Э.Е. Тихонов

Методы прогнозирования в условиях рынка

УЧЕБНОЕ ПОСОБИЕ

Невинномысск, 2006

УДК [338.26+004.67](075.8)

Э.Е. Тихонов. Методы прогнозирования в условиях рынка: учебное пособие. - Невинномысск, 2006. - 221 с.

В учебном пособии рассмотрены вопросы практического применения методов прогнозирования. Особенность данного издания - рассмотрение концепций применения методов прогнозирования одновременно с вопросами их практической реализации в современных программных средствах MS Excel, Statistica, Statistica Neural Networks.

Пособие состоит из четырех частей. В первой части сделан аналитический обзор методов и моделей прогнозирования. Во второй части на примерах разобраны адаптивные методы прогнозирования, модель Уинтерса (экспоненциального сглаживания с мультипликативной сезонностью и линейным ростом) и модель Тейла-Вейджа. В третьей части подробно разобраны вопросы практической реализации линейных и нелинейных многофакторных моделей в пакете Statistica. В четвертой части описаны практические аспекты нейросетевого прогнозирования с использованием пакета Statistica Neural Networks. Каждый раздел заканчивается контрольными вопросами и заданиями для самостоятельной работы.

Пособие предназначено для студентов, изучающих дисциплины: рынок ценных бумаг, менеджмент и компьютерная поддержка принятия решений, прогнозирование и планирование в задачах производственного менеджмента.

Рецензенты: д-р техн. наук, проф. *Н.И. Червяков*
д-р техн. наук, проф. *В.Ф. Минаков*

ISBN 5-89571-077-8

© Северо-Кавказский государственный технический университет, 2006
© Э.Е. Тихонов, 2006

Оглавление

Введение…… 6

Глава 1. Аналитический обзор моделей и методов прогнозирования……. 8
1.1. Прогнозная экстраполяция…… 11
1.1.1. Метод наименьших квадратов……. 13
1.1.2. Метод экспоненциального сглаживания…… 15
1.1.3. Метод вероятностного моделирования……... 19
1.2. Интуитивные (экспертные) методы прогнозирования…... 24
1.3. Корреляционный и регрессионный анализы……... 33
1.4. Модели стационарных временных рядов и их идентификация…….. 40
1.4.1. Модели авторегрессии порядка p (AR(p)-модели)… 41
1.4.2. Модели скользящего среднего порядка q (MA(q)-модели)……. 43
1.4.3. Авторегрессионные модели со скользящими средними в остатках (ARMA(p, q)-модели)……. 45
1.5. Модели нестационарных временных рядов и их идентификация……... 46
1.5.1. Модель авторегрессии-проинтегрированного скользящего среднего (ARIMA(p, k, q)-модель)…….. 46
1.5.2. Модели рядов, содержащих сезонную компоненту.. 48
1.5.3. Прогнозирование на базе ARIMA-моделей……... 49
1.6. Адаптивные методы прогнозирования…… 51
1.7. Метод группового учета аргументов…… 55
1.8. Теория распознавания образов……. 62
1.9. Прогнозирование с использованием нейронных сетей, искусственного интеллекта и генетических алгоритмов…….. 63
Глава 2. Практическая реализация адаптивных методов прогнозирования…… 71
2.1. Общие положения……...……... 71
2.2. Полиномиальные модели временных рядов. Метод экспоненциальной средней…… 72
2.2.1. Адаптивная полиномиальная модель нулевого порядка (p=0)……... 73
2.2.2. Адаптивная полиномиальная модель первого порядка (p=1)……. 76

2.2.3. Адаптивная полиномиальная модель второго порядка (p=2)…….. 80
2.3. Прогнозирование с использованием модели Уинтерса (экспоненциального сглаживания с мультипликативной сезонностью и линейным ростом)…….. 83
2.4. Прогнозирование объема производства по модели Тейла-Вейджа…… 91

Глава 3. Практическая реализация многофакторных моделей прогнозирования……. 97
3.1. Линейные многофакторные модели……. 99
3.2. Нелинейные многофакторные модели……. 111

Глава 4. Нейросетевое прогнозирование экономических показателей в пакете Statistica Neural Networks…….. 122
4.1. Основные возможности пакета Statistica Neural Networks……... 122
4.1.1. Создание набора данных…….. 122
4.1.2. Добавление наблюдений…….. 123
4.1.3. Удаление лишних наблюдений……... 124
4.1.4. Изменение переменных и наблюдений……... 124
4.1.5. Другие возможности редактирования данных……... 124
4.1.6. Создание новой сети……. 126
4.1.7. Создание сети…… 126
4.1.8. Сохранение набора данных и сети…….. 127
4.1.9. Обучение сети……... 127
4.1.10. Оптимизация обучения…….. 129
4.1.11. Выполнение повторных прогонов……. 129
4.1.12. Ошибки для отдельных наблюдений…… 130
4.2. Запуск нейронной сети…….. 131
4.2.1. Обработка наблюдений по одному……. 132
4.2.2. Прогон всего набора данных…….. 133
4.2.3. Тестирование на отдельном наблюдении……... 134
4.3. Создание сети типа Многослойный персептрон……. 134
4.3.1. Создание сети типа многослойный персептрон с помощью мастера…… 135
4.4. Построение нейронной сети без мастера……. 147
4.5. Обучение сети……. 149
4.5.1. Обучение методом обратного распространения ошибки 149
4.5.2. Обучение с помощью метода Левенберга-Маркара.. 155
4.5.3. Алгоритм выполнения обучения сети с помощью 159

метода Левенберга-Маркара……..
4.6. Генетические алгоритмы отбора входных данных……. 159
4.7. Применение нейронных сетей в задачах прогнозирования и проблемы идентификации моделей прогнозирования на нейронных сетях…….……. 166
4.7.1. Сравнительный анализ нейронных сетей.……. 168
4.7.2. Исследование нейросетевых структур для курсов акций ОАО «Ростелеком», ОАО «Лукойл», «Сбербанк»……. 174
4.8. Сравнительная оценка классических и нейросетевых методов прогнозирования……. 182

Литература…… 188
Приложения……. 206

Введение

Развитие прогностики как науки в последние десятилетия привело к созданию множества методов, процедур, приемов прогнозирования, неравноценных по своему значению. По оценкам зарубежных и отечественных систематиков прогностики уже насчитывается свыше ста методов прогнозирования, в связи с чем перед специалистами возникает задача выбора методов, которые давали бы адекватные прогнозы для изучаемых процессов или систем [19, 27].

Для тех, кто не является специалистами в прикладной математике, эконометрике, статистике, применение большинства методов прогнозирования вызывает трудности при их реализации с целью получения качественных и точных прогнозов. В связи с этим, каждый метод рассмотрен очень подробно с приведением рекомендаций по практическому применению.

Особенностью данного пособия является рассмотрение тонкостей применения того или иного метода. Учебное пособие разделено на четыре части, в каждой из которых рассмотрен свой класс методов прогнозирования.

В первой части рассмотрены теоретические аспекты построения и применения методов и алгоритмов прогнозирования. Приведена классификация наиболее распространенных методов.

Во второй части рассмотрены классические адаптивные модели прогнозирования, реализованные в MS Excel. Несмотря на то, что они программно реализованы в некоторых статистических и эконометрических пакетах прикладных программ, предложен именно ручной счет, освоив который гораздо легче понимать принципы и специфику данных методов прогнозирования.

В третьей части основное внимание уделено применению классических нелинейных многофакторных моделей прогнозирования. Совершено очевидно, что сложные нелинейные многофакторные модели невозможно просчитать вручную, поэтому подробно рассматривается возможность применения пакета Statistica для этих целей.

В четвертой части рассмотрены нейросетевые методы прогнозирования и особенности их построения. Многие источники подробно рассматривают теорию нейронных сетей, опуская описание практического их использования.

Для понимания того, какие преимущества дают предлагаемые методы анализа данных и прогнозирования, необходимо указать на

три принципиальные проблемы, возникающие при прогнозировании.

Первая проблема – это определение необходимых и достаточных параметров для оценки состояния исследуемой предметной области.

Вторая проблема заключается в так называемом «проклятье размерности». Желание учесть в модели как можно больше показателей и критериев оценки может привести к тому, что требуемая для ее решения компьютерная система вплотную приблизится к «пределу Тьюринга» (ограничению на быстродействие и размеры вычислительного комплекса в зависимости от количества информации, обрабатываемой в единицу времени).

Третья проблема – наличие феномена «надсистемности». Взаимодействующие системы образуют систему более высокого уровня, обладающую собственными свойствами, что делает принципиально недостижимой возможность надсистемного отображения и целевых функций с точки зрения систем, входящих в состав надсистемы.

Для преодоления перечисленных проблем делаются попытки применения таких разделов современной фундаментальной и вычислительной математики, как нейрокомпьютеры, теория стохастического моделирования (теория хаоса), теория рисков, теория катастроф, синергетика и теория самоорганизующихся систем (включая генетические алгоритмы) [123, 134]. Считается, что эти методы позволят увеличить глубину прогноза за счет выявления скрытых закономерностей и взаимосвязей среди плохо формализуемых обычными методами макроэкономических, политических и глобальных финансовых показателей.

Представленное учебное пособие может быть рекомендовано для студентов, аспирантов и преподавателей, занимающихся проблемами совершенствования методов и моделей прогнозирования, а также вопросами их практической реализации.

ГЛАВА 1. АНАЛИТИЧЕСКИЙ ОБЗОР МОДЕЛЕЙ И МЕТОДОВ ПРОГНОЗИРОВАНИЯ

Как было сказано выше, по оценкам зарубежных и отечественных систематиков прогностики, уже насчитывается свыше 100 методов прогнозирования. Число базовых методов прогностики, которые в тех или иных вариациях повторяются в других методах, гораздо меньше. Многие из этих «методов» относятся скорее к отдельным приемам или процедурам прогнозирования, другие представляют набор отдельных приемов, отличающихся от базовых или друг от друга количеством частных приемов и последовательностью их применения. О проблеме классификации было отмечено во введении.

