

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
**«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ НАЦИОНАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ»**
(Н И У « Б е л Г У »)

ИНСТИТУТ ИНЖЕНЕРНЫХ ТЕХНОЛОГИЙ И ЕСТЕСТВЕННЫХ НАУК
КАФЕДРА ОБЩЕЙ МАТЕМАТИКИ

**ЗАДАЧИ АНАЛИЗА И МОДЕЛИРОВАНИЯ ТЕНДЕНЦИЙ
ВРЕМЕННЫХ РЯДОВ**

Выпускная квалификационная работа

обучающегося по направлению подготовки
01.03.02 Прикладная математика и информатика
очной формы обучения,
группы 07001305

Ткачевой Евгении Валерьевны

Научный руководитель
к.ф.-м.н., доцент
Ерина Т.А.

БЕЛГОРОД 2017

Оглавление

ВВЕДЕНИЕ	3
1. Математические методы анализа временных рядов	7
1.1 Оценка данных.....	7
1.2 Выявление тенденций в развитии временного ряда	8
1.3 Аналитическое сглаживание временного ряда.....	11
2. Прогнозирование на основе временных рядов.....	17
2.1 Прогнозирование на основе кривых роста.....	17
2.2 Прогнозирование на основе адаптивных моделей.....	19
2.2.1 Экспоненциальное сглаживание	21
2.2.2 Модели линейного роста	25
2.2.3 Стохастический процесс Тейла И Вейджа.....	27
2.2.4 Сезонные модели.....	29
2.2.5 Прогнозирование с коэффициентами сезонности.....	30
2.2.6 Аддитивная модель сезонных явлений	32
2.2.7 Модели авторегрессии — скользящего среднего (метод Бокса —Дженкинса).....	35
3. Автоматизация построения моделей	36
3.1 Разработка программного средства	36
3.1.1 Функции системы	36
3.1.2 Технические требования к системе.	38
3.1.3 Описание программы.	39
3.1.4 Результат работы системы.....	43
3.2 Анализ работы программной системы	44
ЗАКЛЮЧЕНИЕ.....	59
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	60
ПРИЛОЖЕНИЕ	63

ВВЕДЕНИЕ

Современное общество постоянно испытывает необходимость в прогнозировании. Решение задачи прогнозирования играет важнейшую роль в процессах как стратегического планирования, так и оперативного управления в различных сферах науки и техники.

Одним из методов, позволяющих обеспечить точное планирование, является прогнозирование. Несмотря на большое количество разработанных методов, все они преследуют одну и ту же цель — предсказать события, которые произойдут в будущем, чтобы учесть их при разработке планов и стратегии развития компании.

Задача прогнозирования будущих значений временного ряда на основе его исторических значений является основой для финансового планирования в экономике и торговле, планирования, управления и оптимизации объемов производства, складского контроля и т.д.

В настоящее время компаниями осуществляется накопление исторических значений экономических и физических показателей в базах данных, что существенно увеличивает объемы входной информации для задачи прогнозирования.

Вместе с тем, развитие аппаратных и программных средств предоставляет все более мощные вычислительные платформы, на которых возможна реализация сложных алгоритмов прогнозирования.

Кроме того, современные подходы к экономическому и техническому управлению предъявляют все более жесткие требования к точности прогнозирования.

Таким образом, задача прогнозирования временных рядов усложняется одновременно с развитием информационных технологий.

В настоящее время задача прогнозирования различных временных рядов актуальна и является неотъемлемой частью ежедневной работы многих

компаний. Задача прогнозирования временного ряда решается на основе создания модели прогнозирования, адекватно описывающей исследуемый процесс.

Существенным недостатком авторегрессионного класса является большое число свободных параметров, идентификация которых неоднозначна и ресурсоемка. Существенным недостатком класса нейросетевых моделей является недоступность промежуточных вычислений, выполняющихся в «черном ящике», и, как следствие, сложность интерпретации результатов моделирования.

Кроме того, еще одним недостатком данного класса моделей является сложность выбора алгоритма обучения нейронной сети.

Применение каких-либо из существующих в настоящее время математических моделей и методов прогнозирования временных рядов тесно связано со спецификой предметной области и классификацией прогнозируемого временного ряда.

Задача прогнозирования финансовых временных рядов была и остается актуальной, поскольку предсказание является необходимым элементом любой инвестиционной деятельности, ведь сама идея инвестирования – вложения денег с целью получения дохода в будущем – основывается на идее прогнозирования будущего

Таким образом, задача анализа и прогнозирования работы предприятия является актуальной.

Существуют два общепринятых подхода к прогнозированию: качественный и количественный. Методы качественного прогнозирования особенно важны, если исследователю недоступны количественные данные. Как правило, эти методы носят весьма субъективный характер. Если имеются статистические данные об объекте исследования за прошедший временной период, следует применять методы количественного прогнозирования. Эти методы позволяют предсказать состояние объекта в будущем на основе данных о его прошлом. Методы количественного прогнозирования

разделяются на две категории: анализ временных рядов и методы анализа причинно-следственных зависимостей.

Цель работы – реализация компьютерной поддержки анализа и моделирования тенденций временных рядов.

Задачи:

1. Провести анализ существующих методов анализа и моделирования тенденций временных рядов;
2. Проанализировать подходы к прогнозированию временных рядов с учетом ограничений каждого подхода;
3. Проанализировать существующий инструментарий;
4. Спроектировать программную систему анализа и моделирования тенденций временных рядов;
5. Разработать программную систему;
6. Протестировать работу программной системы.

Объект исследования – методы анализа и моделирования тенденций временных рядов.

Предмет исследования – программная поддержка анализа и моделирования тенденций временных рядов.

Методы исследования – методы алгоритмического сглаживания, методы аналитического сглаживания, адаптивные методы временных рядов.

Данная ВКР состоит из введения, трех глав, заключения, списка использованных источников.

Во введении представлена актуальность, цель, задачи, предмет и объект исследования.

В первой главе представлено описание математических методов анализа временных рядов, сформулированы цели и задачи исследования.

Во второй главе приведено описание методов прогнозирования на основе кривых роста и на основе адаптивных методов.

В третьей главе приведено описание разработанного в среде Delphi программного средства и проведен анализ его работы путем сравнения результатов расчета в программном средстве с аналогичными показателями, рассчитанными с использованием табличного процессора MS Excel.

ВКР написана на 62 страницах, содержит 1 таблицу, 64 формулы, 21 рисунок, 1 приложение.

1. Математические методы анализа временных рядов

1.1 Оценка данных

Анализ временных рядов, отражающих развитие экономических процессов, начинается с оценки данных. Уровни исследуемого показателя обязательно должны быть сопоставимыми, однородными и устойчивыми, а их число должно быть достаточно велико. Сопоставимость достигается в результате одинакового подхода к наблюдениям на разных этапах формирования динамического ряда. Уровни во временных рядах должны иметь одинаковые:

- единицы измерения;
- шаг наблюдений;
- интервал времени;
- методику расчета;
- элементы, относящиеся к неизменной совокупности.

Однородность данных означает отсутствие сильных изменений тенденций, а также аномальных наблюдений, которые нетипичны для данного ряда. Аномальные наблюдения обычно проявляются в виде сильного изменения уровня – скачка или спада – с последующим приблизительным восстановлением предыдущего уровня. Наличие аномалии резко искажает результаты моделирования. Поэтому аномальные наблюдения необходимо исключить из временного ряда, заменив их расчетными значениями. Устойчивость характеризуется преобладанием закономерности над случайностью в изменении уровней ряда. На графиках устойчивых временных рядов закономерность прослеживается визуально, на графиках неустойчивых рядов изменения последовательных уровней представляются хаотичными, и поэтому поиск закономерностей в формировании значений

уровней таких рядов лишен смысла. Требование полноты данных обуславливается тем, что закономерность может обнаружиться лишь при наличии минимально допустимого объема наблюдений.

Экстраполяционное прогнозирование экономических процессов, представленных одномерными временными рядами, сводится к выполнению следующих основных этапов:

- 1) предварительный анализ данных;
- 2) построение моделей: формирование набора аппроксимирующих функций (кривых роста) и численное оценивание параметров моделей;
- 3) проверка адекватности моделей и оценка их точности;
- 4) выбор лучшей модели;
- 5) расчет точечного и интервального прогнозов.

1.2 Выявление тенденций в развитии временного ряда

При анализе и прогнозировании рядов динамики самой важной задачей является определение основной тенденции развития исследуемого процесса.

Одним из приемов выявления тенденции является метод скользящей средней. Данный метод относится к методам алгоритмического сглаживания временного ряда. После применения данного метода уравнение тренда (неслучайной функции времени, описывающей общее изменение процесса во времени) в явном виде не находят, но ряд приобретает более гладкий вид, по которому проще подобрать подходящую функцию для моделирования тенденции. Метод скользящего среднего состоит в замене абсолютных данных средними арифметическими за отдельные периоды. Расчет средних ведется способом скольжения, то есть постепенным исключением из принятого периода скольжения первого уровня и включением следующего.

Алгоритм скользящего среднего заключается в следующем:

$$\hat{f}(t) = \sum_{k=-m}^m w_k x(t+k), t = m+1, m+2, \dots, n-m, \quad (1.1)$$

где w_k – некоторые весовые коэффициенты, в сумме равные 1, т.е.

$$\sum_{k=-m}^m w_k = 1.$$

при переходе от t к $t+1$ в составе слагаемых происходит замена только одного слагаемого $x(t - m)$ слагаемым $x(t + m + 1)$.

Определение весов w_k основано на следующей процедуре.

Для $2m+1$ элементов временного ряда $x(1), x(2), \dots, x(2m+1)$ строится полином степени p методом наименьших квадратов; значение этого полинома используют для расчёта значений оценки тренда $\hat{f}(t)$ в средней точке этого отрезка ряда $m + 1$, т.е.

$$\hat{f}(m+1) = \bar{x}(m+1). \quad (1.2)$$

Эта же процедура выполняется для отрезка временного ряда $x(2), \dots, x(2m+2)$, что приводит к процедуре взвешенного скользящего суммирования, по которой веса определяются из следующей таблицы.

Таблица 1.1 Веса коэффициентов

m	веса
1	1/3, 1/3, 1/3
2	-3/35, 12/35, 17/35, 12/35, -3/35
3	-2/21, 3/21, 6/21, 7/21, 6/21, 3/21, -2/21

Методы скользящего среднего основываются на том, что все значения временного ряда имеют одинаковую информационную ценность. Однако в задачах прогноза, в которых сглаженная функция $\hat{f}(t)$ используется обычно для формирования прогнозов на несколько шагов вперёд, недавние значения $x(t)$ являются более ценными, чем значения ряда в далёком прошлом, так как ряд далее будет вести себя так, какова сформировавшаяся тенденция в настоящем и недалёком прошлом.

Эта идея реализована в методе экспоненциально взвешенного скользящего среднего Брауна

$$\begin{aligned}\widehat{f}(t+1) &= \frac{1-\lambda}{1-\lambda^{t+1}} \widehat{x}(t+1) + \lambda \frac{1-\lambda^t}{1-\lambda^{t+1}} \widehat{f}(t), t=1,2,\dots \\ \widehat{f}(1) &= x(1).\end{aligned}\tag{1.3}$$

Но наиболее эффективным способом выявления основной тенденции развития является аналитическое выравнивание.

При этом уровни ряда динамики выражаются в виде функции времени:

При таком подходе считается, что влияние других факторов на зависимый показатель несущественно или косвенно сказывается через фактор времени.

Основная тенденция (тренд) показывает, как воздействуют систематические факторы на уровень ряда динамики, а колеблемость уровней около тренда служит мерой воздействия остаточных факторов.

Оценка качества модели сводится к оценке ее точности и адекватности.

Проверка адекватности выбранных моделей реальному процессу (в частности, адекватности полученной кривой роста) строится на анализе остаточной компоненты. Остаточная компонента получается после выделения из исследуемого ряда систематической составляющей (тренда и периодической составляющей, если она присутствует во временном ряду).

Принято считать, что модель адекватна описываемому процессу, если остаточная последовательность (ряд остатков) представляет собой случайную компоненту ряда.

Поэтому при оценке «качества» модели проверяют, удовлетворяет ли остаточная последовательность следующим свойствам:

1. случайности колебаний уровней ряда;
2. соответствию распределения остаточной компоненты нормальному закону с нулевым математическим ожиданием;
3. независимости значений уровней ряда остатков между собой.

1.3 Аналитическое сглаживание временного ряда.

Аналитическое сглаживание предполагает построение аналитической функции в явном виде, описывающей тенденцию изменения процесса во времени.

Для построения трендов чаще всего применяются следующие функции:

- линейный тренд: $y_t = a + b \cdot t$;
- гипербола: $y_t = a + \frac{b}{t}$;
- экспоненциальный тренд: $y_t = e^{a+bt}$ (или $y_t = a \cdot b^t$);
- степенная функция: $y_t = a \cdot t^b$;
- полиномы различных степеней: $y_t = a + b_1 \cdot t + b_2 \cdot t^2 + \dots + b_m \cdot t^m$.

Множество математических функций, используемых для построения уравнения тренда, можно свести в 3 группы:

- Функции с монотонным характером возрастания (убывания) и отсутствием пределов роста (снижения);
- Кривые с насыщением, т.е. устанавливается нижняя или верхняя граница изменения уровней ряда;
- S-образные кривые, т.е. кривые с насыщением, имеющие точку перегиба.

В первую группу функций входят полиномы k-й степени:

$$y_t = a + b_1 \cdot t + b_2 \cdot t^2 + \dots + b_m \cdot t^m \quad (1.4)$$

При $k=1$ получают линейный тренд. При нем уровни динамического ряда изменяются равным абсолютным приростом (параметр b).

При $k=2$ получают параболу второй степени. Данная функция рекомендуется для моделирования тенденции, если временной ряд характеризуется постоянным абсолютным ускорением, т.е. постоянными

являются вторые разности (приросты абсолютных приростов). В этом случае скорость ряда изменяется линейно. Параметр a означает начальный уровень ряда динамики при $t=0$. Параметр b_1 представляет собой средний абсолютный прирост за весь период времени, если t обозначено так, что $\sum t=0$. Параметр b_2 характеризует половину абсолютного ускорения динамического ряда.

К кривым с насыщением можно отнести гиперболы вида:

$$y = a + \frac{b}{t}; \quad (1.5)$$

$$y = a + \frac{b}{c+t}. \quad (1.6)$$

Равносторонняя гипербола ($y = a + \frac{b}{t}$) при $b>0$ означает, что уровни ряда снижаются во времени и асимптотически приближаются к параметру a .

Если $b<0$, то уравнение тренда $y = a + \frac{b}{t}$ характеризует тенденцию к росту уровней ряда с асимптотической границей, равной параметру a .

К числу S-образных кривых можно отнести логистическую кривую и кривую Гомперца.

Логистическую кривую вида

$$y = \frac{1}{c + ab^t} \quad (1.7)$$

называют кривой Перла-Рида. В ней верхняя асимптота составит величину $\frac{1}{c}$.

Точка перегиба у этой кривой равна $t_p = \frac{1}{\ln b} \ln \frac{c}{a}$. Значение y в точке перегиба равно $\frac{1}{2c}$.

Чаще применяется логистическая кривая вида

$$y = \frac{c}{1 + be^{-at}}, \quad (1.8)$$

где c – верхняя асимптота;

b и a – параметры функции;

Кривая Гомперца

$$y = ca^{bt} \quad (1.9)$$

Верхняя асимптота соответствует значению параметра c , а нижняя равна 0, если $\ln a < 0$.

Если $\ln a > 0$, то кривая имеет нижнюю асимптоту, равную величине параметра c .

Точкой перегиба данной кривой будет $t_p = \frac{1}{\ln b} \ln\left(-\frac{1}{\ln a}\right)$ со значением функции $y_{t_p} = \frac{c}{e}$, где e – основание натурального логарифма.

Наиболее простую экономическую интерпретацию имеет линейная функция

$$y_t = a + bt :$$

a – начальный уровень временного ряда в момент времени $t = 0$; b – средний за период абсолютный прирост уровней ряда.

Параметры a и b находятся по формулам:

$$a = \frac{\sum_{i=1}^n y_i - b \sum_{i=1}^n t_i}{n} ; \quad b = \frac{n \sum_{i=1}^n t_i y_i - \sum_{i=1}^n t_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n t_i^2 - \left(\sum_{i=1}^n t_i\right)^2} \quad (1.10)$$

Существует несколько способов определения типа тенденции: качественный анализ изучаемого процесса, построение и визуальный анализ графика зависимости уровней ряда от времени, использование автокорреляционной функции.

Тип тенденции можно определить путем сравнения коэффициентов автокорреляции первого порядка, рассчитанных по исходным и преобразованным уровням ряда. Если временной ряд имеет линейную тенденцию, то его соседние уровни y_t и y_{t-1} тесно коррелируют. В этом случае коэффициент автокорреляции первого порядка уровней исходного ряда должен быть высоким. Если временной ряд содержит нелинейную тенденцию, например, в форме экспоненты, то коэффициент автокорреляции первого порядка по логарифмам уровней исходного ряда будет выше, чем соответствующий коэффициент, рассчитанный по уровням ряда. Чем сильнее выражена нелинейная тенденция в изучаемом временном ряде, тем в большей степени будут различаться значения указанных коэффициентов.

