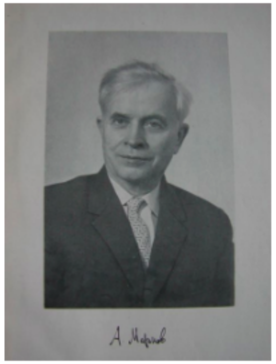
Моделирование марковских случайных процессов. Цепи Маркова.

Андрей Андреевич Марков

- Старший 14.06.1856 - 20.07.1922
- Статистика, Модели Маркова



- Младший 22.09.1903 - 11.10.1979
- Нормальные алгоритмы

Пусть имеется некоторая система S, состояние которой меняется с течением времени (под системой S может пониматься техническое устройство, производственный процесс, вычислительная машина, информационная сеть и т. д.). Если состояние системы S меняется во времени случайным, заранее непредсказуемым образом, говорят, что в системе протекает **случайный процесс.**

Существуют хорошо известные семейства случайных процессов: гауссовы процессы, пуассоновские процессы, авторегрессивные модели, модели скользящего среднего, цепи Маркова и другие. Каждое из этих отдельных случаев имеет определённые свойства, позволяющие нам лучше исследовать и понимать их.

Случайное событие подразумевает, что у некоторого события есть несколько исходов и то, который из исходов произойдет в очередной раз, определяется только его вероятностью. То есть исход выбирается случайно с учетом его вероятности.

Одно из свойств, сильно упрощающее исследование случайного процесса — это **«марковское свойство»**

Случайный процесс, протекающий в системе S, называется **марковским** (или “процессом без последействия”), если он обладает следующим свойством: для каждого момента времени t0 вероятность любого состояния системы в будущем (при t>t0) зависит только от ее состояния в настоящем (при t= t0) и не зависит от того, когда и каким образом система пришла в это состояние (т. е. как развивался процесс в прошлом).

Марковское свойство сообщает нам, что если мы знаем значение, полученное каким-то случайным процессом в заданный момент времени, то не получим никакой дополнительной информации о будущем поведении процесса, собирая другие сведения о его прошлом. Более математическим языком: в любой момент времени условное распределение будущих состояний процесса с заданными текущим и прошлыми состояниями зависит только от текущего состояния, но не от прошлых состояний (свойство отсутствия памяти). Случайный процесс с марковским свойством называется марковским процессом.

$$P(\text{future} \mid \text{present}, \text{past}) = P(\text{future} \mid \text{present}, \cancel{\text{past}})$$

Markov property

Марковское свойство обозначает, что если мы знаем текущее состояние в заданный момент времени, то нам не нужна никакая дополнительная информация о будущем, собираемая из прошлого.

Свойство Маркова называется свойством «отсутствия памяти».

- Для процесса Маркова его будущее (т. е. распределение будущих результатов) зависит только от текущего состояния, но не от его прошлого.
- Нам не нужно знать полную историю состояний, чтобы знать, что произойдет затем, только текущее.
- Свойство Маркова является желаемым свойством в задачах прогнозного моделирования.
- Свойство Маркова приводит к значительному сокращению числа параметров при изучении таких процессов.
- Некоторые немарковские процессы могут быть преобразованы в марковские в пространствах большой размерности

Марковский процесс - это случайный процесс, который удовлетворяет марковскому свойству, что означает, что прошлое и будущее независимы, когда известно настоящее. Это означает, что если кто-то знает текущее состояние процесса, то для создания наилучшего возможного прогноза не требуется никакой дополнительной информации о его прошлых состояниях в будущем.

Почему марковские процессы важны?

Многие аналитические методы и решения разрабатываются только для марковских процессов

Обычный способ решения проблем обработки сигналов - это адаптация их к некоторым марковским моделям

Марковские процессы адекватны многим реальным явлениям

Кроме того, некоторые реальные процессы могут быть аппроксимированы марковскими процессами

Отраслевые приложения цепей Маркова:

- Генерация текста
- Финансовое моделирование и прогнозирование (включая торговые алгоритмы).
- Прогнозирование временных рядов
- Логистика: моделирование будущих поставок или поездок, СМО.
- Поисковые системы: PageRank можно рассматривать как моделирующего случайного интернет-серфера с цепью Маркова.

Марковский случайный процесс можно определить также как последовательность испытаний, в каждом из которых появляется только одно из k несовместных событий Ai из полной группы.

При этом условная вероятность pij(s) того, что в s –ом испытании наступит событие Aj при условии, что в (s – 1) – ом испытании наступило событие Ai, не зависит от результатов предшествующих испытаний.

Независимые испытания являются частным случаем цепи Маркова. События называются **состояниями системы**, а испытания – **изменениями состояний системы.**

Марковский процесс удобно задавать графом переходов из состояния в состояние. Мы рассмотрим два варианта описания марковских процессов **— с дискретным и непрерывным временем.**

Случайный процесс называется процессом с **дискретными состояниями**, если возможные состояния системы S1, S2, S3, ... можно перечислить (перенумеровать) одно за другим, а сам процесс состоит в том, что время от времени система S скачком (мгновенно) переходит из одного состояния в другое.

Кроме процессов с дискретными состояниями существуют **случайные процессы с непрерывными состояниями:** для этих процессов характерен постепенный, плавный переход из состояния в состояние. Например, процесс изменения напряжения в осветительной сети представляет собой случайный процесс с непрерывными состояниями.

Если переходы системы из состояния в состояние возможны только в определенные моменты времени t1, t2, t3,…, то марковский процесс относится к процессам с дискретным временем. В противном случае имеет место процесс с непрерывным временем.

