Untitled

Юрченков Иван Александрович

13 01 2022

## Введение

Данная работа посвящена исследованию возможности построения алгоритма автоматического выявления паттернов медленного движения временного ряда (тренда) на основе линеаризуемых основных моделей динамики природных систем, реализующихся в виде дифференциальных уравнений. Среди таких моделей чаще всего употребимы в классических научных трудах: модели интенсивного роста на основе изменения темпов роста, модели ограниченного роста на основе лимитированной энергии развития системы, комплексные модели динамики систем.

Проблема изучения и выявления характеристик временного ряда данных является классической. В общей постановке, единообразного математического подхода исследования временных рядов с возможностью дальнейшего прогнозирования и верификации наблюдений нет и по сей день. Каждая динамическая система изучается в отдельности и к каждой динамической системе необходим уникальный подход. Однако, на текущий момент существует большое количество общих рекомендаций и решений, которые способствуют выработке правильных выводов о характере динамики и поведении системы как на уровне быстрых движений вокруг общей тенденции, так и на трендовом уровне. Системы таких рекомендаций и решений являются довольно мощными инструментами в вопросах исследования временных рядов и их свойств. Среди таких методов существуют следующие: алгоритм преобразования Фурье, STL-декомпозиция временного ряда, анализ распределений, структурная функция Колмогорова и функция Альтера-Джонсона, корреляционные и автокорреляционные функции.  
Проблематика данной работы была естественным образом поставлена в момент разработки алгоритма выявления почти-периодических компонент динамических систем на основе обобщенной сдвиговой функции Альтера-Джонсона и алгоритма выявления почти пропорций.

Практическая значимость работы состоит в обогащении коллекции алгоритмов выявления значимых характеристик временного ряда данных новым алгоритмом, позволяющим на основе определенных метрик выявить наиболее закономерные участки медленных движений временного ряда данных, отвечающих природным динамическим моделям развития систем.

## Теоретическая часть

### Временные ряды

Временной ряд - собранные в разные моменты времени значения показателей наблюдаемой системы или исследуемого процесса. Каждое измеренное значение системы называется измерением или отсчётом. Во временном ряде для каждого отсчёта должно быть указано время измерения и (или) номер измерения по порядку следования.

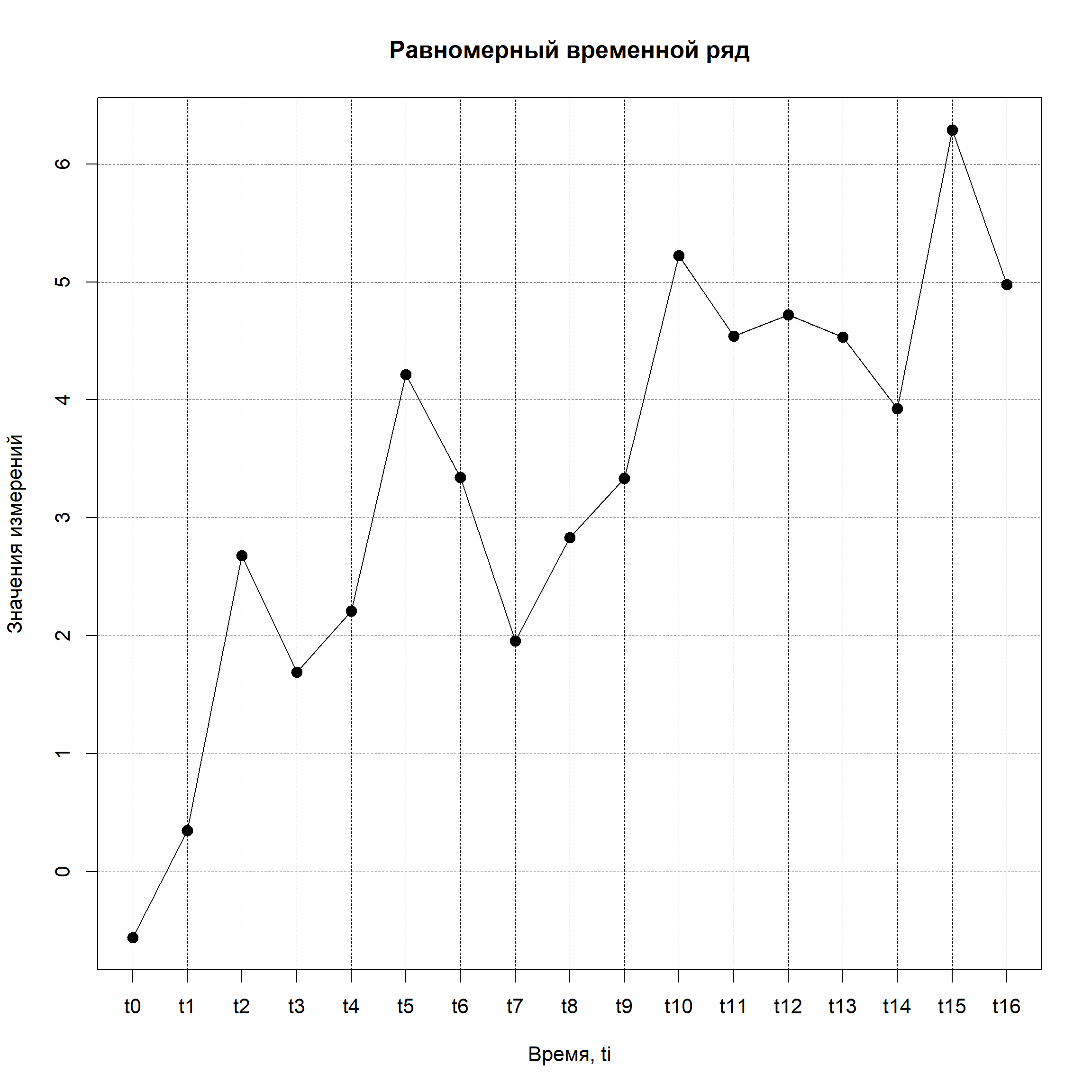
Временной ряд является более сложным понятием в науке о данных по сравнению с понятием *статистической выборки*. Если в статистике статистическая выборка это набор независимых случайно выбранных наблюдений, характеризующих систему в целом, то в понятии *временной ряд* фигурируют такие понятия как зависимость наблюдений друг от друга, зависимость наблюдений от времени и следования. Все эти факторы необходимо учитывать при построении моделей или при анализе данных такого рода.

