УДК 62.50;517.972

О.Н.Граничин, канд.физ.-мат.наук
(Санкт-Петербургский государственный университет)

РОБАСТНОЕ МИНИМАКСНОЕ ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ

Для линейного дискретного объекта управления (ОУ) рассматривается задача о построении стабилизирующего линейного регулятора с подстраиваемыми коэффициентами в условиях априорной неопределенности о параметрах объекта и нерегулярных аддитивных возмущениях. При весьма слабых ограничениях на тип неопределенностей доказано существование стабилизирующего регулятора с определенными оптимальными свойствами. В частности, для доказательства существования не требуется выполнение предположения об управляемости ОУ при любом из возможных значений неизвестных параметров.

1.Введение

В период становления теории систем управления изучались в основном системы со скалярными входами и выходами, хотя в действительности они может быть были и более сложными. Последовательное развитие теории управления началось примерно с начала 60-х годов XX века, когда был заложен фундамент в изучении многомерных систем. Развитие различных программистских технологий и систем для разнообразных операций с матрицами сделало вполне легкими и доступными для применения в практических системах достижений теории. В то время как компьютеры, особенно персональные и рабочие станции становятся все более и более производительными, разработчики систем управления конструируют контроллеры для систем более высоких порядков, используя современную методологию. Все эти разработки естественно поднимают вопрос: в какой степени усложняется решение задачи конструирования регуляторов с ростом размерности задачи, определяемой числом входов/выходов и размерностью пространства состояний. Серьезное изучение сложности задач анализа и синтеза систем управления началось в последнее десятилетие XX века и привело к некоторым неожиданным выводам, в частности, многие из них оказались NP–сложными задачами, см., например, [15].

Рассмотрим для примера задачу о синтезе робастного устойчивого линейного регулятора дискретного объекта управления, с неопределенностью в виде интервальных матриц в задании параметров объекта. Для данного целого числа r обозначим $\mathbf{\Delta}$ некоторый набор пар из множества $\mathbf{R}^{2r^2}$

$$\mathbf{\Delta} = \{\{\alpha_{ij}, \beta_{ij}\},\ -\infty < \alpha_{ij} < \infty,\ -\infty < \beta_{ij} < \infty,\ \alpha_{ij}, \beta_{ij} \in \mathbf{Q},\ i, j = 1, \ldots, r\},$$

где $\mathbf{Q}$ — множество рациональных чисел. Обозначим

$$\mathbf{A}_{\mathbf{\Delta}} := \{A = (a_{ij}) \in \mathbf{R}^{r\times r} :\ \alpha_{ij} \le a_{ij} \le \beta_{ij},\ \forall i, j\}$$

множество называемое *интервальная матрица.* Пусть объект управления описывается разностным уравнением вида

$$x_{t+1} = \mathrm{A}x_t + \mathrm{B}u_t,\ t = 0, 1, 2, \ldots,$$

где $x_t \in \mathbf{R}^r$ — вектор состояния объекта в момент времени t, $u_t \in \mathbf{R}^m$ — набор управляющих воздействий в момент времени t, A — некоторая матрица из множества $\mathbf{A}_{\boldsymbol{\Delta}}$. B — заданная матрица размерности $r \times m$. Задачу о построении робастного стабилизирующего линейного регулятора обычно формулируют в виде: подобрать такую матрицу K размерности $m \times r$, задающую линейную обратную связь вида

$$u_t = \mathrm{K}x_t,$$

чтобы замкнутая система управления была устойчивой при любой матрице A из $\mathbf{A}_{\boldsymbol{\Delta}}$. Эта задача является $\mathcal{NP}$ сложной в том смысле, что при больших размерностях r неизвестно ещ решение за полиномиальное время, но проверка устойчивости замкнутой системы при определенной матрице $\mathrm{A} \in \mathbf{A}_{\boldsymbol{\Delta}}$ и выбранном регуляторе K можно произвести достаточно простым способом. Необходимым условием разрешимости задачи о построении робастного стабилизирующего линейного регулятора является условие управляемости пары матриц A и B при любой $\mathrm{A} \in \mathbf{A}_{\boldsymbol{\Delta}}$. Даже проверка этого более простого условия представляет собой пример $\mathcal{NP}$ сложной задачи. Возможные пути решения этой задачи описываются, например, в [10], [13], [14].

При всей сложности решения описанной выше задачи даже в простейшем скалярном случае $r = 1$ при $\mathbf{A}_{\boldsymbol{\Delta}} = [1; 5]$, $\mathrm{B} = 1$ простой анализ возможности построения робастного стабилизирующего регулятора дает отрицательный ответ. С одной стороны, невозможно подобрать постоянный коэффициент усиления регулятора $u_t = k\,x_t$, стабилизирующий при любом коэффициенте $a \in [1, 5]$ объект управления

$$x_{t+1} = ax_t + u_t,\ t = 0, 1, 2, \ldots.$$

А с другой, при $x_0 \neq 0$ и $u_0 = 0$ регулятор

$$u_t = -\frac{x_1}{x_0}x_t,\ t = 1, 2, \ldots$$

не только стабилизирует систему, но и удерживает $x_t = 0$, $t = 2, 3, \ldots$. Стратегии управления такого типа принято называть *адаптивными*, так как они в процессе функционирования объекта "приспосабливаются" к реально существующей обстановке. Рост производительности используемых на практике вычислительных устройств не только позволяет переходить к решению многомерных сложных задач теории управления, но и позволяет переосмыслить роль адаптивных алгоритмов. Их использование в системах реального времени может оказать существенную помощь в положительном решении некоторых неразрешимых задач.

