Министерство науки и высшего образования РФ

ФГБОУ ВО «Удмуртский государственный университет»

Институт математики, информационных технологий и физики

Кафедра математического анализа

КУРСОВАЯ РАБОТА

НА ТЕМУ «Анализ временных рядов»

Выполнил студент:

Широков Егор Александрович

направление подготовки

Прикладная математика и информатика

группы ОБ-01.03.02.02-31

Научный руководитель:

Латыпова Наталья Владимировна канд. физ.-мат. наук, доцент;

Итоговая оценка по курсовой работе

# Оглавление

Оглавление

[Оглавление 2](#_Toc167736186)

[Введение 3](#_Toc167736187)

[Глава 1. Временные ряды 4](#_Toc167736188)

[1.1 Основные понятия 4](#_Toc167736189)

[1.2 Стационарность временных рядов. 6](#_Toc167736190)

[1.3 Почему ряд должен быть стационарен. 7](#_Toc167736191)

[1.4 Способы проверки ряда на стационарность 7](#_Toc167736192)

[1.5 Автокорреляция 8](#_Toc167736193)

[Глава 2. Методы прогнозирование временных рядов. 11](#_Toc167736194)

[Глава 3. Решение задач. 14](#_Toc167736195)

[3.1 Прогнозирование потребления электроэнергии. 14](#_Toc167736196)

[3.2 Прогнозирование стоимости акции Сбербанка 18](#_Toc167736197)

[Заключение. 21](#_Toc167736198)

[Список литературы 22](#_Toc167736199)

# Введение

Анализ временных рядов является важной областью статистики и анализа данных, которая нашла широкое применение в различных областях, начиная от экономики и финансов до климатологии и медицины. В современном мире оценка и прогнозирование временных рядов является неотъемлемой частью принятия обоснованных решений в различных сферах деятельности.

Целью курсовой работы ставлю такие вещи как:

1.Дать теоретическое описание анализу временных рядов

2.Дать основные определения встречающееся при анализе временных рядов

3.В качестве практической части курсовой работы разобрать на компоненты и спрогнозировать временной ряд.

1-я глава посвящена обзору тому, что такое временной ряд, а также даны основные определения, которые встречаются при анализе временных рядов.

Во 2-й главе был дан обзор основным методам прогнозирования стационарных и нестационарных рядов.

В 3-й главе были разобраны и спрогнозированы два временных ряда: ряд потребления электроэнергии и ряд стоимости акции “Сбербанка” соответственно.

# Глава 1. Временные ряды

## 1.1 Основные понятия

**Временной ряд** – это ряд наблюдений анализируемой случайной величины , произведенных в последовательные моменты времени , ,… .Всякий временной ряд включает в себя два обязательных элемента: время и конкретное значение показателя или уровень ряда.

Пусть вероятностное пространство, на котором задан стационарный процесснаблюдаемый в равноотстоящие моменты времени

Где ряд наблюдений случайной функции неслучайная (долговременная) функция тренда; неслучайная (сезонная) периодическая функция; неслучайная (долговременная, циклическая) функция; нерегулярная компонента (случайная величина, ошибка). Такая модель называется аддитивная. Ее стоит использовать, когда величина изменяется линейно (например, у авиакомпании + 1000 пассажиров в год). Также часто используют другую модель мультипликативную:

Ее стоит использовать, когда величина изменяется нелинейно (например, количество пассажиров увеличивается каждый год в 2 раза).

Анализ временных рядов сводится к выделению регулярных компонент , , и описанию нерегулярной ее части

Как правило, наличие той или иной составляющей можно определить с помощью визуального анализа графика временного ряда.

Рассмотрим основные понятия, связанные с временными рядами.

