

Проблемы
цифровой
реальности

Проектирование
будущего

А.И. Дивеев, Е.А. Софронова,
Е.Ю. Шмалько

**Метод синтезированного
оптимального управления групповым
взаимодействием роботов**

Рекомендуемая форма библиографической ссылки

Дивеев А.И. Софронова Е.А. Шмалько Е.Ю. Метод синтезированного оптимального управления групповым взаимодействием роботов // Проектирование будущего. Проблемы цифровой реальности: труды 3-й Международной конференции (6-7 февраля 2020 г., Москва). — М.: ИПМ им. М.В.Келдыша, 2020. — С. 166-175. — <https://keldysh.ru/future/2020/14.pdf>
<https://doi.org/10.20948/future-2020-14>

Размещено также [видео выступления](#)

Метод синтезированного оптимального управления групповым взаимодействием роботов

А.И. Дивеев, Е.А. Софронова, Е.Ю. Шмалько

ФИЦ «Информатика и управление» РАН

Аннотация. Рассмотрен универсальный подход к разработке робастных оптимальных систем управления. Подход состоит в построении системы стабилизации и обеспечении свойства устойчивости объекта относительно точки в пространстве состояний. Оптимальное управление объектами осуществляется за счет нахождения оптимального расположения точек стабилизации для каждого объекта и их переключения через заданный интервал времени. Показано, что поведение объекта в окрестности устойчивой точки малочувствительно к неточностям математической модели и начальным условиям. Приведен численный пример использования данного подхода к решению задачи оптимального управления группой роботов при наличии фазовых ограничений.

Ключевые слова: синтез системы управления, оптимальное управление, группа роботов

Method of synthesized optimal control for the group of robots

A.I. Diveev, E.A. Sofronova, E.Yu. Shmalko

RAS Federal Research Center "Computer Science and Control"

Abstract. A universal approach to the development of robust optimal control systems is considered. The approach consists in constructing a stabilization system and ensuring the property of stability of an object relative to a point in the state space. Optimal control of objects is carried out by finding the optimal location of the stabilization points for each object and switching them over a given time interval. It is shown that the behavior of an object near the stabilization points is robust to inaccuracies of mathematical model and the initial conditions. A numerical example of using this approach to solving the optimal control problem of a group of robots in the presence of phase constraints is given.

Keywords: control system synthesis, optimal control, group of robots

Введение

Математические модели робототехнических устройств в большинстве случаев описываются системой обыкновенных дифференциальных уравнений, в правых частях которых расположен свободный вектор управления. Любое решение задачи управления состоит в таком определении этого вектора управления, чтобы решение системы дифференциальных уравнений обладало определенными свойствами. Решение задачи оптимального управления в классической постановке [1] приводит к получению управления как функции времени. Подстановка найденной функции управления в правую часть дифференциальных уравнений приводит к тому, что дифференциальные уравнения имеют частное решение, которое из заданного начального состояния достигает заданное терминальное состояние с оптимальным значением критерия качества.

Система дифференциальных уравнений с оптимальной функцией управления не обладает свойствами робастности. При любом малом изменении системы или начальных условий изменяется частное решение, которое уже не будет оптимальным и не достигнет заданного терминального состояния. Чтобы оптимальное решение обладало свойством робастности, необходимо, чтобы это решение стало притягивающим аттрактором как одномерное многообразие в пространстве состояний. Для этой цели на практике предлагается вместе с оптимальным управлением конструировать систему стабилизации оптимальной траектории. Здесь необходимо учитывать, что после решения задачи оптимального управления в правой части системы дифференциальных уравнений не остается свободного вектора управления, так как он заменяется на найденную функцию времени. Для конструирования системы стабилизации оптимальной траектории первоначально полученное оптимальное управление в виде функции времени подставляют в правые части системы и получают оптимальные траектории каждой компоненты вектора пространства состояний во времени. Затем фиксируют моменты времени и определяют множество точек в пространстве состояний. Данное множество описывает оптимальную траекторию в пространстве состояний. Для стабилизации этой траектории затем строят систему управления. В данном подходе существует определенный нюанс, который заключается в том, что математическая модель объекта управления с системой стабилизации будет отличаться от математической модели объекта, для которой решалась задача оптимального управления, причем, если система стабилизации включает пропорционально-интегральные дифференцирующие регуляторы, то отличаться будут даже размерности