Временные ряды в простой классификации бывают *равномерные* и *неравномерные* и различаются характером снятия измерений с динамической системы.

### Равномерный временной ряд

Равномерный временной ряд - ряд данных, измерения или отсчёты в котором производятся через равные интервалы времени, т.е. записываются в таблицу с определенной заранее заданной частотой. Система математически записывается следующим образом:

где - начальный момент времени или время зарождения процесса; - период времени, через который производятся измерения числовых показателей изучаемой системы; - текущий отсчет, соответствующий настоящему моменту времени или отсчет, соответствующий моменту времени завершения процесса в прошлом; - количество снимаемых динамических показателей из системы ; - динамический показатель системы, измеряемый во времени.



Таким образом определяется равномерный временной ряд наблюдаемой системы .

Зачастую равномерные временные ряды это ряды с относительно небольшим возрастом, так как исторические данные в силу технического развития не могли быть сняты через одинаковые довольно короткие промежутки времени. В настоящий момент можно встретить временные ряды с частотой дискретизации наблюдений порядка десяток тысяч наблюдений в секунду.

### Неравномерный временной ряд

### Тренд

### Сезонность

### Ошибки

### STL-разложение

### Модели временных рядов

### Модели интенсивного роста

### Модели ограниченного роста

### Анализ сезонности на основе функции Альтера-Джонсона

### Обобщенная сдвиговая функция

### Анализ почти-пропорций

### Автокорреляционная фукнция

### Линейная регрессия

### Оценочная функция Тейла-Сена

### Коэффициент детерминации

### Распределение остатков

### Быстрое преобразование Хафа

## Основная часть

В основе исследования трендовой зависимости исследуемой динамической системы лежит возможность установления естественных зависимостей в виде моделей дифференциальных уравнений, поскольку именно модели дифференциальных уравнений являются фундаментальными природными моделями развития систем. В связи с этими соображениями появилась идея разработки алгоритма обобщенного исследования временного ряда динамической системы на наличие линеризуемых дифференциальных моделей таких как модели интенсивного и ограниченного роста.

В основе разрабатываемого алгоритма лежит идея линеаризуемости некоторых видов одномерны дифференциальных моделей развития системы или показателя во времени. Такого рода линеаризация производится за счет функционального преобразования координат, которая в линеаризуемых участках ряда идентифицирует одномерную дифференциальную модель.

Существует ряд проблем, которые всё же являются непреодолимыми при разработке такого рода “автоматических” алгоритмов:

1. Невозможно для абсолютно всех классов динамических систем построить общую стройную теорию автоматического исследования характеристик ряда ввиду сложности природных явлений, которые ключевым образом влияют на поведение физических и даже экономических систем.
2. Невозможно адекватно и за приемлемое время работать с длинными и достаточно плотными временными рядами ввиду шума измерительных приборов, числа операций для обработки и т.д.
3. Сложность разработки быстрых эффективных алгоритмов.

