

УДК 517.977

Управление квазилинейными стохастическими системами с неполной информацией на примере механического манипулятора

Румянцев Д.С.^{1*}, Царьков К.А.^{2}**

¹*Институт машиноведения им. А.А. Благонравова РАН,*

Малый Харитоньевский переулок, 4, Москва, 101990, Россия

²*Московский авиационный институт (национальный исследовательский университет),*

Волоколамское шоссе, 4, Москва, А-80, ГСП-3, 125993, Россия

**e-mail: dima_rum@mail.ru*

***e-mail: k6472@mail.ru*

Аннотация

Дано решение задачи стабилизации двухзвенного механического манипулятора. Используются новые результаты, полученные авторами для квазилинейных стохастических управляемых систем при наличии информационных ограничений. Ограничения выражаются в том, что каждая компонента вектора стратегии управления зависит от заранее заданного набора точно измеряемых компонент вектора состояния. Рассматривается алгоритм поиска субоптимального линейного регулятора с полиномиальными по времени коэффициентами. Такой алгоритм позволяет не только решать задачу стабилизации манипулятора, но и конструировать стратегии управления любыми другими системами указанного типа.

Ключевые слова: стохастическое оптимальное управление, квазилинейная динамическая система, неполная обратная связь, субоптимальное управление,

двухзвенный манипулятор.

Введение

Механические манипуляторы используются в различных отраслях промышленности: робототехника, автоматические сборочные линии, глубоководные погружения в батискафах, космические исследования. Для космического корабля «Буран» был создан манипулятор, состоящий из 16 звеньев. Здесь решена задача стабилизации двухзвенного манипулятора, на который во время работы действуют случайные факторы, например, воздушные или подводные течения. Стабилизация осуществляется в условиях информационных ограничений, заключающихся в том, что каждая компонента стратегии управления зависит от своего набора компонент вектора состояния.

Для синтеза стратегии управления используются результаты, детально изложенные в [1]. В данной работе приведены только сведения из [1], необходимые для решения задачи управления манипулятором. В работе даны: общее описание квазилинейных стохастических управляемых систем, информационные ограничения, конечные формулы и конструктивный алгоритм решения задач синтеза субоптимальных стратегий управления такими системами. Задача стабилизации механического манипулятора освещена более подробно. В отличие от [1] здесь рассмотрен случай, когда обе компоненты вектора управления реализуются неточно, выполнено моделирование движения звеньев манипулятора, уточнено количество

итераций градиентного метода, необходимое для решения задачи.

В работе нет теорем и сложных математических выкладок. Приведены результаты, которые можно непосредственно использовать для решения прикладных задач оптимального управления.

Общая постановка задачи

Процесс управления описывается системой уравнений Ито [2]

$$\begin{aligned} dx(t) &= f(t, x(t), u(t, x(t)))dt + g(t, x(t), u(t, x(t)))dw(t), \\ x(t_0) &= x_0, \end{aligned} \quad (1)$$

где $t \in T = [t_0; t_1]$ – время; $x = (x_1, \dots, x_n)^T$ – вектор состояния системы; $w(t)$ – ν -мерный стандартный винеровский процесс; $u = (u_1, \dots, u_m)^T \in U \subseteq R^m$ – вектор управления.

Функция $f(t, x, u)$, где $f = (f_1, \dots, f_n)^T$, линейна по x, u и имеет вид

$$f(t, x, u) = A(t)x + B(t)u.$$

Столбцы $g_l(t, x, u)$, $l = \overline{1, \nu}$, матричной функции $g(t, x, u)$ размера $(n \times \nu)$ также линейны по x, u и имеют вид

$$g_l(t, x, u) = G^{(l)}(t)x + F^{(l)}(t)u + C^{(l)}(t).$$

Здесь $A(t)$, $B(t)$, $G^{(l)}(t)$, $F^{(l)}(t)$, $C^{(l)}(t)$ – матрицы размеров $(n \times n)$, $(n \times m)$, $(n \times n)$, $(n \times m)$, $(n \times 1)$ соответственно, элементы которых являются ограниченными борелевскими функциями на интервале T .

Пусть для рассматриваемого здесь процесса (1) плотность вероятности состояния

$p(t, x) \in C^{1,2}(T \times R^n)$ существует и удовлетворяет уравнению Фоккера–Планка–Колмогорова (ФПК) [2]

$$\frac{\partial p(t, x)}{\partial t} = - \sum_{i=1}^n \frac{\partial}{\partial x_i} [f_i(t, x, u) p(t, x)] + \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} [a_{ij}(t, x, u) p(t, x)], \quad (2)$$

$$a_{ij} = \frac{1}{2} \sum_{l=1}^v g_{il} g_{jl},$$

с начальным условием

$$p(t_0, x) = p_0(x). \quad (3)$$

Множество допустимых стратегий управления образуют функции $u(t, x): T \times R^n \rightarrow U$, удовлетворяющие информационным ограничениям.

