ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
“САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ”
(СПбГУ)

Образовательная программа бакалавриата “Математика”

Отчет о практике

на тему

Марковские цепи на выпуклых компактах

Выполнил студент 3 курса бакалавриата
группа 21.Б01-мкн
Кукушкин Максим Алексеевич

Научный руководитель:
Профессор, Давыдов Юрий Александрович

Санкт-Петербург
2024

Содержание

1. Введение

В данной работе рассматривается замечательный пример *особой цепи Маркова* на выпуклом компакте в Евклидовом пространстве. Задача привлекательна своей многогранностью и прозрачностью.

Будучи поставленной в строго вероятностной формулировке, она имеет интуитивно понятную геометрическую интерпретацию. Более того, самые разные подходы (функциональный, вероятностный, чисто аналитический и даже комбинаторный) позволяют добиться нетривиальных результатов и все вместе дают исчерпывающую картину происходящего.

Приятный бонус состоит в богатом множестве примеров, допускающих компьютерное моделирование и дающих красивые иллюстрации.

2. Постановка задачи

Рассмотрим в качестве образца правильный треугольник T на плоскости, центр которого совпадает с началом координат, и пускай e_1, e_2, e_3 – координаты его вершин. Запустим из точки $X_0 = 0$ следующий процесс.

(1) Соединим точку X_0 с вершинами e_1, e_2, e_3 с помощью отрезков $[X_0, e_j]$, после чего перейдём равновозможно с вероятностью $\frac{1}{3}$ из точки X_0 в середину одного из проведённых отрезков. Обозначим точку, в которую мы попали, за X_1;

(2) Повторим процедуру для точки X_1: соединим её с вершинами e_1, e_2, e_3 и перейдём равновероятно в середину одного из отрезков $[X_1, e_j]$, обозначив полученную точку за X_2;

(3) Продолжим процесс до бесконечности и определим тем самым последовательность случайных векторов $\{X_n\}_{n=1}^{\infty}$.

Каждый новый шаг n индуцирует некоторое распределение μ_n на треугольнике T. Вопрос состоит в следующем: существует ли у μ_n предельное распределение? Положительный ответ на этот вопрос был получен в работе [2]. Примечательно, что предельное распределение оказалось сосредоточено на треугольнике Серпинского.

Естественно попытаться обобщить постановку задачи: взять в качестве носителя цепи некоторый многогранник или даже произвольный выпуклый компакт в $\mathbb{R}^d$. Хочется надеяться, что предельное распределение существует и в этом случае.

Кроме того, мгновенно возникают другие вопросы:

- Можно ли доказать существование и единственность *инвариантного распределения* для Марковской цепи данного вида? Заметим, что в общем случае вопрос этот крайне сложный; мы же получим результат практически бесплатно.
- Что можно сказать о скорости сходимости?
- Инвариантно ли поведение цепи относительно аффинных преобразований $\mathbb{R}^d$?
- Какими симметриями обладает предельное распределение?
- Как можно расширить классы носителей цепи, сохранив при этом стройность теории?
- Есть ли явные формулы для моментов предельного распределения?
- Изменится ли что-то, если делить проводимые отрезки не пополам, а в произвольном фиксированном отношении?

Ответы на эти и другие вопросы будут раскрыты ниже, а пока скажем пару слов о структуре работы. Для удобства читателя заметка разделена на три основных части:

- Постановка задачи в достаточной степени общности (а также разработка ключевых подходов к её решению и формулировка центральных утверждений);
- Последовательное доказательство всех теоретических результатов;
- Исследование конкретных цепей и носителей (красивые иллюстрации и аналитические выкладки для примеров, а также некоторые небольшие обобщения).

Перед тем как перейти к содержательной части, договоримся о несущих определениях и обозначениях, которых мы постараемся придерживаться в течение всего текста.

2.1. **Основные определения и обозначения.**

Определение 1. Пусть $(K,\Sigma),(L,\Delta)$ – измеримые пространства, *стохастическим (марковским) ядром* будем называть отображение $P\colon K\times\Delta\to[0,1]$ такое, что:

(1) Для каждого фиксированного $D\in\Delta$: $x\mapsto P(x,D)$ – Σ-измеримо;
(2) Для каждого фиксированного $x\in K$: $D\mapsto P(x,D)$ – вероятностная мера на (L,Δ).

В дальнейшем будем писать $P(x,D)=P_x(D)$ и $P_x=P_x(\cdot)$ при желании рассмотреть $P(x,\cdot)$ в ипостаси меры на (L,Δ).

Определение 2. Пусть (K,Σ) – измеримое пространство, вероятность $\nu\colon\Sigma\to[0,1]$ считается *инвариантной* (относительно P), если:

$$\int\limits_K P(x,A)\,\nu(dx)=\nu(A)$$

для каждого измеримого множества $A\in\Sigma$.

Определение 3. Говорят, что семейство вероятностных мер $\{\mu_n\}$ на пространстве K *слабо сходится* к мере ν ($\mu_n\Rightarrow\nu$), если:

$$\lim_{n\to\infty}\int\limits_K f(x)\,\mu_n(dx)=\int\limits_K f(x)\,\nu(dx)$$

для любой непрерывной ограниченной функции $f\colon K\to\mathbb{R}$.

Определение 4. Пусть Π – семейство вероятностных мер на (K,Σ). Будем говорить, что Π – *относительно компактно*, если:

$$\forall\{P_n\}\subset\Pi\colon\exists\{P_{n_k}\}\text{ и }Q\text{ – вероятностная мера на }(K,\Sigma)\text{ такая, что }P_{n_k}\Rightarrow Q.$$

Определение 5. Семейство Π вероятностных мер на метрическом пространстве (S,d) называют *плотным*, если:

$$\forall\varepsilon>0\colon\exists K\text{ – компакт такой, что }\forall Q\in\Pi\colon Q(K)>1-\varepsilon.$$

Определение 6. Пусть (K,Σ) – измеримое пространство, *расстоянием по полной вариации* между двумя вероятностными мерами μ,ν на пространстве K называется:

$$\|\mu-\nu\|_{\mathrm{TV}}\overset{\mathrm{def}}{=}\sup_f\Big\{\int\limits_K f\,d\mu-\int\limits_K f\,d\nu\Big\},$$

где f пробегает множество всех измеримых функций из K в $[-1,1]$.

Замечание. Стандартную евклидову норму вектора $x \in \mathbb{R}^d$ будем обозначать через $|x|$, а единичную матрицу размерности n – через I_n.

Будем также писать $X \stackrel{\mathcal{D}}{=} Y$, если случайные вектора X, Y *одинаково распределены.*

2.2. **Функциональный подход.** В дальнейшем нам понадобится одно особое свойство стохастического ядра, которое естественно возникает следующим образом.

Рассмотрим произвольный метрический компакт K, борелевскую σ-алгебру $\mathfrak{B}(K)$ и пространство $\mathcal{P}(K)$ вероятностных мер на нём с топологией слабой сходимости, а также некоторое марковское ядро

$$P(x, A)\colon K \times \mathfrak{B}(K) \to [0, 1].$$

Пусть $\nu_n \in \mathcal{P}(K)$ – распределения на K такие, что $\nu_n \Rightarrow \nu$, т.е.:

$$\forall f \in C(K)\colon \int\limits_K f(x)\, \nu_n(dx) \xrightarrow[n]{} \int\limits_K f(x)\, \nu(dx).$$

Определение 7. Назовём *оператором перехода* $U\colon \mathcal{P}(K) \to \mathcal{P}(K)$ оператор свёртки распределения с марковским ядром:

$$(U\lambda)(B) \stackrel{\text{def}}{=} \int\limits_K P(x, B)\, \lambda(dx).$$

Замечание. Ясно, что мера $\lambda \in \mathcal{P}(K)$ – инвариантна (относительно ядра P) $\Leftrightarrow U\lambda = \lambda$.

Каковы достаточные условия непрерывности оператора U? Нам нужно проверить, что $U\nu_n \Rightarrow U\nu$, т.е. что $\forall f \in C(K)$:

$$\int\limits_K f(\tilde{x})\, U\nu_n(d\tilde{x}) = \int\limits_K \int\limits_K f(\tilde{x})\, P(x, d\tilde{x})\, \nu_n(dx) \xrightarrow[n]{?} \int\limits_K \int\limits_K f(\tilde{x})\, P(x, d\tilde{x})\, \nu(dx) = \int\limits_K f(\tilde{x})\, U\nu(d\tilde{x}).$$

Положим:

$$g(x) \stackrel{\text{def}}{=} \int\limits_K f(\tilde{x})\, P(x, d\tilde{x}).$$

Тогда требуемая сходимость

$$\int\limits_K g(x)\, \nu_n(dx) \xrightarrow[n]{?} \int\limits_K g(x)\, \nu(dx)$$

обеспечивается за счёт слабой сходимости $\nu_n \Rightarrow \nu$, как только $g \in C(K)$.

Определение 8. Будем говорить, что стохастическое ядро P обладает свойством *феллеровости*, если для любой непрерывной функции $f \in C(K)$ отображение $g\colon K \to \mathbb{R}$,

$$g(x) \stackrel{\text{def}}{=} \int\limits_K f(\tilde{x})\, P(x, d\tilde{x})$$

также является непрерывным.

Замечание. Таким образом, оператор перехода, ассоциированный с феллеровым ядром, непрерывен в пространстве $\mathcal{P}(K)$ с топологией слабой сходимости.

Исключительность данного свойства стохастического ядра проявляется в следующем утверждении.

Теорема 1. *Произвольная феллерова марковская цепь в компактном метрическом пространстве обладает инвариантным распределением.*

Доказательство. Рассмотрим метрический компакт (K, d) с феллеровым ядром $P(x, A)$. Пусть μ_0 – распределение цепи в начальный момент времени. Тогда распределение после первого шага имеет вид:

$$\mu_1(A) \stackrel{\text{def}}{=} \int\limits_K P(x, A)\, \mu_0(dx).$$

Пусть μ_n – распределение в момент времени n. Ясно, что $\nu_n \stackrel{\text{def}}{=} \frac{\mu_1 + \ldots + \mu_n}{n}$ также является вероятностной мерой, сосредоточенной на K.

Очевидно, что $\{\nu_n\}$ – плотное семейство мер (ведь K – компакт). Значит, по теореме Прохорова, $\{\nu_n\}$ – относительно компактное семейство в $\mathcal{P}(K)$.

В частности, существует подпоследовательность $\{\nu_{n_k}\}$ такая, что $\nu_{n_k} \Rightarrow \nu$. Рассмотрим

$$\begin{aligned}\nu_{n_k+1} = \frac{1}{n_k+1}\big(\mu_1 + \ldots + \mu_{n_k} + \mu_{n_k+1}\big) &= \frac{n_k}{n_k+1}\Big(\frac{\mu_1 + \ldots + \mu_{n_k}}{n_k}\Big) + \frac{1}{n_k+1}\mu_{n_k+1} \\ &= \frac{n_k}{n_k+1}\nu_{n_k} + \frac{1}{n_k+1}\mu_{n_k+1} \Rightarrow \nu.\end{aligned}$$

Таким образом, $\nu_{n_k} \Rightarrow \nu$ и $\nu_{n_k+1} \Rightarrow \nu$.

Пусть U – оператор перехода, ассоциированный с ядром P. Тогда:

$$U\mu_0 = \mu_1,\, U\mu_1 = \mu_2, \ldots,\, U\mu_{n_k} = \mu_{n_k+1}.$$

Наконец, воспользуемся феллеровостью ядра P:

$$\begin{aligned}\nu_{n_k+1} = \frac{1}{n_k+1}\mu_1 + \frac{n_k}{n_k+1}\Big(\frac{\mu_2 + \ldots + \mu_{n_k+1}}{n_k}\Big) &= \frac{1}{n_k+1}\mu_1 + \frac{n_k}{n_k+1}U\Big(\frac{\mu_1 + \ldots + \mu_{n_k}}{n_k}\Big) \\ &= \frac{1}{n_k+1}\mu_1 + \frac{n_k}{n_k+1}U\nu_{n_k} \Rightarrow U\nu.\end{aligned}$$

Таким образом, $\nu_{n_k+1} \Rightarrow \nu$ и $\nu_{n_k+1} \Rightarrow U\nu$, а значит $\nu = U\nu$ по хаусдорфовости $\mathcal{P}(K)$, т.е. ν – искомое инвариантное распределение. □

Замечание. Добиться того же результата можно было, конечно, и меньшей кровью, если бы мы встали на плечи гигантов.

Теорема Шаудера о неподвижной точке гласит, что в произвольном хаусдорфовом топологическом векторном пространстве V непрерывное отображение $f\colon X \to X$ непустого выпуклого компактного подмножества $X \subset V$ в себя имеет неподвижную точку.

Остаётся только положить V равным пространству регулярных борелевских мер на K с топологией слабой сходимости, в качестве X взять $\mathcal{P}(K)$, а в качестве отображения f – оператор U, который, как мы показали, действительно будет непрерывен в случае феллерости соответствующего ядра.

Неподвижная точка U и будет искомым инвариантным распределением.

Перейдём к формулировке интересующей нас конкретной задачи. Мы будем отталкиваться от явных примеров и обобщать их шаг за шагом.

Пример 1. Возьмём правильный шестиугольник $H \subset \mathbb{R}^2$ с центром в начале координат, обозначим его вершины за $e_1, \ldots, e_6$ и определим марковское ядро P по правилу:

$$P(x, A) \stackrel{\text{def}}{=} \frac{1}{6} \sum_{j=1}^{6} \mathbb{1}_A\Big(\frac{x + e_j}{2}\Big),$$

где $x \in K$ и $A \in \mathfrak{B}(K)$.

Иными словами, начиная из точки $x \in H$, мы равновозможно с вероятностью $\frac{1}{6}$ переходим в середину одного из отрезков $[x, e_j]$.

Пример 2 (Обобщение 1)**.** Возьмём произвольный выпуклый многогранник $K \subset \mathbb{R}^d$ с N вершинами $e_1, \ldots, e_N$. Марковское ядро на нём определим следующим образом:

$$P(x, A) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{j=1}^{N} \mathbb{1}_A\Big(\frac{x + e_j}{2}\Big), \text{ т.е.:}$$

$$P_x = \frac{1}{N} \sum_{j=1}^{N} \delta_{\{\frac{x+e_j}{2}\}},$$

$$P(\cdot, A) = \frac{1}{N} \sum_{j=1}^{N} \mathbb{1}_A\Big(\frac{\cdot + e_j}{2}\Big),$$

где $x \in K$ и $A \in \mathfrak{B}(K)$.

Корректность определения очевидна в силу выпуклости K.

Пример 3 (Обобщение 2)**.** Возьмём произвольный выпуклый компакт $K \subset \mathbb{R}^d$ и зададим на его границе ∂K вероятностную меру μ. В предыдущих случаях μ представляла из себя взвешенную сумму нагрузок в вершинах многогранника:

$$\mu = \frac{1}{N} \sum_{j=1}^{N} \delta_{\{e_j\}}.$$

Теперь же μ – это некоторое распределение на границе компакта K.

Заметим, наконец, что у нас нет препятствий для последнего обощения: ничто не мешает взять меру μ, сосредоточенную не только на границе, но и внутри компакта!

Определение 9. Рассмотрим выпуклый компакт $K \subset \mathbb{R}^d$ и меру $\mu \in \mathcal{P}(K)$, зададим стохастическое ядро $P(x, A)$ по правилу:

$$P(x, A) \stackrel{\text{def}}{=} \int\limits_K \mathbb{1}_A\Big(\frac{x + z}{2}\Big)\, \mu(dz),$$

где $x \in K, A \in \mathfrak{B}(K)$.

Меру μ будем называть *переходным распределением* или *переходной мерой*, а тело K – *носителем* марковской цепи.

