Probability and Statistics 2

Data Science Engineering

Session 4: Stationary distribution. Ergodic Thoerem

Stationary distribution. Ergodic Chains. Ergodic Theorem

One of the key results in the study of Markov chains shows that, under natural conditions, the probability distribution of X_n tends to a limit regardless of the initial distribution of X_0. This belongs to a class of results in random processes known as Ergodic Theorems.

1. Stationary distribution

Throughout X_n denotes a Markov chain with state space S and transition matrix P.

Definition 1.1 (Stationary distribution). *A probability distribution $\pi = (\pi_i, i \in S)$ on S is said to be* stationary *for the chain X_n with transition matrix P if*

$$\pi P = \pi.$$

If a probability distribution on S is stationary for the Markov chain X_n then

$$\pi P^n = \pi \ \text{ for all } n > 0.$$

Therefore, if X_m has distribution π for some m, then X_{m+n} has the same distribution for all $n \geq 0$. This is the reason why it is called stationary. In particular, if the initial distribution is stationary then each variable X_n in the chain has the same distribution.

From the Linear Algebra point of view, a stationary distribution is a left eigenvector of P with eigenvalue 1. We note that, since a transition matrix P is stochastic (all rows add to one), the vector $(1, 1, \ldots, 1)$ is a right eigenvector of P with eigenvalue 1. Therefore P has 1 as a left eigenvalue 1 (left and right eigenvalues of a square matrix are the same). In order that there is a stationary distribution of the chain with transition matrix P, there should be an eigenvector which is also a probability distribution, namely, all entries must be in $[0, 1]$ and the sum of all entries must be 1.

Example 1.2. Computing a stationary distribution for particular cases amounts to solve a linear system of equations. For example, suppose that the transition matrix of a Markov chain is

$$P = \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{pmatrix}.$$

Then a stationary distribution, if it exists, must satisfy the equation $\pi P = \pi$, namely,

$$(3/4)\pi_1 + (1/2)\pi_2 = \pi_1$$
$$(1/4)\pi_1 + (1/2)\pi_2 = \pi_2,$$

which leads to $\pi_1 = 2\pi_2$. It is worth noting that the property of being stochastic shows that P has not full rank and therefore one of the equations is redundant. Since π must be a probability distribution, we have the additional condition that $\pi_1 + \pi_2 = 1$. It follows that $\pi = (2/3, 1/3)$ is a stationary distribution of the chain. □

Example 1.3. Consider the random walk on a circle of length 5 with transition matrix

$$P = \begin{pmatrix} 0 & p & 0 & 0 & q \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ p & 0 & 0 & q & 0 \end{pmatrix}.$$

One can write and try to solve the linear system $\pi P = \pi$. However the solution here is simpler: P is doubly stochastic and therefore the all-ones vector $(1, 1, 1, 1, 1)$ is a left eigenvector. It follows that the uniform distribution $(1/5, 1/5, 1/5, 1/5, 1/5)$ is a stationary distribution of the chain. □

Example 1.4. Consider the Markov chain with transition matrix

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

By solving the linear system $\pi = \pi P$ one finds two stationary distributions $(1/2, 1/2, 0, 0, 0)$ and $(0, 0, 0, 1/2, 1/2)$. This shows that the stationary distribution of a chain might not be unique. □

Example 1.5. In the previous chapter we considered the Markov chain with state space $S = \{0, 1, 2, \ldots\}$ with transition probabilities $p_{i,i+1} = 2/3$ and $p_{i,0} = 1/3$ for all $i \geq 0$. The state space is infinite and the definition of a stationary distribution as an eigenvector of the transition matrix is not appropriate. However, there is still the notion of stationary distribution as the one that, when reached by some variable in the chain, it remains the same for all future times. By the Chapman–Kolmogorov equations, the stationary distribution should satisfy

$$\pi_0 = \sum_{i\geq 0}(1/3)\pi_i$$
$$\pi_i = (2/3)\pi_{i-1},\ i \geq 1,$$

From the second equation we get $\pi_i = (2/3)^i\pi_0$ and then the first one gives no condition $\pi_0 = \frac{1}{3}(\sum_{i\geq 0}(2/3)^i)\pi_0 = \pi_0$. As it has to be a probability distribution, we have $1 = \sum_{i\geq 0}\pi_i = \pi_0\sum_{i\geq 0}(2/3)^i = 3$, so $\pi_i = \frac{1}{3}(2/3)^i$.

