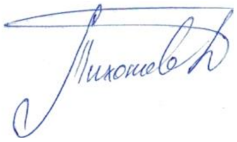
На правах рукописи



Тихонов Денис Максимович

Снижение размерности в задачах мультимоделирования динамических систем

1.2.1 — Искусственный интеллект и машинное обучение

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико-математических наук

Москва — 2026

Работа выполнена в федеральном государственном автономном образовательном учреждении высшего образования «Московский физико-технический институт (национальный исследовательский университет)».

Научный руководитель: **Стрижов Вадим Викторович**
доктор физико-математических наук, старший научный сотрудник ООО «Форексис».

Официальные оппоненты: **Чуличков Алексей Иванович**
доктор физико-математических наук, профессор, федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный университет имени М. В. Ломоносова», физический факультет, заведующий кафедрой математического моделирования и информатики.

Анциперов Вячеслав Евгеньевич
кандидат физико-математических наук, доцент, федеральное государственное бюджетное учреждение науки Институт радиотехники и электроники им. В. А. Котельникова Российской академии наук, ведущий научный сотрудник.

Ведущая организация: Автономная некоммерческая образовательная организация высшего образования «Сколковский институт науки и технологий».

Защита состоится «___» _____ 2026 года в _____ на заседании диссертационного совета 24.1.224.03 при Федеральном исследовательском центре «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН) по адресу: 119333, г. Москва, ул. Вавилова, д. 40.

С диссертацией можно ознакомиться в библиотеке Федерального исследовательского центра «Информатика и управление» Российской академии наук и на сайте https://www.frccsc.ru .

Автореферат разослан «___» _____ 2026 года.

Ученый секретарь
диссертационного совета 24.1.224.03
к.т.н. И. А. Рейер

Общая характеристика работы

Актуальность темы.

В работе исследуется проблема снижения размерности в задачах моделирования динамических систем по наблюдаемым временным рядам. Методы снижения размерности строят подпространство состояний исследуемой системы. В полученном подпространстве модели прогнозирования являются более устойчивыми в сравнении с моделями в исходных пространствах и имеют меньшее число параметров.

Пространство состояния системы восстанавливается с помощью векторизации временных рядов. Под векторизацией подразумевается *метод временных задержек*, который ставит в соответствие значению временного ряда его предысторию конечной длины. Векторное представление временных рядов и их синхронное наблюдение приводит к проблеме мультикорреляции. При высокой мультикорреляции модели прогнозирования оказываются неустойчивыми. Для построения простой, устойчивой и точной модели применяются методы снижения размерности (Pearson: 1901, Hotelling: 1933, Jolliffe: 2016, Benner: 2015).

Методы снижения размерности, применяемые при анализе динамических систем, строят низкоразмерные пространства состояний, сохраняющие основные топологические и динамические свойства исходной системы. Проекция системы на такое подпространство упрощает модели прогнозирования. Метод главных компонент и сингулярное разложение сохраняют дисперсию линейных признаков (Pearson: 1901, Eckart: 1936, Hotelling: 1933, Jolliffe: 2016). Для моделей временных рядов аналогичным образом применяется собственное ортогональное разложение (Sirovich: 1987, Berkooz: 1993, Holmes: 2012). В задачах управления размерность пространств снижается выбором оптимального подмножества координат и проекцией на подпространство (Moore: 1981, Juang: 1985, Antoulas: 2020, Benner: 2015). Переход к подпространству упрощает интерпретацию и анализ моделей (Bellman: 1961, Peherstorfer: 2016).

Восстановление пространства состояний по наблюдаемым временным рядам опирается на теоремы о вложении. Теорема Уитни определяет достаточную размерность пространства для гладкого многообразия (Whitney: 1936). Теорема Такенса показывает, что для *почти всех функций* наблюдения метод задержек одномерного временного ряда строит пространство диффеоморфное истинному пространству состояний (Packard: 1980, Takens: 1981). Для набора временных рядов результат обобщен в многомерном методе задержек (Sauer: 1991, Deyle: 2011). Параметры метода задержек, временная задержка и длина окна, выбираются методом минимума взаимной информации и методом ложных ближайших соседей (Fraser: 1986, Kennel: 1992, Cao: 1997,

Abarbanel: 1996). Условия сохранения геометрических свойств и расстояний получены позднее в работе (Eftekhari: 2018). Эти результаты гарантируют, что восстановленное пространства состояний сохраняет топологические свойства истинной динамической системы.

Линейная проекция координат восстановленного пространства состояний не учитывает нелинейные связи компонент. Метод главных компонент, глобальное вложение по главным компонентам и конкатенация нескольких рядов строят линейные комбинации (Broomhead: 1986, Jolliffe: 2016, Vlachos: 2010). Синхронные временные ряды для восстановления пространства состояний при этом избыточны: каждый ряд повторяет схожие зависимости (Deyle: 2011, Sugihara: 1990). Методы, использующие линейное объединение векторов задержек без учета внутренних зависимостей временных рядов, требуют большую размерность пространства для построения устойчивых модели прогнозирования. Выбор подмножества компонент вектора задержек есть частный случай снижения размерности, снижающий число параметров, но не учитывающий нелинейные зависимости (Tibshirani: 1996).

Динамическую систему в получаемых пространствах описывают линейными и нелинейными моделями. Динамическая модовая декомпозиция и теория оператора Купмана (Koopman: 1931, Mezic: 2005, 2013) заменяют нелинейный *поток* линейным оператором в пространстве признаков большой размерности (Schmid: 2010, Tu: 2014, Schmid: 2022). Нелинейная декомпозиция и нейросетевые автоэнкодеры подбирают это пространство по наблюдаемым временным рядам (Williams: 2015, Kaiser: 2018, Lusch: 2018, Wehmeyer: 2018). Нейросетевые модели строят точный прогноз и не восстанавливают уравнения в исходных переменных (Chen: 2018, Raissi: 2019, Greydanus: 2019, Rubanova: 2019).

Существенной задачей остаётся построение интерпретируемых моделей по наблюдаемым данным при слабых априорных предположениях о структуре системы. Разреженное восстановление нелинейной динамической системы восстанавливает явный интерпретируемый вид системы дифференциальных уравнений, описывающих динамическую систему (Brunton: 2016, Rudy: 2017, Mangan: 2017, Champion: 2019). Нелинейные зависимости учитываются в модели заранее заданным набором нелинейных преобразований. Сложность модели растет экспоненциально при увеличении длины окна метода задержек и увеличению числа исследуемых преобразований. Тензорные разложения (Carroll: 1970, Tucker: 1966) записывают те же произведения меньшим числом параметров (Kruskal: 1977, Kolda: 2009, Oseledets: 2011).

В диссертации решается задача снижения размерности пространства состояний по набору рядов. Предлагается полилинейное отображение векторов задержек, дополненных константой. Линейные задержки сохраняются и

дополняются произведения компонент набора временных рядов. Число параметров снижается тензорным разложением. Для восстановления пространств состояний используются линейные и нелинейные методы. Лемма о полилинейном расширении показывает, что предлагаемое преобразование векторов задержек является вложением ровно тогда, когда вложением является многомерный метод задержек. Теорема о многомерном полилинейном вложении устанавливает достаточные условия на ранг канонического разложения и на размерность восстановленного пространства. Доказаны включение линейной векторной авторегрессии в класс полилинейных моделей и липшицевость проекции, параметризованной статическими параметрами системы.

Цели работы.

1. Исследовать способы решения задачи восстановления пространства состояний динамических систем по набору временных рядов с использованием инструментов тензорной алгебры.
2. Теоретически обосновать применимость полилинейных отображений, снижающих размерность пространств состояний, для многомерных временных рядов.
3. Предложить методы снижения размерности пространств состояний, обобщающие метод анализа сингулярного спектра и учитывающие внутренние зависимости компонент временных рядов.
4. Предложить способ построения интерпретируемой модели, восстанавливающей неизвестные зависимости динамических систем с малым числом параметров.

Основные положения, выносимые на защиту.

1. Исследована проблема восстановления пространств состояний динамической системы по многомерным наблюдаемым временным рядам. Предложен полилинейный метод снижения размерности пространств состояний.
2. Исследованы свойства предлагаемых методов. Доказаны теоремы о полилинейном вложении метода задержек для многомерных временных рядов.
3. Предложены полилинейные методы снижения размерности набора пространств состояний. Предлагаемые методы сохраняют свойства исходных систем и учитывают нелинейные связи наблюдаемых временных рядов.
4. Предложен метод построения интерпретируемой нелинейной модели, восстанавливающий приближённый вид дифференциальных уравнений, описывающих динамическую систему.
5. Разработан алгоритм построения полилинейной модели с малым числом

параметров, оптимизированной с помощью тензорных разложений. Проведены вычислительные эксперименты, подтверждающие эффективность предложенных методов.

Методы исследования. Используются линейные и нелинейные методы восстановления пространств состояний динамической системы. Для восстановления пространства состояний используются методы, основанные на теореме Такенса. Для отбора интерпретируемых нелинейных признаков используются методы тензорной алгебры. Для снижения размерности пространств состояний используются метод главных компонент и нейросетевые методы. Для оптимизации моделей прогнозирования в пространстве состояний используются составные функции потерь.