моделей, а это значит, что в задаче оптимального управления будет другой гамильтониан, другая система уравнений для сопряженных переменных и, следовательно, стабилизируемая траектория уже не будет оптимальной для модели объекта с системой стабилизации. Приблизительно те же проблемы возникают и при использовании метода аналитического конструирования агрегированных регуляторов [2], только в данном случае оптимальную траекторию необходимо представить в виде одномерного многообразия.

В работе рассматривается другой подход к созданию робастных систем управления. Первоначально решается задача синтеза системы стабилизации и в результате обеспечивается стабилизация системы относительно некоторой точки в пространстве состояний. На втором этапе решается задача оптимального расположения точек стабилизации, которые переключаются через заданный интервал времени. В результате объект достигает терминального состояния с оптимальным значением критерия качества. Данный подход мы назвали методом синтезированного оптимального управления. В результате применения этого подхода мы не получаем оптимальную траекторию со свойствами аттрактора, но в каждый момент времени мы имеем систему, в фазовом пространстве которой имеется аттрактор в виде устойчивой точки равновесия, следовательно, система является робастной.

1. Задача оптимального робастного управления

Рассмотрим формальную постановку задачи оптимального робастного управления

Задана модель объекта управления

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

где \mathbf{x} – вектор состояния, \mathbf{u} – вектор управления, $\mathbf{x} \in \mathbf{R}^n$, $\mathbf{u} \in U \subseteq \mathbf{R}^m$, U – компактное множество, $m \leq n$.

Для системы заданы начальные и терминальные условия

$$\mathbf{x}^0, \mathbf{x}^f \in \mathbf{R}^n. \quad (2)$$

Задан критерий качества управления

$$J = \int_0^{t_f} f_0(\mathbf{x}(t), \mathbf{u}(t)) dt \rightarrow \min, \quad (3)$$

где t_f – ограниченное, но не заданное время достижения терминального состояния

$$t_f = \begin{cases} t & t < t^+, \|\mathbf{x}^f - \mathbf{x}(t)\| \leq \varepsilon, \\ t^+ & \text{иначе} \end{cases}, \quad (4)$$

где t^+ и ε – заданные положительные величины.

Необходимо найти управление

$$\mathbf{u} = \mathbf{h}(\mathbf{x}, t) \in U, \quad (5)$$

чтобы система

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{h}(\mathbf{x}, t)), \quad (6)$$

имела частное решение $\mathbf{x}(\mathbf{x}^0, t)$, которое из начального состояния \mathbf{x}^0 достигло терминального состояния

$$\|\mathbf{x}^f - \mathbf{x}(\mathbf{x}^0, t_f)\| = \Delta(\mathbf{h}(\mathbf{x}, t)) \leq \varepsilon, \quad (7)$$

с оптимальным значением критерия качества (3).

$$J = \int_0^{t_f} f_0(\mathbf{x}(\mathbf{x}^0, t), \mathbf{h}(\mathbf{x}(\mathbf{x}^0, t), t)) dt = J(\mathbf{h}(\mathbf{x}, t)), \quad (8)$$

При аддитивных возмущениях модели и начальных условий

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) + \boldsymbol{\xi}(t), \quad \mathbf{x}^0 + \boldsymbol{\xi}^0 \in \mathbf{R}^n, \quad (9)$$

