Лекция № 13

Задачи оптимального управления

12 мая 2014

Содержательная постановка задачи оптимального управления

закон движения фазовой точки (самолета или объекта управления) и закон воздействия управления («рулей») записывается в виде системы дифференциальных уравнений:

$$\frac{dx^i}{dt} = f^i(x,u),\ i = 1,\dots,n,$$

или в векторной форме

$$\frac{dx}{dt} = f(x,u), \qquad (1)$$

где функции f^i непрерывны по переменным x и u, непрерывно дифференцируемы по переменной x. Здесь рассматривается случай, когда система (1) автономна, то есть правые ее части не зависят явно от времени t.

Содержательная постановка задачи оптимального управления

Рассмотрим произвольное допустимое управление $u(t)$. Перепишем уравнение (1) в следующем виде:

$$\frac{dx}{dt} = f(x, u(t)). \tag{2}$$

Тогда при любых начальных условиях $x(t_0) = x_0$ однозначно определяется траектория движения объекта $x = x(t)$, то есть решение этого уравнения, определенное на некотором отрезке времени.

Назовем его решением системы (2), соответствующим управлению $u(t)$ при начальном условии $x(t_0) = x_0$.

Содержательная постановка задачи оптимального управления

Будем говорить, что допустимое управление $u(t)$, $t_0 \le t \le t_1$ переводит фазовую точку x из положения x_0 в x_1, если решение $x(t)$ уравнения (2) с начальным условием $x(t_0) = x_0$ определено на $[t_0, t_1]$ и $x(t_1) = x_1$.

Такую пару $(x(t), u(t))$ назовем управляемым процессом, определенном на отрезке $[t_0, t_1]$.

Постановка задачи

Пусть задана еще одна функция f^0 непрерывная по переменным x и u, непрерывно дифференцируемая по переменной x. Приведем формальную постановку задачи оптимального управления.

Найти среди всех допустимых управлений, переводящих фазовую точку из положения x_0 в положение x_1, такое, для которого функционал

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

принимает наименьшее значение.

Заметим, что при заданных x_0 и x_1 пределы интегрирования t_0, t_1 являются переменными, которые зависят от управления, переводящего x_0 в x_1, и эти пределы определяются из соотношений $x(t_0) = x_0$, $x(t_1) = x_1$.

Определения и соглашения:

Управление $u(\cdot)$, на котором достигается оптимальное значение данной задачи, называется ***оптимальным управлением***, а соответствующая траектория $x(t)$ – ***оптимальной траекторией***.

В этом смысле основная задача – найти оптимальные управления и соответствующие оптимальные траектории, другими словами, найти ***оптимальный управляемый процесс***.

Для $J = t_1 - t_0$ оптимальность управления $u(t)$ эквивалентна минимизации времени перехода из положения x_0 в положение x_1. Задача отыскания оптимальных управлений и траекторий в этом случае называется ***задачей об оптимальном быстродействии***.

Принцип максимума для линейной задачи быстродействия:

Пусть $H(x,u,P)=(P,f(x,u))$ – ***функция Понтрягина***, а

$$\dot{P}_k=-\sum_{i=1}^{n}\frac{\partial f^i}{\partial x_k}(x(t),u(t))P_i,\ k=1,\dots,n, \qquad (3)$$

сопряженная система уравнений для соответствующей пары $(x(t),u(t))$.

Из линейности и однородности системы → (при любых начальных условиях для $P_k,\ k=1,\dots,n,$)

$\exists$ единственное решение этой системы (определенное на всем отрезке, на котором определены управление $u(t)$ и траектория $x(t)$).

Функции $P_1(t),\dots,P_n(t)$ непрерывны и имеют всюду, кроме конечного числа точек разрыва управления $u(t)$, непрерывные производные по t.

Теорема 1 (принцип максимума).

