

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

Белорусский государственный университет

Республиканский конкурс научных работ студентов

Математика. Методы и алгоритмы вычислительной математики  
математического моделирования для решения  
задач экономики, техники и природоведения

## НЕЙРОСЕТЕВЫЕ РЕГУЛЯТОРЫ В СИСТЕМАХ УПРАВЛЕНИЯ С ПРОГНОЗИРУЮЩЕЙ МОДЕЛЬЮ

Павловец Мария Евгеньевна  
выпускник

Наталия Михайловна Дмитрук  
канд. физ.-мат. наук, доцент

Минск, 2019

# РЕФЕРАТ

Работа, 53 с., 23 рис., 5 табл., 24 источника

## ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ, УПРАВЛЕНИЕ С ПРОГНОЗИРУЮЩЕЙ МОДЕЛЬЮ, НЕЙРОННАЯ СЕТЬ, НЕЙРОСЕТЕВОЙ РЕГУЛЯТОР, МАШИННОЕ ОБУЧЕНИЕ

Объектом исследования являются задачи стабилизации нелинейных динамических систем, на управляющие воздействия и траектории которых наложены жесткие ограничения, и связанные с рассматриваемой задачей схемы управления по прогнозирующей модели, реализованные на основе нейронных сетей.

Цель работы — построение нейросетевых регуляторов, гарантирующих асимптотическую устойчивость нелинейной системы.

Для достижения поставленной цели в работе используются результаты теории устойчивости, теории управления, теории управления по прогнозирующей модели, численные методы оптимального управления и оптимизации, методы машинного обучения, в частности, обучения многослойной нейронной сети.

Основной результат работы представлен алгоритмами управления на основе нейронных сетей, позволяющими (по результатам экспериментов) повысить производительность схемы управления с прогнозирующей моделью в 20 раз.

Работа состоит из трех глав. В первой главе определяются основные понятия и результаты в области управления с прогнозирующей моделью. Во второй главе рассматриваются методы оффлайн дизайна обратных связей и их аппроксимаций в системах с прогнозирующей моделью. В третьей главе представлены результаты построения нейросетевого регулятора для нелинейных систем с ограничениями, а также проведено сравнение предлагаемых подходов с классическими алгоритмами, основанными на решении задач оптимального управления онлайн.

Новизна представленной работы состоит в построении нейросетевых регуляторов совместно с построением областей притяжения и областей насыщенных управлений, использовании результатов теории робастного управления для гарантированной стабилизации нелинейной системы, исследовании влияния различных техник сэмплирования точек для построения обучающих выборок, что помогает сократить время обучения без потерь точности.

Областями применения разработанных алгоритмов являются прикладные задачи, решаемые в рамках теории управления с прогнозирующей моделью и возникающие в робототехнике, химической промышленности, транспортных системах, системах управления беспилотными аппаратами и др.

# ОГЛАВЛЕНИЕ

	С.
<b>ВВЕДЕНИЕ</b> . . . . .	5
<b>ГЛАВА 1 Основные понятия и обзор литературы</b> . . . . .	6
1.1 Основные принципы МРС . . . . .	6
1.2 МРС для решения задач стабилизации . . . . .	7
1.3 Базовый алгоритм МРС . . . . .	9
1.4 Терминальные элементы и устойчивость замкнутой системы . . . . .	12
1.5 Выводы . . . . .	16
<b>ГЛАВА 2 Использование методов оффлайн дизайна в системах управления с прогнозирующей моделью</b> . . . . .	17
2.1 Основные понятия нейронных сетей . . . . .	17
2.2 Методы дизайна оффлайн регуляторов в системах управления с прогнозирующей моделью . . . . .	20
2.3 Выводы . . . . .	28
<b>ГЛАВА 3 Построение нейросетевого регулятора</b> . . . . .	30
3.1 Описание иллюстративного примера . . . . .	30
3.2 Подход к решению . . . . .	32
3.3 Базовая реализация нейросетевого регулятора . . . . .	35
3.4 Построение области притяжения и областей насыщения управления с помощью SVM . . . . .	39
3.5 Сравнение методов сэмплирования точек для построения приближенной обратной связи . . . . .	40
3.6 Сравнение результатов производительности . . . . .	43
3.7 Использование нейросетевого регулятора для задачи стабилизации обратного маятника . . . . .	45
3.8 Выводы . . . . .	48
<b>ЗАКЛЮЧЕНИЕ</b> . . . . .	50
<b>СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ</b> . . . . .	51

# ПЕРЕЧЕНЬ УСЛОВНЫХ ОБОЗНАЧЕНИЙ И СОКРАЩЕНИЙ

ОУ — оптимальное управление

MPC — Model Predictive Control, управление с прогнозирующей моделью

SVM — Support Vector Machines, метод опорных векторов

MSE — Mean square error, средняя квадратичная ошибка

RL — Reinforcement learning, метод обучения с подкреплением

MDP — Markov Decision Process, процесс принятия решений Маркова

LBMPC — Learning based model predictive control, управление с прогнозирующей моделью, полученное при применении машинного обучения

$\mathbb{I}_{\geq a}$  — множество целых чисел больше либо равных  $a \in \mathbb{R}$

$\mathbb{I}_{[0, N-1]} = \{0, 1, \dots, N-1\}$

$\|x\|$  — евклидова норма вектора  $x \in \mathbb{R}^n$

$\|x\|_Q^2 = x^T Q x$  — взвешенная норма вектора  $x \in \mathbb{R}^n$ ,  $Q > 0$

$\|x\|_A = \inf \|x - a\|$  — расстояние от точки  $x$  до множества  $A$

$\mathcal{K}_\infty$  — класс функций, где функция  $\alpha : [0, \delta) \rightarrow [0, \infty)$  — непрерывная, строго возрастающая,  $\alpha(0) = 0$  и  $\lim_{\delta \rightarrow \infty} \alpha(\delta) \rightarrow \infty$

$I$  — единичная матрица

$\lambda_{\max}(A)$ ,  $\lambda_{\min}(A)$  — максимальное и минимальное собственное значение матрицы  $A$

$1_p$  — вектор из единиц размерности  $p$

## ВВЕДЕНИЕ

Метод управления с прогнозирующей моделью (МРС) [1, 2] — одна из популярных современных технологий теории управления, основанная на решении в реальном времени задач оптимального управления с конечным горизонтом, аппроксимирующих решение задачи управления с бесконечным временным промежутком (например, задачи стабилизации).

Основными причинами популярности МРС при решении прикладных задач (обзор практических применений можно найти в работе [3]) являются применимость схемы управления к нелинейным системам и возможность учета ограничений на управления и траектории, а также способность работать без экспертного вмешательства в течение длительного времени. С другой стороны, эти же факторы могут стать причиной нереализуемости алгоритма МРС, например, в случае быстрых процессов, для которых решение нелинейной задачи оптимального управления не может быть получено регулятором достаточно быстро, в реальном времени, т.е. в темпе обновления информации о состояниях управляемой системы.

Один из подходов к решению указанной проблемы подразумевает перенос некоторых вычислений оффлайн. В частности, в настоящей работе для реализации функции МРС-регулятора предлагается применить искусственные нейронные сети. Нейронные сети будут использоваться для аппроксимации закона управления.

В настоящей работе будут исследованы вопросы устойчивости и робастности нелинейных динамических систем, замкнутых обратной связью, построенной МРС-регулятором, и проведено сравнение с оптимальным МРС-регулятором для некоторых прикладных задач. В частности, в главе 1 изложены основные принципы МРС, рассмотрена задача стабилизации, описан базовый алгоритм МРС. Там же исследованы вопросы асимптотической устойчивости замкнутой системы вместе с алгоритмами построения терминального множества и терминальной функции для обеспечения устойчивости этой системы. В основной части работы (главы 2 и 3) разрабатываются методы построения нейросетевых регуляторов и областей притяжения для различных нелинейных систем, проводится сравнение производительности и точности при использовании различных техник сэмплирования точек для аппроксимации обратной связи.

# ГЛАВА 1

## ОСНОВНЫЕ ПОНЯТИЯ И ОБЗОР ЛИТЕРАТУРЫ

Управление по прогнозирующей модели — Model Predictive Control (МРС) [1, 2] — современный подход к управлению линейными и нелинейными динамическими системами, основанный на решении в реальном времени последовательности задач оптимального управления (ОУ) с конечным временным горизонтом. Упомянутые задачи ОУ называются прогнозирующими, формулируются в зависимости от целей управления, учитывают текущие измерения состояний объекта управления и ограничения на траектории и управляющие воздействия, а также аппроксимируют исходную задачу управления на бесконечном полуинтервале времени.

В настоящей главе излагаются основные принципы и базовый алгоритм МРС на примере задачи стабилизации управляемых движений динамической системы.

### 1.1 Основные принципы МРС

МРС базируется на следующих основных принципах [2]:

- для предсказания и оптимизации будущего поведения системы используется математическая модель управляемого процесса в пространстве состояний (в отличие от описания в виде передаточной функции и других методов частотной области);
- для выбранной математической модели формулируется прогнозирующая задача ОУ (predictive optimal control problem), которая будет решаться в каждый момент времени; в этой задаче:
  - конечный промежуток управления;
  - начальное состояние математической модели совпадает с измеренным текущим состоянием физического объекта управления;
  - критерий качества отражает цели управления: если целью является стабилизация объекта управления, то критерием качества выступает отклонение траектории объекта от положения равновесия;
  - учтены ограничения на траекторию и управляющие воздействия;

- оптимальное управление прогнозирующей задачи ОУ (предсказанное управляющее воздействие) применяется к объекту в текущий момент времени и до тех пор пока не будет измерено следующее состояние объекта; затем оптимизация повторяется.

Поскольку в каждый момент времени в задаче ОУ учитывается текущее состояние, результирующее управление представляет собой обратную связь.

Популярность МРС в теоретических исследованиях [1, 2] и на практике [3] обусловлена следующими свойствами, которыми не обладают другие методы теории управления:

- критерий качества в прогнозирующей задаче ОУ позволяет учитывать экономические требования к процессу управления (например, минимизацию энергетических затрат);
- учитываются жесткие ограничения на фазовые и управляющие переменные;
- метод применим к нелинейным и многосвязным системам.

Отметим, что поскольку решение задачи ОУ повторяется для каждого текущего момента времени, промежуток, для которого прогнозируется поведение системы, постоянно смещается ("скользит"), в силу чего МРС также иногда называется управлением со скользящим горизонтом — Receding Horizon Control (RHC).

## 1.2 МРС для решения задач стабилизации

Основными и исторически первыми приложениями МРС являются задачи стабилизации и регулирования. Остановимся подробно на результатах, полученных в теории МРС для задачи стабилизации. В этом разделе также вводятся основные обозначения, понятия, определения, базовый алгоритм МРС, свойства МРС-регулятора.

Как было отмечено выше, основная идея МРС состоит в том, чтобы использовать математическую модель процесса в пространстве состояний для предсказания и оптимизации поведения динамической системы в будущем [2]. Далее считаем, что используемая для предсказаний модель точно описывает процесс управления: на объект не действуют возмущения и нет неучтенных различий между моделью и физическим объектом. Такие схемы МРС носят название номинальных (nominal MPC scheme).

Система, которая исследуется в данном разделе, является нелинейной, дискретной, стационарной:

$$x(t+1) = f(x(t), u(t)), \quad x(0) = x_0. \quad (1.1)$$

Здесь  $x(t) \in X \subseteq \mathbb{R}^n$  — состояние системы в момент времени  $t$ ,  $u(t) \in U \subseteq \mathbb{R}^r$  — управляющее воздействие в момент  $t$ ,  $t \in \mathbb{I}_{\geq 0}$  — время, дискретное. Относительно функции  $f : \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}^n$  предполагается, что она непрерывна.

Начальное состояние системы (1.1) задано:

$$x(0) = x_0 \in X.$$

На состояния и управляющие воздействия  $u$  накладываются ограничения вида

$$(x(t), u(t)) \in Z \subseteq X \times U, \quad t \in \mathbb{I}_{\geq 0}, \quad (1.2)$$

которые называются смешанными ограничениями. Понятно, что такая форма задания ограничений включают в себя одновременно и фазовые ограничения на состояния системы, и прямые ограничения на управляющие воздействия. Относительно множества  $Z$  предполагается, что оно компактно.

Цель (стабилизирующего) МРС — построить обратную связь  $u(x)$ , при которой замкнутая система

$$x(t+1) = g(x(t)) = f(x(t), u(x(t))), \quad x(0) = x_0, \quad (1.3)$$

будет устойчива в некотором заданном положении равновесия (заданном множестве), при этом переходный процесс не нарушает ограничения (1.2) при всех  $t \in \mathbb{I}_{\geq 0}$ .

**Определение 1.1** Точка  $x^*$  является положением равновесия для системы (1.3), если выполняется  $x^* = g(x^*)$ .

**Определение 1.2** Множество  $X \subseteq \mathbb{R}^n$  называется положительно инвариантным множеством для системы (1.3), если выполняется  $g(x) \in X \quad \forall x \in X$ .

**Определение 1.3** Пусть  $X \in \mathbb{R}^n$  — положительно инвариантное множество для системы (1.3). Замкнутое, положительно инвариантное множество  $A \subseteq X$  устойчиво для (1.3), если  $\forall \epsilon > 0, \exists \delta > 0$ , что для всех  $\|x_0\|_A \leq \delta, x_0 \in X$ , выполняется  $\|x(t)\|_A \leq \epsilon, \forall t \in \mathbb{I}_{\geq 0}$ .

**Определение 1.4** Множество  $A \subseteq X$ , удовлетворяющее условиям определения 1.3, асимптотически устойчиво с областью притяжения  $X$ , если оно устойчиво и  $\lim_{t \rightarrow +\infty} \|x(t)\|_A = 0 \quad \forall x_0 \in X$ .



**Определение 1.5** Множество — глобально асимптотически устойчиво, если оно асимптотически устойчиво с  $X = \mathbb{R}^n$ .

При  $A = \{x^*\}$  получим классические понятия устойчивости, асимптотической устойчивости и глобальной асимптотической устойчивости решения  $x(t) = x^*$  по Ляпунову.

Далее будет рассматриваться случай стабилизации системы управления (1.1) для заданного положения равновесия, т.е. случай  $A = \{x^*\} \in X$ .

### 1.3 Базовый алгоритм МРС

Как было отмечено в разделе 1.1, идея алгоритма МРС состоит в том, чтобы в каждый момент  $t \in \mathbb{I}_{\geq 0}$  оптимизировать будущее поведение системы (1.1) на конечном горизонте  $N \geq 2$  и использовать первое значение полученного оптимального (программного) управления в качестве значения обратной связи для момента  $t$  (см. рис. 1.1). Под "оптимизацией будущего поведения" понимается решение прогнозирующей задачи ОУ.

