Math-Net.Ru

Общероссийский математический портал

В. К. Доманский, Рандомизированные оптимальные моменты для одного класса игр остановки, *Теория вероятн. и ее примен.*, 2001, том 46, выпуск 4, 770–779

DOI: <https://doi.org/10.4213/tvp3823>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.164.48.24

26 апреля 2023 г., 17:02:55

КРАТКИЕ СООБЩЕНИЯ

© 2001 г.

ДОМАНСКИЙ В. К.*

РАНДОМИЗИРОВАННЫЕ ОПТИМАЛЬНЫЕ МОМЕНТЫ ДЛЯ ОДНОГО
КЛАССА ИГР ОСТАНОВКИ¹⁾

В игровой задаче остановки в формулировке Дынкина ([1]) два игрока наблюдают последовательные состояния однородной цепи Маркова, причем каждый из них может остановить ее в любой момент. После остановки игра заканчивается и игрок 1 выигрывает у игрока 2 сумму, зависящую от того, кто из игроков остановил цепь и от состояния цепи в момент остановки.

В работе рассматриваются игры остановки, для которых множество состояний цепи — множество неотрицательных целых чисел. Из состояния $n > 0$ возможен либо переход в состояние $n + 1$, либо переход в поглощающее состояние 0 с нулевыми выигрышами — обрыв цепи. На выигрыши накладываются условия, приводящие к тому, что уравнения оптимальности не имеют решений в чистых стратегиях. Получены решения для этих игр с использованием рандомизированных моментов остановки. Качественные свойства решений определяются предельным поведением выигрышей.

Ключевые слова и фразы: цепь Маркова, марковский момент остановки, матричная игра, оптимальные стратегии, значение игры.

1. Введение. Игровые задачи остановки являются обобщениями задач оптимальной остановки случайных процессов (см., например, [2]) на случай, когда в принятии решения участвуют несколько лиц с несовпадающими интересами.

В этой работе рассматривается игровая задача остановки случайной последовательности в формулировке, восходящей к Дынкину [1], (см. также [3]–[7]). Два игрока наблюдают последовательные состояния $x_n, n = 1, 2, \dots$, однородной во времени цепи Маркова с пространством состояний X , на котором заданы выигрыши a_{11}, a_{12}, a_{21} , $c: X \rightarrow \mathbf{R}$. Каждый из игроков может остановить цепь в любой момент. После остановки игра заканчивается и игрок 1 получает от игрока 2 сумму, зависящую от того, кто из игроков остановил цепь и от состояния цепи в момент остановки. Если оба игрока останавливают цепь одновременно, то игрок 1 выигрывает у игрока 2 сумму $a_{11}(x)$, где x — состояние цепи в момент остановки. Если первым остановившим является только игрок 1 или только игрок 2, то игрок 1 выигрывает, а игрок 2 проигрывает $a_{12}(x)$ или $a_{21}(x)$ соответственно. Если ни один из игроков не останавливает цепь, то игрок 1 выигрывает $\lim_{n \rightarrow \infty} c(x_n)$. Цель игрока 1 состоит в максимизации математического ожидания своего выигрыша, а цель игрока 2 — в его минимизации.

* С.-Петербургский экономико-математический институт РАН, ул. Чайковского, 1, С.-Петербург, Россия; e-mail: victor@agd.stud.pu.ru

¹⁾ Исследование проводилось при поддержке Российского фонда фундаментальных исследований, проект 99-01-00811.

Чистыми стратегиями игроков для игр остановки являются обычные марковские моменты для наблюдаемой цепи, а смешанными стратегиями поведения — рандомизированные марковские моменты. Рандомизированный момент остановки τ может быть представлен системой измеримых функций $\tau_n: X^n \rightarrow [0, 1]$, выражающих условные вероятности остановить цепь на шаге $n = 1, 2, \dots$, если она не была остановлена ранее, и если наблюдались состояния $x_1, \dots, x_n$:

$$\tau_n(x_1, \dots, x_n) = \mathbf{P}\{\tau = n \mid \tau \geq n; x_1, \dots, x_n\}.$$

Если $\tau_n(x_1, \dots, x_n) = \tau(x_n)$, где $\tau: X \rightarrow [0, 1]$, то стратегия является стационарной.

Пусть T и Σ — множества смешанных стратегий игроков 1 и 2 соответственно. Обозначим $K_x(\tau, \sigma)$ ожидаемый выигрыш игрока 1 в игре с начальным состоянием $x_1 = x$ при применении игроками стратегий $\tau \in T$ и $\sigma \in \Sigma$:

$$K_x(\tau, \sigma) = \mathbf{E}_x(\mathbf{I}_{[\tau=\sigma<\infty]}a_{11}(x_\tau) + \mathbf{I}_{[\tau<\sigma]}a_{12}(x_\tau) + \mathbf{I}_{[\tau>\sigma]}a_{21}(x_\sigma) + \mathbf{I}_{[\tau=\sigma=\infty]} \lim_{n \rightarrow \infty} c(x_n)).$$

Игра имеет значение $v(x)$, если

$$\sup_T \inf_\Sigma K_x(\tau, \sigma) = \inf_\Sigma \sup_T K_x(\tau, \sigma) = v(x). \quad (1.1)$$

Игроки должны на каждом шаге сравнивать выигрыши в случае прекращения игры с ожидаемыми выигрышами в случае ее продолжения. Это приводит к тому, что значение игры остановки $v(x)$ должно удовлетворять уравнению оптимальности, выражающему принцип динамического программирования Беллмана в игровой формулировке. Для того, чтобы выписать это уравнение, рассмотрим двушаговую игру остановки, в которой, если ни один из игроков не останавливает цепь на первом шаге, то на втором шаге игра заканчивается и игрок 1 получает от игрока 2 плату $u(x_2)$, где u — фиксированная функция на X .