### Наивный алгоритм или обобщенный алгоритм выявления трендовых участков

Общий алгоритм исследования временного ряда с использованием обобщенного алгоритма характеризации тренда:

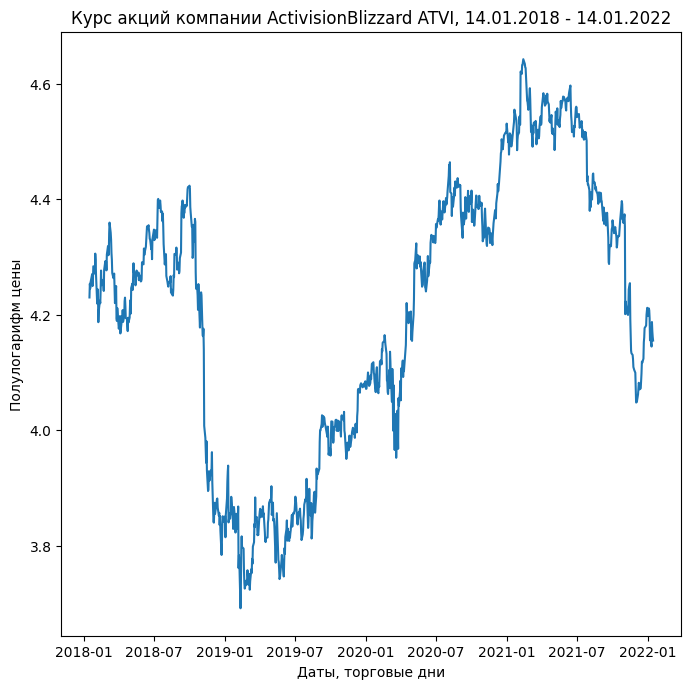
1. Полученный временной ряд предварительно очистить от шума.
2. Изучить быстрые движения относительно тренда на наличие тонких колебательных зависимостей, являющихся сезонными, во избежание нежелательного искажения действительности в трендовых характеристиках.
3. Идентификация почти-периодических зависимостей на основе обобщенной сдвиговой функции Альтера-Джонсона или структурной функции Колмогорова.
4. Выделение тренда на основе сглаживания временного ряда по идентифицированному почти-периодическому значению.
5. Применение алгоритма обобщенного исследования характеристики тренда.
6. Получение метрики уверенности в наличии дифференциальной модели на конкретных участках временного ряда исследуемой динамической системы.

Алгоритм обобщенного исследования характеристики тренда:

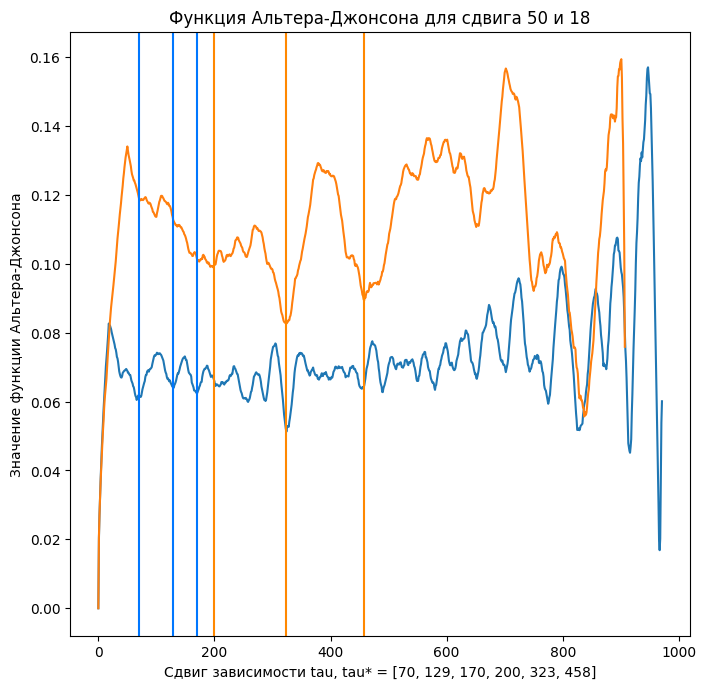
Метрика по которой мы можем судить о достаточном спрямлении данных в аллометрических координатах - *коэффициент детерминации*. Таким образом