В литературе имеется большое количество классификационных схем методов прогнозирования. Однако большинство из них или неприемлемы, или обладают недостаточной познавательной ценностью. Основной погрешностью существующих классификационных схем является нарушение принципов классификации. К числу основных таких принципов, на наш взгляд, относятся: достаточная полнота охвата прогностических методов, единство классификационного признака на каждом уровне членения, открытость классификационной. Предлагаемая нами схема классификации методов прогнозирования показана на рисунке 1.1 [19].

Безусловно, имеют право на существование частные классификационные схемы, предназначенные для определенной цели или задачи.

Каждый уровень детализации определяется своим классификационным признаком: степенью формализации, общим принципом действия, способом получения прогнозной информации.

По степени формализации все методы прогнозирования делятся на интуитивные и формализованные. Интуитивное прогнозирование применяется тогда, когда объект прогнозирования либо слишком прост, либо настолько сложен, что аналитически учесть влияние многих факторов практически невозможно. В этих случаях прибегают к опросу экспертов, Полученные индивидуальные и коллективные экспертные оценки используют как конечные прогнозы или в качестве исходных данных в комплексных системах прогнозирования.

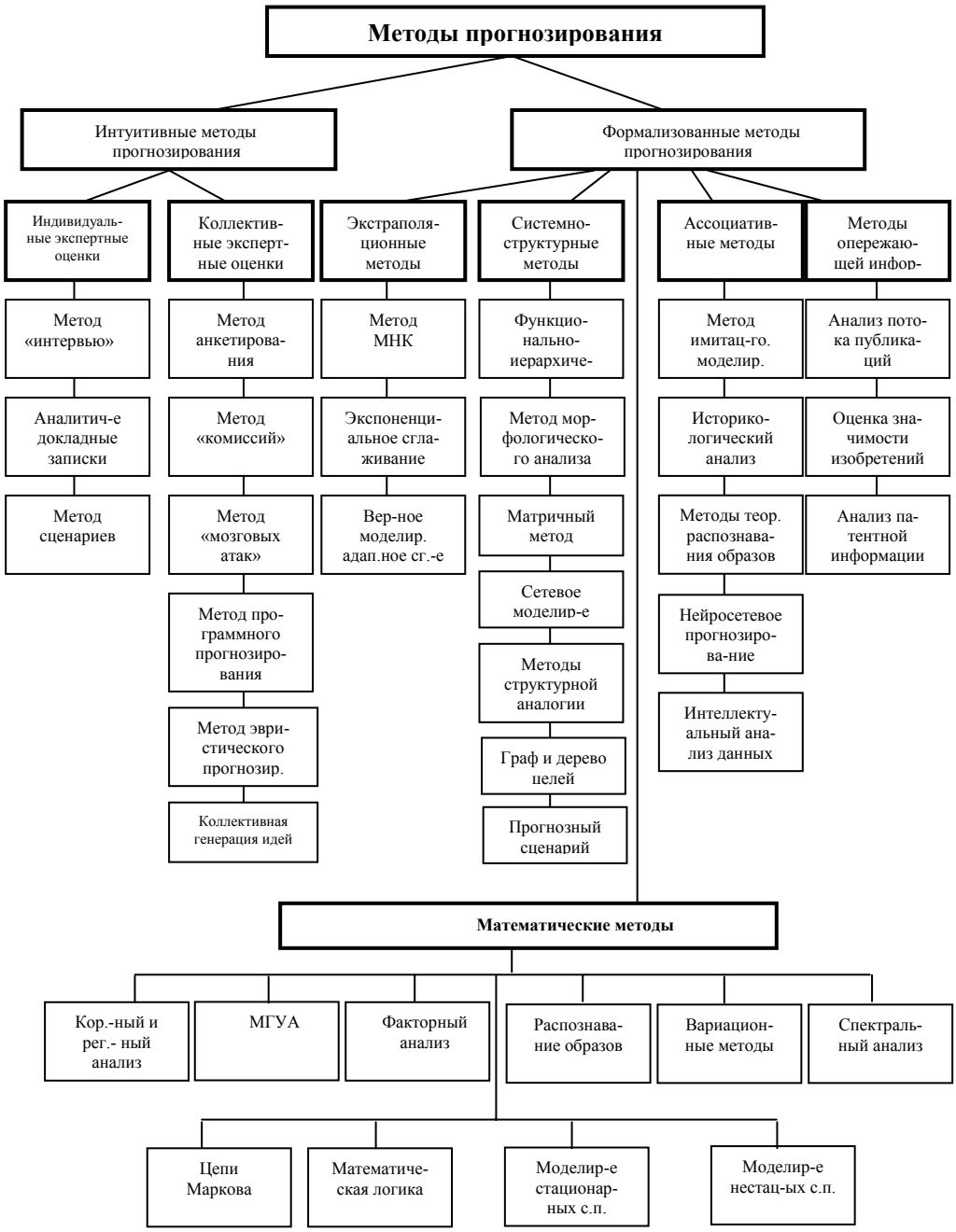


Рисунок 1.1. Классификационная схема методов прогнозирования

В выборе методов прогнозирования важным показателем является глубина упреждения прогноза. При этом необходимо не только знать абсолютную величину этого показателя, но и отнести его к длительности эволюционного цикла развития объекта прогнозирования. Для этого можно использовать предложенный В. Белоконем безразмерный показатель глубины (дальности) прогнозирования (τ)

$$\tau = \Delta t / t ,$$

где Δt – абсолютное время упреждения; t_x – величина эволюционного цикла объекта прогнозирования.

Формализованные методы прогнозирования являются действенными, если величина глубины упреждения укладывается в рамки эволюционного цикла ($\tau << 1$). При возникновении в рамках прогнозного периода «скачка» в развитии объекта прогнозирования ($\tau \approx 1$) необходимо использовать интуитивные методы, как для определения силы «скачка», так и для оценки времени его осуществления, либо теорию катастроф [59]. В этом случае формализованные методы применяются для оценки эволюционных участков развития до и после скачка. Если же в прогнозном периоде укладывается несколько эволюционных циклов развития объекта прогнозирования ($\tau >> 1$), то при комплексировании систем прогнозирования большее значение имеют интуитивные методы.

В зависимости от общих принципов действия интуитивные методы прогнозирования, например, можно разделить на две группы: индивидуальные экспертные оценки и коллективные экспертные оценки.

Методы коллективных экспертных оценок уже можно отнести к комплексным системам прогнозирования (обычно неполным), поскольку в последних сочетаются методы индивидуальных экспертных оценок и статистические методы обработки этих оценок. Но так как статистические методы применяются во вспомогательных процедурах выработки прогнозной информации, на наш взгляд, коллективные экспертные оценки целесообразнее отнести к сингулярным методам прогнозирования.

В группу индивидуальных экспертных оценок можно включить (принцип классификации – способ получения прогнозной информации) следующие методы: метод «интервью», аналитические докладные записки, написание сценария. В группу коллективных эксперт-

ных оценок входят анкетирование, методы «комиссий», «мозговых атак» (коллективной генерации идей).

Класс формализованных методов в зависимости от общих принципов действия можно разделить на группы экстраполяционных, системно-структурных, ассоциативных методов и методов опережающей информации.

В группу методов прогнозной экстраполяции можно включить методы наименьших квадратов, экспоненциального сглаживания, вероятностного моделирования и адаптивного сглаживания. К группе системно-структурных методов – отнести методы функционально-иерархического моделирования, морфологического анализа, матричный, сетевого моделирования, структурной аналогии. Ассоциативные методы можно разделить на методы имитационного моделирования и историко-логического анализа. В группу методов опережающей информации – включить методы анализа потоков публикаций, оценки значимости изобретений и анализа патентной информации.

Представленный перечень методов и их групп не является исчерпывающим. Нижние уровни классификации открыты для внесения новых элементов, которые могут появиться в процессе дальнейшего развития инструментария прогностики.

Некоторые не названные здесь методы являются или разновидностью включенных в схему методов, или дальнейшей их конкретизацией.

1. *На основе каких признаков можно классифицировать методы прогнозирования?*
2. *На какие классы можно разделить методы прогнозирования?*
3. *В чем особенность выбора глубины упреждения прогноза?*

1.1. Прогнозная экстраполяция

В методическом плане основным инструментом любого прогноза является схема экстраполяции. Различают формальную и прогнозную экстраполяцию. Формальная базируется на предположении о сохранении в будущем прошлых и настоящих тенденций развития объекта прогноза. При прогнозной экстраполяции фактическое развитие увязывается с гипотезами о динамике исследуемого процесса с учетом в перспективе его физической и логической сущности.

Основу экстраполяционных методов прогнозирования составляет изучение временных рядов, представляющих собой упорядоченные во времени наборы измерений тех или иных характеристик исследуемого объекта, процесса.

Временной ряд y_t может быть представлен в следующем виде

$$y_t = x_t + S + C + \varepsilon_t \tag{1.1}$$

где x_t – детерминированная неслучайная компонента процесса (тренд); S – сезонная составляющая; C – циклическая составляющая; ε_t – стохастическая компонента процесса.

Если детерминированная компонента (тренд) x_t характеризует существующую динамику развития процесса в целом, то стохастическая компонента ε_t отражает случайные колебания или шумы процесса. Обе составляющие процесса определяются каким-либо функциональным механизмом, характеризующим их поведение во времени. Задача прогноза состоит в определении вида экстраполирующих функций x_t, сезонной и циклической составляющей, и ε_t на основе исходных эмпирических данных.