В этой игре каждый из игроков имеет только две чистых стратегии — «останавливать» и «не останавливать» на первом шаге, а смешанные (рандомизированные) стратегии игроков 1 и 2 соответственно определяются вероятностями t и s использования первой стратегии. Если мы положим

$$a_{22}(x, u) = \mathbf{E}_x u = \mathbf{E}[u(x_2) \mid x_1 = x],$$

то игра с начальным состоянием $x_1 = x$ сводится к матричной игре с матрицей выигрышей $[a_{ij}(x, u)]$ размера 2×2 , и значение двушаговой игры остановки равно значению матричной игры:

$$\text{val} \left\| \begin{array}{cc} a_{11}(x) & a_{12}(x) \\ a_{21}(x) & a_{22}(x, u) \end{array} \right\| = \text{val}[a_{ij}(x, u)].$$

Значение $\text{val}[a_{ij}]$ матричной игры с матрицей выигрышей $[a_{ij}]$ размера 2×2 определяется как $\text{val}[a_{ij}] = \max_{t \in [0, 1]} \min_{s \in [0, 1]} (tsa_{11} + ts'a_{12} + t'sa_{21} + t's'a_{22})$, $t' = 1 - t$, $s' = 1 - s$.

Рассмотрим оператор T , действующий на множестве функций u на X согласно формуле $(Tu)(x) = \text{val}[a_{ij}(x, u)]$. Принцип динамического программирования Беллмана в приложении к игровой задаче остановки приводит к тому, что ее значение $v(x)$, задаваемое формулой (1.1), должно являться неподвижной точкой оператора T . Уравнение оптимальности имеет вид

$$v(x) = (Tv)(x) = \text{val} \left\| \begin{array}{cc} a_{11}(x) & a_{12}(x) \\ a_{21}(x) & a_{22}(x, v) \end{array} \right\|. \quad (1.2)$$

Отметим, что функция c , задающая «выигрыш в бесконечности», не учитывается уравнением (1.2). С другой стороны, неподвижная точка оператора T , вообще говоря, неединственна, и различным функциям c могут соответствовать различные решения уравнения (1.2).

Матричная игра имеет седловую точку (i^*, j^*) , $i^*, j^* = 1, 2$, — решение в чистых (нерандомизированных) оптимальных стратегиях $t^*, s^* = i^*, j^* \pmod{2}$, если $a_{i^*j^*} \leq a_{i^*j^*} \leq a_{i^*j^*}$, для $i \neq i^*, j \neq j^*$. В этом случае $\text{val}[a_{ij}] = a_{i^*j^*}$. Легко проверяется, что, если $a_{11} + a_{22} - a_{12} - a_{21} = 0$, то игра имеет седловую точку.

Если седловой точки нет, то

$$\text{val}[a_{ij}] = \frac{\det[a_{ij}]}{a_{11} + a_{22} - a_{12} - a_{21}}. \quad (1.3)$$

Смешанные оптимальные стратегии определяются соотношениями

$$t^* = \frac{a_{22} - a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}}; \quad s^* = \frac{a_{22} - a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}}. \quad (1.4)$$

Формулы (1.3) и (1.4) остаются верными также и в «пограничном» случае, когда один из игроков имеет единственную чистую выравнивающую оптимальную стратегию, а другой — определяемую формулой (1.4) смешанную выравнивающую оптимальную стратегию. Последняя не является единственной оптимальной стратегией. В этом случае, для $i \neq i^*$, $j \neq j^*$, либо $a_{ij} < a_{i^*j} = a_{i^*j^*} \leq a_{ij^*}$, и i^* — единственная чистая оптимальная стратегия игрока 1, либо $a_{i^*j} \leq a_{i^*j^*} = a_{ij^*} < a_{ij}$, и j^* — единственная чистая оптимальная стратегия игрока 2. В обоих случаях $a_{11} + a_{22} - a_{12} - a_{21} \neq 0$, и

$$\text{val}[a_{ij}] = a_{i^*j^*} = \frac{\det[a_{ij}]}{a_{11} + a_{22} - a_{12} - a_{21}}.$$

В модели Дынкина на каждом шаге лишь один из игроков может остановить цепь. Задается разбиение $X = X_1 \cup X_2$ пространства состояний X , и игрок 1 может остановить цепь на шаге n , если $x_n \in X_1$, а игрок 2 — если $x_n \in X_2$. Если игра продолжается бесконечно игрок 1 выигрывает нуль. Эта модель может быть сведена к вышеописанной, если положить $a_{21}(x) = \infty$ для $x \in X_1$, $a_{12}(x) = -\infty$ для $x \in X_2$ и $c(x) = 0$. Естественно, в этом случае, решение уравнения (1.2) не требует привлечения смешанных стратегий.

Позднее класс рассматриваемых игр остановки был расширен за счет допущения одновременной остановки обоими игроками (см. [5], [6]). В этих работах предполагалось, что выигрыши удовлетворяют соотношениям

$$a_{12}(x) \leq a_{11}(x) \leq a_{21}(x). \quad (1.5)$$

Эти неравенства гарантируют разрешимость уравнения оптимальности с использованием только чистых стратегий. А именно, для любой функции u матрица $[a_{ij}(x, u)]$ имеет седловую точку либо в (2,1), либо в (1,2), либо в (2,2). Вследствие этого такие игры имеют решения в чистых (нерандомизированных) моментах остановки, причем, игроки никогда не останавливают цепь одновременно.

