1995 г. март — апрель *т. 50, вып. 2(302)*

УСПЕХИ МАТЕМАТИЧЕСКИХ НАУК

УДК 519.217

КЛАССИФИКАЦИЯ МАРКОВСКИХ ЦЕПЕЙ, ОПИСЫВАЮЩИХ ЭВОЛЮЦИЮ СЛУЧАЙНЫХ СТРУН

А. С. Гайрат, В. А. Малышев, М. В. Меньшиков, К. Д. Пелих

СОДЕРЖАНИЕ

§1. Введение

Напомним основные определения из [1]. Рассмотрим однородную цепь Маркова $\mathscr{L}$ с дискретным временем и счетным числом состояний

$$\mathscr{A}=\bigcup_{n=0}^{\infty}\{1,\dots,r\}^n.$$

Другими словами, состояния марковской цепи – это упорядоченные r-символьные последовательности (струны) $\alpha=x_n\dots x_1, x_i\in\{1,\dots,r\}$, произвольной длины $n=n(\alpha)=|\alpha|$. Через $\varnothing$ мы обозначим пустую струну (длины 0). Для любых двух струн $\alpha=x_n\dots x_1, \beta=y_m\dots y_1$ мы обозначим их объединение (длины $m+n$) через

$$\alpha\beta=x_n\dots x_1 y_m\dots y_1.$$

Мы полагаем, что вероятности перехода $p_{\alpha,\beta}$ за один шаг $\alpha\to\beta$ удовлетворяют следующим условиям.

Условие (H) (пространственная однородность, ограниченность скачков). Пусть $|\alpha|\geqslant d$ и $\alpha=\gamma\rho$, $\beta=\theta\rho$ для некоторых γ, θ, ρ с длинами $|\gamma|=d, |\theta|\leqslant 2d$. (Тогда

$p_{\alpha,\beta}$ не зависит от ρ, а только от γ и θ. Поэтому мы можем обозначить эту переходную вероятность через $q(\gamma,\theta)$: мы удаляем γ с левого конца струны и ставим вместо нее θ.)

В частности, из условия (H) следует, что $p_{\alpha,\beta} = 0$, если $\big||\alpha| - |\beta|\big| > d$.

Условие (ND) (невырожденности): для любых струн γ, θ: $|\gamma| = d$, $|\theta| \leqslant 2d$, $q(\gamma,\theta) > 0$. Мы предполагаем также положительность всех переходных вероятностей $p_{\alpha,\beta}$ для $|\alpha| < d$, $|\beta| \leqslant |\alpha| + d$.

Из введенных условий следует, что рассматриваемая цепь Маркова неприводима и непериодична. Заметим, что условие (ND) можно было бы ослабить, но при этом существенно усложняются формулировки и доказательства основных результатов.

Цель настоящей работы: дать необходимые и достаточные условия эргодичности, нулевой возвратности и транзиентности для этой марковской цепи. Эти условия даны в терминах максимального собственного числа некоторой конечной матрицы A. Для того чтобы объяснить основные идеи, мы сначала рассмотрим случай $d = 1$ (§2). В этом случае элементы матрицы линейно выражаются через $q(\gamma,\theta)$. В случае $d > 1$ элементы матрицы A даны в терминах положительного решения некоторого фундаментального нелинейного алгебраического уравнения. Мы изучаем это уравнение в §3. Классификация для случая $d > 1$ дана в §§4, 5.

Ранее было опубликовано много работ, посвященных близким задачам. В них рассматривалась свободная группа G, порожденная l генераторами $a_1,\dots,a_l$ ($a_{-i} = a_i^{-1}$). В работах [6]–[9] было рассмотрено произведение независимых случайных величин со значениями в G. Однако в этой ситуации эргодичности быть не может, и рассматривался лишь транзиентный случай.

§2. Линейное уравнение для средних времен: случай $d = 1$

В этой части мы полагаем, что $d = 1$. Пусть $\alpha(t)$, $t = 0,1,2,\dots$ – состояние процесса в момент времени t. Положим

$$\tau_i = \min\{t : |\alpha(t)| < |i\rho| \mid \alpha(0) = i\rho\}, \qquad i \in \{1,\dots,r\}.$$

Распределение τ_i не зависит от ρ.

Если процесс эргодичен, то $E_i = \mathsf{E}\tau_i < \infty$ и $\{E_i\}_{i=1,\dots,r}$ удовлетворяют следующей системе:

$$\begin{aligned}(1)\qquad E_i &= q(i,\varnothing) + \sum_j q(i,j)(1+E_j) + \sum_{j,k} q(i,jk)(1+E_j+E_k)\\ &= 1 + \sum_j q(i,j)E_j + \sum_{j,k} q(i,jk)(E_j+E_k).\end{aligned}$$

Мы можем переписать (1) в матричном виде:

$$E = \bar{1} + AE, \tag{2}$$

где E – вектор с компонентами $(E_i : i = 1,\dots,r)$, $\bar{1}$ – вектор с компонентами 1, A – положительная матрица. По теореме Перрона–Фробениуса A имеет максимальное положительное собственное число λ. Уравнение (2) имеет положительное конечное решение тогда и только тогда, когда $\lambda < 1$. Таким образом, если процесс эргодичен, то $\lambda < 1$. Однако, верен и более сильный результат.

ТЕОРЕМА 2.1. *Процесс $\alpha(t)$ является*

(i) *эргодическим тогда и только тогда, когда $\lambda < 1$,*
(ii) *нулевым возвратным тогда и только тогда, когда $\lambda = 1$,*
(iii) *транзиентным тогда и только тогда, когда $\lambda > 1$.*

ДОКАЗАТЕЛЬСТВО. Мы будем доказывать теорему методом функций Ляпунова (см. [3], [4]).

Пусть

$$\tau(\alpha) = \min\{t : |\alpha(t)| = 0 \mid \alpha(0) = \alpha\}$$

– первый момент достижения пустой струны $\varnothing$, если начальное состояние $\alpha \neq \varnothing$. Тогда, если $\alpha = x_n \dots x_1$,

$$T(\alpha) = \mathsf{E}\tau(\alpha) = \sum_{k=1}^{n} E_{x_k}.$$

На множестве конечных струн рассмотрим следующую неотрицательную функцию $f(\alpha) = f(x_n \dots x_1) = \sum_{k=1}^{n} e_{x_k}$, $f(\varnothing) = 0$, где $\{e_i\}$ – компоненты собственного вектора $e = (e_1, \dots, e_r)$, соответствующего максимальному собственному числу λ матрицы A, т.е.

$$\lambda e_i = \sum_j q(i,j)e_j + \sum_{j,k} q(i,jk)(e_j + e_k). \tag{3}$$

По теореме Перрона–Фробениуса $e_i > 0$ для всех i. Нам понадобится следующая

ЛЕММА 2.1.

(i) *Если $\lambda > 1$, тогда существует $\varepsilon > 0$ такое что*

$$\mathsf{E}\big[f(\alpha(t+1)) - f(\alpha(t)) \mid \alpha(t) = \alpha\big] > \varepsilon$$

для всех α: $|\alpha| \geqslant 1$.

(ii) *Если $\lambda = 1$, то*

$$\mathsf{E}\big[f(\alpha(t+1)) - f(\alpha(t)) \mid \alpha(t) = \alpha\big] = 0$$

для всех α: $|\alpha| \geqslant 1$.

(iii) *Если $\lambda < 1$, то существует $\varepsilon > 0$ такое, что*

$$\mathsf{E}\big[f(\alpha(t+1)) - f(\alpha(t)) \mid \alpha(t) = \alpha\big] < -\varepsilon$$

для всех α: $|\alpha| \geqslant 1$.

