Александр Игоревич Буфетов.

Цепи Маркова.

I. ОПРЕДЕЛЕНИЕ ЦЕПЕЙ МАРКОВА. ЭРГОДИЧЕСКАЯ ТЕОРЕМА ДЛЯ ЦЕПЕЙ МАРКОВА

Рассмотрим связный ориентированный конечный граф (возможно, с петлями), вершины которого занумерованы числами от 1 до n. Припишем каждому его ребру ij (соединяющему вершину i с вершиной j) положительное число p_{ij} (рис. 1, *а*). Каждое такое число p_{ij} мы будем интерпретировать как вероятность перехода из вершины i в вершину j, поэтому эти числа должны удовлетворять условию

$$\sum_j p_{ij}=1$$

для любого i, а суммирование ведётся по всем вершинам j, в которые можно перейти из вершины i. (Это условие означает, что из любой вершины куда-нибудь всегда возможно перейти.)

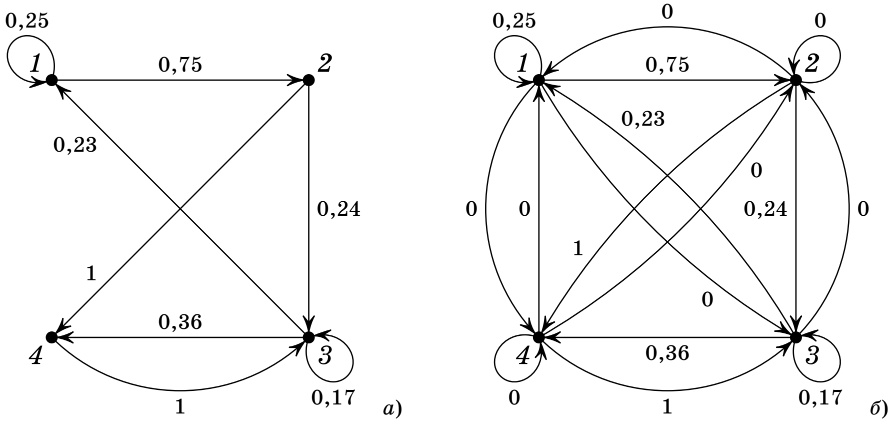

Рис. 1

Можно также считать, что граф у нас полный, всем рёбрам приписаны числа, но некоторые, просто некоторые из этих чисел равны нулю (на рис. 1, *б* показан полный граф, аналогичный графу рис. 1, *а*).

Этот граф с заданными вероятностями перехода от вершины к вершине и называется *марковской цепью*.

Составим из этих вероятностей перехода матрицу $n\times n$ (предполагая, что граф полный, т. е. каждая вершина соединена ребром с каждой, но некоторым рёбрам могут быть приписаны нулевые вероятности)

$$P=\begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix}.$$

Например, для марковской цепи, показанной на рис. 1, *б* матрица P имеет вид

$$\begin{pmatrix} 0{,}25 & 0{,}75 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0{,}23 & 0{,}24 & 0{,}17 & 0{,}36 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Нас интересует поведение системы «за большое время» вот в каком смысле. Пусть в начальный момент времени некоторая частица находилась, скажем, в первой вершине графа, а затем в каждый момент времени частица переходит по ребру в другую вершину (возможно, в ту же самую, т. е. она проходит по петле) в соответствии с приписанными рёбрам вероятностями; при этом мы предполагаем, что вероятность перехода не зависит от того, что было с частицей раньше, а зависит только от того, где частица находится в данный момент: если известно, что частица находится, скажем, в первой вершине, то известны и все переходные вероятности, а где частица была перед этим, никак не учитывается.

Следует заметить, что использованное слово «вероятность» не предполагает знания теории вероятностей: это просто удобное слово, которое вызывает необходимые ассоциации. Вообще же нас будут интересовать элементы матриц P^m, которые мы и будем называть вероятностями.

