Лекция 9
Марковские цепи в дискретном времени

Артемов А. В.
мФТиАД ФКН ВШЭ

12 апреля 2018

1 Основные понятия

Наверное, многие слышали про такое понятие, как марковские цепи. Что это такое? Перед этим дадим несколько необходимых понятий.

Пусть E — это некоторое дискретное (конечное или счётное) множество, которое называют *пространством состояний*. Если система находится в состоянии $i \in E$ в момент времени n, то в момент времени $n+1$ она может перейти в состояние $j \in E$ с *переходной вероятностью* p_{ij}. Это сразу даёт два свойства переходной вероятности:

$$\forall i, j \in E \ p_{ij} \geqslant 0 \text{ и } \forall i \in E \ \sum_{j \in E} p_{ij} = 1.$$

Переходные вероятности образуют *матрицу переходных вероятностей* $P = (p_{ij})_{i,j \in E}$. Теперь можно дать определение марковской цепи.

Определение 1. *Марковская цепь* с пространством состояний E и матрицей переходных вероятностей P — это случайный процесс с дискретным временем $X = (X_n)_{n \in \mathbb{N}}$, $X_n \in E$, для которого

- известны начальные распределения $\alpha_i \equiv \mathsf{P}(X_0 = i)$,
- верно *марковское свойство*: для любого натурального n и любых $i_0, i_1, \ldots, i_{n-1}, i, j$

  $$\mathsf{P}(X_{n+1} = j \mid X_n = i) = \mathsf{P}(X_{n+1} = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) = p_{ij},$$

  если условные вероятности хорошо определены, то есть $\mathsf{P}(X_0 = i_0, \ldots, X_n = i) > 0$.

Неформально говоря, марковское свойство означает, что то, как система будет развиваться в текущий момент, не зависит от того, что было в прошлом и зависит только от настоящего.

Теперь вопрос: допустим, что у нас есть какая-то траектория (последовательность состояний). Какова её вероятность? Ответ на этот вопрос даст одна простая теорема.

Теорема 1 (о состояниях марковской цепи)**.** *Для любого натурального n и любых $i_0, i_1, \ldots, i_{n-1}, i, j$*

$$\mathsf{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \alpha_{i_0} p_{i_0 i_1} \ldots p_{i_{n-1} i_n}.$$

Доказательство. Индукция по количеству состояний. Пусть $n = 0$. Тогда по определению марковской цепи $\mathsf{P}(X_0 = i_0) = \alpha_{i_0}$.

Теперь предположим, что утверждение верно для n состояний. Покажем, что оно верно и для $n + 1$ состояния. Действительно, по определению условной вероятности и марковскому свойству

$$\mathsf{P}(X_0 = i_0, \dots, X_{n+1} = i_{n+1}) = \mathsf{P}(X_0 = i_0, \dots, X_n = i_n)\,\mathsf{P}(X_{i+1} = i_{n+1} \mid X_n = i_n) = \\ = \alpha_{i_0} p_{i_0 i_1} \dots p_{i_{n-1} i_n} p_{i_n i_{n+1}}. \quad \square$$

Следствие. *Для любого натурального n и любого $i_n \in E$*

$$\mathsf{P}(X_n = i_n) = \sum_{i_0, \dots, i_{n-1} \in E} \alpha_{i_0} p_{i_0 i_1} \dots p_{i_{n-1} i_n}.$$

Доказательство. Прямое следствие из формулы полной вероятности и теоремы о состояниях марковской цепи. □

Но обычно нас не интересует полный путь, а лишь начало и конец. Поэтому вводят вероятность перейти из состояния i в состояние j за n шагов:

$$p_{ij}^{(n)} = \mathsf{P}(X_n = j \mid X_0 = i)$$

Чему равна эта вероятность? Воспользуемся теоремой о состояниях:

$$\mathsf{P}(X_n = j \mid X_0 = i) = \frac{\mathsf{P}(X_n = j, X_0 = i)}{\mathsf{P}(X_0 = i)} = \\ = \sum_{i_1, \dots, i_{n-1} \in E} \frac{\mathsf{P}(X_0 = i, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n - j)}{\mathsf{P}(X_0 = i)} = \sum_{i_1, \dots, i_{n-1} \in E} p_{i i_1} \dots p_{i_{n-1} j}.$$

Если мы посмотрим на случай $n = 2$, то полученное выражение очень похоже на скалярное произведение строк матрицы переходной вероятности. Оказывается, что это не так уж и далеко от истины.

