УДК 519.8

Синтез оптимальных детерминированных систем с полной обратной связью методом итерационного динамического программирования

Родионова Д.А.

Московский авиационный институт (национальный исследовательский университет), МАИ, Волоколамское шоссе, 4, Москва, A-80, ГСП-3, 125993, Россия

e-mail: d.arya.rodionova@yandex.ru

Аннотация

Предложен численный метод нахождения оптимального управления нелинейными непрерывными детерминированными динамическими системами с полной обратной связью. Сформированный алгоритм обобщает итерационный метод динамического программирования Лууса на более широкий класс задач. Эффективность алгоритма продемонстрирована на задаче нахождения оптимальной траектории перелета на круговую орбиту максимального радиуса за заданное время.

Ключевые слова: оптимальное управление, обратная связь, метод динамического программирования, метод случайного поиска.

Введение

Задачи синтеза оптимального управления нашли широкое применение при проектировании движения летательных аппаратов в сложных авиационно-космических комплексах. Для нахождения оптимального программного управления обычно применяются соотношения принципа максимума и численные методы решения двухточечных краевых задач, а для определения оптимального управления с полной обратной связью по вектору состояния – уравнение Беллмана и приближенные методы решения уравнений с частными производными.

Одним из подходов к численному решению задач оптимального управления является использование метаэвристических методов глобальной оптимизации [1-4]. К числу таких методов принадлежит метод случайного поиска с последовательной редукцией области исследования (метод Luus-Jakola) [5]. На практике он оказался применимым не только для стандартных задач, но и для невыпуклых или даже недифференцируемых функций. Известны примеры удачного применения данного метода для решения задач оптимального управления [6,7], оптимизации процессов металлообработки и химических технологий. Метод Лууса с использованием метода итерационного динамического программирования (итерационного метода Лууса) обычно используется только для нахождения оптимального программного управления при фиксированном начальном состоянии, поэтому ставится задача его распространения на более широкий класс задач поиска оптимального управления с полной обратной связью, применяемого для заданного множества начальных состояний. При этом как и в классическом методе, реализуется процедура мультитраекторного попятного движения с аппроксимацией законов управления на каждой итерации. Применимость предложенного алгоритма демонстрируется на двух нелинейных примерах управления химическим процессом и движением космического аппарата.

Постановка задачи

Поведение модели объекта управления описывается дифференциальным уравнением

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

где x – вектор состояния системы, $x \in R^n$; u – вектор управления, $u = (u_1, ..., u_q)^T \in U(t) \subseteq R^q$; $U(t)$ – множество допустимых значений управления, для каждого значения t представляющее собой прямое произведение отрезков $[a_i(t), b_i(t)]$, $i = 1, 2, ..., q$; t – непрерывное время, $t \in T = [t_0, t_N]$, начальный t_0 и конечный t_N моменты времени заданы; $f(t, x, u) = \left(f_1(t, x, u), ..., f_n(t, x, u) \right)^T$ – непрерывно-дифференцируемая вектор-функция.

Начальное условие

$$x(t_0) = x_0 \in R^n, \tag{2}$$

где начальное состояние x_0 заранее не задано и может быть произвольным.

Правый конец траектории $x(t_N)$ свободен. Предполагается, что при управлении используется информация о времени t и векторе состояние x, т.е. применяется управление с полной обратной связью.

Множество допустимых процессов $\boldsymbol{D}(t_0, x_0)$ – это множество пар $d = (x(t), u(t))$, включающих траекторию $x(t)$ и кусочно-непрерывное допустимое управление $u(t)$, где $u(t) \in U$, удовлетворяющих дифференциальному уравнению (1) и начальному условию (2).

На множестве допустимых процессов $\boldsymbol{D}(t_0, x_0)$ определен функционал качества управления

$$I(d) = \int_{t_0}^{t_N} f^0(t, x(t), u(t))dt + F\left(x(t_N)\right), \tag{3}$$

где $f^0(t, x, u)$ и $F\left(x\right)$ – заданные непрерывные функции.

