Лекция 4

Конечномерные распределения марковских цепей

Как известно, и случайное блуждание, и процесс восстановления определяются с помощью некоторой последовательности $X_1, X_2, \dots$ независимых одинаково распределенных случайных величин. В начале двадцатого века русский математик А.А. Марков рассмотрел простое обобщение этой последовательности. В новой последовательности случайные величины $\xi_0, \xi_1, \xi_2, \dots$ являются зависимыми, но степень этой зависимости является наиболее слабой из возможных.

Пусть S – некоторое конечное или счетное множество. Элементы S всегда можно пронумеровать, поэтому можно считать, что или $S = \{1, 2, \dots, m\}$, где $m \in \mathbb{N}$, или $S = \mathbb{N}_0$. Рассмотрим последовательность случайных величин $\xi_0, \xi_1, \dots$, принимающих значения в S. По теореме произведения вероятностей при $n \in \mathbb{N}$ и $i_0, \dots, i_n \in S$

$$\mathbf{P}\left(\xi_0 = i_0, \dots, \xi_n = i_n\right) =$$

$$= \mathbf{P}\left(\xi_0 = i_0\right) \prod_{k=0}^{n-1} \mathbf{P}\left(\xi_{k+1} = i_{k+1} \mid \xi_0 = i_0, \dots, \xi_k = i_k\right) \tag{1}$$

(считаем, что все вероятности в правой части определены). В ситуации, когда случайные величины $\xi_0, \xi_1, \dots$ независимы,

$$\mathbf{P}\left(\xi_{k+1} = i_{k+1} \mid \xi_0 = i_0, \dots, \xi_k = i_k\right) = \mathbf{P}\left(\xi_{k+1} = i_{k+1}\right)$$

и формула (1) принимает вид

$$\mathbf{P}\left(\xi_0 = i_0, \dots, \xi_n = i_n\right) = \prod_{k=0}^{n} \mathbf{P}\left(\xi_k = i_k\right).$$

Предположим теперь, что случайная величина ξ_{k+1} зависит от вектора $(\xi_0, \dots, \xi_k)$, но при этом

$$\mathbf{P}\left(\xi_{k+1} = i_{k+1} \mid \xi_0 = i_0, \dots, \xi_k = i_k\right) = \mathbf{P}\left(\xi_{k+1} = i_{k+1} \mid \xi_k = i_k\right) \tag{2}$$

при произвольном $k \in \mathbb{N}_0$ и произвольных $i_0, \dots, i_k \in S$. Если трактовать момент k как "настоящее", моменты, предшествующие k как "прошлое", а моменты, следующие за k, как "будущее", то соотношение (2) говорит о том, что при фиксированном прошлом и настоящем распределение будущего марковской цепи зависит только от настоящего.

Определение 1. *Цепью Маркова* называется последовательность случайных величин $\xi_0, \xi_1, \dots$, принимающих значения в S и удовлетворяющих соотношению (2).

Множество S называется *множеством состояний* марковской цепи, а его элементы – *состояниями*. Марковская цепь называется *конечной* (или *счетной*), если конечно (или счетно) множество S. При $k \in \mathbb{N}$ и $i, j \in S$ положим

$$p_{ij}^{(k)} := \mathbf{P}\left(\xi_k = j \mid \xi_{k-1} = i\right).$$

(предполагается, что $\mathbf{P}\left(\xi_{k-1} = i\right) > 0$). Говорят, что $p_{ij}^{(k)}$ – *переходная вероятность* из состояния i в состояние j на k-ом шаге марковской цепи. Положим $p_i := \mathbf{P}(\xi_0 = i)$, $i \in S$. Совокупность $\{p_i,\ i \in S\}$ называется *начальным распределением* цепи Маркова.

Запишем формулу для конечномерных распределений марковской цепи через переходные вероятности и начальное распределение, используя соотношение (1): при $n \in \mathbb{N}$ и $i_0, \ldots, i_n \in S$

$$\mathbf{P}(\xi_0 = i_0, \ldots, \xi_n = i_n) = p_{i_0} \prod_{k=1}^{n} p_{i_{k-1} i_k}^{(k)}.$$

В дальнейшем рассматриваются *однородные* цепи Маркова, в которых переходные вероятности $p_{ij}^{(k)}$ не зависят от k. Для однородных марковских цепей последняя формула принимает следующий простой вид:

$$\mathbf{P}(\xi_0 = i_0, \ldots, \xi_n = i_n) = p_{i_0} \prod_{k=1}^{n} p_{i_{k-1} i_k} \qquad (3)$$

при произвольном $n \in \mathbb{N}$ и произвольных $i_0, \ldots, i_n \in S$. Заметим, что

$$p_i, p_{ij} \geq 0, \quad \sum_{i \in S} p_i = 1, \ \sum_{j \in S} p_{ij} = 1, \quad i, j \in S. \qquad (4)$$

Соотношения (3) и (4) являются основой для всех последующих рассуждений. Кстати, однородная цепь Маркова может быть определена как последовательность случайных величин $\xi_0, \xi_1, \ldots$, принимающих значения в S и удовлетворяющих соотношению (3) для некоторых чисел p_i, p_{ij}, $i, j \in S$, удовлетворяющих соотношению (4).