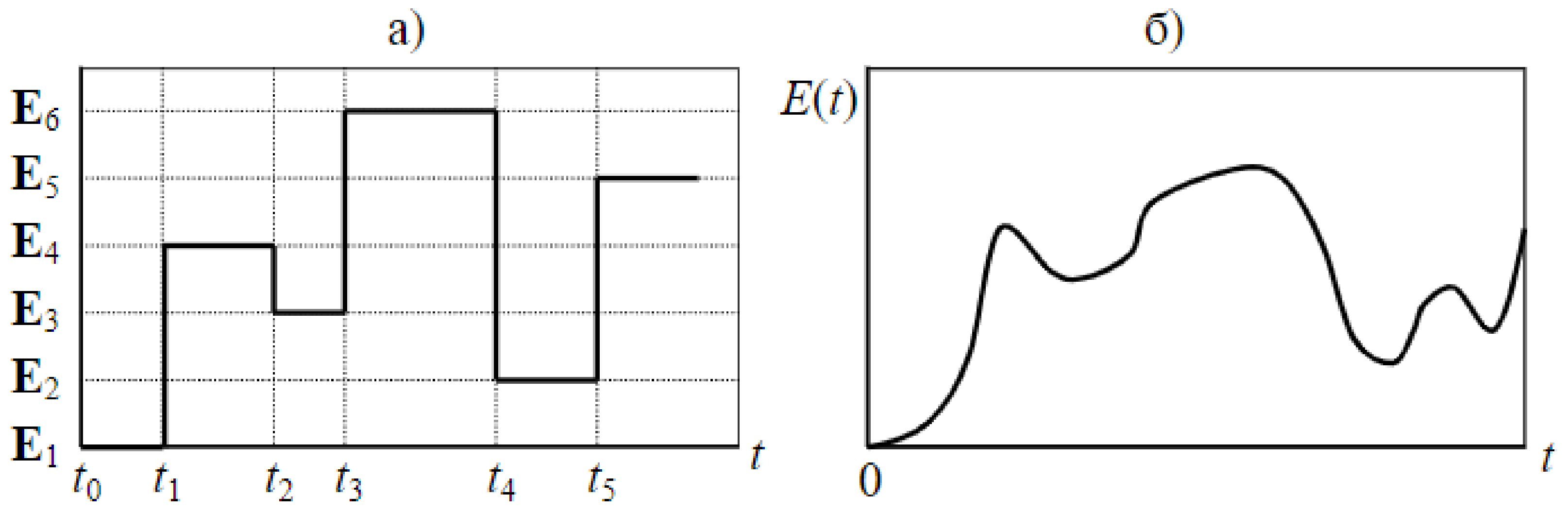
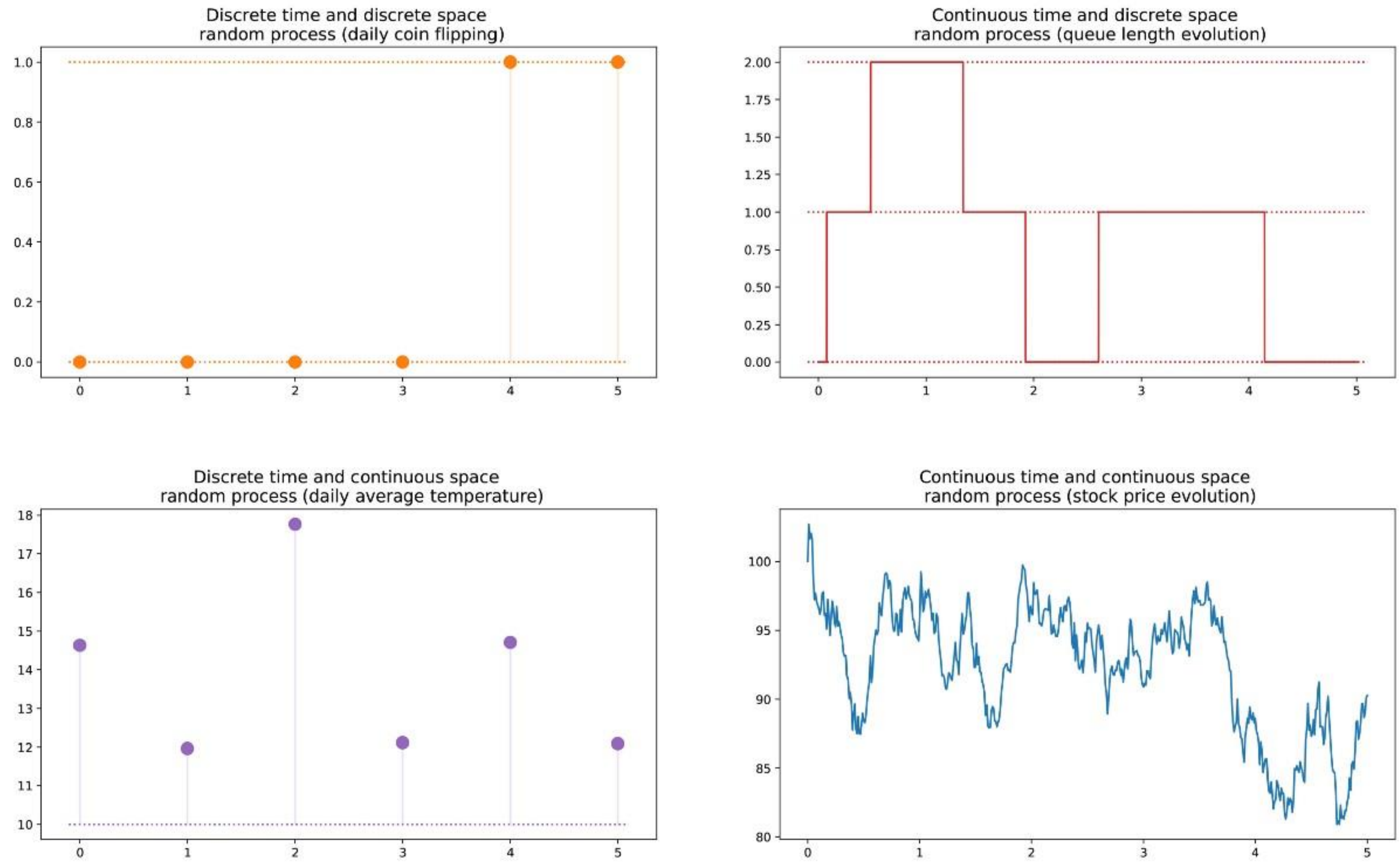
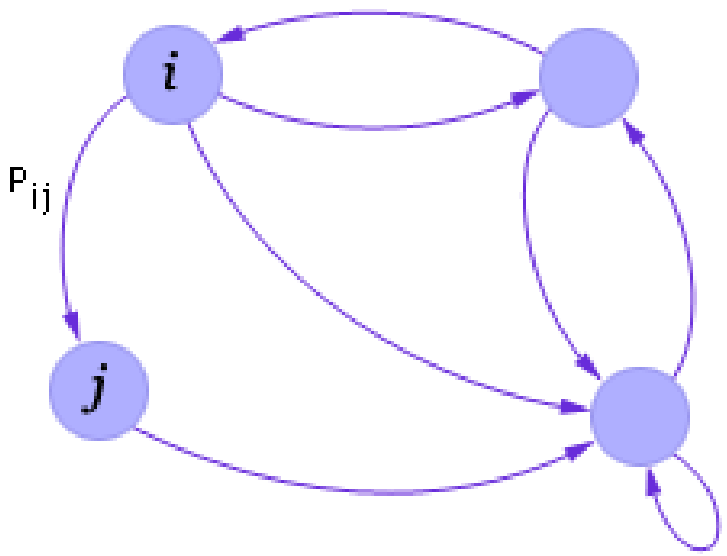