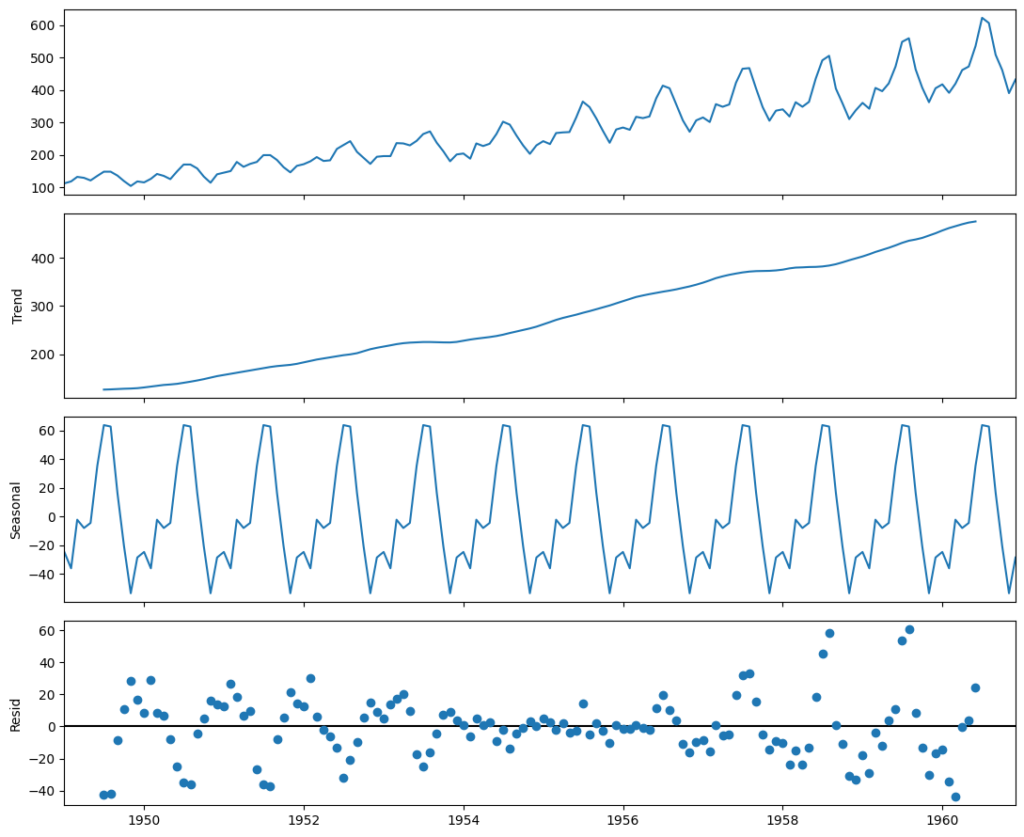


Рис.1. Разложения временного ряд по авиаперевозкам на компоненты.

**Тренд (trend)** — это основная тенденция изменения временного ряда. Тренд может быть направлен вверх (положительный тренд) или вниз (отрицательный тренд). Он отражает общий характер изменений и может быть вызван фундаментальными факторами, такими как экономические тенденции, демографические изменения или другие долгосрочные воздействия. Вид тренда может быть определен с помощью построения функциональной модели статистическими методами.

**Сезонность (seasonality)** отражает периодические колебания в данных. Эти колебания могут быть связаны с сезонными факторами, такими как времена года, месяцы или дни недели. Например, продажи мороженого могут иметь сезонность с пиком в летние месяцы и спадом зимой.

**Остатки (residuals)** представляют собой непрогнозируемую составляющую в данных. Это то, что остается после удаления тренда и сезонности из данных. В идеальном случае остатки выглядят как случайные колебания, то есть математическое ожидание и дисперсия будут постоянны и независимы от времени. Если это не так, значит, мы недостаточно точно выделили тренд или сезонность. Таким образом, мы можем валидировать качество нашей декомпозиции при анализе остатков. При хорошей декомпозиции эти остатки будут выглядеть как случайные колебания.

## 1.2 Стационарность временных рядов.

Одно из самых главных свойств, которым должен обладать, временной ряд — это стационарность. Если рассматривать временные ряды как случайные

процессы, тогда стационарностью будем определять следующую совокупность его свойств:

Математическое ожидание ряда не меняется во времени:

Ковариация смещённого на k отсчётов назад ряда не зависит от времени: ;

Временной ряд, обладающий данными свойствами, будем называть стационарным в широком смысле. Если ряд стационарен, то у него нет ни тренда, ни сезонности.

Ряд называется нестационарным, если в нём есть тренд, сезонность или комбинация этих составляющих. Часто рассматриваемые ряды имеют не стационарный вид, поэтому их приводят такому виду с помощью следующих методов:

1.Взятие разностей между двумя последовательными значениями наблюдений. Такие разности называются первыми. Если в ряду, представляющем первые разности, вновь найти последовательные разности, то получится ряд вторых разностей и так далее. Считается, что нахождение первой-второй разности — достаточная процедура для удаления трендов и приведения большинства временных рядов к стационарному виду.

