

УДК 517.977:519.22

## О ДЕТЕРМИНИРОВАННОМ ПОДХОДЕ К ЗАДАЧЕ СТОХАСТИЧЕСКОГО ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

Н. С. Исмагилов<sup>1</sup>, Ф. С. Насыров<sup>2</sup>

<sup>1</sup>niyaz.ismagilov@gmail.com, <sup>2</sup>farsagit@yandex.ru

ФГБОУ ВПО «Уфимский государственный авиационный технический университет» (УГАТУ)

**Аннотация.** Рассматривается детерминированный подход к задачам стохастического оптимального управления, основанный на формуле разложения решения стохастического дифференциального уравнения, который позволяет рассматривать стохастическую задачу как параметризованное семейство детерминированных задач. Показано, что при модификации соответствующим образом функционала потерь, оптимум детерминированной задачи достигается на неупреждающей функции управления, которая также является оптимальной для стохастической задачи.

**Ключевые слова:** стохастическое оптимальное управление; стохастические дифференциальные уравнения; упреждающие стохастические дифференциальные уравнения; принцип максимума; симметричный интеграл; детерминированный подход.

### ВВЕДЕНИЕ

В настоящей работе исследуется связь между стохастическими и детерминированными задачами оптимального управления. Некоторые результаты из ранних исследований, направленных на установление связи между стохастическими и детерминированными задачами оптимизации, представлены в работах [1, 2] в 70-х гг. В этих работах показано, что в некоторых случаях стохастическая задача может быть рассмотрена с детерминированной точки зрения, сформулирована одна из значимых проблем перехода к детерминированным задачам, заключающаяся в обеспечении неупреждаемости решений, и предложен метод, который путем изменения функционала потерь множителями Лагранжа позволяет накладывать ограничения неупреждаемости.

Позже, в начале 90-х гг. в работах [3, 4], исследовалась связь между линейными и нелинейными стохастическими и детерминированными задачами оптимального управления.

В работе [4] подробно исследована нелинейная задача управления. Доказана возможность представления стохастической

задачи в виде параметризованного семейства детерминированных задач. Приведена явная формула для множителя Лагранжа, используемого для наложения ограничения на неупреждаемость.

Для сведения стохастической задачи к детерминированной авторами работы [4] используется формула для разложения решения СДУ (стохастического дифференциального уравнения) Ито

$$dx_t = b(x_t, u_t)dt + \sigma(x_t)dW_t, \quad x(0) = x_0,$$

которая имеет следующий вид

$$x_t(x) = \xi_t \circ \eta_t(x),$$

$$d\xi_t(y) = \sigma(\xi_t(y)) \circ dW_t,$$

$$\frac{d\eta_t(x)}{dt} = \left(\frac{\partial \xi_t}{\partial x}\right)^{-1}(\eta_t(x))b(\xi_t \circ \eta_t(x), u_t),$$

$$\eta_0(x) = x.$$

Такое разложение позволяет определить решение СДУ для упреждающих управлений. Решение при этом понимается в смысле работы [5]. Определение решения уравнения для упреждающих управлений позволяет авторам работы [4] перейти к детерминированным задачам, при этом неупреждаемость решений достигается введением множителей Лагранжа.

Для анализа детерминированной оптимизационной задачи и поиска подходящих значений множителей Лагранжа используется метод динамического программирования.

В ходе доказательства основных утверждений авторам приходится сталкиваться с уравнением Гамильтона–Якоби–Беллмана, для решения которого, а так же для представления функции выигрыша через фазовую координату, авторам приходится пользоваться результатами работ [6], [7] по приближению решений СДУ решениями ОДУ (обыкновенных дифференциальных уравнений), а также стохастическим методам характеристик [8], которые приводят к довольно сложным вычислениям и сильным ограничениям.

Настоящая работа основана на результатах работы [4], являясь ее дополнением и, в некотором смысле, упрощением. Применение представленного в работе [9] понятия симметричного интеграла позволяет по-новому определить решение задачи с упреждающим управлением. В отличие от работы [4] для сведения стохастической задачи к детерминированной используется формула разложения решения СДУ (см. [10]), которая позволяет существенно упростить технику вычисления и ослабить многие предположения.