Первым этапом экстраполяции тренда является выбор оптимального вида функции, описывающей эмпирический ряд. Для этого проводятся предварительная обработка и преобразование исходных данных с целью облегчения выбора вида тренда путем сглаживания и выравнивания временного ряда, определения функций дифференциального роста, а также формального и логического анализа особенностей процесса. Следующим этапом является расчет параметров выбранной экстраполяционной функции.

Наиболее распространенными методами оценки параметров зависимостей являются метод наименьших квадратов и его модификации, метод экспоненциального сглаживания, метод вероятностного моделирования и метод адаптивного сглаживания.

1. *Назовите виды экстраполяции. В чем разница между экстраполяцией и интерполяцией?*
2. *Назовите основные компоненты временного ряда.*

1.1.1. Метод наименьших квадратов

Сущность метода наименьших квадратов состоит в отыскании параметров модели тренда, минимизирующих ее отклонение от точек исходного временного ряда, т. е.

$$S = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \to \min \qquad (1.2)$$

где $\hat{y}_i$ – расчетные значения исходного ряда; y_i – фактические значения исходного ряда; n – число наблюдений. Если модель тренда представить в виде

$$\hat{y} = f(x_i; a_1, a_2, ..., a_k, t), \qquad (1.3)$$

где $a_1, a_2, ..., a_k$ – параметры модели; t – время; x_i - независимые переменные, то для того, чтобы найти параметры модели, удовлетворяющие условию (1.2), необходимо приравнять нулю первые производные величины S по каждому из коэффициентов a_j. Решая полученную систему уравнений с k неизвестными, находим значения коэффициентов a_j.

Использование процедуры оценки, основанной на методе наименьших квадратов, предполагает обязательное удовлетворение целого ряда предпосылок, невыполнение которых может привести к значительным ошибкам.

1. Случайные ошибки имеют нулевую среднюю, конечные дисперсии и ковариации.

2. Каждое измерение случайной ошибки характеризуется нулевым средним, не зависящим от значений наблюдаемых переменных.

3. Дисперсии каждой случайной ошибки одинаковы, их величины независимы от значений наблюдаемых переменных (гомоскедастичность).

4. Отсутствие автокорреляции ошибок, т. е. значения ошибок различных наблюдений независимы друг от друга.

5. Нормальность. Случайные ошибки имеют нормальное распределение.

6. Значения эндогенной переменной x свободны от ошибок измерения и имеют конечные средние значения и дисперсии.

В практических исследованиях в качестве модели тренда в основном используют следующие функции: линейную y = ax + b; квадратичную y = $ax^2 + bx + c$; степенную y = ax^n; показательную y = a^x; экспоненциальную y = ae^x; логистическую y = $\frac{a}{1+be^{-ex}}$.

Особенно широко применяется линейная, или линеаризуемая, т. е. сводимая к линейной, форма как наиболее простая и в достаточной степени удовлетворяющая исходным данным.

Выбор модели в каждом конкретном случае осуществляется по целому ряду статистических критериев, например по дисперсии, корреляционному отношению и др. Следует отметить, что названные критерии являются критериями аппроксимации, а не прогноза. Однако, принимая во внимание принятую гипотезу об устойчивости процесса в будущем, можно предполагать, что в этих условиях модель, наиболее удачная для аппроксимации, будет наилучшей и для прогноза.

Классический метод наименьших квадратов предполагает равноценность исходной информации в модели. В реальной же практике будущее поведение процесса значительно в большей степени определяется поздними наблюдениями, чем ранними. Это обстоятельство породило так называемое дисконтирование, т. е. уменьшение ценности более ранней информации. Дисконтирование можно учесть путем введения в модель (1) некоторых весов $\beta < 1$. Тогда

$$S = \sum_{i=1}^{n} \beta_i (\hat{y}_i - y_i)^2 \to \min \qquad (1.4)$$

Коэффициенты β, могут задаваться заранее в числовой форме или в виде функциональной зависимости таким образом, чтобы по мере продвижения в прошлое веса убывали, например $\beta_i = a^i$, где a < 1. К сожалению, формальных процедур выбора параметра не разработано, и он выбирается исследователем произвольно.

Метод наименьших квадратов широко применяется для получения конкретных прогнозов, что объясняется его простотой и легкостью реализации на ЭВМ. Недостаток метода состоит в том, что модель тренда жестко фиксируется, и с помощью МНК можно получить надежный прогноз на небольшой период упреждения. Поэтому МНК относится главным образом к методам краткосрочного прогнозирования. Кроме того, существенной трудностью МНК является

правильный выбор вида модели, а также обоснование и выбор весов во взвешенном методе наименьших квадратов.

1. *В чем состоит сущность метода наименьших квадратов?*
2. *Назовите основные предпосылки МНК, уточнение которых является обязательным для получения наилучших оценок параметров временного ряда.*
3. *В чем состоят достоинства МНК?*

1.1.2. Метод экспоненциального сглаживания

Весьма эффективным и надежным методом прогнозирования является экспоненциальное сглаживание. Основные достоинства метода состоят в возможности учета весов исходной информации, в простоте вычислительных операций, в гибкости описания различных динамик процессов. Метод экспоненциального сглаживания дает возможность получить оценку параметров тренда, характеризующих не средний уровень процесса, а тенденцию, сложившуюся к моменту последнего наблюдения. Наибольшее применение метод нашел для реализации среднесрочных прогнозов. Для метода экспоненциального сглаживания основным и наиболее трудным моментом является выбор параметра сглаживания α, начальных условий и степени прогнозирующего полинома [6,64,72,151].

Пусть исходный динамический ряд описывается уравнением

$$y_t = a_0 + a_1 t + \frac{a_2}{2} t^2 + \ldots + \frac{a_p}{p!} t^p + \varepsilon_t \cdot \tag{1.5}$$

Метод экспоненциального сглаживания, являющийся обобщением метода скользящего среднего, позволяет построить такое описание процесса (1.5), при котором более поздним наблюдениям придаются большие веса по сравнению с ранними наблюдениями, причем веса наблюдений убывают по экспоненте. Выражение

$$S_t^{[k]}(y) = \alpha \sum_{i=0}^{n} (1-\alpha)^i S_{t-1}^{[k]}(y) \tag{1.6}$$

называется экспоненциальной средней k-го порядка для ряда y_t, где α – параметр сглаживания.

В расчетах для определения экспоненциальной средней пользуются рекуррентной формулой [151]

$$S_t^{[k]}(y) = \alpha S_t^{[k-1]}(y) + (1-\alpha)^i S_{t-1}^{(k)}(y). \tag{1.7}$$

Использование соотношения (1.7) предполагает задание начальных условий $S_0^{[1]}, S_0^{[2]}, ..., S_0^{[k]}$. Для этого можно воспользоваться формулой Брауна– Мейера, связывающей коэффициенты прогнозирующего полинома с экспоненциальными средними соответствующих порядков

$$S_t^{[k]} = \sum_{p=0}^{n} (-1)^p \frac{\hat{a}_p}{p!} \frac{\alpha\beta}{(k-1)!} \sum_{j=0}^{\infty} j^p \beta^j \frac{(p-1+j)!}{j!}, \tag{1.8}$$

где p = 1, 2, ..., n + 1; $\hat{a}_p$ – оценки коэффициентов; $p = 1-\alpha$. Можно получить оценки начальных условий, в частности, для линейной модели

$$\begin{aligned} S_0^{[1]} &= a_0 - \frac{\beta}{\alpha} a_1; \\ S_0^{[2]} &= a_0 - \frac{2\beta}{\alpha} a_1; \end{aligned} \tag{1.9}$$

для квадратичной модели –

$$\begin{aligned} S_0^{[1]} &= a_0 - \frac{\beta}{\alpha} a_1 + \frac{\beta(2-\alpha)}{2\alpha^2} a_2; \\ S_0^{[2]} &= a_0 - \frac{2\beta}{\alpha} a_1 + \frac{\beta(3-2\alpha)}{2\alpha^2} a_2; \\ S_0^{[3]} &= a_0 - \frac{3\beta}{\alpha} a_1 + \frac{\beta(4-3\alpha)}{2\alpha^2} a_2. \end{aligned} \tag{1.10}$$

Зная начальные условия $S_0^{[k]}$ и значения параметра α, можно вычислить экспоненциальные средние $S_t^{[k]}$.

Оценки коэффициентов прогнозирующего полинома определяются через экспоненциальные средние по фундаментальной теореме Брауна – Мейера. В этом случае коэффициенты и находятся решением системы (p + 1) уравнений с к (p + 1) неизвестными, связывающей параметры прогнозирующего полинома с исходной информацией. Так, для линейной модели получим,

$$\hat{a}_0 = 2S_t^{[1]} - S_t^{[2]}; \qquad (1.11)$$
$$\hat{a}_1 = \frac{\alpha}{\beta}\left(S_t^{[1]} - S_t^{[2]}\right);$$

для квадратичной модели –

$$\hat{a}_0 = 3\left(S_t^{[1]} - S_t^{[2]}\right) + S_t^{[3]};$$
$$\hat{a}_1 = \frac{\alpha}{\beta^2}\left[(6-5\alpha)S_t^{[1]} - 2(5-4\alpha)S_t^{[2]} + (4-3\alpha)S_t^{[3]}\right]; \qquad (1.12)$$
$$\hat{a}_2 = \frac{\alpha^2}{\beta^2}\left[S_t^{[1]} - 2S_t^{[2]} + S_t^{[3]}\right].$$

Прогноз реализуется по выбранному многочлену. Соответственно для линейной модели $\hat{y}_{t+\tau} = \hat{a}_0 + \hat{a}_1\tau$, для квадратичной модели $\hat{y}_{t+\tau} = \hat{a}_0 + \hat{a}_1\tau + \frac{\hat{a}_2}{2}\tau^2$, где τ - период прогноза.

