Лекция № 3

Цепи Маркова

Понятие марковской цепи принадлежит русскому математику А.А.Маркову, чьи первые статьи по этому вопросу при решении лингвистических проблем были опубликованы в 1906-1908 гг.

1. Случайные процессы

Определение 2. Пусть для некоторого опыта задано вероятностное пространство $<\Omega, A, P>$, где Ω — пространство элементарных событий, A — алгебра его подмножеств, P — вероятностная мера на A. Случайным процессом $\xi(t)$, заданным на данном вероятностном пространстве, называется измеримая функция двух переменных $\xi(t,\omega)$, где $\omega\in \Omega$, а t — действительная переменная ($t\in R$), которая часто имеет смысл времени [2].

При фиксированном значении $t = t_i$ случайный процесс представляет собой измеримую функцию $\xi_i =\xi_i(\omega)$, т.е. случайную величину.

При фиксированном элементарном событии ω_i получаем некоторую детерминированную (неслучайную) функцию $x_i(t)$, называемую реализацией (траекторией) случайного процесса.

Случайный процесс можно задавать или как множество реализаций с заданной на нем вероятностной мерой, или как последовательность (упорядоченную совокупность) случайных величин, соответствующих определённым значениям t. В последнем случае его можно рассматривать как случайный вектор и задать с помощью многомерных законов распределения.

Определение 4. Случайный процесс называется *однородным*, если условные вероятности или условные плотности вероятностей зависят не от моментов времени, а от разности моментов времени, т.е.

$$P(x_2,t_2|x_1,t_1) = P(x_2,x_1,t_2-t_1),$$
$$f(x_2,t_2|x_1,t_1) = f(x_2,x_1,t_2-t_1).$$

2. Классификация случайных процессов

Строгой классификации случайных процессов нет, поэтому можно говорить лишь о выделении по тому или иному признаку типов процессов, которые не обязательно в своей совокупности исчерпывают всевозможные типы и не являются несовместимыми друг с другом.

Случайные процессы можно классифицировать по

1) характеру реализаций случайных процессов (характеру пространства состояний случайного процесса и параметра t);
2) виду закона распределения вероятностей;
3) характеру статистической связи между значениями случайного процесса в различные моменты времени.

Классификация по характеру реализаций:

1. *Дискретная последовательность* (дискретный процесс с дискретным временем) — это случайный процесс, у которого область определения и область возможных значений реализаций — дискретные множества. Примеры: процессы в цифровых системах связи, компьютерных сетях, цифровой радиоаппаратуре и т.п.
2. *Случайная последовательность*, или временной ряд (непрерывнозначный процесс с дискретным временем) — это случайный процесс, область возможных значений реализаций которого — непрерывное множество, а область определения — дискретное множество. Примеры: метеорологические наблюдения, телеметрические данные состояния космонавтов и т.п.
3. *Дискретные процессы* (дискретный процесс с непрерывным временем) — это случайный процесс, множество возможных значений реализаций которого — дискретное множество, а область определения — непрерывное множество. Примеры: число абонентов телефонной станции, разговаривающих по телефону, количество автомобилей на автозаправочной стации и т.п.
4. *Непрерывнозначный* случайный процесс — это случайный процесс, у которого область возможных значений и область определения — непрерывные множества. Примеры: различные физические, химические, биологические процесс протекающие в природе, организме человека.

Замечание 3. Случайные процессы с дискретным множеством возможных значений (типы 1 и 3) называются цепями (последовательно переходят от одного состояния к другому, образуя цепочку состояний).

Если рассматривать классификацию случайных процессов по характеру статистической связи между значениями в отдельные моменты времени, можно выделить наиболее простой и хорошо изученный тип — марковский процесс

3. Марковские случайные процессы

Марковский случайный процесс — такой случайный процесс, эволюция которого после любого фиксированного момента t (в будущем) и до момента t (в прошлом) являются независимыми при известном состоянии в момент t (в настоящем) [2]. Это основное свойство марковского процесса, которое можно математически записать по-разному.

