Markov chains

A. Dymov

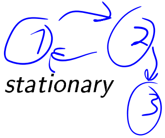
Steklov Mathematical Institute & HSE

23 октября 2020 г.

$\xi_0, \xi_1, \xi_2, \dots$ $\quad p^{(0)} = \pi \Rightarrow \mathbb{P}(\xi_j = k) = \pi_k \ \forall j \quad \pi_j \geq 0, \ \sum_j \pi_j = 1$

Terminology. Markov chain = Mc; Transition probability matrix (матрица переходных вероятностей)=Tpm

Definition. Distribution (распределение) $\pi = (\pi_1, \dots, \pi_L)$ is called *stationary* for a Mc with Tpm Π if $\pi\Pi = \pi$.
That is, π is a left eigenvector of Π with eigenvalue 1.

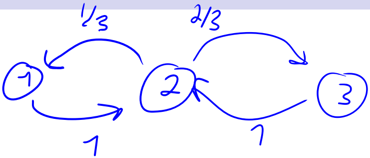


Terminology: Stationary distribution = stationary measure = stationary state.

Remark. We will often consider different Markov chains with the same Tpm (so, these Markov chains will differ only in their initial distribution $p^{(0)}$). Abusing notation, we will often denote them by the same symbols (e.g. as $\xi_0, \dots, \xi_T$). This is convenient since "key information"about a Mc is contained in the Tpm (so, in the graph if the Mc) while initial conditions can be chosen in many different ways (think about the random walk or the Galton-Watson birth-death process).

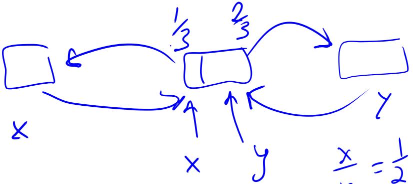
$\Pi \quad p^{(0)} = \pi$ - стац. $\Rightarrow p^{(j)} = \pi \Pi^j = \pi \ \forall j.$

(1)



$$\Pi = \begin{pmatrix} 0 & 1 & 0 \\ 1/3 & 0 & 2/3 \\ 0 & 1 & 0 \end{pmatrix}$$

Найти стац. сост. π? $\quad \pi \Pi = \pi$.



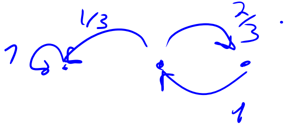
$$\frac{x}{y} = \frac{1}{2}$$

$$2x + 2y = 1.$$

$$x = \frac{1}{6}, \quad y = \frac{1}{3}$$

$$\Rightarrow \pi = \left(\frac{1}{6}, \frac{1}{2}, \frac{1}{3}\right).$$

(2)



$\pi = (1, 0, 0)$ — ! ст. сост.

Let $\mathcal{P}$ denote a space of (probability) distributions, i.e.

$$\mathcal{P} = \{p = (p_1, \dots, p_L) : \sum_{i=1}^{L} p_i = 1, \quad p_i \geq 0 \, \forall i\}.$$

Then $\mathcal{P} \overset{\subseteq}{\cancel{\in}} \mathbb{R}^L$ is a simplex. In particular, $\mathcal{P}$ is a convex compact set in $\mathbb{R}^L$.
For any vector $v = (v_1, \dots, v_L)$ we denote

$$|v|_1 := \sum_{i=1}^{L} |v_j| \quad \text{— the } l^1\text{-norm of } v.$$

Exercise. Check that $(\mathcal{P}, |\cdot|_1)$ is a complete metric space, where $|\cdot|_1$ denotes a metric on $\mathcal{P}$ generated by the l^1-norm.

$d(u,v) = |u-v|_1$ $\qquad \mu_n \in \mathcal{P}$ $\qquad \Rightarrow \nu \in \mathcal{P}$

$|\mu_n - \nu|_1 \to 0$; $\nu \in \mathbb{R}^L$

Lecture 7. Existence of a stationary measure

Theorem 17. *Any homogeneous Markov chain (with finite number of states) has a stationary measure.*

$\pi_j \geq 0 \quad \forall j.$

Corollary 18. Any stochastic matrix has a non-negative left eigenvector corresponding to eigenvalue 1.