Множество допустимых управлений с полной обратной связью U_n - это множество функций $u(t, x): T \times R^n \to U(t)$, которые для любых начальных условий порождают соответствующие пары $d^* = (x^*(t), u^*(t)) \in \boldsymbol{D}(t_0, x_0)$

Требуется найти такую функцию $u^*(t, x) \in U_n$, что

$$I(d^*) = \min_{d \in \boldsymbol{D}(t_0, x_0)} I(d), \quad \forall x_0 \in R^n, \tag{4}$$

где $d^* = d(x^*(\cdot), u^*(\cdot) = u^*(\cdot, x(\cdot)))$.

Заметим, что на практике управление ищется не для всех начальных состояний, а только для принадлежащих некоторому заданному множеству Ω , определяемому спецификой решаемой прикладной задачи.

Стратегия поиска экстремума методом Лууса

Рассмотрим стратегию поиска условного глобального экстремума функции многих переменных, используемую в методе Luus-Jakola [5]. Задается начальная точка на множестве допустимых решений D. Из нее генерируются R точек в случайных направлениях с учетом характерных размеров множества с проверкой принадлежности ему. Среди полученных точек выбирается наилучшая по величине критерия, и из нее процесс продолжается. При этом размер множества поиска сокращается (при помощи параметра γ) от итерации к итерации вплоть до достижения их заданного числа. Как только заданное число итераций выполнено, завершается «проход». При переходе к следующему проходу размер множества поиска восстанавливается (при помощи параметра η), а затем снова выполняется заданное число итераций.

Эта стратегия используется при реализации модифицированного метода итерационного динамического программирования, в котором на каждой итерации путем применения аппроксимации по методу «ближнего соседа» происходит обработка найденного управления для выработки закона, связывающего значения координат вектора состояния со значениями координат вектора управления. В качестве перспективного метода многомерной аппроксимации в алгоритме может использоваться разложение по ортонормированным системам базисных функций, применяемых в спектральном методе анализа систем управления [8].

Алгоритм решения задачи

Шаг 1. Задать множество начальных условий $\Omega \subseteq R^n$.

Шаг 2. Генерировать M точек внутри множества начальных условий Ω, используя равномерное распределение: $x^i(0) = x_0^i,\ i = 1,\ldots,M$.

Шаг 3. Для каждой точки использовать процедуру итерационного динамического программирования.

3.0. Разбить интервал времени $[t_0, t_N]$ на N стадий длиной $L = \dfrac{t_N - t_0}{N}$.

Считать управление кусочно-постоянным на каждой стадии: $u(t_{k-1}) = \mathrm{const}$ при $t \in [t_{k-1}, t_k)$, $k = 1,...,N$. Тогда $I = \sum_{k=1}^{N} \int_{t_{k-1}}^{t_k} f^0(t, x(t), u(t_{k-1}))\, dt + F(x(t_N))$.

3.1. Задать параметры: N - число стадий по времени; P - число проходов; $ITER$ - число итераций, выполняемых за один проход; r_{in} - вектор, характеризующий размер множества допустимых решений; γ - коэффициент уменьшения размера области поиска по управлению; η - коэффициент восстановления размера области поиска по управлению; R - число управлений, генерируемых в текущей точке.

Положить $q = 0$ (число проходов), $j = 1$ (число итераций).

Задать начальный закон управления $u^0(\cdot) = \{u(t_0), u(t_0 + L), ..., u(t_N - L)\}$. Положить $u^{*,j}(\cdot) = u^0(\cdot)$.

3.2. Положить $r^j = \eta^q \cdot r_{in}$ (вектор, характеризующий размеры текущей области поиска по управлению).

3.3. Решить уравнение $\dot{x} = f(t, x(t), u^{*,j}(t_{k-1}))$, $k = 1,...,N$, $x^i(0) = x_0^i$, $i = 1,...,M$. Результатом является траектория $x^*(\cdot)$. Подсчитать значение функционала $I(x^*(\cdot), u^*(\cdot))$. Запомнить координаты вектора состояния $x^{*,j}(t_{N-1}))$.

3.4. Реализовать процедуру поиска наилучшего решения на каждой стадии, выполняя попятное движение от конца (времени t_N) к началу (времени t_0).

