УДК 519.9

А. А. Ибрагимов
(Ин-т кибернетики НПО „Кибернетика" АН Республики Узбекистан, Ташкент)

МАРКОВСКИЕ ИГРЫ С НЕСКОЛЬКИМИ ЭРГОДИЧЕСКИМИ КЛАССАМИ

We consider Markov games of the general form characterized as follows: for all stationary strategies of players, the set of game states is partitioned into several ergodic sets and a transient set that may vary depending on strategies of players. As a criterion, we choose a mean payoff of the first player per unit time. We prove that the general Markov game with finite sets of states and decisions of the both players has a value and both players have ε-optimal stationary strategies. We demonstrate the correctness of this statement by the well-known Blackwell example, i.e., by the game "Big Match".

Розглянуто марковські ігри загального вигляду, які характеризуються тим, що при будь-яких стаціонарних стратегіях гравців множина станів гри розбивається на декілька ергодичних множин і незворотну множину, що можуть змінюватися в залежності від стратегії гравців. За критерій вибрано середній виграш першого гравця за одиницю часу. Доведено, що загальна марковська гра із скінченною множиною станів і розв'язків обох гравців має значення, а обидва гравці мають ε-оптимальні стаціонарні стратегії. Справедливість цього твердження продемонстровано на прикладі Блекуелла — „великий матч".

1. Введение. Постановка задачи. Марковская игра, как последовательность случайных величин $\xi_n = (s_n, a_n, b_n)$, $n = 1, 2, \ldots$, определяется совокупностью пяти объектов:

1) множеством состояний S, $s_n \in S$;
2) множествами решений A_i, $i \in S$, $a_n \in A_{s_n}$ игрока I;
3) множествами решений B_i, $i \in S$, $b_n \in B_{s_n}$ игрока II;
4) переходной функцией $p(s_{n+1} \mid \xi_n)$;
5) платежной функцией $h(\xi_n)$, $|h| \leq C < \infty$.

Игра начинается с некоторого состояния $s_1 = i \in S$ и продолжается до определенного момента времени $T < \infty$ или бесконечно долго. На n-м шаге игроки I и II наблюдают текущее состояние игры $s_n = i \in S$, а затем независимо друг от друга выбирают соответственно решения $a_n = k \in A_i$ и $b_n = r \in B_i$. В результате игрок I выигрывает у игрока II сумму $h_i^{kr} = h(i, k, r)$, и игра с вероятностью $p_{ij}^{kr} = p(j \mid i, k, r)$ переходит в новое состояние $s_{n+1} = j \in S$. Предположим, что $T \to \infty$ и в качестве функции выигрыша выбран средний выигрыш игрока I за единицу времени — предельный средний выигрыш. Игрок I стремится максимизировать свой предельный средний выигрыш, в то время как игрок II стремится его минимизировать. Задача состоит в том, чтобы выбрать для игрока I стратегию, максимизирующую его предельный средний выигрыш, а для игрока II стратегию, минимизирующую предельный средний выигрыш игрока I.

Теория марковских игр, известная еще под названием стохастических игр, берет свое начало с работ Шепли [1], который рассматривал случай, когда игра заканчивается с вероятностью 1. Марковские игры с предельным средним выигрышем впервые были рассмотрены Джиллеттом [2]. Условия существования значения и оптимальных стационарных стратегий игроков I и II для полных сепарабельных метрических пространств S, $A \equiv A_i$ и $B \equiv B_i$, $i \in S$, получены в [3]. Марковские игры с одним эргодическим классом (при любых стратегиях игроков I и II все состояния игры принадлежат единственному эргодическому классу), разыгрывающиеся на конечных S, A_i и B_i, $i \in S$, рассмотрены в [4, 5].

Настоящая работа посвящена марковским играм общего вида, характеризующимся тем, что при любых стационарных стратегиях игроков множество состояний игры разбивается на несколько эргодических множеств и невозвратное множество, которые могут меняться в зависимости от стратегий игроков. Множества S, A_i и B_i, $i \in S$, предполагаются конечными.

Следует отметить, что общая марковская игра в книге Партхасаратхи и Рагхавана [6] описана как игра, не имеющая значения в классе стационарных стратегий. Для подтверждения этого в [6] приводится принадлежащий Джиллетту [2] пример, который под названием „большой матч" рассматривался также Блекуэллом [7]. Однако, в работе [8] показана несостоятельность этих доводов.

Наша цель — доказать, что марковская игра с конечными множествами состояний S и конечными множествами решений A_i и B_i, $i \in S$, имеет значение, а оба игрока — ε-оптимальные стационарные стратегии. Доказательство имеет конструктивный характер и приводит к рекуррентному алгоритму нахождения решения игры. Данная работа является обобщением результата, полученного в [8] для игры „большой матч".

2. Определения и обозначения. Пусть $S = \{1, 2, \ldots, N\}$, $A_i = \{1, 2, \ldots, k_i\}$, $i \in S$, и $B_i = \{1, 2, \ldots, r_i\}$, $i \in S$. Определим $\mathbb{S}^m$ как $(m-1)$-мерный симплекс

$$\mathbb{S}^m = \left\{ p = (p_1, p_2, \ldots, p_m) \in E^m \,\middle|\, \sum_{i=1}^{m} p_i = 1,\ p_i \geq 0,\ 1 \leq i \leq m \right\},$$

где E^m — m-мерное евклидово пространство. Положим $X_i = \mathbb{S}^{k_i}$, $Y_i = \mathbb{S}^{r_i}$, $X = X_1 \times X_2 \times \ldots \times X_N$ и $Y = Y_1 \times Y_2 \times \ldots \times Y_N$.

Предысторией ω^{n-1} марковской игры к моменту времени $n \geq 1$ называется последовательность $\xi_1, \xi_2, \ldots, \xi_{n-1}$. В каждый момент времени $n = 1, 2, \ldots$ игроки I и II независимо друг от друга выбирают свои решения k и r. Эти решения зависят от предыстории ω^{n-1} и текущего состояния $s_n = i \in S$ и определяются с помощью следующих распределений вероятностей:

$$x_n(i) = \left\{ x_i^k(n),\ 1 \leq k \leq k_i \right\} \in X_i, \qquad y_n(i) = \left\{ y_i^r(n),\ 1 \leq r \leq r_i \right\} \in Y_i,$$

где $x_i^k(n) = \mathrm{P}(a_n = k \mid \omega^{n-1},\ s_n = i)$, $y_i^r(n) = \mathrm{P}(b_n = r \mid \omega^{n-1},\ s_n = i)$.

Рандомизированной решающей функцией (на n-м шаге) игрока I называется набор распределений $x_n = \{x_n(i),\ 1 \leq i \leq N\} \in X$, а игрока II — $y_n = \{y_n(i),\ 1 \leq i \leq N\} \in Y$. Стратегией поведения игрока I называется последовательность решающих функций $\pi = (x_1, x_2, \ldots, x_n, \ldots)$, а игрока II — $\varphi = (y_1, y_2, \ldots, y_n, \ldots)$. Стратегия $\pi\,(\varphi)$ называется марковской стратегией, если вероятности $x_i^k(n)$, $\left(y_i^r(n)\right)$, $n \geq 1$, не зависят от предыстории ω^{n-1}, а зависят лишь от текущего состояния игры s_n. Марковские стратегии вида $x^\infty = (x, x, \ldots, x, \ldots)$ и $y^\infty = (y, y, \ldots, y, \ldots)$ называются стационарными.

Множества стратегий поведения, марковских стратегий и стационарных стратегий игрока I (II) обозначаются соответственно через $\tilde{\Pi}$, Π, X^∞ ($\tilde{\Phi}$, Φ, Y^∞). Классы стратегий поведения, марковских стратегий и стационарных стратегий определяются соответственно как $\Sigma_P = \tilde{\Pi} \times \tilde{\Phi}$, $\Sigma_M = \Pi \times \Phi$, $\Sigma_S = X^\infty \times Y^\infty$.