В настоящей работе будет показано, что расширение класса допустимых управлений за счет включения в него линейных регуляторов с подстраиваемыми коэффициентами

позволяет сформулировать слабое достаточное условие на тип нерегулярной неопределенности в описании объекта управления, при котором существует простая стабилизирующая стратегия управления и более того имеет решение задача о построении минимаксного оптимального управления. В частности, как и в [11], нет необходимости накладывать упоминавшееся выше условие об управляемости пары матриц A и B при любой $\mathrm{A} \in \mathbf{A}_{\boldsymbol{\Delta}}$. Кроме априорной неопределенности о действительных параметрах объекта управления ниже будет рассматриваться более общая постановка задачи, учитывающая возможность рассмотрения нерегулярных помех, действующих аддитивно на каждом шаге. Задачи оптимизации в такой постановке рассматривались в [2], [5], [8], [9], [12] и многих других работах при известных и неизвестных параметрах объекта управления.

Рассматриваемый далее подход во многом следует идеям сформулированным в [3].

2. Постановка задачи и основные предположения

Пусть объект управления (ОУ) описывается разностным уравнением вида

$$x_{t+1} = \mathrm{A}x_t + \mathrm{B}u_t + v_{t+1},\ t = 0, 1, 2, \ldots, \tag{1}$$

где $x_t \in \mathbf{R}^r$ — наблюдаемый вектор состояния объекта в момент времени t, $u_t \in \mathbf{R}^m$ — набор управляющих воздействий в момент времени t, $v_t \in \mathbf{R}^r$ — вектор неконтролируемых ограниченных возмущений объекта в момент времени t : $\|v_t\|_1 \leq C_v$, A — матрица размерности $r \times r$ (вообще говоря, неизвестная) из некоторого семейства интервальных матриц $\mathbf{A}_{\boldsymbol{\Delta}}$, B — заданная матрица размерности $r \times m$. Здесь и далее обозначение $\|\cdot\|_1$ используется для нормы равной максимальному из абсолютных значений элементов вектора или матрицы. В дальнейшем будет удобно использовать обозначение $x_0^t = (x_0, \ldots, x_t)^T$ и $u_0^{t-1} = (u_0, \ldots, u_{t-1})^T$ для наборов сформировавшихся к моменту времени t значений выходных и управляющих переменных.

Постановка задачи. Построить линейный регулятор

$$u_t = \mathrm{K}_t x_t,\ t = 0, 1, 2, \ldots \tag{2a}$$

с подстраиваемым матричным размерности $m \times r$ коэффициентом

$$\mathrm{K}_t = K_t(x_0^t, u_0^{t-1}),\ t = 1, 2, \ldots,\ \mathrm{K}_0 = K_0(x_0), \tag{2b}$$

стабилизирующий замкнутую систему управления при любой матрице A из $\mathbf{A}_{\boldsymbol{\Delta}}$

$$\sup_t \|x_t\|_1 + \|u_t\|_1 < \infty \tag{3}$$

и минимизирующий предельный функционал качества с неотрицательной "штрафной" функцией $q(\cdot, \cdot)$

$$\bar{\Phi}(\{K_t(\cdot)\}, x_0) = \sup_{\mathrm{A} \in \mathbf{A}_{\boldsymbol{\Delta}}} \sup_{v_t : \|v_t\| \leq C_v, t > 0} \limsup_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} q(x_t, u_{t-1}). \tag{4}$$

Наряду с задачей о минимизации предельного функционала качества (4) ниже будет рассмотрена и задача с дисконтирующим функционалом качества

$$\Phi_\gamma(\{K_t(\cdot)\}, x_0) = \sup_{\mathrm{A}\in\mathbf{A}_{\mathbf{\Delta}}} \sup_{v_t:\|v_t\|\le C_v, t>0} (1-\gamma)\sum_{t=1}^{\infty}\gamma^t q(x_t, u_{t-1}) \tag{5}$$

где $0 < \gamma < 1$ "забывающий" (дисконтирующий) множитель. Дисконтирующий функционал (5) в отличие от предельного (4) в какой-то мере характеризует качество переходных процессов в системе управления.

Правило формирования набора функций $\{K_t(\cdot)\}_{t=0}^{\infty}$ будем называть *выбором стратегии управления.* Множество всех линейных стратегий управления (2), стабилизирующих ОУ (1), назовем классом линейных стабилизирующих регуляторов $\mathcal{U}_{st}$. Поставленные задачи носят минимаксный характер: ищется правило построения стратегии управления $\{K_t(\cdot)\}$ из класса $\mathcal{U}_{st}$, минимизирующее значение функционала качества (4) или (5) при "наихудщих" значениях неизвестных параметров $\Delta \in \mathbf{\Delta}$ и реализации последовательности возмущений $\{v_n\}$.

Определим новое понятие *ρ–управляемости пары* $(\mathbf{A}_{\mathbf{\Delta}}, \mathrm{B})$, состоящей из интервальной матрицы $\mathbf{A}_{\mathbf{\Delta}}$ (семейства матриц) и матрицы B.

