

УДК 62.50

УТОЧНЕНИЕ ЗАДАЧИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ ДЛЯ ПРАКТИЧЕСКОЙ РЕАЛИЗАЦИИ ЕЕ РЕШЕНИЯ

© 2023 г. А. И. Дивеев^{1,*}

Представлено академиком РАН И.А. Соколовым

Поступило 06.07.2022 г.

После доработки 24.12.2022 г.

Принято к публикации 26.12.2022 г.

Решение задачи оптимального управления в классической постановке представляет собой управление в форме функции времени. Реализация такого решения приводит к разомкнутой системе управления и поэтому не может применяться непосредственно на практике. Считается, что решение классической задачи оптимального управления приводит к получению оптимальной программы управления и программной траектории в пространстве состояний. Для реализации движения объекта управления по программной траектории необходимо построение дополнительной системы стабилизации движения. Задача синтеза системы стабилизации движения по программной траектории и требования, которым должна удовлетворять эта система, не вытекают из классической постановки задачи оптимального управления. Приведена уточненная постановка задачи оптимального управления, которая включает дополнительное требование к оптимальной траектории, и решение которой может быть непосредственно применено на практике в реальном объекте управления.

Ключевые слова: оптимальное управление, синтез управления, сжимающее отображение

DOI: 10.31857/S2686954322600458, **EDN:** CPZONE

Рассмотрим классическую постановку задачи оптимального управления [1].

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}), \quad (1)$$

$$\mathbf{u} \in U \subseteq \mathbb{R}^m, \quad (2)$$

$$\mathbf{x}(0) = \mathbf{x}^0, \quad (3)$$

$$\mathbf{x}(t_f) = \mathbf{x}^f, \quad (4)$$

$$J = \int_0^{t_f} f_0(\mathbf{x}, \mathbf{u}) dt \rightarrow \min_{\mathbf{u} \in U}, \quad (5)$$

где $\mathbf{x}$ — вектор состояния, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u}$ — вектор управления, U — компактное множество, $\mathbf{x}^0$ — заданный вектор начального состояния, $\mathbf{x}^f$ — заданный вектор терминального состояния, t_f — время достижения терминального состояния, как правило, не задано, но ограничено, $t_f \leq t^+$, t^+ — заданная положительная величина, J — значение критерия качества управления.

В классической постановке никаких дополнительных требований к управлению, кроме удовлетворения ограничению на управление (2), не предъявляется, поэтому управление ищется как функция времени

$$\mathbf{u} = \mathbf{v}(t) \in U. \quad (6)$$

Если подставить найденное управление (6) в правую часть модели объекта (1), то полученная система

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{v}(t)) \quad (7)$$

будет иметь частное решение $\mathbf{x}(t, \mathbf{x}^0)$, которое из начального состояния (3) достигнет терминального состояния (4)

$$\mathbf{x}(t_f, \mathbf{x}^0) = \mathbf{x}^f, \quad t_f \leq t^+, \quad (8)$$

с оптимальным значением критерия качества (5).

Математическая модель (7) описывает динамику разомкнутой системы управления, поскольку функция управления (6) не зависит от вектора состояния объекта управления. Математическая модель используется для оценки текущего состояния реального объекта управления. Любая математическая модель дает ошибку при оценке состояния реального объекта. Введем определения.

Определение 1. Математическая модель объекта является качественно подходящей, если в рас-

¹ Федеральный исследовательский центр “Информатика и управление” Российской академии наук, Москва, Россия

*E-mail: aidiveev@mail.ru

смаатриваемой области пространства состояний она описывает качественные свойства реального объекта.

Если объект устойчив относительно некоторой точки в пространстве состояний, то качественно подходящая математическая модель также устойчива.

Определение 2. Математическая модель объекта гарантированно реализуема, если на интервале $\Delta t > 0$ первоначальная ошибка не возрастает со временем

$$\|\tilde{\mathbf{x}}(\Delta t) - \mathbf{x}(\Delta t)\| \leq \|\tilde{\mathbf{x}}(0) - \mathbf{x}(0)\|, \quad (9)$$

где $\tilde{\mathbf{x}}(t)$ — вектор состояния реального объекта в момент t .

Теорема 1. Если качественно подходящая математическая модель реального объекта в области $X \subseteq \mathbb{R}^n$ имеет свойство сжимающего отображения и $\mathbf{x}(t_f, \mathbf{x}^0) = \mathbf{x}^f \in X$, является неподвижной точкой, то этого достаточно, чтобы она была реализуема в этой области.

Доказательство. Согласно определению 1, математическая модель и реальный объект обладают в области $X \subseteq \mathbb{R}^n$ свойством сжимающего отображения.

Пусть

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \quad (10)$$

модель объекта. Пусть в какой-то момент времени t' из-за неучтенных в модели возмущений реального объекта состояние реального объекта $\tilde{\mathbf{x}}(t')$ не совпадает с состоянием, вычисленным по модели

$$\delta(t') = \|\tilde{\mathbf{x}}(t') - \mathbf{x}(t')\|, \quad (11)$$

где δ — ошибка определения состояния объекта по его модели.

Согласно свойству сжимающего отображения $\forall \Delta t > 0$

$$\begin{aligned} \delta(t' + \Delta t) &= \|\tilde{\mathbf{x}}(t' + \Delta t) - \mathbf{x}(t' + \Delta t)\| \leq \\ &\leq \|\tilde{\mathbf{x}}(t') - \mathbf{x}(t')\| = \delta(t'). \end{aligned} \quad (12)$$

Следовательно, ошибка не возрастает во времени. Так как $\mathbf{x}^f \in X$ является неподвижной точкой сжимающего отображения, то

$$\tilde{\mathbf{x}}(\tilde{t}_f, \mathbf{x}^0) = \mathbf{x}^f, \quad (13)$$

где $\tilde{t}_f$ — время достижения неподвижной точки реальным объектом. Тогда в момент

$$t_1 = \max\{t_f, \tilde{t}_f\} \quad (14)$$

выполняются условия

$$\delta(t_1) = \|\tilde{\mathbf{x}}(t_1) - \mathbf{x}(t_1)\| = 0. \quad (15)$$

Ошибка (11) не возрастает во времени и при $\Delta t \geq t_1 - t'$ равна нулю. Модель реализуема согласно определению 2.