Доказательство. Это простая выкладка для $|\alpha| \geqslant 1$. Пусть $\alpha = i\rho$, тогда

$$\begin{aligned}
&\mathsf{E}\big[f(\alpha(t+1)) - f(\alpha(t)) \mid \alpha(t) = \alpha\big] \\
&\qquad = -q(i,\varnothing)e_i + \sum_j q(i,j)(e_j - e_i) + \sum_{j,k} q(i,jk)(e_j + e_k - e_i) \\
&\qquad = -e_i + \sum_j q(i,j)e_j + \sum_{j,k} q(i,jk)(e_j + e_k) = (\lambda - 1)e_i.
\end{aligned}$$

Для доказательства теоремы заметим, что:

a) для всех α

$$f(\alpha) \geqslant 0, \quad f(\alpha) \to +\infty \text{ при } |\alpha| \to \infty,$$

b) существует $L > 0$ такое, что

$$|f(\alpha) - f(\beta)| > L \text{ влечет } p_{\alpha,\beta} = 0.$$

Далее, применяя соответственно теоремы 6.2, 6.1, 6.3, 6.4 (см. Приложение) для введенной функции $f(\alpha)$ и используя лемму 2.1, получаем утверждения теоремы 2.1.

Ассоциированный ветвящийся процесс. В случае $d = 1$ теорему 2.1 можно доказывать, используя некоторый ветвящийся процесс. А именно, введем следующий набор вероятностей $\{p_i\}$,

$$p_i = \mathsf{P}(\tau_i < \infty),$$

которые удовлетворяют следующим уравнениям

$$p_i = q(i,\varnothing) + \sum_j q(i,j)p_j + \sum_{j,k} q(i,jk)p_j p_k.$$

Эти уравнения подсказывают идею ветвящегося процесса ассоциированного с нашим процессом. Рассмотрим ветвящийся процесс с r типами частиц, где частица типа i не дает потомков с вероятностью $q(i,\varnothing)$, дает одного потомка j с вероятностью $q(i,j)$ и дает двух потомков j,k с вероятностью $q(i,jk)$. Тогда последнее уравнение можно интерпретировать, как уравнение для вероятности p_i того, что процесс, начавшись с типа i, когда-либо выродится. Как известно из теории ветвящихся процессов, матрица A (см. [1], [2]) управляет поведением этого процесса: если $\lambda > 1$ (λ – максимальное собственное значение матрицы A), то процесс выродится с вероятностью меньше 1 (это в точности соответствует случаю транзиентности струн); если $\lambda = 1$ то процесс критический (это соответствует нулевой возвратности); если $\lambda < 1$ то процесс вырождается с вероятностью 1 (струна эргодична).

Простейшая очередь. Рассмотрим простейший и хорошо известный в теории массового обслуживания пример системы с непрерывным временем $t \geqslant 0$, состоящей из одного прибора. На прибор за время Δt поступает требование типа i с вероятностью $\lambda_i \Delta t + o(\Delta t)$. Поступившее требование встает в начало очереди. Прибор обслуживает требование, стоящее в начале очереди, с интенсивностью, зависящей от типа требования, т.е. если в начале очереди стоит требование типа i, то за время Δt

оно удаляется из очереди с вероятностью $\mu_i \Delta t + o(\Delta t)$. Хорошо известно условие эргодичности (т.е. условие, когда очередь не накапливается):

$$\sum \frac{\lambda_i}{\mu_i} < 1.$$

Это условие может быть легко объяснено следующими эвристическими аргументами. В течении времени T поступят примерно $\lambda_i T$ требований типа i. Они будут обслуживаться время примерно $\sum_i \lambda_i T(\mu_i)^{-1}$. Но для того чтобы очередь не накапливалась, это время должно быть меньше, чем T, что приводит к условию $\sum \frac{\lambda_i}{\mu_i} < 1$.

Мы покажем, как это условие следует из теоремы 2.1. Вместо процесса с непрерывным временем рассмотрим марковскую цепь с параметрами

$$\begin{aligned} q(i,i) &= 1 - \varepsilon(\mu_i + \lambda_i), \\ q(i,\varnothing) &= \varepsilon\mu_i, \\ q(i,ij) &= \varepsilon\lambda_i, \\ q(i,kj) &= 0, \quad k \neq i, \end{aligned}$$

где ε достаточно мало так, что каждое $q(i,i) > 0$. Тогда

$$A = (a_{ij}), \qquad a_{ij} = \delta_{ij} + \varepsilon(-\mu_i\delta_{ij} + \lambda_j).$$

Заметим, что теорема 2.1 может быть переформулирована в следующем виде.

ТЕОРЕМА 2.1а. *Марковская цепь $\alpha(t)$ (описанная выше) эргодична тогда и только тогда, когда существует положительный вектор $e = (e_1, \ldots, e_n)$ такой, что*

(i) $Ae < e$, *т.е.* $(Ae)_i < e_i$ *для всех* i;

нуль возвратна тогда и только тогда, когда существует положительный вектор $e = (e_1, \ldots, e_n)$ такой, что

(ii) $Ae = e$;

транзиентна тогда и только тогда, когда существует положительный вектор $e = (e_1, \ldots, e_n)$ такой, что

(iii) $Ae > e$, *т.е.* $(Ae)_i > e_i$ *для всех* i.

Действительно, из условий (i), (ii), (iii) следует, что максимальное собственное значение λ матрицы A соответственно меньше, равно или больше 1. Чтобы доказать обратное, достаточно взять в качестве e собственный вектор матрицы A, соответствующий λ.

Пусть e – вектор с компонентами $e_i = 1/\mu_i$. Тогда

$$Ae = e + \varepsilon\left(\sum \frac{\lambda_k}{\mu_k} - 1\right),$$

и, таким образом, из Теоремы 2.1а следует приведенный выше критерий.

Обсудим теперь эвристические аргументы для более общей ситуации. Пусть интенсивность прихода требования типа j зависит от типа требования в начале очереди, т.е. если $\alpha = i\beta$, то j прибывает с интенсивностью λ_{ij}, μ_i – как и прежде, интенсивность обслуживания типа i. Предположим, что соответствующая Марковская цепь транзиентна. Тогда естественно предположить, что существуют предельные вероятности π_i того, что в начале очереди в момент t ($t \to \infty$) находится требование типа i. Следовательно, за время T поступят примерно $\sum_i \pi_i \lambda_{ij} T$ требований типа j. Среднее время их обслуживания примерно $\sum_j \frac{\sum_i \pi_i \lambda_{ij} T}{\mu_j}$ и оно должно быть больше чем T. Отсюда имеем эвристическое доказательство условия накопления очереди: $\sum_j \frac{\sum_i \pi_i \lambda_{ij}}{\mu_j} > 1$. Условие (iii) Теоремы 2.1 может быть сформулировано в близкой форме, если мы возьмем левый собственный вектор $v = (v_1, \dots, v_r)$ матрицы A, соответствующий максимальному собственному значению λ матрицы A, т.е. $vA = \lambda v$. Если $\sum_i v_i = 1$, то легко проверить, что $\lambda > 1$ тогда и только тогда, когда $\sum_j \frac{\sum_i v_i \lambda_{ij} T}{\mu_j} > 1$.

