Лекция 2

2 Метод динамического программирования в задаче оптимального управления с интегральным функционалом

2.1 Постановка задачи

Рассмотрим задачу оптимального управления с интегральным функционалом

$$\begin{cases} \dot{x} = f(x,u), & x \in I\!R^n, u \in U, \quad t_0 \le t \le t_1, \\ x(t_0) = x_0, x(t_1) = x_1, \\ L = \int\limits_{t_0}^{t_1} f^0(x(t),u(t))\,dt \to \min. \end{cases} \tag{P}$$

Здесь $f(x,u)$ — n-мерная векторная функция с координатами $f^1(x,u)$, …, $f^n(x,u)$. Начальный момент времени t_0 и конечный момент времени t_1 считаются заданными, $t_0 < t_1$. Начальная точка x_0 траектории и ее конечная точка x_1 фиксированы. Класс допустимых управлений состоит из кусочно-непрерывных функций времени со значениями в области управления U. Предполагается, что функции $f^i(x,u), i = 0,\dots,n$, и их частные производные $\frac{\partial f^i(x,u)}{\partial x^j}, i = 0,\dots,n, j = 1,\dots,n$, определены и непрерывны на прямом произведении $I\!R^n \times U$.

Ниже обсуждается метод динамического программирования применительно к задаче оптимального управления (P). Основной конструкцией метода динамического программирования в задаче (P) является уравнение Беллмана

$$V'_t(x,t) + \max_{u\in U}\left\{-f^0(x,u) + (V'_x(x,t), f(x,u))\right\} = 0.$$

Это уравнение можно записать в следующей компактной форме

$$\max_{u\in U}\ \mathcal{B}(x,t,u) = 0,$$

где

$$\mathcal{B}(x,t,u) = V'_t(x,t) - f^0(x,u) + (V'_x(x,t), f(x,u)).$$

К уравнению Беллмана присоединяется граничное условие

$$V(x_1,t_1) = 0.$$

Обратим внимание на то, что в этой задаче управления функция Беллмана $V(x,t)$ зависит от двух переменных: $n-$мерной переменной x и одномерной переменной (времени) t. Напомним, что в задаче быстродействия функция Беллмана зависела только от $x \in I\!R^n \setminus \{x_1\}$. При рассмотрении задачи (P) с другими краевыми условиями в конце процесса граничное условие для функции Беллмана определенным образом модифицируется.

2.2 Погружение задачи (P) в семейство аналогичных задач (P_α). Введение функции Беллмана

Наряду с задачей (P) рассмотрим семейство задач управления

$$\left\{\begin{array}{ll} \dot{x} = f(x,u), & x \in I\!R^n, u \in U, \quad t \in [\tau, t_1], \\ x(\tau) = y, x(t_1) = x_1, & \\ L_\alpha = \int\limits_\tau^{t_1} f^0(x(t),u(t))\,dt \to \min\,. & \end{array}\right. \qquad (P_\alpha)$$

Семейство задач (P_α) зависит от параметра $\alpha = (y,\tau)$. Задача (P_α) при $\alpha = (x_0,t_0)$ превращается в задачу (P).

Пусть

$$T(y,\tau) = \min \int_\tau^{t_1} f^0(x(t),u(t))\,dt = \min L_\alpha$$

— минимальное значение функционала в задаче (P_α). Тогда

$$V(y,\tau) = -T(y,\tau)$$

— функция Беллмана.

2.3 Проверка принципа оптимальности ("хвост"оптимального процесса оптимален)

Пусть

$$(x(t),u(t)), \quad t_0 \le t \le t_1, \qquad (1)$$

— оптимальный процесс в задаче (P). Здесь $u(t)$ — допустимое управление, $x(t)$ — отвечающая этому управлению траектория, $x(t_0) = x_0$, $x(t_1) = x_1$. Возьмем любую точку $\tau \in [t_0,t_1)$ и рассмотрим "хвост"процесса (1)

$$(x(t),u(t)), \quad \tau \le t \le t_1, \qquad (2)$$

с начальным условием $x|_{t=\tau} = x(\tau) \equiv y$.

Проверим оптимальность пары (2) в задаче (P_α), где $\alpha = (y = x(\tau),\tau)$. Проверка выполняется методом от противного. Допустим, что это не так. Тогда существует допустимое управление

$$\hat{u}(t), \tau \le t \le t_1,$$

с лучшим значением функционала; отвечающую этому управлению траекторию обозначим

$$\hat{x}(t), \tau \le t \le t_1; \quad \hat{x}(t)|_{t=\tau} = x(\tau) = y, \hat{x}(t)|_{t=t_1} = x_1,$$

при этом

$$\int_\tau^{t_1} f^0(\hat{x},\hat{u})\,dt < \int_\tau^{t_1} f^0(x,u)\,dt.$$

Построим новое управление

$$u_*(t) = \left\{\begin{array}{ll} u(t), & t \in (t_0,\tau), \\ \hat{u}(t), & t \in [\tau,t_1], \end{array}\right.$$

2.4 Формулировка предположений о функции Беллмана. Вывод уравнения Беллмана. Необходимые условия оптимальности

2.4.1 Предположения 2.1, 2.2 о функции Беллмана

Сформулируем предположения о функции Беллмана $V(y, \tau)$.