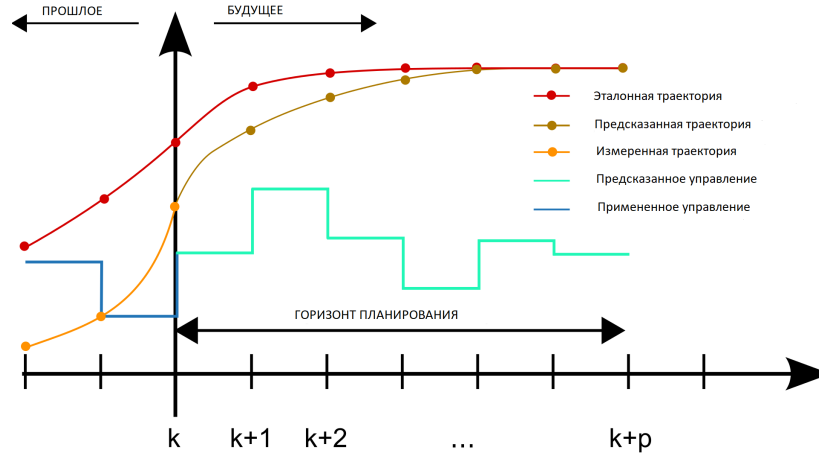


Рис. 1.1: Схема МРС

Понятно, что далее необходимо различать состояния объекта управления  $x(t)$ ,  $t \in \mathbb{I}_{\geq 0}$ , которые измеряются в каждом конкретном процессе управления, и состояния математической модели, которая используется для предсказаний и формулировки прогнозирующей задачи ОУ. Поэтому состояния математической модели будем обозначать  $x(k|t)$ ,  $k \in \mathbb{I}_{[0, N-1]}$ . Они изменяются согласно уравнению

$$x(k+1|t) = f(x(k|t), u(k|t)), \quad x(0|t) = x(t), \quad k \in \mathbb{I}_{[0, N-1]}. \quad (1.4)$$

Здесь аргумент  $t$  после черты подчеркивает зависимость от текущего момента, для которого проводится оптимизация. Начальное состояние — текущее состояние объекта управления  $x(t)$ .

Ограничения (1.2), записанные для состояний математической модели (1.4), имеют вид:

$$(x(k|t), u(k|t)) \in Z, \quad k \in \mathbb{I}_{[0, N-1]}.$$

Кроме приведенных смешанных ограничений, в задачу, как правило, добавляются ограничения в терминальный момент времени. "Терминальные элементы" прогнозирующей задачи более подробно будут рассмотрены ниже после ее формулировки.

Оставшийся элемент прогнозирующей задачи ОУ — критерий качества. В задачах стабилизации критерий качества выбирается исследователем, практиком, и является, скорее, параметром настройки схемы МРС. Например, в задаче стабилизации (см. [2]) критерий качества выбирается из соображений штрафа любого состояния  $x \in X$ , отклоняющегося от состояния равновесия  $x^*$ . Также часто штрафуются отклонения управления  $u \in U$  от значения  $u^*$ . Как отмечается в [2], последнее условие полезно с вычислительной точки зрения, поскольку для численных методов зачастую проще решить задачу, в которой в критерии качества штрафуются управляющие воздействия. С другой стороны [2], с точки зрения реализации управления также желательно избежать значений  $u \in U$ , соответствующих чрезмерным энергетическим затратам.

Критерий качества будет состоять из терминальной стоимости  $V_f(x(N|t))$  и суммарной стоимости переходного процесса, т.е. это будет критерий качества типа Больца. Терминальная стоимость будет рассмотрена ниже, при обсуждении терминальных элементов задачи ОУ. Стоимость переходного процесса для дискретных систем задается суммой стоимостей за каждый этап (для каждого  $k \in \mathbb{I}_{[0, N-1]}$ ):

$$\sum_{k=0}^{N-1} l(x(k|t), u(k|t)).$$

В литературе функция  $l : \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}$  называется стоимостью этапа (stage cost). Предполагается [2], что она непрерывна, а также:

1.  $l(x^*, u^*) = 0$ , т.е. стоимость обращается в нуль в точке равновесия;
2. существует функция  $\alpha_1$  класса  $\mathcal{K}_\infty$ , что выполняется

$$l(x, u) \geq \alpha_1(|x - x^*|) \quad \forall (x, u) \in Z.$$

Таким образом, прогнозирующая задача ОУ для момента времени  $t$  имеет

вид:

$$\mathcal{P}(t) : \quad V(x(t)) = \min_{u(\cdot|t)} \sum_{k=0}^{N-1} l(x(k|t), u(k|t)) + V_f(x(N|t)), \quad (1.5)$$

при условиях

$$x(k+1|t) = f(x(k|t), u(k|t)), \quad k \in \mathbb{I}_{[0, N-1]},$$

$$x(0|t) = x(t),$$

$$(x(k|t), u(k|t)) \in Z, \quad k \in \mathbb{I}_{[0, N-1]},$$

$$x(N|t) \in X_f.$$

В задаче (1.5) обсуждавшиеся выше "терминальные элементы":

- терминальная стоимость  $V_f(x(N|t))$  в критерии качества;
- терминальное ограничение  $x(N|t) \in X_f$ , где  $X_f$  — терминальное множество.

Именно условия на эти элементы обеспечивают устойчивость замкнутой системы несмотря на то, что решается задача с конечным горизонтом (см. раздел 1.4).

Далее используем следующие обозначения:

$u^0(\cdot|t) = \{u^0(0|t), \dots, u^0(N-1|t)\}$  — оптимальное (программное) управление задачи  $\mathcal{P}(t)$ ;

$x^0(\cdot|t) = \{x^0(0|t), \dots, x^0(N|t)\}$  — соответствующая траектория;

$X_N$  — множество всех состояний  $x \in X$ , для которых существует решение задачи (1.5) с  $x(t) = x$ .

**Базовый алгоритм МРС** состоит в следующем:

Для каждого  $t \in \mathbb{I}_{\geq 0}$

1. измерить состояние  $x(t) \in X$  системы (1.1);
2. решить задачу (1.5) с начальным условием  $x(0|t) = x(t)$ , получить ее решение  $u^0(\cdot|t)$ ;
3. подать на вход системы (1.1) управляющее воздействие

$$u_{MPC}(t) := u^0(0|t). \quad (1.6)$$

Таким образом, в каждый момент  $t \in \mathbb{I}_{\geq 0}$  на систему подается управляющее воздействие (1.6), которое неявно зависит от текущего состояния  $x(t)$ . Соответственно, замкнутая система имеет вид

$$x(t+1) = f(x(t), u^0(0|t)), \quad t \in \mathbb{I}_{\geq 0}. \quad (1.7)$$

После 20 лет исследований теория МРС приобрела свои нынешние черты. Согласно [5], все схемы номинального МРС описываются представленным базовым алгоритмом, а все прогнозирующие задачи ОУ в общем виде формулируются как (1.5).

## 1.4 Терминальные элементы и устойчивость замкнутой системы

Схемы МРС зависят от выбора терминальных элементов  $V_f$  и  $X_f$ . Например, самые первые результаты исследований по МРС [2] были получены для состояния равновесия, совпадающего с началом координат  $(x^*, u^*) = (0, 0)$  (т.е.  $f(0, 0) = 0$ ) и предлагали использовать  $X_f = \{0\}$ , т.е. терминальное условие принимало вид  $x(N|t) = 0$ . Понятно, что включение терминальной стоимости в критерий качества в таком подходе не имеет смысла. Прогнозирующая задача ОУ  $\mathcal{P}(t)$  принимает вид

$$\mathcal{P}(t) : \quad V(x(t)) = \min_{u(\cdot|t)} \sum_{k=0}^{N-1} l(x(k|t), u(k|t)),$$

при условиях

$$x(k+1|t) = f(x(k|t), u(k|t)), \quad k \in \mathbb{I}_{[0, N-1]},$$

$$x(0|t) = x(t),$$

$$(x(k|t), u(k|t)) \in Z, \quad k \in \mathbb{I}_{[0, N-1]},$$

$$x(N|t) = 0.$$

Сразу отметим, что ограничения-равенства считаются "плохими" с точки зрения вычислительных методов оптимизации, поэтому дальнейшее развитие теории продолжалось в направлении ослабления этих простейших условий.

В частности, в работе [5] приведены следующие общие требования, которым должны удовлетворять терминальное множество  $X_f$  и терминальная стоимость  $V_f$ . Эти условия дополняют условия 1–2 на функцию стоимости этапа  $l$  и имеют вид:

3. терминальная стоимость  $V_f$  непрерывна на  $X$ ;
4. терминальное множество  $X_f$  замкнуто,  $\{x^*\} \in X_f$ ;
5. существует локальная обратная связь  $k_f : X_f \rightarrow U$ , такая что для  $\forall x \in X_f$  имеет место

- а)  $(x, k_f(x)) \in Z$ , т.е. локальная обратная связь допустима;
- б)  $f(x, k_f(x)) \in X_f$ , т.е. терминальное множество  $X_f$  является положительно инвариантным для системы  $x(t+1) = f(x(t), k_f(x(t)))$ ;
- в)  $V_f(f(x, k_f(x))) - V_f(x) \leq -l(x, k_f(x)) + l(x^*, u^*)$ , откуда следует, что терминальная стоимость  $V_f$  может служить функцией Ляпунова на терминальном множестве  $X_f$ .

В работах [1, 2] для дискретных систем можно найти следующий основной результат:

**Теорема 1.1** Пусть  $x_0 \in X_N$  и выполнены все предположения 1 – 5 относительно функций  $l$ ,  $V_f$  и терминального множества  $X_f$ . Тогда

1. замкнутая система (1.7), полученная в результате применения базового алгоритма МРС, удовлетворяет ограничениям (1.2) для всех  $t \in \mathbb{I}_{\geq 0}$ ;
2. задача (1.5) имеет решение для всех  $t \in \mathbb{I}_{\geq 0}$ ;
3.  $x^*$  — асимптотически устойчивое положение равновесия системы (1.7) с областью притяжения  $X_N$ .

Выбор терминальных элементов  $V_f$  и  $X_f$  зачастую зависит от конкретной задачи, однако существует общий подход к их нахождению и этот способ называется МРС на квази-бесконечном горизонте.

МРС на квази-бесконечном горизонте имеет две отличительные черты:

1. локальная обратная связь является линейной, строится по первому приближению нелинейной системы; она является фиктивной, т.е. для фактического управления системой не используется, однако служит для построения терминальных элементов и при доказательстве основных результатов по асимптотической устойчивости;
2. терминальное ограничение и терминальная стоимость являются квадратичными, их параметры находятся по решению уравнения Ляпунова.

Впервые идея квази-бесконечного МРС появилась в работе [4] для непрерывных нелинейных систем

$$\dot{x} = f(x, u), \quad x(t_0) = x_0, \quad (1.8)$$

с дважды непрерывно дифференцируемой функцией  $f : \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}^n$ , для которых считается, что положение равновесия находится в начале координат, т.е.  $f(0, 0) = 0$ , при этом  $\{0\} \in \text{int } U$ .

В квази-бесконечном МРС все элементы прогнозирующей задачи выбираются квадратичными: функция стоимости этапа задается в виде

$$L(x, u) = \|x\|_Q^2 + \|u\|_R^2 = x^T Q x + u^T R u;$$

терминальная стоимость — квадратичная функция вида:

$$V_f(x) = \|x\|_P^2 = x^T P x; \quad (1.9)$$

терминальное множество — эллипсоид вида:

$$X_f = \{x \in \mathbb{R}^n \mid x^T P x \leq \alpha\}, \quad (1.10)$$

где  $Q \in \mathbb{R}^{n \times n}$ ,  $R \in \mathbb{R}^{r \times r}$ ,  $P \in \mathbb{R}^{n \times n}$  — положительно определенные матрицы. Матрицы  $Q, R$  выбираются исходя из целей процесса управления (скорость сходимости, минимизация затрат и др.) и являются параметрами настройки МРС-регулятора. Матрица  $P$  и константа  $\alpha > 0$  не зависят от выбора  $Q, R$ , а лишь от свойств системы управления. Их выбор описывается ниже.

Будем считать, что первое приближение системы (1.8) в начале координат

$$\dot{x} = Ax + Bu, \quad A = \partial f(0, 0)/\partial x, \quad B = \partial f(0, 0)/\partial u, \quad (1.11)$$

стабилизируемо. Это означает, что существует линейная обратная связь по состоянию  $k_f(x) = Kx$ , что замкнутая система

$$\dot{x} = Ax + Bk_f(x) = (A + BK)x$$

асимптотически устойчива, т.е. матрица  $A_K = A + BK$  является гурвицевой.

Известно [4], что если первое приближение (1.11) системы (1.8) в начале координат стабилизируемо, то

1. уравнение Ляпунова

$$(A_K + kI)^T P + P(A_K + kI) = -(Q + K^T R K) \quad (1.12)$$

имеет единственное положительно определенное решение  $P$ , где параметр  $k \in [0, \infty)$  удовлетворяет неравенству  $k < -\lambda_{\max}(A_K)$ ;

2.  $\exists \alpha \in (0, \infty)$ , определяющая окрестность начала координат

$$X_f^\alpha = \{x \in \mathbb{R}^n \mid x^T P x \leq \alpha\}$$

такую что

- а)  $Kx \in U \quad \forall x \in X_f^\alpha$ , т.е. линейная обратная связь в области  $X_f^\alpha$  не нарушает ограничений на управление;
- б) множество  $X_f^\alpha$  инвариантно для нелинейной системы (1.8), замкнутой локальной линейной обратной связью  $k_f(x) = Kx$ ;
- в) для любого  $x_1 \in X_f^\alpha$  терминальная стоимость  $V_f(x_1)$  ограничивает сверху "хвост" критерия качества для бесконечного горизонта:

$$\int_{t_1}^{\infty} (\|x(t; t_1, x_1, u)\|_Q^2 + \|u(t)\|_R^2) dt \leq x_1^T P x_1.$$

Тогда алгоритм нахождения терминальной функции и терминального множества, согласно [4] состоит в следующем:

1. найти линейную обратную связь  $k_f(x) = Kx$ ;
2. выбрать константу  $k \in [0, \infty)$ , удовлетворяющую неравенству

$$k < -\lambda_{\max}(A_K),$$

и решить уравнение Ляпунова (1.12) для нахождения  $P$ ;

3. найти наибольшее  $\alpha_1$ , для которого имеет место  $Kx \in U \quad \forall x \in X_f^{\alpha_1}$ ;
4. найти наибольшее  $\alpha \in (0, \alpha_1)$ , при котором выполнено неравенство

$$\sup \left( \frac{\|f(x, Kx) - A_K x\|}{\|x\|} \mid x \in X_f^\alpha, x \neq 0 \right) \leq \frac{k \lambda_{\min}(P)}{\|P\|}.$$

В итоге получим матрицу  $P$  и значение  $\alpha$ , которые определяют терминальную функцию (1.9) и терминальное множество (1.10). Эти элементы строятся оффлайн, до начала процесса управления.

