LIETUVOS MATEMATIKOS RINKINYS, VIII Nr. 1, 1968
ЛИТОВСКИЙ МАТЕМАТИЧЕСКИЙ СБОРНИК, VIII № 1, 1968

УДК – 519.21

ОБ УСЛОВИЯХ ЕДИНСТВЕННОСТИ РЕШЕНИЯ УРАВНЕНИЯ БЕЛЛМАНА

Б. ГРИГЕЛИОНИС

1. Пусть $X^d=\{x_t^d,\ \zeta^d,\ \mathcal{M}_t^d,\ \mathbf{P}_x^d\}$, $d\in D$, – набор однородных марковских процессов с дискретным временем $t=0,\ 1,\ \dots$ на полукомпакте $(E, \mathcal{B})$ (по поводу обозначений и терминологии см. монографию [1]). Предположим для простоты, что $D=\{1,\ \dots,\ m\}$ – конечное множество. Управляемым (посредством управления δ) случайным процессом называют набор $X^\delta=\{\xi_t,\ \zeta,\ \delta,\ \mathcal{M}_t^\delta,\ \mathbf{P}_x^\delta\}$, где $\xi_t=x_t^{\delta_t}$, δ_t – индекс наблюдаемого в момент t процесса из набора X^d, $d\in D$, определяемый значениями процесса ξ_s, $s=0,\ 1,\ \dots,\ t-1$, $\zeta=\max\limits_{d\in D}\zeta^d$, $\mathcal{M}_t^\delta$ – σ-алгебра событий, определяемых течением процесса ξ_t до момента времени t, а $\mathbf{P}_x^\delta$ – вероятностные меры, индуцированные процессом ξ_t при начальном состоянии x (точное и подробное определение управляемого процесса см. в [2]). Стратегией называют пару $\Delta=(\delta,\ \tau)$, где δ – управление, а τ – марковский момент относительно σ-алгебр $\mathcal{M}_t^\delta$, называемый моментом остановки управляемого процесса.

Пусть далее задана неотрицательная ограниченная измеримая функция $g(x)$. Определим функцию

$$s(x)=\sup_{\Delta\in\Delta}\mathbf{M}_x^\delta g(\xi_\tau),$$

где Δ – класс всех стратегий, а $\mathbf{M}_x^\delta\eta$ – математическое ожидание случайной величины η по мере $\mathbf{P}_x^\delta$, называемую ценой.

Одной из основных задач теории оптимальной остановки управляемых процессов является построение ε-оптимальных ($\varepsilon\geqslant 0$) стратегий, т. е. стратегий $\Delta_\varepsilon=(\delta_\varepsilon,\ \tau_\varepsilon)$ таких, что

$$\mathbf{M}_x^{\delta_\varepsilon} g(\xi_{\tau_\varepsilon})\geqslant s(x)-\varepsilon$$

для всех $x\in E$. ε-оптимальные стратегии легко построить, когда известна цена $s(x)$ (см., например, работы [3 – 12]). Отсюда видно, что важной является задача нахождения цены.

Известно (см., например, [6], [11], [2]), что цена $s(x)$ является $\hat{T}$-наименьшей эксцессивной мажорантой (н.э.м.) функции $g(x)$. (Неотрицательную измеримую функцию $f(x)$ называем $\hat{T}$-эксцессивной, если $\hat{T}f(x)=\max\limits_{d\in D}\mathbf{M}_x^d f(x_1^d)\leqslant f(x)$ для всех $x\in E$; $f(x)$ называем $\hat{T}$-н.э.м. функции $g(x)$, если $f(x)\geqslant g(x)$ и $f(x)\leqslant f_1(x)$, где $f_1(x)$ – любая $\hat{T}$-э.ф., $f_1(x)\geqslant g(x)$, $x\in E$). Рассуждая аналогично рассуждениям при доказательстве леммы 1 в [10], получаем, что $s(x)$ является решением следующего уравнения Беллмана:

$$s(x)=\max\{g(x),\ \hat{T}s(x)\}. \tag{1}$$

Если мы нашли какое-то решение этого уравнения, то возникает вопрос, когда оно совпадает с ценой. Легко убедиться, что уравнение (1) в общем случае имеет не единственное решение даже среди ограниченных. Например, если процессы X^d, $d \in D$ не обрывающиеся, т. е. $\mathrm{P}_x^d\{\zeta^d = \infty\} = 1$ для всех $x \in E$ и $d \in D$, то любая функция $f(x) \equiv C \geqslant \sup\limits_{x \in E} g(x)$ является решением уравнения (1).

