

УДК 517.977.8 + 517.977.5 + 519.857 + 519.87

ББК 22.18

УРАВНЕНИЕ ГАМИЛЬТОНА-ЯКОБИ-БЕЛЛМАНА В ДИФФЕРЕНЦИАЛЬНЫХ ИГРАХ СО СЛУЧАЙНОЙ ПРОДОЛЖИТЕЛЬНОСТЬЮ

ЕКАТЕРИНА В. ШЕВКОПЛЯС

Факультет прикладной математики –

процессов управления

Санкт-Петербургский государственный университет

Санкт-Петербург

e-mail: katya_shev@mail.ru

В статье рассматривается класс дифференциальных игр со случайной продолжительностью. Показывается, что задача со случайной продолжительностью может быть сведена к стандартной задаче с бесконечным временем. Для нахождения оптимальных решений в дифференциальных играх со случайной продолжительностью выводится уравнение типа Гамильтона-Якоби-Беллмана. Результаты демонстрируются на примере ретико-игровой модели разработки невозобновляемых ресурсов. Задача решается при предположении о том, что случайная величина, соответствующая моменту окончания игры, распределена по закону Вейбулла.

Ключевые слова: дифференциальные игры, уравнение Гамильтона-Якоби-Беллмана, случайная продолжительность, разработка невозобновляемых ресурсов.

1. Введение

В настоящее время в теории игр, в том числе и в теории дифференциальных игр, достаточно ярко выражен интерес к более адекватному моделированию конфликтно-управляемых процессов, происходящих в реальности. При моделировании таких процессов необходимо учитывать стохастический характер окружающей нас действительности. В частности, очевидно, что любой реальный процесс, развивающийся во времени, также как и сама жизнь, имеет случайную продолжительность, поскольку заканчивается в некоторый случайный момент времени. В приложениях дифференциальных игр часто изучаются различные конфликтно-управляемые процессы в экономике, экологии, менеджменте и других сферах человеческой деятельности, которые развиваются во времени, а затем заканчиваются в фиксированный момент времени. Кроме того, в последнее время многие задачи, особенно в экономических приложениях, рассматриваются на бесконечном временном промежутке при условии дисконтирования мгновенных выигрышей во времени. На практике же такие процессы имеют случайный момент окончания, что связано со множеством причин неигрового характера. В том числе, недавно случившийся экономический кризис послужил основанием для досрочного прекращения многих товарно-денежных отношений. По этой причине рассмотрение игр со случайной продолжительностью представляется более реалистичным.

Впервые постановка задачи со случайной продолжительностью дифференциальной игры была сформулирована в работе Петросяна Л.А., Мурзова Н.В. в 1966 году [7]. В данной работе изучалась антагонистическая игра преследования двух лиц с терминальными выигрышами в последний момент времени, который являлся случайной величиной с известной функцией распределения. Для игры в такой постановке впервые было выведено уравнение типа Айзекса-Беллмана.

Затем общая постановка дифференциальной игры со случайной продолжительностью была представлена в работе Петросяна Л.А., Шевкопляс Е.В. в 2000 году [8]. В данной работе изучались кооперативные дифференциальные игры со случайной продолжительностью и акцент был сделан на проблеме динамической устойчивости

принципов оптимальности в новой постановке игры. В то же время стоит отметить, что параллельно и независимо в теории оптимального управления рассматривались неигровые модели со случайной продолжительностью, начиная с работы М.Е. Yaari [17] 1965 года, посвященной теории страхования жизни потребителя при неопределенном времени его смерти.

В данной работе рассматриваются дифференциальные игры со случайной продолжительностью. В разделе 1 приводится постановка задачи, причем интегральные выигрыши игроков при помощи перестановки интегралов приводятся к стандартному виду. В разделе 2 выводится уравнение типа Гамильтона-Якоби-Беллмана для задачи со случайной продолжительностью. Раздел 3 содержит некоторые сведения и понятия из математической теории надежности, которые возникают в дифференциальной игре при условии случайной продолжительности игры. Кроме того, будет выбрано распределение Вейбулла, как распределение, описывающее случайную величину, являющуюся моментом окончания игры. В разделе 4 рассматривается конкретный пример дифференциальной игры со случайной продолжительностью, а именно теоретико-игровая модель разработки невозобновляемых ресурсов.

2. Модель игры

В дифференциальных играх, как правило, используются два основных подхода к тому, на каком промежутке времени рассматривается игра. Согласно первому подходу, игра развивается во времени на фиксированном временном промежутке $[t_0, T]$: момент окончания игры T известен заранее и игра имеет так называемую предписанную продолжительность [6].

Существуют также подход, имеющий большое число экономических приложений и широко применяющийся на практике, согласно которому игра развивается на бесконечном временном промежутке [12]. При таком подходе подынтегральная функция полезности игрока (или функция мгновенного выигрыша), как правило, дисконтируется при помощи экспоненциальной функции, однако фактически такая игра не имеет окончания. В данной работе используется другая постановка дифференциальной игры, а именно дифференциальные игры со случайной продолжительностью.