Выбор начального состояния $i \in S$ и стратегий $\pi \in \tilde{\Pi}$ и $\varphi \in \tilde{\Phi}$ определяет вероятностную меру $\mathrm{P}_{i,T}^{\pi,\varphi}$ в пространстве историй $\{\omega^T\}$ длины $T \geq 1$. Суммарный средний выигрыш игрока I за T шагов игры определяется функцией

$$w_i^T(\pi, \varphi) = \mathrm{M}_{i,T}^{\pi,\varphi} \sum_{n=1}^{T} h(\xi_n), \quad i \in S,$$

где $\mathrm{M}_{i,T}^{\pi,\varphi}$ — математическое ожидание, соответствующее $\mathrm{P}_{i,T}^{\pi,\varphi}$. Средний выигрыш игрока I за переход на бесконечном горизонте оценивается функцией

$$g_i(\pi, \varphi) = \lim_{T \to \infty} \inf \frac{w_i^T(\pi, \varphi)}{T}, \tag{1}$$

если игра начинается из состояния $i \in S$.

Чтобы сделать марковские игры пригодными для изучения, нужно ограничить класс рассматриваемых классов стратегий. Далее, как и в [6], ограничимся классом марковских стратегий Σ_M.

На n-м шаге игры при заданных марковских стратегиях $\pi \in \Pi$ и $\varphi \in \Phi$ вероятность $P_{ij}[x_n(i), y_n(i)]$ перехода из состояния $i \in S$ в состояние $j \in S$ равна

$$\sum_{k \in A_i} \sum_{r \in B_i} x_i^k(n) y_i^r(n) p_{ij}^{kr},$$

а ожидаемый $H_i[x_n(i), y_n(i)]$ выигрыш игрока I равен

$$\sum_{k \in A_i} \sum_{r \in B_i} x_i^k(n) y_i^r(n) h_i^{kr}.$$

При любых стратегиях π и φ игроков изучаемый процесс является, вообще говоря, неоднородной цепью Маркова, для которой матрица вероятностей перехода за n шагов имеет вид

$$P_n(\pi, \varphi) = P(x_1, y_1) P(x_2, y_2) \ldots P(x_n, y_n), \quad n \geq 1,$$

где $P_n(\pi, \varphi)$ — матрица перехода размера $N \times N$, (i, j)-й элемент которой равен $P_{ij}[x_n(i), y_n(i)]$. При $n = 0$ положим $P_0(\pi, \varphi) = I$ (единичная матрица размера $N \times N$). Решающим функциям x_n и y_n соответствует $(N \times 1)$-мерный вектор выигрышей $H(x_n, y_n) = \{H_i[x_n(i), y_n(i)],\ i \in S\}$. Теперь функцию выигрыша (1) можно представить в более конкретном виде — как N-мерный вектор-столбец

$$G(\pi, \varphi) = \lim_{T \to \infty} \inf \frac{1}{T} \sum_{n=0}^{T-1} P_n(\pi, \varphi) H(x_{n+1}, y_{n+1}),$$

i-й элемент $g_i(\pi, \varphi)$ которого соответствует i-му начальному состоянию игры.

Определение 1. *Тройка* $\Gamma = \langle \Pi, \Phi, G(\pi, \varphi) \rangle$ *называется марковской игрой с предельным средним выигрышем. Подыгра* $\gamma = \langle X^\infty, Y^\infty, G(X^\infty, Y^\infty) \rangle$ *игры* Γ *называется марковской игрой с предельным средним выигрышем в стационарном режиме.*

Верхнее и нижнее значения игры Γ определяются следующими N-мерными вектор-столбцами:

$$\overline{v}(\Gamma) = \inf_{\varphi \in \Phi} \sup_{\pi \in \Pi} G(\pi, \varphi), \qquad \underline{v}(\Gamma) = \sup_{\pi \in \Pi} \inf_{\varphi \in \Phi} G(\pi, \varphi).$$

Таким же образом определяются верхнее $\overline{v}(\gamma)$ и нижнее $\underline{v}(\gamma)$ значения игры $\gamma \subset \Gamma$.

Определение 2. *Игра* Γ *имеет значение* $v(\Gamma)$, *если* $v(\Gamma) = \overline{v}(\Gamma) = \underline{v}(\Gamma)$. *Игра* $\gamma \subset \Gamma$ *имеет значение* $v(\gamma)$, *если* $v(\gamma) = \overline{v}(\gamma) = \underline{v}(\gamma)$.

Поскольку рассматриваемая игра $\Gamma = \langle \Pi, \Phi, G(\pi, \varphi) \rangle$ является одной из разновидностей антагонистической игры в нормальной форме [9, с. 29], то справедливо следующее утверждение.

Теорема 1. *Для того чтобы существовало конечное значение игры* $\Gamma = \langle \Pi, \Phi, G(\pi, \varphi)\rangle$, *необходимо и достаточно, чтобы для любого числа* $\varepsilon > 0$ *существовали стратегии* π_ε *и* φ_ε *такие, что*

$$G(\pi_\varepsilon, \varphi) + \varepsilon \mathbf{1} \geq G(\pi_\varepsilon, \varphi_\varepsilon) \geq G(\pi, \varphi_\varepsilon) - \varepsilon \mathbf{1} \qquad (2)$$

для всех $\pi \in \Pi$ *и* $\varphi \in \Phi$, *причем*

$$\lim_{\varepsilon \to 0} G(\pi_\varepsilon, \varphi_\varepsilon) = v(\Gamma).$$

Здесь $\mathbf{1}$ — *N-мерный вектор-столбец, все элементы которого равны* 1.

Доказательство аналогично доказательству теоремы 1, приведенной в [9, с. 29].

Пара стратегий (π, φ) из Σ_M называется ситуацией. Стратегии π_ε и φ_ε, удовлетворяющие неравенству (2), называются ε-оптимальными стратегиями, а пара $(\pi_\varepsilon, \varphi_\varepsilon)$ — ситуацией ε-равновесия; 0-оптимальные стратегии, т. е. стратегии π_0, φ_0 такие, что

$$G(\pi_0, \varphi) \geq G(\pi_0, \varphi_0) \geq G(\pi, \varphi_0)$$

для всех $\pi \in \Pi$ и $\varphi \in \Phi$, называются оптимальными стратегиями. Пара (π_0, φ_0) называется ситуацией равновесия, или седловой точкой функционала G. Заметим, что значение игры Γ может существовать и тогда, когда оптимальные стратегии игроков I и II не существуют.

3. Функциональное уравнение преобразования цен. Рассмотрим марковскую игру, разыгрывающуюся конечное число шагов T. В этом случае функционал выигрыша игры характеризуется $(N \times 1)$-мерным вектором суммарных средних выигрышей игрока I

$$W^T(\pi, \varphi) = \sum_{n=0}^{T-1} P_n(\pi, \varphi) H(x_{n+1}, y_{n+1}). \qquad (3)$$

Здесь стратегии игроков $\pi \in \Pi$ и $\varphi \in \Phi$ удобно представить в виде $(\ldots, x_n, x_{n-1}, \ldots, x_1)$ и $(\ldots, y_n, y_{n-1}, \ldots, y_1)$, где x_n, y_n — рандомизированные решающие функции, принимаемые игроками I и II соответственно за n шагов до окончания игры.

Из (3) получаем рекуррентное соотношение

$$W^n(\pi, \varphi) = H(x_n, y_n) + P(x_n, y_n) W^{n-1}(\pi, \varphi), \quad n \geq 1, \qquad (4)$$

где $W^0(\pi, \varphi) = 0$.