где $\boldsymbol{\xi}(t) = [\xi_1 \ \dots \ \xi_n]^T$, $\boldsymbol{\xi}^0 = [\xi_1^0 \ \dots \ \xi_n^0]^T$ – векторы случайных величин, показатели качества управления существенно не меняются

$$\Delta(\mathbf{h}(\mathbf{x}, t), \boldsymbol{\xi}(t), \boldsymbol{\xi}^0) - \Delta(\mathbf{h}(\mathbf{x}, t)) < \alpha, \quad (10)$$

$$J(\mathbf{h}(\mathbf{x}, t), \boldsymbol{\xi}(t), \boldsymbol{\xi}^0) - J(\mathbf{h}(\mathbf{x}, t)) < \beta, \quad (11)$$

$$\text{если } \|\boldsymbol{\xi}(t)\| < \Omega, \quad \|\boldsymbol{\xi}^0\| < \Omega, \quad 0 \leq t \leq t_f, \quad (12)$$

где $\Delta(\mathbf{h}(\mathbf{x}, t), \boldsymbol{\xi}(t), \boldsymbol{\xi}^0)$, $J(\mathbf{h}(\mathbf{x}, t), \boldsymbol{\xi}(t), \boldsymbol{\xi}^0)$ показатели качества управления возмущенной системы, α , β , Ω – заданные положительные величины.

1. Классический подход к решению

Первоначально решаем задачу оптимального управления и находим управление как функцию времени

$$\mathbf{u} = \mathbf{v}(t). \quad (13)$$

Для решения задачи используем либо прямые методы сведения задачи оптимального управления к задаче нелинейного программирования, либо принцип максимума Л.С. Понтрягина [1].

Моделируем систему (1) с оптимальным управлением (13) и получаем оптимальную траекторию в пространстве состояний $\mathbf{x}(\mathbf{x}^0, \mathbf{v}(t))$.

Далее строим систему стабилизации

$$\mathbf{u} = \mathbf{h}(\mathbf{x}(\mathbf{x}^0, \mathbf{v}(t)) - \mathbf{x}, \mathbf{q}), \quad (14)$$

где \mathbf{q} – вектор постоянных параметров системы стабилизации, который определяем по условию удовлетворения неравенств (10)–(12).

К недостатку метода следует отнести этап построения системы стабилизации траектории. К тому же после реализации системы управления ее довольно сложно перестроить для решения другой задачи оптимального управления, например, с другой терминальной точкой или другим критерием качества.

2. Метод аналитического конструирования агрегированных регуляторов

На первом этапе решаем также задачу оптимального управления (13).

Полученную оптимальную траекторию $\mathbf{x}(\mathbf{x}^0, \mathbf{v}(t))$ представляем в форме одномерного многообразия.

$$\psi_i(\mathbf{x}) = 0, \quad i = 1, \dots, n-1. \quad (15)$$

Записываем устойчивое дифференциальное уравнение относительно агрегированных переменных [2] $\psi_i, i = 1, \dots, n-1$.

$$T\dot{\psi}(\mathbf{x}) + \mathbf{g}(\psi(\mathbf{x})) = 0, \quad (16)$$

где $\psi(\mathbf{x}) = [\psi_1(\mathbf{x}) \ \dots \ \psi_{n-1}(\mathbf{x})]^T$, T – постоянная величина, $\mathbf{g}(\mathbf{y})$ – некоторая векторная функция, обладающая свойством $\mathbf{g}(\mathbf{0}) = \mathbf{0}$.

Раскрываем (16), получаем

$$T \frac{\partial \psi(\mathbf{x})}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}, \mathbf{u}) + \mathbf{g}(\psi(\mathbf{x})) = 0, \quad (17)$$

Разрешаем это уравнение относительно вектора управления при этом подбираем функцию $\mathbf{g}(\psi(\mathbf{x}))$, чтобы функция Ляпунова системы (16) сохраняла знакопостоянство. В результате получаем управление с обратной связью

$$\mathbf{u} = \mathbf{w}(\mathbf{x}), \quad (18)$$

которая согласно теории синергетического управления, должна обеспечить свойство аттрактора многообразием (15).