Замечание. Эквивалентно, можно положить:

$$P(x,A) \stackrel{\text{def}}{=} \mu\Big(\{z \in K : \frac{x+z}{2} \in A\}\Big).$$

Смысл у обоих определений ядра одинаков: $P(x,A)$ выдаёт вероятность за один шаг процесса попасть из точки x в множество A.

Для подсчёта этой вероятности необходимо из точки x провести отрезок к каждой точке $z \in K$ и перейти в середину одного из отрезков в соответствии с распределением μ.

Для окончания формулировки задачи в обобщённом виде, т.е. для корректного задания марковской цепи распределений μ_n остаётся определить *начальное распределение* μ_0. Мы будем считать, что $\mu_0 \in \mathcal{P}(K)$ – некоторая вероятностная мера, сосредоточенная на K.

Определение 10. Пусть $\mu_0, \mu \in \mathcal{P}(K)$, а U – оператор перехода ядра $P(x,A)$, заданного по формуле:

$$P(x,A) \stackrel{\text{def}}{=} \int\limits_K \mathbb{1}_A\Big(\frac{x+z}{2}\Big)\,\mu(dz).$$

Последовательность вероятностей $\{\mu_n\}_{n=0}^{\infty}$ вида $\mu_{n+1} = U\mu_n$ будем называть *особой марковской цепью* задачи в обобщённом виде.

Замечание. Вообще говоря, при различном выборе μ_0 будут получаться различные последовательности μ_n. Однако впоследствии мы покажем, что *поведение* цепи на бесконечности от начального распределения в действительности совершенно не зависит.

Поймём некоторые очевидные свойства особой марковской цепи.

Лемма 1. *Пусть $K \subset \mathbb{R}^d$ – выпуклый многогранник, $e_1, ..., e_N$ – все его вершины. Тогда стохастическое ядро*

$$P(x,A) = \frac{1}{N}\sum_{j=1}^{N} \mathbb{1}_A(\frac{x+e_j}{2})$$

обладает свойством феллеровости.

Доказательство. Так как

$$P_x = P(x,\cdot) = \frac{1}{N}\sum_{j=1}^{N} \delta_{\{\frac{x+e_j}{2}\}},$$

то для любой функции $f \in C(K)$:

$$g(x) \stackrel{\text{def}}{=} \int\limits_K f(\tilde{x})\,P(x,d\tilde{x}) = \frac{1}{N}\sum_{j=1}^{N} f\Big(\frac{x+e_j}{2}\Big).$$

Очевидно, что функция $g(x)$ такого вида – непрерывна. $\square$

Лемма 2. *Марковское ядро задачи в обобщённом виде также удовлетворяет условию феллеровости.*

Доказательство. Пусть $f \in C(K)$. Рассмотрим функцию $g(x)$ из определения феллеровости ядра:

$$g(x) \stackrel{\text{def}}{=} \int\limits_K f(\tilde{x})\, P(x, d\tilde{x}) = \int\limits_K f\Big(\frac{x+z}{2}\Big)\, \mu(dz).$$

Положим:

$$\psi_{x,h}(z) \stackrel{\text{def}}{=} f\Big(\frac{x+z+h}{2}\Big) - f\Big(\frac{x+z}{2}\Big).$$

Пользуясь ограниченностью непрерывной функции f на компакте и тем, что μ – вероятностная мера, найдём для подынтегрального выражения суммируемую мажоранту:

$$|g(x+h) - g(x)| = \Big|\int\limits_K \Big(f\Big(\frac{x+z+h}{2}\Big) - f\Big(\frac{x+z}{2}\Big)\Big)\, \mu(dz)\Big|$$
$$\stackrel{\text{def}}{=} \Big|\int\limits_K \psi_{h,x}(z)\, \mu(dz)\Big| \leqslant \int\limits_K |\psi_{h,x}(z)|\, \mu(dz) \leqslant C \cdot \mu(K) = C.$$

По равномерной непрерывности f:

$$\psi_{x,h}(z) \xrightarrow[h \to 0]{} 0,$$

Пользуясь теоремой Лебега, перейдём к пределу под знаком интеграла, и заключим, что:

$$|g(x+h) - g(x)| \xrightarrow[h \to 0]{} 0 \ \forall x \in K,$$

т.е. $g(x) \in C(K)$, что и означает феллеровость ядра $P(x, A)$. □

Следствие 1. *Особая марковская цепь обобщённой задачи обладает инвариантным распределением.*

Доказательство. Напрямую следует из теоремы о феллеровой цепи и только что доказанной леммы. □

Вместо того, чтобы и дальше следовать намеченному маршруту в функциональном духе, остановимся на минуту и попробуем на примере с многогранником лучше понять структуру получающейся цепи.

В качестве базового случая удобно положить $\mu_0 = \delta_{\{x_0\}}$ для некоторой фиксированный точки $x_0 \in K$. Тогда:

$$\mu_1(A) = (U\mu_0)(A) = \int\limits_K P_x(A)\, \mu_0(dx) = P_{x_0}(A), \text{ т.е.: } \mu_1 = P_{x_0} = \frac{1}{N} \sum_{j=1}^{N} \delta_{\{\frac{x_0+e_j}{2}\}}.$$

Аналогично,

$$\mu_2 = (U\mu_1)(\cdot) = \int\limits_K P_x(\cdot)\, \mu_1(dx) = \frac{1}{N^2} \sum_{i_1,i_2=1}^{N} \delta_{\{\frac{\frac{x_0+e_{i_1}}{2}+e_{i_2}}{2}\}} = \frac{1}{N^2} \sum_{i_1,i_2=1}^{N} \delta_{\{\frac{x_0}{4}+\frac{e_{i_1}}{4}+\frac{e_{i_2}}{2}\}}.$$

Наконец,

$$\mu_n = \frac{1}{N^n} \sum_{i_1,\dots,i_n=1}^{N} \delta_{\{\frac{x_0}{2^n} + \frac{e_{i_1}}{2^n} + \frac{e_{i_2}}{2^{n-1}} + \dots + \frac{e_{i_n}}{2}\}}.$$

Дискретный вид последней формулы наводит на мысль переформулировать задачу в терминах случайных векторов и перейти к анализу экспоненциально быстро сходящихся рядов. Займёмся претворением этой идеи в жизнь.

2.3. **Марковский подход.**

Определение 11. Рассмотрим произвольный выпуклый компакт $K \subset \mathbb{R}^d$, а также начальное распределение μ_0 и переходную меру μ на нём.

Пусть $\xi, \xi_k \overset{i.i.d.}{\sim} \mu$. Зададим *особую марковскую цепь* $\{X_n\}_{n=0}^{\infty}$ по правилу:

$$X_0 \sim \mu_0,$$

$$X_{n+1} = \frac{X_n + \xi_{n+1}}{2}.$$

Будем называть $\xi_k \in \mathbb{R}^d$ *переходными случайными векторами*, а $X_n \in \mathbb{R}^d$ – *векторами распределения цепи на n-ом шаге.*

Последовательность $X_0, X_1, \dots, X_n, \dots$ очевидным образом подчинена вероятностям перехода $P_x^n(A)$ (из точки x в множество A за n шагов), определяемым стохастическим ядром P. Иными словами,

$$\mu_0 = \mathbb{P}X_0^{-1}, \mu_1 = \mathbb{P}X_1{}^{-1}, \dots, \mu_n = \mathbb{P}X_n^{-1}.$$

Распишем теперь подробнее случайные вектора $X_0, X_1, \dots, X_n, \dots$:

$$\begin{aligned}
&X_0 \sim \mu_0,\\
&X_1 = \frac{X_0 + \xi_1}{2},\\
&X_2 = \frac{X_1 + \xi_2}{2} = \frac{X_0}{4} + \frac{\xi_1}{4} + \frac{\xi_2}{2},\\
&X_n = \frac{X_0}{2^n} + \frac{\xi_1}{2^n} + \frac{\xi_2}{2^{n-1}} + \dots + \frac{\xi_n}{2}.
\end{aligned}$$

Такое представление векторов X_n в виде взвешенных сумм ξ_k позволяет легко и непринуждённо найти предельное распределение.

Теорема 2. *Особая марковская цепь обладает предельным распределением. Более того, предельное распределение не зависит от начального распределения на носителе и однозначно определяется переходной мерой.*

Доказательство. Рассмотрим состояние цепи в момент времени n:

$$X_n = \frac{X_0}{2^n} + \frac{\xi_1}{2^n} + \frac{\xi_2}{2^{n-1}} + \dots + \frac{\xi_n}{2} \overset{\mathcal{D}}{=} \frac{X_0}{2^n} + \frac{\xi_1}{2} + \frac{\xi_2}{2^2} + \dots + \frac{\xi_n}{2^n}.$$

Положим:

$$S_n = \frac{\xi_1}{2} + \frac{\xi_2}{2^2} + \cdots + \frac{\xi_n}{2^n},$$
$$Y = \sum_{k=1}^{\infty} \frac{\xi_k}{2^k}.$$

В силу компактности $K \subset \mathbb{R}^d$, множество K – ограничено. Обозначим радиус компакта:

$$M \stackrel{\text{def}}{=} \max_{x \in K} |x| < \infty.$$

Но ведь переходные вектора $\xi_k \in K$ почти наверное. Значит, $|\xi_k| \leqslant M$ и

$$\sum_{k=1}^{\infty} \frac{|\xi_k|}{2^k} \leqslant \sum_{k=1}^{\infty} \frac{M}{2^k} \leqslant M \cdot \sum_{k=1}^{\infty} \frac{1}{2^k} = M \cdot 1 = M < \infty.$$

По полноте объемлющего пространства из абсолютной сходимости немедленно следует

$$\frac{X_0}{2^n} + S_n \xrightarrow[n \to \infty]{a.e.} Y = \lim_{n \to \infty} S_n.$$

Таким образом,

$$X_n \Rightarrow Y.$$

Считая, что $Y \sim \nu \in \mathcal{P}(K)$, получим слабую сходимость к предельному распределению $\mu_n \Rightarrow \nu$. Из определения Y видно, что его распределение ν не зависит от μ_0 и однозначно определяется распределением случайных векторов $\xi_k \sim \mu$, что и требовалось. $\square$

Теорема 3. *Особая марковская цепь обладает единственным инвариантным распределением.*

Доказательство. Предельное распределение ν очевидным образом доставляет пример инвариантного распределения. Единственность легко следует из следующего рассуждения.

Предположим λ – некоторое инвариантное распределение. Положим $\mu_0 = \lambda$. В силу инвариантности: $\mu_1 = U\mu_0 = U\lambda = \lambda, \mu_2 = U\mu_1 = U\lambda = \lambda, \ldots, \mu_n = U\mu_{n-1} = \lambda$ для любого n. Тогда $\mu_n \Rightarrow \lambda$, но $\mu_n \Rightarrow \nu$. По хаусдорфовости $\mathcal{P}(K)$ с топологией слабой сходимости заключаем, что $\lambda = \nu$ – единственное инвариантное распределение. $\square$

Для комфорта читателя приведём список наиболее интересных свойств найденного предельного и инвариантного распределения μ. Доказательства нижеуказанных предложений будут изложены ниже.

2.4. **Формулировки основных утверждений.**

Теорема 4. *Скорость сходимости к предельному распределению ν – экспоненциальная. Она однозначно определяется радиусом компакта и средним переходной меры.*

Теорема 5. *Особая марковская цепь инвариантна относительно аффинных преобразований своего носителя.*

Теорема 6. *Если переходное распределение μ обладает симметрией относительно некоторого невырожденного аффинного преобразования F, то и предельное распределение ν обладает той же симметрией.*

Теорема 7. *Если переходное распределение μ абсолютно непрерывно, то предельное распределение ν также абсолютно непрерывно. Более того, в этом случае распределения цепи на n-ом шаге сходятся к предельному распределению по вариации.*

Теорема 8. *Все доказанные утверждения верны и для особой марковской цепи с произвольными выпуклыми весами перехода $0 \leqslant \alpha, \beta \leqslant 1$.*

3. Теоретические результаты

3.1. **Оценка скорости сходимости.** Пусть μ – переходная мера на компакте K, $X_0 \sim \mu_0$, $\tilde{X}_0 \sim \tilde{\mu}_0$ – некоторые начальные распределения, $X_n \sim \mu_n, \tilde{X}_n \sim \tilde{\mu}_n$ – распределения через n шагов. Обозначим:

$$X_n = \frac{X_0}{2^n} + \frac{\xi_1}{2^n} + \dots + \frac{\xi_n}{2} \stackrel{\text{def}}{=} \frac{X_0}{2^n} + W_n \Rightarrow Y,$$

$$\tilde{X}_n = \frac{\tilde{X}_0}{2^n} + \frac{\xi_1}{2^n} + \dots + \frac{\xi_n}{2^n} \stackrel{\text{def}}{=} \frac{\tilde{X}_0}{2^n} + W_n \Rightarrow Y,$$

$$Y = \sum_{k=1}^{\infty} \frac{\xi_k}{2^k}.$$

Иными словами, для любой непрерывной функции $f \in C(K)$:

$$\lim_{n\to\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(Y)] = \lim_{n\to\infty} \mathbb{E}[f(\tilde{X}_n)].$$

Чтобы формально говорить о расстоянии между мерами μ_n и $\tilde{\mu}_n$ и о скорости сходимости к предельному распределению ν, заведём на $\mathcal{P}(K)$ хорошо известную метрику (Канторовича-Рубинштейна), согласованную со слабой сходимостью.

Для любых вероятностных мер $Q, R \in \mathcal{P}(K)$ положим:

$$d(Q,R) \stackrel{\text{def}}{=} \sup_{f\in L} \Big| \int\limits_K f(x)\,Q(dx) - \int\limits_K f(x)\,R(dx) \Big|, \tag{1}$$

где L есть множество всех равномерно ограниченных Липшицевых функций на K:

$$L = \mathrm{Lip}_1(K) = \Big\{ f\colon K \to \mathbb{R} : |f(t) - f(s)| \leqslant |t-s| \ \forall t,s \in K,\ |f| \leqslant 1 \Big\}.$$

Так как $X_n \sim \mu_n$, т.е. $\mathbb{P}X^{-1} = \mu_n$, и $X_n \in K$ п.н.:

$$\mathbb{E}[f(X_n)] = \int\limits_\Omega f(X_n)\,d\mathbb{P} = \int\limits_{\mathbb{R}^d} f(x)\,\mathbb{P}X^{-1}(dx) = \int\limits_K f(x)\,\mu_n(dx).$$

Оценим расстояние между μ_n и $\tilde{\mu}_n$:

$$d(\mu_n, \tilde{\mu}_n) = \sup_{f \in L} \Big| \int\limits_K f(x)\, \mu_n(dx) - \int\limits_K f(x)\, \tilde{\mu}_n(dx) \Big| = \sup_{f \in L} \Big| \mathbb{E}[f(X_n)] - \mathbb{E}[f(\tilde{X}_n)] \Big| =$$
$$\sup_{f \in L} \Big| \mathbb{E}[f(X_n) - f(\tilde{X}_n)] \Big| \leqslant \sup_{f \in L} \mathbb{E} \Big| f(X_n) - f(\tilde{X}_n) \Big| \quad (2)$$

Воспользуемся следующим нехитрым соображением. Мы знаем, что $X_n, \tilde{X}_n \in K$ почти наверное. Положим $X_n \stackrel{\text{def}}{=} t$ и $\tilde{X}_n \stackrel{\text{def}}{=} s$, тогда $t, s \in K$. В силу определения множества L для функции $f \in L$ имеем:

$$|f(t) - f(s)| \leqslant |t - s|, \text{ т.е. } |f(X_n) - f(\tilde{X}_n)| \leqslant |X_n - \tilde{X}_n|.$$

Тогда мы можем продолжить оценку (2), избавившись от функции f (и следовательно, от взятия супремума по $f \in L$) и написать, что:

$$d(\mu_n, \tilde{\mu}_n) \leqslant \sup_{f \in L} \mathbb{E} \Big| f(X_n) - f(\tilde{X}_n) \Big| \leqslant \mathbb{E} \Big| X_n - \tilde{X}_n \Big| = \mathbb{E} \Big| \Big(\frac{X_0}{2^n} + W_n \Big) - \Big(\frac{\tilde{X}_0}{2^n} + W_n \Big) \Big| =$$
$$\mathbb{E} \Big| \frac{X_0 - \tilde{X}_0}{2^n} \Big| \leqslant \frac{1}{2^n} \Big(\mathbb{E}|X_0| + \mathbb{E}|\tilde{X}_0| \Big) \leqslant \frac{1}{2^n} \cdot 2 \max_{x \in K} |x| = \frac{M}{2^{n-1}}.$$

Аналогичным образом оценивается расстояние между μ_n, ν, где ν – предельное распределение:

$$X_n \stackrel{\mathcal{D}}{=} \frac{X_0}{2^n} + S_n \text{ и } Y = S_n + \sum_{k=n+1}^{\infty} \frac{\xi_k}{2^k}.$$

Положим $m \stackrel{\text{def}}{=} \mathbb{E}|\xi|$. Тогда:

$$d(\mu_n, \nu) \leqslant \dots \leqslant \mathbb{E}|X_n - Y| = \mathbb{E} \Big| \Big(\frac{X_0}{2^n} + S_n \Big) - \Big(S_n + \sum_{k=n+1}^{\infty} \frac{\xi_k}{2^k} \Big) \Big| \leqslant \frac{1}{2^n} \mathbb{E}|X_0| + \sum_{k=n+1}^{\infty} \mathbb{E} \Big| \frac{\xi_k}{2^k} \Big| \leqslant$$
$$\frac{M}{2^n} + \mathbb{E}|\xi| \cdot \sum_{k=n+1}^{\infty} \frac{1}{2^k} \stackrel{\text{def}}{=} \frac{M}{2^n} + m \cdot \frac{1}{2^n} = \frac{M + m}{2^n}.$$

Таким образом, мы доказали следующее утверждение.

