

## Глава 7.3. Избранные темы: Количественные предельные теоремы

Предположим, мы наблюдаем простое случайное блуждание на  $\mathbb{Z}$  и отслеживаем количество раз, которое частица пересекает (направленное) ребро (например, добавляя одну карту в стопку при каждом пересечении). С течением времени на больших временных масштабах мы будем наблюдать формирование случайного профиля высоты стопки карт.

Например, в симметричном случае [мы видели](#), что для этой конкретной цепи Маркова число пересечений ребра  $(n, n + 1)$  (во время экскурсии из 0) является процессом Гальтона-Ватсона (индексированным  $n$ ). Таким образом, пока мы наблюдаем за цепью в течение длительного времени, кривая будет распределена как сумма независимых процессов Гальтона-Ватсона. Более того, в этом примере по положению случайного блуждания в конечный момент времени можно точно определить разность между числом пересечений ребра  $(n, n + 1)$  и числом пересечений  $(n + 1, n)$  (т.е. разность между синими и оранжевыми столбцами на видео). Например, если конечное положение  $X_T = 5$ , то частица пересекла рёбра  $(n, n + 1)$  и  $(n + 1, n)$  при  $n < 0$  или  $n \geq 5$  одинаковое число раз; в то время как при  $0 \leq n \leq 4$  ребро  $(n, n + 1)$  было пересечено на один раз больше, чем  $(n + 1, n)$ .

Более интересным примером является случайное блуждание на  $\mathbb{Z}_N$ .

В этом случае разность между синими и оранжевыми столбцами (см. видео) даст нам число оборотов, совершённых частицей по циклу (петле) — информацию, которую невозможно восстановить, зная только конечное положение.

Если мы посмотрим на цепи Маркова на большом графе, встроенном в многообразия более высокой размерности, то все станет интереснее.

### Предварительные сведения о марковских потоках

В общем случае для цепи Маркова на конечном или счётном пространстве состояний  $S$  мы можем рассмотреть случайную наблюдаемую, которая считает, сколько раз цепь проходит через данное ребро.

### Обозначения для вероятностей на рёбрах

Обозначим через  $E$  пространство рёбер в графе, индуцированном переходными вероятностями  $(p_{x,y})$ , то есть  $E := \{(x, y) : p_{x,y} > 0\}$ . Как обычно, предполагаем, что  $p_{x,y}$  зафиксированы с самого начала, и обозначаем через  $e \in E$  произвольное ребро. Как и вероятности на вершинах, вероятность на пространстве рёбер отождествляется с набором  $(j_e)_{e \in E}$ , где  $j_e \geq 0$  и  $\sum_e j_e = 1$ . Пространство  $\mathcal{P}(E)$  вероятностных мер на  $E$  [снабжено своим собственным расстоянием Леви](#); если  $|E|$  конечно, это обычная топология симплекса  $\{j_e \geq 0, \sum_e j_e = 1\}$ , рассматриваемого как подмножество  $\mathbb{R}^{|E|}$ , или, в более общем случае (если  $E$  счётно), расстояние Леви топологически эквивалентно (после нумерации рёбер  $\{e_0, e_1, \dots\}$ )

$$d(j, j') = \sum_k |j_{e_k} - j'_{e_k}| 2^{-k} \quad (1)$$

То есть последовательность  $j^n$  сходится к  $j$  тогда и только тогда, когда сходятся все значения  $j_e^n$ .

Для  $\mu \in \mathcal{P}(S)$  и  $f: S \rightarrow \mathbb{R}$  мы обычно обозначаем интеграл  $\mu(f) = \sum_x \mu_x f(x)$ . Аналогично для  $j \in \mathcal{P}(E)$  и  $\xi: E \rightarrow \mathbb{R}$  обозначим  $j(\xi) := \sum_{e \in E} j_e \xi(e)$ .