Научная новизна. Разработан полилинейный метод восстановления пространства состояний динамической системы. Предложен метод объединения подпространств состояний нескольких временных рядов на основе полилинейного отображения, учитывающий нелинейные зависимости между рядами. Разработаны методы снижения числа параметров модели на основе тензорных разложений, адаптированные к применению в нейросетевых моделях. Разработано полилинейное расширение метода разреженного восстановления дифференциальных уравнений динамической системы, обеспечивающее линейный рост числа параметров и интерпретируемость модели.

Теоретическая значимость. Доказана теорема о полилинейном вложении метода задержек для многомерного временного ряда в случае канонического разложения. Доказана теорема о включении линейной векторной авторегрессии в класс полилинейных моделей. Доказана теорема о липшицевой непрерывности параметризованного полилинейного отображения.

Практическая значимость. Предложенные в работе методы предназначены для восстановления пространств состояний динамических систем по временным рядам в условиях ограниченных данных и зашумлённых измерений. Разработанные методы используются в задачах прогнозирования состояний динамических систем и строят интерпретируемые модели с малым числом параметров.

Степень достоверности и апробация работы. Достоверность результатов подтверждена математическими доказательствами, экспериментальной проверкой полученных методов на задачах восстановления динамических систем по временным рядам, а также публикациями результатов исследования в рецензируемых научных изданиях, в том числе рекомендованных ВАК. Результаты работы докладывались и обсуждались на следующих научных конференциях:

1. Объединение пространств состояний динамической системы через полилинейное отображение, Международная конференция «Иванниковские чтения», Москва, 2026.
2. Разреженное восстановление функции эволюции связанных динамических систем, 68-я Всероссийская научная конференция МФТИ, Москва, 2026.
3. Dynamic system reconstruction from multivariate time series via multilinear map and time delay embedding, International Conference on Computational Optimization (ICOMP), Иннополис, 2024.
4. Восстановление зашумленных значений многомерных временных рядов с помощью тензорного метода задержек и тензорного разложения, 66-я Всероссийская научная конференция МФТИ, Москва, 2024.
5. Классификация квазипериодических временных сигналов с помощью сферических гармоник, Международная научная конференция студентов, аспирантов и молодых учёных «Ломоносов-2022», Москва, 2022.

Публикации по теме диссертации. Основные результаты по теме диссертации изложены в 3 статьях в изданиях, рекомендованных ВАК, и в 3 тезисах докладов.

Личный вклад. Все приведенные результаты, кроме отдельно оговоренных случаев, получены диссертантом лично при научном руководстве д.ф.-м.н. В. В. Стрижова.

Структура и объем работы. Диссертация состоит из оглавления, введения, пяти разделов, заключения, списка иллюстраций, списка таблиц, перечня основных обозначений и списка литературы из 87 наименований. Основной текст занимает 101 страницу.

Основное содержание работы

Во **введении** обоснована актуальность диссертационной работы, сформулированы цели и методы исследования, обоснованы научная новизна, теоретическая и практическая значимость полученных результатов.

В **главе 1** ставится задача восстановления пространства состояний динамической системы по наблюдаемым временным рядам.

Определение 1. *Классическая динамическая система*[1] $(\mathcal{M}, \mu, \phi_t)$ — набор из гладкого многообразия $\mathcal{M}$, меры μ на $\mathcal{M}$ с непрерывной положительной плотностью и однопараметрической группы ϕ_t диффеоморфизмов многообразия $\mathcal{M}$, сохраняющих меру $\mu(A) = \mu(\phi_t A)$ при всех t и всех измеримых $A \subset \mathcal{M}$. Параметр t принимает значения в $\mathbb{R}$ (непрерывное время) или в $\mathbb{Z}$ (дискретное время).

В локальных координатах при $t \in \mathbb{R}$ действие ϕ_t задается гладкими функциями

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{v}(t) = \mathbf{f}(\mathbf{v}(t), t; \mathbf{c}), \quad \mathbf{v} \in U \subseteq \mathbb{R}^{d_s}, \quad \mathbf{f} : U \to \mathbb{R}^{d_s}, \tag{1}$$

где $\mathbf{v}$ — состояние системы, $\mathbf{f}$ — векторное поле, зависящее от состояния, времени t и вектора параметров $\mathbf{c}$.

Определение 2. *Фазовое пространство или пространство состояний* — это пространство U всех возможных состояний системы $\mathbf{v}$. Каждая точка фазового пространства полностью описывает состояние динамической системы в момент времени t.

Определение 3. *Фазовая траектория* — кривая $\{\mathbf{v}(t, \mathbf{v}_0)\}$ в фазовом пространстве, то есть последовательность состояний $\{\mathbf{v}(t) : t \in \mathcal{T}\}$ на интервале времени $\mathcal{T}$ при начальном условии $\mathbf{v}(0) = \mathbf{v}_0$.

Частным случаем (1) являются *автономные системы*, не зависящие в явном виде от времени:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{v} = \mathbf{f}(\mathbf{v}). \tag{2}$$

Решением (2) является гладкое отображением $\phi : T \to U$ оси времени в фазовое пространство:

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(t) = \mathbf{f}(\phi(t)).$$

Системы (1) и (2) сводятся к дискретному виду с шагом Δt, так что $\mathbf{v}_k = \mathbf{v}(k\Delta t)$:

$$\mathbf{v}_{k+1} = \mathbf{F}_{\Delta t}(\mathbf{v}_k) = \mathbf{v}_k + \int_{t_k}^{t_k+\Delta t} \mathbf{f}(\mathbf{v}(s))\, ds.$$

[1] Определение дано по: Арнольд В. И., Авец А. Эргодические проблемы классической механики. Ижевск: РХД, 1999. Гл. 1, § 1.

Далее $\mathbf{F}_{\Delta t}$ обозначается как $\mathbf{F}$ поскольку шаг дискретизации Δt фиксирован экспериментом.

На практике вектор истинного состояния $\mathbf{v}_t$ редко доступен. Исходное фазовое пространство связано с временным рядом $\{\mathbf{v}_{\mathrm{obs}\,t}\}_{t=1}^{N}$ гладкой функцией наблюдения $\mathbf{h} : \mathcal{M} \to \mathbb{R}^{d_{\mathrm{obs}}}$:

$$\mathbf{v}_{\mathrm{obs}\,t} = [x_{1,t}, \ldots, x_{d_{\mathrm{obs}},t}]^{\mathsf{T}} = [h_1(\mathbf{v}_t), \ldots, h_{d_{\mathrm{obs}}}(\mathbf{v}_t)] = \mathbf{h}(\mathbf{v}_t),$$

где $x_{i,t}$ — скалярные значения временного ряда в момент времени t, $i = 1, \ldots, d_{\mathrm{obs}}$. Число наблюдаемых временных рядов $d_{\mathrm{obs}} = 1$ для одномерного случая, функция наблюдения обозначается $h : \mathcal{M} \to \mathbb{R}$; $d_{\mathrm{obs}} \geq 2$ для многомерного случая.

Определение 4. *Функция лаговых координат* $\Phi_{h,F,\tau} : \mathcal{M} \to \mathbb{R}^n$ в одномерном случае сопоставляет значению наблюдаемого временного ряда его предыдущие во времени значения и определяется как

$$\begin{aligned}\Phi_{h,F,\tau}(\mathbf{v}_k) &= [h(\mathbf{F}^{\circ(-(n-1)\tau)}(\mathbf{v}_k)), \ldots, h(\mathbf{F}^{-\tau}(\mathbf{v}_k)), h(\mathbf{v}_k)]^{\mathsf{T}} \\ &= [x_{k-(n-1)\tau}, \ldots, x_{k-\tau}, x_k]^{\mathsf{T}},\end{aligned}$$

с параметрами τ — задержка (лаг) и n — размерность вложения; операция $\mathbf{F}^{\circ i}$ обозначает i-ю итерацию отображения, например $\mathbf{F}^{\circ 3}(\mathbf{v}_k) = \mathbf{F}(\mathbf{F}(\mathbf{F}(\mathbf{v}_k))) = \mathbf{v}_{k+3}$.

Метод временных задержек отображает скалярный временной ряд $\{x_t\}_{t=1}^{N}$ в пространство высокой размерности функцией лаговых координат $\Phi_{h,F,\tau}$ и формирует вектор временных задержек

$$\mathbf{x}_k = [x_{k-(n-1)\tau}, \ldots, x_{k-\tau}, x_k]^{\mathsf{T}}.$$

Матрица, образованная векторами задержек, называется *траекторной матрицей* временного ряда $\mathbf{x}$:

$$\mathbf{X} = \begin{bmatrix} x_1 & \ldots & x_n \\ x_2 & \ldots & x_{n+1} \\ \vdots & \ddots & \vdots \\ x_m & \ldots & x_N \end{bmatrix}^{\mathsf{T}} = \left[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\right], \quad m = N - n + 1, \tag{3}$$

где n — ширина окна, N — длина временного ряда, а τ равен 1.