Информационные ограничения состоят в зависимости каждой компоненты вектора управления u от своего априори назначаемого набора компонент вектора состояния x . Здесь рассматриваются линейные по состоянию стратегии управления.

Пример информационных ограничений. Пусть управляемая динамическая система имеет вектор состояния $x = (x_1, x_2, x_3, x_4, x_5)^T$ и вектор управления $u = (u_1, u_2, u_3, u_4)^T$. Требуется синтезировать стратегию управления в виде

$$u(t, x(t)) = \begin{pmatrix} u_1(t, x_2(t), x_4(t)) \\ u_2(t, x_3(t), x_4(t)) \\ u_3(t, x_1(t), x_4(t)) \\ u_4(t, x_1(t), x_2(t), x_4(t)) \end{pmatrix}.$$

Все компоненты вектора $u(t, x(t))$ зависят от координаты $x_4(t)$, ни одна

компонента не зависит от $x_5(t)$ и все они по-своему зависят от $x_1(t)$, $x_2(t)$, $x_3(t)$.

Через D обозначим множество допустимых процессов управления $z = (p(\cdot), u(\cdot))$, где функция $p(\cdot)$ является решением уравнения (2) с начальным условием (3), а функция $u(\cdot)$ представляет собой стратегию управления с информационными ограничениями. Предполагается, что для начальной плотности $p_0(x)$ известны математическое ожидание m_0 и ковариационная матрица K_0 .

Для процесса $z \in D$ определён критерий качества управления

$$J(z) = \int_{t_0}^{t_1} \int_{R^n} f^c(t, x, u(t, x)) p(t, x) dx dt + \int_{R^n} F^c(x) p(t_1, x) dx : D \rightarrow R, \quad (4)$$

функции $f^c(t, x, u) : T \times R^n \times R^m \rightarrow R^1$, $F^c(x) : R^n \rightarrow R^1$ которого имеют вид неотрицательных квадратичных форм:

$$f^c(t, x, u) = \frac{1}{2} x^T D(t) x + u^T S(t) x + \frac{1}{2} u^T E(t) u,$$

$$F^c(x) = \frac{1}{2} x^T Q x,$$

где $D(t)$, $S(t)$, $E(t)$, Q – матрицы размеров $(n \times n)$, $(m \times n)$, $(m \times m)$, $(n \times n)$ соответственно, элементы $D(t)$, $S(t)$, $E(t)$ – ограниченные борелевские функции на T . Здесь и далее матрицы квадратичных форм считаются симметрическими. Цель управления состоит в минимизации функционала (4) на множестве D .

Субоптимальное управление

Синтез субоптимальных стратегий управления будем проводить с помощью формул из [3, 4].

1. Уравнения для математического ожидания и ковариационной матрицы случайного процесса $x(t)$

$$\frac{dm}{dt} = A^u m - BL, \quad (5)$$

$$\frac{dK}{dt} = A^u K + KA^{uT} + \sum_{l=1}^{\nu} \Upsilon^{(l)} K \Upsilon^{(l)T} + \tilde{Q}, \quad (6)$$

где $A^u = A - BP$, $\tilde{Q} = (\tilde{C} + \Upsilon(m))(\tilde{C} + \Upsilon(m))^T$, $\Upsilon^{(l)} = G^{(l)} - F^{(l)}P$, $\Upsilon^{(l)}(m) = \Upsilon^{(l)}m$,

$\tilde{C}^{(l)} = -F^{(l)}L + C^{(l)}$, Υ, \tilde{C} – матрицы, составленные из столбцов $\Upsilon^{(l)}, \tilde{C}^{(l)}$ соответственно.

Начальные условия для этих уравнений имеют вид

$$m(t_0) = m_0, \quad K(t_0) = K_0. \quad (7)$$

2. Сопряжённая система уравнений

$$\frac{d\gamma}{dt} = \lambda^T BL - \frac{1}{2} L^T EL - \sum_{l=1}^{\nu} \left(\frac{1}{2} C^{(l)T} M C^{(l)} + \frac{1}{2} L^T F^{(l)T} M F^{(l)} L - L^T F^{(l)T} M C^{(l)} \right), \quad (8)$$

$$\frac{d\lambda}{dt} = -A^{uT} \lambda + S^T L + P^T EL + MBL - \sum_{l=1}^{\nu} \left([F^{(l)}P - G^{(l)}]^T M [F^{(l)}L - C^{(l)}] \right), \quad (9)$$

$$\frac{dM}{dt} = -MA^u - A^{uT} M - D + S^T P + P^T S - P^T EP - \sum_{l=1}^{\nu} \left([F^{(l)}P - G^{(l)}]^T M [F^{(l)}P - G^{(l)}] \right). \quad (10)$$