Теорема 2. *Пусть $P^{(n)} = (p_{ij}^{(n)})_{i,j \in E}$. Тогда $P^{(n)} = P \cdot P \cdot \ldots \cdot P = P^n$.*

Доказательство. Индукция по количеству шагов. База ($n = 1$) очевидна, так как $p_{ij}^{(1)} \equiv p_{ij}$. Теперь предположим, что утверждение выполнено для n шагов. Тогда $P^{(n)} = P^n$. Посмотрим на $P^{(n+1)}$:

$$p_{ij}^{(n+1)} = \mathsf{P}(X_{n+1} = j \mid X_0 = i) = \sum_{i_n \in E} \mathsf{P}(X_{n+1} = j \mid X_n = i_n, X_0 = i)\,\mathsf{P}(X_{n+1} = i_n \mid X_0 = i) = \\ = \sum_{i_n \in E} \mathsf{P}(X_{n+1} = j \mid X_n = i_n)\,\mathsf{P}(X_{n+1} = i_n \mid X_0 = i) = \sum_{i_n \in E} p_{i i_n}^{(n)} p_{i_n j}.$$

Отсюда получаем, что $P^{(n+1)} = P^{(n)} P = P^{n+1}$. □

Однако это доказательство работает не всегда. Почему же? Потому что никто не обещал, что переходная вероятность не зависит от шага. Если она действительно не зависит, то говорят, что марковская цепь *однородна.*

Теперь поговорим про состояния марковских цепей. В зависимости от переходных вероятностей поведение цепи в этом состоянии может кардинально различаться. Поэтому их классифицируют.

Первая классификация связана с важностью состояния. Может оказаться так, что из состояния можно выйти за конечное число шагов, но вернуться назад уже невозможно. Такие состояния не слишком влияют на долговременное поведение марковской цепи, поэтому их считают несущественными. Формализуем это:

Определение 2. Пусть $X = (X_n)_{n\in\mathbb{N}}$ — марковская цепь с матрицей переходных вероятностей P и дискретным множеством состояний E. Будем называть состояние $i \in E$ *несущественным*, если существует состояние j и натуральное n такое, что $p_{ij}^{(n)} > 0$, но $\forall m \in \mathbb{N}\ p_{ij}^{(m)} = 0$. В противном случае состояние называют *существенным.*

Вторая классификация связана с возвращением.

Определение 3. Состояние i марковской цепи называется *возвратным* (reccurent), если

$$\mathsf{P}(X_n = i \text{ для бесконечно многих } n) = 1.$$

Если же эта вероятность равна нулю, то состояние называют *невозвратным* (transient).

Так как на данный момент мы рассматриваем только дискретный случай, то у нас есть роскошь: можно смотреть на марковскую цепь, как на ориентированный граф, где вершины — это события, а вес ребра — переходная вероятность. Теперь вспомним, что в теории графов вводилось такое понятие, как связность. Для марковских цепей вводится похожее понятие.

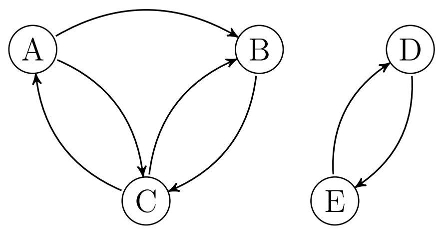

Рис. 1: Пример изображения марковской цепи в виде графа. В данном случае $E = \{A, B, C, D, E\}$. Отсутствие ребра означает, что переходная вероятность равна нулю.

Определение 4. Состояние j марковской цепи называются *достижимым* из состояния i, если существует такое натуральное n, что $p_{ij}^{(n)} > 0$. Обозначение: $i \to j$.

Определение 5. Существенные состояния i и j марковской цепи называются *сообщающимися*, если $i \to j$ и $j \to i$. Обозначение: $i \leftrightarrow j$.