Задача. По рядам $\{x_{i,t}\}$ построить многообразие $\widehat{\mathcal{M}} \subset \mathbb{R}^p$ и отображение $\widehat{\mathbf{f}} : \widehat{\mathcal{M}} \to \widehat{\mathcal{M}}$ такие, что существует диффеоморфизм $\mathcal{E} : \mathcal{M} \to \widehat{\mathcal{M}}$, для которого

$$\mathcal{E} \circ \mathbf{f} = \widehat{\mathbf{f}} \circ \mathcal{E}. \tag{4}$$

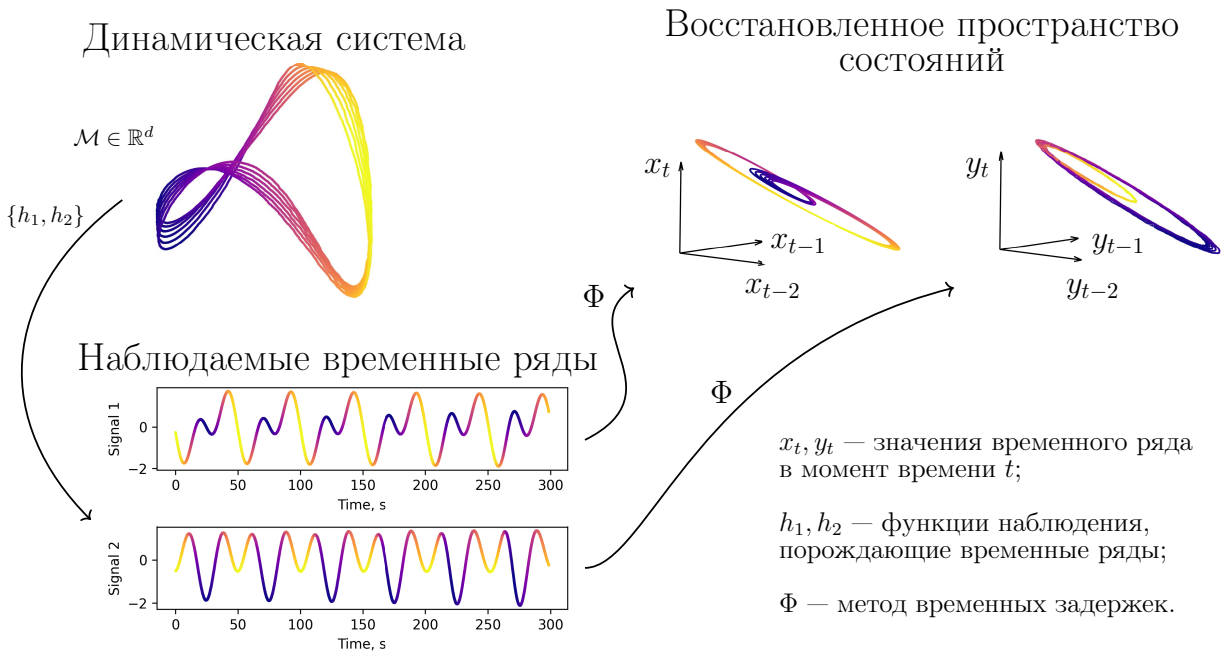


Рис. 1: Восстановление пространства состояний. Цветом показана фаза периодического изменения состояния системы.

Равенство (4) означает, что шаг восстановленной системы совпадает с шагом исходной системы в других координатах. Размерность p мала по сравнению с длиной окна. Схема задачт показана на рис. 1.

По теореме Такенса[2] для почти всех функций наблюдения метод задержек восстанавливает пространство состояний по одному временному ряду, диффеоморфное истинному пространству состояний. Восстановленное пространство задаётся вектором задержек

$$\Phi_{h,\mathbf{F},\tau}(\mathbf{v}_k) = \left[h(\mathbf{F}^{\circ(-(n-1)\tau)}(\mathbf{v}_k)), \ldots, h(\mathbf{v}_k)\right]^{\mathsf{T}}$$

и при длине окна $n \geq 2d_s + 1$ взаимно однозначно вкладывает исходное многообразие. То же относится и к дискретному времени.

Определение 5. *Почти все функции* — это такие функции $h_\alpha = h_0 + \sum_{i=1}^{n_\alpha} \alpha_i h_i$, где $\{h_i\}_{i=1}^{n_\alpha}$ — базис многочленов степени меньше или равной $2n$ и h_0 — произвольная гладкая функция, для которых для почти всех $\alpha \in \mathbb{R}^{n_\alpha}$ по мере Лебега выполняются условия теорем Такенса.

Наборы параметров, для которых теорема неверна, имеют меру Лебега нуль: при случайном выборе коэффициентов с любой абсолютно непрерывной плотностью вероятность получить неподходящую функцию h_α равна нулю.

Теорема связывает наблюдаемые в эксперименте временные ряды с истинным пространством состояний и задаёт пространство высокой размерности.

[2]Takens F. Detecting Strange Attractors in Turbulence // Dynamical Systems and Turbulence, Warwick 1980. Lecture Notes in Mathematics. Vol. 898. Springer, 1981. P. 366–381.

Теорема оценивает минимальную длину окна $2d_s+1$. Размер вектора задержек на практике выбирается избыточный. Для нескольких синхронных временных рядов объединение векторов задержек повторяет одни и те же зависимости.

Для снижения размерности применяется метод главных компонент. Пусть $\mathbf{X} \in \mathbb{R}^{n\times m}$ — центрированная траекторная матрица векторов задержек из (3). Тогда её сингулярное разложение определяется как

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathsf{T}}, \tag{5}$$

где $\mathbf{U}$ — матрица ортонормированных собственных векторов $\mathbf{X}\mathbf{X}^{\mathsf{T}}$, $\mathbf{V}$ — матрица ортонормированных собственных векторов $\mathbf{X}^{\mathsf{T}}\mathbf{X}$, $\mathbf{\Sigma}$ — диагональная матрица сингулярных чисел. С помощью p главных сингулярных чисел на основе матрицы $\mathbf{U}$ строится проекция в подпространство:

$$\mathbf{S} = \mathbf{W}\mathbf{X} = \left[\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_m\right], \quad \mathbf{s}_i \in \mathbb{R}^p, \quad p \ll n.$$

Объединение (5) с методом задержек (3) даёт *вложение глобальных главных компонент.*

В многомерном случае вектор временных задержек содержит каждый из рядов $\{x_{i,t}\}_{t=1}^{N}$, $i = 1, \ldots, d_{\text{obs}}$, и объединяет их в одно фазовое пространство:

$$\mathbf{x}_t = [\mathbf{x}_{1,t}, \ldots, \mathbf{x}_{d_{\text{obs}},t}]^{\mathsf{T}} = [x_{1,t}, \ldots, x_{1,t-(n_1-1)\tau_1}, \ldots, x_{d_{\text{obs}},t}, \ldots, x_{d_{\text{obs}},t-(n_{d_{\text{obs}}}-1)\tau_{d_{\text{obs}}}}]^{\mathsf{T}}.$$

Размерность n_{tot} равна сумме размерностей вложения отдельных рядов: $n_{\text{tot}} = \sum_{i=1}^{d_{\text{obs}}} n_i$.

Методы упорядочены по нарастанию сложности отображения. Основа каждого метода — это набор наблюдаемых временных рядов. Слева направо число параметров модели и сложность вычислений растёт. Схема приведена на рис. 2.

1. *Модели пространства состояний, фильтр Калмана.* Доступны временные ряды, порождаемые наблюдаемой системой. Модель динамической системы известна. Наблюдаемые значения линейны относительно вектора состояний, $x_t = \mathbf{C}\mathbf{s}_t + v_t$, модель прогнозирования линейна, $\mathbf{s}_{t+1} = \mathbf{A}\mathbf{s}_t + w_t$. Требуется задавать матрицы $\mathbf{A}$ и $\mathbf{C}$ до начала наблюдений.
2. *Анализ сингулярного спектра.* Доступны временные ряды, порождаемые наблюдаемой системой. Модель динамической системы неизвестна. По ряду строится траекторная матрица $\mathbf{X}$. К траекторной матрице применяется матричное сингулярное разложение $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathsf{T}}$. Полученные компоненты анализируются, наиболее значимые отбираются для построения моделей прогнозирования. Модель восстановления пространства состояний линейна $\mathbf{s}_t = \mathbf{W}\mathbf{x}_t$.