2.Удаление из ряда тренда. На первом этапе определяют тип тренда. Эта процедура называется также заданием аналитического уравнения тренда.

## 1.3 Почему ряд должен быть стационарен.

Многие методы анализа временных рядов предполагают стационарность, является одной из ключевых причин, почему временные ряды должны быть стационарным. Например, широко используемая модель ARIMA для прогнозирования предполагает, что данные стационарны. Модель не сможет точно отразить лежащие в основе данных закономерности, если она нестационарная, и выводы будут ошибочными.

Это означает, что нестационарность может привести к неправильным или вводящим в заблуждение результатам, является еще одним обоснованием необходимости того, чтобы временные ряды были стационарными перед анализом. В качестве иллюстрации нестационарный временной ряд может создать впечатление, что между двумя переменными существует высокая связь, но на самом деле корреляция существует только из-за тренда или сезонного элемента в данных

1.4 Способы проверки ряда на стационарность.

1.**Тест Дики — Фуллера.**

Для того чтобы, убедиться в том, что после преобразований ряд стал стационарным, проводят тест Дики-Фуллера. Процесс построения теста Дики – Фуллера начинается с рассмотрения простейшего процесса авторегрессии первого порядка AR (1):

, где временной ряд, а ошибка.

Если ,то процесс имеет единичный корень, в этом случае ряд не стационарен, если , то ряд стационарен.

**2.Критерий Квятковский-Филлипс-Шмидт-Шин.**

Если рассматриваемый ряд имеет вид:

Где - некоторый стационарный процесс, некоторый независимый и одинаково распределенный спроцесс с математическим ожиданием 0 и дисперсией .

Выдвигаются две конкурирующие гипотезы: : временной ряд являются стационарным (или, аналогично ), временной ряд не являются стационарным ( . Вычислим статистическое значение ,где -размер выборки, ,оценка дисперсии случайной ошибки.

## 1.5 Автокорреляция

Нестационарные временные ряды характеризуются тем, что значения каждого последующего уровня временного ряда корреляционно зависят от предыдущих значений.

**Автокорреляцией** уровней временного ряда называется корреляционная зависимость между настоящими и прошлыми значениями уровней данного ряда.

**Лагом** называется величина сдвига между рядами наблюдений.

Лаг временного ряда определяет порядок коэффициента автокорреляции. Например, если уровни временного ряда и корреляционно зависимы, то величина временного лага равна единице. Следовательно, данная корреляционная зависимость определяется коэффициентом автокорреляции первого порядка между рядами наблюдений и *.* Если лаг между рядами наблюдений равен двум, то данная корреляционная зависимость определяется коэффициентом автокорреляции второго порядка и т. д.

При увеличении величины лага на единицу число пар значений, по которым рассчитывается коэффициент автокорреляции, уменьшается на единицу. Поэтому максимальный порядок коэффициента автокорреляции рекомендуется брать равным *n/4*, где *n* – количество уровней временного ряда. Автокорреляция между уровнями временного ряда оценивается с помощью выборочного коэффициента автокорреляции, который рассчитывается по формуле:, где ,среднее арифметическое произведение двух рядов наблюдения ,взятых с лагом , значение среднего уровня ряда , средние квадратическое отклонение.

Графиком автокорреляционной функции называется коррелограмм.

Анализ структуры временного ряда с помощью коэффициентов автокорреляции стоится на следующих правилах:

1) исследуемый временной ряд содержит только трендовую компоненту, если наибольшим является значение коэффициента автокорреляции первого порядка

2) исследуемый временной ряд содержит трендовую компоненту и колебания периодом если наибольшим является коэффициент автокорреляции порядка . Эти колебания могут быть как циклическими, так и сезонными;

3) если ни один из коэффициентов автокорреляции не окажется значимым, то делается один из двух возможных выводов:

а) данный временной ряд не содержит трендовой и циклической компонент, а его колебания вызваны воздействием случайной компоненты, т. е. ряд представляет собой модель случайного тренда;

б) данный временной ряд содержит сильную нелинейную тенденцию, для выявления которой необходимо провести его дополнительный анализ.

Графическим способом анализа структуры временного ряда является построение графиков автокорреляционной и частной автокорреляционной функций.