У этого понятия есть одно полезное свойство:

Свойство 1. *Сообщаемость задаёт отношение эквивалентности.*

Доказательство. Для начала нужно показать, что отношение сообщаемости рефлексивно. Пусть $i \in E$ существенно. Это означает, что существуют такие натуральные m и n, что $p_{ij}^{(m)} > 0$ и $p_{ji}^{(n)} > 0$. А это и означает, что $i \leftrightarrow i$.

Коммутативность очевидна. Теперь покажем, что выполнена транзитивность. Для этого достаточно показать, что если $i \to j$ и $j \to k$, то и $i \to k$. Действительно, если $p_{ij}^{(m)} > 0$ и $p_{jk}^{(n)} > 0$, то $p_{ik}^{(m+n)} \geqslant p_{ij}^{(m)} p_{jk}^{(n)} > 0$. □

В итоге мы получаем, что марковскую цепь можно разбить на классы сообщающихся вершин, коих будет не более, чем счётное число. Если такой класс один, то марковскую цепь называют *неприводимой* (или *неразложимой*).

Допустим, что мы разбили множество состояний марковской цепи на классы сообщаемости $S_1, \ldots, S_n, \ldots$ и класс несущественных состояний S_0. Как будет выглядеть матрица переходных вероятностей $P^{(n)}$? Оказывается, что она будет иметь блочный вид:

$$P^{(n)} = \begin{pmatrix} S_0 \to S_0 & S_0 \to S_1 & S_0 \to S_2 & \ldots & S_0 \to S_n & \ldots \\ 0 & S_1 \to S_1 & 0 & \ldots & 0 & \ldots \\ 0 & 0 & S_2 \to S_2 & \ldots & 0 & \ldots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\ 0 & 0 & 0 & \ldots & S_n \to S_n & \ldots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \end{pmatrix},$$

где $S_i \to S_j$ — блочная матрица переходных вероятностей из состояний класса S_i в состояния класса S_j.

Последняя классификация вершин связана с возвращением в состояние. Введём понятие периода состояния.

Определение 6. Пусть $X = (X_n)_{n\in\mathbb{N}}$ — однородная марковская цепь с матрицей переходных вероятностей P и дискретным множеством состояний E. Число $d_i = \gcd\{n \in \mathbb{N} \mid p_{ii}^{(n)} > 0\}$ будем называть *периодом* состояния j. Если $d_j > 1$, то состояние j *периодическое*, иначе же j *апериодическое*.

Осталось ввести одно понятие, которое может понадобиться в дальнейшем.

Определение 7. Состояние i называется *нулевым*, если $p_{ii}^{(n)} \to 0$ при $n \to \infty$.

Теперь посмотрим на какой-нибудь жизненный пример, который можно описать марковской цепью — например, простейшее случайное блуждание. Строится оно так же, как в примере **??**, только берётся распределение, принимающее значения 1, 0 и -1 с какими-то вероятностями. Марковская цепь, симулирующая такой процесс, устроена просто. Множеством состояний является $\mathbb{Z}$, а матрица переходных вероятностей задаётся так:

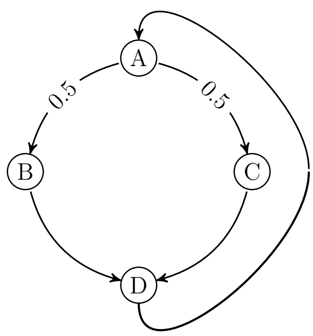

Рис. 2: В данной марковской цепи периодическими будут состояния B и C.

$$p_{ij} = \begin{cases} 0, & |i-j| > 1 \\ p_0, & j = i \\ p_1, & j = i+1 \\ p_2, & j = i-1 \end{cases}$$

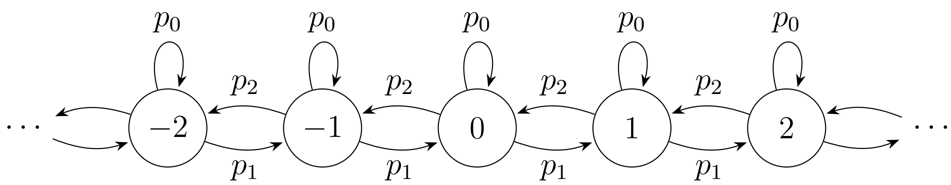

Рис. 3: Графическое изображение марковской цепи, соответствующей простому случайному блужданию.