При наличии во временном ряде тенденции и циклических колебаний значения каждого последующего уровня ряда зависят от предыдущих. Корреляционную зависимость между последовательными уровнями временного ряда количественно можно измерить с помощью линейного коэффициента корреляции между уровнями исходного временного ряда и уровнями этого же ряда, сдвинутыми на несколько шагов во времени.

Формула для расчета коэффициента автокорреляции первого порядка между соседними уровнями ряда y_t и y_{t-1} имеет вид:

$$r_1 = \frac{\sum_{t=2}^n (y_t - \bar{y}_1)(y_{t-1} - \bar{y}_2)}{\sqrt{\sum_{t=2}^n (y_t - \bar{y}_1)^2 \sum_{t=2}^n (y_{t-1} - \bar{y}_2)^2}}, \quad (1.11)$$

где

$$\bar{y}_1 = \frac{1}{n-1} \sum_{t=2}^n y_t, \quad \bar{y}_2 = \frac{1}{n-1} \sum_{t=2}^n y_{t-1}. \quad (1.12)$$

Аналогично можно определить коэффициенты автокорреляции второго и более высоких порядков.

По величине коэффициентов корреляции можно сделать вывод о структуре ряда. Если наиболее высоким оказался коэффициент автокорреляции первого порядка, исследуемый ряд содержит только тенденцию. Если наиболее высоким оказался коэффициент автокорреляции порядка τ , то ряд содержит циклические колебания с периодичностью в τ моментов времени. Если ни один из коэффициентов автокорреляции не является значимым, можно сделать одно из двух предположений относительно структуры этого ряда: либо ряд не содержит тенденции и циклических колебаний, либо ряд содержит сильную нелинейную тенденцию, для выявления которой нужно провести дополнительный анализ.

Последовательность коэффициентов автокорреляции уровней первого, второго и т.д. порядков называют *автокорреляционной функцией* (АКФ) временного ряда.

Для стационарного временного ряда с увеличением величины лага взаимосвязь y_t и $y_{t-\tau}$ ослабевает, и АКФ характеризуется убыванием, что графически представляется затухающей кривой.

АКФ дает представление о внутренней структуре динамического ряда. С помощью АКФ можно определить наличие или отсутствие в ряду динамики периодических колебаний и соответственно величину периода колебаний: она равна той величине лага τ , при которой коэффициент автокорреляции уровней наибольший.

Для динамического ряда с монотонной тенденцией к возрастанию (или уменьшению) уровней АКФ имеет значения, близкие к +1, которые медленно снижаются с увеличением величины лага.

Если ряд характеризуется сменой тенденций, то АКФ примет значения, стремительно уменьшающиеся с увеличением величины лага и сопровождаемые иногда сменой знака коэффициента автокорреляции.

Выбор наилучшего уравнения в случае, когда ряд содержит нелинейную тенденцию, можно осуществить путем перебора основных форм

тренда, расчета по каждому уравнению скорректированного коэффициента детерминации и средней ошибки аппроксимации.

Параметры каждого из перечисленных выше трендов можно определить обычным МНК, используя в качестве независимой переменной время $t = 1, 2, \dots, n$, а в качестве зависимой переменной – фактические уровни временного ряда y_t . Для нелинейных трендов предварительно проводят стандартную процедуру их линеаризации.

Некоторую специфику имеет оценка параметров кривых с насыщением. По этим функциям должна быть сначала определена асимптота. Если она может быть задана исследователем на основе анализа временного ряда, то другие параметры могут быть оценены по МНК.

2. Прогнозирование на основе временных рядов

2.1 Прогнозирование на основе кривых роста

Прогнозирование кривых роста (кривых насыщения) для прогнозирования требует соблюдения определенных условий.

1. Исходный временной ряд должен быть очень длинным (30-40 лет).
2. Исходный временной ряд не должен иметь скачков, и тенденция такого ряда должна описываться достаточно плавной кривой.
3. Использование кривых роста в прогнозировании социально-экономических явлений может давать достаточно хорошие результаты, если предел насыщения будет определен сравнительно точно.

Особенностью кривых роста является то, что абсолютные приращения уменьшаются по мере приближения к пределу. Однако процесс роста идет до конца.

Наиболее распространенными кривыми роста, используемыми в статистической практике прогнозирования, являются кривая Гомперца и кривая Перла-Рида.

Особенностью уравнений этих кривых является то, что их параметры могут быть определены методом наименьших квадратов лишь приближенно. Поэтому для расчета этих кривых используется методы, основанные на разбиении исходного ряда динамики на отдельные группы.

Например, для того чтобы осуществить прогноз на основе кривой Гомперца, необходимо вычислить значения параметров по следующим формулам:

$$c^n = \frac{d_2}{d_1}, \quad (2.1)$$

где n – число уровней ряда в каждой части; $d_1 = S_2 - S_1$; $d_2 = S_3 - S_2$; S_i – сумма элементов i -той части временного ряда, $i=1,2,3$

$$\lg b = \frac{d_1(c-1)}{(c^n - 1)^2}, \quad (2.2)$$

$$\lg a = \frac{1}{n} \left(S_1 - \frac{d_1}{c^n - 1} \right) \quad (2.3)$$

Чтобы использовать данную кривую для экстраполяции за пределы исходного ряда динамики, достаточно подставить соответствующее значение x_t в уравнение кривой.

Наряду с кривой Гомперца достаточно широкое распространение получила также кривая Перла-Рида.

Эта кривая выражает модифицированную геометрическую прогрессию, в которой возрастание затухает по мере приближения к некоторому определенному пределу. Максимальный предел устанавливается, прежде всего, на основании конкретного изучения исследуемого социально-экономического явления.

Кривая Перла-Рида описывается уравнением:

$$\frac{1}{y} = a + bc^x \quad (2.4)$$

Параметры уравнения находятся следующим образом:

$$c^n = \frac{d_2}{d_1}; \quad b = \frac{d_1(c-1)}{(c^n - 1)^2}; \quad a = \frac{1}{n} \left(S_1 - \frac{d_1}{c^n - 1} \right) \quad (2.5)$$

Параметры уравнения кривой Перла-Рида определяются так же, как и параметры кривой Гомперца, за исключением того, что в последнем случае не используется прием логарифмирования. Кроме того, нужно иметь в виду,

что в зависимости от масштаба данных величина $\frac{1}{y}$ умножается на 10000, 100000 или 1000000.

2.2 Прогнозирование на основе адаптивных моделей

Адаптивные методы могут применяться для прогнозирования показателей фондового рынка, денежных потоков, изменений ежедневных остатков на складах, в инструментальных кладовых, магазинах. С помощью этих же методов удастся описать эволюцию изменения технико-экономических характеристик изделий и переменных параметров химических процессов, изучить поведение показателя частоты отказов оборудования в зависимости от его возраста, а также при анализе сезонных явлений. В ряде случаев эти методы могут применяться для прогнозирования макропоказателей. Методы адаптивного прогнозирования применяются там, где основной информацией для прогноза являются временные ряды.

Инструментом прогноза при адаптивном методе служит модель. Первоначальная оценка параметров этой модели основывается на данных базового (исходного) временного ряда. На основе новых данных, получаемых на каждом следующем шаге, происходит корректировка параметров модели во времени, их адаптация к новым, непрерывно изменяющимся условиям развития явления. Таким образом, модель постоянно «впитывает» новую информацию и приспосабливается к ней.

Адаптивные модели изолированных рядов при всей их простоте могут давать более надежные результаты, чем сложные эконометрические системы уравнений. В то же время адаптивные методы пригодны лишь для обработки рядов с умеренными изменениями во времени и при построении краткосрочных прогнозов. Под краткосрочным прогнозом обычно

подразумевается прогноз на один (либо на несколько) интервал времени. Сам интервал может быть любым.

Адаптивные модели способны отражать изменяющиеся во времени условия, учитывать информационную ценность различных членов временной последовательности и давать достаточно точные оценки будущих членов данного ряда.

Реальный анализируемый процесс протекает в изменяющихся условиях, составляющих его внешнюю среду. Он вынужден адаптироваться к внешней среде. А модель адаптируется к ряду, представляющему этот процесс.

При построении адаптивных моделей используются в основном дискретные временные ряды, в которых наблюдения делаются через фиксированный интервал времени, принимаемый за единицу счета.

При этом временной ряд может быть детерминированным, если значения членов временного ряда точно определены какой-либо математической функцией, либо случайным, если значения могут быть описаны только с помощью распределения вероятностей.

Для описания временных рядов используются математические модели. Представим, что временной ряд x_t , генерируемый некоторой моделью, можно представить в виде двух компонент

$$x_t = \xi_t + \varepsilon_t, \quad (2.6)$$

где величина ε_t генерируется случайным неавтокоррелированным процессом с нулевым математическим ожиданием и конечной (не обязательно постоянной) дисперсией, а величина ξ_t может быть генерирована либо детерминированной функцией, либо случайным процессом, либо какой-нибудь их комбинацией. Величины ε_t и ξ_t различаются характером воздействия на значения последующих членов ряда. Переменная ε_t влияет только на значение синхронного ей члена ряда, в то время как величина ξ_t в известной степени определяет значение нескольких или всех последующих

членов ряда. Через величину ζ_t осуществляется взаимодействие членов ряда. Таким образом, в ней содержится информация, необходимая для получения прогнозов.

Величину ζ_t принято называть уровнем ряда в момент t , а закон эволюции уровня во времени — трендом. Тренд может быть выражен как детерминированной, так и случайной функциями, либо их комбинацией. Стохастические тренды имеют, например, ряды со случайным уровнем или случайным скачкообразным характером роста.

Компоненты временного ряда ζ_t и ε_t ненаблюдаемы. Они являются теоретическими величинами. Их выделение и составляет предмет анализа временного ряда в задаче прогнозирования. Оценку будущих членов ряда обычно делают по прогнозной модели. Прогнозная модель - это модель, аппроксимирующая тренд. Прогнозы — это оценки будущих уровней ряда, а последовательность прогнозов для различных периодов упреждения $\tau = 1, 2, \dots, k$ составляет оценку тренда.

При построении прогнозной модели выдвигается гипотеза о динамике величины ζ_t , т. е. о характере тренда.

Однако в связи с тем, что уверенность в гипотезе всегда относительна, рассматриваемые модели наделяются адаптивными свойствами, способностью к корректировке исходной гипотезы или даже к замене ее другой, более адекватно (с точки зрения точности прогнозов) отражающей поведение реального ряда.

Простейшая адаптивная модель основывается на вычислении так называемой экспоненциальной средней.

2.2.1 Экспоненциальное сглаживание

Предположим, что исследуется временной ряд x_t .

Выявление и анализ тенденции динамического ряда часто производится с помощью его выравнивания или сглаживания. Экспоненциальное сглаживание — один из простейших и распространенных приемов выравнивания ряда. В его основе лежит расчет экспоненциальных средних.

Экспоненциальное сглаживание ряда осуществляется по рекуррентной формуле

$$S_t = \alpha x_t + \beta S_{t-1}, \quad (2.7)$$

где S_t — значение экспоненциальной средней в момент t ;

α — параметр сглаживания, $\alpha = \text{const}$, $0 < \alpha < 1$; $\beta = 1 - \alpha$.

Экспоненциальную среднюю S_t можно выразить через значения временного ряда x :

$$\begin{aligned} S_t &= \alpha x_t + \beta S_{t-1} = \alpha x_t + \alpha \beta x_{t-1} + \beta^2 S_{t-2} = \dots = \\ &= \alpha x_t + \alpha \beta x_{t-1} + \alpha \beta^2 x_{t-2} + \dots + \alpha \beta^l x_{t-l} + \dots + \beta^N S_0 = \\ &= \alpha \sum_{l=0}^{N-1} \beta^l x_{t-l} + \beta^N S_0, \end{aligned} \quad (2.8)$$

где N — количество членов ряда;

S_0 — некоторая величина, характеризующая начальные условия для первого применения формулы (3.1) при $t = 1$.

Так как $\beta < 1$, то при $N \rightarrow \infty$ $\beta^N \rightarrow 0$, а сумма коэффициентов

$$\alpha \sum_{l=0}^{N-1} \beta^l \rightarrow 1. \text{ Тогда } S_i = \alpha \sum_{l=0}^{\infty} \beta^l x_{t-l}.$$

Это означает, что величина S_t оказывается взвешенной суммой всех членов ряда. Причем веса падают экспоненциально в зависимости от давности («возраста») наблюдения. Если, например, $\alpha = 0,3$, то текущее

наблюдение будет иметь вес 0,3, а веса предшествующих данных составят соответственно 0,21; 0,147; 0,1029 и т. д.

Рассмотрим, ряд, генерированный моделью

$$x_t = a_1 + \varepsilon_t, \quad (2.9)$$

где a_1 — const;

ε_t — случайные неавтокоррелированные отклонения, или шум, со средним значением 0 и дисперсией σ^2 .

Применим к нему процедуру экспоненциального сглаживания. Тогда

$$S_t = \alpha \sum_{l=0}^{\infty} \beta^l x_{t-l} = \alpha \sum_{l=0}^{\infty} \beta^l (a_1 + \varepsilon_{t-l}) = a_1 + \alpha \sum_{l=0}^{\infty} \beta^l \varepsilon_{t-l}, \quad (2.10)$$

Найдем математическое ожидание

$$M(S_t) = M(x_t) = a_1. \quad (2.11)$$

и дисперсию

$$D(S_t) = M[(S_t - a_1)^2] = M\left[\left(\alpha \sum_{l=0}^{\infty} \beta^l \varepsilon_{t-l}\right)^2\right] = \alpha^2 \sum_{l=0}^{\infty} \beta^{2l} \sigma^2 = \frac{\alpha}{2-\alpha} \sigma^2. \quad (2.12)$$

Так как $0 < \alpha < 1$, $D(S_t) < D(x_t) = \sigma^2$.

Таким образом, экспоненциальная средняя S_t имеет то же математическое ожидание, что и ряд x , но меньшую дисперсию. Как видно из (3.4), при высоком значении α дисперсия экспоненциальной средней незначительно отличается от дисперсии ряда x . Чем меньше α , тем в большей степени сокращается дисперсия экспоненциальной средней.

Следовательно, экспоненциальное сглаживание можно представить, как фильтр, на вход которого в виде потока последовательно поступают члены исходного ряда, а на выходе формируются текущие значения экспоненциальной средней.

И чем меньше α , тем в большей степени фильтруются, подавляются колебания исходного ряда.

Экспоненциальная средняя часто используется для краткосрочного прогнозирования. В этом случае предполагается, что ряд генерируется моделью

$$x_t = a_{1,t} + \varepsilon_t, \quad (2.13)$$

где $a_{1,t}$ — варьирующий во времени средний уровень ряда;

ε_t — случайные неавтокоррелированные отклонения с нулевым математическим ожиданием и дисперсией σ^2 .

Прогнозная модель имеет вид

$$\hat{x}_\tau(t) = \hat{a}_{1,t}, \quad (2.14)$$

где $\hat{x}_\tau(t)$ — прогноз, сделанный в момент t на τ единиц времени (шагов) вперед;

$\hat{a}_{1,t}$ — оценка $a_{1,t}$.

Средством оценки единственного параметра модели служит экспоненциальная средняя $\hat{a}_{1,t} = S_t$. Таким образом, все свойства экспоненциальной средней распространяются на прогнозную модель. В частности, если S_{t-1} рассматривать как прогноз на 1 шаг вперед, то в выражении (2) величина $(x_t - S_{t-1})$ есть погрешность этого прогноза, а новый прогноз S_t получается в результате корректировки предыдущего прогноза с учетом его ошибки. В этом и состоит существо адаптации.

При краткосрочном прогнозировании желательно как можно быстрее отразить изменения $a_{1,t}$ и в то же время как можно лучше «очистить» ряд от случайных колебаний.

Таким образом, с одной стороны, следует увеличивать вес более свежих наблюдений, что может быть достигнуто повышением α , с другой

стороны, для сглаживания случайных отклонений величину α нужно уменьшить.

Эти два требования находятся в противоречии. Поиск компромиссного значения α составляет задачу оптимизации модели.

Экспоненциальное выравнивание всегда требует предыдущего значения экспоненциальной средней. Когда процесс только начинается, должна быть некоторая величина S_0 , которая может быть использована в качестве значения, предшествующего S_1 . Если есть прошлые данные к моменту начала выравнивания, то в качестве начального значения S_0 можно использовать арифметическую среднюю всех имеющихся точек или какой-то их части. Когда для такого оценивания S_0 нет данных, требуется предсказание начального уровня ряда.

Предсказание может быть сделано исходя из априорных знаний о процессе или на основе его аналогии с другими процессами. После k шагов вес, придаваемый начальному значению, равен $(1 - \alpha)^k$. Если есть уверенность в справедливости начального значения S_0 , то можно коэффициент α взять малым. Если такой уверенности нет, то параметру α следует дать большое значение, с таким расчетом, чтобы влияние начального значения быстро уменьшилось. Однако большое значение α может явиться причиной большой дисперсии колебаний S_t . Если требуется подавление этих колебаний, то после достаточного удаления от начального момента времени величину α можно убавить.

