УДК 519.6 EDN: XKJFLR

МАШИННОЕ ОБУЧЕНИЕ УНИВЕРСАЛЬНОЙ СИСТЕМЫ СТАБИЛИЗАЦИИ ДВИЖЕНИЯ В РАСШИРЕННОЙ ЗАДАЧЕ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

А.И. Дивеев aidiveev@mail.ru

ФИЦ ИУ РАН, Москва, Российская Федерация

Аннотация

Рассмотрена расширенная задача оптимального управления, в которой необходимо не только определить функцию управления как функцию времени, обеспечивающую достижение цели управления с оптимальным значением заданного критерия качества, но и обеспечить существование свойства притяжения оптимальной траектории в некоторой ее окрестности. Это свойство позволяет найти такую функцию управления, которую можно реализовать непосредственно в реальном объекте. Для решения задачи синтезируется универсальная система стабилизации движения объекта управления по заданной траектории. Здесь использовано машинное обучение управления с использованием символьной регрессии. Символьная регрессия позволяет найти структуру и параметры функции управления без участия человека. Чтобы обеспечить универсальность синтезированной системы стабилизации, символьная регрессия ищет одну систему стабилизации для некоторых заданных различных траекторий. Синтезированная система стабилизации и исходная математическая модель объекта управления вводятся в систему управления объектом. Полученная математическая модель объекта управления позволяет решить расширенную задачу оптимального управления и получить функцию управления, реализуемую непосредственно в реальном объекте. Представлен пример решения расширенной задачи оптимального управления пространственным движением квадрокоптера

Ключевые слова

Оптимальное управление, символьная регрессия, реализуемая функция управления, система стабилизации, синтез управления

Поступила 01.04.2024
Принята 22.04.2024
© Автор(ы), 2025

Работа выполнена при частичной финансовой поддержке РНФ (проект № 23-29-00339)

Введение. Приведем следующее определение машинному обучению [1]: «Говорят, что компьютерная программа обучается на опыте E относительно некоторого класса задач T и меры качества P, если ее качество на задачах, принадлежащих T, измеренное в соответствии с P, улучшается с увеличением опыта E». Это определение считается классическим и приводится в большинстве научных работ по машинному обучению. Согласно этому определению, машинное обучение — процесс, когда одна компьютерная программа пишет другую компьютерную программу или, когда компьютерная программа изменяется для улучшения своей работы. Искусственная нейронная сеть — универсальный аппроксиматор любой функции. Это функция с заданной структурой и большим числом настраиваемых параметров. Другая обучающая программа находит значения этих настраиваемых параметров, чтобы искусственная нейронная сеть по заданным входным данным возвращала заданные значения выходных данных, т. е. ее работа соответствовала некоторой заданной функции. Искусственная нейронная сеть аппроксимирует любую функцию при наличии хорошей обучающей выборки. Еще одной менее распространенной конструкцией универсальной аппроксимации любой функции является символьная регрессия. Символьная регрессия позволяет находить не только значения параметров, но и структуру искомой функции. Для ее работы не требуется обучающая выборка. Символьная регрессия ищет функцию как решение задачи оптимизации для минимизации некоторой заданной целевой функции. В настоящей работе символьная регрессия использована для решения задачи синтеза системы стабилизации движения объекта управления по оптимальной траектории. Такая проблема возникает, когда необходимо реализовать решение задачи оптимального управления непосредственно в реальном объекте, для которого эта задача решалась. Задача оптимального управления, в которой требуется найти реализуемую функцию управления, называется расширенной задачей оптимального управления. Рассмотрим эту проблему более подробно.

Задача оптимального управления в классической постановке [2] требует нахождения решения в виде функции, зависящей от времени. Полагается, что такая функция управления не может быть реализована непосредственно на реальном объекте, поскольку это управление является разомкнутым и система управления не содержит отрицательной обратной связи. При наличии внешних возмущений или ошибок математической модели объекта управления, используемой в задаче оптимального управления, рассчитанная оптимальная траектория будет отличаться от реального движения объекта управления. Это означает, что объект управления

не достигнет цели управления и значение критерия качества не будет оптимальным.

Для реализации решения задачи оптимального управления в реальном объекте необходимо сначала получить оптимальную траекторию в пространстве состояний, а затем построить систему стабилизации движения объекта управления по оптимальной траектории в пространстве состояний. Оптимальная траектория получается решением системы обыкновенных дифференциальных уравнений модели объекта управления с найденной при решении задачи оптимального управления функцией управления как функции времени. Система стабилизации движения объекта управления по оптимальной траектории представляет собой систему слежения, компенсирующую отклонения в пространстве состояний реального положения объекта управления от заданной во времени оптимальной траектории. Математическая модель объекта управления с системой слежения за оптимальной траекторией представляет собой систему дифференциальных уравнений, имеющую специальное частное решение в виде оптимальной траектории, обладающей свойством притяжения в некоторой ее окрестности. Любая система слежения за траекторией является функцией обратной связи, поэтому она изменяет динамические свойства модели объекта управления. В результате можно сделать вывод, что задача оптимального управления решена для одной математической модели объекта управления и для нее получена оптимальная траектория, а на практике объект управления с другой математической моделью движется по этой оптимальной траектории, т. е. оптимальная траектория, полученная решением задачи оптимального управления, не оптимальна для модели реального объекта, движущегося по этой траектории.

Практически все системы стабилизации движения по заданной траектории строятся по критерию точности отслеживания траектории и, как правило, не учитывают исходный критерий задачи оптимального управления, в результате решения которой была получена данная траектория. Следовательно, если при решении задачи оптимального управления использовать другую математическую модель со следящей функцией в обратной связи, то можно получить иное значение критерия качества.

Здесь система стабилизации движения объекта управления по траектории строится раньше, чем решение задачи оптимального управления и получение оптимальной траектории. Для этого задается набор различных траекторий. Затем на основе машинного обучения методом символьной регрессии решается задача синтеза системы стабилизации движения по всем траекториям из набора. В результате получается универсальная

система стабилизации. В качестве эталонной модели используется исходная модель объекта управления с вектором управления в правой части. Любое частное решение эталонной модели представляет собой заданную траекторию движения, которую отслеживает система стабилизации. Задача оптимального управления состоит в нахождении функции управления как функции времени для эталонной модели. После решения задачи оптимального управления математическая модель объекта управления не меняется.