Вероятность попасть из вершины i в j за один шаг равна p_{ij} (по определению), а какова вероятность попасть из вершины i в вершину j за два шага? Ответ: с вероятностью

$$\sum_{k=1}^{n} p_{ik}p_{kj}.$$

Действительно, с вероятностью p_{ik} можно перейти из вершины i в некоторую вершину k, а затем с вероятностью p_{kj} можно перейти из k в j, таким образом вероятность попасть из i в k, а затем из k в j равна $p_{ik}p_{kj}$. Чтобы получить окончательный ответ, нужно просуммировать эти выражения по всем возможным k.

Выражение для рассмотренной вероятности очень легко получается с использованием матрицы P: эта вероятность равна элементу матрицы P^2, стоящем на пересечении i-й строки и j-го столбца (этот элемент мы будем обозначать через $(P^2)_{ij}$). Вообще, вероятность попасть из вершины i в вершину j за m шагов равна $(P^m)_{ij}$ — это утверждение называется уравнением Колмогорова—Чепмена.

Посмотрим, чему же равны вероятности $(P^m)_{ij}$ при больших m. Для примера возведём матрицу P, соответствующую марковской цепи рис. 1, *б*, в степени $m=2, 3, \ldots, 10$ (табл. 1).

Как видно из таблицы, за длительное время матрицы стабилизируются: величины $(P^m)_{ij}$ перестают зависеть от i и от m. Это

Таблица 1

$P^2=$	0,0625	0,1875	0	0,75
	0	0	1	0
	0,0966	0,2133	0,3889	0,3012
	0,23	0,24	0,17	0,36
$P^3=$	0,015625	0,046875	0,75	0,1875
	0,23	0,24	0,17	0,36
	0,113597	0,165786	0,367313	0,353304
	0,0966	0,2133	0,3889	0,3012
$P^4=$	0,17640625	0,19171875	0,315	0,316875
	0,0966	0,2133	0,3889	0,3012
	0,11288124	0,17335287	0,41574721	0,29801868
	0,113597	0,165786	0,367313	0,353304
$P^5=$	0,1165515625	0,2079046875	0,370425	0,30511875
	0,113597	0,165786	0,367313	0,353304
	0,1238421683	0,1844402604	0,3686957057	0,3230218656
	0,11288124	0,17335287	0,41574721	0,29801868
$P^6=$	0,114335640625	0,176315671875	0,368091	0,3412576875
	0,11288124	0,17335287	0,41574721	0,29801868
	0,115760554386	0,181368595593	0,385700135569	0,317170714452
	0,1238421683	0,1844402604	0,3686957057	0,3230218656
$P^7=$	0,11324484015625	0,17409357046875	0,4038331575	0,308828431875
	0,1238421683	0,1844402604	0,3686957057	0,3230218656
	0,11765116977737	0,17938844832606	0,38273973749873	0,32022064439784
	0,115760554386	0,181368595593	0,385700135569	0,317170714452
$P^8=$	0,1211928362640625	0,1818535879171875	0,37748006865	0,31947350716875
	0,115760554386	0,181368595593	0,385700135569	0,317170714452
	0,1174429320690504	0,1800959143327227	0,3852863997726241	0,3171747538256028
	0,11765116977737	0,17938844832606	0,38273973749873	0,32022064439784
$P^9=$	0,117118624855515625	0,181489843674046875	0,38364511883925	0,3177464126311875
	0,11765116977737	0,17938844832606	0,38273973749873	0,32022064439784
	0,117976604964966143	0,180550934997217584	0,382673441786948897	0,318799018250867376
	0,1174429320690504	0,1800959143327227	0,3852863997726241	0,3171747538256028
$P^{10}=$	0,11751803354690640625	0,17991379716305671875	0,38296608283386	0,319602086456176875
	0,1174429320690504	0,1800959143327227	0,3852863997726241	0,3171747538256028
	0,11750904285223978206	0,18032407975259234253	0,38385350335464868849	0,31831337404051918692
	0,117976604964966143	0,180550934997217584	0,382673441786948897	0,318799018250867376

утверждение называется эргодической теоремой для цепей Маркова и ниже мы точно сформулируем и докажем его, но сначала дадим одно определение.

