ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. ЛОМОНОСОВА»

МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ

КАФЕДРА МАТЕМАТИЧЕСКОЙ ЛОГИКИ И ТЕОРИИ АЛГОРИТМОВ

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
(ДИПЛОМНАЯ РАБОТА)
магистра

ПРИМЕНЕНИЕ ПРИНЦИПА РИССАНЕНА MDL ДЛЯ МАРКОВСКИХ ЦЕПЕЙ

Выполнила студентка
группы М2 ЦТиИИ
Ремизова Анна Петровна

____
подпись студента

Научный руководитель:
д. ф.-м. н. профессор
Верещагин Николай Константинович

____
подпись научного руководителя

Москва

2023

Содержание

1 Введение

Основные термины

Актуальность, научная и практическая значимость работы

Постановка задачи

Есть данные – последовательность 0 и 1, мы хотим подобрать марковскую цепь, для которой наибольшая вероятность получить заданную траекторию. Например, эта последовательность может описывать некоторый текст, набор чисел и так далее. По Риссанену [1], если мы хотим предсказать, что будет дальше, то должны сравнивать друг с другом гипотезы по их сложности, причём даём преимущество простым гипотезам. Выражения для Description length будет выглядеть следующим образом:

$$\mathcal{D}(\mathcal{M}, x) = C(\mathcal{M}) + \log_2 \frac{1}{P_{\mathcal{M}}(x)} \tag{1}$$

где $\mathcal{M}$ – выбранная модель, $C(\mathcal{M})$ – сложность модели (complexity), $P_{\mathcal{M}}(x)$ – вероятность в модели $\mathcal{M}$ получить реализацию x.

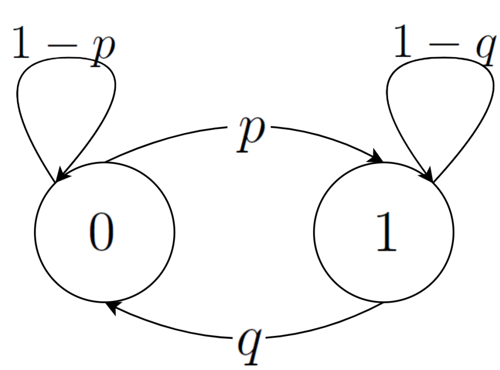
Цели и методы исследования

Обзор литературы

2 Марковские цепи с 2 состояниями

Для начала рассмотрим простые марковские цепи. Пусть марковская цепь состоит из 2 состояний (Рис. (1)). Дана последовательность состояний Марковской цепи из 2 состояний: 0 и 1. Найдём оптимальные переходные вероятности p из 0 в 1 и q из 1 в 0 по принципу Риссанена MDL. В данной задаче рассматриваются однородные цепи Маркова с конечным числом состояний, это означает, что переходные вероятности не зависят от номера шага, а зависят только от того состояния, в котором сейчас находится Марковская цепь.

Рис. 1: Марковская цепь с 2 состояниями



Для решения этой задачи запишем вероятность получения заданной реализации: пусть $n(ij)$ – число переходов из состояния i в состояние j, тогда вероятность получить реализацию x цепи $\mathcal{M}$:

$$P_{\mathcal{M}}(x) = p^{n(01)} \cdot (1-p)^{n(00)} \cdot q^{n(10)} \cdot (1-q)^{n(11)} \to max \tag{2}$$

$$\mathcal{L}(\mathcal{M}, x) = \log_2 \frac{1}{P_{\mathcal{M}}(x)} = -(n(01) \cdot \log_2 p + n(00) \cdot \log_2 (1-p) + n(10) \cdot \log_2 q + n(11) \cdot \log_2 (1-q)) \tag{3}$$

Сложность $C(\mathcal{M})$ будем определять как суммарную длину дробной части записи p и q в двоичной системе счисления, так как $0 \leq p, q \leq 1$. Пусть вероятность p длины k, q – длины l, тогда $C(\mathcal{M}) = k + l$. Далее рассмотрим несколько реализаций Марковских цепей и исследуем, как меняются значения в зависимости от k и l.