Определение. Пара $(\mathbf{A}_{\mathbf{\Delta}}, \mathrm{B})$ называется ρ–управляемой, если в ρ–окрестности любой из матриц $\mathrm{A} \in \mathbf{A}_{\mathbf{\Delta}}$ найдется хотя бы одна $\mathrm{A}_\rho \in \mathbf{A}_{\mathbf{\Delta}}$ управляемая по Калману с матрицей B, т.е.

$$\forall \mathrm{A} \in \mathbf{A}_{\mathbf{\Delta}}\ \exists \mathrm{A}_\rho \in \mathbf{A}_{\mathbf{\Delta}} :\ \|\mathrm{A}_\rho - \mathrm{A}\|_1 < \rho$$

и

$$\mathrm{rank}(\mathrm{B}, \mathrm{A}_\rho\mathrm{B}, \ldots, \mathrm{A}_\rho^{r-1}\mathrm{B}) = r.$$

Основное предположение о структуре параметрической неопределенности, которое будет использоваться в дальнейшем, заключается в том, что пара $(\mathbf{A}_{\mathbf{\Delta}}, \mathrm{B})$ должна быть ρ–управляемой. Стоит отметить, что предположение о ρ–управляемости не является существенным ограничением и выполняется в большинстве типичных случаев.

3. Стабилизирующая стратегия управления

Поставленную выше задачу можно переформулировать в виде: выбрать оптимальную стратегию управления $\{K_t(\cdot)\}_{t=0}^{\infty}$, обеспечивающую минимум функционала качества (4)

$$\{K_t(\cdot)\}_{t=0}^{\infty} = \mathrm{argmin}_{\{K_t(\cdot)\}\in\mathcal{U}} \bar{\Phi}(\{K_t(\cdot)\}, x_0)$$

в классе линейных стабилизирующих стратегий $\mathcal{U}_{st}$.

Для обоснования корректности поставленной задачи надо сформулировать условия, при которых класс линейных стабилизирующих стратегий $\mathcal{U}_{st}$ не пустой.

Теорема 1 *Пусть $\rho < 1$. Если пара $(\mathbf{A}_{\mathbf{\Delta}}, \mathrm{B})$ является ρ–управляемой,* **тогда** *для ОУ (1) существует по крайней мере одна линейная стабилизирующая стратегия управления (2).*

Для доказательства теоремы 1 достаточно привести пример какого–нибудь адаптивного стабилизирующего алгоритма, обеспечивающего неравенство (4) при выполнении условий теоремы и "наихудщих" значениях параметров и возмущений. При большом многообразии выбора рассмотрим стабилизирующий алгоритм "Модифицированная полоска", прототип которого можно найти в [9].

В силу условия ограниченности векторов возмущений v_n неравенства

$$\|x_{t+1} - \hat{\mathrm{A}}x_t - \mathrm{B}u_t\|_1 \le 2C_v + \epsilon\|x_t\|, \ \|x_t\| = \sqrt{x_t^T x_t} \qquad t = 0, 1, \ldots \tag{6}$$

разрешимы относительно $\hat{\mathrm{A}}$ при любом $\epsilon \ge 0$ (им удовлетворяет, например, матрица $\hat{\mathrm{A}} = \mathrm{A}$ истинных параметров ОУ). "Целевые" неравенства (6) порождают алгоритм

$$\hat{\mathrm{A}}_{t+1} = \mathcal{P}_{\mathrm{A}_{\mathbf{\Delta}_\rho}}\left(\hat{\mathrm{A}}_t - \frac{\eta_t x_t^T}{\|x_t\|^2}\right), \tag{7}$$

где $\eta_t = (\eta_t^{(1)}, \ldots, \eta_t^{(r)})^T$, $\delta_t = x_{t+1} - \hat{\mathrm{A}}_t x_t - \mathrm{B}u_t = (\delta_t^{(1)}, \ldots, \delta_t^{(r)})^T$,

$$\eta_t^{(i)} = \begin{cases} \delta_t^{(i)} - C_v - \rho\|x_t\|, & \delta_t^{(i)} > 2C_v + \epsilon\|x_t\|, \\ \delta_t^{(i)} + C_v + \rho\|x_t\|, & \delta_t^{(i)} < -2C_v - \epsilon\|x_t\|, \ i = 1, \ldots, r, \\ 0 \end{cases}$$

а $\mathcal{P}_{\mathrm{A}_{\mathbf{\Delta}_\rho}}$ — проектор на множество $\mathrm{A}_{\mathbf{\Delta}_\rho}$ (сопоставляющий произвольной матрице $\hat{\mathrm{A}}$ размерности $r \times r$ ближайшую к ней матрицу $\mathrm{A}_\rho \in \mathbf{A}_{\mathbf{\Delta}}$, управляемую по Калману с матрицей B).

Пусть в алгоритме (7) параметр ϵ выбран из условия $\rho < \epsilon < 1$. Несложно показать, что алгоритм (7) при задании произвольного начального вектора $\hat{\mathrm{A}}_0$ сходится за конечное число шагов (см. [9], теорема 2.1.8), т.е. существует конечный момент времени $t_* = t_*(\hat{\mathrm{A}}_0, \{v_t\}, \epsilon)$, такой, что $\hat{\mathrm{A}}_t = \hat{\mathrm{A}}_{t_*}$ при $t \ge t_*$. Максимальное количество "переключений" алгоритма равномерно ограничено величиной $\frac{\|\mathrm{A} - \hat{\mathrm{A}}_0\|_1^2}{\epsilon\|x_0\|^2}$. При этом совсем не обязательно $\hat{\mathrm{A}}_{t_*} = \mathrm{A}$, но из вида алгоритма (7) следует, что при $t \ge t_*$ выполнены неравенства

$$\|v'_{t+1}\|_1 \le 2C_v + \epsilon\|x_t\|, \ v'_{t+1} = (\mathrm{A} - \hat{\mathrm{A}}_{t_*})x_t + v_{t+1}.$$

Уравнение объекта управления (1) при $t \ge t_*$ тогда можно переписать в виде

$$x_{t+1} = \hat{\mathrm{A}}_{t_*}x_t + \mathrm{B}u_t + v'_{t+1}, \ t = t_*, \ldots,$$

Так как в каждый момент времени $t = 1, 2, \ldots$ матрицы $\hat{\mathrm{A}}_t$ выбираются управляемыми по Калману с матрицей B, то можно выбрать такие матрицы K_t, задающие обратные связи (2), чтобы для кадратичных форм $\mathrm{D}_t^T\mathrm{D}_t$, $\mathrm{D}_t = \hat{\mathrm{A}}_t + \mathrm{B}\mathrm{K}_t$, выполнялось неравенство

$$\mathrm{D}_t^T\mathrm{D}_t \ < \ (1 - \epsilon)^2\mathrm{I},$$

где I матрица тождественного преобразования.