Целью нашей работы является исследование условий единственности решения уравнения (1) в определенном подмножестве ограниченных измеримых функций.

2. Пусть G — некоторое борелевское подмножество E. Определим оператор T_G равенством:

$$T_G f(x) = \max_{d \in D} \int_G f(y) P^d(x, dy),$$

где $P^d(x, \Gamma) = \mathrm{P}_x^d\{x_1^d \in \Gamma\}$.

Обозначим

$$\rho_n(G) = \sup_{x \in G} T_G^n 1(x),$$

где T_G^n — n-я степень оператора T_G.

Теорема 1. *Пусть $s_1(x)$ и $s_2(x)$ — два ограниченных измеримых решения уравнения (1), совпадающих при $x \in E \setminus G$. Если при некотором n $\rho_n(G) < 1$, то $s_1(x) \equiv s_2(x)$.*

Доказательство. Из равенств

$$s_i(x) = \max\{g(x), \hat{T} s_i(x)\}, \quad i = 1, 2,$$

следует, что

$$r(x) = |s_1(x) - s_2(x)| \leqslant |\hat{T} s_1(x) - \hat{T} s_2(x)|. \tag{2}$$

Но,

$$\hat{T} s_1(x) - \hat{T} s_2(x) \leqslant \hat{T}\big(s_1(x) - s_2(x)\big), \tag{3}$$

и из (2) и (3) получаем, что

$$r(x) \leqslant \hat{T} r(x)$$

или

$$r(x) \leqslant \max_{d \in D} \int_E r(y) P^d(x, dy).$$

Поскольку $r(x) = 0$ при $x \in E \setminus G$, то

$$r(x) \leqslant T_G r(x). \tag{4}$$

В силу монотонности операторов T_G из (4) имеем, что

$$r(x) \leqslant T_G^n r(x).$$

при любом n.

Отсюда

$$\sup_{x \in G} T_G^n r(x) \geqslant \sup_{x \in G} r(x) = \sup_{x \in E} r(x). \tag{5}$$

Снова используя свойство монотонности T_G и то, что $T_G\big(cf(x)\big) = cT_G f(x)$, $c \geqslant 0$, из (5) находим, что

$$\sup_{x \in E} r(x) \leqslant \sup_{x \in E} r(x) \cdot \sup_{x \in G} T_G^n 1(x) = \rho_n(G) \sup_{x \in E} r(x).$$

Если при некотором n $\rho_n(G) < 1$, то $r(x) \equiv 0$, т. е. $s_1(x) \equiv s_2(x)$. Теорема 1 доказана.

Замечание 1. Числа $\rho_n(G)$ монотонно не возрастают при увеличении n. Действительно, поскольку $\max\limits_{d \in D} P^d(x_{n-1}, G) \leqslant 1$ и

$$\rho_n(G) = \sup_{x \in G} \underbrace{\max_{d \in D} \int_G \dots \max_{d \in D} \Big(\int_G}_{n-1 \text{ раз}} \Big(\max_{d \in D} P^d(x_{n-1}, G) \Big) \times$$

$$\times P^d(x_{n-2}, dx_{n-1}) \dots P^d(x, dx_1),$$

то

$$\rho_n(G) \leqslant \sup_{x \in G} \underbrace{\max_{d \in D} \int_G \dots \max_{d \in D} \Big(\int_G}_{n-2 \text{ раз}} \Big(\max_{d \in D} P^d(x_{n-2}, G) \Big) \times$$

$$\times P^d(x_{n-3}, dx_{n-2}) \dots P^d(x, dx_1) = \rho_{n-1}(G).$$

Если условия теоремы 1 выполнены при $n = 1$, то мы получаем следующее утверждение.

Следствие 1. Если $\sup\limits_{x \in E} \max\limits_{d \in D} P^d(x, G) < 1$, то два любых ограниченных измеримых решения $s_1(x)$ и $s_2(x)$ уравнения (1), совпадающих при $x \in E \backslash G$, тождественны.

Следствия 2. Если $\sup\limits_{x \in E} \max\limits_{d \in D} P^d(x, E) < 1$, то уравнение (1) имеет единственное ограниченное измеримое решение (ср. с теоремой 2 гл. IV в [13]).

В самом деле, существование решения этого уравнения следует из существования $\hat{T}$-н.э.м. функции $g(x)$, а единственность – из следствия 1.