2.1 Таблицы с двоичными значениями

В Таблицах (1, 2, 3) в каждой ячейке представлены сначала оптимальные (минимальные, т.к. ищем минимальную описательную длину) значения $\mathcal{L}(\mathcal{M}, x)$ (3), затем сложность по Риссанену, а после - значения p и q, при которых оно достигается, представленные в двоичной системе счисления, для марковских цепей с траекториями, соответствующими 30 первым знакам $\pi, \sqrt{2}, \sqrt{3}$ соответственно. По горизонтали отмечены значения l - длина перебираемых q в двоичной системе, по вертикали - значения k - длина перебираемых p в двоичной системе.

Выводы к Таблице (1) для π: заметим, что при фиксированной длине l (по столбцам) двоичной записи переходной вероятности q оптимальное значение q неизменно, но при этом с увеличением k оптимальное значение логарифма уменьшается. Аналогично для фиксированного k (по строкам).

Выводы к Таблице (2): для $\sqrt{2}$ практически то же, что и для π.

Выводы к Таблице (3): для $\sqrt{3}$ результаты уже отличаются от π, но наблюдаются те же закономерности. Отличие $\sqrt{3}$ от π и $\sqrt{2}$ в количестве диграмм в их двоичной записи, были рассмотрены первые 30 знаков для каждого числа, не считая точки. Если для π и $\sqrt{2}$ распределение количества диграмм близко к равномерному, то для $\sqrt{3}$ оно менее сбалансированно: количество диграмм 00 меньше остальных, а диграмм 11 - больше (см.Таблицу 4).

Утверждение 1 *Оптимальное значение p не зависит от q и наоборот, оптимальное значение q не зависит от p.*

Доказательство. Рассмотрим выражение (3) для логарифма. Значения $n(00), n(01), n(10), n(11)$ – постоянные, и данное выражения можно представить в виде линейной комбинации двух функций $f_1(p) + f_2(q)$. Соотвественно, при максимизации всего выражения (логарифм (3) должен быть маленьким, а так как перед всем выражением стоит минус, то выражение в скобках должно быть большим), так как переменные p и q содержатся в отдельных слагаемых, необходимо найти минимум отдельно для $f_1(p)$ и $f_2(q)$, друг на друга их значения при минимизации не влияют. ■

2.2 Анализ диграмм

В Таблице (4) представлены количества диграмм по рассмотренным примерам - их сумма в каждом случае равна 29, так как рассматриваемые числа округлялись до 30 знаков в двоичной записи суммарно, далее оптимальные значения $k, l, \log_2 \frac{1}{P_{\mathcal{M}}(x)}, MDL$, найденные при $k, l \in [1, 6]$ для минимизации MDL.

Таблица 1: Таблица оптимальных зн-й p и q в двоичной записи для π

k / l	1	2	3	4	5	6
1	31.0 29.0 0.1 0.1	32.0 29.0 0.1 0.10	33.0 29.0 0.1 0.100	33.9891 28.9891 0.1 0.1001	34.9521 28.9521 0.1 0.10001	35.9521 28.9521 0.1 0.100010
2	32.0 29.0 0.10 0.1	33.0 29.0 0.10 0.10	34.0 29.0 0.10 0.100	34.9891 28.9891 0.10 0.1001	35.9521 28.9521 0.10 0.10001	36.9521 28.9521 0.10 0.100010
3	33.0 29.0 0.100 0.1	34.0 29.0 0.100 0.10	35.0 29.0 0.100 0.100	35.9891 28.9891 0.100 0.1001	36.9521 28.9521 0.100 0.10001	37.9521 28.9521 0.100 0.100010
4	34.0 29.0 0.1000 0.1	35.0 29.0 0.1000 0.10	36.0 29.0 0.1000 0.100	36.9891 28.9891 0.1000 0.1001	37.9521 28.9521 0.1000 0.10001	38.9521 28.9521 0.1000 0.100010
5	35.0 29.0 0.10000 0.1	36.0 29.0 0.10000 0.10	37.0 29.0 0.10000 0.100	37.9891 28.9891 0.10000 0.1001	38.9521 28.9521 0.10000 0.10001	39.9521 28.9521 0.10000 0.100010
6	36.0 29.0 0.100000 0.1	37.0 29.0 0.100000 0.10	38.0 29.0 0.100000 0.100	38.9891 28.9891 0.100000 0.1001	39.9521 28.9521 0.100000 0.10001	40.9521 28.9521 0.100000 0.100010