К недостаткам метода следует отнести довольно большое количество строго неформализованных процедур, которые зависят от модели объекта управления, от построения многообразия (15) до решения обратной задачи (17), к тому же при аналитическом решении обратной задачи не учитываются ограничения на управление.

Решение, полученное данным методом, также трудно перестроить для решения другой задачи оптимального управления, с другой терминальной точкой или другим критерием качества.

3. Метод синтезированного оптимального управления

На первом этапе решаем задачу синтеза системы стабилизации

$$\mathbf{u} = \mathbf{h}(\mathbf{x}, \mathbf{x}^*), \quad (18)$$

где \mathbf{x}^* – вектор постоянных параметров размерности n , при этом функция $\mathbf{h}(\mathbf{x}, \mathbf{x}^*)$ и параметры \mathbf{x}^* подбираются таким образом, что система дифференциальных уравнений

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{h}(\mathbf{x}, \mathbf{x}^*)), \quad (19)$$

имеет в пространстве состояний устойчивую точку равновесия $\tilde{\mathbf{x}}(\mathbf{x}^*)$

$$\mathbf{f}(\tilde{\mathbf{x}}(\mathbf{x}^*), \mathbf{h}(\tilde{\mathbf{x}}(\mathbf{x}^*), \mathbf{x}^*)) = 0, \quad (20)$$

Это означает, что характеристический полином Якобиана системы в точке равновесия имеет только устойчивые корни

$$\det\left(\lambda \mathbf{E} - \frac{\partial \mathbf{f}(\tilde{\mathbf{x}}(\mathbf{x}^*), \mathbf{h}(\tilde{\mathbf{x}}(\mathbf{x}^*), \mathbf{x}^*))}{\partial \mathbf{x}}\right) = \lambda^n + p_1 \lambda^{n-1} + \dots + p_0 = \prod_{i=1}^n (\lambda - \lambda_i) = 0 \quad (21)$$

где \mathbf{E} – единичная матрица размерности $n \times n$,

$$\lambda_i = \gamma_i + j\delta_i, \quad \gamma_i < 0, \quad j = \sqrt{-1}. \quad (22)$$

На втором этапе находим оптимальные значения векторов

$$\mathbf{X}^* = (\mathbf{x}^{*,1}, \dots, \mathbf{x}^{*,M}), \quad (23)$$

такие, чтобы частное решение системы

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{h}(\mathbf{x}, \mathbf{x}^*(t))), \quad (24)$$

где $\mathbf{x}^*(t) = \mathbf{x}^{*,i}$, если $(i-1)\Delta t \leq t < i\Delta t$, $i = 1, \dots, M$, Δt – заданный интервал времени, достигнуло из начального состояния терминальное (2) с оптимальным значением критерия качества (3).

К недостаткам метода следует отнести необходимость решения задачи общего синтеза управления на первом этапе. Достоинством метода является то, что его легко перестроить под другую задачу, с другим терминальным состоянием или другим критерием качества. В данном случае повторяется только второй этап, который решается методами нелинейного программирования.

4. Вычислительный эксперимент

Рассмотрим задачу оптимального управления четырьмя мобильными гусеничными роботами при наличии четырех фазовых ограничений.

Модель объекта управления имеет следующее описание

$$\dot{x}_1^j = 0.5(u_1^j + u_2^j)\cos(x_3^j), \quad (25)$$

$$\dot{x}_2^j = 0.5(u_1^j + u_2^j)\sin(x_3^j), \quad (26)$$

$$\dot{x}_3^j = 0.5(u_1^j - u_2^j), \quad (27)$$

где j – номер робота, $j = 1, 2, 3, 4$.