Рис.5.1. Процессы с дискретными (а) и непрерывными (б) состояниями



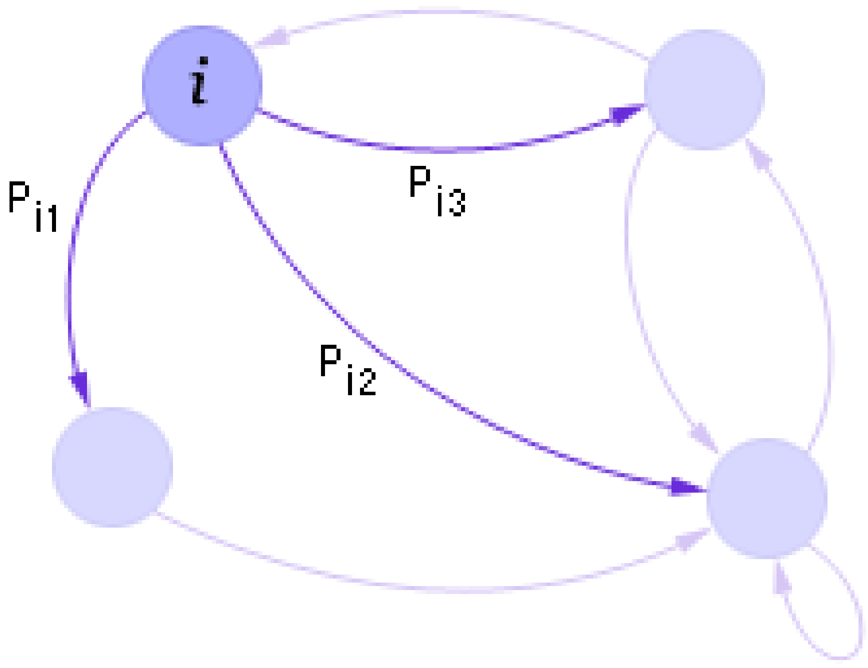
Стоимость опциона на бирже задаёт случайный процесс с непрерывным временем