2.2.2 Модели линейного роста

Экспоненциальная средняя дает систематическую ошибку, когда временной ряд имеет тенденцию линейного роста. Для этого случая разработано несколько вариантов адаптивных моделей, также использующих

процедуру экспоненциального сглаживания. В основе моделей лежит гипотеза о том, что прогноз может быть получен по уравнению

$$\widehat{x}_\tau(t) = \widehat{a}_{1,t} + \tau \widehat{a}_{2,t}, \quad (2.15)$$

где $\widehat{a}_{1,t}, \widehat{a}_{2,t}$ — текущие оценки коэффициентов адаптивного полинома первого порядка.

Одной из первых моделей этого типа была двухпараметрическая модель Ч. Хольта, в которой оценка коэффициентов производится следующим образом:

$$\begin{aligned} \widehat{a}_{1,t} &= \alpha_1 x_t + (1 - \alpha_1)(\widehat{a}_{1,t-1} + \widehat{a}_{2,t-1}), \\ \widehat{a}_{2,t} &= \alpha_2 (\widehat{a}_{1,t} - \widehat{a}_{1,t-1}) + (1 - \alpha_2)\widehat{a}_{2,t-1}, \end{aligned} \quad (2.16)$$

где α_1 и α_2 — параметры экспоненциального сглаживания ($0 < \alpha_1, \alpha_2 < 1$), которые также называют параметрами адаптации.

Эти уравнения могут быть переписаны в виде:

$$\begin{aligned} \widehat{a}_{1,t} &= \widehat{a}_{1,t-1} + \widehat{a}_{2,t-1} + \alpha_1 e_t; \\ \widehat{a}_{2,t} &= \widehat{a}_{2,t-1} + \alpha_1 \alpha_2 e_t, \end{aligned} \quad (2.17)$$

где $e_t = x_t - \widehat{x}_1(t-1)$ — ошибка прогноза.

Частным случаем модели Хольта является модель линейного роста Брауна:

$$\begin{aligned} \widehat{a}_{1,t} &= \widehat{a}_{1,t-1} + \widehat{a}_{2,t-1} + (1 - \beta^2)e_t; \\ \widehat{a}_{2,t} &= \widehat{a}_{2,t-1} + (1 - \beta)^2 e_t, \end{aligned} \quad (2.18)$$

где параметр β — коэффициент дисконтирования, характеризующий обесценение данных наблюдения за единицу времени, $0 < \beta < 1$.

Если модель Хольта усовершенствовать путем включения разности ошибок, то получим полную трехпараметрическую модель прогнозирования Дж. Бокса и Г. Дженкинса:

$$\begin{aligned}\widehat{x}_\tau(t) &= \widehat{a}_{1,t} + \tau \widehat{a}_{2,t}, \\ \widehat{a}_{1,t} &= \alpha_1 x_t + (1 - \alpha_1)(\widehat{a}_{1,t-1} + \widehat{a}_{2,t-1}) + \alpha_3(e_t - e_{t-1}); \\ \widehat{a}_{2,t} &= \alpha_2(\widehat{a}_{1,t} - \widehat{a}_{1,t-1}) + (1 - \alpha_2)\widehat{a}_{2,t-1},\end{aligned}\tag{2.19}$$

где $\alpha_1, \alpha_2, \alpha_3$ являются параметрами модели ($0 < \alpha_1, \alpha_2, \alpha_3 < 1$); $e_t = x_t - \widehat{x}_1(t-1)$ — ошибка прогнозирования.

На основе практических испытаний модели на многих экономических рядах Бокс и Дженкинс пришли к выводу, что включение в модель разности ошибок не является необходимым. Коэффициент α_3 всегда оказывался близким к нулю. Это объясняется стохастическим характером данных, и, в частности, тем, что корреляция ошибок в подобных случаях неустойчива.

К положительным чертам метода Брауна можно отнести следующие: логичная, ясная и легко понимаемая концепция; оптимальное значение единственного параметра можно быстро найти эмпирическим путем; коэффициенты модели прогнозирования оцениваются совместно таким образом, чтобы уменьшить автокорреляцию в остатках. Все это делает модель Брауна легко применимой.

2.2.3 Стохастический процесс Тейла И Вейджа

Г. Тейл и С. Вейдж в целях дальнейшего изучения свойств адаптивных моделей предложили применить двухпараметрический предиктор Хольта для прогнозирования некоторого вероятностного процесса, характеризующегося стохастическим трендом. Они вывели выражения для определения оптимальных параметров адаптации, минимизирующих средний квадрат ошибки прогнозирования.

Процесс Тейла—Вейджа аналитически записывается как:

$$\begin{aligned}
x_t &= a_{1,t} + \varepsilon_t; \\
a_{1,t} &= a_{1,t-1} + a_{2,t}; \\
a_{2,t} &= a_{2,t-1} + v_t,
\end{aligned}
\tag{2.20}$$

где $a_{1,t}$ — значение уровня исследуемого временного ряда x_t в момент t ;
 $a_{2,t}$ — прирост уровня от момента $t-1$ к моменту t
 ε_t , v_t — временные последовательности с нулевым математическим ожиданием, постоянными дисперсиями и отсутствием ковариации, т. е.

$$\begin{aligned}
M(\varepsilon_t) &= M(v_t) = 0; \\
M(\varepsilon_t \varepsilon_{t'}) &= \begin{cases} \sigma_\varepsilon^2 & t = t'; \\ 0 & t \neq t'; \end{cases} \\
M(v_t v_{t'}) &= \begin{cases} \sigma_v^2 & t = t'; \\ 0 & t \neq t'; \end{cases}
\end{aligned}
\tag{2.21}$$

$$M(v_t \varepsilon_{t'}) = 0 \text{ для любой пары } (t, t').$$

Временной ряд x_t не является стационарным и не имеет строго определенной автоковариационной функции. Вторые же разности этого ряда имеют вполне определенную автоковариационную функцию

Схема составления прогноза выглядит следующим образом:

$$\hat{a}_{1,t} = \alpha_1 x_t + (1 - \alpha_1)(\hat{a}_{1,t-1} + \hat{a}_{2,t-1}); \tag{2.22}$$

$$\hat{a}_{2,t} = \alpha_2 (\hat{a}_{1,t} - \hat{a}_{1,t-1}) + (1 - \alpha_2) \hat{a}_{2,t-1}; \tag{2.23}$$

$$\hat{x}_t(t) = \hat{a}_{1,t} + \tau \hat{a}_{2,t};$$

$$0 < \alpha_1, \alpha_2 < 1.$$

Если ошибку прогноза, сделанного в момент t на 1 шаг вперед, обозначить через $e_1(t)$, то уравнения адаптации (2.22) и (2.23) можно записать в виде:

$$\begin{aligned}
\hat{a}_{1,t} &= \hat{a}_{1,t-1} + \hat{a}_{2,t-1} + \alpha_1 e_1(t-1); \\
\hat{a}_{2,t} &= \hat{a}_{2,t-1} + \gamma e_1(t-1);
\end{aligned}
\tag{2.24}$$

$$\gamma = \alpha_1 \alpha_2.$$

Ошибка прогноза:

$$\begin{aligned} e_1(t) &= x_{t+1} - \hat{a}_{1,t} - \hat{a}_{2,t} = (a_{1,t} + a_{2,t} + v_{t+1} + \varepsilon_{t+1}) - \hat{a}_{1,t} - \hat{a}_{2,t} = \\ &= (a_{1,t} - \hat{a}_{1,t}) + (a_{2,t} - \hat{a}_{2,t}) + v_{t+1} + \varepsilon_{t+1} \end{aligned} \quad (2.25)$$

Следовательно, ошибка прогноза является суммой трех компонент: ошибки оценки уровня процесса в момент t , ошибки оценки прироста уровня в момент t и комбинации случайных компонент v и ε в момент $t + 1$.

2.2.4 Сезонные модели

В экономике многие явления характеризуются периодически повторяющимися сезонными эффектами. Соответственно временные ряды, их отражающие, содержат периодические сезонные колебания. Эти ряды и их колебания можно представить, как генерируемые моделями двух основных типов: моделями с мультипликативными и с аддитивными коэффициентами сезонности.

Модели первого типа имеют вид:

$$\begin{aligned} x_t &= \xi_t + \varepsilon_t; \\ \xi_t &= a_{1,t} f_t, \end{aligned} \quad (2.26)$$

где динамика величины $a_{1,t}$ характеризует тенденцию развития процесса;

f_t, f_{t-1}, f_{t-l+1} - коэффициенты сезонности;

l — количество фаз в полном сезонном цикле (если ряд представляет месячные наблюдения, то в экономике обычно $l = 12$, при квартальных данных $l = 4$ и т. п.);

ε_t — неавтокоррелированный шум с нулевым математическим ожиданием.

Модели второго типа записываются как:

$$\begin{aligned}x_t &= \xi_t + \varepsilon_t; \\ \xi_t &= a_{1,t} + g_t,\end{aligned}\tag{2.27}$$

где величина $a_{1,t}$ описывает тенденцию развития процесса;

g_t, g_{t-1}, g_{t-l+1} - аддитивные коэффициенты сезонности;

l — количество фаз в полном сезонном цикле;

ε_t — неавтокоррелированный шум с нулевым математическим ожиданием.

Адаптивная модель с мультипликативной сезонностью была предложена П.Р. Уинтерсом. Аддитивная модель рассмотрена Г. Тейлом и С. Вейджем.

2.2.5 Прогнозирование с коэффициентами сезонности

Модель имеет вид:

$$\begin{aligned}\hat{a}_{1,t} &= \alpha_1 \frac{x_t}{\hat{f}_{t-1}} + (1 - \alpha_1) \hat{a}_{1,t-1} \quad 0 < \alpha_1 < 1; \\ \hat{f}_t &= \alpha_2 \frac{x_t}{\hat{a}_{1,t}} + (1 - \alpha_2) \hat{f}_{t-1} \quad 0 < \alpha_2 < 1.\end{aligned}\tag{2.28}$$

Как видим, $\hat{a}_{1,t}$ является взвешенной суммой текущей оценки $\frac{x_t}{\hat{f}_{t-1}}$, полученной путем очищения от сезонных колебаний фактических данных x_t и предыдущей оценки $\hat{a}_{1,t-1}$.

В качестве коэффициента сезонности f_t берется его наиболее поздняя оценка, сделанная для аналогичной фазы цикла.

Затем величина $\hat{a}_{1,t}$, полученная по первому уравнению, используется для определения новой оценки коэффициента сезонности по второму уравнению.

Прогноз следующего значения ряда:

$$\widehat{x}_1(t) = \widehat{a}_{1,t} \widehat{f}_{t-l+1}. \quad (2.29)$$

Более общим выражением для прогноза на τ шагов вперед будет:

$$\widehat{x}_\tau(t) = \widehat{a}_{1,t} \widehat{f}_{t-l+\tau}, \quad \tau \leq l. \quad (2.30)$$

Величины $\widehat{a}_{1,t}$ и \widehat{f}_t могут быть записаны через прошлые данные и начальные условия:

$$\begin{aligned} \widehat{a}_{1,t} &= \alpha_1 \sum_{n=0}^t (1-\alpha_1)^n \frac{x_{t-n}}{\widehat{f}_{t-l-n}} + (1-\alpha_1)^{t+1} \widehat{a}_{1,0}; \\ \widehat{f}_t &= \alpha_2 \sum_{n=0}^J (1-\alpha_2)^n \frac{x_{t-nl}}{\widehat{a}_{1,t-nl}} + (1-\alpha_2)^{J+1} \widehat{f}_{i,0}, \end{aligned} \quad (2.31)$$

где $\widehat{a}_{1,0}$ — начальное значение a_1 ;

$\widehat{f}_{i,0}$ — начальное значение f в соответствующей i фазе (месяце) цикла (года);

J — наибольшая целая часть $\frac{t}{l}$.

Следовательно, прогноз является функцией всех прошлых значений фактического ряда, параметров α_1 и α_2 и начальных условий $\widehat{a}_{1,0}, \widehat{f}_{1,0}, \widehat{f}_{2,0}, \dots, \widehat{f}_{l,0}$.

Влияние начальных условий на прогноз зависит от величины весов и длины ряда, предшествующего текущему моменту t . Влияние $\widehat{a}_{1,0}$ обычно будет уменьшаться быстрее, чем влияние начальных значений $\widehat{f}_{i,0}$, так как \widehat{a}_1 пересматривается на каждом шаге, а \widehat{f}_i только один раз за цикл.

Если эта сезонная модель прогнозирования, структура которой не содержит элементов для отражения какой-либо тенденции роста, применяется для прогнозирования ряда, характеризующегося ярко выраженной тенденцией, то коэффициенты \widehat{f}_i перестают быть простыми

коэффициентами сезонности и вскоре вбирают в себя в определенной мере эффект роста.

Полная сезонная модель Уинтерса с линейным ростом имеет вид:

$$\begin{aligned}
 \hat{a}_{1,t} &= \alpha_1 \frac{x_t}{\hat{f}_{t-1}} + (1 - \alpha_1)(\hat{a}_{1,t-1} + \hat{a}_{2,t-1}); \\
 \hat{f}_t &= \alpha_2 \frac{x_t}{\hat{a}_{1,t}} + (1 - \alpha_2) \hat{f}_{t-1}; \\
 \hat{a}_{2,t} &= \alpha_3 (\hat{a}_{1,t} - \hat{a}_{1,t-1}) + (1 - \alpha_3) \hat{a}_{2,t-1} \quad 0 < \alpha_1, \alpha_2, \alpha_3 < 1; \\
 \hat{x}_\tau(t) &= (\hat{a}_{1,t} + \tau \hat{a}_{2,t}) \hat{f}_{t-l+\tau}.
 \end{aligned}
 \tag{2.32}$$

Единственным изменением в выражении для $\hat{a}_{1,t}$ является добавление $\hat{a}_{2,t-1}$ — наиболее поздней оценки аддитивного фактора роста, характеризующего изменение среднего за полный сезонный цикл уровня процесса за единицу времени (месяц). Выражение для обновления коэффициента сезонности остается тем же, что и раньше. Оценки $\hat{a}_{2,t}$ модифицируются по аналогичной процедуре экспоненциального сглаживания. Прогноз является здесь функцией прошлых и текущих данных, параметров $\alpha_1, \alpha_2, \alpha_3$ и первоначальных значений $\hat{a}_{1,0}, \hat{a}_{2,0}, \hat{f}_{i,0}$. Качество и точность прогнозов зависит от этих факторов.

Оптимальные параметры $\alpha_1, \alpha_2, \alpha_3$ Уинтерс предлагает находить экспериментальным путем. Критерием сравнения он берет стандартное отклонение ошибки.

2.2.6 Аддитивная модель сезонных явлений

Несмотря на то, что для экономических временных рядов мультипликативная модель обычно оказывается наиболее подходящей,

иногда требуется аддитивная модель. Рассмотрим аддитивную модель сезонных явлений с линейным ростом, предложенную Г. Тейлом и С. Вейджем.

Построение такой модели имеет целью упрощение процедуры прогнозирования, поскольку комбинация мультипликативной сезонной модели с линейным ростом математически громоздка. Кроме того, на практике чаще встречаются экспоненциальные тенденции, чем линейные. Поэтому замена значений первоначального временного ряда их логарифмами преобразует экспоненциальную тенденцию в линейную и одновременно мультипликативную сезонную модель в аддитивную. Тогда временной ряд (исходный или преобразованный) можно представить следующим образом:

$$\begin{aligned} x_t &= a_{1,t} + g_t + \varepsilon_t; \\ a_{1,t} &= a_{1,t-1} + a_{2,t}, \end{aligned} \tag{2.33}$$

где $a_{1,t}$ — величина уровня процесса после элиминирования сезонных колебаний;

$a_{2,t}$ — аддитивный коэффициент роста;

g_t — аддитивный коэффициент сезонности;

ε_t — белый шум.

Сначала рассмотрим адаптивную процедуру обновления значения $\hat{a}_{1,t}$.

В момент t мы располагаем наблюдением x_t , о котором известно, что

$$x_t = a_{1,t} + g_t + \varepsilon_t \tag{2.34}$$

Однако о шуме и сезонном факторе g_t никакой информации нет. Величину ε_t заменим нулем, а в качестве заменителя для g_t возьмем самую последнюю оценку сезонного фактора g_{t-1} , где l — период сезонного цикла. Величину $x_t - \hat{g}_{t-1}$ будем рассматривать как новое «фактическое» значение $a_{1,t}$. Последней оценкой уровня a_1 является $\hat{a}_{1,t-1}$, но она соответствует моменту $t-1$, а не t .

Поэтому необходимо к $\hat{a}_{1,t-1}$ добавить еще $\hat{a}_{2,t}$. Но так как оценку $\hat{a}_{2,t}$ мы еще не можем получить, то вместо нее берем оценку $\hat{a}_{2,t-1}$, полученную на предыдущем шаге.

Это приводит к следующей процедуре адаптации:

$$\hat{a}_{1,t} = \alpha_1 (x_t - \hat{g}_{t-1}) + (1 - \alpha_1)(\hat{a}_{1,t-1} + \hat{a}_{2,t-1}), \quad (2.35)$$

которая при данных весах α_1 и $(1 - \alpha_1)$ оценивает $a_{1,t}$ через наиболее свежее наблюдение x_t и ранее подсчитанные величины $\hat{a}_{1,t-1}$, $\hat{a}_{2,t-1}$, \hat{g}_{t-1} .