### ОСНОВНОЙ РЕЗУЛЬТАТ

Пусть  $(\Omega, F, (F)_{0 \leq t \leq T}, P)$  — полное вероятностное пространство, наделенное естественной фильтрацией одномерного стандартного винеровского процесса,  $W_t, t \in [0, T], T > 0$ . Обозначим через  $A$  множество функций  $u: [0, T] \times \Omega \rightarrow U \subset \mathbf{R}^m$ , измеримых относительно  $\sigma$ -алгебры  $\mathcal{A}[0, T] \times F$ , а через  $N$  — его подмножество, состоящее из неупреждающих функций.

Рассмотрим задачу оптимального управления одномерным процессом

$$dx_t = b(x_t, u_t)dt + \sigma(x_t)dW_t, \quad x(0) = x_0, \quad (1)$$

$$\inf_{u \in N} \mathbf{E} \left[ \int_0^T f(t, x_t, u_t)dt + g(x_T) \right] = \inf_{u \in N} \mathbf{E}J(u), \quad (2)$$

здесь  $x_t: [0, T] \times \Omega \rightarrow \mathbf{R}$  — фазовая координата,  $u_t$  — управляющая функция,  $b: \mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R}$ ,  $\sigma, g: \mathbf{R} \rightarrow \mathbf{R}$ ,  $f: [0, T] \times \mathbf{R} \times \mathbf{R} \rightarrow \mathbf{R}$  — измеримые функции,  $\mathbf{E}$  — символ математического

ожидания,  $\mathbf{E}J(u)$  — функционал качества. Всюду далее будем использовать обозначение

$$J(u) = \int_0^T f(t, x_t, u_t)dt + g(x_T). \quad (3)$$

**Лемма 1.** Пусть функционал качества  $J(u)$  определен и имеет смысл для всех  $u \in A$ , тогда

$$\inf_{u \in A} \mathbf{E}J(u) = \mathbf{E} \inf_{u \in A} J(u) \quad (4)$$

**Доказательство.** Положим  $M = \inf_{u \in A} \mathbf{E}J(u)$  и  $m = m(\omega) = \inf_{u \in A} J(u)$ . Тогда для любого  $u \in A$  имеем

$$\mathbf{E}J(u) \geq M \quad \text{и} \quad J(u) \geq m. \quad (5)$$

Кроме того, для любого  $\varepsilon > 0$  найдутся  $u^* \in A$  и  $\bar{u} \in A$  такие, что

$$\mathbf{E}J(u^*) - M < \varepsilon \quad \text{и} \quad J(\bar{u}) - m < \varepsilon. \quad (6)$$

Комбинируя (5) и (6), имеем

$$|\mathbf{E}m - M| < \varepsilon, \quad (7)$$

что в силу произвольности  $\varepsilon$  равносильно (4).  $\square$

**Следствие 1.** Если точная нижняя грань в первой части (4) достигается на функции  $\bar{u}$ , то на той же функции достигается инфимум и в левой части.

**Доказательство.** Следует из того, что  $\inf_u \mathbf{E}J(u) = \mathbf{E} \inf_u J(u) = \mathbf{E}J(\bar{u})$ .  $\square$

Лемма 1 позволяет рассматривать параметризованную детерминированную задачу минимизации функционала  $J(u)$  функциями  $u_\omega \in A$ . Кроме определения  $J(u)$  для упреждающих управлений, такой подход также требует определить дифференциальное ограничение, накладываемое уравнением (1), так как оно не определено для упреждающих  $u$  и не имеет потраекторного смысла.