Существует обширная литература по играм Дынкина, как с дискретным, так и с непрерывным временем (см. [8]–[10]), дающая достаточные условия для существования значения игры. Во всех этих работах также предполагалось, что выигрыши удовлетворяют соотношениям (1.5).

В работах [11] и [12] снимается предположение (1.5), но, как и ранее, если игра продолжается до бесконечности, то игрок 1 получает нуль. Это позволяет построить значения для рассматриваемых игр и смешанные ε -оптимальные стратегии игроков как пределы значений и ε -оптимальных стратегий для игр с дисконтированными выигрышами.

В данной работе рассматриваются игры остановки для очень простой цепи Маркова. Множество состояний — множество неотрицательных целых чисел $X = \{0, 1, 2, \dots\}$. Из состояния $x > 0$ возможен либо переход в состояние $x + 1$, либо переход в поглощающее состояние 0 с нулевыми выигрышами (обрыв цепи).

Простая структура переходов обуславливает столь же простую структуру чистых стратегий (нерандомизированных марковских моментов) для этой цепи. Наблюдения за цепью не дают игрокам никакой информации. Чистые стратегии игроков определяются номером шага, на котором игрок останавливает цепь, вне зависимости от ее состояния.

В работе опускаются как предположения (1.5), так и предположение нулевого «выигрыша в бесконечности». При $x > 0$ на выигрыши накладываются условия $a_{11}(x) > a_{21}(x) \geq a_{12}(x) = 0$, приводящие к тому, что уравнения оптимальности не имеют решений в чистых стратегиях. Функция c принимает два значения: $c(0) = 0$, $c(x) = d > 0$ при $x > 0$.

Мы рассматриваем «выигрыш в бесконечности» d как переменный параметр и ищем решения для двухпараметрического семейства игр $\Gamma_d(x)$, параметризованных начальными состояниями $x_1 = x$ и числами $0 < d \leq \infty$. Мы получаем для этих задач решения в явном виде с использованием рандомизированных моментов остановки. Качественные свойства решений определяются предельным поведением выигрышей $a_{11}(n)$, $a_{21}(n)$ и «выигрышем в бесконечности» d .

Отметим также, что существуют иные подходы к игровой формулировке задачи об оптимальной остановке. В одной из постановок игровой задачи один из участников выбирает момент остановки, а второй из заданного класса распределений выбирает распределение, определяющее наблюдаемый случайный процесс (см., например, [13] и [14]). В другой постановке игроки независимо наблюдают различные случайные процессы, и выигрыши зависят от моментов остановки выбранных каждым из игроков (см., например, [15], [16] и [17]).

2. Модель и уравнения оптимальности. Для рассматриваемого класса игровых задач множество состояний — множество неотрицательных целых чисел $X = \{0, 1, 2, \dots\}$. Из состояния $x > 0$ возможен либо переход в состояние $x + 1$ с вероятностью $p(x)$, либо переход в поглощающее состояние 0 (обрыв цепи) с вероятностью $1 - p(x)$. Поглощающее состояние 0 имеет нулевые выигрыши $a_{11}(0) = a_{21}(0) = a_{12}(0) = d(0) = 0$.

При $x > 0$ выигрыши удовлетворяют следующим соотношениям:

$$a_{11}(x) > a_{21}(x) \geq a_{12}(x) = 0; \quad d \in (0, \infty]. \quad (2.1)$$

В дальнейшем оказывается удобным отождествлять состояния и номера шагов. Мы обозначаем $\Gamma_d(x)$ игру, начинающуюся в момент $x > 0$ из состояния x с «выигрышем в бесконечности» d . Вероятности, связанные с соответствующей цепью, обозначаются P_x . Так $P_x(k) = P_x\{x_k = k\} = \prod_{l=x}^{k-1} p(l)$ обозначает вероятность достичь состояния k , начиная из состояния x .

Простая структура переходов обуславливает простую структуру чистых стратегий (нерандомизированных марковских моментов) для этой цепи. Наблюдения за цепью не дают игрокам никакой информации. Чистые стратегии игроков в играх $\Gamma_d(x)$ определяются номером шага $k = x, x+1, \dots, \infty$, на котором игрок останавливает цепь.

Смешанные стратегии τ и σ игроков 1 и 2 могут быть представлены последовательностями $(\tau(k))$ и $(\sigma(k))$, $k = x, x+1, \dots$, где $\tau(k), \sigma(k) \in [0, 1]$. Здесь $\tau(k)$ ($\sigma(k)$) — условная вероятность для игрока 1 (для игрока 2) остановить цепь на шаге k , если она не была остановлена им раньше (независимо от поведения цепи). Таким образом, для $k \geq x$

$$\begin{aligned} P_x\{\tau = k\} &= \prod_{l=x}^{k-1} (1 - \tau(l)) \cdot \tau(k); \\ P_x\{x_\tau = k\} &= \prod_{l=x}^{k-1} p(l)(1 - \tau(l)) \cdot \tau(k) = P_x\{\tau = k\} \cdot P_x(k). \end{aligned}$$

Мы полагаем, что выполнено условие $P_1(\infty) = \prod_{l=1}^{\infty} p(l) > 0$, т.е. цепь с начальным состоянием $x = 1$ с положительной вероятностью никогда не оборвется. Отсюда следует, что для любого начального состояния $x > 0$ вероятность обрыва цепи меньше единицы. Отметим, что, если цепь обрывается с вероятностью 1, то стратегия $\sigma = \infty$, предписывающая продолжение наблюдений с вероятностью 1 в любом состоянии $x > 0$, обеспечивает игроку 2 проигрыш 0.