Та же процедура применяется для получения оценки g_t . Новое «фактическое» значение сезонного фактора будет $x_t - \hat{a}_{1,t}$, а старое значение равно \hat{g}_{t-1} , экспоненциально-сглаженное значение

$$\hat{g}_t = \alpha_3 (x_t - \hat{a}_{1,t}) + (1 - \alpha_3)\hat{g}_{t-1} \quad (2.36)$$

Все три параметра сглаживания будут удовлетворять условию $0 < \alpha_1, \alpha_2, \alpha_3 < 1$.

Адаптивное прогнозирование теперь провести сравнительно просто. Предположим, что t — текущий момент времени, так что $\hat{a}_{1,t}$, $\hat{a}_{2,t}$, \hat{g}_t , \hat{g}_{t-1} имеются в нашем распоряжении. Предположим также, что мы хотим получить прогноз величины $x_{t+\tau}$ (прогноз на τ шагов вперед). Экстраполируем тенденцию линейного роста, используя самое последнее значение коэффициента $\hat{a}_{2,t}$, добавляем самую свежую оценку сезонного члена для этой фазы цикла и пренебрегаем шумом. В результате получаем

$$\hat{x}_\tau(t) = \hat{a}_{1,t} + \tau \hat{a}_{2,t} + \hat{g}_{t-l+\tau} \quad (2.37)$$

при условии, что $0 < \tau < l$. Если $l < \tau < 2l$, то необходимо $\hat{g}_{t-l+\tau}$ заменить на $\hat{g}_{t-2l+\tau}$.

Однако на практике удобнее осуществлять адаптивное регулирование $\hat{a}_{1,t}$, $\hat{a}_{2,t}$ и \hat{g}_t с помощью уравнений, связывающих эти величины с ошибкой прогноза, сделанного в конце периода $t - 1$ на один шаг вперед.

2.2.7 Модели авторегрессии — скользящего среднего (метод Бокса — Дженкинса)

Для описания моделей потребуются следующие обозначения:

x_t — значение ряда в момент t ;

ε_t — белый шум с дисперсией σ_ε^2 .

Модель основывается на гипотезе, что изучаемый процесс является выходом линейного фильтра, на вход которого подан процесс белого шума, т. е. что член ряда x_t является взвешенной суммой текущего и предыдущих значений входного потока:

$$x_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots \quad (2.38)$$

где $\mu = \text{const}$ в общем случае является параметром, характеризующим процесс.

Если последовательность ψ_1, ψ_2, \dots конечна или бесконечна, но сходится, то процесс x_t будет стационарным. Тогда μ — среднее значение, вокруг которого процесс варьирует. В противном случае x_t — нестационарен и μ не имеет особого смысла, кроме как некой точки отсчета уровня процесса

3. Автоматизация построения моделей

3.1 Разработка программного средства

Программная система должна осуществлять:

- Проверку ряда на стационарность (критерий, основанный на выборочной медиане);
- Проведение алгоритмического сглаживания, на основе скользящего среднего или взвешенного скользящего среднего для удаления из ряда экстремальных значений и придания ему более гладкого вида;
- Определение структуры ряда на основе расчета коэффициентов автокорреляции;
- Моделирование сезонных колебаний
- Нахождение тренда в виде основных функций (оценка параметров);
- Выбор наиболее подходящей модели для описания эмпирического процесса посредством вычисления коэффициента детерминации, остаточной дисперсии и средней ошибки аппроксимации и их сравнения;
- Прогнозирование на основе построенной модели;
- Прогнозирование на основе кривых роста;
- Прогнозирование на основе адаптивных моделей временных рядов;
- Интерпретацию результатов, выводы и пояснения системы.

3.1.1 Функции системы

Система должна обеспечивать следующие функции:

1. Ввод информации, характеризующей исследуемый процесс.

- Значения независимых переменных (то есть массива времени);
- Значения зависимой переменной;

2. Анализ временного ряда, его структуры, подбор тенденции, исследование ряда на стационарность и наличие сезонных колебаний.

3. Вывод информации о параметрах системы:

- Вывод о наличии тренда;
- Вывод о структуре ряда;
- Параметры трендов;
- Вывод об оптимальном тренде;
- Вывод рассчитанных значений;
- Вывод прогнозных значений по построенным моделям;
- Вывод построенных прогнозов на основе кривых роста;
- Вывод построенных прогнозов на основе адаптивных моделей.

Входной информацией системы является:

1. Информация об исследуемой экономической модели: о значениях исследуемого показателя в соответствующие периоды времени.

Выходной информацией системы является:

1. Информация о наличии или отсутствии тренда.
2. Информация о структуре ряда.
3. Информация о типе выбранной модели.
4. Информация о наличии или отсутствии сезонных колебаний.
2. Информация о параметрах трендов.
3. Построенные прогнозы.
4. Таблицы с расчетными значениями зависимой переменной.

Приложение должно быть разработано в стиле мультимедиа (MDI). В качестве документа здесь используется анализируемая модель – файл Excel с входной информацией.

Доступ к файлам с входной информацией может осуществляться как через локальные диски, так и путем доступа к сетевому окружению. Для функционирования системы дополнительных компонентов (серверной части)

не требуется. Для минимизации трудозатрат на установку и обслуживание программного продукта его необходимо разработать таким образом, чтобы процесс установки и обновления был наиболее простым, не требовалось больших трудозатрат. Поэтому принято решение разработать программу в виде исполнимого модуля (*.exe) и файла конфигурации (*.ini). Для установки программы специальная инсталляция не требуется, установка должна осуществляться копированием вышеуказанных файлов.

3.1.2 Технические требования к системе.

В связи с отсутствием регулярной сетевой нагрузки при работе программы оптимизации сетевой нагрузки не требуется, поэтому в алгоритме работы программы обращение к серверу баз данных не предусмотрено.

Требования к рабочей станции: операционная система Windows XP, объем ОЗУ – не менее 1 ГБ, процессор Celeron 2000 ГГц, жесткий диск не менее 80 ГБ. Должна быть настроена доменная аутентификация пользователей, каждый пользователь должен иметь свою отдельную учетную запись в Active Directory. На клиентской машине должен быть установлен Microsoft Office XP или более новая версия.

Общая структура программного комплекса приведена на рис 3.1.

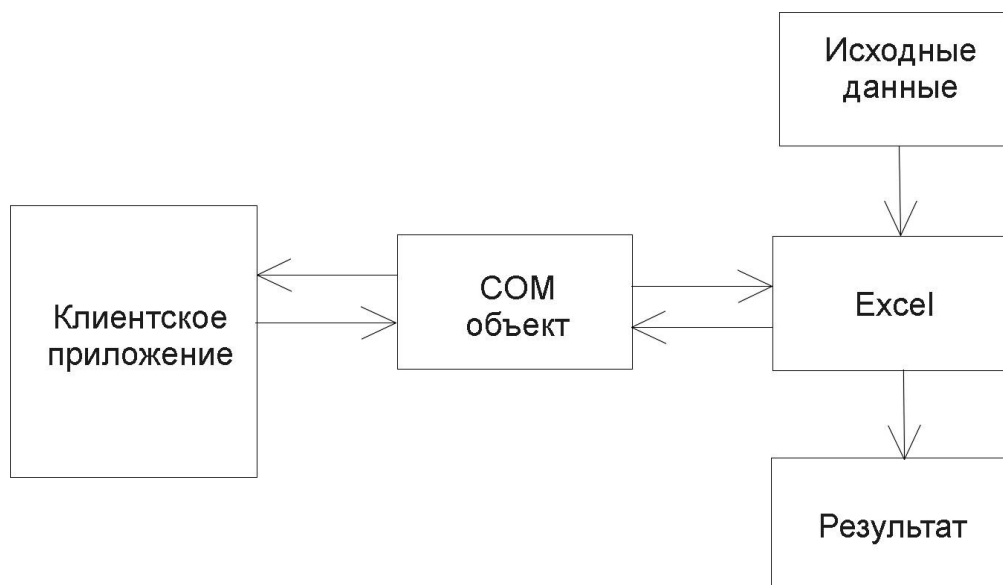


Рис 3.1 Общая структура программного комплекса

Модель приложения построена по следующему принципу:

- Интерфейс программы построен в стиле MDI-приложения.
- Чтение входных данных осуществляется через COM-объект Excel.Application.
- Расчет и анализ выполняется на стороне клиентского приложения.
- Значения расчетных параметров выводятся в текстовом виде с возможностью их копирования в буфер обмена Windows.
- Таблица результирующих данных экспортируется в Excel через COM-объект.
- Диаграмма экспортируется в Excel через COM-объект.

3.1.3 Описание программы.

Основная форма приложения имеет вид:

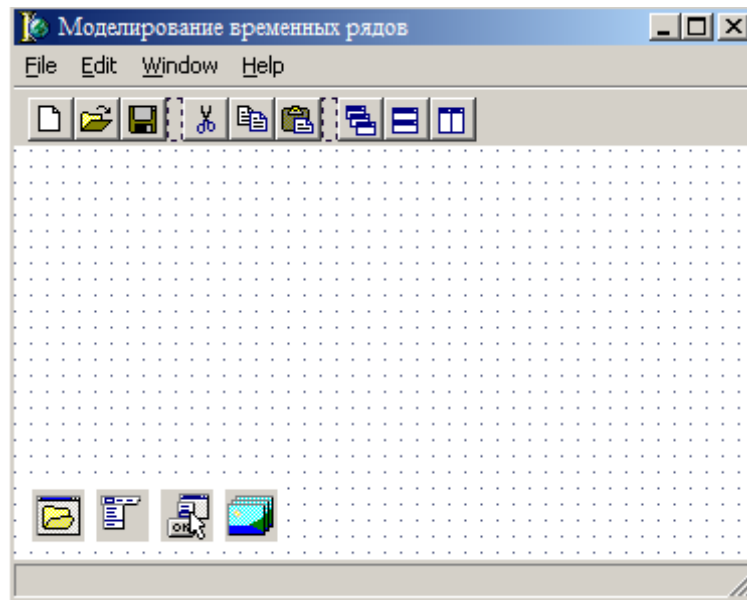


Рис 3.2 Основная форма MDI-приложения

Дочерняя форма приложения создается для анализа входных данных. На первом этапе для ее создания необходимо выбрать файл с входными данными формата Excel (рис 3.3):

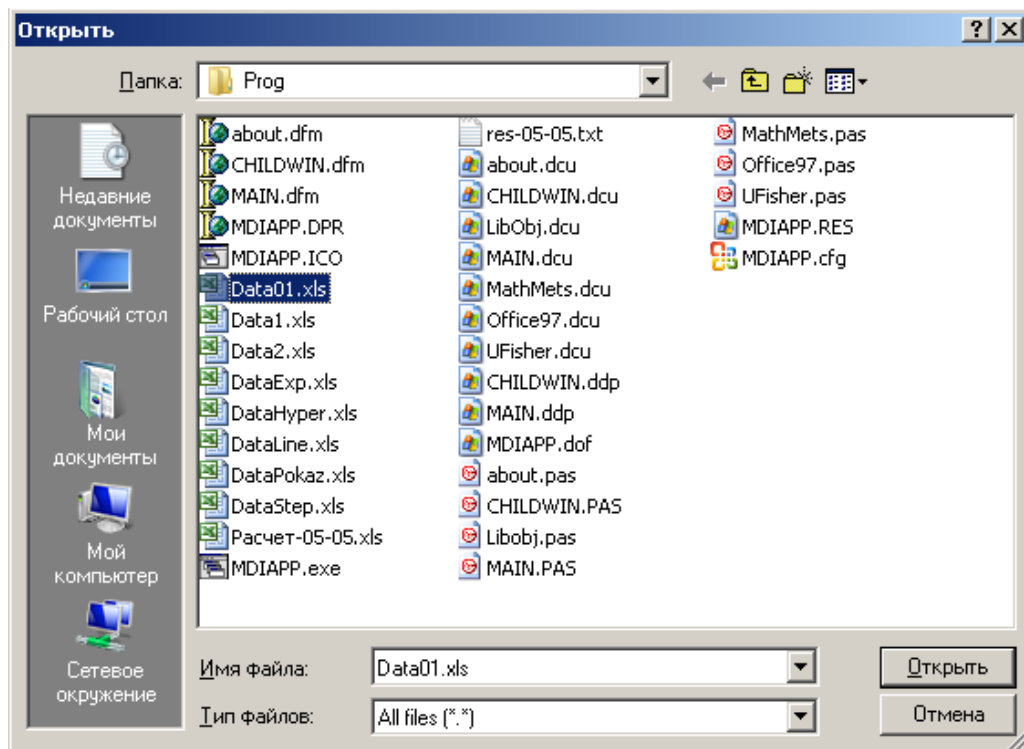


Рис 3.3 Выбор файла с входными данными

После этого будет создана форма с двумя закладками. На первой закладке будут отображены исходные данные (рис 3.4):

t	y(t)
1	1,83
2	2,23
3	2,73
4	3,34
5	4,08
6	4,97
7	6,08

Рис 3.4 Исходные данные после импорта из Excel

На второй закладке выводятся результаты расчета по всем вариантам моделей регрессии и осуществляется выбор оптимальной модели (рис 3.5):

```

Проверка ряда на стационарность: тренд есть
Коэффициент автокорреляции r1=0,685556
Коэффициент автокорреляции r2=0,313459
Коэффициент автокорреляции r3=0,086436
Коэффициент автокорреляции r4=0,279186
Коэффициент автокорреляции r5=0,128698
Коэффициент автокорреляции r6=-0,10727
Коэффициент автокорреляции r7=-0,06252
Коэффициент автокорреляции r8=-0,47922
Вывод: ряд имеет тренд (возможно, нелинейный), возможны сезонные колебания с периодом 8
Линейный тренд:  $y^{\wedge}=2,5511x+3369$ , R2=0,274, A=0,005, Sост=464,22, DW=0,754
Полином второго порядка:  $y^{\wedge}=-0,7307x^2+16,435x+3322,7$ , R2=0,7535, A=0,0025, Sост=157,62, DW=1,94
Полином третьего порядка:  $y^{\wedge}=0,0027x^3-1,4991x^2+22,433x+3312$ , R2=0,7667, A=0,0023, Sост=149,18, DW=2,018
Экспоненциальный тренд:  $y^{\wedge}=3368,8e0,0008x$ , R2=0,2753, A=0,005, Sост=467,57, DW=0,749
Степенной тренд:  $y^{\wedge}=3345,1x0,007$ , R2=0,5384, A=0,004, Sост=297,9, DW=1,09
Логарифмический тренд:  $y^{\wedge}=23,759\ln(x)+3345,2$ , R2=0,5361, A=0,004, Sост=296,63, DW=1,095
Наилучшие тренды:  $y1^{\wedge}=-0,7307x^2+16,435x+3322,7$ ;  $y2^{\wedge}=0,0027x^3-1,4991x^2+22,433x+3312$ 
Прогноз: (t=19):  $y1^{\wedge}(19)= 3371,1823$ ;  $y2^{\wedge}(19)= 3382,2449$ 
Прогнозирование на основе метода Брауна  $y^{\wedge}t+L=3380,96+4,2115t$ ;  $y^{\wedge}t+L=4058,43-5753,84t-1/2*53,3475t^2$ 

```

Рис 3.5 Результаты расчета модели

После этого осуществляется запуск Excel через COM-объект и экспорт в Excel таблицы с расчетными данными (рис 3.6):

The screenshot shows a Microsoft Excel window titled "Microsoft Excel - Лист1". The menu bar includes "Файл", "Правка", "Вид", "Вставка", "Формат", "Сервис", "Данные", "Окно", and "Справка". The toolbar contains various icons for file operations and calculations. The active cell is A1, containing the letter 't'. The formula bar shows "=X". The spreadsheet contains the following data:

	A	B	C	D	E	F	G	H	I
1	t	У-Исх.дан	Линейная	Гипербола	Степенная	Показатель	Экспоненциальная		
2	1	1,83	1,510714	1,14949	1,590025	1,828922	1,828922		
3	2	2,23	2,21	3,102407	2,423737	2,234385	2,234385		
4	3	2,73	2,909286	3,75338	3,101568	2,729737	2,729737		
5	4	3,34	3,608571	4,078866	3,694596	3,334906	3,334906		
6	5	4,08	4,307857	4,274158	4,231614	4,074238	4,074238		
7	6	4,97	5,007143	4,404352	4,72784	4,977476	4,977476		
8	7	6,08	5,706429	4,497348	5,192526	6,080957	6,080957		
9									
10									
11									
12									
13									
14									

The status bar at the bottom shows "Готово" and "NUM".

Рис 3.6 Экспорт в Excel таблицы с расчетными данными

Формирование отчетов: экспорт в Excel путем запуска этого приложения с передачей в него информации по технологии COM.

Под входной информацией понимается вся информация, необходимая для решения задачи и расположенная на различных носителях: первичных документах, машинных носителях, в памяти персонального компьютера.

Входной информацией для системы является:

- Значения независимых переменных анализируемого процесса.
- Значение зависимой переменной. Может быть получена в файле Excel.

В ходе разработки системы результатными показателями являются:

- информация о расчетном значении зависимой переменной;
- информация о выбранном оптимальном методе;
- диаграмма с расчетными значениями зависимой переменной.

В результате расчета получаем значение параметров возможных

моделей и выбираем оптимальную модель.

3.1.4 Результат работы системы

На рис 3.7 представлен результат построения исходного временного ряда и проверки ряда на стационарность.

