



МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«МИРЭА – Российский технологический университет»

РТУ МИРЭА

Институт информационных технологий (ИИТ)

Кафедра прикладной математики (ПМ)

ОТЧЕТ ПО ПРАКТИЧЕСКИМ РАБОТАМ

по дисциплине «Методы анализа данных»

Студент группы *ИНБО-05-19, Меликбекян А.А.*

(подпись)

Преподаватель *Кузьмин В.И.*

(подпись)

Отчет представлен «__» декабря 2021 г.

Москва 2021 г.

Оглавление

1. Исходные данные.....	3
2. Модели интенсивного роста	4
2.1. Темп роста пропорционален размеру системы	4
2.2. Темп роста пропорционален возрасту системы	6
2.3. Темп роста растет по экспоненциальной зависимости от возраста системы.....	8
3. Модели ограниченного роста	9
3.1. Модель Гомперца.....	9
3.2. Логистическая модель.....	11
4. Исключение тренда.....	12
5. Вывод	15
6. Список литературы.....	16

1. Исходные данные

В данных практических работах анализируется курс стоимости акций компании Oracle за период с 1986 года по 2021. Данные с историческими ценами акций взяты с сайта финансовой статистики finance.yahoo.com.

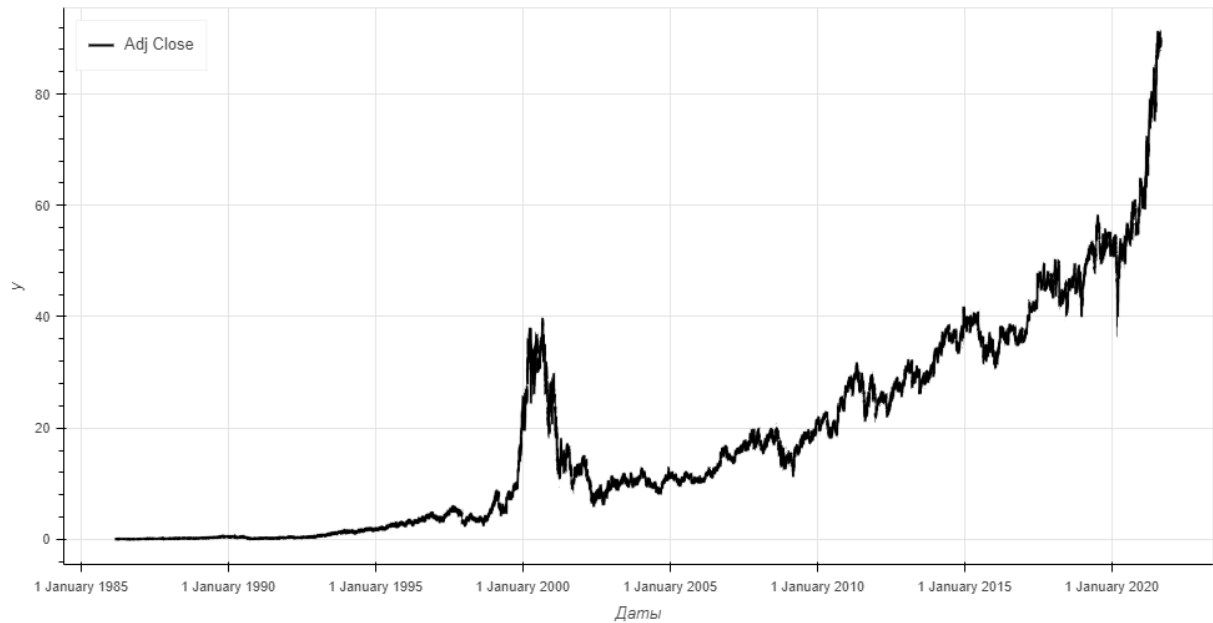


Рисунок 1 – Исходный график

Построим график в полулогарифмическом масштабе. Отметим уровни, расстояние между которыми равно единице, что в обычном масштабе соответствует отношению $e = 2.718$ между уровнями.

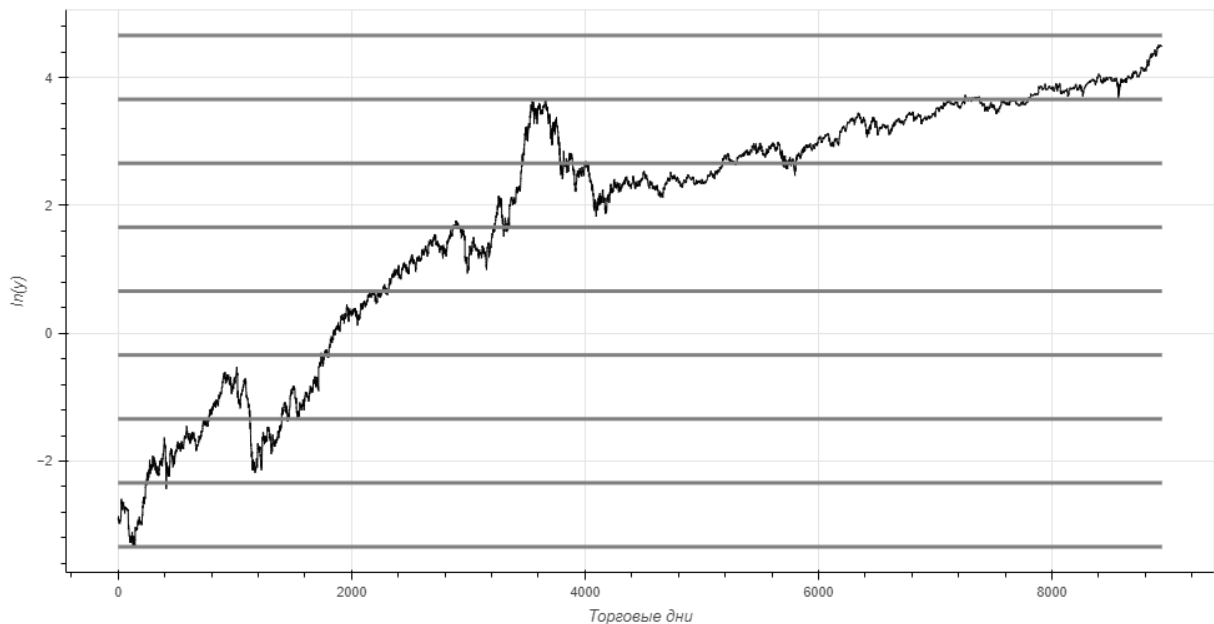


Рисунок 2 – График в полулогарифмическом масштабе

2. Модели интенсивного роста

2.1. Темп роста пропорционален размеру системы

Модель пропорциональная размеру системы характеризуется уравнением

$$k = a \cdot y,$$

где y – размер системы. Тогда

$$\dot{y} = k \cdot y^2$$

$$y = -\frac{1}{a \cdot t + c}$$

Анаморфозой для данного вида модели является:

$$\frac{1}{y} = a \cdot t + c \quad (2.1)$$

Спрямление достигается в координатах:

$$\frac{1}{y} \sim t \quad (2.2)$$

Эта модель является гиперболой и определяет положение вертикальной асимптоты, при которой значения функции достигает бесконечного значения, что для реальных систем невозможно. Поэтому по положению вертикальной асимптоты можно сделать оценку сверху для определения времени завершения процесса интенсивного роста. Для этого линеаризованные данные в координатах (2.2) нужно продлить до пересечения с осью абсцисс.