Если выраженный максимум коррелограммы оказывается для лага, то временной ряд содержит циклическую компоненту с периодом (Рис.2).

Изображение выглядит как линия, текст, диаграмма, снимок экрана

Автоматически созданное описание

Рис.2

Если коррелограмма имеет максимум при, то ряд содержит только тренд (Рис.3).

Изображение выглядит как линия, снимок экрана, График, диаграмма

Автоматически созданное описание

Рис.3

# Глава 2. Методы прогнозирование временных рядов.

**1.Наивное предсказание**

Модель временного ряда, в которой его текущее значение равно предыдущему наблюдаемому значению этого ряда. «Наивная» модель — самый примитивный метод прогнозирования.

Задается выражениемгдепоследнее наблюдаемое значение, прогнозируемое значение. Грубо говоря прогноз на завтра – это то, что мы наблюдаем сегодня

«Наивные» прогнозы полезны для создания базовой линии, т. к. применяются в качестве эталона, с которым можно сравнивать результаты, полученные более сложными моделями. Такой прогноз быстро реализуется без каких-либо затрат. Кроме того, «наивная» модель позволяет работать при отсутствии исторических данных.

Основным недостатком данного метода является низкая точность прогноза. Модель не учитывает закономерности самого процесса, сезонную составляющую и тренды.

**2. Модель AR ()**

Авторегрессионные модели (AR-модели) используют прошлые значения ряда для прогнозирования его будущих значений. Эта модель предполагает, что текущее значение ряда зависит от его предыдущих значений. Наиболее популярной AR-моделью является модель первого порядка (AR (1)), которая предполагает, что текущее значение ряда зависит только от его предыдущего значения. Имеет вид:

Где с— постоянная (часто для упрощения предполагается равной нулю), a белый шум.

**3. Метод скользящего среднего.MA ().**

Модель скользящего среднего является распространенным подходом для моделирования одномерных временных рядов. Согласно модели, оценка прогнозируемых членов ряда линейно зависит от текущего и прошлых значений, а также некоторого стохастического члена, который отражает вероятностный характер модели.

Метод скользящего скользящего среднего имеет вид:

,

где среднее значение ряда,параметры модели и белый шум.

Эта модель использует прошлые значения ошибок - разницу между фактическими значениями ряда и его прогнозируемыми значениями, для прогнозирования будущих значений. Модель q-ого порядка (MA(q)) предполагает, что текущая ошибка зависит только от ее q(смещение).

MA-модели схожи с AR-моделями, однако в отличие от них модель представляет собой не линейную комбинацию прошлых значений, а линейную комбинацию прошлых компонент белого шума. МA-модель “видит” случайные шоки белого шума непосредственно для каждого текущего значения серии. AR-модель же “видит” шоки косвенным образом, через репрессирование к своим предыдущим значениям) предыдущих значений.

4. **Модель ARMA ().**

Модель ARMA –представляет собой комбинацию AR и MA методов.

где константа, белый шум,

Проинтерпретировать модель можно следующим образом: текущее значение зависит от прошлых значений до лага и от текущего и прошлых внешних шоков до лага

5. **Модель ARIMA ().**

ARIMA позволяет моделировать данные, не являющиеся стационарными, как это требует для AR- и MA-модели. ARIMA включает три параметра: параметр авторегрессии (р), параметр скользящего среднего (q) и параметр интегрирования (d).