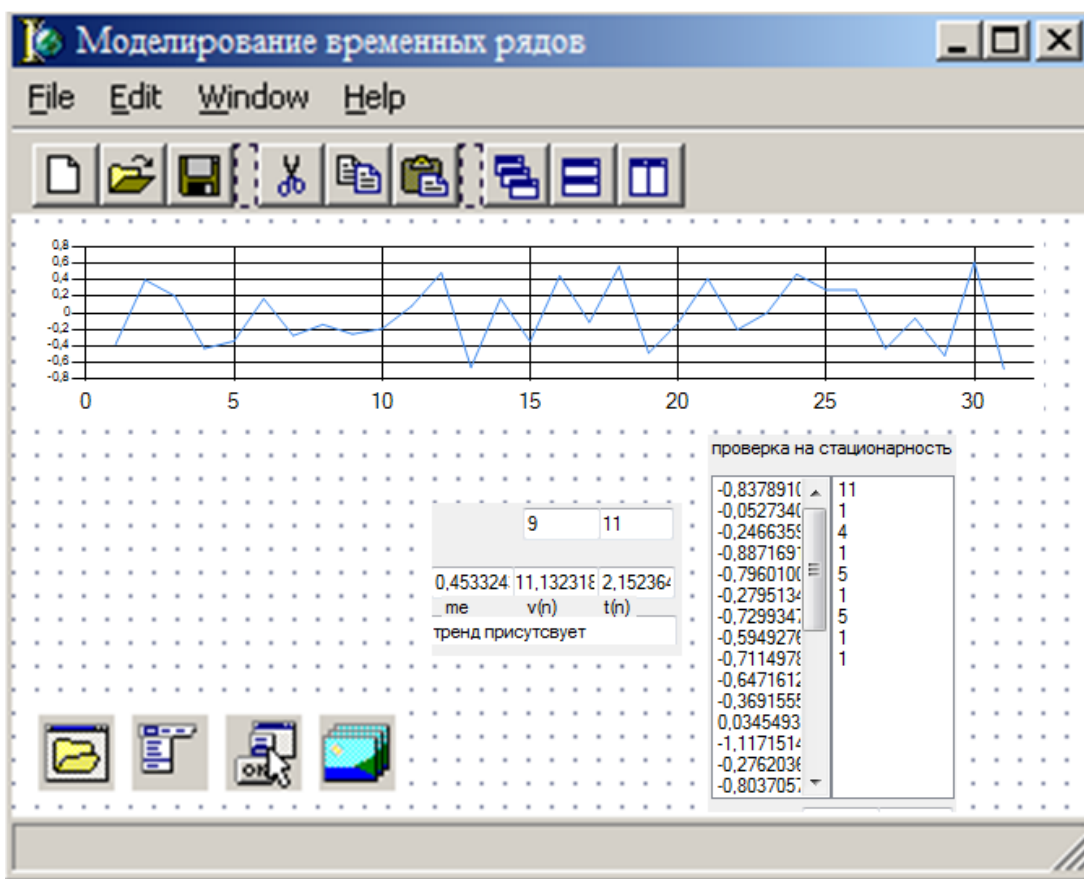


Рис 3.7 Проверка ряда на стационарность и визуальное отображение ряда

На рис. 3.8 представлено визуальное отображение исходного временного ряда и построение аналитического линейного тренда.

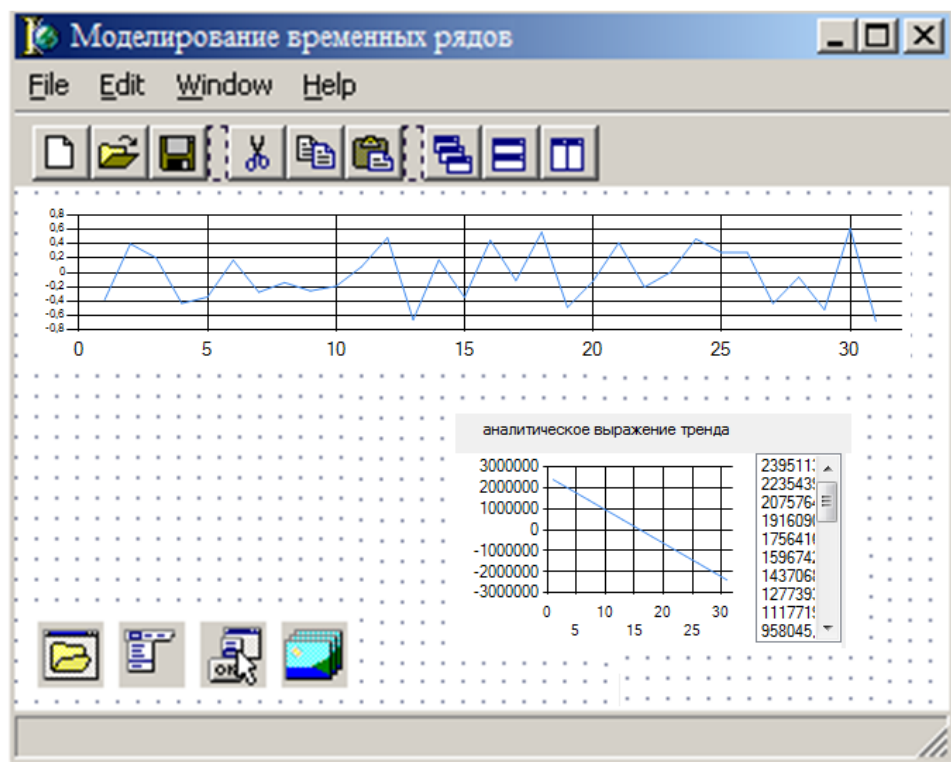


Рис 3.8 Проверка ряда на стационарность и визуальное отображение ряда

На рис. 3.9 представлены значения трендов, рассчитанные методом скользящего среднего при $m=1,2,3$, рассчитанные методом экспоненциально взвешенного скользящего среднего Брауна, а также их визуальное представление.

3.2 Анализ работы программной системы

Проведем расчеты с использованием MS Excel для сравнения результатов

1. исследование ряда на стационарность на основе критерия, основанного на выборочной данные медиане.

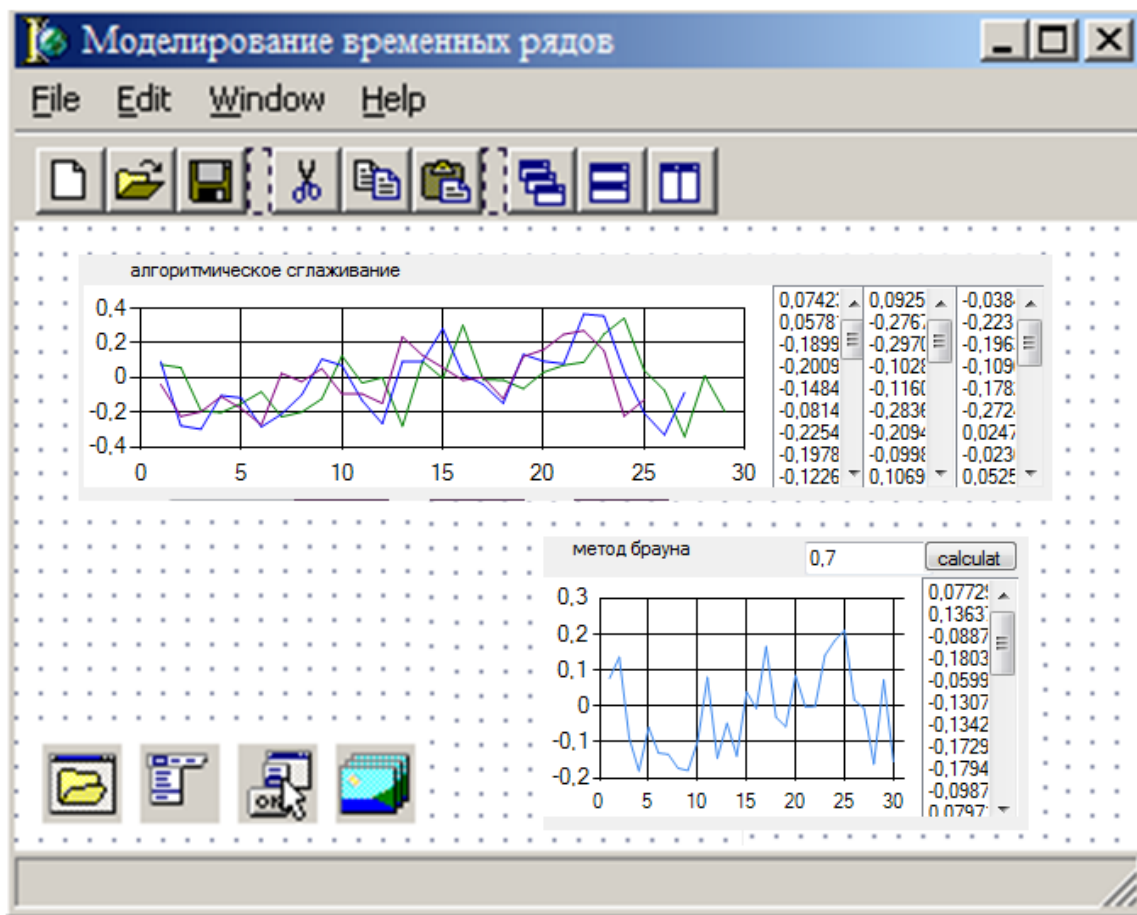


Рис 3.9 Расчет трендов методом скользящего среднего и методом Брауна

По представленному временному ряду вычислили наблюдаемые значения количества серий и длины самой протяженной серии и сравнили выполнение следующих неравенств сравнения с теоретическими значениями.

$$v(n) > 0,5(n + 2 - 1,96\sqrt{n-1}),$$

$$\tau(n) < 1,43\ln(n+1),$$

и сделали вывод о наличии тренда, поскольку неравенства не выполнились (см. рис. 3.10).

Результаты расчета совпали с аналогичными, выполненными в программе.

23	3333	-65	-1	1	3398															
24	3337	-61	-1	2																
25	3364	-34	-1	3																
26	3418	20	1	1																
27	3392	-6	-1	1																
28	3380	-18	-1	2																
29	3406	8	1	1																
30	3394	-4	-1	1																
31	3409	11	1	1	7 >				5											
32	3410	12	1	2	7 >				4											
33	3425	27	1	3																
34	3409	11	1	4																
35	3415	17	1	5	неравенство 2 не выполнено															
36	3416	18	1	6	тренд есть															
37	3402	4	1	7																
38	3387	-11	-1	1																
39	3391	-7	-1	2																
40	3390	-8	-1	3																

$$v(n) > 0,5(n+2-1,96\sqrt{n-1}),$$

$$\tau(n) < 1,43 \ln(n+1),$$

Рис 3.10 Проверка ряда на стационарность

2. Проведение алгоритмического сглаживания, на основе скользящего среднего или взвешенного скользящего среднего для удаления из ряда экстремальных значений и придания ему более гладкого вида (см. рис. 3.11);

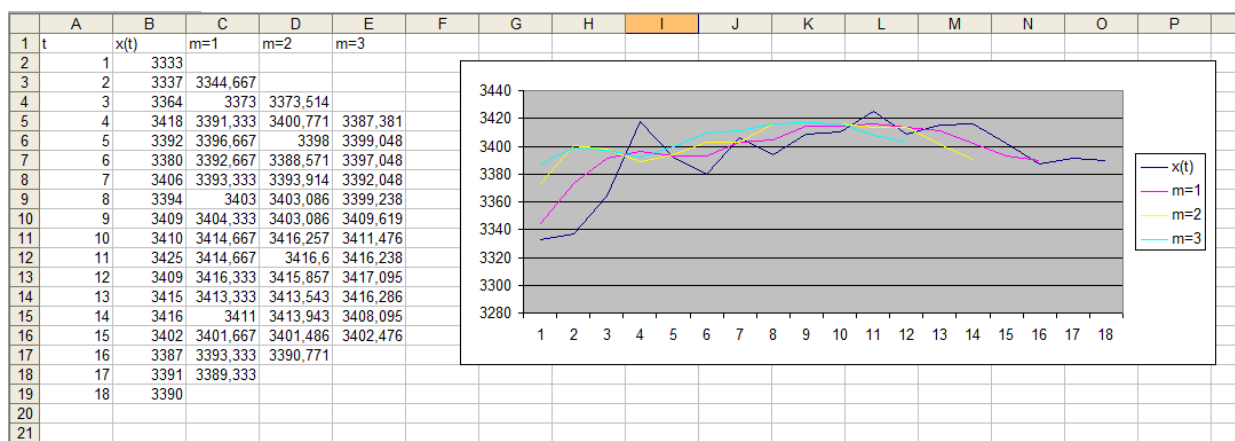


Рис 3.11 Построение тренда методом скользящего среднего

Результаты совпали с аналогичными, полученными с использованием программы.

➤ Определение структуры ряда на основе расчета коэффициентов автокорреляции;

Рассчитали коэффициенты автокорреляции r_1, r_2, \dots, r_8 . Рассчитанные значения совпали с полученными в результате работы программы (см. рис. 3.5). Поскольку самым большим значением является коэффициент

автокорреляции первого порядка, то временной ряд содержит трендовую компоненту, но, возможно, нелинейную. Предварительно можно предположить наличие сезонных колебаний с периодом 8.

	A	B	C	D	E	F	G	H	I
40	3390	-8	-1	3					
41									
42									
43	1	3333	0,685556	ряд имеет трендовую компоненту					
44	2	3337	0,313459	возможны сезонные колебания					
45	3	3364	0,086436	с периодом 8					
46	4	3418	0,279186						
47	5	3392	0,128698						
48	6	3380	-0,10727						
49	7	3406	-0,06252						
50	8	3394	-0,47922						
51	9	3409							
52	10	3410							
53	11	3425							
54	12	3409							
55	13	3415							
56	14	3416							
57	15	3402							
58	16	3387							
59	17	3391							
60	18	3390							
61									

Рис 3.12 Определение структуры временного ряда

После этого были построены основные виды трендов, которые автоматически предлагает для построения MS Excel. Полученные уравнения полностью совпали с рассчитанными с помощью разработанной программы (рис. 3.4, 3.5).

➤ Выбор наиболее подходящей модели для описания эмпирического процесса посредством вычисления коэффициента детерминации, остаточной дисперсии и средней ошибки аппроксимации и их сравнения;

Рассчитывают для каждой модели теоретические значения тренда, коэффициент детерминации, остаточную дисперсию и среднюю ошибку аппроксимации. На основе сравнения этих показателей выбирают оптимальный тренд, рассчитывают статистику Дарбина-Уотсона для

определения адекватности выбранной модели, и делают прогноз по данной модели (рис. 3.14 - 3.17).

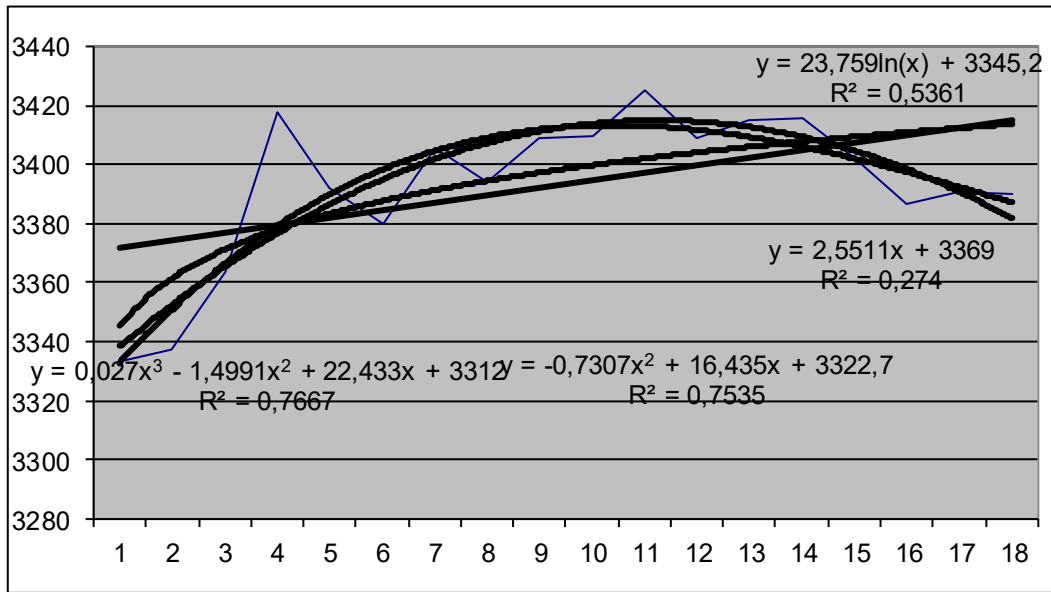


Рис 3.13 Нахождение тренда в виде основных функций (оценка параметров)