Предполагаем, что функция Беллмана $V(y, \tau)$ определена и непрерывна на некотором множестве $W \subset I\!R^{n+1}$ значений (y, τ), $y \in I\!R^n$, $\tau \in [t_0, t_1]$, причем точка

$$(x_1, t_1) \in \partial W$$

— граничная точка множества W. Для функции V выполняется граничное условие

$$V(x_1, t_1) = 0.$$

Аргументы функции Беллмана будем обозначать далее буквами x и t.

Предположение 2.1. *Функция $V(x, t)$ определена и непрерывна на множестве $W \subset I\!R^n \times [t_0, t_1]$ с непустой внутренностью*

$$W^0 = \mathrm{int} W \neq \emptyset,$$

$$(x_1, t_1) \in \partial W, \quad V(x_1, t_1) = 0.$$

Предположение 2.2. *Функция Беллмана имеет в непустом открытом множестве W^0 непрерывные частные производные первого порядка*

$$V_x'(x, t), V_t'(x, t), \quad (x, t) \in W^0.$$

2.4.2 Необходимые условия оптимальности. Вывод уравнения Беллмана

Пусть

$$(x(t), u(t)), \quad t_0 \leq t \leq t_1,$$

— оптимальный процесс в задаче (P) и пусть выполняется включение

$$(x(t), t) \in W^0 \quad \forall t \in (t_0, t_1).$$

Обсудим *необходимые условия оптимальности.* Возьмем точку $t \in (t_0, t_1)$, и малое положительное число Δt, так что $t + \Delta t \in (t_0, t_1)$. Считаем, что t — точка непрерывности управления $u(t)$. На основании принципа оптимальности вдоль оптимального процесса имеем:

$$T(x(t), t) = \int_t^{t+\Delta t} f^0(x(s), u(s))\, ds + \int_{t+\Delta t}^{t_1} f^0(x(s), u(s))\, ds =$$

$$= \int_t^{t+\Delta t} f^0(x(s), u(s))\, ds + T(x(t+\Delta t), t+\Delta t).$$

Отсюда, привлекая функцию V, получаем равенство

$$\frac{V(x(t+\Delta t), t+\Delta t) - V(x(t), t)}{\Delta t} = \frac{1}{\Delta t} \int_t^{t+\Delta t} f^0(x(s), u(s))\, ds.$$

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

$$\frac{d}{dt}\, V(x(t), t) = f^0(x(t), u(t)),$$

или

$$(V_x'(x(t), t), \dot{x}(t)) + V_t'(x(t), t) = f^0(x(t), u(t)),$$

или

$$(V_x'(x(t), t), f(x(t), u(t))) + V_t'(x(t), t) = f^0(x(t), u(t)).$$

Таким образом, вдоль оптимального процесса выполняется равенство

$$(V_x'(x(t), t), f(x(t), u(t))) + V_t'(x(t), t) - f^0(x(t), u(t)) = 0 \quad \forall t \in (t_0, t_1). \tag{3}$$

Заметим, что условие (3) доказано для любой точки t непрерывности управления $u(t)$. Однако оно имеет место и в любой точке $\theta \in (t_0, t_1)$ конечного скачка управления. Действительно, принимая во внимание соглашение $u(\theta - 0) = u(\theta)$ о непрерывности слева допустимых управлений в точке разрыва и выполняя в (3) предельный переход $t \to \theta - 0$, получаем утверждение (3) в полном объеме. Итак, нами получены *необходимые условия оптимальности* в форме (3).

С помощью принципа оптимальности аналогичным образом (соответствующие выкладки приводятся ниже, см. также раздел "Метод динамического программирования в задаче быстродействия") получаем неравенство

$$V_t'(x,t) + (V_x'(x,t), f(x,u)) - f^0(x,u) \le 0 \qquad \forall (x,t) \in W^0, \forall u \in U. \tag{4}$$

Из (3) и (4) следует, что

$$\max_{u \in U} \left[V_t'(x,t) + (V_x'(x,t), f(x,u)) - f^0(x,u)\right] = 0, \quad (x,t) \in W^0, \tag{5}$$

или

$$V_t'(x,t) + \max_{u \in U} \left[-f^0(x,u) + (V_x'(x,t), f(x,u))\right] = 0, \qquad (x,t) \in W^0. \tag{6}$$