3.4.1. На N-й стадии ($t \in [t_{N-1}, t_N]$):

- генерировать R допустимых векторов управления:

$$u^{j+1,m}(t_{N-1}) = u^{*,j}(t_{N-1}) + D^m \cdot r^j, \quad m = 1,...,R$$

с проверкой условия $u_i^{j+1,m}(t_{N-1}) \in [a_i(t_{N-1}), b_i(t_{N-1})], \quad i = 1,...,q,$ где D^m – диагональная матрица со случайными взаимно независимыми элементами, равномерно распределенными на отрезке $[-1,1]$;

- решить R раз уравнение

$$\dot{x}(t) = f(t, x(t), u^{j+1,m}(t_{N-1})), \quad t \in [t_{N-1}, t_N]$$

с начальным условием $x(t_{N-1}) = x^{*,j}(t_{N-1})$. В результате получить решения $x^{j+1,m}(t)), \quad t \in [t_{N-1}, t_N],\ m = 1,...,R;$

- вычислить R соответствующих значений функционала

$$I(x^{*,j}(t_{N-1}),1) = \int_{t_{N-1}}^{t_N} f^0(t, x^{j+1,m}(t), u^{j+1,m}(t_{N-1}))\, dt + F(x^{j+1,m}(t_N))\ ;\ m = 1,...,R;$$

- среди полученных значений выбрать наименьшее, а соответствующее ему управление на N-й стадии обозначить $u^{*,j+1}(t_{N-1})$.

3.4.2. На k-й стадии $(t \in [t_{k-1}, t_k))$:

- генерировать R допустимых векторов управления

$$u^{j+1,m}(t_{k-1}) = u^{*,j}(t_{k-1}) + D^m \cdot r^j, \quad m = 1,...,R$$

с проверкой условия $u_i^{j+1,m}(t_{k-1}) \in [a_i(t_{k-1}), b_i(t_{k-1})],\ i = 1,...,q;$

- решить R раз уравнение (1) с начальным условием $x(t_{k-1}) = x^{*,j}(t_{k-1})$ и управлением $\{u^{(i)j+1,m}(t_{k-1}), \bar{u}^{(i)}(t_k), ..., \bar{u}^{(i)}(t_{N-1})\}$, где управления $\bar{u}^{(i)}(t_k), ..., \bar{u}^{(i)}(t_{N-1})$ определить следующим образом:

- в силу интегрирования уравнения (1) с начальным условием $x(t_{k-1}) = x^{(i)*,j}(t_{k-1})$ и управлением $u^{(i)j+1,m}(t_{k-1})$ на промежутке $[t_{k-1}, t_k]$ получить часть траектории $x^{(i),j+1,m}(t)$, а при $t = t_k$ - вектор состояния $x^{(i)\,j+1,m}(t_k)$;
- среди векторов $x^{(i)\,*,j}(t_k)$, $i = 1,...,M$ найти такой $x^{(s)\,*,j}(t_k)$, евклидово расстояние от которого до вектора $x^{(i),j+1,m}(t_k)$ минимально;
- положить $\bar{u}^{(i)}(t_k) = u^{(s)*,j+1}(t_k)$, т.е. в качестве прикладываемого управления выбрать соответствующее найденному вектору $x^{(s)\,*,j}(t_k)$, полученное на предыдущих стадиях;
- выполнить описанные операции при $k = 2,...,N-1$.

В результате получить решения $x^{(i)j+1,m}(t)),\quad t \in [t_{k-1}, t_N]$, $m = 1,...,R$;

- вычислить R соответствующих значений функционала

$$I(x^{*,j}(t_{k-1}), N-k+1) = \int_{t_{k-1}}^{t_k} f^0(t, x^{j+1,m}(t), u^{j+1,m}(t_{k-1}))dt + \sum_{s=k+1}^{N} \int_{t_{s-1}}^{t_s} f^0(t, x^{j+1,m}(t), u^{*,j+1}(t_{s-1}))dt +$$

$$+F(x^{j+1,m}(t_N));$$

- среди полученных значений выбрать наименьшее, а соответствующее ему управление на k-й стадии обозначить $u^{*,j+1}(t_{k-1})$.