Следствие 3. Пусть $s_1(x)$ – некоторое ограниченное измеримое решение уравнения (1) и $G = \{x : s_1(x) > g(x)\}$. Если при некотором n $\rho_n(G) < 1$, то $s_1(x) \equiv s(x)$, т. е. $s_1(x)$ является ценой.

Из (1) следует, что $s_1(x) \geqslant g(x)$ и $s_1(x) \geqslant \hat{T} s_1(x)$. Значит, $s_1(x)$ является $\hat{T}$-эксцессивной мажорантой функции $g(x)$. Но поскольку $s(x)$ – $\hat{T}$-н. э. м. функции $g(x)$, то $g(x) \leqslant s(x) \leqslant s_1(x)$ и $s_1(x) = s(x)$ при $x \in E \backslash G$. Утверждение следствия 3 вытекает из теоремы 1.

Замечание 2. Аналогичные доказанным результаты можно получить и для более общего уравнения:

$$s(x) = \max \Big\{ g(x), \sup_{d \in D} \Big(c^d(x) + T^d s(x) \Big) \Big\},$$

где $g(x)$, $c^d(x)$ – заданные функции, $T^d f(x) = \mathbf{M}_x^d f(x_1^d)$, а множество D не обязательно конечно.

3. Рассмотрим теперь случай, когда D содержит лишь один элемент, т. е. случай оптимальной остановки марковского процесса.

Уравнение (1) тогда принимает вид:

$$s(x) = \max\{ g(x), Ts(x) \}, \tag{6}$$

где $Tf(x) = \int\limits_E f(y) P(x, dy)$, а $P(x, \Gamma) = \mathbf{P}_x\{ x_1 \in \Gamma \}$ – переходная функция вероятностей за один шаг рассматриваемого марковского процесса

$$X = \{ x_t, \zeta, \mathcal{M}_t, \mathbf{P}_x \}.$$

Обозначим $\tau(G)=\inf\{t : x_t \bar{\in} G\}$ — момент первого выхода процесса X из борелевского множества G.

Теорема 2. *Если* $\lim\limits_{n\to\infty} \inf\limits_{x\in G} \mathsf{P}_x\{\tau(G)\leqslant n\}>0$, *то два любых ограниченных измеримых решения уравнения* (6), *совпадающих при* $x\in E\setminus G$, *тождественны*.

Доказательство. Имеем, что

$$\rho_n(G)=\sup_{x\in G} \underbrace{\int\limits_G \cdots \int\limits_G}_{n-1 \text{ раз}} P(x_{n-1}, G)\, P(x_{n-2}, dx_{n-1}) \cdots P(x, dx_1)=$$

$$=\sup_{x\in G} \mathsf{P}_x\{\tau(G)>n\}\leqslant 1-\inf_{x\in G} \mathsf{P}_x\{\tau(G)\leqslant n\}.$$

Если же выполнено условие теоремы, то при достаточно большом n $\rho_n(G)<1$ и наше утверждение следует из теоремы 1.

Из доказанной теоремы следует теорема 2 в [10]. Из нее также легко вытекает

Следствие 4. Пусть $s_1(x)$ — некоторое ограниченное измеримое решение уравнения (6) и для $G=\{x : s_1(x)>g(x)\}$ $\lim\limits_{n\to\infty} \inf\limits_{x\in G} \mathsf{P}_x\{\tau(G)\leqslant n\}>0$. Тогда $s_1(x)$ является ценой.

4. Пусть $X=\{x_t, \zeta, \mathcal{M}_t, \mathsf{P}_x\}$ — стандартный марковский процесс с непрерывным временем $t\geqslant 0$ на полукомпакте $(E, \mathcal{B})$ (см. [1]). Обозначим $\tau(U)$ момент первого выхода процесса X из множества U, $T_t f(x)=\mathsf{M}_x f(x_t)$.

Аналогом уравнения (6) в этом случае является уравнение

$$s(x)=\max\{g(x), T_{\tau(U)}s(x)\}, \tag{7}$$

верное для любой окрестности U точки x, содержащейся в некоторой достаточно малой окрестности $\tilde{U}(x)$ с компактным замыканием (см. [10], [14]). Условия единственности решения уравнения (7) изучались в работах [10] и [14]. Мы сейчас докажем результат, аналогичный теореме 2.

Пусть G — некоторое открытое множество, $\pi(x)=\mathsf{P}_x\{\tau(G)<\zeta\}$.