Функция выигрыша игры $\Gamma_d(x)$, соответствующая чистой стратегии $k \in [x, \infty)$ игрока 1 и стратегии σ игрока 2, равна

$$K_{x,d}(k, \sigma) = \sum_{l=x}^{k-1} P_x\{x_\sigma = l\} a_{21}(l) + P_x\{x_\sigma = k\} a_{11}(k).$$

Функция выигрыша игры $\Gamma_d(x)$, соответствующая чистой стратегии $k \in [x, \infty)$ игрока 2 и стратегии τ игрока 1, равна

$$K_{x,d}(\tau, k) = P_x\{x_\tau = k\} a_{11}(k) + P_x\{x_k = k; \tau > k\} a_{21}(k).$$

Если игрок 2 использует чистую стратегию $\sigma = \infty$, то он обеспечивает себе проигрыш 0 против любой стратегии игрока 1, останавливающей цепь с вероятностью 1. Лучшим ответом игрока 1 на стратегию $\sigma = \infty$ является стратегия $\tau = \infty$, с выигрышем $\mathbf{P}_x(\infty) \cdot d$. Таким образом, для значения $v_d(x)$ игры $\Gamma_d(x)$ выполняются неравенства $v_d(x) \leq \mathbf{P}_x(\infty) \cdot d$.

Уравнение оптимальности (1.2) для игры $\Gamma_d(x)$ имеет вид

$$v_d(x) = \text{val} \left\| \begin{array}{cc} a_{11}(x) & 0 \\ a_{21}(x) & p(x) \cdot v_d(x+1) \end{array} \right\|. \quad (2.2)$$

Если выполняются неравенства

$$p(x) \cdot v_d(x+1) \geq a_{21}(x), \quad (2.3)$$

то матрица $[a_{ij}(x, v_d)]$ не имеет единственной седловой точки, и согласно (1.3) уравнения оптимальности могут быть записаны как

$$v_d(x) = \frac{a_{11}(x) \cdot p(x) \cdot v_d(x+1)}{a_{11}(x) - a_{12}(x) + p(x) \cdot v_d(x+1)}. \quad (2.4)$$

Вероятности $\tau_d(x)$ и $\sigma_d(x)$ использования первых стратегий в оптимальных стратегиях игроков в игре с матрицей выигрышей $[a_{ij}(x, v_d)]$ удовлетворяют соотношениям:

$$\tau_d(x) = 1 - v_d(x) \cdot (p(x) v_d(x+1))^{-1};$$

$$\sigma_d(x) = v_d(x) \cdot (a_{11}(x))^{-1}.$$

В противном случае, если $p(x) \cdot v_d(x+1) < a_{21}(x)$, то пара стратегий (2,2) является единственной седловой точкой. Уравнения оптимальности приобретают вид:

$$v_d(x) = p(x) \cdot v_d(x+1).$$

3. Решения для игр с нулевыми выигрышами $a_{21}(x)$. Обозначим $\Gamma_{0,d}(x)$ игру, для которой $a_{21}(x) = a_{12}(x) = 0$.

Для игр $\Gamma_{0,d}(x)$ неравенства (2.3) выполняются всегда и уравнения оптимальности для их значений $v_{0,d}(x)$ при всех $x > 0$ могут быть представлены в виде

$$u(x) = \frac{a_{11}(x) \cdot p(x) \cdot u(x+1)}{a_{11}(x) + p(x) \cdot u(x+1)}. \quad (3.1)$$

Вероятности $\tau(x, u)$ и $\sigma(x, u)$ использования первых стратегий в оптимальных стратегиях игроков в игре с матрицей выигрышей $[a_{ij}(x, u)]$ удовлетворяют соотношениям:

$$\tau(x, u) = \sigma(x, u) = u(x) \cdot (a_{11}(x))^{-1}; \quad (3.2)$$

$$1 - \tau(x, u) = 1 - \sigma(x, u) = u(x) \cdot (p(x))^{-1} (u(x+1))^{-1}.$$

Отметим, что функция $u_0(x) = 0$ при всех $x > 0$, удовлетворяет уравнению (3.1), однако стационарные стратегии $\tau = \infty$ и $\sigma = \infty$, порождаемые $\tau(x, u_0) = \sigma_0(x, u_0) = 0$, дают выигрыш $\mathbf{P}_x(\infty) \cdot d > 0$.

Общее решение функционального уравнения (3.1) дается формулой

$$w_{0,d}(x) = \left(S_0(x) + (\mathbf{P}_x(\infty) \cdot d)^{-1} \right)^{-1}, \quad (3.3)$$

где

$$S_0(x) = \sum_{k=x}^{\infty} (\mathbf{P}_x(k))^{-1} (a_{11}(k))^{-1}. \quad (3.4)$$

Ввиду того, что для любого начального состояния $x > 0$ и состояния $k > x$ выполняются неравенства $\mathbf{P}_x(k) \geq \mathbf{P}_1(\infty) > 0$, сходимость рядов, определяющих $S_0(x)$, равносильна сходимости ряда

$$\sum_{k=1}^{\infty} (a_{11}(k))^{-1}. \quad (3.5)$$

Теорема 3.1. Если ряд (3.5) расходится, то значение $v_{0,d}(x)$ игры $\Gamma_{0,d}(x)$ равно нулю при всех d . Игрок 2 не имеет оптимальной стратегии. Для любого $\varepsilon > 0$ стратегия $\sigma_{\varepsilon,x}$, для которой при $k \geq x$

$$\sigma_{\varepsilon,x}(k) = \min \left(1, \varepsilon \cdot (\mathbf{P}_x(k))^{-1} (a_{11}(k))^{-1} \right)$$

обеспечивает игроку 2 проигрыш не превышающий ε .