3.4.3. На 1-й стадии $(t \in [t_0, t_1))$:

- генерировать R допустимых векторов управления:

$u^{j+1,m}(t_0) = u^{*,j}(t_0) + D^m \cdot r^j, \quad m = 1,...,R$

с проверкой условия $u_i^{j+1,m}(t_0) \in [a_i, b_i],\ i = 1,...,q;$

- решить R раз уравнение (1) с начальным условием $x^i(t_0) = x_0^i$ и управлением $\{u^{j+1,m}(t_0), u^{*,j+1}(t_1),...,u^{*,j+1}(t_{N-1})\}$, где управления $u^{*,j+1}(t_1),...,u^{*,j+1}(t_{N-1})$ найдены на предыдущих стадиях. В результате получить решения $x^{j+1,m}(t)), \quad t \in [t_0, t_N], m = 1,...,R;$

- вычислить R соответствующих значений функционала

$$I(x_0, N) = \int_{t_0}^{t_1} f^0(t, x^{j+1,m}(t), u^{j+1,m}(t_0))dt + \sum_{s=2}^{N} \int_{t_{s-1}}^{t_s} f^0(t, x^{j+1,m}(t), u^{*,j+1}(t_{s-1}))dt + F(x^{j+1,m}(t_N));$$

- среди полученных значений выбрать наименьшее, а соответствующее ему управление на 1-й стадии обозначить $u^{(i)*,j+1}(t_0)$; оно соответствует состоянию $x_0^{(i)}$, которому соответствует также управление $\boldsymbol{u}^{*,j+1}(t_0, x^{(i)}(t_0)) = u^{(i)*,j+1}(t_0)$.

Итерация завершена.

3.5. Уменьшить размер области поиска:

$r^{j+1} = \gamma \cdot r^j.$

3.6. Положить $j = j + 1$. Если $j < ITER$, то перейти к шагу 3.3. Иначе перейти к шагу 7.

3.7. Положить $q = q + 1$. Если $q < P$, положить $j = 1$ и перейти к шагу 3.2. Иначе процесс завершить. В качестве решения выбрать последнее найденное управление и соответствующую ему траекторию.

Прикладные примеры и анализ эффективности

На основе изложенного алгоритма сформирована программа, реализующая применение модифицированного итерационного динамического программирования и метода Luus-Jakola к задачам нахождения оптимального управления нелинейными детерминированными системами с полной обратной связью. Среда разработки Microsoft Visual Studio 2010, язык программирования C#.

Программа позволяет выбирать задачу из заданного списка модельных примеров, задавать ограничения на множество начальных условий, задавать параметры решения задачи, число шагов разбиения; для заданного начального условия получать графические результаты (вид оптимального управления и оптимальной траектории), выводить значение оптимального управления и оптимальной траектории на каждом шаге; настраивать параметры изображения результатов; генерировать и сохранять отчет о работе.

Пример 1 (задача управления химическим процессом).

Поведение непрерывной модели объекта управления описывается системой дифференциальных уравнений:

$$\dot{x}_1 = u_4 - qx_1 - 17{,}6x_1x_2 - 23x_1x_6u_3;$$

$$\dot{x}_2 = u_1 - qx_2 - 17{,}6x_1x_2 - 146x_2x_3;$$

$$\dot{x}_3 = u_2 - qx_3 - 73x_2x_3;$$

$$\dot{x}_4 = -qx_4 + 35{,}2x_1x_2 - 51{,}3x_4x_5;$$

$$\dot{x}_5 = -qx_5 + 219x_2x_3 - 51{,}3x_4x_5;$$

$$\dot{x}_6 = -qx_6 + 102{,}6x_4x_5 - 23x_1x_6u_3;$$

$$\dot{x}_7 = -qx_7 + 46x_1x_6u_3;$$

$$\dot{x}_8 = 5{,}8(qx_1 - u_4) - 3{,}7u_1 - 4{,}1u_2 + q(23x_4 + 11x_5 + 28x_6 + 35x_7) - 5u_3^2 - 0{,}099,$$

где $q = u_1 + u_2 + u_4$, $0 \le t \le t_N = 0{,}2$.