Утверждение 2 *Значения* $p = \dfrac{n(01)}{n(01)+n(00))}$ *и* $q = \dfrac{n(10)}{n(10)+n(11)}$ *являются точкой максимума функции* $\log_2 \dfrac{1}{P_{\mathcal{M}}(x)}$ *(3). Их значения при заданных длинах двоичной записи* k *и* l *- это приближения оптимальных значений* p *и* q *числами вида* $x = \dfrac{n}{2^k}$ *и* $x = \dfrac{n}{2^l}$ *соответственно.*

Доказательство

1. Найдём точку максимума функции $f_1(p) = n(01)\log_2 p + n(00)\log_2(1-p)$. Её производная: $f_1'(p) = \dfrac{n(01)}{p\ln 2} - \dfrac{n(00)}{(1-p)\ln 2}$, критические точки $p = \dfrac{n(01)}{n(01)+n(00)}, p = 0, p = 1$. Т.к. $0 \le \dfrac{n(01)}{n(01)+n(00)} \le 1$, то $f_1'(p)$ отрицательна на $p \in (-\infty; 0) \cup \left(\dfrac{n(01)}{n(01)+n(00)}; 1\right)$, положительная на остальных промежутках, а значит точка максимума $p = \dfrac{n(01)}{n(01)+n(00)}$, если это значение отлично от 0 и 1, и $p = 0$ иначе. Аналогично для $f_2(q)$ точкой максимума является $q = \dfrac{n(10)}{n(10)+n(11)}$, либо $q = 0$.

Таблица 2: Таблица оптимальных зн-й p и q в двоичной записи для $\sqrt{2}$

k / l	1	2	3	4	5	6
1	31.0 29.0 0.1 0.1	32.0 29.0 0.1 0.10	32.9148 28.9148 0.1 0.101	33.7965 28.7965 0.1 0.1001	34.7965 28.7965 0.1 0.10010	35.795 28.795 0.1 0.100101
2	32.0 29.0 0.10 0.1	33.0 29.0 0.10 0.10	33.9148 28.9148 0.10 0.101	34.7965 28.7965 0.10 0.1001	35.7965 28.7965 0.10 0.10010	36.795 28.795 0.10 0.100101
3	33.0 29.0 0.100 0.1	34.0 29.0 0.100 0.10	34.9148 28.9148 0.100 0.101	35.7965 28.7965 0.100 0.1001	36.7965 28.7965 0.100 0.10010	37.795 28.795 0.100 0.100101
4	33.9891 28.9891 0.0111 0.1	34.9891 28.9891 0.0111 0.10	35.9039 28.9039 0.0111 0.101	36.7856 28.7856 0.0111 0.1001	37.7856 28.7856 0.0111 0.10010	38.7842 28.7842 0.0111 0.100101
5	34.9521 28.9521 0.01111 0.1	35.9521 28.9521 0.01111 0.10	36.8669 28.8669 0.01111 0.101	37.7485 28.7485 0.01111 0.1001	38.7485 28.7485 0.01111 0.10010	39.7471 28.7471 0.01111 0.100101
6	35.9521 28.9521 0.011110 0.1	36.9521 28.9521 0.011110 0.10	37.8669 28.8669 0.011110 0.101	38.7485 28.7485 0.011110 0.1001	39.7485 28.7485 0.011110 0.10010	40.7471 28.7471 0.011110 0.100101

2. Рассмотрим функцию вероятности $P(x) = x^p(1-x)^{1-p}$. Найдём её вторую производную: $P'(x) = (1-x)^{-p}(p-x)x^{p-1}, P''(x) = (p-1)p(1-x)^{-p-1}x^{p-2}$. При фиксированном p $P''(x)$ имеет нули в точках $x = 0, x = 1$ и $P''(x) \geq 0$ на $x \in [0; 1]$, а значит на этом интервале исходная функция выпукла вверх – см. Рис. (2). Кроме того, её максимальное значение достигается при $x = p$. Так как при фиксированном k мы рассматриваем двоичные числа с k знаками после запятой, то $x = \dfrac{n}{2^k}, n \in \mathbb{N}$. Соответственно, оптимальным будет именно приближение точки максимума функции $P(x) : x = p$, а будет это приближение с избытком или недостатком – зависит от того, какое из чисел будет ближе к p по значению функции $P(x)$.