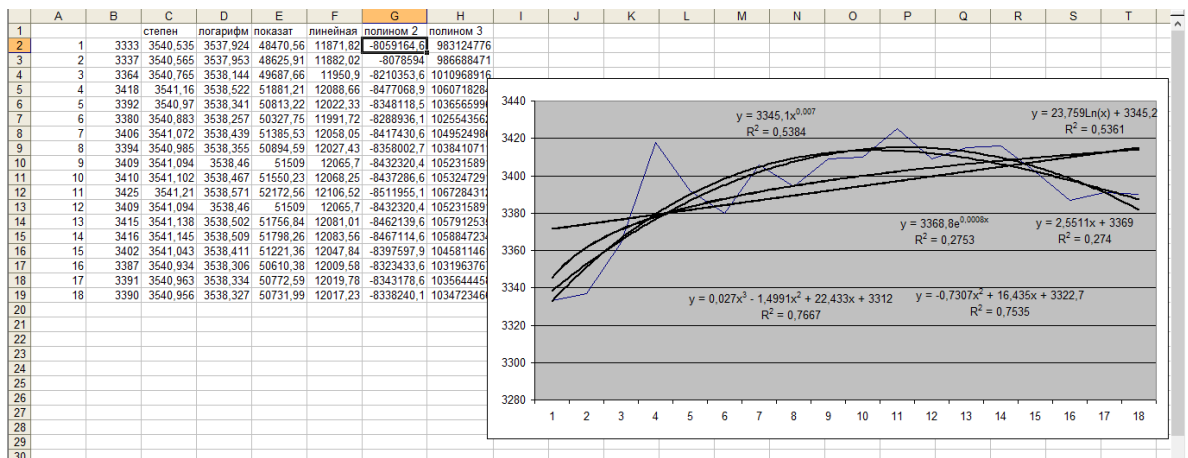


Рис 3.14 Расчет теоретических значений по полученным уравнениям

	A	B	C	D	E	F	G	H
19	18	3390	3413,469	3413,872	3417,662	3414,92	3381,7832	3387,549
20								
21		146,41	148,84	1481,951	1486,187	29,20646	0,00152881	
22		593,8952	608,5341	1383,423	1376,573	244,8349	198,404127	
23		47,94102	53,31817	166,2766	160,106	2,041184	6,43180321	
24		1622,554	1589,061	1474,758	1505,099	1701,662	1484,22186	
25		81,01437	73,29696	94,04759	104,9498	29,07906	3,75390625	
26		53,57529	60,37932	25,09119	18,5468	225,144	340,860214	
27		225,7028	212,201	334,2242	366,4276	16,47792	1,35466321	
28		0,021811	0,366572	12,74875	21,07912	179,9676	235,487439	
29		145,2688	134,4705	251,448	290,365	5,896641	9,94077841	
30		111,2286	101,8662	199,9823	239,9091	15,8404	11,6964	
31		541,8524	521,1361	698,2139	780,5263	98,59894	136,681819	
32		27,09382	22,6681	59,34835	88,11201	32,48088	8,88993856	
33		86,44342	78,48841	120,5969	164,7552	4,550969	29,15892	
34		72,76617	65,58792	85,69867	127,3422	41,3089	93,578537	
35		50,63356	56,86012	55,78785	27,73602	7,938306	0,10400625	
36		559,6162	579,5544	634,9288	520,6429	134,5786	115,5711	
37		445,3748	462,8658	572,58	456,6213	0,005975	3,14033841	
38		550,8123	569,8887	765,1685	620,9964	67,5158	6,00446016	
39	S2ост	297,9002	296,6324	467,5707	464,2208	157,6182	149,182324	
40								

Рис 3.15 Расчет остаточной дисперсии по построенным моделям

	A	B	C	D	E	F	G	H
40								
41		0,00363	0,00366	0,01155	0,011566	0,001621	1,1731E-05	
42		0,007303	0,007392	0,011146	0,011118	0,004689	0,00422104	
43		0,002058	0,002171	0,003833	0,003761	0,000425	0,00075389	
44		0,011785	0,011663	0,011235	0,01135	0,012069	0,01127139	
45		0,002654	0,002524	0,002859	0,00302	0,00159	0,0005712	
46		0,002166	0,002299	0,001482	0,001274	0,004439	0,00546225	
47		0,004411	0,004277	0,005368	0,00562	0,001192	0,00034172	
48		4,35E-05	0,000178	0,001052	0,001353	0,003953	0,00452139	
49		0,003536	0,003402	0,004652	0,004999	0,000712	0,00092488	
50		0,003093	0,00296	0,004147	0,004542	0,001167	0,00100293	
51		0,006796	0,006665	0,007715	0,008157	0,002899	0,00341346	
52		0,001527	0,001397	0,00226	0,002754	0,001672	0,00087463	
53		0,002723	0,002594	0,003216	0,003759	0,000625	0,00158123	
54		0,002497	0,002371	0,00271	0,003303	0,001881	0,00283185	
55		0,002092	0,002217	0,002196	0,001548	0,000828	9,4797E-05	
56		0,006984	0,007108	0,00744	0,006737	0,003425	0,00317402	
57		0,006224	0,006345	0,007057	0,006302	2,28E-05	0,00052259	
58		0,006923	0,007042	0,00816	0,007351	0,002424	0,00072283	
59	A	0,004247	0,004237	0,005449	0,005473	0,002535	0,00234988	
60								

Рис 3.16 Расчет средней ошибки аппроксимации по построенным моделям

➤ Прогнозирование на основе построенной модели;

119								
120	DW	1,090044	1,094596	0,748701	0,754288	1,940427	2,0182406	
121								
122	t							
123	19	3414,762	3415,157			3371,1823	3382,2449	
124	20	3415,988	3416,376			3359,12	3377,02	
125								

Рис 3.17 Расчет статистики Дарбина-Уотсона по построенным моделям и прогноза по наиболее подходящим моделям

Таким образом, наиболее подходящими для описания использованных эмпирических данных являются полиномы второго и третьего порядков. Аналогичные результаты были получены на основе работы программы.

➤ Моделирование сезонных колебаний

Построили аддитивную модель временного ряда.

1. Провели сглаживание исходных уровней ряда методом скользящей средней
2. Вычислили центрированные скользящие средние
3. Нашли оценки сезонной компоненты как разность между фактическими уровнями ряда и центрированными скользящими средними
4. Рассчитали значения сезонной компоненты с учетом того, сезонные воздействия за период должны взаимопогашаться (сумма значений сезонной компоненты по всем кварталам должна быть равна нулю)
5. Исключили влияние сезонной компоненты, вычитая ее значение из каждого уровня исходного временного ряда, в результате полученные данные содержат только тенденцию и случайную компоненту
6. Провели аналитическое выравнивание последнего ряда (с помощью линейного тренда) для получения трендовой компоненты
7. К уровням трендовой компоненты прибавили значения сезонной компоненты
8. Для оценки качества построенной модели рассчитали сумму квадратов полученных абсолютных ошибок
9. Сделали прогноз по полученной аддитивной модели

➤ Прогнозирование на основе кривых роста;

➤ Прогнозирование на основе адаптивных моделей временных рядов;

1) Прогнозирование на основе модели Брауна

Выполним прогнозирование для линейного тренда и для параболы второго порядка.

Для линейного тренда прогнозирование методом простого экспоненциального сглаживания осуществили в следующей последовательности:

1. Определили параметры линейного тренда a_0 и a_1 , описывающего тенденцию исходного временного ряда. Параметры a_0 и a_1 определяются методом наименьших квадратов.

2. Определили начальные условия первого и второго порядков (порядок начальных условий определяется числом параметров уравнения тренда – линейного тренда - a_0 и a_1) по формулам вида:

- начальное условие первого порядка:

$$S_0^{[1]}(y) = a_0 - \frac{1-\alpha}{\alpha} a_1 \quad (3.1)$$

- начальное условие второго порядка:

$$S_0^{[2]}(y) = a_0 - \frac{2(1-\alpha)}{\alpha} a_1 \quad (3.2)$$

где a_0 и a_1 – параметры уравнения тренда, полученные методом наименьших квадратов.

3. Рассчитали экспоненциальные средние первого и второго порядка:

- экспоненциальная средняя первого порядка:

$$S_t^{[1]}(y) = \alpha y_t + (1-\alpha) \times S_0^{[1]}(y); \quad (3.3)$$

где y_t – значение последнего фактического уровня исходного временного ряда;

- экспоненциальная средняя второго порядка:

$$S_t^{[2]}(y) = \alpha \times S_t^{[1]}(y) + (1-\alpha) \times S_0^{[2]}(y) \quad (3.4)$$

Прогноз построили по модели вида:

$$\hat{y}^*_{t+L} = \hat{a}_0 + \hat{a}_1 t, \quad (3.5)$$

Получили: $\hat{y}^*_{t+L} = 3380,96 + 4,2115t$

где оценки коэффициентов модели определяются по следующим формулам:

$$\begin{aligned} \hat{a}_0 &= 2S_t^{[1]}(y) - S_t^{[2]}(y); \\ \hat{a}_1 &= \frac{\alpha}{1-\alpha} [S_t^{[1]}(y) - S_t^{[2]}(y)] \end{aligned} \quad (3.6)$$

4. Ошибку прогноза определили по следующей формуле:

$$\delta_{\hat{y}^*_{t+L}} = \delta_y \times \sqrt{\frac{\alpha}{(2-\alpha)^5} \times [1 - 4(1-\alpha) + 5(1-\alpha)^2 + 2\alpha(4-3\alpha)] \times t + 2\alpha^2 L^2} \quad (3.7)$$

где δ_y - средняя квадратическая ошибка, рассчитанная по отклонениям эмпирических значений признака от теоретических, полученных по уравнению линейного тренда:

$$\delta_y = \sqrt{\frac{\sum (y_i - \hat{y}^*_{t+L})^2}{n - K}}, \quad (3.8)$$

где K - число степеней свободы, определяемое в зависимости от длины исходного временного ряда (n) и числа параметров уравнения тренда.

Основные показатели экспоненциального сглаживания для ряда динамики, описываемого параболой второго порядка, рассчитали по следующим формулам.

Начальные условия:

$$\begin{aligned} & S_o^{[1]}(y) = a_0 - \frac{1-\alpha}{\alpha} a_1 + \frac{(1-\alpha)(2-\alpha)}{2\alpha^2} a_2; \\ \text{- первого порядка} & \\ & S_o^{[2]}(y) = a_0 - \frac{2(1-\alpha)}{\alpha} a_1 + \frac{(1-\alpha)(3-2\alpha)}{\alpha^2} a_2; \quad (3.9) \\ \text{- второго порядка} & \\ & S_o^{[3]}(y) = a_0 - \frac{3(1-\alpha)}{\alpha} a_1 + \frac{3(1-\alpha)(4-3\alpha)}{2\alpha^2} a_2. \\ \text{- третьего порядка} & \end{aligned}$$

Экспоненциальные средние:

- первого порядка $S_t^{[1]}(y) = \alpha \cdot y_t + (1 - \alpha)S_0^{[1]}(y);$
- второго порядка $S_t^{[2]}(y) = \alpha \cdot S_t^{[1]}(y) + (1 - \alpha)S_0^{[2]}(y);$
- третьего порядка $S_t^{[3]}(y) = \alpha \cdot S_t^{[2]}(y) + (1 - \alpha)S_0^{[3]}(y).$

(3.10)

Модель прогноза:

$$\hat{y}_{t+L}^* = \hat{a}_0 + \hat{a}_1 t + \frac{1}{2} \hat{a}_2 t^2 \quad (3.11)$$

получили:

Оценка параметров модели прогноза:

$$\hat{a}_0 = 3[S_t^{[1]}(y) - S_t^{[2]}] + S_t^{[3]}(y)$$

$$\hat{a}_1 = \frac{\alpha}{2(1 - \alpha)^2} [(6 - 5\alpha)S_t^{[1]}(y) - 2(5 - 4\alpha)S_t^{[2]}(y) + (4 - 3\alpha)S_t^{[3]}(y)] \quad (3.12)$$

$$\hat{a}_2 = \frac{\alpha^2}{(1 - \alpha)^2} [S_t^{[1]}(y) - 2S_t^{[2]}(y) + S_t^{[3]}(y)]$$

Ошибка прогноза определяется по формуле:

$$\sigma_{\hat{y}_{t+L}^*} = \sigma_y \sqrt{2\alpha + 3\alpha^2 + 3\alpha^3 t}, \text{ где } \sigma_y = \sqrt{\frac{\sum (y_i - \bar{y}_{t+L}^*)^2}{n - 3}}. \quad (3.13)$$

Полученные результаты представлены на рис. 3.18. Они совпали с рассчитанными в программе.

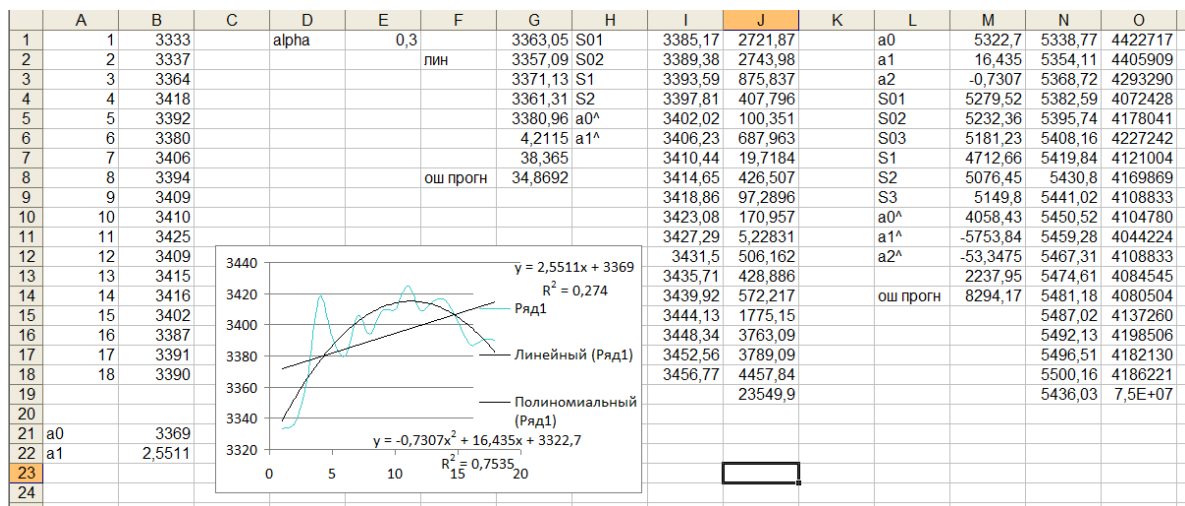


Рис 3.18 Прогнозирование на основе модели Брауна

2) Прогнозирование на основе модели Хольта

Рассчитали коэффициенты модели:

- определили начальные значения коэффициентов модели

$$\hat{a}_{11} = x_1, \quad \hat{a}_{21} = x_2 - \hat{a}_{11} \quad (3.14)$$

и параметров адаптации

$$\alpha_1 = 0,1, \quad \alpha_2 = 0,1.$$

- рассчитали текущие значения коэффициентов регрессии

$$\begin{aligned} \hat{a}_{1t} &= \alpha_1 x_t + (1 - \alpha_1)(\hat{a}_{1t-1} + \hat{a}_{2t-1}); \\ \hat{a}_{2t} &= \alpha_2(\hat{a}_{1t} - \hat{a}_{1t-1}) + (1 - \alpha_2)\hat{a}_{2t-1}, \quad t = \overline{2, 9}. \end{aligned} \quad (3.15)$$

Подобрали оптимальные значения параметров адаптации

$$\alpha_1 = 0,8, \quad \alpha_2 = 0,9.$$

Рассчитали прогнозные значения по адаптивной модели

- рассчитали текущие коэффициенты модели ($t = \overline{2, 12}$) с использованием оптимально настроенных параметров адаптации.

- рассчитали прогнозные значения \hat{y}_t , ($t = 13; 14; 15$) по модели с текущими коэффициентами для момента $t = 12$.

Результаты расчетов представлены на рис. 3.19. Они полностью совпали с рассчитанными с помощью программы.

	A	B	C	D	E	F	G	H	I	J	K
1	год	объем продаж	t								
2	1988	936000	1								
3	1989	945400	2					t	a^1t	a^2t	
4	1990	1058000	3					2	945400	9400	
5	1991	1010500	4					3	965120	10432	
6	1992	1023600	5					4	979047	10781,48	
7	1993	1033200	6					5	993205	11119,1972	
8	1994	1088100	7					6	1007212	11407,9507	
9	1995	1083400	8					7	1025568	12102,7494	
10	1996	1159700	9					8	1042244	12560,0406	
11	1997	1230100	10					9	1065293	13609,0024	
12	1998	1361000	11								
13	1999	1523000	12								
14											
15	a^11=	936000		a^11t=alf1xt+(1-alf1)(a^11t-1+a^22t-1)				0,1	0,1		
16	a^21=	9400		a^22t=alf2(a^11t-a^11t-1)+(1-alf2)a^22t-1							
17	alf1	0,1						1230100	1180639	S(alf1,alf2)	813009
18	alf2	0,1						1361000	1122222		
19	alf3	0,8						1523000	136100		
20	alf4	0,9									
21											
22	период	y	a1	a2	прогноз	(y-y^a)^2					
23	1	936000	936000	9400							
24	2	945400	945400	9400							
25	3	1058000	965120	10432							
26	4	1010500	979047	10781,5							
27	5	1023600	993205	11119,2							
28	6	1033200	1007212	11408							
29	7	1088100	1025568	12102,7							
30	8	1083400	1042244	12560							
31	9	1159700	1065293	13609							
32	контрольная выборка										
33	10	1230100			1201383,466	824639323,5					
34	11	1361000			1214992,468	21318199275					
35	12	1523000			1228601,471	86670493973					

Рис 3.19 Прогнозирование на основе модели Хольта

3) Построение модели Авторегрессии-скользящего среднего (ARMA)

Проверили, какую модель лучше построить по исходным данным - модель авторегрессии первого порядка AR(1) $Y_t = a_0 + a_1 Y_{t-1} + \varepsilon_t$, где $\varepsilon_t = Y_t - \bar{Y}_t$, или авторегрессионную модель скользящей средней ARMA(1,1) $Y_t = a_0 + a_1 Y_{t-1} + b_1 \varepsilon_{t-1} + u_t$, где u_t - ненаблюдаемая ошибка в данном уравнении.

