А. Б. ГОРСТКО

ОБ ОДНОЙ СТОХАСТИЧЕСКОЙ ЗАДАЧЕ ТЕОРИИ УПРАВЛЕНИЯ

Как правило, в теории управления рассматриваются процессы, описываемые обыкновенными дифференциальными уравнениями. Типичная задача формулируется так:

заданы дифференциальное уравнение $\frac{dx}{dt}=f(x,u)$, $x|_{t=t_0}=x_0$ и целевая функция $R=\int_{t_0}^{T} g(x,u)\,dt$. Нужно найти такое управление u_0 из класса допустимых управлений, при котором достигается минимум R.

К настоящему времени изучены уже многочисленные модификации этой задачи. В качестве аппарата для изучения подобных вопросов используются методы динамического программирования, развитые американским математиком Р. Беллманом, и принцип максимума, разработанный Л. Понтрягиным и его учениками.

В настоящей работе рассматривается одна частная проблема теории управления в предположении, что u_0 не может быть выбрано точно.

I. Постановка задачи

I. Задано дифференциальное уравнение

$$\frac{dx}{dt}=ax+u+r, \quad x|_{t=t_0}=x_0, \tag{1}$$

где a - константа;

r - случайная величина с известной плотностью вероятности $f(r)$;

u - управление (произвольное вещественное число).

2. Заданы N чисел $y_1, y_2, \ldots, y_N$.

3. Фиксированы значения $t_i \in [t_0, T]$ $(i=1,2,\ldots,N)$.

4. Требуется выбрать $u(t_i)=u_i$ $(i=0,1,\ldots,N-1)$ так, чтобы достигал минимума функционал

$$R_N = \sum_{i=1}^{N} \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} (x_i - y_i)^2 f(r_1) f(r_2) \ldots f(r_N) dr_1 dr_2 \ldots dr_N + \lambda \sum_{i=0}^{N-1} u_i^2, \tag{2}$$

где $x_i = x(t_i)$, $\lambda > 0$.

Величина $(x_i - y_i)^2$ имеет смысл штрафа за уклонение x_i от y_i; λ - цена управления.

Замечание I. Во многих прикладных задачах последовательность $\{y_i\}$ может задаваться не явно, а, например, как совокупность значений решения некоторого дифференциального уравнения в фиксированных заранее точках.

Замечание 2. Зачастую представляют интерес не абсолютные, а приведенные затраты. Чтобы учесть их, следует в i-ое слагаемое функционала (2) ввести множитель $\varphi(i)$, позволяющий привести все затраты к одному моменту.

Замечание 3. Все результаты могут быть легко обобщены на случай системы линейных дифференциальных уравнений с постоянными коэффициентами.

Поясним теперь более подробно процесс выбора уравнения в (I):

а) в момент t_0 выбирается управление u_0,

б) по (1) вычисляется значение $x_1 = x_1(r)$,

в) определяется величина $\int_{-\infty}^{\infty} [x_1(r) - y_1]^2 f(r) dr + \lambda u_0^2$, дающая два первых слагаемых функционала (2),

г) измеряется значение r, реализованное при $t = t_0$,

д) из "б" и "г" находится значение x_1,

е) в момент t_1 выбирается управление u_1 для $t \in [t_1, t_2]$, ж) по (1), "е" и условию, что $x(t_1)=x_1$, вычисляется $x_2=x_2(r)$. Аналогично процесс продолжается и дальше.

Таким образом, предполагаем, что система, описываемая (1), обладает обратной связью, позволяющей, несмотря на ошибку при выборе u_K, определить точное значение x_{K+1}. Это предположение весьма существенно для дальнейшего.

II. Анализ и метод решения

Очевидно, что всякие N вещественных чисел могут быть выбраны в качестве управлений. Будем называть $\bar{u}_0, \bar{u}_1, \ldots, \bar{u}_{N-1}$ оптимальными, если $R_N(\bar{u}_0, \bar{u}_1, \ldots, \bar{u}_{N-1}) \leq R_N(u_0, u_1, \ldots, u_{N-1})$ для любых $u_0, u_1, \ldots, u_{N-1}$. Далее, введем понятие состояния системы. Под состоянием системы $\mathfrak{M}_i$ будем понимать вектор, первая компонента которого x_i, а вторая $N-i$. Решение поставленной задачи получим, пользуясь принципом оптимальности, сформулированным Р. Беллманом: "Оптимальное поведение обладает тем свойством, что, каковы бы ни были первоначальное состояние и решение в начальный момент, последующие решения должны составлять оптимальное поведение относительно состояния, получающегося в результате первого решения".

Проблема отыскания величин $\bar{u}_0, \bar{u}_1, \ldots, \bar{u}_{N-1}$ в последовательные моменты времени может рассматриваться как многошаговый процесс принятия решений. Принцип оптимальности, в применении к изучаемому вопросу, может формулироваться так: оптимальная последовательность $\bar{u}_0, \bar{u}_1, \ldots, \bar{u}_{N-1}$ обладает тем свойством, что для любого начального состояния $\mathfrak{M}_0$ и начального управления u_0 оставшиеся управления образуют оптимальную последовательность относительно состояния $\mathfrak{M}_1$, полученного в результате выбора начального управления. Таким образом, задача определения $\bar{u}_1$ аналогична определению $\bar{u}_0$ с той лишь разницей, что $\mathfrak{M}_0$ должно быть заменено на $\mathfrak{M}_1$, соответствующее выбору $\bar{u}_0$.