Итак, рассмотрим дифференциальную игру n лиц $\Gamma(x_0)$ со случайной продолжительностью $T - t_0$ [8-10] и начальным состоянием x_0 . Динамика игры задается системой обыкновенных дифференциальных уравнений:

$$\begin{aligned} \dot{x} &= g(x, u_1, \dots, u_n), \quad x \in R^n, u_i \in U \subseteq \text{comp } R^l, \\ x(t_0) &= x_0. \end{aligned} \quad (2.1)$$

Игра начинается в момент t_0 из состояния x_0 , однако момент ее окончания не фиксирован заранее, а является реализацией некоторой случайной величины T . Будем полагать, что для случайной величины T задана функция распределения $F(t)$, которая определена при $t \in [t_0, \infty)$ и удовлетворяет условию нормировки:

$$\int_{t_0}^{\infty} dF(t) = 1.$$

Функцию "мгновенного" выигрыша игрока i в момент времени τ , $\tau \in [t_0, \infty)$ обозначим как $h_i(x(\tau)) \geq 0$. Предполагается, что $h_i(\cdot)$ являются непрерывными функциями на R^m . Тогда ожидаемый интегральный выигрыш игрока i имеет вид:

$$K_i(x_0, t_0, u_1, \dots, u_n) = \int_{t_0}^{\infty} \int_{t_0}^t h_i(x(\tau)) d\tau dF(t), \quad i = 1, \dots, n. \quad (2.2)$$

Пусть игра развивается во времени, тогда в некоторый промежуточный момент времени ϑ , $\vartheta \in (t_0; \infty)$ игроки попадают в подыгру $\Gamma(x(\vartheta))$ с начальным состоянием $x(\vartheta) = x$. Очевидно, что игра может и закончиться до момента ϑ с вероятностью $F(\vartheta)$, а вероятность продолжить игру после момента ϑ равна $(1 - F(\vartheta))$.

Тогда ожидаемый интегральный выигрыш игрока i вычисляется по формуле

$$K_i(x, \vartheta, u_1, \dots, u_n) = \int_{\vartheta}^{\infty} \int_{\vartheta}^t h_i(x(\tau)) d\tau dF_{\vartheta}(t), \quad (2.3)$$

где $F_{\vartheta}(t)$, $t \geq \vartheta$ - это функция распределения момента окончания игры в подыгре $\Gamma(x(\vartheta))$. Не трудно заметить, что $F_{\vartheta}(t)$ является услов-

ной функцией распределения, а именно функцией распределения момента окончания игры при условии, что игра не закончилась до момента ϑ , $\vartheta \in (t_0; \infty)$. Кроме того, необходимо, чтобы $F_\vartheta(t)$ удовлетворяла стандартному условию нормировки при $\vartheta \in (t_0; \infty)$. В данной работе мы рассматриваем только стационарные процессы. Тогда условная функция распределения $F_\vartheta(t)$ вычисляется по следующей формуле:

$$F_\vartheta(t) = \frac{F(t) - F(\vartheta)}{1 - F(\vartheta)}, \quad t \in [\vartheta, \infty). \quad (2.4)$$

Далее мы будем предполагать существование плотности распределения момента окончания игры $f(t) = F'(t)$. Очевидно, что в подыгре $\Gamma(x(\vartheta))$ условная плотность распределения $f_\vartheta(t)$ определяется следующим образом:

$$f_\vartheta(t) = \frac{f(t)}{1 - F(\vartheta)}, \quad t \in [\vartheta, \infty). \quad (2.5)$$

Таким образом, при предположении о существовании плотности $f(t) = F'(t)$ и учитывая равенства (2.3) и (2.5), получаем интегральный выигрыш игрока i , $i = 1, \dots, n$, в подыгре $\Gamma(x(\vartheta))$:

$$K_i(x, \vartheta, u_1, \dots, u_n) = \frac{1}{1 - F(\vartheta)} \int_\vartheta^\infty \int_\vartheta^t h_i(x(\tau)) d\tau f(t) dt. \quad (2.6)$$

2.1. Перестановка интегралов в интегральном функционале

Рассмотрим интегральный выигрыш игрока i , который имеет вид (2.6). Не умаляя общности, в этом разделе положим $t_0 = 0$. Кроме того, введем более компактное обозначение $h_i(\tau) = h_i(x(\tau))$. Итак, рассмотрим интегральный функционал

$$\int_0^\infty \int_0^t h_i(\tau) d\tau f(t) dt.$$

Введем кусочную функцию $a(t, \tau)$ следующим образом:

$$a(t, \tau) = f(t)h_i(\tau) \cdot \chi_{\{\tau \leq t\}} = \begin{cases} f(t)h_i(\tau), & \tau \leq t; \\ 0, & \tau > t. \end{cases}$$

Если функция $a(t, \tau) \geq 0$ неотрицательно определена, то можно использовать теорему Тонелли [2,3] о перестановке интегралов в

повторном интеграле. Тогда справедливо следующее равенство:

$$\begin{aligned} & \int_0^{+\infty} dt \int_0^t f(t)h_i(\tau)d\tau = \int_0^{+\infty} dt \int_0^{+\infty} a(t, \tau)d\tau = \\ & = \iint_{[0,+\infty) \times [0,+\infty)} a(t, \tau)dtd\tau = \int_0^{+\infty} d\tau \int_0^{+\infty} a(t, \tau)dt = \\ & = \int_0^{+\infty} d\tau \int_{\tau}^{+\infty} f(t)h_i(\tau)dt = \int_0^{+\infty} (1 - F(\tau))h_i(\tau)d\tau. \end{aligned}$$

Если же нельзя гарантировать $a(t, \tau) \geq 0$ (а, фактически, неотрицательность функции мгновенного выигрыша $h_i(\tau)$), но при этом выполнено условие абсолютной сходимости кратного интеграла

$$\iint_{[0,+\infty) \times [0,+\infty)} |a(t, \tau)|dtd\tau < +\infty,$$

то можно использовать теорему Фубини [2,3] и также изменить порядок интегрирования.

