

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ НАЦИОНАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

Лабораторная работа №5 по дисциплине
“Статистика и анализ данных”

Семестр 2

Выполнили студенты:

Косарев Илья,
гр. J3110, ИСУ 466304

Капустина Юлия,
гр. J3110, ИСУ 466110

Кащеев Максим,
гр. J3111, ИСУ 466147

Отчет сдан:
19.05.2025

Санкт-Петербург,
2025 г.

Оглавление

Введение	2
1.1 Цель и задачи	2
1.2 Теоретическая подготовка	3
1.2.1 Разностный ряд	3
1.2.2 Тренд, сезонность, лаг, белый шум	3
1.2.3 Стационарность	3
1.2.4 Автокорреляционная функция ряда (АКФ)	3
1.2.5 Экспоненциальное сглаживание	3
Практическая часть	4
Заключение	7

Введение

1.1 Цель и задачи

Цель: познакомиться с понятием временного ряда и применить различные методы анализа к данному.

Задачи:

1. Импортировать временной ряд для анализа (в данной работе в качестве такого ряда был взят временной ряд цен закрытия акций *MSFT*);
2. построить разностные ряды различного порядка;
3. построить автокорреляционную функцию ряда и провести анализ трендов, сезонности и т.д.;
4. применить тест Дики-Фуллера для исходного ряда и для полученных разностных рядов;
5. построить сглаженные ряды с различными параметрами;
6. визуализировать результаты;
7. сделать выводы.

1.2 Теоретическая подготовка

1.2.1 Разностный ряд

Разностный (дифференцированный) ряд — это новая последовательность, полученная из исходного временного ряда $\{y_t\}$ как приращения соседних наблюдений заданного порядка:

$$\Delta y_t = y_t - y_{t-1}$$

Разностевание подавляет тренд/сезонность и приближает ряд к стационарности, что упрощает его моделирование.

1.2.2 Тренд, сезонность, лаг, белый шум

Тренд — это систематическое, долговременное изменение среднего уровня временного ряда, не связанное с сезонными колебаниями или случайным шумом.

Сезонность — это регулярные, повторяющиеся с фиксированным периодом (месяц, квартал, сутки, час и т. д.) колебания временного ряда, возникающие из-за календарных, климатических, технологических или поведенческих циклов.

Лag (lag) — это сдвиг наблюдения временного ряда на k шагов назад. Величина k называется порядком лага.

Белый шум — это процесс нулевого математического ожидания, постоянной конечной дисперсии и некоррелированности любых двух различных моментов времени.

1.2.3 Стационарность

Стационарность (слабая) — свойство временного ряда, которое гласит:

- $\mathbb{E}(X_t) = \mu = \text{const} \quad \forall$ сдвига по лагу
- $\text{Var}(X_t) = \sigma^2 = \text{const} \quad \forall$ сдвига по лагу
- $\text{Cov}(X_t, X_{t-k}) = \text{Cov}(X_{t+h}, X_{t+h-k}) = \gamma(k) \quad \forall t, h$

1.2.4 Автокорреляционная функция ряда (АКФ)

АКФ — функция, которая сопоставляет порядку лага значение, отображающее линейную зависимость элементов временного ряда с предыдущими:

$$\rho(k) = \frac{\text{Cov}(X_t, X_{t-k})}{\text{Var}(X_t)} = \frac{\mathbb{E}[(X_t - \mu)(X_{t-k} - \mu)]}{\sigma^2}, \quad -1 \leq \rho(k) \leq 1 \quad \forall k$$

1.2.5 Экспоненциальное сглаживание

Экспоненциальное сглаживание — процесс построения нового ряда на основе данного по правилу:

$$\tilde{y}_{t+1} = \alpha y_t + (1 - \alpha) \tilde{y}_t,$$

где:

- y_t - последнее наблюдение
- \tilde{y}_t - сглаженное наблюдение
- α - "скорость забывания", при больших значениях результирующий ряд быстро улавливает переломы тренда в исходном, при маленьких - происходит сглаживание шума, из-за чего ряд легче анализировать визуально

Практическая часть

Код находится в публичном репозитории: [лабораторная №5](#)

В качестве ряда были взяты цены закрытия акций компании Microsoft (MSFT) за период с 01.01.2022 по 31.12.2023:



Рис. 2.1: Исходный временной ряд

По данному ряду были построены разностные ряды первого и второго порядков:



Рис. 2.2: Разностные ряды

На основе данных рядов были построены разные АКФ:

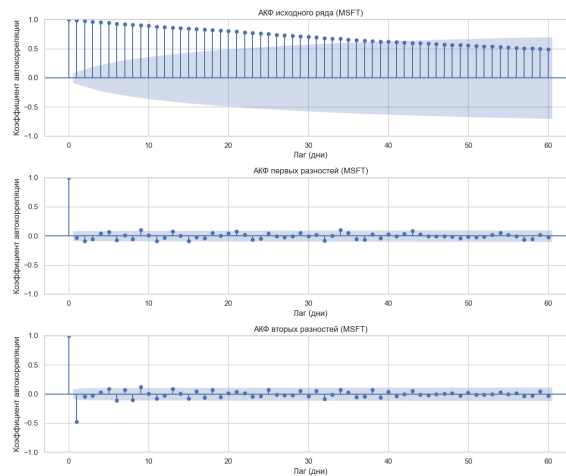


Рис. 2.3: АКФ для исходного ряда, разностного первого порядка и разностного второго порядка

На основе данного графика можно сделать несколько выводов:

- Исходный ряд не стационарен, о чем свидетельствует отсутствие быстрого убывания в ноль столбцов функции;
- у исходного ряда присутствует тренд;
- дифференцирование действительно убрало тренд из результирующих рядов, из-за чего ряды относительно стационарны;
- при построении разностного ряда второго порядка произошло "передифференцирование" из-за чего $\rho(1)$ заметно отличается от разностного ряда первого порядка;
- сезонность выявить трудно: нет резких совпадающих пиков на графике исходного ряда.

На данных рядах был применен тест Дики-Фуллера для анализа стационарности:

```

Анализ стационарности для MSFT

Результаты теста Дики-Фуллера для исходного ряда:
ADF Statistic: -0.4503
p-value: 0.9014
Критические значения:
1%: -3.4435
5%: -2.8674
10%: -2.5699
Вывод: Ряд нестационарен (не отвергаем H0)

Результаты теста Дики-Фуллера для первых разностей:
ADF Statistic: -17.6470
p-value: 0.0000
Критические значения:
1%: -3.4435
5%: -2.8674
10%: -2.5699
Вывод: Ряд стационарен (отвергаем H0)

Результаты теста Дики-Фуллера для вторых разностей:
ADF Statistic: -9.3366
p-value: 0.0000
Критические значения:
1%: -3.4440
5%: -2.8676
10%: -2.5700
Вывод: Ряд стационарен (отвергаем H0)

```

Рис. 2.4: Результаты теста Дики-Фуллера для данных рядов

Как можно увидеть из результатов, ранние предположения оказались верными: исходный ряд нестационарен, что делает его менее пригодным для анализа, чем дифференцированные.

Наконец, были построены экспоненциальные сглаживания исходного ряда с разными коэффициентами памяти:

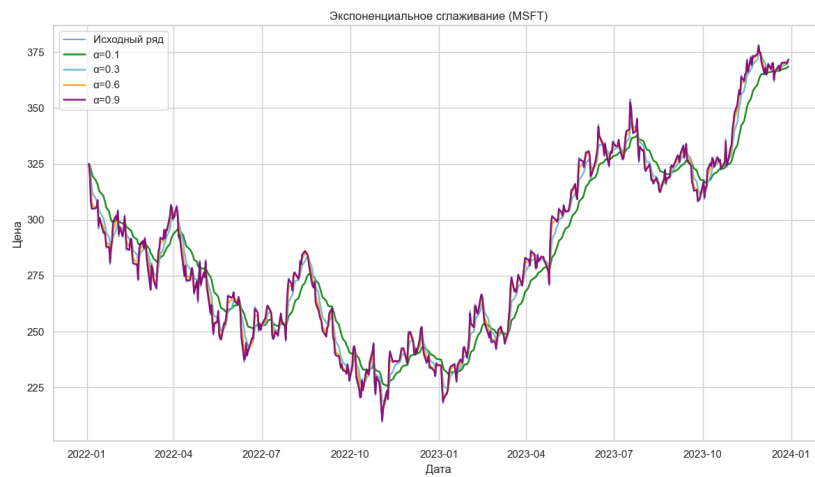


Рис. 2.5: Сглаживание исходного ряда

Как можно заметить, при малых значениях α шум убрался, при этом сохранились общие очертания исходного графика, что может быть полезно для визуализации и быстрого прогнозирования. При высоких значениях сильно заметны резкие колебания, что помогает замечать моментальные изменения, однако из-за этого прогнозирование и визуализация усложняются.

Заключение

В результате выполнения лабораторной работы мы познакомились с временными рядами и способами их анализа и преобразования. Применив теоретические знания на практике, мы провели анализ конкретного временного ряда акций компании и выявили определенные закономерности, а также смогли привести его к наиболее удобному для анализа виду. Знания и практический опыт, полученные во время выполнения работы, определенно понадобятся нам по профессии в будущем.