Определение 3. *Стратегии игроков* π^* *и* φ^* *называются равномерно оптимальными, если для любой ситуации* $(\pi, \varphi) \in \Sigma_M$ *и любого* $n \geq 1$ *выполняется двойное неравенство*

$$W^n(\pi^*, \varphi) \geq W^n(\pi^*, \varphi^*) \geq W^n(\pi, \varphi^*).$$

Таким образом, если стратегии π^* и φ^* равномерно оптимальны при горизонте T, то они также оптимальны при горизонте T, но обратное утверждение неверно.

Для нахождения равномерно оптимальных стратегий игроков воспользуемся принципом оптимальности Беллмана [10] и принципом максимина фон Неймана [6, 11]. Исходя из (4) получаем так называемое функциональное уравнение преобразования цен [11]:

$$w_i^n = \operatorname{val}\left\| h_i^{kr} + \sum_{j=1}^{N} p_{ij}^{kr} w_j^{n-1} \right\|, \quad i \in S, \quad n = 1, 2, \ldots, \tag{5}$$

где $w_j^0 = 0$, $j \in S$; w_i^n — i-я компонента вектора $W^n(\pi^*, \varphi^*)$; $\operatorname{val}\|\cdot\|$ — значение (цена) игры с матрицей $\|\cdot\|$.

Поскольку решение матричной игры в большинстве случаев определяется неоднозначно, при использовании рекуррентного соотношения (5) будем придерживаться *правила выбора стратегий матричных игр*: при каждом $i \in S$, если смешанная стратегия $x_i(n-1) \in X_i$ (соответственно $y_i(n-1) \in Y_i$), определенная на $(n-1)$-м шаге, наряду с $x_i(n)$ (соответственно $y_i(n)$) является решением матричной игры $\|\cdot\|$, то будем полагать $x_i(n) = x_i(n-1)$ (соответственно $y_i(n) = y_i(n-1)$).

Таким образом, марковская игра с конечным числом шагов имеет решение в классе рандомизированных марковских стратегий. Оптимальные стратегии игроков π^* и φ^* и значение игры $W^T(\pi^*, \varphi^*)$ могут быть определены с помощью функционального уравнения преобразования цен (5).

Далее будет рассмотрена марковская игра Γ, продолжающаяся бесконечно долго. При переходе к игре с бесконечным числом шагов естественно рассматривать случай $T = \infty$ как предельный для игры с конечным числом шагов. В связи с этим будут исследованы асимптотические свойства функционала выигрыша $G_n(\pi, \varphi) = W^n(\pi, \varphi)/n$ и стратегий игроков $\pi = (\ldots, x_n, x_{n-1}, \ldots, x_1)$ и $\varphi = (\ldots, y_n, y_{n-1}, \ldots, y_1)$ при $n \to \infty$. Нам понадобятся следующие определения.

Определение 4. *Стратегии вида* $\pi_n = \left(x_n^\infty, x_{n-1}, \ldots, x_1\right)$ *и* $\varphi_n = \left(y_n^\infty, y_{n-1}, \ldots, y_1\right)$ *игроков* I *и* II *называются квазистационарными стратегиями.*

Определение 5. *Стратегия* $\pi = (\ldots, x_n, x_{n-1}, \ldots, x_1)$ *игрока* I *называется асимптотически стационарной, если* $x_n \to x \in X$ *при* $n \to \infty$.

Можно дать аналогичное определение для игрока II.

Определение 6. *Ситуация* $(\pi, \varphi) \in \Sigma_{\mathrm{M}}$ *называется асимптотически устойчивой, если существует предел* $G(\pi, \varphi) = \lim_{n \to \infty} G_n(\pi, \varphi)$.

Обозначим через Σ_{KS} класс квазистационарных стратегий, через Σ_{AS} класс асимптотически стационарных стратегий, через Σ_{AU} класс асимптотически устойчивых ситуаций. Заметим, что $\Sigma_{\mathrm{S}} \subset \Sigma_{\mathrm{KS}} \subset \Sigma_{\mathrm{AS}}$.

4. Вспомогательные результаты. Сначала приведем два определения.

Определение 7. *Расходящаяся последовательность* $\{c_n,\ n \geq 1\}$ *асимптотически периодична с периодом* $d > 1$, *если существует конечный предел* $\lim_{n \to \infty} c_{nd+l} = k_l$ *при каждом* $l = 0, 1, \ldots, d-1$.

Определение 8. *Расходящийся ряд* $\sum_{n=1}^{\infty} c_n$ *асимптотически периодичен с периодом* $d > 1$, *если частичные суммы* C_n, $n \geq 1$, *ряда образуют асимптотически периодичную последовательность с периодом* d.

В классе марковских стратегий Σ_{M} определим

$$\mathrm{W}^n((\pi, \varphi), U) = W^n(\pi, \varphi) + P(x_n, y_n) \ldots P(x_1, y_1) U,$$

где U — произвольный $(N \times 1)$-мерный вещественный вектор. Величину $\mathrm{W}^n((\pi, \varphi), U)$ можно интерпретировать как вектор суммарных средних выигрышей игрока I, получаемых за n шагов, при условии, что в момент окончания игры игроку I выплачивается сумма u_j, равная j-й компоненте вектора U, если при этом игра попадает в состояние $j \in S$. В частности, получаем $\mathbb{W}^n((\pi, \varphi), 0) = W^n(\pi, \varphi)$. Заметим, что если в рекуррентном соотношении (4) положить $W^0(\pi, \varphi) = U$, то на n-м шаге итерации получим величину $\mathbb{W}^n((\pi, \varphi), U)$. В дальнейшем нам потребуется следующая лемма.

Лемма 1. *Зафиксируем ситуацию* $(x^\infty, y^\infty) \in \Sigma_S$ *и рассмотрим марковскую цепь, задаваемую матрицей перехода* $P(x, y)$, *и последовательность* $\left\{\mathbb{W}^n\left(\left(x^\infty, y^\infty\right), U\right) - nG\left(x^\infty, y^\infty\right), n \geq 0\right\}$. *Тогда*:

а) *если цепь апериодична, т. е. если существует* $\lim_{n\to\infty} P^n(x, y)$, *то данная последовательность сходится*;

б) *если цепь неразложима и периодична с периодом* $d > 1$, *то данная последовательность асимптотически периодична с периодом* d.

Доказательство вытекает из результатов Ховарда и Брауна [12, с. 42, 89], полученных относительно управляемых марковских цепей.

Теперь приведем некоторые факты из теории матриц. Обозначим через $\mathcal{M}_N$ множество всех квадратных матриц порядка N, а через $\mathcal{P}_N$ множество всех неотрицательных матриц порядка N.

Определение 9. *Последовательность матриц* $\{P_n, n \geq 0\}$ *суммируема по Чезаро к* P^*, *если предел*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} P_m$$

существует и равен P^*.

Лемма 2. *Пусть* P *— стохастическая матрица размера* $N \times N$. *Тогда последовательность* $\left\{P^n, n \geq 0\right\}$ *суммируема по Чезаро к стохастической матрице* P^* *такой, что*

$$PP^* = P^*P = P^*P^* = P^*. \tag{6}$$

Доказательство см. в [13].

Лемма 3. *Для конечной марковской цепи, задаваемой матрицей перехода* P, *существует единственная стохастическая матрица* P^*, *удовлетворяющая равенствам* (6).

Доказательство. Возможны три случая.

1. Цепь эргодична (неразложима и апериодична). В этом случае существует единственная предельная матрица $P^+ = \lim_{n\to\infty} P^n$, состоящая из одинаковых строк $p = (p_1, \ldots, p_N)$, $p_j > 0$, $\sum_{j\in S} p_j = 1$ [14, с. 132]. С другой стороны, согласно теореме Коши [15] (теорема 1)

$$\frac{I + P + \ldots + P^n}{n+1} \to P^+$$

при $n \to \infty$, так что матрица $P^* \equiv P^+$ единственна.

