УДК 519.642.6

О РАСПРОСТРАНЕНИИ ГРАДИЕНТНОГО ПОДХОДА НА ОНЛАЙН ОПТИМИЗАЦИЮ РОБАСТНЫХ УПРАВЛЯЕМЫХ СИСТЕМ СО СКОЛЬЗЯЩИМ РЕЖИМОМ

А.В. Назин
Институт проблем управления им. В.А. Трапезникова РАН
Московский физико-технический институт (национальный исследовательский университет)
Россия, 117997, Москва, Профсоюзная ул., 65
Россия, 141701, Московская обл., Долгопрудный, Институтский пер., 9
E-mail: nazine@ipu.ru

А.С. Позняк
CINVESTAV-IPN
Mexico, CD Mexico, CP 07360, Av. IPN, 2508
E-mail: apoznyak@ctrl.cinvestav.mx

Ключевые слова: градиентные методы оптимизации, зеркальный спуск, задача онлайн оптимизации, нелинейный динамический объект, ОДУ, неопределенность.

Аннотация: Рассматривается класс управляемых нелинейных объектов, динамика которых определена системой ОДУ с частично неизвестной правой частью и аддитивным управлением. Ставится задача онлайн оптимизации, заключающаяся в нахождении робастной реализуемой стратегии отслеживания желаемой траектории вектора состояния с известными ограничениями на переменные состояния при условии, что эти переменные и их производные по времени наблюдаются. Для ее решения объединяются две ключевые идеи: (1) применение метода инерционного зеркального спуска (ЗС) как нетривиального обобщения метода проекции градиента, и (2) развитие метода интегрального скользящего режима для систем управления с непрерывным временем. Основные результаты включают доказательство того, что «желаемый режим» – нестационарный аналог поверхности скольжения, обеспечивающий ЗС, – может быть достигнут с самого начала процесса управления, а также получение явной верхней границы дефекта функции стоимости (ошибки слежения).

1. Введение

Задачи оптимизации и методы их итеративного решения интенсивно исследуются на протяжении последних десятилетий (см., например, [1–3]). Эти методы также были распространены на процессы оптимизации для класса статических объектов [4–7], где стратегии управления в большинстве публикаций, трактуемые как *методы статической оптимизации* (МСО), в непрерывном времени могут быть представлены в виде $F(x_t) \underset{t\to\infty}{\to} F^* := \min\limits_{x\in X_{adm}\subseteq\mathbb{R}^n} F(x)$, где $F : \mathbb{R}^n \to \mathbb{R}$ – выпуклая целевая функция, X_{adm} – выпуклое допустимое множество, а процесс $(x_t)_{t\geqslant 0}$ порождается ОДУ $\dot{x}_t = u_t,\ t \geqslant 0,\ \ x_0 \in \mathbb{R}^n$. Все известные процедуры МСО отличаются лишь построением управляющего воздействия u_t (или алгоритма оптимизации) в зависимости от текущего состояния x_t (стратегия Маркова) или более глубокой доступной истории, а именно: $u_t = u(t, x_\tau\mid_{\tau\in[0,t]})$.

Далее мы рассматриваем более общую и более сложную ситуацию, когда процесс $(x_t)_{t\geqslant 0}$ генерируется *нелинейным динамическим объектом* [8, 9]

$$\ddot{x}_t = \mathbf{f}\left(t, x_t, \dot{x}_t\right) + u_t, \quad t \geqslant 0, \quad x_0,\ \dot{x}_0 \text{ заданы},\ \ x_t, u_t \in \mathbb{R}^n, \tag{1}$$

где вектор-функция $\mathbf{f}$ в правой части предполагается неизвестной, но принадлежащей некоторому классу нелинейностей $\mathcal{C}$. Эта проблема больше похожа на задачу поиска экстремума [4], [10] , где нелинейная динамика содержит только производные первого порядка; см. также [11–13].