3. Для π оптимальные p и q, вычисленные по указанным формулам, выглядят следующим образом: $p_0 = 0.5_{10} = 0.1_2, q_0 = 0.5(3)_{10} = 0.(1000)_2$, чему удовлетворяют значения из Таблицы (1).

   Для $\sqrt{2}$ имеем: $p_0 = 0.4(6)_{10} = 0.(0111)_2, q_0 = 0.(571428)_{10} = 0.(100)_2$ – по Таблице (2) совпадает q, но не совпадает, на первый взгляд, p. Но значение, представленное в таблице, к примеру, для $k = 6$ – это $0.011110_2 = \dfrac{30}{64}_{10}$, а значение, равное округ-

Таблица 3: Таблица оптимальных зн-й p и q в двоичной записи для $\sqrt{3}$

k / l	1	2	3	4	5	6
1	31.0 29.0 0.1 0.1	32.0 29.0 0.1 0.10	33.0 29.0 0.1 0.100	33.8419 28.8419 0.1 0.0111	34.8419 28.8419 0.1 0.01110	35.8419 28.8419 0.1 0.011100
2	31.9053 28.9053 0.11 0.1	32.9053 28.9053 0.11 0.10	33.9053 28.9053 0.11 0.100	34.7472 28.7472 0.11 0.0111	35.7472 28.7472 0.11 0.01110	36.7472 28.7472 0.11 0.011100
3	32.4067 28.4067 0.101 0.1	33.4067 28.4067 0.101 0.10	34.4067 28.4067 0.101 0.100	35.2486 28.2486 0.101 0.0111	36.2486 28.2486 0.101 0.01110	37.2486 28.2486 0.101 0.011100
4	33.4067 28.4067 0.1010 0.1	34.4067 28.4067 0.1010 0.10	35.4067 28.4067 0.1010 0.100	36.2486 28.2486 0.1010 0.0111	37.2486 28.2486 0.1010 0.01110	38.2486 28.2486 0.1010 0.011100
5	34.4067 28.4067 0.10100 0.1	35.4067 28.4067 0.10100 0.10	36.4067 28.4067 0.10100 0.100	37.2486 28.2486 0.10100 0.0111	38.2486 28.2486 0.10100 0.01110	39.2486 28.2486 0.10100 0.011100
6	35.4029 28.4029 0.101001 0.1	36.4029 28.4029 0.101001 0.10	37.4029 28.4029 0.101001 0.100	38.2448 28.2448 0.101001 0.0111	39.2448 28.2448 0.101001 0.01110	40.2448 28.2448 0.101001 0.011100

Таблица 4: Числа, количество диграмм в них, оптимальные k и l

Число	$n(00)$	$n(01)$	$n(10)$	$n(11)$	k	l	$\log_2 \dfrac{1}{P_{\mathcal{M}(x)}}$	MDL
π	7	7	8	7	1	1	29.0	31.0
$\sqrt{2}$	8	7	8	6	1	1	29.0	31.0
$\sqrt{3}$	4	7	8	10	1	1	29.0	31.0

лению до 6 знаков после запятой найденного по формуле p – это $0.011101_2 = \dfrac{29}{64}_{10}$
И действительно, если обозначить p_0 - найденное по формуле, то $\dfrac{29}{64} < p_0 < \dfrac{30}{64}$ и $0.00019 \approx \left| f_1(p_0) - f_1\left(\dfrac{30}{64}\right)\right| < \left| f_1(p_0) - f_1\left(\dfrac{29}{64}\right)\right| \approx 0.00800.$

Для $\sqrt{3}$ имеем: $p_0 = 0.(63)_{10} = 0.(1010001011)_2, q_0 = 0.(4)_{10} = 0.(011100)_2$ – по Таблице (3) также совпадает q, и также совпадает p с верхним приближением: для $k = 6$, к примеру, $0.101001_2 = \frac{41}{64}_{10}$, $0.101000_2 = \frac{40}{64}_{10}$, $\frac{40}{64} < p_0 < \frac{41}{64}$.