**Теорема 1.2** [4] Если первое приближение (1.11) системы (1.8) в начале координат стабилизируемо и прогнозирующая задача оптимального управления  $\mathcal{P}(t)$  имеет решение в  $t = 0$ , то в отсутствие возмущений замкнутая система асимптотически устойчива. Кроме того, если  $X \subseteq \mathbb{R}^n$  — множество состояний  $x_0$ , для которых  $\mathcal{P}(0)$  имеет решение, то  $X$  — область притяжения замкнутой системы.

## 1.5 Выводы

Основная идея алгоритма МРС состоит в том, чтобы в каждый момент времени оптимизировать будущее поведение системы на конечном горизонте и использовать первое значение полученного оптимального (программного) управления в качестве значения обратной связи для этого момента времени. В связи с тем, что эта идея интуитивно понятна практикам и достаточно проста в реализации, она получила широкое распространение в промышленных приложениях [3].

Метод опирается на решение задач ОУ в режиме реального времени, которое в подавляющем числе случаев может быть получено только численно. Несмотря на то, что с привлечением методов алгоритмического дифференцирования численные методы решения задач ОУ достигли значительной эффективности [18, 19] и развитие вычислительной техники позволяет быстро решать достаточно сложные задачи, для существенно нелинейных систем, систем большой размерности, систем с быстро меняющейся динамикой, существующие методы могут оказаться неэффективными или слишком медленными. В связи с этим недостатком, далее в работе предлагается вынести некоторые вычисления из классического алгоритма МРС оффлайн, что позволит повысить производительность систем управления с прогнозирующей моделью.



## ГЛАВА 2

# ИСПОЛЬЗОВАНИЕ МЕТОДОВ ОФФЛАЙН ДИЗАЙНА В СИСТЕМАХ УПРАВЛЕНИЯ С ПРОГНОЗИРУЮЩЕЙ МОДЕЛЬЮ

Метод МРС опирается на решение задач ОУ в режиме реального времени, однако несмотря на то, что существуют методы эффективного решения этих задач [18, 19] и развитие вычислительной техники позволяет быстро решать достаточно сложные задачи, для многих классов нелинейных систем существующие методы могут оказаться неэффективными или слишком медленными. Для преодоления этих недостатков в настоящей работе предлагается вынести некоторые вычисления из классического алгоритма МРС оффлайн. В данной главе будут рассмотрены основные существующие подходы для задач МРС, а также теория методов машинного обучения, которая будет использоваться далее в численных экспериментах.

### 2.1 Основные понятия нейронных сетей

Нейронная сеть представляет собой серию алгоритмов, которые стремятся распознать базовые отношения в наборе данных посредством процесса, который имитирует работу человеческого мозга.

Нейронная сеть основана на наборе связанных единиц или узлов, называемых искусственными нейронами, которые свободно моделируют нейроны в биологическом мозге. Каждое соединение, подобно синапсам в биологическом мозге, может передавать сигнал от одного искусственного нейрона к другому. Искусственный нейрон, который получает сигнал, может обрабатывать его, а затем сигнализировать дополнительные искусственные нейроны, связанные с ним [16].

Ключевой моделью глубокого обучения являются нейронные сети с прямым распространением (многослойные персептроны). Целью данного вида нейронных сетей является аппроксимация некоторой функции  $f^*$ . Например, для классификатора  $y = f^*(x)$  сеть отображает вход  $x$  в категорию  $y$ . Сеть определяет отображение  $y = f(x; \theta)$  и изучает значение параметров  $\theta$ , которые приводят к приближению функции наилучшим образом.

Нейронные сети называются сетями, потому что они обычно представляются

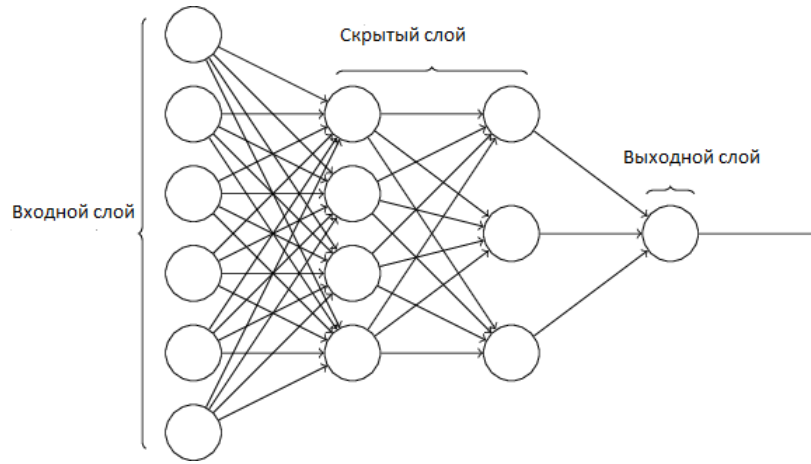


Рис. 2.1: Нейронная сеть со скрытыми слоями.

объединением многих различных функций. Модель связана с ориентированным ациклическим графом (Рис. 2.1), описывающим, как функции состоят вместе. Например, можно было бы иметь три функции  $f^{(1)}$ ,  $f^{(2)}$  и  $f^{(3)}$ , связанные в цепочке, с образованием  $f(x) = f^{(3)}(f^{(2)}(f^{(1)}(x)))$ . Эти цепные структуры являются наиболее часто используемыми структурами нейронных сетей. В этом случае  $f^{(1)}$  называется первым слоем сети,  $f^{(2)}$  называется вторым слоем и т. д. Длина цепочки слоев называется глубиной сети.

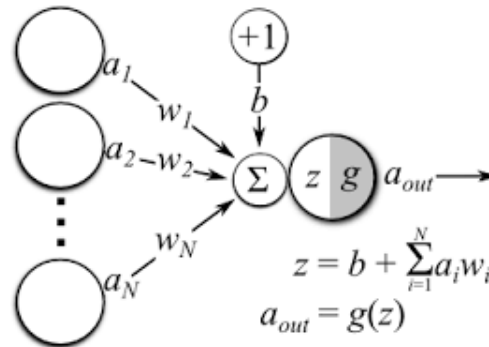


Рис. 2.2: Строение нейрона.

Нейрон обычно получает много одновременных входов. Каждый вход имеет свой собственный относительный вес, который дает входное воздействие, которое ему необходимо для функции суммирования элемента обработки. Эти веса выполняют тот же тип функции, что и различные синаптические силы биологических нейронов. Веса — это адаптивные коэффициенты в сети, которые определяют интенсивность входного сигнала, зарегистрированного искусственным нейроном. На рисунке (2.2) веса обозначены  $w_i$ , значения нейронов предыдущего слоя —  $a_i$ .  $b$  параметр представляет собой смещение для линейного преобразования входных нейронов. Таким образом, получаем значение функции суммирования в виде линейного преобразования  $z = b + \sum_{i=1}^N a_i w_i$ .

Функция  $g$  на рисунке (2.2) — это функция активации нейрона. Цель использования функции активации заключается в том, чтобы позволить суммируемому результату меняться в зависимости от времени. Функцией по умолчанию является нелинейная активационная функция ReLU  $g(x) = \max(0, x)$ , которая рекомендованна для использования с большинством нейронных сетей прямого распространения. Поскольку ReLU почти линейна, она сохраняет многие свойства, которые упрощают оптимизацию линейных моделей с помощью методов, основанных на градиенте. Другими популярными видами функции активации являются сигмоидальная (логистическая) функция  $\sigma(x) = \frac{1}{1+e^{-x}}$ , гиперболический тангенс  $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ .

подавляющее большинство искусственных нейронных сетевых решений проходят обучение с учителем. В этом режиме фактический выход нейронной сети сравнивается с желаемым выходом. Веса, которые обычно начинаются с произвольного начала, затем корректируются сетью, так что следующая итерация или цикл приведут к более близкому совпадению между желаемым и фактическим выходом. Метод обучения пытается минимизировать текущие ошибки всех элементов обработки. Это глобальное сокращение ошибок создается со временем, постоянно изменяя веса ввода до тех пор, пока не будет достигнута приемлемая точность сети. Когда процесс тренировки заканчивается, то уже в онлайн процессах используются эти натренированные параметры и веса.

Алгоритм обучения — алгоритм обратного распространения ошибки, в котором используется стохастический градиентный спуск. Для задачи регрессии чаще всего в качестве функции потерь используется MSE (2.1):

$$L(y_{out}, y_{true}) = \frac{1}{N} \sum_{i=1}^N (y_{out}(i) - y_{true}(i))^2. \quad (2.1)$$

Согласно универсальной теореме аппроксимации — нейронная сеть с одним скрытым слоем может аппроксимировать любую непрерывную функцию многих переменных с любой точностью [17]. Необходимо определить достаточное количество нейронов для достижения этой точности. С любой нелинейностью сеть остаётся универсальным аппроксиматором и при правильном выборе структуры может достаточно точно аппроксимировать функционирование любой непрерывной функции.

## 2.2 Методы дизайна оффлайн регуляторов в системах управления с прогнозирующей моделью

Существует несколько подходов использования методов обучения в МРС системах:

- Явный МРС
- Аппроксимация закона управления
- Использование обучаемой модели для аппроксимации динамики прогнозирующей модели
- Итерационный подход для построения терминального региона и функции из предыдущих итераций

### 2.2.1 Явный МРС

При некоторых слабых предположениях для линейных систем задача оптимизации может быть решена оффлайн, т.е. до начальной реальной процедуры управления [6]. В результате, получается явный закон управления  $u(x)$ . Развитие подхода на нелинейные системы не тривиально, и кроме того даже в линейных случаях существуют проблемы с эффективностью вычислений для метода из [6].

В момент времени  $t$ , вычисляем состояние  $x(t)$ , решаем линейную задачу с линейными ограничениями

$$V(x) = \min_{u(\cdot|t)} \sum_{k=t}^{t+N-1} L(x(k|t), u(k|t)) + F(x(t+N|t)), \quad (2.2)$$

при условиях

$$x(k+1|t) = Ax(k|t) + Bu(k|t), \quad t \leq k \leq t+N-1,$$

$$x(t|t) = x(t), \quad t \leq k \leq t+N-1,$$

$$C_x x(k|t) \leq d_x, \quad t \leq k \leq t+N-1,$$

$$C_u u(k|t) \leq d_u, \quad t \leq k \leq t+N-1,$$

$$C^f x(t+N|t) \leq d^f,$$

с квадратичной функцией стоимости этапа и квадратичной терминальной функ-

цией

$$L(x(t), u(t)) = x(t)^T Q x(t) + u(t)^T R u(t), \quad Q, R > 0, \quad F(x(t)) = x(t)^T P x(t).$$

Можно переписать задачу 2.2 в виде задачи квадратичного программирования. Для этого обозначим

$$X := [x^T(t+1|t), \dots, x^T(t+N|t)]^T,$$

$$U := [u^T(t+1|t), \dots, u^T(t+N-1|t)]^T,$$

Перепишем функцию стоимости

$$F(x(t), U) = x^T(t) Q x(t) + X^T \tilde{Q} X + U^T \tilde{R} U, \quad (2.3)$$

где  $\tilde{Q} = \text{diag}(Q, \dots, Q, P) \in \mathbb{R}^{n \times (N+1)}$ ,  $\tilde{R} = \text{diag}(R, \dots, R) \in \mathbb{R}^{m \times N}$ .

Перепишем динамику системы:

$$x(t+k|t) = A^k x(t) + \sum_{j=0}^{k-1} A^j B u(t+k-j-1|t), \quad k = 1, \dots, N,$$

Тогда

$$X = \begin{bmatrix} A \\ A^2 \\ \vdots \\ A^N \end{bmatrix} x(t) + \begin{bmatrix} B & 0 & \dots & 0 \\ AB & B & 0 & \dots & 0 \\ \ddots & & & & \\ A^{N-1}B & A^{N-2}B & \dots & \dots & B \end{bmatrix} U, \quad (2.4)$$

или, введя обозначения  $\Omega = \begin{bmatrix} A \\ A^2 \\ \vdots \\ A^N \end{bmatrix}$  и  $\Gamma = \begin{bmatrix} B & 0 & \dots & 0 \\ AB & B & 0 & \dots & 0 \\ \ddots & & & & \\ A^{N-1}B & A^{N-2}B & \dots & \dots & B \end{bmatrix}$  сократим

запись  $X = \Omega x(t) + \Gamma U$ .

С помощью (2.4) перепишем (2.3) :

$$J(x(t), U) = \frac{1}{2} x^T(t) Y x(t) + \frac{1}{2} U^T H U + x^T(t) F U,$$

где

$$Y = 2(Q + \Omega^T \tilde{Q} \Omega),$$

$$H = 2(\Gamma^T \tilde{Q} \Gamma + \tilde{R}),$$

$$F = 2\Omega^T \tilde{Q} \Gamma.$$

Аналогично, для всех ограничений можно провести такие же преобразования

$$GU \leq W + Ex(t),$$

где

$$G = \text{diag}(C_x, \dots, C_x)\Gamma, \quad W = [d_x, \dots, d_x]^T, \quad E = \text{diag}(C_x, \dots, C_x)\Omega.$$

Тогда, задача (2.2) примет вид

$$\min_U \frac{1}{2}U^T H U + x^T F U + \frac{1}{2}x^T Y x, \quad (2.5)$$

при условии

$$GU \leq W + Ex(t).$$

Применим замену переменных вида

$$z := U + H^{-1}F^T x,$$

Нетрудно заметить, что  $H^{-1}$  — положительно определенная матрица. Тогда (2.5) примет вид

$$\min_z \frac{1}{2}z^T H z + \frac{1}{2}x^T \tilde{Y} x, \quad (2.6)$$

при условиях

$$\begin{aligned} Gz &\leq W + Sx, \\ \tilde{Y} &:= Y - FH^{-1}F^T, \\ S &:= E + GH^{-1}F^T. \end{aligned}$$

Далее эта задача решается с помощью теории выпуклой оптимизации через условия Каруша-Куна-Такера (ККТ). Так как задача (строго) выпуклая с допустимым множеством с непустой внутренней частью (по предположениям), то условия Слейтера выполняются. Оптимальное решение единственное и характеризуется условиями ККТ [6].

Явный МРС решает задачу для всех состояний, таким образом все пространство состояний делится на области, где в каждой области для состояния есть явная функция управления.