Построим график в координатах (2.2) и сравним с полулогарифмическим на соответствующих периодах времени.

На рисунке 3 показано сравнение графика в координатах (2.2) и в полулогарифмическом масштабе на периоде от 100 до 1100 торгового дня, на данном промежутке два участка с интенсивным ростом. По графику в координатах (2.2) мы можем определить время завершения процессов интенсивного роста, которые будут соответствовать 300 и 910 торговым дням. На полулогарифмическом графике процессы интенсивного роста заканчиваются на 300 и 910 торговые дни.

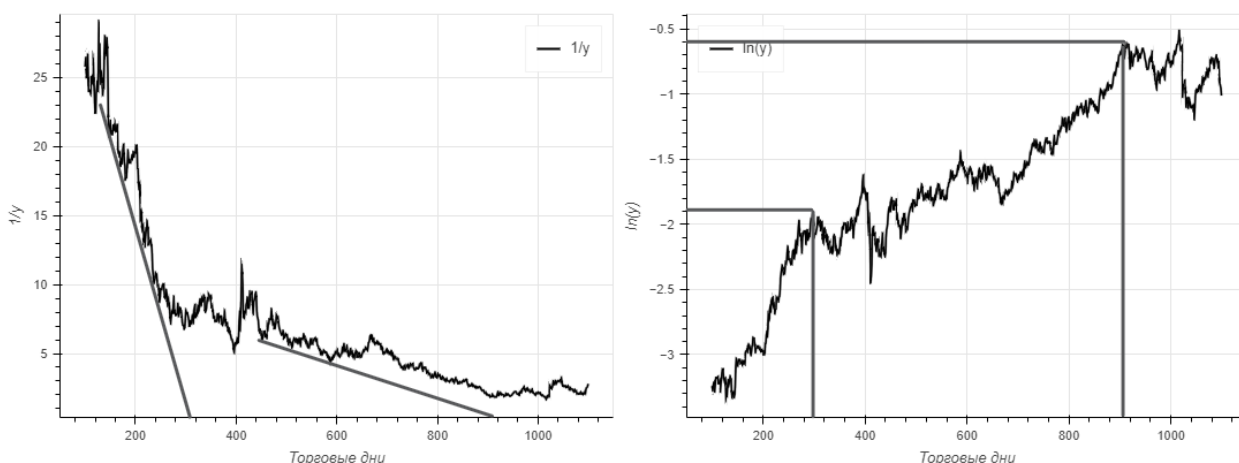


Рисунок 3 – Сравнение графика в координатах (2.2) и в полулогарифмическом масштабе на первом промежутке

На рисунке 4 показано сравнение графика в координатах (2.2) и в полулогарифмическом масштабе на периоде от 2000 до 3000 торгового дня. По графику в координатах (2.2) мы можем определить время завершения процесса интенсивного роста, которое будет соответствовать 2880 торговому дню. На полулогарифмическом графике интенсивный рост заканчивается на 2880 торговый день.

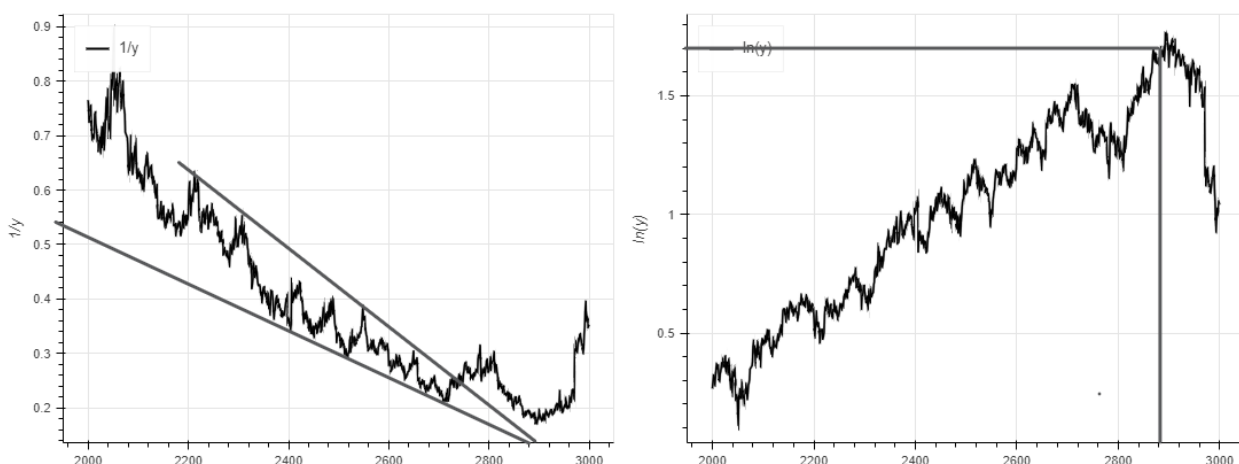


Рисунок 4 – Сравнение графика в координатах (2.2) и в полулогарифмическом масштабе на втором промежутке

На рисунке 5 показано сравнение графика в координатах (2.2) и в полулогарифмическом масштабе на периоде от 3300 до 3700 торгового дня. По графику в координатах (2.2) мы можем определить время завершения процесса интенсивного роста, которое будет соответствовать 3550 торговому дню. На полулогарифмическом графике интенсивный рост заканчивается на 3550 торговый день.