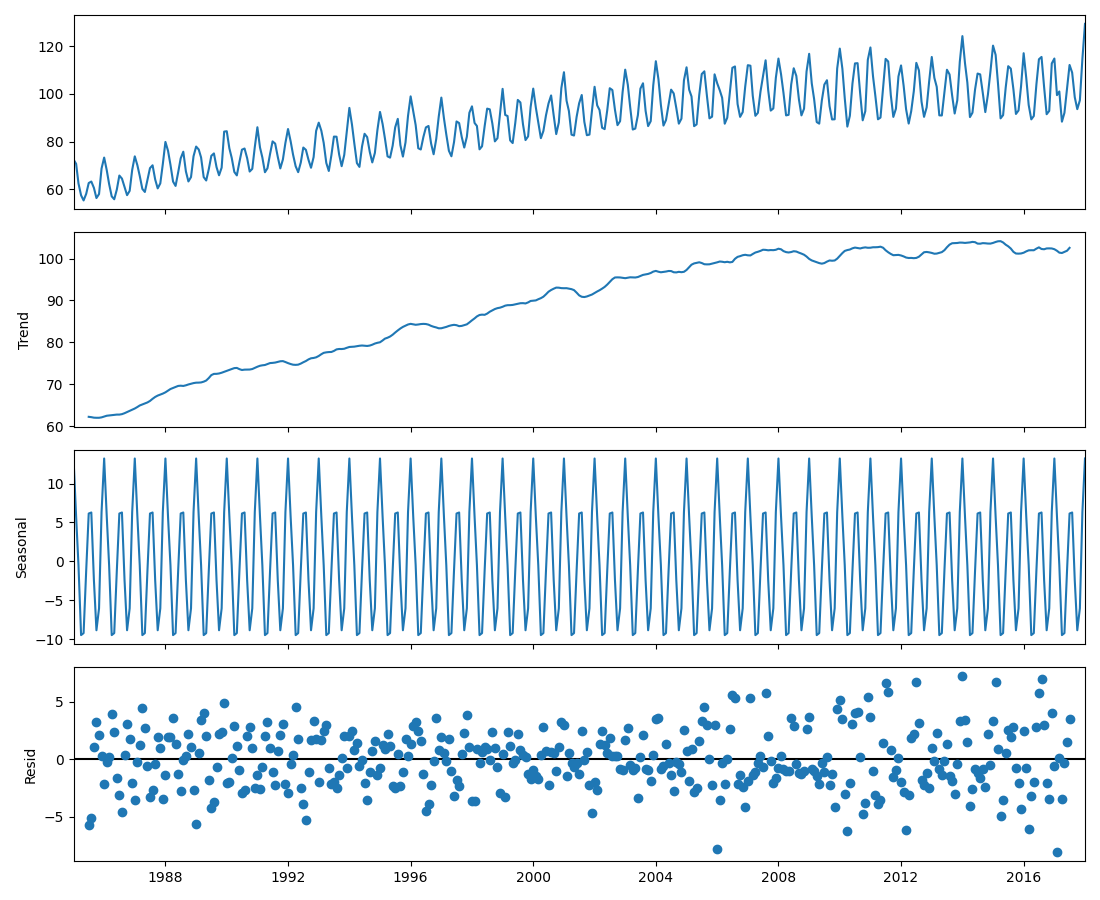
Можем привести ряд к стационарному виду с помощью оператора разности временного ряда порядка (производим последовательное взятие раз разностей первого порядка — сначала от временного ряда, затем от полученных разностей первого порядка, затем от второго порядка и т. д.)

# Глава 3. Решение задач.

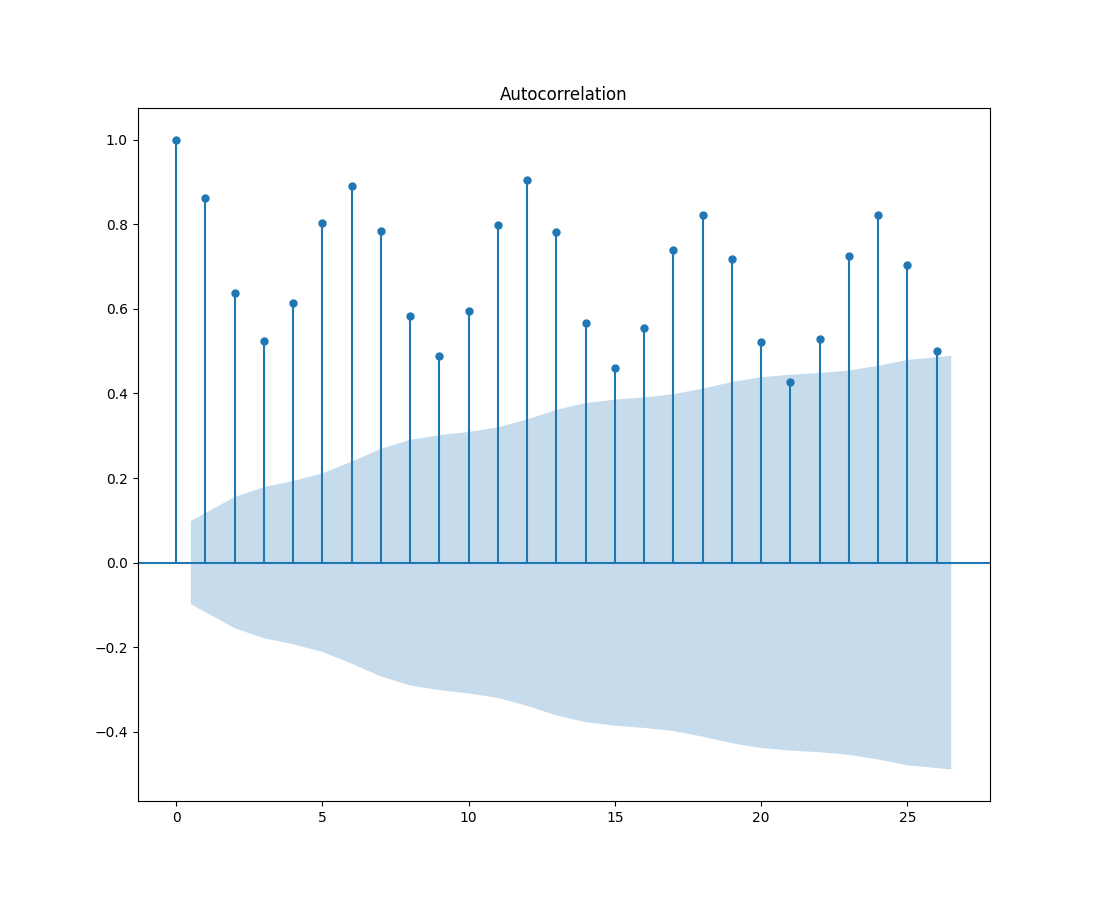
## 3.1 Прогнозирование потребления электроэнергии.