Если  $j$  — вероятность на рёбрах, то обозначим через  $\mu^j \in \mathcal{P}(S)$  вероятность на вершинах, задаваемую формулой

$$\mu_x^j = \sum_{y: (x,y) \in E} j_{x,y} \quad (2)$$

а через  $p^j$  — переходные вероятности

$$p_{x,y}^j = j_{x,y} / \mu_x^j \quad (3)$$

### Потоки как вероятности на рёбрах

Приведённый выше пример побуждает нас рассмотреть меру на рёбрах

$$J^T := \frac{1}{T} \sum_{t=0}^{T-1} \delta_{(X_t, X_{t+1})}$$

$$J_{x,y}^T \equiv \frac{|\{t \leq T-1 : X_t = x, X_{t+1} = y\}|}{T}$$

Другими словами,  $J_e^T$  считает количество раз, которое цепь прошла через (направленное) ребро  $e$  до момента времени  $T$ , нормированное на общее число скачков (то есть  $T$ ).  $J^T$  является вероятностной мерой на пространстве рёбер  $E$  пространства  $S$ . Мы можем рассматривать  $J^T$  как отображение на **пространстве путей Скорохода**

$$J^T: D(S) \rightarrow \mathcal{P}(E)$$

сопоставляющее каждому пути вероятность на рёбрах в соответствии с тем, сколько раз путь  $\mathbf{X}$  пересекал каждое ребро до момента времени  $T$ .

*Примечание.*  $J^T$  — это более богатая наблюдаемая величина, чем мера пребывания  $\pi^T := \sum_{t=0}^{T-1} \delta_{X_t}$ , которая кодирует (нормированное) число посещений цепью каждой вершины. В самом деле

$$\mu_x^{J^T} := \sum_y J_{x,y}^T = \pi_x^T \quad (4)$$

поэтому мы можем восстановить  $\pi^T$  из  $J^T$ . Если рассматривать  $J^T$  как вероятность на  $S \times S$ , это говорит нам о том, что  $\pi^T$  — это просто её проекция на первую компоненту. Из-за сходства с 1-потокками на многообразиях, иногда  $J^T$  называют **потоком** (current).

### Задачи и интересные оценки

**Мы уже знаем**, что  $\pi^T$  сходится к инвариантной мере  $m$  (для неприводимых, положительно возвратных цепей) при  $T \rightarrow \infty$ . Каждый раз, когда цепь посещает вершину  $x$ , существует вероятность  $p_{x,y}$  пересечь ребро  $(x, y)$ ; и при  $T \rightarrow \infty$  (неприводимая, положительно возвратная) цепь Маркова  $\mathbf{X}$  будет посещать  $x$  долю времени  $m_x$ . Таким образом, мы ожидаем, что

$$\lim_T J_{x,y}^T = m_x p_{x,y} \quad (5)$$

что действительно согласуется с Уравнение 4.

В этой главе мы хотим получить более количественный результат о сходимости для  $J^T$ , который, в свою очередь, повлечёт за собой Уравнение 5, а через Уравнение 4 — и аналогичные оценки для  $\pi^T$ . Это глубоко мотивированная задача; рассмотрим пример, который может дать некоторое представление<sup>1</sup>.

**Пример 0.1.** В протоколе управления передачей (TCP) для интернет-коммуникации сервер  $x$  выбирает следующий напрямую подключенный сервер  $y$  для отправки полученного пакета данных в зависимости от конечного пункта назначения пакета (это определяется по маске подсети, так что пакет действительно проходит через множество серверов и не следует по оптимальному геодезическому пути). Справедливо предположить, что пакеты совершают цепь Маркова с некоторыми переходными вероятностями  $p_{x,y}$ . Пара  $(x, y)$  напрямую соединённых серверов будет иметь максимальную пропускную способность  $C_{x,y}$ , определённую для малого промежутка времени, и передача не удастся, если эта способность будет превышена. Если мы наблюдаем за трафиком в течение некоторого периода времени, мы можем очень точно оценить  $p_{x,y}$  и, возможно, захотим узнать, как назначить пропускные способности  $C_{x,y}$ , чтобы минимизировать вероятность сбоев. То есть нам нужно знать для произвольного распределения  $C_{x,y}$

$$\mathbb{P}(J_{x,y}^T \leq C_{x,y}, \forall x, y \in S) \approx ?$$

### Основные результаты: Формулировки

**Определение 0.1.** Вспомним Уравнение 2, Уравнение 3 и определим функцию  $I: \mathcal{P}(E) \rightarrow [0, \infty]$

$$I(j) := \begin{cases} \sum_{x,y} \mu_x^j p_{x,y}^j \log(p_{x,y}^j / p_{x,y}) & \text{если } \sum_y j_{x,y} = \sum_y j_{y,x} \\ +\infty & \text{иначе} \end{cases} \quad (6)$$