Следовательно, при некоторых стандартных ограничениях на функцию $h_i(\cdot)$ мгновенного выигрыша игрока i , интегральный функционал, соответствующий ожидаемому интегральному выигрышу игрока i в игре $\Gamma(x_0)$ может быть приведен к стандартному виду для динамического программирования:

$$\begin{aligned} K_i(x_0, t_0, u_1, \dots, u_n) &= \int_{t_0}^{\infty} \int_{t_0}^t h_i(x(\tau))d\tau dF(t) = \\ &= \int_{t_0}^{\infty} (1 - F(\tau))h_i(x(\tau))d\tau. \end{aligned}$$

Аналогично, для выигрыша игрока i в подыгре $\Gamma(x(\vartheta))$ справедливо следующее представление:

$$\begin{aligned} K_i(x, \vartheta, u_1, \dots, u_n) &= \int_{\vartheta}^{\infty} \int_{\vartheta}^t h_i(x(\tau))d\tau dF_{\vartheta}(t) = \\ &= \frac{1}{1 - F(\vartheta)} \int_{\vartheta}^{\infty} (1 - F(\tau))h_i(x(\tau))d\tau. \quad (2.7) \end{aligned}$$

Таким образом, нестандартный для динамического программирования функционал, а именно функционал с повторным интегрированием (2.6), был приведен к стандартному виду (2.7) при помощи замены порядка интегрирования.

3. Уравнение Гамильтона-Якоби-Беллмана

Рассмотрим кооперативную форму игры $\Gamma(x_0)$. Перед началом игры игроки договариваются об использовании ими таких допустимых программных управлений, которые будут максимизировать совокупный ожидаемый выигрыш игроков:

$$\sum_{i=1, \dots, n} K_i(x_0, t_0, u_1, \dots, u_n) = \sum_{i=1, \dots, n} \int_{t_0}^{\infty} (1 - F(\tau)) h_i(x(\tau)) d\tau. \quad (3.1)$$

Управления $\{u_1^*(t), \dots, u_n^*(t)\}$, доставляющие максимум (3.1) будем называть оптимальными, а траекторию $x^*(t)$, соответствующую оптимальным управлениям, – условно-оптимальной. Очевидно, что при непрерывности функций h_i в (3.1) знак суммирования можно перенести в подынтегральную функцию:

$$\sum_{i=1, \dots, n} K_i(x_0, t_0, u_1, \dots, u_n) = \int_{t_0}^{\infty} (1 - F(\tau)) \sum_{i=1, \dots, n} h_i(x(\tau)) d\tau.$$

Обозначим $\sum_{i=1, \dots, n} h_i(x(\tau))$ как $L(x, \tau)$. В общем случае будем рассматривать $L(x, u, \tau)$.

Для нахождения оптимальных управлений можно использовать и принцип максимума Понтрягина, и уравнение Гамильтона-Якоби-Беллмана. В данной работе используется второй подход, поскольку новое уравнение Гамильтона-Якоби-Беллмана можно будет использовать не только для нахождения оптимальных управлений, но и для нахождения других решений в управлениях с обратной связью для кооперативного и некооперативного вариантов игры.

Итак, рассмотрим следующую задачу максимизации:

$$\begin{aligned} \frac{1}{1 - F(t)} \int_t^{\infty} L(x, u, s)(1 - F(s)) ds, & \quad (3.2) \\ \dot{x} &= g(x, u) \\ x(t) &= x. \end{aligned}$$

Пусть $W(x, t)$ – это функция Беллмана для данной оптимизационной задачи.

Кроме того, рассмотрим другую задачу максимизации, которая отличается от сформулированной выше только отсутствием множителя $\frac{1}{1-F(t)}$ перед интегральным функционалом:

$$\int_t^\infty L(x, u, s)(1 - F(s))ds, \tag{3.3}$$

$$\dot{x} = g(x, u) \quad x(t) = x.$$

Обозначим как $\bar{W}(x, t)$ функцию Беллмана для этой задачи.