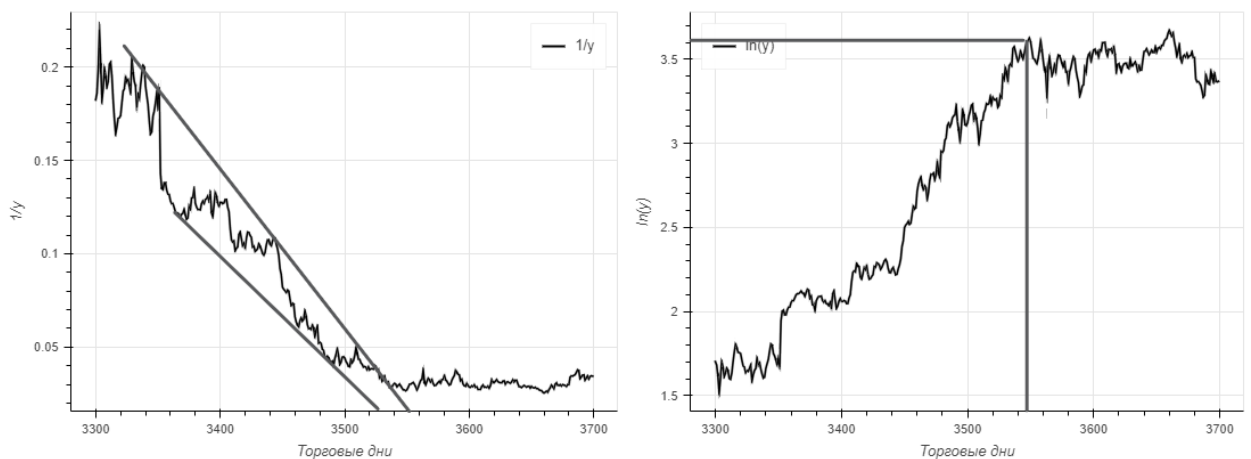


Рисунок 5 – Сравнение графика в координатах (2.2) и в полулогарифмическом масштабе на третьем промежутке от 3300 до 3700 торгового дня

Используя анаморфозу (2.1) можно найти время завершения процесса интенсивного роста.

2.2. Темп роста пропорционален возрасту системы

Модель пропорциональная возрасту системы характеризуется уравнением

$$k = a \cdot t,$$

где t – возраст системы. Тогда,

$$\dot{y} = a \cdot t \cdot y$$

$$y = C \cdot e^{\frac{a(t-t_0)^2}{2}}$$

Анаморфозой для данного вида модели является:

$$\ln(y) = a \cdot \frac{(t-t_0)^2}{2} + \ln C \quad (2.3)$$

Спряmlение достигается в координатах:

$$\ln(y) \sim (t - t_0)^2 \quad (2.4)$$

Определим предел роста значений стоимости акций. Построим графики в координатах (2.4) на промежутках, полученных в пункте 2.1.

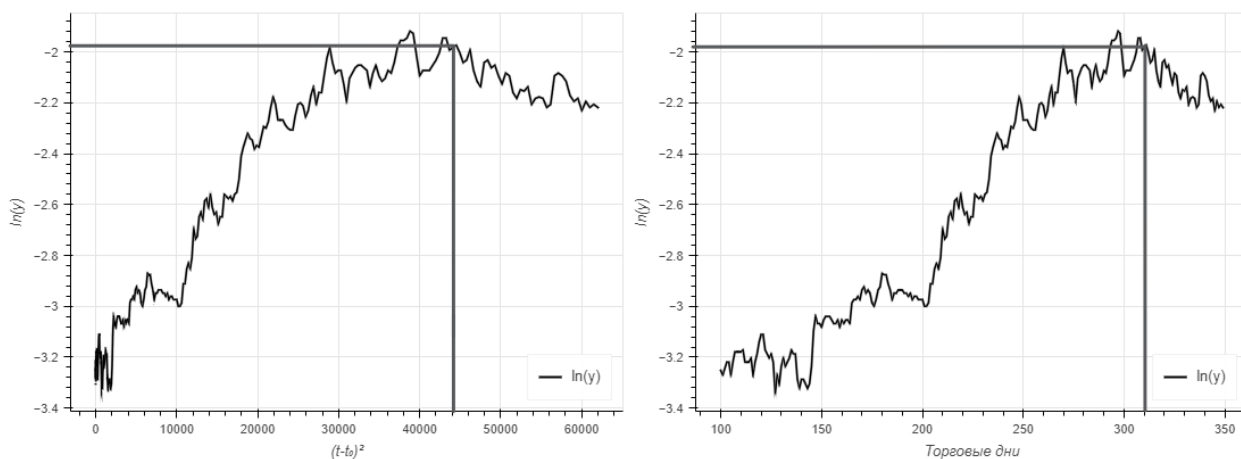


Рисунок 6 – Сравнение графика в координатах (2.4) и в полулогарифмическом масштабе на первом промежутке от 0 до 350 торгового дня

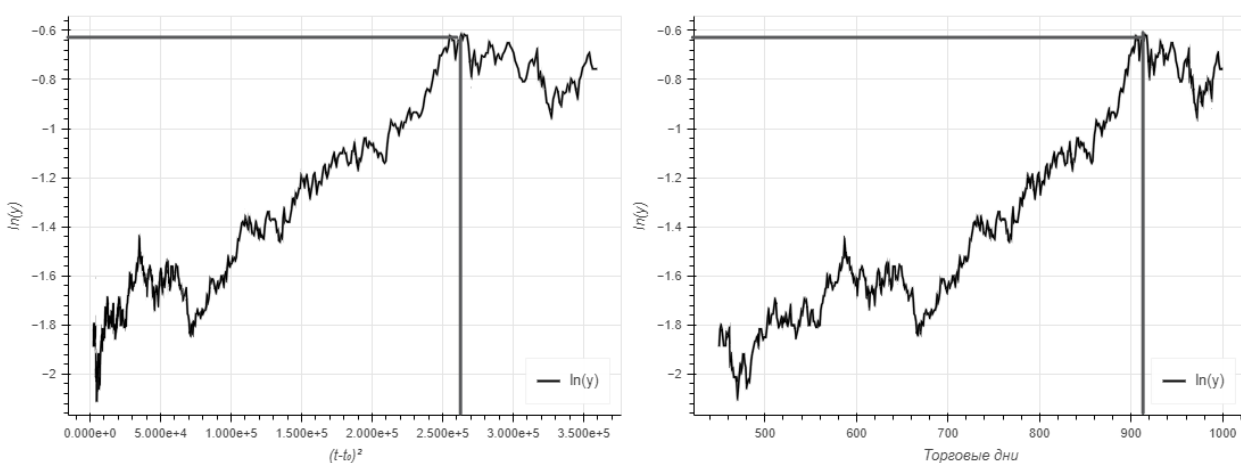


Рисунок 7 – Сравнение графика в координатах (2.4) и в полулогарифмическом масштабе на первом промежутке от 400 до 1100 торгового дня

