



Российская Академия Наук

А АВТОМАТИКА И ДАЛЕКОУПРАВЛЕНИЕ AUTOMATION AND REMOTE CONTROL

Журнал основан в 1936 году

Выходит 12 раз в год

9

сентябрь

Москва

2025

Учредители журнала:

Отделение энергетики, машиностроения, механики и процессов управления РАН,
Институт проблем управления им. В.А. Трапезникова РАН (ИПУ РАН),
Институт проблем передачи информации им. А.А. Харкевича РАН (ИППИ РАН)

Главный редактор:

Галяев А.А.

Заместители главного редактора:

Рубинович Е.Я., Хлебников М.В.

Ответственный секретарь:

Самохин А.С.

Редакционный совет:

Васильев С.Н., Желтов С.Ю., Каляев И.А., Кузнецов О.П., Кулешов А.П.,
Куржанский А.Б., Мартынюк А.А. (Украина), Пешехонов В.Г., Попков Ю.С.,
Федосов Е.А., Черноусько Ф.Л.

Редакционная коллегия:

Алескеров Ф.Т., Арутюнов А.В., Бахтадзе Н.Н., Бобцов А.А., Виноградов Д.В.,
Вишневский В.М., Воронцов К.В., Граничин О.Н., Жилякова Н.Ю.,
Каравай М.Ф., Кибзун А.И., Краснова С.А., Крищенко А.П., Кузнецов Н.В.,
Кушнер А.Г., Лазарев А.А., Ляхов А.И., Маликов А.И., Матасов А.И.,
Меерков С.М. (США), Мещеряков Р.В., Миллер Б.М., Михальский А.И.,
Моржин О.В., Мунасыпов Р.А., Назин А.В., Немировский А.С. (США),
Новиков Д.А., Олейников А.Я., Пакшин П.В., Пальчунов Д.Е.,
Поляков А.Е. (Франция), Протасов В.Ю., Рапопорт Л.Б., Родионов И.В.,
Сельвесюк Н.И., Соболевский А.Н., Степанов О.А., Филимонюк Л.Ю.,
Фрадков А.Л., Хоров Е.М., Цыбаков А.Б. (Франция), Чеботарев П.Ю.,
Чхартишвили А.Г., Щербаков П.С.

Адрес редакции: 117997, Москва, Профсоюзная ул., 65

Тел./факс: 8 (495) 198-17-20, доб. 1443

Электронная почта: ait@pran.ru, redacsia@ipu.ru

Сайт: <https://aitras.ru>

Зав. редакцией *Е.А. Мартехина*

Москва

ФГБУ «Издательство «Наука»

Тематический выпуск (окончание)¹

© 2025 г. С.С. АБЛАЕВ, канд. физ.-мат. наук (seydamet.ablaev@yandex.ru)

(Крымский федеральный университет им. В.И. Вернадского, Симферополь;

Московский физико-технический институт

(национальный исследовательский университет)),

Ф.С. СТОНЯКИН, д-р физ.-мат. наук (fedyor@mail.ru)

(Московский физико-технический институт

(национальный исследовательский университет);

Крымский федеральный университет им. В.И. Вернадского, Симферополь),

М.Н. ФЕДОТОВ (MaximFd-Nk@yandex.ru)

(Национальный исследовательский университет

«Высшая школа экономики», Москва),

М.С. АЛКУСА, канд. физ.-мат. наук (m.alkousa@innopolis.ru)

(Университет Иннополис),

О.С. САВЧУК (oleg.savchuk19@mail.ru)

(Крымский федеральный университет имени В.И. Вернадского, Симферополь;

Адыгейский государственный университет, Майкоп),

А.В. ГАСНИКОВ, д-р физ.-мат. наук (gasnikov@yandex.ru)

(Университет Иннополис, Долгопрудный;

Московский физико-технический институт

(национальный исследовательский университет))

ИССЛЕДОВАНИЕ ГРАДИЕНТНОГО МЕТОДА С НЕТОЧНОЙ ИНФОРМАЦИЕЙ О ГРАДИЕНТЕ НА НЕКОТОРЫХ КЛАССАХ (L_0, L_1)-ГЛАДКИХ НЕВЫПУКЛЫХ ЗАДАЧ²

Работа посвящена исследованию градиентного метода на классе (L_0, L_1) -гладких функций при условии, что на итерациях метода вместо точных значений градиента доступны лишь его приближенные значения, что соответствует ситуациям, возникающим при использовании зашумленных данных. Рассмотрено два класса задач: первый – квазар-выпуклые функции относительно всякого решения, удовлетворяющие условию градиентного доминирования Поляка–Лоясиевича, второй – квазар-выпуклые функции без требования выполнения условия Поляка–Лоясиевича, но с дополнительным ограничением на параметр квазар-выпуклости. Для квазар-выпуклых функций с РЛ-условием доказан результат о близкой к линейной скорости сходимости метода в окрестности точного решения. Если значения неточного градиента достаточно малы (что достигается за конечное число итераций), то метод сходится

¹ Статьи с 3 стр. по 63 стр. являются окончанием тематического выпуска № 8, 2025.

² Исследование выполнено при финансовой поддержке Министерства науки и высшего образования РФ в рамках государственного задания № 075-03-2024-074 по проекту «Исследование асимптотических характеристик колеблемости дифференциальных уравнений и систем, а также оптимизационных методов».

с близкой к линейной скоростью на классе задач с условием Поляка–Лоясиевича без дополнительного предположения о квазар-выпуклости. Для $(0, M)$ -гладких квазар-выпуклых функций предложен адаптивный градиентный метод и получена оценка его скорости сходимости. Показано, что в случае использования точных значений градиента метод сходится с линейной скоростью.

Ключевые слова: градиентный метод, Δ -неточный градиент, (L_0, L_1) -гладкая функция, ρ -квазар-выпуклая функция, условие Поляка–Лоясиевича, логистическая регрессия.

DOI: 10.31857/S0005231025090015, **EDN:** VMPJNF

1. Введение

В связи с активным развитием машинного обучения возрастает актуальность разработки и анализа численных методов оптимизации с эффективными скоростными гарантиями. Для достижения оптимальной скорости сходимости традиционных методов, таких как градиентный спуск и его модификации, ключевым предположением является гладкость целевой функции (т.е. липшицевость ее градиента). В контексте современных задач машинного обучения такое допущение может оказаться слишком жестким [1]. Часто достаточно естественные и простые с точки зрения реализации методы плохо работают для задач в глубоком обучении, где глобальное условие гладкости зачастую не выполняется. Например, методы редукции дисперсии [2–8], как известно, являются более быстрыми с точки зрения теории (для задач минимизации конечных сумм гладких функций), но на практике могут уступать методам, теоретически не снижающим дисперсию [9]. В экспериментах, приведенных в [1], показано, что при обучении нейронных сетей норма гессиана может коррелировать с нормой градиента функции потерь. Таким образом, вышеизложенные факторы стимулируют поиск и разработку альтернативных предположений, ослабляющих стандартное требование гладкости целевой функции.

Одним из таких предположений является так называемая обобщенная (L_0, L_1) -гладкость, впервые предложенная несколько лет назад в [1] для дважды дифференцируемых функций. Это предположение позволяет норме гессиана целевой функции расти линейно с ростом нормы градиента. В частности, (L_0, L_1) -гладкость может сохраняться даже для функций с полиномиально растущими градиентами – типичная ситуация для задач глубокого обучения. Более того, концепция (L_0, L_1) -гладкости охватывает широкий класс задач и допускает обобщение на дифференцируемые, но не обязательно дважды дифференцируемые функции [10, 11].

В последние годы во многих работах исследовались различные методы решения задач (L_0, L_1) -гладкой оптимизации. Однако, несмотря на интерес со стороны научного сообщества, существующие результаты относительно сходимости методов в ряде важных случаев остаются неоптимальными, а теоретический анализ зачастую является недостаточно полным. В то время как

авторы [1], как и далее в [10–16], сосредоточились на максимально общем классе невыпуклых задач, допускающих лишь сходимость к нулю метрик типа нормы градиента, класс (L_0, L_1) -гладких выпуклых функций до совсем недавнего времени (конец 2024 – начало 2025 гг.) был изучен гораздо слабее. В частности, известные результаты о сходимости методов, таких как градиентный спуск с клиппингом [17] и градиентный спуск с шагом Поляка [18], в применении к (L_0, L_1) -гладким выпуклым задачам либо требуют дополнительного предположения об L -гладкости [19, 20], либо накладывают ограничения на размер шага, необходимые для обеспечения ограниченности траектории метода в области, где градиент удовлетворяет условию Липшица в силу (L_0, L_1) -гладкости целевой функции [21].

Приведем обзор нескольких работ последнего времени, посвященных оптимизации (L_0, L_1) -гладких функций.

В [22] авторы сосредоточились на исследовании класса сильно выпуклых (L_0, L_1) -гладких функций и получили новые гарантии скорости сходимости для нескольких существующих методов. В частности, были получены улучшенные оценки скорости сходимости для градиентного спуска с (сглаженным) клиппингом и для градиентного спуска с шагом Б.Т. Поляка. В отличие от уже существующих результатов эти оценки не основаны на стандартном предположении о гладкости и не имеют экспоненциальной зависимости от расстояния от начальной точки до решения.

Параллельно с [22] в [23] независимо были получены аналогичные результаты для градиентного метода с различными стратегиями выбора шага, включая шаг Поляка, а также получены новые оценки сложности для градиентного метода с нормированным шагом на классе (L_0, L_1) -гладких задач.

В [24] представлен более детальный анализ сходимости градиентного спуска и его вариантов в условиях обобщенной (L_0, L_1) -гладкости. В частности, показано, что в выпуклом случае скорость сходимости градиентного метода меняется в зависимости от нормы градиента: при $\|\nabla f(x_k)\| \geq \frac{L_0}{L_1}$ достигается линейная скорость, а при $\|\nabla f(x_k)\| < \frac{L_0}{L_1}$ наблюдается стандартная сублинейная скорость сходимости.

В отличие от рассмотренных выше работ в настоящем исследовании анализируется градиентный метод для (L_0, L_1) -гладких задач в условиях, когда на каждой итерации доступны лишь приближенные (аддитивно неточные) значения градиента. То есть получила развитие в новом направлении статья [25], где аналогичные проблемы исследовались для гладких задач в классическом понимании. При этом рассматривается несколько направлений. Во-первых, исследована сходимость метода для ρ -квазар-выпуклых относительно любого точного решения [26–28] (L_0, L_1) -гладких функций, удовлетворяющих условию Поляка–Лоясиевича

$$(1) \quad f(x) - f(x_*) \leq \frac{1}{2\mu} \|\nabla f(x)\|^2 \quad \forall x \in \mathbb{R}^n, \mu > 0,$$

где $\|\cdot\|$ – евклидова норма, x_* – точка минимума функции f . В работе показано, что при некоторых предположениях метод сходится с линейной скоростью в некоторую окрестность точного решения. Доказано, что начиная с некоторого номера итерации метод сходится на классе задач с условием Поляка–Лоясиевича (без требования квазар-выпуклости). Во-вторых, отдельно рассматривается класс ρ -квазар-выпуклых $(0, M)$ -гладких функций, представляющий интерес в контексте задач машинного обучения, в частности при обучении моделей логистической регрессии. Для такого класса задач предложен адаптивный вариант градиентного метода с использованием неточных градиентов на итерациях, и получены оценки скорости сходимости при некотором ограничении на параметр квазар-выпуклости ρ . Эффективность предложенных методов подтверждена численными экспериментами на задачах логистической регрессии (удовлетворяющей условию Поляка–Лоясиевича на любом компактном множестве [29]) и на некоторой невыпуклой квазар-выпуклой задаче [27].

2. Постановка задачи и необходимые вспомогательные понятия

Будем рассматривать задачи минимизации вида

$$(2) \quad \min_{x \in \mathbb{R}^n} f(x),$$

где $f : \mathbb{R}^n \rightarrow \mathbb{R}$ – выпуклая и (L_0, L_1) -гладкая функция. Всюду в работе x_* будет обозначать точку глобального минимума, а $f^* = f(x_*)$ – глобальное минимальное значение функции, норма $\|\cdot\|$ – евклидова.

Как уже было отмечено, условие (L_0, L_1) -гладкости ($L_0, L_1 > 0$) было первоначально предложено в [1] для дважды дифференцируемых функций $f : \mathbb{R}^n \rightarrow \mathbb{R}$ следующим образом:

$$(3) \quad \|\nabla^2 f(x)\| \leq L_0 + L_1 \|\nabla f(x)\| \quad \forall x \in \mathbb{R}^n,$$

где $\|\cdot\|$ обозначает евклидову норму для векторов и спектральную норму для матриц.

В [10] это понятие было распространено на класс дифференцируемых, но необязательно дважды дифференцируемых, функций.

Определение 1. Функция $f : \mathbb{R}^n \rightarrow \mathbb{R}$ называется (L_0, L_1) -гладкой, если $\forall x, y \in \mathbb{R}^n$ при условии $\|y - x\| \leq \frac{1}{L_1}$ справедливо неравенство

$$(4) \quad \|\nabla f(x) - \nabla f(y)\| \leq (L_0 + L_1 \|\nabla f(y)\|) \|x - y\|.$$

Рассмотрим несколько примеров функций, которые удовлетворяют условию (4).

Пример 1. Функция $f(x) = \|x\|^{2m}$ является $(2m, 2m - 1)$ -гладкой. При этом f не является L -гладкой (напомним, что L -гладкой называется функция, градиент которой удовлетворяет условию Липшица) при $m \geq 2$ для всех $L \geq 0$ (см. [22]).

Пример 2. Функция $f(x) = e^{ax}$, $a \in \mathbb{R}^n$ является $(0, \|a\|)$ -гладкой. При этом $f(x)$ не является L -гладкой при $a \neq 0$ для всех $L \geq 0$ (см. [22]).

Пример 3. Логистическая функция вида $f(x) = \log(1 + \exp(-a^T x))$, где $a \in \mathbb{R}^n$, является (L_0, L_1) -гладкой с $L_0 = 0$ и $L_1 = \|a\|$ (см. [22]).

Понятие (L_0, L_1) -гладкости было обобщено в [11] следующим образом: $f : \mathbb{R}^n \rightarrow \mathbb{R}$ называется α -симметрично обобщенно гладкой функцией, если при некотором $\alpha \in [0, 1] \forall x, y \in \mathbb{R}^n$ выполняется условие

$$(5) \quad \|\nabla f(y) - \nabla f(x)\| \leq \left(L_0 + L_1 \max_{\theta \in [0, 1]} \|\nabla f(\theta y + (1 - \theta)x)\|^\alpha \right) \|y - x\|.$$

В [11] показано, что условие (5) является более общим, чем (3) и (4): из выполнения (3) или (4) следует выполнение (5) (при $\alpha = 1$). При этом для дважды дифференцируемых функций условия (3) и (5) эквивалентны при $\alpha = 1$ [11].

Приведем определение квазар-выпуклой функции, которое является одним из ключевых в данной работе.

Определение 2. Пусть $\rho \in (0, 1]$ и $\mu > 0$. Функция $f : \mathbb{R}^n \rightarrow \mathbb{R}$ называется (ρ, μ) -квазар сильно выпуклой, если для всех $x \in \mathbb{R}^n$ справедливо неравенство

$$(6) \quad f(x_*) \geq f(x) + \frac{1}{\rho} \langle \nabla f(x), x_* - x \rangle + \frac{\mu}{2} \|x_* - x\|^2,$$

где x_* – точка глобального минимума функции f . При $\mu = 0$ функция f называется ρ -квазар выпуклой функцией.

Условие Поляка–Лоясиевича можно рассматривать не только как релаксацию обычной сильной выпуклости, но и сильной квазар-выпуклости [27]: если $f : \mathbb{R}^n \rightarrow \mathbb{R}$ есть (ρ, μ) -квазар сильно выпуклая функция, то она удовлетворяет условию Поляка–Лоясиевича (1) с константой $\mu\rho^2$.

Пример 4 [27]. Пусть $f : \mathbb{R}^n \rightarrow \mathbb{R}$ определяется как

$$f(x) = \|x\|^2 g\left(\frac{x}{\|x\|}\right), \quad \forall x \in \mathbb{R}^n,$$

где

$$g(x) = 1 + \sum_{i=1}^n a_i \sin(b_i x_i)^2, \quad \forall x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Тогда функция f является $(1, 2)$ -квазар сильно выпуклой, т.е. удовлетворяет PL-условию (1) с константой 2.

Вопросы сходимости градиентного метода для выпуклых задач с обобщенным условием гладкости (4) активно исследовались в последнее время, например, в [22–24]. В этих работах получены результаты о сходимости нормированного градиентного метода, градиентного метода с шагом Поляка и

ускоренного градиентного метода исходя из предположения обобщенной гладкости. Также было показано, что градиентный метод сходится с линейной скоростью сходимости до тех пор, пока значения норм градиентов на итерациях остаются достаточно большими. Данная статья посвящена некоторому развитию этих работ.

Здесь сосредоточимся на исследовании градиентного спуска для задач минимизации вида (2) в условиях, когда вместо точных значений градиента на каждой итерации метода доступен лишь градиент, содержащий аддитивную погрешность. Напомним, что вектор $\tilde{\nabla}f(x)$ называется неточным градиентом в точке x при $\Delta > 0$, если

$$\|\nabla f(x) - \tilde{\nabla}f(x)\| \leq \Delta.$$

В случае малых значений норм градиента на итерациях метода становится проблематично получить теоретические оценки о сходимости. Это обусловлено тем, что в отличие от случая использования точных градиентов невозможно гарантировать монотонное убывание нормы градиента на каждой итерации. Эту проблему можно обойти, если дополнительно потребовать выполнение условия Поляка–Лоясиевича для целевой функции f . Как показано в [25], для задач с неточными градиентами из условия Поляка–Лоясиевича следует неравенство

$$(7) \quad \|\tilde{\nabla}f(x)\|^2 \geq \mu(f(x) - f^*) - \Delta^2 \quad \forall x \in \mathbb{R}^n, \mu > 0.$$

В данном разделе приведены постановка задачи и необходимые предварительные сведения. В разделе 3 исследован градиентный метод для (L_0, L_1) -гладких задач при наличии аддитивной погрешности в градиенте на каждой итерации. Получена оценка скорости сходимости метода на классе ρ -квазар-выпуклых функций, удовлетворяющих условию Поляка–Лоясиевича. Показано, что до определенного уровня значений градиента скорость сходимости метода линейная. В разделе 4 исследован адаптивный вариант градиентного метода для $(0, M)$ -гладких ρ -квазар-выпуклых задач. В разделе 5 приведены результаты вычислительных экспериментов для задачи логистической регрессии (удовлетворяющей на всяком компакте условию Поляка–Лоясиевича [29]) и для некоторой квазар-выпуклой задачи [27], которая не выпукла.

3. Градиентный метод для (L_0, L_1) -гладких ρ -квазар выпуклых функций

Пусть $f : \mathbb{R}^n \rightarrow \mathbb{R}$ — (L_0, L_1) -гладкая (4) ρ -квазар выпуклая относительно любого решения x_* функция (см. определение 2, $\mu = 0$):

$$(8) \quad f(x_*) \geq f(x) + \frac{1}{\rho} \langle \nabla f(x), x_* - x \rangle, \quad \rho \in (0, 1].$$

Если f сильно квазар-выпукла (см. определение 2, $\mu > 0$), то решение x_* единственно, и тогда f сильно квазар-выпукла относительно любого решения. Отметим, что задачи с невыпуклыми гладкими (а значит, и (L_0, L_1) -гладкими) ρ -квазар-выпуклыми функциями нередко встречаются в приложениях (см. пример 4) [27].

Для решения задачи (2) исследуем градиентный метод вида

$$(9) \quad x_{k+1} = x_k - \eta_k \widetilde{\nabla} f(x_k)$$

с шагом

$$(10) \quad \eta_k = \frac{\alpha}{L_0 + L_1 \left(\|\widetilde{\nabla} f(x_k)\| + \Delta \right)},$$

где $\alpha \in (0, 1]$, а $\widetilde{\nabla} f(x_k)$ обозначает неточный градиент функции f в точке x_k .

Важно отметить, что, несмотря на использование приближенных значений градиента в алгоритме, в теоретическом анализе предполагается существование точного градиента.

Докажем несколько лемм, необходимых для обоснования оценок скорости сходимости метода.

Лемма 1 [10, Lemma A3]. Пусть функция $f(x)$ (L_0, L_1) -гладкая (4). Тогда $\forall x, y \in \mathbb{R}^n$ при условии $\|x - y\| \leq \frac{1}{L_1}$ справедливо неравенство

$$(11) \quad f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L_0 + L_1 \|\nabla f(x)\|}{2} \|y - x\|^2.$$

Доказательство. Пусть $g(t) = f(x + t(y - x))$ при $t \in [0, 1]$. Тогда

$$f(y) - f(x) = g(1) - g(0) = \int_0^1 g'(t) dt = \int_0^1 \langle \nabla f(x + t(y - x)), y - x \rangle dt.$$

Добавим и вычтем $\langle \nabla f(x), y - x \rangle$ под интегралом:

$$f(y) - f(x) = \int_0^1 \langle \nabla f(x), y - x \rangle dt + \int_0^1 \langle \nabla f(x + t(y - x)) - \nabla f(x), y - x \rangle dt.$$

Вычисляя первый интеграл и применяя неравенство Коши–Буняковского для выражения под вторым интегралом, получим следующее неравенство:

$$f(y) - f(x) \leq \langle \nabla f(x), y - x \rangle + \int_0^1 \|\nabla f(x + t(y - x)) - \nabla f(x)\| \|y - x\| dt.$$

Учитывая, что

$$\|\nabla f(y) - \nabla f(x)\| \leq (L_0 + L_1 \|\nabla f(x)\|) \|y - x\|,$$

имеем неравенство

$$f(y) - f(x) \leq \langle \nabla f(x), y - x \rangle + \|y - x\| \int_0^1 (L_0 + L_1 \|\nabla f(x)\|) \|y - x\| t dt,$$

откуда следует

$$f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L_0 + L_1 \|\nabla f(x)\|}{2} \|y - x\|^2.$$

Замечание 1. При наличии аддитивной неточности в градиенте неравенство (11) принимает вид

$$(12) \quad \begin{aligned} f(y) &\leq f(x) + \langle \tilde{\nabla} f(x), y - x \rangle + \\ &+ \frac{L_0 + L_1 (\|\tilde{\nabla} f(x)\| + \Delta)}{2} \|y - x\|^2 + \Delta \|y - x\|. \end{aligned}$$

Лемма 2. Пусть $f(x)$ — (L_0, L_1) -гладкая функция (4). Предположим, что $\|\nabla f(x_k)\| > \Delta$. Тогда на итерациях алгоритма (9) с шагом (10) значения функции монотонно убывают $f(x_{k+1}) \leq f(x_k)$, причем справедливо неравенство

$$(13) \quad f(x_k) - f(x_{k+1}) \geq \frac{\alpha(2 - \alpha)}{2} \frac{(\|\nabla f(x_k)\| - \Delta)^2}{L_0 + L_1 \|\nabla f(x_k)\| + 2L_1 \Delta}.$$

Доказательство. Положим в неравенстве (11) $y = x_{k+1}$, $x = x_k$. Отметим, что $\|x_{k+1} - x_k\| \leq \frac{1}{L_1}$. Тогда

$$\begin{aligned} f(x_{k+1}) - f(x_k) &\leq \\ &\leq \frac{-\alpha \langle \nabla f(x_k), \tilde{\nabla} f(x_k) \rangle}{L_0 + L_1 (\|\tilde{\nabla} f(x_k)\| + \Delta)} + \frac{L_0 + L_1 (\|\tilde{\nabla} f(x_k)\| + \Delta)}{2 (L_0 + L_1 (\|\tilde{\nabla} f(x_k)\| + \Delta))^2} \alpha^2 \|\tilde{\nabla} f(x_k)\|^2 = \\ &= \frac{\alpha^2 \|\tilde{\nabla} f(x_k)\|^2 - 2\alpha \langle \nabla f(x_k), \tilde{\nabla} f(x_k) \rangle}{2 (L_0 + L_1 (\|\tilde{\nabla} f(x_k)\| + \Delta))} = \\ &= \frac{\alpha^2 \|\tilde{\nabla} f(x_k)\|^2 - 2\alpha \langle \nabla f(x_k), \tilde{\nabla} f(x_k) \rangle + \|\nabla f(x_k)\|^2 - \|\nabla f(x_k)\|^2}{2 (L_0 + L_1 (\|\tilde{\nabla} f(x_k)\| + \Delta))} = \\ &= \frac{\|\alpha \tilde{\nabla} f(x_k) - \nabla f(x_k)\|^2 - \|\nabla f(x_k)\|^2}{2 (L_0 + L_1 (\|\tilde{\nabla} f(x_k)\| + \Delta))}. \end{aligned}$$

Следовательно,

$$f(x_k) - f(x_{k+1}) \geq \frac{\|\nabla f(x_k)\|^2 - \|\alpha \tilde{\nabla} f(x_k) - \nabla f(x_k)\|^2}{2 (L_0 + L_1 (\|\tilde{\nabla} f(x_k)\| + \Delta))}.$$

Воспользуемся следующим соотношением для $\alpha \in (0; 1]$, $a, b \in \mathbb{R}^n$:

$$\|\alpha a - b\| = \|\alpha(a - b) + (\alpha - 1)b\| \leq \alpha\|a - b\| + (1 - \alpha)\|b\|.$$

$$\begin{aligned} f(x_k) - f(x_{k+1}) &\geq \\ &\geq \frac{\|\nabla f(x_k)\|^2 - \left(\alpha\|\tilde{\nabla} f(x_k) - \nabla f(x_k)\| + (1 - \alpha)\|\nabla f(x_k)\|\right)^2}{2\left(L_0 + L_1\left(\|\tilde{\nabla} f(x_k)\| + \Delta\right)\right)} \geq \\ &\geq \frac{\|\nabla f(x_k)\|^2 - (\alpha\Delta + (1 - \alpha)\|\nabla f(x_k)\|)^2}{2\left(L_0 + L_1\left(\|\tilde{\nabla} f(x_k)\| + \Delta\right)\right)} = \\ &= \frac{\alpha(\|\nabla f(x_k)\| - \Delta)((2 - \alpha)(\|\nabla f(x_k)\| - \Delta) + 2\Delta)}{2\left(L_0 + L_1\left(\|\tilde{\nabla} f(x_k)\| + \Delta\right)\right)} > \\ &> \frac{\alpha(2 - \alpha)}{2} \frac{(\|\nabla f(x_k)\| - \Delta)^2}{L_0 + L_1\left(\|\tilde{\nabla} f(x_k)\| + \Delta\right)} \geq \\ &\geq \frac{\alpha(2 - \alpha)}{2} \frac{(\|\nabla f(x_k)\| - \Delta)^2}{L_0 + L_1\|\nabla f(x_k)\| + 2L_1\Delta}. \end{aligned}$$

Лемма 3. Пусть f — (L_0, L_1) -гладкая функция (4), удовлетворяющая условию Поляка–Лоясиевича (1), а x_* — ближайшее к x_{k+1} точное решение задачи (2). Тогда для градиентного метода вида (9) с шагом (10) справедливо неравенство

$$(14) \quad \|x_{k+1} - x_*\| \leq \sqrt{\frac{2}{\mu} (f(x_0) - f^*)}.$$

Доказательство. Как известно ([29, Appendix A]), из условия Поляка–Лоясиевича следует (для ближайшего точного решения x_*), что функция имеет квадратичный рост:

$$\frac{\mu}{2} \|x_{k+1} - x_*\|^2 \leq f(x_{k+1}) - f^*.$$

С учетом монотонности функции f неравенство принимает вид

$$\frac{\mu}{2} \|x_{k+1} - x_*\|^2 \leq f(x_0) - f^*,$$

откуда вытекает искомое неравенство

$$\|x_{k+1} - x_*\| \leq \sqrt{\frac{2}{\mu} (f(x_0) - f^*)}.$$

Введем обозначение $g_k = \|\nabla f(x_k)\| - \Delta$. Тогда из леммы 2 вытекает неравенство

$$(15) \quad f(x_k) - f(x_{k+1}) > \frac{\alpha(2-\alpha)}{2} \frac{g_k^2}{L_0 + L_1 g_k + 3L_1 \Delta}.$$

Предположим, что $g_k \geq 3\Delta$. Будем исследовать сходимость метода в зависимости от величины g_k . Возможны два случая: $g_k > \frac{L_0}{L_1}$ и $g_k \leq \frac{L_0}{L_1}$.

При $g_k > \frac{L_0}{L_1}$ из (15) имеем

$$(16) \quad \begin{aligned} f(x_k) - f(x_{k+1}) &> \frac{\alpha(2-\alpha)}{2} \frac{g_k^2}{L_1 g_k + L_1 g_k + L_1 g_k} > \\ &> \frac{\alpha(2-\alpha)}{2} \frac{g_k}{3L_1} = \frac{\alpha(2-\alpha)}{6L_1} g_k. \end{aligned}$$

Если же $g_k \leq \frac{L_0}{L_1}$, то из (15) следует

$$(17) \quad \begin{aligned} f(x_k) - f(x_{k+1}) &> \frac{\alpha(2-\alpha)}{2} \frac{g_k^2}{L_0 + L_0 + L_0} > \\ &> \frac{\alpha(2-\alpha)}{2} \frac{g_k^2}{3L_0} = \frac{\alpha(2-\alpha)}{6L_0} g_k^2. \end{aligned}$$

Для такого класса задач справедлива следующая теорема.

Теорема 1. Пусть f – ρ -квазар-выпуклая относительно всякого точного решения (L_0, L_1) -гладкая функция, удовлетворяющая условию Поляка–Лоясиевича (1). Пусть также справедливы неравенства $\|\tilde{\nabla} f(x_k)\| \geq 5\Delta$ и $\mu < \min \left\{ \frac{6L_0}{\alpha(2-\alpha)}, \frac{72L_1^2(f(x_0) - f^*)}{\alpha^2(2-\alpha)^2\rho^2} \right\}$. Тогда для алгоритма (9) с шагом (10) справедливо неравенство

$$(18) \quad \begin{aligned} &f(x_{k+1}) - f^* \leq \\ &\leq \max \left\{ 1 - \frac{\alpha(2-\alpha)\rho}{6L_1\sqrt{\frac{2}{\mu}}(f(x_0) - f^*)}; 1 - \frac{\alpha(2-\alpha)\mu}{6L_0} \right\}^{k+1} (f(x_0) - f^*) + \\ &+ \Delta \max \left\{ \frac{1}{L_1}; \frac{\Delta}{L_0} \right\} \max \left\{ \frac{L_1\sqrt{\frac{2}{\mu}}(f(x_0) - f^*)}{\rho}; \frac{L_0}{\mu} \right\}. \end{aligned}$$

Доказательство. Рассмотрим сначала случай $g_k > \frac{L_0}{L_1}$. Из квазар-выпуклости функции $f(x)$ при $x = x_k$ вытекает неравенство

$$(19) \quad \begin{aligned} f(x_k) - f^* &\leq \frac{1}{\rho} \langle \nabla f(x_k), x_k - x_* \rangle \leq \frac{1}{\rho} \|\nabla f(x_k)\| \|x_k - x_*\| = \\ &= \frac{1}{\rho} (g_k + \Delta) \|x_k - x_*\|. \end{aligned}$$

Применяя к (19) лемму 3, получим неравенство

$$f(x_k) - f^* \leq \frac{1}{\rho} \sqrt{\frac{2}{\mu}} (f(x_0) - f^*)(g_k + \Delta),$$

откуда

$$(20) \quad g_k \geq \frac{\rho(f(x_k) - f^*)}{\sqrt{\frac{2}{\mu}} (f(x_0) - f^*)} - \Delta.$$

Из неравенств (16) и (20) вытекает оценка

$$(21) \quad \begin{aligned} f(x_{k+1}) - f^* &\leq f(x_k) - f^* - \frac{\alpha(2-\alpha)}{6L_1} \left(\frac{\rho(f(x_k) - f^*)}{\sqrt{\frac{2}{\mu}} (f(x_0) - f^*)} - \Delta \right) = \\ &= \left(1 - \frac{\alpha(2-\alpha)\rho}{6L_1\sqrt{\frac{2}{\mu}} (f(x_0) - f^*)} \right) (f(x_k) - f^*) + \frac{\alpha(2-\alpha)\Delta}{6L_1}. \end{aligned}$$

При $g_k \leq \frac{L_0}{L_1}$ из (17) следует, что

$$\begin{aligned} f(x_{k+1}) - f(x_k) &\leq -\frac{\alpha(2-\alpha)}{6L_0} g_k^2 = -\frac{\alpha(2-\alpha)}{6L_0} (\|\nabla f(x_k)\| - \Delta)^2 \leq \\ &\leq -\frac{\alpha(2-\alpha)}{6L_0} \left(\frac{1}{2} \|\nabla f(x_k)\|^2 - \Delta^2 \right) = \\ &= -\frac{\alpha(2-\alpha)}{12L_0} \|\nabla f(x_k)\|^2 + \frac{\alpha(2-\alpha)\Delta^2}{6L_0} = \\ &= -\frac{\alpha(2-\alpha)\mu}{6L_0} (f(x_k) - f^*) + \frac{\alpha(2-\alpha)\Delta^2}{6L_0}. \end{aligned}$$

Таким образом,

$$(22) \quad f(x_{k+1}) - f^* \leq \left(1 - \frac{\alpha(2-\alpha)\mu}{6L_0} \right) (f(x_k) - f^*) + \frac{\alpha(2-\alpha)\Delta^2}{6L_0}.$$

Из неравенств (21) и (22) следует справедливость неравенства

$$\begin{aligned} f(x_{k+1}) - f^* &\leq \max \left\{ 1 - \frac{\alpha(2-\alpha)\rho}{6L_1\sqrt{\frac{2}{\mu}} (f(x_0) - f^*)}; 1 - \frac{\alpha(2-\alpha)\mu}{6L_0} \right\} (f(x_k) - f^*) + \\ &+ \max \left\{ \frac{\alpha(2-\alpha)\Delta}{6L_1}; \frac{\alpha(2-\alpha)\Delta^2}{6L_0} \right\}, \end{aligned}$$

или

$$\begin{aligned}
f(x_{k+1}) - f^* &\leq \max \left\{ 1 - \frac{\alpha(2-\alpha)\rho}{6L_1\sqrt{\frac{2}{\mu}}(f(x_0) - f^*)}; 1 - \frac{\alpha(2-\alpha)\mu}{6L_0} \right\} (f(x_k) - f^*) + \\
&+ \frac{\alpha(2-\alpha)\Delta}{6} \max \left\{ \frac{1}{L_1}; \frac{\Delta}{L_0} \right\} \leq \\
&\leq \max \left\{ 1 - \frac{\alpha(2-\alpha)\rho}{6L_1\sqrt{\frac{2}{\mu}}(f(x_0) - f^*)}; 1 - \frac{\alpha(2-\alpha)\mu}{6L_0} \right\}^{k+1} (f(x_0) - f^*) + \\
&+ \frac{\alpha(2-\alpha)\Delta}{6} \max \left\{ \frac{1}{L_1}; \frac{\Delta}{L_0} \right\} \sum_{i=0}^k \max \left\{ 1 - \frac{\alpha(2-\alpha)\rho}{6L_1\sqrt{\frac{2}{\mu}}(f(x_0) - f^*)}; \right. \\
&\qquad \qquad \qquad \left. 1 - \frac{\alpha(2-\alpha)\mu}{6L_0} \right\}^i \leq \\
&\leq \max \left\{ 1 - \frac{\alpha(2-\alpha)\rho}{6L_1\sqrt{\frac{2}{\mu}}(f(x_0) - f^*)}; 1 - \frac{\alpha(2-\alpha)\mu}{6L_0} \right\}^{k+1} (f(x_0) - f^*) + \\
&+ \frac{\alpha(2-\alpha)\Delta}{6} \max \left\{ \frac{1}{L_1}; \frac{\Delta}{L_0} \right\} \frac{1}{1 - \max \left\{ 1 - \frac{\alpha(2-\alpha)\rho}{6L_1\sqrt{\frac{2}{\mu}}(f(x_0) - f^*)}; 1 - \frac{\alpha(2-\alpha)\mu}{6L_0} \right\}}.
\end{aligned}$$

При положительных a, b верна цепочка равенств

$$\frac{1}{1 - \max\{1 - a; 1 - b\}} = \frac{1}{1 - (1 - \min\{a; b\})} = \frac{1}{\min\{a; b\}} = \max \left\{ \frac{1}{a}; \frac{1}{b} \right\},$$

следовательно,

$$\begin{aligned}
f(x_{k+1}) - f^* &\leq \\
&\leq \max \left\{ 1 - \frac{\alpha(2-\alpha)\rho}{6L_1\sqrt{\frac{2}{\mu}}(f(x_0) - f^*)}; 1 - \frac{\alpha(2-\alpha)\mu}{6L_0} \right\}^{k+1} (f(x_0) - f^*) + \\
&+ \frac{\alpha(2-\alpha)\Delta}{6} \max \left\{ \frac{1}{L_1}; \frac{\Delta}{L_0} \right\} \frac{6}{\alpha(2-\alpha)} \max \left\{ \frac{L_1\sqrt{\frac{2}{\mu}}(f(x_0) - f^*)}{\rho}; \frac{L_0}{\mu} \right\} \leq \\
&\leq \max \left\{ 1 - \frac{\alpha(2-\alpha)\rho}{6L_1\sqrt{\frac{2}{\mu}}(f(x_0) - f^*)}; 1 - \frac{\alpha(2-\alpha)\mu}{6L_0} \right\}^{k+1} (f(x_0) - f^*) = \\
&+ \Delta \max \left\{ \frac{1}{L_1}; \frac{\Delta}{L_0} \right\} \max \left\{ \frac{L_1\sqrt{\frac{2}{\mu}}(f(x_0) - f^*)}{\rho}; \frac{L_0}{\mu} \right\}.
\end{aligned}$$

Замечание 2. Количество итераций алгоритма (9), для которых справедливо неравенство $\|\tilde{\nabla}f(x_k)\| \geq 5\Delta$, можно оценить как

$$N \leq \left\lceil \frac{6L_1^2(f(x_0) - f^*)}{\alpha(2 - \alpha)L_0} \right\rceil + \left\lceil \frac{2L_0(f(x_0) - f^*)}{3\alpha(2 - \alpha)\Delta^2} \right\rceil.$$

Доказательство. Оценим число итераций, когда $g_k > \frac{L_0}{L_1}$. Из (16) следует неравенство

$$f(x_k) - f(x_{k+1}) > \frac{\alpha(2 - \alpha)L_0}{6L_1^2}.$$

Пусть i_1, \dots, i_m , $p = \overline{1, m}$ – индексы таких итераций, для которых $g_{i_p} > \frac{L_0}{L_1}$. Тогда после суммирования верно неравенство

$$f(x_0) - f^* \geq f(x_{i_1}) - f(x_{i_m}) \geq \frac{\alpha(2 - \alpha)L_0}{6L_1^2}m,$$

откуда

$$m \leq \left\lceil \frac{6L_1^2(f(x_0) - f^*)}{\alpha(2 - \alpha)L_0} \right\rceil, \quad L_0 \neq 0.$$

Оценим теперь количество итераций, для которых $g_k \leq \frac{L_0}{L_1}$. Принимая во внимание, что $g_k^2 \geq 9\Delta^2$, получаем из неравенства (17) оценку $f(x_k) - f(x_{k+1}) > \frac{9\alpha(2 - \alpha)\Delta^2}{6L_0}$. Пусть j_1, \dots, j_n , $l = \overline{1, n}$ – индексы итераций, для которых $g_{j_l} \leq \frac{L_0}{L_1}$. После суммирования имеем неравенство

$$f(x_0) - f^* \geq f(x_{j_1}) - f(x_{j_n}) \geq \frac{9\alpha(2 - \alpha)\Delta^2}{6L_0}n,$$

откуда следует ограничение на число таких итераций

$$n \leq \left\lceil \frac{2L_0(f(x_0) - f^*)}{3\alpha(2 - \alpha)\Delta^2} \right\rceil.$$

Замечание 3. Неравенство (21) указывает на сходимость метода (9) с шагом (10) со скоростью, близкой к скорости геометрической прогрессии, при достаточно больших значениях норм градиентов, т.е. $g_k > \frac{L_0}{L_1}$.

Замечание 4. Будем считать, что если $\|\tilde{\nabla}f(x_k)\| < 5\Delta$, то приемлемое качество решения уже достигнуто. Действительно, если $\|\tilde{\nabla}f(x_k)\| < 5\Delta$, то $\|\nabla f(x_k)\| < 6\Delta$. Тогда из условия Поляка–Лоясиевича (1) следует

$$f(x_k) - f^* \leq \frac{1}{2\mu} \|\nabla f(x_k)\|^2 < \frac{18\Delta^2}{\mu}.$$

Замечание 5. Поскольку количество итераций, для которых возможно выполнение условия $g_k > \frac{L_0}{L_1}$, ограничено, начиная с некоторого номера N норма градиента будет меньше, чем $\frac{L_0}{L_1}$. Тогда можно утверждать о сходимости метода (9) с шагом (10) на классе задач с условием Поляка–Лоясиевича (без требования ρ -квазар-выпуклости) начиная с этого номера. Таким образом, начиная с некоторого номера справедлива альтернатива:

1) либо из неравенства (22) по рекурсии следует

$$f(x_{k+1}) - f^* \leq \left(1 - \frac{\alpha(2-\alpha)\mu}{6L_0}\right)^k (f(x_N) - f^*) + \frac{\Delta^2}{\mu},$$

2) либо приемлемая точность решения уже достигнута.

4. Адаптивный градиентный метод для квазар-выпуклых $(0, M)$ -гладких функций

Рассмотрим случай $L_0 = 0$, соответствующий, например, задаче логистической регрессии. Для удобства введем обозначение $M = L_1$. Тогда условие (4) принимает вид

$$(23) \quad \|\nabla f(y) - \nabla f(x)\| \leq M \|\nabla f(x)\| \|y - x\| \quad \forall x, y \in \mathbb{R}^n \text{ таких, что } \|y - x\| \leq \frac{1}{M}.$$

Далее исследуется адаптивный вариант градиентного метода с использованием неточной информации о градиенте на классе ρ -квазар выпуклых функций.

Пусть $f - (0, M)$ -гладкая ρ -квазар-выпуклая функция (в отличие от предыдущего раздела уже не обязательно относительно всякого точного решения), причем $\rho \in (\frac{1}{2}; 1]$. Рассмотрим следующий адаптивный вариант градиентного метода с использованием неточных градиентов на итерациях.

Алгоритм 1 (адаптивный градиентный метод для $(0, M)$ -гладких задач).

1. Вход $x_0 \in \mathbb{R}^n$, $\Delta > 0$, $0 < M_0 < 2M$, $\alpha > 0$.
2. $k = 0$.
3. Повторять
4. $M_{k+1} = \max \{M_0/2, M_k/2\}$,
5. $\eta_k = \frac{\alpha}{M_{k+1}(\|\tilde{\nabla} f(x_k)\| + \Delta)}$,
6. $x_{k+1} = x_k - \eta_k \tilde{\nabla} f(x_k)$,
7. Если $f(x_k) - f(x_{k+1}) \geq \frac{\alpha(2-\alpha)}{4M_{k+1}} \left(\|\tilde{\nabla} f(x_k)\| - 2\Delta\right)$ и $f(x_k) - f(x_{k+1}) \geq \alpha \left(1 - \frac{\alpha}{2}\right) \frac{\|\tilde{\nabla} f(x_k)\|^2}{M_{k+1}(\|\tilde{\nabla} f(x_k)\| + \Delta)} - \frac{\Delta\alpha}{M_{k+1}}$,
8. $k := k + 1$.
9. Иначе
10. $M_{k+1} = 2M_{k+1}$ и перейти к 5.
11. Выход $\hat{x} = x_{k+1}$.

Докажем несколько вспомогательных утверждений для $(0, M)$ -гладких функций. В частности, они обоснуют корректность алгоритма 1, т.е. выполнение для $(0, M)$ -гладких функций при достаточно больших M_{k+1} критерия выхода из итерации согласно п. 7 алгоритма 1.