В силу выполнения всех "целевых" неравенств при больших n система управления диссипативна, (это легко устанавливается прямым методом Ляпунова).

4. Минимизация дисконтирующего функционала качества

Для начала рассмотрим задачу оптимизации на конечном промежутке времени $t = 0, 1, \ldots, n$. Для ее решения в классе линейных регуляторов можно воспользоваться идеями метода динамического программирования. Задача заключается в нахождении минимума "укороченного" функционала качества

$$\Phi_{n_\gamma}(\{K_t(\cdot)\}, x_0) = \sup_{\mathrm{A}\in\mathbf{A}_{\mathbf{\Delta}}} \sup_{v_t:\|v_t\|\le C_v, 0<t\le n} (1-\gamma)\sum_{t=1}^{n}\gamma^t q(x_t, u_{t-1}) \tag{8}$$

при фиксированном $n < \infty$ в классе линейных регуляторов с подстраиваемыми коэффициентами. Исследованию возможностей решения примерно такой задачи посвящена работа [4], где был предложен следующий алгоритм синтеза оптимальной для функционала (8) стратегии управления, базирующийся на идеях метода динамического программирования Белмана. Пусть функции $K_0(\cdot), \ldots, K_{n-2}(\cdot)$ в (2b) выбраны каким–либо образом и под действием соответствующих управлений и возмущающих переменных $v_1, \ldots, v_{n-1}$ реализовались выходные переменные $x_1, \ldots, x_{n-1}$. Знание значений выходных переменных до момента времени $n-1$ и управляющих до момента времени $n-2$ позволяет уточнить множество $\mathbf{A}_{n-1}$ возможных значений параметров A ОУ (1). Именно, в силу ограниченности возмущений v_t, $t = 1, \ldots, n-1$ можно утверждать, что матрица A принадлежит множеству

$$\mathbf{A}_{n-1}(x_0^{n-1}, u_0^{n-2}) = \cap_{t=0}^{n-2}\{\mathrm{A}' : \|x_{t+1} - \mathrm{A}'x_t - \mathrm{B}u_t\|_1 \le C_v\} \cap \mathrm{A}_{\mathbf{\Delta}}.$$

В [4] и [1] доказано утверждение, что функционал (8) с учетом введенных обозначений можно переписать в виде

$$\Phi_{n_\gamma}(\{K_t(\cdot)\}, x_0) = \sup_{\mathrm{A}\in\mathbf{A}_{\mathbf{\Delta}}} \sup_{v_t:\|v_t\|\le C_v, 0<t<n} (1-\gamma)\Big(\sum_{t=1}^{n-1}\gamma^t q(x_t, u_{t-1}) +$$

$$+\gamma^n \sup_{\mathrm{A}\in\mathbf{A}_{n-1}(x_0^{n-1}, u_0^{n-2})} \sup_{v_n:\|v_n\|\le C_v} q(\mathrm{A}x_{n-1} + \mathrm{B}u_{n-1} + v_n, u_{n-1})\Big)$$

От управляющей переменной u_{n-1} зависит лишь последнее слагаемое в правой части последнего соотношения, поэтому оптимальное на последнем такте управление определяется из условия его минимизации по u_{n-1}, т.е.

$$\mathrm{K}_{n-1} = \arg\min_{\mathrm{K}_{n-1}} \sup_{\mathrm{A}\in\mathbf{A}_{n-1}(x_0^{n-1}, u_0^{n-2})} \sup_{\|v_n\|\le C_v} q((\mathrm{A} + \mathrm{B}\mathrm{K}_{n-1})x_{n-1} + v_n, \mathrm{K}_{n-1}x_{n-1}).$$

После подстановки в (8) вместо u_{n-1} полученного выше оптимального значения получаем выражение

$$\sup_{\mathrm{A}\in\mathbf{A}} \sup_{v_t:\|v_t\|\le C_v, 0<t\le n} (1-\gamma)\Big(\sum_{t=1}^{n-2}\gamma^t q(x_t, u_{t-1}) + \bar{q}_{n-1}(x_0^{n-1}, u_0^{n-2})\Big),$$

которое по своей сути имеет форму похожую на вид функционала (8) при замене $n \to n-1$. Поэтому описанную выше процедуру минимизации можно продолжить, используя множество $\mathbf{A}_{n-2}(x_0^{n-2}, u_0^{n-3})$, получить вид функции $K_{n-2}^{n_\gamma -}(\cdot)$ и т.д. На шаге с номером n придем к множеству $\mathbf{A}_0(x_0)$. Принимая $\mathbf{A}_0(x_0) = \mathbf{A}_{\mathbf{\Delta}}$, можем полностью определить n_γ–оптимальную стратегию $\{K_t^{n_\gamma - opt}(\cdot)\}$, минимизирующую "укороченный" функционал качества (8).