Для начала спрогнозируем простую задачу для анализа временных рядов. Решим данную задачу на ЭВМ с помощью языка программирования phyton. Используем данные наблюдения потребления энергии с 01.01.1985 года по 01.01.2018 года для прогнозирования. Изображение выглядит как текст, снимок экрана, Шрифт, линия

Автоматически созданное описание

Используя визуальный анализ можно подумать ,что ряд стационарен ,но проведенный тест Дики-Фулера говорит об обратном ,статестическое знаначение равно 16.89032 ,когда пороговое значение равно 0.05.Так что имеет смысл разбить ряд на компоненты .

Можно сразу выделить сильную сезоность и тренд на рост. Остатки выглядят как шум ,значит мы правильно выделили тренд и сезоность.

Посмотрим на график автокореляции.

Автокоряляция также выделяет сезоноостью,так имеются такие лаги ,которые сильнее соседних коррелируют с исходным рядом.То что она положительная также говорит о наличии тренда.Так что можно переходить к прогнозу.

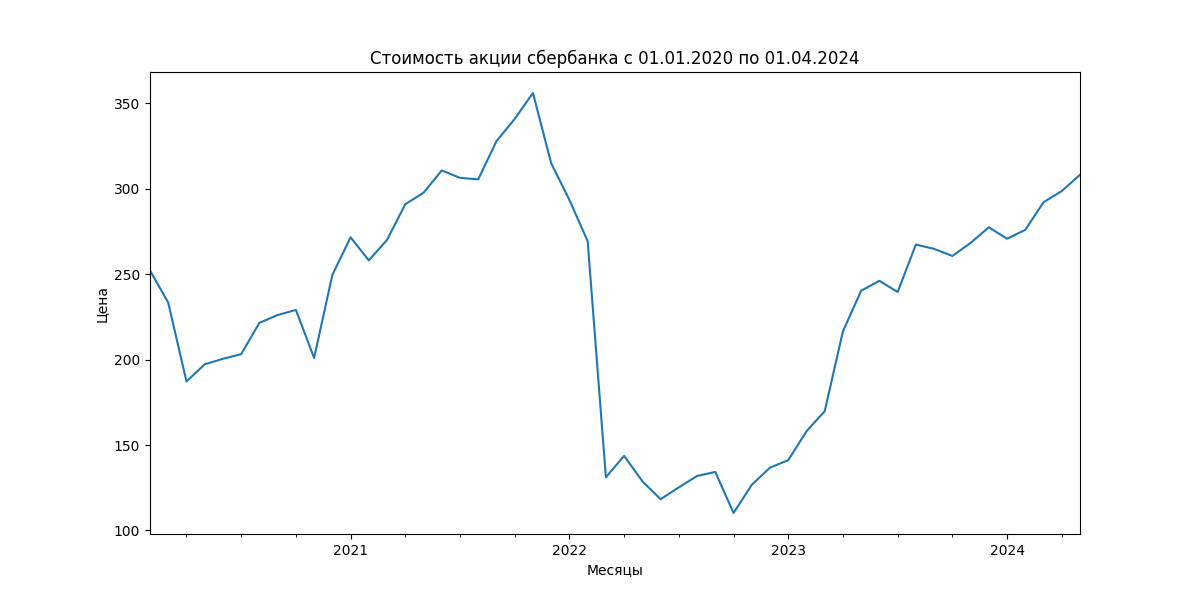
Спрогнозируем с помощью модели ARIMA(1,1,1).