Марковский процесс с дискретным временем



Пример графа переходов

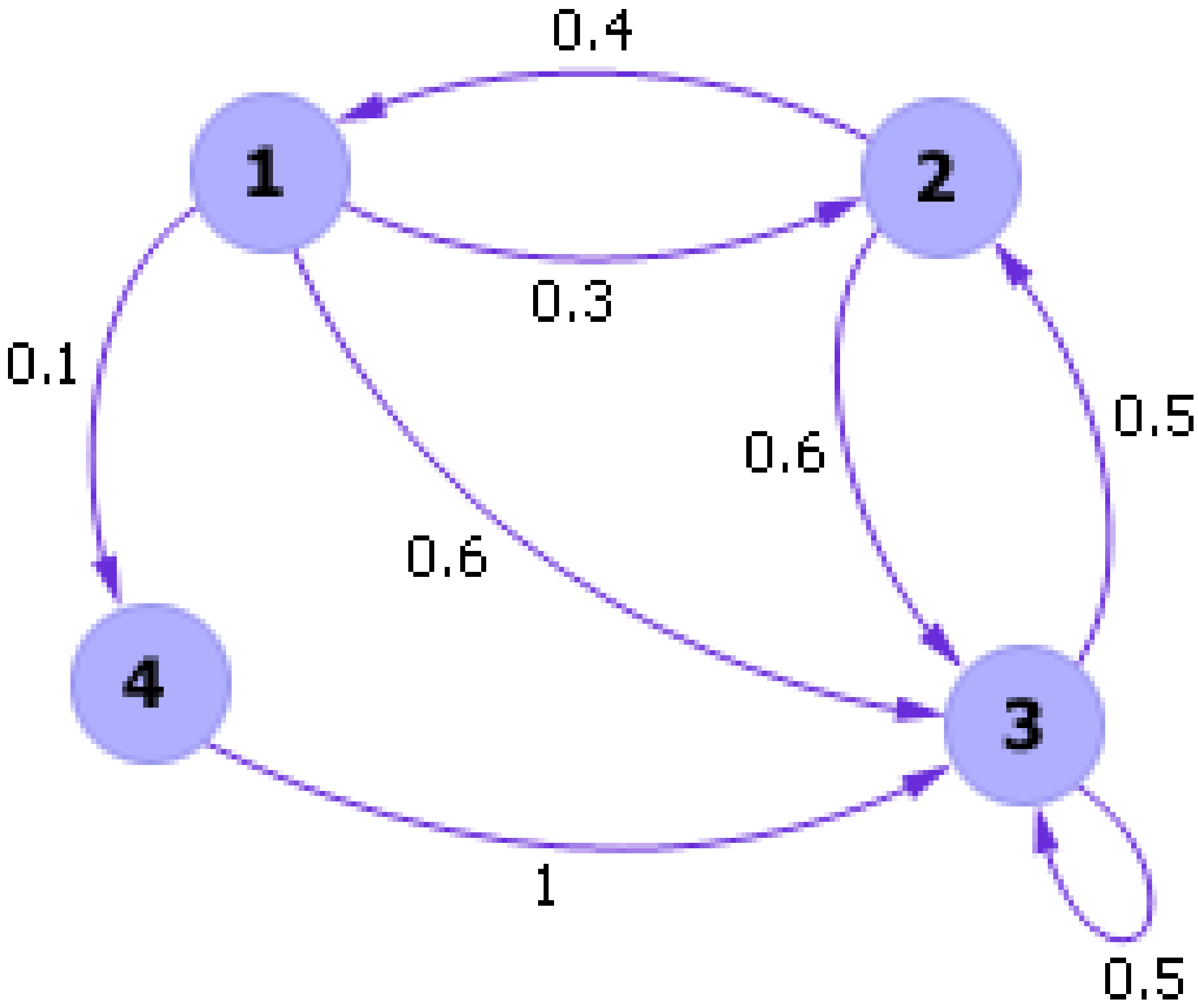
Каждый переход характеризуется вероятностью перехода Pij. Вероятность Pij показывает, как часто после попадания в i-е состояние осуществляется затем переход в j-е состояние. Конечно, такие переходы происходят случайно, но если измерить частоту переходов за достаточно большое время, то окажется, что эта частота будет совпадать с заданной вероятностью перехода.



$$\boxed{P_{i1} + P_{i2} + P_{i3} = 1}$$

Фрагмент графа переходов (переходы из i-го состояния являются полной группой случайных событий)

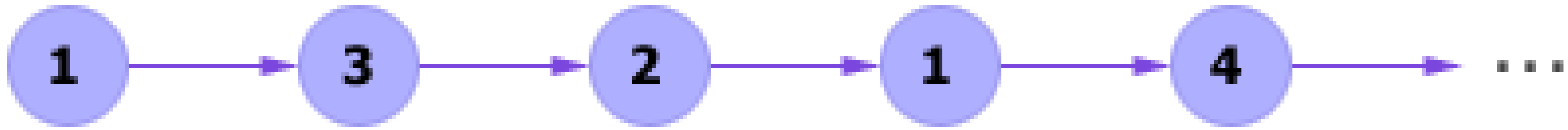
У каждого состояния сумма вероятностей всех переходов (исходящих стрелок) из него в другие состояния должна быть всегда равна 1



Пример марковского графа переходов

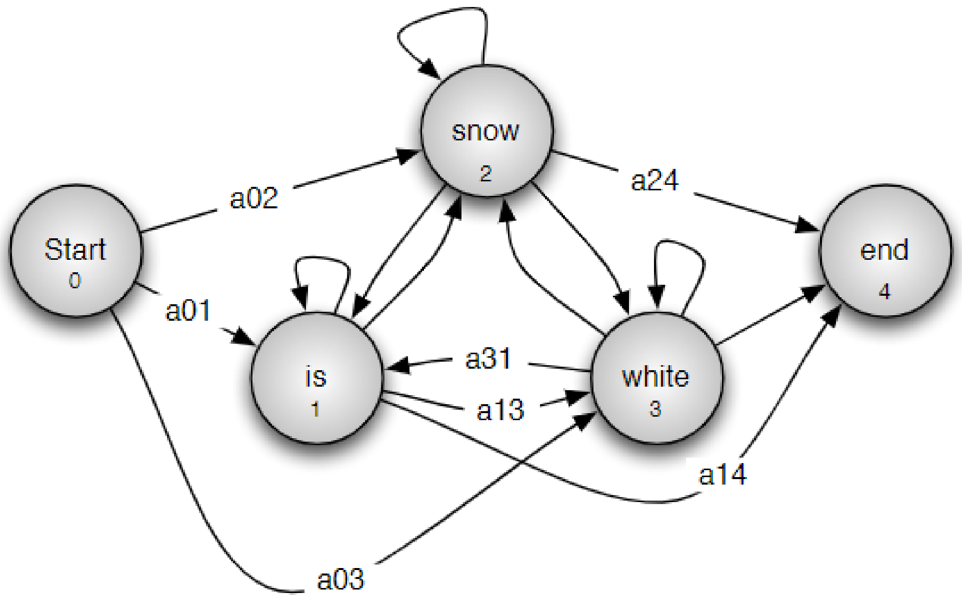
ЦЕПЬ МАРКОВА

Реализация марковского процесса (процесс его моделирования) представляет собой вычисление последовательности (цепи) переходов из состояния в состояние.



Цепь маркова

- Специальный случай взвешенного конечного автомата
- Входная последоваьельность уникально определяет состояния автомата
- $Q=q_1q_2..q_N$
- $A=a_{01}a_{02}\ldots a_{n1}\ldots a_{nn}$
- q_0,q_F



Цепь Маркова — это марковский процесс с дискретным временем и дискретным пространством состояний.

Цепь Маркова — это дискретная последовательность состояний, каждое из которых берётся из дискретного пространства состояний (конечного или бесконечного), удовлетворяющее марковскому свойству.

