**Министерство науки и высшего образования Российской Федерации**

федеральное государственное автономное образовательное учреждение   
высшего образования

«Национальный исследовательский Томский политехнический университет»

**ОТЧЁТ**

|  |  |
| --- | --- |
| Тема | Применение тензорного модального разложения для редукции порядка моделей и оптимального размещения сенсоров |

|  |  |
| --- | --- |
| Выполнил | Саматов Денис Сергеевич |

Томск 2024

Оглавление

[Введение 3](#_Toc180181720)

[Обзор Статьи 3](#_Toc180181721)

[Ключевые вклады 3](#_Toc180181722)

[Методология 4](#_Toc180181723)

[Альтернативные методы модального разложения 5](#_Toc180181724)

[Разложение по Собственным Функциям (Proper Orthogonal Decomposition, POD) 5](#_Toc180181725)

[Динамическое Разложение по Модам (Dynamic Mode Decomposition, DMD) 6](#_Toc180181726)

[Сравнение POD, DMD и TBMD 7](#_Toc180181727)

[Результаты работы TBMD алгоритма 8](#_Toc180181728)

[Сильные и слабые стороны работы 9](#_Toc180181729)

[Перспективы дальнейших исследований 9](#_Toc180181730)

[Заключение 10](#_Toc180181731)

# **Введение**

Современные методы моделирования сложных динамических систем играют важнейшую роль в различных областях науки и техники, включая аэродинамику, гидродинамику, климатологию и обработку изображений. Одной из основных задач в анализе таких систем является сокращение их размерности при сохранении ключевой информации, что позволяет существенно снизить вычислительную сложность и повысить эффективность моделирования. Для решения этой задачи активно применяются методы модального разложения, такие как разложение по собственным функциям (Proper Orthogonal Decomposition, POD) и динамическое разложение по модам (Dynamic Mode Decomposition, DMD). Тем не менее, данные методы имеют ограничения, связанные с потерей пространственной структуры данных и недостаточной эффективностью при работе с высокоразмерными тензорными данными.

В исследовании, рассмотренном в данной статье, авторы предлагают новый подход — метод тензорного модального разложения (Tensor-Based Modal Decomposition, TBMD), который позволяет эффективно извлекать низкоразмерные моды из высокоразмерных динамических систем с минимальной потерей энергии. Более того, авторы разработали метод тензорной QR-факторизации для оптимального размещения датчиков, что способствует сохранению пространственной структуры входных данных и обеспечивает высокую точность реконструкции системы даже при ограниченном количестве измерений.

# **Обзор Статьи**

Статья *"A Novel Tensor-Based Modal Decomposition Method for Reduced Order Modeling and Optimal Sparse Sensor Placement"* под авторством Чжиронга Чжуна и его коллег представляет собой значительный вклад в развитие методологии построения моделей пониженного порядка (Reduced Order Modeling, ROM). Основная цель исследования заключается в разработке инновационного метода тензорного разложения, который эффективно обрабатывает высокоразмерные динамические системы и оптимизирует разреженное размещение датчиков, что особенно актуально при создании цифровых двойников сложных систем.

Авторы акцентируют внимание на ограничениях традиционных методов модального разложения, таких как Proper Orthogonal Decomposition (POD) и Dynamic Mode Decomposition (DMD). Эти методы, будучи широко используемыми для анализа динамических систем, сталкиваются с трудностями при работе с высокоразмерными тензорными данными. Проблема заключается в необходимости преобразования данных в матричную форму, что приводит к утрате пространственной информации и повышению вычислительных затрат. В ответ на эти вызовы авторы статьи предлагают новый подход, основанный на тензорных методах, который позволяет сохранить многомерную структуру данных и обеспечить более эффективное и точное разложение.

