ОБ ОДНОМ СПОСОБЕ ВЫВОДА УРАВНЕНИЯ БЕЛЛМАНА

К. А. Лурье (Ленинград)

Будем рассматривать следующую оптимальную задачу. Дана система обыкновенных дифференциальных уравнений и начальных условий

$$\frac{dx^i}{dt} = g^i(t;\ x,\ u), \qquad x^i(t_0) = x_0{}^i, \qquad i = 1, \ldots, n \tag{1}$$

Здесь $u(t) = (u^1(t), \ldots, u^m(t))$ — управляющая вектор-функция.

Кроме того, задана система связей, выражаемых конечными равенствами [1]

$$R^j(t;\ x,\ u) = 0, \quad j = 1, \ldots, r \leqslant m \tag{2}$$

Предполагается, что матрица $\|\partial R^j / \partial u^i\|$ имеет максимальный ранг.

Требуется определить управляющую функцию $u(t)$, доставляющую минимум функционалу $J = S_T(x(T), T)$, имеющему смысл функции конечной точки $(x(T), T)$ фазовой траектории в пространстве (x, t):

$$S_T(x(T), T) = F(x^1(T), \ldots, x^n(T), T) \tag{3}$$

Как известно, поставленная задача может быть решена при помощи дифференциального уравнения Беллмана [1] (принято обычное условие суммирования)

$$\frac{\partial S_T}{\partial t} = \max_u \left[-\frac{\partial S_T}{\partial x^i} g^i(t;\ x,\ u) \right], \qquad u \in A(t,\ x) \tag{4}$$

Здесь $A(t, x)$ обозначает множество допустимых управлений, т. е. управлений, удовлетворяющих условию (2). Решая уравнение (4) при граничном условии (3), находим функцию $S_T(x, t)$; подстановка ее в выражение

$$u = \gamma(x,\ t,\ \partial S_T / \partial x^i) \tag{5}$$

получающееся из условия максимума

$$\frac{\partial}{\partial u} \left[\frac{\partial S_T}{\partial x^i} g^i(t;\ x,\ u) + \Gamma_j R^j(t;\ x,\ u) \right] = 0 \tag{6}$$

приводит к соотношению

$$u = \varphi(x,\ t) \tag{7}$$

определяющему синтез оптимального управления. Решения системы уравнений

$$dx^i/dt = g^i(t;\ x,\ \varphi(x,\ t)) \tag{8}$$

минимизируют функционал (3) при произвольных начальных данных $x^i(t_0) = x^i{}_0$; связи (2) при этом выполняются тождественно $R^j(t;\ x, \varphi(x, t)) \equiv 0$.

Рассуждение, приводящее к уравнению Беллмана (4), хорошо известно. По существу уравнение (4) есть не что иное, как дифференциальная формулировка принципа оптимальности.

Решение уравнения (4) при данных граничных условиях дает синтезирующую функцию (7); подставляя ее вместо u в правую часть (4), получим

$$\frac{\partial S_T}{\partial t} = -\frac{\partial S_T}{\partial x^i} g^i(t; x,\ \varphi(x,\ t)) \tag{9}$$

Вернемся теперь к поставленной в самом начале задаче и возьмем какое-нибудь допустимое управление $U(x, t)$. Заметим, что функция U выбрана зависящей от двух переменных x, t, которые рассматриваются как независимые.

Подставляя $U(x, t)$ вместо u в систему (1), придем к дифференциальным уравнениям допустимых траекторий; при заданном начальном состоянии (x_0, t_0) допустимая траектория определена единственным образом.

---

[1] Случай связей, заданных конечными неравенствами, легко приводится к рассматриваемому путем введения «вспомогательных управлений».

Рассмотрим функционал $S_T(x_0, t_0)$ на допустимой траектории; его значение определяется лишь конечным (в момент $t = T$) положением фазовой точки, поэтому это значение не зависит от того, какая точка на допустимой траектории будет выбрана в качестве начальной; это обстоятельство выражается равенством

$$\frac{dS_T(x_0, t_0)}{dt_0} = 0 \tag{10}$$

справедливым вдоль допустимой траектории.

Выполняя дифференцирование в (10) и учитывая (1), получим

$$\frac{\partial S_T(x_0, t_0)}{\partial t_0} = -\frac{\partial S_T(x_0, t_0)}{\partial x_0{}^i} g^i(t_0; x_0, U(x_0, t_0))$$

или, отбрасывая нули,

$$\frac{\partial S_T(x, t)}{\partial t} = -\frac{\partial S_T(x, t)}{\partial x^i} g^i(t; x, U(x, t)) \tag{11}$$

До сих пор (11) рассматривалось лишь вдоль допустимой траектории; заметим теперь, что в качестве допустимой может быть выбрана любая траектория системы (1) (где вместо u подставлено $U(x, t)$), так что $S_T(x_0, t_0)$, согласно (10), сохраняется вдоль всякой траектории системы (1). Но это означает [2], что $S_T(x, t)$ как функция двух переменных (x, t) удовлетворяет уравнению (11), рассматриваемому как уравнение в частных производных первого порядка.

Необходимо еще раз подчеркнуть, что аргументом функции $S_T(x, t)$ являются координаты начальной точки допустимой траектории, а сама функция S определяется как некоторая комбинация координат конечной точки траектории. При выбранном управлении значения $S_T(x, t)$ сохраняются при движении вдоль траектории, так что, в частности, минимум значения $S_T(x(T), T)$ совпадает с минимумом значения $S_T(x_0, t_0)$. Но величины $x(T)$ не известны, а x_0, t_0 известны. Поэтому можно вместо $S_T(x,(T), T)$ минимизировать значение $S_T(x_0, t_0)$ в данной точке (x_0, t_0) границы допустимой области плоскости xyt.