2. Нелинейный объект с неопределенностью

2.1. Динамическая модель

Динамическая модель второго порядка (1) (см., например, [8,11,13]) может быть представлена следующим расширенным форматом:

$$\begin{pmatrix} \dot{\mathbf{x}}_{1,t} \\ \dot{\mathbf{x}}_{2,t} \end{pmatrix} = \begin{pmatrix} \mathbf{x}_{2,t} \\ \mathbf{f}\left(t, \mathbf{x}_{1,t}, \mathbf{x}_{2,t}\right) \end{pmatrix} + \begin{pmatrix} 0_{n\times n} \\ I_{n\times n} \end{pmatrix} u_t, \tag{2}$$

где управление $u_t \in \mathbb{R}^n$, а начальные $\mathbf{x}_{1,t_0} = \mathring{\mathbf{x}}_1 \in \mathbb{R}^n$, $\mathbf{x}_{2,t_0} = \mathring{\mathbf{x}}_2 \in \mathbb{R}^n$. Здесь переменные расширенного пространства состояний $\mathbf{x}_{1,t} = x_t$, $\mathbf{x}_{2,t} = \dot{x}_t$ – текущие координаты и их производные по времени, $t \geqslant 0$. Неизвестная функция $\mathbf{f}(t, \mathbf{x}_1, \mathbf{x}_2)$ кусочно-непрерывная по всем переменным $t \geqslant 0$, $\mathbf{x}_1 \in \mathbb{R}^n$, $\mathbf{x}_2 \in \mathbb{R}^n$, и

$$\left\|\mathbf{f}\left(t, \mathbf{x}_1, \mathbf{x}_2\right)\right\|_2 \leqslant k_x\left(\mathbf{x}_1, \mathbf{x}_2\right) := c_0 + c_1\left\|\mathbf{x}_1\right\|_2 + c_2\left\|\mathbf{x}_2\right\|_2 \tag{3}$$

с конечными положительными константами c_0, c_1, c_2. Эта функция может включать неизмеряемые непотенциальные силы, такие как кулоновское трение, гистерезис, кориолис, демпфирование, центростремительные эффекты, а также учитывать ограничения на внешние возмущения.

2.2. Эталонная траектория и формулировка задачи

Цель контроллера (см. ниже) – реализовать отслеживание состояния $\mathbf{x}_t$ относительно заданной эталонной траектории $\{\mathbf{x}_t^*\}_{t\geqslant 0}$. Определим *ошибки отслеживания* как $\delta_{1,t} := \mathbf{x}_{1,t} - \mathbf{x}_{1,t}^*$, $\delta_{2,t} := \dot{\delta}_{1,t} = \mathbf{x}_{2,t} - \mathbf{x}_{2,t}^*$, где $\mathbf{x}_{1,t}^*$ – дважды непрерывно дифференцируемая траектория, которую нужно как можно точнее отслеживать, удовлетворяющую $\dot{\mathbf{x}}_{1,t}^* = \mathbf{x}_{2,t}^* = \varphi\left(t, \mathbf{x}_{1,t}^*\right)$, $t \geqslant 0$; $\mathbf{x}_{1,0}^*$ задано. С учетом этого динамику отслеживания ошибок можно представить как

$$\begin{gathered} \begin{pmatrix} \dot{\delta}_{1,t} \\ \dot{\delta}_{2,t} \end{pmatrix} = \begin{pmatrix} \delta_{2,t} \\ \mathbf{f}_\delta\left(t, \delta_{1,t}, \delta_{2,t}\right) \end{pmatrix} + \begin{pmatrix} 0_{n\times n} \\ I_{n\times n} \end{pmatrix} u_t, \\ \mathbf{f}_\delta\left(t, \delta_{1,t}, \delta_{2,t}\right) := \mathbf{f}\left(t, \delta_{1,t} + \mathbf{x}_{1,t}^*, \delta_{2,t} + \mathbf{x}_{2,t}^*\right) - \dot{\mathbf{x}}_{2,t}^*. \end{gathered} \tag{4}$$

Потребуем, чтобы динамика $\delta_{1,t}$ реализовывалась за время $t \geqslant t_0 \geqslant 0$ в пределах ограниченного допустимого множества $\mathcal{D}_{adm} \subseteq \mathbb{R}^n$.

Формулировка задачи: *Основная цель – разработать робастное управление (что означает успешную работу замкнутой системы на классе неопределенности (3)), минимизирующее ошибку отслеживания $\delta_{1,t}$ в смысле функциональной верхней границы дефекта отслеживания $F(\delta_{1,t}) - \min_{\delta_1 \in \mathcal{D}_{adm}} F(\delta_1) \leqslant O(t^{-1})$.*

Примеры целевых функций:

$$(5)\qquad \begin{gathered} 1)\; F\left(\delta_1\right) = \sum_{i=1}^{n} |\delta_{1,i}| = \|\delta_1\|_1\,, \\ 2)\; F\left(\delta_1\right) = \sum_{i=1}^{n} |\; \delta_{1,i}|_{\varepsilon}^{+}\,, \quad |z|_{\varepsilon}^{+} := \max\{0, |z| - \varepsilon\}. \end{gathered}$$

3. Желаемая динамика и МЗС

Далее используем обозначение допустимого множества $X_{adm} := \mathcal{D}_{adm}$ и введем основные предположения **A1** – **A5**.