Очевидно, что справедливо следующее равенство:

$$\bar{W}(x, t) = W(x, t) \cdot (1 - F(t)). \tag{3.4}$$

Тогда частные производные от функции \bar{W} по ее аргументам вычисляются по следующим формулам:

$$\frac{\partial \bar{W}}{\partial t} = -f(t)W + (1 - F(t))\frac{\partial W}{\partial t}; \tag{3.5}$$

$$\frac{\partial \bar{W}}{\partial x} = (1 - F(t))\frac{\partial W}{\partial x}. \tag{3.6}$$

Для задачи динамического программирования (3.3) для функции Беллмана \bar{W} мы имеем стандартное уравнение Гамильтона-Якоби-Беллмана [12]:

$$\frac{\partial \bar{W}}{\partial t} + \max_u \left(L(x, u, t)(1 - F(t)) + \frac{\partial \bar{W}}{\partial x}g(x, u) \right) = 0. \tag{3.7}$$

Используя (3.4), (3.5), (3.6), из уравнения (3.7) получаем уравнение Гамильтона - Якоби - Беллмана для задачи (3.2) со случайным моментом окончания игры:

$$\frac{f(t)}{1 - F(t)}W = \frac{\partial W}{\partial t} + \max_u \left(L(x, u, t) + \frac{\partial W}{\partial x}g(x, u) \right). \tag{3.8}$$

Уравнение типа Гамильтона-Якоби-Беллмана для дифференциальной игры со случайной продолжительностью впервые было выведено в работе Л.А. Петросяна, Н.В. Мурзова [7] в 1966 году для частного случая дифференциальной игры преследования с терминальными выигрышами. Позднее в диссертационной работе Е.В. Шевкопляс [9,10] в 2004 году было напрямую получено уравнение (3.8) в общей

постановке дифференциальной игры со случайной продолжительностью и интегральными выигрышами без предварительного приведения интегрального функционала к стандартному виду.

Кроме того, независимо уравнение (3.8) было получено и другими авторами различными способами. Как будет показано ниже в разделе 4.3, интегральный функционал, соответствующий выигрышу игрока для случая игры со случайной продолжительностью, приведенный к стандартному виду, эквивалентен функционалу, соответствующему выигрышу игрока в задаче с непостоянным дисконтированием мгновенного выигрыша на бесконечном временном промежутке [11,15,16]. В неигровой постановке задачи страхования жизни с рекурсивным дисконтированием функции полезности, Ф.Р. Чангом в 2004 году [11] также было выведено уравнение (3.8) в форме (4.3). Независимо от работ Чанга, уравнение Гамильтона-Якоби-Беллмана (3.8) в форме (4.3) для теоретико-игровых моделей, в которых функции полезности игроков дисконтируются при помощи функции $e^{-\int_{t_0}^t \lambda(\tau) d\tau}$, было получено в работах [15,16] в 2007-2009 годах для нахождения так называемых предварительных соглашений для наивных агентов.

Заметим, что если случайная величина T распределена по экспоненциальному закону, а именно

$$f(t) = \lambda e^{-\lambda(t-t_0)}; \quad F(t) = 1 - e^{-\lambda(\tau-t_0)}; \quad \frac{f(t)}{1 - F(t)} = \lambda,$$

то интегральный выигрыш $K_i(\cdot)$ игрока i эквивалентен интегральному выигрышу игрока i в игре, рассматриваемой на бесконечном временном промежутке и дисконтированием со ставкой λ мгновенных выигрышей игроков:

$$\begin{aligned} K_i(x_0, t_0, u_1, \dots, u_n) &= \int_{t_0}^{\infty} h_i(x(\tau))(1 - F(\tau))d\tau = \\ &= \int_{t_0}^{\infty} h_i(x(\tau))e^{-\lambda(\tau-t_0)} d\tau. \end{aligned}$$

Полученное уравнение Гамильтона-Якоби-Беллмана, таким образом, для случая с экспоненциальным распределением момента окончания игры, должно сводиться к известному уравнению для игры с дисконтированными мгновенными выигрышами (или функциями полезности).

сти в терминологии зарубежной литературы по теории игр) [12]. Легко проверить, что при $\frac{f(t)}{1-F(t)} = \lambda$ новое уравнение типа Гамильтона-Якоби-Беллмана (3.8) принимает вид уравнения Гамильтона-Якоби-Беллмана (3.9) для игры с дисконтированными мгновенными выигрышами, опубликованного в работе [12] в 2000 году:

$$\lambda W(x, t) = \frac{\partial W(x, t)}{\partial t} + \max_u \left(L(x, u, t) + \frac{\partial W(x, t)}{\partial x} g(x, u) \right). \quad (3.9)$$

Таким образом, задача со случайной продолжительностью с экспоненциальным законом распределения T и стандартная задача с дисконтированными мгновенными выигрышами являются эквивалентными. Этот факт ранее был отмечен в работе А. Нэриэ [13] для другой постановки задачи со случайной продолжительностью игры, а именно игры конфликта поколений, имеющих случайную продолжительность существования.

4. Вероятностные распределения для момента окончания игры

Заметим, что множитель $\frac{f(\vartheta)}{1-F(\vartheta)}$, появившийся в левой части уравнения Гамильтона-Якоби-Беллмана (3.8), является стандартной для теории надежности функцией интенсивности отказов $\lambda(t)$ [1, 4, 14]. В математической теории надежности одной из важнейших изучаемых случайных величин является время T отказа системы элементов, а функция интенсивности отказов $\lambda(t)$ (или функция угрозы) — одна из основных ее характеристик. Таким образом, можно провести аналогию между теорией игр, в которой игроки, как элементы системы, взаимодействуют между собой в некотором процессе, разворачивающемся во времени, и математической теорией надежности. Мы рассматриваем игры со случайным моментом окончания T , к которым можно непосредственно применить основную терминологию из математической теории надежности.