Если существует такое натуральное число m_0, что все элементы матрицы P^{m_0} положительны (далее в таком случае мы будем и сами матрицы называть *положительными*), то матрица P называется *регулярной*. Цепь Маркова с регулярной матрицей обладает тем свойством, что за некоторое число шагов (m_0) из любой вершины можно попасть в любую другую. Заметим, что если из любой вершины можно попасть за m_0 шагов в любую другую, то и за m_0+1 шагов можно попасть из любой вершины в любую другую. Это следует из того, что произведение положительной и неотрицательной матриц является положительной матрицей. То же самое можно объяснить и проще: предположим, что за m_0 шагов из любой вершины можно попасть в любую другую, то на первом шаге перейдём из исходной в любую другую вершину, а затем за m_0 шагов перейдём в искомую.

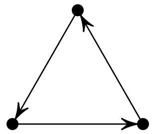

Рис. 2

Может показаться, что для регулярности матрицы цепи Маркова достаточно потребовать, чтобы из любой вершины можно было бы попасть в любую другую вершину за некоторое число шагов (это условие называется *неприводимостью*). Этого оказывается не достаточно: пример графа, для которого неприводимость не означает регулярности, приведён на рис. 2. Однако, если вдобавок к неприводимости добавить условие *непериодичности*, т. е. потребовать выполнения равенства

$$\text{НОД}\{m \mid (P^m)_{ii}>0\}=1,$$

то получится уже регулярность.

Эргодическая теорема для цепей Маркова. Если P — регулярная матрица, то существуют такие положительные числа p_1, p_2, …, p_n ($p_1+p_2+\ldots+p_n=1$), к которым сходятся величины $(P^m)_{ij}$:

$$\lim_{m\to\infty}(P^m)_{ij}=p_i \qquad (1)$$

для любого $i=1, 2, \ldots, n$.

Эту теорему можно интерпретировать следующим образом. Пусть в каждой вершине графа сидит много зайцев, которые прыгают из вершины в вершину согласно некоторой матрице, тогда через некоторое время (как мы увидим, через довольно маленькое время, поскольку сходимость в (1) экспоненциальная) ничего не будет заметно: считая зайцев в некоторой вершине никаких изменений мы не увидим, — сколько зайцев упрыгало, столько и припрыгало.

Представим себе, что мы уже доказали эргодическую теорему. Но как найти числа p_1, p_2, …, p_n? Не будем же мы возводить матрицу P в степени, чтобы их найти? Определить их можно следующим образом.

Набор $p=(p_1, p_2, \ldots, p_n)$ удовлетворяет условию $pP=p$. Более того, если для некоторого набора $q=(q_1, q_2, \ldots, q_n)$ выполнено равенство $qP=q$, причём $q_1+q_2+\ldots+q_n=1$, то $q=p$.

Доказательство эргодической теоремы довольно простое, но является очень важным для многих применений, поскольку мы получим не только само доказательство, но и оценим скорость сходимости в (1); кроме того, идеи, которыми мы будем пользоваться, очень часто встречаются в эргодической теории.

Рассмотрим j-й столбец матрицы P^m и обозначим через Γ_j^m наибольший элемент этого столбца, а через γ_j^m наименьший элемент, т. е.

$$\Gamma_j^m=\max_j(P^m)_{ij}, \qquad \gamma_j^m=\min_j(P^m)_{ij}.$$

Запишем уравнение Колмогорова—Чепмена:

$$(P^{m+1})_{ij}=\sum_{k=1}^n p_{ik}(P^m)_{kj}.$$

Полагая в этой формуле i равным номеру строки, в которой стоит наибольший элемент j-го столбца (Γ_j^{m+1}), получаем:

$$\Gamma_j^{m+1}=\sum_{k=1}^n p_{ik}(P^m)_{kj}\leqslant\sum_{k=1}^n p_{ik}\Gamma_j^m=\Gamma_j^m\sum_{k=1}^n p_{ik}=\Gamma_j^m\cdot 1=\Gamma_j^m.$$

Аналогично, полагая в уравнении Колмогорова—Чепмена i равным номеру строки, в которой стоит наименьший элемент j-го столбца, получаем:

$$\gamma_j^{m+1}=\sum_{k=1}^n p_{ik}(P^m)_{kj}\geqslant\sum_{k=1}^n p_{ik}\gamma_j^m=\gamma_j^m\sum_{k=1}^n p_{ik}=\gamma_j^m\cdot 1=\gamma_j^m.$$

Таким образом, последовательность Γ_j^1, Γ_j^2, …, Γ_j^m, … является убывающей, а последовательность γ_j^1, γ_j^2, …, γ_j^m, … — возрастающей. Но этого ещё мало: обе последовательности действительно сходятся (поскольку являются монотонными и ограниченными), однако может получиться так, что одна из них сходится, скажем к $1/5$, а другая — к $1/10$; нам нужно ещё доказать, что их пределы совпадают. Докажем, что при росте m разность между Γ_j^m и γ_j^m уменьшается до нуля.