Определение 5. Случайный процесс $\xi(t)$ называется *марковским*, если для любых моментов времени, связанных условием $t_k < t_j < t_i$, справедливо соотношение

$$P\left(\xi(t_k) < x_k, \xi(t_i) < x_i \middle| \xi(t_j) = x_j\right) = P\left(\xi(t_k) < x_k \middle| \xi(t_j) = x_j\right) P\left(\xi(t_i) < x_i \middle| \xi(t_j) = x_j\right). \quad (7)$$

Для дискретного случайного процесса можно записать

$$P\left(\xi(t_k) = x_k, \xi(t_i) = x_i \middle| \xi(t_j) = x_j\right) = P\left(\xi(t_k) = x_k \middle| \xi(t_j) = x_j\right) P\left(\xi(t_i) = x_i \middle| \xi(t_j) = x_j\right). \quad (8)$$

Можно дать эквивалентное определение марковского процесса в несколько иной математической форме.

Определение 6. Случайный процесс $\xi(t)$ называется марковским, если

$$P(\xi(t_n) < x_n | \xi(t_1) = x_1, \dots, \xi(t_{n-1}) = x_{n-1}) = P(\xi(t_n) < x_n | \xi(t_{n-1}) = x_{n-1}). \quad (9)$$

Для дискретного случайного процесса имеем

$$P(\xi(t_n) = x_n | \xi(t_1) = x_1, \dots, \xi(t_{n-1}) = x_{n-1}) = P(\xi(t_n) = x_n | \xi(t_{n-1}) = x_{n-1}). \quad (10)$$

В обширном классе марковских случайных процессов можно выделить различные типы по характеру реализаций:

1. Дискретная последовательность (цепь Маркова).
2. Случайная (марковская) последовательность.
3. Дискретный случайный процесс (дискретный марковский процесс).
4. Непрерывнозначный случайный процесс (непрерывнозначный марковский процесс).

В теории массового обслуживания наиболее часто используются марковские цепи и дискретные марковские процессы, последние иногда называют марковскими цепями с непрерывным временем.

4. Цепи Маркова

Определение 7. Цепь Маркова — это марковский случайный процесс с дискретными множествами возможных значений (состояний цепи) $E_1, \dots, E_n$ и значений аргумента $t_0, t_1, t_2, t_3, \dots$.

Если число возможных состояний n конечно, то цепь называется конечной.

Вместо значений аргумента можно указывать их номер. Разность между двумя соседними значениями аргумента $t_{k+1} - t_k$, называется шагом.

Цепь Маркова задается множеством значений $(E_1, \dots, E_n)$ и вероятностями:

1. Начальные вероятности $P_j^0 = P(\xi(0) = E_j)$, которые удовлетворяют условию нормировки $\sum_j P_j^0 = 1$.

2. $P(\xi(n+1) = E_j | \xi(n) = E_i)$ — вероятность перехода из одного состояния в другое за один шаг. Если марковская цепь однородна, то $P\left(\xi(n+1) = E_j \middle| \xi(n) = E_i\right) = P_{ij}$. Условие нормировки $\sum_j P_{ij} = 1$.

3. Вероятность перехода из одного состояния в другое за k шагов $P\left(\xi(n+k) = E_j \middle| \xi(n) = E_i\right)$. Если марковская цепь однородна, то $P\left(\xi(n+k) = E_j \middle| \xi(n) = E_i\right) = P_{ij}(k)$. Условие нормировки $\sum_j P_{ij}(k) = 1$.

4. Вероятность состояния E_i в k-й момент времени:

$P(\xi(k) = E_j) = P_j(k)$. Условие нормировки $\sum_j P_j(k) = 1$.

Если цепь Маркова конечна, то

- Вероятности состояния в начальный момент времени могут быть представлены в виде вектора-строки $\Pi = \{P_j^0\}$.
- Вероятности состояний на k-том шаге образуют вектор-строку $\Pi\{k\} = \{P_j(k)\}$.
- Вероятности перехода из одного состояния в другое за один шаг или k шагов для однородной цепи представляются соответственно в виде квадратных матриц: $P = \{P_{ij}\}$ или $P = \{P_{ij}(k)\}$. У этих матриц сумма элементов в каждой строке равна 1: $\sum_j P_{ij} = 1$, $\sum_j P_{ij}(k) = 1$.