Предложение 1. *Скорость сходимости распределений на n-ом шаге особой марковской цепи к предельному распределению – экспоненциальная, причём:*

$$d(\mu_n, \nu) \leqslant \frac{M + m}{2^n},$$

где M – радиус компакта K, а $m = \mathbb{E}|\xi|$.

3.2. **Инвариантность относительно аффинных преобразований.** Интуитивно кажется понятным, что аффинные преобразования объемлющего пространства никак не должны сказываться на предельном поведении нашей марковской цепи. Попробуем придать этой интуиции формальное выражение на бумаге.

Рассмотрим произвольное аффинное преобразование $F\colon \mathbb{R}^d \to \mathbb{R}^d$. Не умаляя общности, можно считать, что F – линейное (при подходящем выборе начала координат). Для нас ключевое свойство аффинного отображения F состоит в том, что оно заведомо переводит выпуклые компакты в выпуклые компакты. Тем самым, $F(K)$ можно также рассмотреть в качестве носителя особых марковских цепей.

Пусть $\mu_0 \in \mathcal{P}(K)$ – начальное распределение на компакте K. Под действием F оно перейдёт в некоторое распределение $F\mu_0 \in \mathcal{P}(F(K))$ на компакте $F(K)$, определённое следующим образом:

$$(F\mu_0)(B) \stackrel{\text{def}}{=} \mu_0(F^{-1}(B)).$$

Другими словами, $F\mu_0$ есть образ меры μ_0 под действием линейного отображения F.

Пусть, кроме того, на носителе K задано переходное распределение $\mu \in \mathcal{P}(K)$. Под действием F оно перейдёт в соответствующее распределение $F\mu \in \mathcal{P}(F(K))$ на компакте $F(K)$ аналогичным образом.

Мы вправе рассмотреть марковскую цепь на компакте K, которая есть не что иное как последовательность распределений на n-ом шаге $\{\mu_n\}$, однозначно определяемая μ_0 и μ. В то же время на $F(K)$ задаются одновременно две марковские цепи:

- С одной стороны, начальное распределение $F\mu_0 \stackrel{\text{def}}{=} \mu_0^F$ и переходное распределение $F\mu \stackrel{\text{def}}{=} \mu^F$ на $F(K)$ автоматически порождают цепь $\{\mu_n^F\}$;
- С другой стороны, цепь $\{\mu_n\}$ на исходном носителе вместе с отображением F индуцируют последовательность распределений $\{F\mu_n\}$ на $F(K)$ (как образы соответствующих мер).

Геометрический смысл происходящего очень прост: мы либо начинаем свой путь в компакте K, делаем там n шагов согласно распределению μ, после чего переходим с помощью F в новый носитель $F(K)$, либо сначала отображаемся в $F(K)$ посредством F и затем делаем n шагов в новом носителе согласно распределению $F\mu$.

Не трудно видеть, что эти два способа эквивалентны друг другу. Действительно, для удобства вновь перейдём на язык случайных величин:

- $X_0 \sim \mu_0$, $\xi_k \stackrel{i.i.d.}{\sim} \mu$, $X_n \sim \mu_n$;
- $X_0^F \sim \mu_0^F$, $\xi_k^F \stackrel{i.i.d.}{\sim} \mu^F$, $X_n^F \sim \mu_n^F$.

Поскольку $\mu_0^F = F\mu_0$ и $\mu^F = F\mu$ по определению, соответствующие случайные вектора равны по распределению: $X_0^F \stackrel{\mathcal{D}}{=} F(X_0)$ и $\xi_k^F \stackrel{\mathcal{D}}{=} F(\xi_k)$.

Воспользуемся линейностью преобразования F (точнее, тем, что F переводит выпуклые комбинации точек в выпуклые комбинации):

$$X_1^F = \frac{X_0^F + \xi_1^F}{2} \stackrel{\mathcal{D}}{=} \frac{F(X_0) + F(\xi_1)}{2} = F\Big(\frac{X_0 + \xi_1}{2}\Big) = F(X_1).$$

Аналогично,

$$X_n^F = \frac{X_0^F}{2^n} + \frac{\xi_1^F}{2^n} + \dots + \frac{\xi_n^F}{2} \overset{\mathcal{D}}{=} \frac{F(X_0)}{2^n} + \frac{F(\xi_1)}{2^n} \dots + \frac{F(\xi_n)}{2} = F\Big(\frac{X_0}{2^n} + \frac{\xi_1}{2^n} + \dots + \frac{\xi_n}{2}\Big) = F(X_n).$$

Таким образом, $F\mu_n = \mu_n^F$, т.е. распределения на компакте $F(K)$, получаемые двумя способами индуцирования с K с помощью F, просто-напросто совпадают. Но тогда, в частности, совпадают и все их характеристики, а также предельное поведение при больших номерах шагов.

Мы знаем, что $X_n \Rightarrow Y$, или, что то же, $\mu_n \Rightarrow \nu$. Применив к этой слабой сходимости линейное (тем более, непрерывное) отображение F, получим: $F(X_n) \Rightarrow F(Y)$, или, что то же, $F\mu_n \Rightarrow F\nu$.

Применим теорему о существовании предельного распределения уже к индуцированной цепи $\{\mu_n^F\}$ с носителем $F(K)$ и получим: $X_n^F \Rightarrow Y^F$, или, что то же, $\mu_n^F \Rightarrow \nu^F$, где ν^F – некоторое распределение на $F(K)$.

Наконец, из сходимостей $F\nu \Leftarrow F\mu_n = \mu_n^F \Rightarrow \nu^F$ и хаусдорфовости $\mathcal{P}(F(K))$ заключаем, что и предельное распределение индуцированной цепи есть всего-навсего образ предельного распределения цепи исходной: $\nu^F = F\nu$.

До сих пор нам ничего не требовалось от самого отображения F, кроме его аффинности. Чтобы "обратить" наши рассуждения, необходимо потребовать, чтобы линейная часть F была неособой. В таком случае можно легально рассмотреть F^{-1}, отображающее $F(K)$ на K, и повторить наши выкладки для него.

Таким образом, мы доказали следующее утверждение.

Предложение 2. *Неособое аффинное преобразование осуществляет "изоморфизм" пространства особых марковских цепей с исходным носителем на пространство особых марковских цепей с носителем, равным образу исходного, в следующем смысле:*

- *Распределение на n-ом шаге переходит в соответствующее распределение на шаге n;*
- *Предельное распределение исходной цепи переходит в предельное распределение индуцированной цепи.*

3.3. **Сохранение симметрий при предельном переходе.** Предположим, переходная мера μ обладает некоторыми симметриями (например, равномерное распределение на окружности сохраняется при поворотах диска). Естественно ожидать, что предельное распределение ν унаследует их от μ. Давайте формализуем это соображение.

Будем считать, что $\dim(K) = d$, т.е. что внутренность K непуста (если нужно, переходя в подходящую гиперплоскость). Рассмотрим невырожденное аффинное преобразование F пространства $\mathbb{R}^d$ такое, что $F\colon K \to K$, т.е. F отображает K на K. Предположим, что переходная мера μ инвариантна относительно F, т.е. $F\mu = \mu$, или, что то же:

$$\mu(B) = (F\mu)(B) = \mu(F^{-1}(B))$$

для любого борелевского B.

В силу невырожденности F, можно в качестве B рассматривать множества вида $F^{-1}(A)$ по всем борелевским A, поэтому для нас инвариантность μ относительно F эквивалентна также тому, что для любого борелевского B:

$$\mu(B) = \mu(F(B)).$$

Поскольку предельное поведение цепи не зависит от начального распределения, в качестве μ_0 далее будем полагать δ_{x_0} для произвольной $x_0 \in K$. В этом случае $U\mu_0 = U\delta_{x_0} = \mu$, и значит $\mu_n = U^n\mu_0 = U^{n-1}(U\mu_0) = U^{n-1}\mu$.

Заметим, что распределения μ_n также будут инвариантны относительно F. Это напрямую следует из полученного нами ранее соотношения $F\mu_n = \mu_n^F$, а также инвариантности μ относительно F:

$$\mu_n(F^{-1}(B)) \stackrel{\text{def}}{=} (F\mu_n)(B) = \mu_n^F(B) = (U^{n-1}(\mu^F))(B) = (U^{n-1}(F\mu))(B) = (U^{n-1}(\mu))(B) = \mu_n(B).$$

Предельный переход от μ_n к ν обосновывается следующим соображением.

Лемма 3. *Пусть $\mu_n, \nu \in \mathcal{P}(K)$, $\mu_n \Rightarrow \nu$, а F – произвольное измеримое отображение $F\colon K \to K$ такое, что*

$$\mu_n(F^{-1}(B)) = \mu_n(B)\ \forall B \in \mathfrak{B}(K).$$

Тогда

$$\nu(F^{-1}(B)) = \nu(B)\ \forall B \in \mathfrak{B}(K).$$

Иначе говоря, подпространство $\mathcal{M}(K, F) \subset \mathcal{P}(K)$ мер, инвариантных относительно измеримого отображения F, замкнуто относительно слабой сходимости на $\mathcal{P}(K)$.

Доказательство. В силу инвариантности μ_n относительно F для любой непрерывной функции $f \in C(K)$ имеем равенство интегралов:

$$\int\limits_K f\, d\mu_n = \int\limits_K (f \circ F)\, d\mu_n.$$

По слабой сходимости $\mu_n \Rightarrow \nu$ осуществим предельный переход при $n \to \infty$ в обеих частях равенства, чтобы получить:

$$\int\limits_K f\, d\nu = \int\limits_K (f \circ F)\, d\nu.$$

Поскольку последнее равенство выполнено для любой непрерывной функции $f \in C(K)$, мера ν инвариантна относительно отображения F, что и требовалось. □

Таким образом, мы доказали следующее утверждение.

Предложение 3. *Если переходное распределение особой марковской цепи инвариантно относительно некоторого невырожденного аффинного преобразования F, то и предельное распределение инвариантно относительно F.*

Пример 4. В качестве носителя K возьмём n-мерный диск $D^n \subset \mathbb{R}^n$. Равномерное распределение на D^n и равномерное распределение на его границе S^{n-1} инвариантны относительно ортогональной группы $O(n)$.

Значит, если в качестве переходного распределения μ взять одну из данных мер, то предельное распределение ν также будет инвариантно относительно всей группы ортогональных преобразований $O(n)$ пространства $\mathbb{R}^n$.

3.4. **Моменты предельного распределения.** Поскольку последовательность моментов однозначно определяет распределение на компакте, интересно попробовать вычислить моменты предельного распределения. Оказывается, для не слишком высоких порядков, можно выписать явные формулы в терминах моментов переходных случайных векторов.

Для удобства сдвинем компакт так, что $0 \in K$, и примем $X_0 = 0$ почти наверное. Вычислим математическое ожидание предельного случайного вектора Y:

$$\mathbb{E}Y = \mathbb{E}\Big(\sum_{k=1}^{\infty} \frac{\xi_k}{2^k}\Big) = \mathbb{E}\xi \cdot \sum_{k=1}^{\infty} \frac{1}{2^k} = \mathbb{E}\xi.$$

Итак,

$$\mathbb{E}Y = \mathbb{E}\xi.$$

В частности, если центрировать переходные случайные вектора, то $\mathbb{E}Y = 0$. Далее будем рассматривать только такой случай.

Обозначим за $\operatorname{cov}(Y)$ матрицу ковариации $(\operatorname{cov}(Y_i, Y_j))_{i,j=1}^d = (\mathbb{E}(Y_i Y_j))_{i,j=1}^d = \mathbb{E}(YY^\intercal)$. Воспользуемся независимостью ξ_k и билинейностью ковариации:

$$\operatorname{cov}(Y) = \operatorname{cov}\Big(\sum_{k=1}^{\infty} \frac{\xi_k}{2^k}\Big) = \sum_{k=1}^{\infty} \operatorname{cov}\Big(\frac{\xi_k}{2^k}\Big) = \operatorname{cov}(\xi) \cdot \sum_{k=1}^{\infty} \frac{1}{2^{2k}} = \frac{1}{3}\operatorname{cov}(\xi).$$

Итак,

$$\operatorname{cov}(Y) = \frac{1}{3}\operatorname{cov}(\xi).$$

Замечание. Эта формула, конечно, остаётся верной и без предположения центрированности векторов ξ_k.

Найдём моменты высших степеней. Обозначим $\xi = (\xi^1, \ldots, \xi^d) \in \mathbb{R}^d$. Для удобства введём также вектор $\xi_0 = 0$, в частности $\mathbb{E}\xi_0 = 0$.

$$\mathbb{E}Y^2 = \mathbb{E}\Big(\sum_{k=0}^{\infty} \frac{\xi_k}{2^k}\Big)^2 \stackrel{\text{def}}{=} \mathbb{E}\Big(\sum_{k=0}^{\infty} \frac{a_k}{2^k}\Big) = \sum_{k=0}^{\infty} \frac{1}{2^k}\mathbb{E}a_k,$$

где случайные вектора a_k соответствуют свёртке последовательности ξ_k с самой собой.