Пример 1. Пусть блуждание начинается в нуле и $p_0 = p_1 = 1/2$. Отсюда сразу же получаем, что $p_2 = 0$. Что мы можем сказать про такое блуждание? А ничего хорошего. Все состояния несущественны, невозвратны и нулевые.

Пример 2. Теперь изменим вероятности: $p_1 = p_2 = 1/2$. Получится так называемое *симметричное случайное блуждание*. В таком случае мы можем сказать, что нулевое состояние периодично с $d_0 = 2$, так как есть ненулевая вероятность вернуться в начало за $2k$ шагов (и понятно, что вернуться за нечётное число шагов невозможно). Вообще, если и p_1, и p_2 больше нуля, то все состояния существенны и сообщаются.

Пример 3. Сузим множество состояний до $\{-k, \ldots, -1, 0, 1, \ldots, k\}$ и изменим вероятности перехода на границах: $p_{kk} = p_{-k,-k} = 1$. Получится *случайное блуждание с поглощающими экранами.* В таком случае существенными состояниями будут только $\pm k$.

Вернёмся к возвратности. Как можно определить, является ли состояние возвратным? По определению это сдедать весьма непросто. Но есть теорема, которая упрощает жизнь.

Теорема 3. *Пусть $f_i = \mathsf{P}(\exists\, n \in \mathbb{N} : X_n = i)$ — вероятность того, что мы хотя бы раз попали в состояние i. Тогда состояние i будет возвратным тогда и только тогда, когда $f_i = 1$.*

Как посчитать f_i? Пусть $A_i^{(n)} = \{$вернуться в i ровно за n шагов$\}$. Вероятность этого события равна:

$$f_i(n) \equiv \mathsf{P}\left(A_i^{(n)}\right) = \mathsf{P}(X_n = i, X_k \neq i, k \in \{1, \ldots, n-1\} \mid X_0 = i)$$

Отсюда несложно получить, что f_i равна сумме $f_i(n)$ по всем натуральным n.

Доказательство. Зафиксируем состояние i. Пусть $B_k = \{X_n = i$ хотя бы k раз$\}$. Это можно записать по-другому:

$$B_k = \{\exists\, n_1, \ldots, n_k \in \mathbb{N} : X_{n_1} = i, \ldots, X_{n_k} = i\}.$$

Что мы можем сказать про такие события? Во-первых, $B_{k+1} \subseteq B_k$, так как если мы посетили состояние i $k+1$ раз, то мы его посетили и k раз. Далее, марковское свойство даёт нам “отсутствие памяти”. Тогда $\mathsf{P}(B_k) = f_i^k$.

Теперь заметим, что вероятность возвратности можно записать следующим образом:

$$\mathsf{P}(X_n = i \text{ для бесконечно многих } n) = \mathsf{P}\left(\bigcap_{n=1}^{\infty} B_n\right).$$

По непрерывности вероятностной меры

$$\mathsf{P}\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mathsf{P}(B_n) = \lim_{n\to\infty} f_i^n = \begin{cases} 1, & f_i = 1 \\ 0, & f_i < 1 \end{cases}$$ □

Теорема 4. *Пусть i — состояние марковской цепи. Оно будет возвратным тогда и только тогда, когда расходится ряд $\sum p_{ii}^{(n)}$.*

Доказательство. Зафиксируем состояние i и рассмотрим случайную величину

$$V_i = \sum_{k=1}^{\infty} \mathrm{I}\{X_k = i\} = \#\{n \in \mathbb{N} : X_n = i\}.$$

Что мы можем сказать про неё? Во-первых, $\mathsf{P}(V_i \geqslant k) = \mathsf{P}(B_k) = f_i^k$. Далее, мы можем посчитать матожидание двумя способами:

$$\mathsf{E}[V_i] = \mathsf{E}\left[\sum_{k=1}^{\infty} \mathrm{I}\{X_k = i\}\right] = \sum_{k=1}^{\infty} \mathsf{P}(X_k = i) = \alpha_i \sum_{k=1}^{\infty} p_{ii}^{(n)}. \tag{1.1}$$

$$\mathsf{E}[V_i] = \sum_{k=1}^{\infty} k\, \mathsf{P}(V_i = k) = \sum_{k=1}^{\infty} \left(\sum_{j=k}^{\infty} \mathsf{P}(V_i = j)\right) = \sum_{k=1}^{\infty} \mathsf{P}(V_i \geqslant k) = \sum_{k=1}^{\infty} f_i^k. \tag{1.2}$$

Теперь заметим одну вещь: если матожидание конечно, то вероятность того, что $V_i = +\infty$, нулевая. Тогда мы сразу получаем желаемое: если i возвратно, то $f_i = 1$ и ряд $\sum p_{ii}^{(n)}$ расходится. Иначе же ряд $\sum f_i^n$ сходится и матожидание конечно. Следовательно, ряд $\sum p_{ii}^{(n)}$ сходится. □

Теперь применим эту теорему к простейшим случайным блужданиям.