Suppose we change the probabilities to $p_{i,i+1} = 1 - 1/2^i$ and $p_{i,0} = 1/2^i$ for all $i > 0$ and $p_{01} = 1$. Now we have

$$\pi_0 = \sum_{i=1}^{\infty}(1/2^i)\pi_i$$
$$\pi_i = (1/2^{i-1})\pi_{i-1},\ i \geq 1,$$

From the second equation we get $\pi_i = (1/2^{i(i-1)/2})\pi_0$ and then the first one gives $\pi_0 = (\sum_{i=1}^{\infty} 1/2^{i(i+1)/2})\pi_0$. Since the sum is $r < 1$ the only solution is $\pi_0 = 0$ and then $\pi_i = 0$ for all i.

If the chain is irreducible (consists of a single class), then if one state is positive-recurrent, all the states are. Thus, it makes sense to talk about irreducible positive-recurrent chains. Such chains have a stationary distributions.

Theorem 1.6. *An irreducible Markov chain has a stationary distribution π if and only if it is positive recurrent. In this case, the stationary distribution π is unique and given by*

$$\pi_i = \frac{1}{\mathbb{E}(T_i)},$$

where $\mathbb{E}(T_i)$ is the mean return time of the state i.

The first part of the theorem (existence and unicity of stationary distribution) can be obtained in the finite case as a direct consequence of the Perron–Frobenius theorem from the theory of nonnegative matrices in Linear Algebra.

Example 1.7. In Example 1.2, we have seen that if

$$P = \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{pmatrix}.$$

then $\pi = (2/3, 1/3)$ is a stationary distribution. As the chain is irreducible, according to the above theorem, the mean return time of the two states is

$$\mathbb{E}(T_1) = 4/3 \text{ and } \mathbb{E}(T_2) = 2.$$

In the random walk on the circle of length m (Example 1.3), we similarly obtain that the mean return time is m for every state.

The two infinite chains in Example 1.5 are irreducible. The first one has a stationary distribution, so the chain is positive recurrent, as we showed in an exercise from the previous chapter. The second one does not have a stationary distribution and the chain is not positive recurrent. It is not hard to show that in fact, the chain is transient. □

2. Ergodic chains

Stationary distributions are connected to limiting distributions which are of central importance in the theory of random processes.

Proposition 2.1. *Let $\{X_n, n \geq 0\}$ be a Markov chain with a finite space state $S = \{1, 2, \ldots, m\}$. Suppose that, for some $i \in S$,*

$$\lim_{n\to\infty} p_{ij}(n) = \pi_j, \; j = 1, \ldots, m.$$

Then $\pi = (\pi_1, \ldots, \pi_m)$ is a stationary distribution.

Proof. For each j,

$$\pi_j = \lim_{n\to\infty} p_{ij}(n) = \lim_{n\to\infty} \sum_{k=1}^{m} p_{ik}(n-1)p_{kj} = \sum_{k=1}^{m} (\lim_{n\to\infty} p_{ik}(n-1))p_{kj} = \sum_{k=1}^{m} \pi_k p_{kj},$$

so that $\pi P = \pi$. It is clear that π is a probability distribution since all entries are in $[0, 1]$ and

$$\sum_{j=1}^{m} \pi_j = \sum_{j=1}^{m} \lim_{n\to\infty} p_{ij}(n) = \lim_{n\to\infty} \sum_{j=1}^{m} p_{ij}(n) = \lim_{n\to\infty} 1 = 1.$$

□

The converse of the above result, that stationary distributions are limiting ones, is a key result in the theory of Markov chains. Results on the existence of limiting distributions of random processes are known as ergodic theorems. The following examples aim to illustrate the situation.