Постановка задачи и подход к решению. Рассмотрим задачу оптимального управления. Приведена математическая модель объекта управления в виде системы обыкновенных дифференциальных уравнений

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

где $\mathbf{x} = [x_1 \ \ldots \ x_n]^{\mathrm{T}}$ — вектор состояний, $\mathbf{x} \in \Re^n$; $\mathbf{u} = [u_1 \ \ldots \ u_m]^{\mathrm{T}}$ — вектор управления, $\mathbf{u} \in \mathrm{U} \subseteq \Re^m$, U — компактное множество, которое определяет ограничения на значения управления;

$$\mathbf{u}^- \le \mathbf{u} \le \mathbf{u}^+, \tag{2}$$

$\mathbf{u}^-$, $\mathbf{u}^+$ — заданные нижнее и верхнее предельные значения вектора управления.

Заданы начальное

$$\mathbf{x}(0) = \mathbf{x}^0 = \left[x_1^0 \ \ldots \ x_n^0 \right]^{\mathrm{T}} \tag{3}$$

и терминальное состояния

$$\mathbf{x}(t_f) = \mathbf{x}^f = \left[x_1^f \ \ldots \ x_n^f \right]^{\mathrm{T}}. \tag{4}$$

Здесь t_f — время достижения терминального состояния (4) — не задано, но ограничено, $t_f \le t^+$, t^+ — заданное предельное время процесса управления.

Задан критерий качества

$$J_1 = \int_0^{t_f} f_0(\mathbf{x}, \mathbf{u}) dt \to \min_{\mathbf{u} \in \mathrm{U}}. \tag{5}$$

Уравнения (1)–(5) описывают классическую постановку задачи оптимального управления. В классической постановке необходимо найти управление как функцию времени

$$\mathbf{u} = \mathbf{v}(t) \in \mathrm{U}, \ 0 \le t \le t_f. \tag{6}$$

Если подставить найденную функцию управления (6) в правую часть модели (1), то получим систему уравнений

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

имеющую частное решение из начального состояния (3), которое будет достигать заданное терминальное состояние (4) с оптимальным значением критерия качества (5). Это частное решение является оптимальной траекторией движения объекта в пространстве состояний

$$\mathbf{x}^*(t) \in \Re^n. \tag{8}$$

При обычном решении задачи оптимального управления для реализации движения объекта управления по оптимальной траектории строится система стабилизации, но, как было отмечено, эта система меняет математическую модель объекта и к тому же требует дополнительных ресурсов управления, которых может и не быть. В рассматриваемом случае предположим, что класс, которому принадлежит оптимальная траектория (8), известен. Построим еще до решения задачи оптимального управления универсальную систему стабилизации движения объекта по любой траектории из этого класса

$$\mathbf{u} = \mathbf{h}(\mathbf{x}^* - \mathbf{x}) \in \mathrm{U}, \tag{9}$$

где $\mathbf{h}(\mathbf{x}^* - \mathbf{x}) : \Re^n \to \Re^m$.

Добавим эту систему стабилизации (9) в модель объекта управления и запишем расширенную модель объекта управления

$$\dot{\mathbf{x}}^* = \mathbf{f}(\mathbf{x}^*, \mathbf{u}), \tag{10}$$

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{h}(\mathbf{x}^* - \mathbf{x})), \tag{11}$$

где первая подсистема (10) в расширенной модели является эталонной моделью и предназначена для генерации оптимальной траектории в пространстве состояний, вторая подсистема (11) является моделью объекта управления с системой стабилизации движения вдоль заданной траектории в пространстве состояний.

Для эталонной модели (10) начальное состояние совпадает с начальным состоянием (3), заданным в исходной постановке задачи оптимального управления:

$$\mathbf{x}^*(0) = \mathbf{x}^0 = \left[x_1^0 \quad \ldots \quad x_n^0 \right]^{\mathrm{T}}. \tag{12}$$

Получена следующая задача оптимального управления. Математическая модель объекта управления задается в расширенном виде (10), (11).

Даны ограничения на управление (2). Начальное и конечное состояния заданы (3), (4). Необходимо определить функцию управления как функцию времени такую, что после подстановки этой найденной функции управления в правую часть дифференциального уравнения эталонной модели (10) частное решение модели (11) из заданного начальное состояние (3) достигает заданного конечного состояния (4) с минимальным значением заданного критерия качества (5). Для решения задачи оптимального управления для эталонной модели (10) находится функция управления как функция времени (6).

Для решения задачи оптимального управления используем кусочно-линейную аппроксимацию функции управления. Значения компонент управления вычисляем по формуле

$$u_i(t)=\begin{cases} u_i^+, \text{ если } \tilde{u}_i(t)\ge u_i^+; \\ u_i^-, \text{ если } \tilde{u}_i(t)\le u_i^-; \\ \tilde{u}_i(t) \text{ иначе}, \quad i=1,\ldots,m, \end{cases} \tag{13}$$

где

$$\tilde{u}_i(t)=\frac{y_{i+jm}-y_{i+(j-1)m}}{\Delta t}(t-(j-1)\Delta t)+y_{i+(j-1)m}, \tag{14}$$

$(i-1)\Delta t\le t<i\Delta t$, $j=1,\ldots,K$, $i=1,\ldots,m$, Δt — временной интервал, K — число временных интервалов,

$$K=\left\lfloor \frac{t^+}{\Delta t} \right\rfloor. \tag{15}$$

Для каждой компоненты управления определяем $K+1$ значений параметров на границах интервалов:

$$\mathbf{y}=[y_1 \quad \ldots \quad y_{(K+1)m}]^{\mathrm{T}}. \tag{16}$$

В этом случае решение задачи оптимального управления с расширенной моделью (10), (11) объекта управления будет обладать свойством реализуемости, несмотря на то, что найденное решение будет функцией времени, поскольку расширенная модель объекта управления включает в себя систему стабилизации (9) движения объекта управления по оптимальной траектории. Следует отметить, что свойство реализуемости функции управления будет получено, если найденная функция управления для эталонной модели обеспечит формирование оптимальной траектории в классе функций, которые может стабилизировать система стаби-

лизации движения. Отметим также, что система стабилизации (9) в этой задаче синтезирована раньше, чем решалась задача оптимального управления, поэтому при решении задачи синтеза системы стабилизации движения необходимо учитывать класс функций, выбранный для нахождения оптимальной функции управления, т. е. синтезируемая система стабилизации должна обладать свойством универсальности и обеспечивать стабилизацию движения по различным траекториям из некоторого класса. Рассмотрим задачу синтеза универсальной системы стабилизации движения по заданным во времени траекториям.