Формула (3) хорошо объясняет, почему именно слово "цепь" использовано в определении: "сцепление" элементов $i_0, \ldots, i_n$ осуществляется через пары $(i_0, i_1), (i_1, i_2), \ldots, (i_{n-1}, i_n)$, которые являются "звеньями" цепи.

Лемма 1. *Если $\{\xi_n,\ n \in \mathbb{N}_0\}$ – марковская цепь и $\mathbf{P}(\xi_l = i_0) > 0$ при некоторых $l \in \mathbb{N}_0$ и $i_0 \in S$, то для произвольных $n \in \mathbb{N}$ и $i_1, \ldots, i_n \in S$*

$$\mathbf{P}\left(\xi_{l+1} = i_1, \ldots, \xi_{l+n} = i_n \mid \xi_l = i_0\right) = \prod_{k=1}^{n} p_{i_{k-1} i_k}, \qquad (5)$$

т.е. указанная вероятность не зависит от l.

Доказательство. При $l = 0$ утверждение леммы следует из (3) в результате деления на $\mathbf{P}(\xi_0 = i_0) = p_{i_0}$. Пусть $l \geq 1$, тогда ввиду формулы (3)

$$\mathbf{P}\left(\xi_l = i_0, \xi_{l+1} = i_1, \ldots, \xi_{l+n} = i_n\right) =$$

$$= \sum_{j_0 \in S} \cdots \sum_{j_{l-1} \in S} \mathbf{P}\left(\xi_0 = j_0, \ldots, \xi_{l-1} = j_{l-1}, \xi_l = i_0, \xi_{l+1} = i_1, \ldots, \xi_{l+n} = i_n\right) =$$

$$= \sum_{j_0 \in S} \cdots \sum_{j_{l-1} \in S} p_{j_0} \left(\prod_{k=1}^{l-1} p_{j_{k-1} j_k} \right) p_{j_{l-1} i_0} \left(\prod_{k=1}^{n} p_{i_{k-1} i_k} \right) =$$

$$= \left(\prod_{k=1}^{n} p_{i_{k-1} i_k} \right) \sum_{j_0 \in S} \cdots \sum_{j_{l-1} \in S} p_{j_0} \left(\prod_{k=1}^{l-1} p_{j_{k-1} j_k} \right) p_{j_{l-1} i_0} =$$

$$= \left(\prod_{k=1}^{n} p_{i_{k-1} i_k} \right) \mathbf{P} \left(\xi_l = i_0 \right).$$

Откуда делением на $\mathbf{P}(\xi_l = i_0)$ получаем (5) при $l \geq 1$. Лемма доказана.

Следующее утверждение означает, что *при фиксированном прошлом и настоящем распределение будущего марковской цепи зависит только от настоящего.*

Теорема 1. *Если* $\{\xi_n,\ n \in \mathbb{N}_0\}$ *– марковская цепь, то для произвольных натуральных чисел* l и n, *произвольных множеств* $A \subset S^l$ и $B \subset S^n$ *и произвольного состояния* i

$$\mathbf{P} \left(\left(\xi_{l+1}, \ldots, \xi_{l+n} \right) \in B \;\middle|\; \left(\xi_0, \ldots, \xi_{l-1} \right) \in A,\ \xi_l = i \right) =$$

$$= \mathbf{P} \left(\left(\xi_{l+1}, \ldots, \xi_{l+n} \right) \in B \;\middle|\; \xi_l = i \right)$$

(*предполагается, что* $\mathbf{P}\left(\left(\xi_0, \ldots, \xi_{l-1} \right) \in A,\ \xi_l = i \right) > 0$).