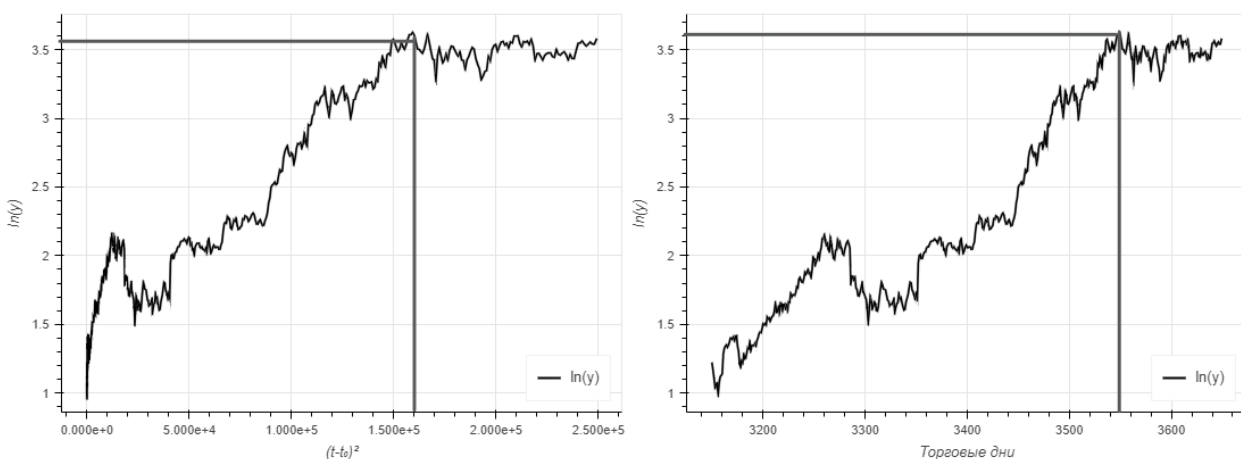


Рисунок 8– Сравнение графика в координатах (2.4) и в полулогарифмическом масштабе на третьем промежутке

Анаморфоза (2.3) позволяет сделать прогноз предела роста значения стоимости акций.

2.3. Темп роста растет по экспоненциальной зависимости от возраста системы

Модель пропорциональная размеру системы характеризуется уравнением

$$k = A \cdot e^{at},$$

где t – возраст системы. Тогда

$$\dot{y} = A \cdot e^{at} \cdot y,$$

$$y = C \cdot e^{\frac{A}{a}e^{at}}$$

Анаморфозой для данного вида модели является:

$$\ln\left(\ln\left(\frac{y}{C}\right)\right) = \ln\left(\frac{A}{a}\right) + a \cdot t \quad (2.5)$$

Спрямление достигается в координатах:

$$\ln\left(\ln\left(\frac{y}{C}\right)\right) \sim t \quad (2.6)$$

Варьированием константы C достигается наилучшее спрямление исходных данных. Для используемых данных наилучшее спрямление при $C = 0,035$.

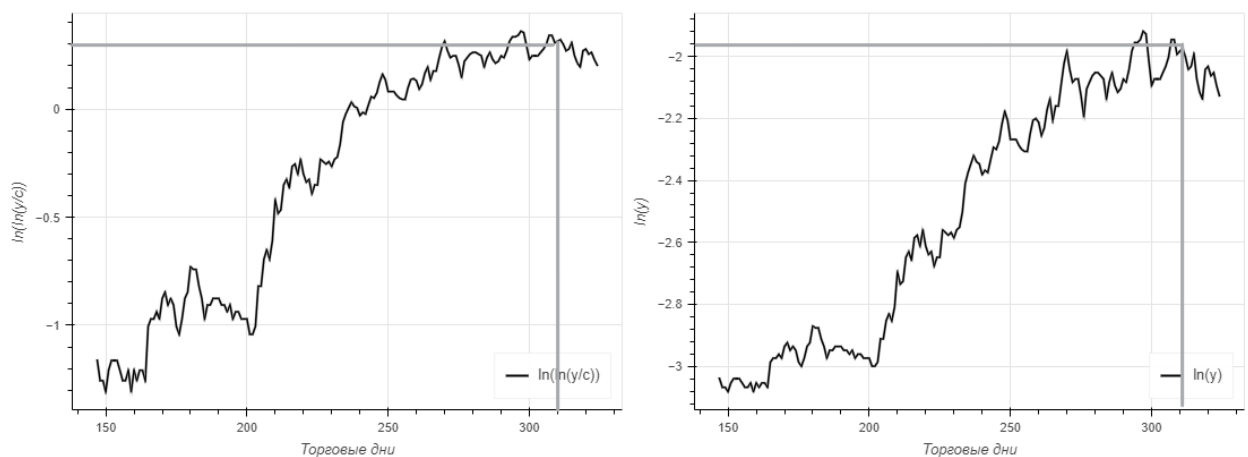


Рисунок 9 – Сравнение графика в координатах (2.6) и в полулогарифмическом масштабе на первом промежутке

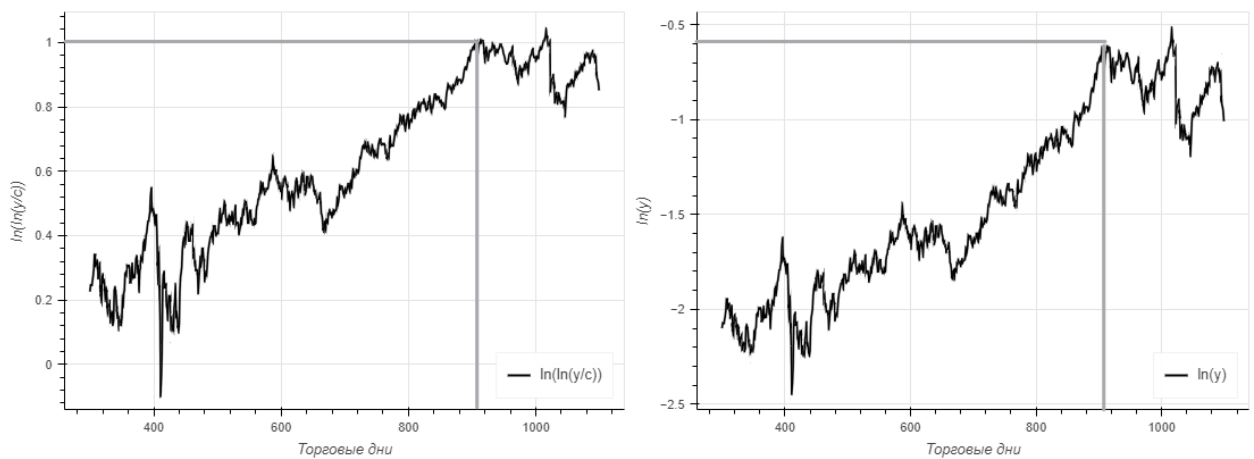


Рисунок 10 – Сравнение графика в координатах (2.6) и в полулогарифмическом масштабе на первом промежутке

3. Модели ограниченного роста

Наибольшее распространение среди моделей ограниченного роста получили модель Гомперца и логистическая модель. Первая модель определяет асимметричную кривую, а вторая – симметричную. На основе этих моделей можно эффективно прогнозировать предел роста системы.