Уравнение (1) для упреждающих управлений может быть определено при помощи симметричного интеграла. Симметричный интеграл, впервые введенный Насыровым в работе [9], определяется для любых непрерывных функций, выступающих в качестве интеграторов. Построенный по траекториям винеровского процесса симметричный интеграл п.н. совпадает с интегралом Стратоновича по тому же процессу

(подробнее см. [10]). При этом он не требует неупреждаемости коэффициентов, и для него может быть определено решение следующего дифференциального уравнения (семейства параметризованных дифференциальных уравнений с симметричным интегралом) в потракторном смысле

$$dx(t, \omega) = \tilde{b}(x(t, \omega), u(t, \omega))dt + \sigma(x(t, \omega)) * dW(t, \omega), \quad x(0) = x_0, \quad (8)$$

где  $u \in A$ , а  $\tilde{b}$  определен ниже. Для каждого  $\omega$  решением (8) называется функция  $x(t, \omega)$ , для которой определены интегралы в правой части и которая обращает интегральное равенство, соответствующее (8), в тождество. Известно (см. [10]), что это решение при неупреждающих коэффициентах п. н. совпадает с решением соответствующего СДУ в форме Стратоновича. Ввиду этого, если уравнение (1) переписать в форме Стратоновича

$$dx_t = \tilde{b}(x_t, u_t)dt + \sigma(x_t) \circ dW_t, \quad x(0) = x_0, \quad (9)$$

$$\tilde{b}(x_t, u_t) = b(x_t, u_t) - \frac{1}{2} \sigma(x_t) \sigma'(x_t),$$

то для упреждающих  $u$  его можно интерпретировать в форме (8). Для неупреждающих же управлений (1), (9) и (8) совпадают.

Таким образом, определив решение для более широкого класса управляющих функций  $A$ , можно корректно сформулировать задачу упреждающего оптимального управления: требуется найти  $u \in A$ , доставляющий минимум функционалу  $J(u)$  при условии (8)

$$\inf_{u \in A} \mathbf{E} \left[ \int_0^T f(t, x_t, u_t) dt + g(x_T) \right] = \inf_{u \in A} \mathbf{E} J(u). \quad (10)$$

Далее нам понадобится теорема о разложении решений дифференциальных уравнений с симметричным интегралом, доказательство которой приведено в [10].

**Теорема 1.** Пусть  $b(t, x), \sigma(t, x): [0, T] \times \mathbf{R} \rightarrow \mathbf{R}$  — непрерывные функции, равномерно липшицевые по  $x$ . Пусть, кроме того,  $\sigma'_t$  непрерывна, и существует константа  $C > 0$ , такая что  $|\sigma(t, x)| \geq C$ . Тогда решение уравнения с симметричным интегралом

$$dx(t) = b(t, x(t))dt + \sigma(t, x(t)) * dW(t), \quad x(0) = x_0,$$

имеет вид  $x(t) = \Phi(t, y(t) + W(t))$ , где  $\Phi(t, v)$  есть явное выражение для  $\phi$  из соотношения

$$\int_0^\phi \frac{d\phi}{\sigma(t, \phi)} = v,$$

а  $y(t)$  является решением следующей задачи Коши

$$\dot{y}(t) = \frac{b(t, \Phi(t, y(t) + W(t))) - \Phi'_t(t, y(t) + W(t))}{\sigma(t, \Phi(t, y(t) + W(t)))},$$

$$y(0) = y_0, \quad y_0 : x_0 = \Phi(0, y_0 + W(0)).$$

Воспользуемся теоремой 1 и докажем другую теорему, позволяющую свести упреждающие стохастические задачи оптимального управления к детерминированным задачам.

**Теорема 2.** Пусть  $b(x, u)$  и  $\sigma(x)$  в уравнении (8) непрерывны вместе со своими производными  $b'_x, b'_u$  и  $\sigma'$ . Пусть, кроме того, существует константа  $C > 0$  такая, что  $|\sigma(x)| \geq C$ . Тогда задача (8), (10) эквивалентна семейству детерминированных задач следующего вида

$$\dot{y}(t) = \frac{b(\Phi(y(t) + W(t, \omega)), u(t))}{\sigma(\Phi(y(t) + W(t, \omega)))}, \quad (11)$$

$$y(0) = \Phi^{-1}(x_0),$$

$$\inf_{u \in A} \int_0^T f(t, \Phi(y(t) + W(t, \omega)), u(t)) dt + g(\Phi(y(T) + W(T, \omega))), \quad (12)$$

где  $\Phi$  — некоторая известная функция,  $\Phi^{-1}$  — функция, обратная к ней.