Иными словами,

$$\Big(\sum_{k=0}^{\infty} \frac{\xi_k}{2^k}\Big)^2 = \sum_{k=0}^{\infty} \frac{\xi_k}{2^k} \cdot \sum_{k=0}^{\infty} \frac{\xi_k}{2^k} = \sum_{k=0}^{\infty}\Big(\sum_{n=0}^{k} \frac{\xi_n}{2^n} \cdot \frac{\xi_{k-n}}{2^{k-n}}\Big) = \sum_{k=0}^{\infty} \frac{1}{2^k}\Big(\sum_{n=0}^{k} \xi_n \xi_{k-n}\Big) \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} \frac{1}{2^k} \cdot a_k.$$

В свою очередь,

$$\mathbb{E}a_k = \mathbb{E}\Big(\sum_{n=0}^{k} \xi_n \xi_{k-n}\Big) = \sum_{n=0}^{k} \mathbb{E}(\xi_n \xi_{k-n}).$$

В силу независимости и центрированности векторов ξ_k:

$$\mathbb{E}(\xi_n \xi_{k-n}) = \begin{cases} \mathbb{E}\xi^2, & n = k-n, \ \text{т.е. при } k = 2n; \\ 0, & n \neq k-n, \ \text{т.е. при } k \neq 2n. \end{cases}$$

Значит, в случае нечётного k в последней сумме все слагаемые будут нулевыми. В случае же чётного k, т.е. при $k = 2m$ для некоторого $m \in \mathbb{N}$, в сумме будет единственное ненулевое слагаемое, равное $\mathbb{E}\xi^2$, т.е.:

$$\mathbb{E}a_k = \begin{cases} \mathbb{E}\xi^2, & k = 2m \text{ для некоторого } m \in \mathbb{N}; \\ 0, & k = 2m+1 \text{ для некоторого } m \in \mathbb{N} \text{ или при } k = 0. \end{cases}$$

Возвращаясь ко второму моменту предельного вектора Y, имеем:

$$\mathbb{E}Y^2 = \sum_{k=0}^{\infty} \frac{1}{2^k}\mathbb{E}a_k = \sum_{k=0, k=2m, m\in\mathbb{N}}^{\infty} \frac{1}{2^k}\mathbb{E}\xi^2 = \mathbb{E}\xi^2 \cdot \sum_{m=1}^{\infty} \frac{1}{2^{2m}} = \frac{1}{3}\mathbb{E}\xi^2,$$

что вполне согласуется с нашим вычислением для матрицы ковариации.

Аналогичным образом распишем третий момент вектора Y:

$$\mathbb{E}Y^3 = \mathbb{E}\Big(\sum_{k=0}^{\infty} \frac{\xi_k}{2^k}\Big)^3 \stackrel{\text{def}}{=} \mathbb{E}\Big(\sum_{k=0}^{\infty} \frac{b_k}{2^k}\Big) = \sum_{k=0}^{\infty} \frac{1}{2^k}\mathbb{E}b_k,$$

где случайные вектора b_k соответствуют двукратной свёртке последовательности ξ_k с самой собой.

Иными словами,

$$\Big(\sum_{k=0}^{\infty} \frac{\xi_k}{2^k}\Big)^3 = \sum_{k=0}^{\infty} a_k \cdot \sum_{k=0}^{\infty} \frac{\xi_k}{2^k} = \sum_{k=0}^{\infty}\sum_{n=0}^{k} \Big(\frac{a_n}{2^n} \cdot \frac{\xi_{k-n}}{2^{k-n}}\Big) = \sum_{k=0}^{\infty} \frac{1}{2^k}\Big(\sum_{n=0}^{k} a_n \xi_{k-n}\Big) \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} \frac{1}{2^k} b_k.$$

Тем самым,

$$b_k = \sum_{n=0}^{k} a_n \xi_{k-n} = \sum_{n=0}^{k}\Big(\sum_{m=0}^{n} \xi_m \xi_{n-m}\Big)\xi_{k-n} = \sum_{i+j+l=k} \xi_i \xi_j \xi_l.$$

В свою очередь,

$$\mathbb{E}b_k = \mathbb{E}\Big(\sum_{i+j+l=k} \xi_i \xi_j \xi_l\Big) = \sum_{i+j+l=k} \mathbb{E}(\xi_i \xi_j \xi_l).$$

В силу независимости и центрированности векторов ξ_k:

$$\mathbb{E}(\xi_i \xi_j \xi_l) = \begin{cases} \mathbb{E}\xi^3, & i = j = l; \\ 0, & \text{иначе.} \end{cases}$$

Значит, в случае $k = 3m$ для некоторого $m \in \mathbb{N}$ в сумме будет единственное ненулевое слагаемое, равное $\mathbb{E}\xi^3$. В случаях же $k = 3m+1$ или $k = 3m+2$ для некоторого $m \in \mathbb{N}$ все слагаемые в сумме будут равны нулю, т.е.:

$$\mathbb{E}b_k = \begin{cases} \mathbb{E}\xi^3, & k = 3m \text{ для некоторого } m \in \mathbb{N}; \\ 0, & \text{иначе.} \end{cases}$$

Возвращаясь к третьему моменту предельного вектора Y, имеем:

$$\mathbb{E}Y^3 = \sum_{k=0}^{\infty} \mathbb{E}b_k = \sum_{k=0, k=3m, m\in\mathbb{N}}^{\infty} \frac{1}{2^k}\mathbb{E}\xi^3 = \mathbb{E}\xi^3 \cdot \sum_{m=1}^{\infty} \frac{1}{2^{3m}} = \frac{1}{7}\mathbb{E}\xi^3.$$

Совершенно аналогично расписывается четвёртый момент:

$$\mathbb{E}Y^4 = \sum_{k=0}^{\infty} \frac{1}{2^k}\mathbb{E}c_k, \text{ где } c_k = \sum_{i+j+l+s=k} \xi_i\xi_j\xi_l\xi_s.$$

Оценка $\mathbb{E}c_k$, однако, становится уже немного сложнее. Возможны следующие случаи:

(1) если хотя бы один из индексов отличен от всех остальных, мы можем выделить множитель $\mathbb{E}\xi = 0$, а значит итоговый вклад будет равен нулю;
(2) $i = j = l = s = \frac{k}{4}$, при $k = 4m$ для некоторого $m \in \mathbb{N}$, даст вклад $\mathbb{E}\xi^4$;
(3) $i = j \neq l = s$ или $i = l \neq j = s$ или $i = s \neq j = l$ даст вклад $\left(\mathbb{E}\xi^2\right)^2$.

Последний случай возможен только при $k = 2v + 2u$, т.е. при $\frac{k}{2} = v + u$. Сколько всего таких вариантов представить число $\frac{k}{2} \overset{\text{def}}{=} m$ в виде суммы двух натуральных чисел? Это в точности *количество композиций* m в два числа, что считается по формуле $\binom{m-1}{2-1} = m-1$.

Просуммировав по всем m, получим:

$$\sum_{k=0, k=2m, m\in\mathbb{N}}^{\infty} \frac{1}{2^k}\cdot(m-1)\cdot 3\cdot\left(\mathbb{E}\xi^2\right)^2 = \left(\mathbb{E}\xi^2\right)^2\cdot\sum_{m=1}^{\infty}\frac{1}{2^{2m}}(m-1)\cdot 3 = \left(\mathbb{E}\xi^2\right)^2\cdot 3\cdot\frac{1}{9} = \frac{1}{3}\left(\mathbb{E}\xi^2\right)^2.$$

Нужно, однако, учесть, что при $k = 2m = 4w$, где $w \in \mathbb{N}$, т.е. при k, кратном четырём, в разбиении $k = 2v + 2u$ или, что то же, в разбиении $m = v + u$ мы в этой сумме трижды посчитали лишний вариант $v = u$, т.е. $i = j = l = s$. Значит, необходимо отнять

$$\sum_{k=0, k=4w, w\in\mathbb{N}}^{\infty} \frac{1}{2^k}\cdot 3\cdot\mathbb{E}\xi^4 = \frac{3}{15}\mathbb{E}\xi^4.$$

Итого, последний случай даст вклад, равный:

$$\frac{1}{3}\left(\mathbb{E}\xi^2\right)^2 - \frac{3}{15}\mathbb{E}\xi^4.$$

В свою очередь, первый случай даст вклад, равный:

$$\sum_{k=0, k=4w, w\in\mathbb{N}}^{\infty} \frac{1}{2^k}\mathbb{E}\xi^4 = \frac{1}{15}\mathbb{E}\xi^4.$$

Итого:

$$\mathbb{E}Y^4 = \frac{1}{15}\mathbb{E}\xi^4 + \frac{1}{3}\left(\mathbb{E}\xi^2\right)^2 - \frac{3}{15}\mathbb{E}\xi^4 = \frac{1}{3}\left(\mathbb{E}\xi^2\right)^2 - \frac{2}{15}\mathbb{E}\xi^4.$$

Заметим, что при анализе пятого момента ненулевые слагаемые будут возникать:

- Во-первых, при $k = 5m$ и $i_1 = i_2 = i_3 = i_4 = i_5 = \frac{k}{5}$, что в конечном итоге даст вклад $\frac{1}{31}\mathbb{E}\xi^5$;
- Во-вторых, при $k = 3u + 2v$, и, соответственно, 10 способах разбить 5 индексов суммирования на одну тройку равных и одну пару равных. Значит, потребуется искать число представлений k в виде $k = 3u + 2v$, при $u, v \in \mathbb{N}$.

Для произвольного же момента $\mathbb{E}Y^p$ потребуется ещё более сложный комбинаторный анализ представлений числа k в виде некоторых сумм.

Другой путь к подсчёту моментов высоких порядков состоит в дифференцировании производящей функции моментов Y в точке 0. Догадливый читатель, однако, поймёт, что для взятия производных у соответствующих сумм произведений придётся применять обобщённое правило Лейбница, которое приведёт к не менее сложному комбинаторному анализу мультиномиальных коэффициентов.

Мы же не будем углубляться в дебри комбинаторики и перейдём к рассмотрению куда более интересного вероятностого вопроса.

3.5. **Сходимость по вариации в абсолютно непрерывном случае.** Любопытно исследовать тот случай, когда переходная мера обладает плотностью. Оказывается, тогда слабая сходимость $\mu_n \Rightarrow \nu$ непременно усилится до сходимости *по вариации*:

$$\|\mu_n - \nu\|_{\mathrm{TV}} \xrightarrow[n\to\infty]{} 0.$$

Для удобства будем считать, что $K \subset \mathbb{R}^d$, $X_0 = x_0 = 0 \in K$, $M = \max\limits_{x \in K} |x| < 1$, $\lambda^d(K) < 1$. Добьёмся этого, предварительно сдвинув и сжав компакт (как мы знаем, это не скажется на предельном поведении цепи).

Обозначим:

$$X_k \stackrel{\mathcal{D}}{=} S_k = \frac{\xi_1}{2} + \dots + \frac{\xi_k}{2^k},$$

$$S_k \xrightarrow[k]{a.e.} Y = \sum_{k=1}^{\infty} \frac{\xi_k}{2^k} \stackrel{\text{def}}{=} S_n + Z_n,$$

$$Z_n = \sum_{k=n+1}^{\infty} \frac{\xi_k}{2^k}.$$

Предположим, что переходные случайные вектора $\xi_k, \xi \stackrel{i.i.d.}{\sim} \mu$ абсолютно непрерывны, т.е. имеют плотность $p(x)$ относительно меры лебега λ^d. Тогда $\frac{\xi_k}{2^k}$ также обладают плотностями $r_k(x)$, ровно как и вектора распределений на k-ом шаге:

$$X_k \stackrel{\mathcal{D}}{=} S_k \sim p_k(x),$$

$$p_{k+1}(x) = (p_k * r_{k+1})(x).$$

Видно, что предельный вектор Y также обладает плотностью $q(y)$, которая выражается в терминах плотности вектора S_n и распределения $\mathbb{P}^{-1} Z_n \stackrel{\text{def}}{=} \mu_{Z_n}$ вектора Z_n:

$$q(y) = \int\limits_{\mathbb{R}^d} p_n(y - z)\, \mu_{Z_n}(dz) = \int\limits_K p_n(y - z)\, \mu_{Z_n}(dz).$$

Напомним, что сходимость вероятностных мер по вариации эквивалентна сходимости их плотностей в L^1 (как только последние существуют). Значит, чтобы доказать, что:

$$\|\mu_n - \nu\|_{\mathrm{TV}} \xrightarrow[n\to\infty]{} 0,$$

достаточно показать, что:

$$\|q - p_n\|_{L^1} \xrightarrow[n\to\infty]{} 0.$$

Займёмся оценкой L^1-нормы разности q и p_n. Ключевая идея состоит в том, чтобы воспользоваться видом распределения μ_{Z_n}, а точнее, его сконцентрированностью в небольшом шаре.

$$\|q-p_n\|_{L^1} = \int\limits_K |q(y)-p_n(y)|\,dy = \int\limits_K \Big|\int\limits_K p_n(y-z)\,\mu_{Z_n}(dz) - p_n(y)\Big|\,dy =$$

$$\int\limits_K \Big|\int\limits_K (p_n(y-z)-p_n(y))\,\mu_{Z_n}(dz)\Big|\,dy \leqslant \int\limits_K \Big(\int\limits_K |p_n(y-z)-p_n(y)|\,\mu_{Z_n}(dz)\Big)\,dy =$$

$$\int\limits_K \Big(\int\limits_{z\in K\cap\{|z|<\delta\}} |p_n(y-z)-p_n(y)|\,\mu_{Z_n}(dz)\Big)\,dy + \int\limits_K \Big(\int\limits_{z\in K\cap\{|z|\geqslant\delta\}} |p_n(y-z)-p_n(y)|\,\mu_{Z_n}(dz)\Big)\,dy.$$

Заметим, что:

$$|Z_n| = \Big|\sum_{k=n=1}^{\infty} \frac{\xi_k}{2^k}\Big| \leqslant \Big(\sum_{k=n+1}^{\infty} \frac{1}{2^k}\Big)\cdot M = \frac{M}{2^n} < \frac{1}{2^n}.$$

Это означает, что хвост ряда сосредоточен на маленьком шаре вокруг нуля внутри K. Иными словами,

$$\mu_{Z_n}\Big(K\cap\{|z|\geqslant\frac{1}{2^n}\}\Big) = 0.$$

Положим $\delta = \delta_n = \frac{1}{2^n}$ и увидим, что:

$$\int\limits_K \Big(\int\limits_{z\in K\cap\{|z|\geqslant\delta\}} |p_n(y-z)-p_n(y)|\,\mu_{Z_n}(dz)\Big)\,dy = 0.$$

Таким образом,

$$\|q-p_n\|_{L^1} \leqslant \int\limits_K \Big(\int\limits_{z\in K\cap\{|z|<\delta\}} |p_n(y-z)-p_n(y)|\,\mu_{Z_n}(dz)\Big)\,dy. \tag{3}$$

Для прозрачности оценки предположим дополнительно, что плотность переходного распределения $p(x)$ непрерывна (в дальнейшем мы безболезненно откажемся от этого). По свойствам свёртки отсюда автоматически следует, что $r_k(x)$ и $p_k(x)$ – непрерывны для каждого k.

Введём для любой непрерывной функции $f\in C(K)$ её модуль непрерывности:

$$\omega_f(\delta) \stackrel{\text{def}}{=} \sup_{|t-s|<\delta} |f(t)-f(s)|,\ \text{где } t,s\in K.$$

Заметим, что в подынтегральном выражении (3)

$$|(y-z)-y| = |z| < \delta,$$

и значит

$$|p_n(y-z)-p_n(y)| \leqslant \omega_{p_n}(\delta),$$

где в качестве δ берётся

$$\delta = \delta_n = \frac{1}{2^n}.$$

Подставляя эту оценку в интеграл, вынося модуль непрерывности и пользуясь тем, что μ_{Z_n} – вероятностная мера, получаем:

$$\|q-p_n\|_{L^1} \leqslant \int\limits_K \Bigg(\int\limits_{z\in K\cap\{|z|<\delta\}} |p_n(y-z)-p_n(y)|\,\mu_{Z_n}(dz)\Bigg)\,dy \leqslant$$
$$\int\limits_K \Bigg(\int\limits_{z\in K\cap\{|z|<\delta\}} \omega_{p_n}\Big(\frac{1}{2^n}\Big)\,\mu_{Z_n}(dz)\Bigg)\,dy \leqslant$$
$$\int\limits_K \omega_{p_n}\Big(\frac{1}{2^n}\Big)\cdot 1\,dy \leqslant \lambda^d(K)\cdot\omega_{p_n}\Big(\frac{1}{2^n}\Big) \leqslant 1\cdot\omega_{p_n}\Big(\frac{1}{2^n}\Big).$$

Итак,

$$\|q-p_n\|_{L^1} \leqslant \omega_{p_n}\Big(\frac{1}{2^n}\Big).$$

Нам бы хотелось заключить, что

$$\omega_{p_n}(\frac{1}{2^n}) \xrightarrow[n\to\infty]{} 0,$$

и затруднение состоит лишь в том, что уменьшение аргумента происходит с одновременным переходом к новой функции p_n.