3.1. Модель Гомперца

Для модели Гомперца

$$\frac{dy}{dt} = Ae^{-kt} \cdot y \quad (3.1)$$

её интеграл дает $\ln(y) = -\frac{A}{k}e^{-kt} + \ln C$

При $t \rightarrow \infty$ постоянная интегрирования $C = y_{\infty}$. Тогда

$$\ln \frac{y}{y_{\infty}} = -\frac{A}{k}e^{-kt} \text{ или } Ae^{-kt} = -k \ln \frac{y}{y_{\infty}}.$$

После подстановки последнего соотношения в уравнение (3.1), получим:

$$\frac{1}{y} \frac{dy}{dt} = k(\ln y_{\infty} - \ln y) = -k \ln \left(\frac{y}{y_{\infty}} \right) \quad (3.2)$$

Из исходного уравнения модели Гомперца получается следующая анаморфоза:

$$\ln \left(\frac{1}{y} \frac{dy}{dt} \right) = \ln(A) - kt \quad (3.3)$$

Спрямление достигается в координатах:

$$\ln \left(\frac{1}{y} \frac{dy}{d\tau} \right) \sim t \quad (3.4)$$

На пересечении продолжения спрямления и прямой k находится момент точки перегиба.

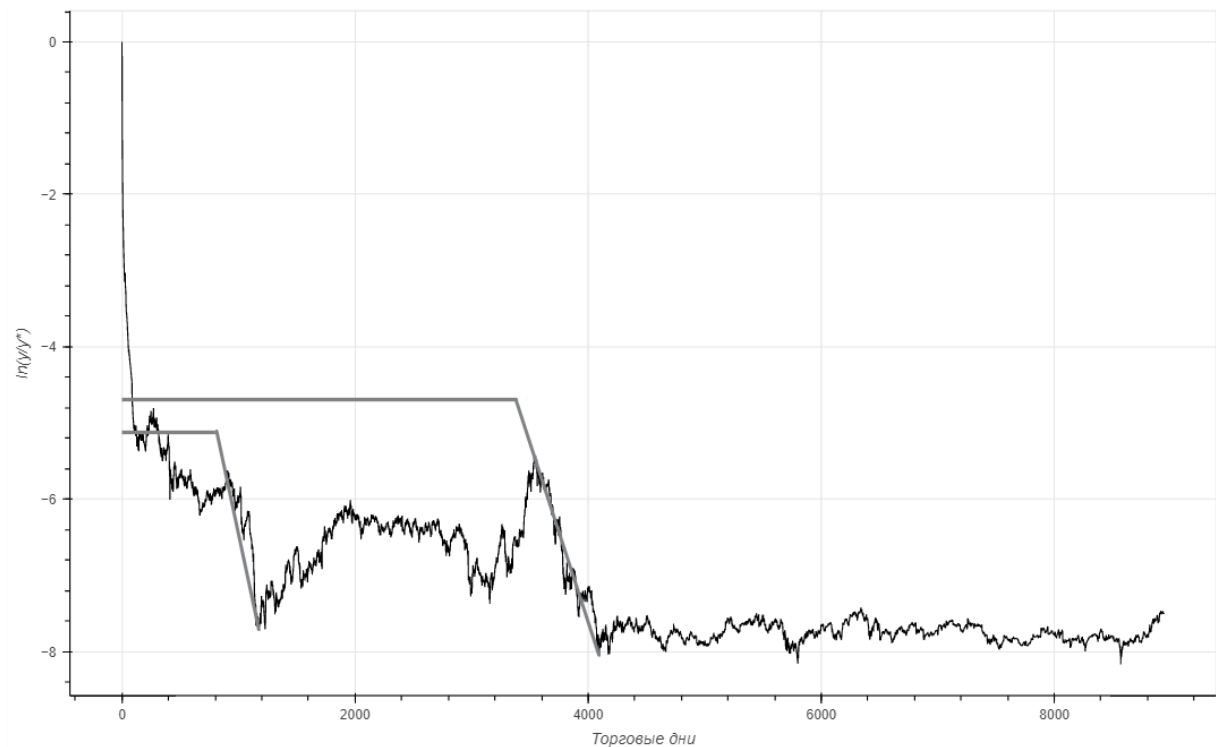


Рисунок 9 – График в координатах (3.4)

Таблица 1 – Коэффициент, полученный из Рис. 9

Промежуток	k	ln(k)
1	0,0059	-5.12
2	0,0092	-4.69

Следующая анаморфоза получается из уравнения (3.2):

$$\frac{1}{y} \frac{dy}{dt} = k(\ln y_{\infty} - \ln y) \quad (3.5)$$

Таким образом, анаморфозой, соответствующей линейной зависимости между характеристиками системы, являются координаты:

$$\frac{1}{y} \frac{dy}{dt} \sim \ln(y) \quad (3.6)$$

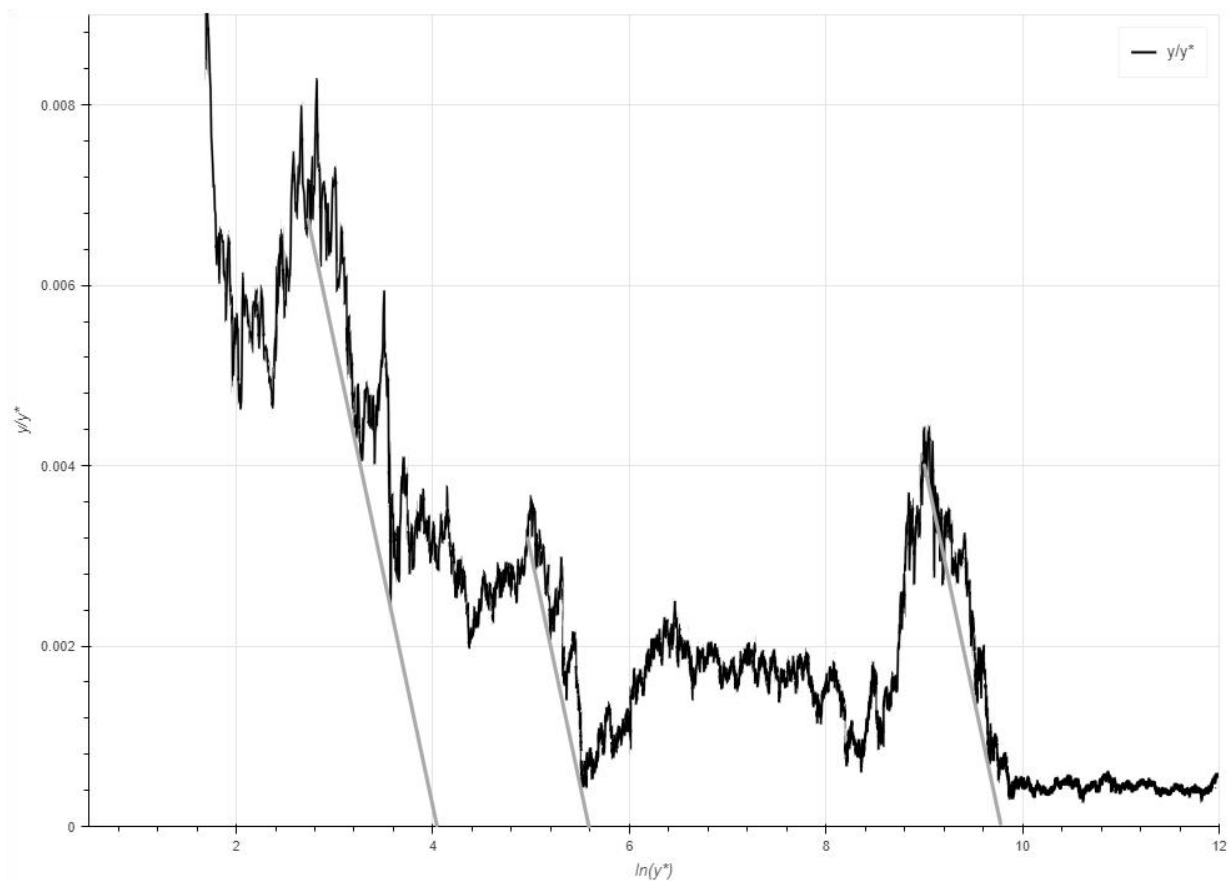


Рисунок 10 – График в координатах (3.6)