Разработанный метод Tensor-Based Modal Decomposition (TBMD) направлен на извлечение низкоразмерных мод с минимальными потерями энергии, а предложенная тензорная QR-факторизация обеспечивает оптимальное размещение датчиков. Такой подход позволяет не только эффективно моделировать высокоразмерные системы, но и уменьшить число необходимых измерений без потери точности реконструкции, что особенно важно для построения цифровых двойников в реальном времени.

# **Ключевые вклады**

Основные научные вклады, представленные в статье, можно охарактеризовать следующим образом:

1. **Разработка нового метода тензорного модального разложения (TBMD):** Авторами предложен метод, который позволяет эффективно извлекать низкоразмерные моды из высокоразмерных динамических систем при сохранении пространственно-временной структуры данных. В основе TBMD лежит модель Такера для тензорного разложения, обеспечивающая более точное и компактное представление исходных данных, что минимизирует потерю информации при уменьшении размерности.
2. **Метод тензорной QR-факторизации для оптимального размещения датчиков:** В статье предложен новый метод тензорной QR-факторизации с выбором трубок (tube-pivot), который учитывает пространственные связи между элементами данных. Это позволяет оптимально размещать датчики в наиболее информативных точках, что значительно уменьшает количество необходимых измерений без потери качества реконструкции.
3. **Построение тензорной модели сжимающего измерения:** Авторами представлен метод сжимающего измерения на основе тензорного представления данных, который решает задачу восстановления полной информации о системе при ограниченном количестве измерений. Для решения этой задачи используется метод множителей переменных направлений (ADMM), что обеспечивает эффективную реконструкцию динамических систем с высокой точностью.
4. **Валидация методов на нескольких наборах данных:** Предложенные методы были апробированы на пяти различных наборах данных, включая задачи в области динамики жидкостей, аэродинамики и климатологии. Экспериментальные результаты демонстрируют превосходство TBMD по точности и вычислительной эффективности по сравнению с традиционными методами, а также его способность уменьшить количество необходимых датчиков для точной реконструкции систем.

# **Методология**

Предложенная методология основана на интеграции нескольких ключевых компонентов, которые обеспечивают высокую эффективность анализа динамических систем:

1. Тензорное модальное разложение (Tensor-Based Modal Decomposition, TBMD):
   * Модель Такера: TBMD использует модель тензорного разложения по Такеру, в которой многомерные данные представляются как произведение ядра тензора и матриц мод по каждому измерению. Это позволяет сохранить пространственно-временную структуру данных и значительно уменьшить размерность системы, сохраняя при этом ключевую информацию.
   * Извлечение мод низкого порядка: Для выделения основных динамических характеристик системы используется метод высшего порядка сингулярного разложения (Higher-Order Singular Value Decomposition, HOSVD). Это разложение позволяет извлечь низкоразмерные моды, которые в наиболее полной мере отражают динамическое поведение системы при минимальной потере информации.
2. Тензорная QR-факторизация для оптимального размещения датчиков:
   * Выбор трубок (tube-pivot): В процессе QR-факторизации используется выбор наиболее энергетически значимых трубок тензора. Этот подход позволяет определить оптимальные позиции для размещения датчиков, обеспечивая сбор наиболее информативных измерений при минимальном количестве сенсоров.
   * Преобразование Хаусхолдера: Алгоритм QR-факторизации основан на итеративном применении преобразований Хаусхолдера, адаптированных для тензорных данных. Это позволяет добиться ортогональности преобразований и стабильности вычислений, что особенно важно для обработки больших массивов данных.
3. Модель сжимающего измерения на основе тензоров:
   * Задача реконструкции данных: В условиях ограниченного количества измерений задача восстановления данных формулируется как задача оптимизации, где минимизируется ошибка между наблюдаемыми и восстановленными данными. Такой подход позволяет восстанавливать полную информацию о системе с высокой точностью при использовании минимального числа датчиков.
   * Решение с помощью метода ADMM: Для решения задачи оптимизации применяется метод множителей переменных направлений (Alternating Direction Method of Multipliers, ADMM), который обеспечивает быструю и эффективную сходимость к оптимальному решению. ADMM позволяет успешно решить задачу сжимающего измерения, восстанавливая полную информацию о системе из разреженных данных.

