УДК 519.281.2

Вестник ЛГУ. Сер. 1, 1986, вып. 1

О. Н. Граничин, В. Н. Фомин

МЕТОД ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ В ЗАДАЧЕ МИНИМАКСНОГО УПРАВЛЕНИЯ

1. Постановка задачи. Рассматривается объект управления (ОУ) со скалярным входом и выходом, описываемый уравнением

$$y_t = Y_t(y_0^{t-1},\ u_0^{t-1},\ \tau) + v_t, \tag{1}$$

где t — дискретное время, $t=1, 2, \ldots, T$; y_t, u_t, v_t — выходная, управляющая и возмущающая переменные; $y_0^{t-1} = (y_{t-1}, \ldots, y_0)$, $u_0^{t-1} = (u_{t-1}, \ldots, u_0)$; y_0 — заданное начальное условие; $Y_t(\cdot)$ — функции, заданные с точностью до векторного параметра τ из компактного множества $\mathbf{T}_0$, $\tau \in \mathbf{T}_0$.

Предполагается, что возмущающие переменные могут принимать произвольные значения из ограниченного множества $\mathbf{V}$,

$$v_t \in \mathbf{V},\ t=1, 2, \ldots, T. \tag{2}$$

Управление u_t формируется по типу обратных связей

$$u_t = U_t(y_0^t,\ u_0^{t-1}),\ t=1, \ldots, T,\ u_0 = U_0(y_0), \tag{3}$$

где $\{U_t(\cdot)\}$ — набор функций, называемый стратегией управления. Предполагается, что выделено некоторое множество $\mathbf{U}$ допустимых стратегий управления. Качество стратегии управления определяется функционалом

$$J[\{U_t(\cdot)\}] = \sup_{\tau \in \mathbf{T}_0} \sup_{v_t \in \mathbf{V}} \sum_{t=1}^{T} q_t(y_0^t,\ u_0^{t-1}), \tag{4}$$

который в силу (1) и (3) может рассматриваться как функция применяемой стратегии управления. Здесь $q_t(\cdot)$ — заданные вещественные функции соответствующих аргументов.

Оптимальное управление u_0^{T-1} определяется стратегией

$$\{U_t^{opt}(\cdot)\} = \arg \min_{\{U_t(\cdot)\} \in \mathbf{U}} J[\{U_t(\cdot)\}]. \tag{5}$$

Таким образом, критерий качества управления носит минимаксный характер. Качество оптимальной стратегии может существенно зависеть от выбора множества $\mathbf{U}$. Ниже для конкретного примера обсуждаются следующие классы неупреждающих стратегий управления:

а) Замкнутые стратегии. Здесь $\mathbf{U} = \mathbf{U}_3$ — множество всевозможных неупреждающих стратегий $\{U_t(\cdot)\}$;

б) Программные стратегии. Это подмножество $\mathbf{U} = \mathbf{U}_п$ множества $\mathbf{U}_3$, в котором стратегии $\{U_t(\cdot)\}$ определяются произвольными функциями $U_t(\cdot)$, не зависящими от выходных переменных.

в) Программно-замкнутое управление. Здесь стратегия строится с помощью серии программных оптимизаций с учетом значений выходной переменной в текущий и предшествующие моменты времени с уточнением на каждом такте управления множества возможных значений неизвестных параметров.

Показывается, как строятся оптимальные стратегии в каждом из перечисленных классов, и вычисляются соответствующие значения функционала качества. Для приводимого ниже примера все построенные стратегии (и их качество) оказываются различными.

2. Построение оптимальных стратегий. Задача оптимизации управления в классе U_3 решается с использованием идей метода динамического программирования.