§ 3. Фундаментальное нелинейное уравнение: случай $d > 1$

Пусть $\tau_{\gamma,n}$ – первый момент, когда длина струны $\alpha(t)$ станет меньше n, т.е.

$$\tau_{\gamma,n} = \min\{t : |\alpha(t)| < n \mid \alpha(0) = \gamma\rho, |\gamma| = d, |\gamma\rho| = n\}.$$

Очевидно, что распределение $\tau_{\gamma,n}$ и $E_\gamma = \mathsf{E}\tau_{\gamma,n}$ не зависит от ρ и n. Введем следующие обозначения:

$$p = \{p(\gamma, \delta) : |\gamma| = d, |\delta| < d\}, \tag{4}$$

$$p(\gamma, \delta) = \mathsf{P}\big(\tau_{\gamma,n} < \infty, \alpha(\tau_{\gamma,n}) = \delta\rho \mid \alpha(0) = \gamma\rho\big). \tag{5}$$

Для возвратной цепи Маркова

$$\sum_\delta p(\gamma, \delta) = 1. \tag{6}$$

Чтобы получить уравнение для вектора p, нам понадобятся следующие определения.

Определение 3.1. Упорядоченная пара струн (α, β) согласована, если $\alpha = \gamma\rho$, $|\gamma| = d$, $\beta = \delta\rho$, $|\delta| < d$, для некоторых ρ, γ, δ. Для таких α и β мы положим $p(\alpha, \beta) = p(\gamma, \delta)$.

Определение 3.2. Последовательность струн $\Gamma = \alpha_1, \dots, \alpha_n$ называется путем, если каждая пара (α_i, α_{i+1}) согласована.

Пусть S – это некоторое множество струн и $\Gamma(\alpha, S)$ – множество путей Γ таких, что:

i) первый элемент α_1 пути Γ есть α;
ii) последний элемент $\alpha_n(\Gamma) = \alpha_n$ пути Γ принадлежит S и является единственным элементом пути Γ, принадлежащим S.

Обозначим через S_n множество путей длины, меньшей n

$$S_n = \{\alpha : |\alpha| < n\},$$

и определим вес пути

$$P_\Gamma = \prod_{i=1}^{n-1} p(\alpha_i, \alpha_{i+1}) \tag{7}$$

для $\Gamma = \alpha_1, \ldots, \alpha_n$. Нетрудно понять, что справедлива следующая система уравнений

$$p(\gamma, \delta) = q(\gamma, \delta) + \sum_{|\theta| \geqslant d} q(\gamma, \theta) \sum_{\Gamma \in \Gamma(\theta, S_d): \alpha_n(\Gamma) = \delta} P_\Gamma. \tag{8}$$

Систему (8) можно записать в виде

$$p = F(p). \tag{9}$$

В следующем параграфе мы будем работать с некоторыми положительными решениями фундаментального уравнения (9). По крайней мере одно такое решение всегда существует. Это вектор вероятностей $p(\gamma, \delta)$, определенный выше. Следующая лемма показывает, что существуют также и другие решения уравнения (8).

Лемма 3.1. *Для любой марковской цепи, удовлетворяющей условиям* (H), (ND), *существует положительное решение* p^* *системы* (8) *такое, что для любой струны* γ

$$\sum_{\delta} p^*(\gamma, \delta) = 1. \tag{10}$$

Замечание 1. Отсюда следует, что в транзиентном случае существует по крайней мере два положительных решения фундаментального уравнения, поскольку в транзиентном случае для вероятностного решения (10) не выполняется.

Замечание 2. В случае $d = 1$ лемма 3.1 тривиальна, поскольку (8) имеет вид (мы положили $p_i = p(i, \varnothing)$)

$$p_i = q(i, \varnothing) + \sum_j q(i, j) p_j + \sum_{j,k} q(i, jk) p_j p_k, \tag{11}$$

и легко видеть, что $p_i^* \equiv 1$ является решением.

Доказательство леммы 3.1. Пусть

$$\Delta = \left\{p^* = \{p^*(\gamma,\delta)\} : |\gamma| = d, |\delta| < d, p^*(\gamma,\delta) \geqslant 0, \forall\gamma \sum_{\delta} p^*(\gamma,\delta) = 1\right\}.$$

Докажем, что функция F отображает множество Δ в себя. Тогда утверждение леммы 3.1 будет следовать из теоремы о неподвижной точке. Если $g = F(p)$, $p \in \Delta$, то

$$\sum_{|\delta|<d} g(\gamma,\delta) = \sum_{|\delta|<d}\left[q(\gamma,\delta) + \sum_{|\theta|\geqslant d} q(\gamma,\theta) \sum_{\Gamma\in\Gamma(\theta,S_d):\alpha_n(\Gamma)=\delta} P_\Gamma\right] = \sum_{|\delta|<d} q(\gamma,\delta) + \sum_{|\theta|\geqslant d} q(\gamma,\theta) \sum_{\Gamma\in\Gamma(\theta,S_d)} P_\Gamma. \tag{12}$$

Поскольку $p \in \Delta$, мы можем интерпретировать $p(\gamma,\delta)$ как вероятность перейти за один шаг из состояния $\gamma\rho$ в состояние $\delta\rho$ для некоторой марковской цепи. В этом случае $\sum_{\Gamma(\theta,S_d)} P_\Gamma$ является вероятностью достигнуть множества S_d из начального состояния θ. Очевидно, эта вероятность равна 1. Из (12) мы получим

$$\sum_{|\delta|<d} g(\gamma,\delta) = \sum_{|\delta|<d} q(\gamma,\delta) + \sum_{|\theta|\geqslant d} q(\gamma,\theta) = 1.$$

Минимальное решение уравнения $F(p) = p$. В этой части нам будет полезно сравнивать поэлементно вектора из $\mathbb{R}^n$. Для этой цели введем следующее частичное упорядочение в $\mathbb{R}^n$: для $p, v \in \mathbb{R}^n$ $p \leqslant v$ тогда и только тогда, когда $p_i \leqslant v_i$, $i = 1,\dots,n$.

Определение 3.3. Пусть Λ – подмножество $\mathbb{R}^n$. Отображение $G\colon \Lambda \subset \mathbb{R}^n \to \mathbb{R}^n$ называется *изотонным на* Λ, если $G(p) \leqslant G(v)$ для любых $p \leqslant v$, $p, v \in \Lambda$.

Ясно, что F (см. (9)) изотонное на $\{p \geqslant 0\} = \mathbb{R}^n_+$, $n = r^d(1 + \dots + r^{d-1})$.

Далее мы будем использовать обозначение $p^k \uparrow p$, $k \to \infty$, если $p^0 \leqslant p^1 \leqslant \dots \leqslant p$, $\lim_{k\to\infty} p^k = p$.

Следующая лемма – вариант леммы Канторовича (см. [5]).

Лемма 3.2. *Пусть* $G\colon \mathbb{R}^n_+ \to \mathbb{R}^n_+$ *такое, что*

$$G \ \text{изотонно и непрерывно на}\ \mathbb{R}^n_+, \tag{13}$$

$$G(u) \leqslant u \ \text{для некоторого}\ u \in \mathbb{R}^n_+. \tag{14}$$

Тогда для последовательности $G^{(k)}(0) = G(G^{(k-1)}(0))$, $k = 0,1,\dots,$ *существует вектор* $\widetilde{p}$ *такой, что*

$$G^{(k)}(0) \uparrow \widetilde{p},$$
$$G(\widetilde{p}) = \widetilde{p},$$

и для любого иного решения уравнения $G(p) = p$, $p \geqslant 0$,

$$\widetilde{p} \leqslant p.$$

Замечание 3. Отметим, что функция F удовлетворяет условию (14), если в качестве u мы возьмем вероятностное решение уравнения (9).