Приходим к следующей оптимальной задаче для уравнения (11), рассматриваемого как уравнение в частных производных первого порядка: определить управляющую функцию $u(x, t)$ двух независимых переменных при ограничениях (2) так, чтобы решение $S_T(x, t)$, удовлетворяющее при $t = T$ условию (3), принимало в точке (x_0, t_0) минимальное возможное значение. Управление разыскивается в классе кусочно непрерывных функций, а решение $S_T(x, t)$ — в классе непрерывных функций двух независимых переменных.

Решение этой задачи получается немедленно по развитому ранее [3] общему методу. Именно, обозначив $S_T(x, t)$ через z, запишем основное уравнение (11) в виде системы (для простоты возьмем случай одной координаты x; случай многих координат рассматривается совершенно аналогично):

$$z_t = -g\zeta, \qquad z_x = \zeta$$

Введем множители Лагранжа ξ, η, Γ; по функции Гамильтона

$$H = -\xi g\zeta + \eta\zeta - \Gamma R$$

составим уравнения Эйлера

$$\partial\xi / \partial t + \partial\eta / \partial x = 0, \quad \eta - g\xi = 0 \quad \xi\zeta\partial g / \partial u + \Gamma\partial R / \partial u = 0$$

исключая из последних переменную η, получим

$$\partial\xi / \partial t + g\partial\xi / \partial x + \xi\partial g / \partial x = 0 \tag{12}$$

На характеристике $x = x(t)$ уравнение (12) принимает вид

$$\frac{d\xi}{dt} = -\xi\left[\frac{\partial g}{\partial x}\right]_{x=x(t)}$$

Отсюда следует, что знак ξ на характеристике определяется знаком начального значения ξ.

Функция Вейерштрасса, составленная с учетом требования непрерывности z на границе полоски варьирования [4], оказывается равной (g соответствует оптимальному управлению, G — допустимому)

$$\Delta H = -\xi \frac{gt_\tau - x_\tau}{Gt_\tau - x_\tau} \zeta [g - G]$$

Здесь t_τ, x_τ — направляющие косинусы полоски варьирования.

Если минимизируется значение z при $t = t_0$, $x = x_0$, то естественное граничное условие [3] дает

$$\xi(\tau) = \delta(\tau - x_0) > 0 \quad (\delta \text{ — дельта-функция})$$

Через τ обозначена длина дуги границы основной области, проходимой в положительном направлении (в данном случае эта область — полоса $t_0 \leqslant t \leqslant T$ плоскости (x, t); естественное условие ставится на линии $t = t_0$).

Написанное условие по существу выделяет только характеристику, начинающуюся в точке (x_0, t_0). Поэтому требование $\Delta H \geqslant 0$ фактически накладывается и выполняется только вдоль этой характеристики, и вдоль нее оно эквивалентно, как будет показано, обычному условию максимума $-\zeta[g - G] \geqslant 0$.

Действительно, из сказанного выше вытекает, что $\xi > 0$ вдоль характеристики; остается показать, что отношение $(gt_\tau - x_\tau) / (Gt_\tau - x_\tau)$ всегда положительно. При $g = G$ это отношение равно единице; при $g \neq G$ найдутся такие направления (t_τ, x_τ), которым соответствуют отрицательные значения этого отношения. Такие направления следует исключить из рассмотрения, так как им соответствуют характеристики, приходящие на разрыв (границу полоски варьирования) с обеих его сторон, и возмущения накапливаются на разрыве, вместо того чтобы передаваться через него. Нормальная ситуация такова, что возмущения, созданные полоской варьирования, проходят через ее границы и передаются во внешние части области, а это, как легко видеть, возможно лишь тогда, когда скорости их распространения (наклоны характеристик) одновременно либо больше, либо меньше скорости распространения разрыва (наклона границы полоски варьирования). Последнее требование эквивалентно положительности упомянутого выше отношения.

Как видим, в данной задаче возникает естественное ограничение наклона полосы варьирования. Условие Вейерштрасса принимает, таким образом, форму неравенства $-\zeta[g - G] \geqslant 0$, справедливого вдоль характеристики, выходящей из начальной точки (x_0, t_0). Уравнение (11) при этом переписывается в виде

$$\frac{\partial S_T}{\partial t} = -\max_U \left[\frac{\partial S_T}{\partial x^i} g^i(t;\ x,\ U(x, t)) \right], \qquad U \in A(t, x)$$

справедливом вдоль оптимальной траектории. Теперь, используя известное рассуждение [5], нетрудно прийти к уравнению Беллмана.

Поступила 1 X 1966

ЛИТЕРАТУРА

1. Berkovitz L. D. Variational Methods in Problems of Control and Programming. J. Math. Anal. and Appl. 1961, No. 3
2. Трикоми Ф. Лекции по уравнениям в частных производных. М. Изд. иностр. лит., 1957.
3. Лурье К. А. Задача Майера — Больца для кратных интегралов и оптимизация поведения систем с распределенными параметрами. ПММ, 1963, т. 27, вып. 5, стр. 842—853.
4. Лурье К. А. Оптимальное управление проводимостью жидкости, движущейся по каналу в магнитном поле. ПММ, 1964, т. 28, вып. 2, стр. 258—267.
5. Розоноэр Л. И. Принцип максимума Л. С. Понтрягина в теории оптимальных систем, III. Автоматика и телемеханика, 1959, т. 20, № 12.