Синтез универсальной системы стабилизации движения. Задача синтеза управления решается с использованием машинного обучения символьной регрессии. Этот численный метод решения задачи синтеза управления позволяет решить задачу в автоматическом режиме с минимальным применением ручных операций. Как правило, для решения задачи синтеза системы отслеживания траектории исследователь анализирует математическую модель объекта управления для определения каналов управления, задает структуру системы управления и устанавливает контроллеры в каналах управления [3–7] либо использует аналитические методы синтеза, обратный обход интегратора (бэкстеппинг), аналитическое конструирование агрегированных регуляторов (АКАР) или для линейных систем с квадратичным критерием качества — аналитическое конструирование оптимальных регуляторов (АКОР) [8–10], которые также предполагают большой объем ручного труда, в частности, решение специальных алгебраических уравнений. Подходы для решения задачи синтеза управления с применением ручных преобразований нельзя отнести к методам машинного обучения. В отличие от перечисленных методов, символьная регрессия позволяет решать задачу синтеза без применения ручных операций. Она автоматически вставляет математическое выражение искомой функции управления в правую часть системы дифференциальных уравнений модели объекта управления, анализирует качество полученной системы управления и подбирает наиболее оптимальную согласно установленному критерию функцию управления.

Основная проблема синтеза универсальной системы стабилизации в автоматическом режиме состоит в построении множества траекторий, которые должна отслеживать система стабилизации. Множество траекторий в [11] определялось как множество решений различных задач оптимального управления. В указанной работе множество траекторий порождается множеством кусочно-линейных функций управления эталонной модели

$$\mathrm{V}=\left\{\mathbf{v}^{*,1}(t),\ldots,\mathbf{v}^{*,D}(t)\right\}, \tag{17}$$

где D — число кусочно-линейных функций управления.

Для всех траекторий, порожденных набором кусочно-линейных функций управления (17), ищется одна функция управления в виде (9), которая должна минимизировать следующий критерий качества:

$$J_1=\sum_{j=1}^{B}\sum_{i=1}^{D}\int_0^{t^+}||\,\mathbf{x}^*\left(t,\mathbf{x}^{*,0}\right)-\mathbf{x}\left(t,\mathbf{x}^{j,0}\right)||\,dt\rightarrow\min_{\mathbf{h}(\mathbf{x}^*-\mathbf{x})\in\mathrm{U}}. \tag{18}$$

Здесь B — число начальных состояний, используемых для решения задачи синтеза управления. В множество начальных состояний задачи синтеза управления входят первое начальное состояние задачи оптимального управления и отклонения от него в обе стороны, в плюс и в минус:

$$\mathrm{X}_0=\left\{\mathbf{x}^{0,j+1}=\mathbf{x}^0-\boldsymbol{\Delta}_0+\boldsymbol{\Delta}_0\odot(j)_3,\ j=1,\ldots,\tilde{n}^3-1\right\}, \tag{19}$$

где $\boldsymbol{\Delta}_0$ — вектор отклонений начальных условий, $\boldsymbol{\Delta}_0=[\Delta_1\ \ldots\ \Delta_n]^{\mathrm{T}}$; $\odot$ — оператор адамарова покомпонентного произведения векторов; $(j)_3$ — троичный код числа j; $\tilde{n}$ — число компонент вектора состояния, для которых рассматриваются отклонения начальных условий. Например, если $j=7$, то $(7)_3=[0\ \ 2\ \ 1]^{\mathrm{T}}$ и $\mathbf{x}^{0,7}=\left[\begin{array}{ccc}x_1^0-\Delta_1 & x_2^0+\Delta_2 & x_3^0\end{array}\right]^{\mathrm{T}}$.

Для решения задачи синтеза управления используется машинное обучение управлению методом символьной регрессии [12]. Такой подход, как было отмечено выше, автоматизирует процесс создания системы автоматического управления и исключает ручную работу. Метод символьной регрессии включает в себя эволюционный алгоритм для поиска математического выражение на пространстве специальных кодов. Код математического выражения определяется выбранным методом символьной регрессии. В качестве эволюционного алгоритма применяется специальный генетический алгоритм, в котором операция скрещивания выполняется в соответствии с формой кода математического выражения так, чтобы после выполнения операции скрещивания получать правильные коды новых математических выражений.

Вычислительный эксперимент. Рассмотрим задачу оптимального управления пространственным перемещением квадрокоптера. Математическая модель объекта управления представляет собой систему из шести обыкновенных дифференциальных уравнений:

$$\dot{x}_1=x_4, \tag{20}$$

$$\dot{x}_2 = x_5, \tag{21}$$

$$\dot{x}_3 = x_6, \tag{22}$$

$$\dot{x}_4 = u_4 \sin u_3 \cos u_2 \cos u_1 - \sin u_1 \sin u_5, \tag{23}$$

$$\dot{x}_5 = u_4 \cos u_3 \cos u_1 - g, \tag{24}$$

$$\dot{x}_6 = u_4 \cos u_2 \sin u_1 + \cos u_1 \sin u_2 \sin u_5, \tag{25}$$

где g — ускорение свободного падения $g = 9{,}80665$.

Значение вектора управления ограничено

$$\mathbf{u}^- \le \mathbf{u} \le \mathbf{u}^+, \tag{26}$$

где $\mathbf{u}^- = [-\pi/12 \;\; -\pi \;\; -\pi/12 \;\; 0]^{\mathrm{T}}$; $\mathbf{u}^- = [\pi/12 \;\; \pi \;\; \pi/12 \;\; 12]^{\mathrm{T}}$.

Заданы начальное

$$\mathbf{x}(0) = \mathbf{x}^0 = [0 \;\; 5 \;\; 0 \;\; 0 \;\; 0 \;\; 0]^{\mathrm{T}}. \tag{27}$$

и терминальное состояния

$$\mathbf{x}(t_f) = \mathbf{x}^f = [10 \;\; 5 \;\; 10 \;\; 0 \;\; 0 \;\; 0]^{\mathrm{T}}. \tag{28}$$

Здесь

$$t_f = \begin{cases} t, \text{ если } t < t^+ \text{ и } \| \mathbf{x}^f - \mathbf{x}(t) \| \le \varepsilon; \\ t^+ \;\; \text{иначе}, \end{cases} \tag{29}$$

$t^+ = 6{,}4$ с, $\varepsilon = 0{,}05$. Предельное время управления определено с учетом ограничений на управление и расстояние между заданными начальным и терминальным состояниями объекта управления.