**Определение 0.2** (Функционал Донскера-Варадхана). Определим функцию<sup>2</sup>  $I_{DV}: \mathcal{P}(E) \rightarrow [0, \infty]$

$$I_{DV}(\mu) := \inf\{I(j), j \in \mathcal{P}(E) : \mu^j = \mu\} \quad (7)$$

В Упражнение 0.3, Упражнение 0.4 и Упражнение 0.5 приведены более явное выражение и некоторые свойства  $I_{DV}$ .

*Примечание.* Принимая во внимание определение относительной энтропии  $H(\cdot|\cdot)$ , функционал  $I(\cdot)$  можно также записать как

$$I(j) := \begin{cases} \sum_x \mu_x^j H(p_{x,\cdot}^j | p_{x,\cdot}) & \text{если } \sum_y j_{x,y} = \sum_y j_{y,x} \\ +\infty & \text{иначе} \end{cases} \quad (8)$$

Наш основной результат заключается в следующем:

**Теорема 0.1.** Рассмотрим неприводимую цепь Маркова на конечном пространстве состояний. Тогда для любого замкнутого множества  $\mathcal{C} \subset \mathcal{P}(E)$  выполняется

$$\mathbb{P}(J^T \in \mathcal{C}) \leq e^{-T \inf_{j \in \mathcal{C}} I(j) + o(T)} \quad (9)$$

Это, в свою очередь, влечёт экспоненциальную версию Уравнение 5 (по вероятности) и, через Уравнение 4, аналогичный результат для  $\pi^T$ .

<sup>1</sup>Флуктуации математических объектов такого типа (например,  $J^T$ ) имеют довольно глубокий смысл, предоставляя универсальное динамическое определение понятия энтропии в сложных системах.

<sup>2</sup>Это называется *функционалом Донскера-Варадхана*, в честь известной серии статей математиков из NYU М.Д.Донскера и С.Р.С.Варадхана в 70-х годах. Статьи в основном посвящены марковским процессам с непрерывным временем (поэтому могут быть довольно технически сложными), но адаптация для дискретного времени проста.

**Следствие 0.1.** Определим типичный поток как  $\bar{j}_{x,y} := m_x p_{x,y}$ . Тогда, при тех же условиях, что и в Теореме 0.1,  $I(j) = 0$  тогда и только тогда, когда  $j = \bar{j}$ . В частности, для каждого  $\varepsilon > 0$  существует константа  $c_\varepsilon > 0$ , такая что

$$\mathbb{P}\left(\sum_{e \in E} |J_e^T - \bar{j}_e| \geq \varepsilon\right) \leq e^{-T(c_\varepsilon + o(1))}$$

**Следствие 0.2.** При тех же предположениях, что и в Теореме 0.1, для любого замкнутого множества  $\mathcal{C} \subset \mathcal{P}(S)$  выполняется

$$\mathbb{P}(\pi^T \in \mathcal{C}) \leq e^{-T \inf_{\mu \in \mathcal{C}} I_{DV}(\mu) + o(T)} \quad (10)$$

Можно доказать (что выходит за рамки данной заметки), что  $I$  оптимальна в следующем смысле: для каждого измеримого  $\mathcal{A} \subset \mathcal{P}(E)$

$$e^{-T \inf_{j \in \text{Interior}(\mathcal{A})} I(j) + o(T)} \leq \mathbb{P}(J^T \in \mathcal{A}) \leq e^{-T \inf_{j \in \text{Closure}(\mathcal{A})} I(j) + o(T)}$$

(и аналогичное утверждение справедливо для  $I_{DV}$ ), где Interior — внутренность, а Closure — замыкание множества.