Доказательство леммы 3.2. Из изотонности имеем

$$0 \leqslant u,$$
$$0 \leqslant G(0) \leqslant G(u) \leqslant u,$$
$$0 \leqslant G(0) \leqslant G^{(2)}(0) \leqslant u, \text{ и т. д.}$$
$$0 \leqslant G(0) \leqslant \dots \leqslant G^{(k)}(0) \leqslant u.$$

Следовательно, монотонная последовательность $\{G^{(k)}(0)\}$ имеет предел (обозначим его $\widetilde{p}$), причем $\widetilde{p} \leqslant u$.

Так как G непрерывно, то

$$\widetilde{p} = G(\widetilde{p}).$$

Допустим, что $p \geqslant 0$ – другое решение. Тогда, полагая $u = p$, получим $\widetilde{p} \leqslant p$.

$\widetilde{p}$ будем называть *минимальным решением уравнения* $G(p) = p$.

Докажем теперь, что минимальное решение совпадает с вероятностным.

Лемма 3.3. *Вектор вероятностей* $p = \{p(\gamma, \delta)\}$ (*определенный в* (5)) *является минимальным решением уравнения* $F(p) = p$.

Доказательство. Пусть $p^k = F^{(k)}(0)$, $p^k = \{p^k(\gamma, \delta)\}$. Мы будем рассматривать p^k как сумму по траекториям, т.е.

$$p^k(\gamma, \delta) = \sum_{(\alpha_0, \dots, \alpha_n)} p_{\alpha_0, \alpha_1} \cdot p_{\alpha_1, \alpha_2} \cdot \dots \cdot p_{\alpha_{n-1}, \alpha_n}.$$

Сумма берется по всем траекториям $(\alpha_0, \dots, \alpha_n)$ таким, что $\alpha_0 = \gamma$, $\alpha_n = \delta$, для всех $i < n$ $|\alpha_i| \geqslant d$ и $\#\{i : |\alpha_i| \leqslant |\alpha_{i+1}|\} \leqslant k$.

Следовательно,

$$p^k(\gamma, \delta) \uparrow p(\gamma, \delta), \qquad k \to \infty.$$

Замечание 4. Можно действовать другим образом. Запишем нашу систему в виде

$$y_0 = q_0 + F_0(y_0), \qquad y_0 = p,$$

где q_0 – константа и $F_0(0) = 0$. Положительные решения фундаментального уравнения соответствуют решениям $y_0 \geqslant q_0$. Полагая $y_0 = y_1 + q_0$, мы получим

$$y_1 = q_1 + F_1(y_1), \qquad y_1 \geqslant 0.$$

По индукции имеем y_n, q_n и F_n. Ясно, что $\sum_{n=0}^{\infty} q_n$ является минимальным решением и любое другое решение может быть представлено как сумма минимального и решения уравнения

$$z = F_\infty(z), \quad \text{где} \quad F_\infty = \lim_{n \to \infty} F_n.$$

Как и прежде, F_∞ – полином той же степени, что и F_0, и $F_\infty(0) = 0$.

Замечание 5. В введении мы отметили связь между марковскими цепями, описывающими динамику случайных струн, и блужданиями на свободных группах. Здесь мы обсудим этот вопрос подробнее. Рассмотрим свободно порожденную группу G с l генераторами $a_1, \ldots, a_l$ и положим $a_{-i} = a_i^{-1}$. В работах [6]–[9] рассматривалось произведение независимых случайных величин со значениями в G. А именно, заданы вероятности u_i, $i = 1, \ldots, l, -1, \ldots, -l$ для переходов $\alpha \to \alpha i$, $\alpha \in G$. Заметим, что a_i и a_{-i} не могут стоять рядом и это частный случай нашей задачи. Действительно, достаточно задать

$$q(i, \varnothing) = u_{-i}, \quad q(i, ij) = u_j, \quad q(i, kj) = 0, \quad \text{если} \quad k \neq i.$$

Эргодический случай в этой ситуации не возникает, и рассмотрен был в основном транзиентный случай. Изучение этой задачи началось в работе [7], там же было получено подобное (9) фундаментальное уравнение следующего вида:

$$p_i = u_{-i} + p_i \sum_{k \neq -i} u_k p_k.$$

Оказалось, что это уравнение точно решается, и во многих работах использовался этот факт. В нашем случае точного решения нет.

Замечание 6. Из леммы 3, главы 5 в [10] следует, что в случае $d = 1$ существует только одно решение фундаментального уравнения, удовлетворяющее условию $0 \leqslant p_i \leqslant 1$. К сожалению, примененный нами метод не позволил доказать единственность решения $p^* \in \Delta$.

Замечание 7. Пусть $\mathscr{Q} = \{q(\gamma, \delta)\}$ – пространство параметров (мы пренебрегаем переходами около границы). Мы хотим доказать, что наша марковская цепь может быть нуль, возвратной только на подмножестве лебеговой меры ноль в пространстве параметров.

Предположим, что для параметров $\{q(\gamma, \delta)\}$ цепь $\mathscr{L}$ нуль-возвратна, и p^* – решение фундаментального уравнения. Рассмотрим следующее малое возмущение параметров q.

Зафиксируем некоторые достаточно малые $\varepsilon(\gamma, \theta) > 0$ для любых γ, δ так, что $|\gamma| = d$, $|\theta| \geqslant d$. Подберем другие $\varepsilon(\gamma, \delta)$ так, чтобы p^* осталось решением фундаментального уравнения для параметров

$$\begin{aligned}
\widetilde{q}(\gamma, \theta) &= q(\gamma, \theta) + \varepsilon(\gamma, \theta), \\
\widetilde{q}(\gamma, \delta) &= q(\gamma, \delta) + \varepsilon(\gamma, \delta) > 0.
\end{aligned}$$

$\varepsilon(\gamma, \delta)$ могут быть найдены из уравнения (определим для краткости $\Gamma(\theta, \delta) = \Gamma(\theta, S_d) : \alpha_n(\Gamma) = \delta$):

$$\begin{aligned}
p(\gamma, \delta) &= q(\gamma, \delta) + \sum_{|\theta| \geqslant d} q(\gamma, \theta) \sum_{\Gamma \in \Gamma(\theta, \delta)} P_\Gamma^* \\
&= q(\gamma, \delta) + \sum_{|\theta| \geqslant d} \big(q(\gamma, \theta) + \varepsilon(\gamma, \theta) - \varepsilon(\gamma, \theta)\big) \sum_{\Gamma \in \Gamma(\theta, \delta)} P_\Gamma^* \\
&= q(\gamma, \delta) - \sum_{|\theta| \geqslant d} \varepsilon(\gamma, \theta) \sum_{\Gamma \in \Gamma(\theta, \delta)} P_\Gamma^* + \sum_{|\theta| \geqslant d} \big(q(\gamma, \theta) + \varepsilon(\gamma, \theta)\big) \sum_{\Gamma \in \Gamma(\theta, \delta)} P_\Gamma^*.
\end{aligned}$$

Итак, мы имеем

$$\varepsilon(\gamma,\delta) = -\sum_{|\theta|\geqslant d}\varepsilon(\gamma,\theta)\sum_{\Gamma\in\Gamma(\theta,\delta)}P^*_\Gamma.$$

По определенным таким образом параметрам $\widetilde{q}(\gamma,\delta)$ построим матрицу $\widetilde{A}(p^*)$ (по аналогии с матрицей $A(p)$, см. уравнение (22) в следующем параграфе). Мы получим $\widetilde{A}(p^*) > A(p^*)$, поскольку $\widetilde{q}(\gamma,\theta) > q(\gamma,\theta)$ для любых γ,θ ($|\gamma| = d$, $|\theta| \geqslant d$). Отсюда следует, что марковская цепь, определенная по $\widetilde{q}$, будет транзиентна.