2. Цепь неразложима и периодична с периодом $d > 1$. В этом случае класс S разбивается на d подклассов $S_1, S_2, \ldots, S_d$ так, что

$$S = \bigcup_{h=1}^{d} S_h, \quad S_l \cap S_h = \varnothing, \quad 1 \le l \ne h \le d,$$

и матрицу P^{nd} можно представить в виде [16]

$$P^{nd} = \begin{pmatrix} P_1^n & 0 & \ldots & 0 \\ 0 & P_2^n & \ldots & 0 \\ \hdotsfor{4} \\ 0 & 0 & \ldots & P_d^n \end{pmatrix},$$

где P_h^n, $1 \le h \le d$, — неразложимая и примитивная стохастическая матрица, элементами которой являются вероятности перехода за d шагов от состояний подкласса S_h в состояния того же подкласса, и порядок матрицы P_h совпадает с числом состояний, входящих в подкласс S_h. На основании изложенного в п. 1 существует единственная предельная матрица $P^+ = \lim_{n\to\infty} P^{nd}$ в виде

$$P^{+} = \begin{pmatrix} P_1^+ & 0 & \ldots & 0 \\ 0 & P_2^+ & \ldots & 0 \\ \hdotsfor{4} \\ 0 & 0 & \ldots & P_d^+ \end{pmatrix},$$

где матрица P_h^+, $1 \le h \le d$, составлена из одинаковых строк $p(h) = \left[p_j(h)\right]$, $j \in S_h$, $p_j(h) > 0$, $\sum_{j\in S} p_j(h) = 1$. Поскольку

$$P^* = \lim_{n\to\infty} \frac{I + P + \ldots + P^{(n+1)d-1}}{(n+1)d} =$$

$$= \lim_{n\to\infty} \left(\frac{I + P^d + P^{2d} + \ldots + P^{nd}}{n+1} \right) \left(\frac{I + P + \ldots + P^{d-1}}{d} \right) =$$

$$= \frac{P^+\left(I + P + \ldots + P^{d-1}\right)}{d},$$

матрица P^* единственна.

3. Цепь с несколькими эргодическими классами и невозвратным множеством. В этом случае матрицу перехода P можно представить в виде

$$P = \begin{pmatrix} P_1 & 0 & \ldots & 0 & 0 \\ 0 & P_2 & \ldots & 0 & 0 \\ \hdotsfor{5} \\ 0 & 0 & \ldots & P_r & 0 \\ L_1 & L_2 & \ldots & L_r & T \end{pmatrix},$$

где P_h, $1 \le h \le r$, — подматрицы, связанные с каждым эргодическим классом S_h, $1 \le h \le d$, соответственно, а оставшиеся состояния образуют множество S_{r+1} невозвратных состояний и характеризуются подматрицами L_h, $1 \le h \le d$, и T, причем

$$S = \bigcup_{h=1}^{r+1} S_h, \quad S_l \cap S_h = \varnothing, \quad 1 \le l \ne h \le r+1.$$

Тогда матрица P^* имеет вид

$$P^* = \begin{pmatrix} P_1^* & 0 & \ldots & 0 & 0 \\ 0 & P_2^* & \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & P_r^* & 0 \\ L_1^* & L_2^* & \ldots & L_r^* & 0 \end{pmatrix}.$$

Здесь согласно пп. 1, 2 подматрицы P_h^*, $1 \le h \le r$, единственны. Единственность подматриц L_h^*, $1 \le h \le r$, вытекает из следующих импликаций:

$PP^* = P^* \Rightarrow P_h P_h^* = P_h^*$, $L_h P_h^* + TL_h^* = L_h^* \Rightarrow L_h^* = (I-T)^{-1} L_h P_h^*$, $1 \le h \le r$.

Лемма доказана.

Определение 10. *Пусть* $\{P_n, n \ge 0\}$ *и* $\{Q_n, n \ge 0\}$ *— две произвольные матричные последовательности,* P_n, $Q_n \in \mathcal{M}_N$. *Новая последовательность* $\{R_n, n \ge 0\}$, *определяемая формулой*

$$R_n = \sum_{m=0}^{n} P_m Q_{n-m}, \tag{7}$$

называется сверткой последовательностей $\{P_n\}$ *и* $\{Q_n\}$, *а сумма* (7) *обозначается через* $R_n = P_n * Q_n$.

Имеют место следующие свойства операции свертки:

$$P_n * Q = (P_n * I)Q,$$

$$P_n * Q_n = Q_n * P_n,$$

$$P_n * (Q_n * D_n) = (P_n * Q_n) * D_n = P_n * Q_n * D_n,$$

$$P_n * (Q_n + D_n) = P_n * Q_n + P_n * D_n.$$

Лемма 4. *Пусть даны две матричные последовательности* $\{P_n, n \ge 0\}$ *и* $\{Q_n, n \ge 0\}$, $P_n \in \mathcal{M}_N$, $Q_n \in \mathcal{P}_N$. *Если* $P_n \to 0$, $Q_n \to 0$ *при* $n \to \infty$ *и выполняется условие*

$$\sum_{m=0}^{n} \|Q_m\| \le K, \quad K = \text{const}, \quad n \ge 0,$$

то $R_n = P_n * Q_n \to 0$ *при* $n \to \infty$. *Здесь* $\|Q_m\|$ *— норма матрицы* Q_m.

Доказательство непосредственно вытекает из теоремы Теплица и ее третьего следствия, приведенных в [15, с. 325 – 327] (теорема 2).

Лемма 5. *Пусть даны две матричные последовательности* $\{P_n, n \ge 0\}$ *и* $\{Q_n, n \ge 0\}$, $P_n \in \mathcal{M}_N$, $Q_n \in \mathcal{P}_N$. *Если* $\{P_n\}$ *сходится к* $P \in \mathcal{M}_N$, *а* $\{Q_n\}$ *суммируема по Чезаро к* $Q \in \mathcal{P}_N$, *то* $\{(P_n * Q_n)/(n+1)\}$ *сходится к* $PQ \in \mathcal{M}_N$.

Доказательство. Имеем

$$\frac{P_n * Q_n}{n+1} = \frac{(P_n - P) * Q_n}{n+1} + \frac{P(I * Q_n)}{n+1}.$$

Первое слагаемое справа стремится к 0 согласно лемме 4, если заменить в ней Q_n на $Q_n/(n+1)$. Второе же слагаемое стремится к PQ. Лемма доказана.

Теорема 2. *В классе асимптотически стационарных стратегий* Σ_{AS} *последовательность* $\{G_n(\pi, \varphi)\}$ *сходится к* $G(x^\infty, y^\infty)$.

Доказательство. Далее для краткости пару (x_n, y_n) обозначим через z_n, а пару (π, φ) — через ζ. По предположению $\zeta \in \Sigma_{\text{AS}}$, так что $z_n \to z = (x, y) \in$

$\in X \times Y$ при $n \to \infty$. Рассмотрим матричную последовательность

$$\mathbb{P}_n(\zeta) = P_n(\zeta) * I = \sum_{m=0}^{n} P_m(\zeta),$$

где $P_0(\zeta) = I$, $P_m(\zeta) = P(z_n)P(z_{n-1})\ldots P(z_{n-m+1})$, $m \geq 1$.