Допустим, что пределы этих последовательностей не совпадают. Поскольку P — регулярная матрица, то $(P^{m_0})_{ik}>0$ для любых i, k. Более того, для некоторого положительного числа ε и для всех i, j выполнено $(P^{m_0})_{ik} > \varepsilon$. Рассмотрим тот элемент

$$\sum_{k=1}^{n} (P^{m_0})_{ik} (P^m)_{kj},$$

на котором достигается максимум $\Gamma_j^{m+m_0}$. Этот элемент есть линейная комбинация вероятностей $(P^m)_{kj}$ для разных k (с положительными коэффициентами $(P^{m_0})_{ik}$, которые в сумме дают 1) в том числе и того элемента, на котором достигается минимум γ_j^m; берётся среднее взвешенное, и вес минимума хотя бы ε, т. е.

$$\Gamma_j^{m+m_0} \leqslant \Gamma_j^m (1-\varepsilon) + \gamma_j^m \varepsilon,$$
$$\gamma_j^{m+m_0} \geqslant \gamma_j^m (1-\varepsilon) + \Gamma_j^m \varepsilon.$$

Осталось вычесть из первого неравенства второе, и мы получаем:

$$\Gamma_j^{m+m_0} - \gamma_j^{m+m_0} \leqslant (1-2\varepsilon)(\Gamma_j^m - \gamma_j^m).$$

Таким образом, разность $\Gamma_j^m - \gamma_j^m$ убывает экспоненциально и мы получили доказательство существования пределов (1) для некоторых m; но, поскольку, как мы установили выше, последовательности $\{\Gamma_j^m\}$ и $\{\gamma_j^m\}$ монотонны, из этого следует существование пределов (1) и экспоненциальная сходимость для всех m.

Нам осталось ещё ответить на следующие вопросы. Почему сумма пределов $p_1 + p_2 + \ldots + p_n$ равна 1? Почему они все положительны? Почему набор $p = (p_1, p_2, \ldots, p_n)$ удовлетворяет равенству $pP = p$? И почему равенства $pP = p$ и $p_1 + p_2 + \ldots + p_n = 1$ однозначно определяют набор p?

Равенство суммы пределов $p_1 + p_2 + \ldots + p_n$ единице получается предельным переходом при $m \to \infty$ из равенства

$$\sum_{j=1}^{n} (P^m)_{ij} = 1.$$

Положительность чисел p_1, p_2, …, p_n следует из (1), поскольку и все $(P^m)_{ij}$ положительны.

Равенство $pP = p$ получается из равенств

$$\sum_{j=1}^{n} (P^m)_{ij} p_{jk} = (P^{m+1})_{ik}$$

($k = 1$, 2, …, n) предельным переходом при $m \to \infty$:

$$\sum_{j=1}^{n} p_j p_{jk} = p_k,$$

таким образом, $pP = p$.

Предположим теперь, что для некоторого набора $q = (q_1, q_2, \ldots, q_n)$ выполнены равенства $qP = q$ и $q_1 + q_2 + \ldots + q_n = 1$. Тогда

$$qP^m = (\ldots((\underbrace{qP)P)\ldots)P}_{m \text{ раз}} = q.$$

Докажем теперь, что qP^m сходится при $m \to \infty$ к p, это и будет означать, что $p = q$. Действительно, j-й элемент qP^m записывается в виде

$$\sum_{k=1}^{n} q_k (P^m)_{kj} \xrightarrow[m\to\infty]{} \sum_{k=1}^{n} q_k p_j = p_j \sum_{k=1}^{n} q_k = p_j,$$

т. е. и вправду qP^m сходится при $m \to \infty$ к p.