Доказательство. В силу соотношения (3) и леммы 1 при $i_0, \ldots, i_{l+n} \in S$

$$\mathbf{P} \left(\xi_0 = i_0, \ldots, \xi_{l+n} = i_{l+n} \right) = p_{i_0} \prod_{k=1}^{l+n} p_{i_{k-1} i_k} = p_{i_0} \prod_{k=1}^{l} p_{i_{k-1} i_k} \prod_{k=l+1}^{l+n} p_{i_{k-1} i_k} =$$

$$= \mathbf{P} \left(\xi_0 = i_0, \ldots, \xi_l = i_l \right) \mathbf{P} \left(\xi_{l+1} = i_{l+1}, \ldots, \xi_{l+n} = i_{l+n} \;\middle|\; \xi_l = i_l \right).$$

Полагая $i_l = i$ и суммируя это равенство по всем $(i_0, \ldots, i_{l-1}) \in A$ и всем $(i_{l+1}, \ldots, i_{l+n}) \in B$, находим, что

$$\mathbf{P} \left(\left(\xi_0, \ldots, \xi_{l-1} \right) \in A,\ \xi_l = i,\ \left(\xi_{l+1}, \ldots, \xi_{l+n} \right) \in B \right) =$$

$$= \mathbf{P} \left(\left(\xi_0, \ldots, \xi_{l-1} \right) \in A,\ \xi_l = i \right) \mathbf{P} \left(\left(\xi_{l+1}, \ldots, \xi_{l+n} \right) \in B \;\middle|\; \xi_l = i \right).$$

Разделив последнее равенство на $\mathbf{P}\left(\left(\xi_0, \ldots, \xi_{l-1} \right) \in A,\ \xi_l = i \right)$, получаем утверждение теоремы.

Наряду с переходными вероятностями за один шаг рассмотрим *вероятность перехода за* n *шагов* из состояния $i \in S$ в состояние $j \in S$

$$p_{ij}(n) := \mathbf{P} \left(\xi_n = j \mid \xi_0 = i \right), \qquad n \in \mathbb{N}$$

(предполагается, что $\mathbf{P}(\xi_0 = i) > 0$). Отметим, что при каждом $l \in \mathbb{N}_0$

$$\mathbf{P} \left(\xi_{l+n} = j \;\middle|\; \xi_l = i \right) = \mathbf{P} \left(\xi_n = j \mid \xi_0 = i \right)$$

и, следовательно,

$$p_{ij}(n) = \mathbf{P} \left(\xi_{l+n} = j \;\middle|\; \xi_l = i \right) \tag{6}$$

(предполагается, что $\mathbf{P}(\xi_l = i) > 0$). Действительно, в силу леммы 1

$$\mathbf{P}\left(\xi_{l+n} = j \mid \xi_l = i\right) =$$

$$= \sum_{i_1,\ldots,i_{n-1}\in S} \mathbf{P}\left(\xi_{l+1} = i_1, \ldots, \xi_{l+n-1} = i_{n-1}, \xi_{l+n} = j \mid \xi_l = i\right) =$$

$$= \sum_{i_1,\ldots,i_{n-1}\in S} \mathbf{P}\left(\xi_1 = i_1, \ldots, \xi_{n-1} = i_{n-1}, \xi_n = j \mid \xi_0 = i\right) =$$

$$= \mathbf{P}\left(\xi_n = j \mid \xi_0 = i\right).$$

Следующее утверждение указывает, как находить конечномерные распределения марковской цепи.

Теорема 2. *Если $\{\xi_n,\ n \in \mathbb{N}_0\}$ – марковская цепь, то для произвольного $r \in \mathbb{N}_0$, произвольных состояний $i_0, i_1, \ldots, i_r$ и произвольных моментов времени $n_1, \ldots, n_r \in \mathbb{N}$* $(0 = n_0 < n_1 < n_2 < \ldots < n_r)$

$$\mathbf{P}\left(\xi_{n_0} = i_0, \xi_{n_1} = i_1, \ldots, \xi_{n_r} = i_r\right) = p_{i_0} \prod_{k=1}^{r} p_{i_{k-1}i_k}\left(n_k - n_{k-1}\right). \tag{7}$$

Доказательство. По теореме умножения вероятностей

$$\mathbf{P}\left(\xi_{n_0} = i_0, \xi_{n_1} = i_1, \ldots, \xi_{n_r} = i_r\right) =$$

$$= \mathbf{P}\left(\xi_{n_0} = i_0\right) \prod_{k=1}^{r} \mathbf{P}\left(\xi_{n_k} = i_k \mid \xi_{n_0} = i_0, \xi_{n_1} = i_1, \ldots, \xi_{n_{k-1}} = i_{k-1}\right).$$

Осталось заметить, что в силу теоремы 1 и соотношения (6)

$$\mathbf{P}\left(\xi_{n_k} = i_k \mid \xi_{n_0} = i_0, \xi_{n_1} = i_1, \ldots, \xi_{n_{k-1}} = i_{k-1}\right) =$$

$$= \mathbf{P}\left(\xi_{n_k} = i_k \mid \xi_{n_{k-1}} = i_{k-1}\right) = p_{i_{k-1}i_k}\left(n_k - n_{k-1}\right).$$

Теорема доказана.