Доказательство. Для чистой конечной стратегии k игрока 1, $x \leq k < \infty$, функция выигрыша не превышает ε :

$$K_{x,d}(k, \sigma_{\varepsilon,x}) = \mathbf{P}_x\{x_{\sigma_{\varepsilon,x}} = k\} a_{11}(k) \leq \sigma_{\varepsilon,x}(k) \mathbf{P}_x(k) a_{11}(k) \leq \varepsilon.$$

Покажем, что вероятность того, что процесс не будет остановлен стратегией $\sigma_{\varepsilon,x}$, равна нулю. Действительно,

$$\mathbf{P}_x\{\sigma_{\varepsilon,x} = \infty\} \leq \prod_{k=x}^{\infty} \left(1 - \varepsilon \cdot (\mathbf{P}_x(k))^{-1} \cdot (a_{11}(k))^{-1} \right) = 0,$$

поскольку ряды (3.4) расходятся. Поэтому функция выигрыша $K_{x,d}(\infty, \sigma_{\varepsilon,x}) = 0$. Таким образом, стратегия $\sigma_{\varepsilon,x}$ обеспечивает игроку 2 общий проигрыш, не превышающий ε против любой чистой стратегии игрока 1, что доказывает теорему.

Если ряд (3.5) сходится, то функция $S_0(x)$, определенная формулой (3.4), конечна, и функция $w_{0,d}(x)$, определена формулой (3.3) для всех $d \in (0, \infty]$.

Теорема 3.2. Пусть ряд (3.5) сходится. Тогда при любом $d \in (0, \infty]$ значение $v_{0,d}(x)$ игры $\Gamma_{0,d}(x)$ равно $w_{0,d}(x)$.

Если $0 < d < \infty$, то стационарные оптимальные стратегии игроков в игре $\Gamma_{0,d}(x)$ определяются оптимальными стратегиями уравнения оптимальности: $\tau_{0,d}(k) = \sigma_{0,d}(k) = w_{0,d}(k) \cdot (a_{11}(k))^{-1}$ при $k \geq x$.

В игре $\Gamma_{0,\infty}(x)$ игрок 1 не имеет оптимальной стратегии. Стратегия $\tau_{0,d}$ при достаточно большом d является ε -оптимальной стратегией.

Оптимальная стратегия игрока 2 в игре $\Gamma_{0,\infty}(x)$ определяется соотношениями

$$\sigma_{0,\infty}(k) = w_{0,\infty}(k) \cdot (a_{11}(k))^{-1} \quad \text{при } k \geq x.$$

Доказательство. Непосредственно проверяется, что функция $w_{0,d}(x)$ при любом $d \in (0, \infty]$ удовлетворяет уравнению (3.1).

Ввиду (3.2) для стратегий $\tau_{0,d}$ и $\sigma_{0,d}$ при любом $0 < d \leq \infty$ и при $1 \leq k < \infty$ выполняются соотношения

$$1 - \tau_{0,d}(k) = 1 - \sigma_{0,d}(k) = w_{0,d}(k) \cdot (p(k))^{-1} (w_{0,d}(k+1))^{-1}.$$

Поэтому при $x \leq k < \infty$

$$\mathbf{P}_x\{\tau_{0,d} = k\} = \mathbf{P}_x\{\sigma_{0,d} = k\} = w_{0,d}(x) \cdot (\mathbf{P}_x(k))^{-1} (a_{11}(k))^{-1};$$

$$\mathbf{P}_x\{\tau_{0,d} = \infty\} = \mathbf{P}_x\{\sigma_{0,d} = \infty\} = w_{0,d}(x) \cdot (\mathbf{P}_x(\infty))^{-1} d^{-1}.$$

Таким образом, для любой конечной чистой стратегии k , $x \leq k < \infty$, при любом d

$$K_{x,d}(\tau_{0,d}, k) = K_{x,d}(k, \sigma_{0,d}) = w_{0,d}(x).$$

Если d конечно, то и $K_{x,d}(\tau_{0,d}, \infty) = K_{x,d}(\infty, \sigma_{0,d}) = w_{0,d}(x)$. Отсюда следует, что $w_{0,d}(x)$ является значением игры $\Gamma_{0,d}(x)$ для $0 < d < \infty$, а $\tau_{0,d}$ и $\sigma_{0,d}$ — оптимальными выравнивающими стратегиями игроков.

С другой стороны, поскольку $\mathbf{P}(\tau_{0,\infty} = \infty) = 0$, функция выигрыша $K_{x,\infty}(\tau_{0,\infty}, \infty)$ также равна нулю и стратегия $\tau_{0,\infty}$ не гарантирует игроку 1 выигрыш $w_{0,\infty}(x)$. Тем не менее, поскольку $\lim_{d \rightarrow \infty} w_{0,d}(x) = w_{0,\infty}(x)$, стратегия $\tau_{0,d}$ при достаточно большом d является ε -оптимальной стратегией. Теорема доказана.

Пример 3.1. Рассмотрим игры $\Gamma_{0,d}(x)$, для которых $p(k) = 1$, то есть вероятность обрыва цепи равна нулю, $a_{11}(k) = k(k+1)$. Тогда $S_0(x) = \sum_{k=x}^{\infty} (k(k+1))^{-1} = 1/x$. Значение $v_{0,d}(x)$ игры $\Gamma_{0,d}(x)$ равно $w_{0,d}(x) = xd/(x+d)$. Оптимальные стратегии задаются соотношениями $\tau_{0,d}(k) = \sigma_{0,d}(k) = d/[(k+1)(k+d)]$ при $k \geq x$. Вероятность того, что цепь не будет остановлена стратегией $\tau_{0,d}$ равна $\prod_{k=x}^{\infty} 1 - d/((k+1)(k+d)) = x/(x+d) = w_{0,d}(x) \cdot d^{-1}$.