Аналогично мы можем зафиксировать $\varepsilon(\gamma,\theta) > 0$ для γ,θ таких, что $|\gamma| = d$, $|\theta| \leqslant d$, и получить эргодичность.

§ 4. Основной критерий. Случай $d > 1$

Подобно случаю $d = 1$, мы получим критерий эргодичности, нулевой возвратности и транзиентности через максимальное собственное значение положительной матрицы A, зависящей, однако, (и в этом существенное отличие от случая $d = 1$) от решения $p^*(\gamma,\delta)$, определенного в лемме 3.1.

Предположим, что наша марковская цепь эргодична. Тогда мы можем получить уравнения, подобные уравнениям (2) § 2. Введем

$$E_\gamma = \mathsf{E}\big[\tau_{\gamma,n} \mid \alpha(0) = \gamma\rho\big], \qquad |\gamma| = d, \quad |\gamma\rho| = n. \tag{15}$$

В § 3 было упомянуто, что E_γ не зависит от ρ.

Для $\alpha = \gamma\rho$, $|\gamma| = d$, мы полагаем $E_\alpha = E_\gamma$ и

$$T(\alpha) \overset{\text{def}}{=} \mathsf{E}\tau(\alpha,d), \tag{16}$$

где $\tau(\alpha,d)$ – первый момент попадания в S_d, при условии что начальное состояние было α. Положим $T(\alpha) = 0$ для α: $|\alpha| < d$. Легко видеть, что

$$T(\gamma\rho) = E_\gamma + \sum_\delta p(\gamma,\delta)T(\delta\rho), \tag{17}$$

или

$$T(\gamma\rho) = \sum_\delta p(\gamma,\delta)(E_\gamma + T(\delta\rho)). \tag{18}$$

Из (18) следует, что

$$T(\alpha) = \sum_{\Gamma\in\Gamma(\alpha,S_d)} P_\Gamma E_\Gamma, \tag{19}$$

где

$$E_\Gamma = \sum_{i=1}^N E_{\alpha_i} \tag{20}$$

для $\Gamma = \alpha_1, \dots, \alpha_N$. Теперь мы можем записать для E_γ:

$$E_\gamma = 1 + \sum_{|\theta| \geqslant d} q(\gamma, \theta) T(\theta) = 1 + \sum_{|\theta| \geqslant d} q(\gamma, \theta) \sum_{\Gamma \in \Gamma(\theta, S_d)} P_\Gamma E_\Gamma, \tag{21}$$

или, в матричной форме,

$$E = \bar{1} + A(p) E. \tag{22}$$

В случае $d > 1$ матрица A от зависит от вектора вероятностей p (см. (5)).

Теперь перейдем от эргодического к общему случаю. Мы можем сформулировать основной критерий.

ТЕОРЕМА 4.1. *Пусть p^* – любое положительное решение уравнения* (9) (*по лемме* 3.1 *существующее и удовлетворяющее* (10)), λ – *максимальное собственное число матрицы* $A(p^*)$. *Процесс $\alpha(t)$ является эргодичным, нулевым возвратным, транзиентным тогда и только тогда, когда соответственно* $\lambda < 1$, $\lambda = 1$, $\lambda > 1$.

ДОКАЗАТЕЛЬСТВО. Мы будем использовать метод функции Ляпунова. Как и в случае $d = 1$, мы будем искать функцию Ляпунова f как среднее время достижения множества S_d (см. (18)). Определим функцию f с помощью рекуррентных соотношений

$$\begin{gathered} f(\gamma\rho) = \sum_\delta p^*(\gamma, \delta)(f(\delta\rho) + e_\gamma), \\ f(\delta) = 0 \quad \text{для} \quad \delta\colon |\delta| < d, \end{gathered} \tag{23}$$

где $\{e_\gamma\}_{|\gamma| = d}$ – компоненты собственного вектора, соответствующего максимальному собственному значению λ, т.е.

$$\lambda e_\gamma = \sum_{|\theta| \geqslant d} q(\gamma, \theta) \sum_{\Gamma \in \Gamma(\theta, S_d)} P_\Gamma^*, \tag{24}$$

и по теореме Перрона–Фробениуса для всех γ

$$e_\gamma > 0. \tag{25}$$

Заметим,что из рекуррентных уравнений (23) следует, что для всех θ, ρ: $|\theta| \geqslant d$

$$f(\theta\rho) = \sum_{\Gamma \in \Gamma(\theta, S_d)} P_\Gamma^* (e_\Gamma + f(\delta\rho)). \tag{26}$$

Здесь e_Γ определяется через e по формулам (подобно определению E_Γ, см. (20)):

$$e_\Gamma = \sum_{i=1}^N e_{\alpha_i} \quad \text{для} \quad \Gamma = \alpha_1, \dots, \alpha_N.$$

Лемма 4.1. *Пусть f задана с помощью* (23). *Тогда справедливы следующие утверждения*:

(i) *если* $\lambda > 1$, *то существует* $\varepsilon > 0$ *такое, что*

$$\mathsf{E}\big[f(\alpha(t+1)) - f(\alpha(t)) \mid \alpha(t) = \alpha\big] > \varepsilon$$

для всех α, $|\alpha| \geqslant d$, *и любого* t;

(ii) *если* $\lambda = 1$, *то*

$$\mathsf{E}\big[f(\alpha(t+1)) - f(\alpha(t)) \mid \alpha(t) = \alpha\big] = 0$$

для всех α, $|\alpha| \geqslant d$, *и любого* t;

(iii) *если* $\lambda < 1$, *то существует* $\varepsilon > 0$ *такое, что*

$$\mathsf{E}\big[f(\alpha(t+1)) - f(\alpha(t)) \mid \alpha(t) = \alpha\big] < -\varepsilon$$

для всех α, $|\alpha| \geqslant d$, *и любого* t.

Доказательство. Доказательство проводится с помощью следующих выкладок:

$$\begin{aligned}
&\mathsf{E}[f(\alpha(t+1)) - f(\alpha(t)) \mid \alpha(t) = \gamma\rho] \\
&= \sum_{|\delta|<d} q(\gamma,\delta) f(\delta\rho) + \sum_{|\theta|\geqslant d} q(\gamma,\theta) f(\theta\rho) - f(\gamma\rho) \\
&\overset{(26)}{=} \sum_{|\delta|<d} q(\gamma,\delta) f(\delta\rho) + \sum_{|\theta|\geqslant d} q(\gamma,\theta) \sum_{|\delta|<d} \sum_{\Gamma\in\Gamma(\theta\rho,\{\delta\rho\})} P^*_\Gamma [e_\Gamma + f(\delta\rho)] - f(\gamma\rho) \\
&= \sum_{|\delta|<d} f(\delta\rho) \Bigg[q(\gamma,\delta) + \sum_{|\theta|\geqslant d} \sum_{\Gamma\in\Gamma(\theta\rho,\{\delta\rho\})} P^*_\Gamma \Bigg] + \sum_{|\theta|\geqslant d} q(\gamma,\theta) \sum_{\Gamma\in\Gamma(\theta,S_d)} P^*_\Gamma * e_\Gamma - f(\gamma\rho)
\end{aligned}$$

так как p^* – решение уравнения (9), мы можем продолжить

$$\begin{aligned}
&= \sum_{|\delta|<d} f(\delta\rho) p^*(\gamma,\delta) + \sum_{|\theta|} q(\gamma,\theta) \sum_{\Gamma\in\Gamma(\theta,S_d)} P^*_\Gamma * e_\Gamma - f(\gamma\rho) \\
&\overset{(24)}{=} \sum_{|\delta|<d} f(\delta\rho) p^*(\gamma,\delta) + \lambda e_\gamma - f(\gamma\rho) \\
&\overset{(23)}{=} \lambda e_\gamma - e_\gamma = (\lambda - 1) e_\gamma.
\end{aligned}$$

Чтобы воспользоваться критериями Приложения, нам нужно, чтобы функция f удовлетворяла свойствам, которые мы сформулируем в виде следующей леммы.