Ещё раз отметим, что смысл эргодической теоремы для цепей Маркова заключается в том, что степени матрицы P (поэлементно) сходятся к матрице

$$\mathscr{P} = \begin{pmatrix} p_1 & p_2 & \ldots & p_n \\ p_1 & p_2 & \ldots & p_n \\ \hdotsfor{4} \\ p_1 & p_2 & \ldots & p_n \end{pmatrix};$$

через довольно большое время теряется зависимость от начального положения частицы (зайцев и т. п.).

Заметим ещё, что если бы мы потребовали в формулировке эргодической теоремы лишь условия неприводимости матрицы (т. е. мы бы разрешили периодичность), надо было бы заменить матрицу P на P^T, где $T = \text{НОД}\{m \mid (P^m)_{ii} > 0\}$ — период. В таком случае степени P^m всё равно стабилизируются, правда, к периодической матрице. Фактически, граф распадается на T регулярных кусков с матрицей перехода P^T. Для неприводимых матриц цепей стационарное распределение всё равно существует и единственно, но уже степени матрицы к нему не сходятся, поскольку степени матрицы вообще не сходятся, к стационарному распределению сходятся *чезаровские средние*:

$$\frac{1}{N} \sum_{n=0}^{N-1} P^n \xrightarrow[N\to\infty]{} \mathscr{P} = \begin{pmatrix} p_1 & p_2 & \ldots & p_n \\ p_1 & p_2 & \ldots & p_n \\ \hdotsfor{4} \\ p_1 & p_2 & \ldots & p_n \end{pmatrix}.$$

II. МАРКОВСКИЕ ЦЕПИ В ЭРГОДИЧЕСКОЙ ТЕОРИИ И ДИНАМИЧЕСКИХ СИСТЕМАХ

Дадим важное для дальнейшего обсуждения марковских цепей в эргодической теории и теории динамических систем определение пространства возможных траекторий. Когда говорят о теории вероятности, обычно говорят о некотором пространстве с мерой, и именно это пространство мы сейчас и определим.

Предположим, что в начальный момент частица находилась в вершине ω_1 нашего графа Γ, затем она переместилась в вершину ω_2, затем — в вершину ω_3, и т. д. Таким образом, путь частицы задаётся последовательностью $\omega = \omega_1\omega_2\omega_3\ldots$ Множество

$$\Omega_m^\Gamma = \{\omega = \omega_1\omega_2\omega_3\ldots\}$$

всех таких последовательностей и является пространством всех возможных траекторий, разрешённых нашим графом Γ (это и обозначает индекс Γ у Ω_n^Γ: Ω_n — это пространство вообще всех последовательностей, составленных из $\omega_i = 1, 2, \ldots, m$, а Ω_n^Γ — это пространство именно допустимых траекторий).

Рассмотрим сначала всё пространство Ω_m. Оно является, конечно же, топологическим пространство, сейчас мы введём на нём топологию и меру. Основное множество пространства последовательностей — это цилиндр. *Цилиндром* называется множество таких последовательностей, у которых некоторые k элементов (соответствующие индексам $i_1, i_2, \ldots, i_k$) фиксированы, т. е. $\omega_{i_l} = \alpha_l$ (каждое из чисел α_l принадлежит нашему алфавиту, т. е. является натуральным числом от 1 до m, а l пробегает все значения от 1 до k), а все остальные элементы последовательностей произвольны.

Заметим, что дополнением всего пространства Ω_m до цилиндра является объединение некоторого конечного числа цилиндров с тем же набором индексов $i_1, i_2, \ldots, i_k$ (проверьте это самостоятельно). Эти цилиндры и будут открытыми множествами нашей топологии, а поскольку дополнение до цилиндра есть объединение конечного числа цилиндров, они будут по совместительству и замкнутыми. В топологии на $\mathbb{R}$ нет такого, чтобы множество было одновременно и открытым, и замкнутым, а здесь таких сколько угодно, и более того, они образуют топологию.