Причиной нереализуемости решения задачи оптимального управления является то, что получаемая в результате решения модель разомкнутой системы управления описывает объект управления с ошибкой, которая может расти во времени из-за возмущений, например по начальным условиям. Для обеспечения реализуемости решения задачи оптимального управления в классическую постановку задачи введены некоторые уточнения. Вместо соотношения (6) управление ищем в форме

$$\mathbf{u} = \mathbf{g}(\mathbf{x}, t) \in U, \quad (16)$$

$$\mathbf{g}(\mathbf{x}, t) : \mathbb{R}^n \times \mathbb{R}_+ \rightarrow \mathbb{R}^m.$$

После подстановки функции управления (16) в правые части модели (1) получаем систему

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{g}(\mathbf{x}, t)). \quad (17)$$

Частное решение системы (17) из начального состояния (3) достигает терминального состояния (4) в момент t_f (8) с оптимальным значением критерия качества (5).

Введем также дополнительные условия, которым должна удовлетворять оптимальная траектория, т.е. частное решения уравнения (17) из начального состояния (3). Для оптимального частного решения $\mathbf{x}^*(t, \mathbf{x}^0)$ должна существовать окрестность, определяемая параметром $\delta^*(t) > 0$, $0 \leq t \leq t^+$, такая, что, если для другого частного решения $\mathbf{x}(t, \mathbf{y})$ системы (17) из другого начального состояния

$$\mathbf{x}(0) = \mathbf{y} \neq \mathbf{x}^0, \quad (18)$$

в момент t' , $0 \leq t' < t^+$, выполняются условия

$$\|\mathbf{x}(t', \mathbf{x}^0) - \mathbf{x}(t', \mathbf{y})\| \leq \delta^*(t'), \quad (19)$$

то $\exists \varepsilon^* > 0$ и $\exists t'' > t'$, $0 \leq t'' \leq t^+ - t'$ такие, что должно выполняться неравенство

$$\|\mathbf{x}(t'', \mathbf{x}^0) - \mathbf{x}(t'', \mathbf{y})\| < \varepsilon^*. \quad (20)$$

Тогда в области, определяемой параметром ε^* , выполняются условия: $\forall \Delta t', \Delta t'', 0 \leq \Delta t' < \Delta t''$

$$\begin{aligned} \|\mathbf{x}(t'' + \Delta t'', \mathbf{x}^0) - \mathbf{x}(t'' + \Delta t', \mathbf{y})\| < \\ < \|\mathbf{x}(t'' + \Delta t', \mathbf{x}^0) - \mathbf{x}(t'' + \Delta t', \mathbf{y})\| \leq \varepsilon^*. \end{aligned} \quad (21)$$

Определение 3. Задача (1)–(5), (16)–(21) называется расширенной задачей оптимального управления.

Цель формулировки расширенной постановки задачи оптимального управления состоит в обес-

печение возможности непосредственного применения ее решения в реальном объекте управления. Параметр $\delta^*(t) > 0$ определяет область в пространстве, из которой объект попадает в другую, возможно, большую область, определяемую параметром $\varepsilon^* > 0$, где выполняются условия сжимающего отображения. Величина $\delta^*(t)$ указывает на сложность решения расширенной задачи оптимального управления. Если $\delta^*(t) = 0$, то получаем задачу оптимального управления в классической постановке. Если $\delta^*(t) = \infty$, то получаем самый сложный случай, когда необходимо решить задачу общего синтеза управления [2] для всего пространства состояний.

Теорема 2. Система дифференциальных уравнений, частные решения которой обладают свойством (16)–(21), при $\delta^*(0) > 0$ обладают свойством сжимающего отображения в области пространства состояний, определяемой параметром ε^* .

Доказательство. Пусть $\mathbf{x}(t, \mathbf{x}^{0,1})$ и $\mathbf{x}(t, \mathbf{x}^{0,2})$ — два частных решения дифференциального уравнения из двух разных начальных условий $\mathbf{x}^{0,1} \neq \mathbf{x}^{0,2} \neq \mathbf{x}^0$. Пусть в момент $t' = 0$ выполняются условия

$$\|\mathbf{x}(0, \mathbf{x}^0) - \mathbf{x}(0, \mathbf{x}^{0,1})\| \leq \delta^*(0), \quad (22)$$

$$\|\mathbf{x}(0, \mathbf{x}^0) - \mathbf{x}(0, \mathbf{x}^{0,2})\| \leq \delta^*(0). \quad (23)$$

Тогда в момент $t'' > t' = 0$

$$\|\mathbf{x}(t'', \mathbf{x}^0) - \mathbf{x}(t'', \mathbf{x}^{0,1})\| \leq \varepsilon^*, \quad (24)$$

$$\|\mathbf{x}(t'', \mathbf{x}^0) - \mathbf{x}(t'', \mathbf{x}^{0,2})\| \leq \varepsilon^*. \quad (25)$$

Согласно условию (21) $\exists \Delta t'' > \Delta t' \geq 0$, что

$$\begin{aligned} & \|\mathbf{x}(t'' + \Delta t'', \mathbf{x}^0) - \mathbf{x}(t'' + \Delta t'', \mathbf{x}^{0,1})\| < \\ & < \|\mathbf{x}(t'' + \Delta t', \mathbf{x}^0) - \mathbf{x}(t'' + \Delta t', \mathbf{x}^{0,1})\| \leq \varepsilon^*, \end{aligned} \quad (26)$$

$$\begin{aligned} & \|\mathbf{x}(t'' + \Delta t'', \mathbf{x}^0) - \mathbf{x}(t'' + \Delta t'', \mathbf{x}^{0,2})\| < \\ & < \|\mathbf{x}(t'' + \Delta t', \mathbf{x}^0) - \mathbf{x}(t'' + \Delta t', \mathbf{x}^{0,2})\| \leq \varepsilon^*. \end{aligned} \quad (27)$$

Поэтому

$$\begin{aligned} & \|\mathbf{x}(t'' + \Delta t'', \mathbf{x}^{0,1}) - \mathbf{x}(t'' + \Delta t'', \mathbf{x}^{0,2})\| < \\ & < \|\mathbf{x}(t'' + \Delta t', \mathbf{x}^{0,1}) - \mathbf{x}(t'' + \Delta t', \mathbf{x}^{0,2})\|. \end{aligned} \quad (28)$$

Система уравнений описывает сжимающее отображение в области, определяемой параметром $\varepsilon^* > 0$.

