Глава 12

Методы оптимального управления

12.1. Постановка задач оптимального управления

Рассмотрим систему с n степенями свободы, движение которой описывают уравнения Лагранжа второго рода

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = Q_j, \quad j = 1, 2, \ldots, n, \tag{1}$$

где T — кинетическая энергия, q_j — обобщенные координаты, Q_j — обобщенные силы.

В механике имеются два вида задач:

1) прямая задача, состоящая в определении обобщенных сил Q_j по заданным функциям $q_j(t)$;

2) обратная задача, в которой надо найти закон движения, зная обобщенные силы Q_j и начальные условия.

В теории управления решаются задачи третьего типа. Предполагается, что $Q_j = P_j + U_j$, где P_j — заданные силы, U_j — искомые управляющие силы. Задача управления состоит в определении функций U_j обеспечивающих переход системы из начального положения $(q_j^0, \dot{q}_j^0)$ в заданное конечное положение $(q_j^1, \dot{q}_j^1)$.

Если не задано никаких дополнительных условий, то задача управления имеет, как правило, бесчисленное множество решений. Естественным условием является предположение об ограниченности управляющих сил. Однако и при выполнении этого условия число решений остается бесконечным.

Конечное число решений, и в частности единственное решение, задача управления имеет в том случае, когда требуется, чтобы управление в каком-то смысле было оптимальным. Например, можно потребовать, чтобы переход системы из начального положения в конечное совершался за наименьшее возможное время. В задачах о космических перелетах оптимальным может быть управление, для которого требуется минимальный расход ракетного топлива. Примером такой задачи является перелет спутника с одной круговой орбиты на другую по эллипсу Гомана.

Рассмотрим систему диффернциальных уравнений первого порядка

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{u}), \tag{2}$$

где $\mathbf{x} = (x_1, x_2, \ldots, x_m)^T$, $\mathbf{u} = (u_1, u_2, \ldots, u_p)^T$, $\mathbf{f} = (f_1, f_2, \ldots, f_m)^T$. Переменные x_j называются фазовыми координатами, а переменные u_j — управлениями. Вектор $\mathbf{x}$ принадлежит m-мерному фазовому пространству R^m, а вектор $\mathbf{u}$ — p-мерному пространству управлений R^p. Уравнения Лагранжа (1) сводятся к системе (2) с $m = 2n$.

Пусть в пространстве R^m заданы точки $\mathbf{x}^0$ и $\mathbf{x}^1$. Вектор $\mathbf{u}(t)$ называется управлением, переводящим систему (2) из положения $\mathbf{x}^0$ в положение $\mathbf{x}^1$, если существует решение $\mathbf{x}(t)$ системы (2) такое, что $\mathbf{x}(t_0) = \mathbf{x}^0$, $\mathbf{x}(t_1) = \mathbf{x}^1$.

Обычно предполагается, что $\mathbf{u} \in U \subset R^p$, где U — область допустимых управлений. В частности, если управляющие силы ограничены, т. е. $|u_k| \leqslant C_k$, $k = 1.2, \ldots p$, то областью управлений будет прямоугольный параллелепипед в пространстве R^p.

Условия оптимальности управления, как правило, сводятся к требованию, чтобы некоторый функционал

$$I = \int_{t_0}^{t_1} F(\mathbf{x}, \mathbf{u})\, dt$$

имел минимальное значения для искомого оптимального управления. В случае $F = 1$ получаем $I = t_1 - t_0$, и задача оптимального управления называется задачей об оптимальном быстродействии. Этот частный случай является весьма важным, так как формально общий случай сводится к нему заменой переменной интегрирования $d\tau = F dt$ в интеграле I. Действительно, в результате такой замены мы получаем

$$I = \int_{\tau_0}^{\tau_1} d\tau = \tau_1 - \tau_0.$$

В дальнейшем мы ограничимся исследованием задач об оптимальном быстродействии. Управление $\mathbf{u}^*$, переводящее систему (2) из положения $\mathbf{x}^0$ в положение $\mathbf{x}^1$ будем называть оптимальным, если для него время перехода $T(t) = t_1 - t_0$ имеет наименьшую величину. Решение $\mathbf{x}^*$ системы

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{u}^*) \tag{3}$$

называют оптимальной траекторией, а пару вектор-функций $(\mathbf{x}^*, \mathbf{u}^*)$ — оптимальным процессом.