Лемма 4. Пусть $f(x)$ – $(0, M)$ -гладкая функция. Предположим, что $\|\tilde{\nabla}f(x_k)\| > 5\Delta$ и $\{M_i\}_{i=0}^k$ – последовательность, причем $M_{i+1} \geq M$. Тогда на итерациях алгоритма

$$x_{k+1} = x_k - \frac{\alpha}{M_{k+1} \left(\|\tilde{\nabla}f(x_k)\| + \Delta \right)} \tilde{\nabla}f(x_k), \quad \alpha \in (0; 1]$$

значения функции монотонно не возрастают $f(x_{k+1}) \leq f(x_k)$, причем справедливо неравенство

$$(24) \quad f(x_k) - f(x_{k+1}) \geq \frac{\alpha(2-\alpha)}{2} \frac{(\|\tilde{\nabla}f(x_k)\| - 2\Delta)^2}{M_{k+1}(\|\tilde{\nabla}f(x_k)\| + \Delta)}.$$

Доказательство. Из неравенства (11) при $L_0 = 0$, $L_1 = M$ следует неравенство

$$f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{M\|\nabla f(x)\|}{2} \|y - x\|^2.$$

Положим в этом неравенстве $y = x_{k+1}$, $x = x_k$. Тогда

$$\begin{aligned} f(x_{k+1}) - f(x_k) &\leq \\ &\leq \frac{-\alpha \langle \nabla f(x_k), \tilde{\nabla}f(x_k) \rangle}{M_{k+1} \left(\|\tilde{\nabla}f(x_k)\| + \Delta \right)} + \frac{M\|\nabla f(x_k)\|}{2M_{k+1}^2 \left(\|\tilde{\nabla}f(x_k)\| + \Delta \right)} \alpha^2 \|\tilde{\nabla}f(x_k)\|^2 \leq \\ &\leq \frac{-\alpha \langle \nabla f(x_k), \tilde{\nabla}f(x_k) \rangle}{M_{k+1} \left(\|\tilde{\nabla}f(x_k)\| + \Delta \right)} + \frac{M \left(\|\tilde{\nabla}f(x_k)\| + \Delta \right)}{2M \left(\|\tilde{\nabla}f(x_k)\| + \Delta \right)} \frac{\alpha^2 \|\tilde{\nabla}f(x_k)\|^2}{M_{k+1} \left(\|\tilde{\nabla}f(x_k)\| + \Delta \right)} = \\ &= \frac{\alpha^2 \|\tilde{\nabla}f(x_k)\|^2 - 2\alpha \langle \nabla f(x_k), \tilde{\nabla}f(x_k) \rangle}{2M_{k+1} \left(\|\tilde{\nabla}f(x_k)\| + \Delta \right)} = \\ &= \frac{\alpha^2 \|\tilde{\nabla}f(x_k)\|^2 - 2\alpha \langle \nabla f(x_k), \tilde{\nabla}f(x_k) \rangle + \|\nabla f(x_k)\|^2 - \|\nabla f(x_k)\|^2}{2M_{k+1} \left(\|\tilde{\nabla}f(x_k)\| + \Delta \right)} = \\ &= \frac{\|\alpha \tilde{\nabla}f(x_k) - \nabla f(x_k)\|^2 - \|\nabla f(x_k)\|^2}{2M_{k+1} \left(\|\tilde{\nabla}f(x_k)\| + \Delta \right)}. \end{aligned}$$

Следовательно,

$$f(x_k) - f(x_{k+1}) \geq \frac{\|\nabla f(x_k)\|^2 - \|\alpha \tilde{\nabla} f(x_k) - \nabla f(x_k)\|^2}{2M_{k+1} (\|\tilde{\nabla} f(x_k)\| + \Delta)}.$$

Воспользуемся следующим соотношением для $\alpha \in (0; 1]$, $a, b \in \mathbb{R}^n$:

$$\|\alpha a - b\| = \|\alpha(a - b) + (\alpha - 1)b\| \leq \alpha\|a - b\| + (1 - \alpha)\|b\|.$$

$$\begin{aligned} f(x_k) - f(x_{k+1}) &\geq \\ &\geq \frac{\|\nabla f(x_k)\|^2 - \left(\alpha\|\tilde{\nabla} f(x_k) - \nabla f(x_k)\| + (1 - \alpha)\|\nabla f(x_k)\|\right)^2}{2M_{k+1} (\|\tilde{\nabla} f(x_k)\| + \Delta)} \geq \\ &\geq \frac{\|\nabla f(x_k)\|^2 - (\alpha\Delta + (1 - \alpha)\|\nabla f(x_k)\|)^2}{2M_{k+1} (\|\tilde{\nabla} f(x_k)\| + \Delta)} = \\ &= \frac{\alpha(\|\nabla f(x_k)\| - \Delta)((2 - \alpha)(\|\nabla f(x_k)\| - \Delta) + 2\Delta)}{2M_{k+1} (\|\tilde{\nabla} f(x_k)\| + \Delta)} > \\ &> \frac{\alpha(2 - \alpha)}{2} \frac{(\|\nabla f(x_k)\| - \Delta)^2}{M_{k+1} (\|\tilde{\nabla} f(x_k)\| + \Delta)} \geq \\ &\geq \frac{\alpha(2 - \alpha)}{2} \frac{(\|\tilde{\nabla} f(x_k)\| - 2\Delta)^2}{M_{k+1} (\|\tilde{\nabla} f(x_k)\| + \Delta)}. \end{aligned}$$

Пусть $\{M_i\}_{i=0}^k$ – последовательность положительных чисел, удовлетворяющих предположениям предыдущей теоремы. Тогда из (24) при $\|\tilde{\nabla} f(x_k)\| \geq 5\Delta$ вытекает

$$(25) \quad \begin{aligned} f(x_k) - f(x_{k+1}) &\geq \frac{\alpha(2 - \alpha)(\|\tilde{\nabla} f(x_k)\| - 2\Delta)^2}{2(M_{k+1}(\|\tilde{\nabla} f(x_k)\| - 2\Delta) + 3M_{k+1}\Delta)} \geq \\ &\geq \frac{\alpha(2 - \alpha)}{4M_{k+1}} (\|\tilde{\nabla} f(x_k)\| - 2\Delta). \end{aligned}$$

Лемма 5. Пусть $f(x)$ – $(0, M)$ -гладкая функция, $\{M_i\}_{i=0}^k$ – некоторая последовательность, причем $M_{i+1} \geq M$. Тогда для алгоритма

$$x_{k+1} = x_k - \frac{\alpha}{M_{k+1} (\|\tilde{\nabla} f(x_k)\| + \Delta)} \tilde{\nabla} f(x_k), \quad \alpha \in (0; 1]$$

справедливо неравенство

$$(26) \quad \frac{\|\tilde{\nabla} f(x_k)\|^2}{M_{k+1} (\|\tilde{\nabla} f(x_k)\| + \Delta)} \leq \frac{2}{\alpha(2 - \alpha)} (f(x_k) - f(x_{k+1})) + \frac{2\Delta}{M_{k+1}(2 - \alpha)}.$$

Доказательство. Из неравенства (12) при $L_0 = 0$, $L_1 = M$ следует

$$f(y) \leq f(x) + \langle \tilde{\nabla} f(x), y - x \rangle + \frac{M(\|\tilde{\nabla} f(x)\| + \Delta)}{2} \|y - x\|^2 + \Delta \|y - x\|.$$

Положим в этом неравенстве $y = x_{k+1}$ и $x = x_k$:

$$\begin{aligned} f(x_{k+1}) &\leq f(x_k) + \langle \tilde{\nabla} f(x_k), x_{k+1} - x_k \rangle + \\ &+ \frac{M(\|\tilde{\nabla} f(x_k)\| + \Delta)}{2} \|x_{k+1} - x_k\|^2 + \Delta \|x_{k+1} - x_k\| \leq \\ &\leq f(x_k) - \frac{\alpha \|\tilde{\nabla} f(x_k)\|^2}{M_{k+1}(\|\tilde{\nabla} f(x_k)\| + \Delta)} + \frac{\alpha^2 \|\tilde{\nabla} f(x_k)\|^2}{2M_{k+1}(\|\tilde{\nabla} f(x_k)\| + \Delta)} + \\ &+ \frac{\Delta \alpha \|\tilde{\nabla} f(x_k)\|}{M_{k+1}(\|\tilde{\nabla} f(x_k)\| + \Delta)} \\ &= f(x_k) - \alpha \left(1 - \frac{\alpha}{2}\right) \frac{\|\tilde{\nabla} f(x_k)\|^2}{M_{k+1}(\|\tilde{\nabla} f(x_k)\| + \Delta)} + \frac{\Delta \alpha \|\tilde{\nabla} f(x_k)\|}{M_{k+1}(\|\tilde{\nabla} f(x_k)\| + \Delta)}. \end{aligned}$$

Оценим слагаемое, содержащее Δ :

$$\frac{\Delta \alpha \|\tilde{\nabla} f(x_k)\|}{M_{k+1}(\|\tilde{\nabla} f(x_k)\| + \Delta)} \leq \frac{\frac{\Delta \alpha}{M_{k+1}} M_{k+1}(\|\tilde{\nabla} f(x_k)\| + \Delta)}{M_{k+1}(\|\tilde{\nabla} f(x_k)\| + \Delta)} = \frac{\Delta \alpha}{M_{k+1}}.$$

Таким образом, получаем неравенство

$$f(x_k) - f(x_{k+1}) \geq \alpha \left(1 - \frac{\alpha}{2}\right) \frac{\|\tilde{\nabla} f(x_k)\|^2}{M_{k+1}(\|\tilde{\nabla} f(x_k)\| + \Delta)} - \frac{\Delta \alpha}{M_{k+1}},$$

или в другом виде

$$\frac{\|\tilde{\nabla} f(x_k)\|^2}{M_{k+1}(\|\tilde{\nabla} f(x_k)\| + \Delta)} \leq \frac{2}{\alpha(2 - \alpha)} (f(x_k) - f(x_{k+1})) + \frac{2\Delta}{M_{k+1}(2 - \alpha)}.$$

Доказанные выше неравенства (25) и (26) означают, что при $M_{k+1} \geq M$ автоматически выполнен критерий выхода из итерации алгоритма 1.

Лемма 6. Пусть $f(x) - \rho$ -квазар-выпуклая функция при $\rho \in (\frac{1}{2}; 1]$, $\{M_i\}_{i=0}^k$ - некоторая последовательность положительных чисел из алгоритма 1. Тогда при $\alpha \leq \frac{2\rho-1}{\rho}$ справедливо неравенство

$$(27) \quad \|x_{k+1} - x_*\| \leq \|x_k - x_*\| + \Delta \eta_k + \frac{1}{M_{k+1}}.$$

Доказательство. Поскольку $x_{k+1} = x_k - \eta_k \tilde{\nabla} f(x_k)$, то верны следующие равенства:

$$\begin{aligned} \|x_{k+1} - x_*\|^2 &= \|x_k - x_* - \eta_k \tilde{\nabla} f(x_k)\|^2 = \\ &= \|x_k - x_*\|^2 - 2\eta_k \langle \tilde{\nabla} f(x_k), x_k - x_* \rangle + \eta_k^2 \|\tilde{\nabla} f(x_k)\|^2. \end{aligned}$$

Из квазар-выпуклости функции f следует, что

$$\begin{aligned}
& \|x_{k+1} - x_*\|^2 \leq \\
& \leq \|x_k - x_*\|^2 - 2\eta_k(\rho(f(x_k) - f^*) - \Delta\|x_k - x_*\|) + \eta_k^2 \|\tilde{\nabla} f(x_k)\|^2 \stackrel{(26)}{\leq} \\
& \stackrel{(26)}{\leq} \|x_k - x_*\|^2 - 2\eta_k(\rho(f(x_k) - f^*) - \Delta\|x_k - x_*\|) + \\
& + \alpha\eta_k \left(\frac{2}{\alpha(2-\alpha)}(f(x_k) - f^*) + \frac{2\Delta}{M_{k+1}(2-\alpha)} \right) = \\
& = \|x_k - x_*\|^2 + \eta_k \left(\frac{2}{2-\alpha} - 2\rho \right) (f(x_k) - f^*) + \\
& + 2\Delta\eta_k \left(\|x_k - x_*\| + \frac{\alpha}{M_{k+1}(2-\alpha)} \right).
\end{aligned}$$

Выберем параметр α так, чтобы $\frac{2}{2-\alpha} - 2\rho \leq 0$. Поскольку $\alpha \in (0; 1]$, получаем, что это возможно тогда и только тогда, когда $\rho \in (\frac{1}{2}; 1]$. В этом случае имеем:

$$\begin{aligned}
\|x_{k+1} - x_*\|^2 & \leq \|x_k - x_*\|^2 + 2\Delta\eta_k\|x_k - x_*\| + \\
& + \Delta^2\eta_k^2 + \frac{2\alpha\Delta\eta_k}{M_{k+1}(2-\alpha)} - \Delta^2\eta_k^2 = \\
& = (\|x_k - x_*\| + \Delta\eta_k)^2 + \Delta\eta_k \left(\frac{2\alpha}{M_{k+1}(2-\alpha)} - \Delta\eta_k \right).
\end{aligned}$$

Рассмотрим функцию $y(t) = -t^2 + \frac{2\alpha}{M_{k+1}(2-\alpha)}t$. В точке $t = \frac{\alpha}{M_{k+1}(2-\alpha)}$ она достигает максимального значения $\frac{\alpha^2}{M_{k+1}^2(2-\alpha)^2}$, причем при $\alpha \leq 1$ справедливо $\frac{\alpha^2}{M_{k+1}^2(2-\alpha)^2} \leq \frac{1}{M_{k+1}^2}$. Таким образом,

$$\|x_{k+1} - x_*\|^2 \leq (\|x_k - x_*\| + \Delta\eta_k)^2 + \frac{1}{M_{k+1}^2}.$$

Заметим, что при $a, b, c > 0$ из неравенства $a^2 \leq b^2 + c^2$ вытекает $a \leq (b+c)^2$, откуда $a \leq b+c$. Следовательно,

$$\|x_{k+1} - x_*\| \leq \|x_k - x_*\| + \Delta\eta_k + \frac{1}{M_{k+1}}.$$

Отметим, что неравенства (25) и (26) заведомо выполняются, если $M_{k+1} \geq \geq M$. Если $\|\tilde{\nabla} f(x_k)\| \geq 5\Delta$ и для некоторого $M_0 > 0$ верно $M_i \geq \frac{M_0}{2} \forall i = \overline{1, k+1}$, то

$$\eta_k = \frac{\alpha}{M_{k+1} \left(\|\tilde{\nabla} f(x_k)\| + \Delta \right)} \leq \frac{\alpha}{6M_{k+1}\Delta} \leq \frac{\alpha}{3M_0\Delta}.$$

Теорема 2. Пусть $f(x)$ – ρ -квазар-выпуклая $(0, M)$ -гладкая функция, где $\rho \in (\frac{1}{2}; 1]$, а $\alpha \in (0; 1]$ такое, что $\alpha \leq \frac{2\rho-1}{\rho}$. Тогда после $\left\lceil \frac{8M(f(x_0)-f^*)}{3\alpha(2-\alpha)\Delta} \right\rceil$ итераций алгоритма 1 найдется k , для которого либо $\|\tilde{\nabla} f(x_k)\| < 5\Delta$, либо будет достигнуто точное решение.

Доказательство. Из неравенства (25) с учетом того, что $\|\tilde{\nabla} f(x_k)\| \geq 5\Delta$, после суммирования по k от 0 до $N-1$ вытекают соотношения

$$f(x_0) - f^* \geq f(x_0) - f(x_N) \geq \frac{3\alpha(2-\alpha)\Delta}{4} \sum_{k=0}^{N-1} \frac{1}{M_{k+1}},$$

откуда

$$\sum_{k=0}^{N-1} \frac{1}{M_{k+1}} \leq \frac{4(f(x_0) - f^*)}{3\alpha(2-\alpha)\Delta}.$$

Учитывая, что $M_i \leq 2M$, оценим количество итераций, для которых $\|\tilde{\nabla} f(x_k)\| \geq 5\Delta$. Имеем $\frac{N}{2M} \leq \frac{4(f(x_0)-f^*)}{3\alpha(2-\alpha)\Delta}$, или

$$N \leq \left\lceil \frac{8M(f(x_0) - f^*)}{3\alpha(2-\alpha)\Delta} \right\rceil.$$

Замечание 6. Будем считать, что если $\|\tilde{\nabla} f(x_k)\| < 5\Delta$, то приемлемое качество решения уже достигнуто. Действительно, из условия ρ -квазар-выпуклости функции f и неравенства (27) следует

$$\begin{aligned} f(x_k) - f^* &\leq \frac{1}{\rho} \left(\|\tilde{\nabla} f(x_k)\| + \Delta \right) \|x_k - x_*\| \leq \\ &\leq \frac{6\Delta}{\rho} \left(\|x_0 - x_*\| + \frac{(\alpha+6)N}{3M_0} \right). \end{aligned}$$

Отсюда следует, что при небольшом количестве итераций можно гарантировать, что качество решения находится в пределах $O(\Delta)$. В случае больших значений N этого утверждать нельзя. Отметим, что если целевая функция удовлетворяет условию Поляка–Лоясиевича, то согласно замечанию 4 малая норма неточного градиента гарантирует достижение приемлемого качества решения.

В частном случае, при $\Delta = 0$, верно следующее утверждение.

Теорема 3. Пусть $f(x)$ – ρ -квазар-выпуклая $(0, M)$ -гладкая функция, где $\rho \in (\frac{1}{2}; 1]$, а $\alpha \in (0; 1]$ такое, что $\alpha \leq \frac{2\rho-1}{\rho}$. Тогда алгоритм 1 в случае $\Delta = 0$ сходится со скоростью геометрической прогрессии.

Доказательство. Рассмотрим случай $\Delta = 0$, т.е. $\tilde{\nabla} f(x) = \nabla f(x)$ для всякого x . Тогда из п. 7 листинга алгоритма 1 следует, что

$$(28) \quad f(x_k) - f(x_{k+1}) \geq \frac{\alpha(2-\alpha)}{2M_{k+1}} \frac{\|\nabla f(x_k)\|^2}{\|\nabla f(x_k)\|} = \frac{(2-\alpha)\eta_k}{2} \|\nabla f(x_k)\|^2.$$

Также в этом случае справедливы соотношения

$$\begin{aligned}
\|x_{k+1} - x_*\|^2 &= \|x_k - x_*\|^2 - 2\eta_k \langle \nabla f(x_k), x_k - x_* \rangle + \eta_k^2 \|\nabla f(x_k)\|^2 \leq \\
&\leq \|x_k - x_*\|^2 - 2\eta_k \rho (f(x_k) - f^*) + \eta_k^2 \|\nabla f(x_k)\|^2 \leq \\
&\leq \|x_k - x_*\|^2 - 2\eta_k \rho (f(x_k) - f^*) + \frac{2\eta_k}{(2-\alpha)} (f(x_k) - f^*) = \\
&= \|x_k - x_*\|^2 + \eta_k (f(x_k) - f^*) \left(\frac{2}{2-\alpha} - 2\rho \right) \leq \\
&\leq \|x_k - x_*\|^2 \leq \|x_0 - x_*\|^2 = R^2,
\end{aligned}$$

где $\alpha \in (0; 1]$, $\rho \in (\frac{1}{2}; 1]$. Далее, из квазар-выпуклости функции f имеем

$$f(x_k) - f^* \leq \frac{1}{\rho} \|\nabla f(x_k)\| \|x_k - x_*\| \leq \frac{1}{\rho} \|\nabla f(x_k)\| \|x_0 - x_*\| = \frac{R}{\rho} \|\nabla f(x_k)\|.$$

Вспомним, что ввиду лемм 4 и 5 при $M_{k+1} \geq M$ заведомо выполнен критерий выхода из итерации алгоритма 1 (п. 7 листинга). Тогда для всякого $k \geq 0$ верно $M_{k+1} \leq 2M$, и из (28) вытекает

$$\begin{aligned}
f(x_{k+1}) - f^* &\leq f(x_k) - f^* - \frac{\alpha(2-\alpha)}{2M_{k+1}} \|\nabla f(x_k)\| \leq \\
&\leq f(x_k) - f^* - \frac{\alpha(2-\alpha)}{2M_{k+1}} \frac{\rho}{R} (f(x_k) - f^*) \leq \\
&\leq \left(1 - \frac{\alpha(2-\alpha)\rho}{2M_{k+1}R} \right) (f(x_k) - f^*) \leq \\
&\leq \left(1 - \frac{\alpha(2-\alpha)\rho}{4MR} \right) (f(x_k) - f^*) \leq \\
&\leq \left(1 - \frac{\alpha(2-\alpha)\rho}{4MR} \right)^{k+1} (f(x_0) - f^*),
\end{aligned}$$

что и завершает доказательство.

5. Вычислительные эксперименты

В данном разделе представлены результаты численных экспериментов, проведенных для оценки эффективности предложенных алгоритмов при минимизации логистической функции вида (29) и квазар-выпуклой функции вида (30).

5.1. Результаты тестирования адаптивного алгоритма 1

Сравним предложенный алгоритм 1 с адаптивным универсальным градиентным методом из [30] на задаче минимизации логистической функции из примера 3:

$$(29) \quad f(x) = \frac{1}{m} \sum_{i=1}^m \log(1 + \exp(-a_i^\top x)), \quad a_i \in \mathbb{R}^n.$$

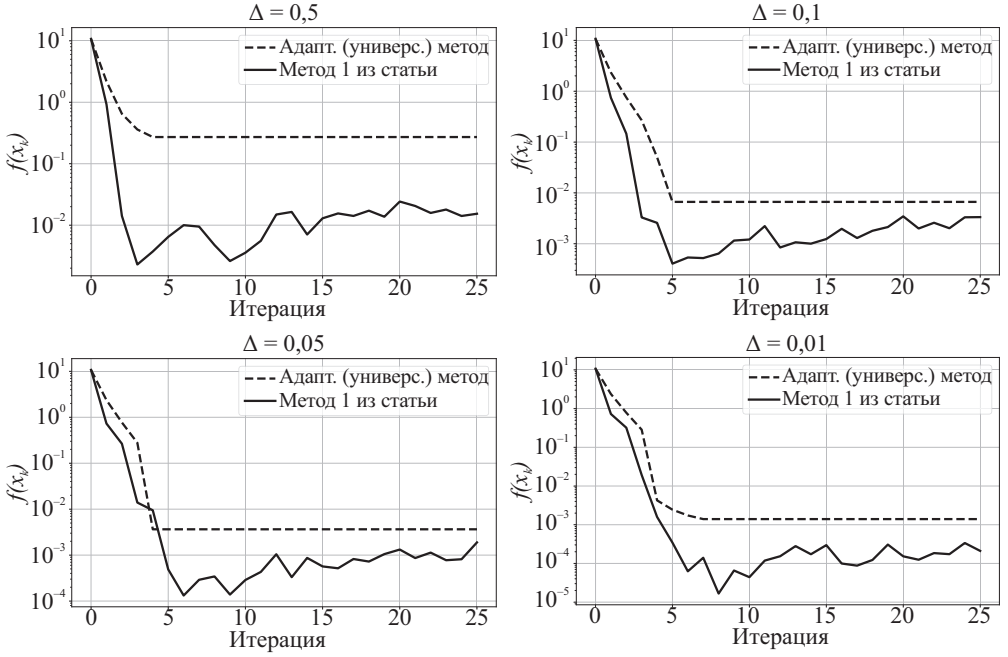


Рис. 1. Результаты алгоритма 1 и адаптивного универсального градиентного метода [30] для задачи логистической регрессии с (29) при $n = 10^3$, $m = 100$.

Известно, что логистическая функция (29) является выпуклой L -гладкой с константой $L = \frac{1}{m} \sum_{i=1}^m \|a_i\|^2$, а также (L_0, L_1) -гладкой с параметрами $L_0 = 0$ и $L_1 = \max_{1 \leq i \leq m} \|a_i\|$.

Алгоритмы запускались с параметрами $n = 1000$, $m = 100$ и начальной точкой $x_0 = (1, \dots, 1) \in \mathbb{R}^n$. Параметр неточности градиента Δ принимал значения из множества $\{0,5; 0,1; 0,05; 0,01\}$. Векторы $\{a_i\}_{i=1}^m$ генерировались случайно из стандартного нормального распределения. На рис. 1 представлена динамика значений целевой функции в зависимости от номера итерации для сравниваемых алгоритмов. Результаты показывают, что предложенный алгоритм 1, использующий (L_0, L_1) -гладкость, демонстрирует более высокую эффективность по сравнению с адаптивным универсальным градиентным методом [30].

5.2. Пример невыпуклой задачи из класса квазар-выпуклых функций

Представлено сравнение предложенного метода (9) с шагом (10) и градиентного метода с постоянным шагом размера 0,001 на примере квазар-выпуклой функции из [27]:

$$(30) \quad f(x) = h(\|x\|_2)g\left(\frac{x}{\|x\|_2}\right),$$

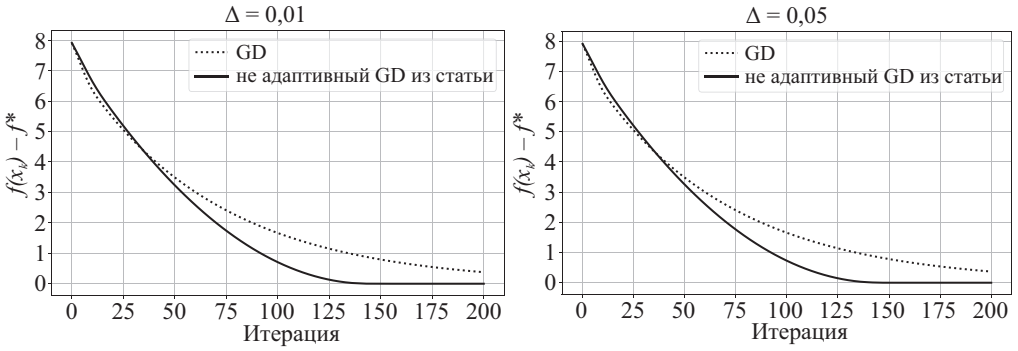


Рис. 2. Результаты применения метода (9) с шагом (10) и градиентного метода с постоянным шагом для минимизации функции (30).

где $h(t) = t^2$ и

$$g(x_1, x_2) = 1 + \frac{1}{4N} \sum_{i=1}^N a_i \sin^2(b_i x_1) + c_i \cos^2(d_i x_2).$$

В вычислительных экспериментах использовались последовательности $\{a_i\}_i$ и $\{c_i\}_i$, элементы которых независимо и равномерно распределены на отрезке $[0, 20)$. Аналогично, последовательности $\{b_i\}_i$ и $\{d_i\}_i$ независимо и равномерно распределены на отрезке $[-25, 25)$. Все методы инициализировались из начальной точки $(1, 1)$ с параметрами $N = 10$ и $\alpha = 0,01$. На рис. 2 представлена динамика изменения значений целевой функции в зависимости от номера итерации. Полученные результаты демонстрируют, что предложенный алгоритм (9) с выбором шага (10), основанным на (L_0, L_1) -гладкости целевой функции, обладает большей эффективностью по сравнению с градиентным методом с постоянным шагом.

6. Заключение

В работе предложен и проанализирован ряд методов градиентного типа, предназначенных для решения задач минимизации (L_0, L_1) -гладких функций при условии аддитивной неточности в значениях градиента на итерациях. Внимание было сосредоточено на исследовании сходимости методов для некоторых классов невыпуклых функций, таких как ρ -квасар-выпуклые функции и функции, удовлетворяющие условию Поляка–Лоясиевича.

В работе представлен обзор современного состояния исследований в этой области. Все ключевые определения и вспомогательные результаты адаптированы на случай Δ -неточного градиента. Сформулированы и доказаны леммы, характеризующие влияние аддитивной неточности градиента на траекторию градиентного метода. Получена теоретическая оценка скорости сходимости градиентного метода на классе квазар-выпуклых функций, удовлетворяющих условию Поляка–Лоясиевича. В частности, показано, что обеспечивает

ся сходимость метода с близкой к линейной скоростью (с точностью до параметров, связанных с погрешностями). Доказано, что начиная с некоторого номера итерации метод сходится в окрестность точного решения на классе функций с условием Поляка–Лоясевича. Отдельное внимание уделено детальному исследованию случая $L_0 = 0$ (класс $(0, M)$ -гладких функций), который представляет значительный интерес для задач машинного обучения, таких как обучение моделей логистической регрессии. Разработан адаптивный вариант градиентного метода, использующий неточные градиенты, для которого получена теоретическая оценка скорости сходимости. Показано, что в случае точных градиентов (при $\Delta = 0$) метод сходится со скоростью геометрической прогрессии. Эффективность предложенных методов подтверждена результатами численных экспериментов, проведенных как на задаче логистической регрессии, так и на некоторой невыпуклой квазар-выпуклой задаче.

В качестве направлений дальнейших исследований представляется интересным изучение ускоренных методов для (L_0, L_1) -гладких задач с неточным градиентом, а также разработка методов с адаптивной настройкой на параметры L_0 и L_1 .

СПИСОК ЛИТЕРАТУРЫ

1. *Zhang J., He T., Sra S., Jadbabaie A.* Why gradient clipping accelerates training: A theoretical justification for adaptivity // International Conference on Learning Representations. 2020. P. 1–14.
2. *Schmidt M., Le Roux N., Bach F.* Minimizing finite sums with the stochastic average gradient // Mathematical Programming. 2017. V. 162. P. 83–112.
3. *Johnson R., Zhang T.* Accelerating stochastic gradient descent using predictive variance reduction // Advances in neural information processing systems. 2013. No. 26. P. 315–323.
4. *Defazio A., Bach F., Lacoste-Julien S.* SAGA: A Fast Incremental Gradient Method With Support for Non-Strongly Convex Composite Objectives // Advances in neural information processing systems. 2014. No. 2. P. 1646–1654.
5. *Nguyen L., Liu J., Scheinberg K., Takač M.* Sarah: A novel method for machine learning problems using stochastic recursive gradient // 34th International Conference on Machine Learning (ICML). 2014. V. 6. P. 4009–4023.
6. *Nguyen L., Scheinberg K., Takač M.* Inexact sarah algorithm for stochastic optimization // Optimization Methods and Software. 2021. V. 36. No. 1. P. 237–258.
7. *Beznosikov A., Takač M.* Random-reshuffled SARAH does not need full gradient computations // Optim Lett. 2024. V. 18. P. 727–749.
8. *Shi Z., Sadiev A., Loizou N., et al.* Ai-sarah: Adaptive and implicit stochastic recursive gradient methods // Transactions on Machine Learning Research. 2023. P. 1–40.
9. *Defazio A., Bottou L.* On the ineffectiveness of variance reduced optimization for deep learning // Advances in Neural Information Processing Systems. 2019. V. 32. P. 1753–1763.

10. *Zhang B., Jin J., Fang C., Wang L.* Improved Analysis of Clipping Algorithms for Non-convex Optimization // Advances in Neural Information Processing Systems. 2020. V. 19. P. 15511–15522.
11. *Chen Z., Zhou Y., Liang Y., Lu Z.* Generalized-smooth nonconvex optimization is as efficient as smooth nonconvex optimization // International Conference on Machine Learning (PMLR). 2023. P. 5396–5427.
12. *Zhao S.-Y., Xie Y.-P., Li W.-J.* On the convergence and improvement of stochastic normalized gradient descent // Science China Information Sciences. 2021. V. 64. P. 1–13.
13. *Faw M., Rout L., Caramanis C., Shakkottai S.* Beyond uniform smoothness: A stopped analysis of adaptive sgd // The Thirty Sixth Annual Conference on Learning Theory (PMLR). 2023. P. 89–160.
14. *Wang B., Zhang H., Ma Z., Chen W.* Convergence of adagrad for non-convex objectives: Simple proofs and relaxed assumptions // The Thirty Sixth Annual Conference on Learning Theory (PMLR). 2023. P. 161–190.
15. *Li H., Rakhlin A., Jadbabaie A.* Convergence of adam under relaxed assumptions // Advances in Neural Information Processing Systems. 2024. P. 1792–1804.
16. *Hubler F., Yang J., Li X., He N.* Parameter-agnostic optimization under relaxed smoothness // International Conference on Artificial Intelligence and Statistics (PMLR). 2024. P. 4861–4869.
17. *Pascanu R., Mikolov T., Bengio Y.* On the difficulty of training recurrent neural networks // International Conference on Machine Learning. 2013. V. 28. P. 1310–1318.
18. *Polyak B.* Introduction to Optimization // Optimization Software. 1987. 438 p.
19. *Koloskova A., Hendrikx H., Stich S.* Revisiting gradient clipping: Stochastic bias and tight convergence guarantees // International Conference on Machine Learning. 2023. P. 17343–17363.
20. *Takezawa Y., Bao H., Sato R. et al.* Polyak meets parameter-free clipped gradient descent // 2024. arXiv:2405.15010
21. *Li H., Qian J., Tian Y., Rakhlin A., Jadbabaie A.* Convex and non-convex optimization under generalized smoothness // Advances in Neural Information Processing Systems. 2024. V. 36. P. 2675–2686.
22. *Gorbunov E., Tupitsa N., Choudhury S., et al.* Methods for Convex (L_0, L_1) -Smooth Optimization: Clipping, Acceleration, and Adaptivity // 2024. arXiv:2409.14989
23. *Vankov D., Rodomanov A., Nedich A., et al.* Optimizing (L_0, L_1) -Smooth Functions by Gradient Methods // 2024. arXiv:2410.10800
24. *Lobanov A., Gasnikov A., Gorbunov E., Takáč M.* Linear Convergence Rate in Convex Setup is Possible! Gradient Descent Method Variants under (L_0, L_1) -Smoothness // 2024. arXiv:2412.17050
25. *Stonyakin F., Kuruzov I., Polyak B.* Stopping Rules for Gradient Methods for Non-convex Problems with Additive Noise in Gradient // Journal of Optimization Theory and Applications. 2022. V. 198. P. 531–551.
26. *Wang J., Wibisono A.* Continuized Acceleration for Quasar Convex Functions in Non-Convex Optimization // 2023. 10.48550/arXiv.2302.07851
27. *Hermant J., Aujol J.F., Dossal C., Rondepierre A.* Study of the behaviour of Nesterov accelerated gradient in a non convex setting: the strongly quasar convex case // 2024. arXiv:2405.19809

28. *Hinder O., Sidford A., Sohoni N.* Near-Optimal Methods for Minimizing Star-Convex Functions and Beyond // 2019. arXiv.1906.11985.
29. *Karimi H., Nutini J., Schmidt M.* Linear convergence of gradient and proximal-gradient methods under the Polyak–Lojasiewicz condition // Joint European Conference on Machine Learning and Knowledge Discovery in Databases. 2016. P. 795–811.
30. *Stonyakin F., Tyurin A., Gasnikov A., et al.* Inexact Relative Smoothness and Strong Convexity for Optimization and Variational Inequalities by Inexact Model // 2024. arXiv preprint arXiv:2402.06319.

Статья представлена к публикации членом редколлегии П.С. Щербаковым.

Поступила в редакцию 03.03.2025

После доработки 20.05.2025

Принята к публикации 27.06.2025

© 2025 г. А. АХАВАН, д-р (arya.akhavan@stats.ox.ac.uk)
(Оксфордский университет, Великобритания),
А.Б. ЦЫБАКОВ, д-р физ.-мат. наук (alexandre.tsybakov@ensae.fr)
(Центр исследований по экономике и статистике (CREST);
Парижская национальная школа статистики
и экономического управления (ENSAE);
Политехнический институт Парижа, Франция)

БЕЗГРАДИЕНТНАЯ СТОХАСТИЧЕСКАЯ ОПТИМИЗАЦИЯ ДЛЯ АДДИТИВНЫХ МОДЕЛЕЙ¹

Рассматривается задача оптимизации нулевого порядка по зашумленным наблюдениям для целевой функции, удовлетворяющей условию Поляка–Лоясевича или условию сильной выпуклости. Кроме того, предполагается, что целевая функция имеет аддитивную структуру и удовлетворяет свойству гладкости высокого порядка, характеризующему гильдеровым семейством функций. Аддитивная модель для гильдеровых классов функций хорошо изучена в литературе по непараметрическому оцениванию функций; в частности, показано, что точность оценивания для такой модели существенно лучше, чем для гильдеровой модели без аддитивной структуры. В данной статье аддитивная модель изучается в задаче безградиентной оптимизации. Предлагается рандомизированная оценка градиента, позволяющая при подключении к алгоритму градиентного спуска достичь минимаксно-оптимальной ошибки оптимизации порядка $dT^{-(\beta-1)/\beta}$, где d – размерность задачи, T – количество пробных точек, а $\beta \geq 2$ – гильдерова степень гладкости. Устанавливается, что, в отличие от непараметрических задач оценивания, использование аддитивных моделей в безградиентной оптимизации не приводит к существенному выигрышу в точности.

Ключевые слова: аддитивная модель, безградиентная оптимизация, минимаксная оптимальность, условие Поляка–Лоясевича.

DOI: 10.31857/S0005231025090025, **EDN:** VMQENP

1. Введение

Аддитивное моделирование является распространенным подходом к снижению размерности в задачах непараметрического оценивания [1–3]. Оно заключается в следующем. Предполагается, что неизвестная функция $f : \mathbb{R}^d \rightarrow \mathbb{R}$, которую необходимо оценить по доступным данным, имеет вид $f(\mathbf{x}) = \sum_{j=1}^d f_j(x_j)$, где x_j – координаты вектора $\mathbf{x} \in \mathbb{R}^d$, а f_j – неизвестные функции одной переменной. Основное свойство, доказанное в литературе по аддитивным моделям непараметрической регрессии, можно сформулировать так.

¹ Исследование Арьи Ахаван поддержано Фондом исследований и инноваций Великобритании (UKRI) в рамках государственной гарантии финансирования “Горизонт Европа” (Horizon Europe) правительства Великобритании (грант № EP/Y028333/1).

Если каждая из функций f_j является β -гельдеровой (см. определение 1 ниже), то минимаксная ошибка оценивания f , поточечно или по норме L_2 , имеет порядок $n^{-\beta/(2\beta+1)}$, где n – число наблюдений [1]. Данный результат сильно отличается от того, что можно получить в задаче оценивания β -гельдеровых функций на \mathbb{R}^d без аддитивной структуры: для таких функций минимаксная ошибка имеет порядок $n^{-\beta/(2\beta+d)}$ [4–7]. Таким образом, при переходе от общих моделей непараметрической регрессии к аддитивным существенно уменьшается ошибка оценивания.

В настоящей работе показывается, что подобное свойство снижения размерности не имеет места в задаче безградиентной оптимизации. Рассматривается задача минимизации неизвестной функции $f : \mathbb{R}^d \rightarrow \mathbb{R}$, удовлетворяющей аддитивной модели, когда доступны только последовательные зашумленные оценки значений f . Предполагается, что функция f либо сильно выпукла, либо удовлетворяет условию Поляка–Лоясевича (условию PL) [8, 9] и допускает аддитивное представление (см. выше) с β -гельдеровыми компонентами f_j . Функции, удовлетворяющие условию Поляка–Лоясевича, будем для краткости называть PL-функциями.

Рассматриваемая постановка относится к семейству задач безградиентной стохастической оптимизации (оптимизации нулевого порядка), по которым имеется богатая литература, см. публикации [10–23] и ссылки в них. В этих работах аддитивная структура f не предполагалась. В [13] доказано, что для β -гельдеровой ($\beta \geq 2$) функции f , удовлетворяющей условию квадратичного роста, минимаксно-оптимальная ошибка оптимизации, как функция количества T последовательно выбранных пробных точек, имеет порядок $T^{-(\beta-1)/\beta}$ с точностью до неопределенного множителя, зависящего от размерности d . Дальнейшие исследования были посвящены зависимости минимаксной ошибки от d в предположении, что функция f является β -гельдеровой с $\beta \geq 2$ и удовлетворяет либо условию сильной выпуклости [14, 15, 18, 21, 23, 24], либо условию PL [21, 25]. В случае PL-функций изучалась минимизация без ограничений, в то время как сильно выпуклый случай анализировался как при наличии ограничений, так и без них. Достигнут значительный прогресс, хотя полностью (для всех $\beta \geq 2$) задача пока не решена. Получена минимаксная нижняя граница для класса β -гельдеровых и сильно выпуклых функций, которая имеет порядок $dT^{-(\beta-1)/\beta}$. Данный результат доказан в [15] для $\beta = 2$ и гауссовского шума и в [19] для всех $\beta \geq 2$ и более общего шума; кроме того, более общая нижняя граница представлена в [21]. С другой стороны, для $\beta = 2$ существует алгоритм, достигающий ошибки такого же порядка при общих условиях на шум (без предположения о независимости или нулевом среднем), см. [19]. Таким образом, для $\beta = 2$ известно, что минимаксная ошибка имеет порядок d/\sqrt{T} . Для $\beta > 2$ в литературе приведены различные зависимости верхних границ от d , определяемые геометрией условия β -гельдеровости. К примеру, для гильдеровых классов, определяемых через приближение многочленом Тейлора, наилучшая известная верхняя граница

имеет порядок $d^{2-1/\beta}T^{-(\beta-1)/\beta}$ в случае сильно выпуклых функций [23, 24]. С другой стороны, для гельдеровых классов, заданных условиями тензорного типа, можно достичь ошибки порядка $d^{2-2/\beta}T^{-(\beta-1)/\beta}$ как для сильно выпуклых функций, так и для PL-функций [21]. Наконец, также в предположении сильной выпуклости, в недавней работе [26] рассмотрен класс функций, у которых гессиан удовлетворяет условию Липшица (разновидность условия Гельдера для $\beta = 3$) и получена верхняя граница порядка $dT^{-2/3}$. Отметим, что нижняя граница из [19] с ошибкой $dT^{-(\beta-1)/\beta}$ справедлива для всех вышеперечисленных гельдеровых классов, поскольку она получена для аддитивных функций, принадлежащих всем этим классам. Таким образом, ошибка $dT^{-(\beta-1)/\beta}$ оказывается минимаксно-оптимальной не только для $\beta = 2$, но и для $\beta = 3$ при подходящем определении 3-гельдероваго класса сильно выпуклых функций.

Основным результатом настоящей работы является верхняя граница $dT^{-(\beta-1)/\beta}$, установленная для ошибки оптимизации в безградиентной постановке для класса β -гельдеровых функций при наличии шума, описываемых аддитивной моделью и удовлетворяющих условию PL или условию сильной выпуклости. Вместе с нижней границей, доказанной в [19], это означает, что $dT^{-(\beta-1)/\beta}$ – минимаксно-оптимальная ошибка оптимизации в данной постановке для всех $\beta \geq 2$. Это заключение является довольно неожиданным, так как оно противоречит интуитивным представлениям на основе упомянутых выше классических результатов по непараметрическому оцениванию. Действительно, по крайней мере для $\beta \in \{2, 3\}$, при переходе от общей β -гельдеровой модели к аддитивной β -гельдеровой модели ошибка не улучшается, ни по T , ни по d . Улучшение может проявляться только в величине множителя, не зависящего от T и d . Подобное свойство можно объяснить сравнительной простотой оптимизационной постановки по сравнению с непараметрическим оцениванием функций в том смысле, что ее целью является оценка функционала от неизвестной функции f (ее минимума), а не самой функции f . Приведем другой, схожий факт из области безградиентной стохастической оптимизации. А именно, нет существенной разницы между сложностью минимизации сильно выпуклой функции с липшицевым градиентом (что соответствует рассмотренному выше случаю $\beta = 2$ с ошибкой порядка d/\sqrt{T}) и сложностью минимизации выпуклой функции без дополнительных свойств, для которой можно построить алгоритм, сходящийся с ошибкой в пределах от $d^{1.5}/\sqrt{T}$ до $d^{1.75}/\sqrt{T}$ (с точностью до логарифмического множителя) [27].

2. Постановка задачи

Пусть Θ – замкнутое выпуклое подмножество пространства \mathbb{R}^d . Рассмотрим задачу минимизации неизвестной функции $f : \mathbb{R}^d \rightarrow \mathbb{R}$ на множестве Θ по наблюдениям значений f на фоне шума в пробных точках, которые могут быть выбраны последовательно в зависимости от предыдущих наблюдений. А именно, предполагаем, что на шаге $t \in \{1, \dots, T\}$ наблюдаются два зашум-

ленных значения f в точках $\mathbf{z}_t, \mathbf{z}'_t \in \mathbb{R}^d$:

$$y_t = f(\mathbf{z}_t) + \xi_t, \quad y'_t = f(\mathbf{z}'_t) + \xi'_t,$$

где ξ_t, ξ'_t – скалярные шумы, а пробные точки $\mathbf{z}_t, \mathbf{z}'_t$ могут быть выбраны в зависимости от $\{\mathbf{z}_i, \mathbf{z}'_i, y_i, y'_i\}_{i=1}^{t-1}$ и от возможной рандомизации.

Далее везде предполагается, что f описывается аддитивной моделью

$$f(\mathbf{x}) = \sum_{j=1}^d f_j(x_j),$$

где x_j – координаты вектора $\mathbf{x} \in \mathbb{R}^d$, а f_j – неизвестные функции одной переменной.

Предполагается, что каждая из функций $f_j : \mathbb{R} \rightarrow \mathbb{R}$, $j = 1, \dots, d$, принадлежит классу β -гельдеровых функций с $\beta \geq 2$, который определяется следующим образом.

Определение 1. Для $\beta > 0$ и $L > 0$ обозначим через $\mathcal{F}_\beta(L)$ множество всех функций $f : \mathbb{R} \rightarrow \mathbb{R}$, которые $\ell = \lfloor \beta \rfloor$ раз дифференцируемы и удовлетворяют для всех $x, z \in \mathbb{R}$ условию

$$(1) \quad \left| f(z) - \sum_{m=0}^{\ell} \frac{1}{m!} f^{(m)}(x)(z-x)^m \right| \leq L|z-x|^\beta,$$

где $f^{(m)}$ – производная функции f порядка m , а $\lfloor \beta \rfloor$ – максимальное целое число, меньшее β . Элементы класса $\mathcal{F}_\beta(L)$ будем называть β -гельдеровыми функциями.

Если $\beta > 2$, то из $f_j \in \mathcal{F}_\beta(L)$ не следует, что $f_j \in \mathcal{F}_2(L)$, однако понадобится и последнее условие. Удобно использовать его в несколько иной форме, задаваемой следующим определением.

Определение 2. Функция $f : \mathbb{R} \rightarrow \mathbb{R}$ называется \bar{L} -гладкой, если она дифференцируема на \mathbb{R} и существует число $\bar{L} > 0$ такое, что для любых $x, x' \in \mathbb{R}$ справедливо

$$|f'(x) - f'(x')| \leq \bar{L}|x - x'|.$$

Обозначим через $\mathcal{F}'_2(\bar{L})$ класс всех \bar{L} -гладких функций.

Также предположим, что f является либо α -сильно выпуклой функцией, либо α -PL функцией (т.е. удовлетворяет условию PL с заданным α); см. два определения ниже.

Определение 3. Пусть $\alpha > 0$. Функция $f : \mathbb{R}^d \rightarrow \mathbb{R}$ называется α -PL функцией, если она дифференцируема на \mathbb{R}^d и

$$2\alpha \left(f(\mathbf{x}) - \min_{\mathbf{z} \in \mathbb{R}^d} f(\mathbf{z}) \right) \leq \|\nabla f(\mathbf{x})\|^2 \quad \text{для всех } \mathbf{x} \in \mathbb{R}^d,$$

где $\|\cdot\|$ обозначает евклидову норму.

Функции, удовлетворяющие условию PL, не обязательно являются выпуклыми. Условие PL является полезным инструментом в задачах оптимизации, поскольку, как показано Б.Т. Поляком [8], оно обеспечивает линейную сходимость алгоритма градиентного спуска без свойства выпуклости. Более подробное обсуждение условия PL можно найти в [28].

Определение 4. Пусть $\alpha > 0$. Функция $f : \mathbb{R}^d \rightarrow \mathbb{R}$ называется α -сильно выпуклой, если она дифференцируема на \mathbb{R}^d и

$$f(\mathbf{x}) - f(\mathbf{x}') \leq \langle \nabla f(\mathbf{x}), \mathbf{x} - \mathbf{x}' \rangle - \frac{\alpha}{2} \|\mathbf{x} - \mathbf{x}'\|^2 \quad \text{для всех } \mathbf{x}, \mathbf{x}' \in \mathbb{R}^d.$$

Для минимизации f применим вариант метода проекции градиента (см. алгоритм 1). Пусть $\{\eta_t\}_{t=1}^T$ – последовательность положительных чисел, $\{\mathbf{g}_t\}_{t=1}^T$ – последовательность случайных векторов и $\mathbf{x}_1 \in \mathbb{R}^d$ – произвольный фиксированный вектор. Определим векторы \mathbf{x}_t , $t = 2, \dots, T$ с помощью рекуррентного соотношения

$$(2) \quad \mathbf{x}_{t+1} = \text{Proj}_{\Theta}(\mathbf{x}_t - \eta_t \mathbf{g}_t),$$

где $\text{Proj}_{\Theta}(\cdot)$ обозначает евклидову проекцию на Θ .

Алгоритм 1

Вход: Множество Θ , функция $K : [-1, 1] \rightarrow \mathbb{R}$, размер шага $\eta_t > 0$ и параметр возмущения $h_t > 0$ для $t = 1, \dots, T$.

Инициализация: Сгенерировать векторы $\mathbf{r}_t = (r_{t,1}, \dots, r_{t,d}) \in \mathbb{R}^d$, $t = 1, \dots, T$, с независимыми равномерно распределенными на $[-1, 1]$ компонентами и выбрать $\mathbf{x}_1 \in \Theta$.

Для $t = 1, \dots, T$ выполнить:

- Получить наблюдения $y_t = f(\mathbf{x}_t + h_t \mathbf{r}_t) + \xi_t$ и $y'_t = f(\mathbf{x}_t - h_t \mathbf{r}_t) + \xi'_t$
- **Для $j = 1, \dots, d$ выполнить:**
 - $g_{t,j} := \frac{d}{2h_t}(y_t - y'_t)K(r_{t,j})$
- $\mathbf{g}_t := (g_{t,1}, \dots, g_{t,d})$ // оценка градиента
- $\mathbf{x}_{t+1} = \text{Proj}_{\Theta}(\mathbf{x}_t - \eta_t \mathbf{g}_t)$ // обновление

В данной работе оценка градиента $\mathbf{g}_t = (g_{t,1}, \dots, g_{t,d}) \in \mathbb{R}^d$ на шаге $t \in \{1, \dots, T\}$ алгоритма определяется следующим образом. Для заданного $\beta \geq 2$ и $\ell = \lfloor \beta \rfloor$, пусть $K : [-1, 1] \rightarrow \mathbb{R}$ – функция такая, что

$$(3) \quad \int u K(u) du = 1, \quad \int u^j K(u) du = 0, \quad j = 0, 2, 3, \dots, \ell,$$

$$\kappa_\beta \equiv \int |u|^\beta |K(u)| du < \infty.$$

Предполагается, что величина $\kappa := \int K^2(r)dr$ конечна. Несложно построить функции K , удовлетворяющие этим условиям. В частности, можно использовать полиномы Лежандра, например см. [13, 18, 29].

На каждом шаге t алгоритма генерируем случайный вектор $\mathbf{r}_t = (r_{t,1}, \dots, r_{t,d}) \in \mathbb{R}^d$, где компоненты $r_{t,j}$ независимы и распределены равномерно на $[-1, 1]$. Взяв $h_t > 0$, произведем два наблюдения

$$y_t = f(\mathbf{x}_t + h_t \mathbf{r}_t) + \xi_t, \quad y'_t = f(\mathbf{x}_t - h_t \mathbf{r}_t) + \xi'_t,$$

и определим величины

$$(4) \quad g_{t,j} = \frac{1}{2h_t}(y_t - y'_t)K(r_{t,j}),$$

где $j \in \{1, \dots, d\}$. Будем рассматривать оценку градиента вида $\mathbf{g}_t = (g_{t,1}, \dots, g_{t,d})$.

Заметим, что применяя другие оценки градиента можно получить результаты, аналогичные представленным ниже, в частности, используя оценки, основанные на конечно-разностной аппроксимации с учетом гладкости более высокого порядка. В отличие от (4), такие конечно-разностные схемы более высокого порядка имеют сложный вид и для них требуется много пробных точек на каждом шаге алгоритма.

Допустим, что переменные шума ξ_t, ξ'_t и рандомизирующие переменные $r_{t,j}$ удовлетворяют следующим условиям.

Предположение 1. Существует величина $\sigma^2 > 0$ такая, что для всех $t \in \{1, \dots, T\}$ справедливо:

- (i) случайные величины $r_{t,j} \sim U[-1, 1]$, $j = 1, \dots, d$, независимы от \mathbf{x}_t и условно независимы от ξ_t, ξ'_t при заданном векторе \mathbf{x}_t ;
- (ii) $\mathbb{E}[\xi_t^2] \leq \sigma^2$, $\mathbb{E}[(\xi'_t)^2] \leq \sigma^2$.