Рассмотрим методы решения расширенной задачи оптимального управления. Одним из подходов к решению расширенной задачи оптимального управления является решение задачи синтеза управления. В задаче синтеза управления задана модель объекта управления (1), ограничения

на управления (2), начальные условия в виде области пространства состояний

$$X_0 \subseteq \mathbb{R}^n. \quad (29)$$

Заданы терминальное состояние (4) и критерий качества управления (5). Необходимо найти управление в форме функции

$$\mathbf{u} = \mathbf{h}(\mathbf{x}^f - \mathbf{x}) \in U \subseteq \mathbb{R}^m. \quad (30)$$

Если подставить найденную функцию управления (30) в модель объекта управления (1), то получим систему уравнений

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{h}(\mathbf{x}^f - \mathbf{x})), \quad (31)$$

любое частное решение которой из любого начального условия из заданной области (29)

$$\mathbf{x}^0 \in X_0 \subseteq \mathbb{R}^n \quad (32)$$

за конечное время $t_f(\mathbf{x}^0) \leq t^+$ попадет в терминальное состояние (4) с оптимальным значением критерия качества (5).

В данном случае после решения задачи синтеза управления получаем не одну оптимальную траекторию, а множество оптимальных траекторий из любого начального состояния из области (29). В этом случае область ε^* со свойствами сжимающего отображения может оказаться малой окрестностью терминального состояния (4). Универсальных методов решения задачи синтеза управления нет. Аналитические методы решения задачи синтеза, метод бэкстеппинг [3] и аналитическое конструирование агрегированных регуляторов (АКАР) [4] привязаны к виду модели объекта управления и направлены на обеспечения его устойчивости по Ляпунову. В настоящее время для решения задачи синтеза управления используют машинное обучение численными методами символьной регрессии [5].

Методы символьной регрессии кодируют математическое выражение в форме специального кода и ищут оптимальное решение на пространстве кодов с помощью специального генетического алгоритма, в котором изменены основные операции скрещивания и мутации для работы с соответствующим кодом символьной регрессии. Для кодирования математического выражения первоначально устанавливается алфавит элементов кода. В алфавит входят элементарные функции и аргументы искомого математического выражения. Каждый элемент алфавита в общем случае кодирует элементарную функцию с помощью целочисленного вектора из двух компонент

$$\mathbf{a} = [a_1 \ a_2]^T, \quad (33)$$

где a_1 — количество аргументов элементарной функции, a_2 — номер функции. Если $a_1 = 0$, то это

код аргумента искомого математического выражения.

Код математического выражения любого метода символьной регрессии представляет собой упорядоченное множество упорядоченных множеств кодов элементов (33) алфавита

$$G = (G_1, \dots, G_L) = (G_1 = (a^{1,1}, \dots, a^{1,M_1}), \dots, G_L = (a^{L,1}, \dots, a^{L,M_L})), \quad (34)$$

где

$$a^{i,j} = [a_1^{i,j} \ a_2^{i,j}]^T, \quad i = 1, \dots, M_j, \quad j = 1, \dots, L. \quad (35)$$

Операция скрещивания кодов математических выражений в форме (34), (35) выполняется в соответствии с правилами кодирования, чтобы полученные в результате скрещивания коды имели правильную запись и могли быть декодированы. Например, код генетического программирования [5] в общем случае представляет собой граф в форме дерева, на листьях которого расположены коды аргументов, которые можно в коде (33) рассматривать как функции без аргументов, $a_1 = 0$. При скрещивании выполняется операция обмена ветвей двух отобранных деревьев. Скрещивание для кодов в форме (34) означает обмен двух подмножеств элементов кода. При этом необходимо найти по первому элементу обмениваемой ветви дерева конечный код подмножества, который соответствует последнему элементу этой ветви дерева, при этом длины кодов после операции скрещивания могут измениться.

Основным недостатком выполнения операции скрещивания для кодов в форме (34) является сильное изменение кодов полученных новых возможных решений. Коды “потомков” часто оказываются не похожими на коды их “родителей”. В результате операция скрещивания становится операцией генерации новых кодов, при этом не соблюдаются свойство наследования и принцип эволюционного поиска. Поиск решения в этом случае становится прямым случайным поиском, что не приемлемо для поиска на нечисловом пространстве.

Для эффективного поиска на пространстве кодов, сохранения свойства наследования и условий эволюции используется принцип малых вариаций базисного решения [5, 6], согласно которому множество возможных решений кодируется с помощью кода в форме (34) одного базисного возможного решения и упорядоченных множеств кодов малых вариаций этого базисного решения.

$$W = \{W_1, \dots, W_H\} = \{W_1 = (w^{1,1}, \dots, w^{1,d}), \dots, W_H = (w^{H,1}, \dots, w^{H,d})\}, \quad (36)$$

где

$$w^{i,j} = [w_1^{i,j} \ \dots \ w_k^{i,j}]^T, \quad i = 1, \dots, H, \quad j = 1, \dots, d, \quad (37)$$

вектор кода малой вариации, H — количество возможных решений, d — длина вариации.