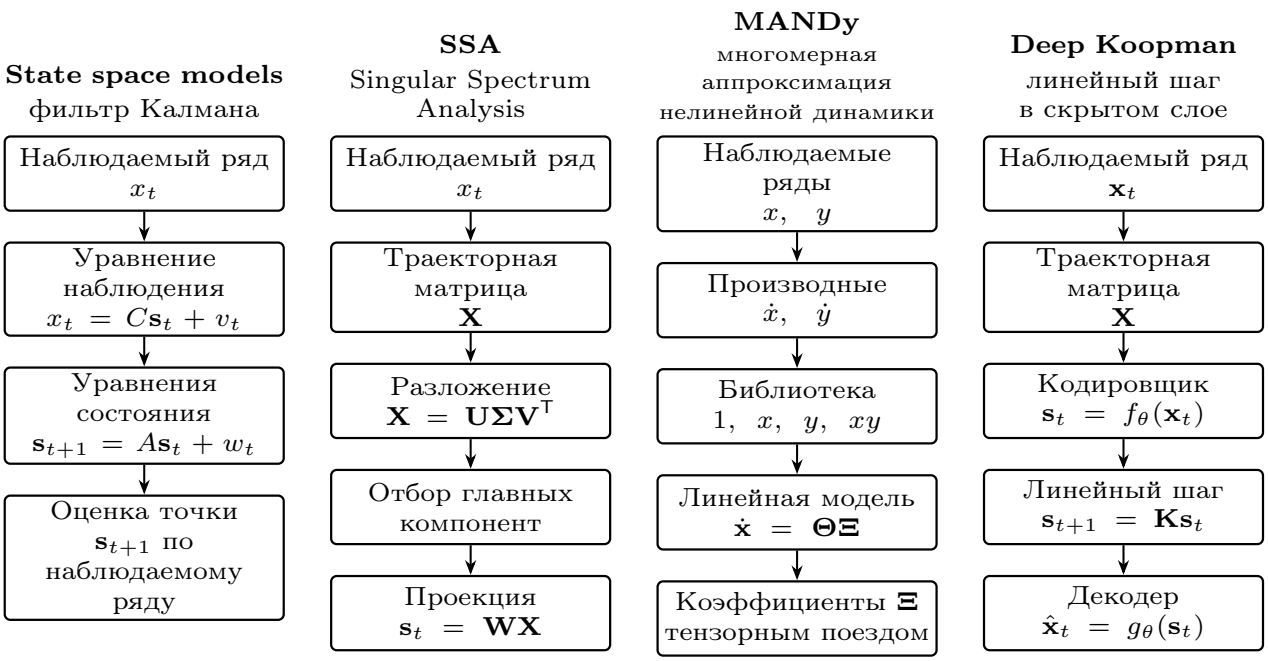


Рис. 2: Методы по нарастанию сложности. Слева направо отображение сложнее. Основа — наблюдаемый ряд.

3. *Многомерная аппроксимация нелинейной динамики (MANDy).* Доступны временные ряды и их производные, порождаемые наблюдаемой системой. Модель динамической системы неизвестна. Правая часть системы обыкновенных дифференциальных уравнений аппроксимируется линейной комбинацией библиотеки нелинейных функций x, y, xy, $\cos x$, $\sin x$. Итоговые уравнения записываются в исходных переменных. Библиотека и функции наблюдения задаются до обучения.
4. *Нейросетевой оператор Купмана.* Доступны временные ряды, порождаемые наблюдаемой системой. Модель динамической системы неизвестна. По ряду строится траекторная матрица $\mathbf{X}$. Кодировщик $\mathbf{s}_t = f_\theta(\mathbf{x}_t)$ строит проекцию в подпространство. Модель прогнозирования в подпространстве линейна, $\mathbf{s}_{t+1} = \mathbf{K}\mathbf{s}_t$. Декодер восстанавливает исходные значения временных рядов. Уравнения в исходных переменных метод не восстанавливает. Число параметров определяется архитектурой сети.

Предлагаемые полилинейные методы, например, T-SINDy, сохраняют интерпретируемость. Исходные переменные системы при этом не наблюдаются: вложение строится по рядам и по теореме о вложении. Каноническое разложение снижает число параметров этого отображения. Метод T-SINDy записывает уравнения в исходных переменных. Экспоненциальный перебор библиотеки заменяется числом параметров, линейным по рангу.

В **главе 2** пространство состояний восстанавливается полилинейным отображением векторов задержек.

Определение 6. *Тензорное произведение по n-ой моде* тензора $\mathsf{A} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ и матрицы $\mathbf{M} \in \mathbb{R}^{J \times I_n}$ определяется формулой $\mathsf{C} \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N}$ с элементами

$$\mathsf{C} = \mathsf{A} \times_n \mathbf{M}, \quad c_{i_1,\ldots,i_{n-1},j,i_{n+1},\ldots,i_N} = \sum_{i_n=1}^{I_n} a_{i_1,\ldots,i_n,\ldots,i_N} m_{j,i_n}.$$

В случае тензора $\mathsf{A} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ и вектора $\mathbf{v} \in \mathbb{R}^{I_n}$ произведение по n-ой моде даёт $\mathsf{C} \in \mathbb{R}^{I_1 \times \cdots \times I_{n-1} \times I_{n+1} \times \cdots \times I_N}$:

$$\mathsf{C} = \mathsf{A} \times_n \mathbf{v}, \quad c_{i_1,\ldots,i_{n-1},i_{n+1},\ldots,i_N} = \sum_{i_n=1}^{I_n} a_{i_1,\ldots,i_n,\ldots,i_N} v_{i_n}.$$

Без ограничения общности рассматривается случай, когда динамическая система описывается тремя временными рядами x, y, z. Для трёх рядов отображение в пространство состояний размерности p имеет вид

$$\mathbf{s}_t = \mathsf{A} \times_2 \begin{bmatrix} 1 \\ \mathbf{x}_t \end{bmatrix} \times_3 \begin{bmatrix} 1 \\ \mathbf{y}_t \end{bmatrix} \times_4 \begin{bmatrix} 1 \\ \mathbf{z}_t \end{bmatrix}, \tag{6}$$

где $\mathsf{A} \in \mathbb{R}^{p \times (n+1) \times (n+1) \times (n+1)}$, а векторы задержек $\mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t$ дополнены константой. Константа включает в модель линейные члены и их произведения. Развёртка (6) по первой моде есть полиномиальная регрессия: свободный член, линейная часть, билинейные произведения и тройные произведения векторов задержек. Полный тензор содержит $p(n+1)^3$ параметров.

То же отображение записывается в классических обозначениях линейной алгебры. Пусть $f_A : \mathbf{X} \times \mathbf{Y} \times \mathbf{Z} \longrightarrow \mathbf{S}$, где $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ и $\mathbf{S}$ — конечномерные векторные пространства размерности вложения $n+1$ и размерности результирующего пространства p. Существуют базисы $\{\mathbf{e}_{k,1}, \ldots, \mathbf{e}_{k,n+1}\}$ для каждого из $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ и базис $\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ для $\mathbf{S}$. Тензор A задаёт полилинейную функцию

$$f_A(\mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t) = \sum_{j=1}^{p} \sum_{i_1=1}^{n+1} \sum_{i_2=1}^{n+1} \sum_{i_3=1}^{n+1} \mathsf{A}_{j,i_1,i_2,i_3} \cdot \tilde{x}_{t,i_1} \tilde{y}_{t,i_2} \tilde{z}_{t,i_3} \cdot \mathbf{b}_j,$$

где $\tilde{\mathbf{x}}_t$, $\tilde{\mathbf{y}}_t$ и $\tilde{\mathbf{z}}_t$ — векторы задержек, дополненные единицей в первой координате. Развёртка полилинейного преобразования в матричной форме имеет вид полиномиальной регрессии

$$\mathbf{s}_t = \mathbf{A}_{(1)} \big[1 \overbrace{\; x_t \;\; \ldots \;\; z_{t-n} \;}^{\text{линейная часть}} \overbrace{\; x_t y_t \;\; \ldots \;\; y_{t-n} z_{t-n} \;}^{\text{билинейная часть}} \overbrace{\; x_t y_t z_t \;\; \ldots \;\; x_{t-n} y_{t-n} z_{t-n} \;}^{\text{полилинейная часть}} \big]^{\mathsf{T}},$$

где $\mathbf{A}_{(1)} \in \mathbb{R}^{p \times (n+1)^3}$ — результат развёртки тензора по первой моде.

Лемма 1 (О полилинейном расширении). *Пусть $\mathcal{M}$ — компактное гладкое многообразие, $\mathbf{F} : \mathcal{M} \to \mathcal{M}$ — диффеоморфизм класса C^2, $\tau \geq 1$ — временной шаг метода задержек и $h_i \in C^2(\mathcal{M}, \mathbb{R})$, $i = 1, \ldots, r$. Положим*

$$\mathbf{s}_i(\mathbf{v}) = \big[1,\, h_i(\mathbf{v}),\, h_i(\mathbf{F}^{-\tau}(\mathbf{v})),\, \ldots,\, h_i(\mathbf{F}^{-(n_i-1)\tau}(\mathbf{v}))\big]$$

и

$$\Phi(\mathbf{v}) = \bigotimes_{i=1}^{r} \mathbf{s}_i(\mathbf{v}) \in \mathbb{R}^{\prod_{i=1}^{r}(n_i+1)},$$

где $\otimes$ — кронекерово произведение. Пусть $\Psi(\mathbf{v})$ — конкатенация линейных задержек,

$$\Psi(\mathbf{v}) = \big[h_1(\mathbf{v}), \ldots, h_1(\mathbf{F}^{-(n_1-1)\tau}(\mathbf{v})),\, \ldots,\, h_r(\mathbf{v}), \ldots, h_r(\mathbf{F}^{-(n_r-1)\tau}(\mathbf{v}))\big] \in \mathbb{R}^{\sum_{i=1}^{r} n_i}.$$

Тогда Φ является гладким вложением $\mathcal{M}$ тогда и только тогда, когда вложением является Ψ.

Проекция на координаты, в которых из одного окна взята задержка, а из остальных — единица, возвращает Ψ и отбрасывает произведения. Если $\sum_i n_i \geq 2d_s + 1$, то для почти всех функций наблюдения вложением является Ψ, а вместе с ним и Φ. Достаточное условие на Ψ устанавливает многомерный метод задержек.