4.1. Функция интенсивности отказов

Итак, функция интенсивности отказов или функция угрозы, определяется следующим образом:

$$\lambda(t) = \frac{f(t)}{1 - F(t)}. \quad (4.1)$$

Фактически, функция интенсивности отказов является условной плотностью распределения, т.е. плотностью распределения при условии, что система не отказала до момента t , в нашей же терминологии – плотностью распределения момента окончания игры при условии, что игра не закончилась до момента t . Стандартными обозначениями являются $\lambda(t)$ или $h(t)$, далее будем использовать обозначение $\lambda(t)$. Функция интенсивности отказов $\lambda(t)$, описывающая жизненный цикл системы, имеет следующий вид: Первая фаза называется фа-

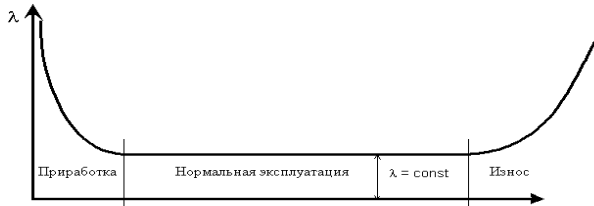


Рисунок 1. Функция интенсивности отказов $\lambda(t)$

зой приработки и отказы в ней, согласно теории надежности, возникают за счет невыявленных перед началом эксплуатации скрытых дефектов [4]. Специфика этой проблемы понятна и не только с точки зрения приложения к техническим системам элементов, не зря в теории надежности в англоязычной литературе для данного периода используется такая терминология как "новорожденный период" "детская смертность" и "ранние отказы". С точки зрения теории игр, ранние отказы могут быть вызваны неопытностью, несогласованностью и не налаженной деятельностью игроков, только что вступивших в игру. Функция интенсивности отказов $\lambda(t)$ в данной фазе является убывающей функцией времени.

Следующим периодом жизненного цикла системы является так называемый период нормальной эксплуатации системы. Функция интенсивности отказов $\lambda(t)$ в этом периоде постоянна (либо постоянна в приближении), а сами внезапные отказы вызваны несовершенством самой системы и обусловлены, как правило, внешними причинами. В англоязычной терминологии этот период называется "взрослым" периодом [14]. Игра в рассматриваемом периоде может прекратиться под воздействием каких-то непредусмотренных обстоятельств

внешнего мира, в частности, в экономических приложениях - под воздействием дефолта, кризиса или других серьезных причин.

После периода нормальной эксплуатации, система переходит в фазу старения. Отказы системы в этом периоде связаны с износом и старением элементов системы, а функция интенсивности отказов $\lambda(t)$ является возрастающей функцией. Очевидно, что этому факту нетрудно дать интерпретацию в игре – наступает износ технической и инструментальной базы игроков (например, в игре разработки природных ресурсов оборудование подвергается износу). Кроме того, сами соглашения, заключенные перед началом игры, в общем случае не являются реализуемыми в долгосрочных проектах. Этот факт, впервые отмеченный в работе Л.А. Петросяна в 1977 году [5,6] и названный динамической неустойчивостью решений, в англоязычной литературе получил название несостоятельности во времени. Таким образом, можно заключить, что сами соглашения игроков, заключенные перед началом игры, тоже в некотором смысле подвергаются старению и износу.

Отметим, что функция интенсивности отказов является константой тогда и только тогда, когда случайное время отказа системы T подчинено экспоненциальному закону распределения, т.е.

$$\frac{f(t)}{1 - F(t)} = \lambda(t) = \lambda = const$$

только при

$$f(t) = \lambda e^{-\lambda(t-t_0)}, \quad t > t_0.$$

Таким образом, только при экспоненциальном распределении случайной величины T выведенное уравнение Гамильтона-Якоби-Беллмана (3.8) сводится к более простому уравнению (3.9).

4.2. Выбор распределения для момента окончания игры

В математической теории надежности для случайной величины T (момента отказа системы технических элементов) используются различные вероятностные распределения, а именно экспоненциальное, Вейбулла, нормальное, логарифмически нормальное, Гамма-распределение и другие [4, 14].

В актуарной математике, а также в геронтологии (науке о продолжительности жизни), для времени существования биологических систем также замечены некоторые закономерности, при этом основными законами являются закон Гомперца-Мейкема и закон Вейбулла [1].

Поскольку в данной статье игра может быть интерпретирована как сочетание различных взаимодействий (как на биологическом, так и техническом уровне), то распределение Вейбулла, используемое для описания продолжительности существования как технических, так и биологических систем, представляется наиболее адекватным распределением для случайного момента окончания игры.

Распределение Вейбулла имеет функцию интенсивности отказов следующего вида [4]:

$$\lambda(t) = \lambda \delta t^{\delta-1}; \quad (4.2)$$

$$t \geq 0; \lambda > 0; \delta > 0.$$

Здесь λ и δ – параметры, определяющие данное распределение.