Предположим, что

$$\lim_{n\to\infty} \frac{\mathbb{P}_n(\zeta)}{n+1} = \mathbb{P}(\zeta)$$

существует. Тогда из соотношения

$$\frac{\mathbb{P}_n(\zeta)}{n+1} = \frac{I}{n+1} + \frac{n}{n+1} P(z_n) \frac{\mathbb{P}_{n-1}(\zeta)}{n}$$

вытекает, что при $n \to \infty$

$$\mathbb{P}(\zeta) = P(z)\mathbb{P}(\zeta). \tag{8}$$

С другой стороны, замечая, что

$$P_m(\zeta) = P_{m-1}(\zeta)P(z_{n-m+1}), \quad m \geq 1,$$

и полагая $P(z_0) = I$, можем записать

$$\mathbb{P}_n(\zeta) = I + \sum_{m=1}^{n+1} P_{m-1}(\zeta)P(z_{n-m+1}) - P_n(\zeta) =$$

$$= I - P_n(\zeta) + \sum_{m=0}^{n} P_m(\zeta)P(z_{n-m}).$$

Следовательно,

$$\frac{\mathbb{P}_n(\zeta)}{n+1} = \frac{I - P_n(\zeta)}{n+1} + \frac{P_n(\zeta) * P(z_n)}{n+1}.$$

Отсюда ввиду леммы 5 при $n \to \infty$ имеем

$$\mathbb{P}(\zeta) = \mathbb{P}(\zeta)P(z). \tag{9}$$

Теперь рассмотрим матричную последовательность вида

$$L_{n-m}(\zeta) = P_m(\zeta)\frac{\mathbb{P}_{n-m}(\zeta)}{n+1}, \quad 0 \leq m \leq n. \tag{10}$$

Заметим, что

$$L_n(\zeta) = \frac{\mathbb{P}_n(\zeta)}{n+1} \to P(\zeta), \quad n \to \infty.$$

Из (10) следует

$$\frac{1}{n+1}\sum_{m=0}^{n} L_{n-m}(\zeta) = \frac{P_n(\zeta) * [\mathbb{P}_n(\zeta)/(n+1)]}{n+1}.$$

Здесь вновь применяя лемму 5, при $n \to \infty$ получаем равенство

$$\mathbb{P}(\zeta) = \mathbb{P}(\zeta)\mathbb{P}(\zeta). \tag{11}$$

Равенства (8), (9) и (11) запишем в объединенном виде

$$P(z)\,\mathbb{P}(\zeta) = \mathbb{P}(\zeta)P(z) = \mathbb{P}(\zeta)\mathbb{P}(\zeta) = \mathbb{P}(\zeta). \tag{12}$$

Сопоставляя (12) с (6) и учитывая лемму 3, получаем

$$\mathbb{P}(\zeta) \equiv P^*(z). \tag{13}$$

Заметим, что полученные результаты сохраняют силу для любой частичной последовательности $\{\mathbb{P}_{n_k}(\zeta)/(n_k+1), k \geq 1\}$, так что предел $\mathbb{P}(\zeta) = \lim_{n\to\infty} \mathbb{P}_n(\zeta)/(n+1)$ существует.

Теперь осталось показать, что $G_n(\zeta) \to G(z^\infty)$ при $n \to \infty$, где $z^\infty = (x^\infty, y^\infty)$. Полагая $H(z_0) = 0$ и применяя еще раз лемму 5, имеем

$$G_n(\zeta) = \frac{1}{n+1}\sum_{m=0}^{n} P_m(\zeta)H(z_{n-m}) =$$

$$= \frac{P_n(\zeta) * H(z_n)}{n+1} \xrightarrow[n\to\infty]{} \mathbb{P}(\xi)H(z) = P^*(z)H(z) = G(z^\infty).$$

Теорема доказана.

***Следствие* 1.** $\Sigma_{\mathrm{AS}} \subset \Sigma_{\mathrm{AU}}$.

5. Существование значения общей марковской игры. Далее нам потребуется установить равномерную сходимость последовательности $\{G_n(\pi, \varphi)\}$ в классе квазистационарных стратегий Σ_{KS}. Для краткости функции на Σ_{M} будем записывать без аргументов π и φ. Вопрос о равномерной сходимости последовательности вектор-функций $\{G_n\}$ можно свести к вопросу о равномерной сходимости векторного функционального ряда

$$\sum_{n=1}^{\infty}(G_n - G_{n-1}), \qquad G_0 = 0,$$

для которого частичными суммами будут G_1, G_2, ...

Полагая $d_n = d_n(\pi, \varphi) = P_{n-1}(\pi, \varphi)H(x_n, y_n)$ и используя равенство $W^n - W^{n-1} = d_n$, имеем $nG_n - nG_{n-1} = d_n - G_{n-1}$, или

$$G_n - G_{n-1} = \frac{d_n - G_{n-1}}{n}, \qquad n \geq 1.$$

В связи с этим рассмотрим следующие векторные функциональные ряды:

$$\sum_{n=1}^{\infty}\frac{d_n - G_{n-1}}{n}, \tag{14}$$

$$\sum_{n=1}^{\infty}(d_n - G_{n-1}). \tag{15}$$

Заметим, что ряд (14) относится к классу рядов Дирихле [15] (теорема 2). Чтобы установить равномерную сходимость функционального ряда (14) в классе Σ_{KS}, необходимо доказать равномерную ограниченность частичных сумм $\{D_n, n \geq 1\}$ ряда (15) в данном классе.

Далее в качестве нормы вектора $a \in E^N$ выберем

$$\|a\| = \max_{1\leq i\leq N}|a_i|.$$

Лемма 6. *В классе асимптотически устойчивых ситуаций* Σ_{AU} *общий член* $c_n = n(G_n - G_{n-1})$ *ряда* (15) *стремится к нулю.*

Доказательство. Допустим сначала, что $h_i^{kr} > 0$ для всех i, k, r. Тогда $W^n > W^{n-1}$ для всех $n \geq 1$. Поскольку $c_n = W^n - W^{n-1} - W^{n-1}/(n-1)$ и существует предел $G_n \to G$ при $n \to \infty$, согласно теореме Штольца [15] (теорема 1) $c_n \to 0$ при $n \to \infty$.

В случае, когда среди выигрышей $\{h_i^{kr}\}$ существуют неположительные величины, прибавляя константу $C > 0$ к каждой величине h_i^{kr}, можно добиться выполнения неравенства $h_i^{kr} + C > 0$ для всех i, k, r. Тогда

$$c_n = (W^n + nC) - (W^{n-1} + (n-1)C) - \left(\frac{W^{n-1}}{n-1} + C\right) \to 0$$

при $n \to \infty$, так как $W^n + C > W^{n-1}$, и тем самым условие теоремы Штольца выполняется. Лемма доказана.

Лемма 7. *Пусть при некоторой ситуации* $(\pi, \varphi) \in \Sigma_{\mathrm{M}}$ *последовательность* $\{G_n(\pi, \varphi)\}$ *не сходится. Тогда частичные пределы этой последовательности расположены всюду плотно между их нижним и верхним пределами*:

$$G^-(\pi, \varphi) = \lim_{n\to\infty} \inf G_n(\pi, \varphi), \qquad G^+(\pi, \varphi) = \lim_{n\to\infty} \sup G_n(\pi, \varphi),$$

т. е. любое число из отрезка $\left[G^-(\pi, \varphi), G^+(\pi, \varphi)\right]$ *является частичным пределом данной последовательности.*

Доказательство вытекает из (14) и утверждения Феджера [17, с. 37] относительно числового ряда такого, что общий член этого ряда стремится к нулю, а сам ряд не сходится, но ограничен.

Наряду с (15) рассмотрим векторный функциональный ряд

$$\sum_{n=1}^{\infty} (d_n - G). \tag{16}$$

Частичная сумма $L_n = L_n(\pi, \varphi)$ данного ряда равна

$$L_n(\pi, \varphi) = W^n(\pi, \varphi) - nG(\pi, \varphi), \qquad n \geq 1.$$

Отсюда и из леммы 1 получаем такое следствие.

Лемма 8. *Зафиксируем ситуацию* $\left(x^\infty, y^\infty\right) \in \Sigma_{\mathrm{S}}$ *и рассмотрим марковскую цепь, задаваемую матрицей перехода* $P(x, y)$. *Тогда*:

а) *если цепь апериодична, то ряд* (16) *сходится*;

б) *если цепь неразложима и периодична с периодом* $d > 1$, *то ряд* (16) *асимптотически периодичен с периодом* d.