Алгоритм нахождения явных функций управления [6]:

1. Взять любой  $x_0 \in \mathbb{X}$ .
2. Решить задачу (2.6) с начальным условием  $x = x_0$ .
3. Определить активные ограничения для оптимизационной задачи (2.6).

4. Вычислить критическую область по активным ограничениям и вычислить функцию управления для этой области.
5. Перейти к новому  $x_0$ .

Главный недостаток этого метода состоит в том, что количество областей может быть достаточно большим, что в онлайн процедуре может плохо сказываться на производительности. Так как в каждый момент времени нужно будет искать к какому из регионов относится текущее состояние, чтобы определить управление для него.

### 2.2.2 Аппроксимация закона управления

Существует несколько подходов к получению аппроксимативного решения для оптимизации МРС. Для линейных систем в [7] алгоритм обучения представлен дополнительными ограничениями для обеспечения стабильности и накладывает ограничения на ошибку аппроксимации. Одним из подходов к аппроксимации МРС является выпуклое многопараметрическое нелинейное программирование [9], где вычисляется субоптимальная аппроксимация закона управления МРС. Другой подход — аппроксимировать МРС с помощью методов машинного обучения. Это делают нейронные сети в [10], [11], [12]. Эти методы не гарантируют устойчивость или удовлетворение ограничениям для аппроксимационного МРС, что особенно важно, если рассматривать жесткие ограничения на состояния. В [8] используется метод опорных векторов (SVM) для аппроксимации МРС. Устойчивость и удовлетворение ограничений могут быть гарантированы для произвольных ошибок малого приближения, основанных на присущих свойствам устойчивости. В [13] аппроксимируется МРС с липшицевым сужающим ограничением, что обеспечивает устойчивость при не исчезающих ошибках аппроксимации. Ошибка аппроксимации, выведенная в [8], [13], обычно не достижима для практического применения.

В [11] находят аппроксимирующий закон управления

$$u^0(t) = \gamma^0(x(t)) \in U,$$

где  $u^0(t)$  — первый вектор последовательности управления, которая минимизирует стоимость

$$J(x(t), u(t)) = \sum_{k=t}^{t+N-1} l(x(k), u(k)) + a\|x(t+N)\|_P^2, \quad t \geq 0.$$

Стоимость формируется из стоимости переходов на горизонте планирования

длины  $N$  и терминальной функции. Аппроксимация закона управления происходит с помощью нейронной сети:  $m$  параллельных сетей с одним выходным параметром, состоящий из одного скрытого слоя с  $v_j$  нейронами на скрытом слое для  $j = 1..m$  сети и линейными активационными функциями.

Для каждой функции  $\hat{\gamma}_j^{(v_j)}$  нужно найти количество нейронов  $v_1^*, \dots, v_m^*$ , такое что

$$\min_{w_j} \max_{x_t \in X} |\gamma_j^0(x(t)) - \gamma_j^{(v_j)}(x(t), w_j)| \leq \frac{\epsilon}{\sqrt{m}}, \quad j = 1, \dots, m. \quad (2.7)$$

Процедура нахождения количества нейронов представляется таким образом: для каждого  $j$  увеличиваем  $v_j$ , пока (2.7) не станет верным. Также приводится в статье теорема, в которой утверждается, что для любой функции управления  $\gamma_j^0(x_t)$  число параметров, необходимых для достижения погрешности приближения  $L_2$  или  $L_\infty$  порядка  $O(\frac{1}{v_j})$ , равно  $O(v_j n)$ , которое растет линейно с размерностью  $n$  вектора состояния.

В работе [13] уже вводятся некоторые гарантии устойчивости метода на основании предположения о непрерывности по Липшицу правой части динамической системы. Система представляется следующим образом:

$$x(t+1) = f(x(t), u(t), \xi_t), \quad t \geq 0, \quad x_0 = \tilde{x},$$

где  $\xi_t \in \mathbb{R}^r$  — возмущение системы. Номинальная система вводится для дизайна управления таким образом:

$$x(t+1) = \hat{f}(x(t), u(t)) + d_t, \quad t \geq 0, \quad x_0 = \tilde{x},$$

где  $d_t = f(x(t), u(t), \xi_t) - \hat{f}(x(t), u(t))$ . Предполагают непрерывность по Липшицу для функции  $\hat{f}$  относительно  $x$  с константой  $L_{f_x}$ , а также относительно управления  $u$  и предполагают, что существует функция  $K$  класса, такая что

$$|\hat{f}(x(t), u(t)) - \hat{f}(x(t), u'(t))| \leq \eta_u(|u(t) - u'(t)|) \quad \forall x(t) \in X, \quad \forall u(t) \in U, \quad u'(t) \in U.$$

Также вводится предположение об ограниченности возмущения и то, что верно  $|d_t| \leq \mu(|\xi_t|)$   $t \geq 0$ , а также  $d_t \in D = B^m(\bar{d})$ ,  $\bar{d} \in \mathbb{R}_{\geq 0} < \infty$ . Предполагается, что для системы существует управление  $k(x(t)) \in U$ , которое является стабилизирующим относительно состояния.

Накладываются дополнительные ограничения на погрешности относительно состояния  $q_t \in Q = B^m(\bar{q})$  и управления  $v_t \in V = B^m(\bar{v})$ , где погрешность аппроксимации состояния  $q_t = \xi_t - x(t)$  и управления  $v_t = k^*(x_t) - k^*(\xi_t)$ . Тогда



система уже переписывается в таком виде:

$$x(t+1) = \hat{f}(x(t), k(x(t) + q_t) + v_t) + w_t, \quad x_0 = \tilde{x}, \quad t \geq 0.$$

И устойчивость системы доказывается, если верно следующее, что погрешности аппроксимации состояния и управления совместно с возмущением ограничены изначальным возмущением системы  $\bar{d}_q + \bar{d}_v + \bar{d}_w \leq \bar{d}$ . В данной статье рассматривались аппроксимации закона управления с помощью нейронной сети, которая показала довольно хорошие результаты.

В работе с использованием результатов [20] исследуются условия, при которых, несмотря на ошибки аппроксимации, гарантируется выполнение ограничений и асимптотическая устойчивость замкнутой системы. На основе этого источника будет продолжаться исследование аппроксимации закона управления с помощью нейронной сети и результаты будут продемонстрированы в главе 3.

### 2.2.3 Аппроксимация динамики системы с прогнозирующей моделью

Существует методы для аппроксимации динамики системы, в работе [14] был предложен доказуемо безопасный и робастный метод управления, основанный на обучении для систем с прогнозирующей моделью, который называется LVMPC. LVMPC изучает динамику системы по предоставленным точкам, также имеется возможность обновления динамики системы с помощью новых измерений, обеспечивает при этом безопасность и устойчивость, используя теорию из робастного MPC, чтобы проверить, примененное управление сохраняет ли номинальную модель устойчивой, когда она подвержена неопределенности.

Динамика системы представляется в таком виде:

$$x(t+1) = Ax(t) + Bu(t) + g(x(t), u(t))$$

где  $g(x(t), u(t))$  описывает несмоделированную динамику, которая по предположению ограничена и лежит в политопе  $W$ .

Данный метод вводит дополнительную систему, которая обучается на данных и имеет такой вид:

$$\tilde{x}(t+1) = A\tilde{x}(t) + B\tilde{u}(t) + O_n(\tilde{x}(t), \tilde{u}(t))$$

где  $O_n$  — зависящая от времени функция, которая обучается с помощью любого из статистических методов.

Вся теория устойчивости строится на том, что наша система представляется

в виде робастного МРС с обученной динамикой на известных вычисленных точках, и может адаптироваться к новым полученным точкам, чтобы улучшать точность обученной динамики системы.

Задача формулируется в таком виде:

$$V_n(x()) = \min_{c, \theta} \phi_n(\theta, \tilde{x}(t), \dots, \tilde{x}(t+N), \tilde{u}(t), \dots, \tilde{u}(t+N-1))$$

при условиях

$$\tilde{x}(t) = x_t, \quad \bar{x}(t) = x_t,$$

$$\tilde{x}(t+i+1) = A\tilde{x}(t+i) + B\tilde{u}(t+i) + O_n(\tilde{x}(t+i), \tilde{u}(t+i)),$$

$$\bar{x}(t+i+1) = A\bar{x}(t+i) + B\tilde{u}(t+i),$$

$$\tilde{u}(t+i) = K\bar{x}(t+i) + c_{n+i},$$

$$\bar{x}(t+i+1) \in X \ominus R_i, \quad \tilde{u}(t+i) \in U \ominus KR_i,$$

$$(\bar{x}(t+N) \in \Omega \ominus (R_N \times \{0\})),$$

где  $\Omega$  — допустимое инвариантное робастное множество таких точек, что любая траектория системы с начальным условием, выбранным из этого множества и с управлением  $u(t)$ , остается в множестве для любой последовательности ограниченного возмущения, удовлетворяя ограничениям на состояние и управление.

Моделируется точка устойчивого состояния  $\bar{x}_s = \Lambda\theta$  и управления  $\bar{u}_s = \Psi\theta$ , где  $\theta \in \mathbb{R}^m$  и  $\Lambda \in \mathbb{R}^{n \times m}$ ,  $\Psi \in \mathbb{R}^{m \times m}$  — параметры моделирования, параметры будут описывать точку равновесия для системы, если  $A + BK$  устойчива по Шуру при управлении

$$\bar{u}(t) = K(\bar{x}(t) - \bar{x}_s) + \bar{u}_s = K\bar{x}(t) + (\Psi - K\Lambda)\theta.$$

Множества  $R_0 = \{0\}$  и  $R_i = \oplus_{j=0}^{i-1} (A + BK)^j W$  необходимы для робастного МРС, они представляют собой сужающиеся ограничения. И тогда с помощью данного неравенства формулируется удовлетворение ограничений:

$$\Omega \subseteq \{(\bar{x}, \theta) : \bar{x} \in X; \Lambda\theta \in X; K\bar{x} + (\Psi - K\Lambda)\theta \in U; \Psi\theta \in U\}.$$

А с помощью данного неравенства инвариантность возмущения:

$$\begin{bmatrix} A + BK & B(\Psi - K\Lambda) \\ 0 & \mathbb{I} \end{bmatrix} \Omega \oplus (W \times \{0\}) \subseteq \Omega.$$

Устойчивость данного метода основана на робастном МРС. Были получены хорошие результаты относительно предсказанных траекторий и быстрой сходимости к точке устойчивого состояния, однако по производительности уступало

нелинейному МРС с нейросетевыми регуляторами.

### 2.2.4 Итерационный подход

Существует итерационный подход для построения МРС регулятора [15]. Данный метод используется для повторяющихся задач, где эталонная траектория неизвестна. Например, системы для гоночных и раллийных машин, где среда и динамика сложны и не совсем известны. Регулятор имеет справочную информацию и способен улучшать свою эффективность, изучая предыдущие итерации. Для обеспечения рекурсивной выполнимости и неубывающей эффективности на каждой итерации используются безопасное терминальное множество и терминальная функция стоимости. Построение данного регулятора обеспечивает тот факт, что функция стоимости убывает с каждой итерацией, также из удовлетворения ограничений на  $j - 1$  итерации следует удовлетворение ограничений на  $j$  итерации и точка равновесия замкнутой системы асимптотически устойчива. Оптимальность траекторий построенных этим регулятором доказывается для выпуклых задач.

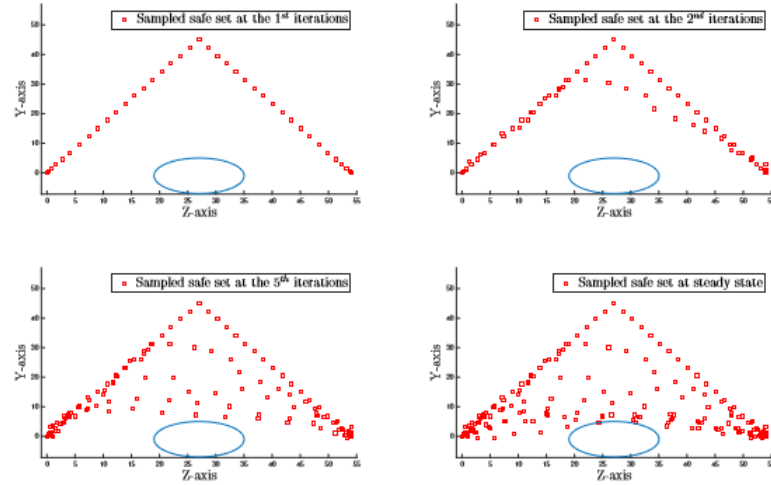


Рис. 2.3: Построение сэмплированного безопасного множества.

Управляемое на  $N$  шагов вперед множество по отношению к  $S$

$$K_j(S) = Pre(K_{j-1}(S)) \cap X, \quad K_0(S) = S, j \in \{1, \dots, N\}, \quad (2.8)$$

где  $Pre(S) = \{x \in \mathbb{R}^n : \exists u \in U \text{ s.t. } f(x, u) \in S\}$ .

Сэмплированное безопасное множество на итерации  $j$

$$SS^j = \{\cup_{i \in M^j} \cup_{t=0}^{\infty} x^i(t)\},$$

где  $M^j = \{k \in [0, j] : \lim_{t \rightarrow \infty} x^k(t) = x_F\}$ .

$SS^j$  — это множество всех траекторий на итерации  $i$  для  $i \in M^j$ . Как выглядит это множество показано на рисунке 2.3.

Стоимость на сэмплированном безопасном множестве вводится таким образом, чтобы штрафовались те состояния, которые не находятся в нашем построенном безопасном множестве на предыдущих итерациях

$$Q^j(x) = \begin{cases} \min_{(i,t) \in F^j(x)} J_{t \rightarrow \infty}^i, & x(t) \in SS^j \\ +\infty, & x \notin SS^j \end{cases},$$

где  $\forall x(t) \in SS_j$ ,  $Q_j(x) = J_{t^* \rightarrow \infty}^{i^*}(x(t)) = \sum_{k=t^*}^{\infty} l(x^{i^*}(k), u^{i^*}(k))$ .

$$J_{t \rightarrow t+N}(x_t^j) = \min_{u_{t|t}, \dots, u_{t+N-1|t}} \left[ \sum_{k=t}^{t+N-1} l(x(k|t), u(k|t)) + Q^{j-1}(x(t+N|t)) \right],$$

при условии

$$x(k+1|t) = f(x(k|t), u(k|t)) \quad \forall k \in [t, \dots, t+N-1],$$

$$x(k|t) \in X, \quad u(k|t) \in U, \quad \forall k \in [t, \dots, t+N-1],$$

$$x(t+N|t) \in SS^{j-1},$$

$$x(t|t) = x^j(t).$$

Данный метод обладает рекурсивной выполнимостью, устойчивостью и сходимостью. Но предлагаемый подход является дорогостоящим с точки зрения вычисления даже для линейной системы, поскольку регулятор должен решать задачу смешанного целочисленного программирования в каждый момент времени. Есть улучшения данного подхода с использованием параллельных вычислений, а также попытки сделать терминальные ограничения более выпуклыми.