Множество начальных условий Ω : $0{,}14 \le x_1 \le 0{,}23$, $0{,}19 \le x_2 \le 0{,}31$, $0{,}035 \le x_3 \le 0{,}058$, $0{,}068 \le x_4 \le 0{,}112$, $0{,}14 \le x_5 \le 0{,}22$, $0{,}1 \le x_6 \le 0{,}17$, $0{,}078 \le x_7 \le 0{,}13$, $x_8 = 0$.

Ограничения на управление:

$$0 \le u_1(t) \le 20,\ 0 \le u_2(t) \le 6,\ 0 \le u_3(t) \le 4,\ 0 \le u_4(t) \le 20.$$

Критерий качества управления: $I = x_8(t_N)$.

Необходимо максимизировать критерий при помощи выбора управления, удовлетворяющего заданным ограничениям.

Выберем следующие параметры метода: число генерируемых решений $R = 100$; число проходов $P = 10$; число итераций за один проход $ITER = 10$; коэффициент уменьшения размера множества поиска $0{,}8 \le \gamma \le 0{,}96$; коэффициент восстановления начального множества поиска $\eta = 0{,}89$, $M = 50$.

Результаты решения задачи для начального условия $x(0)=\left[0{,}1883;0{,}2507;0{,}0467;0{,}0899;0{,}1804;0{,}1394;0{,}1046;0\right]^T$ представлены в табл. 1 и на рис. 1.

Таблица 1.

Оптимальная траектория и управление в примере 1

γ	$N=11$	$N=27$	$N=40$
0,80	21,437	21,378	21,491
0,83	21,470	21,452	21,568
0,87	21,659	21,605	21,671
0,90	21,612	21,597	21,675
0,93	21,632	21,721	21,689
0,96	21,614	21,674	21,701

Для начального условия $x(0)=[0{,}18;0{,}25;0{,}04;0{,}08;0{,}18;0{,}13;0{,}1;0]^T$ оптимальное значение критерия (при $\gamma=0{,}9$) составило $I=20{,}4817$ (значение для соответствующего программного управления $I^*=20{,}5499$).

Для начального условия $x(0)=[0{,}2;0{,}3;0{,}04;0{,}1;0{,}2;0{,}15;0{,}1;0]^T$ оптимальное значение критерия (при $\gamma=0{,}9$) составило $I=22{,}8273$ (значение для соответствующего программного управления $I^*=22{,}4118$).

Таким образом, для заданного начального условия управление с полной обратной связью совпадает по величине критерия с программным управлением с точностью до $\max|I-I^*|=0{,}14$.