Компоненты вектора кода малой вариации указывают тип и место изменения кода базисного решения. Например, w_1 — номер подмножества в коде (34), w_2 — номер элемента этого подмножества, w_3 — номер элемента в коде элементарной функции и w_4 — новое значение элемента,

$$a_{w_3}^{w_1, w_2} = w_4. \quad (38)$$

Операция скрещивания в данном случае выполняется по правилам классического генетического алгоритма над упорядоченными множествами кодов малых вариаций. Отбираются два “родителя”

$$W_\alpha = (w^{\alpha,1}, \dots, w^{\alpha,d}), \quad W_\beta = (w^{\beta,1}, \dots, w^{\beta,d}), \quad (39)$$

находится случайно точка скрещивания $k \in \{1, \dots, d\}$ и обмениваются элементы после точки скрещивания. В результате получаем два новых возможных решения — “потомка”

$$\begin{aligned} W_{H+1} &= (w^{\alpha,1}, \dots, w^{\alpha,k}, w^{\beta,k+1}, \dots, w^{\beta,d}), \\ W_{H+2} &= (w^{\beta,1}, \dots, w^{\beta,k}, w^{\alpha,k+1}, \dots, w^{\alpha,d}). \end{aligned} \quad (40)$$

Задача машинного обучения всегда связана с поиском неизвестной функции,

$$y = \alpha(x, q), \quad (41)$$

где y — вектор значений функции, $y \in Y \subseteq \mathbb{R}^r$, x — вектор аргументов, $x \in \mathbb{R}^n$, q — вектор постоянных параметров, $q \in Q \subseteq \mathbb{R}^p$, $\alpha(x, q) : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^r$, либо эта функция при обучении аппроксимирует некоторый набор данных, который называется обучающей выборкой,

$$J = \sum_{i=1}^N \|\hat{y}^i - \alpha(x^i, q)\|, \quad (42)$$

где $\hat{Y} = \{\hat{y}^1, \dots, \hat{y}^N\}$ — обучающая выборка, либо она непосредственно используется для минимизации некоторого заданного функционала

$$J = \int_0^{t_f} f_0(x(t), \alpha(x(t), q)) dt, \quad (43)$$

t_f — время достижения цели.

Особенность машинного обучения состоит в том, что при обучении не требуется точного достижения минимума критерия (42) или (43)

$$J_1 \leq \min J + \Delta^*, \quad (44)$$

где Δ^* — заданная положительная величина, которая определяет достижимое при обучении значение функционала.

Для критерия (42) минимальное значение равно нулю. Для критерия (43) оно может быть неизвестно, тогда вместо $\min J$ следует использовать предельное значение критерия J^-

$$J_1 \leq J^- + \Delta^*, \quad (45)$$

где $J^- \leq \min J$.

Если в результате обучения найденная функция (41) должна приобрести некоторые свойства, то доказательство наличия этих свойств подтверждается моделированием

$$J_2 = \sum_{i=1}^K \vartheta(\varphi(\alpha(x^i, q))) = 0, \quad (46)$$

где $\vartheta(z)$ – функция Хэвисайда,

$$\vartheta(z) = \begin{cases} 1, & \text{если } z > 0 \\ 0 & - \text{иначе} \end{cases}, \quad (47)$$

$\varphi(\alpha(x, q))$ – условие, которое определяет наличие свойства функции

$$\varphi(\alpha(x, q)) \leq 0, \quad (48)$$

K – количество экспериментов с выполненным условием (48).

Переформулируем задачу синтеза управления так, чтобы при ее решении использовать машинное обучение методом символьной регрессии. Задана математическая модель объекта управления (1), ограничение на управление (2). Задано множество точек начальных состояний

$$\tilde{X}_0 = \{x^{0,1}, \dots, x^{0,K}\}. \quad (49)$$

Задано терминальное состояние (4). Задан критерий качества управления

$$J_3 = \sum_{i=1}^K \left(p_i \|x^f - x(t_{f,i}, x^{0,i})\| + \int_0^{t_{f,i}} f_0(x, u) dt \right) \rightarrow \min_{u \in U}, \quad (50)$$

где p_i – заданный весовой коэффициент,

$$t_{f,i} = \begin{cases} t, & \text{если } t < t^+ \text{ и } \|x^f - x(t, x^{0,i})\| \leq \varepsilon \\ t^+, & \text{иначе} \end{cases}, \quad (51)$$

t^+ , ε – заданные положительные величины. Необходимо найти управление в форме (30), чтобы минимизировать функционал (50) для любого количества K начальных условий из заданной области (49). Постановка задачи включает функционал, который вместе с минимизацией заданного критерия (5) учитывает проверку свойства устойчивости терминального состояния. Решение задачи синтеза управления обеспечивает за счет свойства устойчивости терминального состояния

свойство сжимающего отображения в окрестности этого терминального состояния.

Другим подходом к решению расширенной задачи оптимального управления является метод синтезированного управления [7]. Согласно этому методу, первоначально независимо от вида функционала (5) решается задача синтеза управления для обеспечения устойчивости относительно некоторой точки в пространстве состояний.

В задаче задана модель (1), множество точек начальных условий (49) в некоторой области (29) пространства состояний

$$\tilde{X}_0 \subset X_0, \quad (52)$$

задано терминальное состояние в той же области

$$x^* \in X_0, \quad x^* \notin \tilde{X}_0, \quad (53)$$

задан функционал качества

$$J_4 = \sum_{i=1}^K (p_i \|x^f - x(t_{f,i}, x^{0,i})\| + t_{f,i}) \rightarrow \min_{u \in U}. \quad (54)$$

Необходимо найти управление в форме

$$u = h(x^* - x) \in U, \quad (55)$$

чтобы минимизировать функционал (54).

На втором этапе рассматриваем модель замкнутой системы управления

$$\dot{x} = f(x, h(x^* - x)). \quad (56)$$

Для этой системы решаем задачу оптимального управления с критерием качества (5).

$$J_5 = \int_0^{t_f} f_0(x, h(x^* - x)) dt \rightarrow \min_{z^*}. \quad (57)$$

В качестве управления используем вектор $x^* = [x_1^* \dots x_n^*]^T$, определяющий положение устойчивой точки равновесия в пространстве состояний с ограничением (55). В синтезированном управлении в окрестности оптимальной траектории всегда существует устойчивая точка равновесия, к которой стремится объект управления, поэтому траектория движения объекта всегда находится в области, где имеются свойства сжимающего отображения, т.е. выполняются требования расширенной задачи оптимального управления.