Пункт (i) предположения 1 можно рассматривать не как ограничение, а как часть определения алгоритма, связанную с выбором рандомизирующих переменных $r_{t,j}$. Эти переменные естественным образом выбираются независимо от всех других источников случайности. Для доказательств оказывается достаточным даже более слабое свойство, которое и приведено здесь в виде предположения, чтобы можно было сослаться на него в дальнейшем. Отметим также, что в предположении 1 не требуется, чтобы шумы ξ_t, ξ'_t имели нулевое среднее. Более того, они могут быть неслучайными, и не предполагается независимость этих шумов на разных шагах алгоритма. Тот факт, что сходимость безградиентных алгоритмов может быть достигнута при общих условиях подобного типа на шумы, не требующих независимости и нулевых средних, восходит к работам [30, 31].

3. Основные результаты

В этом разделе представлены верхние границы для ошибки оптимизации алгоритма, определенного в разделе 2. Предположим сначала, что функция f

описывается аддитивной моделью и является α -PL функцией. Заметим, что при наложении этого условия на сумму f все компоненты f_j являются PL-функциями. В классе PL-функций рассмотрим задачу минимизации без ограничений:

$$f^* = \min_{\mathbf{x} \in \mathbb{R}^d} f(\mathbf{x}).$$

Обозначим через $[n]$ множество натуральных чисел, не превышающих заданного натурального числа n .

Теорема 1. Пусть $\alpha > 0$, $\beta \geq 2$, $\bar{L}, L > 0$, и функции $f_j : \mathbb{R} \rightarrow \mathbb{R}$, $j \in [d]$, таковы, что $f_j \in \mathcal{F}'_2(\bar{L}) \cap \mathcal{F}_\beta(L)$. Допустим, что функция $f : \mathbb{R}^d \rightarrow \mathbb{R}$ представима в виде $f(\mathbf{x}) = \sum_{j=1}^d f_j(x_j)$ и является α -PL функцией. Пусть выполнено предположение 1 и пусть $\{\mathbf{x}_t\}_{t=1}^T$ – реализация алгоритма 1 с $\Theta = \mathbb{R}^d$ и

$$\eta_t = \min \left(\frac{4}{\alpha t}, \frac{1}{18\bar{L}d\kappa} \right),$$

$$h_t = \begin{cases} t^{-\frac{1}{2\beta}} & \text{при } \eta_t = \frac{4}{\alpha t}, \\ T^{-\frac{1}{2\beta}} & \text{при } \eta_t = \frac{1}{18\bar{L}d\kappa}. \end{cases}$$

Тогда

$$\mathbf{E} [f(\mathbf{x}_T) - f^*] \leq \frac{A_0}{T} (f(\mathbf{x}_1) - f^*) + A \frac{d}{\alpha} \left(\left(\frac{1}{\alpha T} \right)^{\frac{\beta-1}{\beta}} + \frac{d}{T} \left(\frac{1}{\alpha T} \right)^{\frac{1}{\beta}} \right),$$

где $A_0 = \max(1, 144\bar{L}d\kappa/\alpha)$ и $A > 0$ зависит только от L , \bar{L} , β и σ^2 .

Следствие 1. В условиях теоремы 1, если $T \geq \alpha^{\beta/2-1} d^{\beta/2}$, то

$$\mathbf{E} [f(\mathbf{x}_T) - f^*] \leq \frac{A_0}{T} (f(\mathbf{x}_1) - f^*) + A \frac{d}{\alpha} \left(\frac{1}{\alpha T} \right)^{\frac{\beta-1}{\beta}},$$

где $A > 0$ зависит только от L , \bar{L} , β и σ^2 .

Заметим, что при $2 \leq \beta \leq 3$ условие $T \geq Cd^{\beta/2}$ с некоторой константой $C > 0$ (см. следствие 1) не является ограничительным, поскольку оно слабее, чем задание диапазона для T , при котором граница следствия 1 представляет интерес. Действительно, если $T \leq Cd^{\beta/2}$ и $2 \leq \beta \leq 3$, то данная граница больше некоторой константы, не зависящей от T и d .

Исследуем теперь ошибку оптимизации для алгоритма, определенного в разделе 2, в предположении α -сильной выпуклости целевой функции. В отличие от случая PL-функций, рассмотрим задачу минимизации с ограничениями:

$$\min_{\mathbf{x} \in \Theta} f(\mathbf{x}),$$

где Θ – компактное выпуклое подмножество \mathbb{R}^d .

Теорема 2. Пусть $\alpha > 0$, $\beta \geq 2$, $\bar{L}, L > 0$, пусть Θ – компактное выпуклое подмножество \mathbb{R}^d и функции $f_j : \mathbb{R} \rightarrow \mathbb{R}$, $j \in [d]$, таковы, что $f_j \in \mathcal{F}'_2(\bar{L}) \cap \mathcal{F}_\beta(L)$. Допустим, что функция $f : \mathbb{R}^d \rightarrow \mathbb{R}$ представима в виде $f(\mathbf{x}) = \sum_{j=1}^d f_j(x_j)$, является α -сильно выпуклой функцией и удовлетворяет условию $\max_{\mathbf{x} \in \Theta} \|\nabla f(\mathbf{x})\| \leq G$. Пусть выполнено предположение 1 и пусть $\{\mathbf{x}_t\}_{t=1}^T$ – реализация алгоритма 1 с

$$\eta_t = \frac{4}{\alpha(t+1)}, \quad h_t = \left(\frac{3}{2t} \frac{\kappa \sigma^2}{\kappa_\beta^2 L^2} \right)^{\frac{1}{2\beta}}.$$

Рассмотрим взвешенную оценку

$$\bar{\mathbf{x}}_T = \frac{2}{T(T+1)} \sum_{t=1}^T t \mathbf{x}_t.$$

Тогда для любого $\mathbf{x} \in \Theta$ имеем

$$\mathbf{E}[f(\bar{\mathbf{x}}_T) - f(\mathbf{x})] \leq \frac{36G^2 d \kappa}{\alpha T} + A \frac{d}{\alpha} \left(1 + dT^{-\frac{2}{\beta}}\right) T^{-\frac{\beta-1}{\beta}},$$

где $A > 0$ зависит только от L , \bar{L} , β и σ^2 .

Следствие 2. В условиях теоремы 2, если $T \geq d^{\beta/2}$, то

$$\mathbf{E}\left[f(\bar{\mathbf{x}}_T) - \min_{\mathbf{x} \in \Theta} f(\mathbf{x})\right] \leq A \frac{d}{\alpha} T^{-\frac{\beta-1}{\beta}},$$

где $A > 0$ не зависит от T, d и α .

Так же, как и в связи со следствием 1, можно отметить следующий факт: при $2 \leq \beta \leq 3$ условие $T \geq Cd^{\beta/2}$ с константой $C > 0$ в следствии 2 не является ограничительным, а указывает на значимый диапазон величины T , поскольку при $T \leq Cd^{\beta/2}$ граница следствия 2 больше константы.

Замечание 1. Ввиду сильной выпуклости теорема 2 и следствие 2 немедленно приводят к соответствующим границам для ошибки оценивания $\mathbf{E}[\|\bar{\mathbf{x}}_T - \mathbf{x}^*\|^2]$, где \mathbf{x}^* – точка минимума f на Θ , если в ней достигается глобальный минимум. Таким образом, при условиях следствия 2, если $\nabla f(\mathbf{x}^*) = 0$, то

$$\mathbf{E}[\|\bar{\mathbf{x}}_T - \mathbf{x}^*\|^2] \leq 2A \frac{d}{\alpha^2} T^{-\frac{\beta-1}{\beta}},$$

где $A > 0$ – константа из следствия 2.

Согласно следствию 2 и нижним границам, доказанным в [19, 21], при неограничительных условиях на параметры задачи, ошибка $\frac{d}{\alpha} T^{-\frac{\beta-1}{\beta}}$ является минимаксно-оптимальной в классе аддитивных функций f , удовлетворяющих условиям теоремы 2. Искомая нижняя граница не указана в явном виде

в [19, 21], но непосредственно следует из представленных там доказательств, так как нижние границы в [19, 21] получены на аддитивных функциях. Для полноты изложения приведем здесь утверждение о нижней границе для аддитивных функций, основанное на результатах [21].

Рассмотрим все стратегии выбора пробных точек в виде

$$\mathbf{z}_t = \Phi_t((z_i, y_i)_{i=1}^{t-1}, (z'_i, y'_i)_{i=1}^{t-1}, \tau_t) \text{ и } \mathbf{z}'_t = \Phi'_t((z_i, y_i)_{i=1}^{t-1}, (z'_i, y'_i)_{i=1}^{t-1}, \tau_t) \text{ при } t \geq 2,$$

где Φ_t и Φ'_t – измеримые функции, $\mathbf{z}_1, \mathbf{z}'_1 \in \mathbb{R}^d$ – любые случайные величины, а $\{\tau_t\}$ – последовательность случайных величин со значениями в измеримом пространстве $(\mathcal{Z}, \mathcal{U})$ такая, что τ_t не зависит от $((z_i, y_i)_{i=1}^{t-1}, (z'_i, y'_i)_{i=1}^{t-1})$. Обозначим через Π_T множество всех таких стратегий выбора пробных точек вплоть до шага $t = T$. Класс Π_T содержит последовательную стратегию алгоритма из раздела 2 с оценкой градиента (4). В этом случае $\tau_t = \mathbf{r}_t$, $\mathbf{z}_t = \mathbf{x}_t + h_t \mathbf{r}_t$ и $\mathbf{z}'_t = \mathbf{x}_t - h_t \mathbf{r}_t$.

Используемая здесь нижняя граница из [21] доказана при следующем предположении на шумы (ξ, ξ') . Пусть $H^2(\cdot, \cdot)$ – квадрат расстояния Хеллингера, который определяется для двух вероятностных мер \mathbf{P}, \mathbf{P}' на измеримом пространстве (Ω, \mathcal{A}) по формуле

$$H^2(\mathbf{P}, \mathbf{P}') = \int (\sqrt{d\mathbf{P}} - \sqrt{d\mathbf{P}'})^2.$$

Предположение 2. Для каждого шага $t \geq 1$ справедливо следующее:

- функция распределения $F_t : \mathbb{R}^2 \rightarrow \mathbb{R}$ случайной величины (ξ_t, ξ'_t) такова, что

$$(5) \quad H^2(P_{F_t(\cdot, \cdot)}, P_{F_t(\cdot+v, \cdot+w)}) \leq I_0 \max(v^2, w^2), \quad |v|, |w| \leq v_0,$$

для некоторых $0 < I_0 < \infty$, $0 < v_0 \leq \infty$. Здесь $P_{F(\cdot, \cdot)}$ обозначает вероятностную меру, соответствующую функции распределения $F(\cdot, \cdot)$;

- случайная величина (ξ_t, ξ'_t) не зависит от $((z_i, y_i)_{i=1}^{t-1}, (z'_i, y'_i)_{i=1}^{t-1}, \tau_t)$.

Пусть $\Theta = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| \leq 1\}$. При фиксированных $\alpha, L, \bar{L} > 0$, $G > \alpha$, $\beta \geq 2$ обозначим через \mathcal{F} множество всех функций f , удовлетворяющих условиям теоремы 2 и достигающих своего минимума на \mathbb{R}^d в множестве Θ .

Теорема 3. Пусть $\Theta = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\| \leq 1\}$ и верно предположение 2. Допустим, что $\alpha > T^{-1/2+1/\beta}$ и $T \geq d^\beta$. Тогда для любой оценки $\tilde{\mathbf{x}}_T$ по наблюдениям $((z_t, y_t), (z'_t, y'_t), t = 1, \dots, T)$, где $((z_t, z'_t), t = 1, \dots, T)$ получены с помощью произвольной стратегии из класса Π_T , имеем

$$(6) \quad \sup_{f \in \mathcal{F}} \mathbf{E}[f(\tilde{\mathbf{x}}_T) - \min_{\mathbf{x} \in \Theta} f(\mathbf{x})] \geq C \frac{d}{\alpha} T^{-\frac{\beta-1}{\beta}}$$

с константой $C > 0$, не зависящей от T, d и α .

Теорема 3 непосредственно следует из доказательства в [21, теорема 22], так как семейство используемых там функций принадлежит классу \mathcal{F} . Нижняя граница из [21, теорема 22] имеет вид

$$C \min \left(\max(\alpha, T^{-1/2+1/\beta}), \frac{d}{\sqrt{T}}, \frac{d}{\alpha} T^{-\frac{\beta-1}{\beta}} \right)$$

и сводится к $C \frac{d}{\alpha} T^{-\frac{\beta-1}{\beta}}$ при предположениях на T, d и α , использованных в теореме 3.

Замечание 2. Поскольку сильная выпуклость и свойство PL справедливы для каждой аддитивной компоненты функции f , возможной альтернативой была бы покоординатная процедура (минимизация каждой компоненты f_j отдельно). Однако подобный подход приводит к худшему результату. Действительно, в этом случае на каждом шаге необходимо параллельно выбирать $2d$ пробных точек (по две на каждую компоненту) и, таким образом, можно сделать только $\sim T/d$ шагов при общем количестве пробных точек, равном T . В итоге, применяя теорему 1 или 2 в одномерном случае, для каждой компоненты получим ошибку порядка $(T/d)^{-(\beta-1)/\beta}$. Эта ошибка не может быть улучшена, что следует из одномерного случая теоремы 3. Суммируя по d компонентам, получаем общую ошибку порядка $d(T/d)^{-(\beta-1)/\beta} = d^{2-1/\beta} T^{-(\beta-1)/\beta}$, т.е. ошибка будет зависеть от d неоптимальным образом.

4. Доказательства

Начнем с доказательства вспомогательных лемм.

Лемма 1. Пусть $f : \mathbb{R}^d \rightarrow \mathbb{R}$ – дифференцируемая функция такая, что $\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{x}')\| \leq \bar{L} \|\mathbf{x} - \mathbf{x}'\|$ для всех $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^d$, где $\bar{L} > 0$. Пусть выполнено предположение 1 и пусть $\{\mathbf{x}_t\}_{t=1}^T$ – реализация алгоритма 1 с $\Theta = \mathbb{R}^d$. Тогда для всех $t \in [T]$ справедливо неравенство

$$(7) \quad \mathbf{E}[f(\mathbf{x}_{t+1})|\mathbf{x}_t] \leq f(\mathbf{x}_t) - \frac{\eta_t}{2} \|\nabla f(\mathbf{x}_t)\|^2 + \frac{\eta_t}{2} \mathbf{E}[\|\mathbf{g}_t|\mathbf{x}_t] - \nabla f(\mathbf{x}_t)\|^2] + \frac{\bar{L}\eta_t^2}{2} \mathbf{E}[\|\mathbf{g}_t\|^2|\mathbf{x}_t].$$

Доказательство. Из условия на f следует, что

$$f(\mathbf{x}_{t+1}) = f(\mathbf{x}_t - \eta_t \mathbf{g}_t) \leq f(\mathbf{x}_t) - \eta_t \langle \nabla f(\mathbf{x}_t), \mathbf{g}_t \rangle + \frac{\bar{L}\eta_t^2}{2} \|\mathbf{g}_t\|^2.$$

Прибавляя и вычитая $\eta_t \|\nabla f(\mathbf{x}_t)\|^2$, получаем

$$f(\mathbf{x}_{t+1}) \leq f(\mathbf{x}_t) - \eta_t \|\nabla f(\mathbf{x}_t)\|^2 - \eta_t \langle \nabla f(\mathbf{x}_t), \mathbf{g}_t - \nabla f(\mathbf{x}_t) \rangle + \frac{\bar{L}\eta_t^2}{2} \|\mathbf{g}_t\|^2.$$

Взяв условное математическое ожидание, будем иметь

$$\begin{aligned} \mathbf{E}[f(\mathbf{x}_{t+1})|\mathbf{x}_t] &\leq f(\mathbf{x}_t) - \eta_t \|\nabla f(\mathbf{x}_t)\|^2 - \eta_t (\nabla f(\mathbf{x}_t), \mathbf{E}[\mathbf{g}_t|\mathbf{x}_t] - \nabla f(\mathbf{x}_t)) + \\ &\quad + \frac{\bar{L}\eta_t^2}{2} \mathbf{E}[\|\mathbf{g}_t\|^2|\mathbf{x}_t] \leq \\ &\leq f(\mathbf{x}_t) - \eta_t \|\nabla f(\mathbf{x}_t)\|^2 + \eta_t \|\nabla f(\mathbf{x}_t)\| \mathbf{E}[\|\mathbf{g}_t|\mathbf{x}_t] - \nabla f(\mathbf{x}_t)\| + \\ &\quad + \frac{\bar{L}\eta_t^2}{2} \mathbf{E}[\|\mathbf{g}_t\|^2|\mathbf{x}_t]. \end{aligned}$$

Используя неравенство $2ab \leq a^2 + b^2$, $\forall a, b \in \mathbb{R}$, приходим к требуемому результату. Лемма доказана.

Лемма 2. Пусть функции $f_j : \mathbb{R} \rightarrow \mathbb{R}$, $j \in [d]$, таковы, что $f_j \in \mathcal{F}_\beta(L)$, где $\beta \geq 2$ и $L > 0$. Допустим, что функция $f : \mathbb{R}^d \rightarrow \mathbb{R}$ задается аддитивной моделью $f(\mathbf{x}) = \sum_{j=1}^d f_j(x_j)$ и верно предположение 1(i). Тогда для всех $t \in [T]$ справедливо неравенство

$$\|\mathbf{E}[\mathbf{g}_t|\mathbf{x}_t] - \nabla f(\mathbf{x}_t)\| \leq \kappa_\beta L \sqrt{d} h_t^{\beta-1}.$$

Доказательство. Используя предположение 1(i) для любых $j \in [d]$ и $t \in [T]$ имеем $\mathbf{E}(K(r_{t,j})) = 0$, $\mathbf{E}[\xi_t K(r_{t,j})|\mathbf{x}_t] = 0$ и $\mathbf{E}[\xi_t^2 K(r_{t,j})|\mathbf{x}_t] = 0$. Таким образом,

$$\begin{aligned} \mathbf{E}[g_{t,j}|\mathbf{x}_t] &= \frac{1}{2h_t} \mathbf{E}[(f_j(x_{t,j} + h_t r_{t,j}) - f_j(x_{t,j} - h_t r_{t,j}))K(r_{t,j})|\mathbf{x}_t] + \\ &\quad + \frac{1}{2h_t} \sum_{m \neq j} \mathbf{E}[(f_m(x_{t,m} + h_t r_{t,m}) - f_m(x_{t,m} - h_t r_{t,m}))K(r_{t,j})|\mathbf{x}_t] = \\ &= \frac{1}{2h_t} \mathbf{E}[(f_j(x_{t,j} + h_t r_{t,j}) - f_j(x_{t,j} - h_t r_{t,j}))K(r_{t,j})|\mathbf{x}_t], \end{aligned}$$

где использовано равенство $\mathbf{E}(K(r_{t,j})) = 0$ и независимость $r_{t,j}$ от $r_{t,m}$ при $m \neq j$, а также от \mathbf{x}_t . В силу разложения Тейлора,

$$\begin{aligned} \frac{1}{2h_t} (f_j(x_{t,j} + h_t r_{t,j}) - f_j(x_{t,j} - h_t r_{t,j})) &= f_j'(x_{t,j}) r_{t,j} + \\ &\quad + \frac{1}{h_t} \sum_{1 \leq m \leq \ell, m \text{ нечетно}} \frac{h_t^m}{m!} f_j^{(m)}(x_{t,j}) r_{t,j}^m + \frac{R(h_t r_{t,j}) - R(-h_t r_{t,j})}{2h_t}, \end{aligned}$$

где $|R(-h_t r_{t,j})|, |R(h_t r_{t,j})| \leq L|r_{t,j}|^\beta h_t^\beta$. Умножив обе части этого неравенства на $K(r_{t,j})$ и взяв условное математическое ожидание, получаем

$$|\mathbf{E}[g_{t,j}|\mathbf{x}_t] - f_j'(x_{t,j})| \leq L \mathbf{E}[|r_{t,j}|^\beta K(r_{t,j})] h_t^{\beta-1} = \kappa_\beta L h_t^{\beta-1}.$$

Лемма следует из этого неравенства в сочетании с оценкой

$$\|\mathbf{E}[\mathbf{g}_t|\mathbf{x}_t] - \nabla f(\mathbf{x}_t)\| \leq \sqrt{d} \max_{j=1, \dots, d} |\mathbf{E}[g_{t,j}|\mathbf{x}_t] - f_j'(x_{t,j})|.$$

Лемма 3. Пусть функции $f_j : \mathbb{R} \rightarrow \mathbb{R}$, $j \in [d]$, таковы, что $f_j \in \mathcal{F}'_2(\bar{L})$, где $\bar{L} > 0$. Допустим, что функция $f : \mathbb{R}^d \rightarrow \mathbb{R}$ задается аддитивной моделью $f(\mathbf{x}) = \sum_{j=1}^d f_j(x_j)$ и верно предположение 1. Тогда для всех $t \in [T]$ справедливо неравенство

$$\mathbf{E} [\|\mathbf{g}_t\|^2 | \mathbf{x}_t] \leq \frac{3}{2} \kappa d \left(\frac{3}{4} (d\bar{L}^2 h_t^2 + 8\|\nabla f(\mathbf{x}_t)\|^2) + \frac{\sigma^2}{h_t^2} \right).$$

Доказательство. Для $i \in [d]$ определим

$$G_i = f_i(x_{t,i} + h_t r_{t,i}) - f_i(x_{t,i} - h_t r_{t,i}).$$

Тогда

$$\mathbf{E} [g_{t,j}^2 | \mathbf{x}_t] = \frac{1}{4h_t^2} \mathbf{E} \left[\left(\sum_{i=1}^d G_i + \xi_t - \xi'_t \right)^2 K^2(r_{t,j}) | \mathbf{x}_t \right].$$

Заметим, что $r_{t,i}$ и $-r_{t,i}$ распределены одинаково. Следовательно, $\mathbf{E}[G_i | \mathbf{x}_t] = 0$ и можно записать

$$\begin{aligned} \mathbf{E} [g_{t,j}^2 | \mathbf{x}_t] &\leq \frac{3}{4h_t^2} \mathbf{E} \left[\left(\left(\sum_{i=1}^d G_i \right)^2 + \xi_t^2 + (\xi'_t)^2 \right) K^2(r_{t,j}) | \mathbf{x}_t \right] \leq \\ &\leq \frac{3}{4h_t^2} \mathbf{E} \left[\sum_{i=1}^d G_i^2 K^2(r_{t,j}) + \sum_{i,k=1, i \neq k}^d G_i G_k K^2(r_{t,j}) | \mathbf{x}_t \right] + \frac{3\sigma^2 \kappa}{2h_t^2} = \\ &= \frac{3}{4h_t^2} \mathbf{E} \left[\sum_{i=1}^d G_i^2 K^2(r_{t,j}) | \mathbf{x}_t \right] + \frac{3\sigma^2 \kappa}{2h_t^2}. \end{aligned}$$

Так как $f_i \in \mathcal{F}'_2(\bar{L})$, для всех $i \in [d]$ имеем

$$\begin{aligned} G_i^2 &= \left((f_i(x_{t,i} + h_t r_{t,i}) - f(x_{t,i}) - f'_i(x_{t,i}) h_t r_{t,i}) - \right. \\ &\quad \left. - (f_i(x_{t,i} - h_t r_{t,i}) - f(x_{t,i}) + f'_i(x_{t,i}) h_t r_{t,i}) + 2f'_i(x_{t,i}) h_t r_{t,i}) \right)^2 \leq \\ &\leq 3 \left((f_i(x_{t,i} + h_t r_{t,i}) - f(x_{t,i}) - f'_i(x_{t,i}) h_t r_{t,i})^2 + \right. \\ &\quad \left. + (f_i(x_{t,i} - h_t r_{t,i}) - f(x_{t,i}) + f'_i(x_{t,i}) h_t r_{t,i})^2 + 4(f'_i(x_{t,i}) h_t r_{t,i})^2 \right) \leq \\ &\leq 3 \left(\frac{\bar{L}^2}{2} h_t^4 + 4(f'_i(x_{t,i}))^2 h_t^2 \right). \end{aligned}$$

Таким образом,

$$\mathbf{E} [g_{t,j}^2 | \mathbf{x}_t] \leq \frac{9}{4} \kappa \left(\frac{d\bar{L}^2}{2} h_t^2 + 4 \sum_{i=1}^d (f'_i(x_{t,i}))^2 \right) + \frac{3\sigma^2 \kappa}{2h_t^2},$$

что завершает доказательство леммы.

Лемма 4. Пусть функции $f_j : \mathbb{R} \rightarrow \mathbb{R}$, $j \in [d]$, таковы, что $f_j \in \mathcal{F}'_2(\bar{L}) \cap \mathcal{F}_\beta(L)$, где $\beta \geq 2$ и $\bar{L}, L > 0$. Допустим, что функция $f : \mathbb{R}^d \rightarrow \mathbb{R}$ задается аддитивной моделью $f(\mathbf{x}) = \sum_{j=1}^d f_j(x_j)$. Пусть выполнено предположение 1 и пусть $\{\mathbf{x}_t\}_{t=1}^T$ – реализация алгоритма 1 с $\Theta = \mathbb{R}^d$ и оценками градиента (4). Тогда для всех $t \in [T]$ справедливо неравенство

$$\begin{aligned} \mathbf{E} [f(\mathbf{x}_{t+1})|\mathbf{x}_t] &\leq f(\mathbf{x}_t) - \frac{\eta_t}{2} (1 - 9\bar{L}dk\eta_t) \|\nabla f(\mathbf{x}_t)\|^2 + \\ &+ \frac{\eta_t}{2} d \left(L\kappa_\beta h_t^{\beta-1} \right)^2 + \frac{3\bar{L}\eta_t^2}{4} \kappa d \left(\frac{\sigma^2}{h_t^2} + \frac{3\bar{L}^2 d}{4} h_t^2 \right). \end{aligned}$$

Доказательство. Требуемый результат следует из лемм 1, 2 и 3.

Лемма 5. Пусть функция f α -сильно выпукла с $\alpha > 0$. Предположим, что $\{\mathbf{x}_t\}_{t=1}^T$ – реализация алгоритма 1. Тогда для всех $t \in [T]$ и $\mathbf{x} \in \Theta$ справедливо неравенство

$$\begin{aligned} f(\mathbf{x}_t) - f(\mathbf{x}) &\leq (2\eta_t)^{-1} (\|\mathbf{x}_t - \mathbf{x}\|^2 - \mathbf{E} [\|\mathbf{x}_{t+1} - \mathbf{x}\|^2|\mathbf{x}_t]) + \\ &+ \frac{1}{\alpha} \|\nabla f(\mathbf{x}_t) - \mathbf{E} [\mathbf{g}_t|\mathbf{x}_t]\|^2 + \frac{\eta_t}{2} \mathbf{E} [\|\mathbf{g}_t\|^2|\mathbf{x}_t] - \frac{\alpha}{4} \|\mathbf{x}_t - \mathbf{x}\|^2. \end{aligned}$$

Доказательство. В силу определения евклидовой проекции, для фиксированного вектора $\mathbf{x} \in \Theta$ имеем $\|\mathbf{x}_{t+1} - \mathbf{x}\|^2 \leq \|\mathbf{x}_t - \eta_t \mathbf{g}_t - \mathbf{x}\|^2$. Это неравенство можно записать в эквивалентном виде

$$\langle \mathbf{g}_t, \mathbf{x}_t - \mathbf{x} \rangle \leq (2\eta_t)^{-1} (\|\mathbf{x}_t - \mathbf{x}\|^2 - \|\mathbf{x}_{t+1} - \mathbf{x}\|^2) + \frac{\eta_t}{2} \|\mathbf{g}_t\|^2.$$

Пусть $a_t = \|\mathbf{x}_t - \mathbf{x}\|^2$. Так как функция f α -сильно выпукла,

$$\begin{aligned} f(\mathbf{x}_t) - f(\mathbf{x}) &\leq \langle \nabla f(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x} \rangle - \frac{\alpha}{2} a_t = \\ &= \langle \mathbf{g}_t, \mathbf{x}_t - \mathbf{x} \rangle + \langle \nabla f(\mathbf{x}_t) - \mathbf{g}_t, \mathbf{x}_t - \mathbf{x} \rangle - \frac{\alpha}{2} a_t \leq \\ &\leq (2\eta_t)^{-1} (a_t - a_{t+1}) + \langle \nabla f(\mathbf{x}_t) - \mathbf{g}_t, \mathbf{x}_t - \mathbf{x} \rangle + \frac{\eta_t}{2} \|\mathbf{g}_t\|^2 - \frac{\alpha}{2} a_t. \end{aligned}$$

Взяв условное математическое ожидание при заданном \mathbf{x}_t и воспользовавшись неравенством $ab \leq a^2/\lambda + \lambda b^2/4$, справедливым для всех $a, b \in \mathbb{R}$ и $\lambda > 0$, получим

$$\begin{aligned} f(\mathbf{x}_t) - f(\mathbf{x}) &\leq (2\eta_t)^{-1} (a_t - \mathbf{E} [a_{t+1}|\mathbf{x}_t]) + \langle \nabla f(\mathbf{x}_t) - \mathbf{E} [\mathbf{g}_t|\mathbf{x}_t], \mathbf{x}_t - \mathbf{x} \rangle + \\ &+ \frac{\eta_t}{2} \mathbf{E} [\|\mathbf{g}_t\|^2|\mathbf{x}_t] - \frac{\alpha}{2} a_t \leq \\ &\leq (2\eta_t)^{-1} (a_t - \mathbf{E} [a_{t+1}|\mathbf{x}_t]) + \|\nabla f(\mathbf{x}_t) - \mathbf{E} [\mathbf{g}_t|\mathbf{x}_t]\| \|\mathbf{x}_t - \mathbf{x}\| + \\ &+ \frac{\eta_t}{2} \mathbf{E} [\|\mathbf{g}_t\|^2|\mathbf{x}_t] - \frac{\alpha}{2} a_t \leq \\ &\leq (2\eta_t)^{-1} (a_t - \mathbf{E} [a_{t+1}|\mathbf{x}_t]) + \frac{1}{\alpha} \|\nabla f(\mathbf{x}_t) - \mathbf{E} [\mathbf{g}_t|\mathbf{x}_t]\|^2 + \\ &+ \frac{\eta_t}{2} \mathbf{E} [\|\mathbf{g}_t\|^2|\mathbf{x}_t] - \frac{\alpha}{4} a_t. \end{aligned}$$

Доказательство теоремы 1. Поскольку $\eta_t \leq 1/(18\bar{L}d\kappa)$, из леммы 4 следует, что

$$\begin{aligned} \mathbf{E}[f(\mathbf{x}_{t+1})|\mathbf{x}_t] &\leq f(\mathbf{x}_t) - \frac{\eta_t}{4}\|\nabla f(\mathbf{x}_t)\|^2 + \\ &+ \frac{\eta_t}{2}d\left(L\kappa_\beta h_t^{\beta-1}\right)^2 + \frac{3\bar{L}\eta_t^2}{4}\kappa d\left(\frac{\sigma^2}{h_t^2} + \frac{3\bar{L}^2 d}{4}h_t^2\right). \end{aligned}$$

Взяв математическое ожидание обеих частей этого неравенства и воспользовавшись тем, что f — α -PL функция, получаем

$$(8) \quad \delta_{t+1} \leq \delta_t \left(1 - \frac{\eta_t \alpha}{2}\right) + \frac{\eta_t}{2}d\left(L\kappa_\beta h_t^{\beta-1}\right)^2 + \frac{3\bar{L}\eta_t^2}{4}\kappa d\left(\frac{\sigma^2}{h_t^2} + \frac{3\bar{L}^2 d}{4}h_t^2\right),$$

где $\delta_t = \mathbf{E}[f(\mathbf{x}_t) - f^*]$. Пусть $T_0 = \lceil 72\bar{L}d\kappa/\alpha \rceil$. Заметим, что

$$\eta_t = \begin{cases} \frac{1}{18\bar{L}d\kappa} & \text{при } t \leq T_0, \\ \frac{4}{\alpha t} & \text{при } t \geq T_0 + 1 \end{cases}$$

и

$$(9) \quad \frac{4}{(T_0 + 1)\alpha} \leq \eta_t \leq \frac{4}{T_0\alpha}, \quad \text{при } t \leq T_0.$$

Рассмотрим отдельно случаи $T \geq T_0 + 1$ и $T \leq T_0$.

Используя (8), легко проверить, что результат теоремы имеет место при $T \geq T_0 + 1$ и $T \in \{1, 2\}$.

Рассмотрим теперь случай, когда $T \geq T_0 + 1$ и $T \geq 3$, т.е. $T \geq T'_0 + 1$, где $T'_0 = \max(T_0, 2)$. При всех $t \geq T_0 + 1$ имеем $\eta_t = 4/(\alpha t)$, так что

$$\delta_{t+1} \leq \delta_t \left(1 - \frac{2}{t}\right) + \frac{d}{\alpha t} \left(2\left(L\kappa_\beta h_t^{\beta-1}\right)^2 + \frac{3\bar{L}}{\alpha t}\kappa \left(\frac{\sigma^2}{h_t^2} + \frac{3\bar{L}^2 d}{4}h_t^2\right)\right).$$

Подставляя сюда $h_t = \left(\frac{3\bar{L}}{\alpha t} \frac{\kappa \sigma^2}{2L^2 \kappa_\beta^2}\right)^{\frac{1}{2\beta}}$, получаем, что при $t \geq T'_0 + 1$ выполняется неравенство

$$\delta_{t+1} \leq \delta_t \left(1 - \frac{2}{t}\right) + \mathbf{A} \frac{d}{\alpha t} \left(\left(\frac{1}{\alpha t}\right)^{\frac{\beta-1}{\beta}} + \frac{d}{t} \left(\frac{1}{\alpha t}\right)^{\frac{1}{\beta}} \right).$$

Здесь и далее через \mathbf{A} обозначаются положительные постоянные, не обязательно принимающие одни и те же значения и зависящие только от L, \bar{L}, β , и σ^2 . Так как $T \geq T'_0 + 1 \geq 3$, то в силу [21, лемма 32] справедлива верхняя граница

$$(10) \quad \delta_T \leq \frac{2T'_0}{T} \delta_{T'_0+1} + \mathbf{A} \frac{d}{\alpha} \left(\left(\frac{1}{\alpha T}\right)^{\frac{\beta-1}{\beta}} + \frac{d}{T} \left(\frac{1}{\alpha T}\right)^{\frac{1}{\beta}} \right).$$

Если $T_0 \leq 2$ (т.е. $T'_0 = 2$), то используя (8) при $t = 1, t = 2$, определение T_0 , соотношение (9) и определение h_t , получаем

$$\delta_{T'_0+1} = \delta_3 \leq \delta_1 + \mathbf{A} \frac{d}{\alpha} \left(\left(\frac{1}{\alpha} \right)^{\frac{\beta-1}{\beta}} + d \left(\frac{1}{\alpha} \right)^{\frac{1}{\beta}} \right).$$

Подстановка этого неравенства в (10) завершает доказательство для случая $T'_0 = 2, T \geq T'_0 + 1$.

Пусть теперь $T'_0 \geq 3$ (что влечет $T_0 = T'_0$) и $T \geq T'_0 + 1$. Из неравенств (8), (9) и определения h_t следует, что для всех $t \leq T_0$,

$$(11) \quad \delta_{t+1} \leq \delta_t \left(1 - \frac{2}{T_0 + 1} \right) + \mathbf{A} \frac{d}{\alpha T_0} \left(\left(\frac{1}{\alpha} \right)^{\frac{\beta-1}{\beta}} \left(T^{-\frac{\beta-1}{\beta}} + \frac{T^{\frac{1}{\beta}}}{T_0} \right) + \frac{d}{T_0} \left(\frac{1}{\alpha T} \right)^{\frac{1}{\beta}} \right).$$

Итерируя (11) с упрощением $1 - 2/(T_0 + 1) \leq 1$ и принимая во внимание определение T_0 и неравенство $T \geq T_0 + 1$, получаем

$$\frac{2T_0}{T} \delta_{T_0+1} \leq \frac{144\bar{L}d\kappa}{\alpha T} \delta_1 + \mathbf{A} \frac{d}{\alpha} \left(\left(\frac{1}{\alpha T} \right)^{\frac{\beta-1}{\beta}} + \frac{d}{T} \left(\frac{1}{\alpha T} \right)^{\frac{1}{\beta}} \right).$$

Подстановка этого неравенства в (10) дает:

$$\delta_T \leq \frac{144\bar{L}d\kappa}{\alpha T} \delta_1 + \mathbf{A} \frac{d}{\alpha} \left(\left(\frac{1}{\alpha T} \right)^{\frac{\beta-1}{\beta}} + \frac{d}{T} \left(\frac{1}{\alpha T} \right)^{\frac{1}{\beta}} \right),$$

что завершает доказательство для случая $T \geq T_0 + 1$.

Рассмотрим теперь случай $T \leq T_0$. В этом случае (11) выполняется при всех $t \leq T$, так что

$$\delta_{T+1} \leq \delta_1 \left(1 - \frac{2}{T_0 + 1} \right)^T + \mathbf{A} \frac{d}{\alpha} \left(\left(\frac{1}{\alpha T} \right)^{\frac{\beta-1}{\beta}} + \frac{d}{T} \left(\frac{1}{\alpha T} \right)^{\frac{1}{\beta}} \right).$$

Так как $(1 - \lambda)^T \leq \exp(-\lambda T) \leq 1/(\lambda T)$ для всех $T > 0, \lambda \in (0, 1)$, то

$$\delta_{T+1} \leq \frac{T_0 + 1}{2T} \delta_1 + \mathbf{A} \frac{d}{\alpha} \left(\left(\frac{1}{\alpha T} \right)^{\frac{\beta-1}{\beta}} + \frac{d}{T} \left(\frac{1}{\alpha T} \right)^{\frac{1}{\beta}} \right).$$

Учитывая, что $4TT_0 \geq (T + 1)(T_0 + 1)$, окончательно приходим к неравенству

$$\begin{aligned} \delta_{T+1} &\leq \frac{2T_0}{T + 1} \delta_1 + \mathbf{A} \frac{d}{\alpha} \left(\left(\frac{1}{\alpha(T + 1)} \right)^{\frac{\beta-1}{\beta}} + \frac{d}{T + 1} \left(\frac{1}{\alpha(T + 1)} \right)^{\frac{1}{\beta}} \right) \leq \\ &\leq \frac{144\bar{L}d\kappa}{\alpha(T + 1)} \delta_1 + \mathbf{A} \frac{d}{\alpha} \left(\left(\frac{1}{\alpha(T + 1)} \right)^{\frac{\beta-1}{\beta}} + \frac{d}{T + 1} \left(\frac{1}{\alpha(T + 1)} \right)^{\frac{1}{\beta}} \right). \end{aligned}$$

Доказательство теоремы 2. Фиксируя $\mathbf{x} \in \Theta$, по лемме 5 имеем

$$f(\mathbf{x}_t) - f(\mathbf{x}) \leq \frac{a_t - \mathbf{E}[a_{t+1}|\mathbf{x}_t]}{2\eta_t} + \frac{\|\nabla f(\mathbf{x}_t) - \mathbf{E}[\mathbf{g}_t|\mathbf{x}_t]\|^2}{\alpha} + \\ + \frac{\eta_t}{2} \mathbf{E}[\|\mathbf{g}_t\|^2|\mathbf{x}_t] - \frac{\alpha a_t}{4},$$

где $a_t = \|\mathbf{x}_t - \mathbf{x}\|^2$. В силу лемм 2, 3 и условия $\max_{\mathbf{x} \in \Theta} \|\nabla f(\mathbf{x})\| \leq G$ получим

$$f(\mathbf{x}_t) - f(\mathbf{x}) \leq \frac{a_t - \mathbf{E}[a_{t+1}|\mathbf{x}_t]}{2\eta_t} + \frac{d}{\alpha} (\kappa_\beta L h_t^{\beta-1})^2 + \\ + \frac{3\eta_t}{4} \kappa d \left(\frac{3}{4} (d\bar{L}^2 h_t^2 + 8G^2) + \frac{\sigma^2}{h_t^2} \right) - \frac{\alpha a_t}{4}.$$

Пусть $b_t = \mathbf{E}[\|\mathbf{x}_t - \mathbf{x}\|^2]$. Подставив сюда $\eta_t = 4/(\alpha(t+1))$ и взяв математическое ожидание, приходим к неравенству

$$\mathbf{E}[f(\mathbf{x}_t) - f(\mathbf{x})] \leq \frac{\alpha}{8} ((t+1)(b_t - b_{t+1}) - 2b_t) + \\ + \frac{d}{\alpha} (\kappa_\beta L h_t^{\beta-1})^2 + \frac{3}{\alpha(t+1)} \kappa d \left(\frac{3}{4} (d\bar{L}^2 h_t^2 + 8G^2) + \frac{\sigma^2}{h_t^2} \right).$$

Суммируя обе части этого неравенства от 1 до T и используя тот факт, что

$$\sum_{t=1}^T t((t+1)(b_t - b_{t+1}) - 2b_t) \leq 0,$$

находим

$$\sum_{t=1}^T t \mathbf{E}[f(\mathbf{x}_t) - f(\mathbf{x})] \leq \\ \leq \frac{d}{\alpha} \sum_{t=1}^T \left(t(\kappa_\beta L h_t^{\beta-1})^2 + \frac{3t}{t+1} \kappa \left(\frac{3}{4} (d\bar{L}^2 h_t^2 + 8G^2) + \frac{\sigma^2}{h_t^2} \right) \right).$$

Подставляя сюда $h_t = \left(\frac{3}{2t} \frac{\kappa \sigma^2}{\kappa_\beta^2 L^2} \right)^{\frac{1}{2\beta}}$, приходим к неравенству

$$\sum_{t=1}^T t \mathbf{E}[f(\mathbf{x}_t) - f(\mathbf{x})] \leq \frac{18G^2 d \kappa T}{\alpha} + \mathbf{A} \frac{d}{\alpha} \sum_{t=1}^T \left(t^{\frac{1}{\beta}} + dt^{-\frac{1}{\beta}} \right) \leq \\ \leq \frac{18G^2 d \kappa T}{\alpha} + \mathbf{A}' \frac{d}{\alpha} \left(1 + dT^{-\frac{2}{\beta}} \right) T^{\frac{\beta+1}{\beta}}.$$

Здесь через \mathbf{A}, \mathbf{A}' обозначаются положительные постоянные, зависящие только от L, \bar{L}, β , и σ^2 . Разделив обе части последнего неравенства на $T(T+1)/2$ и применив неравенство Йенсена, получим результат теоремы.

СПИСОК ЛИТЕРАТУРЫ

1. *Stone C.J.* Additive regression and other nonparametric models // *Annals of Statistics*. 1985. V. 13. P. 689–705.
2. *Hastie T., Tibshirani R.* Generalized additive models // *Statistical Science*. 1986. V. 1. No. 3. P. 297–310.
3. *Wood S.N.* Generalized additive models. Chapman and Hall/CRC, 2017.
4. *Stone C.J.* Optimal rates of convergence for nonparametric estimators // *Annals of Statistics*. 1980. V. 8. P. 1348–1360.
5. *Stone C.J.* Optimal global rates of convergence for nonparametric regression // *Annals of Statistics*. 1982. V. 10. P. 1040–1053.
6. *Ibragimov I.A., Khas'minskiĭ R.Z.* Statistical estimation: Asymptotic theory. Springer, 1981.
7. *Ибрагимов И.А., Хасъминский Р.З.* О границах качества непараметрического оценивания регрессии // *Теория вероятн. и ее примен.* 1982. V. 27. No. 1. P. 81–94.
8. *Поляк Б.Т.* Градиентные методы минимизации функционалов // *Ж. вычисл. матем. и матем. физ.* 1963. V. 3. No. 4. P. 643–653.
9. *Lojasiewicz S.* A topological property of real analytic subsets // *Coll. du CNRS, Les équations aux dérivées partielles*. 1963. V. 117. No. 2. P. 87–89.
10. *Kiefer J., Wolfowitz J.* Stochastic estimation of the maximum of a regression function // *Annals of Mathematical Statistics*. 1952. V. 23. P. 462–466.
11. *Fabian V.* Stochastic approximation of minima with improved asymptotic speed // *Annals of Mathematical Statistics*. 1967. V. 38. No. 1. P. 191–200.
12. *Nemirovsky A.S., Yudin D.B.* Problem complexity and method efficiency in optimization. Wiley & Sons, 1983.
13. *Поляк Б.Т., Цыбаков А.Б.* Оптимальные порядки точности поисковых алгоритмов стохастической оптимизации // *Пробл. передачи информ.* 1990. V. 26. No. 2. P. 45–53.
14. *Jamieson K.G., Nowak R., Recht B.* Query complexity of derivative-free optimization // *Advances in Neural Information Processing Systems*. 2012. V. 26. P. 2672–2680.
15. *Shamir O.* On the complexity of bandit and derivative-free stochastic convex optimization // *Proc. 30th Annual Conference on Learning Theory*. 2013. P. 1–22.
16. *Ghadimi S., Lan G.* Stochastic first- and zeroth-order methods for nonconvex stochastic programming // *SIAM Journal on Optimization*. 2013. V. 23(4). P. 2341–2368.
17. *Nesterov Y., Spokoiny V.* Random gradient-free minimization of convex functions // *Foundations of Computational Mathematics*. 2017. V. 17. No. 2. P. 527–566.
18. *Bach F., Perchet V.* Highly-smooth zero-th order online optimization // *Proc. 29th Annual Conference on Learning Theory*. 2016.
19. *Akhavan A., Pontil M., Tsybakov A.B.* Exploiting higher order smoothness in derivative-free optimization and continuous bandits // *Advances in Neural Information Processing Systems*. 2020. V. 33. P. 9017–9027.
20. *Balasubramanian K., Ghadimi S.* Zeroth-order nonconvex stochastic optimization: Handling constraints, high dimensionality, and saddle points // *Foundations of Computational Mathematics*. 2021. P. 1–42.

21. *Akhavan A., Chzhen E., Pontil M., Tsybakov A.B.* Gradient-free optimization of highly smooth functions: improved analysis and a new algorithm // *Journal of Machine Learning Research*. 2024. V. 25. No. 370. P. 1–50.
22. *Гасников А.В., Лагуновская А.А., Усманова И.Н., Федоренко Ф.А.* Безградиентные прокс-методы с неточным оракулом для негладких задач выпуклой стохастической оптимизации на симплексе // *Автомат. и телемех.* 2016. No. 10. P. 57–77.
23. *Novitskii V., Gasnikov A.* Improved exploitation of higher order smoothness in derivative-free optimization // *Optimization Letters*. 2022. V. 16. P. 2059–2071.
24. *Akhavan A., Pontil M., Tsybakov A.B.* Distributed zero-order optimization under adversarial noise // *Advances in Neural Information Processing Systems*. 2021. V. 34. P. 10209–10220.
25. *Gasnikov A.V., Lobanov A.V., Stonyakin F.S.* Highly smooth zeroth-order methods for solving optimization problems under the PL condition // *Computational Mathematics and Mathematical Physics*. 2024. V. 64. P. 739–770.
26. *Yu Q., Wang Y., Huang B., et al.* Stochastic zeroth-order optimization under strong convexity and Lipschitz Hessian: Minimax sample complexity // *Advances in Neural Information Processing Systems*. 2024. V. 37.
27. *Fokkema H., van der Hoeven D., Lattimore T., Mayo J.J.* Online Newton method for bandit convex optimisation // *arXiv preprint arXiv:2406.06506*. 2024.
28. *Karimi H., Nutini J., Schmidt M.* Linear convergence of gradient and proximalgradient methods under the Polyak–Lojasiewicz condition // *Machine Learning and Knowledge Discovery in Databases*. 2016. P. 795–811.
29. *Tsybakov A.B.* Introduction to nonparametric estimation. New York: Springer, 2009.
30. *Граничин О.Н.* Процедура стохастической аппроксимации с возмущением на входе // *Автомат. и телемех.* 1992. No. 2. P. 97–104.
31. *Polyak B.T., Tsybakov A.B.* On stochastic approximation with arbitrary noise (the KW-case) // *Advances in Soviet Mathematics*, vol. 12. 1992. P. 107–113.

Статья представлена к публикации членом редколлегии П.С. Щербаковым.