Далее необходимо построить универсальную систему стабилизации движения объекта управления вдоль любой траектории, порожденной кусочно-линейной аппроксимацией (14) функции управления. Для этого зададим такие фазовые ограничения, что задача оптимального управления как задача нелинейного программирования на пространстве параметров (16) имеет несколько локальных минимумов, причем локально оптимальные траектории движения объекта управления к терминальному состоянию существенно различаются. Пусть заданы следующие фазовые ограничения:

$$\varphi_i(\mathbf{x}) = r_i - \sqrt{(x_{1,i} - x_1)^2 + (x_{3,i} - x_3)^2} \le 0, \;\; i = 1, \ldots, M = 4, \tag{30}$$

где $r_1 = 2$; $x_{1,1} = 2{,}5$; $x_{3,1} = 2{,}5$; $r_2 = 2{,}5$; $x_{1,2} = 2$; $x_{3,2} = 8$; $r_3 = 2{,}5$; $x_{1,3} = 8$; $x_{3,3} = 2$; $r_4 = 2$; $x_{1,4} = 7{,}5$; $x_{3,4} = 7{,}5$.

Для решения задачи оптимального управления зададим критерий качества управления

$$J_2 = t_f + p_1 \left\| \mathbf{x}^f - \mathbf{x}(t_f) \right\| + p_2 \sum_{i-1}^{4} \int_0^{t_f} \vartheta(\varphi_i(\mathbf{x}))dt \to \min_{\mathbf{u} \in \mathrm{U}}, \tag{31}$$

где $p_1 = 1$, $p_2 = 3$, $\vartheta(\alpha)$ — функция Хэвисайда,

$$\vartheta(\alpha) = \begin{cases} 1, \text{ если } \alpha > 0; \\ 0 \text{ иначе.} \end{cases} \tag{32}$$

Для решения задачи используем прямой подход и эволюционные алгоритмы [13, 14], которые, как показали исследования [15, 16], работают для задач оптимального управления с фазовыми ограничениями лучше, чем градиентные и другие алгоритмы. В частности, здесь использован гибридный эволюционный алгоритм [17], который на этапе эволюции каждого возможного решения использует случайно отобранную эволюцию из алгоритмов: генетического [18], «роя-частиц» [19] или оптимизатора «серого волка» [20].

Для решения задачи использован прямой подход, для чего установлен интервал времени $\Delta t = 0{,}4$. Таким образом, ось времени разбита на $K = \left\lfloor t^+ / \Delta t \right\rfloor = \left\lfloor 6{,}4 / 0{,}4 \right\rfloor = 16$ интервалов, на границах которых необходимо найти $p = m(K+1) = 4 \cdot 17 = 68$ значений параметров. Гибридный эволюционный алгоритм имел следующие значения параметров: число возможных решений в начальной популяции 512; число поколений 512; число возможных эволюций в одним поколении 64; число бит для кода Грея в генетическом алгоритме 16, из которых 4 бита отводилось для целой части числа; число лидирующих решений при эволюции алгоритма оптимизатора «серого волка» 3; параметры эволюции алгоритма «роя-частиц» $k_\alpha = 0{,}729$, $k_\beta = 0{,}15$, $k_\gamma = 0{,}85$, $k_\sigma = 1$, число информаторов или случайно отобранных возможных решения 5. Обычно в конечномерных задачах глобальной оптимизации коэффициент при эволюции за счет лучшего текущего возможного решения равен $k_\beta = 0{,}85$, а за счет лучшего среди случайно отобранных — $k_\gamma = 0{,}15$. Однако, как установлено в [15], для задачи оптимального управления или задачи глобальной бесконечномерной оптимизации значения этих коэффициентов необходимо поменять местами.

В результате вычислений гибридный эволюционный алгоритм нашел два различных по форме траекторий решения $\left(\mathbf{y}^1 = \left[\, y_1^1 \ \ldots \ y_{68}^1 \, \right]^{\mathrm{T}} ; \right.$ $\left. \mathbf{y}^2 = \left[\, y_1^2 \ \ldots \ y_{68}^2 \, \right]^{\mathrm{T}} \right)$, которые давали приблизительно одинаковые значения критерия качества (31). Значения найденных векторов приведены

в табл. 1. Эти значения в дальнейшем использованы в качестве обучающей выборки для машинного обучения управления методом символьной регрессии для синтеза универсальной системы стабилизации.

Таблица 1

Параметры функций управлений из обучающей выборки

$\mathbf{y}^1$				$\mathbf{y}^2$			
i	y_i^1	i	y_i^1	i	y_i^2	i	y_i^2
1	–2,638	35	–8,400	1	–0,050	35	–11,345
2	12,616	36	9,862	2	6,493	36	18,995
3	–7,934	37	–9,124	3	–0,000	37	–6,779
4	8,819	38	–7,377	4	–3,644	38	1,749
5	–7,518	39	–0	5	10,727	39	7,731
6	–17,134	40	1,504	6	–13,998	40	9,599
7	19,856	41	10,160	7	–2,832	41	15,313
8	19,921	42	0,265	8	17,577	42	–17,900
9	–8,278	43	–6,8660	9	–17,199	43	–0,450
10	–12,601	44	9,599	10	0,239	44	0,244
11	–8,901	45	–9,447	11	–18,909	45	17,544
12	13,904	46	8,108	12	18,734	46	–10,717
13	9,068	47	15,141	13	0,881	47	–17,910
14	11,869	48	12,242	14	18,862	48	14,167
15	–14,095	49	–1,813	15	–10,055	49	8,477
16	17,657	50	3,233	16	8,066	50	–10,685
17	–11,037	51	1,173	17	–8,356	51	–5,769
18	–8,927	52	18,141	18	18,247	52	14,040
19	–9,413	53	–4,408	19	–1,344	53	–0,253
20	16,046	54	4,166	20	19,923	54	2,405
21	–0,740	55	2,950	21	–11,033	55	–0,003
22	3,643	56	15,591	22	6,399	56	10,922
23	11,488	57	–1,632	23	18,957	57	4,917
24	1,278	58	0,383	24	13,202	58	–14,223
25	5,061	59	1,425	25	–13,661	59	14,670
26	11,613	60	8,649	26	14,151	60	6,251
27	–3,064	61	–5,364	27	7,418	61	1,007
28	17,310	62	1,028	28	17,614	62	–0,595
29	17,623	63	0,490	29	1,228	63	0,020
30	17,514	64	14,119	30	–1,470	64	19,693
31	–11,344	65	–4,730	31	–1,647	65	0,013
32	15,340	66	0,667	32	12,002	66	–0,014
33	–0,652	67	2,895	33	15,594	67	–5,226
34	–2,532	68	15,594	34	8,258	68	13,448

Далее решалась задача синтеза универсальной системы стабилизации в форме (9). Для этого использован метод сетевого оператора [12]. Для решения задачи необходимо определить одну функцию управления в форме (9), чтобы стабилизировать движение объекта управления по пространственным траекториям, порожденным найденным управлением (см. табл. 1). При решении задачи синтеза начальное состояние объекта управления (27) заменено множеством начальных состояний (19), где $\tilde{n}=3$, $\mathbf{\Delta}_0=[0{,}2\ \ 0{,}2\ \ 0{,}2\ \ 0\ \ 0\ \ 0]^{\mathrm{T}}$.