Но прежде чем углубляться в эти рассуждения, введём на пространстве последовательностей функцию расстояния по аналогии с расстоянием на пространстве бесконечных двоичных последовательностей: рассмотрим две последовательности $\omega = \omega_1\omega_2\omega_3\ldots$ и $\omega' = \omega'_1\omega'_2\omega'_3\ldots$; положим расстояние между ними равным

$$d(\omega, \omega') = \sum_{i=1}^{\infty} \frac{|\omega_i - \omega'_i|}{n^i}.$$

В этой метрике близкими являются те последовательности, у которых совпадают начальные куски.

Рассмотрим последовательность $w^{(1)}, w^{(2)}, w^{(3)}, \ldots$ последовательностей из Ω_m. Сходимость в этой метрике устроена так же, как и для действительных чисел: при уменьшении расстояния

от $w^{(i)}$ до предельной последовательности ω сначала стабилизируется первый элемент, затем стабилизируется второй, третий и т. д.

Теперь давайте нарисуем пространство последовательностей и убедимся, что топологически оно выглядит как канторово множество. Вспомним, как оно строится. Берётся отрезок (0, 1) (рис. 3, *а*); затем из него выкидывается средняя треть, левому отрезку приписывается 0, правому — 1 (рис. 3, *б*); затем у каждого из этих отрезков выкидывается средняя треть, четырём получившимся отрезкам приписываются последовательности 00, 01, 10, 11 (рис. 3, *в*); затем и у каждого из полученных отрезок выкидывается средняя треть, полученным восьми отрезкам приписываются 000, 001, 010, 011, 100, 101, 110, 111 (рис. 3, *г*), и т. д.

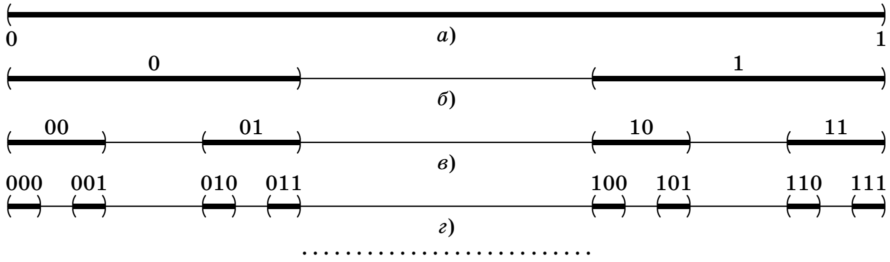

Рис. 3

Получилось взаимно-однозначное отображение из множества двоичных последовательностей Ω_2 в канторово множество. Понятно, что если символов не два, а больше, надо на каждом шаге выделять не два, а n отрезков. На рис. 4 показан пример такой конструкции при $n = 5$.

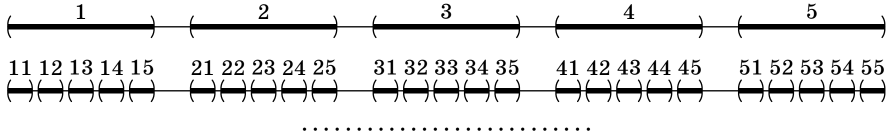

Рис. 4

Но надо заметить, что канторово множество на свете одно, все получающиеся пространства последовательностей при разных n гомеоморфны. Существуют взаимно-однозначные отображения между пространствами последовательностей Ω_n при разных n. Чтобы убедиться в том, что Ω_n гомеоморфно $\Omega_{n'}$ нужно однозначно закодировать n символов при помощи n' символов.

Итак, мы установили, что все пространства последовательностей Ω_n гомеоморфно канторову множеству. Для нас будет важно, что пространство Ω_n является кроме того ещё и динамической системой. Зададим на Ω_n отображение $\sigma_m\colon \Omega_m \to \Omega_m$, задав его формулой $\sigma_m\colon \omega_1\omega_2\omega_3\ldots \mapsto \omega_2\omega_3\ldots$, по существу, σ_m — это сдвиг последовательности на одну позицию влево.