Теорема 5. *Пусть простейшее случайное блуждание задаётся случайными величинами, принимающими значения 1 и* -1 *с вероятностями* p *и* $q \equiv 1-p$ *соответственно. Если* $p = 1/2$, *то все состояния возвратны. Иначе же все состояния невозвратны.*

Доказательство. Без ограничения общности рассмотрим состояние $i = 0$. Как известно, соответствующее случайное блуждание задаётся так:

$$X_n = \xi_1 + \ldots + \xi_n, \quad \{\xi_n\}_{n=1}^{\infty} - \text{iid}, \quad \mathsf{P}(\xi_k = 1) = p, \quad \mathsf{P}(\xi_k = -1) = 1-p.$$

Заметим, что $\mathsf{E}[\xi_k] = \mathsf{P}(\xi_k = 1) - \mathsf{P}(\xi_k = -1) = 2p - 1$. Тогда оно не ноль при $p \neq 1/2$. Далее, по усиленному закону больших чисел

$$\frac{X_n}{n} \xrightarrow{\text{п.н.}} \mathsf{E}[\xi_1] \implies \begin{cases} X_n \xrightarrow{\text{п.н.}} +\infty, & p > 1/2 \\ X_n \xrightarrow{\text{п.н.}} -\infty, & p < 1/2 \end{cases}$$

Из этого следует, что если $p \neq 1/2$, то все состояния невозвратны. Теперь покажем, что если $p = 1/2$, то состояние возвратно. Для этого посчитаем вероятность вернуться в ноль:

$$p_{ii}^{(n)} = \begin{cases} 0, & n \neq 2k, k \in \mathbb{Z} \\ 2^{-2k} C_{2k}^{k}, & n = 2k, k \in \mathbb{Z} \end{cases}$$

Теперь оценим поведение $p_{ii}^{(n)}$. Для этого можно вспомнить формулу Стирлинга:

$$p_{ii}^{(2k)} \sim \frac{1}{2^{2k}} \frac{\sqrt{4\pi k}(2k)^{2k} e^{-2k}}{2\pi k \cdot k^{2k} e^{-2k}} = \frac{1}{\sqrt{\pi k}} \implies \sum_{n=1}^{\infty} p_{ii}^{(n)} = +\infty.$$

Следовательно, состояние $i = 0$ возвратно. □

Пусть у нас есть какая-то марковская цепь и i — её состояние. Можно ли как-то оценить, сколько времени цепь проводит в этом состоянии? Пусть $T_1, T_2, \ldots$ — это упорядоченные моменты попадания в состояние i. Далее, введём "времена вне состояния i" $T^{(1)} = T_1, T^{(2)} = T_2 - T_1, \ldots$. Если мы рассматриваем однородную марковскую цепь, то эти случайные величины независимы в совокупности и одинаково распределены.

Нередко возникает вопрос об относительном времени пребывания в состоянии i, то есть рассматривается отношение n/N, где n — количество шагов, на которых цепь находилась в состоянии i, а N — общее количество шагов.

Теперь рассмотрим матожидание времени вне состояния i для однородной цепи:

$$\mathsf{E}\big[T^{(n)}\big] = \mathsf{E}\big[T^{(1)}\big] = \sum_{k=1}^{\infty} k\, \mathsf{P}\big(T^{(1)} = k\big).$$

Если состояние i возвратно, то по определению $\mathsf{P}\big(T^{(1)} < \infty\big) = 1$ и мы ничего не можем сказать про сходимость полученного ряда. Однако мы можем сказать две вещи:

- Если матожидание конечно, то $\mathsf{P}\big(T^{(1)} < \infty\big) = 1$ и состояние i возвратно.
- Если состояние невозвратно, то ряд точно расходится. Действительно, в таком случае $\mathsf{P}\big(T^{(1)} = \infty\big) = 1 - \mathsf{P}\big(T^{(1)} < \infty\big) > 0$.