В частности, $v_{0,1}(x) = x/(x+1)$, $\tau_{0,1}(k) = \sigma_{0,1}(k) = (k+1)^{-2}$.

Значение игры $\Gamma_{0,\infty}(x)$ равно x , а оптимальные стратегии задаются соотношениями $\tau_{0,\infty}(k) = \sigma_{0,\infty}(k) = (k+1)^{-1}$.

Вероятность того, что цепь не будет остановлена стратегией $\tau_{0,\infty}$ равна

$$\prod_{k=x}^{\infty} \left(1 - \frac{1}{k+1}\right) = 0,$$

и $\tau_{0,\infty}$ не является оптимальной стратегией. Стратегия $\tau_{0,d}$ при $x - xd/(x+d) < \varepsilon$ ε -оптимальна для игры $\Gamma_{0,\infty}(x)$.

4. Решения для игр с положительными выигрышами $a_{21}(x)$. В этом пункте мы находим решения для игр $\Gamma_d(x)$ в случае, когда для всех $k \geq x$ игра с входящей в уравнение (2.3) матрицей выигрышей $[a_{ij}(x, v_d)]$ имеет решение в смешанных оптимальных стратегиях игроков.

Допустим, что при всех $x > 0$ выполняются неравенства (2.4). Тогда значение $v_d(x)$ игры $\Gamma_d(x)$ при всех $x > 0$ удовлетворяет уравнениям оптимальности (2.5). Функция $v_d(x)$, $x > 0$, должна являться решением функционального уравнения

$$u(x) = \frac{a_{11}(x) \cdot p(x) \cdot u(x+1)}{a_{11}(x) - a_{21}(x) + p(x) \cdot u(x+1)}. \quad (4.1)$$

Общее решение этого функционального уравнения дается формулой

$$w_d(x) = (S(x) + \pi(\infty, x) (\mathbf{P}_x(\infty) \cdot d)^{-1})^{-1}, \quad (4.2)$$

где для $k = x, x+1, \dots, \infty$

$$\pi(k, x) = \prod_{l=x}^{k-1} \left(1 - \frac{a_{21}(l)}{a_{11}(l)}\right); \quad (4.3)$$

$$S(x) = \sum_{k=x}^{\infty} \pi(k, x) \cdot (\mathbf{P}_x(k) a_{11}(k))^{-1}. \quad (4.4)$$

Отметим что, если $\pi(\infty, x) > 0$, то

$$\lim_{x \rightarrow \infty} w_d(x) = d. \quad (4.5)$$

Если $\pi(\infty, x) = 0$, то уравнение (4.1) имеет единственное решение

$$w_d(x) = (S(x))^{-1}, \quad (4.6)$$

и $\lim_{x \rightarrow \infty} w_d(x)$ не зависит от d .

Теорема 4.1. Если для всех $k \geq x$ выполняются неравенства

$$a_{21}(k) \leq p(k) \cdot w_d(k+1) \leq \mathbf{P}_k(\infty) \cdot d, \quad (4.7)$$

то значение $v_d(x)$ игры $\Gamma_d(x)$ равно $w_d(x)$, где $w_d(x)$ определено формулами (4.2)–(4.4).

Оптимальные стратегии игроков определяются соотношениями

$$\begin{aligned} \tau_d(k) &= 1 - w_d(k) \cdot (p(k))^{-1} (w_d(k+1))^{-1}; \\ \sigma_d(k) &= w_d(k) \cdot (a_{11}(k))^{-1}. \end{aligned}$$

Д о к а з а т е л ь с т в о. Непосредственно проверяется, что функция $w_d(x)$ при любом $d \in [0, \infty]$ удовлетворяет уравнению (4.1).

Для стратегии τ_d при $x \leq k < \infty$

$$\mathbf{P}_x\{\tau_d \geq k\} = w_d(x) \cdot (\mathbf{P}_x(k))^{-1} (w_d(k))^{-1},$$

$$\mathbf{P}_x\{\tau_d = \infty\} = w_d(x) \cdot (\mathbf{P}_x(\infty))^{-1} \left(\lim_{k \rightarrow \infty} w_d(k) \right)^{-1} \geq w_d(x) \cdot (\mathbf{P}_x(\infty))^{-1} d^{-1}.$$

Таким образом, для любой чистой стратегии $k \in [x, \infty)$ игрока 2

$$K_{x,d}(\tau_d, k) = \mathbf{P}_x\{\tau_d \geq k\} \cdot \mathbf{P}_x(k) \cdot w_d(k) = w_d(x).$$

Для стратегии ∞ игрока 2

$$K_{x,d}(\tau_d, \infty) = \mathbf{P}_x\{\tau_d = \infty\} \cdot \mathbf{P}_x(\infty) \cdot d \geq w_d(x).$$

Итак, стратегия τ_d обеспечивает игроку 1 в игре $\Gamma_d(x)$ выигрыш $w_d(x)$.