Сформировали ряды Y_t и Y_{t-1} . Проверили временной ряд на стационарность с помощью критерия Дики-Фуллера.

Оценили параметры модели $Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \varepsilon_t$ с помощью МНК.

Рассчитали статистики $DF_{расч}$ для проверки значимости корня: $DF_{расч} = \beta_1 / S_{\beta_1}$

где $\beta_1 = \alpha_1 - 1$, а S_{β_1} - стандартная ошибка β_1 и сравнили ее с критическим значением расширенного критерия Дики-Фуллера при 95%-ном уровне значимости

$EDF = \varphi_0 + \frac{\varphi_1}{T} + \frac{\varphi_2}{T^2}$, ($\varphi_0 = -2,57$ (1%) или $-1,94$ (5%); $\varphi_1 = -1,96$ (1%) или $-0,398$ (5%); $\varphi_2 = -10,04$ (1%) или 0 (5%)).

Получили, что $DF_{расч} > EDF$, то ряд нестационарен для данного уровня значимости.

Далее сформировали ряд $\Delta Y_t = Y_t - Y_{t-1}$, оценили параметры модели

$\Delta Y_t = \alpha_0 + \alpha_1 \Delta Y_{t-1} + \varepsilon_t$, рассчитали статистику $DF_{расч} = \frac{\beta}{S_\beta}$ и сравнили ее с критическим значением расширенного критерия Дикки-Фуллера на 95%-ном уровне значимости,

$\varphi_0 = -3,43$ (1%) или $-2,86$ (5%); $\varphi_1 = -6,00$ (1%) или $-2,74$ (5%); $\varphi_2 = -29,25$ (1%) или $-8,36$ (5%),

так как модель содержит свободный член.

Получили, что $DF_{расч} < EDF$, таким образом, ряд стационарен, то есть представляет собой процесс $I(1)$.

После этого определили порядок авторегрессии для преобразованного ряда, рассчитав частные коэффициенты автокорреляции. Их анализ показал, что имеет смысл строить модель $ARMA(1,1,0)$.

После этого рассчитали прогнозные значения по формулам:

$$\widehat{Y}_{t+1} = \alpha_0 + (\beta + 1)Y_t - \beta Y_{t-1}$$

$$\widehat{Y}_{t+2} = \alpha_0 + (\beta + 1)\widehat{Y}_{t+1} - \beta Y_t$$

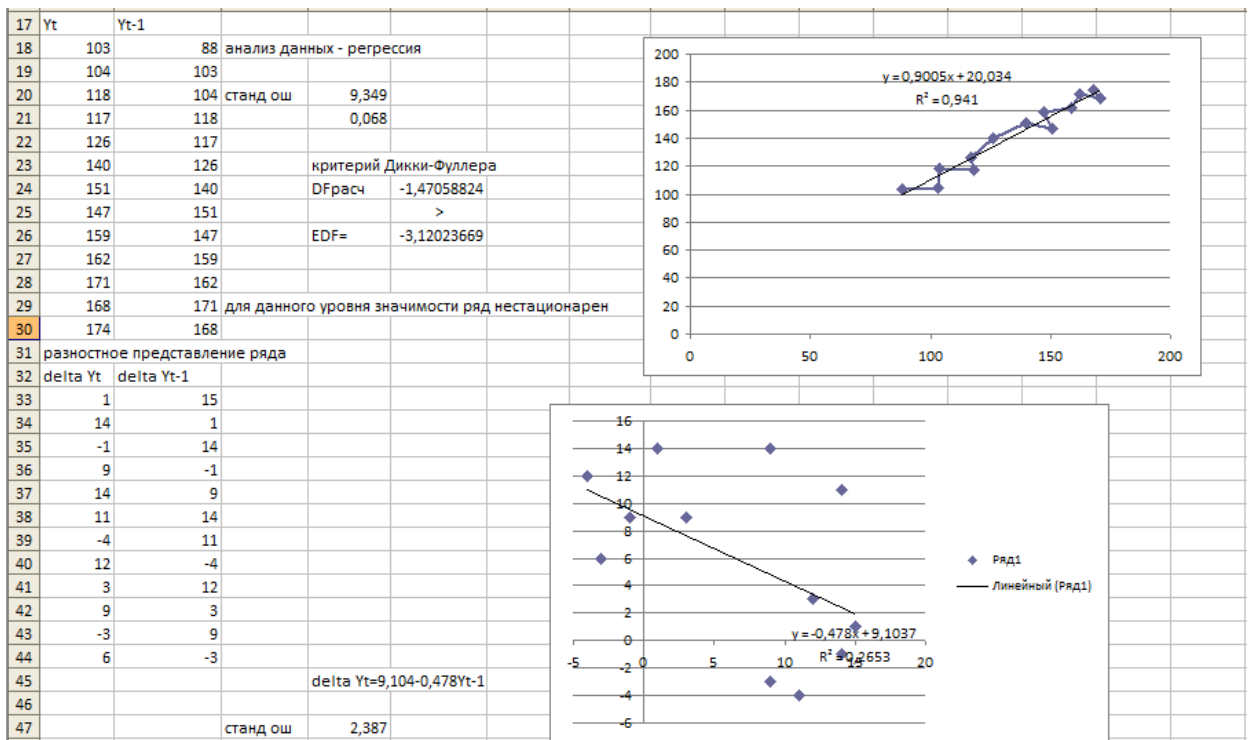


Рис 3.20 Исследование ряда

48															
49	DFрасч=	-5,865079365													
50	EDF=	-3,146388889													
51															
52	т.к DFрасч<EDF, то процесс I(1)														
53															
54	rho1=r1=-0,478														
55															
56	delta Yt	delta Yt-1	delta Yt-2												
57	14	1	15												
58	-1	14	1												
59	9	-1	14												
60	14	9	-1												
61	11	14	9												
62	-4	11	14												
63	12	-4	11												
64	3	12	-4												
65	9	3	12												
66	-3	9	3												
67	6	-3	9												
68															
69	rho2=-0,036														
70															
71	так как значение частного коэффициента автокорреляции резко падает, то для преобразованного врем ряда имеет смысл строить модель ARMA(1,1,0)														
72															
73	Прогноз														
74	delta Yt=9,104-0,478deltaYt-1														
75	Yt-Yt-1=9,104-0,478(Yt-1-Yt-2)														
76	Yt=9,104+(1-0,478)Yt-1+0,478Yt-2														
77	Y^*t+1=	9,104+0,522Yt+0,478Yt-1			180,236				180						
78	Y^*t+2=	9,104+0,522Yt+1+0,478Yt			186,35919				186						

Рис 3.21 Построение модели ARMA(1,1,0) и построение прогноза

➤ Интерпретация результатов, выводы и пояснения системы.

Таким образом, анализ показал работоспособность системы и адекватность полученных результатов.

Основные выводы и результаты

Проведен анализ методов анализа и моделирования временных рядов, а также прогнозирования на основе данных моделей. Разработана программная система, автоматизирующая моделирование и прогнозирование процессов с использованием аппарата временных рядов. Были проведены аналитические расчеты реализованных в программе моделей на основе известных в теории формул и использованием табличного процессора MS Excel. Сравнение результатов показало работоспособность программы и правильность проводимых ею расчетов и формулируемых выводов.

ЗАКЛЮЧЕНИЕ

Цель работы достигнута. Реализована компьютерная поддержка анализа и моделирования тенденций временных рядов.

Поставленные задачи решены:

1. Проведен анализ существующих методов анализа и моделирования тенденций временных рядов;
2. Проанализированы подходы к прогнозированию временных рядов с учетом ограничений каждого подхода;
3. Проанализирован существующий инструментарий;
4. Спроектирована программная система анализа и моделирования тенденций временных рядов;
5. Разработана программная система;
6. Протестирована работа программной системы.

В дальнейшем планируется доработка программы: добавление в нее моделей финансовой эконометрики, автоматизации построения и анализа моделей авторегрессии и скользящего среднего первого порядка.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Айвазян, С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики: учебник для вузов – М.: ЮНИТИ, 1998. – 1022 с.
2. Берндт Э. Практика эконометрики: классика и современность. - М.: ЮНИТИ-ДАНА, 2005. - 863 с.
3. Вендров А.М. Проектирование программного обеспечения экономических информационных систем: учебник – М.: «АСТ», 2004. – 456 с.
4. Елисеева И.И. Эконометрика. – М.: «Финансы и статистика», 2003, - 344 с.
5. Замков О.О., Толстопятенко А.В., Черемных Ю.Н. Математические методы в экономике: Учебник / Под общ. ред. д.э.н., проф. А.В. Сидоровича; МГУ им. М.В. Ломоносова. – 3-е изд., перераб. – М.: Издательство «Дело и Сервис», 2001. – 368 с.
6. Карп Д.Б. Эконометрика: основные формулы с комментариями. Учебно-методическое пособие. Владивосток, 2004. - 50 с.
7. Практикум по эконометрике - Елисеева И.И. - Учебное пособие. – М.: «Финансы и статистика», 2003 - 192 с.
8. Кендэл М. Временные ряды. - М.: Финансы и статистика, 1981.
9. Колеников С.. Прикладной эконометрической анализ в статистическом пакете Stata, 2003.
10. Кнут Д. Искусство программирования: В 3 т. / Д. Кнут. – М.: Основные алгоритмы : Вильямс, 2006. – Т. 1. – 720 с.
11. Кнут Д. Искусство программирования: В 3 т. / Д. Кнут. – М.: Получисленные методы: Вильямс, 2007. – Т. 2. – 832 с.
12. Кругов В. В. Методы прогнозирования многомерных временных рядов // Приборы и системы. Управление, контроль, диагностика. – 2005. - № 2. – С. 62-66.

13. Лукашин Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов: Учеб. пособие. - М.: Финансы и статистика, 2003.- 416с: ил.
14. Магнус, Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс: учебник – 6-е изд., перераб. и доп. – М.: Дело, 2004. – 576 с.
15. Минашкин В.Г. Садовникова Н.А. Шмойлова Р.А. Бизнес-статистика и прогнозирование /Московский государственный университет экономики, статистики и информатики. - М., 2008. – 154с.
16. Осипов Д.Л. Базы данных и Delphi. Теория и практика - БХВ-Петербург, 2011, 752 с.
17. Осипов Д.Л. Delphi. Программирование для Windows, OS X, iOS и Android - БХВ-Петербург, 2014, 464 с.
18. Проблемы прогнозирования некоторых макроэкономических показателей / Р.М. Энтов [и др.] 28.11.2002 - <http://www.iet.ru/publication.php>.
19. Разработка учебно-методических материалов для преподавания курсов по применению количественного инструментария к решению социально-экономических задач. Исследовательский проект инициативной группы под руководством Елены Вакуленко. // Экономическая социология, 2010 - Т. 11. № 4.
20. Ратникова Т.А. Введение в эконометрический анализ панельных данных. Москва: Издательский дом ГУ-ВШЭ, 2010.
21. Ржаницына В. Анализ полугодовой отчетности // Консультант, № 15, август 2005 г. // Справочно-правовая система «Гарант»
22. Стрижов В.В., Шакин В.В. Прогноз и управление в авторегрессионных моделях/ В. В. Стрижов, В. В. - <http://www.ccas.ru>
23. Тихомиров, Н.П., Дорохина Е.Ю. Эконометрика: учебник – М.: Издательство «Экзамен», 2003. – 512 с.
24. Шанченко, Н. И. Лекции по эконометрике: учебное пособие -

Ульяновск : УлГТУ, 2008. – 139 с.

25. Швайко П. Эконометрические модели анализа и прогнозирования емкости первичного рынка ГКО // Эковест. – 2002. - № 2. – С. 111 – 153.

26. Шкрыль А. Разработка клиент-серверных приложений в Delphi - БХВ-Петербург, 2006, 480 с.

Листинг процедур программы

```

Const
    zero=1.0E-15;
    MaxVec=128;
type
    TMDIChild = class(TForm)
        PageControl1: TPageControl;
        TSParams: TTabSheet;
        TSResult: TTabSheet;
        GridXY: TStringGrid;
        Panel1: TPanel;
        BitBtn1: TBitBtn;
        MemoRes: TMemo;
        procedure FormClose(Sender: TObject; var Action: TCloseAction);
    private
    public
        xyInp,yRes:array[1..16000,1..5] of real;
        xInp,yInp:array[1..16000] of real;
        xls:TMyExcel;
        procedure LoadXls(fname:string);
        procedure Single_Regression;
        procedure Multi_Regression;
        procedure Method_Line(var S_ost:real;FuncType:integer);
        function MainFunc(a,b,x:real; FuncType:integer):real;
        function MainFunc1(a,b,x:real; FuncType:integer):real;
    end;
    TVec=array [1..MaxVec] of float;
    TMatr=array [1..MaxVec] of TVec;

    function Opred(Nx,Ny:integer;const A:TMatr):float;

implementation

{$R *.dfm}
uses UFisher;

function AlgDop(x,y,Nx,Ny:integer;const A:TMatr):float;
VAR r:float;
    m:TMatr;
    i,j,ii,jj:integer;
Begin

```

```

ii:=0;
for i:=1 to Ny do
if i<>y then
begin
inc(ii);
jj:=0;
for j:=1 to Nx do
if j<>x then
begin
inc(jj);
m[ii,jj]:=A[i,j];
end;
end;
r:=Opred(Nx-1,Ny-1,m);
if Odd(x+y) then r:=-r;
Result:=r;
End;
function Opred(Nx,Ny:integer;const A:TMatr):float;
VAR i,n:integer;
Begin
Result:=0;
if Nx<>Ny then
begin
Application.MessageBox('Opred: Матрица не квадратная !', 'Ошибка!');
exit;
end;
n:=Nx;
case n of
0:Result:=0;
1:Result:=A[1,1];
2:Result:=A[1,1]*A[2,2]-A[1,2]*A[2,1];
Else
for i:=1 to n do
Result:=Result+A[1,i]*AlgDop(i,1,Nx,Ny,A);
End;
End;
function Kramer(Nx,Ny:integer;const A:TMatr; const H:TVec; var X:TVec):boolean;
VAR
i,j,n:integer;
ta:TMatr;
d:float;
Begin
Result:=false;
if Nx<>Ny then
begin

```



```

    Application.MessageBox('Opred: Матрица не квадратная!', 'Ошибка!');
    exit;
end;
n:=Nx;
d:=Opred(Nx,Ny,A);
if abs(d)<zero then
begin
    Application.MessageBox('Определитель равен 0, решений нет!', 'Ошибка!');
    exit;
end;
Result:=true;
for i:=1 to n do
begin
    ta:=A;
    for j:=1 to n do
        ta[j,i]:=H[j];
    X[i]:=Opred(Nx,Ny,ta)/d;
end;
End;

```

```

procedure TMDIChild.FormClose(Sender: TObject; var Action: TCloseAction);
begin
    Action := caFree;
end;

```

```

procedure TMDIChild.LoadXls(fname: string);
var
    xs:array[1..16000,1..5] of string;
    i,j,a,row,rowcnt,colcnt:integer;
    s:string;
begin
    xls:=TMyExcel.Create;
    xls.RunExcel;
    xls.FileOpen(fname);
    xls.SetSheet(1);
    row:=1;
    colcnt:=0;
    rowcnt:=0;
    for j:=1 to 5 do
    begin
        s:=xls.Cells[row,1+j-1];
        if s="" then break;
        inc(colcnt);
    end;
    if colcnt<2 then

```

```

begin
  xls.CloseExcel;
  Application.MessageBox('Недостаточно колонок в таблице исходных данных!',
    'Ошибка в исходном файле!');
  exit;
end;
GridXY.ColCount:=colcnt;
inc(row);
for i:=1 to 16000 do
begin
  for j:=1 to colcnt do
  begin
    xs[i,j]:=xls.Cells[row+i-1,j];
  end;
  if xs[i,1]=" then break;
  inc(rowcnt);
end;
if rowcnt<7 then
begin
  xls.CloseExcel;
  Application.MessageBox('Недостаточно строк в таблице исходных данных!',
    'Ошибка в исходном файле!');
  exit;
end;
GridXY.RowCount:=rowcnt+1;
for j:=1 to colcnt do
  GridXY.Cells[j-1,0]:=xls.Cells[1,j];
for i:=1 to rowcnt do
  for j:=1 to colcnt do
  begin
    s:=xs[i,j];
    GridXY.Cells[j-1,i]:=s;
    a:=Pos(',',s);
    if a>0 then s[a]:='.';
    val(s,xyInp[i,j],a);
  end;
  xls.CloseExcel;
  xls.Free;

  if colcnt=2 then Single_Regression;
  if colcnt>2 then Multi_Regression;
end;

procedure TMDIChild.Multi_Regression;
var