Важную роль в методе экспоненциального сглаживания играет выбор оптимального параметра сглаживания α, так как именно он определяет оценки коэффициентов модели, а, следовательно, и результаты прогноза [72, 103, 215].

В зависимости от величины параметра прогнозные оценки поразному учитывают влияние исходного ряда наблюдений: чем больше α, тем больше вклад последних наблюдений в формирование тренда, а влияние начальных условий быстро убывает. При малом α прогнозные оценки учитывают все наблюдения, при этом уменьшение влияния более «старой» информации происходит медленно.

Известны два основных соотношения, позволяющие найти приближенную оценку α. Первое соотношение Брауна, выведенное из условия равенства скользящей и экспоненциальной средней $\alpha = \frac{2}{N+1}$, где N – число точек ряда, для которых динамика ряда считается однородной и устойчивой (период сглаживания). Вторым является соотношение Мейера $\alpha \approx \frac{\sigma_n}{\sigma_\varepsilon}$, где σ_n - среднеквадратическая ошибка модели; σ_ε - среднеквадратическая ошибка исходного ряда. Однако использование последнего соотношения затруднено тем, что

достоверно определить σ_n и σ_ε из исходной информации очень сложно.

Выбор параметра α целесообразно связывать с точностью прогноза, поэтому для более обоснованного выбора α можно использовать процедуру обобщенного сглаживания, которая позволяет получить следующие соотношения, связывающие дисперсию прогноза и параметр сглаживания [103, 129].

Для линейной модели –

$$\sigma^2_{\hat{x}_\tau} = \frac{\alpha}{(1+\beta)^2}\left[1+4\beta+2\beta^2+2\alpha(1+3\beta)\tau+2\alpha^2\tau^3\right]\sigma^2_\varepsilon . \quad (1.13)$$

Для квадратичной модели –

$$\sigma^2_{\hat{x}_\tau} \approx \left[2\alpha+3\alpha^3+3\alpha^2\tau\right]\sigma^2_\varepsilon . \quad (1.14)$$

Для обобщенной модели вида

$$y(t) = \sum_{i=1}^{n} a_i f_i(t) + \varepsilon_t . \quad (1.15)$$

Дисперсия прогноза имеет следующий вид

$$\sigma^2_{\hat{x}_\tau} = \sum_{j=1}^{n}\sum_{k=1}^{n} f_i(\tau)\operatorname{cov}(a_j, a_k) f_k(\tau) = \vec{f}^T V f(\tau)\sigma^2_\varepsilon , \quad (1.6)$$

где σ_ε – среднеквадратическая ошибка аппроксимации исходного динамического ряда; $f_i(t)$ – некоторая известная функция; V– матрица ковариации коэффициентов модели.

Отличительная особенность этих формул состоит в том, что при $\alpha = 0$ они обращаются в нуль. Это объясняется тем, что, чем ближе к нулю α, тем больше длина исходного ряда наблюдений $t \to \infty$ и, следовательно, тем меньше ошибка прогноза. Поэтому для уменьшения ошибки прогноза необходимо выбирать минимальное α.

В то же время параметр α определяет начальные условия, и, чем меньше α, тем ниже точность определения начальных условий, а следовательно, ухудшается и качество прогноза. Ошибка прогноза растет по мере уменьшения точности определения начальных условий [103].

Таким образом, использование формул (1.13)-(1.16) приводит к противоречию при определении параметра сглаживания: с уменьшением α уменьшается среднеквадратическая ошибка, но при этом возрастает ошибка в начальных условиях, что в свою очередь влияет на точность прогноза.

Кроме того, при использовании соотношений (1.13)-(1.16) необходимо принимать во внимание следующие обстоятельства, а именно: эти выражения получены для бесконечно длинных рядов без учета автокорреляции наблюдений. На практике мы имеем дело с конечными рядами, характеризующимися внутренней зависимостью между исходными наблюдениями. Все это снижает целесообразность использования соотношений (1.13)-(1.16).

В ряде случаев параметр α выбирается таким образом, чтобы минимизировать ошибку прогноза, рассчитанного по ретроспективной информации.

Весьма существенным для практического использования является вопрос о выборе порядка прогнозирующего полинома, что во многом определяет качество прогноза. Превышение второго порядка модели не приводит к существенному увеличению точности прогноза, но значительно усложняет процедуру расчета [40, 53].

Рассмотренный метод является одним из наиболее надежных и широко применяется в практике прогнозирования. Одно из наиболее перспективных направлений развития данного метода представляет собой метод разностного прогнозирования, в котором само экспоненциальное сглаживание рассматривается как частный случай [129, 138].

1. *Определение какого параметра в методе экспоненциального сглаживания является основным?*
2. *Как правильно выбрать параметр сглаживания?*

1.1.3. Метод вероятностного моделирования

Прогнозирование с использованием вероятностных моделей базируется на методе экспоненциального сглаживания. Вероятностные модели по своей сути отличны от экстраполяционных моделей временных рядов, в которых основой является описание изменения во времени процесса.

Во временных рядах модели представляют собой некоторую функцию времени с коэффициентами, значения которых оценивают-

ся по наблюдениям. В вероятностных моделях оцениваются вероятности, а не коэффициенты.

Пусть определено n взаимно независимых и исключающих событий. В каждом случае наблюдения измеряются в единой шкале, помещаются в (n + 1) ограниченный класс и обозначаются так: x_1, x_2, …, x_n. Событие, связанное с наблюдением x(t), соответствует числу интервалов, в которое это событие попадает, т. е. существует единственное значение k, такое, что $x_{k-1} < x(t) \le x_k$. И поэтому k-е событие связывается с наблюдением *x(t)*.

Рассмотрим метод оценивания вероятностей $\hat{p}_n(t)$, связанных с различными событиями $x_{k-1} < x(t) \le x_k$. На первом этапе задаются начальные значения различных вероятностей: $\hat{p}_k(0)$; $k \ne 1,2,...,n$. Наблюдение *x(t)* связано с *k*-м событием следующим образом: если $x_{k-1} < x(t) \le x_k$, то строится единичный вектор $\vec{u}_k$, (k - 1) компонент которого равен 0 и *k*-й компонент равен 1. Это может быть *k*-м столбцом единичной матрицы ранга *k*.

Процесс, реализующий оценки вероятностей, описывается вектором сглаживания по формуле

$$\hat{\vec{p}}(t) = \alpha \vec{u}_k + (1-\alpha)\vec{p}(t-1). \qquad (1.17)$$

Каждая компонента вектора меняется по закону простого экспоненциального сглаживания между нулем и единицей. Если вектор $\vec{p}(t-1)$ вероятностный, то все его компоненты должны быть неотрицательными, и их сумма должна быть равна 1. Значение оценки $\vec{p}_k(t)$ есть результат экспоненциального сглаживания, и если распределение вероятностей наблюдений *x(t)* не меняется, то получаемые вероятности и будут действительными вероятностями k-го события. Если существует достаточно длительная реализация процесса, то начальные оценки со временем перестанут оказывать влияние (будут достаточно «взвешены»), и вектор сглаживания будет в среднем описывать вероятности и взаимно исключающих, и независимых событий. Значения компонент вектора $\vec{u}(t)$ представляют собой выборку с биномиальным распределением, поэтому дисперсии k-й компоненты будут $\vec{p}_k(1-\vec{p}_k)$. Дисперсия оценок k-и вероятности определяется соотношением

$$\sigma_k^2 = \frac{\alpha}{2-\alpha}\vec{p}_k(1-\vec{p}_k), \qquad (1.18)$$

где α – константа сглаживания $(0 \le \alpha \le 1)$, используемая для получения оценок вектора вероятностей.

Возможны два варианта. В первом случае пределы классов заданы так, что p_k может быть или очень большим (около 1), или очень малым (около 0). Тогда дисперсия компонент вектора вероятностей будет небольшой. Если форма распределения меняется со временем, большое значение константы сглаживания может быть использовано, чтобы устранить влияние «старой» информации.

Во втором случае распределение вероятностей постоянно во времени, нет необходимости «взвешивать» старую информацию. Малое значение константы сглаживания, может быть, позволит уменьшить дисперсию оценок. Тогда можно использовать меньшие интервалы классов с не очень большими вероятностями.

Автоматизированные системы прогнозирования требуют постоянного добавления новых значений информации. Некоторые системы могут просто накапливать информацию, затем использовать ее для прогноза. Если мы имеем дело с поступающей информацией, то система может практически бездействовать в течение значительного промежутка времени. Если информация достаточно важна, следует рассматривать ее как непрерывный во времени поток наблюдений или предсказывать распределение поступлений наблюдений. Очевидно, для таких прогнозов следует использовать модель, изложенную выше. Если в какой-то период нет никаких наблюдений, можно перестроить систему на другой вид информации. Кроме того, оценки коэффициентов (или других параметров) в модели прогноза не изменяются, если наблюдения равны нулю; соответственно и прогноз будет тем же [54, 72].

Вероятностная модель оперирует последовательностью наблюдений с учетом их распределения и игнорирует последовательность этой информации уже непосредственно во времени. Поэтому вектор вероятностей $\vec{p}(t)$, который служит текущей оценкой вероятностей отдельных событий, является оценкой этих вероятностей в будущем. Последовательность наблюдений может быть представлена как временной ряд *x(t)*, где *x* измерен по некоторой шкале $x_0 \le x \le x_n$ а x_0 и x_n есть минимум и максимум возможных значений наблюдений.