Лемма 4.2.

(i) $f(\alpha) \to \infty$, *если* $|\alpha| \to \infty$.

(ii) *Существует* $L > 0$ *такое, что неравенство*

$$|f(\alpha) - f(\beta)| > L \quad \text{влечет} \quad p_{\alpha,\beta} = 0. \tag{27}$$

Доказательство. Первое свойство легко следует из (23) и (10). Второе свойство доказывается сложнее. Запишем (27) в эквивалентной форме: существует $L > 0$ такое, что для любых γ, θ, ρ: $|\gamma| = d$, $|\theta| \leqslant 2d$

$$|f(\gamma\rho) - f(\theta\rho)| \leqslant L. \tag{28}$$

Мы докажем, что существует $K > 0$ такое, что для всех $\alpha, \delta_1, \delta_2$: $|\delta_1| < d$, $|\delta_2| < d$

$$|f(\delta_1\alpha) - f(\delta_2\alpha)| < K. \tag{29}$$

Тогда (28) нетрудно вывести из (29) с помощью рекуррентных соотношений (23).

Определим

$$M_\alpha = \max_{|\delta|<d} f(\delta\alpha), \quad m_\alpha = \min_{|\delta|<d} f(\delta\alpha), \quad \Delta_\alpha = M_\alpha - m_\alpha.$$

Чтобы получить оценку (29), достаточно показать, что существуют $a > 0, 1 > b > 0$ такие, что для всех $\alpha = \gamma\rho$, $|\gamma| = d$,

$$\Delta_\alpha \leqslant a + b\Delta_\rho, \tag{30}$$

и, следовательно,

$$\Delta_\alpha \leqslant a + b(a + b(\ldots)) = a(1 + b + b^2 + \ldots) = K < \infty$$

Докажем (30). Из (26) имеем

$$f(\delta\alpha) = f(\delta\gamma\rho) = \sum_{\Gamma\in\Gamma(\delta\alpha, S_{|\alpha|})} P^*_\Gamma(e_\Gamma + f(\varkappa_\Gamma)), \tag{31}$$

где $\varkappa_\Gamma$ – последний элемент Γ.

Из условий $\Gamma \in \Gamma(\delta\alpha, S_{|\alpha|})$ следует, что $\varkappa_\Gamma = \widetilde{\delta}\rho$ для некоторого $\widetilde{\delta}$: $|\widetilde{\delta}| < d$. Поэтому (31) можно переписать в виде

$$f(\delta\alpha) = a(\delta\gamma) + \sum_{|\widetilde{\delta}|<d} p_\gamma(\delta, \widetilde{\delta}) f(\widetilde{\delta}\rho),$$

где

$$a(\delta\gamma) = \sum_{\Gamma\in\Gamma(\delta\alpha, S_{|\alpha|})} P^*_\Gamma e_\Gamma,$$

$$p_\gamma(\delta, \widetilde{\delta}) = \sum_{\Gamma\in\Gamma(\delta\alpha, \{\widetilde{\delta}\rho\})} P^*_\Gamma.$$

Заметим, что

$$2\max_{\delta\gamma} a(\delta\gamma) = a < \infty,$$

$$\sum_{|\widetilde{\delta}|<d} p_\gamma(\delta, \widetilde{\delta}) = 1.$$

Это можно показать аналогично доказательству леммы 3.1. Кроме того, из условия (ND) (невырожденности) следует, что

$$\min_{\delta,\gamma,\widetilde{\delta}} p_\gamma(\delta,\widetilde{\delta}) = \varepsilon > 0.$$

Из сказанного выше получаем

$$|f(\delta_1\gamma\rho) - f(\delta_2\gamma\rho)| \leqslant a + \left| \sum_{|\widetilde{\delta}|<d} p_\gamma(\delta_1,\widetilde{\delta}) f(\widetilde{\delta}\rho) - \sum_{|\widetilde{\delta}|<d} p_\gamma(\delta_2,\widetilde{\delta}) f(\widetilde{\delta}\rho) \right|$$

$$= a + \left| \sum_{|\widetilde{\delta}|<d} \left[p_\gamma(\delta_1,\widetilde{\delta}) - \varepsilon\right] f(\widetilde{\delta}\rho) - \sum_{|\widetilde{\delta}|<d} \left[p_\gamma(\delta_2,\widetilde{\delta}) - \varepsilon\right] f(\widetilde{\delta}\rho) \right|$$

$$= a + \Big| \sum_{|\widetilde{\delta}|<d} \left[p_\gamma(\delta_1,\widetilde{\delta}) - \varepsilon\right] \{f(\widetilde{\delta}\rho) - m_\rho\} - \sum_{|\widetilde{\delta}|<d} \left[p_\gamma(\delta_2,\widetilde{\delta}) - \varepsilon\right] \{f(\widetilde{\delta}\rho) - m_\rho\} \Big|$$

$$\leqslant a + \max_{\delta_1,\delta_2} \sum_{|\widetilde{\delta}|<d} \left[p_\gamma(\delta_i,\widetilde{\delta}) - \varepsilon\right] \{f(\widetilde{\delta}\rho) - m_\rho\}$$

$$\leqslant a + \left[1 - \sum_{|\widetilde{\delta}|<d} \varepsilon\right] \{M_\rho - m_\rho\}.$$

Полагая $b = 1 - \sum_{|\widetilde{\delta}|<d} \varepsilon$, получаем (30).

Теперь утверждение теоремы 4.1 вытекает из лемм 4.1, 4.2 и теорем Приложения.

§ 5. Классификация цепи Маркова с помощью минимального решения

ТЕОРЕМА 5.1.

(i) *Марковская цепь $\alpha(t)$ возвратна тогда и только тогда, когда минимальное решение $\widetilde{p}(\gamma,\delta)$ удовлетворяет свойству*

$$\sum_\delta \widetilde{p}(\gamma,\delta) = 1$$

для всех γ.

(ii) *Пусть марковская цепь $\alpha(t)$ возвратна. Тогда она нуль-возвратна, только если спектральный радиус (максимальное собственное значение) матрицы $F' = \left(\dfrac{\partial F_i}{\partial p_j}(\widetilde{p})\right)$ равен 1.*

(i) очевидно. Доказательство (ii) содержится в следующих леммах.

ЛЕММА 5.1. *Пусть $G\colon \mathbb{R}_+^n \to \mathbb{R}_+^n$, удовлетворяет условиям* (13), (14),

$$G(p) > 0, \qquad p \geqslant 0, \tag{32}$$

и G имеет производную $D = \left(\dfrac{\partial G_i}{\partial p_j}(\widetilde{p})\right)$ такую, что

$$D > 0, \quad \text{т.е.} \quad \frac{\partial G_i}{\partial p_j}(\widetilde{p}) > 0 \quad \text{для всех } i, j. \tag{33}$$

Тогда спектральный радиус $R(D)$ матрицы D не превосходит 1: $R(D) \leqslant 1$.