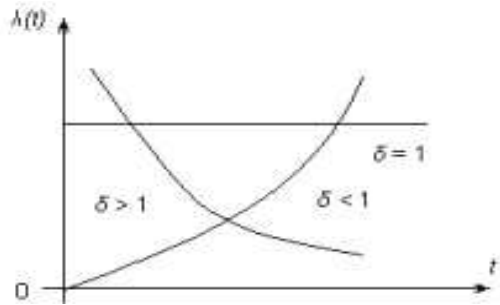


Рисунок 2. Функция интенсивности отказов $\lambda(t)$ для распределения Вейбулла

λ – это параметр масштаба, а параметр формы δ соответствует одной из трех фаз, в которой может находиться система (в данном контексте игра). Значение $\delta < 1$ соответствует "новорожденному" периоду, здесь функция интенсивности отказов $\lambda(t)$ является убывающей функцией. При $\delta = 1$ система находится в режиме нормальной эксплуатации, $\lambda(t)$ равна константе λ . Отметим, что при

$\delta = 1$ распределение Вейбулла соответствует экспоненциальному распределению. При $\delta > 1$ система находится в состоянии износа, $\lambda(t)$ является возрастающей функцией. Частным случаем распределения Вейбулла для "стареющей" системы является распределение Рэлея, которому соответствует $\delta = 2$.

4.3. Уравнение Гамильтона-Якоби-Беллмана для дифференциальной игры со случайным моментом окончания, распределенным по закону Вейбулла

Согласно определению функции интенсивности отказов $\lambda(t)$ (), выведенное в разделе 3 уравнение Гамильтона-Якоби-Беллмана (3.8) может быть переписано в следующем виде:

$$\lambda(t)W(x, t) = \frac{\partial W(x, t)}{\partial t} + \max_u \left(L(x, u, t) + \frac{\partial W(x, t)}{\partial x} g(x, u) \right). \quad (4.3)$$

Как было сказано выше в разделе 4.1 функция $\lambda(t)$ для экспоненциального распределения является константой, а именно $\lambda(t) = \lambda$. Следовательно, для экспоненциального распределения момента окончания игры из уравнения (4.3) непосредственно следует стандартное уравнение Гамильтона-Якоби-Беллмана [12] для игр с дисконтированными мгновенными выигрышами (3.9).

Кроме того, отметим, что уравнение (4.3) имеет тот же вид, что и уравнение Гамильтона-Якоби-Беллмана для игр с непостоянными дисконтированными мгновенными выигрышами, а именно, когда дисконтирование выигрыша h_i производится при помощи умножения его на $e^{-\int_{t_0}^{\tau} \lambda(t) dt}$ [15,16]. Это непосредственно следует из того, что в приведенном к стандартному виду функционале (2.6) при помощи функции интенсивности отказов $\lambda(t)$ (4.1), выражение $1 - F(\tau)$ путем несложных преобразований можно записать как $1 - F(\tau) = e^{-\int_{t_0}^{\tau} \lambda(t) dt}$.

Пусть момент окончания игры T распределен по закону Вейбулла. Тогда функция интенсивности отказов $\lambda(t)$ вычисляется по формуле (4.2). Следовательно, уравнение (4.3) Гамильтона-Якоби-Беллмана принимает следующий вид:

$$\lambda \delta t^{\delta-1} W(x, t) = \frac{\partial W(x, t)}{\partial t} + \max_u \left(L(x, u, t) + \frac{\partial W(x, t)}{\partial x} g(x, u) \right). \quad (4.4)$$

Очевидно, что при значении $\delta = 1$, которое соответствует экспоненциальному распределению, уравнение (4.4) принимает стандартный вид (3.9).

5. Пример

В качестве примера рассмотрим теоретико-игровую модель разработки невозобновляемых ресурсов симметричными игроками, опубликованную в работе [12]. Однако в данной работе будем предполагать, что мгновенные выигрыши игроков не дисконтируются и, кроме того, игра заканчивается в случайный момент времени T , распределенный по закону Вейбулла. Итак, согласно модели [12], в игре участвуют n игроков – фирмы или страны, которые разрабатывают некоторый невозобновляемый природный ресурс, например, нефть. Множество всех игроков обозначим как $I = \{1, 2, \dots, n\}$. Пусть $x(t)$ – это поток невозобновляемого ресурса. Управлениями игроков являются темпы разработки ресурса, которые обозначим как $\{c_i(t)\}$. Динамика изменений потока ресурса $x(t)$ описывается следующим дифференциальным уравнением:

$$\begin{aligned} \dot{x}(t) &= - \sum_{i=1}^n c_i(t); \\ \lim_{t \rightarrow \infty} x(t) &\geq 0; \\ x(t_0) &= x_0. \end{aligned}$$