Далее, введём обозначение $m_T \equiv \mathsf{E}\big[T^{(1)}\big]$. Что мы можем сказать про предел отношения n/N при $N \to \infty$?

- Пусть $m_T < \infty$. В таком случае состояние i возвратно и n неограниченно возрастает вместе с ростом $T_n \equiv N$. Далее, по усиленному закону больших чисел:

$$\frac{N}{n} = \frac{T_n}{n} = \frac{1}{n}\sum_{k=1}^{n} T^{(k)} \xrightarrow{\text{п.н.}} \mathsf{E}\big[T^{(1)}\big] = m_T \implies \frac{n}{N} \xrightarrow{\text{п.н.}} \frac{1}{m_T}.$$

  Как известно, если $Z_n \xrightarrow{\mathsf{P}} c$, где c — неслучайная величина, то $\mathsf{E}[Z_n] \to c$. Тогда $\mathsf{E}[n/N] \to 1/m_T$.

- Теперь предположим, что $m_T = \infty$. В таком случае мы уже не можем сказать, что состояние возвратно или не возвратно. Поэтому рассмотрим оба случая. Пусть i невозвратно. Тогда всё просто: мы посещаем состояние конечное число раз с вероятностью 1 и $n/N \xrightarrow{\mathsf{P}} 0$. Оказывается, что и для возвратного состояния выполнена та же сходимость.

Теперь докажем одну теорему, связанную с классами сообщаемости. Оказывается, у состояний в одном классе весьма немало общего.

Теорема 6. *Пусть X — это неразложимая однородная марковская цепь с множеством состояний E и матрицей переходных вероятностей P. Тогда*

- *Если одно из состояний цепи нулевое, то все состояния нулевые.*
- *Если одно из состояний цепи возвратное, то все состояния возвратные.*
- *Если одно из состояний цепи имеет период d, то все состояния имеют тот же период.*

Доказательство. Пусть $i \leftrightarrow j$, то есть существуют натуральные M и N такие, что $\alpha \equiv p_{ij}(M) > 0$ и $\beta \equiv p_{ji}(N) > 0$. Теперь возьмём произвольное натуральное n и распишем $p_{ii}(M+n+N)$:

$$p_{ii}(M+n+N) = \sum_{k,l\in E} p_{ik}(M)p_{kl}(n)p_{li}(N) = \sum_{\substack{k,l\in E\\ k\neq j, l\neq j}} p_{ik}(M)p_{kl}(n)p_{li}(N) + \\ + p_{ij}(M)p_{jj}(n)p_{ji}(N) \geqslant \alpha\beta p_{jj}(n).$$

Теперь можно приступить к доказательству:

1. Пусть i нулевое. Тогда и j нулевое:

$$0 = \lim_{n\to\infty} p_{ii}(M+n+N) = \lim_{n\to\infty} \alpha\beta p_{jj}(n) \implies \lim_{n\to\infty} p_{jj}(n) = 0.$$

2. Пусть i возвратное. По теореме 4:

$$\infty = \sum_{n=1}^{\infty} p_{ii}(M+n+N) = \alpha\beta\sum_{n=1}^{\infty} p_{jj}(n) \implies \sum_{n=1}^{\infty} p_{jj}(n) = \infty.$$

   Отсюда получаем, что и j является возвратным состоянием.

3. Теперь предположим, что i и j имеют периоды d_i и d_j соответсвенно. Докажем, что $d_i = d_j$. Введём два множества:
$$W_i = \{n \in \mathbb{N} : p_{ii}(n) > 0\}, \quad d_i = \gcd W_i, \tag{1.3}$$
$$W_j = \{n \in \mathbb{N} : p_{jj}(n) > 0\}, \quad d_j = \gcd W_j. \tag{1.4}$$
Так как $p_{ii}(M+N) \geqslant \alpha\beta > 0$, то $M+N \in W_i$. Аналогично, $M+N \in W_j$. Теперь построим новое множество:
$$W = \{M+N+n \mid n \in W_i\} \subset W_i.$$
По построению W имеет общий делитель d_i. Однако для любого $n \in W_i$ $p_{jj}(M+n+M) \geqslant \alpha\beta p_{ii}(n) > 0$. Тогда $W \subseteq W_j$ и любой элемент из W делится на d_j. Тогда можно скзаать, что $d_j \leqslant d_i$. Рассуждая аналогично, получим желаемое.