Пусть для фиксированных функций $U_0(\cdot), \ldots, U_{T-2}(\cdot)$ под действием соответствующих управлений (а также возмущения v_1^{T-1}) реализовались выходные переменные $y_1, \ldots, y_{T-1}$. Знание выходных переменных y_0^{T-1} и управляющих переменных u_0^{T-2} позволяет уточнить множество возможных значений параметра τ. Именно, в силу (2) можно утверждать, что τ принадлежит множеству

$$\mathrm{T}_{T-1}(y_0^{T-1}, u_0^{T-2}) = \bigcap_{t=1}^{T-1} \{\tau' : y_t - Y_t(y_0^{t-1}, u_0^{t-1}, \tau') \in \mathrm{V}\} \cap \mathrm{T}_0. \qquad (6)$$

Теорема. *Функционал* (4) *может быть представлен в виде*

$$J = \sup_{\tau\in \mathrm{T}_0} \sup_{v_t \in \mathrm{V}} \left[\sum_{t=1}^{T-1} q_t(y_0^t, u_0^{t-1}) + \right.$$

$$\left. + \sup_{\tau' \in \mathrm{T}_{T-1}(y_0^{T-1}, u_0^{T-2})} \sup_{v_T\in \mathrm{V}} q_T(Y_T(y_0^{T-1}, u_0^{T-1}, \tau') + v_T, y_0^{T-1}, u_0^{T-1}) \right], \qquad (7)$$

где множество $\mathrm{T}_{T-1}(y_0^{T-1}, u_0^{T-2})$ *определяется формулой* (6).

Доказательство. В (4) исключим y_T в силу (1). Произведя максимизацию по v_T, получим

$$J = \sup_{\tau\in \mathrm{T}_0} \sup_{v_t\in \mathrm{V}} Q(x_{T-1}, \tau),$$

где $x_t = \mathrm{col}(y_0, \ldots, y_t, u_0, \ldots, u_t)$ и $Q(x_{T-1}, \tau) = \sum_{t=1}^{T-1} q_t(y_0^t, u_0^{t-1}) + \sup_{v_T\in\mathrm{V}} q_T[Y_T(y_0^{T-1}, u_0^{T-1}, \tau) + v_T, y_0^{T-1}, u_0^{T-1}]$.

Поскольку $\tau \in \mathrm{T}_{T-1}(y_0^{T-1}, u_0^{T-2})$ (см. (6)), то

$$J \leqslant \sup_{\tau\in \mathrm{T}_0} \sup_{v_t\in \mathrm{V}} \left[\sup_{\tau' \in \mathrm{T}_{T-1}(y_0^{T-1}, u_0^{T-2})} Q(x_{T-1}, \tau') \right]. \qquad (8)$$

Правые части в (8) и (7) с учетом введенных обозначений совпадают. Установим неравенство, обратное (8). Вектор x_{T-1} в силу (1) и (3) является функцией τ и v_1^{T-1}, $x_{T-1} = X_{T-1}(\tau, v_1^{T-1})$. Способ определения множества (6) таков, что при любых векторе $\tau \in \mathrm{T}_0$, векторе $\tau' \in \mathrm{T}_{T-1}(y_0^{T-1}, u_0^{T-2})$ и возмущающих воздействиях $v_t \in \mathrm{V}$, $t = 1, \ldots, T-1$, найдется последовательность $v_1', \ldots, v_{T-1}'$, удовлетворяющая (2), такая, что $X_{T-1}(\tau, v_1^{T-1}) = X_{T-1}(\tau', (v')_1^{T-1})$, а потому правая часть в (8) не превосходит величины

$$\sup_{\tau'\in \mathrm{T}_0} \sup_{v_t'\in \mathrm{V}} Q[X_{T-1}(\tau', (v')_1^{T-1}), \tau'].$$

Последняя величина равна значению функционала (4). Теорема доказана.