Поэтому p-й процентилью будет такое число x_p, что для p процентов времени реализуется условие $x_0 \le x(t) \le x_p$, а (100 – p) процентов $-x_p \le x(t) \le x_n$. Кумулятивная вероятность того, что наблюдение попадает левее по шкале, для данного класса записывается так

$$p_r\{x \le x_k\} = P_k = \sum_{i=1}^{k} P_i \,. \tag{1.19}$$

Очевидно, $p_0 = 0$ и $p_k = 1, 0$.

Для связи с вероятностью дается несколько иное представление, нежели процентное. Если $p = p_k$ для класса k, то в- этом случае [54]

$$\sum_{i=1}^{k} p_i(t) = \hat{p}_{k-1}(t) < p < \hat{p}_k(t) = \sum_{i=1}^{k} p_i(t) \,. \tag{1.20}$$

Кумулятивная вероятность $\hat{p}_{k-1}(t)$ для класса x_{k-1} меньше, чем желаемая вероятность, которая в свою очередь меньше, чем кумулятивная вероятность $\hat{p}_k(t)$ для следующего класса. Простейшая оценка необходимой процентили может быть получена по линейной интерполяции

$$x_p(t) = \frac{[\hat{p}_k(t) - p]x_{k-1} + (p - \hat{p}_{k-1}(t))x_k}{\hat{p}_k(t) - \hat{p}_{k-1}(t)} \tag{1.21}$$

Если классы очень малы (или p_k близко к p_{k-1}), линейная интерполяция достаточно хороша. Возможно и интерполирование по полиномам более высокой степени. Около хвостов распределения можно ожидать, что кумулятивная вероятность ведет себя как

$$p(x) = 1 - j^x \text{ или } p(x) = 1 - \delta^{x^2} \,, \tag{1.22}$$

где j или δ - числа меньше единицы. Такие функции могут быть оценены на основании имеющейся информации.

Определим дисперсию вероятностей модели следующим образом

$$\sigma_x^2 = \bar{x}^2 - x^2 \,, \tag{1.23}$$

где величина $\overline{x}^2$ может быть оценена экспоненциальным сглаживанием квадратов наблюдений. Можно считать, что наблюдения почти всегда распределены нормально. Тогда вероятностная модель может быть применена непосредственно к этим наблюдениям.

Пусть x – случайная величина с ожиданием m и конечной дисперсией σ. Тогда сумма случайных выборок x будет нормально распределена со средним и дисперсией $n^{-1}\sigma^2$, и вероятности как суммы точек наблюдений будут распределены нормально.

Пусть случайная величина x распределена между нулем и единицей. Введем функцию

$$f(x) = \begin{cases} 1, если \quad 0 \le x \le 1, \\ 0, если \quad x < 0 или \quad x > 1. \end{cases} \tag{1.24}$$

Пусть y_N – сумма N случайных выборок, тогда функция распределения этих сумм будет

$$f_N(y) = \frac{1}{(N-1)!}\left[y^{N-1} - \left(\frac{N}{1}\right)(y-1)^{N-1} + \left(\frac{N}{2}\right)(y-2)^{N-2} + \ldots \right] \tag{1.25}$$

где 0 < y < N. Находим среднее значение и дисперсию для величины y

$$\overline{y} = \frac{N}{2}; \qquad \sigma_y^2 = \frac{N}{12}. \tag{1.26}$$

Тогда можно выразить точку y_p через среднюю и дисперсию распределения

$$y_p = \overline{y} + k_p \sigma_y = \frac{N}{2} + k_p \sqrt{\frac{N}{12}}, \tag{1.27}$$

где k_p – некоторый множитель, учитывающий число степеней свободы распределения.

Данное соотношение может служить основой оценок для вероятностной модели. При достаточном количестве исходной информации вероятностная модель может дать вполне надежный прогноз. Кроме того, эта модель отличается большой простотой и наглядностью. Оценки, получаемые с помощью этой модели, имеют вполне конкретный смысл. Недостатком модели является требование большого количества наблюдений и незнание начального распределения,

что может привести к неправильным оценкам. Тем не менее, при определении процедуры начального распределения или с помощью байесовского метода, корректируя его, можно рассматривать вероятностную модель как эффективный метод прогноза.

1. *В чем состоит основная особенность метода вероятностного моделирования?*
2. *В чем отличие методов экспоненциального сглаживания и вероятностного моделирования?*

1.2. Интуитивные (экспертные) методы прогнозирования

Прогнозные экспертные оценки отражают индивидуальное суждение специалистов относительно перспектив развития объекта и основаны на мобилизации профессионального опыта и интуиции. Методы экспертных оценок используются для анализа объектов и проблем, развитие которых либо полностью, либо частично не поддается математической формализации, т. е. для которых трудно разработать адекватную модель. Применяемые в прогнозировании методы экспертной оценки разделяют на, индивидуальные и коллективные.

Индивидуальные экспертные методы основаны на использовании мнений экспертов-специалистов соответствующего профиля независимо друг от друга. Наиболее часто применимыми являются следующие два метода формирования прогноза: интервью и аналитические экспертные оценки.

Метод интервью предполагает беседу прогнозиста с экспертом, в ходе которой прогнозист в соответствии с заранее разработанной программой ставит перед экспертом вопросы относительно перспектив развития прогнозируемого объекта. Успех такой оценки в значительной степени зависит от способности интервьюируемого эксперта экспромтом давать заключения по самым различным фундаментальным вопросам. Аналитические экспертные оценки предполагают длительную и тщательную самостоятельную работу эксперта над анализом тенденций, оценкой состояния и путей развития прогнозируемого объекта. Этот метод дает возможность эксперту использовать всю необходимую ему информацию об объекте прогноза. Свои соображения эксперт оформляет в виде докладной записки.

Основными преимуществами рассматриваемых методов являются возможность максимального использования индивидуальных способностей эксперта и незначительность психологического давления, оказываемого на отдельного работника. Однако эти методы мало пригодны для прогнозирования наиболее общих стратегий из-за ограниченности знаний одного специалиста-эксперта о развитии смежных областей науки.

Методы коллективных экспертных оценок основываются на принципах выявления коллективного мнения экспертов о перспективах развития объекта прогнозирования. В основе применения этих методов лежит гипотеза о наличии у экспертов умения с достаточной степенью достоверности оценить важность и значение исследуемой проблемы, перспективность развития определенного направления исследований, времени свершения того или иного события, целесообразности выбора одного из альтернативных путей развития объекта прогноза и т. д. В настоящее время широкое распространение получили экспертные методы, основанные на работе специальных комиссий, когда группы экспертов за круглым столом обсуждают ту или иную проблему с целью согласования мнений и выработки единого мнения. Этот метод имеет недостаток, заключающийся в том, что группа экспертов в своих суждениях руководствуется в основном логикой компромисса.

В свою очередь в **методе Дельфи** вместо коллективного обсуждения той или иной проблемы проводится индивидуальный опрос экспертов обычно в форме анкет для выяснения относительной важности и сроков свершения гипотетических событий. Затем производится статистическая обработка анкет и формируется коллективное мнение группы, выявляются, обобщаются аргументы в пользу различных суждений. Вся информация сообщается экспертам. Участников экспертизы просят пересмотреть оценки и объяснить причины своего несогласия с коллективным суждением. Эта процедура повторяется 3–4 раза. В результате происходит сужение диапазона оценок. Недостатком этого метода является невозможность учета влияния, оказываемого на экспертов организаторами опросов при составлении анкет.

Как правило, основными задачами при формировании прогноза с помощью коллектива экспертов являются: формирование репрезентативной экспертной группы, подготовка и проведение экспертизы, статистическая обработка полученных документов.

При формировании группы экспертов основными являются вопросы определения ее качественного и количественного состава. Отбор экспертов начинается с определения вопросов, которые охватывают решение данной проблемы; затем составляется список лиц, компетентных в этих областях.

Для получения качественного прогноза к участникам экспертизы предъявляется ряд требований, основными из которых являются: высокий уровень общей эрудиции; глубокие специальные знания в оцениваемой области; способность к адекватному отображению тенденции развития исследуемого объекта; наличие психологической установки на будущее; наличие академического научного интереса к оцениваемому вопросу при отсутствии практической заинтересованности специалиста в этой области; наличие производственного и (или) исследовательского опыта в рассматриваемой области.

Для определения соответствия потенциального эксперта перечисленным требованиям используется анкетный опрос. Дополнительно к этому часто используют способ самооценки компетентности эксперта. При самооценке эксперт определяет степень своей осведомленности в исследуемом вопросе также на основании анкеты. Обработка данных дает возможность получить количественную оценку компетентности потенциального эксперта по следующей формуле

$$K = 0{,}5\left(\frac{\sum_{j=3}^{m} V_j}{\sum_{j=1}^{m} V_{j_{\max}}} + \frac{\lambda}{P}\right), \qquad (1.28)$$

где V_j – вес градации, перечеркнутой экспертом по j-й характеристике в анкете в баллах; $V_{j_{\max}}$ – максимальный вес (предел шкалы) j-й характеристики в баллах; т – общее количество характеристик компетентности в анкете; λ – вес ячейки, перечеркнутой экспертом в шкале самооценки в баллах; р – предел шкалы самооценки эксперта в баллах.

Установить оптимальную численность группы экспертов довольно трудно. Однако в настоящее время разработан ряд формализованных подходов к этому вопросу. Один из них основан на установлении максимальной и минимальной границ численности группы. При этом исходят из двух условий: высокой средней компетент-

ности групп экспертов и стабилизации средней оценки прогнозируемой характеристики.