Теорема 1 (О включении линейной векторной авторегрессии). *Пусть имеется m наблюдаемых временных рядов, удовлетворяющих линейной векторной авторегрессии порядка p_{ar}:*

$$\mathbf{x}_{t+1} = \sum_{k=1}^{p_{\mathrm{ar}}} \mathbf{A}_k \mathbf{x}_{t-k+1} + \boldsymbol{\varepsilon}_t, \qquad \mathbf{x}_t = (x_{1,t}, \ldots, x_{m,t})^\top \in \mathbb{R}^m, \quad \mathbf{A}_k \in \mathbb{R}^{m \times m},$$

где $\boldsymbol{\varepsilon}_t$ — центрированный шум с конечной дисперсией. Вектор задержек упорядочим по лагам и по компонентам ряда и дополним его константой:

$$\mathbf{z}_t = \big[1,\, \mathbf{x}_t^\top,\, \mathbf{x}_{t-1}^\top,\, \ldots,\, \mathbf{x}_{t-p_{\mathrm{ar}}+1}^\top\big]^\top \in \mathbb{R}^q, \qquad q = mp_{\mathrm{ar}} + 1.$$

Тогда матрица

$$\mathbf{W} = \big[\mathbf{0}_{m\times 1},\, \mathbf{A}_1,\, \ldots,\, \mathbf{A}_{p_{\mathrm{ar}}}\big] \in \mathbb{R}^{m \times q}$$

удовлетворяет $\mathbf{W}\mathbf{z}_t = \sum_{k=1}^{p_{\mathrm{ar}}} \mathbf{A}_k \mathbf{x}_{t-k+1}$. Рассмотрим полилинейную модель порядка $K \geq 1$

$$\mathbf{s}_t = \mathsf{B} \times_2 \mathbf{z}_t \times_3 \mathbf{z}_t \times_4 \cdots \times_{K+1} \mathbf{z}_t$$

с тензором B, у которого первая мода имеет размер p, а каждая из следующих K мод — размер q. Существует такой тензор, что полилинейная

модель совпадает с детерминированной частью векторной авторегрессии с точностью до перенормировки. При $p = m$

$$\mathbf{s}_t = \mathbf{W}\mathbf{z}_t = \mathbf{x}_{t+1} - \boldsymbol{\varepsilon}_t.$$

При $p > m$ *то же равенство выполнено для первых* m *компонент* $\mathbf{s}_t$*, а остальные компоненты равны нулю.*

Число параметров снижается низкоранговым разложением. Каноническое разложение ранга R записывает (6) через фактор-матрицы:

$$\mathbf{s}_t \approx \left(\sum_{r=1}^{R} \mathbf{a}_{r,1} \otimes \mathbf{a}_{r,2} \otimes \mathbf{a}_{r,3} \otimes \mathbf{a}_{r,4}\right) \times_2 \begin{bmatrix}1\\ \mathbf{x}_t\end{bmatrix} \times_3 \begin{bmatrix}1\\ \mathbf{y}_t\end{bmatrix} \times_4 \begin{bmatrix}1\\ \mathbf{z}_t\end{bmatrix} = \\ = \mathbf{A}_1\Big(\mathbf{A}_2^T \begin{bmatrix}1\\ \mathbf{x}_t\end{bmatrix} \odot \mathbf{A}_3^T \begin{bmatrix}1\\ \mathbf{y}_t\end{bmatrix} \odot \mathbf{A}_4^T \begin{bmatrix}1\\ \mathbf{z}_t\end{bmatrix}\Big). \quad (7)$$

Здесь $\odot$ — поэлементное произведение Адамара. Число параметров равно $pR + 3(n+1)R$ и линейно по длине окна и по числу рядов. Сравнение числа параметров методов тензорного разложения показано в табл. 1. Для одного ряда и единичного тензора запись совпадают с методом задержек при $n \geq 2d_s + 1$. Веса (7) обучаются обратным распространением ошибки совместно с линейным шагом $\mathbf{s}_{t+1} = \mathbf{W}_{\text{ds}}\mathbf{s}_t$. Ошибка восстановления рядов

$$\mathcal{L}_{\text{rec}} = \frac{1}{T}\sum_{t=1}^{T}\big(\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_2^2 + \|\mathbf{y}_t - \hat{\mathbf{y}}_t\|_2^2 + \|\mathbf{z}_t - \hat{\mathbf{z}}_t\|_2^2\big)$$

минимизируется по фактор-матрицам и ядрам разложения. Ошибка шага в пространстве состояний

$$\mathcal{L}_{\text{pred}} = \frac{1}{T-1}\sum_{t=1}^{T-1}\|\mathbf{s}_{t+1} - \mathbf{W}_{\text{ds}}\mathbf{s}_t\|_2^2$$

минимизируется по матрице $\mathbf{W}_{\text{ds}}$.

Таблица 1: Число параметров полилинейного отображения трёх рядов

Разложение	Число параметров
Без разложения	$p(n+1)^3$
Каноническое	$R\big(p + 3(n+1)\big)$
Таккера	$pR^3 + 3(n+1)R$
Тензорный поезд	$pR + (n+1)(2R^2 + R)$

Теорема 2 (О многомерном полилинейном вложении). *Пусть* $\Psi : \mathcal{M} \to \mathbb{R}^{n_1+1} \times \cdots \times \mathbb{R}^{n_r+1}$ *— гладкое вложение компактного многообразия размерности* d_s,

$$\Psi(\mathbf{v}) = [s_1(\mathbf{v}), \ldots, s_r(\mathbf{v})], \qquad s_i(\mathbf{v}) = \begin{bmatrix} 1 \\ \mathbf{x}_i(\mathbf{v}) \end{bmatrix}.$$

Каноническая проекция ранга R *задаётся фактор-матрицами* $\mathbf{A}_1 \in \mathbb{R}^{p \times R}$ *и* $\mathbf{A}_{i+1} \in \mathbb{R}^{(n_i+1) \times R}$:

$$\mathbf{g}(\mathbf{v}) = \mathbf{A}_2^{\mathsf{T}} s_1(\mathbf{v}) \odot \cdots \odot \mathbf{A}_{r+1}^{\mathsf{T}} s_r(\mathbf{v}), \qquad \mathbf{s}(\mathbf{v}) = \mathbf{A}_1 \mathbf{g}(\mathbf{v}).$$

Если $R \geq 2d_s + 1$, *то множество матриц* $\mathbf{A}_2, \ldots, \mathbf{A}_{r+1}$, *для которых* $\mathbf{g}$ *является гладким вложением* $\mathcal{M}$ *в* $\mathbb{R}^R$, *непусто и открыто. Если дополнительно* $p \geq 2d_s + 1$, *то множество наборов* $(\mathbf{A}_1, \ldots, \mathbf{A}_{r+1})$, *для которых* $\mathbf{s}$ *является гладким вложением* $\mathcal{M}$ *в* $\mathbb{R}^p$, *непусто и открыто. Для погружения необходимы* $R \geq d_s$ *и* $p \geq d_s$.

Открытость означает, что вложение сохраняется при малом отклонении фактор-матриц. Теорема обосновывает класс канонических моделей. Для системы Лоренца $d_s = 3$, поэтому достаточное условие требует $R \geq 7$ и $p \geq 7$.

В **главе 3** несколько рядов объединяются в одно пространство состояний. Сравниваются четыре способа: вложение глобальных компонент траекторной матрицы, объединение исходных векторов задержек полносвязной сетью, объединение скрытых представлений отдельных сетей и полилинейное объединение этих представлений.

Для анализа нелинейных динамическиз систем используются нейросетевые методы. Первый способ объединяет исходные векторы задержек $[\mathbf{x}_t, \mathbf{y}_t, \mathbf{z}_t]^{\mathsf{T}}$. Второй объединяет уже полученные представления $[\mathbf{h}_{\mathrm{x},t}, \mathbf{h}_{\mathrm{y},t}, \mathbf{h}_{\mathrm{z},t}]^{\mathsf{T}}$. Задача представима как минимизация ошибки восстановления исходных временных рядов:

$$\hat{\theta}, \hat{\phi} = \arg\min_{\theta,\phi} \frac{1}{n} \sum_{i=1}^{n} \mathcal{L}\big(\mathbf{X}, g_\phi(f_\theta(\mathbf{X}))\big),$$

где θ — параметры сети снижения размерности $f_\theta : \mathbb{R}^n \to \mathbb{R}^p$, ϕ — параметры сети восстановления ряда $g_\phi : \mathbb{R}^p \to \mathbb{R}^n$.

Линейное объединение $[\mathbf{h}_{x,t}, \mathbf{h}_{y,t}, \mathbf{h}_{z,t}]$ не включает произведения компонент разных рядов. Полилинейное объединение задаётся тензором связей

$$\mathbf{s}_t = \mathsf{W} \times_2 \begin{bmatrix} 1 \\ \mathbf{h}_{x,t} \end{bmatrix} \times_3 \begin{bmatrix} 1 \\ \mathbf{h}_{y,t} \end{bmatrix} \times_4 \begin{bmatrix} 1 \\ \mathbf{h}_{z,t} \end{bmatrix}. \tag{8}$$

Каноническое разложение представляет тензор суммой R внешних произведений $\mathsf{W} \approx \sum_{r=1}^{R} \mathbf{w}_{r,1} \otimes \mathbf{w}_{r,2} \otimes \mathbf{w}_{r,3} \otimes \mathbf{w}_{r,4}$ и записывает (8) в виде

$$\mathbf{s}_t = \mathbf{W}_1\big(\mathbf{W}_2^{\mathsf{T}} \begin{bmatrix} 1 \\ \mathbf{h}_{\mathrm{x},t} \end{bmatrix} \odot \mathbf{W}_3^{\mathsf{T}} \begin{bmatrix} 1 \\ \mathbf{h}_{\mathrm{y},t} \end{bmatrix} \odot \mathbf{W}_4^{\mathsf{T}} \begin{bmatrix} 1 \\ \mathbf{h}_{\mathrm{z},t} \end{bmatrix}\big).$$

Каноническое разложение заменяет W фактор-матрицами по той же схеме, что (7). Вид нейросетевой модели с каноническим разложением показан на рис. 3.