Example 2.2. *Let $\{X_n, n \geq 0\}$ be a Markov chain with transition matrix*

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then

$$P^n = \begin{cases} P & n \text{ odd} \\ Id & n \text{ even} \end{cases}$$

so there is no limiting distribution. However, $(1/2, 1/2)$ is a stationary distribution. □

Example 2.3. *Let $\{X_n, n \geq 0\}$ be a Markov chain with transition matrix*

$$P = \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{pmatrix}.$$

We can check that

$$P^2 = \begin{pmatrix} 0.687 & 0.3125 \\ 0.625 & 0.375 \end{pmatrix}, \; P^3 = \begin{pmatrix} 0.672 & 0.328 \\ 0.656 & 0.344 \end{pmatrix}, \; P^4 = \begin{pmatrix} 0.668 & 0.332 \\ 0.664 & 0.336 \end{pmatrix}, \dots$$

which suggest a limiting distribution $\pi = (2/3, 1/3)$, which is a stationary distribution as we saw.

The matrix A can be diagonalized,

$$A = B \begin{pmatrix} 1 & 0 \\ 0 & 1/4 \end{pmatrix} B^{-1}$$

for some inversible matrix B and

$$A^n = B \begin{pmatrix} 1 & 0 \\ 0 & 1/4 \end{pmatrix}^n B^T = B \begin{pmatrix} 1 & 0 \\ 0 & (1/4)^n \end{pmatrix} B^{-1},$$

which shows that the limit indeed exists. □

Recall that an irreducible Markov chain is aperiodic if all states in its single class have period one. We also recall that a recurrent state i is positive if the expected time to return is finite: $\mathbb{E}(T_i) < \infty$, and null otherwise.

Definition 2.4 (Ergodic chain)**.** *A Markov chain is* ergodic *if it is irreducible, aperiodic and positive-recurrent.*

If a chain has finite state, every irreducible chain is positive-recurrent so we do not need to check it.

Theorem 2.5 (Ergodic Theorem)**.** *Let $\{X_n, n \geq 0\}$ be an ergodic Markov chain with state space S. Then, for each $i, j \in S$, the limit*

$$\lim_{n\to\infty} p_{ij}(n) = \pi_j$$

exists.

If the chain is not aperiodic, then we do not have convergence, as Example 2.2 showed. Recall that the stationary distribution, can be related to the mean return time as in Theorem 1.6.

3. Exercises and Problems

(1) A Markov chain has transition matrix

$$P = \begin{pmatrix} p_1 & 1-p_1 \\ 1-p_2 & p_2 \end{pmatrix}$$

Determine conditions on p_1 and p_2 so that the chain is

(a) not irreducible.

(b) irreducible but not ergodic.

(c) ergodic.

(2) Consider the random walk with reflecting barriers at states 0 and 3 with diagram

$$\begin{tikzcd} 0 \arrow[r, bend left, "1"] & 1 \arrow[l, bend left, "1/2"] \arrow[r, bend left, "1/2"] & 2 \arrow[l, bend left, "1/2"] \arrow[r, bend left, "1/2"] & 3 \arrow[l, bend left, "1"] \end{tikzcd}$$

Find a stationary distribution. Determine the mean recurrence of the states.

(3) (Lazy chains) Let P be the transition matrix of a irreducible positive recurrent chain. Show that:

$$P_{erg} = \frac{1}{2}\left(Id + P\right)$$

is ergodic and has the same stationary distribution of P.