Первое условие используется для определения максимальной численности группы экспертов n_{max}: $CK_{max} \le \frac{\sum_{i=1}^{n} k_i}{n_{max}}$, где С - константа; K_{max} – максимально возможная компетентность по используемой шкале компетентности; K_i – компетентность i-го эксперта. Это условие предполагает, что если имеется группа экспертов, компетентность которых максимальна, то среднее значение их оценок можно считать «истинным». Для определения константы используется практика голосования, т. е. группа считается избранной, если за нее подано 2/3 голосов присутствующих. Исходя из этого, принимается, что С=2/3. Таким образом, максимальная численность экспертной группы устанавливается на основании неравенства

$$n_{max} \le \frac{3\sum_{i=1}^{n} K_i}{2k_{max}}. \qquad (1.29)$$

Далее определяется минимальная численность экспертной группы n_{min}. Это осуществляется посредством использования условия стабилизации средней оценки прогнозируемой характеристики, которое формулируется следующим образом: включение или исключение эксперта из группы незначительно влияет на среднюю оценку прогнозируемой величины

$$\frac{B - B'}{B_{max}} < \varepsilon, \qquad (1.30)$$

где В – средняя оценка прогнозируемой величины в баллах, данная экспертной группой; В' – средняя оценка, данная экспертной группой, из которой исключен (или в которую включен) один эксперт; B_{max} – максимально возможная оценка прогнозируемой величины в принятой шкале оценок; ε – заданная величина изменения средней ошибки при включении или исключении эксперта.

Величина средней оценки наиболее чувствительна к оценке эксперта, обладающего наибольшей компетентностью и поставившего наибольший балл при $B \le B_{max}$ и минимальный – при $B \ge B_{max}/2$.

Поэтому для проверки выполнения условия (1.30) предлагается исключить из группы одного эксперта.

В литературе приводится правило расчета минимального числа экспертов в группе в зависимости от заданной (допустимой) величины изменения средней оценки ε

$$n_{\min} = 0{,}5\left(\frac{3}{\varepsilon} + 5\right). \qquad (1.31)$$

Таким образом, правила (1.29)-(1.31) дают возможность получить оценочные значения максимального и минимального числа экспертов в группе. Окончательная численность экспертной группы формируется на основании последовательного исключения малокомпетентных экспертов, при этом используется условие $\left(K_{\max} - K_i\right) \le \eta$, где η – задаваемая величина границы допустимого отклонения компетентности 1-го эксперта от максимальной. Одновременно могут включаться в группу новые эксперты. Численность группы устанавливается в пределах $n_{\min} \le n \le n_{\max}$.

Кроме описанных выше процедур в методах коллективных экспертных оценок используется подробный статистический анализ экспертных заключений, в результате которого определяются качественные характеристики группы экспертов. В соответствии с этими характеристиками в процессе проведения экспертизы качественный и количественный состав экспертной группы может корректироваться.

Подготовка к проведению экспертного опроса включает разработку анкет, содержащих набор вопросов по объекту прогноза. Структурно-организационный набор вопросов в анкете должен быть логически связан с центральной задачей экспертизы. Хотя форма и содержание вопросов определяются спецификой объекта прогнозирования, можно установить общие требования к ним: вопросы должны быть сформулированы в общепринятых терминах, их формулировка должна исключать всякую смысловую неопределенность, все вопросы должны логически соответствовать структуре объекта прогноза, обеспечивать единственное толкование.

По форме вопросы могут быть открытыми и закрытыми, прямыми и косвенными. Вопрос называют открытым, если ответ на него не регламентирован. Закрытыми считаются вопросы, в формулировке которых содержатся альтернативные варианты ответов, и эксперт

должен остановить свой выбор на одном (или нескольких) из них. Косвенные вопросы используют в тех случаях, когда требуется замаскировать цель экспертизы. К подобным вопросам прибегают тогда, когда не уверены, что эксперт, давая информацию, будет вполне искренен или свободен от посторонних влияний, искажающих объективность ответа. Рассмотрим основные группы вопросов, используемых при проведении коллективной экспертной оценки:

- вопросы, предполагающие ответы в виде количественной оценки: о времени свершения событий, о вероятности свершения событий, об оценке относительного влияния факторов. При определении шкалы значений количественных характеристик целесообразно пользоваться неравномерной шкалой. Выбор конкретного масштаба неравномерности определяется характером зависимости ошибки прогноза от периода упреждения;
- вопросы, требующие содержательного ответа в свернутой форме: дизъюнктивные, конъюнктивные, импликативные;
- вопросы, требующие содержательного ответа в развернутой форме: в виде перечня сведений об объекте; в виде перечня аргументов, подтверждающих или отвергающих тезис, содержащийся в вопросе.

Эти вопросы формируются в два этапа. На первом этапе экспертам предлагается сформулировать наиболее перспективные и наименее разработанные проблемы. На втором – из названных проблем выбираются принципиально разрешимые и имеющие непосредственное отношение к объекту прогноза.

Процедура проведения экспертизы может быть различной, однако здесь также можно выделить три основных этапа. На первом этапе эксперты привлекаются для уточнения формализованной модели объекта прогноза, формулировки вопросов в анкетах, уточнения состава группы. На втором этапе осуществляется непосредственная работа экспертов над вопросами в анкетах. На третьем этапе после предварительной обработки результатов прогноза эксперты привлекаются для консультаций по недостающей информации, необходимой для окончательного формирования прогноза.

При статистической обработке результатов экспертных оценок в виде количественных данных, содержащихся в анкетах, определяются статистические оценки прогнозируемых характеристик и их доверительные границы, статистические оценки согласованности мнений экспертов.

Среднее значение прогнозируемой величины определяется по формуле

$$B = \sum_{\varepsilon=1}^{n} B_i / n \,, \qquad (1.32)$$

где B_i - значение прогнозируемой величины, данное i-м экспертом; *n* – число экспертов в группе.

Кроме того, определяется дисперсия

$$D = \left[\sum_{i=1}^{n} \left(B_i - B\right)^2 / \left(n-1\right) \right]$$

и приближенное значение доверительного интервала

$$j = t\sqrt{\frac{D}{n-1}} \,,$$

где t – критерий Стьюдента для заданного уровня доверительной вероятности и числа степеней свободы k = (n – 2).

Доверительные границы для значения прогнозируемой величины вычисляются по формулам: для верхней границы A_B = B +j, для нижней границы A_H=B-j.

Коэффициент вариации оценок, данных экспертами, определяется по зависимости $v = \frac{\sigma}{B}$, где σ - среднеквадратическое отклонение.

При обработке результатов экспертных оценок по относительной важности направлений среднее значение, дисперсия и коэффициент вариации вычисляются для каждого оцениваемого направления. Кроме того, вычисляется коэффициент конкордации, показывающий степень согласованности мнений экспертов по важности каждого из оцениваемых направлений, и коэффициенты парной ранговой корреляции, определяющие степень согласованности экспертов друг с другом.

Для этого производится ранжирование оценок важности, данных экспертами. Каждая оценка, данная i-м экспертом, выражается числом натурального ряда таким образом, что число 1 присваивается максимальной оценке, а число *n* - минимальной. Если все оценки различны, то соответствующие числа натурального ряда есть ранги

оценок i-го эксперта. Если среди оценок, данных; i-м экспертом, есть одинаковые, то этим оценкам назначается одинаковый ранг, равный средней арифметической соответствующих чисел натурального ряда.

Сумма рангов S_j, назначенных экспертами направлению j (j = 1,..., т; х - число исследуемых направлений), определяется по формуле

$$S_j = \sum_{i=1}^{n} R_{ij}, \tag{1.33}$$

где R_{ij} – ранг оценки, данной i-м экспертом j-му направлению. Среднее значение суммы рангов оценок по всем направлениям равно $\overline{S} = \sum_{j=1}^{m} S_j / m$. Отклонение суммы рангов, полученных j-м направлением, от среднего значения суммы рангов определяется как $d_j = S_j - \overline{S}$. Тогда коэффициент конкордации, вычисленный по совокупности всех направлений, составляет

$$W = \frac{12\sum_{j=1}^{m} d_j^2}{n^2\left(m^3 - m\right) - n\sum_{i=1}^{n} T_i}. \tag{1.34}$$

Величина $T_i = \sum_{i=1}^{n} t_i^3 - t_e$, рассчитывается при наличии равных рангов (n - количество групп равных рангов, t_e – количество равных рангов в группе).

Коэффициент конкордации принимает значение в пределах от 0 до 1. W=l означает полную согласованность мнений экспертов, при W=O - полную несогласованность. Коэффициент конкордации показывает степень согласованности всей экспертной группы. Низкое значение этого коэффициента может быть получено как при отсутствии общности мнений всех экспертов, так и из-за наличия противоположных мнений между подгруппами экспертов, хотя внутри подгруппы согласованность может быть высокой.

Для выявления степени согласованности мнений экспертов используется коэффициент парной ранговой корреляции

$$\rho_{i,i+1} = \frac{\sum_{j=1}^{m} \psi_j^2}{\frac{1}{\sigma}\left(m^3 - m\right) - \frac{1}{12}\left(\pi - T_j - 1\right)}, \qquad (1.35)$$

где ψ_j – разность (по модулю) величин рангов оценок j-го направления, назначенных экспертами i и i + 1,

$$\psi_j = \left|R_i - R_{i+1}\right|. \qquad (1.36)$$

Коэффициент парной ранговой корреляции может принимать значения от +1 до -1. Значение $\rho = 1$ соответствует полной согласованности мнений двух экспертов. Значение $\rho = -1$ показывает, что мнение одного эксперта противоположно мнению другого.

Для определения уровня значимости значений коэффициентов W и $\rho_{i,i+1}$ можно использовать критерий χ^2. Для этого вычисляется величина

$$\chi^2 = \frac{12\sum_{j=1}^{m} \alpha_j^2}{m \cdot n(m+1) - \frac{1}{m-1}\sum_{i=1}^{n} T_i} \qquad (1.37)$$

(число степеней свободы k = т – 1) и по соответствующим таблицам определяется уровень значимости полученных значений.