Таким образом, на представленных примерах всё согласуется с утверждением. ■

Заметим по Таблице (4), что для рассмотренных трёх случаев высокая точность переходных вероятностей p и q не выгодна по Риссанену, Minimal description length достигается при $k = l = 1$. Это будет не так, если при увеличении на 1 бит точности переходной вероятности логарифм будет уменьшаться больше, чем на 1. Т.е. количество диграмм $n(00)$ и $n(01)$ должно быть сильно не сбалансированно.

Рассмотрим различные значения $n(01)$ для $n(00) = 1$ и найдём, при каком k достигается MDL. Алгоритм: берём $p = \frac{n(01)}{n(01) + n(00))}$, переводим в двоичную систему с k знаками после запятой, считаем для каждого $Descriptionlength = k - (n(01)\log_2 p + n(00)\log_2(1-p))$ и ищем минимальное такое при различных $k \in [1, 100]$ В Таблицах (5) видно, что результаты k уже нетривиальные – мы нашли те примеры последовательностей, для которых оптимальная модель подразумевает достаточно точные значения переходных вероятностей. Симметричная ситуация будет наблюдаться и для l.

Таблица 5: Оптимальные k для разных n(01) – отличие от n(00) в 2^x раз

n(00)	n(01)	k	p двоичная	p десятичная	$\log_2 \frac{1}{P_{\mathcal{M}}(x)}$	MDL
1	4	2	0.11	0.8	3.6601	5.6601
1	8	2	0.11	0.8889	5.3203	7.3203
1	16	3	0.111	0.9412	6.0823	9.0823
1	32	4	0.1111	0.9697	6.9795	10.9795
1	64	5	0.11111	0.9846	7.9314	12.9314
1	128	6	0.111111	0.9922	8.9082	14.9082
1	256	7	0.1111111	0.9961	9.8967	16.8967
1	512	8	0.11111111	0.9981	10.891	18.891
1	1024	9	0.111111111	0.999	11.8882	20.8882

Заметим, что для всех кодов в общем случае верно соотношение $\sum_b n(ab) = \sum_b n(ba)$ для всех букв a, кроме первой и последней (для них левая и правая части могу отличаться на 1) [2, стр. 147]. Тогда над двоичным алфавитом условие будет выглядеть несколько проще, т.е. будет выполнено одно из соотношений:

а) $n(10) = n(01)$, если код начинается и заканчивается одной и той же буквой;

б) $n(10) = n(01) - 1$, если код начинается с 0, а заканчивается 1;

в) $n(10) = n(01) + 1$, если код начинается с 1, а заканчивается 0.

При этом $n(00), n(11)$ произвольны.

Так как в рассмотренных выше примерах кодами, реализациями Марковской цепи являлись первые 30 знаков двоичного представления чисел: $\pi \approx 11.001001000011111101101010100 0_2$, $\sqrt{2} \approx 1.0110101000001001111001100110 0_2$, $\sqrt{3} \approx 1.1011101101100111101011101000 0$ – все они начинаются с 1, а заканчиваются 0, то для них как раз выполнено соотношение (в) (Табл. (4)).

Таблица 6: Оптимальные k для разных n(01) – отличие от n(00) в x раз

n(00)	n(01)	k	p двоичная	p десятичная	$\log_2 \frac{1}{P_{\mathcal{M}}(x)}$	MDL
100000	200000	9	0.101010101	0.6667	275489.1627	275498.1627
100000	300000	2	0.11	0.75	324511.2498	324513.2498
100000	400000	8	0.11001101	0.8	360965.426	360973.426
100000	500000	9	0.110101011	0.8333	390014.7766	390023.7766
100000	600000	9	0.110110111	0.8571	414171.2664	414180.2664
100000	700000	3	0.111	0.875	434851.5546	434854.5546
100000	800000	9	0.111000111	0.8889	452932.8105	452941.8105
100000	900000	9	0.111001101	0.9	468996.8194	469005.8194
100000	1000000	10	0.1110100011	0.9091	483446.7613	483456.7613

Соответственно, при задании реализации Марковской цепи с помощью количеств диграмм каждого вида мы вправе задать 3 свободные переменные $n(00), n(01), n(11)$, а $n(10)$ задать одним из 3 перечисленных выше способов. Для простоты будем считать, что $n(10) = n(01)$, и построим график значений MDL в зависимости от $k + l$ при $n(00) = n(11) = 1$.