*Примечание.* Условие  $\sum_y J_{x,y} = \sum_y J_{y,x}$  означает, что мера  $\mu^j$  инвариантна для индуцированных переходных вероятностей  $p^j$ . Таким образом, Теорема 0.1 можно интерпретировать следующим образом: вероятность того, что наблюдаемый поток  $J^T$  флуктуирует вблизи данного  $j$ , убывает экспоненциально как  $\exp(-TI(j))$ . Функция уклонений  $I(j)$  представляет собой относительную энтропию модифицированной динамики  $p^j$  по отношению к исходной динамике  $p$ , усреднённую по  $p^j$ -инвариантной мере  $\mu^j$ .

Аналогично, Следствие 0.2 количественно характеризует редкость наблюдения эмпирической меры  $\pi^T$ , близкой к  $\mu$  (где  $\mu$  не обязательно является инвариантной мерой). Стоимость задаётся величиной  $I_{DV}(\mu)$ , которая количественно выражает «отсутствие  $P$ -инвариантности» у  $\mu$ , см., например, оценку Упражнение 0.4.

## Основные результаты: Доказательства

Этот раздел посвящён доказательству Теорема 0.1, Следствие 0.1, Следствие 0.2.

**Определение 0.3.** Функция  $F: \mathcal{P}(E) \rightarrow [0, \infty]$  называется **полунепрерывной снизу** (или **пнс** для краткости), если для каждого  $c \geq 0$  множество  $\{j \in \mathcal{P}(E) : F(j) \leq c\}$  замкнуто.

Следующая лемма утверждает, что если мы докажем неравенство типа Уравнение 9, заменив  $I$  некоторой функцией  $I_\alpha$ , зависящей от параметра, то мы можем получить то же неравенство, оптимизируя по  $\alpha$ . Это не просто следствие взятия  $\inf$  по  $\alpha$  с обеих сторон неравенства, так как нам нужно поменять местами  $\sup_\alpha$  (появляющийся из-за знака минус в экспоненте) и  $\inf_j$ .

**Лемма 0.1.** Пусть  $\mathcal{U}$  — произвольное множество индексов, и для  $\alpha \in \mathcal{U}$  пусть  $I_\alpha: \mathcal{P}(E) \rightarrow [0, \infty]$  — пнс функция. Предположим, что для каждого  $\alpha \in \mathcal{U}$  выполняется Уравнение 9 с заменой  $I$  на  $I_\alpha$ , то есть

$$\overline{\lim}_T \frac{1}{T} \log \mathbb{P}(J^T \in \mathcal{C}) \leq - \inf_{j \in \mathcal{C}} I_\alpha(j), \quad \text{для всех замкнутых } \mathcal{C} \quad (11)$$

Тогда Уравнение 9 выполняется с заменой  $I$  на

$$\hat{I}(j) := \sup_{\alpha \in \mathcal{U}} I_\alpha(j)$$

*Доказательство.* Согласно общим принципам теории больших уклонений, существует оптимальная полунепрерывная снизу функция  $\bar{I}$ , такая что Уравнение 11 выполняется (с  $\bar{I}$  вместо  $I_\alpha$ ) на компактных множествах. Поскольку мы предположили, что  $E$  конечно,  $\mathcal{P}(E)$  компактно, и, следовательно, каждое замкнутое множество компактно. Поэтому Уравнение 11 читается как

$$\bar{I} \geq I_\alpha \quad \text{для всех } \alpha \in \mathcal{U}$$

откуда  $\bar{I} \geq \sup_{\alpha \in \mathcal{U}} I_\alpha$ . □

**Лемма 0.2.** Для  $\xi \in C(E)$  определим  $\phi_\xi \in C(S)$  как

$$\phi_\xi(x) := \log \left( \sum_y e^{\xi(x,y)} p_{x,y} \right)$$

Тогда

$$\mathbb{E} \left[ \exp \left( \sum_{t=0}^{T-1} \xi(X_t, X_{t+1}) - \phi_\xi(X_t) \right) \right] = 1 \quad (12)$$

