tour on a spanning tree of G. The expected time to go through the vertices $v_0, v_1, \ldots v_{2|V|-2} = v_0$ upper bounds the cover time.

Parrondo's Paradox (since 1996)

A question that seems silly at the first glance

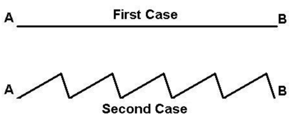
Can you combine two losing games to get a winning one?
Yes!

The magic example

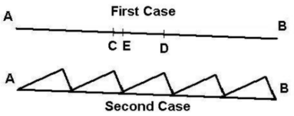
- Game G_1: flip coin a with head probability $p_a < \frac{1}{2}$. You win a dollar if you get Head, otherwise lose a dollar.
- Game G_2: Let l be the number of losses so far and w be that of wins. You have coins b and c. Flip b if $w - l = 0(mod\ 3)$, and flip c otherwise. You win a dollar if you get Head, otherwise lose a dollar.
- Game G_3: repeatedly flip a fair coin d. If you get Head, proceed as in game G_1; otherwise proceed to G_2.

When $p_a = 0.49, p_b = 0.09, p_c = 0.74$, A and B are losing games while C is a winning one.

An intuitive interpretation



Both cases get a tie



- In both cases, B wins
- If the cases appear alternately, A can win

Randomness is not necessary

G_1': lose 1. G_2': lose 5 for odd capital, win 3 otherwise.
G_3': Play alternatively, beginning with G_2'

Why?

The difficulty lies in analyzing G_2

Try to determine the relative probability of reaching -3 or $+3$ first, or study the probability of wins in stationary distribution.

Game G_3 is like G_2, except that the head probabilities are slightly different.

References

- Lecture Notes of Stochastic Processes, by Glen Takahara `http://www.mast.queensu.ca/~stat455/`
- Introductory Lecture Notes on Markov Chains And Random Walks, by Takis Konstantopoulos `http://www2.math.uu.se/~takis/L/McRw/mcrw.pdf`
- Section 2, Lecture 16 of Lecture notes on Probability and Computing by Ryan O'Donnell
- Section 7.4&7.5 of the textbook *Probability and Computing*

Thank you! Happy the year of pig!