Обозначим $\mathcal{I}_{n_\gamma}(x_0), n = 1, 2, \ldots$ минимальные значения "укороченных" функционалов качества (8). Зная алгоритм выбора оптимальной стратегии для "укороченного" функционала качества (8) можно на бесконечном промежутке времени построить γ–оптимальную стратегию $\{K_t^{\gamma - opt}(\cdot)\}_{t=0}^{\infty}$ для функционала (5). Если предположить, что штрафная функция $q(\cdot, \cdot)$ неограниченно возрастает при неограниченном росте нормы любого из ечц аргументов и $\mathcal{I}_{n_\gamma}(x_0)$ равномерно ограничены, то последовательность $\{K_0^{n_\gamma - opt}(x_0)\}_{n=1}^{\infty}$ при фиксированном x_0 будет ограниченной. Следовательно, из нечц можно выбрать сходящуюся подпоследовательность. Обозначим $K_0^{\gamma - opt}(x_0)$ предел этой подпоследовательности. Далее, последовательность $\{K_1^{n_\gamma - opt}(x_0^1, K_0^{\gamma - opt}(x_0)x_0)\}_{n=1}^{\infty}$ в силу тех же условий также ограничена. Рассматривая подпоследовательность этой последовательности, соответствующую выбранной ранее сходящейся подпоследовательности, можно выбрать из нечц сходящуюся подпоследовательность, предел которой обозначим $K_1^{\gamma - opt}(x_0^1, u_0)$. Повторив многократно описанную процедуру, можно получить стратегию управления $\{K_t^{\gamma - opt}(\cdot)\}_{t=0}^{\infty}$, которая называется слабым пределом последовательности стратегий $\{K_t^{n_\gamma - opt}(\cdot)\}_{t=0}^{\infty}, n = 1, 2, \ldots$.

Лемма 1 *Пусть $\rho < 1$ и функция $q(\cdot, \cdot)$ неограниченно возрастает при неограниченном росте нормы любого из ечц аргументов. Если пара $(\mathbf{A}_{\mathbf{\Delta}}, \mathrm{B})$ является ρ–управляемой,* **тогда** *для ОУ (1) задача минимизации дисконтирующего функционала (5) в классе линейных регуляторов с подстраиваемыми коэффициентами имеет решение. В частности, слабый предел последовательности стратегий $\{K_t^{n - opt}(\cdot)\}_{t=0}^{n}$ при $n \to \infty$ минимизирует функционал качества (5) и его соответствующее минимальное значение равно*

$$\mathcal{I}_\gamma(x_0) = \Phi_\gamma(\{K_t^{\gamma - opt}(\cdot)\}, x_0) = \lim_{n \to \infty} \mathcal{I}_{n_\gamma}(x_0).$$

Доказательство леммы 1 приведено в приложении.

5. Существование предельно–оптимальной стратегии управления

Приведенное выше решение задачи о минимизации дисконтирующего функционала не очень хорошее, так как полученная стратегия, вообще говоря, не обязана принадлежать классу линейных стабилизирующих регуляторов $\mathcal{U}_{st}$. Интуитивно должно быть понятно, что при $\gamma \approx 1$ значения предельного (4) и дисконтирующего (5) функционалов качества становятся близкими и оптимальные стратегии управления тоже должны быть близкими. Кроме того, если существует оптимальная стратегия для предельного функционала (4), то она обязательно будет стабилизирующей.

Теорема 2 *Пусть $\rho < 1$ и функция $q(\cdot,\cdot)$ неограниченно возрастает при неограниченном росте нормы любого из еҷ аргументов. Если пара $(\mathbf{A}_{\mathbf{\Delta}}, \mathrm{B})$ является ρ–управляемой,* **тогда** *для ОУ (1) задача минимизации предельного функционала (4) в классе линейных стабилизирующих регуляторов с подстраиваемыми коэффициентами $\mathcal{U}_{st}$ имеет решение. В частности, слабый предел последовательности стратегий $\{K_t^{\gamma-opt}(\cdot)\}_{t=0}^{\infty}$ при $\gamma \to 1$ является стабилизирующей стратегией, минимизирующей функционал (4) и соответствующее минимальное значение функционала качества (4) равно*

$$\bar{\mathcal{I}}(x_0) = \bar{\Phi}(\{K_t^{opt}(\cdot)\}, x_0) = \lim_{N\to\infty} \mathcal{I}_\gamma(x_0).$$

Доказательство теоремы 2 также приведено в приложении.

Замечание. Необходимо подчеркнуть, что из вида полученной оптимальной стратегии управления следует, что текущие коэффициенты регулятора обратной связи определяются получающимися текущими рассчитываемыми множествами возможных значений неизвестных параметров. В работе [7] был отмечен тот факт, что если выбрать чуть-чуть завышенный уровень внешнего возмущения $C_v' > C_v$, то при выборе любой стабилизирующей стратегии управления может произойти только конечное число изменений множеств $\mathbf{A_t}$, т.е. начиная с некоторого момента времени новые данные наблюдений перестают влиять на вид предельного множества, содержащего набор неизвестных параметров.