Рассуждая таким образом для всех состояний, получаем желаемое. □

Эргодичность вводится и для марковских цепей, хоть и немного по-другому.

Определение 8. Пусть $X = (X_n)_{n\in\mathbb{N}}$ — марковская цепь с множеством состояний E и матрицей переходных вероятностей $\mathbf{P}$. Будем называть её *эргодической*, если существует независимое от начального распределения ненулевое предельное распределение вероятностей и состояния $\{p_j^*\}_{j\in E}$ такие, что
$$\forall i,j \in E \lim_{n\to\infty} p_{ij}(n) = p_j^* > 0.$$

Когда марковская цепь является эргодической и как определить предельное распределение? Для этого введём понятие стационарного распределения:

Определение 9. Набор состояний $\{\overline{p}_j\}_{j\in E}$ называется *стационарным распределением вероятностей* (или *инвариантной мерой*) дискретной марковской цепи, если он не изменяется со временем.

Найти стационарное распределение $\mathbf{p}$ не так уж и сложно. Заметим, что
$$p_j(n+1) = \sum_{i\in E} p_i(n)p_{ij} \implies \overline{p}_j = \sum_{i\in E} \overline{p}_i p_{ij} \implies \mathbf{p} = \mathbf{p}\mathbf{P}.$$

Теперь можно сформулировать теорему (к сожалению, без доказательства):

Теорема 7 (первая эргодическая)**.** *Марковская цепь является эргодической тогда и только тогда, когда она неразложима и периодична. При этом её предельное распределение равно стационарному распределению, которое единственно.*

2 Пример применения марковских сетей: модель системы массового обслуживания

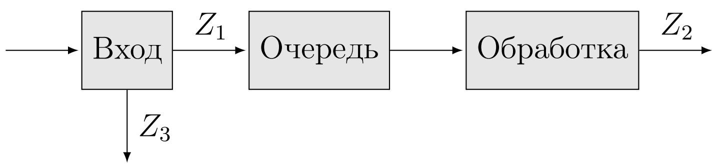

Рис. 4: Графическое изображение модели системы массового обслуживания. Z_1 — поток объектов, принятых на обработку, Z_2 — поток обработанных объектов, Z_3 — поток объектов, отклонённых от обработки из-за переполненности очереди.

Рассмотрим модель некоторой организационной системы, предназначенной для массовой обработки однотипных объектов и состоящей из *устройства обработки* и *очереди*, способной хранить N объектов, ожидающих обработки.

- Будем считать, что время дискретно: что-либо происходит только в моменты времени $t + k\tau$, где $k \in \mathbb{Z}$, а t и τ — какие-то фиксированные отрезки времени.
- В каждый момент времени с вероятностью p появляется новый вызов.
- Если в момент времени t ещё не была закончена обработка, то она заканчивается в момент времени $t + \tau$ с вероятностью q.
- Процессы поступления и обработки независимы.
- Объект проходит три стадии: приём (в случае, если в очереди есть место), ожидание и обработка, после чего направляется на выходной поток.
- Очередь имеет объём $N < \infty$. Из этого следует, что система всего вмещает в себя $N+1$ объект и множество состояний можно представить в виде $E = \{0, 1, \ldots, N, N+1\}$. $N+1$ возникает из-за того, что на обработке может находиться 0 или 1 объект.

Теперь введём три события:

$$A_k(t) = \{\text{в момент времени } t \text{ в системе ровно } k \text{ объектов}\} \tag{2.1}$$

$$B = \{\text{в момент времени } t \text{ в системе появился новый объект}\} \tag{2.2}$$

$$C = \{\text{в момент времени } t+\tau \text{ система закончила обрабатывать объект}\} \tag{2.3}$$

Теперь, как выразить $A_k(t+\tau)$ через них? Начнём с нуля. Заметим, что в системе может не оказаться объектов в двух случаях:

1. В системе в предыдущий момент времени не было объектов. Тогда либо в систему не подали новый объект, либо подали, но система закончила обрабатывать другой обьект.
2. В системе был один объект, обработка которого закончилась. При этом нового объекта не поступило.