Пусть $((x_*(t), u_*(t)),\ t \in [t_0, t_1]$ – оптимальный управляемый процесс. Тогда существует ненулевая непрерывная вектор-функция $P(t) = (P_1(t), \ldots, P_n(t))$ такая, что справедливы следующие утверждения:

а) $\dot{P}_k = -\sum_{i=1}^{n} \frac{\partial f^i}{\partial x_k}(x_*(t), u_*(t)) P_i,\ k = 1, \ldots, n;$

б) $H(x_*(t), u_*(t), P(t)) = \max_{u \in U} H(x_*(t), u, P(t)),\ t \in [t_0, t_1];$

в) $H(x_*(t_1), u_*(t_1), P(t_1)) \geq 0.$

Соглашения:

Пусть f линейна $\longrightarrow$ система (1) записывается:

$$\text{в виде } \dot{x} = Ax + Bu \quad \text{или} \quad \dot{x}^i = \sum_{\alpha=1}^{k} a^i_\alpha x_\alpha + \sum_{\beta=1}^{r} b^i_\beta u_\beta .$$

В дальнейшем предполагается, что U – выпуклый многогранник в R^r,

$0 \in U$, и 0 не является вершиной U.

Теорема 2 (принцип максимума для линейной задачи быстродействия).

Пусть $(x_*(t), u_*(t))$, $t \in [t_0, t_1]$ – оптимальный управляемый процесс. Тогда существует такое непрерывное нетривиальное решение $P(t)$ сопряженной системы $\dot{P} = -PA$, что справедливо

$$P(\tau)Bu_*(\tau) = \max_{u \in U} P(\tau)Bu, \ \tau \in [t_0, t_1]. \tag{4}$$

Управление $u_*(t)$ удовлетворяет принципу максимума, если существует нетривиальное решение сопряженной системы (3) и выполняется равенство (4).

Задача синтеза оптимальных управлений

Пример:

Рассмотрим уравнение $\dfrac{d^2x}{dt^2} = u$, где u – вещественный управляющий параметр, удовлетворяющий ограничению $|u| \le 1$. В фазовых координатах $x^1 = x$, $x^2 = \dfrac{dx}{dt}$ это уравнение переписывается в виде следующей системы:

$$\frac{dx^1}{dt} = x^2, \frac{dx^2}{dt} = u. \tag{5}$$

Задача синтеза оптимальных управлений

Рассмотрим для фазовой точки, движущейся по закону (5), задачу о наискорейшем попадании в начало координат $x_1 = (0,0)$ из заданного начального состояния x_0.

Функция H в данном случае имеет вид

$$H = \psi_1 x^2 + \psi_2 u. \qquad (6)$$

Далее, для вспомогательных переменных ψ_1, ψ_2 получается система уравнений (см. (3), (6))

$$\frac{d\psi_1}{dt} = 0, \quad \frac{d\psi_2}{dt} = -\psi_1,$$

откуда $\psi_1 = c_1$, $\psi_2 = c_2 - c_1 t$, где c_1, c_2 – постоянные.

Задача синтеза оптимальных управлений

С учетом (6) и условия $|u| \le 1$, из соотношения (4) следует

$$u(t) = sign\,\psi_2(t) = sign\,(c_2 - c_1 t). \tag{7}$$

Откуда получим, что каждое оптимальное управление $u(t)$, $t_0 \le t \le t_1$, является кусочно-постоянной функцией, принимающей значения ± 1 и имеющей не более двух интервалов постоянства, так как линейная функция $c_2 - c_1 t$ не более одного раза меняет знак на отрезке $[t_0, t_1]$.

Обратно, любая такая функция $u(t)$ может быть получена из соотношения (7) при некоторых значениях постоянных c_1, c_2.

Задача синтеза оптимальных управлений

Для отрезка времени, на котором $u \equiv 1$, в силу системы (5) справедливо

$$x^2 = t + s_2,\; x^1 = \frac{t^2}{2} + s_2 t + s_1 = \frac{1}{2}(t + s_2)^2 + \left(s_1 - \frac{s_2{}^2}{2}\right),$$

где s_1, s_2 – постоянные интегрирования, откуда следует

$$x^1 = \frac{1}{2}(x^2)^2 + s, \tag{8}$$

где $s = s_1 - \frac{1}{2}s_2{}^2$ – постоянная. Таким образом, часть фазовой траектории, для которой $u \equiv 1$, представляет собой дугу параболы (8).

Семейство парабол (8) показано на рис. 1.