$$

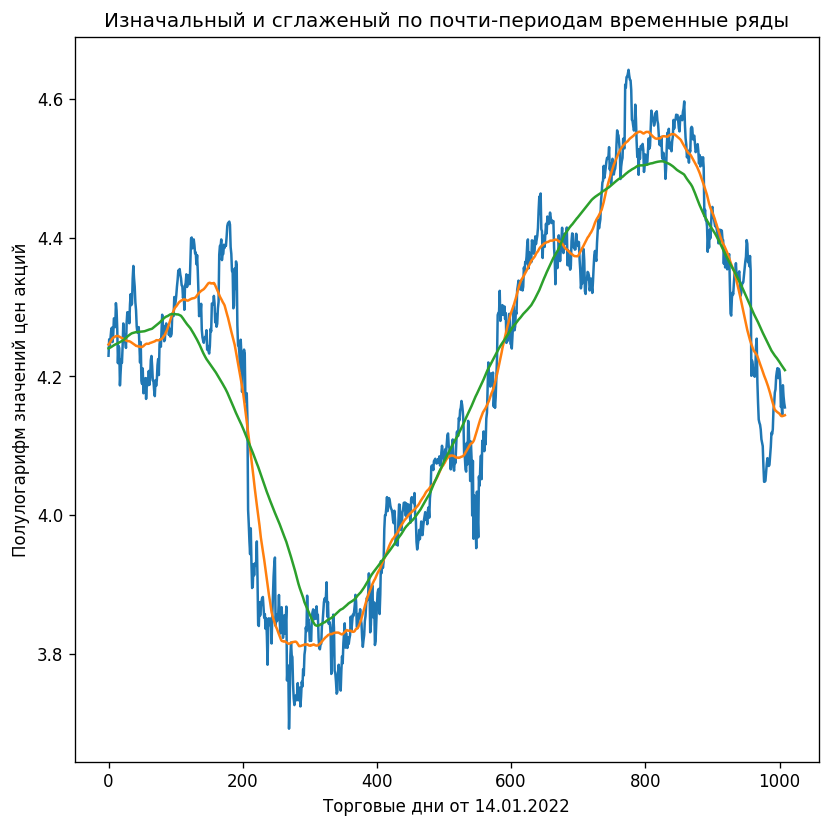
$$



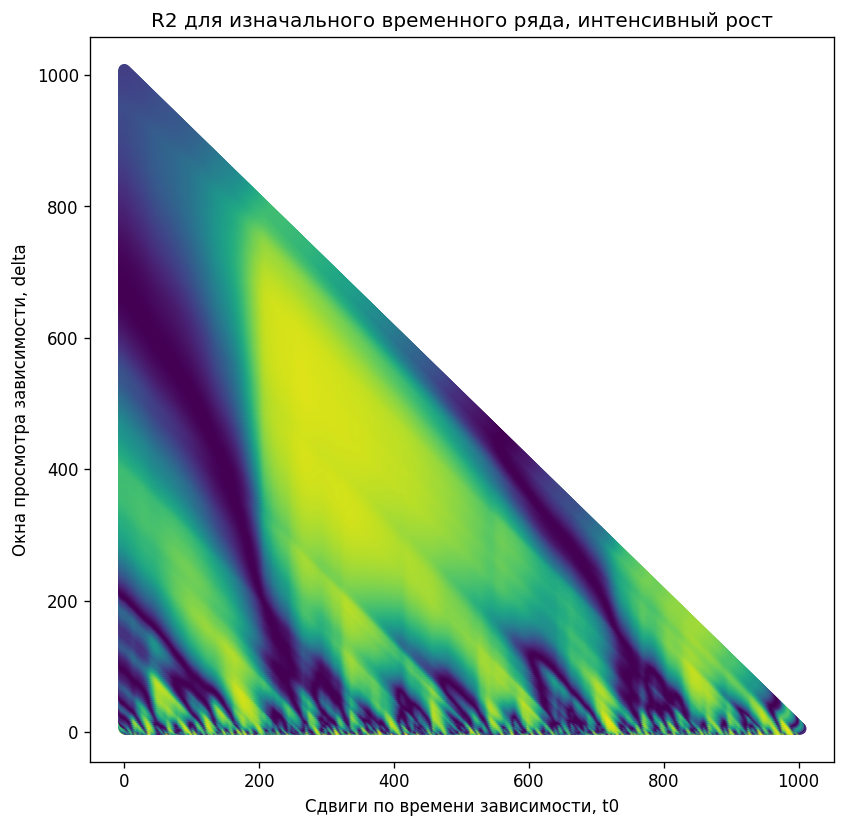
pic1



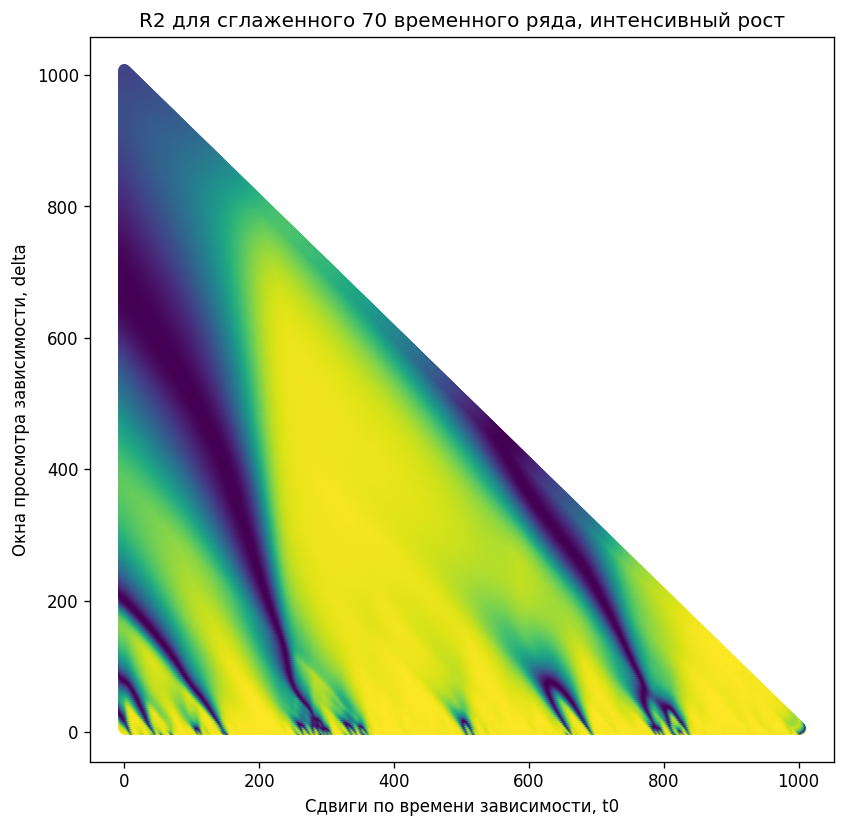