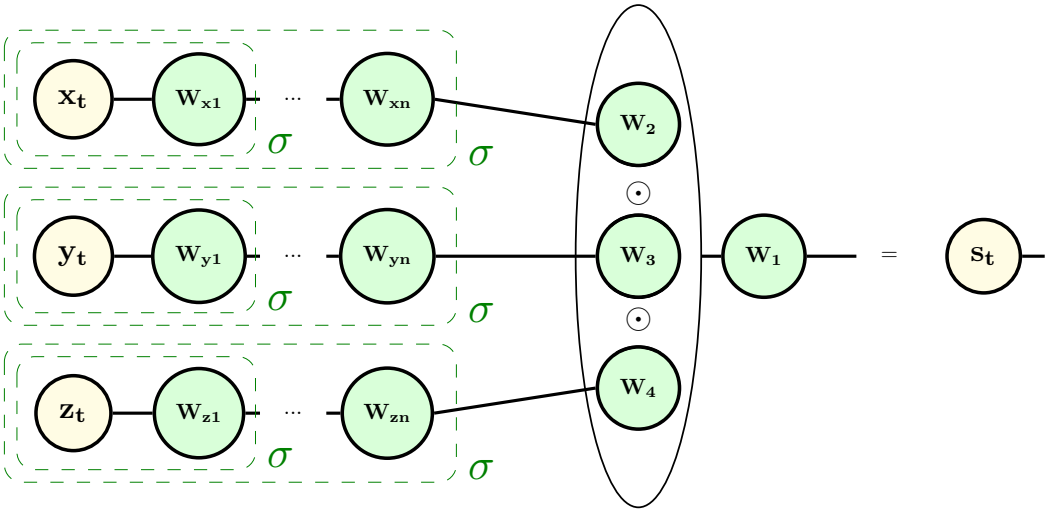


Рис. 3: Полилинейное объединение подпространств с каноническим разложением

В **главе 4** восстанавливается аналитический вид дифференциальных уравнений по неполным наблюдениям.

Рассматривается динамическая система, задаваемая системой обыкновенных дифференциальных уравнений

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{v}(t) = \mathbf{f}_{\mathrm{ds}}(\mathbf{v}(t)),$$

где $\mathbf{f}_{\mathrm{ds}} : \mathbb{R}^{d_s} \to \mathbb{R}^{d_s}$ — гладкая функция, $\mathbf{v}(t) \in \mathbb{R}^{d_s}$ — точка пространства состояний. Истинные переменные $\mathbf{v}(t)$ недоступны. Наблюдаемый ряд $\mathbf{x}(t) \in \mathbb{R}^{d_{\mathrm{obs}}}$ связан с $\mathbf{v}(t)$ неизвестной функцией наблюдения $\mathbf{h} : \mathbb{R}^{d_s} \to \mathbb{R}^{d_{\mathrm{obs}}}$, причём $d_{\mathrm{obs}} < d_s$. Восстановление аналитического вида по наблюдениям есть поиск функции $\mathbf{f}_{\mathrm{obs}}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = \mathbf{f}_{\mathrm{obs}}(\mathbf{x}(t)), \quad \mathbf{x}(t) = \mathbf{h}(\mathbf{v}(t)). \tag{9}$$

Наблюдения дискретны: $\mathbf{x}_1, \ldots, \mathbf{x}_T$ в моменты $t_1 < \cdots < t_T$ с шагом Δt. Конечно-разностная схема связывает $\mathbf{f}_{\mathrm{obs}}$ с рядом:

$$\frac{\mathbf{x}_{t+1} - \mathbf{x}_t}{\Delta t} \approx \mathbf{f}_{\mathrm{obs}}(\mathbf{x}_t).$$

Метод SINDy строит библиотеку функций-кандидатов $\boldsymbol{\Theta}(\mathbf{X}_t)$ и ищет разреженную матрицу коэффициентов как решение задачи минимизации

$$\min_{\boldsymbol{\Xi}} \frac{1}{T-1} \sum_{t=1}^{T-1} \left\| \mathbf{x}_{t+1} - \widehat{\mathbf{f}}(\mathbf{x}_t, \boldsymbol{\Xi}) \right\|_2^2 + \mathcal{R}(\boldsymbol{\Xi}).$$

Для разреженной регрессии строятся матрица данных $\mathbf{X}_t \in \mathbb{R}^{(T-1)\times d_{\text{obs}}}$ и матрица прогнозируемых значений $\mathbf{X}_{t+1} \in \mathbb{R}^{(T-1)\times d_{\text{obs}}}$. Библиотека функций-кандидатов $\{\theta_1, \ldots, \theta_{n_\theta}\}$ состоит из нелинейных функций и их комбинаций и преобразует матрицу данных:

$$\boldsymbol{\Theta}(\mathbf{X}_t) = \left[\theta_1(\mathbf{X}_t),\ \ldots,\ \theta_{n_\theta}(\mathbf{X}_t)\right] \in \mathbb{R}^{(T-1)\times(n_\theta d_{\text{obs}})}.$$

Решение SINDy есть линейная комбинация кандидатов:

$$\mathbf{X}_{t+1} = \boldsymbol{\Theta}(\mathbf{X}_t)\,\boldsymbol{\Xi},$$

где $\boldsymbol{\Xi} \in \mathbb{R}^{(n_\theta d_{\text{obs}})\times d_{\text{obs}}}$ — матрица коэффициентов. Малое число ненулевых компонент $\boldsymbol{\Xi}$ находится последовательным пороговым методом наименьших квадратов.

Таблица 2: Сравнение подходов к восстановлению уравнений динамической системы

Метод	Преимущества	Недостатки
SINDy	Уравнения в исходных переменных; разреженный отбор; малый объём данных	Заданная библиотека; чувствительность к шуму; нужны истинные переменные; число параметров растёт экспоненциально по размеру библиотеки
Нейросетевые (Neural ODE, PINN)	Аппроксимация сложных систем; устойчивость к шуму	Нет уравнений в исходных переменных; подбор архитектуры; больший объём данных
DMD / eDMD / Koopman	Линейные уравнения в поднятом пространстве; малый объём данных	Уравнения в скрытом пространстве; при нелинейном расширении те же ограничения, что у SINDy
T-SINDy	Уравнения в исходных переменных; линейный рост числа параметров по размеру библиотеки	Заданная библиотека; умеренный объём данных; зависимость от ранга разложения

Если доступны только ряды из (9), пространство состояний строится методом задержек. При $d_{\text{obs}} > 1$ строится объединённый вектор задержек

$$\mathbf{z}_t = \left[x_{1,t}, x_{1,t-\tau}, \ldots, x_{1,t-(n_1-1)\tau},\ \ldots,\ x_{d_{\text{obs}},t-(n_{d_{\text{obs}}}-1)\tau}\right]^\top \in \mathbb{R}^{n_{\text{tot}}},$$

где $n_{\text{tot}} = \sum_{i=1}^{d_{\text{obs}}} n_i$. Отображение из $\mathbb{R}^{n_{\text{tot}}}$ в подпространство $\mathbb{R}^{d_{\text{lat}}}$ задаётся автоэнкодером

$$\mathbf{s}_t = E(\mathbf{z}_t; \mathbf{W}_{\text{enc}}) \in \mathbb{R}^{d_{\text{lat}}}, \qquad \tilde{\mathbf{z}}_t = D(\mathbf{s}_t; \mathbf{W}_{\text{dec}}) \in \mathbb{R}^{n_{\text{tot}}}.$$

В подпространстве аналитический вид восстанавливается методом SINDy:

$$\mathbf{S}_{t+1} = \boldsymbol{\Theta}(\mathbf{S}_t)\,\boldsymbol{\Xi}.$$

Уравнения зависят от архитектуры автоэнкодера и не записаны в исходных переменных.

Число столбцов библиотеки растёт экспоненциально с числом кандидатов и с длиной окна. Уравнения SINDy записаны в исходных переменных, если эти переменные наблюдаются. Нейросетевые модели и оператор Купмана при неполном наблюдении переносят уравнения в скрытое пространство. Сравнение методов показано в табл. 2.

