ВЕСТНИК УДМУРТСКОГО УНИВЕРСИТЕТА

МАТЕМАТИКА 2008. Вып. 2

УДК 517.977

© *Д. А. Серков*

СТРАТЕГИЯ МИНИМАКСНОГО РИСКА (СОЖАЛЕНИЯ) ДЛЯ ЗАДАЧ УПРАВЛЕНИЯ В УСЛОВИЯХ ДИНАМИЧЕСКИХ ПОМЕХ [1]

Для задачи управления в условиях динамических помех обсуждается возможность использования критерия минимаксного риска (сожаления) Сэвиджа. На примере системы с простыми движениями проводится сравнение соответствующих стратегий с оптимальными позиционными стратегиями.

Ключевые слова: оптимальное управление, позиционные стратегии, критерий минимаксного риска Сэвиджа.

Введение

Рассматривается управляемая система, описываемая обыкновенными дифференциальными уравнениями и подверженная действию неопределенных динамических помех. Управление и помеха стеснены геометрическими ограничениями. Сторона, формирующая управление, минимизирует величину терминального показателя качества. В теории дифференциальных игр (см. [1–4] и библ. в этих работах) рассмотрена ситуация эффективного противодействия помехи. Для этих задач типичны формирование помехи исходя из целей, противоположных целям управления, и зависимость помехи от состояния управляемой системы или от действий управляющей стороны. Такие предположения о характере помехи и её поведении определили оценку стратегии, называемую гарантированным результатом [1–4] и отвечающую минимаксному критерию.

Имеются также и задачи, в которых поведение помехи не связано со значениями рассматриваемого показателя качества и не зависит от состояния управляемой системы или действий управляющей стороны. В этих обстоятельствах бывает целесообразно перейти от минимаксной конструкции к оценке управления, основанной на критерии минимаксного риска (сожаления) Сэвиджа [5]. При этом в качестве неопределенностей рассматривать программные помехи, а в качестве альтернатив — неупреждающие (в частности, позиционные) стратегии.

[1] Работа выполнена при финансовой поддержке РФФИ (проект 06-01-00436).

Поясним конструкцию этой оценки. Выберем произвольную допустимую (программную) помеху и найдем результат в нашей задаче оптимального управления при заданной помехе. Затем вычислим значение показателя качества для этой же помехи и какой-либо стратегии управления. Превышение второй величины над первой характеризует наш риск при (сожаление о) выборе данной стратегии и реализации данной помехи. Стратегию, у которой максимум риска (сожаления) по всем помехам минимален, назовем оптимальной по риску, а величину соответствующего риска — оптимальным риском [6]. Отдельный интерес представляют задачи, обладающие нулевым оптимальным риском [7, 8]. В этом случае оптимальная по риску стратегия действует так, как если бы помеха была известна заранее.

Возникает вопрос, не будут ли решать задачу в предлагаемой постановке уже известные способы синтеза оптимального управления, исходящие из минимизации гарантированного результата. Частичный ответ дается в этой работе: приводится пример задачи управления для системы с простыми движениями, в которой оптимальные по риску стратегии отличаются от оптимальных позиционных стратегий [1, 2].

Пример

Рассмотрим управляемую систему

$$\begin{cases} \dot{x}(\tau) = u[\tau] + v[\tau], \quad \tau \in [t,2] \subseteq [0,2] \equiv T, \\ x(t) \equiv (x_1(t), x_2(t)) = (z_1, z_2) \equiv z, \quad (t,z) \in T \times \mathbb{R}^2, \end{cases}$$

где измеримые реализации управления и помехи при почти всех $\tau \in T$ стеснены ограничениями:

$$u[\tau] \equiv (u_1[\tau], u_2[\tau]) \in \mathcal{P} \equiv \left\{ (u_1, u_2) \in \mathbb{R}^2 \;\middle|\; \frac{|u_1|}{\alpha+\beta} + \frac{|u_2|}{\alpha} \leqslant 1 \right\},$$
$$\alpha > 0, \ \beta > 0, \ \alpha + \beta < 1,$$
$$v[\tau] \equiv (v_1[\tau], v_2[\tau]) \in \mathcal{Q} \equiv \left\{ (v_1, v_2) \in \mathbb{R}^2 \;\middle|\; |v_1| + |v_2| \leqslant 1 \right\}$$