Для однородных цепей Маркова вероятность P_{ij} перехода системы из состояния E_i в состояние E_j за один шаг зависит только от того, из какого состояния в какое осуществлялся переход.

Если число состояний цепи ограничено, то цепь Маркова удобно представлять в виде графа, вершины которого — возможные состояния. А дуги имеют направление и вес, соответствующие вероятностям перехода из одного состояния в другое за один шаг.

5. Классификация состояний марковской цепи

Для удобства обозначим i-ое состояние его номером: $E_i = i$.

1. Исходное состояние. Состояние i называется исходным, если $P_j^0 = 0$, $P_i^0 = 0$, $i \neq j$.
2. Поглощающее состояние. Состояние i называется поглощающим, если $P_{ii} = 1$, $P_{ij} = 0$, $i \neq j$. Если цепь Маркова содержит хотя бы одно поглощающее состояние, то такая цепь называется поглощающей.
3. Достижимое состояние. Состояние j называется достижимым из состояния i (обозначение: $i \to j$), если существует такое m, что $P_{ij}(m) = 0$.
4. Сообщающееся состояние. Состояние называется сообщающимся (обозначение: $i \leftrightarrow j$), если $i \to j$, $j \to i$.
5. Несущественное состояние. Состояние i называется несущественным, если $\exists j : i \to j,\ j \nrightarrow i$. Все остальные состояния называются существенными.

Множество всех существенных состояний разбивается на непересекающиеся классы сообщающихся состояний так, что любые два состояния из одного класса сообщаются между собой.

Если из множества всех состояний марковской цепи *исключить несущественное*, то оставшиеся будут сообщающимися состояниями. Они могут составлять один класс сообщающихся состояний или несколько классов состоящих из сообщающихся внутри класса состояний и имеющие нулевую вероятность перехода из одного класса в другой. (матрицы разбивается на блоки — подматрицы).

Несущественные состояния не играют роли при изучении долговременного поведения цепи Маркова, а потому их чаще всего игнорируют.

Если цепь Маркова содержит один неразложимый класс существенных состояний, то такая цепь называется *неразложимой*.

Если цепь состоит из нескольких классов существенных по отдельности, то каждый класс анализируется в отдельности.

6. Периодическое состояние. Состояние i называется периодическим с периодом d_i, если возвращение в это состояние возможно за число шагов кратное d_i, $P_{ii}(m) > 0$, $m = kd_i$.
7. Апериодическое состояние. Состояние, для которого $d_i = 1$.

Пример 3. Пусть марковская цепь имеет состояния $E = (1,2,3,4)$,

$$P = \begin{pmatrix} 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{pmatrix}.$$

Граф данной цепи представлен на рис. 3. Все состояния существенные, поскольку $P_{ii}(n) > 0$, $n = 2m$, и имеют один и тот же период равный двум $d_i = 2$. Рассматриваемая цепь имеет особенность:

если выделить два подкласса $C_0 = (1,2)$ и $C_1 = (3,4)$, то за один шаг цепь переходит из одного класса состояний, в какое-либо состояние другого класса. C_0, C_1 — циклические подклассы.

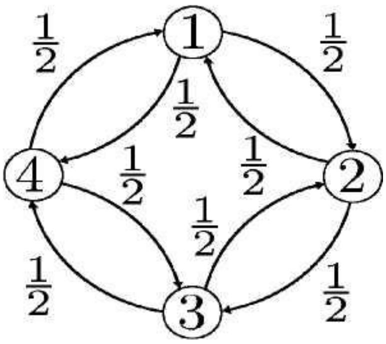

Рис. 3.

***Теорема* 3.** *Для конечной марковской цепи, состоящей из одного неразложимого класса существенных состояний, периоды всех состояний одинаковы:* $d_i = d\ \forall i$.

5. Циклические подклассы и матрица вероятности перехода для периодической цепи

Пусть цепь Маркова состоит из одного класса существенных состояний, период которых (т.е. период цепи) — d.