В результате для задачи синтеза универсальной системы стабилизации использован критерий оптимизации (18) с параметрами $B=\tilde{n}^3=27$, $D=2$. В качестве метода символьной регрессии выбран метод сетевого оператора, который нашел следующее решение:

$$u_i=h_i(\mathbf{x}^*-\mathbf{x})=\begin{cases}u_i^+, \text{ если } \tilde{h}_i(\mathbf{x}^*-\mathbf{x})\ge u_i^+;\\ u_i^-, \text{ если } \tilde{h}_i(\mathbf{x}^*-\mathbf{x})\le u_i^-; \quad i=1,\ldots,m=4,\\ \tilde{\tilde{h}}_i(\mathbf{x}^*-\mathbf{x}) \text{ иначе},\end{cases} \tag{33}$$

где

$$\tilde{h}_1(\mathbf{x}^*-\mathbf{x})=\mu(E), \tag{34}$$

$$\tilde{h}_2(\mathbf{x}^*-\mathbf{x})=\mu(E)-\mu^3(E), \tag{35}$$

$$\tilde{h}_3(\mathbf{x}^*-\mathbf{x})=\operatorname{sgn}(V-\mu(E))+\cos G+\rho_{17}(H), \tag{36}$$

$$\tilde{h}_4\left(\mathbf{x}^*-\mathbf{x}\right)=\exp\left(\tilde{h}_3(\mathbf{x}^*-\mathbf{x})\right)+\ln\left(|\tilde{h}_3(\mathbf{x}^*-\mathbf{x})|\right)+$$

$$+\operatorname{sgn}(U)\sqrt{|U|}+\exp\left(d_2\left(x_2^*-x_2\right)\right)+$$

$$+\ \rho_{19}(V)+\operatorname{arctg}(B+\operatorname{tg}C)+\operatorname{sgn}C+\operatorname{arctg}D+\rho_{18}\left(d_5\left(x_5^*-x_5\right)\right),$$

$$U=V+\mu(E),\ \ V=W+\operatorname{tg}B+\exp E+d_5-d_6^3,\ \ W=A+\operatorname{sgn}(B)\sqrt{|B|},$$

$$A=\exp(B+\operatorname{tg}C)+\cos\left(d_6\left(x_6^*-x_6\right)\right),$$

$$B=C+\sqrt[3]{D}+\sin H,\ \ C=-DEG\operatorname{arctg}E,$$

$$D=E+\sqrt[3]{G}+\operatorname{sgn}\left(x_5^*-x_5\right)+\left(x_2^*-x_2\right)^3,$$

$$E=\rho_{17}(F)+G^3+H+\vartheta\left(d_6\left(x_6^*-x_6\right)\right)+\left(x_5^*-x_5\right)^2,$$

$$F = d_6\left(x_6^* - x_6\right) + d_3\left(x_3^* - x_3\right),$$

$$G = \sin\left(d_6\left(x_6^* - x_6\right)\right) + d_5\left(x_5^* - x_5\right) + d_2\left(x_2^* - x_2\right) + \cos d_1 + \vartheta\left(x_2^* - x_2\right),$$

$$H = d_4\left(x_4^* - x_4\right) + \operatorname{sgn}\left(d_5\left(x_5^* - x_5\right)\right)\sqrt{d_5\left(x_5^* - x_5\right)} +$$

$$+ d_1\left(x_1^* - x_1\right) + \sin d_5 + \operatorname{sgn}\left(x_3^* - x_3\right),$$

$$\mu(\alpha) = \begin{cases} \alpha, & \text{если } |\alpha| < 1; \\ \operatorname{sgn}\alpha & \text{иначе,} \end{cases}$$

$$\rho_{17}(\alpha) = \operatorname{sgn}\alpha \ln(|\alpha| + 1), \quad \rho_{18}(\alpha) = \operatorname{sgn}\alpha(\exp(|\alpha|) - 1),$$

$$\rho_{19}(\alpha) = \operatorname{sgn}\alpha \exp(-|\alpha|),$$

$$d_1 = 6{,}07617, \quad d_2 = 7{,}98145, \quad d_3 = 13{,}84058, \quad d_4 = 15{,}89258,$$

$$d_5 = 4{,}36548, \quad d_6 = 3{,}35059.$$

Оптимальные траектории (линия синего цвета), порожденные управлением из табл. 1, и траектории (линия черного цвета) движения объекта управления с синтезированной универсальной системой стабилизации управления (33)–(36) приведены на рис. 1. Здесь и далее окружности красного цвета указывают фазовые ограничения, в частности на рис. 1 ограничения (30).

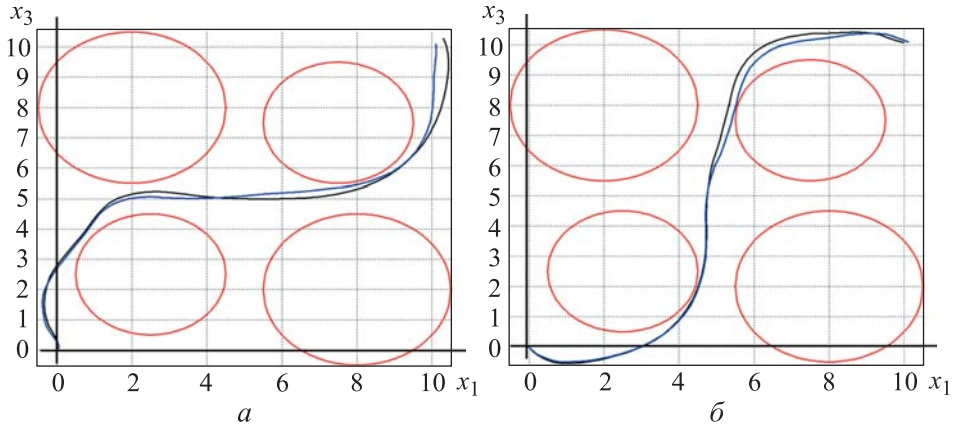


Рис. 1. Оптимальная и реальная траектории движения объекта для управлений $\mathbf{y}^1$ (*а*) и $\mathbf{y}^2$ (*б*)

Оптимальные (линия синего цвета) и реальные (линии черного цвета) траектории движения объекта управления из восьми начальных условий из множества (19) приведены на рис. 2.