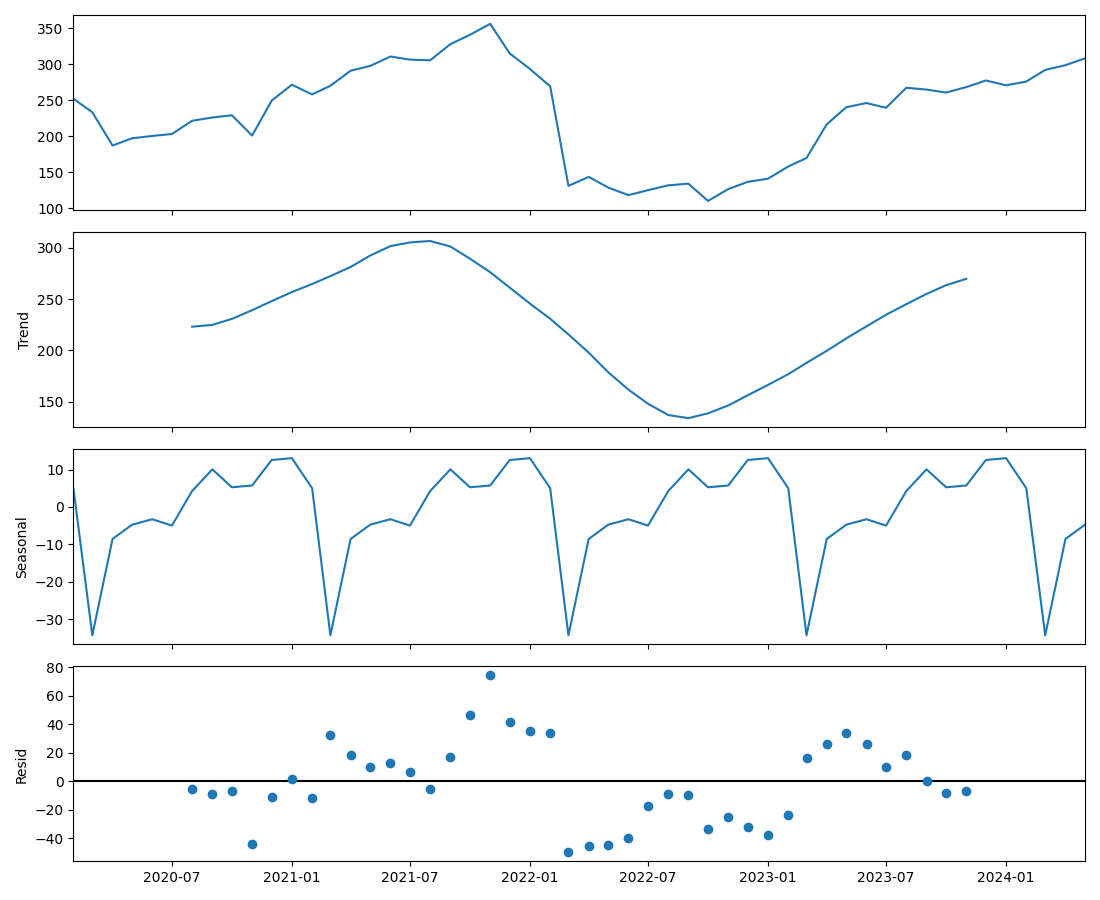
Изображение выглядит как текст, График, линия, диаграмма

Автоматически созданное описание

Данный прогноз идет с ошибкой RMSE= 5.1675326520493785, что является не плохим результатом. Данные такого типа мало зависят от внешних факторов, что позволяет прогнозировать их в долгосрочный период.