Поступила в редакцию 03.03.2025

После доработки 20.05.2025

Принята к публикации 27.06.2025

© 2025 г. Р.С. БИРЮКОВ, канд. физ.-мат. наук (biryukovrs@gmail.com)
(Нижегородский государственный университет им. Н.И. Лобачевского),
М.М. КОГАН, д-р физ.-мат. наук (mkogan@nngasu.ru)
(Научно-технологический университет “Сириус”, Сочи)

СИНТЕЗ РОБАСТНЫХ ЗАКОНОВ УПРАВЛЕНИЯ ПО ЗАШУМЛЕННЫМ ТЕКУЩИМ ДАННЫМ¹

В статье излагается новый подход к построению робастных законов управления неопределенной динамической системой с обобщенной H_∞ -нормой в качестве критерия. Параметры закона управления настраиваются в процессе реального функционирования системы по мере поступления текущих измерений состояния и управления, получаемых с ошибкой. Результаты численного моделирования активного гашения колебаний сооружений при сейсмических воздействиях иллюстрируют эффективность рассматриваемого подхода.

Ключевые слова: текущие данные, робастное управление, обобщенная H_∞ -норма, двойственные системы, линейные матричные неравенства.

DOI: 10.31857/S0005231025090035, EDN: VMTKUI

1. Введение

В статье разрабатывается новый подход к построению оптимальных законов управления линейными динамическими объектами с неизвестными математическими моделями при неизмеряемых возмущениях и неопределенных начальных условиях, использующий априорную и текущую информацию в online режиме. Как известно, один из самых распространенных методов построения робастного управления основан на минимаксном подходе: синтезируется закон управления, который обеспечивает наилучшую оценку оптимизируемого функционала на всем множестве значений неопределенных параметров, определяемом на основе априорной информации (см., например, [1, 2]). При всех достоинствах такой подход имеет существенный недостаток: гарантируемая верхняя оценка функционала из-за большой величины множества неопределенности оказывается достаточно грубой, а при отсутствии точной априорной информации множество неопределенности может оказаться таким большим, что робастного закона управления не существует.

В последнее время в теории управления наблюдается значительный интерес к развитию методов, использующих данные, получаемые предварительно на стадии эксперимента. В англоязычной литературе получаемое таким образом управление называется “data-driven control” [3, 4]. Первоначально в осно-

¹ Работа выполнена при финансовой поддержке Российского научного фонда (проект № 24-11-20023).

ву этих методов был положен фундаментальный результат в [5], показывающий, что для полной характеристики линейной стационарной динамической системы без возмущений может быть использована единственная траектория, полученная при выполнении так называемого условия исчезающего возбуждения. В [6] было показано, что параметризация замкнутой системы управления на основе экспериментальных данных позволяет синтезировать законы управления, используя линейные матричные неравенства. В [7] было установлено, что для построения законов управления по экспериментальным данным достаточно выполнения менее ограничительного условия информативности данных по отношению к изучаемому свойству. Затем эти методы были распространены для систем с возмущениями [8–11], в [12] – для линейных нестационарных систем, а в [13] осуществлен синтез стабилизирующего управления на основе измерений, полученных с ошибкой.

Результаты математического моделирования, приведенные в этих работах, показывают, что уже при сравнительно небольших амплитудах возмущений гарантированная оценка показателя качества управления оказывается достаточно завышенной. Это происходит по причине того, что с ростом амплитуды помехи множество объектов, согласованных с экспериментальными данными, значительно расширяется. Кроме того, к экспериментальным данным в большинстве из этих работ предъявляются определенные требования: матрица, составленная из измерений состояний и управлений по траектории системы, должна быть максимального строчного ранга. Для выполнения этого рангового условия требуется, чтобы входные сигналы в экспериментах обеспечили исчезающее возбуждение в системе, необходимое для идентифицируемости неизвестных параметров. И хотя формально для обеспечения информативности данных не требуется исчезающее возбуждение в системе, множество неопределенности может стать неограниченным даже при небольших амплитудах помех, если для экспериментальных данных не выполняется ранговое условие.

Указанные обстоятельства наводят на мысль о синтезе управления, использующем как априорные, так и экспериментальные данные. В [14–20] параметры обратной связи по состоянию для линейных стационарных систем на бесконечном горизонте, нестационарных систем на конечном горизонте, а также некоторых классов нелинейных систем находятся при совместном использовании априорных и экспериментальных данных путем решения минимаксной задачи на множестве неизвестных параметров, согласованных с априорной информацией и экспериментальными данными. Такие законы управления, получаемые на основе решения линейных матричных неравенств, обеспечивают заведомо меньшее значение гарантированной оценки функционала, чем классическое робастное управление и чем закон управления, построенный только по экспериментальным данным.

Вместе с тем, экспериментальный этап, который предшествует синтезу закона управления и на котором входные сигналы выбираются исходя из требований информативности получаемых данных, не всегда возможно реализо-

вать в силу, например, неустойчивости объекта. Кроме того, имеются объекты управления, параметры которых зависят от состояния окружающей среды, например, от температуры или давления. В этих случаях управление, синтезированное по данным, полученным на предварительном этапе, может оказаться не совсем подходящим при других условиях. Вместо этого было бы желательно на каждом шаге реального функционирования системы по мере получения новых текущих данных перестраивать параметры закона управления на основе решения минимаксной задачи для множества моделей, согласованных с имеющейся к данному моменту времени информацией. Поскольку это множество на каждом шаге сокращается или, в крайнем случае, не возрастает, то определяемые при этом обратные связи будут обеспечивать монотонно убывающие или невозрастающие гарантируемые значения функционала для неизвестного объекта. Процесс настройки параметров прекращается, когда гарантируемое значение функционала становится меньше желаемого. Именно такой подход разработан в настоящей работе. При этом не требуется обеспечивать выполнение условий информативности и снимается существенный вопрос о количестве измерений, которое необходимо сделать. Иллюстративный пример активного гашения колебаний сооружений при сейсмических воздействиях демонстрирует эффективность синтезируемого закона управления.

2. Формулировка задачи

Рассмотрим неопределенную систему вида

$$(2.1) \quad \begin{aligned} x_{t+1} &= A_{real}x_t + B_{real}u_t + B_w w_t, & x(0) &= x_0, \\ z_t &= Cx_t + Du_t, & t &= 0, 1, \dots, \end{aligned}$$

где $x_t \in \mathbb{R}^{n_x}$ – состояние, $u_t \in \mathbb{R}^{n_u}$ – управление, $w_t \in \mathbb{R}^{n_w}$ – возмущение, $z_t \in \mathbb{R}^{n_z}$ – целевой (управляемый) выход. Предполагается, что начальное состояние x_0 может быть произвольным, возмущение $\{w_t, t \geq 0\}$ имеет ограниченную l_2 -норму, т.е. $\|w\| = (\sum_{t=0}^{\infty} |w_t|^2)^{1/2} < \infty$, где $|\cdot|$ обозначает евклидову норму соответствующего вектора, и состояние доступно для управления в контуре обратной связи. Системные матрицы, собранные в матрицу $\Omega_{real} = (A_{real} \ B_{real})$, неизвестны, но предполагается, что пара (A_{real}, B_{real}) стабилизируема, а матрица B_w , которая, в частности, может быть единичной, известна.

Так как системные матрицы неизвестны, то стандартные методы построения оптимальных законов управления не применимы. В связи с этим для синтеза закона управления будут привлечены априорные данные, как это принято в робастном управлении, и текущие данные, что характерно для адаптивного управления. А именно, допустим, что возмущение w_t в системе имеет ограниченную евклидову норму, а состояние и управление измеряются с некоторыми ограниченными помехами так, что

$$(2.2) \quad \hat{x}_t = x_t + \xi_t, \quad \hat{u}_t = u_t + \eta_t,$$

где \hat{x}_t и \hat{u}_t – измерения состояния и управления в момент времени t , и

$$(2.3) \quad |\xi_t| \leq \varepsilon_\xi, \quad |\eta_t| \leq \varepsilon_\eta, \quad |w_t| \leq \varepsilon_w \quad \forall t \geq 0.$$

Заметим, что измерения состояния \hat{x}_t и управления \hat{u}_t , которые будут входить в текущие данные, содержат указанные ограниченные ошибки, а управлению доступно состояние x_t .

В общем плане требуется синтезировать законы управления вида линейной обратной связи по состоянию с настраиваемыми параметрами на основе априорной и текущей информации в ходе реального функционирования системы так, чтобы через некоторый промежуток времени прийти к стационарным значениям, при которых минимизируется влияние начальных условий и внешнего возмущения на целевой выход замкнутой стационарной системы.

3. Характеризация нормы в терминах двойственной системы

Прежде, чем перейти к решению задачи, напомним основные факты относительно обобщенной H_∞ -нормы от входа w_t к выходу z_t заданной замкнутой системы при управлении $u_t = \Theta x_t$, описываемой уравнениями

$$(3.1) \quad \begin{aligned} x_{t+1} &= (A + B\Theta)x_t + B_w w_t, \\ z_t &= (C + D\Theta)x_t. \end{aligned}$$

Обобщенная H_∞ -норма характеризует влияние начальных условий и внешнего возмущения на l_2 -норму целевого выхода, т.е. на переходной процесс, и определяется как

$$(3.2) \quad \|H\|_{g\infty}(\Omega, \Theta) = \sup_{\{x_0, w_t \in l_2\}} \frac{\|z\|}{(x_0^T R^{-1} x_0 + \|w\|^2)^{1/2}},$$

где $\Omega = \begin{pmatrix} A & B \end{pmatrix}$ и $R = R^T \succ 0$ – весовая матрица, отражающая меру сравнительной важности для целевого выхода неопределенности в начальных условиях и внешнем возмущении, $S \succ T$ означает, что матрица $S - T$ положительно определенная, а $S \succeq T$ – что матрица $S - T$ неотрицательно полуопределенная. В этом определении предполагается, что знаменатель правой части в (3.2) не обращается в ноль и

$$H(s) = (C + D\Theta)[sI - (A + B\Theta)]^{-1} B_w$$

передаточная матрица замкнутой системы от возмущения к целевому выходу. В частности, обобщенная H_∞ -норма при нулевых начальных условиях в системе совпадает со стандартной H_∞ -нормой, а при отсутствии внешнего возмущения она характеризует максимальное значение квадратичного функционала целевого выхода при начальном состоянии, принадлежащем эллипсоиду $x^T R^{-1} x \leq 1$. Имеют место следующие утверждения относительно вычисления этих норм, которые будут применяться далее.

Лемма 3.1 [21]. *Обобщенная H_∞ -норма системы (3.1) удовлетворяет условию $\|H\|_{g_\infty} < \gamma$ тогда и только тогда, когда существует положительно определенная квадратичная форма $V(x) = x^T Y x$, $Y \prec \gamma^2 R^{-1}$, для которой при всех ненулевых x_t и w_t выполняется неравенство*

$$(3.3) \quad \Delta V(x_t) + |z_t|^2 - \gamma^2 |w_t|^2 < 0,$$

где $\Delta V(x_t) = V(x_{t+1}) - V(x_t)$ обозначает приращение функции $V(x)$ по траектории соответствующей системы.

Лемма 3.2 [17]. *Обобщенная H_∞ -норма системы (3.1) удовлетворяет условию $\|H\|_{g_\infty} < \gamma$ тогда и только тогда, когда существует положительно определенная квадратичная форма $V_d(x) = x^T P x$, $P \succ R$, для приращения которой по траектории двойственной системы*

$$(3.4) \quad \begin{aligned} x_{t+1}^{(d)} &= (A + B\Theta)^T x_t^{(d)} + (C + D\Theta)^T w_t^{(d)}, \\ z_t^{(d)} &= B_w^T x_t^{(d)} \end{aligned}$$

при всех ненулевых $x_t^{(d)}$ и $w_t^{(d)}$ выполняется неравенство

$$(3.5) \quad \Delta V_d(x_t^{(d)}) + |z_t^{(d)}|^2 - \gamma^2 |w_t^{(d)}|^2 < 0.$$

Замечание 3.1. Для обобщенной H_∞ -нормы матрицы квадратичных форм $V(x) = x^T Y x$ и $V_d(x^{(d)}) = x^{(d)T} P x^{(d)}$ прямой и двойственной систем связаны соотношением $P = \gamma^2 Y^{-1}$.

Замечание 3.2. При выполнении условий каждой из перечисленных выше лемм соответствующие системы являются асимптотически устойчивыми при отсутствии возмущений.

4. Априорные и текущие данные

Пусть, как это принято в традиционных методах синтеза робастного управления, имеется априорная информация о том, что реальная матрица $\Omega_{real} = (A_{real} \ B_{real})$ находится строго внутри области, определяемой неравенством

$$(4.1) \quad (\Omega - \Omega_*)(\Omega - \Omega_*)^T \preceq \rho^2 I, \quad \Omega_* = (A_* \ B_*),$$

где Ω_* и ρ характеризуют номинальный объект и размер области неопределенности. Запишем неравенство (4.1) в виде

$$(4.2) \quad F_a(\Omega) := (\Omega \ I) \Psi_a (\Omega \ I)^T \preceq 0,$$

где

$$(4.3) \quad \Psi_a = \left(\begin{array}{c|c} I & \star \\ \hline -\Omega_* & \Omega_* \Omega_*^T - \rho^2 I \end{array} \right),$$

★ заменяет соответствующий блок симметрической матрицы. Назовем $\Delta^{(a)} := \{\Omega \in \mathbb{R}^{n_x \times (n_x + n_u)} : F_a(\Omega) \leq 0\}$ множеством матриц, согласованных с априорной информацией. Согласно предположению $\Omega_{real} \in \Delta^{(a)}$.

Текущие данные, полученные в момент времени $t + 1$, в силу (2.1) и (2.2) удовлетворяют уравнению

$$(4.4) \quad \hat{x}_{t+1} = \Omega_{real} \hat{\varphi}_t + (I \ \Omega_{real} \ B_w) \begin{pmatrix} \xi_{t+1} \\ \zeta_t \\ w_t \end{pmatrix},$$

в котором

$$\hat{\varphi}_t := \begin{pmatrix} \hat{x}_t \\ \hat{u}_t \end{pmatrix}, \quad \zeta_t := - \begin{pmatrix} \xi_t \\ \eta_t \end{pmatrix}.$$

Множество матриц Ω , которые удовлетворяют уравнению (4.4) при некотором возмущении и помехах, для которых выполняются ограничения (2.3), назовем множеством матриц, согласованных с данными \hat{x}_t, \hat{x}_{t+1} и \hat{u}_t , и обозначим $\Delta_t^{(p)}$. Ясно, что $\Omega_{real} \in \Delta_t^{(p)}$ и что множество $\Delta_t^{(p)}$ соответствует всем системам, которые могли бы генерировать эти данные. Данное множество характеризуется следующим образом.

Лемма 4.1. Матрицы $\Omega \in \Delta_t^{(p)}$, согласованные с данными \hat{x}_t, \hat{x}_{t+1} и \hat{u}_t , удовлетворяют неравенству

$$(4.5) \quad (\Omega \ I) \begin{pmatrix} \hat{\varphi}_t \hat{\varphi}_t^T - d^2 I & \star \\ -\hat{x}_{t+1} \hat{\varphi}_t^T & \hat{x}_{t+1} \hat{x}_{t+1}^T - (I + B_w B_w^T) d^2 \end{pmatrix} (\Omega \ I)^T \leq 0,$$

где $d^2 = 2\varepsilon_\xi^2 + \varepsilon_\eta^2 + \varepsilon_w^2$.

Доказательство этого и последующих утверждений приведены в Приложении. Пусть к фиксированному моменту времени k получены измерения состояний $\hat{x}_0, \dots, \hat{x}_k$ и выбранных ранее управлений $\hat{u}_0, \dots, \hat{u}_{k-1}$ системы (2.1) при некоторых неизвестных возмущениях w_0, \dots, w_{k-1} . Матрицы Ω , согласованные со всеми данными $\hat{x}_0, \dots, \hat{x}_k, \hat{u}_0, \dots, \hat{u}_{k-1}$, полученными к моменту времени $k \geq 1$, принадлежат множеству $\Delta_{[0, k-1]}^{(p)} = \bigcap_{t=0}^{k-1} \Delta_t^{(p)}$ и определяются неравенствами (4.5) для $t = 0, \dots, k-1$. Обозначим $\Delta_0 = \Delta^{(a)}$ и $\Delta_k = \Delta^{(a)} \cap \Delta_{[0, k-1]}^{(p)}$, $k \geq 1$ множество матриц Ω , согласованных с априорной информацией и полученными к моменту времени k данными $\hat{x}_0, \dots, \hat{x}_k, \hat{u}_0, \dots, \hat{u}_{k-1}$ для $k \geq 1$. Заметим, что $\Delta_0 \supseteq \Delta_1 \supseteq \dots$ и что $\Omega_{real} \in \Delta_k$ для любого $k \geq 0$.

Однако так как проверка того, что пересечение эллипсоидов включено в заданный эллипсоид, не может быть представлена как задача полуопределенного программирования (см. [22]), заменим множество Δ_k его внешней эллипсоидальной аппроксимацией. Это позволит находить закон управления, минимизирующий верхнюю границу рассматриваемой нормы замкнутой системы для всех моделей, согласованных с имеющимися данными.

5. Внешняя аппроксимация множества Δ_k

Суммируя неравенства (4.2) и (4.5) с некоторыми неотрицательными множителями $\mu \geq 0$, $\tau_0 \geq 0, \dots, \tau_{k-1} \geq 0$, получим, что все матрицы Ω , согласованные с априорными и текущими данными к моменту времени k , удовлетворяют неравенству

$$(5.1) \quad \widehat{F}_{\mu, \tau_0^{k-1}}(\Omega) := (\Omega \quad I)\Psi(\mu, \tau_0^{k-1})(\Omega \quad I)^T \preceq 0, \quad k \geq 1,$$

где

$$(5.2) \quad \Psi(\mu, \tau_0^{k-1}) = \begin{pmatrix} \mu I + \sum_{t=0}^{k-1} \tau_t (\widehat{\varphi}_t \widehat{\varphi}_t^T - d^2 I) & \star \\ -\mu \Omega_* - \sum_{t=0}^{k-1} \tau_t \widehat{x}_{t+1} \widehat{\varphi}_t^T & \mu (\Omega_* \Omega_*^T - \rho^2 I) + \sum_{t=0}^{k-1} \tau_t \widehat{x}_{t+1} \widehat{x}_{t+1}^T - M_k \end{pmatrix},$$

$$\tau_0^{k-1} = (\tau_0, \dots, \tau_{k-1}), \quad M_k := \sum_{t=0}^{k-1} \tau_t (I + B_w B_w^T) d^2.$$

Определим $\widehat{\Delta}_k(\mu, \tau_0^{k-1}) = \left\{ \Omega \in \mathbb{R}^{n_x \times (n_x + n_u)} : \widehat{F}_{\mu, \tau_0^{k-1}}(\Omega) \preceq 0 \right\}$ как множество матриц Ω , удовлетворяющих неравенству (5.1) при фиксированных $\mu \geq 0$, $\tau_0^{k-1} \geq 0$. Таким образом, $\Delta_k \subseteq \widehat{\Delta}_k(\mu, \tau_0^{k-1})$ для любых $\mu \geq 0$, $\tau_0^{k-1} \geq 0$ и $\Omega_{real} \in \widehat{\Delta}_k(\mu, \tau_0^{k-1})$. Это множество при определенных значениях $\mu \geq 0$ и $\tau_0^{k-1} \geq 0$ будет далее служить внешней аппроксимацией множества Δ_k .

Выясним, при каких условиях множество $\widehat{\Delta}_k(\mu, \tau_0^{k-1})$ является ограниченным. Обозначим соответствующие блоки матрицы (5.2) через $\Psi_{ij}(\mu, \tau_0^{k-1})$, $i, j = 1, 2$, аргументы которых будут иногда опускаться.

Лемма 5.1. Если

$$(5.3) \quad \Psi_{11} = \mu I + \sum_{t=0}^{k-1} \tau_t (\widehat{\varphi}_t \widehat{\varphi}_t^T - d^2 I) \succ 0,$$

то множество $\widehat{\Delta}_k(\mu, \tau_0^{k-1})$ является ограниченным матричным эллипсоидом, определяемым неравенством

$$(5.4) \quad (\Omega + \Psi_{12}^T \Psi_{11}^{-1}) \Psi_{11} (\Omega + \Psi_{12}^T \Psi_{11}^{-1})^T \preceq \Gamma_k,$$

где $\Gamma_k = \Psi_{12}^T \Psi_{11}^{-1} \Psi_{12} - \Psi_{22} \succeq 0$.

Замечание 5.1. Условие (5.3) означает, что “энергия” измеряемого сигнала на всем интервале с учетом априорной информации должна превышать суммарную “энергию” возмущения и помехи в измерениях. Если это условие не выполнено, то множество $\widehat{\Delta}_k(\mu, \tau_0^{k-1})$ не ограничено (см. лемму 2.1 в [19]).

6. Робастный закон управления на основе текущих данных

Для построения робастного управления с матрицей в обратной связи, настраиваемой по текущим данным, требуется предварительно решить следующую вспомогательную задачу: определить матрицу Θ , при которой будет минимальна верхняя граница обобщенной H_∞ -нормы неопределенной замкнутой стационарной устойчивой системы вида (3.1) для всех $\Omega \in \hat{\Delta}_k(\mu, \tau_0^{k-1})$ при некоторых $\mu \geq 0$ и $\tau_0^{k-1} \geq 0$. Для каждой модели с матрицей Ω , вообще говоря, может существовать своя функция Ляпунова, которая согласно лемме 3.1 характеризует границу рассматриваемой нормы. Однако далее будут получены условия существования единой функции Ляпунова для всех $\Omega \in \hat{\Delta}_k(\mu, \tau_0^{k-1})$. Это вносит некоторый консерватизм, но позволяет получить необходимые и достаточные условия существования такой функции и найти соответствующие обратные связи. Заметим, что аналогичный подход применялся в робастном управлении при определении так называемой квадратичной устойчивости и стабилизации в условиях неопределенности (см., например, [23]).

Определение 6.1. Закон управления $u_t = \Theta x_t$ назовем гарантирующим обобщенным H_∞ -управлением с уровнем γ для неопределенной системы (3.1) при априорных и текущих данных к моменту времени k , если существует функция $V_k(x) = x^T Y_k x$, $0 < Y_k < \gamma^2 R^{-1}$ и такие $\mu \geq 0$ и $\tau_0^{k-1} \geq 0$, что неравенство (3.3) выполняется при всех $\Omega \in \hat{\Delta}_k(\mu, \tau_0^{k-1})$.

Согласно лемме 3.1 и замечанию 3.2 замкнутая система с гарантирующим обобщенным H_∞ -управлением с уровнем γ является асимптотически устойчивой и $\|H\|_{g\infty} < \gamma$ для всех моделей, согласованных с априорными и текущими данными к моменту времени k .

Теорема 6.1. Закон управления $u_t = \Theta x_t$ является гарантирующим обобщенным H_∞ -управлением с уровнем γ для неопределенной системы (3.1) при текущих данных к моменту времени k тогда и только тогда, когда $\Theta = Q_k P_k^{-1}$ и следующее линейное матричное неравенство разрешимо относительно $P_k = P_k^T \succ R$, Q_k , $\mu \geq 0$ и $\tau_0^{k-1} \geq 0$:

$$(6.1) \quad \begin{pmatrix} -P_k & * & * & * \\ \begin{pmatrix} P_k \\ Q_k \end{pmatrix} & -\Psi_{11} & * & * \\ 0 & -\Psi_{21} & -\Psi_{22} - P_k + B_w B_w^T & * \\ (C \ D) \begin{pmatrix} P_k \\ Q_k \end{pmatrix} & 0 & 0 & -\gamma^2 I \end{pmatrix} < 0,$$

где Ψ_{ij} , $i, j = 1, 2$ обозначают соответствующие блоки матрицы $\Psi(\mu, \tau_0^{k-1})$, заданной в (5.2).

Замечание 6.1. Если в неравенстве (6.1) положить $\mu = 0$, то априорная информация не будет включена в синтез управления, а если положить $\tau_0^{k-1} = 0$, то не используются текущие данные.

Замечание 6.2. Если матрица $\Psi(\mu, \tau_0^{k-1})$ состоит из данных, полученных при эксперименте на интервале $[0, k-1]$ с разомкнутой неопределенной системой при управлениях, определяемых экспериментатором, то теорема 6.1 остается верна. Заметим, что в этом случае при $\mu = 0$ и $\tau_0 = \dots = \tau_{k-1} = 1$ соответствующее множество $\widehat{\Delta}_k(\mu, \tau_0^{k-1})$ совпадает с множеством матриц, согласованных с полученными данными, как это определено в [6, 7, 16]. Это означает, что рассмотренный здесь подход приводит к результатам, менее консервативным, чем в указанных работах.

Назовем минимальное значение $\gamma > 0$, при котором разрешимо неравенство (6.1), гарантированной обобщенной H_∞ -нормой неопределенной системы по данным к моменту времени k и обозначим ее γ_k^* , а соответствующую матрицу обратной связи обозначим как Θ_k . Так как неравенство (6.1) для момента времени k при $\tau_{k-1} = 0$ обращается в неравенство (6.1) для момента времени $k-1$, то гарантированные обобщенные H_∞ -нормы по мере поступления текущей информации образуют невозрастающую последовательность, т.е.

$$(6.2) \quad \gamma_0^* \geq \gamma_1^* \geq \dots$$

Теперь перейдем непосредственно к синтезу робастного управления. В качестве матрицы параметров управления в начальный момент времени возьмем матрицу, получаемую на основе только априорной информации, т.е. $\Theta_0 = Q_0 P_0^{-1}$, где P_0 и Q_0 – решение неравенства (6.1) при $\tau_0^{k-1} = 0$ и $\gamma = \gamma_0^*$. При $t = 1, \dots, N-1$ матрицы Θ_t вычисляются как в теореме 6.1 для минимального $\gamma = \gamma_t^*$. При $t \geq N$ положим $\Theta_t = \Theta_{N-1}$.

Теорема 6.2. Для неопределенной системы (2.1) при законе управления $u_t = \Theta_t x_t$, где Θ_t при $t = 0, \dots, N-1$ вычисляются как в теореме 6.1 при минимальном $\gamma > 0$ и $\Theta_t = \Theta_{N-1}$ при $t \geq N$, после настройки параметров выполняется

$$(6.3) \quad \sum_{t=N-1}^{\infty} |z_t|^2 < \gamma_{N-1}^{*2} \left(x_{N-1}^T R^{-1} x_{N-1} + \sum_{t=N-1}^{\infty} |w_t|^2 \right) \quad \forall w_t.$$

Как следует из (6.2) и (6.3), при синтезе закона управления на основе текущих данных требуется разрешить компромисс между качеством замкнутой системы, получаемой после окончания настройки, и длительностью интервала настройки, т.е. затратами: чем больше N , тем меньше γ_{N-1}^* и наоборот.

В том случае, когда радиус неопределенности достаточно большой, т.е. априорная информация грубая, неравенство (6.1) при $\tau_0^{k-1} = 0$ может не иметь решений. Это означает, что не существует единого регулятора, обеспечивающего устойчивость для всех моделей из начального множества Δ_0 . В этом случае в качестве начального управления выбирается случайный вектор. Если на следующем шаге по-прежнему не существует единого регулятора для нового множества моделей $\widehat{\Delta}_1(\mu, \tau_0) \subseteq \Delta_0$, то также выбирается случайное управление, и этот процесс продолжается до момента времени t_* , пока

у неравенства (6.1) не появится решение. Гарантированная оценка качества управления, данная в теореме 6.2, в этом случае остается в силе, если за начальный момент времени принять $t = t_*$.

7. Иллюстративный пример

Рассмотрим модель колебаний трехэтажного здания при сейсмическом воздействии на фундамент. На рис. 1 изображена общая схема указанного процесса. Этажи здания представляются материальными точками, последовательно связанными между собой и с фундаментом линейными упругими и диссипативными элементами. Колебания, порождаемые сейсмическим воздействием (движением фундамента), происходят в горизонтальной плоскости. Будем считать, что здание является однородным, т.е. массы материальных точек, коэффициенты упругости и демпфирования упругих и диссипативных элементов одинаковые. При указанном предположении динамика рассматриваемой конструкции (в безразмерных переменных и параметрах) описывается системой дифференциальных уравнений

$$(7.1) \quad \begin{aligned} \ddot{x}_1 &= -\beta(2\dot{x}_1 - \dot{x}_2) - (2x_1 - x_2) + w, \\ \ddot{x}_2 &= -\beta(2\dot{x}_2 - \dot{x}_1 - \dot{x}_3) - (2x_2 - x_1 - x_3) + u + w, \\ \ddot{x}_3 &= -\beta(\dot{x}_3 - \dot{x}_2) - (x_3 - x_2) + w, \\ z &= -x_1 - \beta\dot{x}_1 + \alpha u, \end{aligned}$$

где x_i – координаты материальных точек относительно подвижного фундамента, u – демпфирующая сила, порожденная установленной в здании системой активной виброзащиты, w – с точностью до знака ускорение фундамента,

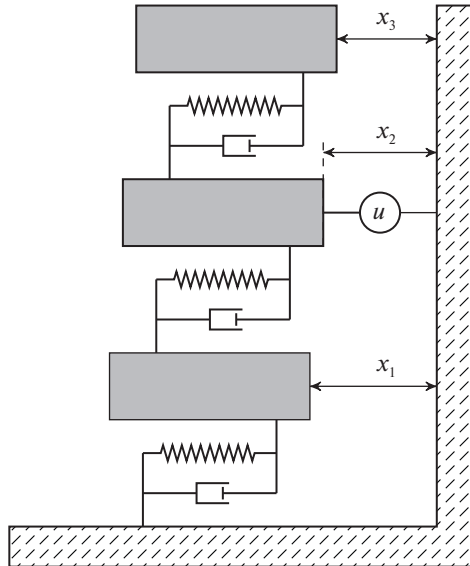


Рис. 1. Схематическое изображение здания как многомассовой упругой системы.

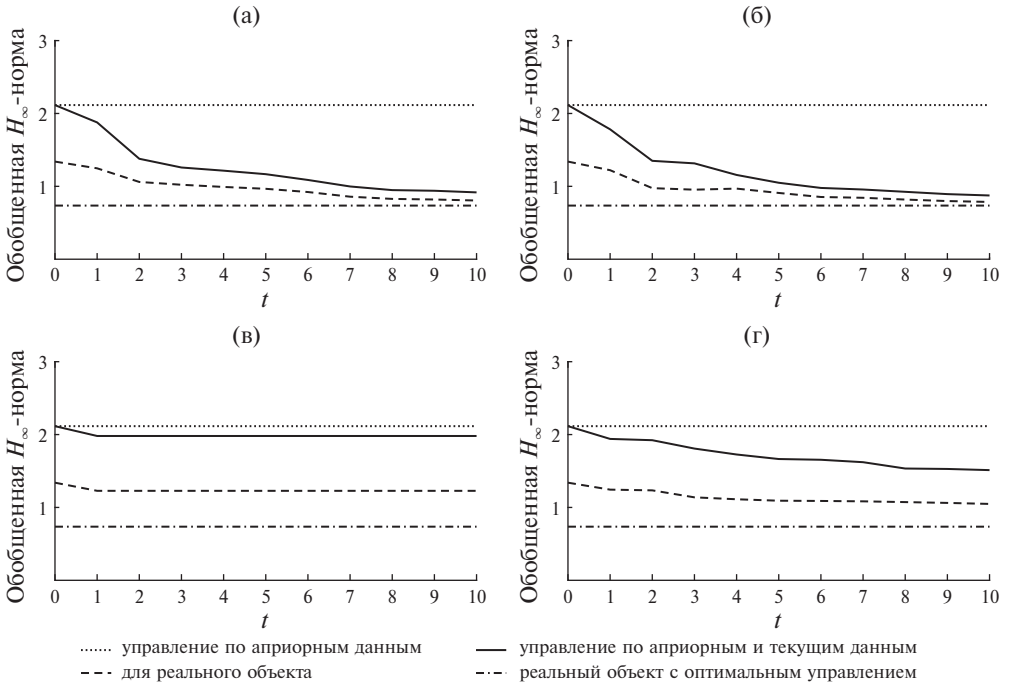


Рис. 2. Эволюции гарантированных обобщенных H_∞ -норм при использовании различной информации с $\varepsilon_\xi = 0,01$ в следующих случаях: а - $\varepsilon_w = 0,01$, $\varepsilon_0 = 2$, б - $\varepsilon_w = 0,01$, $\varepsilon_0 = 5$, в - $\varepsilon_w = 0,1$, $\varepsilon_0 = 2$, г - $\varepsilon_w = 0,1$, $\varepsilon_0 = 5$.

z – целевой скалярный выход, определяющий максимальную силу, противодействующую смещению упругого объекта относительно основания, β – параметр демпфирования. Задача виброизоляции состоит в поиске управления, минимизирующего в некотором смысле z .

Перейдем от непрерывной системы к дискретной с шагом дискретизации $h = 0,5$. Зададим следующие числовые значения параметров: $\alpha = 0,1$, $\beta = 0,1$, $R = 0,01I$. Центр матричной сферы $\Delta^{(a)}$ состоит из матриц дискретной модели системы (7.1) при $\beta = 0,2$, и неравенство (4.1) выполняется при $\rho = 0,07$. Помеха в измерении состояния, возмущение и начальное состояние выбирались случайно на сферах радиусов ε_ξ , ε_w и ε_0 соответственно, помеха при измерении управления отсутствовала, т.е. $\varepsilon_\eta = 0$. Всего неопределенная система (7.1) в канонической форме содержит 49 неизвестных параметров.

На рис. 2 для $\varepsilon_\xi = 0,01$ показаны эволюции гарантированной обобщенной H_∞ -нормы γ_t^* неопределенной системы (7.1) при законе управления $u_t = \Theta_t x_t$ на основе априорных и текущих данных (сплошная кривая); обобщенная H_∞ -норма $\|H\|_{g_\infty}(\Omega_{real}, \Theta)$ для замкнутой системы, состоящей из реального объекта и стационарной обратной связи с $\Theta = \Theta_t$ (пунктирная кривая); гарантированная обобщенная H_∞ -норма неопределенной системы (7.1) с робастным управлением $u_t = \Theta_0 x_t$, полученным на основе только априорных данных (штриховая линия); оптимальная H_∞ -норма при оптимальном управ-

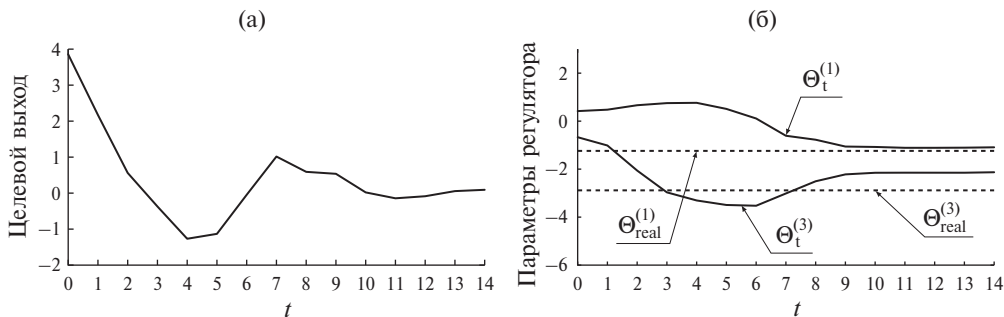


Рис. 3. Эволюция *a* – целевого выхода и *б* – двух настраиваемых параметров для реальной системы при $u_t = \Theta_t x_t$.



Рис. 4. Эволюция гарантированной обобщенной H_∞ -нормы для $\varepsilon_\xi = 0,01$ и $\varepsilon_w = 0,1$ в случае, когда не существует единого робастного регулятора по априорной информации, при $a - \varepsilon_0 = 3$ и $b - \varepsilon_0 = 5$.

лении для реальной замкнутой системы (штрих-пунктирная линия). Как видно из сравнения случаев *a* и *б* или *в* и *г* на этом рисунке, увеличение начального состояния ведет к “увеличению” блока Ψ_{11} и соответственно к уменьшению матричного эллипсоида $\hat{\Delta}_k(\mu, \tau_0^{k-1})$, что в конечном итоге приводит к ускорению сходимости обобщенной H_∞ -нормы к минимальному значению. Сравнение случаев *a* и *б* со случаями *в* и *г* показывает, что с увеличением амплитуды возмущения скорость сходимости показателя замедляется вплоть до возможной остановки. Это связано с тем, что увеличивается множество моделей, согласованных с полученными данными. Рисунок 3 иллюстрирует типичные траектории целевого выхода и двух из шести параметров управления на основе априорных и текущих данных для реальной системы.

Рисунок 4 демонстрирует, что даже когда априорная информация при $\rho = 0,2$ не точная и гарантированная обобщенная H_∞ -норма неопределенной системы (7.1) с робастным управлением $u_t = \Theta_0 x_t$, полученным на основе только априорных данных, не может быть определена, к восьмому шагу при

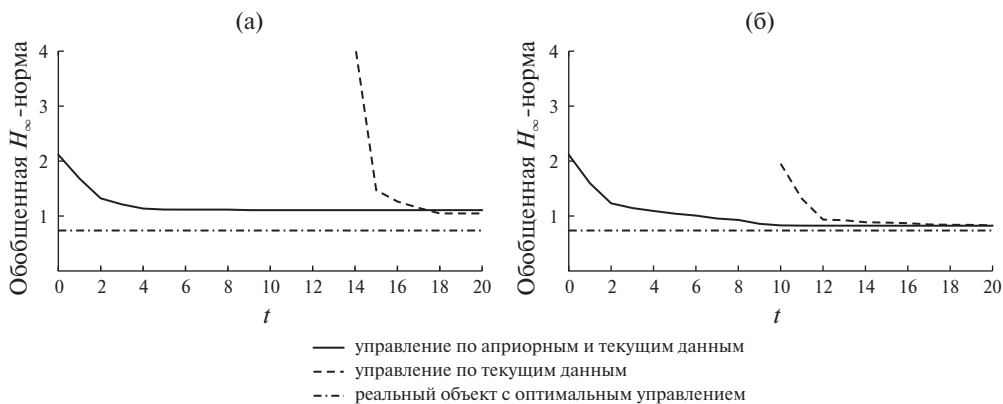


Рис. 5. Эволюции гарантированных обобщенных H_∞ -норм при использовании совместно априорной и текущей информации и только текущей информации для $\varepsilon_\xi = 0,01$, $\varepsilon_w = 0,1$ при $a - \varepsilon_0 = 3$ и $b - \varepsilon_0 = 5$.

$\varepsilon_0 = 3$ и к десятому шагу при $\varepsilon_0 = 5$ текущей информации становится достаточно для того, чтобы неравенство (6.1) стало разрешимым и регулятор быстро настроился.

Рисунок 5, на котором для сравнения показаны эволюции гарантированных обобщенных H_∞ -норм при использовании совместно априорных и текущих данных (сплошная кривая) и только текущих данных (пунктирная кривая), демонстрирует позитивную роль априорной информации в рассматриваемом синтезе управления: управление только по текущим данным “включается” только на 14 или 10 шагах, тогда как управление по совместной информации с первого шага обеспечивает приемлемый уровень качества системы.

8. Заключение

В статье разработан новый метод синтеза законов управления неизвестной динамической системой на основе априорной информации и зашумленных текущих данных, получаемых в процессе реального функционирования системы. На каждом шаге матрица параметров обратной связи находится путем решения линейных матричных неравенств так, чтобы минимизировать верхнюю границу обобщенной H_∞ -нормы замкнутой системы для всех моделей, согласованных с данными, полученными к этому шагу. Доказано, что последовательность получаемых при этом гарантированных значений соответствующей нормы монотонно не возрастает и что при остановке процесса настройки полученный регулятор обеспечивает определенное качество замкнутой системы. Отличительная особенность предлагаемого подхода состоит в том, что для его реализации не требуется выполнять условие неисчезающего возбуждения в системе для ее идентификации. Кроме того, в отличие от адаптивного управления, где основным вопросом является установление сходимости настраиваемых параметров к их истинным значениям и отсутствует

возможность оценить качество получаемой системы, здесь имеется возможность в реальном времени отслеживать значение оптимизируемого функционала. Численное моделирование активной защиты сооружений от сейсмических воздействий демонстрирует эффективность предложенного подхода.

ПРИЛОЖЕНИЕ

Доказательство леммы 4.1. Для матриц Ω , согласованных с данными \hat{x}_t, \hat{x}_{t+1} и \hat{u}_t , для некоторых возмущений и помех, удовлетворяющих ограничениям (2.3), верно равенство (4.4). Отсюда следует, что выполняется неравенство

$$\begin{aligned}
 & (\hat{x}_{t+1} - \Omega \hat{\varphi}_t)(\hat{x}_{t+1} - \Omega \hat{\varphi}_t)^T = \\
 \text{(П.1)} \quad & = (I \ \Omega \ B_w) \begin{pmatrix} \xi_{t+1} \\ \zeta_t \\ w_t \end{pmatrix} \begin{pmatrix} \xi_{t+1} \\ \zeta_t \\ w_t \end{pmatrix}^T (I \ \Omega \ B_w)^T \preceq \\
 & \preceq d^2 (I \ \Omega \ B_w) (I \ \Omega \ B_w)^T.
 \end{aligned}$$

Записывая это эквивалентно в виде

$$\Omega(\hat{\varphi}_t \hat{\varphi}_t^T - d^2 I) \Omega^T - \Omega \hat{\varphi}_t \hat{x}_{t+1}^T - \hat{x}_{t+1} \hat{\varphi}_t^T \Omega^T + \hat{x}_{t+1} \hat{x}_{t+1}^T - (I + B_w B_w^T) d^2 \preceq 0,$$

получим (4.5). В свою очередь, если выполнено неравенство в (П.1), то, как показано в [13], для некоторого возмущения и помех, удовлетворяющих ограничениям (2.3), имеет место равенство (4.4).

Доказательство леммы 5.1. Записывая неравенство (5.1) как

$$\Omega \Psi_{11} \Omega^T + \Omega \Psi_{12} + \Psi_{12}^T \Omega^T + \Psi_{22} \preceq 0$$

и выделяя полный квадрат, приведем его к виду (5.4). Так как $\Omega_{real} \in \hat{\Delta}_k(\mu, \tau_0^{k-1})$, то это множество не пусто и, следовательно, $\Gamma_k \succeq 0$. Умножая (5.4) слева на a^T и справа на a и учитывая, что $\Psi_{11} > 0$, получим

$$\begin{aligned}
 |(\Omega + \Psi_{12}^T \Psi_{11}^{-1})^T a|^2 & \leq \lambda_{\min}^{-1}(\Psi_{11}) \lambda_{\max}(\Gamma_k) \quad \forall a : |a| = 1 \\
 \Rightarrow \|(\Omega + \Psi_{12}^T \Psi_{11}^{-1})\|^2 & \leq \lambda_{\min}^{-1}(\Psi_{11}) \lambda_{\max}(\Gamma_k),
 \end{aligned}$$

т.е. множество $\hat{\Delta}_k(\mu, \tau_0^{k-1})$ ограничено.

Доказательство теоремы 6.1. Пусть неравенство (6.1) разрешимо при некоторых $P_k = P_k^T \succ R$, Q_k , $\mu \geq 0$ и $\tau_0^{k-1} \geq 0$. Заменим в нем $Q_k = \Theta P_k$ и, применяя лемму Шура относительно отрицательно определенного блока,

стоящего в первой строке и первом столбце, придем к эквивалентному неравенству

$$\begin{pmatrix} \Upsilon - \Psi_{11} & \star & \star \\ -\Psi_{21} & -\Psi_{22} - P_k + B_w B_w^T & \star \\ (C \ D)\Upsilon & 0 & (C \ D)\Upsilon(C \ D)^T - \gamma^2 I \end{pmatrix} \prec 0,$$

$$\Upsilon = \begin{pmatrix} I \\ \Theta \end{pmatrix} P_k \begin{pmatrix} I \\ \Theta \end{pmatrix}^T.$$

Непосредственная подстановка показывает, что последнее неравенство выполняется тогда и только тогда, когда для функции $V_k(x^{(a)}) = x^{(a)T} P_k x^{(a)}$ с $P_k \succ R$ при всех ненулевых $x_t^{(a)}$, $w_t^{(a)}$ и w_t^Δ выполняется неравенство

$$(II.2) \quad \Delta V_k(x_t^{(a)}) + |z_t^{(a)}|^2 - \gamma^2 |w_t^{(a)}|^2 - \begin{pmatrix} w_t^\Delta \\ x_t^{(a)} \end{pmatrix}^T \Psi(\mu, \tau_0^{k-1}) \begin{pmatrix} w_t^\Delta \\ x_t^{(a)} \end{pmatrix} < 0,$$

в котором $\Delta V_k(x_t^{(a)}) = V_k(x_{t+1}^{(a)}) - V_k(x_t^{(a)})$ обозначает приращение функции $V_k(x^{(a)})$ по траектории системы

$$(II.3) \quad \begin{aligned} x_{t+1}^{(a)} &= \begin{pmatrix} I \\ \Theta \end{pmatrix}^T w_t^\Delta + \begin{pmatrix} I \\ \Theta \end{pmatrix}^T (C \ D)^T w_t^{(a)}, \\ z_t^{(a)} &= B_w^T x_t^{(a)}. \end{aligned}$$

Далее, если в (II.3) положить $w_t^\Delta = \Omega^T x_t^{(a)}$, то эти уравнения примут вид

$$(II.4) \quad \begin{aligned} x_{t+1}^{(d)} &= \begin{pmatrix} I \\ \Theta \end{pmatrix}^T \Omega^T x_t^{(d)} + \begin{pmatrix} I \\ \Theta \end{pmatrix}^T (C \ D)^T w_t^{(d)}, \\ z_t^{(d)} &= B_w^T x_t^{(d)} \end{aligned}$$

и совпадают с уравнениями двойственной системы для исходной системы

$$(II.5) \quad \begin{aligned} x_{t+1} &= \Omega \begin{pmatrix} I \\ \Theta \end{pmatrix} x_t + B_w w_t, \\ z_t &= (C \ D) \begin{pmatrix} I \\ \Theta \end{pmatrix} x_t. \end{aligned}$$

Отсюда следует, что для функции $V_k(x^{(d)}) = x^{(d)T} P_k x^{(d)}$ с $P_k \succ R$ при всех ненулевых переменных $x_t^{(d)}$, $w_t^{(d)}$ системы (II.4) выполняются неравенства

$$(II.6) \quad \Delta V_k(x_t^{(d)}) + |z_t^{(d)}|^2 - \gamma^2 |w_t^{(d)}|^2 - x_t^{(d)T} \begin{pmatrix} \Omega^T \\ I \end{pmatrix}^T \Psi(\mu, \tau_0^{k-1}) \begin{pmatrix} \Omega^T \\ I \end{pmatrix} x_t^{(d)} < 0.$$

Так как для всех $\Omega \in \widehat{\Delta}_k(\mu, \tau_0^{k-1})$ имеем $(\Omega I)\Psi(\mu, \tau_0^{k-1})(\Omega I)^T \preceq 0$, то для функции $V_k(x^{(d)}) = x^{(d)T} P_k x^{(d)}$ с $P_k \succ R$ при всех ненулевых переменных $x_t^{(d)}$, $w_t^{(d)}$ и всех $\Omega \in \widehat{\Delta}_k(\mu, \tau_0^{k-1})$ выполняются неравенства

$$(II.7) \quad \Delta V_k(x_t^{(d)}) + |z_t^{(d)}|^2 - \gamma^2 |w_t^{(d)}|^2 < 0.$$

С учетом леммы 3.2 отсюда следует, что $u_t = \Theta x_t$ является гарантирующим обобщенным H_∞ -управлением с уровнем γ .