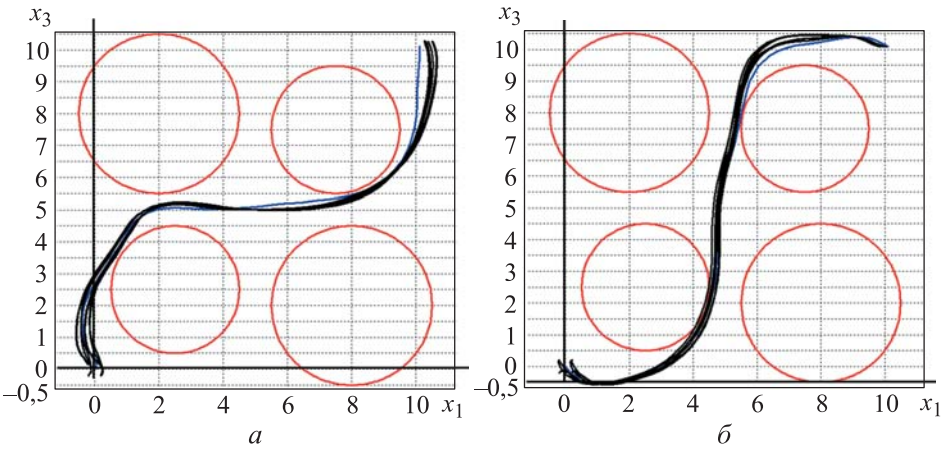


Рис. 2. Оптимальная и реальные траектории движения объекта из восьми начальных состояний для управлений $\mathbf{y}^1$ (*а*) и $\mathbf{y}^2$ (*б*)

Далее для проверки качества обучения универсальной системы стабилизации рассмотрена задача оптимального управления (20)–(29), (31) со следующими фазовыми ограничениями:

$$\varphi_i(\mathbf{x}) = r_i - \sqrt{(x_{1,i} - x_1)^2 + (x_{3,i} - x_3)^2} \le 0, \;\; i = 1, \ldots, M = 4,$$

где $r_1 = 1$; $x_{1,1} = 2$; $x_{3,1} = 2$; $r_2 = 2{,}5$; $x_{1,2} = 2$; $x_{3,2} = 8$; $r_3 = 2{,}5$; $x_{1,3} = 8$; $x_{3,3} = 2$; $r_4 = 1$; $x_{1,4} = 5$; $x_{3,4} = 5$; $r_5 = 1$; $x_{1,5} = 8$; $x_{3,5} = 8$.

Для решения задачи оптимального управления использованы прямой подход и гибридный эволюционный алгоритм. Найденное решение представлено в табл. 2.

Таблица 2

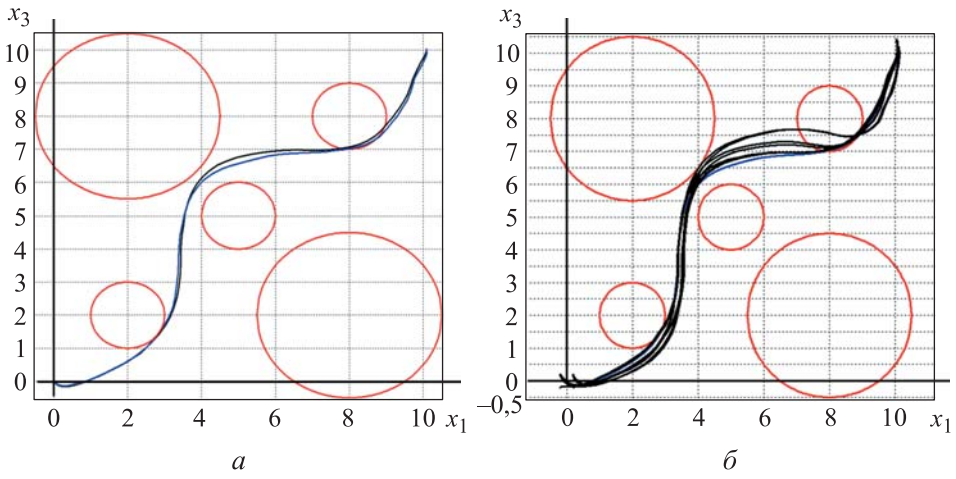
Параметры оптимальной функции управления $\mathbf{y}^3$

i	y_i^3	i	y_i^3	i	y_i^3
1	4,299	24	18,049	47	0,119
2	11,162	25	–0,201	48	8,167
3	–6,456	26	15,400	49	–10,873
4	11,689	27	–0,014	50	12,562
5	–3,611	28	13,402	51	9,280
6	18,718	29	11,516	52	13,718
7	–9,944	30	–6,657	53	0,065
8	19,922	31	–0,775	54	–0,057
9	–3,813	32	19,153	55	–8,474
10	–18,614	33	14,908	56	6,139
11	–17,651	34	–13,264	57	0,021
12	9,031	35	–15,456	58	–0,572

Окончание табл. 2

i	y_i^3	i	y_i^3	i	y_i^3
13	–16,209	36	19,012	59	–0,002
14	3,166	37	–2,018	60	14,148
15	17,663	38	–2,586	61	5,778
16	4,652	39	–0,348	62	5,034
17	0,071	40	7,468	63	3,201
18	3,918	41	–2,931	64	10,048
19	1,694	42	0,161	65	–7,487
20	12,604	43	–9,679	66	–0,892
21	–2,681	44	7,707	67	7,798
22	–13,642	45	1,394	68	17,551
23	0,857	46	2,068		

Оптимальная (линия синего цвета) и реальная (линия черного цвета) траектории движения объекта управления при той же системе стабилизации движения, что и для ранее найденных оптимальных решений, приведена на рис. 3, *а*.



а *б*

Рис. 3. Оптимальная и реальная траектории движения объекта (*а*), а также реальные траектории движения объекта из восьми начальных состояний (*б*) для управления $\mathbf{y}^3$

Оптимальная и реальные траектории движения объекта управления из восьми начальных условий из множества (19) приведены на рис. 3, *б*. Полученная универсальная система стабилизации (33)–(36) обеспечивает стабилизацию движения объекта управления вдоль заданной траектории из некоторого класса. Малые возмущения начальных условий не приводят к существенному изменению характера траектории движения.