(4) (Uniform stationary distribution) Let $\{X_n, n \geq 0\}$ Markov chain on $\{1, 2, \ldots, m\}$ with a doubly stochastic transition matrix (each row and each column add to 1). Show that the uniform distribution $\pi = (1/m, 1/m, \ldots, 1/m)$ is stationary.

(5) Let Y_n be a random binary sequence: each Y_i is a Bernoulli random variable $B(1/2)$ and the are all independent. We seek to find the mean time since the first appearance of the pattern 010. For this we build a Markov chain with states

$$0, 1, 00, 01, 10, 100, 010, 001, 110, 101, 011, 111$$

where X_n is the outcome of the last three digits in $Y_1, ..., Y_n$ if $n \geq 3$ or the last two when $n = 2$ or the first one if $n = 1$.

(a) Draw a diagram of the chain. Is it irreducible? aperiodic? recurrent?

(b) Find the limiting probability of the state 010. Compute the mean recurrence time of the state 010.

(c) Use the above to find the expected time to see the pattern 010 for the first time.

(6) (Card shuffling) Use the exercise on doubly stochastic matrices to show that the top-to-random and riffle shuffling defined in the exercises of Chapter 2 have uniform stationary distributions. They give a correct shuffling method!

(7) (Positivity of stationary distribution) In an irreducible chain with stationary distribution π, show that $\pi_i > 0$ for every $i \in S$.

(8) (Random walks on graphs) Consider a graph $G = (V, E)$, where V is the set of vertices and E the set of edges. A random walk on G starting at v is a Markov chain X_n with

$X_0 = v$ and transitions defined by

$$\Pr(X_n = w | X_{n-1} = u) = \begin{cases} \frac{1}{\deg(u)}, & uw \in E \\ 0, & \text{otherwise} \end{cases},$$

Assuming that the graph is connected, find the stationary distribution of X_n.

(9) Suppose we have a knight in a corner of a chess board. At each step, the knight chooses one of the allowed moves uniformly at random and executes it. Compute the expected time it will take the knight to return to the starting cell. Is there another initial cell that would reduce the expected return time?

(10) Let π be a probability distribution that satisfies for all $i, j \in S$

$$p_{ij}\pi_i = p_{ji}\pi_j. \qquad (*)$$

Prove that π is stationary.

Give an example of a Markov chain with stationary distribution that does not satisfy $(*)$.

(11) (Ehrenfest urn) Recall the definition of the Ehrenfest urn in Chapter 2. Show that its stationary distribution is $\pi = (\pi_0, \dots, \pi_n)$ with $\pi_i = \binom{m}{i} 2^{-m}$. Is the chain ergodic?

(12) Consider a homogeneous Markov chain on states $\{0, 1, 2, \dots\}$ with transition matrix

$$P = \begin{pmatrix} p_0 & p_1 & p_2 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

for some $p_m > 0$ such that $\sum_{m \geq 1} p_m = 1$.

(a) Show that the chain is recurrent.
(b) What is the condition on the values p_m so it is null-recurrent?
(c) Apply the condition in (b) to determine if the origin is null-recurrent in each distribution
  i) $p_0 = 0$ and $p_m = \frac{6}{(\pi m)^2}$ for $m \geq 1$.
  ii) $p_0 = 0$ and $p_m = 2^{-m}$ for $m \geq 1$.
(d) Find the stationary distribution for the distribution (c.ii).
(Hint: $\sum_{m \geq 1} m x^{m-1} = \frac{1}{(1-x)^2}$.)

(13) (Random walks on digraphs) Let $\vec{G} = (V, \vec{E})$ be a directed graph, so $\vec{E}$ is a set of arcs (*ordered* pairs of elements in V). Suppose that for every $x \in V$, there are $d(x)$ out-going arcs from x and $d(x)$ in-going arcs to x (*Eulerian digraph*). Consider the random walk on $\vec{G}$, where at each step we traverse a random out-going arc from the current state. Show that $\pi(x) = \frac{d(x)}{|\vec{E}|}$ is a stationary distribution.