## 2.3 Выводы

Существует ряд подходов для вынесения вычислений оффлайн для систем с прогнозирующей моделью. Каждый метод используется для своего типа задач. Для линейных задач с небольшим количеством областей подходит явный МРС. Для задач с нелинейной динамикой будет более эффективным метод аппроксимации закона управления. Для задач с неизвестной динамикой, представленной в виде набора точек, лучше использовать аппроксимация динамики системы. Для задач с неизвестной динамикой или изменяющей средой подходят итеративные

методы управления МРС.

В данной работе вынести вычисления оффлайн предлагается для нелинейных систем, поэтому более эффективное направление — использование аппроксимации закона управления. Для этого будут использоваться нейронные сети, как универсальный аппроксиматор непрерывных функций.

## ГЛАВА 3

### ПОСТРОЕНИЕ НЕЙРОСЕТЕВОГО РЕГУЛЯТОРА

В настоящей главе демонстрируются потенциальные возможности вынесения вычисления оффлайн для нелинейных систем с прогнозирующей моделью. Как было показано в предыдущей главе, наиболее перспективным направлением для рассматриваемого типа задач является использование аппроксимации закона управления. Для этого будут использоваться нейронные сети, как универсальный аппроксиматор непрерывных функций.

В этой главе описаны эксперименты и результаты использования нейросетевых регуляторов, а также методы построения областей насыщения управления с помощью метода опорных векторов. Проводится сравнение производительности для стандартного МРС-регулятора, решающего онлайн прогнозирующие задачи оптимального управления, и нейросетевого регулятора, использующего нейронную сеть, которая заменит процедуру решения прогнозирующей задачи. А также рассматриваются способы сэмплирования точек для уменьшения времени на сбор данных для обучения нейронной сети.

#### 3.1 Описание иллюстративного примера

Рассмотрим нелинейную дискретную динамическую систему

$$x(t+1) = f(x(t), u(t)), \quad (3.1)$$

где  $x(t) \in \mathbb{R}^n$  — вектор состояний,  $u(t) \in \mathbb{R}^m$  — вектор управления. Предполагается, что  $f(0, 0) = 0$ .

Для системы (3.1) рассматриваются ограничения вида

$$X = \{x \in \mathbb{R}^n | Hx \leq 1_p\}, \quad U = \{u \in \mathbb{R}^m | Lu \leq 1_q\}, \quad (3.2)$$

поэтому в каждый момент времени требуется выполнение включения

$$(x(t), u(t)) \in X \times U \quad \forall t \geq 0.$$

Будем минимизировать функцию стоимости

$$\sum_{k=0}^{N-1} l(x, u) + V_f(X(N)),$$

содержащую сумму стоимостей этапов и терминальную стоимость, которая, напомним, строится специальным образом, обеспечивая вместе с терминальным ограничением асимптотическую устойчивость замкнутой системы.

Предлагаемые подходы будут демонстрироваться на примере из [10], только будут изменены ограничения на управление, чтобы лучше продемонстрировать насыщенные управления.

Нелинейная система в дискретном виде представляет собой

$$\dot{x}_1(t) = -x_2(t) + 0.5(1 + x_1(t))u(t), \quad (3.3)$$

$$\dot{x}_2(t) = x_1(t) + 0.5(1 - 4x_2(t))u(t),$$

с ограничениями на фазовые и управляющие переменные

$$x \in X = \{-2 \leq x_1 \leq 0.5, -1 \leq x_2 \leq 2\}, \quad (3.4)$$

$$u \in U = \{u \in \mathbb{R} \mid |u| \leq 1\},$$

Шаг дискретизации  $\delta = 0.1$ . Данная система имеет стабилизируемое первое приближение в начале координат, которое имеет вид

$$x(t+1) = Ax(t) + Bu, \quad A = \begin{bmatrix} 1 & -0.05 \\ 0.05 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0.025 \\ 0.025 \end{bmatrix}.$$

Для системы (3.3) выбрана квадратичная функция стоимости этапа:

$$l(x, u) = ||u||^2 + 0.1||x||^2,$$

а терминальная стоимость и терминальное множество найдены методом квази-бесконечного МРС из главы 1:

$$V_f(x(t)) = x(t)^T P x(t),$$

$$X_f = \{x \in \mathbb{R}^n \mid x(t)^T P x \leq \alpha\},$$

где  $\alpha = 0.2129$ ,

$$P = \begin{bmatrix} 48.3043 & -2.53 \\ -2.53 & 70.8191 \end{bmatrix}$$

Линейная локальная обратная связь для терминального множества имеет вид

$$k_f(x(t)) = Kx(t) = 0.3799x_1(t) - 0.4201x_2(t).$$

Тогда прогнозирующая задача оптимального управления

$$V(x(t)) = \min_{u(\cdot|t)} \sum_{k=0}^{N-1} (||u(k|t)||^2 + 0.1||x(k|t)||^2) + V_f(x(t)), \quad (3.5)$$

при условиях

$$x_1(k+1|t) = x_1(k|t) + \delta(-x_2(k|t) + 0.5(1 + x_1(k|t))u(k|t)),$$

$$x_2(k+1|t) = x_2(k|t) + \delta(0.5(1 - 4x_2(k|t))u(k|t)),$$

$$-2 \leq x_1(k|t) \leq 0.5, \quad -1 \leq x_2(k|t) \leq 2,$$

$$|u(k|t)| \leq 1,$$

$$x(N|t) \in X_f.$$

Таким образом, для системы (3.3) определены все параметры МРС-регулятора.

Задача (3.5) будет решаться онлайн при стандартной реализации алгоритма МРС, или использоваться для подготовки датасета для нейрорегулятора.

## 3.2 Подход к решению

Алгоритм МРС для стабилизации положения равновесия  $x = 0$  системы (3.3) состоит в следующем: Для каждого  $t = 0, 1, \dots$

1. измерить состояние  $x(t) \in X$  системы (3.3);
2. решить задачу (3.5) с начальным условием  $x(0|t) = x(t)$ , получить  $u^0(\cdot|t)$ ;
3. подать на вход системы (3.3) управляющее воздействие  $u_{MPC}(x(t)) := u^0(0|t)$ .

Поскольку для построенных функций  $l$ ,  $V_f$  и множества  $X_f$  выполнены требования теоремы 1.1, замкнутая система

$$x(t+1) = f(x(t), u_{MPC}(x(t))), \quad t = 0, 1, \dots,$$



асимптотически устойчива с областью притяжения

$$X_0 = \{x_0 : V(x_0) < +\infty\} \subseteq X,$$

состоящей из всех точек  $x_0 \in X$ , для которых задача (3.5) с  $x(0|0) = x_0$  имеет решение [2].

Как отмечено выше, шаг 2 приведенного алгоритма может оказаться достаточно трудоемким. В связи с этим в настоящей работе предлагается вместо онлайн решения прогнозирующей задачи (3.5) численными методами оптимального управления использовать до начала процесса управления ряд методов машинного обучения, которые на основе обучающей выборки  $(x, u_{MPC}(x)) \in X \times U$  будут строить приближенные значения  $\bar{u}_{MPC}(x(t))$  обратной связи  $u_{MPC}(x)$ ,  $x \in X$ , для текущих состояний  $x(t)$ .

Далее будем рассматривать систему (3.3) в контексте робастного MPC [?], поскольку ошибка аппроксимации  $\bar{u}_{MPC} - u_{MPC}$  может рассматриваться как возмущение специального вида в динамике замкнутой системы, которая теперь имеет вид:

$$x(t+1) = f(x(t), u_{MPC}(x(t)) + d(t)),$$

где  $d(t) \in D$ ,  $D = \{d \in \mathbb{R}^m \mid \|d\|_\infty \leq \eta\}$  — множество допустимых ошибок реализации управления,  $\eta > 0$ .

В работе с использованием результатов [20] исследуются условия, при которых, несмотря на ошибки аппроксимации, гарантируется выполнение ограничений и асимптотическая устойчивость замкнутой системы. При этом необходимые для выполнения этих свойств предположения следующие:

1. Локальная инкрементируемая стабилизуемость: Существует такая обратная связь  $k : X \times X \times U \rightarrow \mathbb{R}^m$ ,  $\delta$ -функция Ляпунова  $V_\delta : X \times X \times U \rightarrow \mathbb{R}_{\geq 0}$ , непрерывная по первому аргументу и удовлетворяющая  $V_\delta(x, x, v) = 0 \ \forall x \in X, \forall v \in U$ , и параметры  $c_{\delta, l}, c_{\delta, u}, \delta_{loc}, k_{\max} \in \mathbb{R} > 0, \rho \in (0, 1)$ , такие что следующие неравенства выполняются для всех  $(x, z, v) \in X \times X \times U, (z^+, v^+) \in X \times U$  с  $V_\delta(x, z, v) \leq \delta_{loc}$ :

$$c_{\delta, l} \|x - z\|^2 \leq V_\delta(x, z, v) \leq c_{\delta, u},$$

$$\|k(x, z, v) - v\| \geq k_{\max} \|x - z\|,$$

$$V_\delta(x^+, z^+, v^+) \geq \rho V_\delta(x, z, v),$$

где  $x^+ = f(x, k(x, z, v)), z^+ = f(z, v)$ .

2. Локальная непрерывность по Липшицу. Существует  $\lambda \in \mathbb{R}$ , такая что при

$\forall x \in X, \forall u \in U, \forall u + d \in U$  выполняется неравенство

$$\|f(x, u + d) - f(x, u)\| \leq \lambda \|d\|_\infty.$$

3. Радиус множества  $D$  удовлетворяет неравенству  $\eta \leq \frac{1}{\lambda} \sqrt{\frac{\delta_{loc}}{c_{\delta,u}}}$ .

Для робастного выполнения ограничений используются сужение ограничений в виде увеличивающейся трубки: чем ближе к началу координат, тем более увеличивающеесяся трубка для ограничений. Поэтому заменим ограничения (3.2) на "сужение" вида:

$$\bar{X}_k = (1 - \epsilon_k)X = \{x \in \mathbb{R}^n : Hx \leq (1 - \epsilon_k)1_p\},$$

$$\bar{U}_k = (1 - \epsilon_k)U = \{u \in \mathbb{R}^m : Lu \leq (1 - \epsilon_k)1_q\}.$$

где

$$\epsilon_k = \epsilon \left( \frac{1 - \sqrt{\rho^k}}{1 - \sqrt{\rho}} \right), \quad k = 0, 1, \dots, N;$$

$$\epsilon = \eta \lambda \sqrt{\frac{c_{\delta,u}}{c_{\delta,l}}} \max\{\|H\|_\infty, \|L\|_\infty k_{\max}\}.$$

В настоящей работе для построения аппроксимации  $\bar{u}_{MPC}(x(t))$  применяются метод опорных векторов (SVM) и нейронные сети.

Метод опорных векторов с радиальной базисной функции Гаусса [8] используется, во-первых, для выделения и аппроксимации области притяжения  $X_0$  системы (3.1). Это позволяет эффективно обрабатывать текущие состояния динамической системы и не допускать выход за пределы области притяжения при управлении с помощью приближенных обратных связей. Во-вторых, метод опорных векторов применяется для многоклассовой классификации с целью выделения областей насыщения управления.

В областях, в которых обратная связь  $u_{MPC}$  принимает промежуточные значения, она аппроксимируется с использованием нейронной сети.

Обучение нейронной сети и классификация на основе SVM дает приближенное управления типа обратной связи  $\bar{u}_{MPC}(x)$ ,  $x \in X_0$ . Система, замкнутая обратной связью  $\bar{u}_{MPC}(x)$ ,  $x \in X_0$ , имеет вид

$$x(t + 1) = f(x(t), \bar{u}_{MPC}(x(t))), \quad t = 0, 1, \dots$$

Далее в численных экспериментах проводится сравнение приближенных законов управления, обученных на равномерной сетке, на сетке, полученной на основе равномерно распределенных последовательностей, и на случайной сетке с

увеличением объема обучающей выборки в окрестности положения равновесия.

### 3.3 Базовая реализация нейросетевого регулятора

Для численных экспериментов в рассматриваемом примере стабилизации системы (3.3) был выбран горизонт планирования  $N = 30$ . В качестве начального состояния выбрана точка  $x_0 = (0.4; 1.5)$ .

После реализации стандартной процедуры МРС для системы (3.3), описанной в главе 1, получена траектория замкнутой системы, представленная на рисунке 3.1. Соответствующая реализация МРС управления  $u_{MPC}(t)$  Изображена на рисунке 3.2.

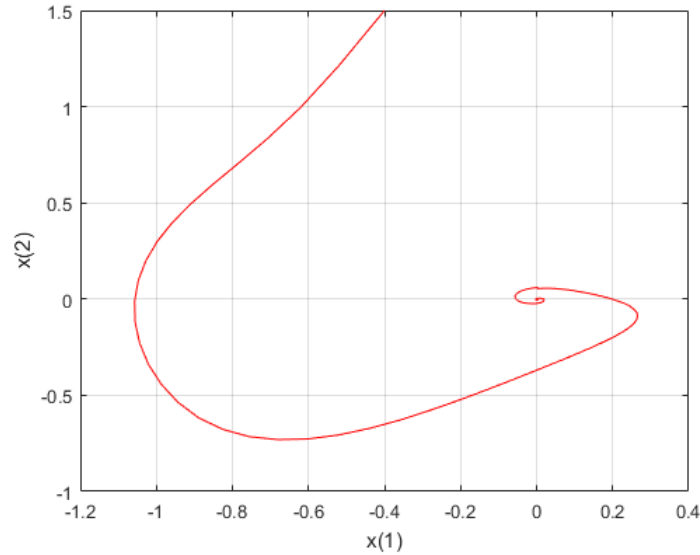


Рис. 3.1: Траектория замкнутой системы в примере (3.3) при применении стандартной процедуры МРС

Базовая реализация нейрорегулятора включает создание нейронной сети для аппроксимации обратной связи  $\bar{u}_{MPC}(x)$  для системы (3.3).

Для построения аппроксимации обратной связи необходимо создать обучающую выборку для нейронной сети. Для этого в первом варианте реализации алгоритма нейросетевого управления построим равномерную сетку на множестве (3.4) допустимых состояний. Шаг равномерной сетки выбран равным 0.05.

В узлах равномерной сетки  $x^i, i \in I$ , вычислим решение прогнозирующей задачи (3.5) с начальным условием  $x(0|0) = x^i$ . Запомним значение  $u^i := u^0(0|0, x^i)$ .