Выводы

3 Марковские цепи с 4 состояниями

Рассмотрим более сложную марковскую цепь с четырьмя состояниями, но по-прежнему над двоичным алфавитом. Пусть в цепи четыре состояния a, b, c, d, при посещении части их них печатается 1, при посещении остальных – 0. Для некоторых последовательностей, в частности для тех, которые имеют период 4, например $x = 00010001\ldots0001$, такая модель будет давать более оптимальный с точки зрения MDL результат: возьмём Марковскую цепь $\mathcal{M}$ с состояниями $a = b = c = 0, d = 1$, (Рис. (3)) тогда при переходных вероятностях $p(a, b) = p(b, c) = p(c, d) = p(d, a) = 1$, остальных $p(i, j) = 0$ получаем $P_{\mathcal{M}}(x) = 1, \log_2 \frac{1}{P_{\mathcal{M}}(x)} = 0$, а значит сложность $\mathcal{D}(\mathcal{M}, x) = C(\mathcal{M}) = 16$ – столько бит нужно для хранения информации и всех 16 переходных вероятностях за 1 шаг (это 0 и 1 – по 1 биту), и это, очевидно, MDL для данной реализации цепи. Вообще говоря, для описания данной модели также нужно дополнительно 4 бит информации на описание чисел для каждого из 4 состояний, также 2 бит на начальное состояние (номер состояния – число от 0 до 4), но этот размер памяти одинаков для всех моделей из 4 состояний, поэтому при сравнении моделей его можно не учитывать.

Рассмотрим обратную задачу: зададим Марковскую цепь, случайным блужданием по ней получим некоторую реализацию и для неё найдём с помощью перебора оптимальную с точки зрения MDL модель. Сравним полученную подель с исходной. В случае Марковской цепи с 4 состояниями вероятность в модели $\mathcal{M}$ получить реализацию x вычисляется по рекурсивной формуле:

$$P_a(x) = \begin{cases} \frac{1}{k}\sum_a P_b(x[1:]) & \text{, a - первый символ} \\ \sum_b p(a,b) & \text{, a - предпоследний символ} \\ \sum_b p(a,b)\cdot P_b(x[1:]) & \text{, иначе} \end{cases} \quad (4)$$

где k – количество состояний Марковской цепи, выдающих первый символ строки, т.е. начальное состояние выбираем равновероятно, b – состояние Марковской цепи, выдающее символ, следующий за символом, выдаваемым состоянием a, p(a,b) – переходная вероятность из состояния a в состояние b, x[1:] – строка x без первого символа.

Рассмотрим произвольную нетривиальную матрицу P переходных вероятностей для 4 состояний (Табл. (7)), в ней элемент p_{ij} – вероятность перехода из состояния i в состояние j. Такая матрица обладает следующим свойством: сумма элементов по строкам равна 1. Пусть в состояниях a и b печатается 0, в состояниях c и d – 1. Обозначим такую модель за $\mathcal{M}$.

Таблица 7: Матрица переходных вероятностей для 4 состояний

	a	b	c	d
a	1/4	1/4	1/2	0
b	1/2	1/4	0	1/4
c	1/4	1/2	0	1/4
d	1/4	0	1/4	1/2

Пусть $k = 20$ – количество диграмм в слове, случайным блужданием по представленной Марковской цепи была получена реализация $x = 011110100010000111110$ длины $k+1$, вероятность получить её в модели $\mathcal{M} : P(x) \approx 1.97 \cdot 10^{-6}$.

Переберём модели класса: в состояниях a и b печатается 0, в состояниях c и d – 1, с разными матрицами переходных вероятностей. Пусть вероятности - числа от 0 до 1 включительно с шагом 0.1, в каждой из 4 строк мы задаём 3 переходных вероятности, четвёртую вычисляем из соображения о том, что в сумме по строке 4 вероятности дают 1. Таким образом, в матрице переходных вероятностей P у нас 12 свободных переменных.

Выводы

4 Заключение

Список литературы

[1] Rissanen J. MDL denoising //IEEE Transactions on Information Theory. – 2000. – Т. 46. – №. 7. – С. 2537-2543.