Лемма 9. *Утверждения леммы* 8 *относительно ряда* (16) *справедливы и для ряда* (15).

Доказательство. а) Положим $q_n = G_n - G_{n-1}$. В силу леммы 6 $n q_n \to 0$ при $n \to \infty$, так что найдется конечное положительное число M такое, что

$$M\mathbf{1} > n q_n > -M\mathbf{1} \tag{17}$$

для всех $n = 1, 2, \ldots$.

Имеем

$$\sum_{m=n+1}^{n+p} mq_m = (n+p)\left(G_{n+p} - G_n\right) + pG_n - \sum_{m=n}^{n+p-1} G_m. \tag{18}$$

Ввиду (17) для любого G_m, $n+p > m \geq n$, справедлива оценка

$$G_n + \frac{p}{n}M\mathbf{1} > G_m > G_n - \frac{p}{n}M\mathbf{1},$$

откуда, суммируя по m, находим

$$\frac{p^2}{n}M\mathbf{1} > pG_n - \sum_{m=n}^{n+p-1} G_m > -\frac{p^2}{n}M\mathbf{1}. \tag{19}$$

С другой стороны,

$$(n+p)\left(G_{n+p} - G_n\right) = (n+p)\left(G_{n+p} - G\right) - n(G_n - G) - p(G_n - G) =$$

$$= L_{n+p} - L_n - \frac{p}{n}L_n. \tag{20}$$

Теперь будем произвольно увеличивать n до бесконечности, а изменение p подчиним требованию, чтобы отношение p^2/n стремилось к наперед заданному числу $\varepsilon/M > 0$. Тогда левая и правая части неравенства (19) будут стремиться к пределу $\varepsilon\mathbf{1}$ и $-\varepsilon\mathbf{1}$ соответственно, а правая часть равенства (20) в силу леммы 8 стремится к нулю, так что согласно (18) для достаточно больших значений n

$$\left\| \sum_{m=n+1}^{n+p} mq_m \right\| < \varepsilon.$$

б) Имеем

$$\sum_{m=\nu+1}^{\nu+pd} mq_m = (\nu+pd)\left(G_{\nu+pd} - G_\nu\right) + pdG_\nu - \sum_{m=\nu}^{\nu+pd-1} G_m,$$

где $\nu = nd + l$. Неравенство

$$\left\| \sum_{m=\nu+1}^{\nu+pd} mq_m \right\| < \varepsilon$$

при каждом $l = 0, 1, \ldots, d-1$, для достаточно больших значений n устанавливается аналогично случаю а). Лемма доказана.

Лемма 10. *Зафиксируем ситуацию* $(\pi_m, \varphi_m) \in \Sigma_{KS}$ *и рассмотрим марковскую цепь, задаваемую матрицей перехода* $P(x_m, y_m)$. *Тогда:*

а) *если цепь апериодична, то ряд* (15) *сходится;*

б) *если цепь неразложима и периодична с периодом* $d > 1$, *то ряд* (15) *асимптотически периодичен с периодом* d.

Доказательство. Из доказательства леммы 9 видно, что для установления справедливости утверждения а) достаточно доказать сходимость ряда (16) в классе Σ_{KS}.

Учитывая теорему 2, имеем

$$L_{n+m}(\pi_m, \varphi_m) = W^{n+m}((\pi_m, \varphi_m), 0) - (n+m)G(\pi_m, \varphi_m) =$$

$$= W^n\left(\left(x_m^\infty, y_m^\infty\right), U\right) - nG\left(x_m^\infty, y_m^\infty\right),$$

где $U = W^m((\pi_m, \varphi_m), 0) - mG(\pi_m, \varphi_m)$. В силу леммы 1 последовательность $\left\{W^n\left(\left(x_m^\infty, y_m^\infty\right), U\right) - nG\left(x_m^\infty, y_m^\infty\right), n \ge 0\right\}$ сходится. Случай б) доказывается аналогично. Лемма доказана.

В дальнейшем важную роль играет следующая теорема.

Теорема 3. *В классе квазистационарных стратегий* $\Sigma_{KS} = \{\pi_m, \varphi_m \mid 1 \le \le m < \infty\}$ *сходимость* $\lim_{n\to\infty} G_n(\pi_m, \varphi_m) = G\left(x_m^\infty, y_m^\infty\right)$ *равномерна.*

Доказательство. Из леммы 10 следует, что для всех $(\pi_m, \varphi_m) \in \Sigma_{KS}$ частичные суммы ряда (15) удовлетворяют условию

$$\| D_n(\pi_m, \varphi_m) \| \le K, \quad K = \text{const}, \quad 1 \le m < \infty, \quad n \ge 1.$$

С другой стороны, последовательность $\{1/n, n \ge 1\}$, вне зависимости от $(\pi_m, \varphi_m) \in \Sigma_{KS}$, монотонно сходится к 0. Тогда согласно признаку Дирихле [15] (теорема 2) функциональный ряд (14) равномерно сходится в классе Σ_{KS}. Теорема доказана.

Лемма 11. *В игре* Γ *для каждого* $n \ge 1$ *существует такая ситуация* $(\pi_n, \varphi_n) \in \Sigma_{KS}$, *что при всех* $\pi \in \Pi$ *и* $\varphi \in \Phi$

$$G_n(\pi_n, \varphi) \ge G_n(\pi_n, \varphi_n) \ge G_n(\pi, \varphi_n).$$

Доказательство. Пусть $\pi^* = \left(\ldots, x_n^*, x_{n-1}^*, \ldots, x_1^*\right)$ и $\varphi^* = \left(\ldots, y_n^*, y_{n-1}^*, \ldots \ldots, y_1^*\right)$ — равномерно оптимальные стратегии игроков, полученные с помощью функционального уравнения преобразования цен (5). Стратегии π_n и φ_n, соответствующие условиям леммы, могут быть образованы из π^* и φ^* следующим образом:

$$x_m = \begin{cases} x_m^*, & \text{если } m \le n; \\ x_n^*, & \text{если } m > n, \end{cases} \qquad y_m = \begin{cases} y_m^*, & \text{если } m \le n; \\ y_n^*, & \text{если } m > n, \end{cases}$$

где $m = 1, 2, \ldots$. Лемма доказана.

Теорема 4. *Пусть множества* S, A_i *и* B_i, $i \in S$, *конечны. Тогда общая марковская игра* Γ *имеет значение и существуют* ε*-оптимальные стационарные стратегии для каждого игрока.*

Доказательство. Согласно лемме 11 для любых стационарных стратегий $x^\infty \in X^\infty$ и $y^\infty \in Y^\infty$ существуют квазистационарные стратегии $\pi_n \in \Pi$ и $\varphi_n \in \in \Phi$ такие, что

$$G_n\left(\pi_n, y^\infty\right) \ge G_n(\pi_n, \varphi_n) \ge G_n\left(x^\infty, \varphi_n\right), \quad n \ge 1.$$

Исходя из теоремы 3, для любого $\varepsilon > 0$ найдется такое $n < \infty$, что для всех $m \ge n$ и $\left(x^\infty, y^\infty\right) \in \Sigma_S$

$$G_m\left(\pi_n, y^\infty\right) + \varepsilon\mathbf{1} \ge G_m(\pi_n, \varphi_n) \ge G_m\left(x^\infty, \varphi_n\right) - \varepsilon\mathbf{1}.$$

Здесь переходя к пределу по m и используя теорему 2, имеем

$$G\left(x_n^\infty, y^\infty\right) + \varepsilon\mathbf{1} \ge G\left(x_n^\infty, y_n^\infty\right) \ge G\left(x^\infty, y_n^\infty\right) - \varepsilon\mathbf{1} \quad \forall \left(x^\infty, y^\infty\right) \in \Sigma_S.$$

Полагая в этих неравенствах $x_n^\infty = x_\varepsilon^\infty$ и $y_n^\infty = y_\varepsilon^\infty$, получаем неравенства

$$G\left(x_\varepsilon^\infty, y^\infty\right) + \varepsilon\mathbf{1} \ge G\left(x_\varepsilon^\infty, y_\varepsilon^\infty\right) \ge G\left(x^\infty, y_\varepsilon^\infty\right) - \varepsilon\mathbf{1} \quad \forall \left(x^\infty, y^\infty\right) \in \Sigma_S,$$

из которых, в свою очередь, ввиду теоремы 1 следует, что общая марковская игра Γ в классе стационарных стратегий Σ_S имеет значение

$$v(\gamma) = \lim_{\varepsilon \to 0} G\left(x_\varepsilon^\infty, y_\varepsilon^\infty\right). \tag{21}$$

Теперь покажем, что ε-оптимальные стационарные стратегии игроков x_ε^∞ и y_ε^∞ являются ε-оптимальными стратегиями и в классе марковских стратегий Σ_M.