A1 Вектор состояний и их производные $(\mathbf{x}_t, \dot{\mathbf{x}}_t) \in \mathbb{R}^n \times \mathbb{R}^n$ объекта (2) предполагаются измеряемыми онлайн и непрерывно дифференцируемыми при любом $t \geqslant 0$.

A2 Функция $\mathbf{f} : \mathbb{R}^{1+2n} \to \mathbb{R}^n$ с учетом (3) кусочно непрерывна и неизвестна.

A3 Вектор состояний эталонной траектории и его производная $(\mathbf{x}_t^*, \dot{\mathbf{x}}_t^*) \in \mathbb{R}^n \times \mathbb{R}^n$ также доступны онлайн и непрерывно дифференцируемы для каждого $t \geqslant 0$.

A4 Субградиент $F' : \mathbb{R}^n \to \mathbb{R}^n$ функции потерь $F : \mathbb{R}^n \to \mathbb{R}$ доступен онлайн на ошибке слежения $\delta_{1,t}$ $\forall\, t \geqslant 0$, и что множество оптимизаторов δ_1^* функции $F(\cdot)$ на $\mathcal{D}_{adm}$ включает в себя начало координат $\delta_1^* = 0$, т.е. $0 \in Arg \min\limits_{\delta_1 \in \mathcal{D}_{adm}} F(\delta_1)$.

A5 Допустимое множество $\mathcal{D}_{adm}$ является непустым выпуклым компактом.

Приведем инерционный МЗС в непрерывном времени [14]. Для этого определим динамику сопряженной переменной $\zeta_t \in \mathbb{R}^n$ и исходной переменной $\delta_{1,t} \in \mathbb{R}^n$ как

$$(6)\qquad \begin{gathered} \dot{\zeta}_t = -F'(\delta_{1,t}), \quad F'(\delta_{1,t}) \in \partial F(\delta_{1,t}), \quad \zeta_{t_0} = 0, \\ (t+\theta)\,\dot{\delta}_{1,t} + \delta_{1,t} = \nabla U_*\left(\zeta_t - \eta\right), \quad t \geqslant t_0 \geqslant 0,\ \eta \in \mathbb{R}^n,\ \theta > 0. \end{gathered}$$

Важно отметить, что интегрирование второго уравнения в (6) на интервале $[t_0, t]$ показывает, что $\delta_{1,t} \in \mathcal{D}_{adm}$ $\forall\, t \geqslant t_0$ из-за свойств выпуклости. Следующая теорема объясняет, почему динамику $\delta_{1,t}$ в (6) можно считать желаемой.

Теорема 1. *Если в предположениях **A1**–**A5** прокси-функция $U(\cdot)$ выбрана так, что ее минимум на допустимом множестве $\mathcal{D}_{adm}$ достигается в начале координат, то на траекториях $\delta_{1,t}$, сгенерированных* (6), *выполняется*

$$(7)\qquad F(\delta_{1,t}) - \min_{\delta_1 \in \mathcal{D}_{adm}} F(\delta_1) \leqslant \left[F(\delta_{1,t_0}) - F(0)\right] \frac{t_0+\theta}{t+\theta}, \quad \forall\, t \geqslant t_0 \geqslant 0.$$

Замечание 1. *Отметим, что результат* (7) *не зависит явно от прокси функции $U(\cdot)$. Действительно, определив $\delta_1^*(\eta) := \arg \min\limits_{\delta_1 \in \mathcal{D}_{adm}} \{-\eta^\intercal \delta_1 + U\left(\delta_1\right)\} = \nabla U_*\left(\eta\right)$, получим $\delta_1^*\left(\eta\right)|_{\eta=0} = \arg \min\limits_{\delta_1 \in \mathcal{D}_{adm}} U\left(\delta_1\right) = 0$.*