На рисунке 3.3 показано векторное поле сетки с шагом 0.1, а не 0.05, для более хорошей демонстрации, в котором видно как происходит переход из состояния  $x(0|0) = x^i$  в состояние  $x(1|0) = f(x^i, u^i)$ , используя полученные значения управ-

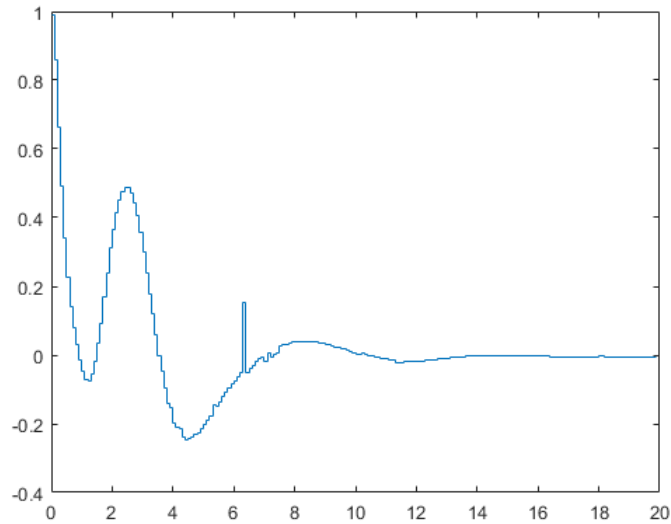


Рис. 3.2: Реализация обратной связи для системы (3.3) при применении стандартной процедуры МРС

ления (обратной связи). На рисунке видно, что для большинства точек следующее состояние находится в пределах допустимых значений. Однако, существуют также точки, где следующее состояние выходит за пределы области обучения, и в этом случае можно получить неустойчивое решение.

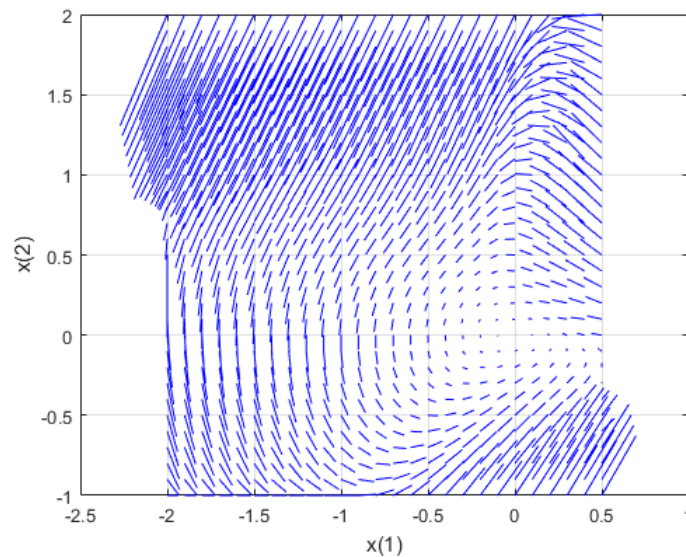


Рис. 3.3: Векторное поле для сетки с шагом 0.1.

Продолжая анализ, на рисунке 3.4 представлен график  $u^i$  в зависимости от  $x^i$ . Можно увидеть множество одинаковых управлений для состояний.

Важно отметить, что сетку значений состояния в окрестности начала координат нужно сделать достаточно детальной, чтобы не получалось закливание регулятора, т.е. не происходил выбор одного и того же управления около

нуля, которое не способствует приближению регулятора к началу координат.

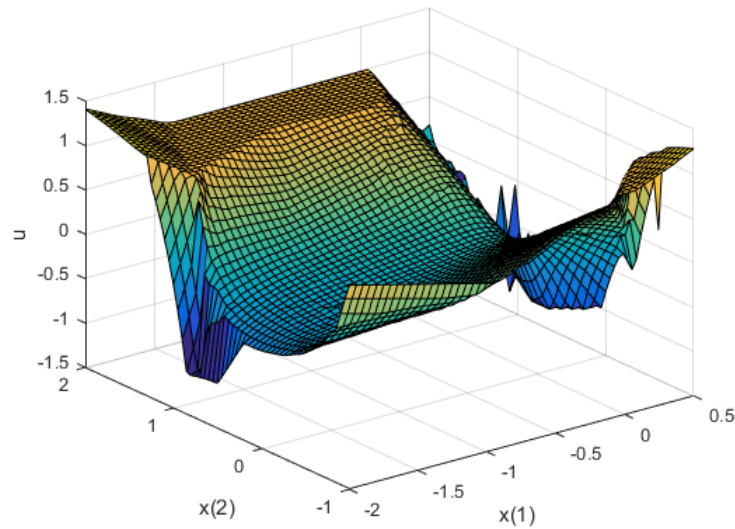


Рис. 3.4: Управление для равномерной сетки значений состояний для обучения нейронной сети

Перейдем к построению нейронной сети. Для выбора количества нейронов на скрытом слое примем подход работы [11], где использовался итеративный процесс для поиска оптимального количества нейронов. Алгоритм выбирает некоторое начальное значение количества нейронов, строит сеть, которая аппроксимирует функцию, а затем для каждого узла сетки вычисляет значение приближенной обратной связи и обратной связи, полученной из стандартной процедуры МРС, пока не будет достигнута необходимая точность

$$\|\bar{u}(t) - u(t)\| < \frac{\epsilon}{\sqrt{m}},$$

где, напомним,  $m$  — размерность вектора управления.

В таблице 3.1 представлена зависимость точности от количества нейронов на скрытом слое. Ошибка аппроксимации вычислялась как средняя ошибка для точек из валидационного множества. Валидационное множество состояло из точек, которые не участвовали в обучении. Для достижения точности  $5 \cdot 10^{-3}$  понадобилось 25 нейронов.

На рисунке 3.5 представлены траектории замкнутой системы (3.3), полученные при помощи стандартной процедуры МРС (красная кривая) и при помощи построенного нейросетевого регулятора (синяя кривая). Как видно из рисунка, получились достаточно близкие траектории, асимптотически стремящиеся к началу координат.

Подведем итоги представленной базовой реализации нейрорегулятора. В представленном подходе необходимо сначала вычислить значения управления в

Таблица 3.1: Таблица зависимости точности аппроксимации от количества нейронов на скрытом слое

Количество нейронов	Ошибка аппроксимации
5	0.021
8	0.0092
10	0.008
15	0.0075
20	0.0063
25	0.0049

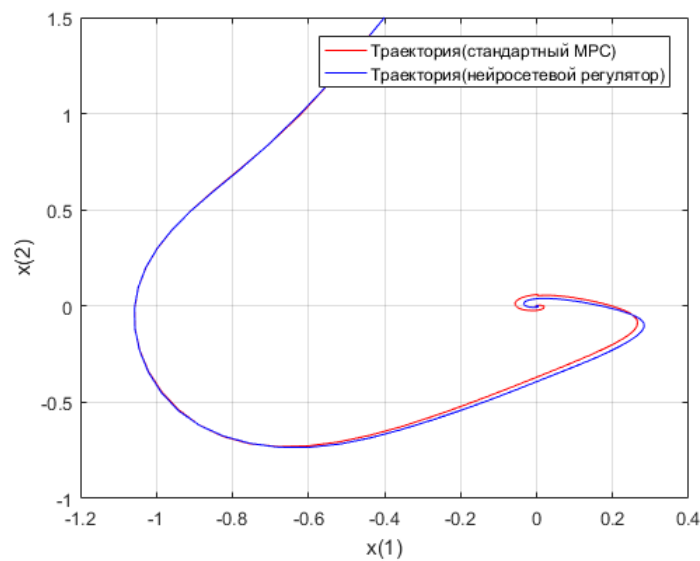


Рис. 3.5: Траектории, полученные при помощи стандартной процедуры МРС и нейросетевого регулятора

каждой точке равномерной сетки на множестве допустимых значений состояний. Затем на полученной выборке, состоящей из пар состояние-значение обратной связи, обучается нейронная сеть, как задача обучения с учителем.

Отметим основные недостатки данного подхода:

- Для хорошей аппроксимации необходима выборка, содержащая достаточно много точек, хотя некоторые области имеют одинаковые значения управления и для них его не надо вычислять заново.
- Обучение и валидация нейронной сети проходят достаточно долго.

Предполагаемые улучшения для этого подхода:

- Уменьшить количество генерируемых точек для построения аппроксимации, рассмотреть генерирование точек с помощью стохастического подхода

и стохастического с увеличением плотности точек у начала координат.

- Построить  $X_0$  — область притяжения, множество состояний, из которых система достигает терминального множества за  $N$  шагов.
- Для ускорения поиска управления, выделить множества точек, для которых применяется насыщенное управление.
- Рассмотреть возможность использования обучения без учителя и обучение с подкреплением для построения множества допустимых значений и закона управления.

### 3.4 Построение области притяжения и областей насыщения управления с помощью SVM

При построении базового регулятора в разд. 3.3 было замечено, что в выбранной области допустимых значений состояний можно выделить подмножество точек, из которых можно достичь начало координат — область притяжения, и подмножества, в которых замкнутая система неустойчива. Предлагается аппроксимировать область притяжения и только по точкам из нее строить приближение для обратной связи.

Было сгенерировано 2000 точек внутри области (3.4) допустимых значений состояний с использованием квази-случайного метода построения с помощью множеств Хальтона [21]. Для всех этих точек определяем возможность достижения из данного состояния точки начала координат. Помечаем состояние, из которого достижимо начало координат, с помощью метки 1, а неустойчивые состояния как  $-1$ . Для классификации применяется метод опорных векторов (SVM) [22] для аппроксимации области притяжения. В качестве функции ядра для метода SVM применяет радиальная функция Гаусса.

После применения SVM для аппроксимации области притяжения, предсказанные пометки для точек из области достижимых значений представлены на рисунке 3.6. Также на рисунке отмечены опорные вектора, которые определяют область устойчивых состояний.

Для определения областей с насыщенными управлениями также воспользуемся методом SVM.

Метод SVM хорошо работает для разграничения двух классов, а в рассматриваемой задаче у будет 3 класса: 1) насыщение на верхней границе,  $u = 1$ , 2) насыщение на нижней границе,  $u = -1$ , 3) промежуточные значения управления,  $|u| < 1$ . В связи с этим создадим две отдельные модели для аппроксимации

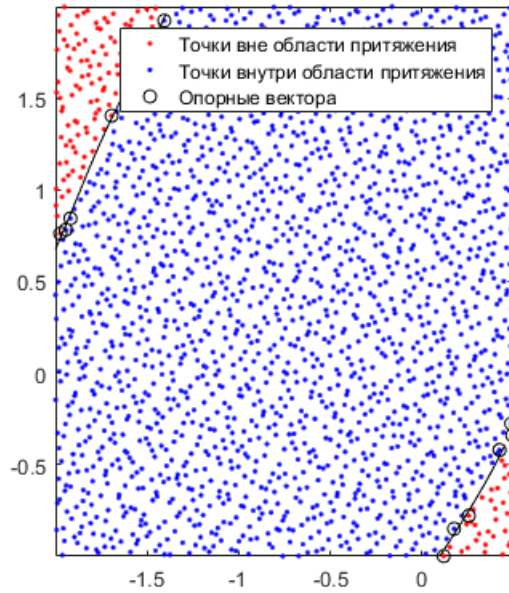


Рис. 3.6: Аппроксимация области притяжения для задачи (3.3)

областей: первую модель для определения области с управлением  $u = 1$  и вторую модель для определения области с управлением  $u = -1$ . Для каждой модели помечаем те состояния, которые соответствуют насыщенному управлению, как 1, остальные состояния как  $-1$ .

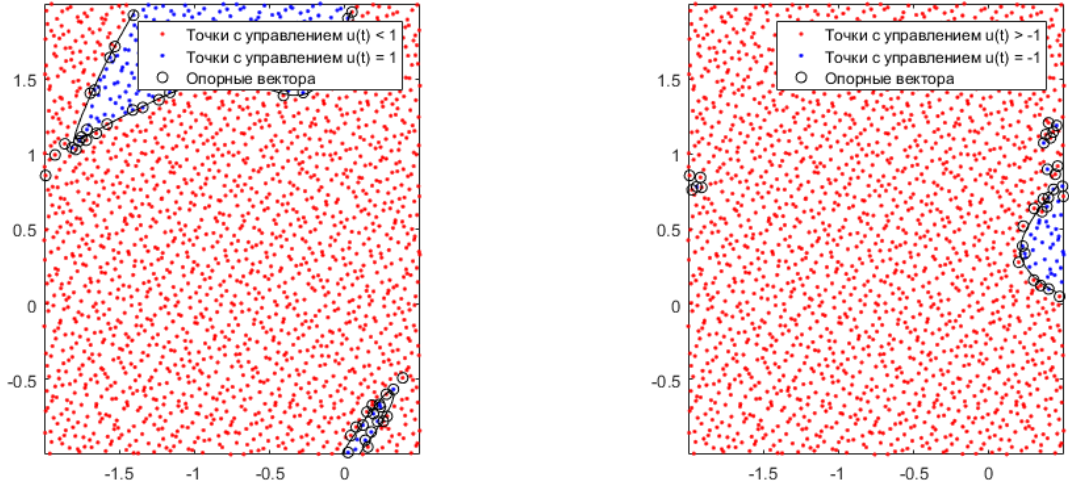
Здесь нужно принять во внимание, что в робастной версии используются сужения ограничений, поэтому насыщение управления будем определять с запасом в  $\epsilon$ , т.е. пометим все состояния из которых будет управление  $u(t) \geq 1 - \epsilon$  либо  $u(t) \leq -1 + \epsilon$ . Далее обучаем модели с помощью SVM с радиальной ядерной функцией Гаусса.

На рисунке 3.7 представлены области состояний, для которых необходимо применять насыщенные управления. Данные области помогут ускорить онлайн процедуру, так как теперь не нужно будет вычислять даже с помощью нейросетевого регулятора обратную связь для данных состояний, нужно будет только применять соответствующее насыщенное управление и не контролировать выход аппроксимированного управления за границы допустимых значений.

### 3.5 Сравнение методов сэмплирования точек для построения приближенной обратной связи

В базовой реализации используется равномерная сетка для построения множества точек для приближения обратной связи. Однако для равномерной сетки





(а) Аппроксимированная область применения  $u(t) = 1$  для задачи (3.3).

(б) Аппроксимированная область применения  $u(t) = -1$  для задачи (3.3).

Рис. 3.7: Аппроксимированные области с насыщенным управлением.