Но из невозрастания модулей непрерывности по n,

$$\omega_{p_1} \geqslant \omega_{p_2} \geqslant \ldots \geqslant \omega_{p_n}, \tag{4}$$

мгновенно вытекает желаемое

$$\omega_{p_n}(\frac{1}{2^n}) \leqslant \omega_{p_1}(\frac{1}{2^n}) \xrightarrow[n\to\infty]{} 0$$

по непрерывности p_1 на компакте K.

Осталось оправдать (4).

Лемма 4. *Для любого n $\omega_{p_n} \leqslant \omega_{p_{n-1}}$.*

Доказательство.

$$\omega_{p_n}(\delta) = \omega_{p_{n-1}*r_n}(\delta) = \sup_{|t-s|<\delta} |(p_{n-1}*r_n)(t) - (p_{n-1}*r_n)(s)| =$$
$$\sup_{|t-s|<\delta}\Big|\int\limits_K p_{n-1}(t-y)r_n(y)\,dy - \int\limits_K p_{n-1}(s-y)r_n(y)\,dy\Big| \leqslant$$
$$\sup_{|t-s|<\delta}\int\limits_K |r_n(y)|\cdot|p_{n-1}(t-y)-p_{n-1}(s-y)|\,dy.$$

Заметим, что в последнем выражении

$$|(t-y)-(s-y)| = |t-s| < \delta,$$

и значит

$$|p_{n-1}(t-y)-p_{n-1}(s-y)| \leqslant \omega_{p_{n-1}}(\delta).$$

Таким образом,

$$\sup_{|t-s|<\delta}\int_K |r_n(y)|\cdot|p_{n-1}(t-y)-p_{n-1}(s-y)|\,dy \leqslant$$

$$\sup_{|t-s|<\delta}\int_K |r_n(y)|\cdot\omega_{p_{n-1}}(\delta)\,dy \leqslant$$

$$\omega_{p_{n-1}}(\delta)\cdot\int_K |r_n(y)|\,dy = \omega_{p_{n-1}}(\delta)\cdot\int_K r_n(y)\,dy = \omega_{p_{n-1}}(\delta)\cdot 1 = \omega_{p_{n-1}}(\delta),$$

что и требовалось. □

Далее нам понадобится аналог модуля непрерывности для интегрируемой функции. Напомним, что сдвиги – непрерывны в L^1. Точнее, если $f\in L^1(\mathbb{R}^d)$, $h\in\mathbb{R}^d$ – фиксированный вектор и $(\tau_h f)(x) \stackrel{\text{def}}{=} f(x-h)$ – оператор сдвига на h, то:

- $\tau_h f\in L^1$,
- $\|\tau_h f - f\|_{L^1} = o(1)$ при $|h|\to 0$,
- $\|\tau_{v+h} f - \tau_v f\|_{L^1} = o(1)$ при $|h|\to 0$ равномерно по $v\in\mathbb{R}^d$.

В силу этого замечательного свойства сдвига для любой функции $f\in L^1(\mathbb{R}^d)$ определён *модуль непрерывности в* L^1, т.е. такая возрастающая функция $\omega_f\colon [0,+\infty)\to[0,+\infty)$, что:

- $\|\tau_h f - f\|_{L^1}\leqslant \omega_f(|h|)$,
- $\lim\limits_{t\to 0}\omega_f(t)=\omega_f(0)=0$.

Попробуем теперь отказаться от непрерывности $p(x)$. При этом p, r_k, p_k, q всё ещё останутся плотностями соответствующих распределений. В частности, все они лежат в L^1.

Заметим также, что оценка (3) останется верной, ибо в её выводе мы не пользовались непрерывностью $p(x)$.

Применим свойства модуля непрерывности в L^1 и доведём её до конца.

$$\|q-p_n\|_{L^1}\leqslant \int_K\left(\int_{|z|<\delta}|p_n(y-z)-p_n(y)|\,\mu_{Z_n}(dz)\right)dy =$$

$$\int_K\left(\int_{|z|<\delta}|(\tau_z p_n)(y)-p_n(y)|\,\mu_{Z_n}(dz)\right)dy =$$

$$\int_{|z|<\delta}\left(\int_K|(\tau_z p_n)(y)-p_n(y)|\,dy\right)\mu_{Z_n}(dz) =$$

$$\int_{|z|<\delta}\|\tau_z p_n - p_n\|_{L^1}\,\mu_{Z_n}(dz)\leqslant \int_{|z|<\delta}\omega_{p_n}(|z|)\,\mu_{Z_n}(dz)\leqslant$$

$$\int_{|z|<\delta}\omega_{p_n}(\delta)\,\mu_{Z_n}(dz)\leqslant \omega_{p_n}(\delta)\cdot 1 = \omega_{p_n}\left(\frac{1}{2^n}\right)\leqslant \omega_{p_1}\left(\frac{1}{2^n}\right)\xrightarrow[n\to\infty]{} 0.$$

Последний переход оправдаем отдельно.

Лемма 5. *Для любого n $\omega_{p_n} \leqslant \omega_{p_{n-1}}$.*

Доказательство.

$$\begin{aligned}
\omega_{p_n}(\delta) = \omega_{p_{n-1}*r_n}(\delta) = &\sup_{|h|<\delta} \|\tau_h(p_{n-1} * r_n) - p_{n-1} * r_n\|_{L^1} = \\
&\sup_{|h|<\delta} \int\limits_K \Big|(\tau_h(p_{n-1} * r_n))(x) - (p_{n-1} * r_n)(x)\Big|\,dx = \\
&\sup_{|h|<\delta} \int\limits_K \Big|(p_{n-1} * r_n)(x-h) - (p_{n-1} * r_n)(x)\Big|\,dx = \\
&\sup_{|h|<\delta} \int\limits_K \Big|\int\limits_K p_{n-1}(x-h-y) r_n(y)\,dy - \int\limits_K p_n(x-y) r_n(y)\,dy\Big|\,dx = \\
&\sup_{|h|<\delta} \int\limits_K \Big|\int\limits_K r_n(y)\Big(p_{n-1}(x-h-y) - p_{n-1}(x-y)\Big)\,dy\Big|\,dx \leqslant \\
&\sup_{|h|<\delta} \int\limits_K \left(\int\limits_K \Big|p_{n-1}(x-y-h) - p_{n-1}(x-y)\Big|\,dx\right) \cdot |r_n(y)|\,dy = \\
&\sup_{|h|<\delta} \int\limits_K \left(\int\limits_K \Big|p_{n-1}(u-h) - p_{n-1}(u)\Big|\,du\right) \cdot |r_n(y)|\,dy = \\
&\sup_{|h|<\delta} \int\limits_K \|\tau_h p_{n-1} - p_{n-1}\|_{L^1} \cdot |r_n(y)|\,dy \leqslant \\
&\int\limits_K \omega_{p_{n-1}}(\delta) \cdot |r_n(y)|\,dy \leqslant \omega_{p_{n-1}}(\delta) \cdot 1 = \omega_{p_{n-1}}(\delta),
\end{aligned}$$

что и требовалось. □

Таким образом, мы доказали следующее утверждение.

Предложение 4. *Если переходное распределение μ абсолютно непрерывно, то предельное распределение ν также абсолютно непрерывно. Более того, в этом случае распределения цепи на n-ом шаге сходятся к предельному распределению по вариации.*

3.6. **Произвольные выпуклые комбинации.** Вернёмся к изначальной постановке задачи. Напомним её геометрический смысл: из текущей точки проводится отрезок в одну из точек компакта в соответствии с переходным распределением, после чего осуществляется переход *в середину* проведённого отрезка. Оказывается, ничего не изменится, если разбивать отрезок не пополам, а в произвольном отношении!

Как и раньше, мера μ_0 соответсвует некоторому начальному распределению на выпуклом компакте $K \subset \mathbb{R}^d$, а $\xi_k \overset{i.i.d.}{\sim} \mu$ обозначают переходные случайные вектора.

Пусть $0 \leqslant \alpha, \beta \leqslant 1$ и $\alpha + \beta = 1$.

Определение 12. Зададим на K *особую марковскую цепь* $\{X_n\}_{n=0}^{\infty}$ *с выпуклыми весами* α, β по правилу:

$$X_0 \sim \mu_0,$$
$$X_{n+1} = \alpha X_n + \beta \xi_{n+1}.$$

Распишем теперь подробнее случайные вектора $X_0, X_1, \ldots, X_n, \ldots$:

$$
\begin{aligned}
&X_0 \sim \mu_0,\\
&X_1 = \alpha X_0 + \beta \xi_1 \in [X_0, \xi_1] \subset K,\\
&X_2 = \alpha X_1 + \beta \xi_2 = \alpha(\alpha X_0 + \beta \xi_1) + \beta \xi_2 = \alpha^2 X_0 + \alpha\beta\xi_1 + \beta\xi_2,\\
&X_3 = \alpha X_2 + \beta \xi_3 = \alpha(\alpha^2 X_0 + \alpha\beta\xi_1 + \beta\xi_2) + \beta\xi_3 = \alpha^3 X_0 + \alpha^2\beta\xi_1 + \alpha\beta\xi_2 + \beta\xi_3,\\
&X_n = \alpha^n X_0 + \alpha^{n-1}\beta\xi_1 + \alpha^{n-2}\beta\xi_2 + \ldots + \alpha^2\beta\xi_{n-2} + \alpha^1\beta\xi_{n-1} + \beta\xi_n.
\end{aligned}
$$

Таким образом,

$$X_n = \alpha^n X_0 + \sum_{k=1}^{n} \beta\alpha^{n-k}\xi_k \overset{\mathcal{D}}{=} \alpha^n X_0 + \sum_{k=1}^{n} \beta\alpha^{k-1}\xi_k,$$

что вполне согласуется с изученным ранее вдоль и поперёк случаем $\alpha = \beta = \frac{1}{2}$.

Рассмотрим отдельно два симметричных вырожденных случая:

- Если $\alpha = 1, \beta = 0$, то: $X_n = X_0 \sim \mu_0$ для каждого n. Тогда предельное распределение ν есть начальное распределение μ_0: $\nu = \mu_0$.
- Если же $\alpha = 0, \beta = 1$, то: $X_n = \xi_n \sim \mu$ для каждого n. Тогда предельное распределение ν есть переходное распределение μ: $\nu = \mu$.

Далее считаем, что $0 < \alpha, \beta < 1$ и цепь не вырождена. Доказательство существования предельного распределения слово в слово повторяет вывод для $\alpha = \beta = \frac{1}{2}$.

Обозначим:

$$X_n \overset{\mathcal{D}}{=} \alpha^n X_0 + \beta \sum_{k=1}^{n} \alpha^{k-1}\xi_k \overset{\text{def}}{=} \alpha^n X_0 + S_n.$$

Поскольку $0 < \alpha < 1$, первое слагаемое всюду стремится к 0.

В силу компактности $K \subset \mathbb{R}^d$, множество K – ограничено. Обозначим радиус компакта:

$$M \overset{\text{def}}{=} \max_{x \in K} |x| < \infty.$$

Но ведь переходные вектора $\xi_k \in K$ почти наверное. Значит, $|\xi_k| \leqslant M$ и

$$\sum_{k=1}^{\infty} |\beta\alpha^{k-1}\xi_k| = \sum_{k=1}^{\infty} \beta\alpha^{k-1}|\xi_k| \leqslant \beta M \cdot \sum_{k=1}^{\infty} \alpha^{k-1} = \beta M \cdot \frac{1}{1-\alpha} = \beta M \cdot \frac{1}{\beta} = M.$$

По полноте объемлющего пространства из абсолютной сходимости немедленно следует

$$\alpha^n X_0 + S_n \xrightarrow[n\to\infty]{a.e.} Y = \lim_{n\to\infty} S_n.$$

Таким образом,

$$X_n \Rightarrow Y = \sum_{k=1}^{\infty} \beta\alpha^{k-1}\xi_k.$$

Считая, что $Y \sim \nu \in \mathcal{P}(K)$, получим слабую сходимость к предельному распределению $\mu_n \Rightarrow \nu$. Из определения Y видно, что его распределение ν не зависит от μ_0 и однозначно определяется распределением случайных векторов $\xi_k \sim \mu$.

А значит, для особой марковской цепи с произвольными выпуклыми весами:

- Существует единственное предельное распределение, не зависящее от начального распределения μ_0.
- Существует единственное инвариантное распределение.

Оценка скорости сходимости подвергнется лишь косметическим изменениям. Обозначим: $m \stackrel{\text{def}}{=\!=} \mathbb{E}|\xi_1|$. Тогда:

$$d(\mu_n, \nu) \leqslant \ldots \leqslant \mathbb{E}|X_n - Y| = \mathbb{E}\Big|\Big(\alpha^n X_0 + W_n\Big) - \Big(W_n + \sum_{k=n+1}^{\infty} \beta\alpha^{k-1}\xi_k\Big)\Big| \leqslant$$

$$\alpha^n \cdot \mathbb{E}|X_0| + \sum_{k=n+1}^{\infty} \mathbb{E}|\beta\alpha^{k-1}\xi_k| \leqslant$$

$$\alpha^n \cdot M + \mathbb{E}|\xi_1| \cdot \sum_{k=n+1}^{\infty} \beta\alpha^{k-1} \leqslant \alpha^n M + m \cdot \beta\frac{\alpha^n}{1-\alpha} = \alpha^n(M+m).$$

Поскольку $0 < \alpha < 1$, мы вновь имеем экспоненциально быструю сходимость к предельному распределению.

Заметим, что инвариантность задачи относительно аффинных преобразований объемлющего пространства получается автоматически, ибо замена выпуклой комбинации с коэффициентами, равными степеням двойки, на коэффициенты, зависящие от параметров $0 < \alpha, \beta < 1$, ничего не изменит: аффинное отображение всё ещё будет переводить выпуклые комбинации в выпуклые комбинации.

Останется верным и такой результат: если переходное распределение инвариантно относительно некоторого невырожденного линейного преобразования, то и предельное распределение инвариантно относительно него.

Пересчитаем формулы для моментов. Не умаляя общности, $X_0 = 0 \in K$. Тогда:

$$\mathbb{E}Y = \mathbb{E}\Big(\sum_{k=1}^{\infty} \beta\alpha^{k-1}\xi_k\Big) = \beta \cdot \mathbb{E}\xi_1 \cdot \sum_{k=1}^{\infty} \alpha^{k-1} = \mathbb{E}\xi_1 \cdot \frac{\beta}{1-\alpha} = \mathbb{E}\xi_1.$$

$$\operatorname{cov}(Y) = \operatorname{cov}\Big(\sum_{k=1}^{\infty} \beta\alpha^{k-1}\xi_k\Big) = \sum_{k=1}^{\infty} \operatorname{cov}(\beta\alpha^{k-1}\xi_k) =$$

$$\operatorname{cov}(\xi_1) \cdot \sum_{k=1}^{\infty} \alpha^{2(k-1)}\beta^2 = \operatorname{cov}(\xi_1) \cdot \frac{\beta^2}{1-\alpha^2} = \operatorname{cov}(\xi_1) \cdot \frac{1-\alpha}{1+\alpha}.$$

Новая формула, конечно же, согласуется со старой при $\alpha = \beta = \frac{1}{2}$.

Покажем, что абсолютная непрерывность переходного распределения μ, как и прежде, гарантирует сходимость распределений μ_n к ν по вариации.