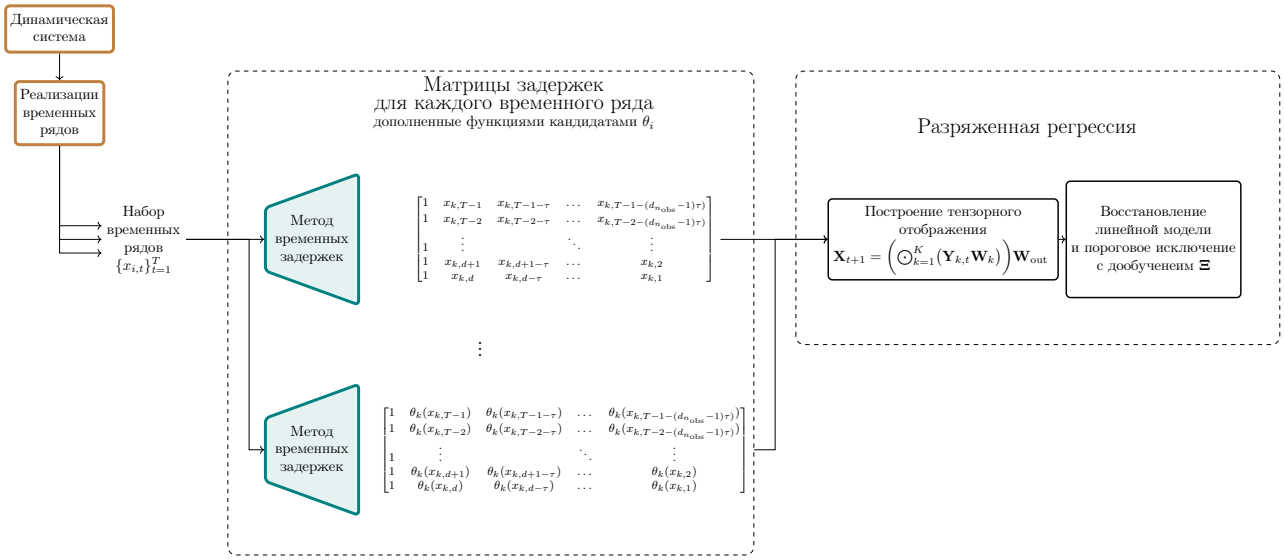


Рис. 4: Схема метода T-SINDy.

Метод T-SINDy заменяет перебор комбинаций тензором. Схема метода показана на рис. 4. Для каждой компоненты и каждого кандидата строится матрица признаков: траекторная матрица, дополненная столбцом единиц, либо то же вектор задержек после нелинейного преобразования. Число матриц складывается из d_{obs} матриц задержек и $n_\theta d_{\text{obs}}$ матриц, преобразованных кандидатами:

$$K = d_{\text{obs}} + n_\theta \, d_{\text{obs}}.$$

При $k = 1, \ldots, d_{\text{obs}}$

$$\mathbf{Y}_{k,t} = \begin{bmatrix} 1 & x_{k,T-1} & x_{k,T-1-\tau} & \cdots & x_{k,T-1-(n_{d_{\text{obs}}}-1)\tau} \\ 1 & x_{k,T-2} & x_{k,T-2-\tau} & \cdots & x_{k,T-2-(n_{d_{\text{obs}}}-1)\tau} \\ \vdots & \vdots & & \ddots & \vdots \\ 1 & x_{k,n} & x_{k,n-\tau} & \cdots & x_{k,1} \end{bmatrix},$$

а при $k = d_{\text{obs}} + 1, \ldots, K$ та же матрица строится по преобразованию θ_k. Та же регрессия записывается тензором W:

$$x_{i,t+1} = \sum_{j_1=1}^{n_1+1} \cdots \sum_{j_K=1}^{n_K+1} \mathsf{W}_{j_1,\ldots,j_K,i} \prod_{k=1}^{K} y_{k,j_k,t}, \quad \mathsf{W} \in \mathbb{R}^{(n_1+1)\times\cdots\times(n_K+1)\times d_{\text{obs}}},$$

где $y_{k,j_k,t}$ — j_k-й элемент строки признаков k-й модальности в момент t. Прогноз i-й компоненты есть свёртка тензора W со строками этих матриц. Полный тензор содержит $d_{\text{obs}} \prod_k (n_k+1)$ элементов. Каноническое разложение ранга R записывает прогноз как

$$\mathbf{X}_{t+1} = \left(\bigodot_{k=1}^{K} \left(\mathbf{Y}_{k,t} \mathbf{W}_k \right) \right) \mathbf{W}_{\text{out}}.$$

Число параметров равно $R\big(\sum_k (n_k+1)+d_{\text{obs}}\big)$ и линейно по числу модальностей K и по рангу R. Построение библиотеки показано на рис. 5. Повтор одного окна на нескольких модах строит полином: степень равна числу повторов. Нелинейные преобразования ряда, например sin и $1/x$, входят в свёртку отдельными модами. Таблица всех мономов для такого подхода не строится.

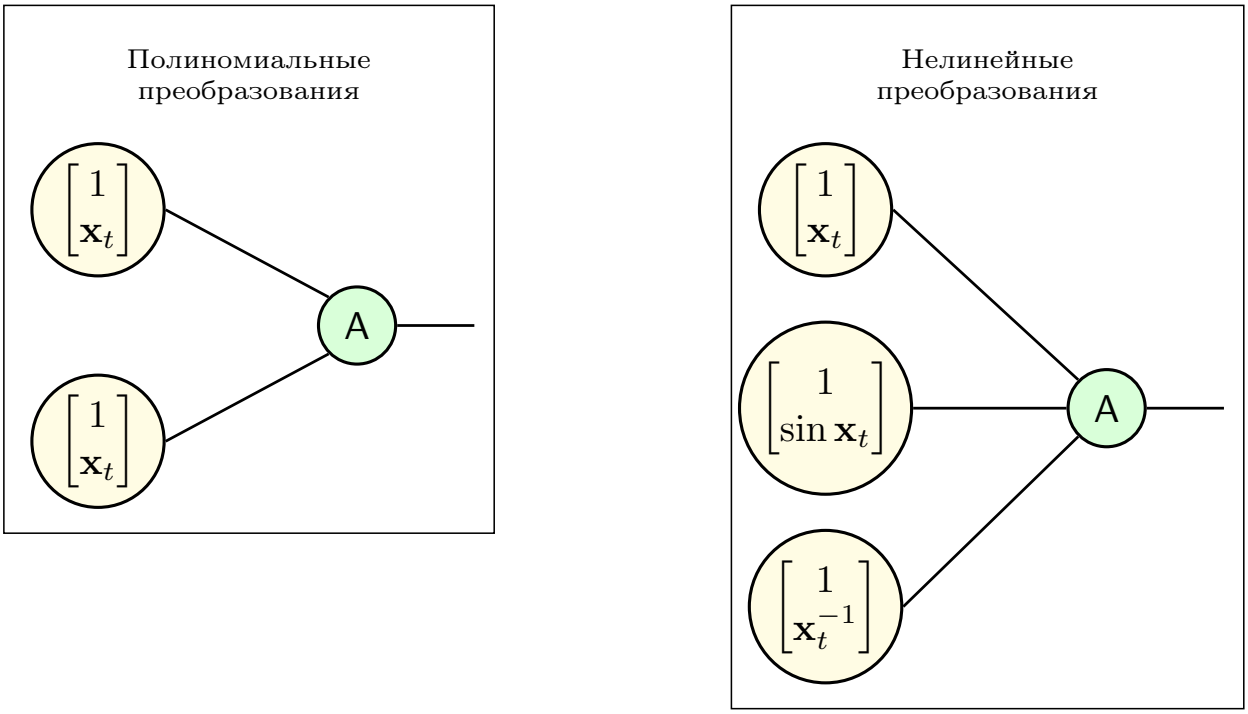


Рис. 5: Библиотека метода T-SINDy. Слева повтор одного окна, справа преобразования ряда.

Слева на рис. 5 одно и то же окно $\mathbf{x}_t$ подаётся на несколько мод тензора A. Произведение мод даёт полиномиальные комбинации задержек: x_t^2, $x_t x_{t-1}$, x_{t-1}^2 и далее по длине окна. Таким образом итоговые уравнения остаются в исходных переменных.

В **главе 5** пространство состояний новой системы класса строится по короткому ряду и вектору статических параметров. Классом называются системы с общим законом эволюции, различающиеся вектором $\mathbf{c} \in \mathcal{C} \subset \mathbb{R}^{d_c}$. Рассматривается конечный набор динамических систем $\{\mathbf{E}_i\}_{i=1}^{N_E}$ одного класса.