12.2. Уравнение Беллмана

Зафиксируем точку $\mathbf{x}^1$ фазового пространства R^m и обозначим $T(\mathbf{x})$ время оптимального перехода из произвольной точки $\mathbf{x} \in R^m$ в точку $\mathbf{x}^1$. Предположим, что функция $T(\mathbf{x})$ определена для любой точки фазового пространства.

Пусть время перехода из точки $\mathbf{x}$ в точку $\mathbf{x}^0$ под действием произвольного управления $\mathbf{u} \in U$ равно $t_0 - t$. Переведем теперь систему из положения $\mathbf{x}^0$ в положение $\mathbf{x}^1$ по оптимальной траектории. Тогда время перехода из $\mathbf{x}$ в $\mathbf{x}^1$ будет равно $t_0 - t + T(\mathbf{x}^0)$ (см. диаграмму).

$$\begin{array}{ccccccc} & & & T(\mathbf{x}) & & & \\ & \mathbf{x} & & \longrightarrow & & \mathbf{x}^1 & \\ t_0 - t & & \searrow & & \nearrow & & T(\mathbf{x}^0) \\ & & & \mathbf{x}^0 & & & \end{array}$$

Учитывая, что наименьшее возможное время перехода из точки $\mathbf{x}$ в точку $\mathbf{x}^1$ равно $T(\mathbf{x})$, получаем

$$T(\mathbf{x}) \leqslant t_0 - t + T(\mathbf{x}^0).$$

Введем функцию $\omega(\mathbf{x}) = -T(\mathbf{x})$. Тогда

$$\frac{\omega(\mathbf{x}^0) - \omega(\mathbf{x})}{t_0 - t} \leqslant 1.$$

Переходя к пределу при $t_0 - t \to 0$, получим

$$\frac{d\omega(\mathbf{x})}{dt} \leqslant 1.$$

По правилу дифференцирования сложной функции

$$\frac{d\omega(\mathbf{x})}{dt} = \sum_{i=1}^{m} \frac{\partial \omega}{\partial x}_i \frac{dx_i}{dt}.$$

Вектор-функция $\mathbf{x}(t)$ является решение системы (2), поэтому

$$\frac{dx_i}{dt} = f_i(\mathbf{x}, \mathbf{u}).$$

Следовательно, для произвольных $\mathbf{x}$ и $\mathbf{u}$ имеет место неравенство

$$\frac{d\omega(\mathbf{x})}{dt} = B(\mathbf{x}, \mathbf{u}) = \sum_{i=1}^{m} \frac{\partial \omega}{\partial x}_i f_i(\mathbf{x}, \mathbf{u}) \leqslant 1. \tag{4}$$

Рассмотрим теперь переход из положения $\mathbf{x}^0$ в положение $\mathbf{x}^1$ по оптимальной траектории $\mathbf{x}^*$ за время $T(\mathbf{x}^0) = t_1 - t_0$. Пусть $t \in (t_0, t_1)$. Тогда

$$T(\mathbf{x}^0) = (t - t_0) + (t_1 - t).$$

Покажем, что $t_1 - t = T(\mathbf{x}^*(t))$, т. е. $t_1 - t$ является наименьшим временем перехода из точки $\mathbf{x}^*(t)$ в точку $\mathbf{x}^1$ (см. диаграмму).

$$\mathbf{x}^0 \xrightarrow{t - t_0} \mathbf{x}^*(t) \xrightarrow{t_1 - t} \mathbf{x}^1$$

Действительно, пусть $\mathbf{u}$ — управление, переводящее из положения $\mathbf{x}^*(t)$ в положение $\mathbf{x}^1$ за время $T_1 < t_1 - t$. Тогда имеется переход их $\mathbf{x}^0$ в $\mathbf{x}^1$ за время $T_1 + (t - t_0) < T(\mathbf{x}^0)$, что противоречит определению $T(\mathbf{x}^0)$. Следовательно,

$$T(\mathbf{x}^0) = (t - t_0) + T(\mathbf{x}^*(t)).$$

Дифференцирование последнего равенства по времени t дает формулу

$$\frac{dT(\mathbf{x}^*(t))}{dt} = -1.$$

Подставим в нее $\omega = -T$. Тогда

$$\frac{d\omega(\mathbf{x}^*)}{dt} = 1.$$

Используя правило дифференцирования сложной функции и учитывая, что вектор-функция $\mathbf{x}^*(t)$ является решение системы (3), получаем

$$B(\mathbf{x}^*, \mathbf{u}^*) = \sum_{i=1}^{m} \frac{\partial\omega}{\partial x}_i f_i(\mathbf{x}^*, \mathbf{u}^*) = 1. \tag{5}$$