Под эквивалентностью в теореме 2 понимается равенство решений двух задач.

**Доказательство.** Из теоремы 1 следует, что решение (8) имеет вид

$$x(t, \omega) = \Phi(y(t, \omega) + W(t, \omega)), \quad (13)$$

где  $\Phi(v)$  — функция определяемая из соотношения

$$\int_0^\phi \frac{d\phi}{\sigma(\phi)} = v,$$

а  $y(t)$  — решение задачи Коши для ОДУ

$$\dot{y}(t) = \frac{b(\Phi(y(t) + W(t, \omega)), u(t))}{\sigma(\Phi(y(t) + W(t, \omega)))}, \quad (14)$$

$$y(0) = \Phi^{-1}(x_0).$$

Функционал (3), записанный с использованием (13), имеет вид (12). Лемма 1 и его следствие гарантируют, что если измеримая функция  $\bar{u}$  доставляет минимум функционалу (12), то она же доставляет минимум функционалу (10), и выполнено

$$\mathbf{E}J(\bar{u}) = \mathbf{E} \int_0^T f(t, \Phi(y(t) + W(t, \omega)), \bar{u}(t)) dt + g(\Phi(y(T) + W(T, \omega))).$$

Теперь, из того, что уравнение (8) эквивалентно паре (13), (14), следует, что задача (11), (12) эквивалентна задаче (8), (10).

□

Теорема 2 позволяет проследить связь между стохастическими и детерминированными задачами оптимального управления, которая заключается в том, что упреждающая стохастическая задача может быть сведена к параметризованному семейству детерминированных задач. Если в стохастическую задачу (8), (10) добавить ограничение, которое гарантирует достижение минимума на неупреждающей функции  $u \in N$ , то теорема 2 может быть использована для решения задачи (1), (2). Такое ограничение неупреждаемости может быть наложено путем добавления к функционалу качества множителей Лагранжа и решения несколько измененной задачи:

$$\inf_{u \in A} \mathbf{E} \left[ \int_0^T (f(t, x_t, u_t) + \lambda u_t) dt + g(x_T) \right]. \quad (15)$$

Основным результатом работы является следующая теорема, которая утверждает, что такой множитель Лагранжа существует, и предоставляет явную формулу для нее.

**Теорема 3** Пусть функции  $b, \sigma, f, g$  непрерывны вместе со своими производными  $b'_x, b'_u, b''_{uu}, \sigma', f'_x, f'_u, f''_{uu}, g'$ . Пусть, кроме того, существует константа  $C > 0$  такая, что  $|\sigma(x)| \geq C$ . Далее, предположим, что существует решение  $(x_t^*, u_t^*)$  задачи (1), (2), и для всех  $(t, x, u) \in [0, T] \times \mathbf{R} \times \mathbf{U}$  выполнено

$$f''_{uu}(t, x, u) - \psi \frac{b''_{uu}(x, u)}{\sigma(x)} > 0, \quad (16)$$

где  $\psi(t)$  есть решение задачи Коши

$$\dot{\psi} = -\psi \frac{b'_x \sigma - b \sigma'}{\sigma} + \psi \frac{b''_{xu} \sigma - b'_u \sigma'}{\sigma} - \sigma f''_{ux},$$

$$\psi(T) = -g'(x_T^*) \sigma(x_T^*).$$

Тогда следующая задача детерминированного оптимального управления

$$\dot{y}(t) = \frac{b(\Phi(y(t) + W(t, \omega)), u(t))}{\sigma(\Phi(y(t) + W(t, \omega)))}, \quad (17)$$

$$y(0) = \Phi^{-1}(x_0),$$

$$\inf_u \int_0^T [f(t, \Phi(y(t) + W(t, \omega)), u(t)) + \lambda u(t)] dt + g(\Phi(y(T) + W(T, \omega))),$$

(18)

где

$$\lambda = \lambda(t, y, \omega) = \frac{\psi(t) b'_u(\Phi(y + W(t, \omega)), u^*(t))}{\sigma(\Phi(y + W(t, \omega)))} - f'_u(t, \Phi(y + W(t, \omega)), u^*(t)),$$

имеет решение, совпадающее с  $u^*$ .