Каждая система $\mathbf{E}_i$ характеризуется вектором статических параметров $\mathbf{c}_i \in \mathcal{C} \subset \mathbb{R}^{d_c}$ и временным рядом $\{x_{i,t}\}_{t=1}^{T_i}$. Функция эволюции $\mathbf{E}$ неизвестна и задаёт будущее состояние системы

$$x_{i,t+1} = \mathbf{E}(x_{i,t}, \mathbf{c}_i).$$

По конечному ряду строится отображение в пространство состояний

$$\mathbf{s}_{i,t} = \tilde{\mathbf{E}}(x_{i,1}, \ldots, x_{i,T_i}, \mathbf{c}_i, \mathbf{w}_i),$$

где $\mathbf{w}_i \in \mathbb{R}^{N_w}$ — настраиваемые параметры модели, $\mathbf{s}_{i,t} \in \mathbb{R}^p$. Предполагается, что существует гладкое отображение $\mathbf{P} : \mathbb{R}^p \to \mathbb{R}^p$, связывающее пространства состояний двух систем класса:

$$\mathbf{s}_{j,t} = \mathbf{P}(\mathbf{s}_{i,t}, \mathbf{c}_i, \mathbf{c}_j, \mathbf{w}_P) + \varepsilon_t.$$

Задача: для новой системы $\mathbf{E}_{N_E+1}$ с известным вектором статических параметров $\mathbf{c}_{N_E+1}$ и коротким временным рядом $\{x_{N_E+1,t}\}_{t=1}^{T_{\text{short}}}$, недостаточным для построения пространства состояний, найти отображение $\tilde{\mathbf{E}}_{N_E+1}$ такое, что восстановленное пространство состояний $\hat{\mathbf{S}}_{N_E+1}$ аппроксимирует пространство $\mathbf{S}_{N_E+1}$, которое было бы получено по полному ряду:

$$\min_{\mathbf{w}_P} \sum_{i \neq j} \|\mathbf{S}_j - \mathbf{P}(\mathbf{S}_i, \mathbf{c}_i, \mathbf{c}_j, \mathbf{w}_P)\|^2.$$

Первый способ задаёт проекцию всех систем одним тензором. Вектор задержек и вектор параметров дополняются константой:

$$\mathbf{s}_{i,t} = \mathsf{B} \times_2 \begin{bmatrix} 1 \\ \mathbf{x}_{i,t} \end{bmatrix} \times_3 \begin{bmatrix} 1 \\ \mathbf{c}_i \end{bmatrix} = \mathsf{B}(\mathbf{c}_i) \begin{bmatrix} 1 \\ \mathbf{x}_{i,t} \end{bmatrix}.$$

Тензор $\mathsf{B} \in \mathbb{R}^{p \times (n+1) \times (d_c+1)}$ один на весь класс. Свёртка по моде параметров строит матрицу проекции $\mathsf{B}(\mathbf{c}_i)$. На этапе обучения внешние произведения векторов задержек и параметров всех систем класса собираются в одну матрицу:

$$\mathbf{z}_{i,t} = \operatorname{vec}\left(\begin{bmatrix} 1 \\ \mathbf{x}_{i,t} \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{c}_i \end{bmatrix}^{\mathsf{T}} \right) \in \mathbb{R}^{(n+1)(d_c+1)},$$

и метод главных компонент строит развёртку $\mathbf{B}_{(1)} \in \mathbb{R}^{p \times (n+1)(d_c+1)}$ тензора B. На этапе восстановления для нового $\mathbf{c}$ матрица проекции строится одной свёрткой

$$\mathsf{B}(\mathbf{c}_{N_E+1}) = \mathsf{B} \times_3 \begin{bmatrix} 1 \\ \mathbf{c}_{N_E+1} \end{bmatrix},$$

короткий ряд отображается этой матрицей, обратный переход задаётся псевдообратной матрицей $\mathsf{B}(\mathbf{c}_{N_E+1})^\dagger$.

Теорема 3 (Липшицева непрерывность параметризованного полилинейного отображения)**.** *Пусть* $\mathcal{C} \subset \mathbb{R}^{d_c}$ *— множество допустимых статических параметров. Рассмотрим семейство полилинейных отображений*

$$\{\, \mathsf{B}(\mathbf{c}) : \mathbb{R}^{n+1} \to \mathbb{R}^{p} \,\}_{\mathbf{c}\in\mathcal{C}},$$

заданных как

$$\mathsf{B}(\mathbf{c}) = \mathsf{B}_0 \times_3 \begin{pmatrix} 1 \\ \mathbf{c} \end{pmatrix},$$

где $\mathsf{B}_0 \in \mathbb{R}^{p\times(n+1)\times(d_c+1)}$ *— фиксированный тензор. Тогда отображение*

$$\Psi : \mathcal{C} \to \mathbb{R}^{p\times(n+1)}, \qquad \mathbf{c} \mapsto \mathsf{B}(\mathbf{c})$$

липшицево. Существует константа $L > 0$ *такая, что для любых* $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C}$

$$\|\mathsf{B}(\mathbf{c}_1) - \mathsf{B}(\mathbf{c}_2)\|_F \le L\,\|\mathbf{c}_1 - \mathbf{c}_2\|_2,$$

где $\|\cdot\|_F$ *— норма Фробениуса матрицы* $p \times (n+1)$*, а* $\|\cdot\|_2$ *— евклидова норма вектора параметров.*

Малое изменение статических параметров меняет матрицу проекции не сильнее, чем на эту константу, умноженную на норму приращения. Пространства состояний близких по параметрам систем остаются близкими.

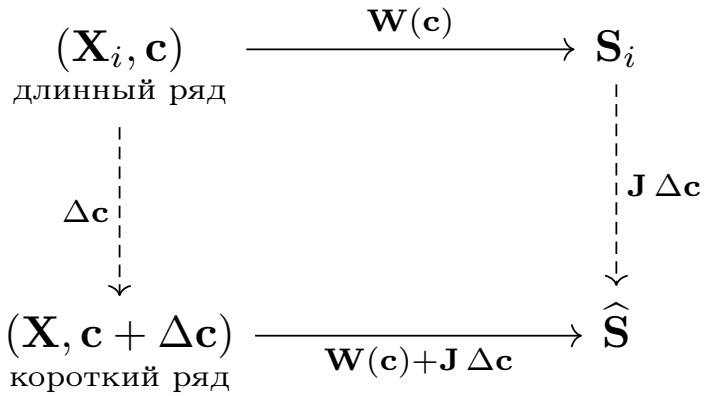


Рис. 6: Переход к системе с коротким временным рядом рядом.

Второй способ не хранит общий тензор класса. Проекция новой системы получается из проекции ближайшей по параметрам системы. Линейная часть отображения $\mathbf{P}$ в окрестности точки $\mathbf{s}_{1,0}$ задаётся матрицей Якоби

$$\mathbf{s}_{2,t} = \mathbf{P}(\mathbf{s}_{1,0}) + \mathbf{J}\,(\mathbf{s}_{1,t} - \mathbf{s}_{1,0}) + o(\|\mathbf{s}_{1,t} - \mathbf{s}_{1,0}\|), \quad \mathbf{J} = \begin{bmatrix} \frac{\partial P_1}{\partial s_{1,1}} & \cdots & \frac{\partial P_1}{\partial s_{1,p}} \\ \vdots & \ddots & \vdots \\ \frac{\partial P_p}{\partial s_{1,1}} & \cdots & \frac{\partial P_p}{\partial s_{1,p}} \end{bmatrix} \in \mathbb{R}^{p\times p}.$$

Для аффинного отображения $\mathbf{P}$ матрица Якоби не зависит от точки, а для центрированных траекторий $\mathbf{s}_{2,t} = \mathbf{J}\mathbf{s}_{1,t}$. Строится линейная по приращению параметров модель матрицы Якоби

$$\mathbf{J}_{ij} = \mathbf{I} + \mathsf{J} \times_3 (\mathbf{c}_j - \mathbf{c}_i), \qquad \mathbf{s}_{j,t} = \mathbf{J}_{ij}\mathbf{W}_i\mathbf{x}_{j,t}.$$

При $\mathbf{c}_j = \mathbf{c}_i$ переход тождествен. Тензор $\mathsf{J} \in \mathbb{R}^{p\times p\times d_c}$ обучается так, чтобы переход в пространство другой системы класса и обратно восстанавливал траекторную матрицу:

$$\min_{\mathbf{W},\mathsf{J}} \sum_{i=1}^{N_E}\sum_{j=1}^{N_E} \|\mathbf{X}_i - \mathbf{W}_i^{\dagger}\mathbf{J}_{ji}\mathbf{J}_{ij}\mathbf{W}_i\mathbf{X}_i\|^2.$$

Для новой системы выбирается ближайшая обученная система, по приращению параметров строится $\mathbf{J}$, и короткий ряд отображается преобразованной проекцией. Схема показана на рис. 6.

Публикации соискателя по теме диссертации

Публикации в журналах из списка ВАК.

1. Тихонов Д. М., Стрижов В. В. Объединение пространств состояний динамической системы через полилинейное отображение // Труды ИСП РАН, 2026. Т. 38, вып. 6, ч. 1. С. 63–72.
2. Тихонов Д. М., Стрижов В. В. Тензорное разреженное восстановление дифференциальных уравнений нелинейных динамических систем // Моделирование и анализ информационных систем, 2026. Т. 33, № 3. С. 344–357.
3. Тихонов Д. М. Полилинейное отображение для мультимоделирования динамических систем // Труды МФТИ, 2026. Т. 18, № 2(70). С. 88–97.

Прочие публикации.

4. Tikhonov D. M., Strijov V. V. Dynamic system reconstruction from multivariate time series via multilinear map and time delay embedding // International Conference on Computational Optimization, Иннополис, 2024.
5. Тихонов Д. М. Восстановление зашумленных значений многомерных временных рядов с помощью тензорного метода задержек и тензорного разложения // 66-я Всероссийская научная конференция МФТИ: тезисы докладов. Прикладная математика и информатика, 2024. С. 128–130.
6. Тихонов Д. М. Классификация квазипериодических временных сигналов с помощью сферических гармоник // Материалы Международного молодежного научного форума «ЛОМОНОСОВ-2022», 2022.