Доказательство. Предположим, что $R(D) > 1$. Тогда для некоторого $v \in \mathbb{R}^n$, $v > 0$, имеем $Dv > v$, и, следовательно, существует $\varepsilon > 0$ такое, что $\widetilde{p} > u = \widetilde{p} - \varepsilon v > 0$ и

$$G(u) = \widetilde{p} - \varepsilon Dv + o(\varepsilon) = u + \varepsilon(v - Dv) + o(\varepsilon) < u.$$

Таким образом,

$$0 < G(u) < u < \widetilde{p}.$$

С другой стороны, $\widetilde{p} \leqslant u$ (см. доказательство леммы 3.2). Полученное противоречие доказывает, что $R(D) \leqslant 1$.

Лемма 5.2. *Пусть G удовлетворяет условиям* (13), (14) (32), (33), *и*

$$\text{для всех } i,j,s \quad \frac{\partial^2 G_i}{\partial p_j \partial p_s}(p) > 0 \quad \text{при} \quad p \geqslant 0, \tag{34}$$
$$\text{и хотя бы для одного } i \quad p_i > 0.$$

Тогда

a) $R(G'(p)) < 1$, *если* $p \leqslant \widetilde{p}, p \neq \widetilde{p}$,

b) $R(D) = 1$, *если существует единственное решение уравнения* $G(p) = p$ (*тогда этим решением является* $\widetilde{p}$).

Доказательство. a) Из (34) и (13) следует, что

$$0 \leqslant G'(p) < G'(\widetilde{p}) = D.$$

Значит,

$$R(G'(p)) < R(D) \leqslant 1.$$

b) Пусть $p \geqslant 0$ – другое решение уравнения $G(p) = p$, тогда $p \geqslant \widetilde{p}, p \neq \widetilde{p}$ (см. лемму 3.2). Согласно условию (34) G – строго порядково-выпукло (см. 13.3.2 [5]), т.е.

$$G(z) \geqslant G(y) + G'(y)(z - y), \quad \text{если} \quad z - y \geqslant 0 \quad \text{или} \quad z - y \leqslant 0,$$

и неравенство строгое, если $z \neq y$. Поэтому

$$p = G(p) > G(\widetilde{p}) + D(p - \widetilde{p}) \quad \text{или} \quad p - \widetilde{p} > D(p - \widetilde{p}).$$

Следовательно, $R(D) < 1$.

Лемма 5.3. *Пусть G удовлетворяет свойствам* (13), (14), (32)-(34). *Тогда*

a) *если* $R(D) < 1$, *то существуют* $c_1, c_2 > 0$ *такие, что*

$$\left\|\widetilde{p} - G^{(k)}(0)\right\| < c_1 e^{-c_2 k},$$

b) *если* $R(D) = 1$, *то существуют* c_1, $c_2 > 0$ *такие, что*

$$\frac{c_1}{k} < \left\|\widetilde{p} - G^{(k)}(0)\right\| < \frac{c_2}{k}. \tag{35}$$

Доказательство. a) следует из леммы 5.2.

b) Рассмотрим случай $n=1$. Из условий (13)–(14) вытекает, что

$$G(p)=\widetilde{p}+(p-\widetilde{p})+b(p-\widetilde{p})^2+o(|p-\widetilde{p}|^2),\qquad b>0.$$

Обозначим $G^{(k)}(0)$ через p^k. Тогда

$$p^k=p^{k-1}+b(p^{k-1}-\widetilde{p})^2+o(|p^{k-1}-\widetilde{p}|^2),$$

$$\frac{1}{\widetilde{p}-p^k}-\frac{1}{\widetilde{p}-p^{k-1}}=\frac{b(p^{k-1}-\widetilde{p})^2+o(|p^{k-1}-\widetilde{p}|^2)}{(p^{k-1}-\widetilde{p})^2+o(|p^{k-1}-\widetilde{p}|^2)}=b+O(|p^{k-1}-\widetilde{p}|).$$

Следовательно,

$$\frac{1}{\widetilde{p}-p^k}=b\cdot k+o(k),$$

и немедленно получаем (35).

Пусть теперь $n>1$. Как и в случае $n=1$,

$$G(p)=\widetilde{p}+D(p-\widetilde{p})+\frac{1}{2}B(p-\widetilde{p})\cdot(p-\widetilde{p})+o(\|p-\widetilde{p}\|^2), \tag{36}$$

где $B(y)=(b_{i,j}(y))$, $b_{i,j}(y)=\sum_s\dfrac{\partial^2 G_i}{\partial p_j\partial p_s}(\widetilde{p})y_s$. Пусть π – собственный вектор матрицы D с собственным значением 1. Тогда

$$\pi>0 \tag{37}$$

и

$$\pi p^k=\pi p^{k-1}+\pi B(p^{k-1}-\widetilde{p})\cdot(p^{k-1}-\widetilde{p})+o(\|p-\widetilde{p}\|^2).$$

Из условий (34) и (37) вытекает, что существуют $b_1,b_2>0$ такие, что

$$b_1\left(\pi(p^{k-1}-\widetilde{p})\right)^2\leqslant\pi B(p^{k-1}-\widetilde{p})\cdot(p^{k-1}-\widetilde{p})\leqslant b_2\left(\pi(p^{k-1}-\widetilde{p})\right)^2.$$

Тогда, рассуждая так же, как и в случае $n=1$, получим (35).

Теорема 5.2. *Марковская цепь $\mathscr{L}$ является*

a) *эргодичной или транзиентной тогда и только тогда, когда существуют $c_1,c_2>0$ такие, что*

$$\left\|\widetilde{p}-F^{(k)}(0)\right\|<c_1e^{-c_2k}, \tag{38}$$

b) *нуль-возвратной тогда и только тогда, когда существуют $c_1,c_2>0$ такие, что*

$$\frac{c_1}{k}<\left\|\widetilde{p}-F^k(0)\right\|<\frac{c_2}{k}. \tag{39}$$

Доказательство. Прежде всего, покажем, что

$$R(F'(\widetilde{p})) = R(A(\widetilde{p})), \quad \text{если} \quad p > 0 \tag{40}$$
$$\text{и} \quad \sum_{\delta} p(\gamma,\delta) = 1 \quad \text{для всех} \quad \gamma, \ |\gamma| = d.$$

Пусть $v = \{v(\gamma,\delta)\}$ – собственный вектор матрицы $F'(p)$, соответствующий максимальному собственному значению $0 < \lambda = R(F'(p))$. В координатной форме

$$\big(F'(p)v\big)\,(\gamma,\delta) = \sum_{\theta} q(\gamma,\theta) \sum_{\Gamma\in\Gamma(\theta,\delta)} P_{\Gamma} v^*_{\Gamma} = \lambda v(\gamma,\delta),$$

где

$$v^*_{\Gamma} = \sum_{i=0}^{N} \frac{v(\gamma_i,\delta_i)}{p(\gamma_i,\delta_i)}, \quad \text{если} \quad \Gamma = \big((\gamma_0,\delta_0),\dots,(\gamma_N,\delta_N)\big).$$

Если $\widetilde{v}(\gamma) = \sum_{\delta} v(\gamma,\delta)$, то из свойства $\sum_{\delta} p(\gamma,\delta) = 1$ получаем

$$\sum_{\delta} \big(F'(p)v\big)\,(\gamma,\delta) = \sum_{\theta} q(\gamma,\theta) \sum_{\Gamma\in\Gamma(\theta,S_d} P_{\Gamma}\widetilde{v}_{\Gamma} = \lambda\widetilde{v}(\gamma), \tag{41}$$

где

$$\widetilde{v}_{\Gamma} = \sum_{i=0}^{N} \widetilde{v}(\gamma_i), \quad \text{если} \quad \Gamma = \big((\gamma_0,\delta_0),\dots,(\gamma_N,\delta_N)\big).$$

Запишем (41) в матричной форме:

$$A(p)\widetilde{v} = \lambda\widetilde{v},$$

откуда вытекает (40).