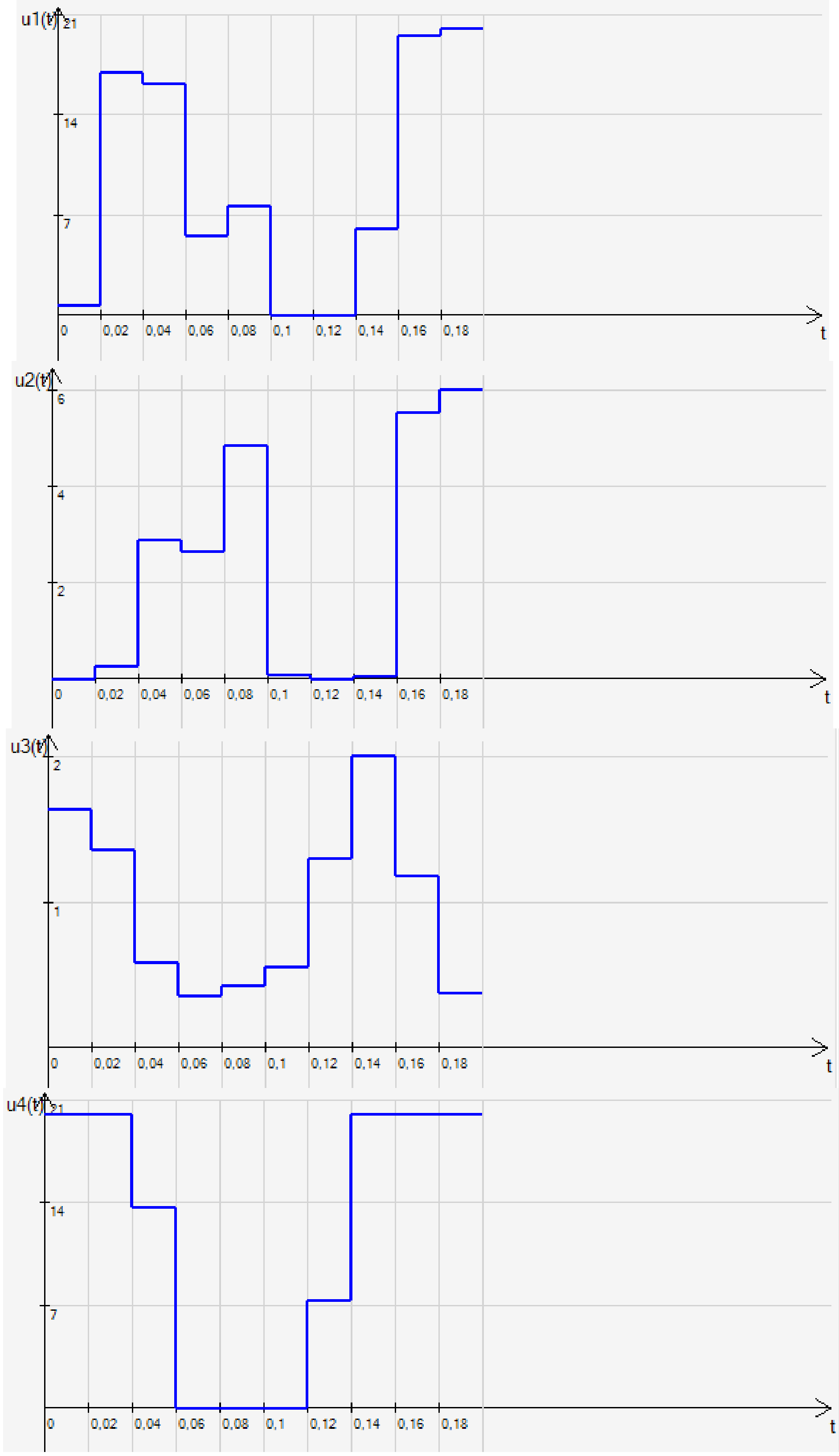

Рис. 1. Оптимальное управление в примере 1

Пример 2 (задача нахождения оптимальной траектории перелета на круговую орбиту максимального радиуса за заданное время).

Найти оптимальный угол направления тяги ракеты для перелета с заданной начальной круговой орбиты на круговую орбиту максимально возможного радиуса. Ракетный двигатель развивает постоянную тягу T, время работы двигателя задано.

Модель системы управления описывается системой дифференциальных уравнений:

$$\dot{r} = u;$$

$$\dot{u} = \frac{v^2}{r} - \frac{\mu}{r^2} + \frac{T\sin\phi}{m_0 - |\dot{m}|t};$$

$$\dot{v} = -\frac{uv}{r} + \frac{T\cos\phi}{m_0 - |\dot{m}|t};$$

где r - радиальное расстояние от космического корабля до центра притяжения, u - радиальная компонента скорости, v - тангенциальная компонента скорости, $\dot{m}$ - массовый расход топлива, ϕ - угол направления тяги, μ - гравитационная постоянная притягивающего центра.

Параметры задачи связаны между собой соотношениями:

$$\frac{T/m_0}{\mu/r_0^2} = 0{,}1405, \quad \frac{\dot{m}\sqrt{\mu/r_0}}{T} = 0{,}533, \quad \frac{t_N}{\sqrt{r_0^3/\mu}} = 3{,}32.$$

Зададим следующие значения параметров: $m_0 = 0{,}25$, $\mu = 11$, $T = 1{,}55$, $r_0 = 1$, $\dot{m} = 0{,}25$. Конечные условия: $u(t_N) = 0$, $v(t_N) = \sqrt{\dfrac{\mu}{r(t_N)}}$, $0 \le t \le t_N = 1$.