В качестве примера рассмотрим задачу оптимального управления группой из двух мобильных роботов с дифференциальным приводом [8]. Два одинаковых мобильных робота должны поменяться местами, не столкнувшись друг с другом в процессе движения за минимальное время. Математические модели роботов имеют следующий вид:

$$\dot{x}_1^j = 0.5(u_1^j + u_2^j) \cos(x_3^j), \quad (58)$$

$$\dot{x}_2^j = 0.5(u_1^j - u_2^j) \sin(x_3^j), \quad (59)$$

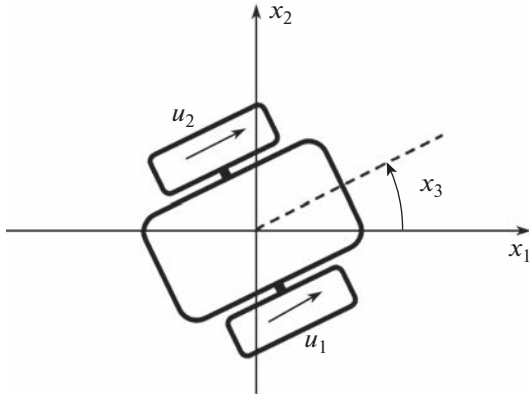


Рис. 1. Схема мобильного робота с дифференциальным приводом.

$$\dot{x}_3^j = 0.5(u_1^j - u_2^j), \quad (60)$$

$$-10 = u_i^- \leq u_i^j \leq u_i^+ = 10, \quad i = 1, 2, \quad (61)$$

$j = 1, 2$.

Схема мобильных роботов с дифференциальным приводом с указанием компонент вектора состояний приведена на рис. 1 [8].

Для системы (58)–(61) заданы начальные условия

$$\begin{aligned} \mathbf{x}^1(0) = \mathbf{x}^{1,0} &= [0 \ 0 \ 0]^T, \\ \mathbf{x}^2(0) = \mathbf{x}^{2,0} &= [10 \ 10 \ \pi]^T. \end{aligned} \quad (62)$$

Заданы терминальные состояния

$$\begin{aligned} \mathbf{x}^1(t_f) = \mathbf{x}^{1,f} &= [10 \ 10 \ 0]^T, \\ \mathbf{x}^2(t_f) = \mathbf{x}^{2,f} &= [0 \ 0 \ \pi]^T, \end{aligned} \quad (63)$$

где $t_{j,f}$ — время достижения терминального состояния роботом j не задано, но ограничено и вычисляется из следующего соотношения

$$t_{j,f} = \begin{cases} t, & \text{если } t < t^+ \text{ и } \|\mathbf{x}^{j,f} - \mathbf{x}(t)\| \leq \varepsilon_0, \\ t^+ & \text{иначе} \end{cases}, \quad (64)$$

$$t^+ = 2.8, \varepsilon_0 = 0.01, j = 1, 2.$$

Задан критерий качества

$$\begin{aligned} J_6 = & p_1 \sum_{j=1}^2 \|\mathbf{x}^{j,f} - \mathbf{x}^j(t_{j,f})\| + t_f + \\ & + p_2 \sum_{j=1}^2 \int_0^{t_f} \sum_{i=1}^2 \vartheta(\varphi_i(\mathbf{x}^j)) dt + \\ & + p_3 \int_0^{t_f} \theta(\chi(\mathbf{x}^1, \mathbf{x}^2)) dt \rightarrow \min_{\mathbf{u}^1, \mathbf{u}^2 \in U}, \end{aligned} \quad (65)$$

где $p_1 = 1, p_2 = 3, p_3 = 3, t_f = \max\{t_{1,f}, t_{2,f}\}$,

$$\varphi_i(\mathbf{x}^j) = r_i - \sqrt{(x_1^j - x_{1,i})^2 + (x_2^j - x_{2,i})^2}, \quad i = 1, 2, \quad (66)$$

$$r_1 = 1, r_2 = 1, x_{1,1} = 1, x_{1,2} = 2.5, x_{2,1} = 4, x_{2,2} = 2.5, j = 1, 2,$$

$$\chi(\mathbf{x}^1, \mathbf{x}^2) = r_0 - \sqrt{(x_1^1 - x_1^2)^2 + (x_2^1 - x_2^2)^2}, \quad (67)$$

$$r_0 = 1.$$

Первоначально была решена задача оптимального управления в классической постановке. Для решения задачи был использован прямой подход, в котором функция управления была аппроксимирована кусочно-линейной функцией времени

$$u_i^j = \begin{cases} u_i^+, & \text{если } \tilde{u}_i^j \geq u_i^+ \\ u_i^-, & \text{если } \tilde{u}_i^j \leq u_i^- \\ \tilde{u}_i^j & \text{иначе} \end{cases}, \quad i = 1, 2, \quad j = 1, 2, \quad (68)$$

где

$$\begin{aligned} \tilde{u}_i^j = & (q_{k+1+M(i-1)+(j-1)2(M+1)} - q_{k+M(i-1)+(j-1)2(M+1)}) \times \\ & \times \frac{t - k\Delta t}{\Delta t} + q_{k+M(i-1)+(j-1)2(M+1)}, \end{aligned} \quad (69)$$

$i = 1, 2, j = 1, 2, k = 1, \dots, M + 1, M$ — количество интервалов, Δt — временной интервал, $\Delta t = 0.2$, $M = \lfloor t^+ / \Delta t \rfloor = \lfloor 2.8 / 0.2 \rfloor = 14$.

В каждом интервале необходимо найти по два значения на границах интервала для каждого управления. Итого необходимо найти $(2 + 2) \times (14 + 1) = 4 \times 15 = 60$ значений параметров, $\mathbf{q} = [q_1 \dots q_{60}]^T$. Многочисленные исследования по сравнению градиентных алгоритмов с эволюционными для решения сложных задач оптимального управления с фазовыми ограничениями показали преимущество последних [9], так как фазовые ограничения рассматриваемого типа приводят к неунимодальности функционала. Для поиска решения был использован эволюционный гибридный алгоритм [10]. В результате были получены кусочно-линейные аппроксимации функций управления для каждого робота, определяемые значением вектора параметров. Найденное значение вектора параметров приведено в Приложении 1.

На рис. 2 представлены оптимальные траектории движения роботов на горизонтальной плоскости. На рисунке сплошная линия — траектория первого робота, пунктирная линия — траектория второго робота. Красные окружности обозначают фазовые ограничения. Как видно из рис. 2, оба робота достигают терминального состояния и не нарушают фазовых ограничений. Значение функционала (65) для найденного оптимального решения составило величину $J_6 = 2.8101$.