*Доказательство.* Определим  $q_{x,y} := e^{\xi(x,y) - \phi_\xi(x)} p_{x,y}$ . Тогда  $q_{x,y} \geq 0$  и  $\sum_y q_{x,y} = 1$ . Следовательно,  $(q_{x,y})$  являются марковскими переходными вероятностями, и мы можем построить меру  $\mathbb{Q}$  на  $D(S)$ . Имеем

$$\mathbb{Q}(X_0 = x_0, \dots, X_T = x_T) = \exp \left( \sum_{t=0}^{T-1} \xi(x_t, x_{t+1}) - \phi_\xi(x_t) \right) \mathbb{P}(X_0 = x_0, \dots, X_T = x_T)$$

Таким образом, Уравнение 12 есть просто математическое ожидание постоянной функции  $\mathbf{1}$  по мере  $\mathbb{Q}$ . □

**Лемма 0.3.** Для  $\xi, \phi_\xi$  как в Лемма 0.2, и  $\mathcal{A} \subset \mathcal{P}(E)$

$$\mathbb{P}(J_T \in \mathcal{A}) \leq \exp(-\inf_{j \in \mathcal{A}} j(\xi) - \mu^j(\phi_\xi))$$

*Доказательство.* Показатель экспоненты в Уравнение 12 можно записать как  $J^T(\xi) - \pi^T(\phi_\xi)$ . Вспоминая Уравнение 4 и просто умножая и деля на один и тот же множитель, получаем

$$\begin{aligned} \mathbb{P}(J^T \in \mathcal{A}) &= \mathbb{E} \left[ e^{-T(j(\xi) - \mu^j(\phi_\xi))} e^{\sum_{t=0}^{T-1} \xi(X_t, X_{t+1}) - \phi_\xi(X_t)} \mathbf{1}_{\mathcal{A}}(J^T) \right] \\ &\leq \sup_{j \in \mathcal{A}} e^{-T(j(\xi) - \mu^j(\phi_\xi))} \mathbb{E} \left[ e^{\sum_{t=0}^{T-1} \xi(X_t, X_{t+1}) - \phi_\xi(X_t)} \mathbf{1}_{\mathcal{A}}(J^T) \right] \\ &\leq \exp(-T \inf_{j \in \mathcal{A}} (j(\xi) - \mu^j(\phi_\xi))) \mathbb{E} \left[ e^{\sum_{t=0}^{T-1} \xi(X_t, X_{t+1}) - \phi_\xi(X_t)} \right] \end{aligned}$$

что совпадает с утверждением в силу Лемма 0.2. □

Для функции  $f: S \rightarrow \mathbb{R}$  пусть  $df(x, y) := f(y) - f(x)$ . Определим

$$\mathcal{P}_*(E) := \{j \in \mathcal{P}(E) : j(df) = 0, \text{ для всех } f \in C(S)\} = \left\{ j \in \mathcal{P}(E) : \sum_y j_{x,y} = \sum_z j_{z,x}, \text{ для всех } x \in S \right\}$$

**Лемма 0.4.** Для любого компактного множества  $\mathcal{K} \subset \mathcal{P}(E)$  такого, что  $\mathcal{K} \cap \mathcal{P}_*(E) = \emptyset$ , выполняется

$$\mathbb{P}(J^T \in \mathcal{K}) = 0, \quad \text{для всех } T \text{ достаточно больших.}$$

*Доказательство.* Для данного  $T \geq 1$  пусть  $y_0 = X_T, \dots, y_N = X_0$  — путь в  $D(S)$ , соединяющий  $X_T$  с  $X_0$  за минимальное число шагов (следовательно, не более  $|E|$  шагов). Определим поток

$$\tilde{J}^T = \frac{1}{T+N} \sum_{t=0}^{T+N-1} \delta_{(X_t, X_{t+1})}$$

Телескопируя сумму, получаем

$$J^T(df) = \frac{1}{T} \sum_{t=0}^{T-1} f(X_{t+1}) - f(X_t) = (f(X_T) - f(X_0))/T$$

$$\tilde{J}^T(df) = 0$$

С другой стороны, согласно Уравнение 1,  $d(J^T, \tilde{J}^T) \leq |E|/T$ . В частности,  $J^T$  находится на расстоянии не более  $|E|/T$  от  $\mathcal{P}_*(E)$  с вероятностью 1. Поскольку  $\mathcal{K}$  компактно, а  $\mathcal{P}_*(E)$  замкнуто (на самом деле компактно, так как мы предполагаем конечность  $S$ ), если эти два множества не пересекаются, они находятся на строго положительном расстоянии. Следовательно,  $J^T \notin \mathcal{K}$  с вероятностью 1 при достаточно малом  $|E|/T$ .  $\square$