С учетом изложенного расширенная математическая модель объекта управления (10), (11) с подсистемой эталонной модели (10) для генерации траектории и подсистемой (11) модели объекта управления универсальной системой стабилизации (9) позволяет решать задачу оптимального управления в классической постановке и получать решения, реализуемые в реальном объекте управления.

Заключение. Рассмотрен подход к решению расширенной задачи оптимального управления, в которой необходимо получить решение в классе реализуемых практически функций управления. Предложено для решения задачи оптимального управления использовать расширенную математическую модель объекта управления. Эта модель включает в себя эталонную модель объекта управления, решение которой с управлением в виде функции времени позволяет получить траекторию движения объекта в пространстве состояний, и модель объекта управления с универсальной системой стабилизации движения в окрестности заданной траектории. Если универсальная система стабилизации движения синтезирована ранее и обеспечивает стабилизацию движения объекта вдоль любой заданной траектории из некоторого класса, то решение задачи оптимального управления возможно в классической постановке, т. е. нахождение управления в виде функции времени, при этом полученное решение является реализуемым в реальном объекте. Дополнительное преимущество такого подхода состоит в том, что задачу оптимального управления для объекта с системой стабилизации движения можно решать в режиме реального времени на бортовом компьютере объекта управления.

ЛИТЕРАТУРА

[1] Mitchell T.M. Machine learning. McGraw-Hill, 1997.

[2] Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В. и др. Математическая теория оптимальных процессов. М., Наука, 1983.

[3] Shen X., Shi W. Adaptive trajectory tracking control of wheeled mobile robot. *Proc. CCDC*, 2019, pp. 5161–5165. DOI: https://doi.org/10.1109/CCDC.2019.8833019

[4] Samir A., Hammad A., Hafez A., et al. Quadcopter trajectory tracking control using state-feedback control with integral action. *Int. J. Comput. Appl.*, 2017, vol. 168, no. 9, pp. 1–7.

[5] Uddin N. A robot trajectory tracking control system design using pole domination approach. *IEEE IEMCON*, pp. 506–512.
DOI: https://doi.org/10.1109/IEMCON.2018.8614899

[6] Andreev A., Peregudova O., Sutyrkina K., et al. On global output feedback trajectory tracking control of a wheeled mobile robot. *Proc. ICMT*, *2019*.
DOI: https://doi.org/10.1109/ICMECT.2019.8932122

[7] Su T., Liang X., He G., et al. Robust trajectory tracking of delta parallel robot using sliding mode control. *IEEE SSCI*, pp. 508–512.
DOI: https://doi.org/10.1109/SSCI44817.2019.9003125

[8] Колесников А.А., Колесников Ал.А., Кузьменко А.А. Методы АКАР и АКОР в задачах синтеза нелинейных систем управления. *Мехатроника, автоматизация, управление*, 2016, т. 17, № 10, с. 657–669.
DOI: https://doi.org/10.17587/mau.17.657-669

[9] Колесников А.А., Колесников Ал.А., Кузьменко А.А. Методы АКАР и бэкстеппинг в задачах синтеза нелинейных систем управления. *Мехатроника, автоматизация, управление*, 2016, т. 17, № 7, с. 435–445.
DOI: https://doi.org/10.17587/mau.17.435-445

[10] Nguyen A.T., Xuan-Mung N., Hong S-K. Quadcopter adaptive trajectory tracking control: a new approach *via* backstepping technique. *Appl. Sci.*, 2019, vol. 9, iss. 18, art. 3873. DOI: https://doi.org/10.3390/app9183873

[11] Diveev A.I., Sofronova E.A. Universal stabilisation system for control object motion along the optimal trajectory. *Mathematics*, 2023, vol. 11, iss. 16, art. 3556.
DOI: https://doi.org/10.3390/math11163556

[12] Diveev A.I., Shmalko E.Yu. Machine learning control by symbolic regression. Cham, Springer, 2021. DOI: https://doi.org/10.1007/978-3-030-83213-1

[13] Карпенко А.П. Современные алгоритмы поисковой оптимизации. Алгоритмы, вдохновленные природой. М., Изд-во МГТУ им. Н.Э. Баумана, 2014.

[14] Саймон Д. Алгоритмы эволюционной оптимизации. М., ДМК Пресс, 2020.

[15] Дивеев А.И., Константинов С.В. Исследования практической сходимости эволюционных алгоритмов оптимального программного управления колесным роботом. *Известия РАН. Теория и системы управления*, 2018, № 4, с. 80–106.
DOI: https://doi.org/10.31857/S000233880002513-3

[16] Diveev A.I., Sofronova E.A., Konstantinov S.V. Approaches to numerical solution of optimal control problem using evolutionary computations. *App. Sci.*, 2021, vol. 11, iss. 15, art. 7096. DOI: http://dx.doi.org/10.3390/app11157096

[17] Diveev A. Hybrid evolutionary algorithm for optimal control problem. In: Arai K. (ed). *Intelligent Systems and Applications. IntelliSys 2022. Lecture Notes in Networks and Systems*, vol. 543. Cham, Springer, 2022, pp. 726–738.
DOI: https://doi.org/10.1007/978-3-031-16078-3_50

[18] Davis L. Handbook of genetic algorithms. Van Nostrand Reinhold, 1991.

[19] Kennedy J., Eberhart R. Particle swarm optimization. *IEEE ICNN*, 1995, vol. 4, pp. 1942–1948. DOI: https://doi.org/10.1109/ICNN.1995.488968

[20] Mirjalili S., Mirjalili S.M., Lewis A. Grey wolf optimizer. *Adv. Eng. Softw.*, 2014, vol. 69, pp. 46–61. DOI: https://doi.org/10.1016/j.advengsoft.2013.12.007

Дивеев Асхат Ибрагимович — д-р техн. наук, профессор, главный научный сотрудник, начальник отдела «Управление робототехническими устройствами» ФИЦ ИУ РАН (Российская Федерация, 119333, Москва, ул. Вавилова, д. 44, корп. 2).