# **Альтернативные методы модального разложения**

В области построения моделей пониженного порядка и анализа сложных динамических систем широкое распространение получили методы Разложения по Собственным Функциям (Proper Orthogonal Decomposition, POD) и Динамического Разложения по Модам (Dynamic Mode Decomposition, DMD). Эти методы являются базовыми подходами для модального разложения и находят широкое применение в различных научных и инженерных задачах. В данном разделе будут рассмотрены математические основы этих методов, их ключевые преимущества и ограничения, а также проведено их сравнение с предложенным в статье методом Tensor-Based Modal Decomposition (TBMD).

## **Разложение по Собственным Функциям (Proper Orthogonal Decomposition, POD)**

Разложение по собственным функциям (POD) является одним из наиболее распространенных методов уменьшения размерности данных, позволяющим выделить энергетически значимые режимы системы. Этот метод основывается на представлении сложного процесса через суперпозицию ортогональных базисных функций (мод), которые упорядочены по степени их вклада в общую энергию системы.

Предположим, что имеется набор временных снимков системы , где – пространственные координаты, а – моменты времени. Основная цель POD заключается в нахождении набора ортогональных функций и соответствующих временных коэффициентов , таких что:

где – число сохраняемых мод (обычно , где N — общее количество временных снимков).

Для нахождения мод выполняется сингулярное разложение (SVD) матрицы данных либо решается задача собственных значений для ковариационной матрицы:

Собственные функции являются собственными векторами матрицы , а собственные значения отражают энергетический вклад каждой моды.

Применение POD:

* Гидродинамика и аэродинамика: Анализ турбулентных потоков, выделение крупных вихревых структур.
* Механика конструкций: Модальный анализ колебаний конструкций и систем.
* Обработка изображений и сжатие данных: Уменьшение размерности изображений, фильтрация шумов и извлечение ключевых признаков.

Преимущества:

* Эффективное сжатие данных: POD позволяет свести сложную систему к меньшему числу параметров, сохраняя при этом ключевую информацию.
* Выделение основных структур: Метод сосредоточен на наиболее энергетически значимых режимах, что важно для упрощения анализа.
* Простота реализации: Основан на применении сингулярного разложения (SVD), которое является хорошо изученной и проверенной математической техникой.

Ограничения:

* Отсутствие явной временной информации в модах: Временная динамика системы содержится только в коэффициентах aj(t)a\_j(t)aj​(t), что ограничивает понимание временной эволюции мод.
* Чувствительность к энергетически доминирующим режимам: Мелкомасштабные или малоэнергетические явления могут быть проигнорированы, что может привести к потере важных деталей системы.
* Статичность мод: POD может оказаться неэффективным для систем с сильно изменяющейся динамикой, так как выделенные моды остаются фиксированными и не отражают временные вариации системы.

## **Динамическое Разложение по Модам (Dynamic Mode Decomposition, DMD)**

Динамическое разложение по модам (DMD) – это мощный метод для анализа динамических систем, позволяющий выделять пространственно-временные структуры, которые характеризуются конкретными частотами и темпами роста или затухания. В основе метода лежит предположение, что динамика системы может быть аппроксимирована линейным оператором, который действует на последовательные временные снимки системы.

Предположим, что у нас имеется последовательность временных снимков системы Предполагается, что существует линейный оператор , который связывает два последовательных состояния:

Основная задача DMD заключается в нахождении оператора и его спектральных свойств. Для этого формируются матрицы данных:

Затем решается задача минимизации разности X' - AX в смысле наименьших квадратов. Для вычисления оператора используется сингулярное разложение (SVD) матрицы , что позволяет получить приближенную матриц меньшей размерности. Собственные значения и собственные векторы матрицы дают динамические моды DMD, а также соответствующие им частоты и темпы роста или затухания.