Условия для этих уравнений заданы при $t = t_1$:

$$\gamma(t_1) = \gamma_1 = 0, \quad \lambda(t_1) = \lambda_1 = 0, \quad M(t_1) = M_1 = Q. \quad (11)$$

Функции $P(t), L(t)$ определяют стратегию управления в форме линейного регулятора [3]:

$$u(t, x) = -(P(t)x + L(t)). \quad (12)$$

Выберем $P(t), L(t)$ в виде полиномиальных по времени функций. При этом управление (12) уже не будет в общем случае оптимальным по критерию (4), но его применение на практике может быть значительно выгоднее в силу его простоты. Чтобы u удовлетворяло информационным ограничениям, зафиксируем значения $P_{ij}(t) \equiv 0$, если компонента u_i стратегии управления не зависит от компоненты x_j вектора состояния.

Значение функционала (4) на допустимом процессе $z = (p(\cdot), u(\cdot))$, где стратегия $u(\cdot)$ имеет вид (12), плотность $p(\cdot)$ характеризуется математическим ожиданием $m(t)$ и ковариационной матрицей $K(t)$, а функции $m(t), K(t), P(t), L(t)$ удовлетворяют системе (5) - (11), вычисляется по формуле [3]

$$J = \frac{1}{2} \text{tr}(M_0 K_0) + \frac{1}{2} m_0^T M_0 m_0 + \lambda_0^T m_0 + \gamma_0, \quad (13)$$

где $m_0 = m(t_0)$, $K_0 = K(t_0)$, $M_0 = M(t_0)$, $\lambda_0 = \lambda(t_0)$, $\gamma_0 = \gamma(t_0)$. Здесь и далее за tr обозначен оператор следа квадратной матрицы.

Введём набор параметров $s = (s_1, \dots, s_N)$. Обозначим за $\tilde{P}_{ij}(t, s), \tilde{L}_i(t, s)$, $i = \overline{1, m}, j = \overline{1, n}$, полиномы от t с коэффициентами s_r , вообще говоря различными для

всех \tilde{P}_{ij} и \tilde{L}_i . Здесь N определяется информационными ограничениями и порядками полиномов.

Составим матрицы P, L из элементов

$$P_{ij}(t) = \begin{cases} \tilde{P}_{ij}(t, s) & , u_i = u_i(\cdot, x_j) \\ 0 & , u_i \neq u_i(\cdot, x_j), \end{cases} \quad (14)$$

$$L_i(t) = \tilde{L}_i(t, s).$$

Например, если используются полиномы нулевого порядка, а информационные ограничения отсутствуют, то матрицы P и L будут иметь вид

$$P = \begin{pmatrix} s_1 & s_2 & \dots & s_n \\ s_{n+1} & s_{n+2} & \dots & s_{2n} \\ \dots & \dots & \dots & \dots \\ s_{(m-1)n+1} & s_{(m-1)n+2} & \dots & s_{mn} \end{pmatrix},$$

$$L = \begin{pmatrix} s_{mn+1} \\ s_{mn+2} \\ \dots \\ s_{m(n+1)} \end{pmatrix}.$$

Коэффициенты $s_r, r = \overline{1, N}$, подбираются так, чтобы минимизировать критерий качества управления J .

Таким образом, получена задача безусловной минимизации критериальной функции $J(z(s)) = \hat{J}(s)$, зависящей неявно от $s = (s_1, \dots, s_N)$. Для поиска значений $\bar{s}_r, r = \overline{1, N}$, доставляющих минимум $\hat{J}(s)$, будем использовать численный метод градиентного спуска[5].

На каждой итерации метода компоненты градиента $\partial \hat{J} / \partial s_r$ могут быть вычислены по формуле

$$\begin{aligned} \left. \frac{\partial \hat{J}}{\partial s_r} \right|_{s^{(k)}} = & \int_{t_0}^{t_1} \left\{ \text{tr} \left(P_r'^{(k)T} \left[(E + \Theta^{(k)}) P^{(k)} - B^T M^{(k)} - \Pi^{(k)} - S \right] K^{(k)} \right) + \right. \\ & + \left(P_r'^{(k)} m^{(k)} + L_r'^{(k)} \right)^T \left(\left[(E + \Theta^{(k)}) P^{(k)} - B^T M^{(k)} - \Pi^{(k)} - S \right] m^{(k)} + \right. \\ & \left. \left. + \left[(E + \Theta^{(k)}) L^{(k)} - B^T \lambda^{(k)} - \Lambda^{(k)} \right] \right) \right\} dt, \end{aligned} \quad (15)$$