нужно вычислить много точек, но не все они одинаково важны для построения обратной связи. Поэтому предлагается проверить сэмплирование точки стохастически и стохастически с увеличением плотности точек в окрестности точки начала координат. Последний подход объясняется тем, что в окрестности точки начала координат значения управления малы, и хорошее приближение крайне важно для сохранения свойства асимптотической.

На равномерной сетке было взято 3000 точек, поэтому будем исследовать другие способы сэмплирования на этом же количестве точек.

На рисунке 3.8 представлены сэмплированные точки при стохастическом выборе 3000 точек для области (3.4) внутри области притяжения.

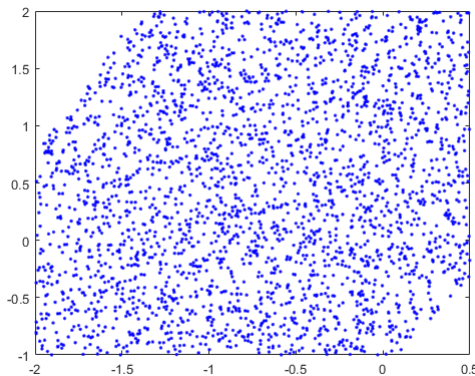


Рис. 3.8: Сэмплированные точки стохастически

На рисунке 3.9 представлены сэмплированные точки при стохастическом выборе 2600 точек для области (3.4) и 400 точек сэмплированных из окрестности

начала координат  $X_{zero} = \{(x_1, x_2) : \|x_1\| \leq 0.2, \|x_2\| \leq 0.2\}$  внутри области притяжения.

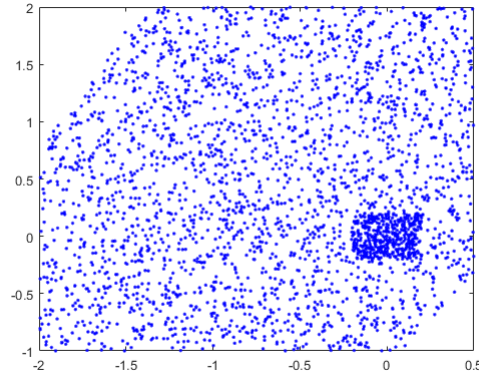


Рис. 3.9: Сэмплированные точки стохастически с увеличенной плотностью около нуля

Ошибка аппроксимации для этих двух способов для выбранного количества нейронов на скрытом слое  $n = 25$  представлена в таблице 3.2.

Таблица 3.2: Зависимость точности аппроксимации от типа сэмплирования

Тип сэмплирования	Ошибка аппроксимации
Равномерная сетка	0.0049
Стохастически сэмплированные точки	0.0062
Стохастически сэмплированные точки и увеличенная плотность в окрестности начала координат	0.0039

Как видно из таблицы 3.2 при способе сэмплирования стохастически с увеличенной плотностью в окрестности начала координат, увеличивается точность аппроксимации по сравнению с равномерной сеткой и простым стохастическим способом.

Поскольку для исследуемой задачи достаточно получить точность  $5 \cdot 10^{-3}$ , то можно уменьшить количество генерируемых точек. В таблице 3.3 представлены результаты аппроксимации для разного количества точек в обучающей выборке. Понятно, что для получения заданной точности, достаточно 2500 точек, а это значит можно сократить количество точек на 17%.

Таблица 3.3: Зависимость точности аппроксимации от количества точек при сэмпировании стохастически и увеличенной плотности точек вокруг начала координат

Количество точек	Ошибка аппроксимации
3000	0.0039
2800	0.0044
2600	0.0047
2500	0.0049

## 3.6 Сравнение результатов производительности

Сравним результаты производительности при реализации классического алгоритма MPC и при реализации нейросетевого регулятора с обучением оффлайн для решения задачи стабилизации нелинейной системы (3.3).

Траектории отдельно для  $x_1$  и  $x_2$  представлены на рисунках 3.10.

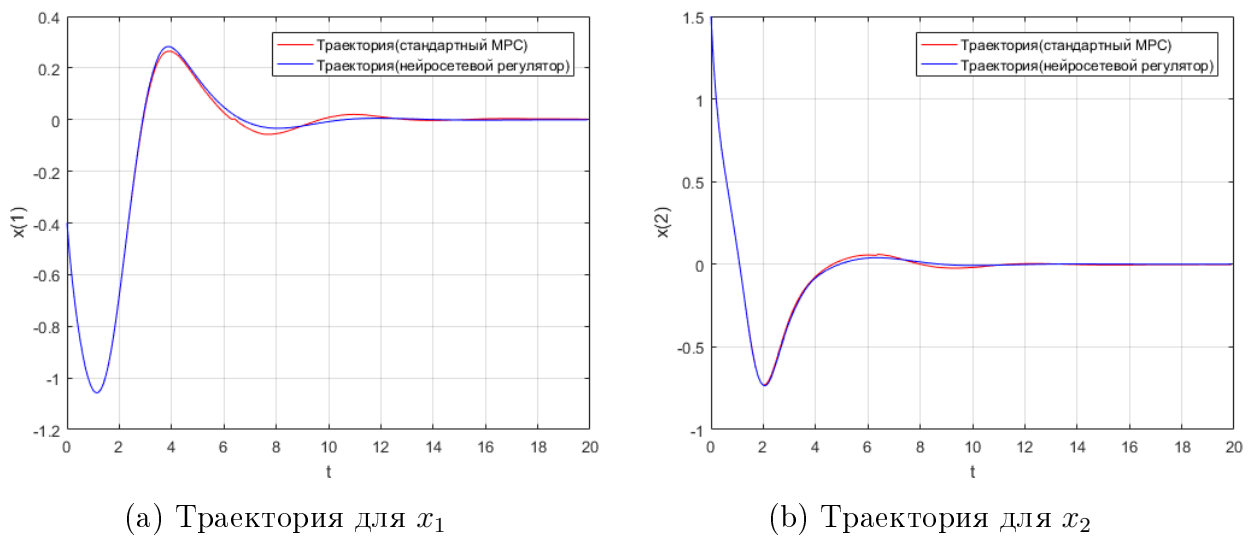


Рис. 3.10: Траектории для  $x_1$  и  $x_2$  для системы (3.5)

Управления, построенные используя стандартный MPC и при помощи нейросетевого регулятора, показаны на рисунке 3.11

В таблице 3.4 представлены значения времени в среднем для проведения одного шага алгоритма MPC онлайн. Для вычисления среднего значения времени была проведена 1000 вычислений первого шага процедуры MPC для различных начальных точек при помощи представленных выше методов.

Как видно из таблицы, последний метод, использующий дополнительно SVM для выделения областей насыщения управления, незначительно выиграл у ней-

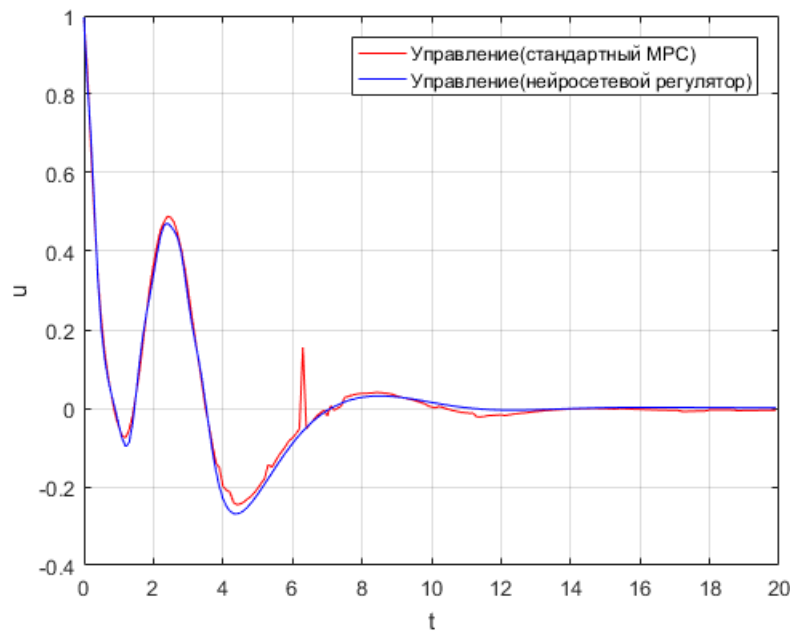


Рис. 3.11: Функция управления для задачи (3.5)

Таблица 3.4: Зависимость времени проведения одного шага МРС от использованного метода

Использованный метод	Время работы(мс)
Стандартная процедура МРС	1240
Нейросетевой регулятор	55
Нейросетевой регулятор с областями насыщенных управлений	52

росетевого регулятора, не различающего граничные и промежуточные значения управления, однако по сравнению со стандартной процедурой МРС время решения уменьшено в 23 раза.

В предыдущем разделе было рассмотрено проведение обучения на меньшем количестве точек за счет сэмплирования стохастически и увеличения плотности сэмплированных точек в окрестности начала координат. Данный метод уменьшил время для обучения на 17%.

Необходимо отметить, что применение нейросетевых регуляторов (также в комбинации с SVM) приводит к траекториям замкнутой системы, которые могут существенно отличаться от "эталонной", где под "эталонной" в данном случае понимается траектория, полученная при помощи стандартной процедуры МРС. Было подсчитано, что расхождения траекторий для 1000 запусков из различных начальных состояний относительно эталонных траекторий составили не более 16%.

Однако приведенные потери можно считать несущественными в связи с целью управления системой, которая состоит в стабилизации системы и гарантированном выполнении ограничений.

В то же время, существенный выигрыш по времени вычисления  $\bar{u}_{MPC}(x(t))$  в сравнении с  $u_{MPC}(x(t))$  позволяет применять нейросетевые регуляторы в тех случаях, когда дискретизация непрерывной системы осуществлена с маленьким шагом  $h$ . Например, в рассматриваемом примере стандартная процедура MPC вычисляет управление за 1240 мс, а нейросетевой регулятор за 55 мс.

Таким образом, нейросетевые регуляторы предпочтительнее в использовании в ряде производственных задач для систем с быстро меняющейся динамикой.

### 3.7 Использование нейросетевого регулятора для задачи стабилизации обратного маятника

Рассмотрим задачу стабилизации обратного маятника на подвижном основании (тележке) [23]. Линеаризация уравнений в окрестности верхнего устойчивого положения маятника имеет вид

$$\begin{bmatrix} \dot{x} \\ \ddot{x} \\ \dot{\phi} \\ \ddot{\phi} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & \frac{-(I+ml^2)b}{I(M+m)+Mml^2} & \frac{m^2gl^2}{I(M+m)+Mml^2} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{-mlb}{I(M+m)+Mml^2} & \frac{mgl(M+m)}{I(M+m)+Mml^2} & 0 \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \\ \phi \\ \dot{\phi} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{I+ml^2}{I(M+m)+Mml^2} \\ 0 \\ \frac{ml}{I(M+m)+Mml^2} \end{bmatrix} u, \quad (3.6)$$

где  $x$  — положение тележки на платформе, а  $\phi$  — угол наклона тележки, и параметры выбраны следующим образом:

$$M = 0.5, \quad m = 0.2, \quad b = 0.1, \quad I = 0.006, \quad g = 9.8, \quad l = 0.3,$$

$$q = (M + m) \cdot (I + m \cdot l^2) - (m \cdot l)^2.$$

Ограничения накладываются только на угол  $|\phi| \leq \frac{\pi}{3}$  и на управление  $|u| \leq 200$ .

Определены терминальные элементы для формулировки задачи в рамках квазибесконечного MPC. Линейная обратная связь в терминальном множестве

имеет вид:

$$u_{loc}(t) = Kx(t) = [10.0000, 15.8365, -79.1980, -18.5048]x(t).$$

Терминальная функция — квадратичная функция вида  $x(t)^T Px(t)$ , где

$$P = \begin{bmatrix} 1.5979 & 0.7505 & -1.8801 & -0.3226 \\ 0.7505 & 0.9811 & -2.6173 & -0.4276 \\ -1.8801 & -2.6173 & 9.2988 & 1.2221 \\ -0.3226 & -0.4276 & 1.2221 & 0.2119 \end{bmatrix}.$$

Точность аппроксимации, как и в предыдущей задаче, выбрана равной  $\eta = 5 \cdot 10^{-3}$ .

Для построения обучающей выборки и дальнейшего анализа будем строить сетку значений состояний в области

$$-\begin{bmatrix} 2 \\ 2 \\ \frac{\pi}{3} \\ \frac{\pi}{3} \end{bmatrix} \leq \begin{bmatrix} x \\ \dot{x} \\ \phi \\ \dot{\phi} \end{bmatrix} \leq \begin{bmatrix} 2 \\ 2 \\ \frac{\pi}{3} \\ \frac{\pi}{3} \end{bmatrix} \quad (3.7)$$

Используя равномерную сетку значений с шагом 0.1 получим 705600 точек.

Используя стохастическое сэмплирование, создадим выборку, состоящую из 585000 точек, при этом 50000 точек будем сэмплировать из множества

$$-\begin{bmatrix} 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \end{bmatrix} \leq \begin{bmatrix} x \\ \dot{x} \\ \phi \\ \dot{\phi} \end{bmatrix} \leq \begin{bmatrix} 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \end{bmatrix}$$

т.е. окрестности начала координат, в которой необходимо увеличить плотность точек для более точной аппроксимации. Оставшиеся точки будем сэмплировать из (3.7).

На этих двух датасетах обучается нейронная сеть для аппроксимации закона управления с точностью  $\eta = 5 \cdot 10^{-3}$ . Для достижения указанной точности понадобилось 40 нейронов на скрытом слое как для равномерной сетке, так и для стохастического сэмплирования.

В таблице (3.5) показаны значения ошибки в зависимости от количества нейронов на скрытом слое для двух сэмплированных датасетов.

Область притяжения имеет вид, представленный на рисунке 3.12. Насыщен-

Таблица 3.5: Таблица зависимости точности аппроксимации от количества нейронов на скрытом слое

Количество нейронов	Ошибка аппроксимации (равномерная сетка)	Ошибка аппроксимации (стохастическое сэмплирование)
5	1.5362	1.2145
10	0.2873	0.2712
15	0.1267	0.0934
20	0.0544	0.0511
25	0.0219	0.013
30	0.0081	0.0078
40	0.0049	0.0047

ных управлений в данном датасете нет, поэтому области насыщенных управлений для этой задачи не строятся.