В возвратном случае мы можем положить $p = \widetilde{p}$. Тогда из теоремы 4.1 следует, что $R(F'(\widetilde{p})) = R(A(\widetilde{p})) = 1$, если $\mathscr{L}$ нуль-возвратна, $R(F'(\widetilde{p})) = R(A(\widetilde{p})) < 1$, если $\mathscr{L}$ эргодична.

Рассмотрим теперь транзиентный случай.

Лемма 3.1 и лемма 3.2 дают два разных решения уравнения $F(p) = p$. А именно p^* и $\widetilde{p}$. Если бы (34) выполнялось для $F(p)$, тогда из леммы 5.2 мы получили бы $R(F'(\widetilde{p})) < 1$. На самом деле, (34) не выполняется для F. Но это условие справедливо для

$$G(p) = F^{(2)}(p).$$

p^* и $\widetilde{p}$ также являются решениями уравнения $G(p) = p$, поэтому

$$R(G'(\widetilde{p})) < 1, \tag{42}$$

но $G'(\widetilde{p}) = F'(\widetilde{p}) \cdot F'(\widetilde{p})$. Значит,

$$R(F'(\widetilde{p})) < 1.$$

Поэтому в эргодическом и транзиентном случае

$$R(F'p)) \leqslant R(F'(\widetilde{p})) < 1 \text{ для } 0 \leqslant p \leqslant \widetilde{p},$$

откуда немедленно получаем a).

Докажем b). Так как $F^{(k)}(0)$ – монотонная последовательность, достаточно доказать (39) для последовательности $F^{(2k)}(0) = G^{(k)}(0)$. А для нее (39) следует из леммы 5.3.

ЛЕММА 5.4.

(a) *Если марковская цепь $\alpha(t)$ эргодична или транзиентна, то $R(F') < 1$.*

(b) *Если марковская цепь $\alpha(t)$ нуль-возвратна, то $R(F') = 1$.*

(c) *В возвратном случае положительное решение фундаментального уравнения, удовлетворяющее условию $\sum_\delta p(\gamma, \delta) = 1$, единственно.*

ДОКАЗАТЕЛЬСТВО. Случаи (a) и (b) были доказаны выше. Докажем (c).

В возвратном случае минимальное решение $\tilde{p}$ удовлетворяет этому условию. Для любого другого p, $p \geqslant \tilde{p}$, $p \neq \tilde{p}$. Поэтому $\sum_\delta p(\gamma, \delta) > 1$.

§ 6. Приложение

Здесь мы напомним теоремы, которые были использованы в § 2 и 3. Мы даем их в более простой форме, чем в [3], [4].

Пусть $\mathscr{L}$ – однородная цепь Маркова со счетным числом состояний $\mathscr{A} = \{\alpha_i, i = 1, 2, \dots\}$ и вероятностями перехода за один шаг $p_{\alpha,\beta}$. Предположим, что $\mathscr{L}$ неприводима и непериодична. Состояние цепи в момент времени t обозначим через ξ_t.

ТЕОРЕМА 6.1. *Марковская цепь $\mathscr{L}$ возвратна тогда и только тогда, когда существует неотрицательная функция $f(\alpha)$, $\alpha \in \mathscr{A}$, и конечное множество A такие, что*

(i) $\mathsf{E}\big[f(\xi_{t+1}) - f(\xi_t) \mid \xi_t = \alpha_i\big] \leqslant 0 \quad \forall \alpha_i \notin A$,

(ii) $f(\alpha_i) \to \infty$, *если* $i \to \infty$.

ТЕОРЕМА 6.2. *Марковская цепь $\mathscr{L}$ эргодична тогда и только тогда, когда существуют положительная функция $f(\alpha)$, $\alpha \in \mathscr{A}$, $\varepsilon > 0$ и конечное множество A такие, что*

(i) $\mathsf{E}\big[f(\xi_{t+1}) - f(\xi_t) \mid \xi_t = \alpha_i\big] \leqslant -\varepsilon \quad \forall \alpha_i \notin A$,

(ii) $\mathsf{E}\big[f(\xi_{t+1}) \mid \xi_t = \alpha_i\big] < \infty \quad \forall \alpha_i \in A$.

ТЕОРЕМА 6.3. *Для того чтобы марковская цепь $\mathscr{L}$ была неэргодична, достаточно, чтобы существовали функция $f(\alpha)$, $\alpha \in \mathscr{A}$, и константы C, L такие, что*

(i) $\mathsf{E}\big[f(\xi_{t+1}) - f(\xi_t) \mid \xi_t = \alpha_i\big] \geqslant 0 \quad \forall \alpha_i \in \{f(\alpha) > C\}$, *где множества* $\{f(\alpha) > C\}$ *и* $\{f(\alpha) \leqslant C\}$ *не пусты,*

(ii) $\mathsf{E}\big[f(\xi_{t+1}) - f(\xi_t) \mid \xi_t = \alpha_i\big] \leqslant L \quad \forall \alpha_i \in A$.

ТЕОРЕМА 6.4. *Для того чтобы марковская цепь $\mathscr{L}$ была транзиентной, достаточно, чтобы существовали положительная функция $f(\alpha)$, $\alpha \in \mathscr{A}$, и константы $\varepsilon, C, L > 0$ такие, что*

(i) $\mathsf{E}\big[f(\xi_{t+1}) - f(\xi_t) \mid \xi_t = \alpha_i\big] \geqslant \varepsilon \quad \forall \alpha_i \in \{f(\alpha) > C\}$, *где множество* $\{f(\alpha) > C\}$ *не пусто,*

(ii) *из неравенства* $|f(\alpha) - f(\beta)| > L$ *следует* $p_{\alpha,\beta} = 0$.

СПИСОК ЛИТЕРАТУРЫ

[1] Malyshev V. A. Stabilization laws for processes with a localised interaction // Rapports de Recherche INRIA. 1992. № 1635.

[2] Malyshev V. A. Evolution of a random string: stabilization laws. Preprint: INRIA, 1992.

[3] Малышев В. А., Меньшиков М. В. Эргодичность, непрерывность и аналитичность счетных цепей Маркова // Труды ММО. 1979. Т. 39. С. 3–48.

[4] Fayolle G., Malyshev V. A., Menshikov M. V. Topics in Constructive Theory of Countable Markov Chains: Cambridge Univ. Press, 1993.

[5] Ortega J., Rheinboldt W. C. Iterative Solution of Nonlinear Equations in Several Variables. New York, London: Academic Press, 1971.

[6] Sawyer S., Steger T. The rate of escape for anisotropic random walks in a tree // Probab. Theory Rel. Fields. 1987. V. 76. P. 207–230.

[7] Дынкин Е. Б., Малютов М. Случайные блуждания на группах с конечным числом образующих // ДАН СССР. 1961. Т. 2. С. 399–402.

[8] Левит Б., Молчанов С. А. Инвариантные цепи на группах с конечным числом образующих // Вест. МГУ. 1971. Т. 6. С. 80–88.

[9] Derriennic Y. Marche aléatoire sur le groupe libre et frontière de Martin // Z. Wahrscheinlichkeitstheor. Verw. Geb. 1975. V. 32. P. 261–276.

[10] Севастьянов Б. А. Ветвящиеся процессы. М.: Наука, 1971.

Московский государственный
университет им. М.В. Ломоносова

Поступила в редакцию
15.02.1995