Таблица 2 – Коэффициенты, полученные из Рис. 10

Промежуток	$\ln(y_{\infty})$	y_{∞}
1	4,1	60
2	5,6	270
3	9,8	18033

3.2. Логистическая модель

Простейшей широко распространенной моделью ограниченного роста является логистическое уравнение:

$$\frac{dy}{dt} = k \cdot y \left(1 - \frac{y}{y_{\infty}}\right) \quad (3.7)$$

Из уравнения (3.7) получается следующая анаморфоза:

$$\frac{dy}{y \, dt} = k \cdot \left(1 - \frac{y}{y_{\infty}}\right) \quad (3.8)$$

Это уравнение определяет линейную зависимость в координатах:

$$\frac{dy}{y dt} \sim y \quad (3.9)$$

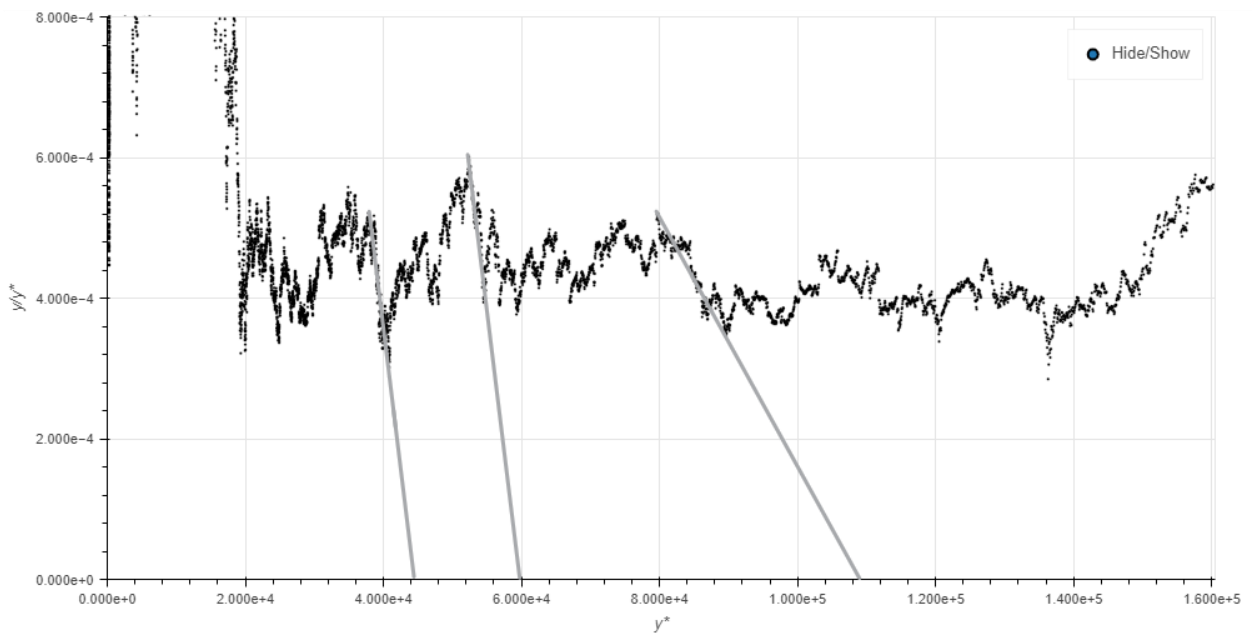


Рисунок 11 – График в координатах (3.9)

Таблица 3 – Коэффициенты, полученные из Рис. 11

Промежуток	y_{∞}
1	44000
2	60000
3	180000

4. Исключение тренда

Для исключения тренда воспользуемся формулой (4.1)

$$S = \frac{y_{t-\Delta t} + y_{t+\Delta t}}{2 \cdot y_t} \quad (4.1)$$

Параметр Δt определяется минимальным значением функции на графике $|m| \sim t$, где m – это среднее значение. Определим Δt равным 205.

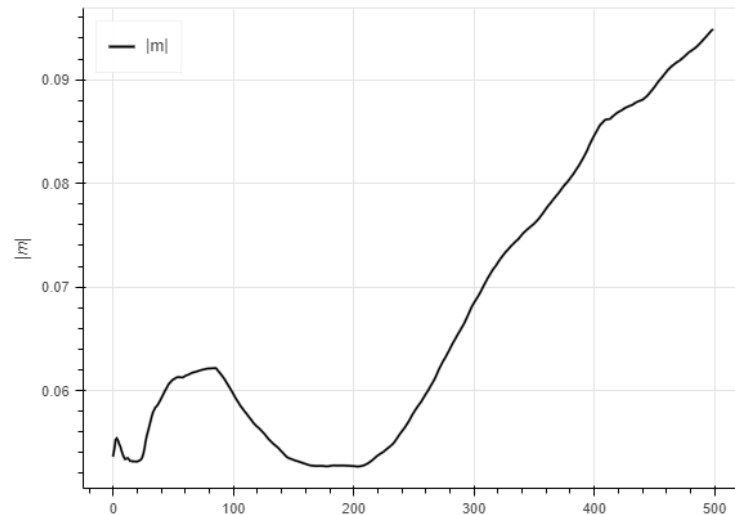


Рисунок 12 – Определение Δt

Исключим тренд из исходных данных.