Формально это можно записать так:

$$A_0(t+\tau) = (A_0(t) \cap (\overline{B} \cup (B \cap C))) \cup (A_1(t) \cap \overline{B} \cap C).$$

Теперь посмотрим на крайний случай — $A_N(t+\tau)$. Рассуждая аналогичным образом, получаем, что

$$A_N(t+\tau) = (A_{N-1}(t) \cap B \cap \overline{C}) \cup (A_N(t) \cap ((B \cap C) \cup (\overline{B} \cap \overline{C}))) \cup (A_{N+1}(t) \cap C).$$

Отсутсвие $\overline{B}$ в последней скобке объясняется тем, что $A_{N+1}(t)$ уже влечёт это событие. В остальных случаях же

$$A_k(t+\tau) = (A_{k-1}(t) \cap B \cap \overline{C}) \cup (A_k(t) \cap ((B \cap C) \cup (\overline{B} \cap \overline{C}))) \cup (A_{k+1}(t) \cap \overline{B} \cap C).$$

Осталось заметить, что $A_0(t) \cup A_1(t) \cup \ldots A_{N+1}(t) = V$, где V — достоверное событие.

Теперь можно записать вероятности этих событий, пользуясь независимостью:

$$\begin{cases} \mathsf{P}_0(t+\tau) = \mathsf{P}_0(t)(1-p+pq) + \mathsf{P}_1(t)(1-p)q \\ \mathsf{P}_k(t+\tau) = \mathsf{P}_{k-1}(t)p(1-q) + \mathsf{P}_k(t)(pq+(1-p)(1-q)) + \mathsf{P}_{k+1}(t)(1-p)q \\ \mathsf{P}_N(t+\tau) = \mathsf{P}_{N-1}(t)p(1-q) + \mathsf{P}_N(t)(pq+(1-p)(1-q)) + \mathsf{P}_{N+1}(t)q \\ \mathsf{P}_0(t+\tau) + \ldots + \mathsf{P}_{N+1}(t+\tau) = \mathsf{P}_0(t) + \ldots + \mathsf{P}_{N+1}(t) = 1. \end{cases}$$

Теперь предположим, что мы хотим найти стационарное состояние. Тогда система превращается в не очень элегантную систему линейных уравнений:

$$\begin{cases} \mathsf{P}_0 = \mathsf{P}_0(1-p+pq) + \mathsf{P}_1(1-p)q \\ \mathsf{P}_k = \mathsf{P}_{k-1}\,p(1-q) + \mathsf{P}_k(pq+(1-p)(1-q)) + \mathsf{P}_{k+1}(1-p)q \\ \mathsf{P}_N = \mathsf{P}_{N-1}\,p(1-q) + \mathsf{P}_N(pq+(1-p)(1-q)) + \mathsf{P}_{N+1}\,q \\ \mathsf{P}_0 + \ldots + \mathsf{P}_{N+1} = 1. \end{cases}$$

В качестве упражнения оставим то, что решение этой системы выглядит так:

$$\forall k \in \{1,2,\ldots,N\}\; \mathsf{P}_k = \left[\frac{p(1-q)}{q(1-p)}\right]^k \mathsf{P}_0, \quad \mathsf{P}_{N+1} = (1-p)\left[\frac{p(1-q)}{q(1-p)}\right]^{N+1} \mathsf{P}_0 \tag{2.4}$$

$$\mathsf{P}_0 = \left(\sum_{k=1}^{N}\left[\frac{p(1-q)}{q(1-p)}\right]^k + (1-p)\left[\frac{p(1-q)}{q(1-p)}\right]^{N+1}\right)^{-1}. \tag{2.5}$$

Оказывается, что данная модель задаёт эргодичечкую марковскую цепь, поэтому стационарные вероятности являются предельными и неплохо описывают поведение системы, устаканивающееся после достаточно большого отрезка времени. Эти соотношения позволяют решать такие практические задачи, как вычисление вероятности отказа от обслуживания поступающего объекта (вероятность $P_{N+1}(t)$), среднего времени ожидания обслуживания, среднего числа объектов, находящихся в бункере и так далее.