Существенной особенностью полученной выше оптимальной стратегии управления является то, что при еҷ построении даже в пределе, фактически, не используется информация о действительном наборе неизвестных параметров A, так как ничто в постановке задачи не может гарантировать еҷ получение. Это свойство оптимальной стратегии управления отличает еҷ от часто используемого идентификационного подхода в синтезе адаптивного управления. Вместо сходимости к точке, соответствующей вектору неизвестных параметров, построение оптимальной стратегии при достаточно больших t базируется на описании некоторого предельного множества, содержащего вектор неизвестных параметров. В ряде работ по синтезу адаптивного управления ранее уже рассматривались стратегии с подобными свойствами (см., например, [6]). По–существу, похожего типа стратегии используются и при построении замкнутых систем управления с помощью нейронных сетей. В контур обратной связи включается блок с нейронной сетью, коэффициенты которой можно подстраивать. На начальном этапе (при обучении) подстраиваются коэффициенты, потом замкнутая система работает без перенастройки.

6. Пример и заключение

Пусть ОУ линейным скалярным уравнением ($r = 1$)

$$x_{t+1} = ax_t + u_t + v_{t+1},\ t = 0, 1, 2, \ldots, x_0 = 1.$$

при неизвестном $a \in \mathbf{A_\Delta} = [1;5]$ и ограниченном возмущении: $|v_t| \le 1$. Рассмотрим функционал качества

$$\Phi_2(K_0(\cdot), K_1(\cdot), x_0) = \sup_{a\in[1;5]} \sup_{|v_1|\le 1, |v_2|\le 1} |x_1| + \frac{1}{2}|x_2|. \tag{9}$$

а) Программное управление. Предположим, что управление в моменты времени $t = 0, 1$ формируется по правилам

$$u_0 = k_0 x_0,\ u_1 = k_1 x_1$$

Задача минимизации функционала (9) в классе программных стратегий состоит в нахождении коэффициентов k_0 и k_1, минимизирующих

$$\sup_{a\in[1;5]} \sup_{|v_1|\le 1, |v_2|\le 1} |a + k_0 + v_1| + \frac{1}{2}|(a + k_1)(a + k_0 + v_1) + v_2|.$$

Решение этой задачи дает $k_0 = k_1 = -3$ и соответствующее минимальное значение функционала качества (9) равно $\mathcal{I}_2^{pr} = 6,5$.

б) Замкнутое управление. При фиксированном u_0 определим интервал

$$\mathbf{A_1}(x_1, u_0) = [x_1 - u_0 - 1; x_1 - u_0 + 1] \cap [1;5] = [a_1; a_2],$$

концы которого вычисляются по формулам: $a_1 = \max\{x_1 - u_0 - 1; 1\}$ и $a_2 = \min\{x_1 - u_0 + 1; 5\}$. Функционал (9) можно переписать в виде

$$\Phi_2(K_0(\cdot), K_1(\cdot), x_0) = \sup_{a\in[1;5]} \sup_{|v_1|\le 1} \Big(|x_1| + \frac{1}{2} \sup_{a'\in[a_1;a_2]} \sup_{|v_2|\le 1} |a' x_1 + u_1 + v_1| \Big).$$

Второе слагаемое является функцией от u_1. Минимизация по u_1 приводит к формуле

$$u_1 = K_1^{2-opt}(x_1, u_0) x_1 = \frac{1}{2}(a_1 + a_2) x_1. \tag{10}$$

Подставляя (10) в (9) получаем

$$\sup_{a\in[1;5]} \sup_{|v_1|\le 1} \Big(|x_1| + \frac{1}{4}(a_2 - a_1)|x_1| + 0,5 \Big).$$

Исключая теперь x_1 и производя соответствующие операции максимизации, можно переписать последнее выражение в виде функции от u_0. Минимизация по u_0 дает выражение $u_0 = K_0^{2-opt}(x_0) x_0 = -3$ и соответствующее минимальное значение функционала качества (9) равно $\mathcal{I}_2^{cl} = 3\frac{9}{16}$.

Зависимость качества управления от задания множества оптимизации (класса неупреждающих стратегий) осознана 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.

Полученное в этой работе решение задачи оптимизации носит в большей степени теоретический характер и достаточно сложно реализуемо на практике. Вместе с тем возможность получить аналитические формулы для построения оптимальной стратегии управления представляется маловероятной. В рассмотренном упрощенном примере уже при оптимизации функционала качества на конечном промежутке времени ($n = 3$) попытки выписать аналитическое решение наталкиваются на существенные трудности.

ПРИЛОЖЕНИЕ

Доказательство леммы 1. Как не трудно убедиться при выполнении условий леммы 1 выполняются и условия теоремы 1. Следовательно, существует хотя бы одна стабилизирующая стратегия управления $\{K_t^{st}(\cdot)\}_{t=0}^{\infty}$ из класса $\mathcal{U}_{st}$.

Рассмотрим числовую последовательность $\{\mathcal{I}_{n_\gamma}(x_0)\}$. Эта последовательность монотонная и ограниченная

$$\mathcal{I}_{1_\gamma}(x_0) \leq \ldots \leq \mathcal{I}_{n_\gamma}(x_0) \leq \ldots \leq \Phi_\gamma(\{K_t^{st}(\cdot)\}, x_0) < \infty,$$

а значит, существует предел. Обозначим его $\mathcal{I}_\gamma^\infty(x_0)$.