Рассмотрим отображение сдвига на примере двоичных записей, рассмотрим двоичные записи чисел из отрезка $(0, 1)$, забыв на время о неприятных двоично-рациональных числах, у которых двоичная запись неоднозначна. Нарисуем график сдвига. Разделим отрезок $(0, 1)$ на две части: часть, в которой последовательности начинаются с 0, и часть, в которой последовательности начинаются с 1. В первой части этот 0 просто вычёркивается, т. е. число удваивается, в другой части вычёркивается 1, т. е. это отображение $x \mapsto 2(x-1/2)=2x-1$, другими словами, сдвиг — это отображение $x \mapsto \{2x\}$ (рис. 5).

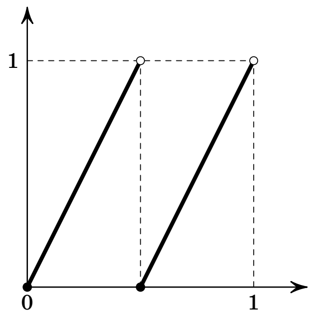

Рис. 5

Ещё интереснее рассмотреть это отображение на окружности $\mathbb{R}/\mathbb{Z}$ (множестве дробных частей вещественных чисел), отождествив 0 с 1. Тогда сдвиг будет непрерывным отображением. Подчеркнём, что сдвиг на окружности — растягивающее отображение (оно растягивает отрезки в два раза). Это отображение теряет информацию, у каждой точки два прообраза. В теории динамических систем говорят, что это отображение является *хаосом*. Даже если мы знаем положение точки в настоящий момент с большой точностью (скажем, с точностью 1/100), через некоторое время (а именно, через семь итераций), точка может оказаться в любом месте окружности $\mathbb{R}/\mathbb{Z}$. Примером другого рода является поворот окружности $x \mapsto x+\alpha$. Если мы знаем положение точки с некоторой точностью, то и в любой момент времени мы будем знать её положение с той же самой точностью, поскольку у этого отображения лишь одна орбита: все точки делают одно и то же, они поворачиваются на некоторый угол α; никакой потери информации не происходит. Ещё раз заметим, что сдвиг ведёт себя совсем по другому: если известно положение точки с точностью ε, то этой точности хватает на очень маленькое время $|\log_2 \varepsilon|$. Это отображение является самым простым примером хаоса в динамических системах.

Хотя сдвиг — это и хаос, он поддаётся некоторому изучению. Мы не можем ничего конкретного сказать о какой-нибудь фиксированной точке, зато мы можем делать утверждения о траекториях в целом, например, почти любая траектория попадает в любой подотрезок отрезка $(0, 1)$, почти любая траектория равномерно распределена на отрезке $(0, 1)$: она попадает в некоторый отрезочек

с такой частотой, какова его длина (в обоих утверждениях именно «почти любая» траектория, но не любая).

Вернёмся теперь к определению меры на пространстве последовательностей Ω_m и определим марковскую меру на пространстве допустимых последовательностей Ω_m^Γ. Когда мы говорим о вероятности, на самом деле, то, о чём мы говорим, — это мера на пространстве последовательностей Что такое мера на отрезке? Мера — это способ измерять длины подотрезочков этого отрезка (каждому отрезочку сопоставляется неотрицательное число). Способы бывают разные, например, можно просто брать длину подотрезочка, этот способ называется мерой Лебега; есть и более примитивный способ: если отрезочек содержит левый край большого отрезка, считать его меру, равной 1, а если не содержит, — равной 1.

Что же такое мера на пространстве последовательностей? Чтобы определить меру на Ω_m^Γ, нужно определить меру на цилиндрах, причём очень специального вида: у которых набор индексов i_1, i_2, …, i_k — это просто 1, 2, …, k. (Как мы доказали, дополнение до цилиндра является конечным объединением цилиндров, а мера всегда предполагается аддитивной, даже счётно-аддитивной: мера объединения счётного числа непересекающихся отрезков равна сумме их мер.) Такие цилиндры мы будем обозначать через $C^{1\,2\,\ldots\,k}_{\alpha_1\alpha_2\ldots\alpha_k}$.