Введем последовательность управляющих функций $\{\Phi_K(\mathfrak{M}_i)\}$ $(K=1,\ldots,N;\ i=1,\ldots,m)$. Здесь $\Phi_K(\mathfrak{M}_i)$ — минимум R_K при условии, что система находится в состоянии $\mathfrak{M}_i$. Члены последовательности управляющих функций связаны между собой рекуррентными соотношениями

$$\begin{cases} \Phi_1(\mathfrak{M}_0) = \min\limits_{u_0} \left\{ \int\limits_{-\infty}^{\infty} (x_1 - y_1)^2 f(r)\,dr + \lambda u_0^2 \right\} \\ \text{-------------------------------} \\ \Phi_N(\mathfrak{M}_0) = \min\limits_{u_0} \left\{ \Phi_1(\mathfrak{M}_0) + \overline{\Phi_{N-1}(\mathfrak{M}_1)} \right\}. \end{cases} \tag{3}$$

Здесь $\mathfrak{M}_1$ - состояние, полученное в результате выбора u_0 и ошибки, сделанной при этом, т. е. это случайная величина. Значение $\overline{\Phi_{N-1}(\mathfrak{M}_1)}$ определяется формулой

$$\overline{\Phi_{N-1}(\mathfrak{M}_1)} = \int\limits_{-\infty}^{\infty} \Phi_{N-1}(\widetilde{\mathfrak{M}}_1 + Gr) f(r)\,dr, \tag{4}$$

где $\widetilde{\mathfrak{M}}_1$ - состояние, полученное в результате выбора u_0 без ошибки;

G - изменение в $\widetilde{\mathfrak{M}}_1$, вызываемое единичной ошибкой.

Установим теперь некоторые свойства оптимальных управлений. Для одношагового процесса величина u_0 находится из условия

$$\frac{d}{du_0} \left\{ \int\limits_{-\infty}^{\infty} (x_1 - y_1)^2 f(r)\,dr + \lambda u_0^2 \right\} = 0. \tag{5}$$

Используя (I), получим

$$\frac{d}{du_0} \left\{ \int\limits_{-\infty}^{\infty} \left[\left(x_0 + \frac{u_0 + r}{a} \right) e^{at_1} - \frac{u_0 + r}{a} - y_1 \right]^2 f(r)\,dr + \lambda u_0^2 \right\} = 0 \tag{6}$$

(для простоты полагаем $t_0 = 0$). Дифференцируя и приводя подобные члены, получаем

$$\bar{u}_0 = \bar{r} A_{t_1} + x_0 B_{t_1} + y_1 C_{t_1}, \tag{7}$$

где $\bar{r}$ - среднее значение случайной величины r. Коэффициенты A_{t_1}, B_{t_1} и C_{t_1} определяются формулами:

$$A_{t_1}=\frac{2\frac{e^{at_1}}{a^2}-\frac{e^{2at_1}}{a^2}-\frac{1}{a^2}}{\frac{e^{2at_1}}{a^2}-\frac{2e^{at_1}}{a^2}+\frac{1}{a^2}+\lambda}; \tag{8}$$

$$B_{t_1}=\frac{\frac{e^{at_1}}{a}-\frac{e^{2at_1}}{a}}{\frac{e^{2at_1}}{a^2}-\frac{2e^{at_1}}{a^2}+\frac{1}{a^2}+\lambda}; \tag{9}$$

$$C_{t_1}=\frac{\frac{e^{at_1}}{a}-\frac{1}{a}}{\frac{e^{2at_1}}{a^2}-\frac{2e^{at_1}}{a^2}+\frac{1}{a^2}+\lambda}. \tag{10}$$

Справедлива следующая теорема: для N-шагового процесса выбора управлений со случайной ошибкой r, с плотностью вероятности $f(r)$, средним $\bar{r}$ и дисперсией σ^2 оптимальное начальное управление $\bar{u}_0$ определяется формулой

$$\bar{u}_0=\bar{r}A_{t_N}+x_0B_{t_N}+\sum_{i=1}^{N}y_i C^i_{t_N}, \tag{11}$$

где $A_{t_N}, B_{t_N}, C^1_{t_N}, \dots, C^N_{t_N}$ - коэффициенты, зависящие от t_1, $t_2, \dots, t_N$. Доказательство легко проводится методом математической индукции.

Итак, поочередно используя (11) и (1), можно найти всю последовательность $\bar{u}_0, \bar{u}_1, \dots, \bar{u}_{N-1}$, так как

$$\begin{cases} \bar{u}_1=\bar{u}_0(\mathfrak{M}_1), \\ \bar{u}_2=\bar{u}_0(\mathfrak{M}_2), \\ \dots\dots\dots \\ \bar{u}_{N-1}=\bar{u}_0(\mathfrak{M}_{N-1}). \end{cases} \tag{12}$$

З а м е ч а н и е. Аналогичным методом может быть решена задача минимизации функционала

$$\int_{t_0}^{T}\int_{-\infty}^{\infty}[x(t,r)-y(t)]^2 f(r)\,dr\,dt+\lambda\int_{t_0}^{T}[u(t)]^2 dt,$$

где $u(t)$ - управление;

$y(t)$ - заданная функция времени;

$x(t,r)$ - решение исходного дифференциального уравнения.

В этом случае последовательность рекуррентных соотношений заменится одним функциональным уравнением, из которого можно найти $u(t)$.

В заключение считаю своим приятным долгом выразить благодарность В. А. Булавскому за обсуждение этой статьи.

Л И Т Е Р А Т У Р А

Б е л л м а н Р. Динамическое программирование. М., 1960.