1.2 Постановка задачи быстродействия. Погружение. Проверка принципа оптимальности. Формулировка Предположений 1.1, 1.2 о функции Беллмана

Рассмотрим задачу быстродействия

$$\begin{cases} \dot{x} = f(x,u), \qquad x \in I\!R^n, \qquad u \in U, \\ x(t_0) = x_0, \qquad x(t_1) = x_1, \\ T \equiv t_1 - t_0 \to \min, \end{cases} \tag{P}$$

в которой допустимыми управлениями являются кусочно-непрерывные функции времени t со значениями в области управления U, $f(x,u)$ — n-мерная векторная функция с координатами $f^1(x,u)$, $\dots$, $f^n(x,u)$, причем функции $f^i(x,u), i = 1,\dots,n$, и их частные производные $\frac{\partial f^i(x,u)}{\partial x^j}, i,j = 1,\dots,n$, определены и непрерывны на прямом произведении $I\!R^n \times U$.

При построении семейства задач (P_α) в качестве параметра погружения α выбирается начальное состояние x_0 управляемого объекта. Задача изучается при определенных условиях (**Предположения 1.1, 1.2**).

Предположение 1.1. *При любом начальном состоянии $x_0 \in I\!R^n$ существует оптимальное управление, которое осуществляет перевод объекта из x_0 в конечную точку x_1 со временем быстродействия $T(x_0)$, причем*

$$T(x_1) = 0, \qquad T(x_0) > 0 \ \text{ при } \qquad x_0 \neq x_1.$$

Таким образом, функция Беллмана

$$V(x) = -T(x)$$

определена при всех $x \in I\!R^n$, причем

$$V(x_1) = 0, \qquad V(x) < 0, \ \ x \neq x_1.$$

Предположение 1.2. *Функция $T(x)$ непрерывна при всех $x \in I\!R^n$ и имеет непрерывный градиент $T'(x)$ при $x \in I\!R^n \setminus \{x_1\}$.* Таким образом, функция Беллмана $V(x)$ непрерывна в $I\!R^n$ и имеет непрерывный градиент $V'(x)$ в $I\!R^n \setminus \{x_1\}$.

Замечание 1.2. Предположение о гладкости функции Беллмана может не выполняться. Например, в задаче быстродействия для тележки [1] функция оптимального времени $T(x)$ недифференцируема на линии переключения AOB, [1], §5, Пример 1.

Для задачи быстродействия (P) легко выполнить обоснование принципа оптимальности. Покажем, что любой "хвост"оптимальной траектории оптимален. Пусть

$$x(t), \ t_0 \le t \le t_1; \qquad x(t_0) = x_0, \ x(t_1) = x_1, \tag{1}$$

— оптимальная траектория, отвечающая оптимальному управлению

$$u(t), \ t_0 \le t \le t_1; \qquad u(t) \in U. \tag{2}$$

Возьмем любое число $\tau \in (t_0, t_1)$ и проверим, что траектория

$$x(t),\ \tau \le t \le t_1,$$

("хвост"траектории (1)) оптимальна по быстродействию из точки $x(\tau)$ в момент времени τ в точку x_1 в момент времени t_1. Если допустить противное, то существуют допустимое управление

$$\hat{u}(t),\ \tau \le t \le \hat{t}_1,$$

и отвечающая ему траектория

$$\hat{x}(t),\ \tau \le t \le \hat{t}_1,$$

такие, что

$$\hat{x}(t)|_{t=\tau} = x(\tau), \quad \hat{x}(t)|_{t=\hat{t}_1} = x_1, \qquad \hat{t}_1 < t_1.$$

Тогда допустимое управление

$$u_*(t) = \begin{cases} u(t),\ t_0 \le t \le \tau, \\ \hat{u}(t),\ \tau < t \le \hat{t}_1, \end{cases}$$

порождает траекторию

$$x_*(t) = \begin{cases} x(t),\ t_0 \le t \le \tau, \\ \hat{x}(t),\ \tau \le t \le \hat{t}_1, \end{cases}$$

вдоль которой переход из точки x_0 в точку x_1 происходит быстрее, чем вдоль траектории (1), что противоречит ее оптимальности. Принцип оптимальности обоснован. Заметим, что управление $u_*(t)$ допустимо, так как класс допустимых управлений состоит из кусочно-непрерывных функций и возможный скачек управления в точке τ не выводит управление $u_*(t)$ из класса допустимых управлений.