где k – номер итерации, $m^{(k)}(t), K^{(k)}(t), M^{(k)}(t), \lambda^{(k)}(t)$ – функции, найденные решением двух задач Коши в прямом и обратном времени для системы уравнений (5)-(7), (8)-(11), $P^{(k)}(t), L^{(k)}(t), P_r'^{(k)}(t), L_r'^{(k)}(t)$ – соответствующие функции, взятые в точке $s^{(k)}$, $P_r'(t), L_r'(t)$ – производные функций $P(t), L(t)$ вида (14) по s_r ,

$$\Pi^{(k)} = \sum_{l=1}^{\nu} F^{(l)T} M^{(k)} G^{(l)}, \Lambda^{(k)} = \sum_{l=1}^{\nu} F^{(l)T} M^{(k)} C^{(l)}, \Theta^{(k)} = \sum_{l=1}^{\nu} F^{(l)T} M^{(k)} F^{(l)}.$$

Алгоритм поиска субоптимальной стратегии управления.

Шаг 1. Произвольным образом или из дополнительных соображений задать Δs – шаг метода градиентного спуска, ε – требуемую максимальную погрешность приближения, $s^{(0)}$ – начальную точку приближения вектора параметров s , и положить номер итерации $k=0$, количество успешных итераций $i=0$. Вычислительный опыт показывает, что в качестве начальной точки целесообразно брать точку $s^{(0)} = 0$.

Шаг 2. Решить (численно) задачи Коши сперва в прямом, а затем и в обратном

времени для системы уравнений (5),(6),(8)-(10) с условиями (7), (11), используя матрицы $P^{(k)}, L^{(k)}$, имеющие вид (14).

Шаг 3. Вычислить значение критерия $J^{(k)}$ по формуле (13). Если $k = 0$, перейти к шагу 5. В противном случае проверить выполнение условия $J^{(k)} < J^{(k-1)}$: если условие выполнено, увеличить i на единицу и перейти к шагу 4, иначе положить $i = 0$, $s^{(k)} = s^{(k-1)}$, уменьшить Δs вдвое, уменьшить k на единицу и перейти к шагу 6.

Шаг 4. Если $i = 2$, увеличить Δs вдвое, положить $i = 0$.

Шаг 5. Для всех $r = \overline{1, q}$ вычислить $\partial \hat{J} / \partial s_r$ по формуле (15).

Шаг 6. Проверить выполнение условия

$$\sum_{r=1}^N \left| \frac{\partial \hat{J}}{\partial s_r} \right| < \varepsilon :$$

если условие выполнено, искомое значение \bar{s} положить равным $s^{(k)}$ и закончить расчет, иначе положить $s_r^{(k+1)} = s_r^{(k)} - \Delta s \cdot \partial \hat{J} / \partial s_r, r = \overline{1, N}$, и перейти к шагу 7.

Шаг 7. Увеличить k на единицу и перейти к шагу 2.

Задача оптимального управления механическим манипулятором

Рассмотрим следующую задачу [6]. На горизонтальной плоскости находится двухзвенный механический манипулятор, каждое звено которого представляет собой абсолютно жёсткий стержень длиной $l_i, i = 1, 2$ (см. рисунок 1).

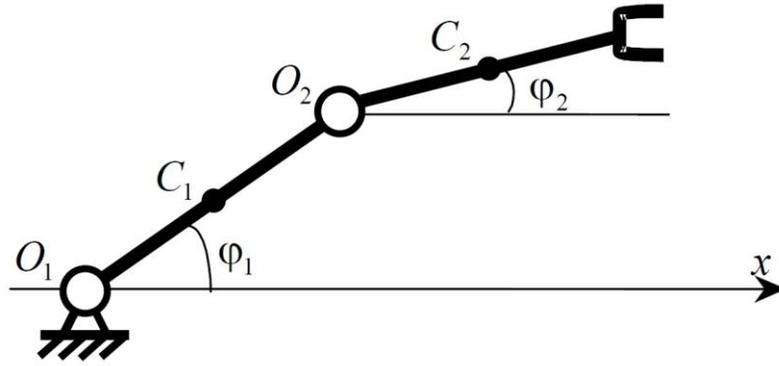


Рисунок 1. Механический манипулятор

Первое звено соединено с неподвижным основанием манипулятора вращательной парой O_1 , а со вторым звеном – вращательной парой O_2 . Масса схвата манипулятора – m , центр i -го звена находится в середине стержня – точке C_i , его масса – m_i , момент инерции i -го звена относительно своего центра масс – $I_i, i = 1, 2$. В соединительных парах могут развиваться управляющие вращательные моменты: соответственно u_1 и u_2 . Известно положение $\varphi_1 = \varphi_1^*, \varphi_2 = \varphi_2^*$, в котором требуется стабилизировать манипулятор.