и показатель качества вида

$$\gamma(x(\cdot)) = \sigma(x(2)) \equiv |x_1(2)| + x_2(2).$$

Нетрудно проверить, что в дифференциальной игре [1, § 18; 2, гл. I, § 6; 4, § 11] для приведенных системы и показателя качества существует цена игры [1, § 8; 2, гл. I, § 6; 4, § 11.3] $\rho(\cdot) : G \to \mathbb{R}$ и она удовлетворяет равенству

$$\rho(t,z) = \begin{cases} |z_1| + z_2 + (1-\alpha-\beta)(\vartheta - t), & |z_1| \geqslant (\alpha+\beta)(\vartheta - t), \\ \dfrac{\alpha}{\alpha+\beta}|z_1| + z_2 + (1-\alpha)(\vartheta - t), & |z_1| \leqslant (\alpha+\beta)(\vartheta - t). \end{cases}$$

Оптимальная позиционная стратегия [1, § 6; 2, гл. I, § 3; 4, § 11.2] для указанной дифференциальной игры может быть построена, например, в форме экстремального сдвига на сопутствующую точку [3]:

$$\hat{U}(\tau, x, \varepsilon) = \begin{cases} (-\operatorname{sign}(x_1)(\alpha+\beta), 0), & |x_1| > \alpha\varepsilon/\sqrt{\alpha^2 + (\alpha+\beta)^2}, \\ (0, -\alpha), & |x_1| \leqslant \alpha\varepsilon/\sqrt{\alpha^2 + (\alpha+\beta)^2}. \end{cases}$$

Определим стратегию управления $\bar{U}(\cdot) : G \to \mathcal{P}$ условиями

$$\bar{U}(\tau, x) = \begin{cases} (-\operatorname{sign}(x_1)(\alpha+\beta), 0), & |x_1| > 2-\tau, \\ (0, -\alpha), & |x_1| \leqslant 2-\tau. \end{cases}$$

В приведенной задаче ε-оптимальные по риску стратегии [6] (в частности, стратегия $\bar{U}$) не являются оптимальными позиционными стратегиями.

Стратегия $\bar{U}$ при произвольной программной помехе дает результат [6], отклоняющийся от результата в указанной задаче оптимального управления (а значит, и от результата для оптимальной позиционной стратегии) на величину, не превосходящую $\beta(2-t)$.

Кроме того, стратегия $\bar{U}$ имеет гарантированный результат [2, гл. I, §3], отклоняющийся от цены игры на величину, не превосходящую $\frac{\beta}{\alpha+\beta}(2-t)$.

Эта же стратегия при начальных условиях $x(0) = (-1, -1)$ и помехе $\bar{v}[\cdot] \equiv (1, 0)$ дает результат, улучшающий на величину порядка $\alpha(2-t)$ аналогичные результаты для стратегий экстремального сдвига на сопутствующую точку и произвольных чистых универсальных оптимальных стратегий [1, 2].

Таким образом, если параметр β пренебрежимо мал по сравнению с параметром α, при переходе к стратегии $\bar{U}$ происходит существенное улучшение результата на отдельных помехах при пренебрежимо малом ухудшении на множестве всех остальных помех.

СПИСОК ЛИТЕРАТУРЫ

1. Красовский Н. Н., Субботин А. И. Позиционные дифференциальные игры. М.: Наука, 1974.
2. Субботин А. И., Ченцов А. Г. Оптимизация гарантии в задачах управления. М.: Наука, 1981.
3. Красовский Н. Н. Управление динамической системой. М.: Наука, 1985.
4. Субботин А. И. Обобщенные решения уравнений в частных производных первого порядка. Перспективы динамической оптимизации. М.; Ижевск: Институт компьютерных исследований. 2003.

5. Savage L. J. The theory of statistical decision // J. American Statistical Association. 1951. № 46. P. 55–67.

6. Серков Д. А. Стратегии минимаксного риска (сожаления) в системе с простыми движениями // Труды ИММ УрО РАН. 2007. Т. 13, № 3.

7. Серков Д. А. Сильно оптимальные стратегии // Докл. АН СССР. 1991. Т. 321, № 2. С. 258–262.

8. Серков Д. А. О равномерных стратегиях // Теория управления и теория обобщенных решений уравнений Гамильтона-Якоби (CGS'2005). Тр. Междунар. семинара посвящ. 60-летию акад. А. И. Субботина: Екатеринбург: Изд-во Урал. ун-та, 2006. Вып. 1. С. 273–284.

Поступила в редакцию 14. 02. 08

D. A. Serkov
Minimax risk (regret) strategy for control problems for the system under dynamic disturbances

The paper presents an approach based on the minimax risk (regret) criterion of Savage to the solution of the control problem for the system under dynamic disturbances. For comparison of the strategies arising in the approach with the optimal feedback control (in the sense of the differential games theory) the system with simple motions is considered.

Серков Дмитрий Александрович
Институт математики
и механики УрО РАН
620219, Россия, Екатеринбург
ул. С.Ковалевской, 16
E-mail: serkov@nm.ru