[2] Верещагин Н., Щепин Е. Информация, кодирование и предсказание. – М.: МЦНМО, 2012

Рис. 2: График функции $P(x)$ при различных значениях p

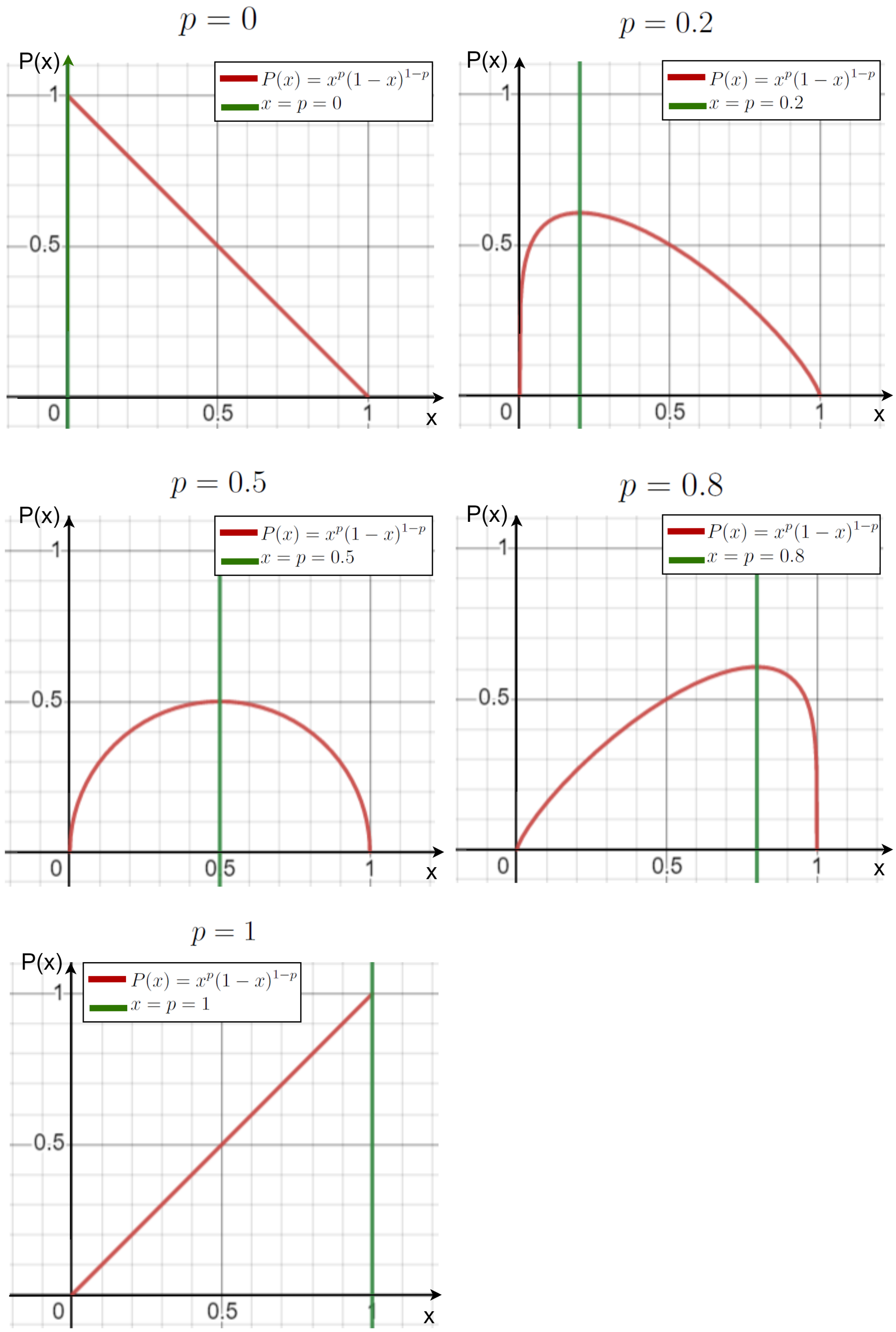


Рис. 3: Оптимальная Марковская цепь с 4 состояниями для последовательности $0001 \cdots 0001$