**Доказательство.** Из теоремы 2 следует, что задача (8), (15) эквивалентна задаче (17), (18), для решения которой будем использовать принцип максимума Понтрягина. Функция Гамильтона–Понтрягина имеет вид

$$H = -(f + \lambda u) + \psi \frac{b}{\sigma},$$

где  $\psi = \psi(t)$  — сопряженная переменная, подчиняющаяся уравнению

$$\dot{\psi} = -\frac{\partial H}{\partial y} = -f'_x \Phi' - \lambda'_y u + \psi \frac{b'_x \Phi' \sigma - b \sigma' \Phi'}{\sigma^2}$$

с начальным условием  $\psi(T) = -g'(\Phi(y(T) + W(T, \omega))) \Phi'(y(T) + W(T, \omega))$ .

При выполнении (16) вторая частная производная  $H$  по переменной  $u$  отрицательна, следовательно, функция  $H$  является выпуклой вверх. Равенство нулю первой производной по  $u$  при  $u = u^*$

$$H'_u|_{u=u^*} = -a_0(f'_u(t, \Phi, u) + \frac{\psi(t)b'_u(\Phi, u^*)}{a_0\sigma(\Phi)} - f'_u(t, \Phi, u^*)) + \psi \frac{b'_u(\Phi, u)}{\sigma(\Phi)}|_{u=u^*} = 0$$

гарантирует, что максимум Понтрягиана, а следовательно, и оптимум достигается на  $u^*$ .

□

**Замечание.** Расширенное интерпретирование уравнения (1) в виде (8) позволяет распространить результаты работы на задачи, в которых дифференциальное ограничение вместо винеровского процесса  $W_t$  содержит произвольный случайный процесс  $X_t(\omega)$  с траекториями неограниченной вариации, т. е. имеет вид

$$dx_t = b(x_t, u_t)dt + \sigma(x_t) * dX_t(\omega).$$

Последнее уравнение интерпретируется в смысле симметричного интеграла, и для задачи минимизации функционала

$$J(u) = \int_0^T f(t, x_t, u_t)dt + g(x(T))$$

справедливы результаты, аналогичные приведенным в теоремах 2 и 3.

#### СПИСОК ЛИТЕРАТУРЫ

1. Rockafellar R. T., Wets R. J. B. Nonanticipativity and  $L^1$  martingales in stochastic optimization problems // *Mathematical Programming Study*. 1973. Т. 6. С. 170–187.
2. Wets R. J. B. On relation between stochastic and deterministic optimization // *Lecture Notes in Economics and Mathematics Systems*. 1975. Т. 107. С. 350–361.
3. Davis M. H. A. Anticipative LQG Control, I // *IMA Journal of Mathematical Control and Information*. Т. 6. 1989. С. 259–265.
4. Davis M. H. A., Burstein G. A deterministic approach to stochastic optimal control, with application to anticipative control // *Stochastics and Stochastics Reports*. 1992. Т. 40. С. 203–256.
5. Ocone D., Pardoux E. A generalized Ito-Ventzell formula. Application to a class of anticipating stochastic differential equations // *Ann. Inst. Henry Poincare*. 1989. Т. 25. С. 39–71.
6. Wong E., Zakai M. On the relation between ordinary and stochastic differential equations // *Int. J. Engineering Science*. 1965. Т. 3. С. 213–229.
7. Sussman H. J. On the gap between deterministic and stochastic differential equations // *Annals of Probability*. 1978. Т. 6. С. 19–41.
8. Kunita H. *Stochastic Flows and Stochastic Differential Equations*. Cambridge: Cambridge University Press, 1990. 346 с.

9. Насыров Ф. С. Симметричные интегралы и стохастический анализ. // *Теория вероятностей и ее применение*. 2006. Т. 51, 3. С. 496–517.

10. Насыров Ф. С. Локальные времена, симметричные интегралы и стохастический анализ. М.: Физматлит, 2011. 212 с.