Не умаляя общности, $X_0 = 0 \in K$. Введём обозначения:

$$X_n \overset{\mathcal{D}}{=} S_n = \sum_{k=1}^{n} \beta\alpha^{k-1}\xi_k \xrightarrow[n\to\infty]{} Y = \sum_{k=1}^{\infty} \beta\alpha^{k-1}\xi_k \overset{\text{def}}{=} S_n + Z_n,$$

$$Z_n = \sum_{k=n+1}^{\infty} \beta\alpha^{k-1}\xi_k.$$

Проделаем те же выкладки, что и раньше, с единственным отличием:

$$|Z_n| = \Big|\sum_{k=n+1}^{\infty} \beta\alpha^{k-1}\xi_k\Big| \leqslant M \cdot \sum_{k=n+1}^{\infty} \beta\alpha^{k-1} = M \cdot \beta\frac{\alpha^n}{1-\alpha} = M\alpha^n.$$

Для удобства (после сжатия компакта K), считаем, что $M < 1$. Тогда в качестве $\delta = \delta_n$ достаточно взять $\delta = \alpha^n$. Поскольку $0 < \alpha < 1$,

$$\delta_n = \alpha^n \xrightarrow[n\to\infty]{} 0,$$

а значит,

$$\|p - p_n\|_{L^1} \leqslant \omega_{p_n}(\alpha^n) \leqslant \omega_{p_1}(\alpha^n) \xrightarrow[n\to\infty]{} 0.$$

4. Исследование конкретных примеров

4.1. **Случай отрезка** $[0,1]$**.** Рассмотрим $K = [0,1] \subset \mathbb{R}^1$.

- Поскольку предельное распределение не зависит от выбора μ_0, для удобства положим $\mu_0 = \delta_{\{\frac{1}{2}\}}$,
- В качестве переходной меры возьмём $\mu = \frac{1}{2}\delta_{\{0\}} + \frac{1}{2}\delta_{\{1\}}$.

Иными словами,

- $X_0 \overset{a.e.}{=} \frac{1}{2}$;
- $\xi_k \overset{i.i.d.}{\sim} B(\frac{1}{2})$, т.е. $\xi_k = \begin{cases} 1, & p = \frac{1}{2}; \\ 0, & q = \frac{1}{2}; \end{cases}$
- $X_n \Rightarrow Y = \displaystyle\sum_{k=1}^{\infty} \frac{\xi_k}{2^k} = 0,\xi_1\xi_2\xi_3\ldots$

В действительности, Y представляет из себя двоичную запись случайного числа на отрезке $[0,1]$ со средним числом единиц в записи, равным $\frac{1}{2}$. Из Закона Больших Чисел для случайных величин ξ_k вытекает, что:

$$\mathbb{P}Y^{-1} = U[0,1], \text{ т.е. } \nu = \lambda^1[0,1].$$

Пусть теперь $0 < p < 1, p \neq \frac{1}{2}$,

- $X_0 \overset{a.e.}{=} \frac{1}{2}$;
- $\xi_k \overset{i.i.d.}{\sim} B(p)$, т.е. $\xi_k = \begin{cases} 1, & p; \\ 0, & q = 1 - p; \end{cases}$
- $X_n \Rightarrow Y_p \overset{\text{def}}{=} \displaystyle\sum_{k=1}^{\infty} \frac{\xi_k}{2^k} = 0,\xi_1\xi_2\xi_3\ldots$

Теперь уже Y_p представляет из себя двоичную запись случайного числа на отрезке $[0,1]$ со средним число единиц в записи, равным $p \neq \frac{1}{2}$. Из Закона Больших Чисел для случайных величин ξ_k вытекает, что:

$$\frac{\xi_1 + \ldots + \xi_n}{n} \xrightarrow[n\to\infty]{a.e.} \mathbb{E}\xi = p,$$

т.е. доля единиц среди первых n знаков после запятой в двоичной записи случайного числа Y_p стремится к $p \neq \frac{1}{2}$.

Введём обозначения:

$$A_p \overset{\text{def}}{=} \{x = 0,x_1x_2x_3\ldots \in [0,1] : \lim_{n\to\infty} \frac{x_1 + \ldots + x_n}{n} = p\}.$$

Пусть $Y_p \sim \nu_p$. Видно, что $\nu_p(A_p) = \mathbb{P}Y_p^{-1}(A_p) = 1$, в то время как $\lambda^1(A_p) = 0$, т.е. предельное распределение ν_p сингулярно относительно меры Лебега на отрезке: $\nu_p \perp \lambda^1[0,1]$. Более того, для любых $0 < p \neq \tilde{p} < 1$: $\nu_p \perp \nu_{\tilde{p}}$, поскольку множества A_p попарно дизъюнктны при различных p.

Таким образом, мы доказали следующее утверждение.

Предложение 5. *Варьирование нагрузкок переходной меры в концевых точках отрезка даёт континуальное семейство попарно сингулярных предельных распределений*

$$\{\nu_p\}_{p\in(0,1)},$$

сосредоточенных на множествах A_p Лебеговой меры 0. Исключением является случай $p = \frac{1}{2}$, при котором предельное распределение есть в точности $\lambda^1[0,1]$.

4.2. **Случай куба I^d.** Рассмотрим для начала квадрат $K = I^2 \subset \mathbb{R}^d$, $I = [0,1]$ с равными нагрузками в своих вершинах:

- $X_0 \overset{a.e.}{=} (\frac{1}{2}, \frac{1}{2})$;
- $\xi_k \overset{i.i.d.}{\sim} \mu = \frac{1}{4}(\delta_{\{(0,0)\}} + \delta_{\{(0,1)\}} + \delta_{\{(1,0)\}} + \delta_{\{(1,1)\}})$;
- $X_n \Rightarrow Y = \sum_{k=1}^{\infty} \frac{\xi_k}{2^k}$.

Покажем, что $\mathbb{P}Y^{-1} = \lambda^2[0,1]$.

Обозначим компоненты вектора верхними индексами: $X_n = (X_n^1, X_n^2), \xi_k = (\xi_k^1, \xi_k^2)$. Тогда для $j = 1,2$:

$$X_n^j \overset{\mathcal{D}}{=} \frac{X_0^j}{2^n} + \frac{\xi_1^j}{2} + \cdots + \frac{\xi_n^j}{2^n}.$$

Из симметрий квадрата следует (в чём читатель может убедиться самостоятельно), что X_n^1, X_n^2 – независимы при каждом $n = 1,2,\ldots$ и, кроме того, для каждого $j = 1,2$:

$$\xi_k^j \overset{i.i.d.}{\sim} B\Big(\frac{1}{2}\Big).$$

Значит, распределения μ_n, определяемые случайными векторами $X_n \in \mathbb{R}^2$, распадаются в произведения одномерных распределений соответствующих случайных компонент:

- $\mu_n = \mu_n^1 \times \mu_n^2$, где $X_n \sim \mu_n, X_n^j \sim \mu_n^j$,
- $I^2 = [0,1] \times [0,1]$,
- $\lambda^2[0,1] = \lambda^1[0,1] \times \lambda^1[0,1]$,

- $\mu_n^1 \Rightarrow \lambda^1[0,1]$, $\mu_n^2 \Rightarrow \lambda^1[0,1]$ (случай отрезка).

В силу сепарабельности метрического пространства I^2, можно воспользоваться теоремой 3.2 [1] и заключить, что:

$$\mu_n \Rightarrow \lambda^2[0,1].$$

Рассмотрим, наконец, d-мерный куб $K = I^d \subset \mathbb{R}^d$ с равными нагрузками во всех 2^d вершинах e_j, $j = 1, \dots, 2^d$:

- $X_0 \overset{a.e.}{=} (\frac{1}{2}, \dots, \frac{1}{2})$;
- $\xi_k \overset{i.i.d.}{\sim} \mu = \dfrac{1}{2^d} \displaystyle\sum_{j=1}^{2^d} \delta_{\{e_j\}}$;
- $X_n \Rightarrow Y = \displaystyle\sum_{k=1}^{\infty} \dfrac{\xi_k}{2^k}$.

Покажем, что $\mathbb{P}Y^{-1} = \lambda^d[0,1]$.

Из симметрий d-мерного единичного куба следует, что $X_n^1, \dots, X_n^d$ – независимы при каждом $n = 1, 2, \dots$ и, кроме того, для каждого $j = 1, \dots, d$:

$$\xi_k^j \overset{i.i.d.}{\sim} B\Big(\frac{1}{2}\Big).$$

Значит, распределения μ_n вновь распадаются в произведения одномерных распределений соответствующих случайных компонент:

- $\mu_n = \mu_n^1 \times \dots \times \mu_n^d$, где $X_n \sim \mu_n, X_n^j \sim \mu_n^j$,
- $I^d = [0,1] \times \dots \times [0,1]$,
- $\lambda^d[0,1] = \lambda^1[0,1] \times \dots \times \lambda^1[0,1]$,
- $\mu_n^1 \Rightarrow \lambda^1[0,1], \cdots, \mu_n^d \Rightarrow \lambda^1[0,1]$ (случай отрезка).

Пользуясь сепарабельностью метрического пространства I^d, заключаем, что:

$$\mu_n \Rightarrow \lambda^n[0,1].$$

Таким образом, мы доказали следующее утверждение.

Предложение 6. *Предельным распределением особой марковской цепи на d-мерном единичном кубе в $\mathbb{R}^d$ с переходным распределением, равным сумме одинаковых нагрузок в вершинах куба, является равномерныое распределение $\lambda^d[0,1]^d$.*

4.3. **Случай диска D^2.** Рассмотрим плоский диск $K = D^2 \subset \mathbb{R}^2$,

$$D^2 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant 1\},\ \partial D^2 = S^1.$$

В качестве переходной меры на нём кажется разумным взять одно из следующих распределений:

(1) $\mu = \frac{H^1}{2\pi} \overset{\text{def}}{=} U(S^1)$ – равномерное распределение на окружности;

(2) $\mu = \frac{\lambda^2}{\lambda^2(D^2)} \overset{\text{def}}{=} U(D^2)$ – равномерное распределение на всём диске.

Начальное распределение, не изменяя традиции, положим равным $\mu_0 = \delta_{\{0\}}$.

Займёмся первым случаем: пусть $\mu \overset{\text{def}}{=} U(S^1)$ – равномерное распределение на единичной окружности. Проанализируем поведение соответствующей марковской цепи.

- Поскольку $X_1 = \frac{0+\xi_1}{2} = \frac{\xi_1}{2}$, распределение после одного шага будет сосредоточено на центральной окружности:
$$X_1 \in \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = \frac{1}{2}\} \stackrel{\text{def}}{=} \mathrm{Ann}\Big(\frac{1}{2}, \frac{1}{2}\Big).$$
- Проделаем ещё один шаг: $X_2 = \frac{\xi_1}{4} + \frac{\xi_2}{2}$. Воспользуемся симметричностью распределения и оценим $|X_2|$ с помощью неравенства треугольника:
$$\frac{1}{4} \leqslant \Big|\frac{1}{4} - \frac{1}{2}\Big| = \Big||\frac{\xi_1}{4}| - |\frac{\xi_2}{2}|\Big| \leqslant \Big|\frac{\xi_1}{4} - \frac{\xi_2}{2}\Big| \stackrel{\mathcal{D}}{=} \Big|\frac{\xi_1}{4} + \frac{\xi_2}{2}\Big| \leqslant \Big|\frac{\xi_1}{4}\Big| + \Big|\frac{\xi_2}{2}\Big| = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}.$$
Другими словами, распределение на втором шаге будет сосредоточено на центральном кольце:
$$X_2 \in \{(x,y) \in \mathbb{R}^2 : \frac{1}{4} \leqslant x^2 + y^2 \leqslant \frac{3}{4}\} = \mathrm{Ann}\Big(\frac{1}{4}, \frac{3}{4}\Big).$$
- Наконец, распределение на n-ом шаге будет сосредоточено на центральном кольце:
$$X_n \in \mathrm{Ann}\Big(\frac{1}{2^n}, 1 - \frac{1}{2^n}\Big).$$

На следующих иллюстрациях представлено приблизительное поведение предельного распределения ν. Графики были получены следующим образом:

(1) Из равномерного распределения на S^1 (слева) и равномерного распределения на D^2 (справа) было сэмплировано $N = 10^5$ точек z_j.
(2) Для каждой из N точек z_j проделано $n = 100$ шагов марковской цепи, так что z_j перешли в $h(z_j) = \sum_{k=1}^{n=100} \frac{1}{2^k} z_j$. Благодаря экспоненциальной скорости сходимости, уже $n = 100$ шагов дают отличное приближение.
(3) По полученным точкам $h(z_j)$ построено их распределение: рис. 1, 2.
(4) По полученным точкам $h(z_j)$ построена гистограмма их маргинальной плотности (по первой координате): рис. 3, 4.
(5) По полученным точкам $h(z_j)$ построена гистограмма их плотности: рис. 5, 6.

Видно, что в случае $\mu = U(S^1)$ предельное распределение ν "размазано" по всему диску с большим сосредоточением вокруг центральных окружностей и постепенным спадом от них как к нулю, так и к границе диска. На иллюстрации это подчёркивается "дыркой" вокруг нуля и пробелом около внешней окружности.

В то же время, в случае $\mu = U(D^2)$ предельное распределение ν значительно более сконцентрированно около нуля с довольно скорым спадом к внешней окружности.

Кроме того, эмпирический подсчёт дисперсии по первой координате выдаёт для случаев $\mu = U(S^1)$ и $\mu = U(D^2)$ нечто около $\frac{1}{6}$ и $\frac{1}{12}$, соответственно. Ниже мы аналитически подтвердим данные результаты, а пока перейдём к многомерному случаю.