Множество начальных состояний Ω : $0,8 \le x_1 \le 1,2$, $x_2 = 0$, $3 \le x_3 \le 3,7$.

Ограничения на управление: $0 \le u \le 2\pi$.

Критерий качества управления: $I = r(t_N)$.

Необходимо максимизировать критерий при помощи выбора управления, удовлетворяющего заданным ограничениям.

Выберем следующие параметры метода: число генерируемых решений $R = 100$; число проходов $P = 10$; число итераций за один проход $ITER = 10$; коэффициент уменьшения размера множества поиска $0,8 \le \gamma \le 0,96$; коэффициент восстановления начального множества поиска $\eta = 0,8$, $M = 150$.

Результаты решения задачи для начального условия $x(0) = [1; 0; \sqrt{11}]^T$ представлены в табл. 2 и на рис. 2.

Таблица 2.

Оптимальная траектория и управление в примере 2

γ	$N = 20$	$N = 30$
0,80	1,5147	1,5120
0,83	1,5167	1,5147
0,87	1,5143	1,5163
0,90	1,5207	1,5183
0,93	1,5188	1,5192
0,96	1,5175	1,5214

Для начального условия $x(0)=[0,9;0;3,2]^T$ оптимальное значение критерия (при $\gamma=0,93$) составило $I=1,1276$ (значение для соответствующего программного управления $I^*=1,1547.$).