Математически мы можем обозначить цепь Маркова так:

$$X = (X_n)_{n \in \mathbb{N}} = (X_0, X_1, X_2, \ldots)$$

где в каждый момент времени процесс берёт свои значения из дискретного множества E, такого, что

$$X_n \in E \qquad \forall n \in \mathbb{N}$$

Тогда марковское свойство подразумевает, что у нас есть

$$\mathbb{P}(X_{n+1}=s_{n+1}|X_n=s_n, X_{n-1}=s_{n-1}, X_{n-2}=s_{n-2}, ...) = \mathbb{P}(X_{n+1}=s_{n+1}|X_n=s_n)$$

Последняя формула отражает тот факт, что для хронологии (где я нахожусь сейчас и где я был раньше) распределение вероятностей следующего состояния (где я буду дальше) зависит от текущего состояния, но не от прошлых состояний.

Пространство состояний Цепи Маркова может быть набором произвольных объектов

$S = \{\text{sleep, eat, excercise}\}$

$S = \{\text{sunny, rainy}\}$

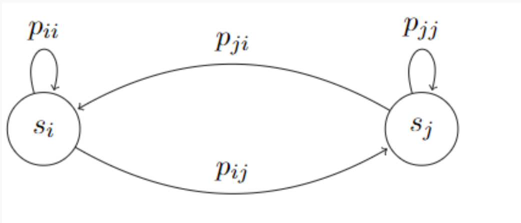
$S = \{\text{bear market, bull market}\}$

$$S = \{s_1, ..., s_k\},$$

Вероятности переходов :

$$p_{ij} = P(X_{n+1} = s_j \mid X_n = s_i), \quad i, j = 1, ..., k$$

Вероятность перехода P_{ij} из состояния si в состояние sj зависит только от i И j и не зависит от шага по времени n



Вероятности перехода упорядочены в матрицу перехода состояний или матрицу вероятностей перехода.

Матрицей перехода системы называется матрица, которая содержит все переходные вероятности этой системы:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{bmatrix}.$$

Так как эти состояния образуют *полную систему* событий, то сумма вероятностей каждой строки этой матрицы равна 1,

то есть $$\sum_{j=1}^{k} p_{ij} = 1 \qquad (i = 1,\ldots, k).$$

Такие матрицы называются *стохастическими*.

Пример: Пусть A_1, A_2, A_3 – три состояния, в котором может находится система.

Матрица перехода $P = \begin{bmatrix} 0{,}5 & 0{,}2 & 0{,}3 \\ 0{,}4 & 0{,}5 & 0{,}1 \\ 0{,}6 & 0{,}3 & 0{,}1 \end{bmatrix}$.

Здесь $p_{11} = 0{,}5$ вероятность перехода $i=1 \rightarrow j=1$.

$p_{21}=0{,}4$ – вероятность перехода $i=2 \rightarrow j=1$.

Markov State Diagram

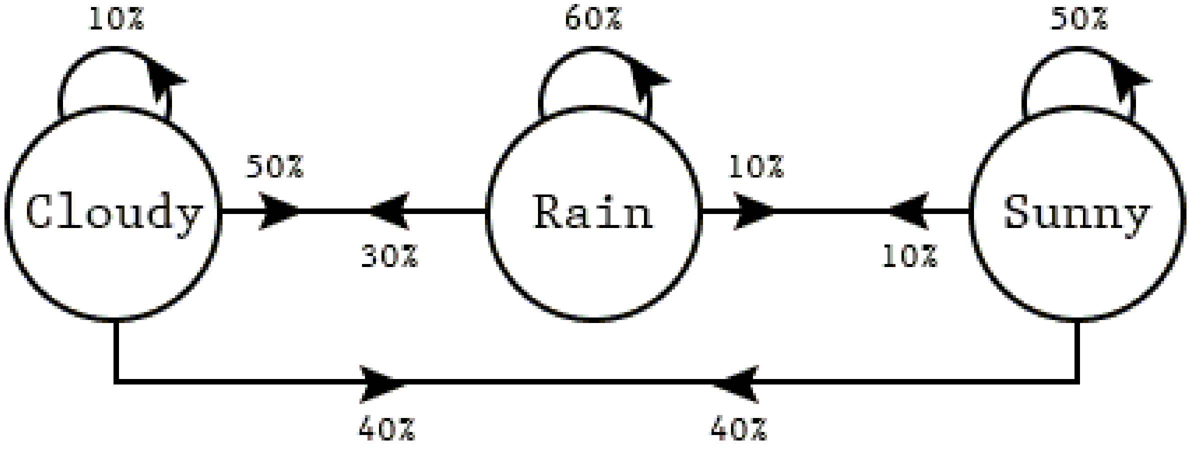


Figure 2

Transition Matrix

	C	R	S
C	0.1	0.5	0.4
R	0.3	0.6	0.1
S	0.4	0.1	0.5

Figure 3

ВИЗУАЛЬНЫЙ ПРИМЕР:

https://setosa.io/ev/markov-chains/

Определение. Пусть $p_{ij}(n)$ – вероятность того, что в результате n шагов система перейдёт из состояния i в состояние j.

Например. $P_{25}(10)$ – вероятность перехода за 10 шагов из второго состояния в пятое.

В частности, при n=1 получаем $p_{ij}(1)= p_{ij}$

Задача. Зная переходные вероятности p_{ij} , найти вероятности p_{ij} (n) перехода за n шагов из состояния i в состояние j.

Оказывается, что зная все породные вероятности p_{ij} , т.е. зная матрицу P перехода из состояния i в состояние j за один шаг, можно найти вероятность $p_{ij}(n)$ перехода из состояния i в состояние j за n шагов, n=2, 3, 4,…

Используя теорему о полной вероятности, получаем:

Теорема. Матрица перехода *P(n)* из состояния в состояние за *n* шагов равна:

$P(n) = P^n$. Мы полагаем, что *P(1)=P*.