Эти подклассы обладают тем свойством, что цепь Маркова за один шаг переходит из одного подкласса в другой, и других переходов быть не может.

6. Стационарные и эргодические цепи Маркова

Определение 8. Цепь Маркова называется стационарной, если для любых n, l выполняется равенство

$$P(\xi(k_1) = E_{i1}, \dots, \xi(k_n) = E_{in}) = P(\xi(k_1 + l) = E_{i1}, \dots, \xi(k_n + l) = E_{in}). \qquad (21)$$

То есть многомерные вероятности не меняются при сдвиге вдоль оси времени (причем сдвиг может быть как вправо, так и влево, т.е. $l < 0$, $l > 0$). Тогда для вероятности некоторого состояния в k-й момент времени:

$$P_{ij}(k) = P(\xi(k) = E_i) = P(\xi(k-k) = E_i) = P(\xi(0) = E_i) = P_i^0.$$

То есть для стационарной цепи распределение вероятностей состояний в произвольный момент времени совпадает с распределением вероятностей в начальный момент. Часто стационарный режим устанавливается не сразу, а по истечении какого-то времени, т.е.

$$P_{ij}(k) \to P_j^*,\ k \to \infty,\ \textstyle\sum_j P_j^* = 1. \qquad (22)$$

где P_j^* — *финальные вероятности*. Финальные вероятности могут как зависеть от начального распределения вероятностей, так и не зависеть. В последнем случае цепь называется *эргодической*.

Определение 9. Цепь Маркова называется *эргодической*, если

$$P_{ij}(k) \to P_j^*,\ k \to \infty,\ P_j > 0,\ \textstyle\sum_j P_j^* = 1. \qquad (23)$$

Для конечной эргодической марковской цепи

$$P(k) \to \begin{pmatrix} P^* \\ \vdots \\ P^* \end{pmatrix},$$

где $P(k)$ — матрица вероятностей перехода за k шагов, $P^* = \{P_j^*\}$ — вектор финальных вероятностей.

Под *эргодической* ДМЦ понимается цепь, не имеющая невозвратных состояний. Таким образом, в такой цепи возможны любые переходы между состояниями.

Установить, является ли цепь эргодической можно на основе следующей теоремы.

***Теорема* 4** (Эргодическая теорема). *Если для конечной марковской цепи найдется такое* m_0, *для которого вероятность перехода*

$P_{ij}(m_0) > 0, \forall i, j,$ (24)

то такая цепь является эргодической, т.е. существуют такие P_j^**, что* $\lim_{n\to\infty} P_{ij}^*(n) = P_j^*, \sum_j P_j^* = 1$ и все $P_j^* > 0$.

Финальные вероятности P_j^* *будут удовлетворять системе уравнений Колмогорова*

$P_j^* = \sum_i P_i^* P_{ij}, \forall i, j.$ (25)

Условие (24) означает, что на шаге m_0 матрица вероятности перехода будет состоять из положительных элементов.

Пример 4. Пусть цепь Маркова имеет два возможных состояния $E = (1,2)$. Матрица вероятностей перехода за один шаг имеет вид

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

Определить, является ли данная цепь эргодической.

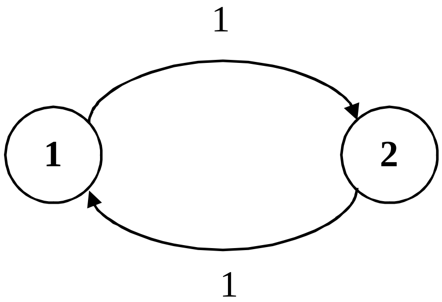

Рис. 5.

Решение. Нарисуем граф, соответствующий данной цепи (рис. 5). Предположим, что в начальный момент времени цепь находилась в состоянии 1, т.е. вектор начальных вероятностей имеет вид: П = (1,0). В соответствии с уравнением Колмогорова-Чепмена в матричном виде (13)

Всё время для любых P^n имеем $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

Таким образом, наблюдаем периодичность в поведении *P(n)* и П(*n*). Цепь не эргодическая.