Последнее уравнение называется *уравнением Беллмана* . К этому уравнению присоединяется граничное условие [1]

$$V(x_1, t_1) = 0. \tag{7}$$

Таким образом, доказана

Теорема 2.1. *В Предположениях 2.1, 2.2 функция Беллмана $V(x,t)$ удовлетворяет условиям (6), (7). Для оптимальной пары $(x(t), u(t)), t_0 \le t \le t_1$, имеет место равенство (3).*

В заключение приведем обоснование неравенства (4). При любых $v \in U$, малом $\Delta t > 0$, $(x,t) \in W^0$, t — точка непрерывности управления $u(t)$, принцип оптимальности позволяет записать неравенство

$$T(x,t) \le \int_t^{t+\Delta t} f^0(y(s), v)\, ds + T(y(t+\Delta t), t+\Delta t),$$

где $y(s), t \le s \le t + \Delta t, y(s)|_{s=t} = x$, — траектория, отвечающая постоянному управлению v. При малых $\Delta t > 0$ будет $(y(t+\Delta t), t+\Delta t) \in W^0$. Записанное выше неравенство в терминах функции V принимает вид

$$\frac{V(y(t+\Delta t), t+\Delta t) - V(x,t)}{\Delta t} \le \frac{1}{\Delta t} \int_t^{t+\Delta t} f^0(y(s), v)\, ds \qquad \forall v \in U$$

и, в результате предельного перехода $\Delta t \to +0$, приводит к утверждению (4).

2.5 Уравнение Беллмана и достаточные условия оптимальности

Теорема 2.2. *Пусть*

1) $V(x,t)$ — гладкое решение уравнения Беллмана (6), удовлетворяющее условию (7),

2) вдоль допустимого процесса

$$(x(t), u(t)), t_0 \le t \le t_1, \quad x(t_0) = x_0, x(t_1) = x_1, \ (x(t), t) \in W^0, t_0 < t < t_1, \tag{8}$$

выполняется условие (3).

Тогда процесс (8) оптимален. Оптимальное значение функционала в задаче (P) равно $-V(x_0, t_0)$.

Доказательство теоремы 2.2. Нужно доказать, что значение функционала L для процесса (8)

$$L_{(8)} = \int_{t_0}^{t_1} f^0(x(t), u(t))\, dt$$

при условии (3) является минимальным. Возьмем любой другой допустимый процесс

$$(\hat{x}(t), \hat{u}(t)), t_0 \le t \le t_1, \quad \hat{x}(t_0) = x_0, \hat{x}(t_1) = x_1, \ (\hat{x}(t), t) \in W^0, t_0 < t < t_1, \tag{9}$$

[1] Условие (7) понимается в следующем смысле: для числа $\Delta > 0$ и любого допустимого процесса $(y(t), v(t)), t_1 - \Delta \le t \le t_1, y(t_1) = x_1; (y(t), t) \in W^0, t_1 - \Delta < t < t_1$;

$$\lim_{t \to t_1 - 0} V(y(t), t) = 0 \equiv V(x_1, t_1).$$

и запишем неравенство (4), полагая в нем $x = \hat{x}(t), u = \hat{u}(t)$:

$$V_t'(\hat{x}(t), t) + (V_x'(\hat{x}(t), t), f(\hat{x}(t), \hat{u}(t))) \leq f^0(\hat{x}(t), \hat{u}(t)).$$

Считая t точкой непрерывности управления $\hat{u}(t)$, последнее неравенство можно записать в форме

$$\frac{d}{dt}\ V(\hat{x}(t), t) \leq f^0(\hat{x}(t), \hat{u}(t)). \tag{10}$$

Проведем, наконец, заключительную выкладку:

$$L_{(8)} = \int\limits_{t_0}^{t_1} f^0(x(t), u(t))\, dt = \{(3)\} =$$

$$= \int\limits_{t_0}^{t_1} \frac{d}{dt}\ [V(x(t), t)]\ dt = \{\text{формула Ньютона-Лейбница}\} =$$

$$= V(x(t), t)|_{t=t_0}^{t=t_1} = V(x_1, t_1) - V(x_0, t_0) = \{(9)\} = V(\hat{x}(t), t)|_{t=t_0}^{t=t_1} =$$

$$= \{\text{формула Ньютона-Лейбница}\} = \int\limits_{t_0}^{t_1} \frac{d}{dt}\ [V(\hat{x}(t), t)]\ dt \leq \{(10)\} \leq$$

$$\leq \int\limits_{t_0}^{t_1} f^0(\hat{x}(t), \hat{u}(t))\, dt = L_{(9)} \equiv L(\hat{x}, \hat{u}).$$

Полученное неравенство доказывает оптимальность процесса (8). В силу (7) из предыдущей выкладки получаем, что оптимальное значение функционала $L_{(8)} = -V(x_0, t_0)$. Теорема 2.2 доказана.

Таким образом, метод динамического программирования применительно к задаче управления (P) рассмотрен как с точки зрения *необходимых условий оптимальности* , так и с точки зрения *достаточных условий оптимальности* .