Далее данная задача была решена методом синтезированного управления. Первоначально

решена задача синтеза управления для каждого объекта с целью обеспечения устойчивости относительно точки в пространстве состояний. Для этой цели сформируем множество из 27 начальных состояний

$$\hat{X}_0 = \{\mathbf{x}^{0,i+(j-1)3+(k-1)9} = [d_1 + (i-1)\Delta_1 \ d_2 + (j-1)\Delta_2 \ d_3 + (k-1)\Delta_3]^T, \quad (70)$$

где $d_1 = -2.5$, $d_2 = -2.5$, $d_3 = -5\pi/12$, $\Delta_1 = 2.5$, $\Delta_2 = 2.5$, $\Delta_3 = 5\pi/12$, $i, j, k = 1, 2, 3$.

Задаем одно терминальное состояние

$$\mathbf{x}^* = [0 \ 0 \ 0]^T. \quad (71)$$

Необходимо найти функцию управления вида (55), обеспечивающую минимум следующего критерия

$$J_7 = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 (p_i \|\mathbf{x}^* - \mathbf{x}(t_{f,r})\| + t_{f,r}) \rightarrow \min_{u \in U}, \quad (72)$$

где $t_{f,r}$ – время достижения терминального состояния (71) определяется из соотношения (51) с $t^+ = 2$, $\varepsilon_0 = 0.01$, $r = i + (j-1)3 + (k-1)9$, $i, j, k = 1, 2, 3$.

Для решения задачи использован метод сетевого оператора [5], который нашел следующую функцию управления:

$$u_i = \begin{cases} u_i^+, & \text{если } \hat{u}_i \geq u_i^+ \\ u_i^-, & \text{если } \hat{u}_i \leq u_i^-, \quad i = 1, 2, \\ \hat{u}_i & - \text{иначе} \end{cases} \quad (73)$$

где

$$\hat{u}_1 = A^{-1} + \sqrt[3]{A} + \operatorname{sgn}(q_3(x_3^* - x_3)) \times \quad (74)$$

$$\times \exp(-|q_3(x_3^* - x_3)|) + \operatorname{sgn}(x_3^* - x_3) + \mu(B),$$

$$\hat{u}_2 = \hat{u}_1 + \sin(\hat{u}_1) + \arctan(H) + \mu(B) + C - C^3, \quad (75)$$

$$A = \tanh(D) + \left(B + \sqrt[3]{x_1^* - x_1} \right)^3 +$$

$$+ C + \sin(q_3(x_3^* - x_3)),$$

$$B = \operatorname{sgn}(\operatorname{sgn}(x_1^* - x_1)q_2(x_2^* - x_2)) \times$$

$$\times \exp(-|\operatorname{sgn}(x_1^* - x_1)q_2(x_2^* - x_2)|) +$$

$$+ G + \sin(x_1^* - x_1) + \tanh(G) + x_1^* - x_1,$$

$$C = \operatorname{sgn}(\operatorname{sgn}(x_1^* - x_1)q_2(x_2^* - x_2)) \times$$

$$\times \exp(-|\operatorname{sgn}(x_1^* - x_1)q_2(x_2^* - x_2)|) +$$

$$+ \sin(x_1^* - x_1),$$

$$D = H + C - C^3 + \operatorname{sgn}(q_1(x_1^* - x_1)) +$$

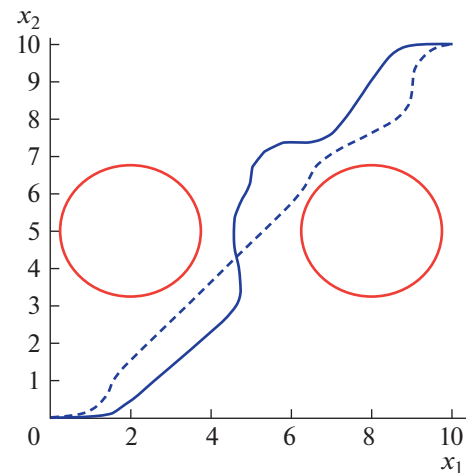


Рис. 2. Оптимальные траектории движения роботов по горизонтальной плоскости.

$$+ \arctan(q_1) + \vartheta(x_3^* - x_3),$$

$$G = \operatorname{sgn}(x_1^* - x_1)q_2(x_2^* - x_2) +$$

$$+ q_3(x_3^* - x_3) + \tanh(q_1(x_1^* - x_1)),$$

$$H = \arctan(q_1(x_1^* - x_1)) + \operatorname{sgn}(W)\sqrt{|W|} +$$

$$+ W + V + 2\operatorname{sgn}(W + \tanh(V)) +$$

$$+ \sqrt[3]{W + \tanh(V)} + \sqrt[3]{x_1^* - x_1} +$$

$$+ \operatorname{sgn}(x_1^* - x_1)\sqrt{|x_1^* - x_1|} + \sqrt[3]{x_1^* - x_1} + \tanh(V),$$

$$W = \operatorname{sgn}(x_1^* - x_1) + \operatorname{sgn}(q_2(x_2^* - x_2)) \times$$

$$\times \operatorname{sgn}(x_1^* - x_1) \tanh(x_1^* - x_1),$$

$$V = q_3(x_3^* - x_3) + \operatorname{sgn}(x_1^* - x_1)q_2(x_2^* - x_2) +$$

$$+ \tanh(x_1^* - x_1),$$

$$\mu(\alpha) = \operatorname{sgn}(\alpha) \min\{1, |\alpha|\}, \quad \tanh(\alpha) = \frac{1 - \exp(-2\alpha)}{1 + \exp(-2\alpha)},$$

$$q_1 = 14.7288, q_2 = 2.0271, q_3 = 4.0222.$$

На втором этапе решаем задачу оптимального управления, в которой необходимо найти кусочно-постоянную функцию времени

$$x_i^{j*,k} = q_{i+(k-1)3+(j-1)M}, \quad (k-1)\Delta t \leq t < k\Delta t, \quad (76)$$