Пример. Пусть $P = \begin{bmatrix} 0,4 & 0,6 \\ 0,3 & 0,7 \end{bmatrix}$. Тогда

$$P(2) = \begin{bmatrix} 0,4 & 0,6 \\ 0,3 & 0,7 \end{bmatrix}^2 = \begin{bmatrix} 0,34 & 0,66 \\ 0,33 & 0,67 \end{bmatrix}.$$

n-step transition matrices:

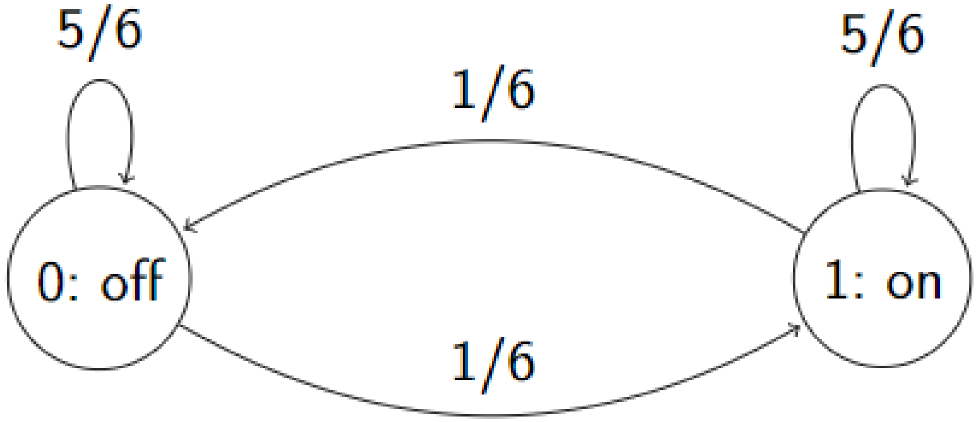
$$P^{(2)} = P^{(1)}P^{(1)} = PP = P^2$$

$$P^{(3)} = P^{(2)}P^{(1)} = P^2P = P^3$$

$$\dots$$

$$P^{(n)} = P^n, \quad n = 1, 2, \dots$$

State transition diagram:



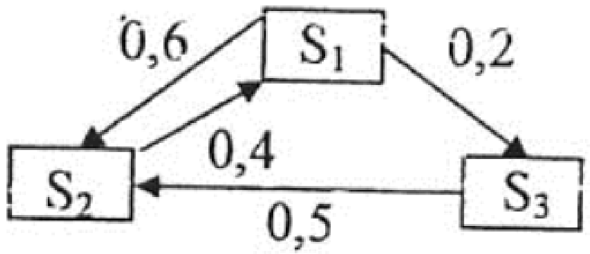
Transition probability matrix:

$$P = \begin{pmatrix} 5/6 & 1/6 \\ 1/6 & 5/6 \end{pmatrix} \approx \begin{pmatrix} 0.83 & 0.17 \\ 0.17 & 0.83 \end{pmatrix}$$

n-step transition matrices:

$$P^{(2)} \approx \begin{pmatrix} 0.72 & 0.28 \\ 0.28 & 0.72 \end{pmatrix}, \, P^{(3)} \approx \begin{pmatrix} 0.65 & 0.35 \\ 0.35 & 0.65 \end{pmatrix}, \, P^{(\infty)} = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$$

Задача. Дан размеченный граф состояний системы.



Найти:

А) матрицы перехода за один и два шага,

Б) вероятности состояний системы после первого, второго, третьего шага, если в начальный момент система находилась в состоянии S_1,

В) финальные вероятности.

Решение. Имеем вероятности (по графу состояний):

$p_{12} = 0,6$, $p_{13} = 0,2$, тогда $p_{11} = 1 - 0,6 - 0,2 = 0,2$.

$p_{21} = 0,4$, $p_{23} = 0$, тогда $p_{22} = 1 - 0,4 - 0 = 0,6$.

$p_{32} = 0,5$, $p_{31} = 0$, тогда $p_{33} = 1 - 0,5 - 0 = 0,5$.

Получаем матрицу перехода за один шаг:

$$P = \begin{pmatrix} 0,2 & 0,6 & 0,2 \\ 0,4 & 0,6 & 0 \\ 0 & 0,5 & 0,5 \end{pmatrix}.$$

Матрица перехода за два шага:

$$P_2 = P^2 = \begin{pmatrix} 0,2 & 0,6 & 0,2 \\ 0,4 & 0,6 & 0 \\ 0 & 0,5 & 0,5 \end{pmatrix} \cdot \begin{pmatrix} 0,2 & 0,6 & 0,2 \\ 0,4 & 0,6 & 0 \\ 0 & 0,5 & 0,5 \end{pmatrix} = \begin{pmatrix} 0,28 & 0,58 & 0,14 \\ 0,32 & 0,60 & 0,08 \\ 0,20 & 0,55 & 0,25 \end{pmatrix}.$$

Матрица перехода за три шага:

$$P_3 = P^3 = \begin{pmatrix} 0,2 & 0,6 & 0,2 \\ 0,4 & 0,6 & 0 \\ 0 & 0,5 & 0,5 \end{pmatrix} \cdot \begin{pmatrix} 0,28 & 0,58 & 0,14 \\ 0,32 & 0,60 & 0,08 \\ 0,20 & 0,55 & 0,25 \end{pmatrix} = \begin{pmatrix} 0,288 & 0,586 & 0,126 \\ 0,304 & 0,592 & 0,104 \\ 0,260 & 0,575 & 0,165 \end{pmatrix}.$$

Начальное состояние: $p_0 = (S_1) = (1;0;0)$.

Тогда

1) вероятности состояний системы после первого шага: $p_1 = p_0 P = (0,2;0,6;0,2)$ (первая строка матрицы P).