1.3 Вывод уравнения Беллмана. Необходимые условия оптимальности

Рассмотрим оптимальную пару $(x(t), u(t)),\ t_0 \le t \le t_1$. На основе принципа оптимальности можно записать равенство

$$T(x(t)) = \Delta t + T(x(t + \Delta t))$$

при любом $\Delta t > 0;\ t, t + \Delta t \in [t_0, t_1)$. Отсюда вытекает соотношение

$$\frac{V(x(t + \Delta t)) - V(x(t))}{\Delta t} = 1.$$

Переходя в нем к пределу при $\Delta t \to 0$, получаем

$$\frac{d}{dt} V(x(t)) = 1,$$

или

$$(V'(x(t)), \dot{x}(t)) = 1.$$

При выводе последнего соотношения мы использовали **Предположение 1.2** о гладкости функции Беллмана и считали, что в точке $t \in [t_0, t_1)$ управление $u(t)$ непрерывно, что влечет существование производной $\dot{x}(t)$. Привлекая дифференциальное уравнение управляемого движения, из последнего равенства получаем *необходимое условие оптимальности* пары $(x(t), u(t)),\ t_0 \le t \le t_1$:

$$(V'(x(t)), f(x(t), u(t))) = 1,\ t_0 \le t < t_1. \tag{3}$$

Момент времени $t = t_1$ здесь исключен, так как в точке x_1 существование градиента $V'(x)$ функции Беллмана не предполагается. Справедливость равенства (3) в точках $\tau_j \in (t_0, t_1)$ скачков управления проверяется предельным переходом $t \to \tau_j - 0$ (с учетом соглашения $u(\tau_j - 0) = u(\tau_j)$). Соотношение (3) при $t = t_0$ влечет

$$(V'(x_0), f(x_0, u(t_0))) = 1 \qquad \forall x_0 \in I\!R^n \setminus \{x_1\}. \tag{4}$$

Пусть теперь v — произвольная точка области управления U. Рассмотрим движение $y(t)$, $t_0 \leq t \leq t_0 + \Delta t$, $\Delta t > 0$, фазовой точки из начального состояния $x_0 \neq x_1$ под воздействием постоянного управления, равного v. При малых Δt точка $y(t_0 + \Delta t) \neq x_1$. Из точки $y(t_0 + \Delta t)$ дальнейшее движение фазовой точки выполняем оптимальным образом до попадания в конечную точку x_1. Имеет место неравенство

$$T(x_0) \leq \Delta t + T(y(t_0 + \Delta t)),$$

из которого следует

$$\frac{V(y(t_0 + \Delta t)) - V(x_0)}{\Delta t} \leq 1.$$

Но $y(t_0 + \Delta t) = y(t_0) + \dot{y}(t_0)\Delta t + o(\Delta t) = x_0 + f(x_0, v)\Delta t + o(\Delta t)$, поэтому предельный переход при $\Delta t \to +0$ в последнем соотношении приводит к неравенству

$$(V'(x_0), f(x_0, v)) \leq 1 \qquad \forall v \in U, \ \forall x_0 \neq x_1. \tag{5}$$

Сравнение соотношений (4) и (5) дает:

$$\max_{v \in U} \ (V'(x_0), f(x_0, v)) = 1 \qquad \forall x_0 \neq x_1. \tag{6}$$

Итак, для функции $V(x)$ получено дифференциальное уравнение Беллмана

$$\left\{ \begin{array}{ll} \max\limits_{u \in U} \ (V'(x), f(x, u)) = 1, & \quad x \neq x_1, \\ V(x_1) = 0, & \end{array} \right. \tag{7}$$