где $j = 1, 2$, $i = 1, 2, 3$, $k = 1, \dots, M$, M – число интервалов времени, $\Delta t = 0.4$, $M = \lceil t^+ / \Delta t \rceil = \lceil 2.8 / 0.4 \rceil = 7$. Всего необходимо найти вектор из $6 \times 7 = 42$ параметров для оптимизации функционала

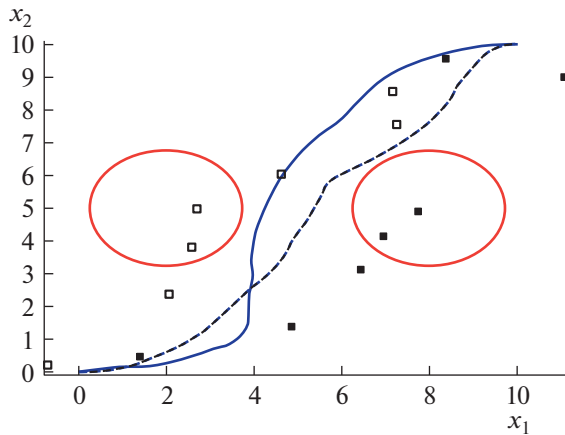


Рис. 3. Оптимальные траектории для синтезированного управления.

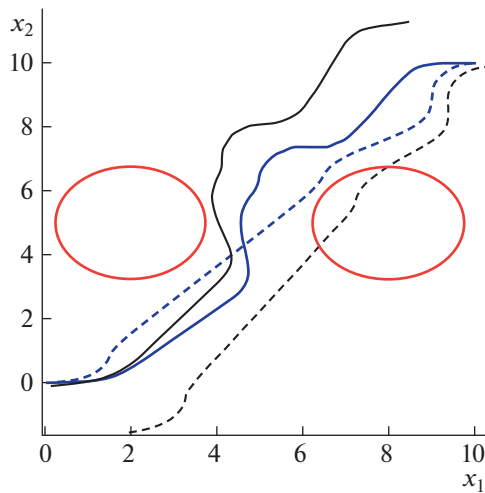


Рис. 4. Возмущенные и оптимальные траектории для прямого управления.

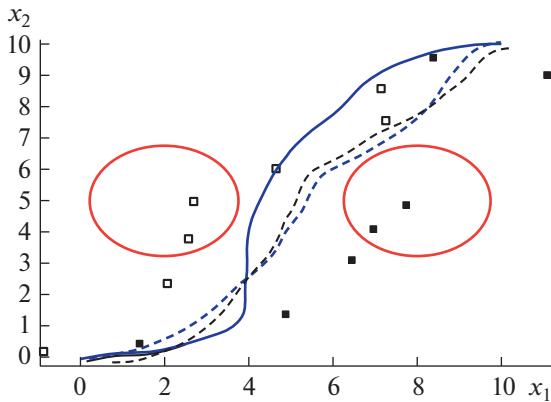


Рис. 5. Возмущенные и оптимальные траектории для синтезированного управления.

$$J_8 = p_1 \sum_{j=1}^2 \|x^{j,f} - x^j(t_{j,f})\| + t_f +$$

$$+ p_2 \sum_{j=1}^2 \int_0^{t_f} \int_{i=1}^2 \vartheta(\varphi_i(x^j)) dt +$$

$$+ p_3 \int_0^{t_f} \theta(\chi(x^1, x^2)) dt \rightarrow \min_{x^{1,*}, x^{2,*}}. \quad (77)$$

Для решения задачи был использован тот же гибридный эволюционный алгоритм [10].

На рис. 3 приведены оптимальные траектории движения мобильных роботов на плоскости для найденного оптимального управления. Значение функционала для полученного решения равно $J_8 = 2.9283$. Значения найденных оптимальных параметров приведены в приложении 2. На рис. 3 в виде малых квадратов приведены проекции на плоскость $\{x_1, x_2\}$ найденных оптимальных значений векторов управлений, черные квадраты — проекции точек стабилизации первого робота $x^{1,*k} = [x_1^{1,*k} \ x_2^{1,*k} \ x_3^{1,*k}]^T$, белые квадраты — проекции точек стабилизации второго робота $x^{2,*k} = [x_1^{2,*k} \ x_2^{2,*k} \ x_3^{2,*k}]^T$, $k = 1, \dots, 7$.

Для сравнения полученных результатов начальные состояния объектов управления были подвергнуты случайным возмущениям. Во всех испытаниях траектории, полученные синтезированным управлением, сохраняли характер движения. Значение функционала при возмущениях синтезированного управления не сильно отличалось от оптимального. Те же возмущения при прямом управлении существенно меняли характер движения и увеличивали значение функционала.

На рис. 4 и 5 приведены возмущенные траектории движения для прямого и синтезированного управлений для одних и тех же начальных условий

$$x^{0,1} = [0.15 \ -0.1 \ 0.15]^T, \quad x^{0,2} = [10.2 \ 9.88 \ 3.3]^T.$$

На рис. 4, 5 приведены синим цветом для сравнения также оптимальные невозмущенные траектории. Как видно из результатов эксперимента, возмущенные траектории прямого управления имеют существенную ошибку достижения терминальных состояний. Значение функционала для возмущенного решения для прямого управления составило $J_6 = 7.625$, т.е. увеличилось более, чем на 170%. Возмущенные траектории для синтезированного управления незначительно отличаются от оптимальных. Для первого робота возмущенная и оптимальная траектории практически совпали. Значение функционала для возмущенного решения для синтезированного управления

составило $J_8 = 4.11$, т.е. увеличилось на 40%, по-видимому, из-за нарушений динамических фазовых ограничений (67).

В результате решения расширенной задачи оптимального управления была найдена функция управления в форме (16), которая является кусочно-постоянной функцией времени и зависит от координат вектора пространства состояний. Найденная функция управления на каждом определенном временном интервале обеспечивает для модели замкнутой системы управления

наличие устойчивой точки равновесия в пространстве состояний, что приводит к выполнению условий (20), (21), наличию окрестности у оптимального решения со свойствами сжимающего отображения. Экспериментально подтверждено, что найденное решение расширенной задачи оптимального управления существенно менее чувствительно к возмущениям модели, чем прямое решение классической задачи оптимального управления.