2) вероятности состояний системы после второго шага: $p_2 = p_0 P_2 = (0,28;0,58;0,14)$ (первая строка матрицы P_2).

3) вероятности состояний системы после третьего шага: $p_3 = p_0 P_3 = (0,288;0,586;0,126)$ (первая строка матрицы P_2).

Финальные вероятности найдем из условий: $p \cdot P = p,\ p = (p_1, p_2, p_3),\ \sum p_i = 1$. Получаем систему:

$$\begin{cases} 0,2p_1 + 0,4p_2 = p_1, \\ 0,6p_1 + 0,6p_2 + 0,5p_3 = p_2, \\ 0,2p_1 + 0,5p_3 = p_3, \\ p_1 + p_2 + p_3 = 1. \end{cases}$$

$$\begin{cases} -8p_1 + 4p_2 = 0, \\ 6p_1 - 4p_2 + 5p_3 = 0, \\ 2p_1 - 5p_3 = 0, \\ p_1 + p_2 + p_3 = 1. \end{cases}$$

$$\begin{cases} p_1 = 5/17, \\ p_2 = 10/17, \\ p_3 = 2/17. \end{cases}$$

ПРИМЕР

Рассмотрим повседневное поведение вымышленного посетителя сайта. В каждый день у него есть 3 возможных состояния: читатель не посещает сайт в этот день (N), читатель посещает сайт, но не читает пост целиком (V) и читатель посещает сайт и читает один пост целиком (R). Итак, у нас есть следующее пространство состояний:

$$E = \{N, V, R\}$$

Допустим, в первый день этот читатель имеет вероятность 50% только зайти на сайт и вероятность 50% посетить сайт и прочитать хотя бы одну статью. Вектор, описывающий исходное распределение вероятностей (n=0) тогда выглядит так:

$$q_0 = (0.0, 0.5, 0.5)$$

Также представим, что наблюдаются следующие вероятности:

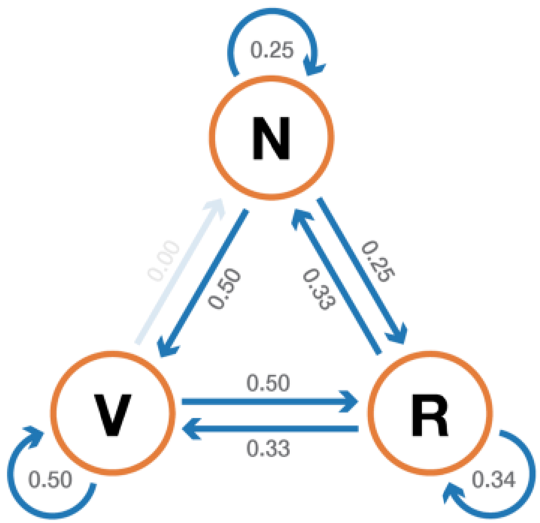
- когда читатель не посещает один день, то имеет вероятность 25% не посетить его и на следующий день, вероятность 50% только посетить его и 25% — посетить и прочитать статью
- когда читатель посещает сайт один день, но не читает, то имеет вероятность 50% снова посетить его на следующий день и не прочитать статью, и вероятность 50% посетить и прочитать
- когда читатель посещает и читает статью в один день, то имеет вероятность 33% не зайти на следующий день (надеюсь, этот пост не даст такого эффекта!), вероятность 33% только зайти на сайт и 34% — посетить и снова прочитать статью
- Тогда у нас есть следующая переходная матрица:

$$p = \begin{pmatrix} 0.25 & 0.50 & 0.25 \\ 0.00 & 0.50 & 0.50 \\ 0.33 & 0.33 & 0.34 \end{pmatrix}$$

Как вычислить для этого читателя вероятность каждого состояния на следующий день (n=1)?

$$q_1 = q_0 p = (0.0, 0.5, 0.5) \begin{pmatrix} 0.25 & 0.50 & 0.25 \\ 0.00 & 0.50 & 0.50 \\ 0.33 & 0.33 & 0.34 \end{pmatrix} = (0.165, 0.415, 0.420)$$

Вероятностную динамику этой цепи Маркова можно графически представить так:

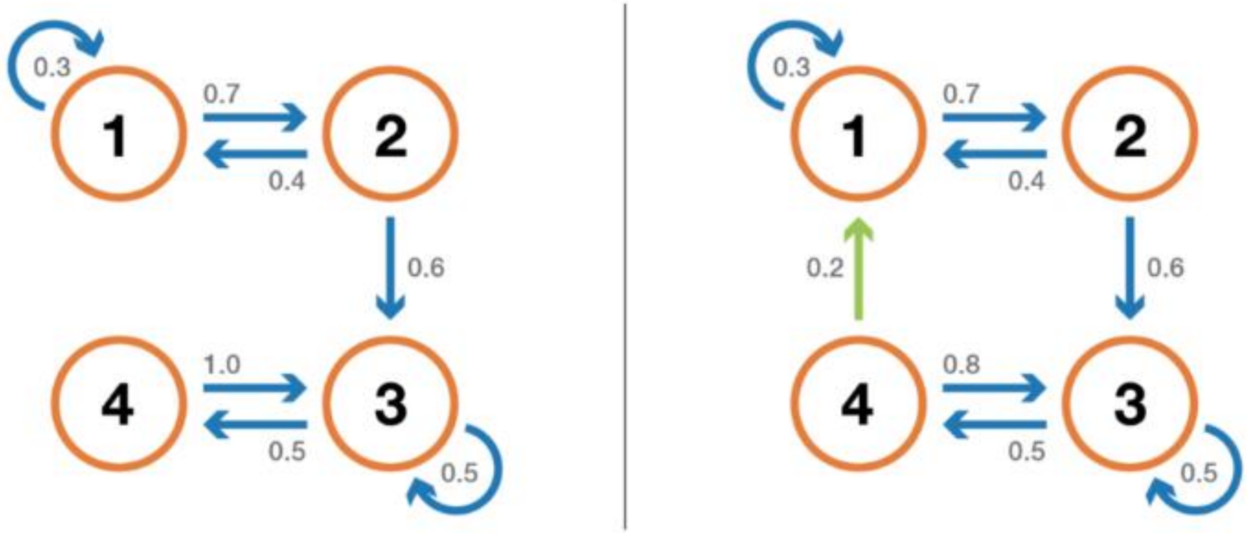


N
V
R
0.25
0.00
0.50
0.33
0.25
0.50
0.33
0.50
0.34

Свойства цепей Маркова

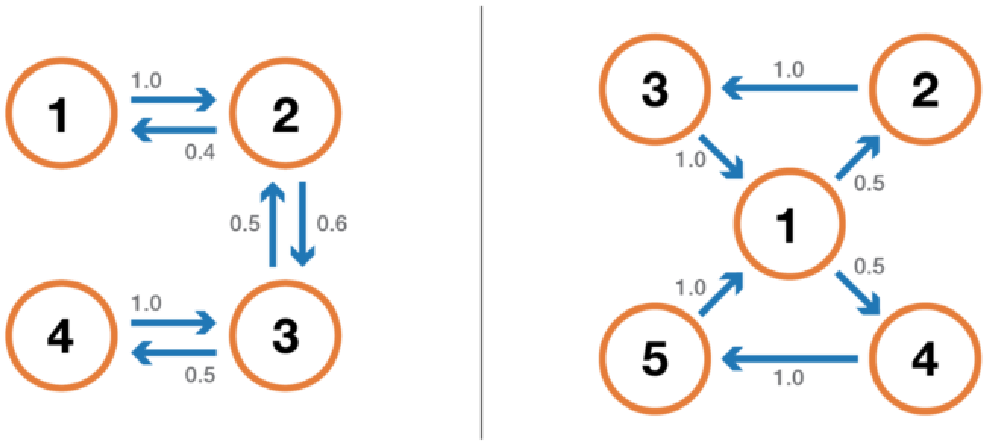
Разложимость, периодичность, невозвратность и возвратность

Цепь Маркова **неразложима**, если можно достичь любого состояния из любого другого состояния (необязательно, что за один шаг времени). Если пространство состояний конечно и цепь можно представить в виде графа, то мы можем сказать, что граф неразложимой цепи Маркова сильно связный (теория графов).



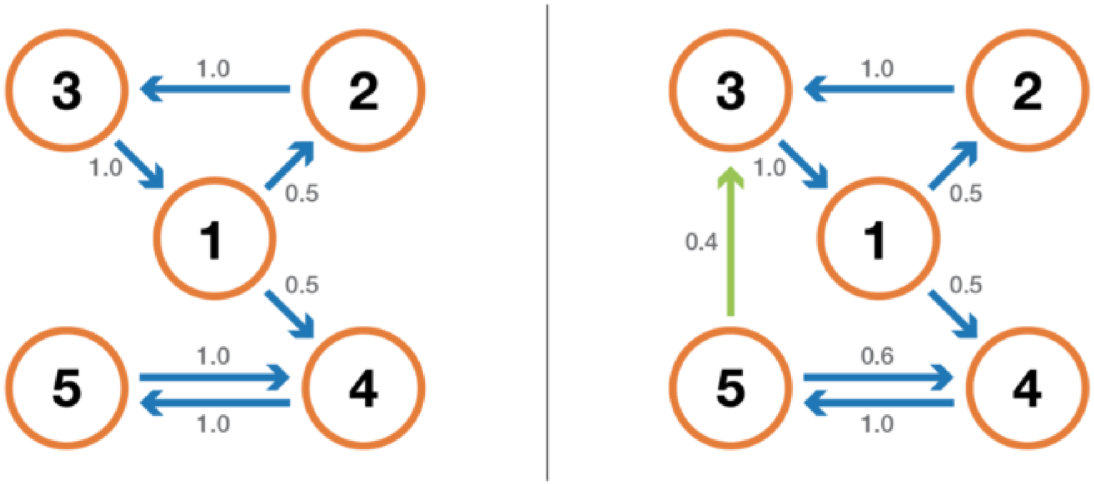
Цепь слева нельзя сократить: из 3 или 4 мы не можем попасть в 1 или 2. Цепь справа (добавлено одно ребро) можно сократить: каждого состояния можно достичь из любого другого.

Состояние имеет период k, если при уходе из него для любого возврата в это состояние нужно количество этапов времени, кратное k (k — наибольший общий делитель всех возможных длин путей возврата). Если k = 1, то говорят, что состояние является апериодическим, а вся цепь Маркова является **апериодической**, если апериодичны все её состояния. В случае неприводимой цепи Маркова можно также упомянуть, что если одно состояние апериодическое, то и все другие тоже являются апериодическими.



Цепь слева периодична с k=2: при уходе из любого состояния для возврата в него всегда требуется количество шагов, кратное 2. Цепь справа имеет период 3.

Состояние является **невозвратным**, если при уходе из состояния существует ненулевая вероятность того, что мы никогда в него не вернёмся. И наоборот, состояние считается **возвратным**, если мы знаем, что после ухода из состояния можем в будущем вернуться в него с вероятностью 1 (если оно не является невозвратным).



Цепь слева имеет такие свойства: 1, 2 и 3 невозвратны (при уходе из этих точек мы не можем быть абсолютно уверены, что вернёмся в них) и имеют период 3, а 4 и 5 возвратны (при уходе из этих точек мы абсолютно уверены, что когда-нибудь к ним вернёмся) и имеют период 2. Цепь справа имеет ещё одно ребро, делающее всю цепь возвратной и апериодической.