4. Синтез робастного контроллера

4.1. Вспомогательная скользящая переменная и ее динамика

Введем вспомогательную *скользящую переменную* $s_t = (t+\theta)\,\delta_{2,t} + \delta_{1,t} - \nabla U_*\,(\zeta_t - \eta)$, $t \geqslant t_0 \geqslant 0$. Заметим, что s_t измеряема онлайн, и что ситуация, когда $s_t = 0$ для всех $t \geqslant t_0$, точно соответствует желаемому режиму (6) с момента t_0. Тогда для $V\,(s_t) = \frac{1}{2}\left\|s_t\right\|_2^2$ в силу уравнений динамического объекта (4) и первого уравнения в (6) имеем

$$\begin{aligned}\frac{d}{dt}V\,(s_t) = s_t^\intercal \dot{s}_t = s_t^\intercal \left[2\dot{\delta}_{1,t} + (t+\theta)\,\dot{\delta}_{2,t} - \frac{d}{dt}\nabla U_*\,(\zeta_t-\eta)\right] = \\ = (t+\theta)\,s_t^\intercal \mathbf{f}\left(t,\delta_{1,t}+x_{1,t}^*,\delta_{2,t}+x_{2,t}^*\right) + \\ + (t+\theta)\,s_t^\intercal \underbrace{\left[\frac{2}{t+\theta}\delta_{2,t} - \dot{x}_{2,t}^* + u_t + \frac{1}{t+\theta}\nabla^2 U_*\,(\zeta_t-\eta)\,F'(\delta_{1,t})\right]}_{-k_t \mathrm{Sign}(s_t)} \leqslant \\ \leqslant (t+\theta)\left(\left\|s_t\right\|_2 \underbrace{\left(c_0 + c_1\left\|\delta_{1,t}+x_{1,t}^*\right\|_2 + c_2\left\|\delta_{2,t}+x_{2,t}^*\right\|_2\right)}_{k_{x,t}:=k_x(\delta_{1,t}+x_{1,t}^*,\delta_{2,t}+x_{2,t}^*)} - k_t s_t^\intercal \mathrm{Sign}\,(s_t)\right).\end{aligned}$$

Здесь вектор $\mathrm{Sign}\,(s_t) = \left(\mathrm{sign}\,(s_{1,t})\,,...,\mathrm{sign}\,(s_{n,t})\right)^\intercal$, где $\mathrm{sign}\,(s_{i,t}) = +1$ при $s_{i,t} > 0$, $\mathrm{sign}\,(s_{i,t}) = -1$ при $s_{i,t} < 0$, и $\mathrm{sign}\,(s_{i,t}) \in [-1,+1]$ при $s_{i,t} = 0$.

4.2. Структура робастного управления и основной результат

Поскольку $s_t^\intercal \mathrm{Sign}\,(s_t) = \sum\limits_{i=1}^{n}|s_{i,t}| \geqslant \left\|s_t\right\|_2$ и, вводя $k_t = k_{x,t} + \rho$, $\rho > 0$, получаем при $s_t \neq 0$, т.е. $V(s_t) > 0$,

$$\frac{d}{dt}V\,(s_t) \leqslant (t+\theta)\left\|s_t\right\|_2 (k_{x,t} - k_t) = -\,(t+\theta)\,\rho\sqrt{2V\,(s_t)},$$

что дает после интегрирования на интервале $[t_0, t]$

$$\begin{aligned}2\left(\sqrt{V\,(s_t)} - \sqrt{V\,(s_{t_0})}\right) \leqslant -\frac{\sqrt{2}}{2}\rho\left[(t+\theta)^2 - (t_0+\theta)^2\right], \\ 0 \leqslant \sqrt{V\,(s_t)} \leqslant \sqrt{V\,(s_{t_0})} - \frac{\sqrt{2}}{4}\rho\left[(t+\theta)^2 - (t_0+\theta)^2\right].\end{aligned}$$

Это означает, что $V(s_t) = 0$ для всех $t \geqslant t_{reach} = \sqrt{\frac{2}{\rho}\left\|s_{t_0}\right\|_2 + (t_0+\theta)^2} - \theta$. Наконец, получаем робастное управление $u_t = u_{comp,t} + u_{disc,t}$ с двумя компонентами, компенсационной и разрывной:

$$u_{comp,t} = -\frac{2}{t+\theta}\delta_{2,t} + \dot{x}_{2,t}^* - \frac{1}{t+\theta}\nabla^2 U_*\,(\zeta_t-\eta)\,F'(\delta_{1,t}),\ u_{disc,t} = -k_t \mathrm{Sign}\,(s_t)\,. \tag{8}$$

Замечание 2. *Желание получить* $t_{reach} = t_0 = 0$ *приводит к равенству*

$$s_0 = \theta\delta_{2,0} + \delta_{1,0} - \nabla U_*\,(-\eta) = \theta\delta_{2,0} + \delta_{1,0} - \delta_1^*\,(\eta) = 0.$$

Поскольку $\delta_1^*\,(\eta) \in \mathcal{D}_{adm}$, *мы заключаем, что параметры* $\theta > 0$, η *и начальные условия* $(\delta_{1,0},\delta_{2,0})$ *должны быть согласованы в смысле включения* $\theta\delta_{2,0} + \delta_{1,0} \in \mathcal{D}_{adm}$.