Обратно, пусть $u_t = \Theta x_t$ является гарантирующим обобщенным H_∞ -управлением с уровнем γ . Тогда из определения и леммы 3.2 следует, что для двойственной системы (II.4) существует функция $V_k(x^{(d)}) = x^{(d)T} P_k x^{(d)}$ с $P_k \succ R$ и такие $\widehat{\mu} \geq 0$ и $\widehat{\tau}_0^{k-1} \geq 0$, что при всех $\Omega \in \widehat{\Delta}_k(\widehat{\mu}, \widehat{\tau}_0^{k-1})$ выполняется неравенство (II.7). Покажем, что для функции $V_k(x^{(a)}) = x^{(a)T} P_k x^{(a)}$ с $P_k \succ R$ по траектории системы (II.3) выполняется неравенство

$$(II.8) \quad \Delta V_k(x_t^{(a)}) + |z_t^{(a)}|^2 - \gamma^2 |w_t^{(a)}|^2 < 0$$

при всех ненулевых $x_t^{(a)}$, $w_t^{(a)}$ и w_t^Δ , удовлетворяющих неравенству

$$(II.9) \quad \begin{pmatrix} w_t^\Delta \\ x_t^{(a)} \end{pmatrix}^T \Psi(\widehat{\mu}, \widehat{\tau}_0^{k-1}) \begin{pmatrix} w_t^\Delta \\ x_t^{(a)} \end{pmatrix} \leq 0.$$

Действительно, выберем $x_t^{(a)} = x_t^{(d)}$, $w_t^{(a)} = w_t^{(d)}$ и для каждого w_t^Δ , удовлетворяющего (II.9), определим Ω как решение линейного матричного уравнения $x_t^{(a)T} \Omega = w_t^\Delta T$, в котором для каждого столбца матрицы Ω имеется одно уравнение. При этом из (II.9) следует, что $x_t^{(a)T} (\Omega I) \Psi(\widehat{\mu}, \widehat{\tau}_0^{k-1}) (\Omega I)^T x_t^{(a)} \leq 0$, т.е. $\Omega \in \widehat{\Delta}_k(\widehat{\mu}, \widehat{\tau}_0^{k-1})$. Таким образом, при таком Ω уравнения (II.3) совпадают с уравнениями (II.4), а неравенство (II.8) совпадает с неравенством (II.7).

В силу неущербности S -процедуры при одном ограничении выполнение неравенства (II.8) при ограничении (II.9) эквивалентно выполнению неравенства

$$\Delta V_k(x_t^{(a)}) + |z_t^{(a)}|^2 - \gamma^2 |w_t^{(a)}|^2 - \nu \begin{pmatrix} w_t^\Delta \\ x_t^{(a)} \end{pmatrix}^T \Psi(\widehat{\mu}, \widehat{\tau}_0^{k-1}) \begin{pmatrix} w_t^\Delta \\ x_t^{(a)} \end{pmatrix} < 0$$

при некотором $\nu \geq 0$. Заметим, что по предположению Ω_{real} находится строго внутри $\widehat{\Delta}_k(\widehat{\mu}, \widehat{\tau}_0^{k-1})$, поэтому при $w_t^\Delta = \Omega_{real}^T x_t^{(a)}$ неравенство (II.9) является строгим. Так как $\Psi(\mu, \tau_0^{k-1})$ линейно зависит от μ и τ_0^{k-1} , то, обозначая $\nu \widehat{\mu} = \mu$ и $\nu \widehat{\tau}_0^{k-1} = \tau_0^{k-1}$, получим неравенство (II.2), которое эквивалентно линейному матричному неравенству (6.1) при $Q_k = \Theta P_k$.

Доказательство теоремы 6.2. Как и в доказательстве теоремы 6.1 при $k = N - 1$, получим, что для функции $V_{N-1}(x_t^{(d)}) = x_t^{(d)T} P_{N-1} x_t^{(d)}$ с

$P_{N-1} > R$ по траектории системы (П.4) при всех $\Omega \in \widehat{\Delta}_{N-1}(\mu, \tau_0^{N-1})$ выполняется неравенство (П.6). Следовательно, согласно леммам 3.1, 3.2 и замечанию 3.1 для функции $V_{N-1}(x_t) = x_t^T Y_{N-1} x_t$, где $Y_{N-1} = \gamma_{N-1}^{*2} P_{N-1}^{-1}$, по траектории системы (2.1), $u_t = \Theta_{N-1} x_t$ для всех $\Omega \in \widehat{\Delta}_{N-1}(\mu, \tau_0^{N-1})$ при $t \geq N-1$ выполняются неравенства

$$\Delta V_{N-1}(x_t) + |z_t|^2 - \gamma_{N-1}^{*2} |w_t|^2 < 0.$$

Суммируя эти неравенства, начиная с $t = N-1$, и учитывая, что $\lim_{t \rightarrow \infty} V_{N-1}(x_t) = 0$ и $Y_{N-1} < \gamma_{N-1}^{*2} R^{-1}$, получим (6.3).

СПИСОК ЛИТЕРАТУРЫ

1. Поляк Б.Т., Щербаков П.С. Робастная устойчивость и управление. М.: Наука, 2002.
2. Petersen I.R., Tempo R. Robust Control of Uncertain Systems: Classical Results and Recent Developments // Automatica. 2014. V. 50. No. 5. P. 1315–1335.
3. Markovskiy I., Huang L., Dörfler F. Data-driven control based on the behavioral approach: from theory to applications in power systems // IEEE Control Syst. 2023. V. 43. No. 5. P. 28–68.
4. Waarde H.J., Eising J., Camlibel M.K., Trentelman H.L. Data-driven control based on the behavioral approach: from theory to applications in power systems // IEEE Control Syst. 2023. V. 43. No. 6. P. 32–66.
5. Willems J.C., Rapisarda P., Markovskiy I., De Moor B. A note on persistency of excitation // Syst. Control Lett. 2005. V. 54. P. 325–329.
6. De Persis C., Tesi P. Formulas for Data-Driven Control: Stabilization, Optimality and Robustness // IEEE Trans. Automat. Control. 2020. V. 65. No. 3. P. 909–924.
7. Waarde H.J., Eising J., Trentelman H.L., Camlibel M.K. Data Informativity: a New Perspective on Data-Driven Analysis and Control // IEEE Trans. Automat. Control. 2020. V. 65. No. 11. P. 4753–4768.
8. Berberich J., Koch A., Scherer C.W., Allgower F. Robust data-driven state-feedback design // Proc. Amer. Control Conf. 2020. P. 1532–1538.
9. Bisoffi A., De Persis C., Tesi P. Trade-offs in learning controllers from noisy data // Syst. Control Lett. 2021. V. 154. Article 104985.
10. Waarde H.J., Camlibel M.K., Mesbahi M. From Noisy Data to Feedback Controllers: Nonconservative Design via a Matrix S-Lemma // IEEE Trans. Automat. Control. 2022. V. 67. No. 1. P. 162–175.
11. Bisoffi A., De Persis C., Tesi P. Data-driven Control via Petersen’s Lemma // Automatica. 2022. V. 145. Article 110537.
12. Nortmann B., Mylvaganam T. Direct data-driven control of LTV systems // IEEE Trans. Automat. Control. 2023. V. 68. P. 4888–4895.
13. Bisoffi A., De Persis C., Monshizadeh N. Controller synthesis for input-state data with measurement errors // IEEE Control Syst. Lett. 2024. V. 8. P. 1571–1576.
14. Berberich J., Scherer C.W., Allgower F. Combining Prior Knowledge and Data for Robust Controller Design // IEEE Trans. Automat. Control. 2023. V. 68. No. 8. P. 4618–4633.

15. Коган М.М., Степанов А.В. Синтез субоптимальных робастных регуляторов на основе априорных и экспериментальных данных // *АиТ*. 2023. № 8. С. 24–42.
16. Li L., De Persis C., Tesi P., Monshizadeh N. Data-based Transfer Stabilization in Linear Systems // *IEEE Trans. Automat. Control*. 2024. V. 69. P. 1866–1873.
17. Коган М.М., Степанов А.В. Синтез обобщенного H_∞ -субоптимальных управления по экспериментальным и априорным данным // *АиТ*. 2024. № 1. С. 3–20.
18. Niknejad N., Modares H. Physics-Informed Data-Driven Safe and Optimal Control Design // *IEEE Control Syst. Lett.* 2024. V. 8. P. 285–290.
19. Коган М.М., Степанов А.В. Как улучшить робастное управление линейной нестационарной системой с помощью экспериментальных данных // *АиТ*. 2024. № 6. С. 115–139.
20. Kogan M.M., Stepanov A.V. Robust H_∞ Control with Transients for LTV Systems Based on Prior Knowledge and Data // *Automatica*. 2025. V. 173. Article 112083.
21. Balandin D.V., Biryukov R.S., Kogan M.M. Pareto suboptimal H_∞ controls with transients // *Proc. Eur. Control Conf.* 2021, Rotterdam. P. 542–547.
22. Boyd S., El Ghaoui L., Feron E., Balakrishnan V. *Linear Matrix Inequalities in System and Control Theory*. Philadelphia: SIAM, 1994.
23. Petersen I.R., McFarlane D.C. Optimal Guaranteed Cost Control and Filtering for Uncertain Linear Systems // *IEEE Trans. Automat. Control*. 1994. V. 30. No. 9. P. 1971–1977.
24. Якубович В.А. *S-процедура в нелинейной теории управления* // *Вестн. Ленинград. ун.-та. Математика*. 1977. Т. 4. С. 73–93.

Статья представлена к публикации членом редколлегии П.С. Щербаковым.

Поступила в редакцию 03.03.2025

После доработки 20.05.2025

Принята к публикации 27.06.2025

© 2025 г. Е.Л. КУЛИДА, канд. техн. наук (elena-kulida@yandex.ru),
В.Г. ЛЕБЕДЕВ, д-р техн. наук (lebedev-valentin@yandex.ru)
(Институт проблем управления им. В.А. Трапезникова РАН, Москва)

ОБЗОР МЕТОДОВ ПРЕДУПРЕЖДЕНИЯ КОНФЛИКТОВ ПРИ УПРАВЛЕНИИ ВОЗДУШНЫМ ДВИЖЕНИЕМ С ПОМОЩЬЮ ГЛУБОКОГО ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

Представлен обзор развития современных подходов к предупреждению конфликтов между воздушными судами на основе глубокого обучения с подкреплением. Рассмотрена базовая концепция обучения с подкреплением и некоторые основные алгоритмы, используемые для предупреждения конфликтов воздушных судов. Представлены модели с дискретными и непрерывными действиями по предупреждению конфликтов в двумерном и трехмерном воздушном пространстве при движении по фиксированным траекториям или в свободном полете. Рассмотрены различные подходы к представлению информации о состоянии воздушного пространства (с помощью вектора состояния и в виде графа) и разные типы взаимодействия между воздушными судами (на основе информации о состоянии окружающих воздушных судов или при помощи обмена сообщениями).

Ключевые слова: управление воздушным движением, предупреждение конфликтов, маневр воздушного судна, глубокое обучение с подкреплением.

DOI: 10.31857/S0005231025090048, EDN: VMUGQH

1. Введение

Важнейшей функцией системы управления воздушным движением является поддержание безопасного расстояния разделения как по вертикали, так и по горизонтали между любыми двумя самолетами в любое время [1]. Нарушение безопасного разделения между воздушными судами (ВС) называется конфликтом. Растущая плотность воздушного движения приводит к увеличению числа потенциальных конфликтов, методы их предупреждения имеют важное значение для снижения рисков столкновений. Стратегическое предупреждение потенциальных конфликтов на основе глобальной модели оптимизации, которая заблаговременно генерирует бесконфликтные четырехмерные траектории для всех ВС, не позволяет справиться с неопределенностью, возникающей в динамике полетов в реальном времени [2]. Tактическое предупреждение потенциальных конфликтов в режиме реального времени критически важно для обеспечения безопасного управления воздушным движением, поскольку позволяет лучше справляться с неопределенностью, возникающей в динамике полетов.

В настоящее время движение ВС в основном осуществляется по фиксированным трассам, и ответственность за предупреждение конфликтов лежит на авиадиспетчерах. Тактические решения по-прежнему принимаются авиадиспетчерами с очень небольшими изменениями по сравнению с решениями, принимавшимися 50 лет назад [3]. С ростом интенсивности воздушного движения нагрузка авиадиспетчеров постоянно увеличивается и может презойти человеческие возможности. Развитие существующих методов оценки динамической воздушной обстановки авиадиспетчером с целью сокращения факторов опасности, включая информационную перегрузку и дефицит времени на функциональные операции, рассматривается в [4]. В [5–7] предложена концепция разработки средств автоматизации для увеличения пропускной способности и безопасности эксплуатации воздушного пространства с целью эффективной интеллектуальной поддержки принятия решений авиадиспетчерами для предупреждения конфликтов. Перспективная организация воздушного движения предполагает использование свободных полетов, когда ВС будут двигаться по произвольным траекториям, и предупреждение конфликтов будет обеспечиваться автономной системой управления воздушным движением. Теоретические исследования подтверждают, что свободный полет обладает потенциалом для повышения безопасности [8] и снижения расхода топлива [9]. Для реализации концепции свободного полета необходимо «обеспечивать контроль за эшелонированием ВС с помощью бортовых систем в дополнение к наземным системам. Более высокая надежность такой структуры позволит внести ясность в целый ряд сценариев моделирования риска столкновения ВС и будет способствовать созданию более безопасных, гибких и емких условий организации воздушного движения» [10]. Основополагающее значение для реализации концепции свободного полета имеет сложнейшая задача разработки децентрализованных автономных средств предупреждения конфликтов [11]. Методы и алгоритмы обнаружения и предотвращения опасных сближений в воздухе в рамках перспективной системы организации воздушного движения с учетом требований по безопасности и эффективности выполнения полетов исследовались в [12, 13].

Разработаны методы, позволяющие предоставить авиадиспетчерам рекомендации по разрешению конфликтов, на основе оптимального управления [14], математического программирования [15–18], геометрической оптимизации [19–21], эволюционных алгоритмов [22], алгоритма поиска по дереву по методу Монте-Карло [23]. Эти методы работают при существующей плотности воздушного движения, но при большей плотности проблемой является их недостаточная вычислительная эффективность. Требуются десятки или даже сотни секунд для получения решения. Учет неопределенности, присущей воздушному движению, значительно увеличивает время вычислений, тем самым снижая способность методов быстро генерировать решения. Большинство традиционных подходов к обеспечению эшелонирования не справляются со стохастическими средами и высокой плотностью воздушного движения [24]. Необходимы новые подходы, которые смогут эффективно реагиро-

вать на динамику внешней среды в режиме реального времени, например, на основе нейронных сетей и машинного обучения [25–27].

В последнее время, благодаря способности решать проблемы принятия решений, которые раньше были недоступны из-за сочетания нелинейности и высокой размерности, широкое применение в различных областях авиации получило глубокое обучение с подкреплением [28]. Использование глубокого обучения с подкреплением позволит автоматически получать безопасные и эффективные варианты решений по предупреждению конфликтов для поддержки принятия решений авиадиспетчерами и снизит их рабочую нагрузку [29]. В будущем полностью автоматизированная система управления станет окончательным решением для обработки высокоплотного, сложного и динамичного воздушного движения [30].

Применение методов обучения с подкреплением состоит из двух этапов: этап обучения модели и этап применения обученной модели на практике. Обучение модели может потребовать значительного времени, зато потом с помощью обученной модели решения получаются очень быстро. Скорость принятия решений – показатель эффективности и преимущество глубокого обучения с подкреплением по сравнению с традиционными алгоритмами. Более быстрое принятие решений означает более раннее обнаружение конфликтов и выработку инструкций для снижения рабочей нагрузки авиадиспетчеров и пилотов. В [31] приводятся следующие цифры (с оговоркой, что данные получены в разных условиях расчета): алгоритму на основе смешанно-целочисленного линейного программирования для генерации решения требуется 49 с, среднее время решения с помощью генетического алгоритма составляет 37,6 с, а обученному агенту глубокого обучения с подкреплением требуется менее 0,2 с. Методы обучения с подкреплением по сравнению с традиционными методами имеют явное преимущество в скорости вычисления решений и способности адаптироваться к динамике внешней среды, что критически важно при разрешении конфликтов в воздушном движении.

Исследования по разрешению конфликтов в воздушном движении с использованием глубокого обучения с подкреплением постоянно проводятся с 2018 г., и за это время было предложено множество моделей и алгоритмов [31]. Рассматриваются модели разрешения конфликтов и для случая движения по трассам, и для свободного полета, модели как с дискретными, так и с непрерывными действиями. В настоящее время в большинстве случаев предлагаются двумерные модели для разрешения конфликтов с помощью горизонтальных маневров и значительно реже трехмерные – для горизонтальных и вертикальных маневров [32]. Двумерные модели игнорируют вертикальные маневры из-за потенциальной нестабильности, которую они могут вызвать в воздушном движении [33]. Различается количество конфликтующих ВС: рассматриваются конфликты между двумя ВС, в группах ВС с фиксированным или с переменным количеством ВС. В [34–36] разработан гибридный подход, объединяющий для разрешения конфликтов сильные стороны геометрических методов и методов обучения с подкреплением. Утверждает-

ся, что большой диапазон различных оптимальных решений, найденных методом обучения с подкреплением, показывает, что правила геометрического метода должны быть расширены, учитывая различные геометрии конфликта. В РФ задача управления движением ВС на основе обучения с подкреплением в настоящее время активно исследуется специалистами ГосНИИАС, и уже имеется значительный научно-технический задел [37].

Поскольку существует большое разнообразие постановок задачи разрешения конфликтов, во многих случаях построение модели обучения с подкреплением начинается с разработки интерактивной среды, в которой агент может обучаться различным стратегиям. Агент искусственного интеллекта учится при помощи алгоритмов глубокого обучения с подкреплением методом проб и ошибок, применяя различные возможные действия и получая обратную связь от среды в виде вознаграждения. Цель агента заключается в изучении такой стратегии выбора действий, которая будет максимизировать математическое ожидание суммарных дисконтированных вознаграждений за продолжительное время. Сходимость модели обучения с подкреплением к нужному для решения поставленной задачи результату определяется выбором функции вознаграждения, с помощью которой агент учится оптимизировать стратегию выбора действий в различных ситуациях. Функция вознаграждения оказывает влияние на скорость обучения, сходимость и производительность агентов. Главное, что должно учитываться в функции вознаграждения – это удача или неудача разрешения конфликта. Кроме этого, для повышения эффективности модели вознаграждение должно учитывать количество маневров и время, необходимые для разрешения конфликта. При помощи функции вознаграждения можно приблизить поведение агента к существующим правилам разрешения конфликтов авиадиспетчерами [38, 39].

Несмотря на успешность обучения с подкреплением в исследованиях по решению задач управления воздушным движением, для применения этого метода в реальных условиях остаются две существенные проблемы. Первая проблема заключается в уязвимости глубоких нейронных сетей к состязательным атакам, вторая – проблема объяснимости моделей «черного ящика» – в непонимании пилотами и авиадиспетчерами механизма принятия моделями тех или иных решений [40, 41]. В [42] предлагаются подходы к решению этих проблем при автономном разрешении конфликтов ВС.

2. Обучение с подкреплением

2.1. Базовая концепция [43]

Машинное обучение подразделяется на контролируемое обучение, неконтролируемое обучение и обучение с подкреплением. Контролируемое обучение позволяет аппроксимировать любую функцию, но требует наличия наборов размеченных данных, которые доступны далеко не всегда. Неконтролируемое обучение предполагает наличие наборов неразмеченных данных. Обучение с подкреплением позволяет реализовать последовательный процесс

принятия решений методом проб и ошибок, синтезируя данные для обучения в процессе взаимодействия агента со средой. При использовании обучения с подкреплением каждое ВС моделируется как интерактивный агент, действиями которого являются маневры по предупреждению конфликтов.

В основе обучения с подкреплением лежит модель марковского процесса принятия решений, при котором состояние системы и действия агента не зависят от того, каким образом система пришла в это состояние. Для марковского процесса принятия решений справедливы уравнения Беллмана. Основной модели марковского процесса принятия решений являются среда и действующий в ней агент. Среда характеризуется набором параметров, состояние среды s – это определенный набор значений этих параметров. Агент – это программа, которая может анализировать состояние среды и выполнять в каждом состоянии определенный набор действий $a \in A(s)$. В результате выполнения агентом действия среда переходит из состояния s в новое состояние s' и получает обратную связь от среды в виде вознаграждения $r = R(s, a, s')$. При многошаговом взаимодействии агента со средой, начиная с состояния s на шаге t до конца эпизода на шаге T , определяется суммарная дисконтированная выгода:

$$G_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots + \gamma^{T-1} r_T = r_{t+1} + \gamma G_{t+1},$$

$\gamma \in [0, 1]$ – коэффициент дисконтирования, который определяет, насколько убывает ценность награды, полученной на более поздних шагах.

Функция перехода $p(s'|s, a)$ определяет вероятность перехода в состояние s' на шаге t при условии, что на шаге $t - 1$ в состоянии s было выбрано действие a :

$$p(s'|s, a) = P(S_t = s' | S_{t-1} = s, A_{t-1} = a), \quad \sum_{s' \in S} p(s'|s, a) = 1, \forall s \in S, \forall a \in A(s).$$

Стратегией (или политикой) называется функция $\pi(a|s)$, сопоставляющая действие агента каждому нетерминальному состоянию среды.

Ожидаемая выгода при следовании агентом стратегии π в состоянии s называется функцией ценности состояния:

$$V_\pi(s) = E_\pi[r_{t+1} + \gamma G_{t+1} | S_t = s].$$

Однако функция ценности состояния не позволяет узнать ожидаемую выгоду от выполнения агентом действия a в состоянии s при следовании агентом стратегии π , она определяется функцией ценности действия:

$$Q_\pi(s, a) = E_\pi[r_{t+1} + \gamma G_{t+1} | S_t = s, A_t = a].$$

Базовая концепция обучения с подкреплением – обобщенная итерация по стратегиям – это итеративная процедура. Шаг этой процедуры включает два процесса: оценка текущей стратегии для уточнения текущего приближения функции ценности и затем улучшение стратегии в соответствии с изменившейся функцией ценности.

Функция ценности может быть представлена в виде

$$V_{\pi}(s) = \sum_a \pi(a|s) \sum_{s',r} p(s', r|s, a)[r + \gamma V_{\pi}(s')], \forall s \in S.$$

Шаг итерации для уточнения функции ценности заключается в ее вычислении для действия с наиболее высокой ценностью:

$$V_{k+1}(s) = \max_a \sum_{s',r} p(s', r|s, a)[r + \gamma V_k(s')].$$

Действие, доставляющее максимальное значение Q -функции, называется жадным. Стратегия может быть улучшена при помощи алгоритма оптимизации стратегий, который заключается в выборе жадного действия относительно Q -функции:

$$\pi'(s) = \operatorname{argmax}_a \sum_{s',r} p(s', r|s, a)[r + \gamma V_{\pi}(s')].$$

В процессе изучения функции ценности очень важно соблюдать баланс между выбором жадного действия и выбором случайного действия для исследования. Существует много различных подходов к решению этой проблемы: эпсилон-жадная стратегия, при которой случайное действие выбирается с вероятностью эпсилон; затухающая эпсилон-жадная стратегия, при которой эпсилон уменьшается по мере обучения агента; стратегии выбора действий, использующие полученные к текущему шагу знания о ценности и изученности действий, и др.

Рассмотренные два процесса стабилизируются, когда функция ценности будет соответствовать стратегии, а стратегия будет жадной по отношению к функции ценности. Стратегия и функция ценности оптимальны, если:

$$V^*(s) = \max_{\pi} V_{\pi}(s).$$

2.2. Алгоритмы глубокого обучения с подкреплением [44, 45]

Для простых примеров функция ценности действий представляется в виде таблицы. Для важных практических задач табличное представление функций ценности реализовать невозможно из-за большого числа состояний, непрерывных переменных или действий. В этом случае используется глубокое обучение с подкреплением, при котором для аппроксимации функций ценности используются алгоритмы глубокого обучения с учителем на основе выборок (s, a, r, s') , сформированных в процессе взаимодействия агента со средой.

В настоящее время в области предупреждения авиационных конфликтов этот подход бурно развивается, предлагаются и исследуются различные алгоритмы аппроксимации функции ценности и оптимизации стратегий. Широко

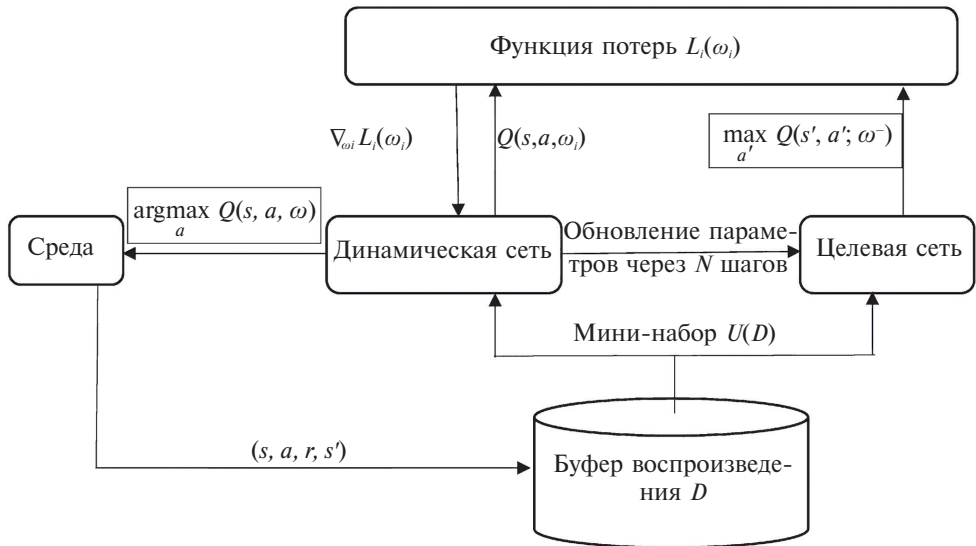


Рис. 1. Принцип работы алгоритма глубокого Q -обучения.

применяется алгоритм глубокого Q -обучения (deep Q -network – DQN), который использует две ключевые технологии: воспроизведение опыта и двойную сетевую структуру. На рис. 1 показан принцип работы алгоритма DQN [46].

Воспроизведение опыта заключается в создании буфера воспроизведения D , в котором накапливается большое количество выборок. Мини-наборы $U(D)$ для обучения сети выбираются из накопленного буфера воспроизведения равномерно случайно и, таким образом, соответствуют различным траекториям и политикам, что повышает стабильность обучения сети. Двойная сетевая структура основана на использовании одной и той же сети с разными наборами параметров. Динамическая сеть используется для аппроксимации текущего значения $Q(s, a; \omega_i)$, параметры этой сети ω_i обновляются на каждом временном шаге i . Целевая сеть используется для получения более устойчивого целевого значения Q , параметры целевой сети ω^- обновляются через N временных шагов. Функция потерь имеет вид

$$L_i(\omega_i) = E_{(s,a,r,s') \sim U(D)} \left[\left(r + \gamma \max_{a'} Q(s', a'; \omega^-) - Q(s, a; \omega_i) \right)^2 \right].$$

Функция потерь оптимизируется методом градиентного спуска:

$$\begin{aligned} \nabla_{\omega_i} L_i(\omega_i) &= \\ &= E_{(s,a,r,s') \sim U(D)} \left[\left(r + \gamma \max_{a'} Q(s', a'; \omega^-) - Q(s, a; \omega_i) \right) \nabla_{\omega_i} Q(s, a; \omega_i) \right]. \end{aligned}$$

Отличие метода двойной глубокой Q -сети (Double Deep Q -network – DDQN) [47] от метода DQN заключается в том, что для выбора действия в уравнении вместо параметров целевой сети используются параметры динамической

сети:

$$\nabla_{\omega_i} L_i(\omega_i) = E_{(s,a,r,s') \sim U(D)} \left[\left(r + \gamma Q(s', \operatorname{argmax}_{a'} Q(s', a'; \omega_i); \omega^-) - Q(s, a; \omega_i) \right) \nabla_{\omega_i} Q(s, a; \omega_i) \right].$$

Алгоритмы DQN и DDQN используются в моделях предупреждения конфликтов с дискретным пространством действий [48–52].

Предлагаются также методы, использующие для разрешения конфликтов между ВС непрерывное пространство действий [34, 53–55]. В этом случае стратегии представляются параметризованными стохастическими функциями $\pi_\theta(a, s)$, которые оптимизируются при помощи алгоритмов актер–критик. В этих алгоритмах, помимо нейронной сети для оценки стратегии (критик), используется вторая нейронная сеть для формирования стратегии на основе оптимизации функции ценности (актер).

Для траектории $\tau = S_0, A_0, R_1, S_1, \dots, S_{T-1}, A_{T-1}, R_T, S_T$ функция $G(\tau)$ – полная дисконтированная выгода, $\pi_\theta(A_t|S_t)$ – вероятность выбора действия A_t в состоянии S_t на шаге t . Сеть актера обновляется в соответствии с градиентом функции ценности:

$$\nabla_\theta E_{\tau \sim \pi_\theta} [G(\tau)] = E_{\tau \sim \pi_\theta} \left[\sum_{t=0}^T G_t(\tau) \nabla_\theta \log \pi_\theta(A_t|S_t) \right].$$

3. Модели предупреждения конфликтов между воздушными судами

3.1. Модели предупреждения конфликтов между двумя воздушными судами

Стратегия разрешения конфликтов может включать как двумерные (2D) маневры – изменение курса и скорости в плоском воздушном пространстве, так и трехмерные (3D) маневры с изменением еще и высоты полета. Хотя 2D-модель не так эффективна в разрешении конфликтов, как 3D-модель, маневры в двумерном воздушном пространстве вызывают меньший дискомфорт у пассажиров и не искажают вертикально стратифицированную структуру воздушного пространства [56]. Слишком большие нейронные сети могут быть пагубными для обучения агента из-за слишком большого числа параметров. Поэтому меньшие нейронные сети для 2D-моделей имеют больший потенциал для дальнейшего расширения модели, чтобы учитывать больше факторов реального мира [30].

Первые исследования по предотвращению конфликтов между ВС при помощи обучения с подкреплением рассматривали разрешение конфликтов между двумя ВС в двумерном воздушном пространстве. В одной из первых работ по предотвращению конфликтов между ВС рассматривались полеты по маршрутам и была предложена иерархическая структура глубокого обучения с подкреплением [48]. В качестве среды обучения использовалось программное обеспечение под названием NASA Sector 33, содержащее 35 задач по

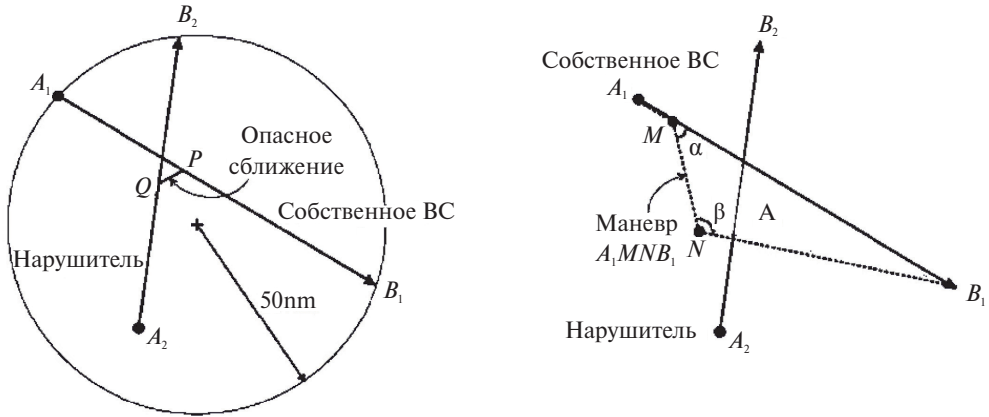


Рис. 2. Сценарий конфликта двух ВС и маневр по его предупреждению.

управлению воздушным движением с участием от двух до пяти самолетов. Иерархическая структура включает родительский агент, предназначенный для решения задачи выбора маршрутов самолетов в начале эпизода, затем дочерний агент управляет действиями по изменению скоростей на выбранных маршрутах. Иерархическая структура позволяет разделить действия по выбору маршрутов, выполняемых в начале эпизода, и действия по управлению скоростями в процессе эпизода. Для обучения агентов используется алгоритм обучения с подкреплением на основе двойной глубокой Q -сети. Первая нейронная сеть (целевая) используется для выбора действий, жадных по отношению к текущей Q -функции, вторая нейронная сеть (динамическая) используется для корректировки Q -функции на основе оценки успешности выполненных действий. В [48] показано, что иерархический агент глубокого обучения с подкреплением может выбрать оптимальные комбинации маршрутов и скоростей, чтобы избежать конфликта двух самолетов при полетах по маршрутам.

В одной из первых работ по обучению с подкреплением для разрешения конфликта при свободном полете рассматривается случай двух самолетов с учетом неопределенности [53]. Разработана среда для моделирования потенциальных конфликтов для обучения и тестирования агента. На рис. 2 представлен конфликт двух ВС в круговом воздушном пространстве с радиусом R , равным 50 морских миль, и маневр по его предотвращению. Траектория собственного ВС – A_1B_1 , траектория нарушителя – A_2B_2 , QP – ближайшее расстояние между ВС, на котором они теряют безопасное разделение, если ни одно из ВС не совершит маневра. В качестве действия по предупреждению конфликта используется однократный маневр по изменению курса в непрерывном двумерном пространстве. Маневр A_1MNB_1 на рис. 2 представляет ряд действий, выполняемых собственным ВС: отклонение от первоначального пути в точке M изменением курса на угол α , затем движение по вектору MN , в точке N поворот к точке B_1 .

Вознаграждение обучающегося агента рассчитывается в виде суммы:

$$R_{final} = 5 + R_{conflict} + R_{maneuver} + R_{deviation},$$

т.е. к вознаграждению 5 добавляются отрицательные вознаграждения: $R_{conflict} = -3$, если маневр не разрешает конфликт; $R_{maneuver} = -2$, в случае выхода за границы области или если угол $\beta > 120^\circ$; $R_{deviation} = -S_{\Delta MNB_1}$, отклонение от исходной траектории оценивается площадью между траекторией маневра и исходной траекторией.

Используется метод глубокого градиента по детерминированным политикам (Deep Deterministic Policy Gradient – DDPG) – один из передовых методов глубокого обучения с подкреплением для задач управления с непрерывным пространством действий [44]. Алгоритм использует две нейронные сети: сеть критика для изучения функции полезности пар состояние–действие $Q(s, a)$ и сеть актера для отображения состояния в детерминированное действие на основе градиента политики. Производительность алгоритма DDPG для предупреждения конфликтов в воздушном движении близка к производительности традиционных методов, но время расчетов при этом значительно сокращается [57].

В [54, 56, 58, 59] постановка задачи предупреждения конфликтов обобщается на случай, когда помимо собственного ВС и нарушителя в области находится еще несколько ВС. В этом случае при маневрировании для предупреждения конфликта могут возникнуть вторичные потенциальные конфликты (эффект домино) [60].

В [54] разработана двумерная среда для моделирования свободных полетов, которая может быть применена для нескольких ВС в секторе (не более 5) (рис. 3).

Предполагается, что между этими ВС нет конфликтов, конфликты могут быть вызваны входящим в сектор ВС, цель которого – пролететь от начальной точки до конечной за минимальное время без столкновений с другими ВС. Предлагается алгоритм «актер–критик» с фиксированным количеством K действий (циклов управления). Чтобы избежать конфликтов, генерируются действия по изменению курсового угла входящего ВС. На каждом временном шаге агент выбирает действие $A = \{\rho, \varphi | \rho \in [0, L], \varphi \in [-\pi, \pi]\}$, которое описывается двумерной полярной координатой, где ρ, φ – полярный радиус и угол.

Функция вознаграждения

$$R_t = \begin{cases} -1, & \text{если есть конфликт,} \\ 1 - \frac{1}{K} \times \frac{|\Delta\varphi_t|}{\pi} & \text{в противном случае.} \end{cases}$$

Функция ценности аппроксимируется с помощью нейронной сети $\hat{V}(S_t, \omega) \approx V_\pi(S)$, где ω – веса нейронов.

Для сети критика определяется δ :

$$\delta_t = R_t + \gamma \hat{V}(S_{t+1}, \omega) - \hat{V}(S_t, \omega),$$

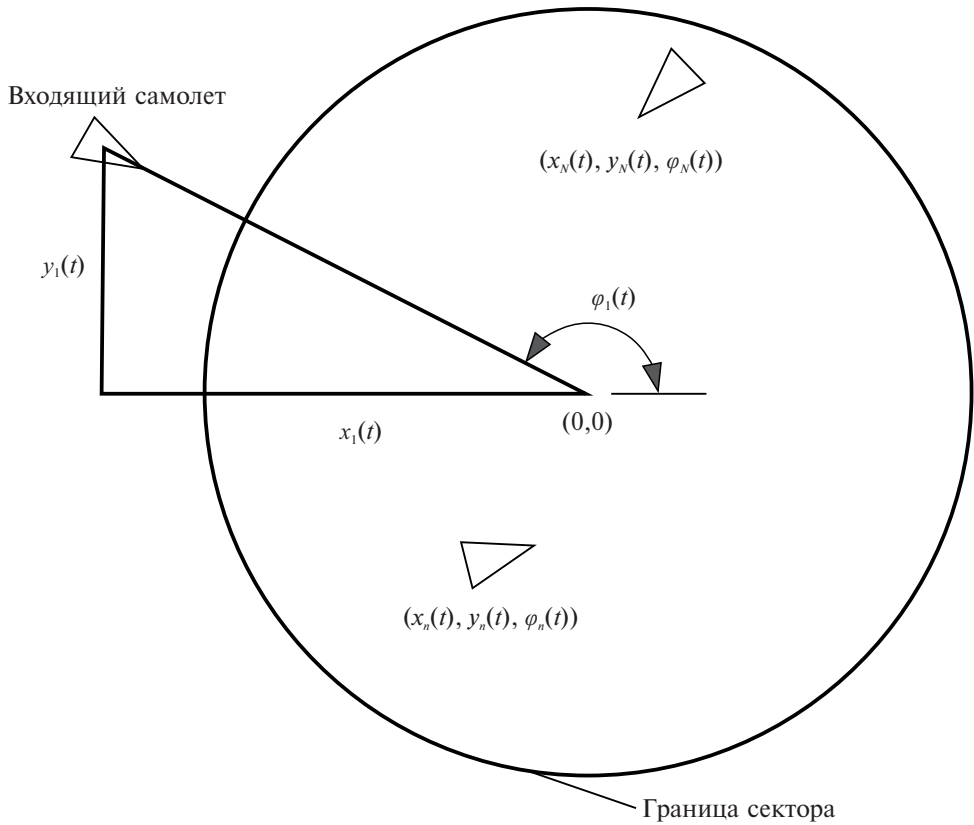


Рис. 3. Сценарий конфликта с несколькими ВС.

где R_t – немедленное вознаграждение, $\hat{V}(S_{t+1}, \omega)$ – значение ценности следующего состояния, $\hat{V}(S_t, \omega)$ – значение ценности текущего состояния. Для обновления параметров применяется метод наименьших квадратов: $\omega \leftarrow \omega + \alpha \nabla \delta^2$, α – скорость обучения.

Для сети актера используется метод градиента политики. Для действия (ρ, φ) применяется уравнение:

$$\ln \pi(\rho_t, \varphi_t | S_t, \theta) = \ln \pi(\rho_t | S_t, \theta) + \ln \pi(\varphi_t | S_t, \theta),$$

где $\pi(\rho_t, \varphi_t | S_t, \theta)$ – вероятность выбора ρ и φ в состоянии S_t с параметрами θ , $\pi(\rho_t | S_t, \theta)$ – вероятность выбора ρ в состоянии S_t с параметрами θ , $\pi(\varphi_t | S_t, \theta)$ – вероятность выбора φ в состоянии S_t с параметрами θ . Параметры θ обновляются по формуле:

$$\theta \leftarrow \theta + \alpha^\theta \delta_t \nabla \ln \pi(\rho_t, \varphi_t | S_t, \theta).$$

Переменное количество циклов управления не отражено в функции вознаграждения, и его нельзя корректировать динамически. Поскольку хорошо обученный агент обладает очень высокой вычислительной эффективностью, предлагается обучить несколько агентов с разными значениями K , например $K = 1, 2, 3$. В реальном процессе управления воздушным движением мож-

но быстро рассчитать несколько вариантов управления и выбрать лучшее решение из нескольких вариантов. Результаты моделирования подтвердили возможность применения глубокого обучения с подкреплением для разрешения конфликтов и его преимущество в вычислительной эффективности по сравнению с известными методами.

В [61] предложена и проанализирована модель для разрешения конфликтов между двумя ВС, учитывающая неопределенность, связанную с ветром. Предложенный метод предупреждения конфликтов применим в случае, когда неопределенность, связанная с направлением и скоростью ветра, непостоянна (стохастична) на протяжении всего моделирования.

В [46] рассматривается стратегия разрешения конфликтов двух самолетов в трехмерном пространстве на основе глубокого обучения с подкреплением. Из сценария воздушного движения, который может содержать несколько конфликтов, для разрешения выбирается один конфликт двух самолетов. Модель разрешения конфликтов моделируется как дискретный по времени марковский процесс принятия решений. Агент использует команды корректировки высоты, корректировки скорости или корректировки курса для разрешения конфликта. Предпочтения авиадиспетчера по выбору маневров разрешения конфликтов передаются агенту при помощи корректировки функции вознаграждения.

3.2. Модели предупреждения конфликтов между фиксированным количеством воздушных судов

Многоагентное обучение с подкреплением рассматривает набор агентов, которые взаимодействуют с одной и той же средой [62]. Каждый агент пытается достичь своих целей, которые неизвестны другим агентам. Одной из стратегий решения проблем в многоагентной среде является независимое Q -обучение, при котором между агентами нет связи и другие агенты считаются частью среды [63]. Однако, когда агент меняет свою политику, это влияет на политику других агентов, что приводит к нестабильности обучения [64].

В [3, 65, 66] для обеспечения коммуникации между агентами предлагается, чтобы состояние для каждого агента включало информацию о состоянии N -ближайших агентов. Пространство состояний для агента имеет постоянный размер, поскольку оно зависит только от N ближайших агентов и не масштабируется с увеличением числа агентов в среде. При этом отмечается, что определение того, какие N ближайших агентов следует учитывать, очень важно для получения хорошего результата, поскольку добавление нерелевантной информации в пространство состояний усложняет обучение [3].

В [67] для решения проблемы самоподдержания безопасного разделения ВС при высоких потоках движения в секторе представлен алгоритм обучения с подкреплением в сочетании с поиском по дереву Монте-Карло (MCTS-UCT). Все агенты (ВС) находятся на одном и том же эшелоне полета, стратегии действий включают изменения курса и крейсерской скорости.

Значение вознаграждения для агента в зависимости от состояния s определяется следующим образом:

$$r(s) = \begin{cases} 1, & \text{если } s \text{ – целевое состояние,} \\ 0, & \text{если } s \text{ – } LOS \text{ или вне границы сектора,} \\ 1 - \frac{d(o, g)}{\max d(o, g)} & \text{в противном случае,} \end{cases}$$

где состояние LOS – потеря безопасного разделения, $d(o, g)$ – расстояние между текущим положением агента и его целью, $\max d(o, g)$ – наибольшее расстояние между агентом и его целью.

Каждое состояние агента рассматривается как узел в дереве, развертывание по дереву выполняется на основе значений состояния, рассчитываемых по формуле

$$UCT(S_j) = \bar{r}_j + 2C \sqrt{\frac{2 \ln N}{n_j}},$$

где \bar{r}_j – среднее значение вознаграждения действия j для текущего агента, N – счетчик посещений узла, n_j – счетчик выбора действия j , $C = 1/\sqrt{2}$.

В процессе совместного принятия решений все n агентов $\{A_1, \dots, A_n\}$ должны делиться своим намерением при выборе каждого отдельного действия. Одна итерация алгоритма мультиагентного марковского процесса принятия решений заключается в следующем. Сначала n агентов $\{A_1, \dots, A_n\}$ инициализируются на уровне $L - 1$. Все агенты продолжают выполнять политику совместных действий по умолчанию $a_{-j} = \{a_i \text{ из политики действий по умолчанию } | i = 1, \dots, n, i \neq j\}$. Агент A_j с минимальным индексом на уровне $L - 1$ при помощи алгоритма MCTS-UCT выбирает свою оптимальную стратегию действий a_j^* в алгоритме, используя следующее уравнение:

$$a_j^* = \operatorname{argmax}_{a_j} r_j^*(s, a_j, a_{-j}), \quad j = 1, \dots, n,$$

где $r_j^*(s, a_j, a_{-j})$ – это значение вознаграждения агента A_j в состоянии s при выполнении действия a_j , в то время как стратегии действий других агентов представлены как a_{-j} . При вычислении a_j^* другие агенты продолжают следовать стратегии действия по умолчанию, установленной как $a_{-j} = \{a_i \text{ из политики действий по умолчанию } | i = 1, \dots, n, i \neq j\}$. Когда агент A_j получает свою оптимальную стратегию действий a_j^* , A_j повышается до уровня L и сохраняет стратегию действия a_j^* для обновления стратегии совместного принятия решений по умолчанию. Затем следующий агент, с минимальным индексом на уровне $L - 1$, выбирает свою оптимальную стратегию действий. Этот процесс будет продолжаться до тех пор, пока все агенты получают свою оптимальную стратегию действий $\{a_i^*, i = 1, \dots, n\}$. Полученные стратегии совместного принятия решений используются для следующего временного шага Δt для всех агентов. Итерации повторяются до тех пор, пока все агенты достигнут целевого состояния.

В [68] предложен метод глубокого ансамблевого многоагентного обучения с подкреплением для динамической корректировки скорости ВС в режиме реального времени. Авторы утверждают, что обширные эмпирические результаты, полученные с помощью модели управления воздушным движением с открытым исходным кодом, разработанной Eurocontrol и основанной на реальных данных, включающих тысячи ВС, демонстрируют, что предложенный метод значительно превосходит другие эталонные подходы.

В [69] предлагается подход к многоагентному обучению с подкреплением для трехмерного разрешения конфликтов в свободном маршрутном пространстве, в котором агенты используют общую нейронную сеть. Обученная сеть развертывается на каждом самолете для формирования распределенной системы принятия решений в реальном времени. При этом коммуникация между агентами сводится к информированию других агентов о выбранных действиях. Введение трехмерного пространства приводит к взрывному росту масштаба нейронной сети и, как следствие, к увеличению сложности обучения. Для преодоления этой проблемы предлагается для трехмерных маневров вместо трехмерного пространства рассматривать конфликты в трех плоскостях: плоскости текущего эшелона полета и двух смежных плоскостях на уровнях выше и ниже. Это значительно снижает сложность обучения агента. Предлагается модель буфера конфликта, в которой каждому ВС назначается защищенная зона и зона предупреждения. Рисунок 4 [69] иллюстрирует модель буфера конфликта, где R^W и R^P обозначают радиусы зоны предупреждения и защищенной зоны соответственно. Нарушители, обнаруженные внутри защищенной зоны, всегда получают большой штраф, нарушители, обнаруженные в зоне предупреждения, получают небольшой штраф.

Количество конфликтующих ВС, за которыми может наблюдать агент, фиксировано в структуре нейронной сети и не должно изменяться. Для решения этой проблемы разработана модель частичного наблюдения, при которой рассматривается только фиксированное количество ВС, представляющих наибольшую угрозу, например наиболее близких к агенту. Для реализации обучения с подкреплением используется алгоритм DQN с улучшениями, называемый Rainbow [70]. Для обучения и оценки предлагаемого подхода создана среда моделирования, учитывающая неопределенность полета (возникающую, например, в результате механических и навигационных ошибок и ветра). Экспериментальные результаты показывают [69], что предложенный метод может разрешать конфликты в сценариях с гораздо более высокой плотностью трафика, чем в сегодняшних реальных ситуациях. Сравниваются 2D и 3D модели. Время обучения 2D модели составляет менее 1% от времени обучения 3D модели, однако это не влияет на производительность модели во время решения из-за природы методов обучения с подкреплением. В некоторых сценариях 3D модель может разрешать конфликты легче, изменяя эшелон полета. Показано, что 3D-модель превосходит 2D-модель по показателю успешности и показателю снижения дополнительной дальности полета.