От переменной u_{T-1} зависит лишь последнее слагаемое в правой части (7), поэтому оптимальное на последнем такте управление определяется из условия минимизации по u_{T-1} функции

$$\tilde{q}_T(y_0^{T-1}, u_0^{T-1}) = \sup_{\tau'\in \mathbf{T}_{T-1}(y_0^{T-1}, u_0^{T-2})} \sup_{v_T\in V} q_T[Y_T(y_0^{T-1}, u_0^{T-1}, \tau') + v_T, y_0^{T-1}, u_0^{T-1}],$$

т. е.

$$u_{T-1}^{\mathrm{opt}} = U_{T-1}^{\mathrm{opt}}(y_0^{T-1}, u_0^{T-2}) = \arg\min_{u_{T-1}} \tilde{q}_T(y_0^{T-1}, u_0^{T-1}). \tag{9}$$

При этом функционал (7) примет значение

$$J = \sup_{\tau\in \mathbf{T}_*} \sup_{v_t\in V} \left[\sum_{t=1}^{T-2} q_T(y_0^t, u_0^{t-1}) + \hat{q}_{T-1}(y_0^{T-1}, u_0^{T-2})\right], \tag{10}$$

где

$$\hat{q}_{T-1}(y_0^{T-1}, u_0^{T-2}) = q_{T-1}(y_0^{T-1}, u_0^{T-2}) + \tilde{q}_T(y_0^{T-1}, U_{T-1}^{\mathrm{opt}}(y_0^{T-1}, u_0^{T-2}), u_0^{T-2}). \tag{11}$$

Функционал (10) имеет ту же форму, что и (4), пр[illegible]1. Поэтому описанную процедуру можно продолжить, вы[illegible]множество $\mathbf{T}_{T-2}(y_0^{T-2}, u_0^{T-3})$, функцию $U_{T-2}^{\mathrm{opt}}(y_0^{T-2}, u_0^{T-3})$ и т. [illegible]те с номером T придем к множеству $\mathbf{T}_0(y_0)$ и функции U_0^{opt} ([illegible]мая $\mathbf{T}_0(y_0) = \mathbf{T}_0$, полностью определим стратегию $\{U_t^{\mathrm{opt}}(\cdot)\}$[illegible], как несложно убедиться, минимизирует в $\mathbf{U}_2$ функционал [illegible]

3. **Зависимость решения задачи оптимизации от выбора [illegible]пустимых стратегий управления.** Пусть ОУ описывается урав[illegible]

$$y_t + \tau y_{t-1} = u_{t-1} + v_t, \quad t = 1, 2, \tag{12}$$

в котором все переменные скалярны и $|v_t| \leqslant 1$, y_0 [illegible] $\in [1;$ [illegible] Примем

$$J = \sup_{\tau\in[1;5]} \sup_{|v_t|<1} |y_2|. \tag{13}$$

а) Програмное управление. Перепишем [illegible] с учетом (12)

$$J = \sup_{\tau\in[1;5]} \sup_{|v_t|<1} |\tau^2 - \tau u_0 + \tau v_1 + u_1 + v_2|.$$

Производя указанную максимизацию по v_t и τ, получим фу[illegible]менных u_0, u_1. Минимизация этой функции по u_0, u_1 приводит [illegible]чинам $u_0 = 7$, $u_1 = 12{,}25$, причем $J_{\mathrm{opt}} = 8{,}25$.

б) Замкнутое управление. При фиксирован[illegible]еделим интервал

$$\mathbf{T}_1(y_1, u_0) = [1; 5] \cap [-1 - y_1 + u_0; 1 + u_0 - y_1] = [\tau_1; \tau_2],$$

концы которого вычисляются по формулам $\tau_1 = \max(-1 + u_0 - y_1; 1)$, $\tau_2 = \min(1 + u_0 - y_1; 5)$. Функционал (13) перепишется тогда в виде

$$J = \sup_{\tau\in[1;5]} \sup_{|v_1|<1} \left[\sup_{\tau'\in[\tau_1;\tau_2]} \sup_{|v_2|<1} |-\tau' y_1 + u_1 + v_1|\right]$$

и после проведения указанных операций становится функцией u_0. Минимизация по u_1 приводит к формуле

$$u_1 = U_1(y_1, u_0) = \frac{1}{2}(\tau_1 + \tau_2) y_1. \tag{14}$$

Подставляя (14) в (13), получим $J = \sup_{\tau \in [1;5]} \sup_{|\varphi_1| < 1} \left[\frac{1}{2}(\tau_2 - \tau_1)\right]|y_1| + 1]$.