С другой стороны, для стратегии σ_d выполняются соотношения $1 - \sigma_d(k) = (1 - a_{21}(k)/a_{11}(k)) \cdot w_d(k) \cdot (p(k))^{-1}(w_d(k+1))^{-1}$. Поэтому, для $x \leq k < \infty$

$$\mathbf{P}_x\{\sigma_d \geq k\} = \pi(k, x) \cdot w_d(x) \cdot (\mathbf{P}_x(k))^{-1}(w_d(k))^{-1},$$

$$\mathbf{P}_x\{\sigma_d = \infty\} = \pi(\infty, x) \cdot w_d(x) \cdot (\mathbf{P}_x(\infty))^{-1} \left(\lim_{k \rightarrow \infty} w_d(k) \right)^{-1}.$$

Таким образом, для любой чистой стратегии $k \in [x, \infty)$ игрока 1

$$K_{x,d}(k, \sigma_d) = w_d(x) \cdot \left(\sum_{l=x}^{k-1} \pi(l, x) \frac{a_{21}(l)}{a_{11}(l)} + \pi(k, x) \right) = w_d(x),$$

поскольку $\sum_{l=x}^{k-1} \pi(l, x) a_{21}(l) (a_{11}(l))^{-1} = 1 - \pi(k, x)$. Для стратегии $\tau = \infty$ игрока 1

$$\begin{aligned} K_{x,d}(\infty, \sigma_d) &= w_d(x) \left(\sum_{l=x}^{\infty} \pi(l, x) \frac{a_{21}(l)}{a_{11}(l)} + \mathbf{P}_x\{\sigma_d = \infty\} \cdot \mathbf{P}_x(\infty) \cdot d \right) \\ &= w_d(x) \cdot (1 - \pi(\infty, x)) + w_d(x) \cdot \pi(\infty, x) \cdot \left(\lim_{k \rightarrow \infty} w_d(k) \right)^{-1} d = w_d(x). \end{aligned}$$

Последнее равенство является следствием соотношений (4.5) и (4.6). Итак, стратегия σ_d обеспечивает игроку 2 в игре $\Gamma_d(x)$ выигрыш $w_d(x)$.

Отсюда следует, что $w_d(x)$ является значением игры $\Gamma_d(x)$, а τ_d и σ_d — оптимальными стратегиями игроков. Теорема доказана.

Далее приводятся достаточные условия, обеспечивающие выполнение условий (4.7) теоремы 4.1.

Следствие 4.1. Если для всех $k \geq x$ выполняются неравенства

$$a_{21}(k) \leq p(k) v_{0,d}(k+1), \quad (4.8)$$

то $w_d(x)$ является значением игры $\Gamma_d(x)$.

Следствие 4.2. Если для всех $k \geq x$ выполняются неравенства

$$a_{21}(k) \leq p(k) a_{21}(k+1) \leq \mathbf{P}_k(\infty) d, \quad (4.9)$$

то $w_d(x)$ является значением игры $\Gamma_d(x)$.

В рассматриваемых ниже примерах игр $\Gamma_d(x)$ вероятность обрыва цепи равна нулю, то есть $p(k) = 1$.

Пример 4.1. Пусть для игры $\Gamma_d(x)$

$$a_{11}(x) = x(x+1), \quad a_{21}(x) = \frac{x}{x+1}, \quad d \geq 1.$$

Тогда выполнены как условия (4.8), так и условия (4.9). Мы имеем

$$\begin{aligned} \pi(k, x) &= \prod_{l=x}^{k-1} \left(1 - \frac{1}{(l+1)^2} \right) = \frac{x(k+1)}{(x+1)k}, \quad \pi(\infty, x) = \frac{x}{x+1}; \\ S(x) &= \frac{x}{x+1} \sum_{k=x}^{\infty} \frac{1}{k^2}. \end{aligned}$$

Таким образом, при $d \geq 1$, значение игры $\Gamma_d(x)$ равно

$$w_d(x) = \frac{x+1}{x} \left(\sum_{k=x}^{\infty} \frac{1}{k^2} + \frac{1}{d} \right)^{-1}.$$

В частности,

$$w_d(x+1) > w_1(x+1) = \frac{x+2}{x+1} \left(\sum_{k=x+1}^{\infty} \frac{1}{k^2} + 1 \right)^{-1} > \frac{x}{x+1} = a_{21}(x),$$

т.е. условия (4.7) теоремы 4.1 действительно выполнены.

П р и м е р 4.2. Пусть теперь для игры $\Gamma_d(x)$

$$a_{11}(x) = x, \quad a_{21}(x) = \frac{x}{x+1}, \quad d \geq 1.$$

В этом случае выполнены только условия (4.9) следствия 4.2. Мы получаем

$$\pi(k, x) = \prod_{l=x}^{k-1} \left(1 - \frac{1}{l+1} \right) = \frac{x}{k},$$

и, следовательно, $\pi(\infty, x) = 0$.

Таким образом, при любом $d \geq 1$, значение игры $\Gamma_d(x)$ равно

$$w_d(x) = (S(x))^{-1} = \left(x \cdot \sum_{k=x}^{\infty} \frac{1}{k^2} \right)^{-1}.$$

Как и в примере 4.1, $w_d(x+1) > x/(x+1) = a_{21}(x)$, т.е. условия (4.7) теоремы 4.1 действительно выполнены.

З а м е ч а н и е. Если в вышеприведенных примерах «выигрыш в бесконечности» $d < 1$, то при $x > d/(1-d) > d$ выполнены неравенства $x/(x+1) > d$ и значение игры $\Gamma_d(x)$ равно d . Значение игры при $x \leq d/(1-d)$ может быть вычислено посредством обратной индукции.