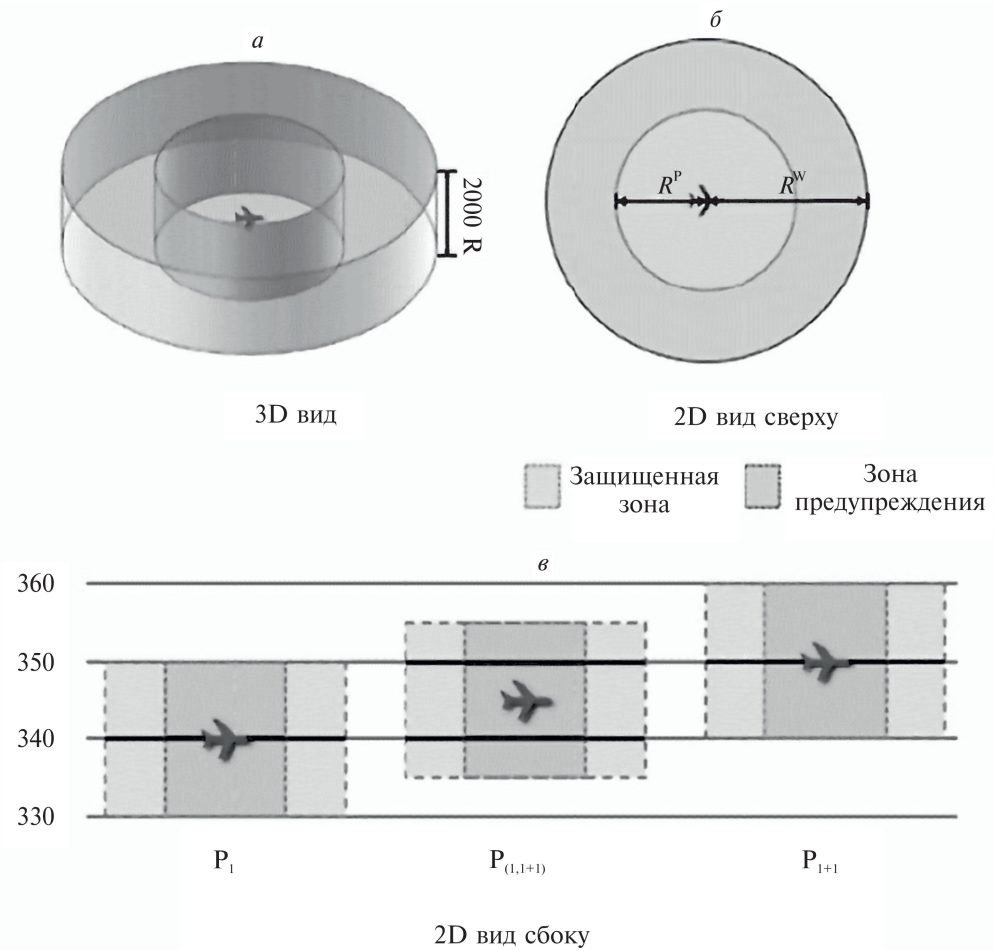


Рис. 4. Модель буфера конфликта.

При централизованном управлении воздушным движением, когда авиадиспетчер передает директивы пилотам, действия должны быть достаточно редкими. Однако в моделях с непрерывным пространством действий агенты могут вносить небольшие корректировки своих траекторий на каждом шаге. В [71] предлагается многоагентный метод глубокого обучения с подкреплением с непрерывным пространством действий, при котором количество действий значительно сокращается с помощью механизма приоритетов. На каждом временном шаге максимум один самолет с наивысшим приоритетом может выполнять действия. Такой подход значительно сокращает количество предпринимаемых действий, сохраняя при этом высокий уровень производительности предупреждения конфликтов. Получаемые решения подходят для централизованного управления воздушным движением, где количество директив, которые могут быть переданы пилоту, ограничено. В [72] механизм приоритетов на основе динамической оценки близости конфликтов между ВС используется в модели с дискретным пространством действий.

3.3. Модели предупреждения конфликтов между произвольным количеством воздушных судов

Выше рассмотрены модели, в которых агент имеет доступ к информации о состоянии N ближайших ВС, где N – гиперпараметр, выбираемый в процессе экспериментов, однако это ограничивает переносимость модели. Использование параметра N является недостатком, поскольку небольшое изменение местоположения самолета может привести к изменению множества N ближайших самолетов и, таким образом, к изменению входных данных для нейронной сети. Нейронная сеть должна понимать, что несмотря на перестановки ВС, это почти одно и то же состояние воздушного пространства, но это может быть сложной задачей [49]. Решение этой проблемы заключается в использовании более продвинутых архитектур нейронных сетей, поиске других способов представления данных, использовании алгоритмов улучшения данных, выборе наиболее релевантных соседних ВС [71]. Одним из способов решения проблемы переменного числа ВС является графическое кодирование информации в изображения фиксированного размера и использование сверточных нейронных сетей (CNN) для извлечения полезной информации аналогично экранам авиадиспетчеров [73]. Другая возможность заключается в использовании рекуррентных нейронных сетей (RNN) с ячейками долгой краткосрочной памяти (LSTM) [74] или управляемыми рекуррентными блоками (GRU) [75], которые работают со всем набором самолетов в среде, кодируя соответствующую информацию в скрытое состояние фиксированного размера.

В [30, 76] рассматривается многоагентное обучение с подкреплением для разрешения конфликтов на маршрутах и пересечениях в структурированном двумерном воздушном пространстве между переменным количеством ВС. Информация о состоянии кодируется при помощи нейронной сети LSTM в вектор фиксированной длины. Агент имеет доступ к закодированной информации обо всех ВС в секторе, в этом случае нет необходимости определять значение N для каждой новой среды. В качестве среды для обучения используется симулятор управления воздушным движением BlueSky [77]. Используется централизованное обучение и децентрализованная схема выполнения, при которой обучается одна нейронная сеть. Эта сеть используется всеми агентами для получения рекомендаций по скорости, при этом в зависимости от состояния действия агентов могут быть разными. Среда является стохастической, поскольку существует неопределенность в действиях других агентов, поэтому используется алгоритм «актер–критик» с названием «проксимальная оптимизация политики» [78].

Для обеспечения требований безопасного разделения вводится идентичная функция вознаграждения для всех агентов:

$$r_t = \begin{cases} -1, & \text{если } d_o^c < d^{LOS}, \\ -\alpha + \delta d_o^c, & \text{если } d_o^c < 10 \ \& \ d_o^c \geq d^{LOS}, \\ 0 & \text{в противном случае,} \end{cases}$$

где d^{LOS} – минимальное безопасное расстояние разделения в морских милях ($d^{LOS} = 3$), d_c^c – расстояние от собственного ВС до ближайшего самолета в морских милях, α и δ – маленькие положительные константы для штрафования агентов по мере приближения к потере безопасного расстояния разделения. На трех практических сценариях воздушного движения показана способность решать задачи принятия решений с переменным числом агентов и неопределенностью [30].

Рекуррентные нейронные сети обрабатывают входные данные последовательно, и выходные данные зависят от этой последовательности. Это может привести к нежелательному результату в ситуациях, когда входная последовательность значения не имеет. Трансформеры были введены как альтернатива рекуррентным нейронным сетям для последовательной обработки входных данных, чтобы обеспечить параллельное обучение [79]. Трансформеры вычисляют относительную важность так называемых токенов, содержащих информацию о состояниях ВС, используя механизм внимания. В [80] для наблюдения используются абсолютные состояния – координаты и скорости ВС в системе отсчета, связанной со средой. В [81] используются относительные состояния – координаты и скорости ВС в системе отсчета собственного ВС с положительным направлением оси абсцисс в направлении полета. Однако исследования пока не показали превосходства производительности архитектуры сетей трансформеров по сравнению с нейронными сетями с прямой связью и с рекуррентными нейронными сетями для обеспечения безопасной навигации ВС [81].

4. Кооперативные стратегии предупреждения конфликтов на основе нейронных сетей передачи сообщений

Коммуникация является ключевой способностью кооперативных многоагентных систем, в которых агенты могут получать значительную выгоду от обмена информацией перед выполнением совместных действий [82]. Модель, основанная на нейронных сетях передачи сообщений, позволяющая агентам обмениваться информацией через протокол связи, может позволить агентам разрабатывать кооперативные стратегии совместных действий по предотвращению конфликтов [83].

В [80] самолеты в воздушном пространстве моделируются как агенты кооперативной многоагентной системы. Состояние $s_i = [x_i, y_i, v_i, \chi_i]$ каждого агента $i \in N$ состоит из координат в евклидовом пространстве x_i, y_i , скорости v_i и курса χ_i . Состояние изменяется по формулам:

$$\begin{aligned} x_i(t+1) &= x_i(t) + v_i(t) \sin \chi_i(t) \Delta t, \\ y_i(t+1) &= y_i(t) + v_i(t) \cos \chi_i(t) \Delta t, \\ v_i(t+1) &= v_i(t) + \Delta v_i, \\ \chi_i(t+1) &= \chi_i(t) + \Delta \chi_i, \end{aligned}$$

где Δv_i и $\Delta \chi_i$ – приращения скорости и курса, Δt – шаг моделирования.

Взаимодействие между агентами представлено в виде графа $G = (V, E)$, каждый узел соответствует одному агенту $i \in N$, ребрами e_{ij} соединены агенты, которые могут обмениваться сообщениями. Вектор наблюдения о текущем состоянии i -го агента состоит из пяти элементов:

$$o_i = \begin{bmatrix} d_i/D \\ \cos(\chi_i - \psi_i) \\ \sin(\chi_i - \psi_i) \\ \bar{v}_i \\ \bar{v}_{ei} \end{bmatrix}^T,$$

где d_i – кратчайшее расстояние до точки выхода из зоны наблюдения, D – нормирующий коэффициент, ψ_i – угол пеленга к точке выхода, нормализованная скорость и отклонение скорости определяются следующим образом:

$$\bar{v}_i = \frac{v_i - v_{\min_i}}{v_{\max_i} - v_{\min_i}}; \quad \bar{v}_{ei} = \frac{v_i - v_{opt_i}}{v_{\max_i} - v_{\min_i}}.$$

На каждом шаге агент кодирует свое состояние o_i в скрытое состояние с помощью нейронной сети:

$$h_i^{(0)} = f_h(o_i).$$

Затем начинается фаза передачи сообщений, состоящая из C раундов связи. На каждом раунде $c = 0, 1, \dots, C - 1$ сообщение каждого агента вычисляется как взвешенная сумма ребер, которые соединяют его с соседями, на основе механизма внимания:

$$m_i^{(c+1)} = \sum_{j \in N_i} a_{ij}^{(c+1)} e_{ij}^{(c+1)},$$

где N_i – множество узлов, связанных ребрами с узлом i .

Значения ребер рассчитываются с помощью нейронной сети с учетом скрытых состояний агентов:

$$e_{ij}^{(c+1)} = f_e^{(c)} \left(\left[h_i^{(c)}, h_j^{(c)}, e_{ij}^{(c)} \right] \right).$$

Веса внимания вычисляются по формуле

$$a_{ij}^{(c+1)} = \frac{\exp \left(v_a^{(c)} f_a^{(c)} \left(\left[h_i^{(c)}, h_j^{(c)}, e_{ij}^{(c)} \right] \right) \right)}{\sum_{j \in N_i} \exp \left(v_a^{(c)} f_a^{(c)} \left(\left[h_i^{(c)}, h_j^{(c)}, e_{ij}^{(c)} \right] \right) \right)},$$

где $v_a^{(c)}$ – вектор параметров.

Затем каждый узел обновляет свое состояние с помощью функции обновления:

$$h_i^{(c+1)} = U^{(c)} \left(h_i^{(c)}, m_i^{(c+1)} \right).$$

После C раундов связи между узлами генерируется распределение вероятностей по всем возможным действиям для каждого агента:

$$a_i = f_a \left(\left[h_i^{(0)}, h_i^{(C)} \right] \right).$$

Ожидаемая награда, одинаковая для всех агентов, вычисляется при помощи функции считывания:

$$V^\pi = f_v \left(\sum_{i \in N} f_y \left(\left[h_i^{(0)}, h_i^{(C)} \right] \right) \right),$$

$f_h, f_e^{(c)}, f_a^{(c)}, f_a, f_v, f_y$ – нейронные сети прямого распространения.

На каждом временном шаге каждый агент выбирает действие a_i , после этого среда дает коллективное вознаграждение команде.

В результате экспериментов по обучению представленной модели было показано, что общая награда за эпизод увеличивается, а количество ожидаемых конфликтов уменьшается с ростом числа эпизодов, т.е. агенты могут улучшать свою политику на основе взаимодействия со средой.

5. Использование глубокого обучения с подкреплением на основе графов для предупреждения конфликтов

Во многих исследованиях в многоагентной формулировке векторы состояний нескольких ВС объединяются в многомерный вектор совокупного состояния с помощью операции конкатенации [30, 84]. Однако агрегирование состояний всех ближайших соседних ВС, независимо от наличия потенциальных конфликтов между ними, может привести к обработке избыточных данных и снижению эффективности модели. Такие векторы не могут инкапсулировать пространственно-временную динамику и различать разные уровни риска и срочности в сценариях конфликтов. Обучение с подкреплением на основе графов (Graph RL) предназначено для обработки данных, структурированных в виде графов [85]. Использование присущих графам свойств позволяет улучшить масштабируемость, эффективность и адаптивность при работе с многомерными и динамическими средами [86]. Обучение с подкреплением на основе графов позволяет использовать свойства графов для представления взаимосвязей между самолетами [87–89]. Методы глубокого обучения с подкреплением на основе графов инвариантны к порядку и количеству самолетов.

В [49] предлагается метод глубокого обучения с подкреплением на основе графов при управлении воздушным движением в трехмерном воздушном пространстве. Для предупреждения конфликтов подбираются высота, курс и скорость ВС. Самолеты представляются вершинами графа, узлы в этом графе соединены, если расстояние между парой самолетов ниже определенного порога. Сравняются два подхода: графовые нейронные сети со сверточными слоями (GCN) [90] и графовые нейронные сети с механизмом внимания (GAT – Graph Attention Network) [91], который используется для эффективного агрегирования информации из соседних узлов в графе. При нормальной плотности движения модель с механизмом внимания способна предотвратить 100% потенциальных столкновений и 89,8% потенциальных конфликтов. Однако производительность ухудшается при увеличении плотности движения.

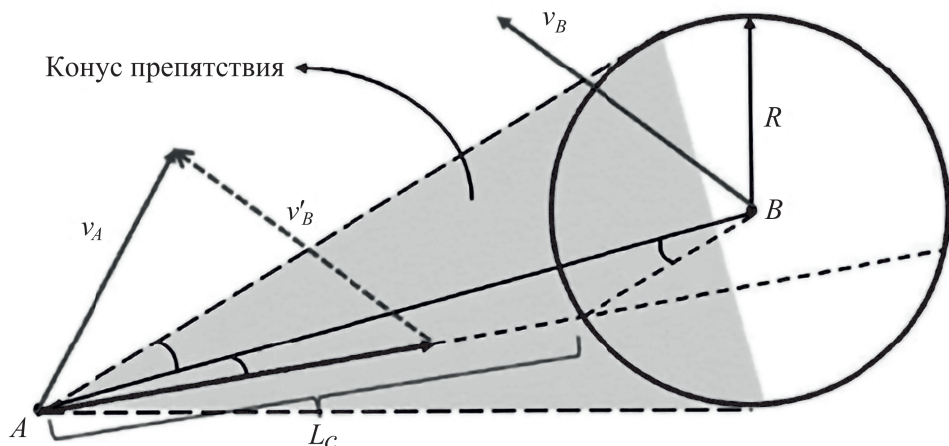


Рис. 5. Геометрия конфликта.

При увеличении плотности потока движения оба метода испытывают трудности с преодолением перегруженного воздушного пространства.

В [92] для сбора пространственно-временных зависимостей полетных данных используется графовая сверточная сеть с ячейками LSTM, а для концентрации на информационных характеристиках ключевых узлов применяется графовая нейронная сеть с повышенным вниманием.

В [93] предложено использовать развивающийся во времени граф конфликтов, в котором самолеты представляются узлами, а связи между ними обозначают срочность конфликта. Срочность конфликта определяется временем до конфликта при сохранении самолетами, находящимися в точках A и B (рис. 5), текущего курса и скорости. Наблюдение ведется из точки A . Точка B является центром окружности с радиусом R (радиус защитной зоны). Из точки A проведены касательные к этой окружности, образующие конус препятствия; v_A и v_B – скорости самолетов A и B . Потенциальный конфликт существует, если вектор относительной скорости $v_r = v_A - v_B$ находится внутри конуса препятствия. Время до конфликта определяется на основе геометрической модели относительного положения и скоростей пары самолетов A и B :

$$t_c = \frac{L_c}{|v_r|},$$

где L_c – отрезок прямой от точки A вдоль относительной скорости v_r до точки пересечения с защитной зоной самолета B . Вес ребра графа конфликта ω_{AB} нормализуется в диапазоне $[0, 1]$ следующим образом:

$$\omega_{AB} = e^{-t_c}.$$

Если потенциального конфликта не существует, то $\omega_{AB} = 0$, если самолеты столкнулись, то $\omega_{AB} = 1$.

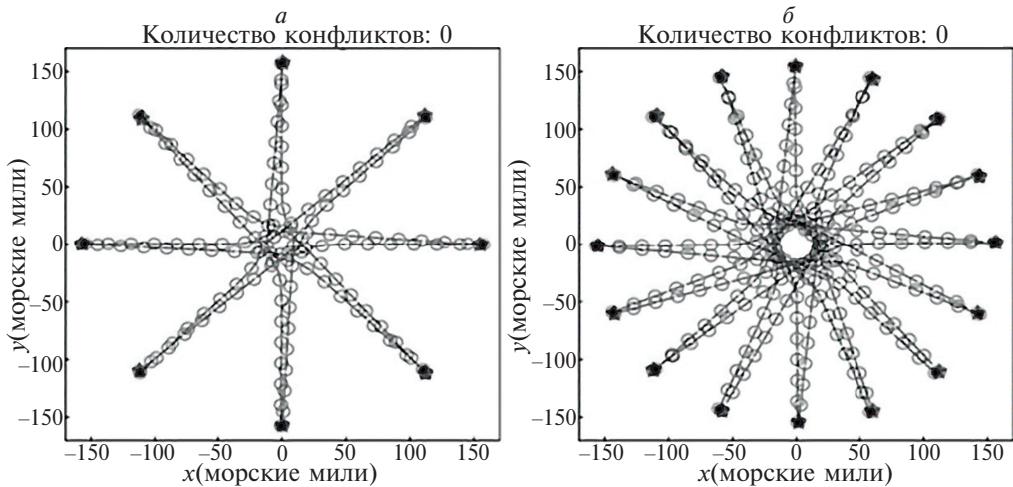


Рис. 6. Сценарий кругового конфликта.

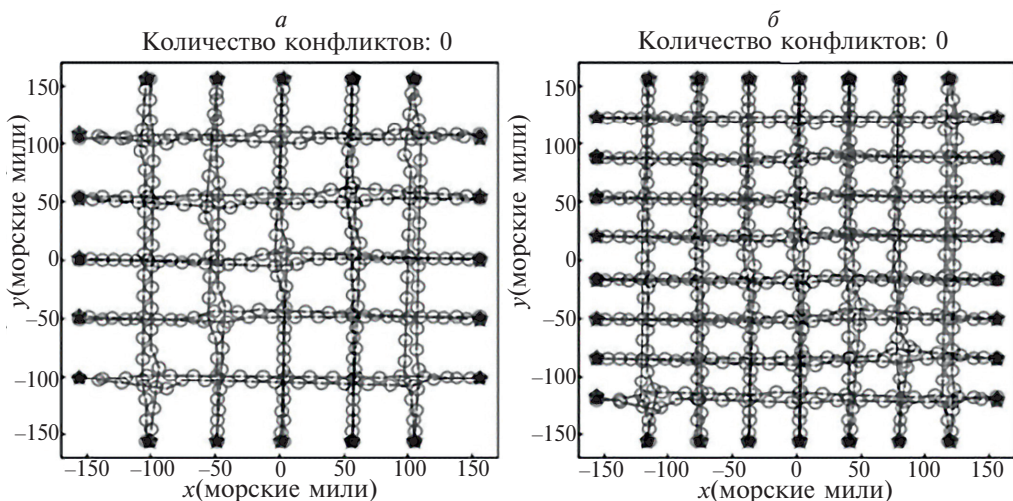


Рис. 7. Сценарий конфликта пересечения.

Затем на основе графа конфликта информация агрегируется при помощи многоголовой нейронной сети внимания. Для повышения стабильности обучения используется механизм временной регуляризации. Эффективность предложенного алгоритма демонстрируется в том числе на двух наглядных сценариях [93].

1) Сценарий кругового конфликта: в этой настройке самолеты начинают полет на точках окружности радиусом 160 морских миль и летят в противоположных направлениях. Такая конфигурация приводит к тому, что каждый самолет конфликтует со всеми остальными в центре окружности. Экспериментальные результаты для сценариев с участием 8 и 16 самолетов показаны на рис. 6, *a* и 6, *б* соответственно. Эти результаты демонстрируют способность

предложенного метода управлять характерными круговыми потенциальными конфликтами и предупреждать все потенциальные столкновения.

2) Сценарий конфликта пересечения: самолеты делятся на две группы, каждая из которых содержит равное количество самолетов. Эти группы летят по пересекающимся траекториям, вызывая конфликты на каждом пересечении. Проведено тестирование с 20 и 30 самолетами. Как видно на рис. 7,а и 7,б, для 20 и 30 самолетов соответственно, с помощью предложенного метода определяются бесконфликтные точки маршрутов на каждом временном шаге с минимальными отклонениями от исходных траекторий.

Несмотря на значительные успехи исследований, имеются серьезные препятствия для практического применения методов обучения с подкреплением в сфере воздушного транспорта из-за строгих стандартов сертификации в авиационной отрасли. Существующая нормативно-правовая база не обеспечивает адекватных и приемлемых средств соответствия требованиям для приложений обучения с подкреплением, и, таким образом, пока не существует правовой базы для их безопасного использования. Необходимо разработать рекомендации по сертификации моделей обучения с подкреплением, предназначенных для воздушного транспорта, чтобы эти многообещающие методы стало возможно использовать в реальных условиях [94].

6. Заключение

Из обзора литературы следует, что методы предупреждения конфликтов в воздушном движении на основе глубокого обучения с подкреплением направлены на решение двух основных задач: автоматическая генерация вариантов решений для поддержки авиадиспетчеров при централизованном управлении воздушным движением и поддержка систем автономного разрешения конфликтов в свободном полете. Для поддержки принятия решений авиадиспетчерами предлагаются, главным образом, модели с дискретными действиями. Модели с непрерывными действиями предназначены для автономного разрешения конфликтов при свободном полете и позволяют всем агентам выполнять действия по корректировке траекторий на каждом временном шаге.

Наиболее перспективным представляется подход обучения с подкреплением на основе графов, поскольку представление информации в виде развивающегося во времени графа конфликта позволяет сократить объем обрабатываемой избыточной информации и обеспечивает масштабируемость моделей на различное количество ВС. Механизм внимания позволяет выделить наиболее срочную информацию, содержащуюся в графе конфликта, обеспечивая улучшенные с точки зрения безопасности и эффективности стратегии предупреждения конфликтов.

По мере развития исследований потенциал практического применения методов предупреждения конфликтов между ВС на основе глубокого обучения с подкреплением становится все более очевидным. Обзор литературы свидетельствует, что исследуемые методы обучения с подкреплением показывают

многообещающие результаты за счет адаптивного принятия решений в режиме реального времени для предупреждения конфликтов в воздушном движении. Однако для их применения на практике управления воздушным движением, где критически важна безопасность, еще существуют серьезные нерешенные проблемы. Обучать модели в реальных условиях нельзя из-за потенциального ущерба, при этом идеально имитировать реальность невозможно. Сдвиг распределения между смоделированной средой и реальностью может ограничивать эффективность моделей обучения с подкреплением. Сертификация таких моделей должна стать одним из направлений исследований в этой области.

СПИСОК ЛИТЕРАТУРЫ

1. *International Civil Aviation Association*. Doc 4444: Air Traffic Management // Procedures for Air Navigation Services, 16th ed. ICAO: Montreal, QC, Canada. 2016.
2. *Кулида Е.Л., Лебедев В.Г.* Методы решения задач планирования и регулирования потоков воздушного движения. Ч. 1 // Проблемы управления. 2023. № 1. С. 3–14. <https://doi.org/10.25728/pu.2023.1.1>
Kulida E.L., Lebedev V.G. Methods for Solving Some Problems of Air Traffic Planning and Regulation. PART I: Strategic Planning of 4D Trajectories // Control Sciences. 2023. No. 1. P. 2–11. <https://doi.org/10.25728/cs.2023.1.1>
3. *Brittain M., Wei P.* Autonomous Separation Assurance in an High-Density En Route Sector: a Deep Multi-Agent Reinforcement Learning Approach // The 22nd IEEE Intelligent Transportation Systems Conference (ITSC). Auckland, New Zealand, 2019. <https://doi.org/10.109/ITSC.2019.8917217>
4. *Пономарев К.Ю.* Метод оценки динамической воздушной обстановки на конфликтность посредством полихромного отображения объектов в информационном обеспечении диспетчера управления воздушным движением // Автореф. дис. канд. техн. наук. СПб.: ФГБОУ ВО СПбГУ ГА им. А.А. Новикова, 2023. 24 с.
5. *Erzberger H.* Automated Conflict Resolution for Air Traffic Control // 25th International Congress of the Aeronautical Sciences (ICAS). Germany, Hamburg. 2006.
6. *Farley T., Field M., Erzberger H.* Fast-time Simulation Evaluation of a Conflict Resolution Algorithm Under High Air Traffic Demand. 2007. <https://www.researchgate.net/publication/255062615>
7. *Erzberger H., Heere K.* 2010. Algorithm and Operational Concept for Resolving Short-range Conflicts // Proc. Inst. Mechan. Engin., Part G: J. Aerospac. Engin. 2010. V. 224. No. 2. P. 225–243. <https://doi.org/10.1243/09544100JAERO546>
8. *Hoekstra J.M., van Gent R.N.H.W., Ruigrok R.C.J.* Designing for Safety: The ‘free flight’ Air Traffic Management Concept // Reliab. Engin. Syst. Safet. 2002. V. 75. No. 2. P. 215–232. [https://doi.org/10.1016/S0951-8320\(01\)00096-5](https://doi.org/10.1016/S0951-8320(01)00096-5)
9. *Clari M.V., Ruigrok R.C.J., Hoekstra J.M., Visser H.G.* Cost-benefit Study of Free Flight with Airborne Separation Assurance // Air Traffic Control Quarterly. 2001. V. 9. No. 4. P. 287–309. <https://doi.org/10.2514/atcq.9.4.287>
10. *Марьин Н.П.* Перспектива внедрения концепции «свободный полет» // Проблемы безопасности полетов. 2009. № 5. С. 42–55.

11. *Yang Y., Zhang J., Cai K., Prandini M.* Multi-aircraft Conflict Detection and Resolution Based on Probabilistic Reach Sets // *IEEE Transactions on Control Systems Technology*. 2016. V. 25. No. 1. P. 309–316. <https://doi.org/10.1109/TCST.2016.2542046>
12. *Орлов В.С.* Разработка и исследование алгоритмов обнаружения и предотвращения опасных сближений в воздухе в рамках перспективной системы ОрВД // *Дисс... канд. техн. наук. М. ФГБОУ ВПО «Московский авиационный институт (национальный исследовательский университет)»*. 2015. 116 с.
13. *Буржин В.С.* Синтез алгоритмов обнаружения и разрешения конфликтов-столкновений воздушных судов по информации системы автоматического зависимого наблюдения в условиях неопределенности // *Изв. РАН. Теория и системы управления*. 2017. № 3. С. 157–169. <https://doi.org/10.7868/S0002338817030088>
14. *Кумков С.И., Пятко С.Г.* Быстрые алгоритмы обнаружения конфликтных ситуаций между воздушными судами / «Теория оптимального управления и приложения (ОСТА 2022)». Материалы международной конференции. Ин-т мат. и механики им. Н.Н. Красовского (ИММ УрО РАН). Екатеринбург. 2022. С. 126–131.
15. *Pelegri M., D'Ambrosio C.* Aircraft Deconfliction via Mathematical Programming: Review and Insights // *Transportation Science*. 2022. V. 56. No. 1. P. 118–140. <https://doi.org/10.1287/trsc.2021.1056>
16. *Cafieri S., Conn A.R., Mongeau M.* Mixed-integer Nonlinear and Continuous Optimization Formulations for Aircraft Conflict Avoidance via Heading and Speed Deviations // *Eur. J. Oper. Res.* 2023. V. 310. No. 2. P. 670–679. <https://doi.org/10.1016/j.ejor.2023.03.002>
17. *Dias F., Rey D.* Aircraft Conflict Resolution with Trajectory Recovery Using Mixed-integer Programming // *J. Global Optim.* 2024. V. 90. P. 1031–1067. <https://doi.org/10.1007/s10898-024-01393-1>
18. *Cecen R.K., Cetek C.* Conflict-free En-route Operations with Horizontal Resolution Manoeuvres Using a Heuristic Algorithm // *Aeronaut. J.* 2020. V. 124. P. 767–785. <https://doi.org/10.1017/aer.2020.5>
19. *Eby M.* A Self-organizational Approach for Resolving Air Traffic Conflicts // *Lincoln Lab. J.* 1994. V. 7. No. 2. P. 239–254.
20. *Balasooryan S.* Multi-aircraft Conflict Resolution Using Velocity Obstacles // *MSc thesis. Delft University of Technology*. 2017. 126 с.
21. *Durand N.* Constant Speed Optimal Reciprocal Collision Avoidance // *Transportation Research. Part C, Emerging Technologies*. 2018. P. 366–379. <https://doi.org/10.1016/j.trc.2018.10.004>
22. *Pan W., Qin L., He Q., Huang Y.* Three-Dimensional Flight Conflict Detection and Resolution Based on Particle Swarm Optimization // *Aerospace*. 2023. V. 10. No. 9. <https://doi.org/10.3390/aerospace10090740>
23. *Sui D., Zhang K.* A Tactical Conflict Detection and Resolution Method for En Route Conflicts in Trajectory-Based Operations // *J. Advanc. Transport.* 2022. No. 2. P. 1–16. <https://doi.org/10.1155/2022/9283143>
24. *Brittain M., Wei P.* Scalable Autonomous Separation Assurance with Heterogeneous Multi-agent Reinforcement Learning // *IEEE Transactions on Automation Science and Engineering*. 2022. V. 19. No. 4. P. 2837–2848. <https://doi.org/10.1109/TASE.2022.3151607>

25. *Самойлов В.А., Дюенко Д.В.* Возможность применения нейронных сетей для поиска и решения потенциальных конфликтных ситуаций между воздушными судами при полетах в верхнем воздушном пространстве / Транспорт России: проблемы и перспективы – 2022. Материалы Международной научно-практической конференции. Институт проблем транспорта им. Н.С. Соломенко РАН. Санкт-Петербург. 2022. С. 180–184.
26. *Wang Z., Liang M., Delahaye D.* Data-driven Conflict Detection Enhancement in 3D Airspace with Machine Learning // 2020 International Conference on Artificial Intelligence and Data Analytics for Air Transportation (AIDA-AT). Singapore. 2020. <https://doi.org/10.1109/AIDA-AT48540.2020.904.9180>
27. *Pinto Neto E.C., Baum D., Almeida J.R., et. al.* Deep Learning in Air Traffic Management (ATM): Applications, Opportunities, and Open Challenges // Aerospace 2023. V. 10. No. 4. <https://doi.org/10.3390/aerospace10040358>
28. *Razzaghi P., Tabrizian A., Guo W., et. al.* A Survey on Reinforcement Learning in Aviation Applications // Engineering Applications of Artificial Intelligence. 2024. V. 136. No. 3. <https://doi.org/10.1016/j.engappai.2024.108911>
29. *Кулида Е.Л., Лебедев В.Г.* Методы решения задач планирования и регулирования потоков воздушного движения Ч. 2. // Проблемы управления. 2023. № 2. С. 3–18. <https://doi.org/10.25728/пу.2023.2.2>
Kulida E.L., Lebedev V.G. Methods for Solving Some Problems of Air Traffic Planning and Regulation. PART II: Application of Deep Reinforcement Learning // Control Sciences. 2023. No. 2. P. 2–14. <https://doi.org/10.25728/cs.2023.2.2>
30. *Brittain M.W., Wei P.* One to Any: Distributed Conflict Resolution with Deep Multi-agent Reinforcement Learning and Long Short-term Memory // AIAA Scitech 2021 Forum. Nashville, Tennessee, USA. <https://doi.org/10.2514/6.2021-1952>
31. *Wang Z., Pan W., Li H., et. al.* Review of Deep Reinforcement Learning Approaches for Conflict Resolution in Air Traffic Control // Aerospace. 2022. V. 9. No. 6. <https://doi.org/10.3390/aerospace9060294>
32. *Groot J., Ribeiro M., Ellerbroek J., et. al.* Improving Safety of Vertical Manoeuvres in a Layered Airspace with Deep Reinforcement Learning // International Conference on Research in Air Transportation (ICRAT). Tampa, Florida, USA. 2022. P. 19–23.
33. *Ribeiro M., Ellerbroek J., Hoekstra J.* Review of Conflict Resolution Methods for Manned and Unmanned Aviation // Aerospace. 2020. V. 7. No. 6. <https://doi.org/10.3390/aerospace7060079>
34. *Ribeiro M., Ellerbroek J., Hoekstra J.* Distributed Conflict Resolution at High Traffic Densities with Reinforcement Learning // Aerospace. 2022. V. 9. No. 9. <https://doi.org/10.3390/aerospace9090472>
35. *Ribeiro M.* Conflict Resolution at High Traffic Densities with Reinforcement Learning // Thesis for PhD. 2023. <https://doi.org/10.4233/uuid:a2979919-cb01-41d1-bbba-fefa9079463b>
36. *Ribeiro M., Ellerbroek J., Hoekstra J.* Improving Algorithm Conflict Resolution Manoeuvres with Reinforcement Learning // Aerospace. 2022. V. 9. No. 12. <https://doi.org/10.3390/aerospace9120847>
37. *Визильтер Ю.В., Вишняков Б.В., Желтов С.Ю.* Современные технологии искусственного интеллекта и их применение в авиационных комплексах // XVI Всероссийская мультиконференция по проблемам управления (МКПУ–2023). Волгоград. Материалы мультиконференции в 4 т. Т. 3. С. 13–16.

38. *Sui D., Ma C., Wei C.* Tactical Conflict Solver Assisting Air Traffic Controllers Using Deep Reinforcement Learning // *Aerospace*. 2023. V. 10. No. 2. <https://doi.org/10.3390/aerospace10020182>
39. *Bastas A., Vouros G.* Data-Driven Modeling of Air Traffic Controllers' Policy to Resolve Conflicts // *Aerospace*. 2023. V. 10. No. 6. <https://doi.org/10.3390/aerospace10060557>
40. *Кулида Е.Л., Лебедев В.Г.* Проблемы при применении методов машинного обучения в авиации // Труды 16-й Международной конференции «Управление развитием крупномасштабных систем» (MLSD'2023, Москва). М.: ИПУ РАН, 2023. С. 1315–1320. <https://doi.org/10.25728/mlsd.2023.1315>
41. *Degas A., Islam M.R., Hurter C., et. al.* A Survey on Artificial Intelligence (AI) and eXplainable AI in Air Traffic Management: Current Trends and Development with Future Research Trajectory // *Appl. Sci.* 2022. V. 12. Iss. 3. <https://doi.org/10.3390/app12031295>
42. *Wang L., Yang H., Lin Y., et. al.* Enhancing Air Traffic Control: A Transparent Deep Reinforcement Learning Framework for Autonomous Conflict Resolution // *Expert Systems with Applications*. 2024. V. 260(2). <https://doi.org/10.06/j.eswa.2024.125389>
43. *Саттон Р.С., Барто Э.Г.* Обучение с подкреплением. М.: ДМК-Пресс, 2020.
44. *Грессер Л., Кенг В.Л.* Глубокое обучение с подкреплением: теория и практика на языке Python. СПб.: Питер, 2022.
45. *Моралес М.* Грокаем глубокое обучение с подкреплением. СПб.: Питер, 2023.
46. *Sui D., Ma C., Dong J.* Conflict Resolution Strategy Based on Deep Reinforcement Learning for Air Traffic Management // *Aviation*. 2023. V. 27. Iss. 3. P. 177–186. <https://doi.org/10.3846/aviation.2023.19720>
47. *Hasselt H.V.* Double q-learning // 24th Annual Conference on Neural Information Processing Systems. Vancouver, Canada. 2010. P. 2613–2621.
48. *Brittain M., Wei P.* Autonomous Aircraft Sequencing and Separation with Hierarchical Deep Reinforcement Learning // International Conference for Research in Air Transportation. Casteldeifeil, Spain. 2018.
49. *Mollinga J., Hoof H.* An Autonomous Free Airspace En-route Controller Using Deep Reinforcement Learning Techniques // 9th International Conference on Research in Air Transportation (ICRAT). Tampa, Florida, USA. 2020.
50. *Sui D., Xu W., Zhang K.* Study on the Resolution of Multi-aircraft Flight Conflicts Based on an IDQN // *Chin. J. Aeronaut.* 2021. V. 35. No. 11. P. 195–213. <https://doi.org/10.06/j.cja.2021.03.015>
51. *Li S., Egorov M., Kochenderfer M.J.* Optimizing Collision Avoidance in Dense Airspace Using Deep Reinforcement Learning // Proceedings of the 13th USA/Europe Air Traffic Management Research and Development Seminar. Vienna, Austria. 17–21 June 2019. <https://doi.org/10.48550/arXiv.1912.10146>
52. *Hermans M.C.* Towards Explainable Automation for Air Traffic Control Using Deep Q-Learning from Demonstrations and Reward Decomposition. Master's Thesis, Delft University of Technology, Delft, The Netherlands, May 2020.
53. *Pham D.-T., Tran N.P., Goh S.K., et. al.* Reinforcement Learning for Two-aircraft Conflict Resolution in the Presence of Uncertainty // IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF). Singapore. 2019. <https://doi.org/10.1109/RIVF.2019.8713624>

54. Wang Z., Li H., Wang J., et. al. Deep Reinforcement Learning Based Conflict Detection and Resolution in Air Traffic Control // IET Intelligent Transport Systems. 2019. V. 13. Iss. 6. P. 1041–1047. <https://doi.org/10.1049/iet-its.2018.5357>
55. Badea C.A., Groot J., Morfin Veytia A., et. al. Lateral and Vertical Air Traffic Control Under Uncertainty Using Reinforcement Learning // Proceedings of the 12th SESAR Innovation Days. Budapest, Hungary. 2022.
56. Pham D., Tran N., Alam S., et. al. A Machine Learning Approach for Conflict Resolution in Dense Traffic Scenarios with Uncertainties // 13th USA/Europe Air Traffic Management Research and Development Seminar. Vienne, Austria. Jun. 2019.
57. Wen H., Li H., Wang Z. Application of DDPG-based Collision Avoidance Algorithm in Air Traffic Control // 12nd International Symposium on Computational Intelligence and Design. Hangzhou, China. 2019. P. 130–133. <https://doi.org/10.1109/ISCID.2019.00036>
58. Pham D.-T., Tran P.N., Alam S., et. al. Deep Reinforcement Learning Based Path Stretch Vector Resolution in Dense Traffic with Uncertainties // Transportation Research, Part C. 2022. V. 135. No. 3. <https://doi.org/10.1016/j.trc.2021.103463>
59. Mukherjee A., Guleria Y., Alam S. Deep Reinforcement Learning for Air Traffic Conflict Resolution Under Traffic Uncertainties // 41st Digital Avionics Systems Conference (DASC). Portsmouth, USA. 2022. <https://doi.org/10.1109/DASC55683.2022.9925772>
60. Sunil E., Ellerbroek J., Hoekstra J.M. Camda: Capacity Assessment Method for Decentralized Air Traffic Control // International Conference on Research in Air Transportation (ICRAT). Barcelona, Spain. 2018. P. 26–29.
61. Dudoit A., Rimsa V., Bogdevicius M. Investigation of Aircraft Conflict Resolution Trajectories under Uncertainties // Sensors. 2024. V. 24. No. 18. <https://doi.org/10.3390/s24185877>
62. Busoniu L., Babuska R., De Schutter B. A Comprehensive Survey of Multi-agent Reinforcement Learning // IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews). 2008. V. 38. No. 2. P. 156–172. <https://doi.org/10.1109/TSMCC.2007.913919>
63. Tan M. Multi-agent Reinforcement Learning: Independent vs. Cooperative Agents // 10th International Conference on Machine Learning (ICML). 1993. P. 330–337. <https://doi.org/10.1016/B978-1-55860-307-3.50049-6>
64. Matignon L., Laurent G.J., Le Fort-Piat N. Independent Reinforcement Learners in Cooperative Markov Games: a Survey Regarding Coordination Problems // Knowledge Engineering Review. 2012. V. 27. No. 1. <https://doi.org/10.1017/S0269888912000057>
65. Everett M., Chen Y.F., How J.P. Motion Planning Among Dynamic, Decision-making Agents with Deep Reinforcement Learning // IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2018. P. 3052–3059. <https://doi.org/10.48550/arXiv.1805.01956>
66. Chen. Y., Hu M., Yang L., et. al. General Multi-agent Reinforcement Learning Integrating Adaptive Manoeuvre Strategy for Real-time Multi-aircraft Conflict Resolution // Transportation Research. Part C. Emerging Technologies. 2023. V. 151. <https://doi.org/10.1016/j.trc.2023.104125>

67. *Xu Q., Chen Z., Li F., et. al.* An Efficient Aircraft Conflict Detection and Resolution Method Based on an Improved Reinforcement Learning Framework // Int. J. Aerospac. Engin. 2023. V. 1. P. 1–16. <https://doi.org/10.1155/2023/6643903>
68. *Ghosh S., Laguna S., Lim S.H., et. al.* A Deep Ensemble Method for Multi-agent Reinforcement Learning: A Case Study on Air Traffic Control // 31st International Conference on Automated Planning and Scheduling. SuiGuangzhou, China. 2021. P. 468–476. <https://doi.org/10.1609/icaps.v31i1.15993>
69. *Chen Y., Xu Y., Yang L., et. al.* General real-time three-dimensional multi-aircraft conflict resolution method using multi-agent reinforcement learning // Transportation Research. Part C. Emerging Technologies. V. 157. <https://doi.org/10.1016/j.trc.2023.104367>
70. *Hessel M., Modayil J., Van Hasselt H., et. al.* Rainbow: Combining Improvements in Deep Reinforcement Learning // Thirty-Second AAAI Conference on Artificial Intelligence. 2018. V. 32. No. 3. <https://doi.org/10.1609/aaai.v32i1.11796>
71. *Nilsson J., Unger J., Eilertsen G.* Self-Prioritizing Multi-Agent Reinforcement Learning for Conflict Resolution in Air Traffic Control with Limited Instructions // Aerospace. 2025. V. 12. No. 2. <https://doi.org/10.3390/aerospace12020088>
72. *Sui D., Zhou Z., Cui X.* Priority-based Intelligent Resolution Method of Multi-aircraft Flight Conflicts // Aeronaut. J. 2024. V. 129. P. 326–350. <https://doi.org/10.1017/aer.2024.75>
73. *Zhao P., Liu Y.* Physics Informed Deep Reinforcement Learning for Aircraft Conflict Resolution // IEEE Transactions on Intelligent Transportation Systems. 2021. V. 23. No. 7. P. 8288–8301.
74. *Hochreiter S., Schmidhuber J.* Long Short-Term Memory // Neural Computation. 1997. V. 9. No. 8. P. 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
75. *Cho K., Merriënboer B., Gulcehre C., et. al.* Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation // Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar. 2014. P. 1724–1734. <https://doi.org/10.3115/v1/D14-1179>
76. *Brittain M.W., Yang X., Wei P.* Autonomous Separation Assurance with Deep Multi-agent Reinforcement Learning // J. Aerospac. Inform. Syst. 2021. V. 18. No. 12. P. 890–905. <https://doi.org/10.254/1.1010973>
77. *Hoekstra J.M., Ellerbroek J.* BlueSky ATC Simulator Project: an Open Data and Open Source Approach // 7th International Conference on Research in Air Transportation. Philadelphia, USA. 2016. V. 131. P. 132.
78. *Schulman J., Wolski F., Dhariwal P., et. al.* Proximal Policy Optimization Algorithms // ArXiv. 2017. <https://doi.org/10.48550/arXiv.1707.06347>
79. *Vaswani A., Shazeer N., Parmar N., et. al.* Attention is All You Need // 31st Conference on Neural Information Processing Systems (NIPS). Long Beach, USA. 2017. <https://doi.org/10.48550/arXiv.1706.03762>
80. *Dalmau R., Allard E.* Air Traffic Control Using Message Passing Neural Networks and Multi-agent Reinforcement Learning // Conference: SESAR Innovation Days (SID), Virtual Event. 2020. P. 7–10.
81. *Groot J., Ellerbroek J., Hoekstra J.* Using Relative State Transformer Models for Multi-Agent Reinforcement Learning in Air Traffic Control // Conference: SESAR Innovation days (SID). Seville, Spain. 2023.

82. *Wollkind S., Valasek J., Ioerger T.* Automated Conflict Resolution for Air Traffic Management Using Cooperative Multi-agent Negotiation // AIAA Guidance, Navigation, and Control Conference and Exhibit. 2004.
<https://doi.org/10.2514/6.2004-4992>
83. *Pritchett R., Genton A.* Negotiated Decentralized Aircraft Conflict Resolution // IEEE Transactions on Intelligent Transportation Systems. 2017. V. 19. No. 1. P. 81–91. <https://doi.org/10.1109/TITS.2017.2693820>
84. *Lai J., Cai K., Liu Z., et. al.* A Multi-agent Reinforcement Learning Approach for Conflict Resolution in Dense Traffic Scenarios // IEEE/AIAA 40th Digital Avionics Systems Conference (DASC). San Antonio, USA. 2021.
<https://doi.org/10.1109/DASC52595.2021.9594437>
85. *Wu Z., Pan S., Chen F., et. al.* A comprehensive Survey on Graph Neural Networks // 2019. <https://doi.org/10.48550/arXiv.1901.00596>
86. *Mendonca M., Ziviani A., Barreto A.* Graph-Based Skill Acquisition for Reinforcement Learning // ACM Computing Surveys (CSUR). 2019. V. 52. No. 1.
<https://doi.org/10.1145/3291045>
87. *Papadopoulos G., Bastas A., Vouros G.A., et. al.* Deep Reinforcement Learning in Service of Air Traffic Controllers to Resolve Tactical Conflicts // Expert Systems with Applications. 2024. V. 236. No. 1. <https://doi.org/10.1016/j.eswa.2023.121234>
88. *Isufaj R., Aranega Sebastia D., Piera M.A.* Toward Conflict Resolution with Deep Multi-agent Reinforcement Learning // J. Air Transport. 2022. V. 30. No. 3. P. 71–80. <https://doi.org/10.2514/1.DO296>
89. *Vouros G., Papadopoulos G., Bastas A., et. al.* Automating the Resolution of Flight Conflicts: Deep Reinforcement Learning in Service of Air Traffic Controllers // PAIS 2022. P. 72–85. <https://doi.org/10.48550/arXiv.2206.07403>
90. *Kipf T.N., Welling M.* Semi-supervised Classification with Graph Convolutional Networks // 2017. <https://doi.org/10.48550/arXiv.1609.02907>
91. *Velickovic P., Cucurull G., Casanova A., et. al.* Graph Attention Networks // 2018. <https://doi.org/10.48.550/arXiv.170.0903>
92. *Zhang Y., Xu S., Zhang L., et. al.* Short-term Multi-step-ahead Sector-Based Traffic Flow Prediction Based on the Attention-enhanced Graph Convolutional LSTM Network (AGC-LSTM) // Neural Computing and Applications. 2024.
<https://doi.org/10.1007/s00521-024-09827-3>
93. *Li Y., Zhang Y., Guo T., et. al.* Graph Reinforcement Learning for Multi-Aircraft Conflict Resolution // IEEE Transactions on Intelligent Vehicles. 2024.
<https://doi.org/10.1109/TIV.2024.3364652>
94. *Ribeiro M., Tseremoglou I., Santos B.* Certification of Reinforcement Learning Applications for Air Transport Operations Based on Criticality and Autonomy // AIAA Science and Technology Forum and Exposition. Orlando, Florida, USA. 2024.
<https://doi.org/10.2514/6.2024-1463>

Статья представлена к публикации членом редколлегии Н.И. Сельвесюком.