Из формул (4) и (5) вытекает, что

$$\max_{\mathbf{u}\in U} B(\mathbf{x}, \mathbf{u}) = \max_{\mathbf{u}\in U} \sum_{i=1}^{m} \frac{\partial\omega}{\partial x}_i f_i(\mathbf{x}, \mathbf{u}) = 1, \tag{6}$$

причем максимум достигается при $\mathbf{u} = \mathbf{u}^*$. Уравнение (6) называется уравнением Беллмана.

12.3. Принцип максимума

Используя уравнение Беллмана, можно получить принцип максимума, который позволяет найти оптимальные процессы при решении задач управления.

Функция $B(\mathbf{x}, \mathbf{u})$ имеет максимум для оптимального процесса, поэтому при $\mathbf{x} = \mathbf{x}^*$, $\mathbf{u} = \mathbf{u}^*$

$$\frac{\partial B}{\partial x_k} = \sum_{i=1}^{m} \frac{\partial^2\omega}{\partial x_k \partial x_i} f_i + \sum_{i=1}^{m} \frac{\partial\omega}{\partial x_i}\frac{\partial f_i}{\partial x_k} = 0, \quad k = 1, 2, \dots m. \tag{7}$$

Учитывая, что

$$\frac{d}{dt}\left(\frac{\partial\omega}{\partial x_k}\right) = \sum_{i=1}^{m} \frac{\partial^2\omega}{\partial x_i \partial x_k}\frac{dx_i}{dt} = \sum_{i=1}^{m} \frac{\partial^2\omega}{\partial x_i \partial x_k} f_i,$$

и предполагая непрерывность вторых частных производных функции ω по переменным x_i, равенство (7) запишем в виде

$$\frac{d}{dt}\left(\frac{\partial\omega}{\partial x_k}\right) + \sum_{i=1}^{m} \frac{\partial\omega}{\partial x_i}\frac{\partial f_i}{\partial x_k} = 0.$$

Обозначим $\Psi_k = \partial\omega/\partial x_k$. Тогда

$$\frac{d\Psi_k}{dt} + \sum_{i=1}^{m} \Psi_i \frac{\partial f_i}{\partial x_k} = 0. \tag{8}$$

Для функции

$$B = \sum_{i=1}^{m} \Psi_i f_i(\mathbf{x}, \mathbf{u})$$

как функции трех векторных переменных $\mathbf{x}$, $\mathbf{u}$ и $\boldsymbol{\Psi} = (\Psi_1, \Psi_2, \dots, \Psi_m)^T$ введем новое обозначение

$$H(\mathbf{x}, \mathbf{u}, \boldsymbol{\Psi}) = \sum_{i=1}^{m} \Psi_i f_i(\mathbf{x}, \mathbf{u}).$$

Тогда соотношение (8) можно представить в виде

$$\frac{d\Psi_k}{dt} = -\frac{\partial H}{\partial x_k}, \quad k = 1, 2, \dots m.$$

Принцип максимума

Если $(\mathbf{x}^*, \mathbf{u}^*)$ — оптимальный процесс для системы уравнений (2), то существует неривиальное решение Ψ_1, Ψ_2,…,Ψ_m уравнений

$$\frac{d\Psi_k}{dt} = -\frac{\partial H}{\partial x_k}, \quad H = \sum_{i=1}^{m} \Psi_i f_i$$

такое, что

$$\max_{\mathbf{u}\in U} H(\mathbf{x}, \mathbf{u}, \mathbf{\Psi}) = H(\mathbf{x}^*, \mathbf{u}^*, \mathbf{\Psi}) = 1.$$

Проинцип максимума можно использовать для отыскания оптимальных процессов, однако он является необходимым, но не достаточным условием оптимальности управления. Если найден процесс, удовлетворяющий принципу максимума, то его оптимальность еще нужно доказать.