Предположим, что последовательность $\left\{G_n(\pi^*, \varphi^*)\right\}$, соответствующая равномерно оптимальным стратегиям игроков π^* и φ^*, расходится, т. е. $G^+(\pi^*, \varphi^*) > G^-(\pi^*, \varphi^*)$. Тогда согласно лемме 7 из последовательности $\left\{G_n(\pi^*, \varphi^*)\right\}$ можно выделить такую частичную последовательность $\left\{G_{n_k}(\pi^*, \varphi^*), k \geq 1\right\}$, что

$$\lim_{k \to \infty} G_{n_k}(\pi^*, \varphi^*) = G^+(\pi^*, \varphi^*).$$

Иначе говоря, для любого $\varepsilon > 0$ найдется такое конечное k', что для всех $k \geq k'$

$$G_{n_k}\left(\pi_{n_k}, \varphi_{n_k}\right) + \frac{\varepsilon}{2}\mathbf{1} \geq G^+(\pi^*, \varphi^*) \geq G_{n_k}\left(\pi_{n_k}, \varphi_{n_k}\right) - \frac{\varepsilon}{2}\mathbf{1}. \tag{22}$$

С другой стороны, согласно теореме 3 число k' может быть выбрано так, что при каждом $k \geq k'$ для всех $m \geq n_k$

$$G_m\left(x_{n_k}^{*\infty}, y_{n_k}^{*\infty}\right) + \frac{\varepsilon}{2}\mathbf{1} \geq G_m\left(\pi_{n_k}, \varphi_{n_k}\right) \geq G_m\left(x_{n_k}^{*\infty}, y_{n_k}^{*\infty}\right) - \frac{\varepsilon}{2}\mathbf{1}.$$

Отсюда и из (22) следует

$$G_m\left(x_{n_k}^{*\infty}, y_{n_k}^{*\infty}\right) + \varepsilon\mathbf{1} \geq G^+(\pi^*, \varphi^*) \geq G_m\left(x_{n_k}^{*\infty}, y_{n_k}^{*\infty}\right) - \varepsilon\mathbf{1}.$$

Стратегии $x_{n_k}^{*\infty}$ и $y_{n_k}^{*\infty}$ можно использовать в качестве ε-оптимальных стационарных стратегий, так что если здесь положить $x_\varepsilon^\infty = x_{n_k}^{*\infty}$, $y_\varepsilon^\infty = y_{n_k}^{*\infty}$ и перейти к пределу при $m \to \infty$ и $\varepsilon \to \infty$, то в силу (21) получим $G^+(\pi^*, \varphi^*) = v(\gamma)$. Аналогичное рассуждение приводит к равенству $G^-(\pi^*, \varphi^*) = v(\gamma)$. Следовательно,

$$G^+(\pi^*, \varphi^*) = G^-(\pi^*, \varphi^*) = v(\gamma).$$

С другой стороны,

$$\lim_{n \to \infty} G_n(\pi^*, \varphi^*) = v(\Gamma).$$

Таким образом, $v(\Gamma) = v(\gamma)$. Теорема доказана.

Следствие 2. $\lim_{n \to \infty} G_n(\pi^*, \varphi^*) = v(\gamma)$.

Следствие 3. *Стратегии* $\pi^* = \left(\ldots, x_n^*, x_{n-1}^*, \ldots, x_1^*\right)$ *и* $\varphi^* = \left(\ldots, y_n^*, y_{n-1}^*, \ldots, y_1^*\right)$, *порожденные функциональным уравнением преобразования цен (5), являются равномерно оптимальными и асимптотически устойчивыми. Кроме того, если на каждом шаге итерации будет соблюдено правило выбора стратегий матричных игр, то они асимптотически стационарны, т. е. имеет место сходимость* $x_n^* \to x^*$ *и* $y_n^* \to y^*$ *при* $n \to \infty$.

Замечание. В случае, когда один из игроков пассивен, общая марковская игра сводится к марковскому процессу принятия решений с критерием средн выигрыша за единицу времени. В этом плане следствие 2 включает в себя частный случай результат Брауна (см. [12], следствие 1 теоремы 4.1), а след зие 3 — результаты В. В. Баранова (см. [18], теоремы 2 и 3).

6. Решение общей марковской игры. Из полученных результатов след что равномерно оптимальные стратегии игроков $\pi^* = \left\{x_n^*, n \geq 1\right\}$ и $= \left\{y_n^*, n \geq 1\right\}$ за конечное число итераций выходят на ситуацию ε-равнове классе стационарных стратегий Σ_S. Тем не менее, функциональное урав преобразования цен (5) нельзя использовать для решения общей марко игры, так как последовательность $\left\{W^n(\pi^*, \varphi^*)\right\}$ в большинстве случаев ется неограниченной. Здесь требуются иные методы последовательны олижений, приспособленных к критерию среднего выигрыша игрока I ду времени.

Рассмотрим рекуррентное соотношение (5). Разделим обе части зенства на n и представим его в виде

$$\frac{W^n(\pi, \varphi)}{n} = \frac{1}{n} H(x_n, y_n) + \frac{n-1}{n} P(x_n, y_n) \frac{W^{n-1}(\pi, \varphi)}{n-1}.$$

Отсюда, в свою очередь, получаем рекуррентное соотношение

$$G_n(\pi, \varphi) = \frac{1}{n} H(x_n, y_n) + \frac{n-1}{n} P(x_n, y_n) G_{n-1}(\pi, \varphi), \qquad n \geq 1,$$

$$G_0(\pi, \varphi) = 0.$$

В соответствии с (23) функциональное уравнение преобразования цен представить в виде

$$g_i^n = \operatorname{val} \left\| \frac{1}{n} h_i^{kr} + \frac{n-1}{n} \sum_{j=1}^{N} p_{ij}^{kr} g_j^{n-1} \right\|, \qquad i \in S,$$

$$g_j^0 = 0, \qquad j \in S.$$

Заметим, что в силу следствия 2 теоремы 4 данный итерационны сходится к значению игры $\gamma \subset \Gamma$. С его помощью ε-оптимальные ст стратегии игроков $x_\varepsilon^\infty = x_n^{*\infty}$ и $y_\varepsilon^\infty = y_n^{*\infty}$ достигаются за кон шагов n.

7. Пример: игра „большой матч”. Джиллетт [2], чтобы пок кую игру, не имеющую цену, приводит пример со следующими вероятностями и выигрышами, приведенными в таблице.

Состояние	Ситуация	Вероятность перехода		
i	(k, r)	p_{i1}^{kr}	p_{i2}^{kr}	p_{i3}^{kr}
1	(1, 1)	1	0	0
	(1, 2)	0	1	0
	(2, 1)	1	0	0
	(2, 2)	0	0	1
2	(1, 1)	0	1	0
3	(1, 1)	0	0	1

Согласно результатам Джиллетта

$$\min_{y^\infty}\max_{x^\infty} g_1\left(x^\infty, y^\infty\right) = 1, \qquad \max_{y^\infty}\min_{x^\infty} g_1\left(x^\infty, y^\infty\right) = \frac{1}{2},$$

т. е. игра в классе стационарных стратегий Σ_S не имеет значения, причем функция $g_1\left(x^\infty, y^\infty\right)$ разрывна по $y = \{y_1, y_2, y_3\}$ при $y_1 = (1, 0)$.