## 3.2 Прогнозирование стоимости акции Сбербанка

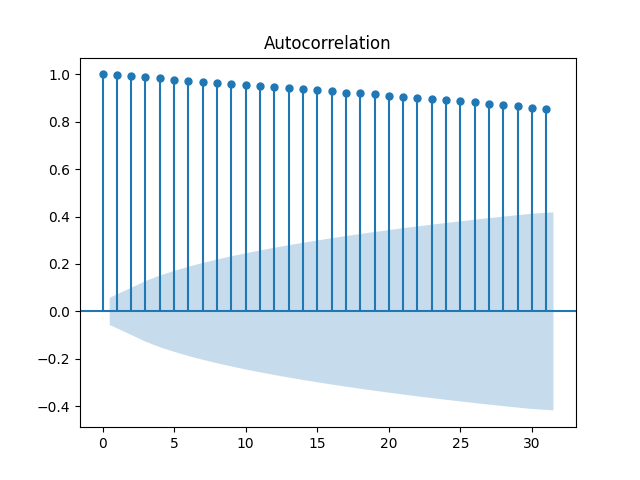
Спрогнозируем стоимость акции сбербанка. Для это используем данные с 1 января 2020 по 1 мая 2024 года для прогнозирования.

Визуально можно сразу увидеть тренд на рост стоимости. Разобьем наш ряд на компоненты и дадим им оценку.

После сильного падения в начале 2022 мы видим, что после июля 2022 держится тренд на рост цены.

Ряд обладает небольшой сезонностью, возможно это связанно с выплатами дивидендов или выпуском новых акции на рынок.

График остатков выглядит как случайные колебания, за исключением периода сильного падения, что говорит о том, что мы успешно разбили ряд на компоненты.

Теперь посмотрим на автокорреляцию и выразим пару мыслей.

Как мы видим корреляция положительная, что говорит о наличии тренда. Синяя граница позволяет оценить статистическую значимость корреляции. Если столбец выходит за ее пределы, то автокорреляция достаточно сильна и ее можно использовать при построении модели, в нашем случаи статистическая значимость очень сильная. А вот насчет сезонность, а именно ее отсутствие, указывает нам то, что лаги одинокого коррелируют с исходным рядом.

Теперь попробуем дать прогноз используя модель ARIMA (0,2,1).

Предсказанная цена на начало июня 322 рубля с ошибкой RSE= 11.56239, что говорит предсказанный диапазон стоимость примерно от 311 до 333 рублей.

Стоит сказать, что использовать модели прогнозирования экономических рядов в долгий срок не имеет смысла из-за невозможности предсказать внешних факторах, но использования на короткий срок дает хороший ориентир куда может пойти цена, что является очень важным аспектов.

# Заключение.

В результате проведенного исследования были достигнуты следующие цели:

1)Были рассмотрены основные термины анализа временных рядов.

2)Данны конкретные примеры существующих моделей, с помощью которых можно спрогнозировать временной ряд

2)Были проанализированы и спрогнозированы два временных ряда.

# Список литературы

Введение в анализ временных рядов: учебное пособие для вузов / Н. В. Артамонов, Е. А. Ивин, А. Н. Курбацкий, Д. Фантаццини ; Московский государственный университет имени М. В. Ломоносова, Московская школа экономики, Кафедра эконометрики и математических методов экономики. – Вологда: ВолНЦ РАН, 2021. – 134 с.

Анализ временных рядов: учебное пособие / В. А. Алексеева. – Ульяновск: УлГТУ, 2020. – 147с.

Анализ временных рядов и прогнозирование: учебник /В. Н. Афанасьев, М.М. Юзбашев. Изд. 2-е, перераб. и доп. –М.: Финансы и статистика, 2010. – 320 с.: ил.

<https://habr.com/ru/companies/otus/articles/732080/> Анализ временных рядов (Дата обращения: 27.05.2024).

<https://www.dmitrymakarov.ru/intro/time-series-20/> Временные ряды (Дата обращения: 27.05.2024).

<https://education.yandex.ru/handbook/ml/article/analitika-vremennyh-ryadov> Аналитика временных рядов (Дата обращения: 27.05.2024).

Медведев Г. А. Практикум на ЭВМ по анализу временных рядов: Учеб. пособие / Г.А.Медведев, В.А.Морозов. — Мн. Университетское, 2001. — 192 с.