$\pi A = \pi.$

$\lambda = 1$ — собств. $\quad A\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

Proof 1 of the theorem: short but specific for our simple case.
Brouwer fixed point theorem: a continuous mapping of a convex compact set to itself has a fixed point.

P — выпуклый компакт в $\mathbb{R}^k$.

$\Pi : P \circlearrowleft, \quad \mu \to \mu\Pi$ — непр.-но $\Rightarrow$

$\exists \pi : \pi \Pi = \pi.$

Proof 2 (Bogoliubov-Krylov method): more complicated but VERY general.

Let $p^{(0)} \in \mathcal{P}$ be an arbitrary initial distribution. Then $p^{(i)} = p^{(0)}\Pi^i$.

Denote

$$p^{(0)} \xrightarrow{\Pi} p^{(1)} \xrightarrow{\Pi} p^{(2)} \dots$$

$$\pi^k = \frac{1}{k}\sum_{i=0}^{k-1} p^{(i)} \in \mathcal{P}. \qquad P_{3_i}$$

Since the sequence π^k is bounded in $\mathbb{R}^L$, it has a convergent subsequence π^{k_j}. Let

$$\pi = \lim_{j\to\infty} \pi^{k_j}.$$

$$p^{(i)}\Pi = p^{(i+1)}$$

Since $(\mathcal{P}, |\cdot|_1)$ is complete, $\pi \in \mathcal{P}$.

Let us show that π is a stationary measure. Indeed,

$$\pi\Pi = (\lim_{j\to\infty} \pi^{k_j})\Pi = \lim_{j\to\infty}(\pi^{k_j}\Pi) = \lim_{j\to\infty}\frac{1}{k_j}\sum_{i=0}^{k_j-1} p^{(i+1)} = \lim_{j\to\infty}\frac{1}{k_j}\sum_{i=1}^{k_j} p^{(i)}$$

$$= \lim_{j\to\infty}\frac{1}{k_j}\Big(\cancel{-p^{(0)}} + \cancel{p^{k_j}} + \sum_{i=0}^{k_j-1} p^{(i)}\Big) = \lim_{j\to\infty}\pi^{k_j} = \pi.$$

0

$A^s = (a_{ij}^{(s)})$ $a_{ij}^{(s)} > 0.$

Definition. A stochastic matrix A is called *ergodic* if there exists $s \in \mathbb{N}$ such that all elements of the matrix A^s are strictly positive.

Lemma 19. *A Tpm Π of a Mc is ergodic if and only if there is $s \in \mathbb{N}$ such that the transition probabilities in s steps $p_{ij}^{(s)} > 0$ for any i, j.*

Proof. Accordingly to Lemma 11 the transition probabilities $p_{ij}^{(s)}$ are elements of the matrix Π^s.

?

Remark 20. The property of a Tpm $\Pi = (p_{ij})$ to be ergodic is NOT related to the explicit values p_{ij} but is related only to the structure of the graph of the corresponding Mc. That is, with positions of positive elements in Π.

$\xi_0, \xi_1, \xi_2, \ldots$

$\mathbb{P}(\xi_1 = i \mid \xi_0 = j) = p_{ji}$

L11: $\xi_0, \xi_s, \xi_{2s}, \ldots$ - Mc

$\mathbb{P}(\xi_s = i \mid \xi_0 = j) = p_{ji}^{(s)}$

Lecture 7. Ergodic theorem (main result of the course, achtung!)