Каждый игрок i имеет функцию полезности (функцию мгновенного выигрыша) $h(c_i)$, определенную для всех $c_i > 0$. Рассмотрим только логарифмический вид функции полезности игрока:

$$h(c_i) = \ln(c_i). \quad (5.1)$$

Пусть игра начинается в момент времени $t_0 = 0$, заканчивается в момент времени T , который является случайной величиной, распределенной по закону Вейбулла. Тогда ожидаемый выигрыш игрока i имеет следующий вид:

$$K_i(x_0, u_1, \dots, u_n) = \int_0^\infty \int_0^t h_i(\cdot) d\tau f(t) dt, \quad (5.2)$$

$$f(t) = \lambda \delta t^{\delta-1} e^{-\lambda t^\delta}.$$

Согласно разделу 2.1, интегральный выигрыш (5.2) можно представить в виде (2.7). Для распределения Вейбулла $1 - F(t) = e^{-\lambda t^\delta}$ [4], следовательно, интегральный выигрыш игрока i вычисляется по следующей формуле:

$$K_i(x_0, u_1, \dots, u_n) = \int_0^\infty h_i(\cdot) e^{-\lambda t^\delta} dt. \quad (5.3)$$

Тогда общий ожидаемый выигрыш игроков вычисляется по формуле

$$\sum_{i=1, \dots, n} K_i(x_0, u_1, \dots, u_n) = \int_0^\infty \sum_{i=1, \dots, n} h_i(\cdot) e^{-\lambda t^\delta} dt.$$

Пусть $V(N, x_0)$ – максимальное значение общего ожидаемого выигрыша в игре $\Gamma(x_0)$.

Очевидно, что в данной модели рассматриваются только симметричные игроки, поэтому положим $c_i = c_j = c$. Кроме того, на данном этапе представляется возможным найти решение для произвольного распределения с функцией интенсивности отказов $\lambda(t)$. Итак, согласно (4.3) имеем уравнение Гамильтона-Якоби-Беллмана следующего вида:

$$\lambda(t)W(x, \vartheta) = \frac{\partial W(x, \vartheta)}{\partial \vartheta} + \max_c \left[n \ln(c(\vartheta)) + \frac{\partial W(x, \vartheta)}{\partial x} (-nc(\vartheta)) \right]. \quad (5.4)$$

Будем искать решение уравнения (5.4) в виде $W(x, t) = A(t) \ln x + B(t)$. Тогда частные производные $W(x, t)$ вычисляются по формуле

$$\frac{\partial W(x, t)}{\partial x} = \frac{A(t)}{x}; \quad \frac{\partial W(x, t)}{\partial t} = \dot{A}(t) \ln(x) + \dot{B}(t). \quad (5.5)$$

Учитывая (5.5), из условия максимизации правой части уравнения (5.4) следует, что оптимальные управления имеют вид $c = \frac{x}{A}$. Применяя метод неопределенных коэффициентов в уравнении (5.4), получаем следующую систему уравнений для коэффициентов $A(t), B(t)$:

$$\dot{A}(t) - \lambda(t)A(t) + n = 0;$$

$$\dot{B}(t) - \lambda(t)B(t) - n \ln(A(t)) - n = 0.$$

Окончательно имеем следующие оптимальные управления для задачи разработки невозобновляемых ресурсов со случайной продолжительностью:

$$c_i^* = c^* = \frac{x \cdot e^{-\lambda(t)t}}{n \int_t^\infty e^{-\lambda(s)s} ds}. \quad (5.6)$$

Предположим, что случайный момент окончания разработки ресурсов подчинен закону Вейбулла. Тогда функция $\lambda(t)$ имеет вид (4.2). Тогда при $\delta = 1$, соответствующем экспоненциальному распределению момента окончания игры, фактически рассматривается уже изученная модель с дисконтированными выигрышами на бесконечном временном промежутке. Непосредственно из (5.6) следует, что при $\delta = 1$ оптимальными стратегиями игроков являются

$$c_i^* = c^* = \frac{\lambda}{n} x, \quad i = 1, \dots, n.$$

Тогда оптимальные управления и траектория вычисляются по формуле

$$\begin{aligned} x^*(t) &= x_0 * e^{-\lambda t}; \\ c_i^*(t) &= \frac{\lambda}{n} x_0 * e^{-\lambda t}. \end{aligned}$$

Этот результат совпадает с результатом, полученным в работе Докнера и др. [12] для случая дисконтированных выигрышей на бесконечном временном промежутке для единичной эластичности маргинальной полезности игроков. Отметим, что траектория $x^*(t)$ удовлетворяет условию устойчивости по Ляпунову. Получаем следующее значение общего ожидаемого выигрыша:

$$W(x(\vartheta), \vartheta) = \frac{\ln(x(\vartheta))}{\lambda} + \frac{(-\ln A - n)}{\lambda}.$$

Далее, для $\delta = 2$, соответствующем распределению Рэлея для стареющей системы, из (5.6) получаем

$$c_i^* = \frac{x \cdot e^{-2\lambda t^2}}{n \int_t^\infty e^{-2\lambda s^2} ds}.$$

Тогда оптимальный способ поведения при разработки ресурса должен определяться согласно следующей формуле

$$\begin{aligned} c_i^* &= \frac{2\sqrt{2}\sqrt{\lambda} \cdot e^{-2\lambda t^2}}{n(1 - \operatorname{erf}(\sqrt{2\lambda}t))}x = \\ &= \frac{2\sqrt{2}\sqrt{\lambda} \cdot e^{-2\lambda t^2}}{n(1 - 2\Phi_0(2\sqrt{\lambda}t))}x, \quad \text{где} \\ \operatorname{erf}(t) &= \frac{2}{\sqrt{\pi}} \int_0^t e^{-s^2} ds, \end{aligned}$$

$\Phi_0(t)$ – интегральная функция Лапласа.