4.4. **Случай диска D^n.** Рассмотрим теперь n-мерный диск $K = D^n \subset \mathbb{R}^n$,
$$D^n = \{x \in \mathbb{R}^n : |x| \leqslant 1\},\ \partial D^n = S^{n-1}.$$

Продолжим работу с первым случаем: пусть $\mu = U(S^1)$ – равномерное распределение на единичной сфере, т.е.:
$$\mu = \sigma^{n-1} = \frac{H^{n-1}}{H^{n-1}(S^{n-1})}.$$

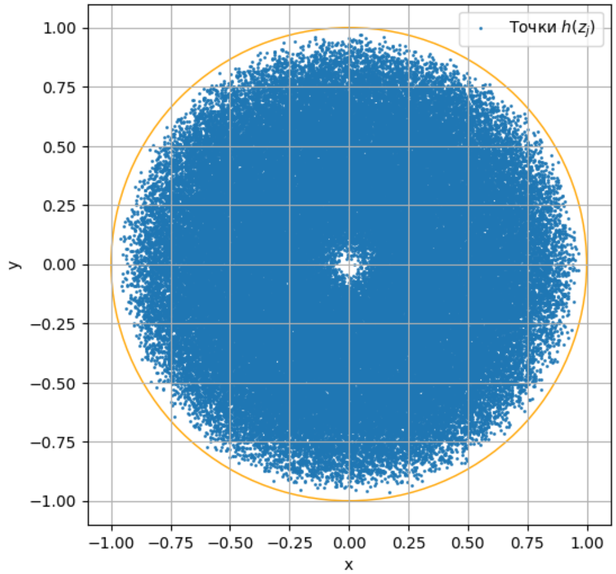

Рис. 1. Предельное распределение ν при $\mu = U(S^1)$

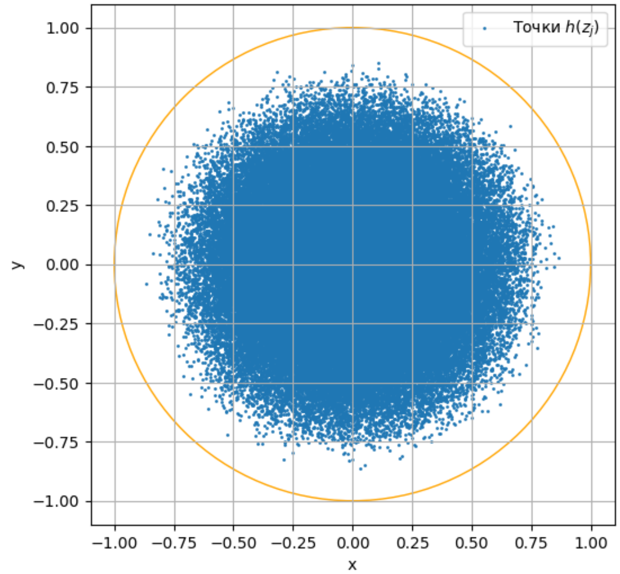

Рис. 2. Предельное распределение ν при $\mu = U(D^2)$

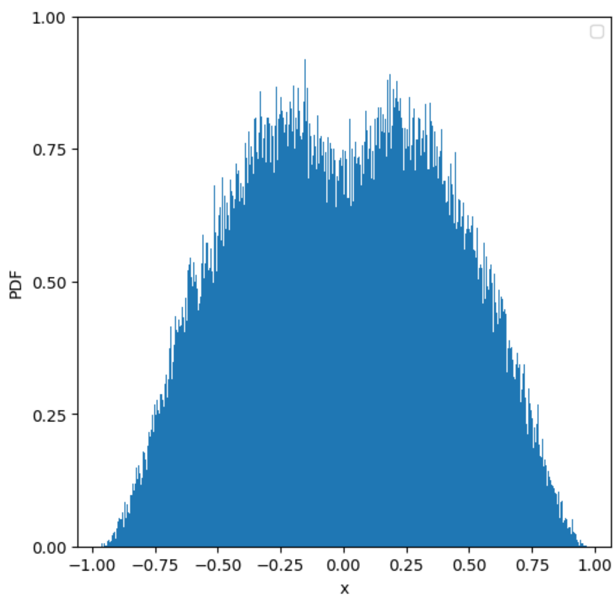

Рис. 3. Маргинальная плотность ν при $\mu = U(S^1)$

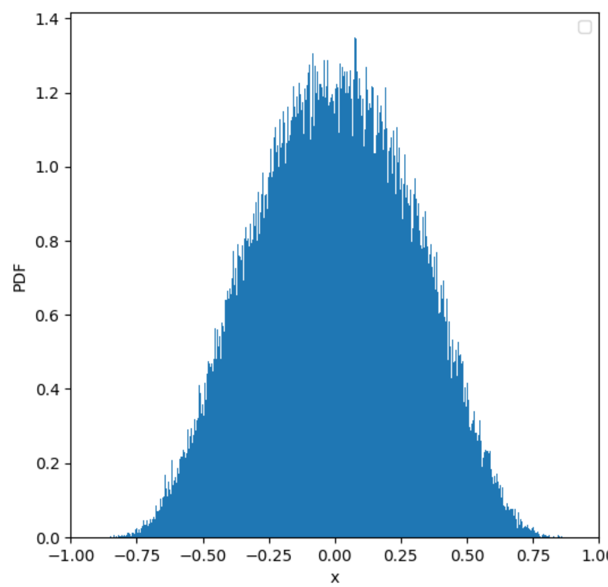

Рис. 4. Маргинальная плотность ν при $\mu = U(D^2)$

Лемма 6. *Матрица ковариации стандартной поверхностной меры σ^{n-1} на сфере S^{n-1} имеет вид:*

$$\operatorname{cov}(\sigma^{n-1}) = \frac{1}{n} \cdot I_n.$$

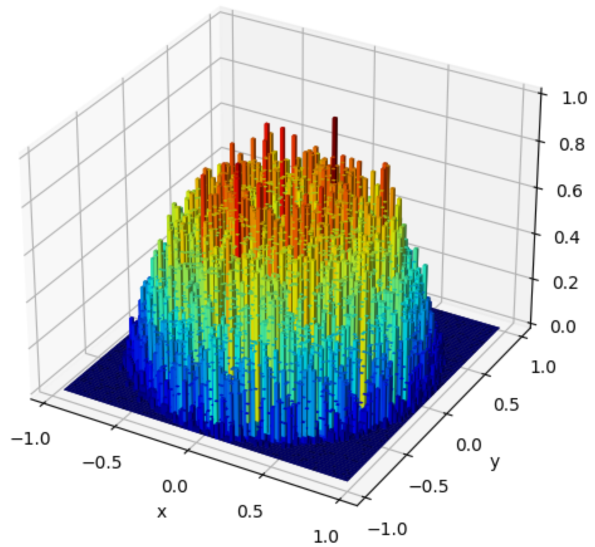

Рис. 5. Плотность ν при $\mu = U(S^1)$

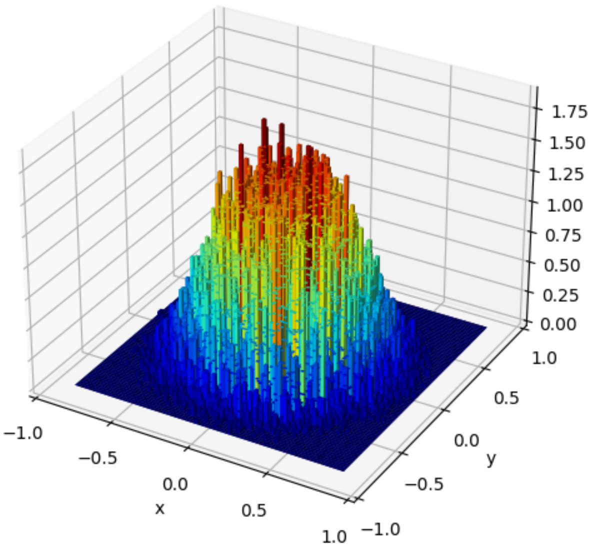

Рис. 6. Плотность ν при $\mu = U(D^2)$

Доказательство. Рассмотрим случайный вектор $Z = (Z_1, \ldots, Z_n)$, $Z \in S^{n-1}$, $Z \sim \sigma^{n-1}$. Благодаря инвариантности меры σ^{n-1} относительно соответствующих симметрий, имеем:

$$\begin{gathered}(Z_1, \ldots, Z_i, \ldots, Z_n) \overset{\mathcal{D}}{=} (Z_1, \ldots, -Z_i, \ldots, Z_n),\\ \mathbb{E}Z_i = -\mathbb{E}Z_i,\\ \mathbb{E}(Z_iZ_j) = -\mathbb{E}(Z_iZ_j),\ i \neq j.\end{gathered}$$

Но тогда $\mathbb{E}Z = 0$ и $\mathbb{E}(Z_iZ_j) = 0$ для любых $i \neq j$. Значит, все недиагональные члены в матрице $\operatorname{cov}(Z)$ равны нулю.

Кроме того, $|Z|^2 = 1$, поскольку $Z \in S^{n-1}$. В частности, $\mathbb{E}|Z|^2 = 1$. Заметим, наконец, что Z_i взаимозаменяемы в силу симметрий, а значит:

$$1 = \mathbb{E}(|Z|^2) = \mathbb{E}(Z_1^2 + \ldots + Z_n^2) = \mathbb{E}Z_1^2 + \ldots + \mathbb{E}Z_n^2 = n \cdot \mathbb{E}Z_1^2, \text{ откуда } \mathbb{E}Z_i^2 = \frac{1}{n}.$$

□

Вернёмся к двумерному диску и применим лемму к случайному вектору $\xi \sim \mu$, распределённому равномерно на единичной окружности $S^1 \subset \mathbb{R}^2$.

$$\operatorname{cov}(\xi) = \frac{1}{2} \cdot I_2, \text{ и } \operatorname{cov}(Y) = \frac{1}{3} \operatorname{cov} \xi = \frac{1}{6} \cdot I_2.$$

В случае произвольного n, получаем:

$$\operatorname{cov}(Y) = \frac{1}{3n} \cdot I_n.$$

Лемма 7. *Матрица ковариации равномерного распределения $U(D^n)$ на диске имеет вид:*

$$\operatorname{cov}(U(D^n)) = \frac{1}{n+2} \cdot I_n.$$

Доказательство. Рассмотрим случайный вектор $Z = (Z_1, \ldots, Z_n)$, $Z \in D^n$, $Z \sim U(D^n)$. Как и для сферы, воспользуемся инвариантностью равномерного распределения относительно симметрий диска:

$$\mathbb{E}Z = 0, \text{ и } \mathbb{E}(Z_i Z_j) = 0, \; i \neq j.$$

В частности, матрица ковариации вектора Z диагональна.

Помимо этого, имеем:

$$\mathbb{E}Z_i^2 = \frac{1}{n}\mathbb{E}|Z|^2.$$

Для вычисления диагональных элементов матрицы ковариации введём случайную величину $R = |Z|$, $R \in [0,1]$. Выразим её функцию распределения и плотность:

$$F_R(r) = \mathbb{P}(R \leqslant r) = \frac{\lambda^n(B(0,r))}{\lambda^n(B(0,1))} = r^n,$$

$$p_R(r) = \frac{dF_R}{dr} = \frac{dr^n}{dr} = nr^{n-1}.$$

Значит,

$$\mathbb{E}Z_i^2 = \frac{1}{n}\mathbb{E}|Z|^2 = \frac{1}{n}\int_0^1 r^2 \cdot nr^{n-1}\, dr = \frac{1}{n+2}.$$

□

Вернёмся к двумерному диску и применим лемму к случайному вектору $\xi \sim \mu$, распределённому равномерно на всём диске $D^2 \subset \mathbb{R}^2$.

$$\operatorname{cov}(\xi) = \frac{1}{4} \cdot I_2, \text{ и } \operatorname{cov}(Y) = \frac{1}{3}\operatorname{cov}\xi = \frac{1}{12} \cdot I_2.$$

В случае произвольного n, получаем:

$$\operatorname{cov}(Y) = \frac{1}{3(n+2)} \cdot I_n.$$

В частности, подсчёт матрицы ковариации даёт формальное обоснование различности получающихся предельных распределений, что на интуитивном уровне представлялось очевидным уже после знакомства с иллюстрациями.

4.5. **Обобщение для банаховых пространств.** Естественно попытаться расширить класс носителей особых марковских цепей. Первый путь для этого состоит в рассмотрении более разнообразных объемлющих пространств, содержащих носитель K.

Заметим, что в доказательствах всех утверждений касательно марковской цепи от носителя K нам нужны были только следующие три свойства:

(1) *Выпуклость* – чтобы брать выпуклую комбинацию текущей точки X_n и точки ξ_{n+1} на отрезке $[X_n, \xi_{n+1}]$ с коэффициентами $0 < \alpha, \beta < 1$ и не выходить при этом за пределы носителя K.

(2) *Ограниченность* – чтобы гарантировать абсолютную сходимость ряда

$$Y = \sum_{k=1}^{\infty} \frac{\xi_k}{2^k}.$$

(3) *Полнота* – чтобы из абсолютной сходимости ряда вывести сходимость его частичных сумм $S_n \xrightarrow[n\to\infty]{} Y$.

Поэтому можно свободно предполагать, что K – произвольное выпуклое ограниченное подмножество некоторого *Банахового пространства* X с нормой $\|\cdot\|\colon X \to \mathbb{R}_+$.

Пример 5 (Гильбертов кирпич)**.** В качестве занимательного примера рассмотрим сепарабельное гильбертово пространство l^2 и в нём подпространство

$$H \stackrel{\text{def}}{=} \Big\{\{x_n\}_{n=1}^{\infty} \in l^2 : 0 \leqslant x_n \leqslant \frac{1}{n}\Big\},$$

именуемое *Гильбертовым кирпичом.*

Хорошо известны следующие свойства H:

- Гильбертов кирпич гомеоморфен Тихоновскому кубу: $H \cong [0,1]^{\infty}$, где гомеоморфизм осуществляется посредством отображения $h(a)_n = n \cdot a_n$;
- H – компактное хаусдорфово пространство;
- $H \subset l^2$ – метрическое (причём топология, наследуемая из l^2, совпадает с Тихоновской топологией);
- H – выпуклое подмножество l^2, $H^{\circ} = \emptyset$, $\operatorname{span} H = l^2$.

Таким образом, H – компактное, выпуклое подмножество сепарабельного Гильбертова пространства l^2, а значит особая марковская цепь на нём корректно определена и обладает всеми доказанными для неё свойствами.

4.6. **Обобщение для неограниченных подмножеств $\mathbb{R}^d$.** Сконцентрируемся теперь на самом носителе K, а не на объемлющем пространстве, его содержащем. От выпуклости множества K отказаться, конечно, нельзя, а вот с ограниченностью дело обстоит интереснее. Давайте рассмотрим выпуклое, *неограниченное* подмножество евклидова пространства $K \subset \mathbb{R}^d$. Как и раньше, будем считать, что $X_0 = 0 \in K$. В таком случае

$$X_n = \frac{\xi_1}{2^n} + \cdots + \frac{\xi_n}{2} \stackrel{\mathcal{D}}{=} \frac{\xi_1}{2} + \cdots + \frac{\xi_n}{2^n} = S_n.$$

Нам бы хотелось найти прозрачные достаточные условия на переходное распределение $\xi_k \sim \mu$, при которых

$$S_n \xrightarrow[n\to\infty]{a.e.} \sum_{k=1}^{\infty} \frac{\xi_k}{2^k},$$

т.е. такие условия, при которых корректно определён предельный случайный вектор Y,

$$Y \stackrel{\text{def}}{=} \sum_{k=1}^{\infty} \frac{\xi_k}{2^k} \stackrel{a.e.}{=} \lim_{n\to\infty} S_n, \tag{5}$$

и, тем самым, определено предельное распределение ν. Ранее сходимость частичных сумм ряда обеспечивалась за счёт ограниченности K. Ныне же нам предстоит найти этому достойную замену.

Легко заметить, что достаточным условием для выполнения (5) является наличие первого абсолютного момента у переходной меры. Действительно, если $\mathbb{E}|\xi| < \infty$, то:

$$\mathbb{E}\Big(\sum_{k=1}^{\infty} \frac{|\xi_k|}{2^k}\Big) = \sum_{k=1}^{\infty} \frac{\mathbb{E}|\xi_k|}{2^k} = \mathbb{E}|\xi| \cdot \sum_{k=1}^{\infty} \frac{1}{2^k} = \mathbb{E}|\xi| < \infty.$$

Обозначим
$$\tilde{Y} \overset{\text{def}}{=} \sum_{k=1}^{\infty} \frac{|\xi_k|}{2^k}.$$

Значит, в терминах $\tilde{Y}$ мы показали:
$$\int\limits_{\Omega} \tilde{Y}\, d\mathbb{P}(\omega) < \infty.$$

Отсюда следует, что для почти всех $\omega \in \Omega$ функция $\tilde{Y} : \Omega \to [0, +\infty]$ меньше бесконечности: $\tilde{Y} < \infty$.

Иначе говоря, почти всюду абсолютно сходится ряд (5), а значит, сходятся и его частичные суммы: $S_n \xrightarrow[n\to\infty]{a.e.} Y$, и $X_n \Rightarrow Y$.

Заметим, однако, что условие $\mathbb{E}|\xi| < \infty$ чересчур сильное. Нам бы хотелось ослабить его, и в этом нам поможет теорема Колмогорова о трёх рядах.

Теорема. *Пусть $\{\zeta_k\}_{k=1}^{\infty}$ – независимые случайные величины. Тогда для сходимости $\sum_{k=1}^{\infty} \zeta_k$ почти наверное необходимо, чтобы для любого $c > 0$, и достаточно, чтобы для некоторого $c > 0$, сходились бы три ряда:*
$$\sum_{k=1}^{\infty} \mathbb{P}(|\zeta_k| \geqslant c) < \infty, \tag{6}$$
$$\sum_{k=1}^{\infty} \mathbb{E}(\mathbb{1}_{\{|\zeta_k| \leqslant c\}} \zeta_k) < \infty, \tag{7}$$
$$\sum_{k=1}^{\infty} \mathbb{D}(\mathbb{1}_{\{|\zeta_k| \leqslant c\}} \zeta_k) < \infty. \tag{8}$$

В нашем случае $\zeta_k \overset{\text{def}}{=} \frac{|\xi_k|}{2^k}$, и мы хотим обеспечить сходимость почти наверное ряда
$$\tilde{Y} = \sum_{k=1}^{\infty} \zeta_k.$$

Поскольку нам интересны лишь достаточные условия, положим для удобства $c = 1$.