1. *В чем состоит особенность метода «Интервью»?*
2. *В чем заключаются преимущества применения методов экспертной оценки?*
3. *Объясните суть метода «Дельфы».*
4. *Каковы основные критерии формирования группы экспертов?*
5. *Как определяется численность группы экспертов и на основе каких показателей?*
6. *На основе каких показателей формируется статистическая оценка мнений экспертов?*

1.3. Корреляционный и регрессионный анализы

Одним из наиболее распространенных способов получения многофакторных прогнозов является упоминавшийся ранее классический метод наименьших квадратов и построение на его основе модели множественной регрессии [16]. Для линейного случая модель множественной регрессии записывается в виде

$$y_j = \sum_{i=1}^{n} \alpha_i x_{ij} + \varepsilon_j \tag{1.38}$$

где α_i - коэффициенты модели; y_j, x_{ij} - соответственно значения j-й функции (зависимой переменной) и i-й независимой переменной; $i = \overline{0,n}$; j =1, $\overline{N}$, ε_j – случайная ошибка; n - число независимых переменных в модели (в ряде случаев полагается, что α_i, – свободный член и $x_{0j} = 1$).

В векторном виде эта модель записывается так [131]

$$Y = X\alpha + \varepsilon, \tag{1.39}$$

где вектор $Y^T = (y_1, y_2, ..., y_n)$ и вектор $\alpha^T = (\alpha_1, \alpha_2, ..., \alpha_n)$ - соответственно векторы значений зависимой переменной и коэффициентов модели, матрица порядка(n×N);

$X = \begin{bmatrix} x_{11}, x_{12}, ...x_{1n} \\ ..... \\ x_{N1}, x_{N2}, ...x_{Nn} \end{bmatrix}$ – матрица независимых переменных;

$\varepsilon^T = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)$ – вектор случайных ошибок.

Неизвестные коэффициенты модели находятся из условия минимума функционала рассогласований, который представляет собой сумму квадратов рассогласований реальных значений зависимой переменной и значений. В векторном виде функционал рассогласований записывается как.

$$\Phi = (Y - X\alpha)^T (Y - X\alpha) \to \min.$$

Условия минимума Φ реализуются при равенстве нулю первых производных функционала по неизвестным коэффициентам, т. е.

$$\frac{\partial \Phi}{\partial \alpha_1} = 0; i = \overline{0,n} \tag{1.41}$$

Данное условие эквивалентно выполнению векторного соотношения $X^T X\alpha = X^T Y$, что дает значения оценок коэффициентов модели

$$\hat{\alpha} = \left(X^T X\right)^{-1} X^T Y. \tag{1.42}$$

Надежность получаемой с помощью оценок α модели определяется с помощью величины остаточной дисперсии, которая вычисляется по формуле

$$\sigma^2 = \frac{\varepsilon^T \varepsilon}{N-n} = \frac{1}{N-n}\left[YY^T - Y^T X\left(X^T X\right)^{-1} X^T Y\right], \tag{1.43}$$

и коэффициента множественной корреляции, вычисляемого по формуле

$$R = \sqrt{1 - \frac{D}{D_{(n+1)(n+1)}}} \tag{1.44}$$

$D_{(n+1)(n+1)}$- алгебраическое дополнение определителя корреляционной матрицы $r = r\left[x_i x_j\right], \left(i, j = \overline{0,n}\right)$ к элементу $r_{x_{n=1} x_{n+1}}$ Величина R^2 – множественный коэффициент детерминации; она показывает, какая доля дисперсии функции объясняется изменениями входящих в уравнение регрессии независимых переменных при полученных значениях коэффициентов модели. Надежность коэффициента множественной корреляции определяется по критерию Фишера

$$F = \frac{R^2 (N-n-1)}{\left(1-R^2\right)(n-1)} \tag{1.45}$$

при заданном уровне надежности и степени свободы ν_1 = n, v_1 = N – *n*. Наличие связи между зависимой переменной y_i и независимыми переменными x_{ij} определяется с помощью коэффициентов корреляции

$$r_{yxi} = \frac{S^2_{yxi}}{\sqrt{S^2_{yy} S^2_{xixi}}} \qquad (1.46)$$

где $S^2_{yxi}; S^2_{yy}; S^2_{xixk}$.,, - коэффициенты ковариации, определяемые по формуле

$$S^2_{xixk} = \frac{1}{N-1}\sum_{j=1}^{N}\left(x_{ij} - \bar{x}_i\right)\left(x_{kj} - \bar{x}_k\right), \qquad (1.47)$$

где $\bar{x}_i, \bar{x}_k$ - средние значения независимых переменных x_i и x_k . Доверительные интервалы Z_1, Z_2 для коэффициента корреляции определяются как

$$Z_{1,2} = arcth \pm \frac{U\alpha/2}{\sqrt{N-3}} - \frac{r}{2(N-1)}, \qquad (1.48)$$

где $U\frac{\alpha}{2} - \frac{\alpha}{2} - r$ – процентиль нормального распределения N (0, 1) с нулевым средним и единичной дисперсией, z - преобразование Фишера, определяемое по формуле [16]

$$z_{xixk} = \frac{1}{2}\ln\frac{1 + r_{xixk}}{1 - r_{xixk}}, \qquad (1.49)$$

th z - гиперболический тангенс аргумента z, вычисленный по формуле

$$thz = \frac{e^z - e^{-z}}{e^z + e^{-z}}. \qquad (1.50)$$

Истинное значение коэффициента корреляции заключено в пределах $\text{th}\, z_1 \le r_{xixk} \le thz_2$. Для определения надежности оценок строится доверительный интервал для полученных оценок $\hat{\alpha}$ коэффициентов модели

$$\left|\hat{\alpha} - \alpha_i\right| \le t_{p,v}\sigma\sqrt{c_{ii}}, \qquad (1.51)$$

где $t_{p,v}$ – значения критерия Стъюдента с уровнем надежности p и степенями свободы $\nu = (N - n - 1); \sigma = \sqrt{\sigma^2}$ (σ^2 определяется по формуле (6)), c_{ii}– i-й диагональный элемент матрицы $\left(X^T X\right)^{-1}$. Поэтому

истинное значение коэффициента а, модели будет лежать в интервале

$$\hat{\alpha}_i - t_{p,\nu}\sigma\sqrt{c_{ii}} < \alpha_1 < \hat{\alpha}_i + t_{p,\nu}\sigma\sqrt{c_{ii}} \qquad (1.52)$$

Использование вычислительной процедуры по методу наименьших квадратов с целью получения оценок коэффициентов модели $\hat{\alpha}_i\left(i = \overline{0,n}\right)$, которые удовлетворяли бы условиям несмещенности, состоятельности, эффективности, предполагает выполнение ряда условий. Рассмотрим эти условия:

– независимые переменные представляют собой неслучайный набор чисел, их средние значения и дисперсии конечны;

– случайные ошибки ε_j - имеют нулевую среднюю и конечную дисперсию

$$M\left(\varepsilon_j\right) = 0; M\left(\varepsilon_j\varepsilon_j\right) = \sigma_\varepsilon^2 < \infty \qquad (1.53)$$

– между независимыми переменными отсутствует корреляция и автокорреляция;

– случайная ошибка не коррелированна с независимыми переменными;

– случайная ошибка подчинена нормальному закону распределения.

Кроме того, можно выделить условие отсутствия мультиколлинеарности, когда несколько независимых переменных связаны между собой линейной зависимостью, и условие гомоскедастичности, т. е. одинаковой дисперсии для всех случайных ошибок. Важным является условие линейной формы связи между зависимой и независимой переменными. Зависимость должна быть именно линейной или сводимой к линейной с помощью некоторых преобразований.

Но иногда исследуемый процесс не может быть сведен к линейной зависимости никакими преобразованиями, как, например, в случае логистической зависимости. Тогда используется ряд методов, например, метод симплексов. Данный метод отличается сравнительной простотой, легкой реализуемостью на ЭВМ, эффективностью при определении оценок коэффициентов модели.

Важной характеристикой реализованной модели является оценка ошибки прогноза. Так, в [55] предлагается следующая оценка дисперсии прогноза

$$\sigma_{t+\tau}^2 = \sigma^2 \left[1 + X_{t+\tau}^T (X^T X)^{-1} X_{t+\tau} \right], \qquad (1.54)$$

где $X_{t+\tau}$ - вектор значений независимых переменных в момент $(t+\tau)$. Поэтому доверительный интервал для значений зависимой переменной определяется в момент t как

$$\left\{ Y_t \pm t_{p,v} \sigma_t \sqrt{I + \frac{I}{t-1} X_t^T \left(X_t^T X_t \right)^{-1} X_t} \right\} \qquad (1.55)$$

где I - единичный столбец; $t_{p,v}$ - значение критерия Стьюдента. В [14] находится более эффективная оценка доверительного интервала для прогнозных значений

$$\left[Y_{t+\tau} \pm t_{p,v} \sigma_{t+\tau} \sqrt{I + \frac{1-r_1}{t+\tau-1} \frac{X_{t+r}^T X_{t+r}}{X_t^T X_t}} \right]. \qquad (1.56)$$

Важным условием получения надежных оценок для модели по методу наименьших квадратов является отсутствие автокорреляции. Оценка автокорреляции для полученной по МНК модели осуществляется по критерию Дарбина – Уотсона [73, 16]

$$d = \frac{\sum_{t=1}^{T-1} (\varepsilon_{t+1} - \varepsilon_t)}{\sum_{t=1}^{T} \varepsilon_t^2}, \qquad (1.57)$$

где T - длина временного ряда.