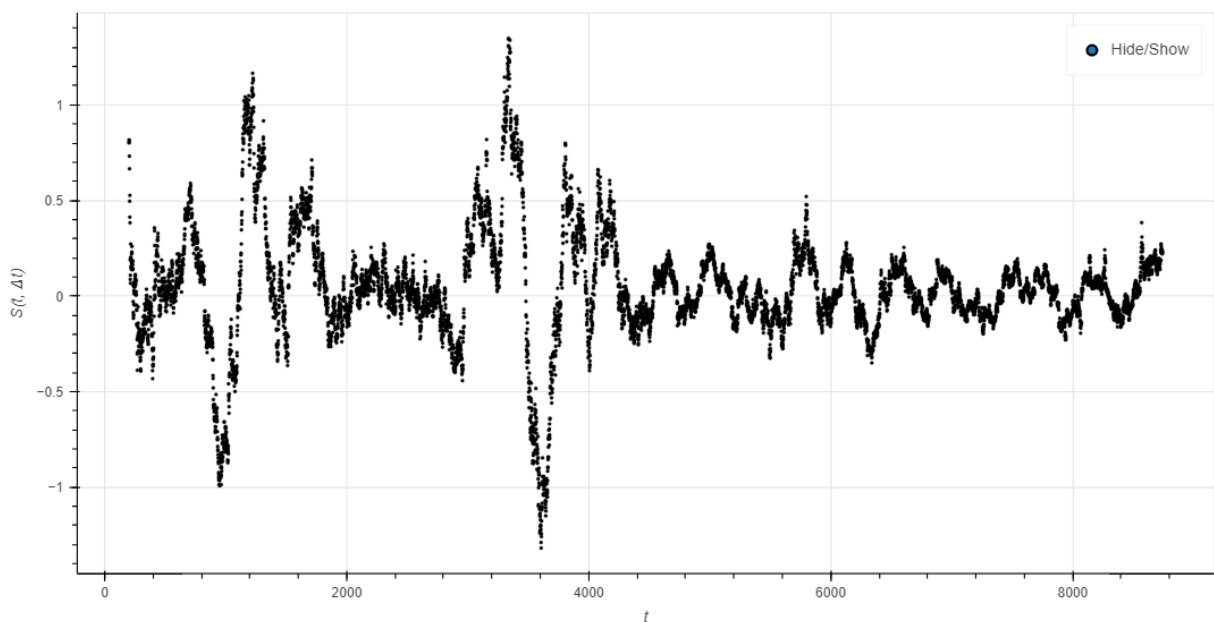


Рисунок 13 – Данные после исключения тренда

Для построения графика функции Альтера-Джонсона воспользуемся формулой (4.2):

$$\alpha(\tau_k) = \frac{1}{n-\tau} \sum_{i=1}^{n-\tau} |f(t_i + \tau_k) - f(t_i)| \quad (4.2)$$

Минимумы на графике функции Альтера-Джонсона определяют совокупность почти-периодов. На графике выделен почти-период равный 684 торговым дням.

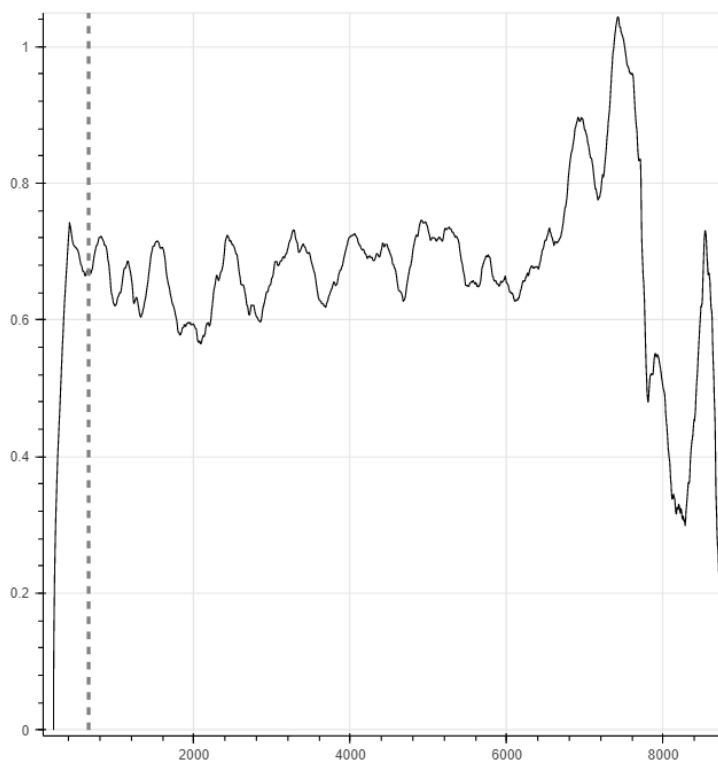


Рисунок 14 - График функции Альтера-Джонсона

Отобразим значения почти-периодов на полулогарифмическом графике.

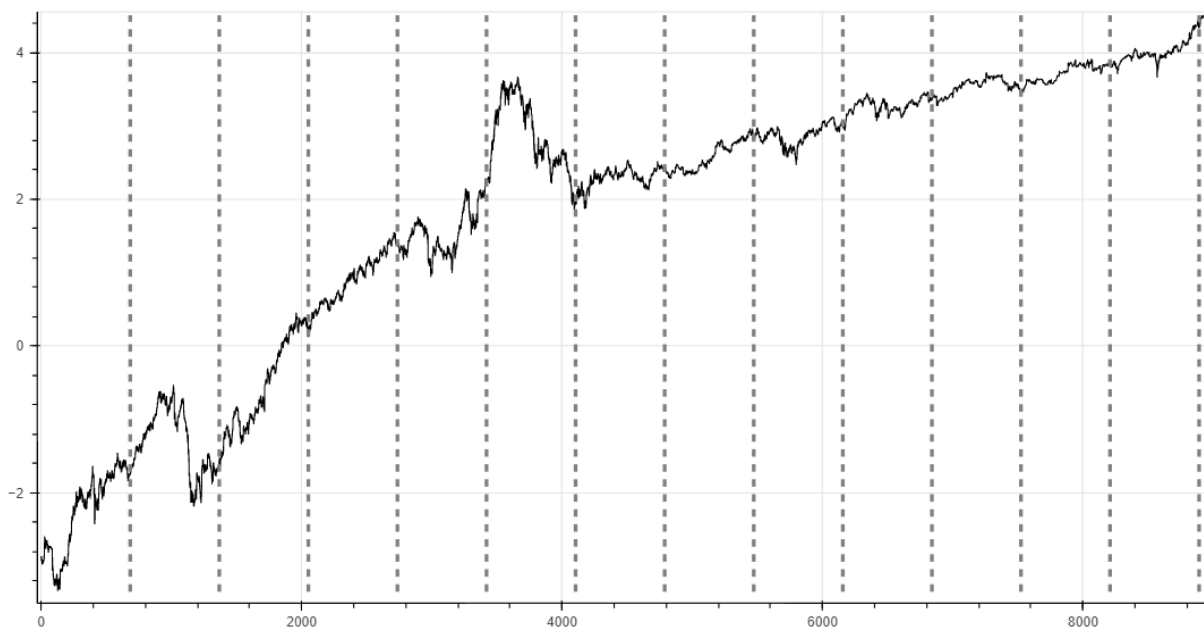


Рисунок 15 – График в полулогарифмическом масштабе с отмеченными почти-периодами

Отобразим значения почти-периодов на графике с исключенным трендом.