Поступила в редакцию 04.03.2025

После доработки 17.03.2025

Принята к публикации 25.03.2025

Управление в социально-экономических системах

- © 2025 г. СИНЬБИНЬ ЧЖАН, канд. техн. наук (xinbinzhang@hit.edu.cn)
(Харбинский политехнический университет, Харбин),
ЧЖО ЧЭНЬ, магистр (986052851@qq.com),
ХАОСЯН СУ, магистр (463099610@qq.com)
(Юго-Западный университет науки и технологий, Мянъян),
ЛИГО ТАНЬ, канд. техн. наук (tanliguo@hit.edu.cn)
(Харбинский политехнический университет, Харбин),
ХУНВЭЙ ЛЮ, канд. техн. наук (liuhongwei@swust.edu.cn),
МАНЬЛУ ЛЮ, канд. техн. наук (liumanlu@swust.edu.cn),
ЦЗЯНЬВЭНЬ ХО, канд. техн. наук (huojianwen2008@hotmail.com)
(Юго-Западный университет науки и технологий, Мянъян),
С.В. НОВИКОВА, д-р. техн. наук (SVNovikova@kai.ru)
(Казанский национальный исследовательский технический университет
им. А.Н. Туполева)

РАЗРАБОТКА КОНСТРУКЦИИ И СИСТЕМЫ ПЛАНИРОВАНИЯ ТРАЕКТОРИИ ГИБКОГО ЗМЕЕПОДОБНОГО РОБОТА-МАНИПУЛЯТОРА НА ОСНОВЕ ЭМПИРИЧЕСКОГО ПОДХОДА¹

Предложена новая конструкция пневматического гибкого манипулятора на основе последовательно-параллельной композитной структуры. Гибкий скелет устройства создан по принципу сочленений позвоночника змеи. Смежные позвонки снабжены расположенными с двух сторон воздушными резервуарами-суставами, соединенными параллельно с позвоночником. Пневматический сустав может быть согнут влево или вправо в плоскости движения робота путем изменения давления в пневмокамере. Разработана кинематическая модель, позволяющая рассчитать положение рабочего органа в зависимости от углов изгиба сочленений манипулятора, которые, в свою очередь, рассчитываются исходя из давления в пневмокамерах-суставах. Для разработанной конструкции предложен способ расчета планируемой траектории манипулятора на основе решения обратных задач кинематики. Приведены несколько примеров расчета оптимальных траекторий движения, отвечающих различным требованиям задачи. Результаты натуральных экспериментов показывают, что разработанный гибкий змееподобный манипулятор точно следует заранее спланированной траектории и демонстрирует широкие возможности перемещения на плоскости.

¹ Работа выполнена при поддержке Фонда Ключевой лаборатории робототехники для особых условий провинции Сычуань (21kftk04) и Фонда естественных наук провинции Хэйлуцзян (LN2023E037), Китай.

Ключевые слова: змееподобный манипулятор, гибкий манипулятор, мягкий манипулятор, пневматический привод, синтез системы планирования траектории, обратная задача кинематики, эмпирический подход, движение в ограниченном пространстве.

DOI: 10.31857/S0005231025090055, **EDN:** VMXHTP

1. Введение

С развитием наукоемких технологий конструкции изделий и устройств становятся, с одной стороны, все более сложными, а с другой – все более компактными. В результате их сборка, техническое обслуживание и другие операции часто должны выполняться в ограниченном пространстве сложной конфигурации. Для автоматизации рутинных процессов обслуживания подобных агрегатов требуется оборудование, способное подстраиваться под топологию пространства, избегать препятствия и выполнять функции обслуживания, не повреждая основную конструкцию.

Традиционные жесткие манипуляторы, например промышленные манипуляционные роботы, как правило, обладают способностью движения для удовлетворения разнообразных требований к конечной позе лишь в открытых пространствах. Однако относительно небольшое количество их сочленений и, соответственно, степеней свободы, существенно ограничивают их возможности при движении в узком ограниченном пространстве. Одним из способов решения данной проблемы является применение избыточных континуальных манипуляторов, которые обладают возможностями изгибаться при движении и обходить имеющиеся препятствия [1–7]. Например, в [1] представлен многозвенный кабельно-приводной манипулятор для автомобильной краскопультной установки, повышающий возможности движений манипулятора и его способность избегания препятствий. В [2, 3] использованы линейные или гибридные приводы воздушных шлангов для создания серии гибких манипуляторов, имитирующих хобот слона. В [4, 5] разработана серия гибких манипуляторов, способных обходить препятствия и достигать определенной точки в ограниченном пространстве. Континуальные манипуляторы повышают способность движения в трехмерном пространстве за счет увеличения степени свободы своих сочленений. Однако увеличение степеней свободы приводит к росту размеров и веса приводного механизма из-за использования дополнительных шлангов и кабелей, повышению сложности кинематических вычислений, что в свою очередь повышает сложность управления и вычислительные затраты. Все это в совокупности ограничивает их применение на практике. Кроме того, компоненты таких манипуляторов полностью состоят из жестких материалов, из-за чего могут возникнуть проблемы безопасности при их взаимодействии с окружающей средой.

В последние годы благодаря появлению новых пластичных материалов появляются так называемые мягкие манипуляторы, способные без труда проникать в сложные и узкие пространства за счет возможности активного и

пассивного изменения геометрической формы. При этом высокая податливость позволяет им безопасно контактировать с элементами окружающего пространства, в случае столкновений нанося минимальный ущерб как элементам среды, так и самим роботам. Мягкие манипуляторы имеют перспективы применения в промышленной сборке, уборке сельскохозяйственной продукции, малоинвазивной хирургии, бытовом обслуживании и других областях.

Мягкие манипуляторы часто имитируют бионические стратегии, наблюдаемые в природе. Например, в [8] используется имитация хобота слона, в [9] – движения червя, в [10–12] авторы имитируют способы фиксации выюнковых растений, в [13] – движение щупальцев осьминога и др. Подобные манипуляторы в качестве структурной опоры используют воздушные подушки, пневматические мышцы, пружины и др. и приводятся в действие с помощью пневматических приводов, сплавов с памятью формы или электроактивных полимеров [14–17]. Уникальные свойства мягких манипуляторов привлекают все большее внимание исследователей, стремящихся преодолеть ограничения существующих конструкций как жестких, так и гибких манипуляторов [18–27]. В качестве примера мягких манипуляторов можно привести континуальный манипулятор OctArm [18], который приводится в движение пневматическими мышцами и устанавливается на гусеничный мобильный робот. В [19] разработан мягкий манипулятор, использующий сотовую пневматическую сетку, которая обеспечивает стабильное управление движением и высокую грузоподъемность. В [20] авторы разработали мягкий манипулятор с использованием искусственных пневматических мышц, который имитирует движение хобота слона и обладает гибкими возможностями пространственного перемещения. В [21] описан мягкий манипулятор с жидкостными приводами, а авторы [22] собрали многосуставной мягкий манипулятор, в котором используются модули всенаправленного изгиба и кручения для эффективного перемещения в пространстве. В [23] предложен мягкий манипулятор для малоинвазивной хирургии, в котором всенаправленное изгибающее движение достигается за счет трех пневматических модулей. В [24] разработан самоорганизующийся мягкий манипулятор, созданный на основе биологических механизмов, таких как нейроны, пыльцевые трубки и толстостенные клетки тканей. Самоорганизующийся манипулятор, по описанию авторов, способен менять форму в зависимости от условий окружающей среды. В [25] разработан трехсуставной мягкий манипулятор, пригодный для работы под водой и позволяющий осуществлять точный захват предметов с помощью механизма противодействия изгибу и вытягиванию. В [26] предложен модульный мягкий робот-запястье с гидравлическим приводом, состоящий из вращающегося и изгибающегося модулей, способный выполнять тонкие и точные манипуляции в сложных условиях глубоководной среды. В [27] разработан мягкий манипулятор с сухожильным приводом, используя три модуля на основе шаблонов Креслинга, который может переключаться между двумя стабильными состояниями: гибкими суставами с низкой жесткостью на изгиб и жесткими звеньями с высокой жесткостью.

Несмотря на многочисленные исследования, посвященные мягким манипуляторам, эти работы все еще находятся в стадии становления, а их практическое применение сопряжено со многими трудностями. Одна из проблем заключается в том, что низкая жесткость мягких роботизированных систем снижает их управляемость и стабильность движения под действием гравитационных и инерционных сил, хотя их легко изменяемая геометрия и высокая пластичность позволяют им адаптироваться к сложным и узким пространствам. Одним из подходов к решению этой проблемы является разработка механизмов с изменяемой жесткостью, что может быть достигнуто за счет использования материалов с регулируемой жесткостью, таких как сплавы с низкой температурой плавления [28], или механизмов заклинивания [29, 30]. Другой подход заключается во внедрении гибких конструкций из твердых материалов для повышения жесткости робота в целом [31]. Первый подход позволяет изменять пластичность и геометрическую форму робота путем регулировки жесткости, однако использование механизмов с переменной жесткостью влияет не только на гибкость, присущую манипулятору, но и значительно усложняет конструкцию и управление. При втором подходе часть возможностей по изменению геометрии мягкого манипулятора будет потеряна.

В данной статье реализован второй подход к повышению управляемости и стабильности мягких манипуляторов. Разработан плоский гибкий змееподобный мягкий манипулятор, вдохновленный строением тела змеи, для движения в узком пространстве с препятствиями. Благодаря использованию жестко-гибкой композитной структуры несущей конструкции (позвоночника), манипулятор демонстрирует хорошую способность к изменению геометрической формы с одновременной стабильностью движения и простотой управления.

Для спроектированного змееподобного манипулятора необходимо разработать алгоритм расчета траектории, соответствующий особенностям движения предложенной конструкции. В настоящее время наибольшее распространение к планированию траектории манипуляторов получил подход, основанный на решении обратной задачи кинематики (Inverse Kinematics). Этот подход применяется как для жестких [32–34], так и мягких многозвенных манипуляторов [21, 35]. При этом обратная задача кинематики решается с использованием различных стратегий: в [33] расчет траектории опирается на решение кинематических дифференциальных уравнений, а в [32] решение предложено искать аналитически, без использования производных. Широкое распространение сегодня получают эвристические подходы. Так, в алгоритме FABRIK (Forward and Backward Reaching Inverse Kinematics) [36] координаты узлов многозвенного манипулятора вычисляются путем повторяющихся линейных сдвигов в сторону целевого положения рабочего органа. Алгоритмы расчета траектории без использования обратной кинематической задачи также описаны в литературе. Так, в [37] для расчета траектории используется специфическая модификация генетического алгоритма, в [38] авторы предлагают ис-

пользовать для планирования траектории аппарат нечеткой логики, а в [39] – гибридные нейро-нечеткие сети. Необходимо отметить, что подобные подходы, как правило, требуют больше вычислительных ресурсов и сложнее в реализации, чем алгоритмы, основанные на решении обратных кинематических задач. Предложенный в данной статье алгоритм расчета траектории базируется на обратном-кинематическом подходе. Кинематическая модель движения робота построена на основе сочетания теории и экспериментов. Прямая задача кинематики гибкого манипулятора формируется на основе результатов предварительно проведенных натуральных экспериментов, что позволяет адаптировать алгоритм под конкретный манипулятор с высокой точностью. Решение обратной задачи получено аналитически. Проведены натурные эксперименты по проверке подвижности гибкого манипулятора с предложенным алгоритмом планирования траектории, которые подтвердили его высокую эффективность при обходе препятствий.

2. Проектирование и сборка гибкого робота

Богатый диапазон движений и маневренность змей в сложной узкой среде можно объяснить уникальным строением их тела. В общем случае тело змеи представляет собой композитную структуру, состоящую как из жестких, так и гибких элементов. Надежной опорой тела является центральный позвоночник, покрытый мягкой мышечной тканью. Позвоночник змеи состоит из сотен позвонков, расположенных последовательно. Манипулируя мышцами в различных направлениях, соседние позвонки могут совершать горизонтальные колебания до 10 градусов и вертикальные до 3 градусов [40]. Хотя угол сгибания одного сустава ограничен, змея может осуществлять сильные изгибы всего тела за счет накопления множества сгибаний в отдельных суставах. Исходя из этого, манипулятор должен иметь аналогичную жестко-гибкую композитную структуру, состоящую из плоского гибкого позвоночника и приводных воздушных резервуаров.

Весь змееподобный манипулятор разделяется на шесть гибких секторов-суставов для независимого управления, при этом каждый гибкий сустав последовательно соединен с остальными, как показано на рис. 1. Каждый сектор-сустав, в свою очередь, состоит из шести последовательно соединенных небольших воздушных резервуаров-позвонков. Позвонки объединены в сустав с помощью плоских гибких шарниров, позволяющих сектору изгибаться в пределах плоскости сустава, при этом ограничивая продольный изгиб. Кроме того, на внешней стороне шарнира закреплены высокопрочные тросы. Их количество увеличивается от хвоста к голове, тем самым усиливая общую продольную прочность гибких суставов.

Между позвонками с обеих сторон расположены две пневмокамеры, выполняющие роль мышц для движения. Каждая пневмокамера содержит шесть внутренних воздушных резервуаров, что соответствует работе шести позвонков. Под действием изменения давления воздуха внутри пневмокамер

a

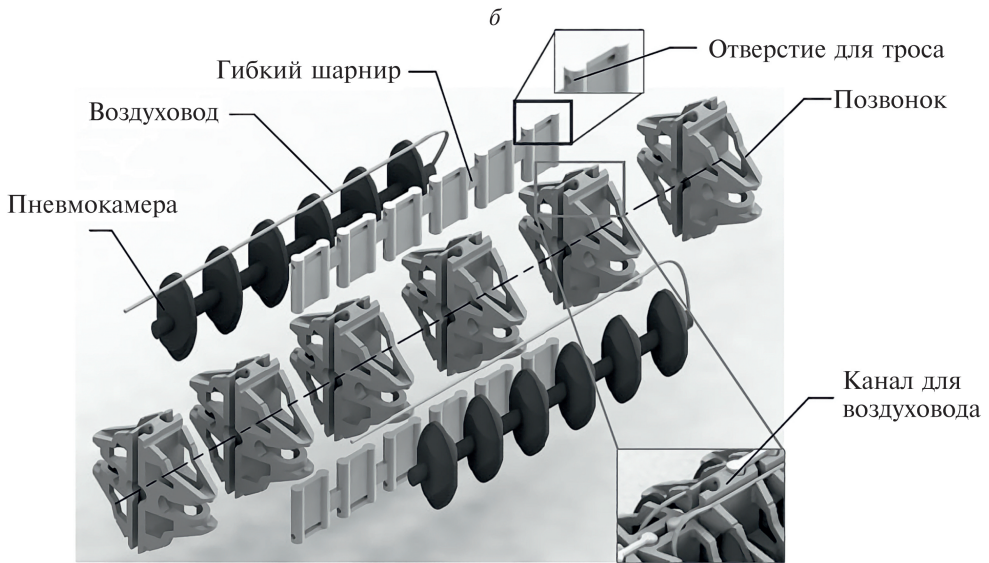
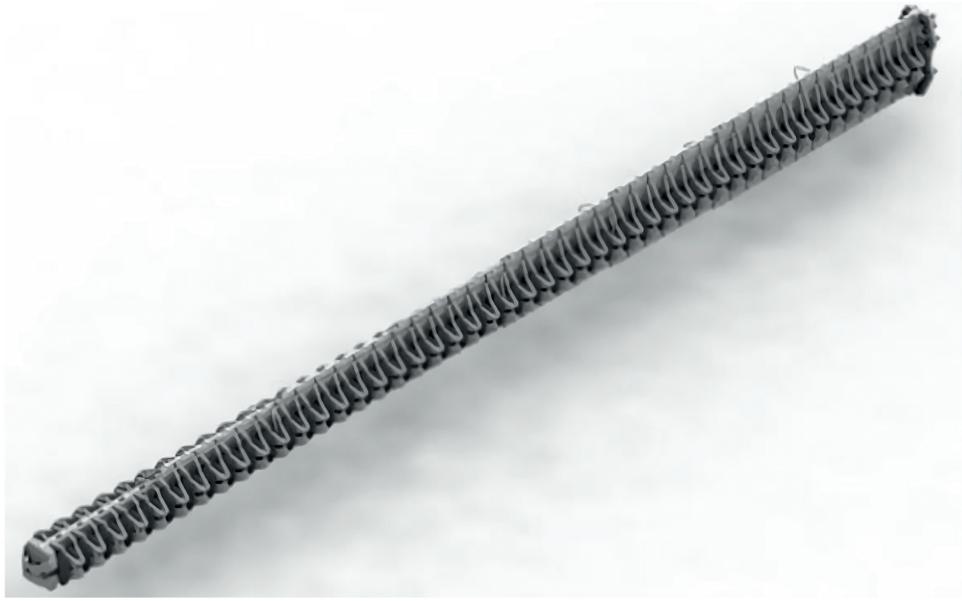


Рис. 1. Гибкий манипулятор в форме змеи: *a* – общий вид, *б* – структура гибкого сустава.

с обеих сторон конструкции они расширяются и удлиняются, тем самым регулируя угол изгиба всего гибкого сустава.

Для упрощения конструкции, за исключением конечного гибкого сустава, нижняя («брюшная») и верхняя («спинная») поверхности позвонков снабжены канавками. Глубина канавок увеличивается по мере продвижения от головы к хвосту робота, в результате чего манипулятор приобретает коническую биоинспирированную форму, напоминающую щупальце осьминога.

Для снижения веса и стоимости робота его позвоночник и гибкие шарниры изготавливались способом 3D-печати с использованием полилактидной кислоты (PLA) и термопластичного полиуретана (TPU) соответственно. Воздушные резервуары изготовлены с использованием силиконовой резины (E640), известной своей высокой эластичностью, прочностью на разрыв и удлинением до 470%. Воздушные резервуары изготавливались методом литья из силикона. Матрица для формовки резервуаров состоит из двух частей. Литье осуществляется путем вставки железного стержня в верхнюю полость для создания двух половин резервуара. Вакуумная сушильная печь создает пониженное давление для предотвращения образования пузырей внутри литых резервуаров в процессе их изготовления. Готовый змееподобный манипулятор состоит из 6 секторов-суставов. Общая длина робота составляет 900 мм, вес около 2 кг.

3. Построение траектории перемещения змееподобного манипулятора на основе эмпирического подхода

В данном разделе разрабатывается модель, позволяющая управлять положением конечного эффектора (рабочего органа) манипулятора в двумерном пространстве. Будем считать, что эффектор располагается на «голове» робота. Прямая кинематическая модель гибкого змееподобного манипулятора при этом включает в себя два основных компонента: зависимость положения конечного эффектора от угла изгиба каждого гибкого сустава и зависимость углов изгиба механических суставов от давления в управляющих пневмокамерах.

3.1. Зависимость положения конечного эффектора манипулятора от угла изгиба суставов

Модель движения отдельного сектора-сустава гибкого змееподобного манипулятора иллюстрирует рис. 2. Здесь представлена упрощенная модель с постоянной кривизной, в результате чего получается плавная непрерывная

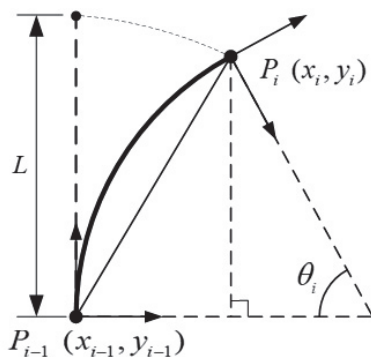


Рис. 2. Изгиб одного гибкого сустава.

кривая. Форма кривой определяется двумя переменными, связанными с гибким суставом: длиной сустава (L) и углом изгиба (θ_i).

На основе кинематической модели, представленной на рис. 2, можно определить матрицу преобразования T из начальной системы координат $P_{i-1}(x_{i-1}y_{i-1})$ в конечную систему координат $P_i(x_iy_i)$. Матрица T , таким образом, будет представлять собой кинематическое отображение из пространства положений суставов робота в пространство положения робота во внешнем пространстве. Начальная система координат $P_{i-1}(x_{i-1}y_{i-1})$ преобразуется в $P_i(x_iy_i)$ путем однократного смещения и однократного вращения:

1. Система координат $P_{i-1}(x_{i-1}y_{i-1})$ смещается вдоль двух координатных осей для перемещения точки P_{i-1} в точку P_i . На основе геометрических отношений получаем матрицу смещения

$$(1) \quad \text{Trans} \left(\frac{L}{\theta_i}(1 - \cos \theta_i), \frac{L}{\theta_i} \sin \theta_i \right) = \begin{bmatrix} 1 & 0 & \frac{L(1 - \cos \theta_i)}{\theta_i} \\ 0 & 1 & \frac{L \sin \theta_i}{\theta_i} \\ 0 & 0 & 1 \end{bmatrix};$$

2. Система координат вращается вокруг оси z на угол θ_i , матрица вращения имеет вид

$$(2) \quad \text{Rot}(z_i, \theta_i) = \begin{bmatrix} \cos \theta_i & \sin \theta_i & 0 \\ -\sin \theta_i & \cos \theta_i & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Перемножив эти две матрицы, получаем матрицу преобразования от начальной к конечной системе координат для одного сустава в виде

$$(3) \quad \begin{aligned} {}^{i-1}T_i &= \text{Trans} \left(\frac{L}{\theta_i}(1 - \cos \theta_i), \frac{L}{\theta_i} \sin \theta_i \right) \times \text{Rot}(z_i, \theta_i) = \\ &= \begin{bmatrix} \cos \theta_i & \sin \theta_i & -\frac{L(\cos \theta_i - 1)}{\theta_i} \\ \sin \theta_i & \cos \theta_i & \frac{L \sin \theta_i}{\theta_i} \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Так как манипулятор состоит из нескольких последовательно соединенных суставов с одной степенью свободы изгиба, то положение концевой эффектора зависит от угла изгиба каждого сустава. Если все суставы используют модель постоянной кривизны круговой дуги, то в соответствии с формулой преобразования координат (3) можно получить однородную матрицу преобразования от системы координат сустава $P_n(x_n, y_n)$ к базовой системе координат $P_0(x_0, y_0)$ для n сегментов:

$$(4) \quad {}^0T_n = {}^0T_1 \times {}^1T_2 \times \dots \times {}^{n-2}T_{n-1} \times {}^{n-1}T_n.$$

Кроме того, если манипулятор установлен на ползуне, который может перемещаться вдоль оси y , то матрица однородных преобразований из неподвижной системы координат в базовую систему координат имеет следующий вид:

$$(5) \quad {}^bT_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & y_b \\ 0 & 0 & 1 \end{bmatrix}.$$

В уравнении параметр y_b представляет собой расстояние, на которое перемещается ползун. Умножение матрицы bT_0 на матрицу преобразований bT_n дает матрицу однородных преобразований от неподвижного конца манипулятора к концевому эффектору манипулятора:

$$(6) \quad {}^bT_n = {}^bT_0 \times {}^0T_n \quad (n = 6).$$

При этом векторы положения и направления концевого эффектора задаются соотношениями

$$(7) \quad \begin{cases} P_n = {}^bT_n [0 \ 0 \ 1]^T, \\ D_n = {}^bT_n [0 \ 0 \ 1]^T - P_n. \end{cases}$$

3.2. Зависимость угла изгиба сустава от давления в управляющих пневмокамерах

При постоянной нагрузке существует однозначная зависимость между деформацией управляющих пневмокамер и давлением внутри них. В процессе работы угол изгиба сустава зависит от следующих факторов: давление, прикладываемое к воздушным резервуарам с обеих сторон позвоночника; физические характеристики расширения резервуаров; структура соединительных шарниров. Различное давление и разная степень расширения резервуаров в разных секторах-суставах приводит к различным углам изгиба. Уникальные характеристики изгиба каждого сустава гибкого змееподобного манипулятора определялись в ходе физических экспериментов.

Гистерезис – это свойство систем, мгновенный отклик которых на приложенные к ним воздействия зависит в том числе и от их текущего состояния. Гистерезис наблюдается в тех случаях, когда состояние системы в данный момент времени определяется внешними условиями не только в тот же, но и в предшествующие моменты времени. Для рассматриваемой задачи это означает, что на деформацию пневмокамеры влияют ее исторические состояния. Если в совместном движении участвуют две приводные пневмокамеры, то зависимость усложняется и решение задачи расчета угла изгиба сустава становится затруднительным. Поэтому в разработанном гибком манипуляторе используется метод управления одной пневмокамерой, который описывается следующим образом:

1) при увеличении угла изгиба сустава с одной стороны давление в пневмокамере с противоположной стороны увеличивается;

2) при уменьшении угла изгиба сустава с одной стороны пневмокамера с противоположной стороны разгерметизируется и сжимается.

Это означает, что одна пневмокамера приводит к одностороннему изгибу сустава. Этот метод управления устанавливает прямую зависимость между углом изгиба сустава и давлением в управляющей пневмокамере, что упрощает расчеты и облегчает последующее планирование траектории манипулятора.

Установка состоит из микроконтроллера STM32 для регулирования клапана, обеспечивающего управление давлением в воздушных камерах для вызова изгиба в змееподобном манипуляторе. На первом позвонке каждого сегмента установлен датчик угла изгиба высокой точности WT61C для измерения угла изгиба сустава. Эксперименты проводились при накачивании пневмокамеры с одной стороны, в то время как противоположная сторона оставалась ненагруженной. Диапазон давления воздуха устанавливался от 0 до 0,064 МПа. Было проведено несколько экспериментов, полученные кривые аппроксимировались полиномом четвертой степени. В результате были получены соотношения «давление–угол изгиба» для всех шести суставов, как показано на рис. 3. Результаты экспериментов с двумя зеркально расположенными пневмокамерами одного и того же сустава практически идентичны, поэтому в данной статье представлены результаты только для одной стороны робота. Расхождение между аппроксимирующей зависимостью «давление–угол изгиба» и фактическими экспериментальными значениями составляет менее 5%, отражая кинематические характеристики сустава с высокой точностью. Полученные кривые служат основой для решения прямых и обратных задач кинематики робота при расчете положения концевой эффектора манипулятора в зависимости от угла изгиба суставов.

Как видно из рис. 3, при одинаковых уровнях давления воздуха угол изгиба сустава в процессе нагнетания воздуха меньше по сравнению с процессом выкачивания. Максимальное угловое запаздывание составляет около 5 градусов, что свидетельствует о гистерезисных характеристиках пневмокамеры. Более того, из-за структурных отличий между гибкими суставами наблюдаются небольшие различия в соответствующих углах изгиба при одинаковом давлении воздуха с максимальным отклонением около 3 градусов.

Исходя из этого, при решении прямой (обратной) задачи кинематики между давлением воздуха и углом изгиба сустава необходимо сначала определить значение давления воздуха (угла изгиба сустава) на предыдущем сочленении, выбрать одну из двух кривых экспериментальных функций подгонки (для нагнетания либо для выкачивания воздуха) и затем получить решение.

Рабочее пространство робота включает пространство, которое может быть охвачено базисной точкой запястья (последнего звена) змееподобного манипулятора с установленным на нем концевым эффектором. Рабочее пространство является ключевым критерием для оценки динамической производительности манипулятора. Исходя из анализа предыдущего раздела диапазон

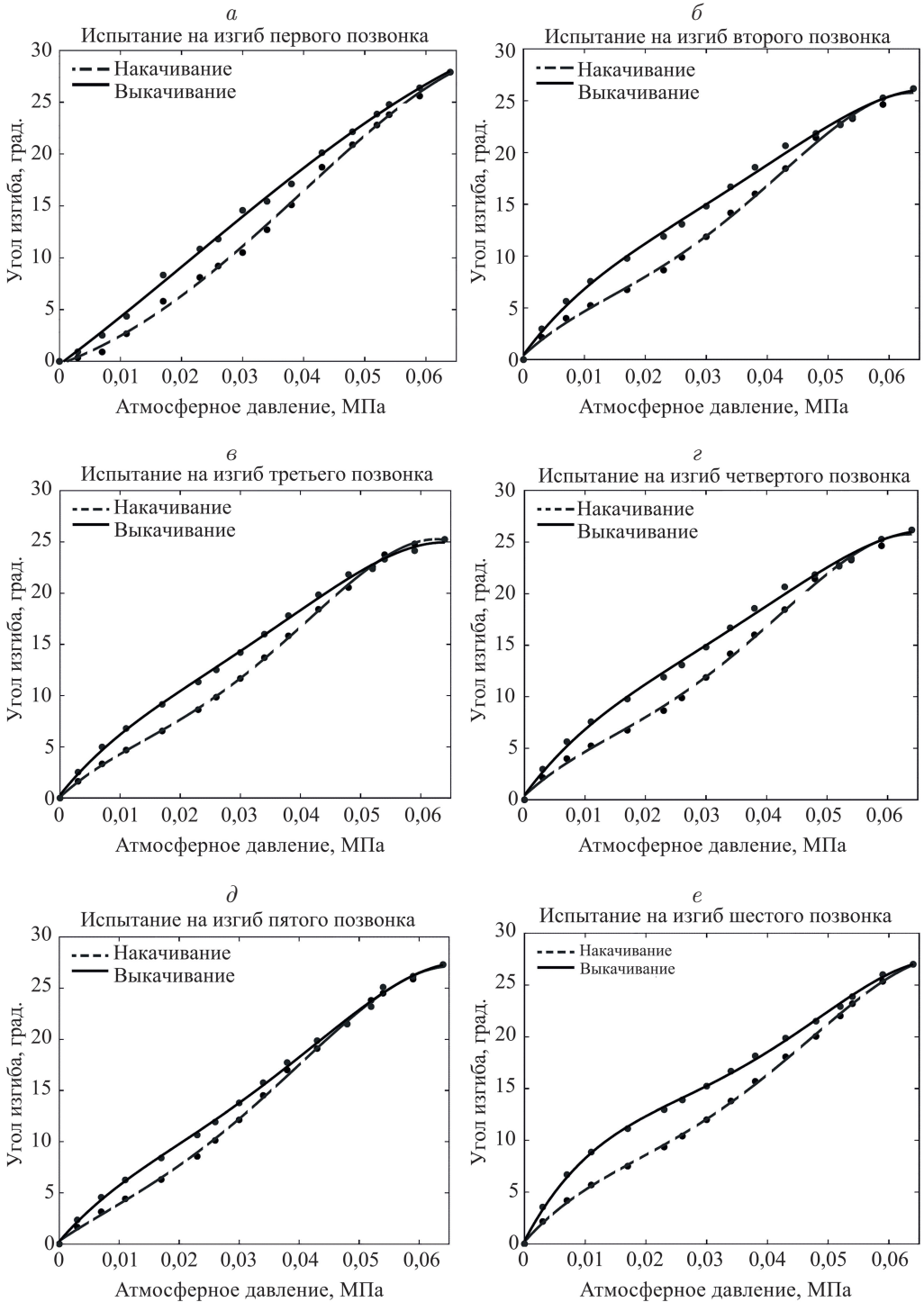


Рис. 3. Отображение зависимости между углами изгиба суставов и давлением в управляющей пневмокамере.

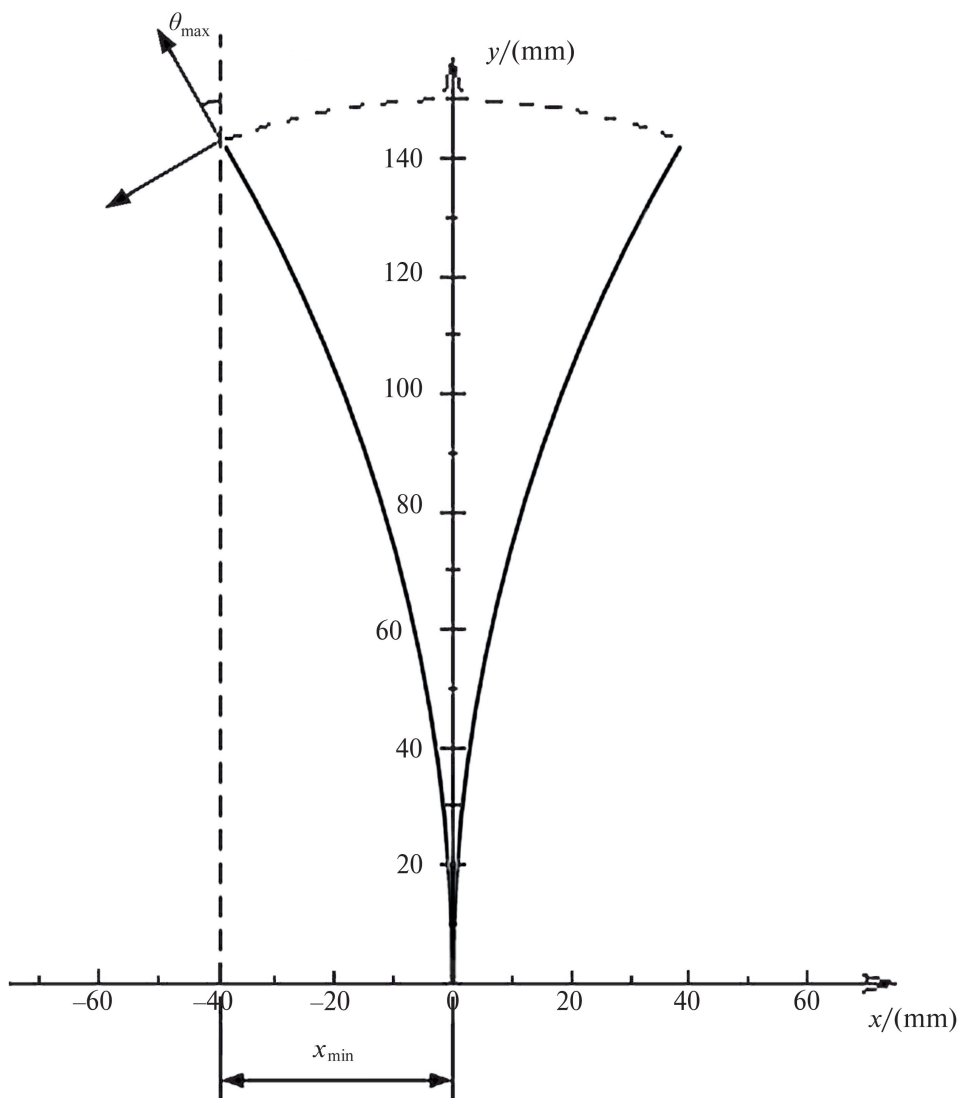


Рис. 4. Диапазон достижимости конечного сустава.

вращения гибких суставов составляет $[-30^\circ, 30^\circ]$, достижимый диапазон для конечного позвонка последнего сустава (головы) показан на рис. 4. Для моделирования рабочего пространства робота, перемещающегося на подвижном ползуне, был применен метод Монте-Карло с диапазоном для перемещения ползуна 1000 мм. В каждом диапазоне генерировалось 20 000 наборов случайных чисел для расчета положения конечного эффектора с использованием уравнений прямой кинематики. Результирующее рабочее пространство демонстрирует рис. 5.

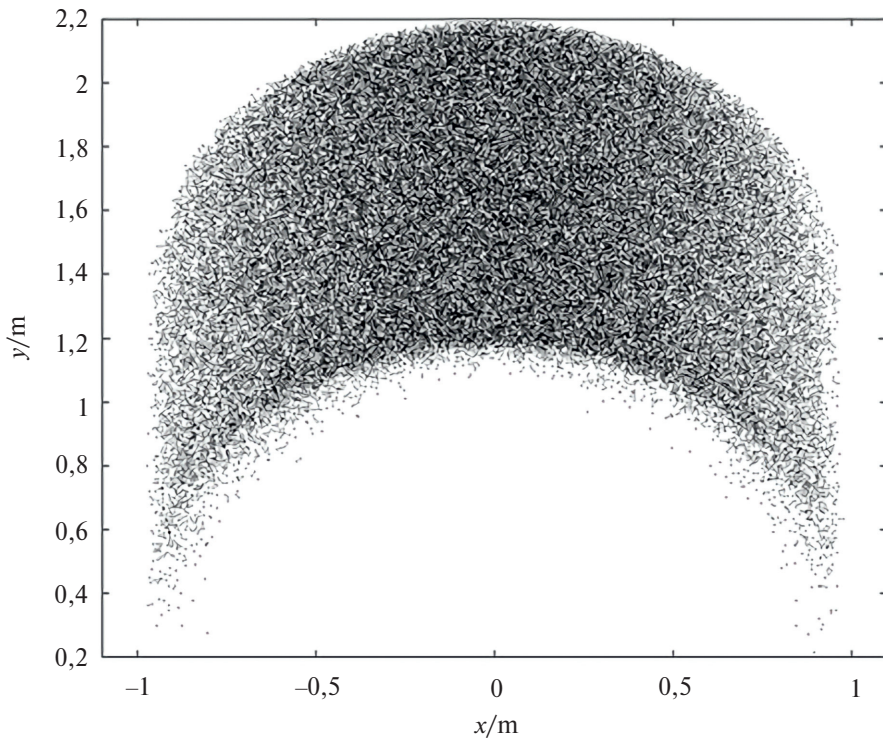


Рис. 5. Рабочее пространство гибкого змееподобного манипулятора.

3.3. Расчет углов изгиба последовательности суставов

Для решения обратной задачи кинематики по расчету потребного изгиба каждого сустава использовался прием равного изгиба суставов и компенсации линейного направления. Для перемещения конечного эффектора на величину x_l в направлении x (перпендикулярно направлению ползуна) и величину y_l в направлении y (вдоль направления ползуна) соответственно необходимо выполнить следующие шаги:

- 1) рассчитать требуемое смещение одного сустава в направлении x делением величины общего смещения x_l на общее количество суставов;
- 2) рассчитать угол изгиба сустава исходя из полученной величины смещения по оси x ;
- 3) рассчитать смещение одного сустава по оси y при изгибе на рассчитанную на предыдущем шаге величину;
- 4) рассчитать общее смещение конечного эффектора при изгибе по оси y путем умножения смещения одного сустава на их общее количество;
- 5) рассчитать разницу между смещением эффектора вследствие изгиба и требуемым смещением y_l ;
- 6) переместить направляющий ползун на полученное значение, компенсируя недостаток смещения робота по оси y .

4. Планирование траектории на базе эмпирической модели

4.1. Планирование на основе отслеживания положения концевого эффектора

С учетом большого количества суставов в пневматическом гибком манипуляторе индивидуальное планирование траектории с избеганием препятствий для каждого положения каждого сустава значительно снизит эффективность расчетов обратной кинематики. В данной работе применена стратегия слежения за головой, когда отслеживается только траектория конечного позвонка последнего сустава. Это значительно повышает эффективность планирования траектории, сохраняя гибкость манипулятора.

Для разработанного гибкого змееподобного манипулятора основной проблемой при использовании стратегии слежения за головой является определение фактической траектории движения рабочего органа. По сути, это позиция гибкого змееподобного манипулятора при достижении рабочей цели. При этом должны быть выполнены следующие условия:

1. Максимальная кривизна траектории должна быть меньше кривизны, соответствующей максимальному углу сгибания сустава.

2. Направляющий ползун должен являться касательной к траектории первого сустава манипулятора, а конец траектории должен достигать целевой точки.

3. Общая длина траектории является целым числом, кратным длине сустава.

4. Траектория не должна пересекать препятствия.

Исходя из этих требований, в данной статье рассматривается построение траекторий движения конечных эффекторов для трех различных сценариев задач.

4.1.1. Достижение указанной позиции и ориентации концевого эффектора

На основе моделирования и анализа из предыдущих разделов становится очевидно, что существует связь между положением конечных точек гибких секторов-суставов и углами изгибов суставов. При решении обратной задачи кинематики через интерполяцию пути существует вероятность столкнуться с точками, где решение отсутствует. Для преодоления данной проблемы в статье используется геометрический метод поиска пригодных путей, как показано на рис. 6. Метод состоит в следующем.

1. Поскольку целевая точка и направление движения определены заранее, целевую точку T можно принять в качестве начальной точки, а направление, противоположное целевому направлению $(-y_e)$, – в качестве начального направления. Поиск пути выполняется в обратном порядке, начиная с конечного эффектора и продвигаясь к основанию робота.

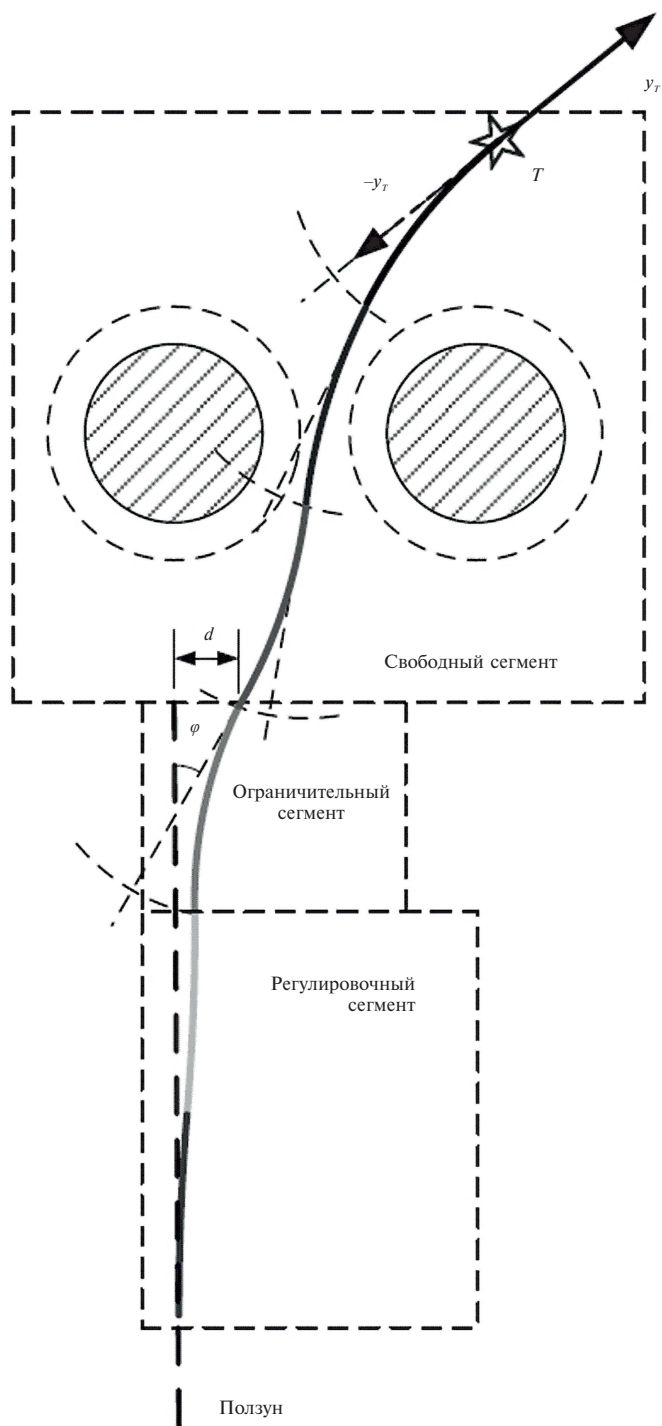


Рис. 6. Метод поиска траектории робота для заданной позиции и направления конечного эффектора.

2. Направление обратного движения y (см. рис. 4) выбирается как направление, противоположное y_T (см. рис. 6). Это позволяет определить рабочее пространство для концевой точки первого от головы сектора-сустава. В этой области определяется соответствующая позиция для создания эскиза первого сегмента гибкого сустава. Положение конечной точки для следующего сустава рассчитывается как решение следующей эмпирической оптимизационной задачи, состоящей из этапов А–D:

А. Конечная точка сустава не пересекается с препятствиями, т.е.

$$(8) \quad P_1 \notin R_0.$$

Здесь: R_0 – запрещенная область (множество точек модели препятствий). На рис. 6 примером запрещенной области являются две окружности R_1 и R_2 .

В. Расстояние d от конечной точки до линии направляющего ползуна должно быть минимально.

С. Если в результате применения двух вышеназванных принципов расстояние d между конечной точкой и линией направляющей оказывается больше, чем $2x_{\max}$, критерий оптимальности B меняется на D .

Д. Угол φ между направлением конца позвонка и линией направляющей должен быть минимальным.

3. Позиции последующих сегментов-суставов определяются аналогичным образом при движении от конца к основанию каждого сустава. Выбор кривых для отображения позиций суставов осуществляется с учетом приближения к прямой линии – направляющей движения ползуна без столкновений с препятствиями. Данные сегменты здесь названы «свободными сегментами» (см. рис. 6).

Для того чтобы последующие суставы могли достичь направляющей подвижного ползуна, на котором крепится основание всего робота, и при этом направляющая располагалась по касательной к первому позвонку сустава-основания, последний сустав свободного сегмента должен удовлетворять следующим условиям:

1) продолжение направления основания последнего сустава должно пересекаться с направляющей;

2) расстояние d от точки основания до линии направляющего ползуна должно быть не больше x_{\max} ;

3) угол φ между направлением основания сустава и направляющей ползуна должен быть не больше θ_{\max} .

Если эти условия не могут быть выполнены, то возникает необходимость перерасчета углов сочленений для суставов свободного сегмента. Если требование к расстоянию d не может быть выполнено, приемлемого решения не существует. Если требование к углу отклонения φ не может быть выполнено, ограничение на угол отклонения следует ввести в оптимизационную задачу при перепланировании, чтобы гарантировать, что конец сегмента находится

достаточно близко к направляющей, и при этом удовлетворяет требованию по максимальному отклонению угла.

Следующий участок назовем «ограничительный сегмент». Он состоит из одного гибкого сустава. Угол его изгиба ограничивается углом φ , при этом направление его основания параллельно направляющей.

Последний участок назовем «регулируемый сегмент». Он включает в себя два сустава. Задав углы этих двух суставов равными по величине, но противоположными по направлению, можно добиться того, что ориентация первого сустава останется неизменной при изменении расстояния от направляющей. В качестве угла изгиба необходимо выбрать угол, обеспечивающий прохождение первого позвонка первого сустава робота по касательной к направляющей. На этом поиск траектории движения концевой эффектора завершен.

В данном методе поиска траектории критичным является использование трех первых суставов для начальной ориентации робота, так как это минимальное количество, способное обеспечить правильное положение манипулятора относительно направляющей ползуна. Следовательно, лишь три сустава свободного сегмента способны обходить препятствия. Поэтому при данных условиях разработанный гибкий змееподобный манипулятор не может выполнять сложные маневры по обходу препятствий.

4.1.2. Подгонка под заданную траекторию

Метод подгонки положения суставов под заданную траекторию аналогичен предыдущему сценарию. Для свободного сегмента непосредственно выбираются те концы суставов, которые пересекаются с заданной траекторией. Метод подгонки, по сути, выстраивает суставы свободного сегмента по траектории. Поэтому фактическая траектория может иметь достаточно большую погрешность по сравнению с запланированной. Для сложных кривых, если конец свободного сегмента не удовлетворяет требованиям, указанным выше, подгонка траектории может оказаться недостижимой.

4.1.3. Достижение указанной позиции концевой эффектора

Описанные выше методы планирования траектории для достижения целевой точки имеют ограничения в сохранении гибкости манипулятора и способности к адаптации к окружающей среде. Это обуславливает необходимость разработки нового метода поиска траектории концевой эффектора.

В соответствии с ранее описанными этапами планирования траектории, очевидно, что основной проблемой является достижение такого положения для первого сустава всего робота, чтобы направляющая ползуна являлась касательной к кривой сустава в точке крепления робота. Во многих случаях для этого требуется не менее трех сочленений. Учитывая нефиксированную ориентацию конца каждого сегмента, планирование достижения заданного положения, начиная с направляющей, можно производить от основания к

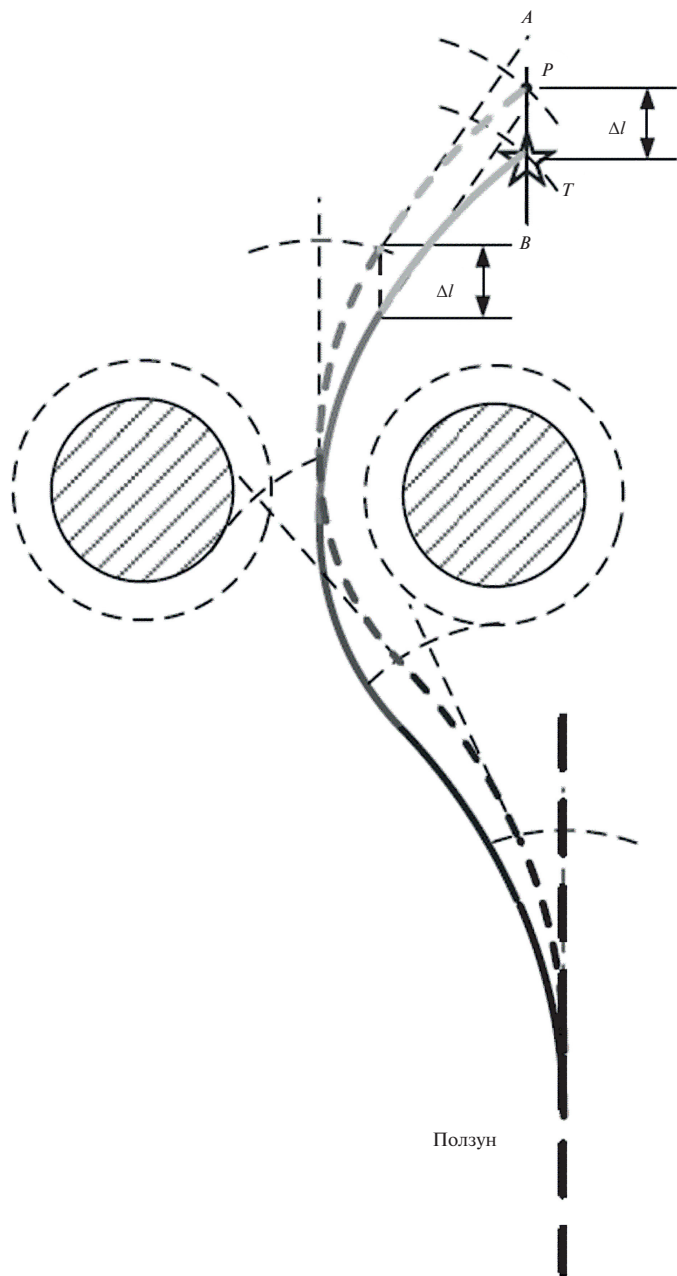


Рис. 7. Метод поиска траектории робота по допустимым областям.