Заданы ограничения на управление

$$-10 = u_i^- \leq u_i^j \leq u_i^+ = 10, \quad i = 1, 2, \quad j = 1, 2, 3, 4. \quad (28)$$

Заданы начальные и терминальные условия

$$\mathbf{x}^{0,1} = [0 \ 0 \ 0]^T, \quad \mathbf{x}^{0,2} = [0 \ 10 \ 0]^T, \quad \mathbf{x}^{0,3} = [10 \ 10 \ 0]^T, \quad \mathbf{x}^{0,4} = [10 \ 0 \ 0]^T. \quad (29)$$

$$\mathbf{x}^{f,1} = [10 \ 10 \ 0]^T, \quad \mathbf{x}^{f,2} = [10 \ 0 \ 0]^T, \quad \mathbf{x}^{f,3} = [0 \ 0 \ 0]^T, \quad \mathbf{x}^{f,4} = [0 \ 10 \ 0]^T \quad (30)$$

Задан критерий качества управления, который включает в себя штраф за нарушение фазовых ограничений

$$J = t_f + s_1 \sum_{j=1}^4 \sum_{k=1}^4 \mathfrak{G}(r_k^2 - (x_{1,k} - x_1^j)^2 - (x_{2,k} - x_1^j)^2) +$$

$$s_2 \sum_{i=1}^3 \sum_{j=i+1}^4 \mathfrak{G}(r^2 - (x_1^i - x_1^j)^2 - (x_2^i - x_2^j)^2) \rightarrow \min \quad (30)$$

$$\text{где } t_f = \max\{t_{f,j} : j = 1, 2, 3, 4\} \quad (31)$$

$t_{f,j}$ – время достижения роботом j терминального состояния $\mathbf{x}^{f,j}$ (30), определяется из соотношения (4) с $\varepsilon = 0.01$, $t^+ = 2.8$ с., s_1, s_2 – коэффициенты штрафа, $s_1 = s_2 = 2.5$, $\vartheta(a)$ – функция Хэвисайда,

$$\vartheta(a) = \begin{cases} 1 & \text{если } a > 0 \\ 0 & \text{иначе} \end{cases},$$

$r_k, x_{k,1}, x_{k,2}$ – параметры фазовых ограничений, $r_k = 1.5, k = 1, 2, 3, 4$, $x_{1,1} = 2.5, x_{1,2} = 2.5, x_{2,1} = 2.5, x_{2,2} = 7.5, x_{3,1} = 7.5, x_{3,2} = 7.5, x_{4,1} = 7.5, x_{4,2} = 2.5$.

На первом этапе решаем задачу общего синтеза управления синтеза. Поскольку все роботы идентичны, то решаем задачу для одного робота. Устанавливаем множество из N начальных значений

$$\tilde{\mathbf{X}}_0 = \{\tilde{\mathbf{x}}^{0,1}, \dots, \tilde{\mathbf{x}}^{0,N}\} \quad (32)$$

и одно терминальное условие

$$\tilde{\mathbf{x}}^f = [0 \ 0 \ 0]^T. \quad (33)$$

В качестве критерия используем функционал быстродействия

$$\tilde{J} = \tilde{t}_f \rightarrow \min, \quad (34)$$

где

$$\tilde{t}_f = \max\{\tilde{t}_{f,i} : i = 1, \dots, N\} \quad (35)$$

$\tilde{t}_{f,i}$ – время достижения решением терминального состояния (33) из начального состояния $\tilde{\mathbf{x}}^{0,i}, i = 1, \dots, N$.

Для решения задачи используем численный метод символьной регрессии, метод Декартового генетического программирования [3].