Задача синтеза оптимальных управлений

Аналогично, для отрезка времени, на котором $u \equiv -1$, имеем

$$x^2 = -t + s_2',$$

$$x^1 = -\frac{t^2}{2} + s_2' t + s_1' = -\frac{1}{2}(-t + s_2')^2 + \left(s_1' + \frac{1}{2}(s_2')^2 \right),$$

откуда получим

$$x^1 = -\frac{1}{2}(x^2)^2 + s'. \qquad (9)$$

Семейство парабол (9) показано на рис. 2. По параболам (8) фазовые точки движутся снизу вверх, так как $\dfrac{dx^2}{dt} = u = +1$, а по параболам (9) – сверху вниз, так как $\dfrac{dx^2}{dt} = -1$.

Задача синтеза оптимальных управлений

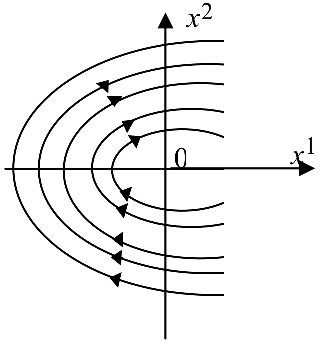

Рис. 1

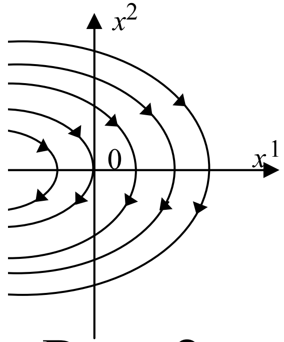

Рис. 2

Если управление $u(t)$ в течение некоторого времени равно $+1$, а затем равно -1, то фазовая траектория состоит из частей двух парабол (рис. 3), примыкающих друг к другу, причем одна из этих частей лежит на той из парабол (9), которая проходит через начало координат, так как искомая траектория должна вести в начало координат. Если же, наоборот, сначала $u = -1$, а затем $u = +1$, то фазовая кривая заменяется центрально симметричной (рис. 4).

Задача синтеза оптимальных управлений

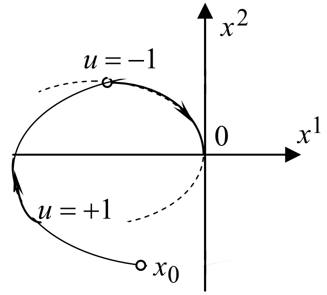

Рис. 3

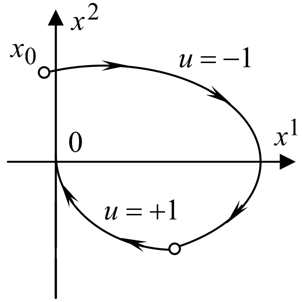

Рис. 4

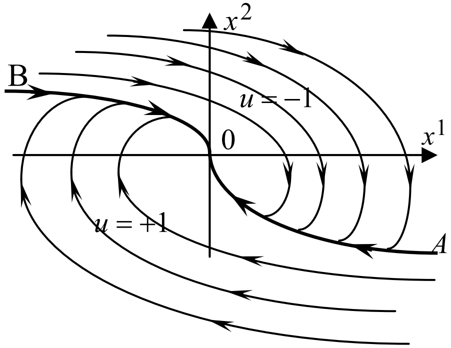

Рис. 5

На рис. 3, 4 на дугах парабол надписаны соответствующие значения управляющего параметра u. На рис. 5 изображено все семейство полученных таким образом фазовых траекторий (AO – дуга параболы

Задача синтеза оптимальных управлений

$x^1 = \frac{1}{2}(x^2)^2$, расположенная в нижней полуплоскости; BO – дуга параболы $x^1 = -\frac{1}{2}(x^2)^2$, расположенная в верхней полуплоскости).

Итак, если начальное положение x_0 расположено выше линии AOB, то фазовая точка должна двигаться под воздействием управления $u = -1$ до тех пор, пока она не попадет на дугу AO; в момент попадания на дугу AO значение u переключается и становится равным $+1$ вплоть до момента попадания в начало координат.

Если же начальное положение x_0 расположено ниже линии AOB, то u должно быть равно $+1$ до момента попадания на дугу BO, а в момент попадания на дугу BO значение u переключается и становится равным -1.