Просьба ссылаться на эту статью следующим образом:

Дивеев А.И. Машинное обучение универсальной системы стабилизации движения в расширенной задаче оптимального управления. *Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение*, 2025, № 1 (150), с. 71–90. EDN: XKJFLR

MACHINE LEARNING OF A UNIVERSAL MOTION STABILIZATION SYSTEM IN THE ADVANCED OPTIMAL CONTROL PROBLEM

A.I. Diveev aidiveev@mail.ru

FRC CSC RAS, Moscow, Russian Federation

Abstract

This paper discusses the advanced task of optimal control. In this task, it is necessary not only to find a control function as a function of time to achieve a control goal with an optimal value of a given quality criterion, but also to ensure the existence of an attraction property of an optimal trajectory in some of its vicinity. This property allows you to find a control function that can be implemented directly in a real object. To solve the problem, a universal system for stabilizing the movement of a control object along a given trajectory is first synthesized. Here machine learning of control by symbolic regression is used. Symbolic regression allows you to find the structure and parameters of the control function without human participation. To ensure the versatility of the synthesized stabilization system, symbol regression looks for a single stabilization system for some given of different trajectories. The synthesized stabilization system and the initial mathematical model of the control object are entered into the object control system. The obtained mathematical model of the control object allows solving an extended problem of optimal control and obtaining a control function implemented directly in a real object. Presented is a method for solving an extended problem of optimal control of spatial motion of a quadcopter

Keywords

Optimal control, symbolic regression, implemented control function, stabilization system, control synthesis

Received 01.04.2024
Accepted 22.04.2024
© Author(s), 2025

The work was partly supported by the Russian Science Foundation (project no. 23-29-00339)

REFERENCES

[1] Mitchell T.M. Machine learning. McGraw-Hill, 1997.

[2] Pontryagin L.S., Boltyanskiy V.G., Gamkrelidze R.V., et al. Matematicheskaya teoriya optimalnykh protsessov [Mathematical theory of optimal processes]. Moscow, Nauka Publ., 1983.

[3] Shen X., Shi W. Adaptive trajectory tracking control of wheeled mobile robot. *Proc. CCDC*, 2019, pp. 5161–5165. DOI: https://doi.org/10.1109/CCDC.2019.8833019

[4] Samir A., Hammad A., Hafez A., et al. Quadcopter trajectory tracking control using state-feedback control with integral action. *Int. J. Comput. Appl*., 2017, vol. 168, no. 9, pp. 1–7.

[5] Uddin N. A robot trajectory tracking control system design using pole domination approach. *IEEE IEMCON*, pp. 506–512.
DOI: https://doi.org/10.1109/IEMCON.2018.8614899

[6] Andreev A., Peregudova O., Sutyrkina K., et al. On global output feedback trajectory tracking control of a wheeled mobile robot. *Proc. ICMT*, *2019*.
DOI: https://doi.org/10.1109/ICMECT.2019.8932122

[7] Su T., Liang X., He G., et al. Robust trajectory tracking of delta parallel robot using sliding mode control. *IEEE SSCI*, pp. 508–512.
DOI: https://doi.org/10.1109/SSCI44817.2019.9003125

[8] Kolesnikov A.A., Kolesnikov Al.A., Kuzmenko A.A. The ADAR method and theory of optimal control in the problems of synthesis of nonlinear control systems. *Mekhatronika, avtomatizatsiya, upravlenie*, 2016, vol. 17, no. 10, pp. 657–669 (in Russ.).
DOI: https://doi.org/10.17587/mau.17.657-669

[9] Kolesnikov A.A., Kolesnikov Al.A., Kuzmenko A.A. Backstepping and ADAR method in the problems of synthesis of the nonlinear control systems. *Mekhatronika, avtomatizatsiya, upravlenie*, 2016, vol. 17, no. 7, pp. 435–445 (in Russ.).
DOI: https://doi.org/10.17587/mau.17.435-445

[10] Nguyen A.T., Xuan-Mung N., Hong S-K. Quadcopter adaptive trajectory tracking control: a new approach *via* backstepping technique. *Appl. Sci*., 2019, vol. 9, iss. 18, art. 3873. DOI: https://doi.org/10.3390/app9183873

[11] Diveev A.I., Sofronova E.A. Universal stabilisation system for control object motion along the optimal trajectory. *Mathematics*, 2023, vol. 11, iss. 16, art. 3556.
DOI: https://doi.org/10.3390/math11163556

[12] Diveev A.I., Shmalko E.Yu. Machine learning control by symbolic regression. Cham, Springer, 2021. DOI: https://doi.org/10.1007/978-3-030-83213-1

[13] Karpenko A.P. Sovremennye algoritmy poiskovoy optimizatsii. Algoritmy, vdokhnovlennye prirodoy [Modern search engine optimization algorithms. Algorithms inspired by nature]. Moscow, BMSTU Publ., 2014.

[14] Simon D. Evolutionary optimization algorithms. Wiley, 2013.

[15] Diveev A.I., Konstantinov S.V. Study of the practical convergence of evolutionary algorithms for the optimal program control of a wheeled robot. *J. Comput. Syst. Sci. Int.*, 2018, vol. 57, no. 4, pp. 561–580. DOI: https://doi.org/10.1134/S106423071804007X

[16] Diveev A.I., Sofronova E.A., Konstantinov S.V. Approaches to numerical solution of optimal control problem using evolutionary computations. *App. Sci.*, 2021, vol. 11, iss. 15, art. 7096. DOI: http://dx.doi.org/10.3390/app11157096

[17] Diveev A. Hybrid evolutionary algorithm for optimal control problem. In: Arai K. (ed). *Intelligent Systems and Applications. IntelliSys 2022. Lecture Notes in Networks and Systems*, vol. 543. Cham, Springer, 2022, pp. 726–738.
DOI: https://doi.org/10.1007/978-3-031-16078-3_50

[18] Davis L. Handbook of genetic algorithms. Van Nostrand Reinhold, 1991.

[19] Kennedy J., Eberhart R. Particle swarm optimization. *IEEE ICNN*, 1995, vol. 4, pp. 1942–1948. DOI: https://doi.org/10.1109/ICNN.1995.488968

[20] Mirjalili S., Mirjalili S.M., Lewis A. Grey wolf optimizer. *Adv. Eng. Softw.*, 2014, vol. 69, pp. 46–61. DOI: https://doi.org/10.1016/j.advengsoft.2013.12.007

Diveev A.I. — Dr. Sc. (Eng.), Professor, Chief Research Fellow, Head of the Department of Robotics Control, FRC CSC RAS (Vavilova ul. 44, korp. 2, Moscow, 119333 Russian Federation).

Please cite this article in English as:

Diveev A.I. Machine learning of a universal motion stabilization system in the advanced optimal control problem. *Herald of the Bauman Moscow State Technical University, Series Instrument Engineering*, 2025, no. 1 (150), pp. 71–90 (in Russ.).
EDN: XKJFLR