Сформулированный принцип максимума годится только для непрерывных управлений. Однако, в целом ряде прикладных задач оптимальными оказываются кусочно-непрерывные управления, имеющие конечное число разрывов первого рода.

Принцип максимума Понрягина

Пусть управления $u_k(t)$, $k = 1, 2, \ldots, p$ являются кусочно-непрерывными функциями, заданными в интервале $[t_0, t_1]$. Для оптимальности процеса $(\mathbf{x}^*, \mathbf{u}^*)$ в смысле быстродействия необходимо существование такого нетривиального решения системы

$$\frac{d\Psi_k}{dt} = -\frac{\partial H}{\partial x_k}, \quad H = \sum_{i=1}^{m} \Psi_i f_i,$$

что во всякой точке непрерывности управления $\mathbf{u}^*(t)$ выполняется условие

$$H(\mathbf{x}^*, \mathbf{u}^*, \mathbf{\Psi}) = \max_{\mathbf{u}\in U} H(\mathbf{x}, \mathbf{u}, \mathbf{\Psi}),$$

а в конечный момент времени t_1

$$H(\mathbf{x}^*(t_1), \mathbf{u}^*(t_1), \mathbf{\Psi}(t_1)) \geqslant 0.$$

Этот принцип был сформулирован в качестве гипотезы Л.С. Понтрягиным. Его доказательство принадлежит В.Г. Болтянскому и приведено в его книге [1].

12.4. Пример применения принципа максимума

Рассмотрим прямолинейное движение материальной точки массой m по оси Ox под действием управляющей силы $F(t)$ такой, что $|F| \leqslant F_0$. Проектируя уравнение второго закона Ньютона на ось Ox, получим

$$m\frac{d^2x}{d\tau^2} = F(\tau), \tag{9}$$

где τ — время. Введем безразмерные переменные z, t по формулам $x = Lz$, $\tau = Tt$ и выберем постоянные L, T так, чтобы

$$\frac{F_0T^2}{mL} = 1.$$

В безразмерных переменных уравнение (9) принимает вид

$$\ddot{z} = u(t), \tag{10}$$

где

$$u = \frac{FT^2}{mL}, \quad |u| \leqslant 1 \quad \dot{z} = \frac{dz}{dt}.$$

Обозначим $x_1 = z$, $x_2 = \dot{z}$ и сведем уравнение (10) к эквивалентной системе

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = u, \quad -1 \leqslant u \leqslant 1. \tag{11}$$

Переменные x_1 и x_2 задают положение точки и ее скорость.

Найдем управление $u(t)$ под действием которого точка из заданного начального состояния x_1^0, x_2^0 попадет в положение $x_1^1 = x_2^1 = 0$ за наименьшее время. Для решения этой задачи используем принцип максимума Понтрягина. Функция H для рассматриваемого примера имеет вид

$$H = \Psi_1 f_1 + \Psi_2 f_2 = \Psi_1 x_2 + \Psi_2 u.$$

Решив систему уравнений

$$\dot{\Psi}_1 = -\frac{\partial H}{\partial x_1} = 0, \quad \dot{\Psi}_2 = -\frac{\partial H}{\partial x_2} = -\Psi_1,$$

получим

$$\Psi_1 = D_1, \quad \Psi_2 = -D_1 t + D_2,$$

где D_1 и D_2 — произвольные постоянные.

Если $\Psi_2 > 0$, то H имеет максимальное значение при $u = 1$. В противном случае максимум достигается при $u = -1$. Следовательно, оптимальное управление

$$u_* = \begin{cases} 1, & \Psi_2 > 0, \\ -1, & \Psi_2 < 0. \end{cases}$$

При изменении t функция Ψ_2 меняет знак не более одного раза, поэтому управление u_* имеет не более двух интервалов постоянства. Возможные четыре типа оптимальных управлений представлены на рис. 12.1.