Теорема 3. *Пусть* $s_1(x)$ *и* $s_2(x)$ — *два измеримых* C_0 — *непрерывных* (*т. е. непрерывных в естественной топологии, связанной с процессом* X) *решения уравнения* (7), *совпадающих на множестве* $E\setminus G$, *и* $|s_1(x)-s_2(x)|\leqslant K\pi(x)$ *для всех* $x\in E$, *где* K — *некоторая константа. Если* $\lim\limits_{t\to\infty} \inf\limits_{x\in G} \mathsf{P}_x\{\tau(G)\leqslant t\}>0$, *то* $s_1(x)\equiv s_2(x)$.

Доказательство. Обозначим $r(x)=|s_1(x)-s_2(x)|$. В работе [14] при условиях теоремы показано (см. доказательство теоремы 3 в [14]), что при любом $t\geqslant 0$

$$r(x)\leqslant \int\limits_G r(y)\,\tilde{P}(t, x, dy), \tag{8}$$

где $\tilde{P}(t, x, \Gamma)=\mathsf{P}_x\{x_t\in\Gamma, \tau(G)>t\}$.

Тогда из (8) имеем, что

$$\sup_{x\in E} r(x)\leqslant \sup_{x\in E} r(x)\cdot \sup_{x\in G} \mathsf{P}_x\{\tau(G)>t\}. \tag{9}$$

Поскольку $\lim\limits_{t\to\infty} \inf\limits_{x\in G} \mathsf{P}_x\{\tau(G)\leqslant t\}>0$, то найдется такое достаточно большое t, что $\sup\limits_{x\in G} \mathsf{P}_x\{\tau(G)>t\}<1$ и из (9) получим, что $r(x)\equiv 0$. Теорема 3 доказана.

Предположим далее, что функция $g(x)$ ограничена и непрерывна. Тогда существует цена $s(x)$ в задаче об оптимальной остановке процесса X, и если, кроме того, $s(x)$ непрерывна снизу, то она является решением уравнения (7) (см. [14]). В этих предположениях верно следующее утверждение.

Следствие 5. Пусть $s_1(x)$ — какое-то измеримое C_0 — непрерывное решение уравнения (7), такое что множество $G=\{x: s_1(x)>g(x)\}$ открыто. Если $\inf\limits_{x\in G}\pi(x)>0$ и $\lim\limits_{t\to\infty}\inf\limits_{x\in G}\mathbf{P}_x\{\tau(G)\leqslant t\}>0$, то $s_1(x)\equiv s_2(x)$.

В самом деле, легко находим, что $s_1(x)=s(x)$ при $x\in E\setminus G$ (см. [14]). Далее из условий следствия вытекает существование такой достаточно большой константы K, что $|s_1(x)-s_2(x)|\leqslant K\pi(x)$. Остается применить теорему 3.

5. Хотя условия единственности решения уравнения Беллмана, полученные в п.п. 2–4, имеют достаточно общий характер, однако существуют важные примеры, где эти условия неприменимы. Пусть, например, E — некоторое неограниченное подмножество евклидового пространства R^n, подмножество G — такое, что $E\setminus G$ ограничено, а рассматриваемый марковский процесс X удовлетворяет условию: для всех $x\in G$ существует константа K, такая, что $\mathbf{P}_x\{|x_1-x|\leqslant K\}=1$. В этом случае, очевидно, $\inf\limits_{x\in G}\mathbf{P}_x\{\tau(G)\leqslant t\}=0$ при каждом конечном t и условия теорем 2 и 3 не выполняются, в то же время во многих случаях, как это легко следует из оценок, полученных в п.п. 2–4, единственность решения в рассматриваемом классе решений имеет место.

Сформулируем поэтому соответствующие утверждения, аналогичные теоремам 1–3.

Теорема 1′. *Пусть $s_1(x)$ и $s_2(x)$ — два ограниченных измеримых решения уравнения (1), совпадающих при $x\in E\backslash G$. Если при каждом $x\in G$ $\lim\limits_{n\to\infty}T_G^n 1(x)=0$, то $s_1(x)\equiv s_2(x)$.*

Доказательство. Поскольку при любом $x\in E$ и n $|s_1(x)-s_2(x)|=r(x)\leqslant T_G^n r(x)$, то $r(x)\leqslant\sup\limits_{x\in E}r(x)\cdot\lim\limits_{n\to\infty}T_G^n 1(x)=0$ при каждом x. Следовательно, $s_1(x)\equiv s_2(x)$.