Применение DMD:

* Анализ нестационарных течений: Выявление колебательных и нестабильных структур в турбулентных и переходных потоках.
* Обработка видео: Разделение движущихся объектов и статического фона в видеопоследовательностях.
* Системы электроснабжения, нейробиология, финансы: Анализ временных рядов для выявления характерных динамических паттернов, что может быть полезным в прогнозировании и обнаружении аномалий.

Преимущества:

* Учет временной динамики: DMD напрямую анализирует временную динамику системы, предоставляя информацию о частотах и темпах роста или затухания динамических мод.
* Выделение транзиентных процессов: Метод особенно эффективен для систем с ярко выраженными временными изменениями, такими как переходные процессы или нестабильные режимы.
* Модельно-независимый подход: DMD не требует явных знаний о физических законах, управляющих системой, что делает его универсальным методом для анализа различных динамических процессов.

Ограничения:

* Предположение о линейности: Основной недостаток DMD заключается в том, что он основан на линейном приближении между последовательными снимками, что может быть неточным для сильно нелинейных систем.
* Чувствительность к шуму и параметрам: Метод чувствителен к выбору временных шагов и может быть подвержен влиянию шумов в данных, что требует тщательной предварительной обработки и настройки.
* Высокие вычислительные затраты: Для обработки больших наборов данных, особенно при высоком разрешении временных снимков, могут потребоваться значительные вычислительные ресурсы.

# **Сравнение POD, DMD и TBMD**

Основные различия и сходства:

* Учет времени:
  + POD: Временная динамика системы представлена через временные коэффициенты , что означает, что временная информация не встроена непосредственно в моды, а лишь сопутствует им.
  + DMD: В DMD временные характеристики встроены в моды, каждая из которых ассоциируется с определенной частотой и темпом роста или затухания, что позволяет непосредственно учитывать динамику системы во времени.
  + TBMD: Метод TBMD сохраняет пространственно-временную структуру данных, работая с тензорами, что позволяет одновременно анализировать как временные, так и пространственные моды, не теряя информации о корреляциях между ними.
* Структура данных:
  + POD и DMD: Оба метода работают с матрицами, что может привести к потере многомерной структуры данных при преобразовании тензоров в матричную форму. Это ограничение особенно проявляется при анализе высокоразмерных систем, где пространственные и временные корреляции играют важную роль.
  + TBMD: TBMD работает напрямую с тензорными представлениями данных, что позволяет сохранить многомерную структуру и учитывать корреляции между различными измерениями (пространственные, временные и другие).
* Вычислительная эффективность:
  + POD и DMD: Несмотря на их широкое использование, эти методы могут оказаться вычислительно затратными при работе с большими наборами данных, поскольку требуют выполнения матричных операций высокого порядка, таких как сингулярное разложение (SVD).
  + TBMD: TBMD использует тензорные разложения, что делает его более подходящим для анализа высокоразмерных данных. Однако вычислительная сложность метода также может возрастать при работе с очень большими системами, требуя дополнительных оптимизаций для повышения производительности.

Преимущества TBMD по сравнению с POD и DMD:

* Сохранение структуры данных: TBMD сохраняет пространственно-временные корреляции в тензорных данных, что позволяет более точно представлять динамические процессы в системе, не теряя важную информацию при переходе к матричным операциям.
* Эффективное сжатие: Тензорное разложение позволяет более компактно представлять данные, сводя многомерные данные к низкоразмерным представлениям без значительных потерь информации.
* Оптимальное размещение датчиков: Важное преимущество TBMD заключается в использовании тензорной QR-факторизации, которая позволяет решать задачу оптимального разреженного размещения датчиков, что особенно важно для систем с ограниченными ресурсами или требующих высокой точности мониторинга.

Ограничения TBMD:

* Сложность реализации: Тензорные методы, такие как TBMD, могут быть более сложными в реализации и понимании по сравнению с матричными методами (POD и DMD). Это требует глубоких знаний в области тензорной алгебры и специализированных вычислительных инструментов.
* Вычислительные требования: Хотя TBMD эффективен для высокоразмерных данных, его вычислительная сложность может стать существенной проблемой для очень больших систем, что требует разработки ускоренных алгоритмов и применения параллельных вычислений для повышения производительности.