В результате получили следующую функцию управления

$$u_i = \begin{cases} u_i^- & \text{если } \tilde{u}_i < u_i^- \\ u_i^+ & \text{если } \tilde{u}_i > u_i^+, \\ \tilde{u}_i & \text{иначе} \end{cases}, \quad (36)$$

$$\text{где } \tilde{u}_1 = A + B + \text{sgn}(A)(\exp(A) - 1), \quad (37)$$

$$\tilde{u}_2 = B - A - \text{sgn}(A)(\exp(A) - 1), \quad (38)$$

$$A = c_1(x_3^* - x_3) + \operatorname{sgn}((x_1^* - x_1)(x_2^* - x_2))\sqrt{|(x_1^* - x_1)(x_2^* - x_2)|},$$

$$B = 2(x_1^* - x_1) + \operatorname{sgn}(x_1^* - x_1)c_2,$$

$$c_1 = 3.109, \quad c_2 = 3.109.$$

На втором этапе находим параметрические векторы по три для каждого робота, $\mathbf{x}^{*,1,j}$, $\mathbf{x}^{*,2,j}$, $\mathbf{x}^{*,3,j}$, $j = 1, 2, 3, 4$, чтобы все роботы достигли из своих начальных состояний (29) свои терминальные положения (30) с оптимальным значением критерия качества (31), при этом переключения точек осуществляем через интервал $\Delta t = 0.7$ с. Оптимальные значения параметров должны были находиться в диапазонах

$$-2 \leq x_i^{*,k,j} \leq 12, \quad k = 1, 2, 3, \quad i = 1, 2, \quad j = 1, 2, 3, 4, \quad (39)$$

$$-0.5\pi \leq x_3^{*,k,j} \leq 0.5\pi, \quad k = 1, 2, 3, \quad j = 1, 2, 3, 4, \quad (40)$$

Для решения этой задачи использовался алгоритм «роя-частиц» [4]. В результате были получены следующие параметрические векторы:

$$\mathbf{x}^{*,1,1} = [7.869 \ 9.348 \ -0.705]^T, \quad \mathbf{x}^{*,2,1} = [9.969 \ 0.95 \ 0.655]^T,$$

$$\mathbf{x}^{*,3,1} = [3.55 \ 8.891 \ -0.015]^T, \quad \mathbf{x}^{*,1,2} = [3.665 \ 4.978 \ 0.276]^T,$$

$$\mathbf{x}^{*,2,2} = [7.419 \ 1.331 \ -0.301]^T, \quad \mathbf{x}^{*,3,2} = [4.993 \ 7.304 \ -0.926]^T,$$

$$\mathbf{x}^{*,1,3} = [7.472 \ 8.335 \ 0.486]^T, \quad \mathbf{x}^{*,2,3} = [7.304 \ 0.104 \ -0.057]^T,$$

$$\mathbf{x}^{*,3,3} = [1.601 \ 10.157 \ 0.219]^T, \quad \mathbf{x}^{*,1,4} = [1.684 \ 0.724 \ -0.061]^T$$

$$\mathbf{x}^{*,2,4} = [4.166 \ 2.190 \ 0.139]^T, \quad \mathbf{x}^{*,3,4} = [1.581 \ 4.512 \ -0.480]^T.$$

Графики оптимальных траекторий приведены на рис. 1.

На рисунке черными кривыми линиями обозначены траектории движения роботов на плоскости $\{x_1, x_2\}$, красные круги обозначают препятствия, черные маленькие квадраты оптимальные положения параметрических точек. Как видим из рисунка каждый робот достиг своей терминальной точки без столкновения с препятствиями. Значение функционала (30) составил величину $J = 2.77$.