Для периода приработки (раннего периода) возьмем $\delta = \frac{1}{2}$. Тогда из уравнения (5.6) получаем

$$c_i^* = \frac{x \cdot e^{-\frac{\lambda}{2}t^{1/2}}}{n \int_t^\infty e^{-\frac{\lambda}{2}s^{1/2}} ds}.$$

Следовательно, получаем оптимальные стратегии в управлениях с обратной связью:

$$c_i^* = \frac{\lambda^2}{4n(\lambda\sqrt{t} + 2)}x.$$

Таким образом, для модели разработки невозобновляемых ресурсов удалось получить оптимальные решения для всех трех фаз игры. Анализ полученных решений может являться предметом дальнейшего изучения.

6. Благодарности

Данное исследование было инициировано профессором Л.А. Петросяном в 1998 году. Кроме того, автор выражает благодарность профессору Дж. Заккуру и господину Д. Громову за информационную помощь во время ознакомления с проблематикой задачи. Ценные замечания и комментарии из области математического анализа и теории оптимального управления во время подготовки статьи были высказаны доцентом Д.С. Челкаком и профессором Дж. Марин-Солано.

СПИСОК ЛИТЕРАТУРЫ

1. Гаврилов Л.А., Гаврилова Н.С. *Биология продолжительности жизни*. М.: Наука, 1991.
2. Данфорд Н., Шварц Дж. Т. *Линейные операторы. Общая теория*. М: Изд-во иностранной литературы, 1962.
3. Колмогоров А.Н., Фомин С.В. *Элементы теории функций и функционального анализа*. М: "Наука 1976.
4. Матвеевский В.Р. *Надежность технических систем*. Учебное пособие. МГУ электроники и математики. Москва, 2002.
5. Петросян Л.А. *Дифференциальные игры преследования*. Л.: Изд-во Ленингр. ун-та, 1977.
6. Петросян Л.А., Данилов Н.В. *Кооперативные дифференциальные игры и приложения*. Томск: Изд-во Томского университета, 1985.
7. Петросян Л.А., Мурзов Н.В. *Теоретико-игровые проблемы в механике*// Литовский математический сборник. 1966. № VI-3. С. 423-433.
8. Петросян Л.А., Шевкопляс Е.В. *Кооперативные дифференциальные игры со случайной продолжительностью*// Вестник СПбГУ. 2000. Сер. 1. Вып. 4. С. 18-23.
9. Шевкопляс Е.В. *Кооперативные дифференциальные игры со случайной продолжительностью*. Автореф. дис. канд. физ.-мат. наук. СПб.: ООП НИИХ СПбГУ, 2004.
10. Шевкопляс Е.В. *О построении характеристической функции в кооперативных дифференциальных играх со случайной продолжительностью*// Труды Межд. семинара "Теория управления и теория обобщенных решений уравнений Гамильтона-Якоби", посвященного 60-летию академика А.И.Субботина. изд-во Уральского ун-та. Екатеринбург. 2006. № 1. С. 285-293.

11. Chang F.R. *Stochastic Optimization in Continuous Time*. Cambridge Univ. Press, 2004.
12. Dockner E.J., Jorgensen S., van Long N., Sorger G. *Differential Games in Economics and Management Science*. Cambridge Univ. Press, 2000.
13. Haurie A. *A Multigenerational Game Model to Analyze Sustainable Development*// *Annals of Operations Research*. 2005. V. 137. № 1. P. 369-386.
14. Henley E.J., Kumamoto H. *Reliability engineering and risk assessment*. Prentice-Hall, Inc., 1981.
15. Karp L. *Non-constant discounting in continuous time*// *Journal of Economic Theory*. 2007. V. 132. P. 557-568.
16. Marín-Solano J., Navas J. *Non-constant discounting in finite horizon: the free terminal time case*// *Journal of Economic Dynamics and Control*. 2009. V. 33. P. 666-675.
17. Yaari M.E. *Uncertain Lifetime, Life Insurance, and the Theory of the Consumer*// *The Review of Economic Studies*. 1965. V. 32. № 2. P. 137-150.

THE HAMILTON-JACOBI-BELLMAN EQUATION FOR A CLASS OF DIFFERENTIAL GAMES WITH RANDOM DURATION

Ekaterina Shevkoplyas, Faculty of Applied Mathematics and Control Processes, Saint-Peterburg State University, Cand. Sc.

(katya_shev@mail.ru).

Abstract: The class of differential games with random duration is studied. It turns out that the problem with random duration of the game can be simplified to the standard problem with infinite time horizon. The Hamilton-Jacobi-Bellman equation which help us to find the optimal solution under condition of random duration of the processes is derived. The results are illustrated with a game-theoretical model of non-renewable resource extraction. The problem is analyzed under condition of Weibull distribution for the random terminal time of the game.

Keywords: differential games, Hamilton-Jacobi-Bellman equation, random duration, non-renewable resource extraction.