# **Результаты работы TBMD алгоритма**

Авторами были проведены эксперименты на пяти различных наборах данных, что позволило продемонстрировать эффективность предложенных методов в разнообразных областях:

1. **Цилиндрическое течение:** Применение TBMD к задаче обтекания цилиндра потоком жидкости показало, что метод способен точно реконструировать ключевые характеристики потока, используя меньшее количество мод и датчиков по сравнению с традиционными методами, такими как POD и DMD.
2. **Аэродинамическое течение вокруг профиля NACA 0012:** В задаче моделирования обтекания аэродинамического профиля TBMD эффективно извлекает аэродинамические моды, соответствующие различным условиям обтекания. Это позволяет точно восстановить поля давления и скорости, используя ограниченное количество датчиков.
3. **Температура поверхности моря (Sea Surface Temperature, SST):** TBMD был успешно применен для анализа климатических данных, демонстрируя высокую точность при реконструкции температурных полей. Важным аспектом является использование меньшего числа датчиков при сохранении точности моделирования.
4. **Набор лиц Yale B:** В задачах обработки изображений TBMD показал превосходные результаты при сжатии и восстановлении изображений лиц при различных условиях освещения. По сравнению с традиционными методами, TBMD обеспечил более высокую точность и стабильную сходимость.
5. **Турбулентное течение в канале:** В моделировании сложных турбулентных течений TBMD продемонстрировал способность сохранять высокую точность при значительном снижении вычислительной нагрузки, что подтверждает его применимость для сложных задач моделирования потоков.

Основные выводы, сделанные по итогам экспериментов, включают:

* **Точность реконструкции:** TBMD обеспечивает более высокую точность восстановления данных по сравнению с POD и DMD, особенно при использовании малого числа мод. Это подтверждается как в задачах моделирования течений, так и при обработке изображений.
* **Сокращение количества датчиков:** Разработанный метод тензорной QR-факторизации позволяет значительно уменьшить количество датчиков, необходимых для точной реконструкции данных. Это является важным результатом для задач, где количество измерений ограничено.
* **Вычислительная эффективность:** TBMD демонстрирует высокую вычислительную эффективность за счет использования тензорных операций и оптимизационных алгоритмов. Предложенные методы обладают быстрой сходимостью и меньшими вычислительными затратами по сравнению с традиционными подходами.

# **Сильные и слабые стороны работы**

Сильные стороны:

* Сохранение структуры данных: Метод TBMD учитывает пространственно-временную структуру данных, что позволяет моделировать сложные динамические системы с более высокой точностью. Это особенно важно при анализе высокоразмерных данных, где традиционные методы могут терять значимые связи между элементами.
* Эффективное размещение датчиков: Тензорная QR-факторизация обеспечивает оптимальное размещение датчиков, что позволяет значительно уменьшить количество необходимых измерений. Это преимущество становится критически важным в системах с ограниченными вычислительными и сенсорными ресурсами.
* Универсальность применения: Разработанные методы демонстрируют широкую применимость в различных областях, включая аэродинамику, климатологию, обработку изображений и моделирование сложных физических процессов. Это делает их универсальными инструментами для решения разнообразных задач.
* Подтверждение на реальных данных: Проведенные эксперименты на пяти различных наборах данных подтверждают эффективность предложенных методов и их практическую значимость. Это свидетельствует о надежности и адаптивности подхода в реальных условиях.