Линеаризованные уравнения движения манипулятора в окрестности φ_1^*, φ_2^* имеют вид

$$\dot{x} = Ax + Bu,$$

где $x = (x_1, x_2, x_3, x_4)^T, x_1 = \varphi_1 - \varphi_1^*, x_2 = \varphi_2 - \varphi_2^*, x_3 = \dot{\varphi}_1, x_4 = \dot{\varphi}_2, u = (u_1, u_2)^T,$

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, B = \frac{1}{ab - c^2} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ b & -c \\ -c & a \end{pmatrix},$$

$$a = 1/4 [l_1^2(m_1 + 4m_2 + 4m) + 4I_1],$$

$$b = 1/4 [l_2^2(m_2 + 4m) + 4I_2], c = 1/2(2m + m_2)l_1l_2.$$

Предполагается, что на движение манипулятора и процесс управления оказывают влияние вектор случайных воздействий, имеющий две компоненты. Будем также считать, что управление осуществляется в условиях информационных ограничений, отражающих невозможность получения полной информации о состоянии. Выбраны следующие характеристики манипулятора: $m_1 = m_2 = m = 30$ кг; $l_1 = l_2 = 1$ м; $I_1 = I_2 = 5/2$ кг·м².

Цель управления состоит в переводе манипулятора из начального состояния $(\varphi_{1_0}, \varphi_{2_0}, \dot{\varphi}_{1_0}, \dot{\varphi}_{2_0})$, определяемого математическим ожиданием $m_0 = (0, 0, 0, 0)^T$ и матрицей ковариаций $K_0 = \text{diag}(0.09, 0.09, 0.09, 0.09)$ вектора x_0 , в состояние $(\varphi_1^*, \varphi_2^*, 0, 0)$. При этом требуется минимизировать затраты на управление. Цель должна быть достигнута за 5 с.

Таким образом, задача принимает вид

$$dx(t) = (Ax(t) + Bu(t, x))dt + g(u(t, x))dw(t), x(0) = x_0,$$

$$J = \int_0^5 \int_{R^4} \left(\frac{1}{2} u^T E u \right) p(t, x) dx dt + \int_{R^4} \left(\frac{1}{2} x^T Q x \right) p(5, x) dx \rightarrow \min,$$

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, B = \frac{1}{155} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 8 & -9 \\ -9 & 14 \end{pmatrix},$$

$$g(u) = \begin{pmatrix} 0.05 & 0.01 \\ 0.01 & 0.05 \\ 0.025u_1 & 0.015u_2 \\ 0.015u_1 & 0.025u_2 \end{pmatrix}, E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, Q = \begin{pmatrix} 15 & 0 & 0 & 0 \\ 0 & 15 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 10 \end{pmatrix}.$$

Для получения стратегии управления воспользуемся постоянными по времени матрицами $P(t), L(t)$ и предложенным алгоритмом. Результаты вычислений представлены в таблице 1. Запись вида $u_2(t, x_2, x_3)$ определяет информационные ограничения на компоненту вектора управления u_2 и означает, что u_2 зависит от x_2, x_3 , но не зависит от x_1, x_4 .

Результаты решения задачи Таблица 1

Номер варианта	Информационные ограничения	Значение критерия	Норма градиента	Количество итераций
1	$u_1(t); u_2(t)$	36.17	0.065	10
2	$u_1(t); u_2(t, x_2, x_3)$	22.84	0.098	61
3	$u_1(t, x_1, x_2, x_4); u_2(t)$	23.76	0.095	93
4	$u_1(t, x_2, x_4); u_2(t, x_1, x_2, x_4)$	21.18	0.098	138
5	$u_1(t, x_1, x_2, x_3, x_4); u_2(t, x_1, x_2, x_3, x_4)$	19.02	0.094	99

Пусть теперь точно известно начальное состояние $x_0 = (0.807, 0.087, -0.206, 0.152)^T$, тогда, применяя найденные для разных случаев информированности стратегии управления, получим следующие результаты.