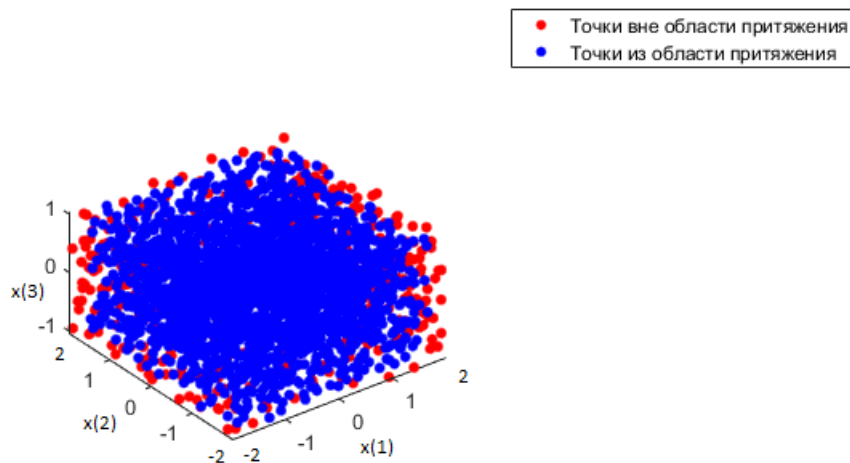


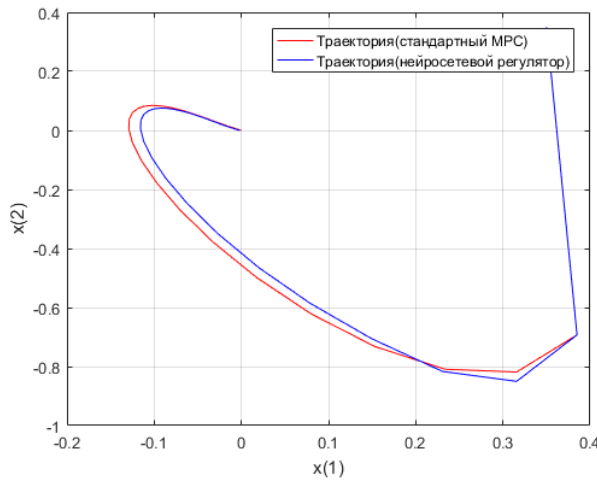
Рис. 3.12: Область притяжения для задачи (3.6)

Результаты представлены на рисунке 3.13. На рисунке приведены траектории стандартного МРС-регулятора и нейросетевого регулятора, начиная из точки  $x(t) = [0.35, 0.35, 0.35, 0.35]$ .

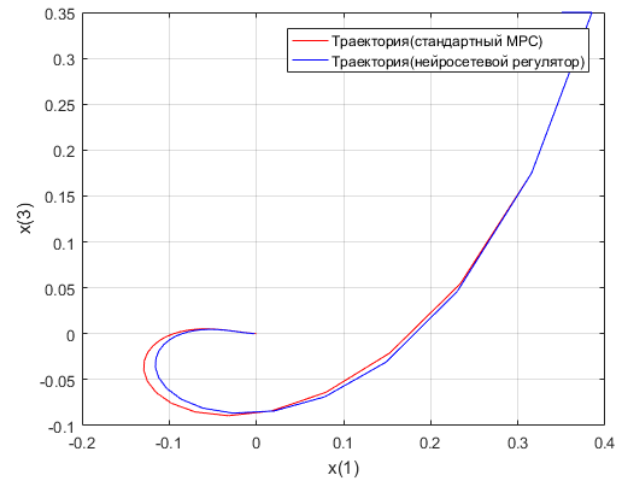
Траектории отдельно для  $x_1$  и  $x_3$  представлены на рисунках 3.14.

Управления, построенные используя стандартный МРС и при помощи нейросетевого регулятора, показаны на рисунке 3.15.

Для вычисления среднего значения времени было проведено 1000 процедур вычисления первого шага процедуры МРС из различных начальных точек при помощи вышепредставленных методов. При использовании стандартной процедуры МРС среднее значение времени вычисления одного равняется 1633 мс, в то

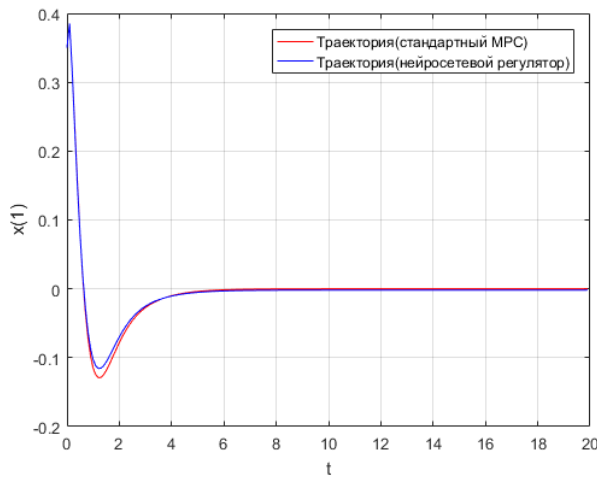


(a) Траектория для обратного маятника на срезе для координат  $x_1$  и  $x_2$

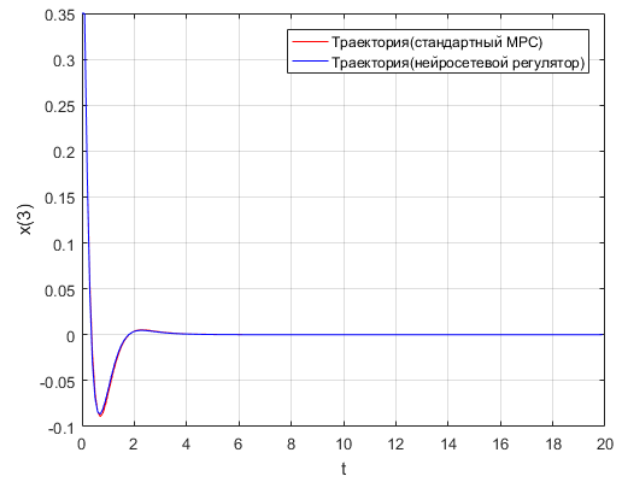


(b) Траектория для обратного маятника на срезе для координат  $x_1$  и  $x_3$

Рис. 3.13: Срезы траектории для стабилизации обратного маятника (3.6)



(a) Траектория  $x_1$



(b) Траектория  $x_3$

Рис. 3.14: Траектории для стабилизации обратного маятника (3.6)

время как применение нейросетевого регулятора сокращает время выполнения одного шага до 78 мс.

## 3.8 Выводы

В данной работе было исследовано применение нейронных сетей для аппроксимации закона управления для нелинейных систем, которые повысили производительность онлайн вычислений в 20 раз. Также были исследованы разные техники сэмплирования, и было показано, что можно сократить количество точек



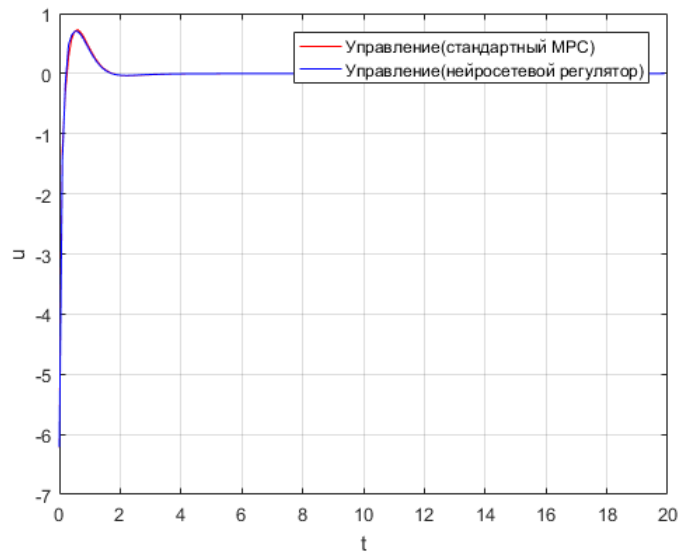


Рис. 3.15: Функция управления для задачи (3.6)

тренировочного набора на 17% при использовании стохастического сэмплирования и повышения плотности точек у начала координат. Были построены области притяжения с помощью SVM для более качественного сэмплирования.

## ЗАКЛЮЧЕНИЕ

МРС представляет собой семейство регуляторов, которое позволяет явно использовать модель объекта управления для получения управляющего сигнала. Несмотря на эффективность и широкое распространение на практике, данный метод не подходит для быстрых процессов, для которых решение нелинейной задачи оптимального управления не может быть получено регулятором за короткий период квантования. Один из подходов к решению указанной проблемы подразумевает перенос некоторых вычислений оффлайн. Другой недостаток метода состоит в том, что МРС предполагает наличие динамической модели, поэтому в условиях неизвестной динамики МРС не сможет находить эффективное управление и необходимо использовать другие методы.

В настоящей работе для построения аппроксимации управления в МРС регуляторе использовались нейронные сети. Были проведены исследования по решению указанной проблемы путем переноса некоторых вычислений оффлайн. Предоставлены результаты базовой реализации алгоритма. В частности, для реализации функции МРС-регулятора были применены искусственные нейронные сети. Были проведены эксперименты для построения области притяжения с помощью SVM. Проанализированы различные техники сэмплирования точек для аппроксимации обратной связи. Для исследования применения нейросетевого регулятора использовались задачи с двумерными и четырехмерными множествами состояний. Техника стохастического сэмплирования с увеличением плотности вокруг начала координат уменьшила количество точек, необходимых для построения тренировочного набора на 17%, а значит и сократила время на обучения. Среднее время выполнения одного шага МРС для стандартной процедуры составляло 1240 мс, а при использовании нейросетевого регулятора составляет 52 мс. Производительность системы с нейросетевым регулятором уменьшила время выполнения одного шага МРС в 20 раз.

Дальнейшие направления исследований включают применение различных архитектур нейронных сетей для обеспечения необходимых параметров МРС-регулятора, а также построение регуляторов при отсутствии для рассматриваемого объекта математической модели в пространстве состояний.

## СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

- 1 Grune, L. Nonlinear model predictive control / L. Grune, J. Pannek // Springer London. – 2011.
- 2 Rawlings, J.B. Model Predictive Control: Theory and Design / J.B. Rawlings, D.Q. Mayne // Madison: Nob Hill Publishing. – 2009.
- 3 Badgwell, T.A. Model-Predictive Control in Practice / T.A. Badgwell, S.J. Qin // Encyclopedia of Systems and Control. Springer, London. – 2015.
- 4 Chen, H. A Quasi-Infinite Horizon Nonlinear Model Predictive Control Scheme with Guaranteed Stability / H. Chen, F. Allgower // Automatica. – 1998. – Vol. 34, No. 10. – P. 1205-1217.
- 5 Fontes, F.A.C.C. A General Framework to Design Stabilizing Nonlinear Model Predictive Controllers / F.A.C.C. Fontes // Systems & Control letters. – 2000. – Vol. 42, No. 2. – P. 127-143.
- 6 Bemporad, A. The explicit linear quadratic regulator for constrained systems / A. Bemporad et al. // Automatica. – 2002. – Vol. 38, No. 1. – P. 3-20.
- 7 Domahidi, A. Learning a feasible and stabilizing explicit model predictive control law by robust optimization / A. Domahidi et al. // Proceedings of the IEEE Conference on Decision & Control. – 2011. – No. EPFL-CONF-169723.
- 8 Chakrabarty, A. Support Vector Machine Informed Explicit Nonlinear Model Predictive Control Using Low-Discrepancy Sequences / A. Chakrabarty et al. // IEEE Transactions on Automatic Control. – 2017. – Vol. 62, No. 1. – P. 135-148.
- 9 Johansen, T. A. Approximate explicit receding horizon control of constrained nonlinear systems // Automatica. – 2004. – Vol. 40, No. 2. – P. 293-300.
- 10 Parisini, T. A receding-horizon regulator for nonlinear systems and a neural approximation / T. Parisini, R. Zoppoli // Automatica. – 1995. – Vol. 31, No. 10. – P. 1443-1451.
- 11 Parisini, T. Nonlinear stabilization by receding-horizon neural regulators / T. Parisini, M. Sanquineti, R. Zoppoli // International Journal of Control. – 1998. – Vol. 70, No. 3. – P. 341-362.
- 12 Akesson, B. M. A neural network model predictive controller / B.M. Akesson, H.T. Toivonen // Journal of Process Control. – 2006. – Vol. 16, No. 9. – P. 937-946.
- 13 Pin, G. Approximate model predictive control laws for constrained nonlinear

discrete-time systems: analysis and offline design / G. Pin et al. // International Journal of Control. – 2013. – Vol.86, No.5. – P. 804-820.

14 Aswani, A. Provably safe and robust learning-based model predictive control / A. Aswani et al. // Automatica. – 2013. – Vol. 49, No. 5. – P. 1216-1226.

15 Rosolia, U. Learning model predictive control for iterative tasks. a data-driven control framework / U.Rosolia, F. Borrelli // IEEE Transactions on Automatic Control. – 2018. – Vol. 63, No. 7.

16 Goodfellow, I. Deep learning / I.Goodfellow et al. // Cambridge : MIT press. – 2016. – Vol. 1.

17 Csaji, B. C. Approximation with artificial neural networks // Faculty of Sciences, Eötvös Loránd University, Hungary. – 2001. – Vol. 24. – P. 48.

18 Andersson, J.A.E. CasADi: a software framework for nonlinear optimization and optimal control / J.A.E. Andersson et al. // Mathematical Programming Computation. – 2018. – P. 1-36.

19 Diehl, M. Efficient numerical methods for nonlinear MPC and moving horizon estimation /M. Diehl, H. J. Ferreau, N. Haverbeke // Nonlinear model predictive control. Springer, Berlin, Heidelberg. – 2009. – P. 391-417.

20 Hertneck, M. Learning an approximate model predictive controller with guarantees / M.Hertneck et al. // IEEE Control Systems Letters. – 2018. – Vol. 2. No. 3. – P. 543-548.

21 Kocis, L. Computational investigations of low-discrepancy sequences / L. Kocis, W.J. Whiten // ACM Transactions on Mathematical Software (TOMS). – 1997. – No. 23. Vol. 2. – P. 266-294.

22 Cristianini, N. An introduction to support vector machines and other kernel-based learning methods // Cambridge university press. – 2000.

23 Liberzon, D. Switching in systems and control // Springer Science and Business Media. – 2003.

24 Павловец, М.Е. Применение методов машинного обучения в системах управления по прогнозирующей модели / М.Е. Павловец, Н.М. Дмитрук // XIX Международная научная конференция по дифференциальным уравнениям "Еругинские чтения-2019": материалы международной научной конференции. Могилев, 14-17 мая 2019 г. – С. 127-129.

## СПИСОК ПУБЛИКАЦИЙ

Павловец, М.Е. Применение методов машинного обучения в системах управления по прогнозирующей модели / М.Е. Павловец, Н.М. Дмитрук // XIX Международная научная конференция по дифференциальным уравнениям "Еругинские чтения-2019": материалы международной научной конференции. Могилев, 14-17 мая 2019 г. – С. 127-129.