Markov chains

A. Dymov

Steklov Mathematical Institute & HSE

23 октября 2020 г.

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

Definition. Distribution (распределение) $\pi = (\pi_1, \ldots, \pi_L)$ 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 $\xi_0, \ldots, \xi_T$). 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).

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

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

Then $\mathcal{P} \subset \mathbb{R}^L$ is a simplex. In particular, $\mathcal{P}$ is a convex compact set in $\mathbb{R}^L$.
For any vector $v = (v_1, \ldots, v_L)$ 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.

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

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

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.

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

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

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

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

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

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

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

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

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.

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*

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

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

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

⇒ Markov chain "forgets" initial conditions

⇒ Exponential convergence to equilibrium

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* $|p_{ij}^{(n)} - \pi_j| \leq C\lambda^n$ *for any* i, j, n *and some* $C > 0$, $0 < \lambda < 1$.

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

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,

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

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