pic2



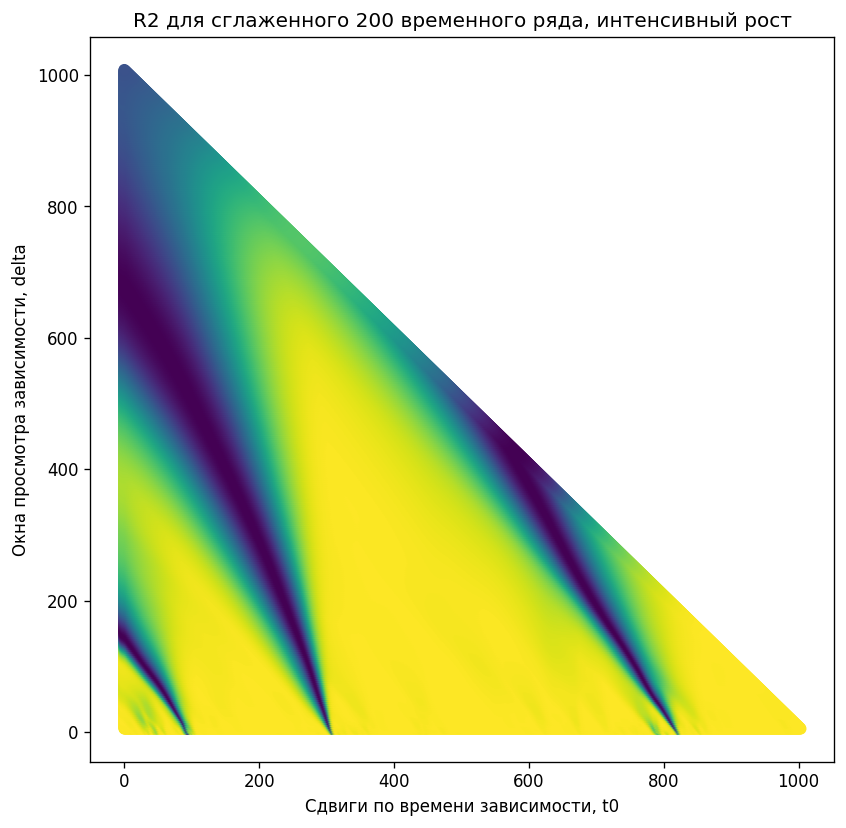
pic3



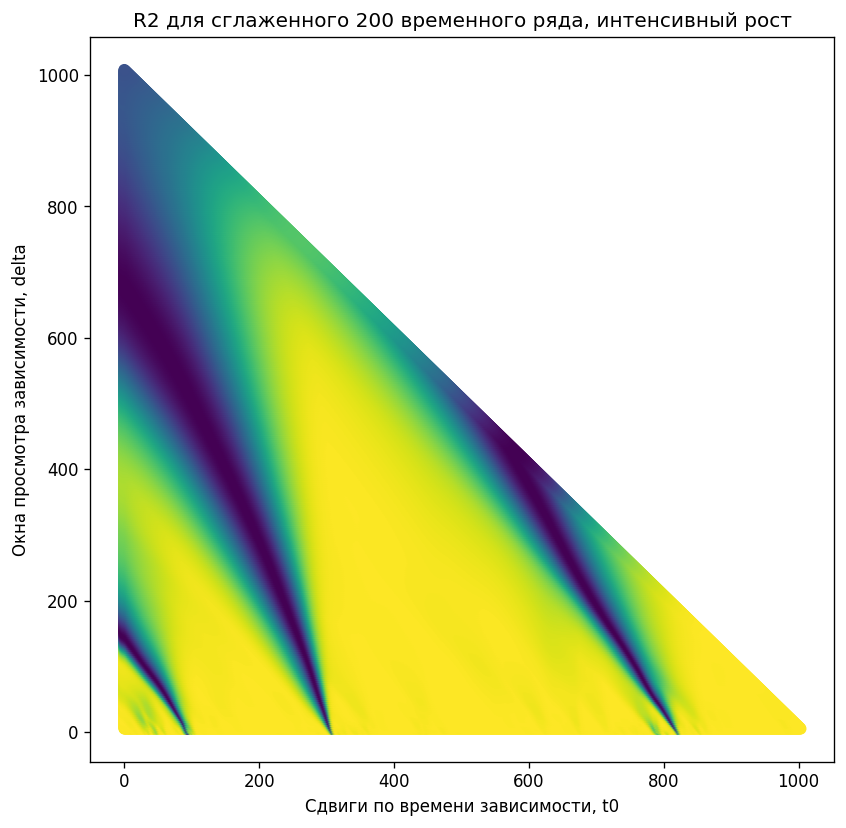
pic4