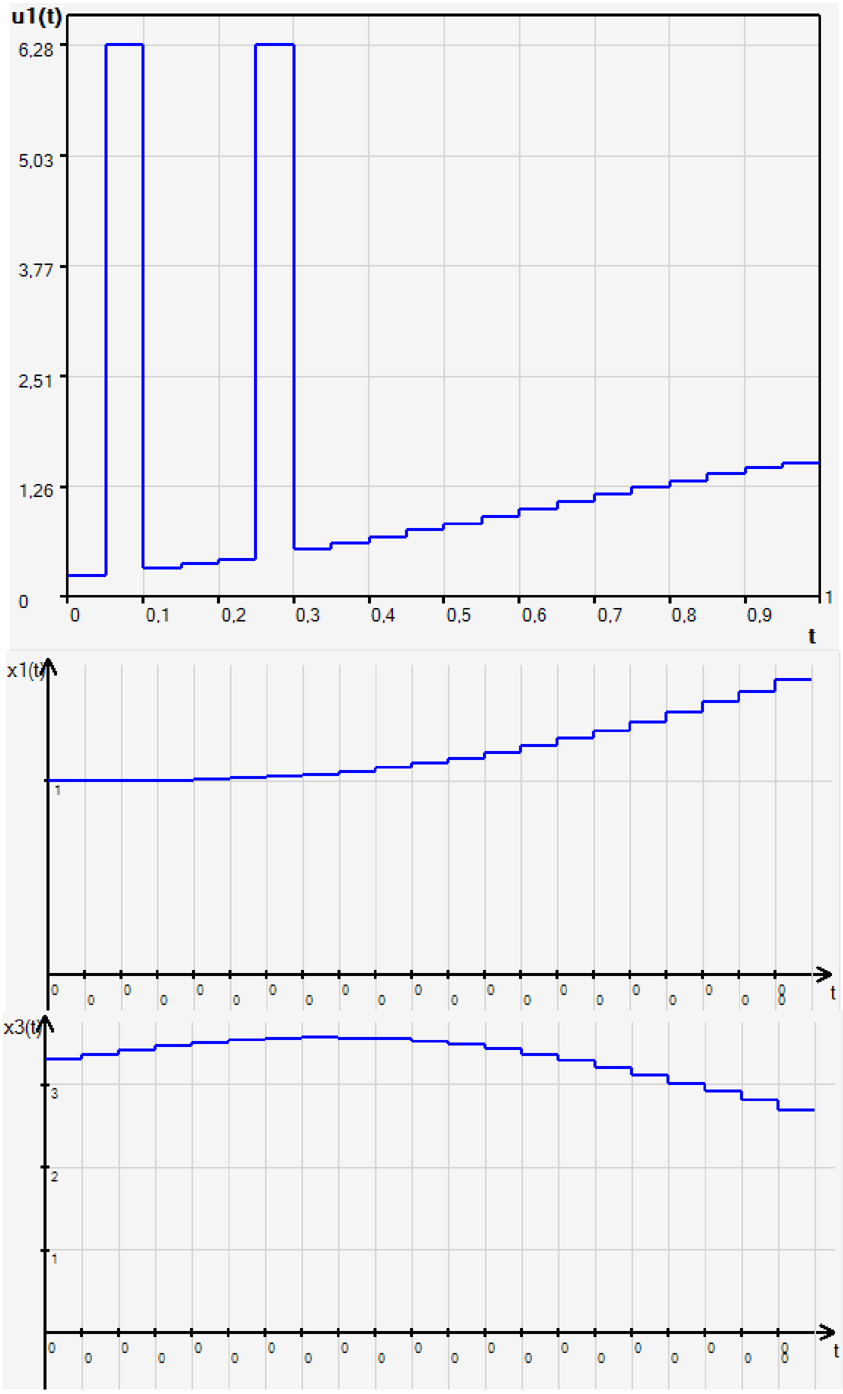

Рис. 2. Оптимальное управление и траектория в примере 2 ($x(0)=[1;0;\sqrt{11}]^T$)

Для начального условия $x(0) = [1,1; 0; 3,5]^T$ оптимальное значение критерия (при $\gamma = 0,93$) составило $I = 1,8016$ (значение для соответствующего программного управления $I^* = 1,909$).

Таким образом, для заданного начального условия управление с полной обратной связью совпадает с программным управлением по величине критерия с точностью до $\max\left|I - I^*\right| = 0,097$.

Приведенные примеры показывают эффективность применения предложенного алгоритма модифицированного итерационного динамического программирования и метода Luus-Jakola для решения задачи синтеза оптимального управления с полной обратной связью.

Заключение

В данной статье было разработано алгоритмическое и программное обеспечение применения метода случайного поиска с последовательной редукцией области исследования и принципа итерационного динамического программирования в задачах синтеза оптимального управления непрерывными детерминированными системами с полной обратной связью. Были решены две прикладные задачи, в том числе задача поиска оптимальной траектории перелета на круговую орбиту максимального радиуса за заданное время, демонстрирующие эффективность применения данного метода.

Библиографический список

1. Пантелеев А.В., Метлицкая Д.В., Алешина Е.А. Методы глобальной оптимизации. Метаэвристические стратегии и алгоритмы. – М.: Вузовская книга, 2013. - 244 с.

2. Пантелеев А.В. Применение эволюционных методов глобальной оптимизации в задачах оптимального управления детерминированными системами. - М.: Изд-во МАИ, 2013. - 160 с.

3. Пантелеев А.В., Метлицкая Д.В. Применение генетических алгоритмов с бинарным и вещественным кодированием для приближенного синтеза субоптимального управления детерминированными системами // Автоматика и телемеханика. 2011. №11. С. 117–129.

4. Пантелеев А.В., Метлицкая Д. В. Применение генетических алгоритмов к задаче оптимального управления дальностью полета летательного аппарата типа воздух-воздух // Вестник Московского авиационного института. 2011. Т.18. №4. С. 102–113.

5. Luus R., Jaakola T.H.I. Optimization by direct search and systematic reduction of the size of search region// American Institute of Chemical Engineers Journal (AIChE) V. 19 (4). 1973. P.760–766.

6. Luus R. Iterative Dynamic Programming. - CRC Press.-2000, 344 p.

7. Bojkov R., Hansel B., Luus R. // Application of direct search optimization to optimal control problems // Hungarian Journal of Industrial Chemistry. – V.21. 1993. P.177–185.

8. Пантелеев А.В., Рыбаков К.А. Прикладной вероятностный анализ нелинейных систем управления спектральным методом.- М.: Изд-во МАИ-ПРИНТ, 2010. - 160 с.

9. Пантелеев А.В., Родионова Д.А. Применение метода случайного поиска с последовательной редукцией области исследования в задачах оптимального управления детерминированными системами // Известия института инженерной физики. 2014. №3(33). С.17-22.