Исключая теперь y_1 в силу (12) и производя указанные операции максимизации, найдем J как функцию u_0, при этом $u_0^{opt} = \arg\min_{u_0} J(u_0) = 3$.

Итак, оптимальная стратегия в U_3 имеет вид

$$u_0 = 3, \quad u_1 = \frac{1}{2}[\min(4 - y_1;\ 5) + \max(2 - y_1;\ 1)]\, y_1,$$

причем $J_{opt} = 2{,}125$.

в) Программно-замкнутое управление. На первом шаге ($t=0$) строится программное управление. В момент времени $t=1$ строится новое программное управление с учетом того, что выходная переменная y_1 известна. Как было показано раньше, $u_0 = 7$. На втором шаге программно-замкнутое управление определяется формулой $u_1 = U_1(y_1, 7)$, где $U_1(\cdot)$ — функция (14). При этом $J = 6$.

Результаты вычисления оптимальных стратегий в разных классах стратегий имеют, таким образом, следующий вид:

а) Программное управление:

$$u_0 = 7, \quad u_1 = 12{,}25, \quad \min_{U_\pi} J = 8{,}25.$$

б) Замкнутое управление:

$$u_0 = 3, \quad u_1 = \frac{1}{2}[\min(4 - y_1;\ 5) + \max(2 - y_1;\ 1)]\, y_1, \quad \min_{U_3} J = 2{,}125.$$

в) Программно-замкнутое управление:

$$u_0 = 7, \quad u_1 = \frac{1}{2}[\min(8 - y_1;\ 5) + \max(6 - y_1;\ 1)]\, y_1, \quad J = 6.$$

Библиографические замечания. Зависимость качества управления [illegible] множества оптимизации (класса неупреждающих стратегий) [illegible]на сравнительно недавно. Дело в том, что в классических задачах управления, в которых все параметры ОУ известны и помехи отсутствуют, множества программных и замкнутых стратегий управления оказываются совпадающими. В стохастическом случае ситуация усложняется: замкнутые стратегии могут обладать лучшим качеством, чем программные [1]. Эта точка зрения подчеркивается в работах Дрейфуса, ссылки и примеры можно найти в [2] (см. также [3]).

Программно-замкнутая стратегия управления рассматривалась как адаптивная в [4—6]. При этом подразумевалось, что качество управления такой стратегии оптимально в классе произвольных обратных связей. Приведенный выше пример показывает, что это не так.

Summary

A discrete scalar plant subjected to additive noise which is assumed to be constrained and otherwise arbitrary is considered. Some of the plant parameters are unknown. The quality criteria of the control function is minimax. Design of the optimal control is based on the method of dynamic programming. At the end the qualities of the optimization for program, close and program-close control functions are compared.

Литература

1. Острем К. Введение в стохатическую теорию управления. М., 1973. 321 с. 2. Соренсон Г. Обзор методов фильтрации и стохастического управления.— В кн.: Фильтрация и стохастическое управление в динамических системах. М., 1980, с. 11—48. 3. Фомин В. Н. Методы управления линейными дискретными объектами. Л., 1985.

340 с. 4. Кунцевич В. М., Лычак М. М. Об оптимальном и адаптивном управлении динамическими объектами в условиях неопределенности. — Автоматика и телемеханика, 1979, № 1, с. 19—83. 5. Кунцевич В. М., Лычак М. М. Игровые системы экстремального управления. — Автоматика, 1982, № 6, с. 74—86. 6. Кунцевич В. М., Лычак М. М. Синтез оптимальных и адаптивных систем управления: игровой подход. Киев, 1985. 20 с.

Статья поступила в редакцию 27 февраля 1985 г.