Задача синтеза оптимальных управлений

Итак, согласно теореме 2, только описанные выше траектории могут быть оптимальными, причем из проведенного исследования видно, что из каждой точки фазовой плоскости исходит только одна траектория, ведущая в начало координат, которая может быть оптимальной, то есть задание начальной точки x_0 однозначно определяет соответствующую траекторию.

Из теорем существования $\longrightarrow$ в данном примере для любой начальной точки x_0 существует оптимальная траектория. Таким образом, найденные траектории (рис. 5) являются оптимальными, и других оптимальных траекторий, ведущих в начало координат, не существует.

Задача синтеза оптимальных управлений

Полученное в рассмотренном примере решение оптимальной задачи можно истолковать следующим образом. Обозначим через $v(x^1, x^2) = v(x)$ функцию, заданную на плоскости x^1, x^2:

$$v(x) = \begin{cases} +1 \text{ ниже линии } AOB \text{ и на дуге } AO, \\ -1 \text{ выше линии } AOB \text{ и на дуге } BO. \end{cases}$$

Тогда на каждой оптимальной траектории значение $u(t)$ управляющего параметра в произвольный момент t равно $v(x(t))$, то есть равно значению функции v в той точке, в которой в момент t находится фазовая точка, пробегающая оптимальную траекторию $u(t) = v(x(t))$. Это означает, что, заменив в системе (5) величину u функцией $v(x)$, получим систему

$$
\begin{cases}
\dfrac{dx^1}{dt} = x^2, \\
\dfrac{dx^2}{dt} = v(x^1, x^2),
\end{cases}
\qquad (10)
$$

решение которой при произвольном начальном состоянии x_0 дает оптимальную фазовую траекторию, ведущую в начало координат. Иначе говоря, система (10) представляет собой систему дифференциальных уравнений с разрывной правой частью для нахождения оптимальных траекторий, ведущих в начало координат.

Доказательство принципа максимума для линейной задачи быстродействия

Пусть $T > 0$ – верхняя граница на длины интервалов, на которых будут рассматриваться управления. Будем говорить, что точка $\bar{x}$ принадлежит сфере достижимости (см. рис. 6), если на интервале $[t_0, t_1]$ существует допустимое управление $u(t)$ и соответствующая ему траектория $x(t)$ такие, что $x(t_0) = \bar{x}$, $x(t_1) = 0$, $t_1 - t_0 \le T$.

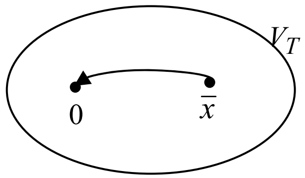

Рис. 6

Доказательство принципа максимума для линейной задачи быстродействия

Лемма 1. Сфера достижимости V_T является выпуклым множеством.

Лемма 2. Если x_0 – внутренняя точка V_T, то из x_0 можно перейти в точку 0 за время строго меньше T.

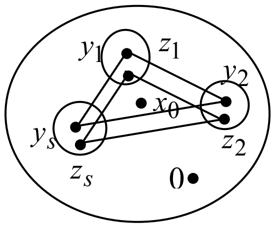

Рис. 8

Доказательство принципа максимума для линейной задачи быстродействия

Лемма 3. Пусть $u(t)$ – допустимое управление на интервале $[t_0, t_1]$, $x(t)$ – соответствующее решение, $P(t)$ – произвольное решение сопряженной системы $\dot{P} = -PA$ на данном интервале. Тогда во всех точках непрерывности управления $u(t)$ справедливы следующие равенства:

$$\frac{d}{dt}(P(t)x(t)) = P(t)Bu(t),$$

$$P(t_1)x(t_1) - P(t_0)x(t_0) = \int_{t_0}^{t_1} P(t)Bu(t)dt.$$

Доказательство принципа максимума для линейной задачи быстродействия

Доказательство.

$$\frac{d}{dt}(P(t)x(t)) = \dot{P}(t)x(t) + P(t)\dot{x}(t) = -P(t)Ax(t) + P(t)(Ax(t) + Bu(t)) =$$
$$= P(t)Bu(t). \blacksquare$$