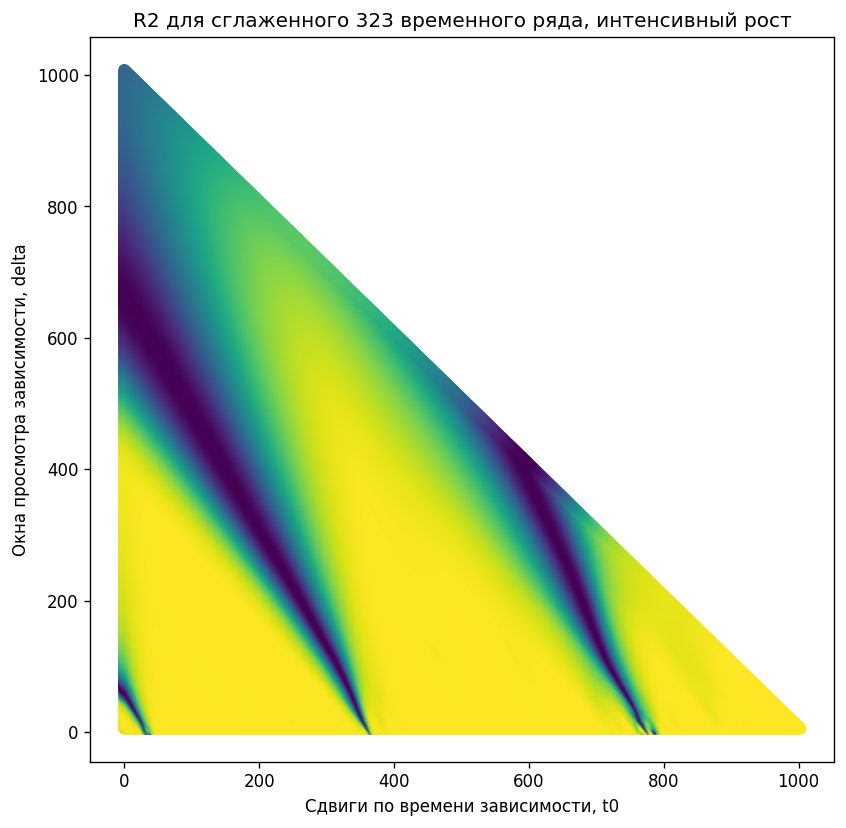
pic5



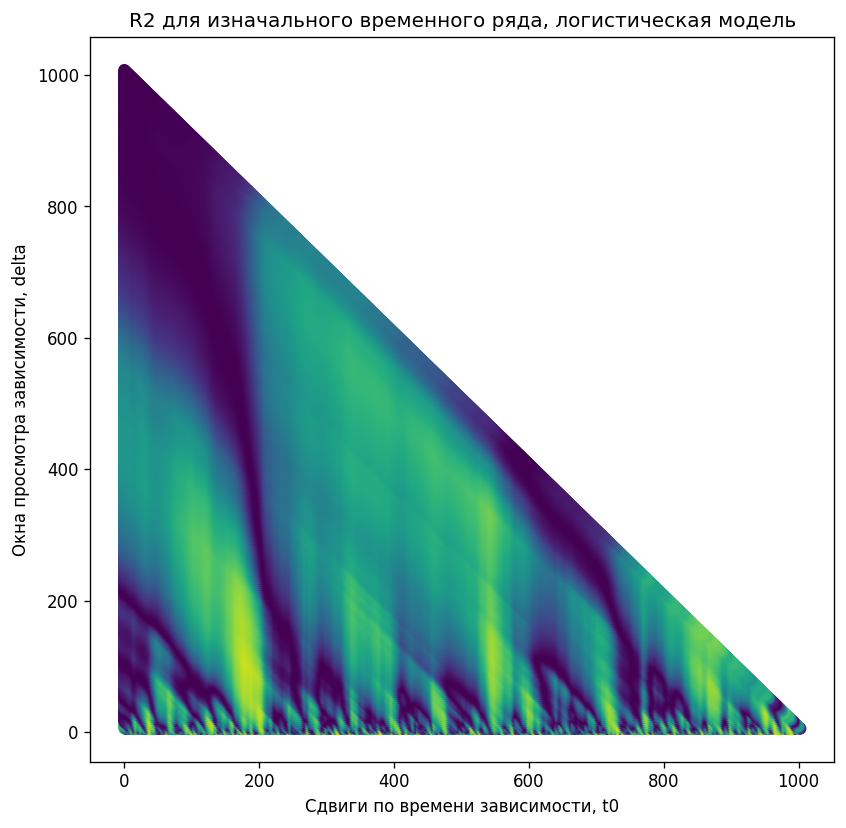
pic6



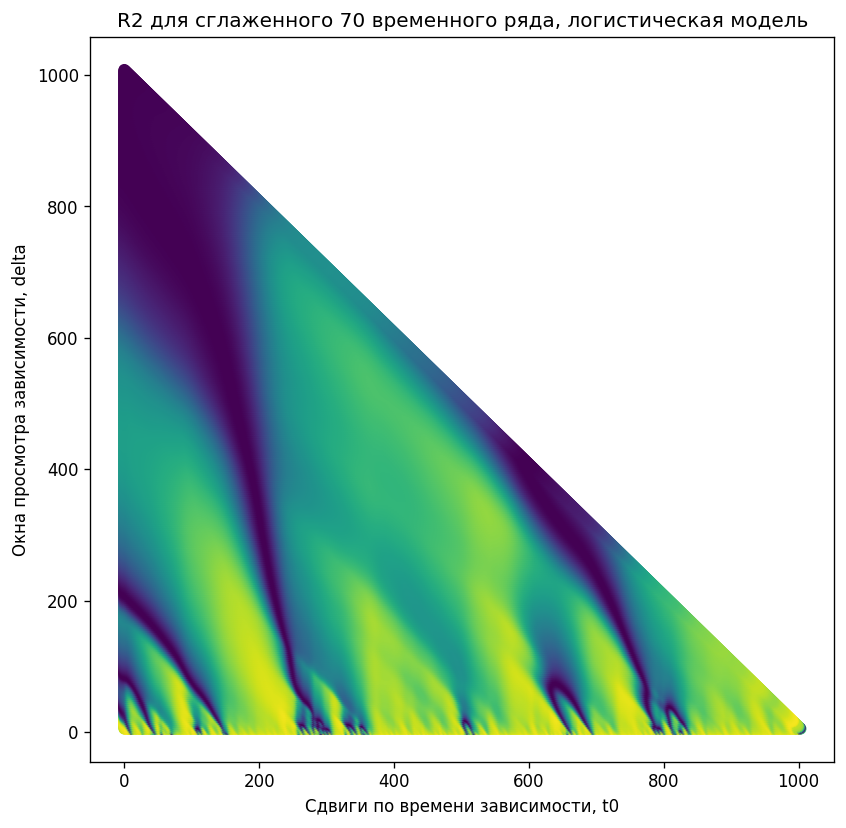