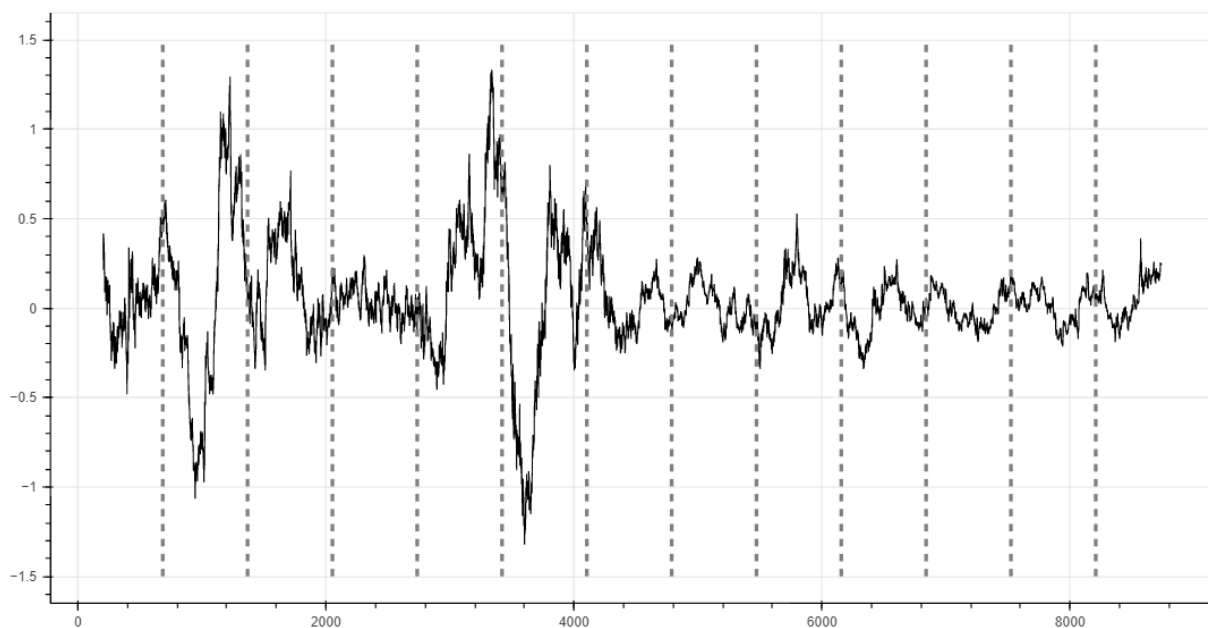


Рисунок 16 – График с исключенным трендом с отмеченными почти-периодами

Выделим ячейки развития на полупологарифмическом графике. Длину такта примем за $T_k = 2080$ – минимум функции Альтера-Джонса.

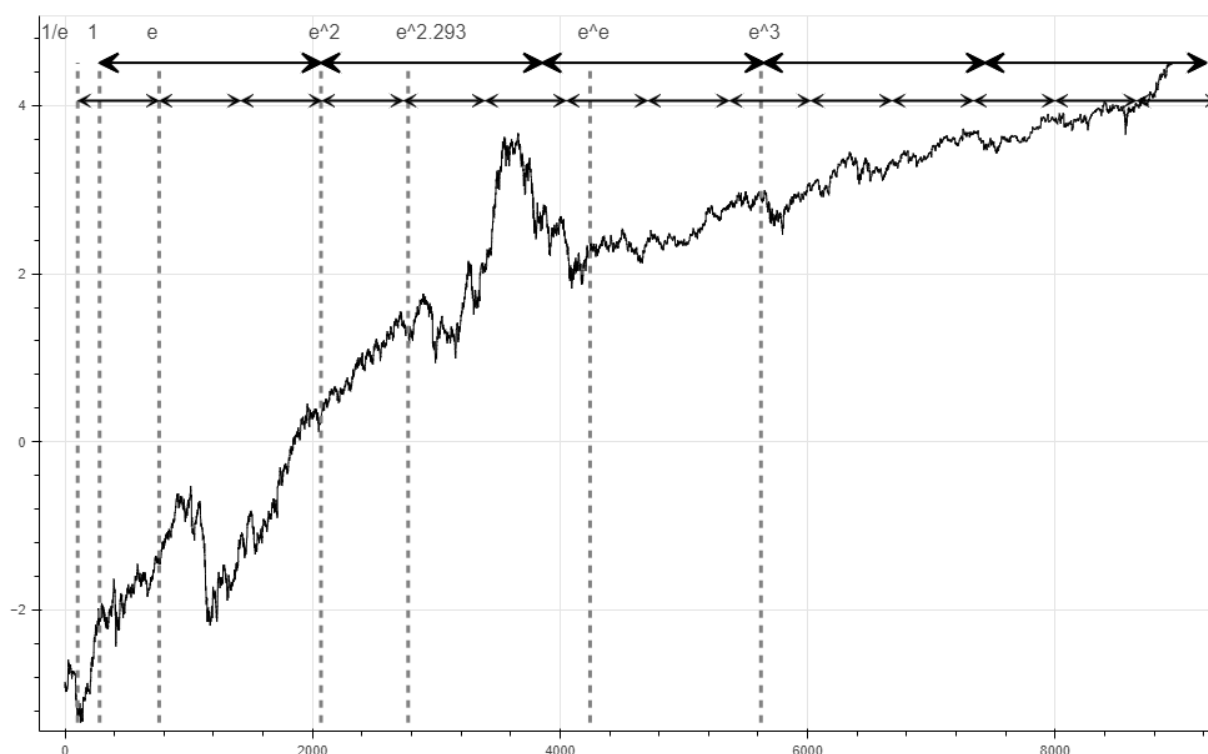


Рисунок 17 – Ячейки развития

5. Вывод

По результатам практических работ был проведен анализ динамики системы, представляющей собой курс акций компании Oracle. С помощью систем нелинейных преобразований (анаморфоз), обеспечивающих

линеаризацию исходного ряда, были выделены последовательности линейных или кусочно-линейных интервалов, используя которые были определены точки перегиба и пределы роста системы. По сдвиговой функции Альтера-Джонсона были определены почти-периоды и ячейки развития.

Исходя из графика в полулогарифмическом масштабе, можно сделать вывод, что в скором времени курс акций дойдет до очередной линии сопротивления. После чего стоимость акций либо пробьет уровень, либо отскочит от него.

6. Список литературы

1. Провайдер финансовой информации «Yahoo! Finance» [Электронный ресурс] / Исторические данные. – Режим доступа: <https://finance.yahoo.com/quote/ORCL/history?p=ORCL>
2. Кузьмин В.И., Гадзаов А.Ф. Методы построения моделей по эмпирическим данным: Учебное пособие / Московский государственный институт радиотехники, электроники и автоматики (технический университет) – М., 2012. – 94 с.