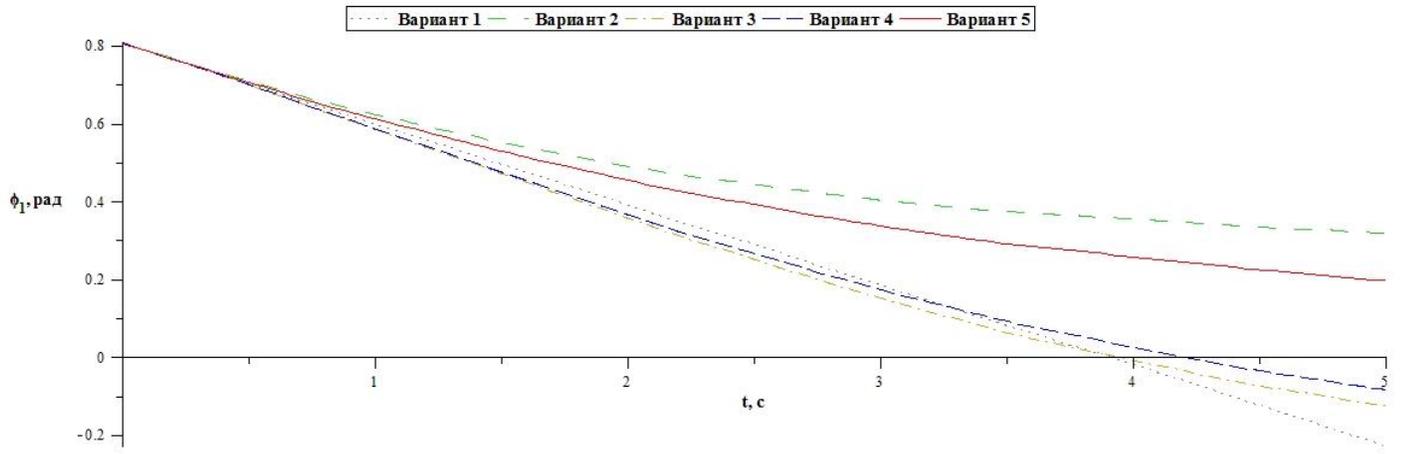


Рисунок 2. Графики $\phi_1(t)$

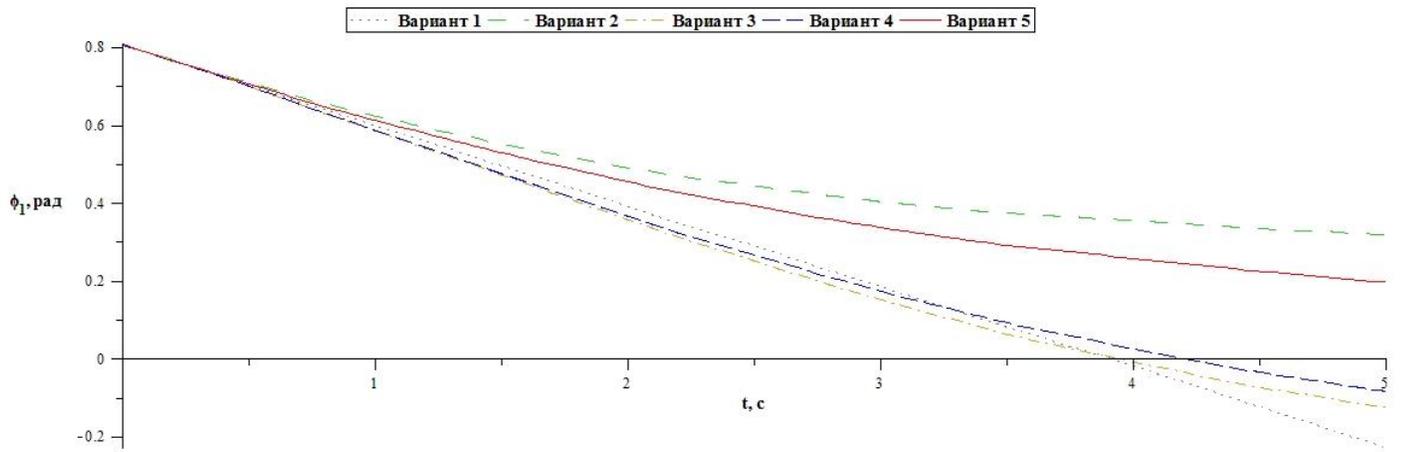


Рисунок 3. Графики $\phi_2(t)$

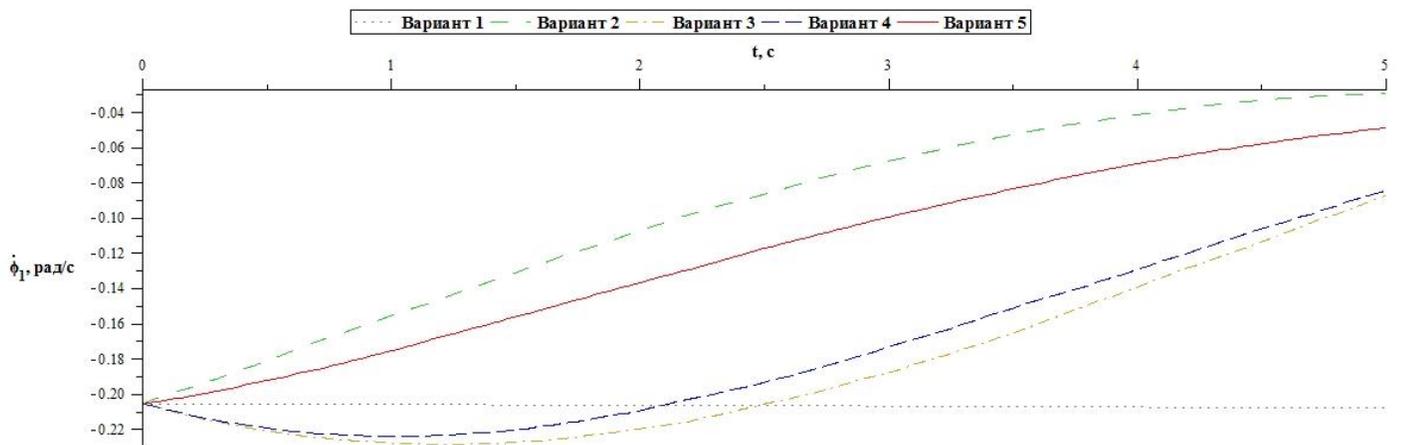


Рисунок 4. Графики $\dot{\phi}_1(t)$

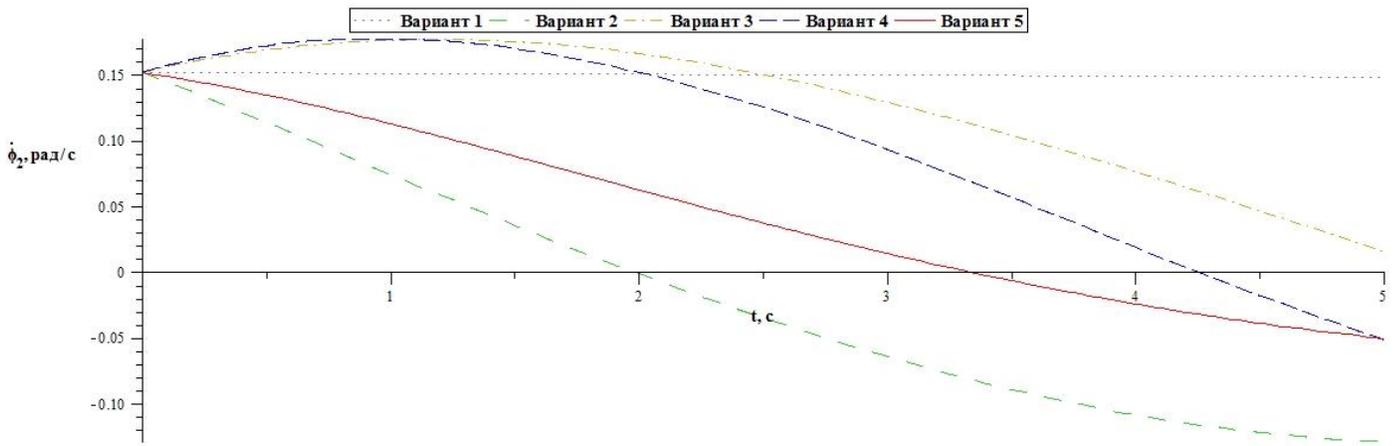


Рисунок 5. Графики $\dot{\phi}_2(t)$

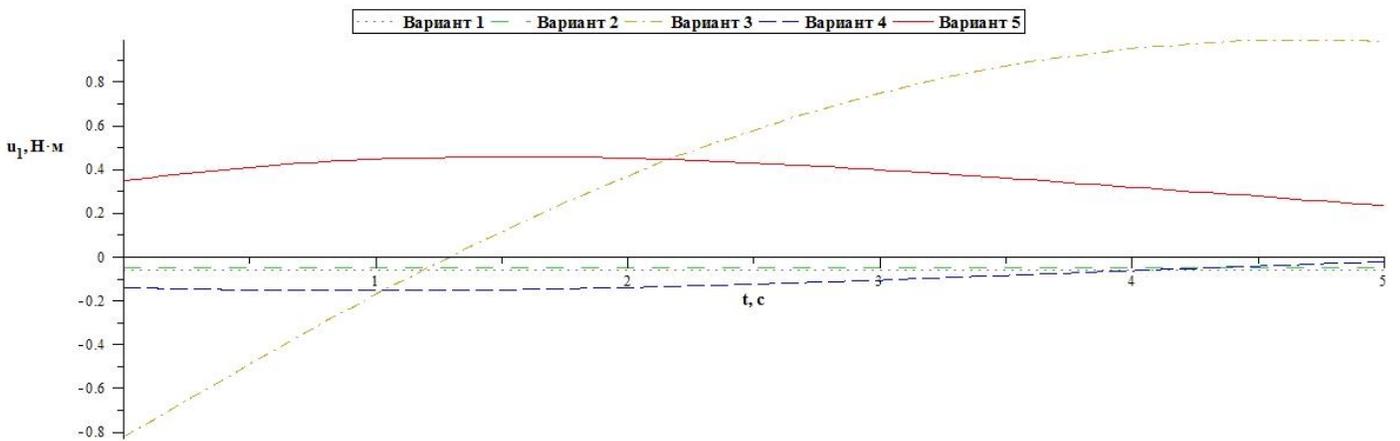


Рисунок 6. Графики $u_1(t)$

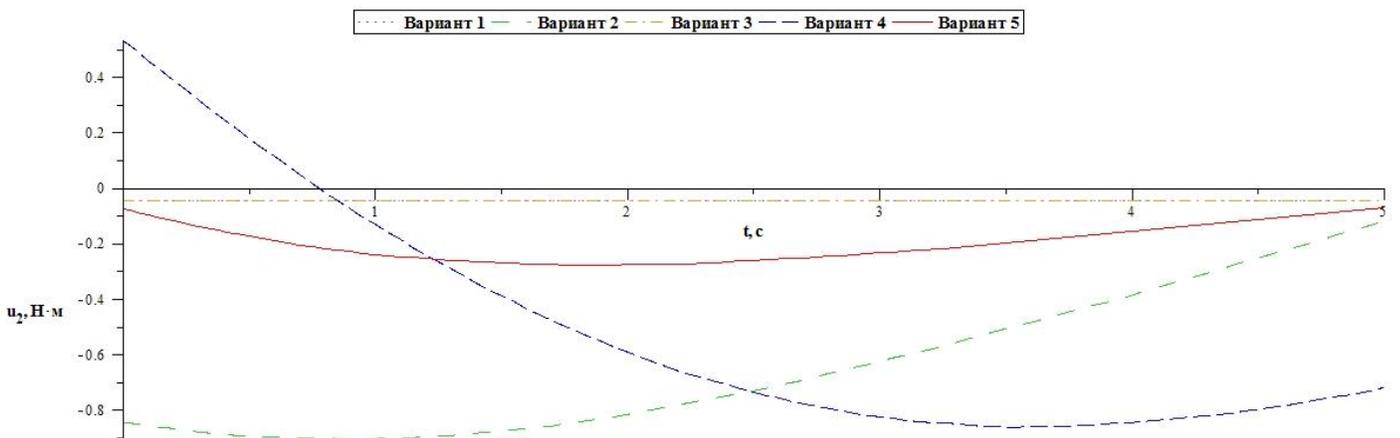


Рисунок 7. Графики $u_2(t)$

На манипулятор действуют небольшие случайные возмущения, управление реализуется со случайными ошибками. На графиках видно, что с помощью достаточно малых вращательных моментов можно получить незначительные отклонения звеньев манипулятора от требуемого положения в конечный момент времени. Это вполне соответствует целям управления.

Заключение