pic7



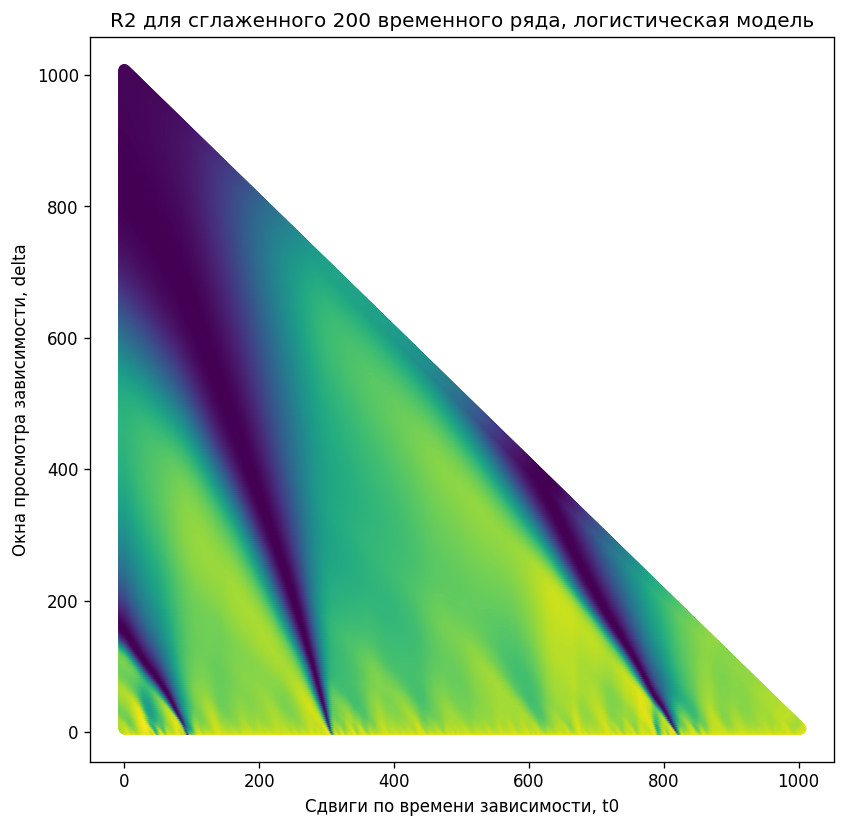
pic8



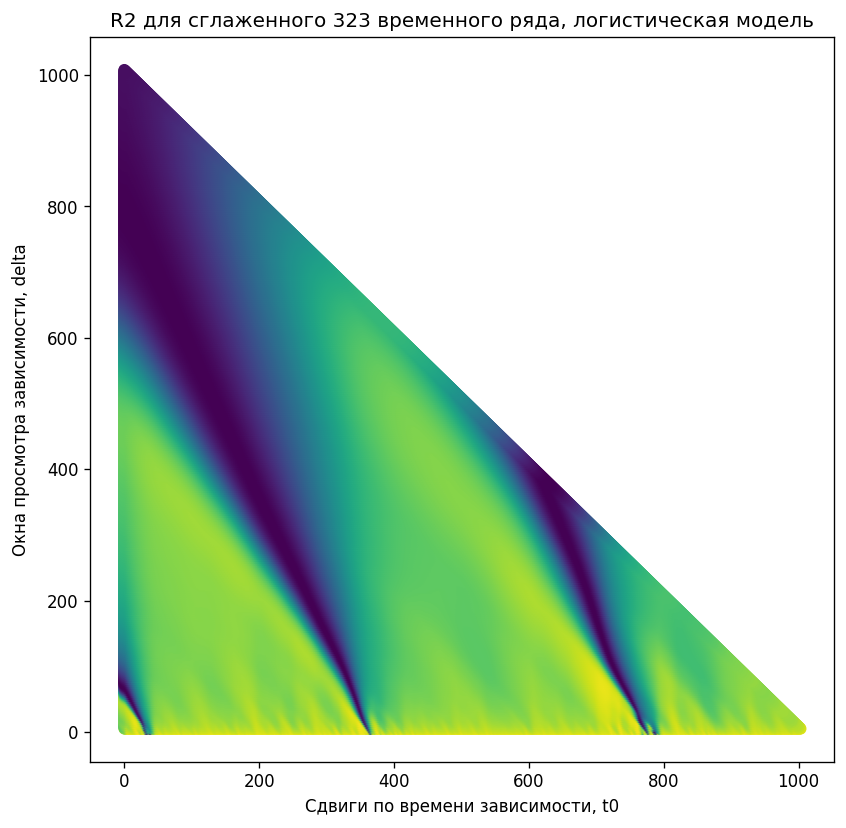
pic9



pic10



pic11



pic12

### Эффективный алгоритм выявления трендовых участков

## Заключение

## Список использованной литературы