Положим 1) $p_{ii}(0) = 1$ при $i \in S$ и 2) $p_{ij}(0) = 0$ при разных $i, j \in S$.

Следствие 1. *Пусть $i, j \in S$ и $n \in \mathbb{N}$. Для каждого целого l, такого что $0 \le l \le n$, справедливо равенство*

$$p_{ij}(n) = \sum_{k\in S} p_{ik}(l)\, p_{kj}(n - l). \tag{8}$$

Доказательство. События $\{\xi_l = k\}$ при разных $k \in S$ попарно несовместны, поэтому

$$\mathbf{P}\left(\xi_n = j \mid \xi_0 = i\right) = \sum_{k\in S} \mathbf{P}\left(\xi_l = k,\ \xi_n = j \mid \xi_0 = i\right).$$

Осталось учесть, что ввиду соотношения (7) при $r = 2$

$$\mathbf{P}\left(\xi_l = k,\ \xi_n = j \mid \xi_0 = i\right) = p_{ik}(l)\, p_{kj}(n - l).$$

Утверждение доказано.

Замечание 1. Соотношение (9) называется *уравнением Колмогорова-Чепмена.* Оно позволяет вычислять переходные вероятности за n шагов, зная переходные вероятности за меньшее число шагов.

Переходные вероятности удобно представлять в матричном виде. Пусть марковская цепь $\{\xi_n, \, n \in \mathbb{N}_0\}$ является конечной и $S = \{1, \ldots, m\}$, где $m \in \mathbb{N}$. Назовем *матрицей переходных вероятностей* матрицу P размера $m \times m$ с элементами p_{ij}, $i, j \in \{1, \ldots, m\}$, т.е.

$$P = \left(\begin{array}{cccc} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{array} \right).$$

Элементы этой матрицы неотрицательны и сумма элементов каждой строки равна 1. Такие матрицы называются *стохастическими.*

В случае, когда $S = \mathbb{N}_0$, матрица переходных вероятностей определяется аналогично (она в этом случае будет бесконечной).

Аналогично матрице переходных вероятностей введем для каждого $n \in \mathbb{N}_0$ матрицу переходных вероятностей за n шагов $P(n)$ (в случае, когда $S = \{1, \ldots, m\}$, это матрица размера $m \times m$ элементами которой являтся числа $p_{ij}(n)$, $i, j \in \{1, \ldots, m\}$). Заметим, что эта матрица также является стохастической. В силу уравнения Колмогорова-Чепмена при $n \in \mathbb{N}$

$$P(n) = P(n-1) P = P(n-2) P^2 = P^n.$$

Рассмотрим примеры марковских цепей.

а) *Случайное блуждание.* Пусть $X_1, X_2, \ldots$ – последовательность независимых одинаково распределенных случайных величин, принимающих два значения: $\mathbf{P}(X_1 = 1) = p$, $\mathbf{P}(X_1 = -1) = q$, $p + q = 1$. Положим $S_0 = 0$, $S_n = \sum_{i=1}^{n} X_i$, $n \in \mathbb{N}$. Последовательность $\{S_n, \, n \in \mathbb{N}_0\}$ называется *простым* случайным блужданием. Множеством состояний является $\mathbb{Z}$. Покажем, что $\{S_n, \, n \in \mathbb{N}_0\}$ образует марковскую цепь. Действительно, пусть $i_0 = 0, i_1, \ldots, i_n$ – возможный вариант траектории случайного блуждания $S_0, S_1, \ldots, S_n$, тогда

$$\mathbf{P}(S_0 = i_0, S_1 = i_1, \ldots, S_n = i_n) = \mathbf{P}(X_k = i_k - i_{k-1}, \; k \in \{1, \ldots, n\}) =$$

$$= \prod_{k=1}^{n} \mathbf{P}(X_1 = i_k - i_{k-1}).$$

Это означает, что $\{S_n, \, n \in \mathbb{N}_0\}$ образует цепь Маркова с начальным распределением $p_0 = 1$, $p_1 = p_2 = \ldots = 0$ и переходными вероятностями $p_{ij} = \mathbf{P}(X_1 = j - i)$, $i, j \in \mathbb{Z}$, т.е.

$$p_{ij} = \begin{cases} p, & j = i + 1; \\ q, & j = i - 1; \\ 0, & j \notin \{i-1, i+1\}. \end{cases} \tag{9}$$

Таким образом, марковская цепь (иногда говорят "движущаяся частица") совершает переходы только в соседние состояния, причем переход направо совершается с вероятностью p, а налево – с вероятностью q.