*Теорема 0.1.* Лемма 0.3 можно интерпретировать так: Уравнение 9 выполняется при замене  $I$  на

$$I_\xi(j) := j(\xi) - \mu^j(\phi_\xi)$$

Лемма 0.4 можно интерпретировать так: Уравнение 9 выполняется при замене  $I$  на

$$I_*(j) := \begin{cases} 0 & \text{если } j \in \mathcal{P}_*(E). \\ +\infty & \text{иначе.} \end{cases}$$

$I_\xi$  непрерывна (в частности, пнс),  $I_*$  пнс на  $\mathcal{P}(E)$ . Следовательно, по Лемма 0.1 неравенство Уравнение 9 выполняется с функцией уклонений

$$\hat{I}(j) = \begin{cases} \sup_\xi I_\xi(j) & \text{если } j \in \mathcal{P}_*(E). \\ +\infty & \text{иначе.} \end{cases}$$

Теперь возьмём  $\xi(x, y) = \log(j_{x,y}/p_{x,y})$  (определение  $\xi(x, y)$  при  $j_{x,y} = 0$  несущественно). Получим

$$\begin{aligned} I_\xi(j) &= \sum_{x,y} j_{x,y} \log(j_{x,y}/p_{x,y}) - \sum_x \mu_x^j \log \left( \sum_z j_{x,z} \right) \\ &= \sum_{x,y} \mu_x^j p_{x,y}^j \log(\mu_x^j p_{x,y}^j / p_{x,y}) - \sum_x \mu_x^j \log \mu_x^j \\ &= \sum_{x,y} \mu_x^j p_{x,y}^j \log(p_{x,y}^j / p_{x,y}) = I(j) \end{aligned}$$

Следовательно, для  $j \in \mathcal{P}_*(E)$  имеем  $\hat{I}(j) = \sup_\xi I_\xi(j) \geq I(j)$ . Поскольку вне  $\mathcal{P}_*(E)$  неравенство выполняется тривиально ( $+\infty = +\infty$ ), то Уравнение 9 справедливо.  $\square$

*Следствие 0.1.* Для двух вероятностей  $\mu, \nu \in \mathcal{P}(S)$  функция  $\nu \mapsto \sum_y \nu_y \log \nu_y / \mu_y$  достигает минимума при  $\nu = \mu$  (где она обращается в ноль). Поэтому в Уравнение 6 каждое слагаемое в сумме по  $x$  неотрицательно.

Если  $I(j) = 0$ , то должно выполняться  $p_{x,y}^j = p_{x,y}$  для всех  $y$  и всех  $x$ , таких что  $\mu_x^j > 0$  ( $p^j$  не определено, если  $\mu_x^j = 0$ , поэтому мы можем произвольно положить  $p_{x,y}^j = p_{x,y}$  везде). Но тогда условие  $j \in \mathcal{P}_*(E)$  означает, что  $\mu^j$  инвариантна для  $p_{x,y}^j \equiv p_{x,y}$ , и так как (в силу неприводимости)  $m$  является единственной инвариантной вероятностью для  $(p_{x,y})$ , то  $\mu^j = m$ . Таким образом,  $J_{x,y} = \mu_x^j p_{x,y}^j = m_x p_{x,y} = \bar{J}_{x,y}$ .