Меру каждого такого цилиндра можно положить, например, равной $\mu(C^{1\,2\,\ldots\,k}_{\alpha_1\alpha_2\ldots\alpha_k})=2^{-k}$. Это равномерная мера (как при бросании монетки): последовательность 0 имеет вероятность 1/2, последовательность 01 — вероятность 1/4 и т. д. Другая мера получается из рассмотрения предельной матрицы $\mathscr{P}$ цепи Маркова. Положим

$$\mu(C^{1\,2\,\ldots\,k}_{\alpha_1\alpha_2\ldots\alpha_k})=p_1p_2\ldots p_k.$$

Эта мера называется мерой Бернулли. Но хотелось бы определить такую меру, для которой вероятность перехода из вершины i в вершину j равна p_{ij}. Таковой является мера

$$\mu(C^{1\,2\,\ldots\,k}_{\alpha_1\alpha_2\ldots\alpha_k})=q_{\alpha_1}p_{\alpha_1\alpha_2}p_{\alpha_2\alpha_3}\ldots p_{\alpha_{n-1}\alpha_n},$$

где q_{α_1} — вероятность того, что в начальный момент частица находится в вершине α_1, эти величины (для разных значений α_1 нужно задавать отдельно, но так, чтобы выполнялось условие

$$\sum_{i=1}^{n} q_{\alpha_1}|_{\alpha_1=i}=\sum_{i=1}^{n} q_i=1.$$

Таким образом, цепь Маркова задаёт не одну меру, а целое семейство мер, запараметризованных q_{α_1}. Среди этих мер (с разными q_{α_1}) особо выделяется мера с $q_{\alpha_1}=p_{\alpha_1}$, потому что такая мера инвариантна относительно сдвига. Что это означает? При сдвиге теряется информация, множество $\{\omega \mid \omega_1 = 1\}$ при сдвиге переходит во всё пространство Ω_m^Γ,

$$\mu(\{\omega \mid \omega_1 = 1\}) < 1, \quad \text{а} \quad \mu(\Omega_m^\Gamma) = 1.$$

Что же означает инвариантность?

Интуитивно понятно, что равномерная мера инвариантна, она не зависит сдвига в том смысле, что неважно когда начинать бросать монетку сегодня, завтра, или послезавтра, — делать будешь одно и то же. Отметим на графике сдвига отрезочек, отвечающий, скажем, 15 фиксированным начальным символам последовательностей (длина этого отрезка равна 2^{-15}). При сдвиге длина этого отрезка удваивается (рис. 6), что противоречит нашей интуиции.

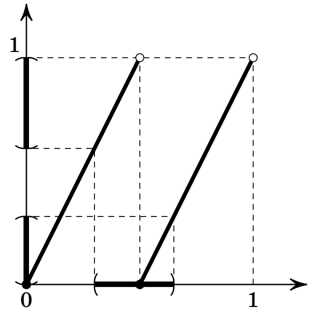

Рис. 6

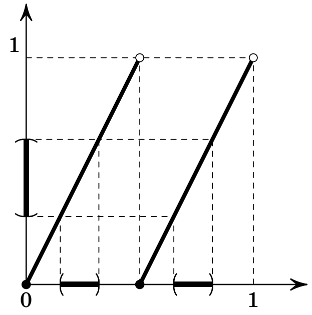

Рис. 7

Инвариантность меры относительно сдвига означает вот что: совпадают длина отрезочка и длина полного прообраза этого отрезочка под действием сдвига. Для равномерной метрики это показано на рис. 7. Аналогично и для марковской метрики

$$\mu(C^{1\ 2\ \dots\ k}_{\alpha_1 \alpha_2 \dots \alpha_k}) = p_{\alpha_1} p_{\alpha_1 \alpha_2} p_{\alpha_2 \alpha_3} \dots p_{\alpha_{n-1} \alpha_n}$$

инвариантность означает равенство $\mu(\sigma_m^{-1} A) = \mu(A)$ для любого отрезочка A.

Вообще, мера μ инвариантна относительно отображения φ, если условие $\mu(\varphi^{-1} A) = \mu(A)$ выполняется для любого отрезочка A.

Все меры, которые мы ввели, — равномерную меру, меру Бернулли и семейство марковских мер — мы называли мерами, хотя мы и не проверили основное свойство меры — аддитивность: мера объединения нескольких непересекающихся отрезочков равна сумме мер самих отрезочков. Это предоставляется читателю в качестве упражнения.