б) *Случайное блуждание с отражением* в точке 0. В отличие от прошлого примера предположим, что частица стартует из точки $l \in \mathbb{N}$. Переходные вероятности p_{ij} при $i \in \mathbb{N}$ имеют вид (9). Но попав в состояние 0, частица в следующий целый момент времени либо переходит в состояние 1 с вероятностью p, либо остается в состоянии 0 с вероятностью q. Следовательно,

$$p_{0j} = \begin{cases} p, & j = 1; \\ q, & j = 0; \\ 0, & j \notin \{0, 1\}. \end{cases}$$

Множеством возможных состояний является $\mathbb{N}_0$. Матрица переходных вероятностей имеет вид

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

(первым состоянием в $\mathbb{N}_0$ является нуль, вторым – единица и т.д.).

в) *Случайное блуждание с поглощением* в точке 0. В отличие от прошлого примера частица, попав в состояние 0, остается там навсегда. Итак, переходные вероятности p_{ij} при $i \in \mathbb{N}$ имеют вид (9), но $p_{00} = 1$ и $p_{0j} = 0$ при $j \neq 0$. Множеством возможных состояний является $\mathbb{N}_0$. Матрица переходных вероятностей имеет вид

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

г) *Стохастически рекурсивная последовательность.* Пусть $X_1, X_2, \dots$ – последовательность независимых одинаково распределенных случайных величин. Пусть или $S = \{1, 2, \dots, m\}$, где $m \in \mathbb{N}$, или $S = \mathbb{N}_0$. Пусть f – измеримое отображение $S \times \mathbb{R}$ в S. Положим при $n \in \mathbb{N}_0$

$$\xi_{n+1} = f(\xi_n, X_{n+1}),$$

при этом будем считать, что ξ_0 является случайной величиной, принимающей значения в S, причем ξ_0 и последовательность $\{X_n\}$ независимы.

Оказывается, последовательность $\{\xi_n,\ n \in \mathbb{N}_0\}$ является марковской цепью. Действительно, при $i_0, \dots, i_{k+1} \in S$, $k \in \mathbb{N}_0$

$$\mathbf{P}\left(\xi_{k+1} = i_{k+1} \mid \xi_0 = i_0, \dots, \xi_k = i_k\right) =$$

$$= \mathbf{P}\left(f(\xi_k, X_{k+1}) = i_{k+1} \mid \xi_0 = i_0, \dots, \xi_k = i_k\right) =$$

$$= \mathbf{P}\left(f(i_k, X_{k+1}) = i_{k+1} \mid \xi_0 = i_0, \dots, \xi_k = i_k\right).$$

Произошло или нет событие $\{\xi_0 = i_0, \dots, \xi_k = i_k\}$, целиком и полностью определяется значениями случайных величин $\xi_0, X_1, \dots, X_k$, поэтому это событие не зависит от случайной величины X_{k+1}. Следовательно,

$$\mathbf{P}\left(f(i_k, X_{k+1}) = i_{k+1} \mid \xi_0 = i_0, \dots, \xi_k = i_k\right) = \mathbf{P}\left(f(i_k, X_{k+1}) = i_{k+1}\right).$$

Итак,

$$\mathbf{P}\left(\xi_{k+1} = i_{k+1} \;\middle|\; \xi_0 = i_0, \ldots, \xi_k = i_k\right) = \mathbf{P}\left(f\left(i_k, X_{k+1}\right) = i_{k+1}\right). \tag{10}$$

Аналогично

$$\mathbf{P}\left(\xi_{k+1} = i_{k+1} \;\middle|\; \xi_k = i_k\right) = \mathbf{P}\left(f\left(\xi_k, X_{k+1}\right) = i_{k+1} \mid \xi_k = i_k\right) =$$

$$= \mathbf{P}\left(f\left(i_k, X_{k+1}\right) = i_{k+1} \mid \xi_k = i_k\right) = \mathbf{P}\left(f\left(i_k, X_{k+1}\right) = i_{k+1}\right). \tag{11}$$

Из соотношений (10), (11) следует, что

$$\mathbf{P}\left(\xi_{k+1} = i_{k+1} \;\middle|\; \xi_0 = i_0, \ldots, \xi_k = i_k\right) = \mathbf{P}\left(\xi_{k+1} = i_{k+1} \;\middle|\; \xi_k = i_k\right),$$

что и требовалось доказать.

Задача. Представьте марковские цепи из примеров а), б) и в) в виде стохастически рекурсивных последовательностей.