концу сегмента, что увеличивает количество свободных суставов и упрощает процесс планирования траектории.

Пусть для основания каждого сегмента определены допустимые области расположения точки основания. Тогда шаги планирования выглядят следующим образом (см. рис. 7).

1. Обозначим целевую точку траектории символом T . Через точку T проводится прямая AB параллельно направляющей ползуна.

2. На расстоянии, не большем чем длина сустава L от прямой AB , выбирается точка из допустимой области в качестве временной точки начала траектории изгиба последнего сустава.

3. Определяется область достижимости для конца первого сустава из временной начальной точки.

4. Рассчитывается положение конечной точки сустава в области достижимости на основании построения оптимизационной задачи $A-D$, при этом расстояние d заменяется на расстояние между конечной и целевой точками.

5. Расчеты для последующих суставов повторяются аналогичным образом, продвигаясь в направлении конца каждого сустава. Когда конец сустава находится близко к целевой точке ($d \leq 2L$), критерий оптимальности B меняется на D : угол φ между конечным направлением и линией, соединяющей конечную точку и целевую точку, должен быть минимальным.

После завершения работы алгоритма будет получена траектория изгиба суставов робота, обеспечивающая достижение альтернативной целевой точки P . Расстояние между альтернативной P и реальной T целевыми точками вдоль прямой AB составляет Δl . Для достижения истинной целевой точки необходимо сдвинуть всю кривую траектории на величину Δl вдоль направления AB . Если после сдвига траектория пересеклась с моделями препятствий, то требуется спланировать траекторию заново с выбором на каждом шаге иных точек из допустимых областей в качестве временных точек основания суставов. Если пересечения отсутствуют, планирование считается завершенным.

Данный метод по сравнению с предыдущими методами позволяет робототехническому манипулятору использовать больше суставов для непосредственного выполнения поставленных задач, когда все суставы используются для достижения целевой точки и избегания препятствий. Следовательно, гибкий робот может эффективно исполнять задачи в более сложных условиях, что приводит к увеличению процента успешно выполненных задач.

4.2. Планирование траектории на основе пассивного изменения геометрической формы

Учитывая ограниченность прилагаемой для деформации пневмокамер силы, отметим, что диапазон движения концевой эффектора манипулятора при активных изгибах суставов часто ограничен. В то же время змееподобный манипулятор обладает гибкостью и широкими возможностями пассивного изменения геометрической формы. При соприкосновении сегментов-суставов с препятствиями во внешней среде манипулятор может самостоятельно автономно обходить эти препятствия за счет собственной гибкости. Это позволяет манипулятору адаптироваться к разнообразным сложным задачам и условиям работы.

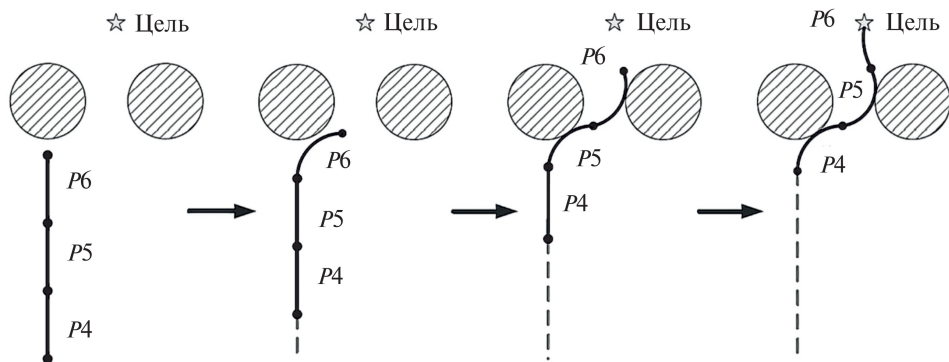


Рис. 8. Планирование траектории на основе пассивного изменения геометрической формы.

На примере, приведенном на рис. 8, во время движения гибкого манипулятора активное искривление потребовалось только для конечного сустава (головы). Остальные суставы изгибаются при контакте с препятствиями и завершают движение только за счет пассивного искривления.

Формула для расчета угла активного искривления последнего i -го соединения имеет вид

$$(9) \quad \theta_i = \frac{\|CP_{i-1}\|}{L} \times \arccos\left(\frac{FD_{i-1}}{\|F\| \times \|D_{i-1}\|}\right),$$

где C – точка контакта с препятствием, F – вектор касательной к контактной поверхности в базовой метке.

Эта стратегия эффективно использует гибкость змееподобного манипулятора, упрощает процесс планирования траектории и, соответственно, решение поставленной задачи, но может привести к частым столкновениям с элементами окружающей среды. Следовательно, такой подход эффективен в сценариях, где робототехнический манипулятор не вызывает повреждений при контакте с препятствием или где требования к безопасности относительно низки.

5. Эксперименты

5.1. Подготовка к эксперименту

Змееподобный манипулятор оснащен 12 приводными пневмокамерами для 6 суставов и 12 пропорциональными пневматическими клапанами управления, подключенными к компрессору. Робот установлен на линейном ползуне для обеспечения его выдвигания и втягивания посредством управления передвижением линейной подставки ползуна. На рис. 9 представлен состав системы управления робототехническим манипулятором. Персональный компьютер (PC) служит хостом для основной управляющей программы, написанной на языке C. Для передачи данных он взаимодействует с модулем микрокон-

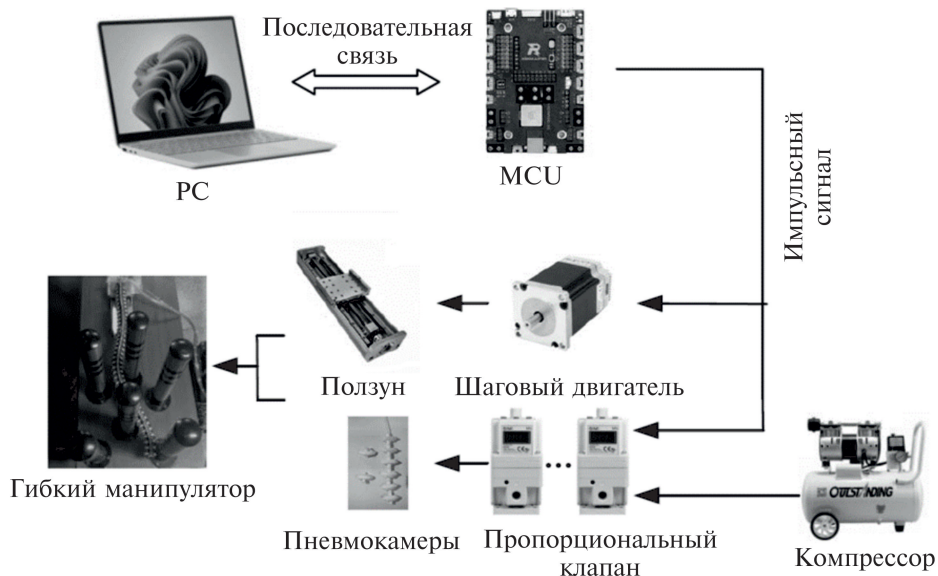


Рис. 9. Принципиальная диаграмма компоновки системы управления гибкого манипулятора.

троллера (MCU) по последовательному каналу связи. Исходя из заранее заданной стратегии движения манипулятора, компьютер рассчитывает углы изгиба каждого гибкого сустава в разные моменты времени и определяет перемещение линейной направляющей. Необходимые значения давления воздуха в 12 управляющих воздушных камерах определяются на основе зависимости между давлением воздуха и углами изгиба суставов. Затем MCU преобразует полученные расчетные значения в соответствующие электрические сигналы и передает их на пропорциональные клапаны и шаговые двигатели. В результате змееподобный манипулятор выполняет различные действия по перемещению и обходу препятствий.

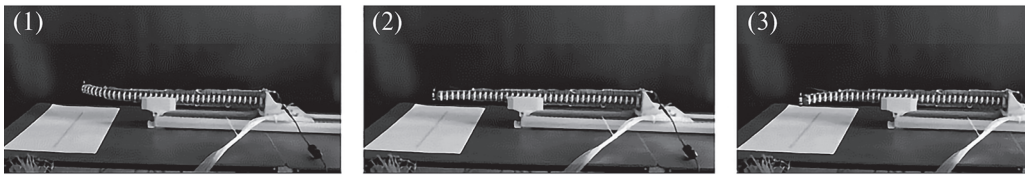
Были проведены две серии экспериментов: движение концевой эффектора вдоль заданной траектории без препятствий для исследования его производительности и движение концевой эффектора в заданную точку с препятствиями.

5.2. Эксперименты по движению по заданной траектории без препятствий

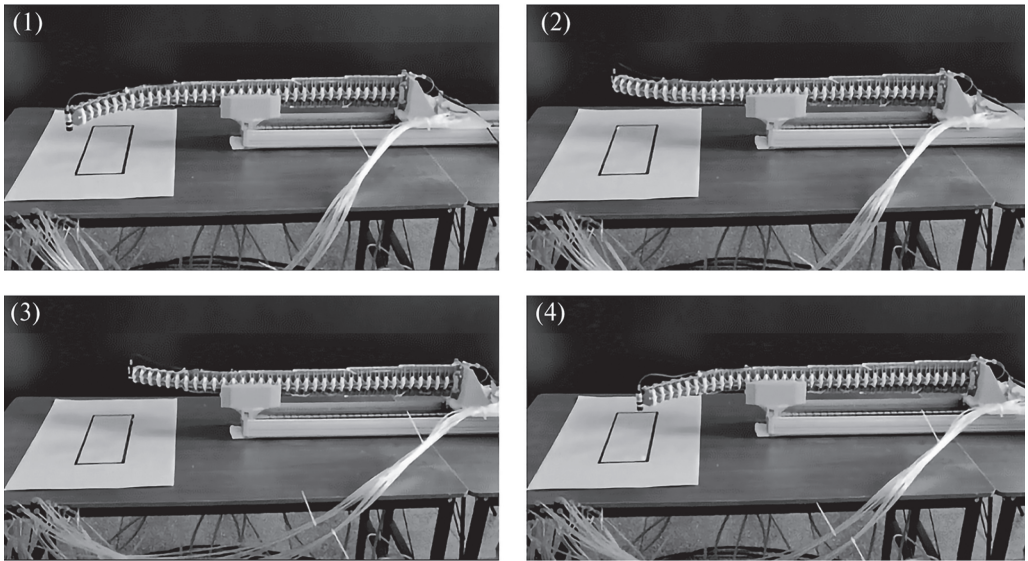
Эксперимент включает в себя создание определенных траекторий движения конечного эффектора для тестирования общей производительности гибкого змееподобного манипулятора, анализ возможностей его движения и оценку потенциального применения в двумерных операциях.

В данных экспериментах было протестировано движение по трем траекториям: прямая линия, прямоугольник и окружность.

a Отслеживание линейной траектории движения манипулятора



б Отслеживание составной линейной траектории манипулятора в виде прямоугольника



в Отслеживание круговой траектории манипулятора

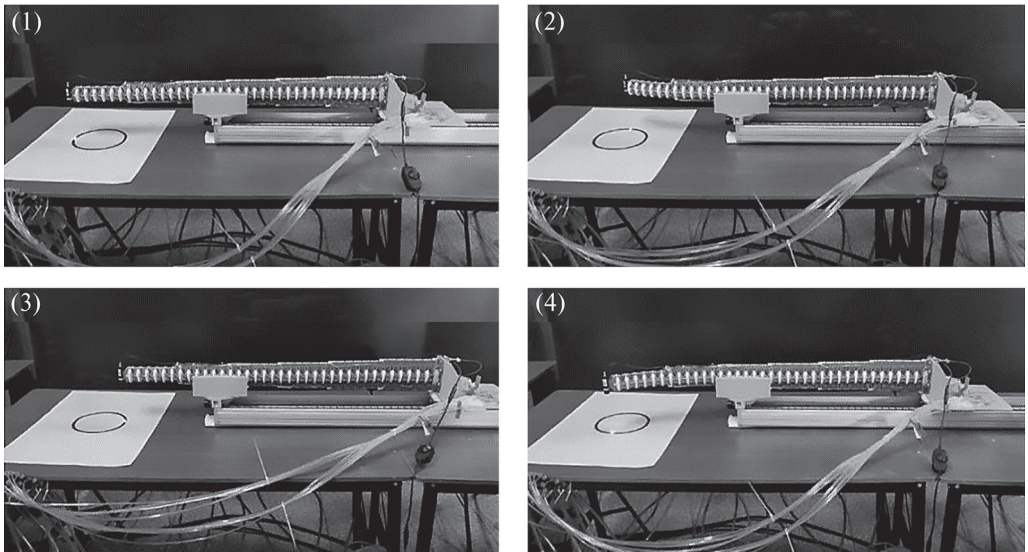


Рис. 10. Результаты эксперимента по управлению движением конечного эф- фектора без препятствий по заданной траектории.

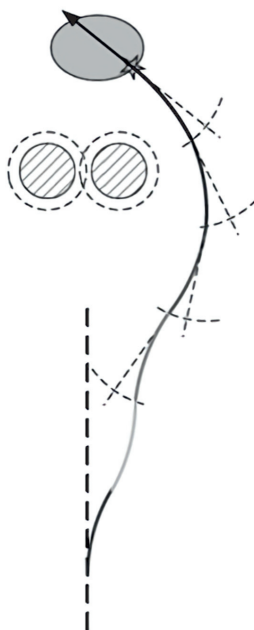
На конечный сустав манипулятора устанавливается лазер для отслеживания траектории движения конечного эффектора и расчета ошибки между реальной и заданной траекториями. Рисунок 10 демонстрирует ход трех экспериментов без препятствий: *а* – отслеживание линейной траектории, *б* – отслеживание составной линейной траектории в виде прямоугольника, *в* – отслеживание круговой траектории. Ошибка определялась как отклонение заданной траектории от реальной в конечный момент времени (конечная точка прямой линии, угол прямоугольника и конечная точка круга). В результате максимальное отклонение конечной позиции мягкого манипулятора не превышает 1%, а отклонение угла сустава в реальном эксперименте от расчетного не превышает 2%, что подтверждает высокую эффективность и точность движения гибкого змееподобного манипулятора при перемещении в двухмерном пространстве.

5.3. Эксперименты по активному избеганию препятствий на основе отслеживания положения конечного эффектора

Эксперимент по избеганию препятствий, изображенный на рис. 11, включает в себя два параллельно установленных опорных столба в качестве препятствий, а эффектор представляет собой стальную иглу, прикрепленную к



а – сцена эксперимента



б – планирование траектории конечного эффектора при заданном положении и ориентации



в – планирование траектории конечного эффектора при заданном положении

Рис. 11. Эксперимент по активному избеганию препятствий.

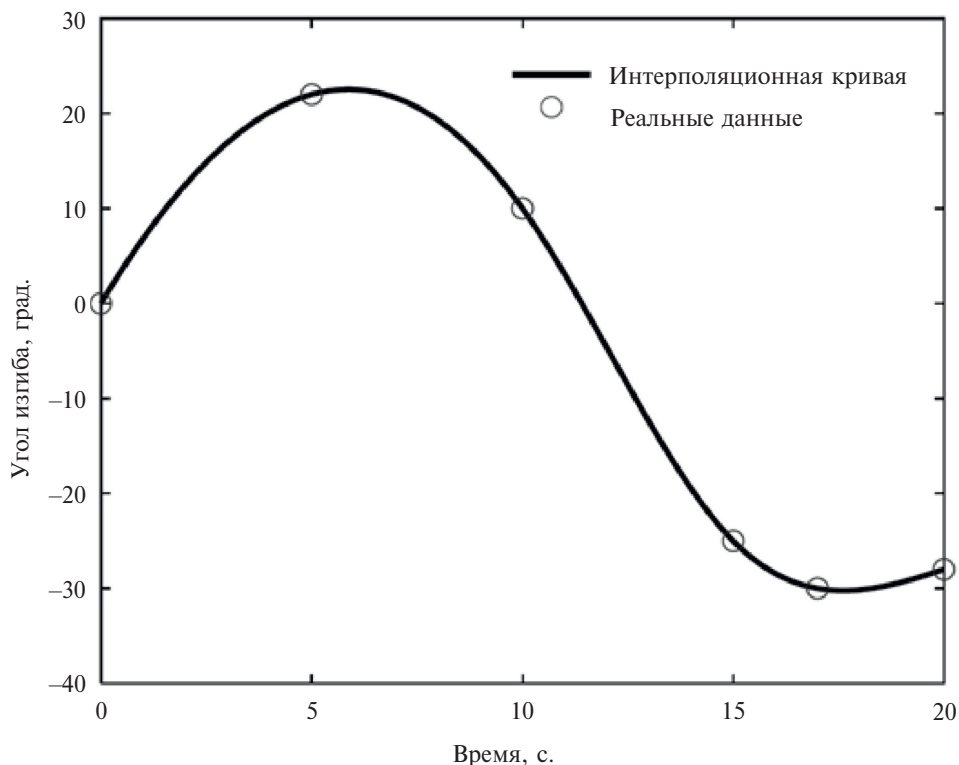


Рис. 12. Расчет траектории изменения угла конечного сустава во время огибания препятствия.

последнему позвонку манипулятора. Целью является воздушный шарик, расположенный позади препятствий. На основе предложенного ранее алгоритма планирования пути эффектор манипулятора должен обойти препятствия, чтобы достичь положения шарика и проколоть его.

Заданная экспериментальная сцена соответствует описанным ранее сценариям задач. На основе методов планирования траектории конечного эффектора, описанных в предыдущих разделах, построены траектории, изображенные на рис. 11.

Анализируя построенные траектории, можно сделать вывод, что при использовании метода с достижением заданного положения и ориентации требуется более длинный направляющий ползун и задействовано большее количество суставов. Так как в данном эксперименте нет строгих требований к положению конечного эффектора, была выбрана траектория, соответствующая методу достижения заданного положения.

Скорость подачи движущегося ползуна установлена на уровне 25 мм/с, время, необходимое для огибания препятствий, установлено в 20 секунд. Для расчета изменения во времени кривизны конечного сустава, несущего эффектор, применена интерполяция кубическим сплайном по четырем рассчитанным угловым положениям конечного эффектора, как показано на рис. 12.

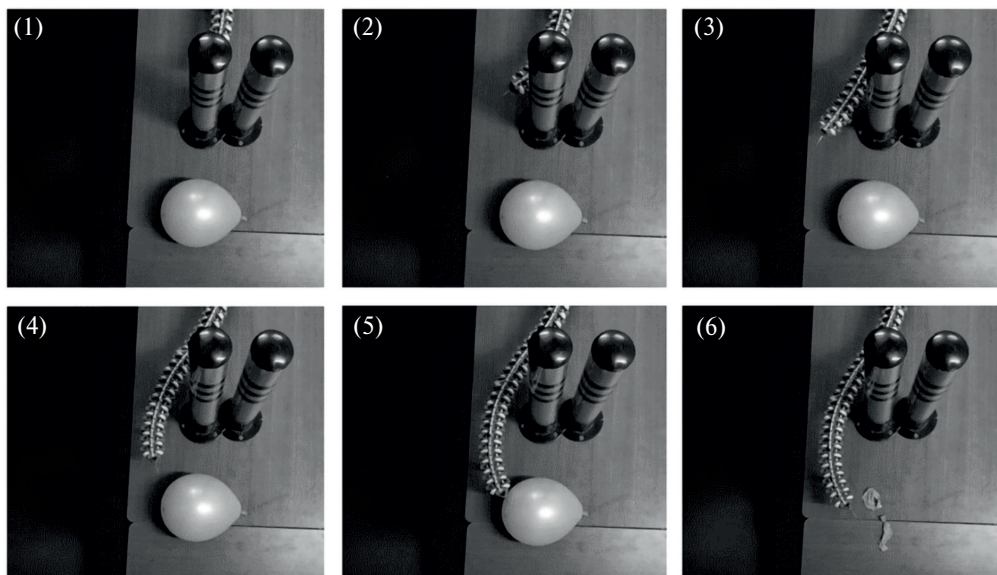


Рис. 13. Эксперимент по движению гибкого змееподобного манипулятора за счет активного изменения геометрической формы с обходом одного препятствия.

На 17-й секунде была добавлена дополнительная точка 30° , чтобы предотвратить превышение допустимого изгиба сустава.

Избегание препятствий гибким манипулятором достигается на основе полученной кривой изменения угла конечного сустава с использованием кинематической модели и стратегии отслеживания конечного эффектора. Рисунок 13 демонстрирует вид экспериментальной установки для движения за счет активного изменения геометрической формы гибкого манипулятора в процессе избегания препятствий. Манипулятор изначально располагается перед препятствием, перемещается вокруг препятствия, огибая его справа по ходу движения, и прокалывает воздушный шар, расположенный позади препятствия. Видно, что змееподобный манипулятор успешно завершил движение, обогнув препятствия по predetermined траектории, с отклонениями углов суставов от расчетных не более чем на 5%.

5.4. Эксперимент по избеганию препятствий на основе пассивного изменения геометрической формы

Для сценариев избегания препятствий на основе пассивного изменения геометрической формы, предложенного выше, контроль требуется только для расчета траектории движения головного сустава с эффектором, чтобы гарантировать движение манипулятора в нужном направлении между препятствиями. На рис. 14 представлен ход эксперимента по движению мягкого манипулятора на основе пассивного изменения геометрической формы для об-

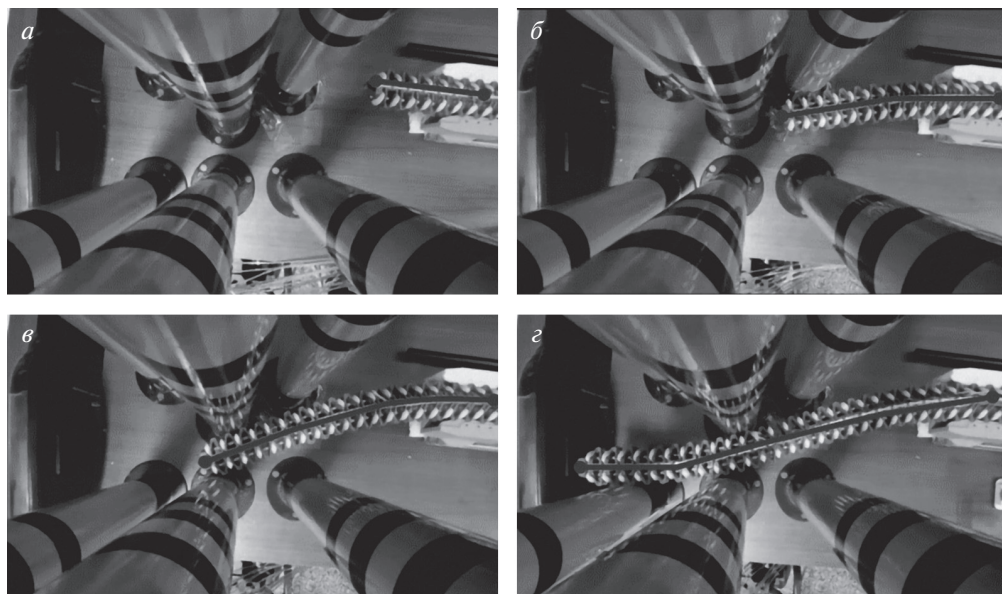


Рис. 14. Эксперимент по движению гибкого змееподобного манипулятора на основе пассивного изменения геометрической формы с обходом группы препятствий.

хода группы препятствий. Изменения углов изгибания суставов в процессе движения также проиллюстрированы на рис. 14. В результате эксперимента змееподобный манипулятор успешно завершил движение с обходом нескольких препятствий, и достиг установленной целевой позиции, используя свои активные и пассивные возможности по изменению геометрии.

6. Сравнительный анализ предложенного подхода с известными методами

В настоящем исследовании предложена оригинальная конструкция гибкого змееподобного манипулятора и методика построения системы планирования его траекторий на основе реальных экспериментов. Для анализа эффективности разработанной конструкции и соответствующей системы планирования траектории проведено сравнение предложенного подхода с известными конструкциями и методами расчета траекторий гибких манипуляторов.

6.1. Сравнение конструкций

Сравнение особенностей разработанной конструкции гибкого змееподобного манипулятора с известными аналогами представлено в табл. 1.

Как следует из табл. 1, представленная в данной работе конструкция гибкого змееподобного манипулятора обладает рядом преимуществ перед анало-

Таблица 1. Характеристики гибких манипуляторов различных конструкций

Тип манипулятора	Тип конструкции	Тип привода	Кол-во степеней свободы	Способ изменения геометрической формы	Способ крепления/перемещения	Метод расчета траектории
Промышленная роботизированная рука (KUKA LBR IIWA, Rozum PULSE и т.п.)	с постоянной жесткостью (высокая жесткость)	механический	3–7	активная	на стационарной подставке	Построение траектории на основе промышленных программ
Многозвеньевая роботизированная рука [1]	с постоянной жесткостью (средняя жесткость)	кабельный	5–20	активная	на стационарной подставке	Нет информации
Сверхмягкий манипулятор из жидкого эластомера [21]	с постоянной жесткостью (низкая жесткость)	гидравлический	15	пассивная/активная	на стационарном подвесе	Расчет траектории на основе отслеживания положения при помощи компьютерного зрения
Мягкий манипулятор на основе заклинивания [29]	с изменяемой жесткостью	затвердевающий под действием давления градулат	–	активная	на подвижном носителе/ползуне	Построение траектории на основе импульсных команд
Предложенная модель гибкого эмсеобразного манипулятора	с изменяемой жесткостью	пневматический	16	пассивная/активная	на ползуне	Расчет траектории на основе экспериментальных исследований

гами, в частности возможностью пассивного изменения геометрии, изменяемой жесткостью и большим количеством степеней свободы. Ключевой особенностью является специфический алгоритм расчета траектории, разрабатываемый на основе физических данных конкретного змееподобного манипулятора путем проведения серии экспериментов.

6.2. Сравнение подходов к построению траектории перемещения

Рассмотрим несколько подходов для планирования траектории, описанных в современной литературе.

– Планирование траектории многозвенного редундантного манипулятора на основе генетического алгоритма [37]: траектория манипулятора представляется координатами его узлов, конечный узел ассоциирован с концевым эффектором. Траектория должна обеспечить минимальное расхождение между положением концевого эффектора и целевой точкой. Траектория формируется как решение оптимизационной задачи, где целевой функцией выступает отклонение координат концевого эффектора от заданных координат цели, ограничениями выступают требования непересечения звеньев манипулятора с препятствиями и друг с другом и допустимости углов поворота звеньев. Оптимизация производится на основе генетического алгоритма, где траектории выступают в качестве хромосом, а координаты узлов в качестве генов. Алгоритм повторяется многократно для получения «генеалогического древа» – последовательности пригодных траекторий, переводящих манипулятор из заданного начального положения в конечное.

– Расчет траектории манипулятора на основе аналитического решения обратной задачи кинематики [32]: определяются углы поворотов трехзвенного манипулятора (плечо–предплечье–схват) исходя из требуемого положения последнего звена (схвата). Метод заключается в последовательном составлении и решении систем нелинейных уравнений, отражающих зависимости между углами поворотов последующего и предыдущего звеньев. За счет относительно простой конфигурации и ограничения по количеству звеньев задача имеет единственное решение.

– Расчет траектории многозвенного манипулятора на основе алгоритма FABRIK (Forward and Backward Reaching Inverse Kinematics) [36]: траектория многозвенного манипулятора рассчитывается согласно многократному повторению прямых и обратных итераций – проходов. Начальный узел считается жестко закрепленным, на последнем узле условно располагается концевой эффектор. На прямом проходе движение узлов осуществляется от концевого эффектора к начальному узлу. Последний узел, с концевым эффектором, перемещается на позицию, указанную в качестве цели. Затем каждый последующий узел передвигается в позицию, лежащую на прямой, соединяющей его и предыдущий, уже сдвинутый, узел, на расстояние, равное длине звена. В результате после того, как все узлы последовательно сдвинуты, начальный узел окажется смещенным относительно своей первоначальной позиции. Так

Таблица 2. Особенности методов построения траекторий для многозвенных манипуляторов

	Группа методов	Тип алгоритма	Кол-во звеньев манипулятора	Размерность рабочего пространства	Учет особенностей конкретного манипулятора	Обход препятствий	Перемещение конструкции
Расчет траектории на основе генетических алгоритмов	Случайный поиск	Однопроходный, итерационный	Не ограничено	3D	нет	да	требуется модификация алгоритма
Аналитический расчет траектории	Аналитическая модель	Однопроходный, без итераций	2–3	3D	частично	нет	нет
Forward and Backward Reaching Inverse Kinematics (FABRIK)	Эвристический алгоритм	Двухпроходный, итерационный	Не ограничено	3D	нет	требуется модификация алгоритма	нет
Расчет траектории на базе эмпирического подхода	Эмпирический алгоритм	Двухпроходный, итерационный	Не ограничено	2D	да	да	да

как начальный узел считается неподвижным и его позиция не может изменяться, производится коррекция положения начального узла – он возвращается в начальную позицию. После чего осуществляется обратный проход – все узлы манипулятора вновь сдвигаются, но уже от начального к конечному узлу по тому же принципу, что и в прямом проходе. Прямой и обратный проходы повторяются до тех пор, пока конечной эффектор не окажется достаточно близко к заданной целевой позиции.

Сравнение характеристик рассмотренных методов, а также предложенного метода на базе эмпирического подхода приведено в табл. 2.

Предложенный метод, в отличие от других существующих методов построения траектории, ориентирован на конкретный манипулятор, так как расчет осуществляется на основе эмпирических данных, полученных в ходе экспериментов. Аналитический метод, также предполагающий расчеты для конкретного манипулятора, основан на теоретических, но не на практически измеренных характеристиках устройства. Предложенный алгоритм не предназначен для изменения положения манипулятора в трехмерном пространстве, однако способен учитывать перемещение начального узла за счет подвижного ползуна.

6.3. Сравнение вычислительных затрат

Для сравнения предложенного алгоритма с другими алгоритмами построения траекторий манипуляторов с точки зрения вычислительных затрат оценим вычислительную временную сложность для соответствующих алгоритмов. Вычислительная временная сложность оценивается функцией зависимости количества требуемых операций от размера входных данных n . В данном разделе приведены значения временной сложности при $n \rightarrow \infty$, т.е. асимптотической вычислительной временной сложности [41].

– *Оценка вычислительной сложности генетического алгоритма.* При обычном выборе способов преобразования (точечная мутация, односточный кроссовер, выбор по принципу рулетки) сложность генетических алгоритмов составляет $O(g(nt + nt + n))$, где g – количество поколений, n – размер популяции, t – размер особей. Таким образом, вычислительная сложность алгоритма может быть оценена как $O(n^2)$.

– *Оценка вычислительной сложности метода аналитического расчета траектории.* Вычислительная сложность аналитического метода самая низкая. Так как алгоритм не предполагает итерационных вычислений, объем вычислений будет оставаться постоянным, и сложность составляет $O(1)$.

– *Оценка вычислительной сложности алгоритма Forward and Backward Reaching Inverse Kinematics.* Алгоритм FABRIK состоит из внутреннего и внешнего циклов. На внутреннем цикле последовательно производятся прямой и обратный проходы алгоритма, каждый такой проход в вычислительном плане порождает дерево разбиения глубиной $\log n$. Следовательно, сложность внутреннего цикла с двойным проходом составляет $O(\log n + \log n)$,

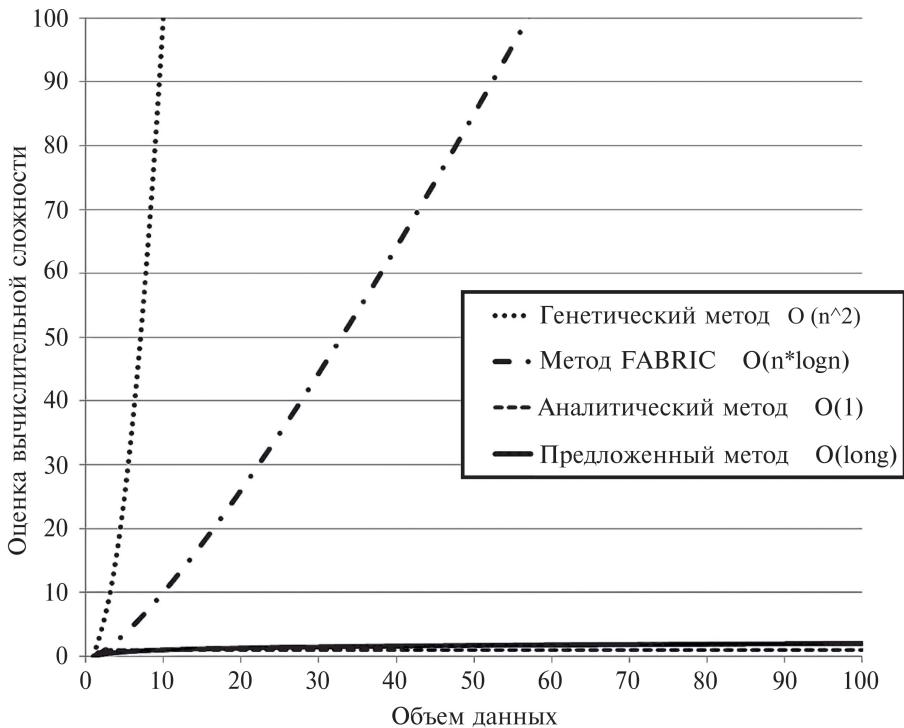


Рис. 15. Рост вычислительной сложности с увеличением объема данных для рассмотренных методов расчета траектории гибких манипуляторов.

что при $n \rightarrow \infty$ эквивалентно сложности $O(\log n)$. Внешний цикл алгоритма повторяет внутренний цикл число раз, пропорциональное объему входных данных n . Следовательно, вычислительная сложность оценивается как $O\left(\sum_{j=0}^n O(\log n)\right) = O(n \times \log n)$.

– Оценка вычислительной сложности алгоритма расчета траектории на базе эмпирического подхода. Предложенный алгоритм, в отличие от рассмотренного выше алгоритма FABRIC, состоит из единственного цикла. Внешний цикл отсутствует, так как конечное звено расположено на ползуне и может менять свое местоположение. Следовательно, сложность предложенного алгоритма составляет $O(\log n)$.

Графически сравнение оценок вычислительной сложности представлено на рис. 15.

7. Заключение

В данной статье на основе результатов международных исследований в области мягких роботов-манипуляторов представлены проектирование и разработка змееподобного гибкого манипулятора с 6, 13-ю степенями свободы на основе формы строения позвоночника змей. Общая конфигурация представляет собой последовательно-параллельную композитную структуру, управ-

ляемую 12 комплектами пневмокамер с обеих сторон манипулятора и линейной направляющей – ползуном. В работе выполнено кинематическое отображение пространства суставов манипулятора в рабочее пространство с помощью модели постоянной кривизны. На основе экспериментов была установлена зависимость между давлением воздуха в пневмокамере и углом изгиба сустава манипулятора. Для моделирования кинематики робота в направлении движения вперед был использован метод одностороннего движения по линейному ползуну.

Для удовлетворения требований практического инженерного применения гибкого манипулятора были разработаны методы планирования траектории для обхода препятствий. На основе подхода отслеживания положения концевой эффектора были предложены два метода расчета траектории движения робота. Также был проанализирован метод обхода препятствий на основе пассивного изменения геометрической формы мягкого робота. Проведенные физические эксперименты подтвердили реализуемость и корректность разработанных методов.

Сравнение предложенного подхода к проектированию гибкого манипулятора с известными проектными решениями продемонстрировало оригинальность разработанной конструкции. Преимуществом разработанного змееподобного гибкого манипулятора является возможность пассивного изменения его геометрической формы во время движения для обхода препятствий, а также изменяемая жесткость с большим количеством степеней свободы. Манипулятор предназначен для передвижения и работы на плоскости, что делает его применение предпочтительным в ряде практических задач. Разработанный для расчета траектории перемещения манипулятора алгоритм основан на экспериментальных измерениях характеристик пневматических камер-суставов, обеспечивающих сгибание–разгибание гибкого змееподобного манипулятора. Алгоритм имеет низкую вычислительную сложность, уступая лишь аналитическому моделированию, которое, в свою очередь, не может обеспечить широких возможностей по изменению траектории.

Планируется дальнейшее расширение исследований по конструированию и планированию траектории гибкого змееподобного манипулятора, а также по его управлению. В частности, будут исследованы методы управления роботом, позволяющие добиться переменной жесткости манипулятора. Также манипулятор может быть оснащен несколькими встроенными датчиками для отслеживания положения в реальном времени, что повысит скорость реакции и точность управления роботом.

СПИСОК ЛИТЕРАТУРЫ

1. *Larson O., Davidson C.* Flexible arm, particularly a robot arm: U.S. Patent No. 4393728. 19 Jul. 1983.
2. *Hannan M.W., Walker I.D.* Kinematics and the implementation of an elephant's trunk manipulator and other continuum style robots // *J. Robot.* 2003. No. 20(2). P. 45–63.

3. McMahan W., Jones B., Walker I.D. Design and implementation of a multi-section continuum robot: Air-Octor // 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems. 2005. P. 2578–2585.
4. Robert B. Robots in the nuclear industry: a review of technologies and applications // *Indust. Robot: Int. J.* 2011. No. 38(2). P. 113–118.
5. Rob B., Andrew G. Nuclear snake-arm robots // *Indust. Robot: Int. J.* 2012. No. 39(1). P. 6–11.
6. Camarillo D.B., Milne C.F., Carlson C.R., et al. Mechanics Modeling of Tendon-Driven Continuum Manipulators // *IEEE Transactions on Robotics*. 2008. No. 24(6). P. 1262–1273.
7. Xu W., Liu T., Li Y. Kinematics, Dynamics, and Control of a Cable-Driven Hyper-Redundant Manipulator // *IEEE/ASME Transactions on Mechatronics*. 2018. No. 23(4). P. 1693–1704.
8. Rolf M., Steil J. Constant curvature continuum kinematics as fast approximate model for the bionic handling assistant // *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE Publ. 2012. P. 3440–3446.
9. Yang G.Z., Bellingham J., Dupont P.E., et al. The grand challenges of Science Robotics // *Sci. Robot.* 2018. No. 3(14). P. eaar7650.
10. Dumais J., Forterre Y. Vegetable Dynamicks: the role of water in plant movements // *Ann. Rev. Fluid. Mech.* 2012. No. 44. P. 453–478.
11. Argiolas A., Puleo G.L., Sinibaldi E., et al. Osmolyte cooperation affects turgor dynamics in plants // *Sci. Rep.* 2016. No. 6. P. 30139.
12. Chen Z.C., Yamaji N., Fujii-Kashino M., et al. A cationchloride cotransporter gene is required for cell elongation and osmoregulation in rice // *Plant. Physiol.* 2016. No. 171. P. 494–507.
13. Margheri L., Laschi C., Mazzolai B. Soft manipulator inspired by the octopus: I. From biological functions to artificial requirements // *Bioinspir. Biomim.* 2012. No. 7(2). P. 1–12.
14. Wirekoh J., Park Y.L. Design of flat pneumatic artificial muscles // *Smart Materials and Structures*. 2017. No. 26(3). P. 035–009.
15. Rodrigue H., Wang W., Han M.W., et al. An overview of shape memory alloy-coupled actuators and robots // *Soft Robot.* 2017. No. 4. P. 3–15.
16. Miriyev A., Stack K., Lipson H. Soft material for soft actuators // *Nature Communications*. 2017. No. 8(1). P. 596–605.
17. Mirvakili S.M., Sim D., Hunter I.W., et al. Actuation of untethered pneumatic artificial muscles and soft robots using magnetically induced liquid-to-gas phase transitions // *Science Robotics*. 2020. No. 5(41). P. eaaz4239.
18. Neppalli S., Jones B., McMahan W., et al. OctArm – A soft robotic manipulator // 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems. 2007. P. 2569–2569.
19. Hao J., Wang Z., Jin Y., Chen X., Peijin L., Gan Y., Chen X. Hierarchical control of soft manipulators towards unstructured interactions // *Int. J. Robot. Res.* 2021. No. 40. P. 0278364–92097936.
20. Escande C., Chettibi T., Merzouki R., Coelen V., Pathak P.M. Kinematic calibration of a multisection bionic manipulator // *IEEE/ASME Transactions on Mechatronics*. 2015. No. 20(2). P. 663–674.

21. *Marchese A.D., Rus D.* Design, kinematics, and control of a soft spatial fluidic elastomer manipulator // *Int. J. Robot. Res.* 2015. No. 35(7). P. 840–869.
22. *Yan J., Zhang X., Xu B., Zhao J.* A new spiral-type inflatable pure torsional soft actuator // *Soft Robotics.* 2018. No. 5(5). P. 527–540.
23. *Thuruthel T.G., Shih B., Laschi C., Tolley M.T.* Soft robot perception using embedded soft sensors and recurrent neural networks // *Science Robotics.* 2019. No. 4(26). P. eaav1488.
24. *Greer J.D., Blumenschein L.H., Alterovitz R., Hawkes E.W., Okamura A.M.* Robust navigation of a soft growing robot by exploiting contact with the environment // *Int. J. Robot. Res.* 2020. No. 39(14). P. 1724–1738.
25. *Gong Z., Fang X., Chen X., Cheng J., Xie Z., Liu J., Wen L.* A soft manipulator for efficient delicate grasping in shallow water: Modeling, control, and real-world experiments // *Int. J. Robot. Res.* 2021. No. 40(1). P. 449–469.
26. *Kurumaya S., Phillips B.T., Becker K.P., Rosen M.H., Gruber D.F., Galloway K.C., Suzumori K., Wood R.J.* A modular soft robotic wrist for underwater manipulation // *Soft Robot.* 2018. No. 5. P. 399–409.
27. *Kaufmann J., Bhovad P., Li S.* Harnessing the multistability of kresling origami for reconfigurable articulation in soft manipulators // *Soft Robotics.* 2022. No. 9(2). P. 212–223.
28. *Schubert B.E., Floreano D.* Variable stiffness material based on rigid low-melting-point-alloy microstructures embedded in soft poly (dimethylsiloxane) (PDMS) // *Rsc. Adv.* 2013. No. 3(46). P. 24671–24679.
29. *Brown E., Rodenberg N., Amend J., et al.* Universal robotic gripper based on the jamming of granular material // *P. Natl. Acad. Sci.* 2010. No. 107(44). P. 18809–18814.
30. *Kim Y.J., Cheng S., Kim S., et al.* A novel layer jamming mechanism with tunable stiffness capability for minimally invasive surgery // *IEEE T. Robot.* 2013. No. 29(4). P. 1031–1042.
31. *Hao J., Wang Z., Jin Y., Chen X., Peijin L., Gan Y., Chen X.* Hierarchical control of soft manipulators towards unstructured interactions // *Int. J. Robot. Res.* 2021. No. 40. P. 027836492097936.
32. *Братчиков С.А., Абрамова Е.А., Федосов Ю.В.* Решение обратной задачи кинематики манипулятора // *Вестник – Томский государственный университет. Управление, вычислительная техника и информатика.* 2021. № 56. С. 4–11. <https://doi.org/10.17223/19988605-56-1>
33. *Молотков А.В., Челноков Ю.Н.* Решение обратной задачи кинематики робота-манипулятора «Пума» с использованием бикватернионной теории кинематического управления // *Математика. Механика.* 2002. № 4. С. 204–206.
34. *Анципорович П.П., Акулнич В.К., Дубовская Е.М.* Решение обратной задачи кинематики манипулятора // *Теоретическая и прикладная механика: международный научно-технический сборник.* 2017. № 32. С. 306–309.
35. *Svinin M.M., Uchiyama M.* A new compensation scheme for the inverse kinematics tasks of flexible robot arms // *Proceedings of the 1994 IEEE International Conference on Robotics and Automation.* 1994. V. 1. P. 315–320. <https://doi.org/10.1109/ROBOT-1994-351276>

36. *Lamb M., Lee S., Billing E., Hogberg D., Yang J.* Forward and Backwards Reaching Inverse Kinematics (FABRIK) solver for DHM: A pilot study // Proceedings of the 7th International Digital Human Modeling Symposium (DHM 2022) and Iowa Virtual Human Summit. 2022. P. dhm.31772. <https://doi.org/10.17077/dhm-31772>
37. *Камильянов А.Р.* Планирование траекторий движения многозвенного манипулятора в сложном трехмерном рабочем пространстве на основе эволюционных методов // Дис. . . канд. техн. наук. Уфа. 2007.
38. *Плотникова Н.В.* Управление манипуляционными роботами на основе нечеткой логики // Наука ЮУрГУ: материалы 62 науч. конф. 2010. Т. 2. С. 170–174.
39. *Рахим Ф.А.* Нейро-нечеткая структура планирования перемещения робота-манипулятора в режиме онлайн в неизвестной динамической среде // Известия вузов. Северо-Кавказский регион. Технические науки. 2008. № 6. С. 41–49.
40. *Song T.L., Lu Y.P., Li Z.Y.* Structural design and research of the bionic snake-like robot // Advanced Materials Research. Trans Tech Publications Ltd. 2012. No. 538. P. 3034–3037.
41. *Oluwade B., Uwadia C., Ayeni J.* Asymptotic Time Complexity of an Algorithm for Finding the Error Pattern of a Uniform Digital Code // J. Sci. Res. Develop. 2001. No. 6. P. 127–134.

Статья представлена к публикации членом редколлегии Р.А. Мунасыповым.

Поступила в редакцию 14.05.2024

После доработки 22.04.2025

Принята к публикации 29.04.2025

СОДЕРЖАНИЕ

Тематический выпуск (окончание)

Аблаев С.С., Стонякин Ф.С., Федотов М.Н., Алкуса М.С., Савчук О.С., Гасников А.В. Исследование градиентного метода с неточной информацией о градиенте на некоторых классах (L_0, L_1) -гладких невыпуклых задач	3
Ахаван А., Цыбаков А.Б. Безградиентная стохастическая оптимизация для аддитивных моделей	28
Бирюков Р.С., Коган М.М. Синтез робастных законов управления по зашумленным текущим данным	46

Обзоры

Кулида Е.Л., Лебедев В.Г. Обзор методов предупреждения конфликтов при управлении воздушным движением с помощью глубокого обучения с подкреплением	64
--	----

Управление в технических системах

Синьбинь Чжан, Чжо Чэнь, Хаосян Су, Лиго Тань, Хунвэй Лю, Маньлю Лю, Цзяньвэнь Хо, Новикова С.В. Разработка конструкции и системы планирования траектории гибкого змееподобного робота-манипулятора на основе эмпирического подхода	93
--	----

CONTENTS

Topical issue

Ablaev S.S., Stonyakin F.S., Fedotov M.N., Alkousa M.S., Savchuk O.S., Gasnikov A.V. Study of Gradient Method with Inexact Gradient Information on Some Classes of (L_0, L_1) -Smooth Non-Convex Problems	3
Akhavan A., Tsybakov A.B. Gradient-Free Stochastic Optimization for Additive Models	28
Biryukov R.S., Kogan M.M. Synthesis of Robust Control Based on Noisy Data	46

Surveys

Kulida E.L., Lebedev V.G. An Overview of Conflict Prevention Methods in Air Traffic Control Using Deep Reinforcement Learning	64
--	----

Control in Technical Systems

Xinbin Zhang, Zhuo Chen, Haoxiang Su, Liguo Tan, Hongwei Liu, Manlu Liu, Jianwen Huo, Novikova S.V. Development of the Design and Trajectory Planning of a Snake-Like Manipulator Robot Based Empirical Approach	93
---	----