Теорема 2. *Если в предположениях* ***A1***–***A5*** *прокси-функция* $U(\cdot)$ *выбрана так, что ее минимум на множестве* $\mathcal{D}_{adm}$ *достигается в начале координат, то на траекториях* $\delta_{1,t}$ *в объекте* (4), *управляемом* $u_t = u_{comp,t} + u_{disc,t}$ (8), *выполняется*

$$F(\delta_{1,t}) - \min_{\delta_1 \in \mathcal{D}_{adm}} F(\delta_1) \leqslant [F(\delta_{1,0}) - F(0)] \frac{\theta}{t+\theta}, \quad \forall\, t \geqslant t_0 = 0.$$

5. Заключение

Таким образом, задача оптимизации слежения за эталонной траекторией при заданном допустимом выпуклом компакте решается с использованием дифференциально управляемого многомерного объекта 2-го порядка с неизвестной ограниченной правой частью модели. Желаемая динамика переменных ошибки отслеживания рассчитывается на основе инерционного МЗС. Установлена сходимость целевой функции к минимуму и получена связанная с ней неасимптотическая верхняя граница. Доказано, что данный робастный регулятор при определенной связи его параметров с начальными условиями обеспечивает определенный скользящий режим с начала процесса управления. Этот метод может иметь несколько применений при разработке робастного управления в механических системах [8, 9].

Работа выполнена при частичной финансовой поддержке Российского научного фонда (проект № 21-71-30005), https://rscf.ru/project/21-71-30005/.

Список литературы

1. Поляк Б.Т. Введение в оптимизацию. М.: Наука, 1983. 384 с.
2. Нестеров Ю.Е. Введение в выпуклую оптимизацию. М.: МЦНМО, 2010. 278 с.
3. Boyd S., Vandenberghe L. Convex optimization. Cambridge University Press, 2004. 727 с.
4. Растригин Л.А. Системы экстремального управления. М.: Наука, 1974. 632 с.
5. Dechter R. Constraint Processing. Morgan Kaufmann Publisher, 2003. 481 с.
6. Leader J.J. Numerical Analysis and Scientific Computation. Pearson Addison Wesley, 2004. 590 с.
7. Sieniutycz S., Jezowski J. Brief review of static optimization methods // Energy Optimization in Process Systems. Elsevier, 2009. P. 1–43.
8. Poznyak A.S., Nazin A.V., Alazki H. Integral Sliding Mode Convex Optimization in Uncertain Lagrangian Systems Driven by PMDC Motors: Averaged Subgradient Approach // IEEE Trans. Autom. Control. 2021. Vol. AC-66, No. 9. P. 4267–4273.
9. Nazin A.V., Alazki H., Poznyak A.S. Robust Tracking as Constrained Optimization by Uncertain Dynamic Plant: Mirror Descent Method and ASG–Version of Integral Sliding Mode Control // Mathematics. 2023. Vol. 11. P. 4112.
10. Solis C.U., Clempner J.B., Poznyak A.S. Extremum seeking by a dynamic plant using mixed integral sliding mode controller with synchronous detection gradient estimation // Int. J. of Robust and Nonlinear Control. 2018. Vol. 29, No. 3. P. 702–714.
11. Utkin V. Sliding Modes in Control Optimization. Berlin: Springer, 1992. 286 p.
12. Fridman L., Poznyak A., Bejarano F.J. Robust Output LQ Optimal Control via Integral Sliding Modes. New York: Birkhäuser, 2014. 161 p.
13. Utkin V., et al. Road Map for Sliding Mode Control Design. Springer Briefs in Mathematics. Cham, Switzerland: Springer, 2020. 127 p.
14. Назин А.В. Алгоритмы инерционного зеркального спуска в выпуклых задачах стохастической оптимизации // Автомат. и телемех. 2018. № 1. С. 100–112.
15. Юдицкий А.Б., Назин А.В., Цыбаков А.Б., Ваятис Н. Рекуррентное агрегирование оценок методом зеркального спуска с усреднением // Пробл. передачи информ. 2005. Т. 41, № 4. С. 78–96.