Сначала покажем, что для любой стратегии $\{K_t(\cdot)\}$ значение функционала (5) не меньше величины $\mathcal{I}_\gamma^\infty(x_0)$, т.е. $\mathcal{I}_\gamma^\infty(x_0) \leq \Phi_\gamma(\{K_t(\cdot)\}, x_0)$. Докажем от противного, сведя к противоречию. Если это не так, то существует такое $\delta > 0$, что $\mathcal{I}_\gamma^\infty(x_0) > \Phi_\gamma(\{K_t(\cdot)\}, x_0) + \delta$. По определению предела существует такое n_δ, что для любых $n \geq n_\delta$ $\mathcal{I}_{n_\gamma}(x_0) > \mathcal{I}_\gamma^\infty(x_0) - \delta/2$. В силу двух последних неравенств имеем

$$\mathcal{I}_\gamma^\infty(x_0) > \Phi_\gamma(\{K_t(\cdot)\}, x_0) + \delta \geq \Phi_{n_\gamma}(\{K_t(\cdot)\}, x_0) + \delta \geq \mathcal{I}_{n_\gamma}(x_0) + \delta > \mathcal{I}_\gamma^\infty(x_0) + \delta/2,$$

что является противоречивым неравенством.

Теперь также методом от противного покажем, что значение функционала (5) не меньше величины $\mathcal{I}_\gamma(x_0) \leq \mathcal{I}_\gamma^\infty(x_0)$. Пусть это не так. Тогда существует такие $\delta > 0$, последовательность $\{v_n'\}$ ($\|v_n'\| \leq C_v$) и матрица $\mathrm{A}' \in \mathbf{A}_{\mathbf{\Delta}}$, что для переменных $x_1', x_2', \ldots$ и $u_0', u_1', \ldots$, реализовавшихся при использовании стратегии $\{K_t^{\gamma-opt}(\cdot)\}$, справедливо неравенство $\mathcal{I}_\gamma(x_0) \geq (1-\gamma)\sum_{t=1}^{\infty}\gamma^t q(x_t', u_{t-1}') > \mathcal{I}_\gamma^\infty(x_0) + \delta$. Следовательно, найдется такое натуральное число k, для которого $(1-\gamma)\sum_{t=1}^{k}\gamma^t q(x_t', u_{t-1}') > \mathcal{I}_\gamma^\infty(x_0) + \delta/2$. Вместе с тем, в силу построения стратегии $\{K_t^{\gamma-opt}(\cdot)\}$ найдется такое достаточно большое $n > k$, что при $t = 0, 1, \ldots, n$ траектории системы $\{x_t'\}$ и $\{u_t'\}$ будут незначительно отличаться от траекторий $\{x_t''\}$ и $\{u_t''\}$, получающихся при использовании стратегии $\{K_t^{n_\gamma-opt}(\cdot)\}$:

$$(1-\gamma)\sum_{t=1}^{n}\gamma^t|q(x_t', u_{t-1}') - q(x_t'', u_{t-1}'')| < \delta/4.$$

А значит,

$$(1-\gamma)\sum_{t=1}^{k}\gamma^t q(x_t', u_{t-1}') < \mathcal{I}_{n_\gamma}(x_0) + \delta/4 < \mathcal{I}_\gamma^\infty(x_0) + \delta/4.$$

Это неравенство противоречит полеченному ранее. Следовательно, сделанное выше предположение неверно и $\mathcal{I}_\gamma(x_0) = \mathcal{I}_\gamma^\infty(x_0)$. Таким образом, лемма 1 доказана.

Доказательство теоремы 2. При выполнении условий теоремы 2 выполняются и условия теоремы 1. Следовательно, существует хотя бы одна стабилизирующая стратегия управления $\{K_t^{st}(\cdot)\}_{t=0}^\infty$ из класса $\mathcal{U}_{st}$. Рассмотрим произвольную числовую последовательность $\{\gamma_i\}$, $0 < \gamma_i < 1$, сходящуюся к единице $\gamma_i \to 1$ при $i \to \infty$, и соответствующую ей последовательность $\{\mathcal{I}_{\gamma_i}(x_0)\}$, которая ограничена в силу существования стабилизирующей стратегии управления. Из любой ограниченной последовательности можно выбрать сходящуюся подпоследовательность. Пусть $\{\gamma_{i_k}\}$ некоторая возрастающая подпоследовательность $\{\gamma_i\}$, для которой подпоследовательность $\{\mathcal{I}_{\gamma_{i_k}}(x_0)\}$ сходится к некоторому пределу, который обозначим $\{\mathcal{I}_1'(x_0)\}$. Далее будет показано, что для любой стратегии управления $\{K_t(\cdot)\}$ значение функционала (4) не меньше величины $\mathcal{I}_1'(x_0)$, а для приведенной в формулировке теоремы 2 $\bar{\Phi}(\{K_t^{opt}(\cdot)\}, x_0) = \mathcal{I}_1'(x_0)$, что в свою очередь в силу произвольности выбора последовательности $\{\gamma_i\}$ и подпоследовательности $\{\gamma_{i_k}\}$ влечет заключение теоремы 2. Стабилизируемость полученной предельно оптимальной стратегии легко устанавливается из вида функционала качества (4). Доказательство обоих утверждений будет проведено от противного.

Пусть для некоторой стратегии $\{K_t(\cdot)\}$ значение функционала (4) меньше величины $\mathcal{I}_1'(x_0)$, т.е. для любых матрицы $\mathrm{A}' \in \mathbf{A}_{\mathbf{\Delta}}$ и последовательности возмущений $\{v_n'\}$ ($\|v_n'\| \le C_v$) найдутся такие $\delta > 0$ и натуральное N_δ, что для любого $N > N_\delta$ для соответствующих траекторий $x_1', x_2', \ldots$ и $u_0', u_1', \ldots$, реализовавшихся при использовании стратегии $\{K_t(\cdot)\}$, справедливо неравенство

$$\mathcal{I}_1'(x_0) > \frac{1}{N}\sum_{t=1}^{N} q(x_t', u_{t-1}') + \delta$$

По определению предела существует такое n_δ, что для любых $n \ge n_\delta$ $\mathcal{I}_{n_{\gamma_{i_n}}}(x_0) > \mathcal{I}_1'(x_0) - \delta/2$. Выбрав n настолько большим, что $\gamma_{i_n} > 1 - \frac{1}{N}$, учитывая неравенство

$$\frac{1}{N}\sum_{t=1}^{N} q(x_t', u_{t-1}') \ge (1-\gamma_{i_n})\sum_{t=1}^{N} \gamma_{i_n}^t q(x_t', u_{t-1}')$$

получаем противоречивое неравенство $\mathcal{I}_1'(x_0) > \mathcal{I}_1'(x_0) + \delta/2$.