Theorem 21. (Ergodic Theorem) *Assume that a homogeneous Markov chain (with finite number of states) has ergodic Tpm* Π. *Then it has a* **unique** *stationary measure* π *and there are constants* $C > 0$ *and* $0 < \lambda < 1$ *such that for any initial distribution* $p^{(0)}$ *we have*

$$|p^{(n)} - \pi|_1 \le C\lambda^n \quad \text{for any } n \ge 1.$$

$p^{(n)} = p^{(0)}\Pi^n$

Moreover, $\pi_j > 0$ *for every* $1 \le j \le L$.

In particular, for any initial distribution $p^{(0)}$ we have $p^{(n)} \to \pi$ as $n \to \infty$.

- $\Rightarrow$ Markov chain "forgets" initial conditions
- $\Rightarrow$ Exponential convergence to equilibrium

$\mathbb{P}(\xi_n = s) \approx \pi_s$

$n = 1000$

Π^n

$\parallel$

$(p_{ij}^{(n)})$

Corollary 22. *Transition probability in n steps* $p_{ij}^{(n)}$ *satisfy* $p_{ij}^{(n)} \to \pi_j$ *as* $n \to \infty$. *This convergence is exponential, that is* $|p_{ij}^{(n)} - \pi_j| \le C\lambda^n$ *for any* i, j, n *and some* $C > 0$, $0 < \lambda < 1$.

Proof. Let $p^{(0)} = (0, \dots, 0, 1, 0, \dots, 0)$. Then $p^{(n)} = p^{(0)}\Pi^n = (p_{i1}^{(n)}, \dots, p_{iL}^{(n)}) \to \pi$. $\leftarrow i$

$i \xrightarrow{n} j \sim \pi_j$

Letcure 7. Proof of the ergodic theorem. Preliminaries

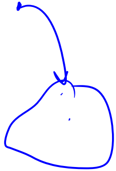
$(v_1, \dots, v_j)$

Notation: For a vector $v \in \mathbb{R}^L$ we denote $\sum_+ v_j := \sum_{j:\, v_j > 0} v_j$.

Lemma 23. *Let $p, q \in \mathcal{P}$. Then $|p - q|_1 = 2\sum_+ (p_j - q_j)$.*

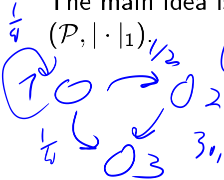
Proof. Since $\sum_{j=1}^{L} (p_j - q_j) = 1 - 1 = 0$, we have $\sum_+ (p_j - q_j) = -\sum_- (p_j - q_j)$. Consequently,

$$|p - q|_1 = \Big(\sum_+ - \sum_-\Big)(p_j - q_j) = 2\sum_+ (p_j - q_j).$$



Proof of the ergodic theorem:

The main idea is to show that the mapping Π^s is a contraction on $(\mathcal{P}, |\cdot|_1)$.



$3_1, 3_1, 3_2$

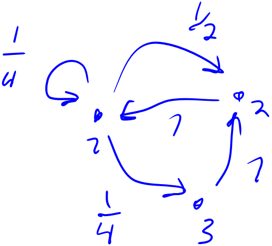
$p^{(0)}, p^{(1)}, p^{(2)} \in P$

$\Pi : P \circlearrowleft$ — лин. оператор

$\mu \to \mu\Pi$

$\Pi \sim B_1^*$

Примеры эргодич. и не эргодич. матриц.



$$\Pi = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

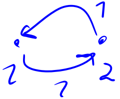
$p_{ij}^{(s)} > 0$

Π^s

Эргодична ли Π?

$p_{22}^{(1)} = 0 \quad s > 1, \quad p_{33}^{(2)} = 0 \Rightarrow s > 2.$

$p_{11}^{(3)} > 0, \quad p_{22}^{(3)} > 0, \quad p_{33}^{(3)} > 0, \quad p_{12}^{(3)} > 0, \ \dots \Rightarrow \Pi$ - эргод.



$$\Pi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

s - неч $\Rightarrow p_{11}^{(s)} = 0$

s - чёт $\Rightarrow p_{12}^{(s)} = 0$