Слабые стороны:

* Вычислительная сложность: Несмотря на явные улучшения по сравнению с традиционными методами, тензорные методы, такие как TBMD, могут оставаться вычислительно затратными для очень больших систем. Это ограничение может препятствовать их применению в реальном времени или в задачах с ограниченными вычислительными ресурсами.
* Чувствительность к параметрам: Эффективность предложенных методов может сильно зависеть от правильного выбора гиперпараметров, таких как ранги разложения и количество используемых мод. Это требует дополнительной настройки, что может увеличивать трудоемкость их применения на практике.
* Ограниченная обработка шума: В статье не рассматривается влияние шума или неточностей в данных, что является важным аспектом для многих практических приложений. В условиях реальных измерений присутствие шума может ухудшить точность моделей, что требует разработки дополнительных методов для его учета и фильтрации.

# **Перспективы дальнейших исследований**

Предложенные в статье методы Tensor-Based Modal Decomposition (TBMD) и тензорной QR-факторизации представляют собой значительный шаг вперед в области моделирования пониженного порядка и оптимального размещения датчиков. Они демонстрируют высокую точность и эффективность в реконструкции сложных динамических систем с использованием ограниченного числа измерений. Это открывает перспективы для широкого применения методов в задачах создания цифровых двойников и мониторинга систем в режиме реального времени. TBMD, за счет сохранения пространственно-временной структуры данных, особенно актуален для задач, связанных с высокоразмерными данными, что делает его универсальным инструментом для различных прикладных областей.

Несмотря на достигнутые результаты, дальнейшие исследования могут быть направлены на следующие перспективные направления:

* Оптимизация вычислительной эффективности: Для расширения практического применения TBMD необходимо сосредоточиться на разработке ускоренных алгоритмов, а также использовании параллельных и распределенных вычислений. Это позволит применять методы в реальном времени и при обработке больших объемов данных.
* Учет шума и неточностей: Для улучшения применимости методов в реальных условиях требуется разработка алгоритмов, устойчивых к шуму и неточностям данных. Включение механизмов фильтрации или подавления шума станет важным шагом в адаптации TBMD для различных практических задач.
* Автоматический выбор гиперпараметров: Исследование автоматических методов настройки гиперпараметров разложения, таких как ранги и количество мод, позволит упростить применение методов и повысить их адаптивность в различных сценариях использования.
* Применение в новых областях: TBMD и связанные с ним методы могут найти широкое применение в таких областях, как биомедицина, финансы и другие, где анализ высокоразмерных данных играет ключевую роль. Это открывает новые возможности для междисциплинарных исследований и приложений.
* Интеграция с машинным обучением: Комбинирование тензорных методов с современными подходами машинного и глубокого обучения может значительно улучшить возможности моделирования и прогнозирования сложных систем, что расширит границы их применения в разнообразных научных и инженерных задачах.

Таким образом, TBMD представляет собой мощный инструмент для анализа динамических систем, и дальнейшие улучшения его производительности и адаптивности способны еще больше расширить его использование в различных отраслях науки и техники.

# **Заключение**

При выборе метода модального разложения для анализа динамических систем необходимо учитывать характеристики данных, цели исследования и ограничения каждого подхода:

* POD является подходящим выбором для задач, требующих уменьшения размерности данных и выделения основных энергетически значимых структур, без необходимости явного учета временной динамики.
* DMD эффективен для анализа систем с ярко выраженными динамическими процессами, так как он позволяет изучать временные характеристики мод и отслеживать частоты и темпы роста или затухания.
* TBMD объединяет преимущества обоих методов, работая напрямую с тензорными данными и сохраняя их пространственно-временную структуру. Это делает его мощным инструментом для анализа сложных высокоразмерных систем, где требуется высокая точность и сохранение многомерных корреляций.

В представленной статье метод TBMD демонстрирует превосходство над традиционными методами POD и DMD в точности реконструкции и вычислительной эффективности при работе с высокоразмерными тензорными данными. Тем не менее, выбор метода всегда должен основываться на специфике решаемой задачи, доступных ресурсах и требованиях к точности и скорости вычислений. Важно учитывать, что каждый из методов имеет свои сильные стороны и ограничения, которые необходимо сопоставлять с целями исследования и характеристиками системы.