Приложение 1

Оптимальные значения параметров для прямого метода

$$\begin{aligned} \mathbf{q} = & [15.7490 \ 15.4717 \ 14.3902 \ 6.0419 \ 10.3588 \ 5.9336 \ -0.1257 \ 8.7810 \\ & -16.5382 \ -0.0040 \ 3.5222 \ 11.1793 \ 15.4434 \ -18.1304 \ 19.9967 \ 12.0334 \\ & -0.5850 \ 17.2859 \ 5.7742 \ -4.0897 \ 14.5804 \ 6.9570 \ -17.6313 \ 15.2778 \\ & 18.8768 \ -16.5075 \ 0.3253 \ 16.7574 \ 3.4849 \ 19.9379 \ -0.4417 \ 19.9998 \\ & 3.7046 \ -6.5899 \ 10.1436 \ 10.6777 \ -2.8380 \ 18.2116 \ 15.2737 \ 19.2273 \\ & 7.6933 \ -16.2777 \ -0.9072 \ 0.0478 \ 5.5856 \ -0.4617 \ -19.6399 \ -4.8071 \\ & 15.3430 \ 11.5493 \ 0.0603 \ 4.9952 \ 16.3146 \ 15.9367 \ 19.9823 \ -4.5805 \\ & 5.9542 \ -15.3081 \ 19.5269 \ 4.9039]^T \end{aligned}$$

Приложение 2

Оптимальные значения параметров для синтезированного управления

$$\begin{aligned} \mathbf{x}^{1*,1} &= [1.4057 \ 0.4520 \ -0.1663]^T, \\ \mathbf{x}^{2*,1} &= [7.2674 \ 7.5531 \ 3.1760]^T, \\ \mathbf{x}^{1*,2} &= [4.8571 \ 1.3882 \ -0.1167]^T, \\ \mathbf{x}^{2*,2} &= [7.1637 \ 8.5559 \ 4.1323]^T, \\ \mathbf{x}^{1*,3} &= [6.9637 \ 4.1181 \ 0.6035]^T, \\ \mathbf{x}^{2*,3} &= [4.6354 \ 6.0315 \ 3.3282]^T, \\ \mathbf{x}^{1*,4} &= [6.4364 \ 3.12121 \ 1.5037]^T, \\ \mathbf{x}^{2*,4} &= [2.6965 \ 4.9781 \ 4.1777]^T, \\ \mathbf{x}^{1*,5} &= [7.7471 \ 4.9075 \ 1.5244]^T, \\ \mathbf{x}^{2*,5} &= [2.5864 \ 3.8087 \ 4.1210]^T, \\ \mathbf{x}^{1*,6} &= [8.3708 \ 9.5678 \ 0.1377]^T, \\ \mathbf{x}^{2*,6} &= [2.0666 \ 2.3788 \ 4.3185]^T, \end{aligned}$$

$$\mathbf{x}^{1*,7} = [11.1035 \ 9.0000 \ 0.4429]^T,$$

$$\mathbf{x}^{2*,7} = [-0.7062 \ 0.2167 \ 3.2446]^T.$$

ИСТОЧНИК ФИНАНСИРОВАНИЯ

Работа выполнена при частичной финансовой поддержке гранта Российского научного фонда № 23-29-00339 “Исследование расширенной задачи оптимального управления и разработка численных методов ее решения”.

СПИСОК ЛИТЕРАТУРЫ

1. Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. М. Наука. Главное издание физико-математической литературы. 1983. 392 с.
2. Болтянский В.Г. Математические методы оптимального управления. М.: Наука. Главное издание физико-математической литературы. 1969. 408 с.
3. Халил Х.К. Нелинейные системы. Ижевск.: Институт компьютерных исследований, 2009. 814 с.
4. Колесников А.А. Синергетические методы управления сложными системами. Теория системного анализа. М.: Книжный дом “ЛИБРОКОМ” 2012. 237 с.
5. Diveev A.I., Shmalko E.Yu. Machine learning Control by Symbolic Regression. Cham, Switzerland, Springer,

2021. 155 p. <https://link.springer.com/book/10.1007/978-3-030-83213-1>
6. Diveev A.I., Sofronova E.A. Universal Approach to Solution of Optimization Problems by Symbolic Regression // Applied Science. 2021. V. 11. P. 5081.
 7. Diveev A., Shmalko E. Yu., Serebrenny V. Zentay P. Fundamentals of Synthesized Optimal Control // Mathematics. 2021. V. 9. P. 21.
 8. Dhaouadi R., Abu Hatab A. Dynamic Modelling of Differential-Drive Mobile Robots using Lagrange and Newton-Euler Methodologies: A Unified Framework // Advances in Robotics A & Automation. 2013. V. 2. Issue 2. P. 1–7.
 9. Дивеев А.И., Константинов С.В. Исследование практической сходимости эволюционных алгоритмов оптимального программного управления колесным роботом // Известия РАН. Теория и системы управления. 2018. Т. 57. № 4. С. 80–106.
 10. Дивеев А.И. Гибридный эволюционный алгоритм для решения задачи оптимального управления // Вопросы теории безопасности и устойчивости систем. 2021. № 23. С. 3–12.

REFINEMENT OF OPTIMAL CONTROL PROBLEM FOR PRACTICAL IMPLEMENTATION OF ITS SOLUTION

A. I. Diveev^a

^a Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences,
Moscow, Russian Federation

Presented by Academician of the RAS I.A. Sokolov

The solution of the optimal control problem in the classical formulation is control in the form of a function of time. The implementation of such a solution leads to an open control system and therefore cannot be applied directly in practice. It is believed that solving the classical optimal control problem leads to an optimal control program and program trajectory in state space. To implement the movement of the control object along the program trajectory, it is necessary to build an additional movement stabilization system. The problem of synthesizing a system for stabilizing movement along a program trajectory and the requirements that this system should meet do not arise from the classical setting of the optimal control problem. An updated statement of the optimal control problem is given, which includes an additional requirement for an optimal trajectory, and the solution of which can be directly applied in practice in a real control object.

Keywords: optimal control, control synthesis, contracting mapping