Известно, что аналитически построить оптимальную стратегию управления даже для линейных систем удаётся не всегда. Различные численные методы также не могут гарантировать получение оптимального решения. Поэтому очень важно иметь возможность синтезировать субоптимальную стратегию. Предложенный в работе градиентный метод позволяет такое субоптимальное решение получить. Это существенно в практических приложениях. Представленные здесь конкретные сведения подходят именно для применения на практике. Они могут быть непосредственно использованы для решения прикладных задач синтеза оптимального управления квазилинейными динамическими стохастическими системами с информационными ограничениями.

Библиографический список

1. Хрусталёв М.М., Румянцев Д.С., Царьков К.А. Алгоритм поиска субоптимальных стратегий управления квазилинейными динамическими стохастическими системами диффузионного типа // Изв. РАН. ТиСУ. 2014. № 1. с. 74-86.
2. Пугачев В.С., Синицин И.Н. Стохастические дифференциальные системы. Анализ и фильтрация, М.: Наука, 1990.
3. Румянцев Д.С., Хрусталев М.М. Оптимальное управление квазилинейными системами диффузионного типа при неполной информации о состоянии // Изв. РАН. ТиСУ. 2006. № 5. С. 43-51.
4. Румянцев Д.С., Хрусталев М.М. Численные методы синтеза оптимального управления для стохастических динамических систем диффузионного типа // Изв. РАН. ТиСУ. 2007. № 3. С. 27-38.
5. Пантелеев А.В., Летова Т.А. Методы оптимизации в примерах и задачах, М.: Высш. шк., 2005. 544 с.
6. Лутманов С.В. Линейные задачи оптимизации. Ч.2. Оптимальное управление линейными динамическими объектами, Пермь: Изд. Пермск. ун-та, 2005. 195 с.