Теорема 2′. *Если при каждом $x\in G$ $\mathbf{P}_x\{\tau(G)=\infty\}=0$, то два любых ограниченных измеримых решения уравнения (6), совпадающих при $x\in E\backslash G$, тождественны.*

Действительно, в этом случае $T_G^n 1(x)=\mathbf{P}_x\{\tau(G)>n\}$ и утверждение теоремы 2′ следует из теоремы 1′.

Теорема 3′. *Пусть $s_1(x)$ и $s_2(x)$ — два измеримых C_0 — непрерывных решения уравнения (7), совпадающих на множестве $E\backslash G$, и $|s_1(x)-s_2(x)|\leqslant K\leqslant\pi(x)$ для всех $x\in E$, где K — некоторая константа. Если при каждом $x\in G$ $\mathbf{P}_x\{\tau(G)=\infty\}=0$, то $s_1(x)\equiv s_2(x)$.*

Из (8) находим, что при каждом $t\geqslant 0$

$$r(x)\leqslant\sup_{x\in E}r(x)\cdot\mathbf{P}_x\{\tau(G)>t\}.$$

Отсюда и из условий теоремы 3 следует, что $r(x)\equiv 0$, т. е. $s_1(x)\equiv s_2(x)$. Аналогично можно видоизменить и формулировки следствий 3–5.

В заключение выражаю искреннюю признательность А. Н. Ширяеву за ряд ценных замечаний.

Институт физики и математики
Академии наук Литовской ССР

Поступило в редакцию
21.VIII.1967

Литература

1. Е. Б. Дынкин, Марковские процессы, М., Физматгиз, 1963.

2. Б. И. Григелионис, А. Н. Ширяев, Об управляемых марковских процессах и задаче Стефана, Проблемы передачи информации, VI, 1 (1968), 60–72.

3. J. L. Snell, Applications of martingale system theorems, Trans. Amer. Math. Soc., **73** (1952), 293–312.

4. Y. S. Chow, H. Robbins, On optimal stopping rules, Z. Wahrscheinlichkeitstheorie, **2**, 1 (1963), 33–49.

5. Е. Б. Дынкин, Оптимальный выбор момента остановки марковского процесса, ДАН СССР, **150**, 2(1963), 238–240.

6. Н. В. Крылов, Об оптимальной остановке управляемой цепи. Сб. „Оптимальное управление и теория информации" (тезисы докладов на VII Всесоюзном совещании по теории вероятн. и матем. статист., Тбилиси, 1963), Изд-во ин-та математики АН УССР, Киев, 1963, 11–15.

7. Н. В. Крылов, О существовании ε-оптимальных однородных марковских стратегий для управляемых цепей, ДАН СССР, **155**, 4(1964), 747–750.

8. Н. В. Крылов, Построение оптимальной стратегии для конечной управляемой цепи, Теория вероятн. и ее применен., **X**, 1(1965), 51–60.

9. А. Н. Ширяев, Последовательный анализ и управляемые случайные последовательности (дискретное время), Кибернетика, 3(1965), 1–24.

10. Б. И. Григелионис, А. Н. Ширяев, О задаче Стефана и оптимальных правилах остановки марковских процессов, Теория вероятн. и ее применен., **XI**, 4(1966), 612–631.

11. G. W. Haggstrom, Optimal stopping and experimental design, Ann. Math. Statist., 37, 1(1966), 7–29.

12. А. Н. Ширяев, Некоторые новые результаты в теории управляемых случайных процессов, Trans. IV Prague Conference inform. theory..., Prague, 1967, 131–203.

13. Р. Беллман, Динамическое программирование, ИЛ, М., 1960.

14. Б. Григелионис, Об оптимальной остановке марковских процессов, Лит. мат. сб., **VII**, 2(1967), 265–279.

APIE BELMANO LYGTIES SPRENDINIO VIENATINUMO SĄLYGAS

B. GRIGELIONIS

(*Reziumė*)

Darbe gautos bendros sąlygos, kai apibrėžtoje išmatuojamų aprėžtų funkcijų klasėje (1) lygtis turi vienintelį sprendinį.

ON THE UNIQUENESS CONDITIONS OF SOLUTION OF BELLMAN'S EQUATION

B. GRIGELIONIS

(*Summary*)

The general conditions for the equation (1) to have the unique solution in the determined class of the measurable bounded functions are obtained.