Для начала преобразуем отдельно каждое слагаемое из условия (6):
$$\mathbb{P}\Big(\frac{|\xi_k|}{2^k} \geqslant 1\Big) = \mathbb{P}\Big(\frac{|\xi|}{2^k} \geqslant 1\Big) = \mathbb{P}(|\xi| \geqslant 2^k) = \mu\{x \in K : |x| \geqslant 2^k\} = \mu\{x : |x| \geqslant 2^k\}.$$

Значит, условие (6) превращается в
$$\sum_{k=1}^{\infty} \mu\{x : |x| \geqslant 2^k\} < \infty. \tag{I}$$

Аналогично, для условия (7) каждое слагаемое имеет вид:
$$\mathbb{E}\Big(\mathbb{1}_{\{\frac{|\xi_k|}{2^k} \leqslant c\}} \frac{|\xi_k|}{2^k}\Big) = \mathbb{E}\Big(\mathbb{1}_{\{\frac{|\xi|}{2^k} \leqslant c\}} \frac{|\xi|}{2^k}\Big) = \frac{1}{2^k}\mathbb{E}(|\xi| \mathbb{1}_{\{|\xi| \leqslant 2^k\}}).$$

Значит, условие (7) превращается в

$$\sum_{k=1}^{\infty} \frac{1}{2^k}\mathbb{E}(|\xi|\mathbb{1}_{\{|\xi|\leqslant 2^k\}}) < \infty. \tag{II}$$

Проделаем ту же работу для последнего условия (8):

$$\mathbb{D}\Big(\mathbb{1}_{\{\frac{|\xi_k|}{2^k}\leqslant c\}}\frac{|\xi_k|}{2^k}\Big) = \mathbb{D}\Big(\mathbb{1}_{\{\frac{|\xi|}{2^k}\leqslant c\}}\frac{|\xi|}{2^k}\Big) = \frac{1}{2^{2k}}\mathbb{D}(|\xi|\mathbb{1}_{\{|\xi|\leqslant 2^k\}}).$$

Значит, условие (8) превращается в

$$\sum_{k=1}^{\infty} \frac{1}{2^{2k}}\mathbb{D}(|\xi|\mathbb{1}_{\{|\xi|\leqslant 2^k\}}) < \infty. \tag{III}$$

Попробуем тривиальным образом оценить м.о.:

$$\mathbb{E}(|\xi|\mathbb{1}_{\{|\xi|\leqslant 2^k\}}) \leqslant 2^k \cdot \mathbb{P}(|\xi| \leqslant 2^k) = 2^k\mu\{x : |x| \leqslant 2^k\}.$$

Тогда условие (II) преобразуется в

$$\sum_{k=1}^{\infty} \mu\{x : |x| \leqslant 2^k\} < \infty. \tag{II*}$$

Однако в таком виде (I) и (II*) не выполнимы совместно, ибо их сумма даст

$$\sum_{k=1}^{\infty} \mu\{x : |x| \leqslant 2^k \text{ или } |x| \geqslant 2^k\} = \sum_{k=1}^{\infty} 1 = \infty.$$

Видно, что тривиальная оценка плоха, а также видно, чем именно она плоха: фактически, мы лишаем распределение μ “свободы”, надо дать оценке “дышать”. Попробуем реализовать эту идею в следующем рассуждении.

Условие I мы заменим на достаточное для сходимости соответствующего ряда:

$$\mu\{x : |x| \geqslant 2^k\} \leqslant \frac{1}{k^2}. \tag{I}$$

Условие II мы также заменим на достаточное для сходимости соответствующего ряда:

$$\mathbb{E}(|\xi| \cdot \mathbb{1}_{\{|\xi|\leqslant 2^k\}}) \leqslant \frac{2^k}{k^2}. \tag{II}$$

Последнее неравенство имеет вид:

$$\int\limits_{|x|\leqslant 2^k} |x|\,\mu(dx) \leqslant \frac{2^k}{k^2}. \tag{II}$$

Выпишем оценку на возникший интеграл:

$$\int\limits_{|x|\leqslant 2^k} |x|\,\mu(dx) = \int\limits_{|x|\leqslant k} |x|\,\mu(dx) + \int\limits_{k\leqslant|x|\leqslant 2^k} |x|\,\mu(dx) \leqslant \mu\{x : |x| \leqslant k\}\cdot k + \mu\{x : k \leqslant |x| \leqslant 2^k\}\cdot 2^k.$$

Значит, для выполнения (II) достаточно (с точностью до мультипликативной константы):

$$\begin{cases} \mu\{x : |x| \leqslant k\} \leqslant \frac{2^k}{k^3}, \\ \mu\{x : k \leqslant |x| \leqslant 2^k\} \leqslant \frac{1}{k^2}. \end{cases}$$

Заметим, что условие $\mu\{x : |x| \leqslant k\} \leqslant \frac{2^k}{k^3}$ будет выполнено автоматически, начиная с некоторого k, поскольку μ – вероятностная мера, и $\frac{2^k}{k^3} \geqslant 1$ для достаточно больших k.

Итого, условие (II) заменяется на достаточное:

$$\mu\{x : k \leqslant |x| \leqslant 2^k\} \leqslant \frac{1}{k^2}. \tag{II}$$

Оценив дисперсию вторым моментом ($\mathbb{D}X \leqslant \mathbb{E}X^2$), заменим условие (III) на достаточное:

$$\mathbb{E}(|\xi|^2 \cdot \mathbb{1}_{\{|\xi| \leqslant 2^k\}}) \leqslant \frac{2^{2k}}{k^2}. \tag{III}$$

Последнее неравенство имеет вид:

$$\int\limits_{|x| \leqslant 2^k} |x|^2 \, \mu(dx) \leqslant \frac{2^{2k}}{k^2}. \tag{III}$$

Выпишем оценку на последний интеграл:

$$\int\limits_{|x| \leqslant 2^k} |x|^2 \, \mu(dx) = \int\limits_{|x| \leqslant k} |x|^2 \, \mu(dx) + \int\limits_{k \leqslant |x| \leqslant 2^k} |x|^2 \, \mu(dx) \leqslant \mu\{x : |x| \leqslant k\} \cdot k^2 + \mu\{x : k \leqslant |x| \leqslant 2^k\} \cdot 2^{2k}.$$

Значит, для выполнения (III) достаточно (с точностью до мультипликативной константы):

$$\begin{cases} \mu\{x : |x| \leqslant k\} \leqslant \frac{2^{2k}}{k^4}, \\ \mu\{x : k \leqslant |x| \leqslant 2^k\} \leqslant \frac{1}{k^2}. \end{cases}$$

Опять же, условие $\mu\{x : |x| \leqslant k\} \leqslant \frac{2^{2k}}{k^4}$ будет выполнено автоматически, начиная с некоторого k. В свою очередь условие $\mu\{x : k \leqslant |x| \leqslant 2^k\} \leqslant \frac{1}{k^2}$ в точности повторяет полученное выше из выкладки для математического ожидания условие (II).

Таким образом, для сходимости почти наверное ряда

$$\tilde{Y} = \sum_{k=1}^{\infty} \frac{|\xi_k|}{2^k}$$

достаточно выполнения лишь двух условий:

$$\begin{cases} \mu\{x : |x| \geqslant 2^k\} \leqslant \frac{1}{k^2}, \\ \mu\{x : k \leqslant |x| \leqslant 2^k\} \leqslant \frac{1}{k^2}. \end{cases}$$

Их, в свою очередь, можно, не ослабив, заменить одним-единственным общим достаточным условием

$$\mu\{x : |x| \geqslant k\} \leqslant \frac{1}{k^2}. \tag{*}$$

Геометрический смысл полученного условия состоит в том, что хвост распределения должен быть контролируемым образом мал, что вполне согласуется с нашей интуицией.

Сделаем ещё одно наблюдение. Видно, что без ущерба для доказательства можно:

(1) Заменить $\frac{1}{k^2}$ на $\frac{1}{k^{1+\varepsilon}}$ для произвольного фиксированного $\varepsilon > 0$.

(2) Заменить $\mu\{x : k \leqslant |x| \leqslant 2^k\}$ на $\mu\{x : k^p \leqslant |x| \leqslant 2^k\}$ для произвольного фиксированного $p > 0$.

Значит, условие (*) превращается в следующее: найдутся такие $\varepsilon > 0$ и $p > 0$, что:

$$\mu\{x : |x| \geqslant k^p\} \leqslant \frac{1}{k^{1+\varepsilon}}. \tag{*}$$

Ясно, что в этом неравенстве можно избавиться от одного лишнего параметра, не потеряв при этом общности. Действительно, положим $n = k^p$. Тогда $k^{1+\varepsilon} = n^{\frac{1+\varepsilon}{p}}$, где $\varepsilon > 0, p > 0$ – произвольные фиксированные числа. При замене $\delta = \frac{1+\varepsilon}{p}$, условие (*) трансформируется в следующее: найдётся такое $\delta > 0$, что:

$$\mu\{x : |x| \geqslant n\} \leqslant \frac{1}{n^{\delta}}. \tag{*}$$

Таким образом, мы доказали следующее утверждение.

Предложение 7. *Для существования предельного распределения у особой марковской цепи с неограниченным носителем достаточно, чтобы переходная мера μ удовлетворяла условию:*

$$\mu\{x : |x| \geqslant n\} \leqslant \frac{1}{n^{\delta}}. \tag{*}$$

с каким-нибудь фиксированным $\delta > 0$.

5. Заключение

В заключении хочется отметить, что читатель может при желании самостоятельно рассмотреть конкретные примеры особых цепей и промоделировать их поведение на компьютере. Так, для тетраэдра $\Delta^3 \subset \mathbb{R}^3$ с равными нагрузками в вершинах в качестве предельного распределения, подобно треугольнику, получится нечто вроде ковра Серпинского.

Кроме того, интересно было бы дать задаче физическую интерпретацию. Автору она чем-то отдалённо напоминает движение частицы под действием электромагнитного поля, однако дать более точную формулировку представляется затруднительным.

Список литературы

[1] *Convergence of probability measures* / Patrick Billingsley, - 2nd ed., 1999.

[2] Тадевосян А.А. *Несколько примеров Марковских цепей на компактах*: курс.раб...бак. / Санкт-Петербург, Санкт-Петербургский гос. университет, 2019.

[3] Миллер Г.Д. *Исследование предельного распределения Марковской цепи на компакте*: курс.раб...бак. / Санкт-Петербург, Санкт-Петербургский гос. университет, 2022.

Приложение A. Иллюстрации предельных распределений

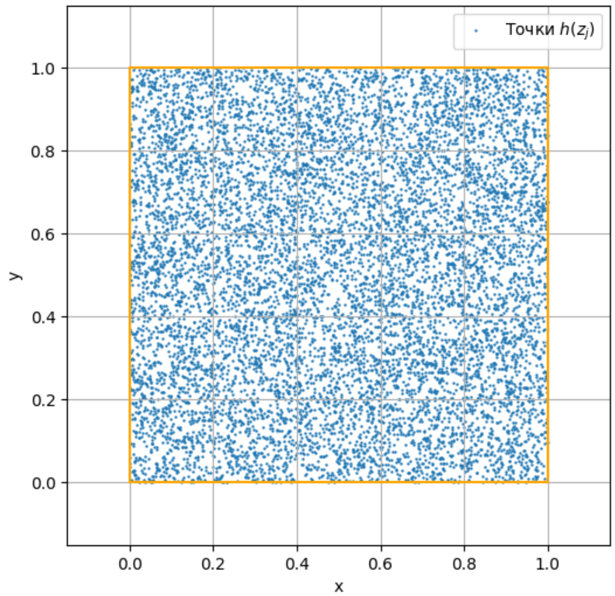

Рис. 7. Предельное распределение ν для квадрата

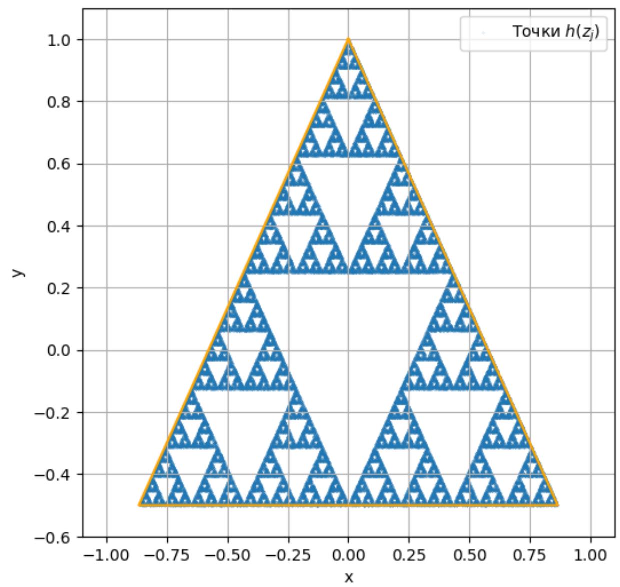

Рис. 8. Предельное распределение ν для треугольника

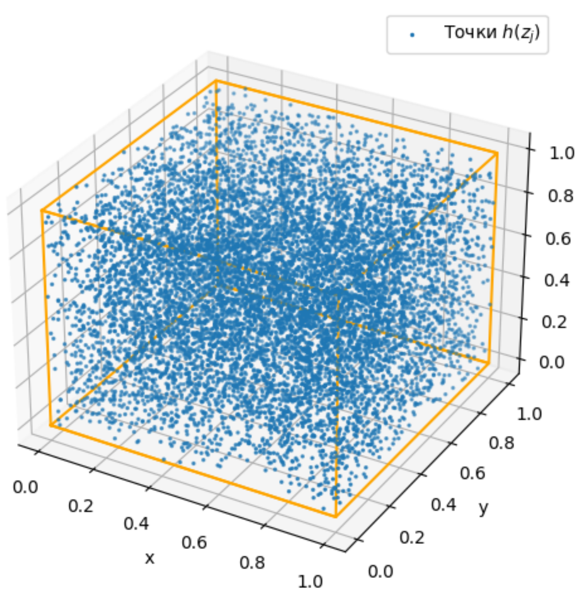

Рис. 9. Предельное распределение ν для куба

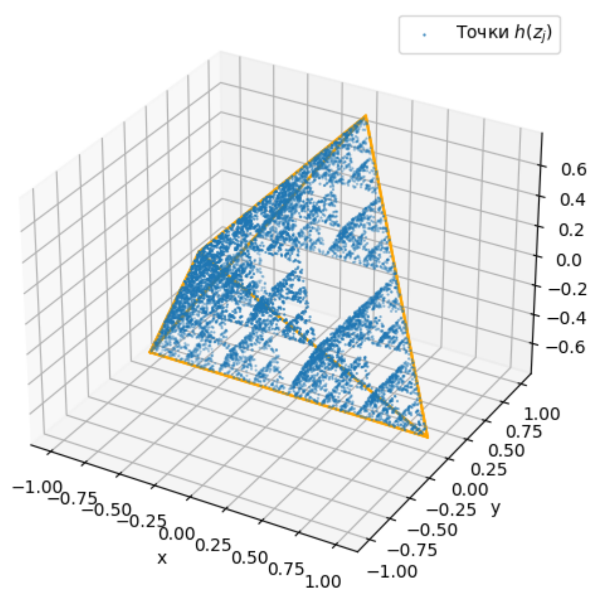

Рис. 10. Предельное распределение ν для тетраэдра