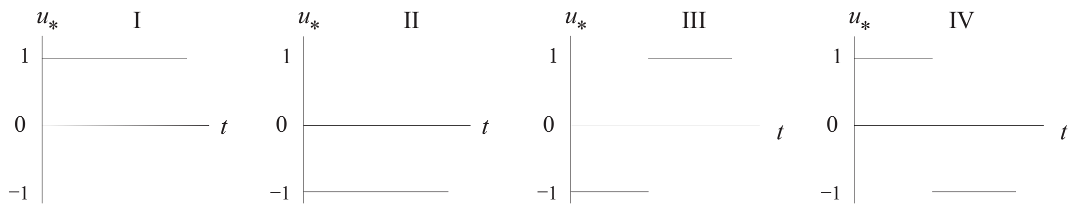

Рис. 12.1. Типы управлений.

Тип оптимального управления зависит от начального положения x_1^0, x_2^0. Для того, чтобы определить эту зависимость найдем оптимальные траектории для управлений $u = 1$ и $u = -1$. При $u = 1$ система (11) принимает вид

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = 1,$$

и имеет общее решение

$$x_2 = t + C_2, \quad x_1 = (t + C_2)^2/2 + C_1,$$

где C_1 и C_2 — призвольные постоянные. Оптимальные фазовые траектории задаются уравнением

$$x_1 = x_2^2/2 + C_1.$$

Они являются параболами на фазовой плоскости (x_1, x_2) и изображены на рис. 12.2.

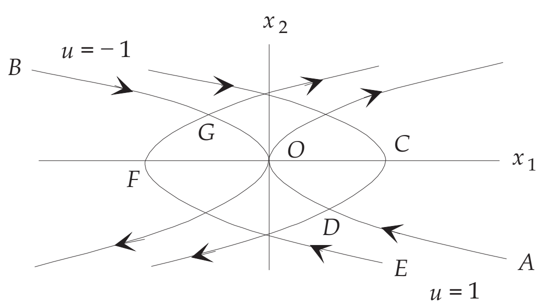

Рис. 12.2. Оптимальные фазовые траектории.

Направление движения изображающей точки (x_1, x_2) по фазовой траектории показано стрелками на рис. 12.2. Оно определяется с помощью формулы $\dot{x}_1 = x_2$. При $x_2(t) > 0$ функция $x_1(t)$ возрастает, а при $x_2(t) < 0$ убывает.

Для управления $u = -1$ имеем систему уравнений

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = -1,$$

фазовые траетории которой определяются по формуле

$$x_1 = -x_2^2/2 + C_1$$

и являются зеркальными отражения относительно вертикали фазовых траекторий в случае $u = 1$ (см. рис 12.2).

Если начальная точка $M_0(x_1^0, x_2^0)$ лежит на кривой AO, и, в частности, $M_0 = A$, то управление $u = 1$ типа I переведет ее в точку $O(0, 0)$ за наименьшее время. Начальную точку M_0, лежащую на кривой BO, за наименьшее время переведет в точку O управление $u = -1$ типа II.

Для оптимального перевода в начало координат точки M_0, лежащей выше кривой AB, следует использовать управление типа III. Так, например, в случае $M_0 = C$ движение изображающей точки $M(x_1, x_2)$ сначала происходит по кривой CD под действием управления $u = -1$, а затем по кривой DO под действием управления $u = 1$.

Точка M_0, лежащая ниже кривой AB перейдет в точку O за наименьшее время под действием управления типа IV. Если, например, $M_0 = E$, то на кривой EG будет действовать управление $u = 1$, а на кривой GO — управление $u = -1$.

Движению изображающей точки $M(x_1, x_2)$ по фазовой траектории соответствует движение материальной точки по прямой. При движении точки M по кривым AO и BO происходит максимально возможное торможение материальной точки, так как сила и скорость направлены в разные стороны. На участке траектории CD происходит максимально возможный разгон материальной точки, а на участке DO — максимально возможное торможение. При движения точки M по кривой EG скорость материальной точки меняет знак в точке F, поэтому одно и то же управление $u = 1$ на участке EF уменьшает абсолютную величину скорости, а на участке FG ее увеличивает. Наконец, на участке GO снова происходит торможение, необходимое для того чтобы материальная точка попала в начало координат с нулевой скоростью.

Доказательство оптимальности полученных траекторий имеется в книге [1].

ЛИТЕРАТУРА

1. *Болтянский В.Г.* Математические методы оптимального управления. М., Наука, 1966.

Оглавление