Эту игру можно сформулировать в терминах игры „большой матч" [7], которая состоит в следующем. Игроки I и II независимо друг от друга выбирают герб (Г) или решку (Р). Если игрок II выбирает герб, то он платит игроку I единицу в ситуации (Г, Г) и ничего не платит в ситуации (Р, Г), и игра начинается сначала. Если же игрок II выбирает решку, то в случае возникновения ситуации (Г, Р) игра заканчивается и никто ничего не платит, но если возникает ситуация (Р, Р), то игра продолжается и игрок II в каждой партии платит игроку I единицу. Игру „большой матч" можно представить тремя игровыми компонентами Γ_1, Γ_2 и Γ_3 следующим образом (по поводу представления марковских игр в виде игр-компонент см. [11]):

$$\Gamma_1 = \begin{bmatrix} 1+\Gamma_1 & \Gamma_2 \\ \Gamma_1 & \Gamma_3 \end{bmatrix}, \quad \Gamma_2 = [\Gamma_2], \quad \Gamma_3 = [1+\Gamma_3].$$

Применим рекуррентные соотношения (24) к играм-компонентам Γ_1, Γ_2, Γ_3:

$$g_1^n = \mathrm{val}\begin{bmatrix} \frac{1}{n}+\frac{n-1}{n}g_1^{n-1} & \frac{n-1}{n}g_2^{n-1} \\ \frac{n-1}{n}g_1^{n-1} & \frac{n-1}{n}g_3^{n-1} \end{bmatrix},$$

$$g_2^n = \frac{n-1}{n}g_2^{n-1},$$

$$g_3^n = \frac{1}{n}+\frac{n-1}{n}g_3^{n-1},$$

$$g_i^0 = 0, \qquad i = 1, 2, 3.$$

Из 2- и 3-го соотношений получаем $g_2^n = 0$, $g_3^n = 1$ для всех $n \geq 1$. Подставляя эти значения в 1-е соотношение, имеем

$$g_1^n = \mathrm{val}\begin{bmatrix} \frac{1}{n}+\frac{n-1}{n}g_1^{n-1} & 0 \\ \frac{n-1}{n}g_1^{n-1} & \frac{n-1}{n} \end{bmatrix}.$$

Решая стандартным путем [19, с. 45] эту (2×2)-игру, находим

$$g_1^n = \frac{n-1}{n}\left(\frac{1}{n}+\frac{n-1}{n}g_1^{n-1}\right), \tag{25}$$

$$x_1^1(n) = \frac{n-1}{n}\left(1-g_1^{n-1}\right), \tag{26}$$

$$y_1^1(n) = \frac{n-1}{n}. \tag{27}$$

Последовательно находя из (25)

$$g_1^1 = 0, \qquad g_1^2 = \frac{1}{2\cdot 2}, \qquad g_1^3 = \frac{2}{2\cdot 3}, \qquad g_1^4 = \frac{3}{2\cdot 4}, \ldots,$$

а затем применяя метод математической индукции, получаем

$$g_1^n = \frac{n-1}{2n}. \tag{28}$$

Отсюда и из (26) следует

$$x_1^1(n) = \frac{1}{2}. \tag{29}$$

При $n \to \infty$ (28) дает значения игры „большой матч" $g_1 = 1/2$, а (29) и (27) — оптимальных стационарных стратегий игроков: $x_1^* = (1/2, 1/2)$, $y_1^* = (1, 0)$. Поскольку $g_1((1,0),(1,0)) = 1$, стратегия $y_1^* = (1, 0)$ не может быть оптимальной, поэтому игроку II приходится воспользоваться ε-оптимальной стационарной стратегией — $y_1(\varepsilon) = (1-\varepsilon, \varepsilon)$.

Таким образом, игра „большой матч" имеет значение $v = (g_1, g_2, g_3) = (1/2, 0, 1)$. Игрок I может применять оптимальную стационарную стратегию $x^* = \left\{x_1^*, x_2^*, x_3^*\right\} = \{(1/2, 1/2), 1, 1\}$, а игрок II — ε-оптимальную стационарную стратегию — $y_\varepsilon = \{y_1(\varepsilon), y_2(\varepsilon), y_3(\varepsilon)\} = \{(1-\varepsilon, \varepsilon), 1, 1\}$.

В заключение заметим, что общая марковская игра (с несколькими эргодическими классами) в случае, когда множество состояний S и множества решений A_i и B_i, $i \in S$, конечны, имеет значение, а оба игрока — ε-оптимальные стационарные стратегии. Решение игры определяется с помощью функционального уравнения преобразования цен (24). Игра „большой матч" как представительница общей марковской игры имеет значение.

1. *Shapley L. S.* Stochastic games // Proc. Nat. Acad. Sci. USA. – 1953. – № 39. – P. 1095–1100.
2. *Gillette D.* Stochastic games with zero stop probabilities// Contrib. Theory Games (Ann. Math. Stud. – № 39.) – 1957. – 3. – P. 179–187.
3. *Губенко Л. Г.* О многошаговых стохастических играх // Теория вероятностей и мат. статистика. – 1973. – Вып. 8. – С. 35–49.
4. *Hoffman A. J., Karp R. M.* On nonterminating stochastic games // Manag. Sci. – 1966. – № 12. – P. 359–370.
5. *Ибрагимов А. А.* Итеративный метод решения марковской игры с одним эргодическим классом // Проблемы информатики и энергетики. – 1999. – № 4. – С. 10–14.
6. *Партхасаратхи Т., Рагхаван Т.* Некоторые вопросы теории игр двух лиц. – М.: Мир, 1974. – 296 с.
7 *Blackwell D.* The big match // SIAM J. Appl. Math. – 1970. – № 19. – P. 473 – 476.
8. *Ибрагимов А. А.* О марковской игре „большой матч" // Автоматика и телемеханика. – 2000. – № 11. – С. 104–113.
9. *Петросян Л. А., Томский Г. В.* Динамические игры и их приложения. – Л.: Изд-во Ленингр. ун-та, 1982. – 252 с.
10. *Беллман Р.* Динамическое программирование. – М.: Изд-во иностр. лит., 1960. – 400 с.
11. *Льюис Р. Д., Райфа Х.* Игры и решения. Введение и критический обзор. – М.: Изд-во иностр. лит., 1961. – 642 с.
12. *Майн Х., Осаки С.* Марковские процессы принятия решений. – М.: Наука, 1977. – 176 с.
13. *Кемени Дж., Снелл Дж.* Конечные цепи Маркова. – М.: Наука, 1970. – 272 с.
14. *Ширяев А. Н.* Вероятность. – М.: Наука, 1980. – 576 с.
15. *Фихтенгольц Г. М.* Курс дифференциального и интегрального исчисления. – М.: Наука, 1970. – Т. 1, 2.
16. *Сарымсаков Т. А.* Основы теории процессов Маркова. – Ташкент: Фан, 1988. – 248 с.
17. *Полиа Г., Сеге Г.* Задачи и теоремы из анализа. Часть первая: Ряды. Интегральное исчисление. Теория функций. – М.: Наука, 1978. – 392 с.
18. *Баранов В. В.* Модель и методы равномерно оптимального стохастического управления // Автоматика и телемеханика. – 1992. – № 5. – С. 42–52.
19. *Дюбин Г. Н., Суздаль В. Г.* Введение в прикладную теорию игр. – М.: Наука, 1981. – 336 с.

Получено 11.05.98,
после доработки — 25.09.2001