Теперь множество  $\mathcal{K}_\varepsilon := \{j : \sum_e |J_e - \bar{J}_e| \geq \varepsilon\}$  является компактным множеством, находящимся на положительном расстоянии от  $\bar{j}$ . Следовательно,  $c_\varepsilon := \inf_{j \in \mathcal{K}_\varepsilon} I(j) > 0$ , поскольку  $I$  полунепрерывна снизу.  $\square$

*Доказательство.* Для замкнутого  $\mathcal{C} \subset \mathcal{P}(S)$  определим  $\mathcal{C}' := \{j : \mu^j \in \mathcal{C}\}$  (это просто прообраз  $\mathcal{C}$  относительно отображения  $j \mapsto \mu^j$ ). Множество  $\mathcal{C}'$  замкнуто в  $\mathcal{P}(E)$ , и поэтому, **так как**  $\pi^T = \mu^{J^T}$ ,

$$\mathbb{P}(\pi^T \in \mathcal{C}) = \mathbb{P}(J^T \in \mathcal{C}') \leq e^{-T \inf_{j \in \mathcal{C}'} I(j) + o(T)} = e^{-T \inf_{\mu \in \mathcal{C}} I_{DV}(\mu) + o(T)} \quad (13)$$

$\square$

## Дополнительные Задачи

**Упражнение 0.1.** Найдите явную оценку (аналогичную Уравнение 9) для  $h^T$  — числа оборотов (намоток) ближайшего случайного блуждания на  $\mathbb{Z}_N$ . Подсказка: выразите число оборотов  $h^T$  через  $J^T$ .

**Упражнение 0.2.** Пусть  $F: \mathcal{P}(E) \rightarrow \mathbb{R}$  непрерывна (и ограничена, что следует из конечности  $E$ ). Докажите, что

$$\mathbb{E} \left[ e^{TF(J^T)} \right] \leq e^{T \sup_j (F(j) - I(j)) + o(T)}$$

**Упражнение 0.3.** Докажите формулу

$$I_{DV}(\mu) = \sup_{u \in C(S), u > 0} \sum_x \mu_x \log \left( \frac{u(x)}{(Pu)(x)} \right) \quad (14)$$

### Решение

В принципе, эту задачу можно решить с помощью множителей Лагранжа и громоздких, но элементарных вычислений. Однако несколько более абстрактное применение техники позволяет найти более быстрый подход.

Инфимум в определении Уравнение 7 можно брать по  $j \in \mathcal{P}_*(E)$ , так как иначе  $I(j) = +\infty$ . Поэтому можно записать  $J_{x,y} = \mu_x q_{x,y}$ , где марковские переходные вероятности  $Q = (q_{x,y})_{x,y}$  удовлетворяют  $\mu Q = \mu$  (из  $j \in \mathcal{P}_*(E)$ ).

Мы можем использовать произвольную функцию  $f \in C(S)$  в качестве множителя Лагранжа, чтобы наложить ограничение  $\mu Q = \mu$ . Действительно, для  $Q = (q_{x,y})_{x,y}$

$$I_{DV}(\mu) = \inf_Q \sup_{f \in C(S)} \sum_{x,y} \mu_x q_{x,y} \log \left( \frac{q_{x,y}}{p_{x,y}} \right) + \mu((I - Q)f)$$

где теперь инфимум берётся по произвольным переходным вероятностям  $Q = (q_{x,y})_{x,y}$ , так как супремум по  $f$  равен  $+\infty$ , если не выполнено условие  $\mu = \mu Q$ .

Ключевой момент заключается в следующем. Выражение в правой части, для которого мы берём  $\inf_Q \sup_f$ , является:

- выпуклым (следовательно, непрерывным) по  $Q$ , для  $Q$  из выпуклого компактного множества

$$\sum_y q_{x,y} = 1, q_{x,y} \geq 0.$$

- аффинным (следовательно, непрерывным) по  $f$ , для  $f \in C(S)$ .

Поэтому мы можем использовать принцип минимакса Сиона и поменять порядок  $\inf$  и  $\sup$ , чтобы получить

$$I_{DV}(\mu) = \sup_{f \in C(S)} \inf_Q \sum_{x,y} \mu_x q_{x,y} \log \left( \frac{q_{x,y}}{p_{x,y}} \right) + \mu((I - Q)f) = \sup_{f \in C(S)} \mu(f) + \inf_Q \sum_{x,y} \mu_x q_{x,y} \log \left( \frac{q_{x,y}}{p_{x,y} e^{f(y)}} \right)$$

Теперь инфимум по  $q_{x,y}$  вычисляется легко (так как нам не нужно накладывать «сложное» условие  $\mu = Q\mu$ , только  $\sum_y q_{x,y} = 1$ ), и мы получаем

$$I_{DV}(\mu) = \sup_{f \in C(S)} \sum_x \mu_x f(x) - \mu_x \log \left( \sum_y p_{x,y} e^{f(y)} \right)$$

что даёт искомое выражение, если положить  $u = e^f$ .