Далее предположим, что $\bar{\Phi}(\{K_t^{opt}(\cdot)\}, x_0) > \mathcal{I}_1'(x_0)$, тогда существует такие $\delta > 0$, натуральное N_δ, последовательность $\{v_n'\}$ ($\|v_n'\| \le C_v$) и матрица $\mathrm{A}' \in \mathbf{A}_{\mathbf{\Delta}}$, что для переменных $x_1', x_2', \ldots$ и $u_0', u_1', \ldots$, реализовавшихся при использовании стратегии $\{K_t^{opt}(\cdot)\}$, справедливо при любом $N > N_\delta$ неравенство

$$\frac{1}{N}\sum_{t=1}^{N} q(x_t', u_{t-1}') > \mathcal{I}_1'(x_0) + \delta$$

Вместе с тем, в силу построения стратегии $\{K_t^{opt}(\cdot)\}$ найдется такое достаточно большое n, что при $k > n$ траектории системы $\{x_t'\}$ и $\{u_t'\}$ будут незначительно

отличаться от траекторий $\{x''_t\}$ и $\{u''_t\}$, получающихся при использовании стратегии $\{K_t^{\gamma_{i_k}-opt}(\cdot)\}$, т.е. при любом $N > N_\delta$

$$\frac{1}{N}\sum_{t=1}^{N}|q(x'_t,u'_{t-1})-q(x''_t,u''_{t-1})|<\delta/4.$$

С другой стороны в силу определения предела существует натуральное n_δ такое, что при достаточно больших $k > n_\delta$ $|\mathcal{I}_{\gamma_{i_k}}(x_0)-\mathcal{I}'_1(x_0)|<\delta/4$. Выбрав k и N из условий $|1-\gamma_{i_k}^N|\sup_{t\le N} q(x''_t,u''_{t-1})<\delta/4$, $k>n_\delta$ и $\gamma_{i_k}<1-\frac{1}{N}$, получаем

$$\frac{1}{N}\sum_{t=1}^{N}q(x'_t,u'_{t-1})\le\frac{1}{N}\sum_{t=1}^{N}q(x''_t,u''_{t-1})+\delta/4\le$$

$$\le(1-\gamma_{i_k})\sum_{t=1}^{k}\gamma_{i_k}{}^t q(x''_t,u''_{t-1})+\delta/2\le\mathcal{I}_{\gamma_{i_k}}(x_0)+\delta/2\le\mathcal{I}'_1(x_0)+0.75\delta.$$

Последнее неравенство противоречит сделанному предположению. Таким образом, теорема 2 доказана.

СПИСОК ЛИТЕРАТУРЫ

[1] *Бакан . .,* "" // АиТ. 198 , No . с. – .

[2] *Барабанов А.Е., Граничин О.Н.* "Оптимальный регулятор линейного объекта с ограниченной помехой." // АиТ. 1984, No 5. с.39–46.

[3] *Граничин О.Н.,Фомин В.Н.* "Минимаксная оптимизация управления объектом на бесконечном промежутке времени" // Рукопись деп. в ВИНИТИ, 1985, No 4851-85, 29с.

[4] *Граничин О.Н.,Фомин В.Н.* "Метод динамического программирования в задаче минимаксного управления" // Вестник Ленингр. ун–та. Сер. 1., 1986, вып.1, с.26–30

[5] *Граничин О.Н.* "Построение дискретного субоптимального регулятора непрерывного объекта с нерегулярной ограниченной помехой" // АиТ. 2001, No 3. С.86–94.

[6] *Коган М..* "" // С. – .

[7] *Соколов В.Ф.* ""// С. – .

[8] *Соколов В.Ф.* "Синтез ℓ_1–субоптимального регулятора для линейного скалярного объекта с неструктурированной неопределенностью" // АиТ. 2001, No 1. С.150–163.

[9] *Фомин В.Н., Фрадков А.Л., Якубович В.А.* "Адаптивное управление динамическими объектами." М.: Наука, 1981. 448 с.

[10] *Calafiore G., Polyak B.* "Fast algorithms for exact and approximate feasibility of robust LMIs"//

[11] *Chen H.F., Cao X.R.* "Controllability is Not Necessary for Adaptive Pole Placement Control"// IEEE Trans.on AC, v.42, 1997, No.9, pp.1222–1229.

[12] *Khammash M.* "A New Approach to the Solution of the l_1 Control Problem: The Scaled-Q Method" //IEEE Trans.on AC, 2000, v.45, No.2, pp.180-187

[13] *Polyak B., Tempo R.* "Probabilistic Robust Design with Linear Quadratic Regulators" //

[14] *Tempo R., Dabbene F.,* "Probailistic Robustness Analysis and Design of Uncertain Systems" //Progress in Systems and Control Theory, v. 25, 1999, pp. 263–282.

[15] *Vidyasagar M.* "Statistical Learning Theory and Randomized Algorithms for Control" //IEEE Control Systems, No 12, 1998, pp.69–85