#### ОБ АВТОРАХ

**ИСМАГИЛОВ Нияз Салаватович**, асп. каф. математики. Дипл. спец. по прикл. математике (УГАТУ, 2009). Готовит дисс. о стохаст. опт. управлении.

**НАСЫРОВ Фарит Сагитович**, проф. той же каф. Дипл. математик (ЛГУ, 1976). Д-р физ.-мат. наук по теории вероятн., мат. статистике и мат. анализу (ИМ им. Соболева, Новосибирск, 2002). Иссл. в обл. теории случ. процессов, теории функций, фин. математики.

#### METADATA

**Title:** On deterministic approach to stochastic optimal control.

**Authors:** N. S. Ismagilov<sup>1</sup>, F. S. Nasyrov<sup>2</sup>

**Affiliation:** Ufa State Aviation Technical University (UGATU), Russia.

**Email:** <sup>1</sup>niyaz.ismagilov@gmail.com, <sup>2</sup>farsagit@yandex.ru.

**Language:** Russian.

**Source:** Vestnik UGATU (scientific journal of Ufa State Aviation Technical University), vol. 17, no. 5 (58), pp. 38-43, 2013. ISSN 2225-2789 (Online), ISSN 1992-6502 (Print).

**Abstract:** We consider a deterministic approach to stochastic optimal control problem based on decomposition of solution of stochastic differential equation. The decomposition formula allows to consider stochastic problem as parameterized family of deterministic control problems. It is shown that value function of deterministic problem can be modified in order to obtain non-anticipating solutions, which are also optimal for stochastic problem.

**Key words:** Stochastic optimal control; stochastic differential equation; anticipating stochastic differential equation; maximum principle; symmetric integral; deterministic approach.

#### References (English Transliteration):

1. R. T. Rockafellar and R. J. B. Wets, "Nonanticipativity and  $L^1$  martingales in stochastic optimization problems," *Mathematical Programming Study*, vol. 6, pp. 170-187, 1973.
2. R. J. B. Wets, "On relation between stochastic and deterministic optimization," *Lecture Notes in Economics and Mathematics Systems*, vol. 107, pp. 350-361, 1975.
3. M. H. A. Davis, "Anticipative LQG control, I," *IMA J. Math. Control and Inform.*, vol. 6, pp. 259-265, 1989.
4. M. H. A. Davis and Burstein G., A deterministic approach to stochastic optimal control, with application to anticipative

- control," *Stochastics and Stochastics Reports*, vol. 40, pp. 203–256, 1992.
5. D. Ocone and E. Pardoux, "A generalized Ito-Ventzell formula. Application to a class of anticipating stochastic differential equations," *Ann. Inst. Henry Poincare*, vol. 25, pp. 39-71, 1989.
  6. E. Wong and M. Zakai, "On the relation between ordinary and stochastic differential equations," *Int. J. Engineering Science*, vol. 3, pp. 213-229, 1965.
  7. H. J. Sussman, "On the gap between deterministic and stochastic differential equations," *Annals of Probability*, vol. 6, pp. 19-41, 1978.
  8. H. Kunita, *Stochastic Flows and Stochastic Differential equations*, Cambridge: Cambridge University Press, 1990.
  9. F. S. Nasyrov, "Symmetric integrals and stochastic analysis," (in Russian), *Teoriya Veroyatnostei i ee Primeneniya*, vol. 14, no. 3, pp. 496-517, 2006.
  10. F. S. Nasyrov, *Local Times, Symmetric Integrals and Stochastic Analysis*, (in Russian). Moscow: Fizmatlit, 2011.

**About authors:**

**ISMAGILOV, Niyaz Salavatovich**, Postgrad. (PhD) Student, Dept. of Mathematics. Master of Applied Mathematics (UGATU, 2009).

**NASYROV, Farit Sagitovich**, Prof., Dept. of Mathematics. Dipl. Mathematician (Leningrad State Univ., 1976). Doctor of Sciences in Physics and Mathematics (Sobolev Institute of Mathematics, 2002).