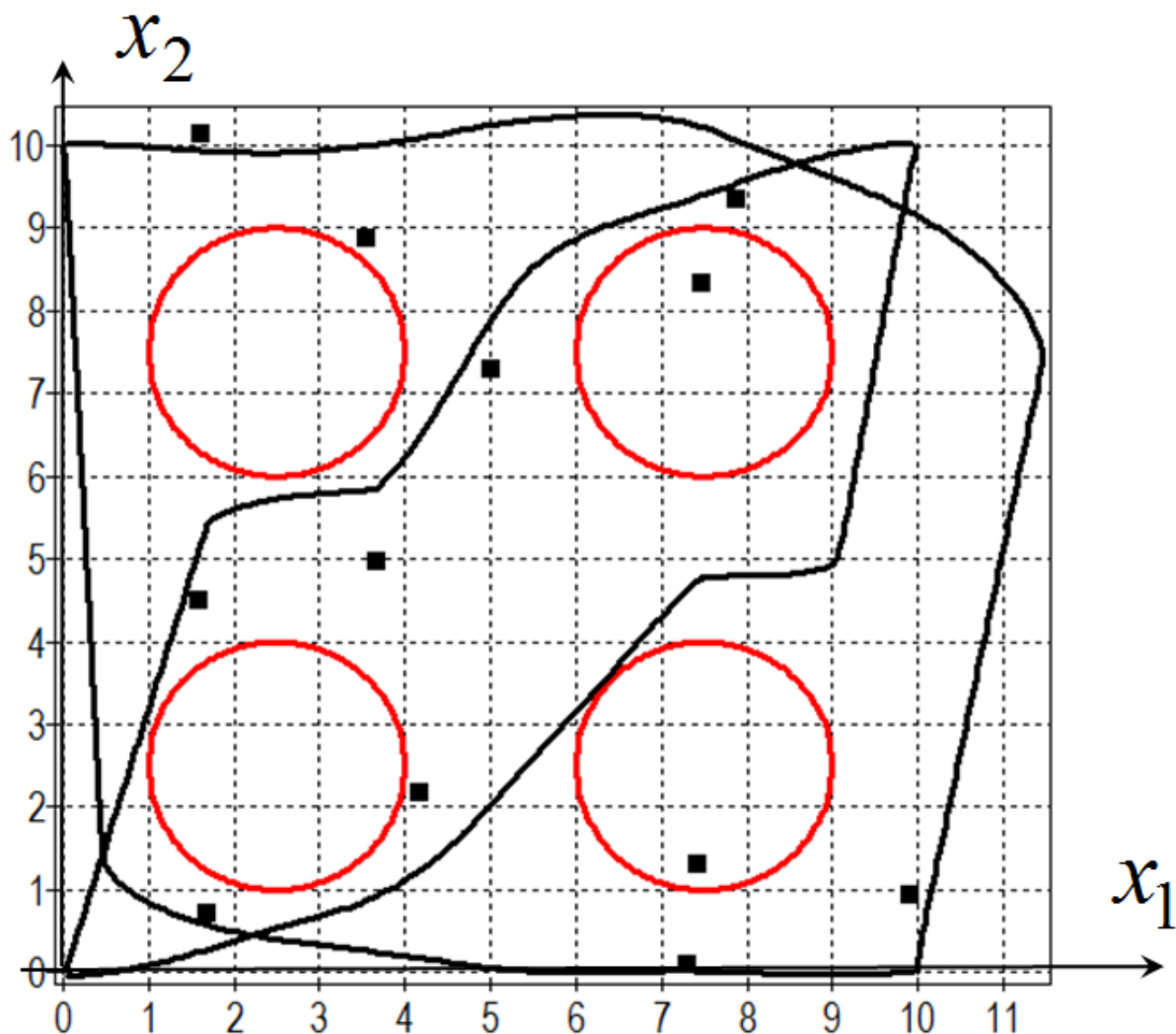


Рис. 1. Оптимальные траектории роботов

Работа выполнена при частичной поддержке Российского научного фонда, проект 19-11-00258.

Литература

1. Понтрягин Л.С., Болтянский В.Г., Гамкредидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. – М.: Наука, 1983. – 392 с.
2. Колесников А.А. Синергетическая теория управления. – Таганрог: ТРТУ, М.: Энергоатомиздат, 1984. – 344 с.
3. Дивеев А.И. Численные методы решения задачи синтеза управления. – М.: РУДН, 2019. – 192 с.
4. Дивеев А.И., Константинов С.В. Исследование практической сходимости эволюционных алгоритмов оптимального программного управления колесным роботом // Известия РАН Теория и системы управления. 2018, № 4, с. 80-106. Doi: 0.1134/S106423071804007X.