СПИСОК ЛИТЕРАТУРЫ

1. Дынкин Е. Б. Игровой вариант задачи об оптимальной остановке. — ДАН СССР, 1969, т. 185, № 1, с. 16–19.
2. Ширяев А. Н. Статистический последовательный анализ. М.: Наука, 1976, 272 с.
3. Фрид Е. Б. Оптимальная остановка цепи Маркова двумя лицами с противоположными интересами. — Теория вероятн. и ее примен., 1969, т. XIV, в. 4, с. 746–749.
4. Кицфер Ю. Н. Игры с оптимальной остановкой. — Теория вероятн. и ее примен., 1971, т. XVI, № 1, с. 183–188.
5. Neveu J. Discrete-Parameter Martingales. Amsterdam: North-Holland, 1975.
6. Элбакидзе Н. В. Построение цены и оптимальных политик в игровой задаче остановки марковского процесса. — Теория вероятн. и ее примен., 1976, т. XXI, в. 1, с. 164–169.
7. Доманский В. К. Игровая задача остановки цепи Маркова. Резюме докладов, сделанных на заседаниях семинара по теор. вер. и мат. стат. в ЛОМИ АН СССР, 1977. — Теория вероятн. и ее примен., 1978, т. XXIII, в. 4, с. 863–865.
8. Bismuth J. M. Sur un problem de Dynkin. — Z. Wahrscheinlichkeitstheor. verw. Geb., 1977, B. 39, S. 31–53.
9. Alario-Lazaret M., Lepeltier J. P., Marshal B. Dynkin Games. — Stochastic Differential Systems (Bad Honnef), 23–32, Lecture Notes in Control and Information Sciences, 1982, v. 43, Springer-Verlag.
10. Lepeltier J. P., Maingueneau M. A. Le jeu de Dynkin en theorie generale sans l'hypothese de Mokobodski. — Stochastics, 1984, v. 13, p. 25–44.
11. Yasuda M. On a randomized strategy in Neveu's stopping problem. — Stochastic Processes Appl., 1985, v. 21, p. 159–166.
12. Rosenberg D., Solan E., Vieille V. Stopping games with randomized strategies. — Laboratoire d'Econometrie de l'ecole polytechnique, Paris, France, preprint, 2000.
13. Sakaguchi M. Sequential deception games. — Math. Jap., 1992, v. 37, № 5, p. 813–826.

14. Irle A. Games of stopping with infinite horizon. — Math. Methods Oper. Res., 1995, v. 42, p. 345–359.
15. Доманский В. К. Об одном классе игровых задач остановки последовательности сумм случайных величин. — Современные направления теории игр. Вильнюс: Мокслас, 1976, с. 86–93.
16. Прессман Э. Л., Соппин И. М. Игровые задачи оптимальной остановки. Существование и единственность точек равновесия. — Вероятностные проблемы управления в экономике. М.: Наука, 1977, с. 115–144.
17. Мазалов В. В., Кочетов Э. А. Игры с оптимальной остановкой случайных блужданий. — Теория вероятн. и ее примен., 1997, т. 42, в. 4, с. 820–826.

Поступила в редакцию
21.III.2000

© 2001 г.

КРУГЛОВ В. М.* , ПЕТРОВСКАЯ Г. Н.*

СЛАБАЯ СХОДИМОСТЬ ОДНОГО ФУНКЦИОНАЛА¹⁾

Изучается функционал $T_n = (S_1^2 + \dots + S_n^2)/(nV_n^2)$, построенный по последовательности $\{X_n\}_{n \geq 1}$ независимых одинаково распределенных случайных величин, где $S_k = X_1 + \dots + X_k$, $V_n^2 = X_1^2 + \dots + X_n^2$. Пусть G — функция распределения случайной величины $\int_0^1 W^2(t) dt$, где $W(t)$, $t \in [0, 1]$, — винеровский процесс. Доказано, что функция распределения T_n слабо сходится к G при $n \rightarrow \infty$ тогда и только тогда, когда функция распределения случайной величины X_1 принадлежит области притяжения нормального закона и $EX_1 = 0$.

Ключевые слова и фразы: слабая сходимость, сходимость по вероятности, случайная величина, функция распределения.

Пусть $\{X_n\}_{n \geq 1}$ — последовательность независимых одинаково распределенных случайных величин, определенных на вероятностном пространстве $(\Omega, \mathcal{F}, \mathbf{P})$. Обозначим $S_n = X_1 + \dots + X_n$, $V_n^2 = X_1^2 + \dots + X_n^2$. Положим $S_n/V_n = 0$, если $V_n = 0$. В статье [1] доказано, что последовательность случайных величин $\{S_n/V_n\}_{n \geq 1}$ слабо сходится к случайной величине со стандартной нормальной функцией распределения Φ тогда и только тогда, когда $EX_1 = 0$ и распределение случайной величины X_1 принадлежит области притяжения нормального закона. Упомянутая статья [1] подводит определенный итог многочисленным исследованиям об асимптотическом поведении последовательности $\{X_n/V_n\}_{n \geq 1}$. Заинтересованный читатель может обратиться к списку литературы, приведенному в [1]. Отметим, что там не упомянута интересная, на наш взгляд, статья [2], в которой среди прочего указано необходимое и достаточное условие слабой сходимости функций распределения случайных величин S_n/V_n к Φ для независимых симметричных (необязательно одинаково распределенных) случайных величин.

Цель настоящей заметки показать, что существуют функционалы, отличные от S_n/V_n , слабая сходимость которых равносильна принадлежности функции распределения случайной величины X_1 области притяжения нормального закона. Мы подробно обсудим один такой функционал. Если нормирование функционалов производить не функциями от V_n , а положительными числами, то примеры таких функционалов и историю их возникновения можно найти в [3].

* МГУ, факультет вычислительной математики и кибернетики, Воробьевы горы, 119899 Москва, Россия; e-mail: vkrug@dataforce.net

¹⁾ Работа выполнена при поддержке Российского фонда фундаментальных исследований, гранты 99-01-00846, 99-01-00847, 01-01-00644 и INTAS, грант 99-01317.