Это доказательство в конечном итоге наводит на мысль, что если бы мы повторили доказательство Теорема 0.1 для  $\pi^T$  вместо  $J^T$ , было бы достаточно оптимизировать по  $\xi(x, y) = f(y)$  (зависящей только от  $y$ , а не от произвольной функции  $x$  и  $y$ ), что согласуется с тем фактом, что оценка для  $\pi^T$  является более слабым результатом и, по сути, может быть выведена из экспоненциальной оценки для  $J^T$ .

**Упражнение 0.4.** Докажите явные оценки

$$\mathcal{E}(\sqrt{\varrho}, \sqrt{\varrho}) \leq I_{DV}(\mu) \leq \sum_{x,y} \mu_x \mu_y \log \left( \frac{\mu_y}{p_{x,y}} \right) \quad (15)$$

где  $\mathcal{E}$  — форма Дирихле, а  $\varrho_x := \mu_x / m_x$  — плотность  $\mu$  относительно инвариантной меры  $m$ .

#### Решение

Следуя началу Упражнение 0.3, мы можем записать  $I_{DV}$  как инфимум по марковским переходным вероятностям  $Q = (q_{x,y})$ , оставляющим  $\mu$  инвариантной:  $\mu = \mu Q$ . Выбор  $q_{x,y} = \mu_y$  оставляет  $\mu$  инвариантной, что легко проверяется, и даёт указанную верхнюю оценку для  $I_{DV}$ .

Что касается нижней оценки, в Уравнение 14 возьмём  $u = \sqrt{\varrho}$  (случай, когда  $\varrho$  обращается в ноль в некоторой точке, легко разбирается, так как эти точки несущественны, и можно положить  $u(x) = 1$ , если  $\varrho(x) = 0$ ). Выражение в Уравнение 14 для этого конкретного выбора  $u$  не больше супремума по  $u$ , следовательно

$$I_{DV}(\mu) \geq \sum_x \mu_x \log \left( \frac{\sqrt{\varrho}(x)}{(P\sqrt{\varrho})(x)} \right) = - \sum_x m_x \log \left( 1 - \frac{((I - P)\sqrt{\varrho})(x)}{\sqrt{\varrho}(x)} \right)$$

Применим  $-\log(1 - z) \geq z$  к каждому слагаемому суммы и получим

$$I_{DV}(\mu) \geq \sum_x m_x \varrho(x) \frac{((I - P)\sqrt{\varrho})(x)}{\sqrt{\varrho}(x)}$$

что упрощается до требуемого неравенства.

**Упражнение 0.5.** Докажите, что  $I_{DV}(\mu) = 0$  тогда и только тогда, когда  $\mu = m$  является инвариантной мерой (вспомните, что мы предполагали неприводимость цепи). Выведите, как и в Следствие 0.1, что для некоторого  $c_\varepsilon > 0$

$$\mathbb{P}\left(\sum_{x \in S} |\pi_x^T - m_x| \geq \varepsilon\right) \leq e^{-T(c_\varepsilon + o(1))}$$

где  $o(1)$  стремится к нулю при  $T \rightarrow \infty$ .

### Решение

Поскольку  $m = \mu^{\bar{j}}$  и  $I(\bar{j}) = 0$ ,

$$0 \leq I_{DV}(\mu) \leq I(\bar{j}) = 0$$

Обратно, если  $I_{DV}(\mu) = 0$ , то  $\mathcal{E}(\sqrt{\varrho}, \sqrt{\varrho}) = 0$ . Так как мы предполагали неприводимость цепи, это означает, что  $\varrho$  постоянна, [согласно пункту с. Предложения о формах Дирихле](#) при  $f = g$ . Это эквивалентно  $\mu = m$ , так как они обе являются вероятностями.

Экспоненциальная оценка следует из того же рассуждения, что и в Следствие 0.1, так как используется только то, что  $I(\cdot)$  имеет единственный корень.