к которому мы присоединили граничное условие; напомним, что $V(x) < 0$ при $x \neq x_1$. Таким образом, доказана

Теорема 1.1. *В Предположениях 1,2 функция Беллмана $V(x)$ удовлетворяет условиям (7). Для оптимальной пары $(x(t), u(t))$, $t_0 \leq t \leq t_1$, имеет место равенство (3).*

Уравнение Беллмана (7) можно переписать в форме уравнения

$$\max_{u \in U} \ \sum_{i=1}^{n} \frac{\partial V(x)}{\partial x^i} f^i(x, u) = 1$$

с частными производными первого порядка неизвестной функции $V(x)$, причем в его левой части содержится операция взятия максимума. Оказывается, что уравнение Беллмана позволяет также описать и достаточные условия оптимальности в задаче быстродействия (P).

1.4 Уравнение Беллмана. Достаточные условия оптимальности

Теорема 1.2. *Пусть функция $V(x)$*
1) определена и непрерывна при всех $x \in I\!R^n$, $V(x_1) = 0$,$V(x) < 0, x \neq x_1$,

2) имеет непрерывный градиент $V'(x)$ при $x \in I\!R^n \setminus \{x_1\}$.
Тогда, если допустимый процесс

$$(x(t), u(t)), t_0 \le t \le t_1, \quad x(t_0) = x_0, x(t_1) = x_1, \; x(t) \ne x_1, t_0 \le t < t_1,$$

удовлетворяет условию

$$(V'(x(t)), f(x(t), u(t))) = 1, \quad t_0 < t < t_1, \tag{9}$$

то этот процесс оптимален по быстродействию в задаче (P), а время быстродействия $T(x_0) \equiv t_1 - t_0 = -V(x_0)$.

Доказательство. Покажем сначала, что для времени перехода при условии (9) имеет место равенство

$$t_1 - t_0 = -V(x_0).$$

Находим:

$$-V(x_0) = 0 - V(x_0) = V(x_1) - V(x_0) = V(x(t))|_{t=t_0}^{t=t_1} = \int\limits_{t_0}^{t_1} \frac{d}{dt} V(x(t))\; dt =$$

$$= \int\limits_{t_0}^{t_1} (V'(x(t)), \dot{x}(t))\; dt = \int\limits_{t_0}^{t_1} (V'(x(t)), f(x(t), u(t)))\; dt = \{(9)\}$$

$$= \int\limits_{t_0}^{t_1} 1\, dt = t_1 - t_0.$$

Любой другой допустимый процесс

$$\hat{x}(t), \hat{u}(t), t_0 \le t \le \hat{t}_1, \quad \hat{x}(t_0) = x_0, \hat{x}(\hat{t}_1) = x_1, \hat{x}(t) \ne x_1, t_0 \le t < \hat{t}_1,$$

имеет время перехода

$$\hat{t}_1 - t_0 \ge t_1 - t_0.$$

Действительно,

$$t_1 - t_0 = -V(x_0) = V(x_1) - V(x_0) = V(\hat{x}(t))|_{t=t_0}^{t=\hat{t}_1} = \int\limits_{t_0}^{\hat{t}_1} \frac{d}{dt} V(\hat{x}(t))\, dt =$$

$$= \int\limits_{t_0}^{\hat{t}_1} (V'(\hat{x}(t)), \dot{\hat{x}}(t))\, dt = \int\limits_{t_0}^{\hat{t}_1} (V'(\hat{x}(t)), f(\hat{x}(t), \hat{u}(t)))\, dt \le$$

$$\le \int\limits_{t_0}^{\hat{t}_1} \max_{v \in U} (V'(\hat{x}(t)), f(\hat{x}(t), v))\, dt =$$

$$= \{p\} = \int\limits_{t_0}^{\hat{t}_1} 1\, dt = \hat{t}_1 - t_0.$$

Теорема 1.2 доказана.