```

```

i,j,k,ny,nx:integer;
MAstd,MA:TMatr;
MHstd,MXstd,MH,MX,AvgX,SigmaX:TVec;
AvgY,SigmaY:real;
sxd,syd,snd:string;
begin
ny:=GridXY.RowCount-1;
nx:=GridXY.ColCount-1;
for j:=1 to nx do
begin
  AvgX[j]:=0;
  SigmaX[j]:=0;
end;
for i:=1 to ny do
begin
  for j:=1 to nx do
  begin
    AvgX[j]:=AvgX[j]+xyInp[i,j];
    SigmaX[j]:=SigmaX[j]+sqr(xyInp[i,j]);
  end;
  AvgY:=AvgY+xyInp[i,nx+1];
  SigmaY:=SigmaY+sqr(xyInp[i,nx+1]);
end;
AvgY:=AvgY/ny;
SigmaY:=SigmaY/ny;
for j:=1 to nx do
  SigmaX[j]:=SigmaX[j]-sqr(AvgX[j]);

MHstd[1]:=0;
for i:=1 to nx do
  MAstd[1,i+1]:=0;
MAstd[1,1]:=ny;
for j:=1 to ny do
begin
  MHstd[1]:=MHstd[1]+(xyInp[j,nx+1]-AvgY)/SigmaY;
  for i:=1 to nx do
    MAstd[1,i+1]:=MAstd[1,i+1]+(xyInp[j,i]-AvgX[i])/SigmaX[i];
  end;
for i:=1 to nx do
begin
  MHstd[i+1]:=0;
  MAstd[i+1,1]:=MAstd[1,i+1];
  for j:=1 to ny do
  begin
    MHstd[i+1]:=MHstd[i+1]+(xyInp[j,i]-AvgX[i])/SigmaX[i]*

```

```

    (xyInp[j,nx+1]-AvgY)/SigmaY;
  for k:=1 to nx do
    MAstd[i+1,k+1]:=MAstd[i+1,k+1]+(xyInp[j,i]-AvgX[i])/SigmaX[i]*
      (xyInp[j,k]-AvgX[k])/SigmaX[k];
  end;
end;
if not Kramer(nx+1,nx+1,MAstd,MHstd,MXstd) then
begin
  MemoRes.Lines.Add('Решений нет!');
  exit;
end;
MemoRes.Lines.Add('Вектор коэффициентов Beta:');
for j:=1 to nx+1 do
  MemoRes.Lines.Add('B['+IntToStr(j-1)+']='+FloatStr0(MXstd[j],6));

MH[1]:=0;
for i:=1 to nx do
  MA[1,i+1]:=0;
MA[1,1]:=ny;
for j:=1 to ny do
begin
  MH[1]:=MH[1]+xyInp[j,nx+1];
  for i:=1 to nx do
    MA[1,i+1]:=MA[1,i+1]+xyInp[j,i];
  end;
for i:=1 to nx do
begin
  MH[i+1]:=0;
  MA[i+1,1]:=MA[1,i+1];
  for j:=1 to ny do
begin
  MH[i+1]:=MH[i+1]+xyInp[j,i]*xyInp[j,nx+1];
  for k:=1 to nx do
    MA[i+1,k+1]:=MA[i+1,k+1]+xyInp[j,i]*xyInp[j,k];
  end;
end;
end;
if not Kramer(nx+1,nx+1,MA,MH,MX) then
begin
  MemoRes.Lines.Add('Решений нет!');
  exit;
end;
MemoRes.Lines.Add('Вектор коэффициентов B:');
for j:=1 to nx+1 do
  MemoRes.Lines.Add('B['+IntToStr(j-1)+']='+FloatStr0(MX[j],6));

```

```

MemoRes.Lines.Add('Частные коэффициенты эластичности Э:');
for j:=1 to nx do
  MemoRes.Lines.Add('Э['+IntToStr(j)+']='+
    FloatStr0(MX[j]*AvgX[j]/AvgY,6));

for i:=1 to ny do
begin
  yRes[i,1]:=MX[1];
  for j:=1 to nx do
    yRes[i,1]:=yRes[i,1]+MX[j+1]*xyInp[i,j];
  end;
  xls:=TMyExcel.Create;
  xls.RunExcel;
  xls.XISheet.Name:='Результат';

  for j:=1 to nx do
    xls.Cells[1,j]:='X'+IntToStr(j);
  xls.Cells[1,nx+1]:='Y-Исх.данные';
  xls.Cells[1,nx+2]:='Расчет';
  for i:=1 to ny do
  begin
    for j:=1 to nx do
      xls.Cells[i+1,j]:=xyInp[i,j];
      xls.Cells[i+1,nx+1]:=xyInp[i,nx+1];
      xls.Cells[i+1,nx+2]:=yRes[i,1];
    end;

    sxd:='C1';
    syd:='C'+IntToStr(nx+1);
    snd:='C'+IntToStr(nx+2);
    xls.XIApp.Charts.Add;
    xls.XIApp.ActiveChart.ChartType := xlLineMarkers;
    xls.XIApp.ActiveChart.Location(Where:=xlLocationAsNewSheet,
Name:='ДиаграммаГрафики');
    xls.XIApp.ActiveChart.HasDataTable := true;
    xls.SetRange(2, nx+1, ny+1, nx+1);
    xls.XIApp.ActiveChart.SetSourceData (Source:=xls.Range, PlotBy:=xlColumns);
    xls.XIApp.ActiveChart.SeriesCollection(1).XValues := '=Результат!R2'+sxd+'!R'+
      IntToStr(ny+1)+sxd;
    xls.XIApp.ActiveChart.SeriesCollection(1).Values := '=Результат!R2'+syd+'!R'+
      IntToStr(ny+1)+syd;
    xls.XIApp.ActiveChart.SeriesCollection(1).Name := 'Исх.данные';
    xls.XIApp.ActiveChart.SeriesCollection.NewSeries;
    xls.XIApp.ActiveChart.SeriesCollection(2).XValues := '=Результат!R2'+sxd+'!R'+
      IntToStr(ny+1)+sxd;

```

```

xls.XlApp.ActiveChart.SeriesCollection(2).Values := '=Результат!R2'+snd+'R'+
  IntToStr(ny+1)+snd;
xls.XlApp.ActiveChart.SeriesCollection(2).Name := 'Расчет';
end;

procedure TMDIChild.Single_Regression;
var
  i,j,rowcnt,colcnt:integer;
  S_ost:array[1..5] of real;
begin
  colcnt:=GridXY.ColCount;
  rowcnt:=GridXY.RowCount-1;
  if colcnt<>2 then exit;

  for i:=1 to 5 do
    Method_Line(S_ost[i],i);

  xls:=TMyExcel.Create;
  xls.RunExcel;
  xls.XlSheet.Name:='Результат';

  xls.SetRangeArray(1,1,'X','Y-Исх.данные','Линейная','Гипербола',
    'Степенная','Показательная','Экспоненциальная');
  for i:=1 to rowcnt do
  begin
    xls.Cells[i+1,1]:=xyInp[i,1];
    xls.Cells[i+1,2]:=xyInp[i,2];
    for j:=1 to 5 do
      xls.Cells[i+1,j+2]:=yRes[i,j];
    end;
  xls.XlApp.Charts.Add;
  xls.XlApp.ActiveChart.ChartType := xlLineMarkers;
  xls.XlApp.ActiveChart.Location(Where:=xlLocationAsNewSheet,
  Name:='ДиаграммаГрафики');
  xls.XlApp.ActiveChart.HasDataTable := true;
  xls.SetRange(2, 1, rowcnt+1, 6);
  xls.XlApp.ActiveChart.SetSourceData (Source:=xls.Range, PlotBy:=xlColumns);
  xls.XlApp.ActiveChart.SeriesCollection(1).XValues := '=Результат!R2C1:R'+
    IntToStr(rowcnt+1)+'C1';
  xls.XlApp.ActiveChart.SeriesCollection(1).Values := '=Результат!R2C2:R'+
    IntToStr(rowcnt+1)+'C2';
  xls.XlApp.ActiveChart.SeriesCollection(1).Name := 'Исх.данные';
  //xls.XlApp.ActiveChart.SeriesCollection.NewSeries;
  xls.XlApp.ActiveChart.SeriesCollection(2).XValues := '=Результат!R2C1:R'+
    IntToStr(rowcnt+1)+'C1';

```

```

xls.XlApp.ActiveChart.SeriesCollection(2).Values := '=Результат!R2C3:R'+
  IntToStr(rowcnt+1)+'C3';
xls.XlApp.ActiveChart.SeriesCollection(2).Name := 'Линейная';
xls.XlApp.ActiveChart.SeriesCollection(3).XValues := '=Результат!R2C1:R'+
  IntToStr(rowcnt+1)+'C1';
xls.XlApp.ActiveChart.SeriesCollection(3).Values := '=Результат!R2C4:R'+
  IntToStr(rowcnt+1)+'C4';
xls.XlApp.ActiveChart.SeriesCollection(3).Name := 'Гипербола';
xls.XlApp.ActiveChart.SeriesCollection(4).XValues := '=Результат!R2C1:R'+
  IntToStr(rowcnt+1)+'C1';
xls.XlApp.ActiveChart.SeriesCollection(4).Values := '=Результат!R2C5:R'+
  IntToStr(rowcnt+1)+'C5';
xls.XlApp.ActiveChart.SeriesCollection(4).Name := 'Степенная';
xls.XlApp.ActiveChart.SeriesCollection(5).XValues := '=Результат!R2C1:R'+
  IntToStr(rowcnt+1)+'C1';
xls.XlApp.ActiveChart.SeriesCollection(5).Values := '=Результат!R2C6:R'+
  IntToStr(rowcnt+1)+'C6';
xls.XlApp.ActiveChart.SeriesCollection(5).Name := 'Показательная';
xls.XlApp.ActiveChart.SeriesCollection(6).XValues := '=Результат!R2C1:R'+
  IntToStr(rowcnt+1)+'C1';
xls.XlApp.ActiveChart.SeriesCollection(6).Values := '=Результат!R2C7:R'+
  IntToStr(rowcnt+1)+'C7';
xls.XlApp.ActiveChart.SeriesCollection(6).Name := 'Экспоненциальная';
{
  Charts.Add
  ActiveChart.ChartType = xlLineMarkers
  ActiveChart.SetSourceData Source:=Sheets("Лист2").Range("A2:A10"), PlotBy:= _
    xlColumns
  ActiveChart.SeriesCollection.NewSeries
  ActiveChart.SeriesCollection(1).XValues = "=Лист2!R2C1:R10C1"
  ActiveChart.SeriesCollection(1).Name = """"R""""
  ActiveChart.SeriesCollection(2).XValues = "=Лист2!R2C1:R10C1"
  ActiveChart.SeriesCollection(2).Values = "=Лист2!R2C2:R10C2"
  ActiveChart.Location Where:=xlLocationAsObject, Name:="Лист1"
  With ActiveChart
    .HasTitle = True
    .ChartTitle.Characters.Text = "Диагр"
    .Axes(xlCategory, xlPrimary).HasTitle = True
    .Axes(xlCategory, xlPrimary).AxisTitle.Characters.Text = "X"
    .Axes(xlValue, xlPrimary).HasTitle = True
    .Axes(xlValue, xlPrimary).AxisTitle.Characters.Text = "Y"
  End With
  ActiveChart.HasDataTable = False
}
xls.Free;

```

end;

```
function TMDIChild.MainFunc(a, b, x: real; FuncType:integer): real;
```

```
begin
```

```
  Result:=a+b*x;
```

```
end;
```

```
function TMDIChild.MainFunc1(a, b, x: real; FuncType:integer): real;
```

```
begin
```

```
  Result:=0;
```

```
  case FuncType of
```

```
    1:Result:=a+b*x;
```

```
    2:Result:=a+b/x;
```

```
    3:Result:=a*exp(b*ln(x));
```

```
    4:Result:=a*exp(x*ln(b));
```

```
    5:Result:=exp(a+b*x);
```

```
  end;
```

end;

```
procedure TMDIChild.Method_Line(var S_ost: real;FuncType:integer);
```

```
var
```

```
  KoefA,KoefB,CovXY,Xavg,Yavg,X2avg,{ YXavg, } Y2avg,Sigma2X,Sigma2Y,Rxy:real;
```

```
  Aavg,S_fact,F_Fisher,F_FisherTab,M_a,M_b,tb_Student,ta_Student:real;
```

```
  i,n:integer;
```

```
  sform,stesnota:string;
```

```
begin
```

```
  n:=GridXY.RowCount-1;
```

```
  Xavg:=0;
```

```
  Yavg:=0;
```

```
  X2avg:=0;
```

```
  //YXavg:=0;
```

```
  Y2avg:=0;
```

```
  sform:='';
```

```
  case FuncType of
```

```
    1:sform:='a+b*x';
```

```
    2:sform:='a+b/x';
```

```
    3:sform:='a*x^b';
```

```
    4:sform:='a*b^x';
```

```
    5:sform:='exp(a+b*x)';
```

```
  end;
```

```
  MemoRes.Lines.Add('Формула аппроксимации: '+sform);
```

```
  for i:=1 to n do
```

```
  begin
```

```
    xInp[i]:=xyInp[i,1];
```



```

yInp[i]:=xyInp[i,2];
case FuncType of
  2:xInp[i]:=1/xInp[i]; //a+b/x
  3:begin //a*x^b
    xInp[i]:=ln(xInp[i]);
    yInp[i]:=ln(yInp[i]);
  end;
  4:begin //a*b^x
    yInp[i]:=ln(yInp[i]);
  end;
  5:begin //exp(a+b*x)
    yInp[i]:=ln(yInp[i]);
  end;
end;
Xavg:=Xavg+xInp[i];
X2avg:=X2avg+sqr(xInp[i]);
Yavg:=Yavg+yInp[i];
Y2avg:=Y2avg+sqr(yInp[i]);
//YXavg:=YXavg+xInp[i]*yInp[i];
end;
Xavg:=Xavg/n;
X2avg:=X2avg/n;
Yavg:=Yavg/n;
Y2avg:=Y2avg/n;
//YXavg:=YXavg/n;
CovXY:=0;
for i:=1 to n do
  CovXY:=CovXY+(xInp[i]-Xavg)*(yInp[i]-Yavg);
CovXY:=CovXY/n;
Sigma2X:=X2avg-sqr(Xavg);
Sigma2Y:=Y2avg-sqr(Yavg);
KoefB:=CovXY/Sigma2X;
KoefA:=Yavg-KoefB*Xavg;
Rxy:=KoefB*sqrt(Sigma2X/Sigma2Y);
stesnota:="";
case abs(trunc(Rxy*100)) of
  0..9:stesnota:='Весьма слабая';
  10..30:stesnota:='Слабая';
  31..50:stesnota:='Умеренная';
  51..70:stesnota:='Заметная';
  71..90:stesnota:='Высокая';
  91..100:stesnota:='Весьма высокая';
end;
MemoRes.Lines.Add('Теснота линейной связи между переменными: '+
stesnota+ifs(Rxy<0,' обратная',' прямая'));

```

```

MemoRes.Lines.Add('Линейный коэффициент корреляции:'+FloatStrT(Rxy,6));
Aavg:=0;
S_ost:=0;
S_fact:=0;
for i:=1 to n do
begin
  yRes[i,FuncType]:=MainFunc(KoefA,KoefB,xInp[i],FuncType);
  if abs(yInp[i])>0.000000001 then
    Aavg:=Aavg+abs((yInp[i]-yRes[i,FuncType])/yInp[i]);
    S_ost:=S_ost+sqr(yInp[i]-yRes[i,FuncType]);
    S_fact:=S_fact+sqr(yRes[i,FuncType]-Yavg);
  end;
S_ost:=S_ost/(n-2);
Aavg:=Aavg/n*100;
MemoRes.Lines.Add('Средняя ошибка аппроксимации:'+FloatStrT(Aavg,6)+'%');
MemoRes.Lines.Add('Факторная дисперсия на одну степень свободы:'+
  FloatStr0(S_fact,6));
MemoRes.Lines.Add('Остаточная дисперсия на одну степень свободы:'+
  FloatStr0(S_ost,6));
F_Fisher:=S_fact/S_ost;
F_FisherTab:=GetFisherKoef(1,n-2);
MemoRes.Lines.Add('F-критерий Фишера:'+FloatStr0(F_Fisher,6)+
  ' Табличное значение =' +FloatStr0(F_FisherTab,6));
M_b:=sqrt(S_ost/Sigma2X/n);
tb_Student:=KoefB/M_b;
M_a:=sqrt(S_ost*X2avg*n/Sigma2X)/n;
ta_Student:=KoefA/M_a;
MemoRes.Lines.Add('Критерий Стьюдента: ta=' +FloatStr0(ta_Student,6)+
  ' tb=' +FloatStr0(tb_Student,6));
MemoRes.Lines.Add('Ma=' +FloatStr0(m_a,6)+ ' Mb=' +FloatStr0(m_b,6));
case FuncType of
3:KoefA:=exp(KoefA);
4:begin
  KoefA:=exp(KoefA);
  KoefB:=exp(KoefB);
end;
5:begin
  //KoefA:=exp(KoefA);
end;
end;
for i:=1 to n do
begin
  yRes[i,FuncType]:=MainFunc1(KoefA,KoefB,xyInp[i,1],FuncType);
end;
MemoRes.Lines.Add('Коэффициенты A=' +FloatStrT(KoefA,6)+ ' B=' +FloatStrT(KoefB,6));

```

end;

end.