Полученное расчетное значение d сравнивается с нижней и верхней границей d_1 и d_2, критерия. Если $d < d_1$, то гипотеза отсутствия автокорреляции отвергается; если $d > d_2$, то гипотеза отсутствия автокорреляции принимается; если $d_1 \leq d \leq d_2$, то необходимо дальнейшее исследование. Одним из известных способов уменьшения автокорреляции является авторегрессионное преобразование для исходной информации или переход к разностям, т. е. $\Delta Y_t = Y_{t+1} - Y_t$; $\Delta X_t = X_{t+1} - X_t$. Если же автокорреляцию устранить не удается, то полученные оценки считаются состоятельными, и среднеквадратическое отклонение корректируется на величину Δj для j-го коэффициента.

$$\Delta_j = \sqrt{\frac{1 + r_1 R_{1j}}{1 - r_1 R_{1j}}}, \qquad (1.58)$$

где r_1 - коэффициент автокорреляции случайных слагаемых первого порядка; R_{1j} - коэффициент автокорреляции для j-й независимой переменной первого порядка.

Другим условием, необходимым для получения состоятельных оценок, является отсутствие мультиколлинеарности. Действительно, при наличии мультиколлинеарности определитель квадратной матрицы [$X^T X$] равен или близок нулю, следовательно, матрица вырождена, и поэтому решения системы нормальных уравнений не существует.

Эффективный подход к определению мультиколлинеарности предполагает следующую последовательность расчетов. Пусть рассматривается уравнение регрессии $y=f(x_1, \ldots, x_n)$. Тогда для выявления существования мультиколлинеарности предлагается критерий

$$\chi^2 = -\left[N - 1 - \frac{1}{6}(2n+5)\right] \lg\left[\left|\widetilde{X}^T \widetilde{X}\right|\right] \qquad (1.59)$$

где $\left|\widetilde{X}^T X\right|$ - определитель матрицы $\left[\widetilde{X}^T X\right]$), имеющий асимптотическое распределение Пирсона с $\frac{1}{2}n(n-1)$ степенями свободы. В формуле (16) N – число наблюдений по каждому переменному, n – число независимых переменных, матрица X включает значения переменных, преобразованных по формуле

$$\widetilde{X}_{ik} = \frac{x_{ik} - \overline{x}_i}{S_i \sqrt{N}}, \qquad (1.60)$$

где S_i, $\overline{x}_i$ – соответственно оценки среднеквадратического отклонения и среднее значение для i-й независимой переменной. Далее вычисляются величины

$$d_{ii} = \frac{\left(\widetilde{X}^T \widetilde{X}\right)_{ii}}{\left|\widetilde{X}^T \widetilde{X}\right|}, \qquad (1.61)$$

которые при неколлинеарности переменных близки единице, а при наличии мультиколлинеарности близки к бесконечности, что дает основание оставить или отбросить показатель x_i, что определяется

статистикой $w_i = (d_{ii} - 1)\frac{N-n}{n-1}$, имеет распределение Фишера с $v_1 = N - n$ и $v_2 = n - 1$ степенями свободы. Существует еще ряд способов определения мультиколлинеарности, В целях устранения или уменьшения ее можно переходить к разностям для исходной информации, использовать метод факторного анализа или метод главных компонент.

Получение прогнозов с помощью многофакторных регрессионных моделей предполагает неизменность значений коэффициентов этих моделей во времени. Тем не менее, в процессе исследования объекта возможно появление новой информации, что позволяет с помощью рекуррентного оценивания корректировать значения оценок коэффициентов моделей. В то же время исходная информация может содержать в себе различные динамики изменения независимых переменных, которые возникают в результате различных «режимов» функционирования исследуемого объекта. В этом случае важным является, как сам факт установления различия динамик процессов на разных временных интервалах, так и выбор такого интервала для построения на нем модели прогнозирования, который был бы наиболее адекватным будущему поведению объекта. Если оказывается, что для одного интервала времени построена многофакторная модель $y_1 = \sum_{i=1}^{m} \alpha_{i1} x_i$, а для другого интервала - модель $y_2 = \sum_{i=1}^{m} \alpha_{i2} x_i$, где $\alpha_{i1} \neq \alpha_{i2}$, то прогноз будет смещен, а, следовательно, резко возрастает дисперсия прогноза.

Построение адекватных регрессионных моделей для целей прогнозирования с помощью метода наименьших квадратов предъявляет к исходной информации весьма жесткие требования. В ряде случаев эти требования для реальных наблюдений оказываются невыполненными, поэтому получаемые оценки оказываются неэффективными, а прогноз – недостоверным. Действительно, требование нормальности распределения ошибок, предъявляемое к исходной информации процедурой метода наименьших квадратов, в большом числе случаев оказывается невыполненным. Так, говорится: «Нормальность – это миф. В реальном мире никогда не было и никогда не будет нормального распределения». Поэтому в последнее время интенсивно разрабатывается новое направление в статистике – так называемая робастная статистика, задача которой в том и состоит,

чтобы получать эффективные оценки в случаях невыполнения некоторых предпосылок, например, нормальности распределения, наличия аномальных наблюдений. Использование робастных методов получения статистических оценок для информации, содержащей аномальные «выбранные» наблюдения, позволяет значительно повысить надежность получаемых оценок по сравнению с обычным методом наименьших квадратов.

1. *В чем состоит суть корреляционного анализа?*
2. *Какую роль в корреляционном анализе играет оценка показателей F-статистики Фишера и t-статистики Стьюдента?*
3. *В чем состоит проблема мультиколлинеарности?*
4. *Чем затруднен процесс построения адекватных прогнозов на основе регрессионных моделей?*

1.4. Модели стационарных временных рядов и их идентификация

Здесь рассматривается набор линейных параметрических моделей. **Речь здесь идет не о моделировании временных рядов, а о моделировании их случайных остатков ε_t, получающихся после элиминирования (вычитания) из исходного временного ряда x_t его неслучайной составляющей (тренда)**. Следовательно, в отличие от прогноза, основанного на регрессионной модели, игнорирующего значения случайных остатков, в прогнозе временных рядов существенно используется взаимозависимость и прогноз самих случайных остатков.

Введем обозначения. Так как здесь описывается поведение случайных остатков, то моделируемый временной ряд обозначим ε_t, и будем полагать, что при всех t его математическое ожидание равно нулю, т.е. $E\varepsilon_t \equiv 0$. Временные последовательности, образующие «белый шум», обозначим δ_t.

Описание и анализ рассматриваемых ниже моделей формулируется в терминах общего линейного процесса, представимого в виде взвешенной суммы настоящего и прошлых значений белого шума, а именно

$$\varepsilon_t = \sum_{k=0}^{\infty} \beta_k \delta_{t-k}, \tag{1.62}$$

где $\beta_0 = 1$ и $\sum_{k=0}^{\infty} \beta_k^2 < \infty$.

Таким образом, белый шум представляет собой серию импульсов, в широком классе реальных ситуаций генерирующих случайные остатки исследуемого временного ряда.

Временной ряд ε_t можно представить в эквивалентном (1.62) виде, при котором он получается в виде классической линейной модели множественной регрессии, в которой в качестве объясняющих переменных выступают его собственные значения во все прошлые моменты времени

$$\varepsilon_t = \sum_{k=1}^{\infty} \pi_k \varepsilon_{t-k} + \delta_t. \tag{1.63}$$

При этом весовые коэффициенты $\pi_1, \pi_2, \ldots$ связаны определенными условиями, обеспечивающими стационарность ряда ε_t. Переход от (1.63) к (1.62) осуществляется с помощью последовательной подстановки в правую часть (1.63) вместо $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots$ их выражений, вычисленных в соответствии с (1.63) для моментов времени $t-1$, $t-2$ и т.д.

Рассмотрим также процесс смешанного типа, в котором присутствуют как авторегрессионные члены самого процесса, так и скользящее суммирование элементов белого шума

$$\varepsilon_t = \sum_{k=1}^{p} \pi_k \varepsilon_{t-k} + \delta_t + \sum_{j=1}^{q} \beta_j \delta_{t-j}. \tag{1.64}$$

Будем подразумевать, что p и q могут принимать и бесконечные значения, а также то, что в частных случаях некоторые (или даже все) коэффициенты π или β равны нулю.

1.4.1. Модели авторегрессии порядка p (AR(p)-модели)

Рассмотрим сначала простейшие частные случаи.

Модель авторегрессии 1-го порядка – AR(1) (марковский процесс). Эта модель представляет собой простейший вариант авторегрессионного процесса типа (1.63), когда все коэффициенты кроме

первого равны нулю. Соответственно, она может быть определена выражением

$$\varepsilon_t = \alpha\varepsilon_{t-1} + \delta_t, \qquad (1.65)$$

где α – некоторый числовой коэффициент, не превосходящий по абсолютной величине единицу ($|\alpha| < 1$), а δ_t – последовательность случайных величин, образующая белый шум. При этом ε_t зависит от δ_t и всех предшествующих δ, но не зависит от будущих значений δ. Соответственно, в уравнении (1.65) δ_t не зависит от ε_{t-1} и более ранних значений ε. В связи с этим, δ_t называют инновацией (обновлением).

Последовательности ε, удовлетворяющие соотношению (1.65), часто называют также марковскими процессами.

Модели авторегрессии 2-го порядка – AR(2) (процессы Юла). Эта модель, как и AR(1), представляет собой частный случай авторегрессионного процесса, когда все коэффициенты π_j в правой части (1.63) кроме первых двух, равны нулю. Соответственно, она может быть определена выражением

$$\varepsilon_t = \alpha_1\varepsilon_{t-1} + \alpha_2\varepsilon_{t-2} + \delta_t, \qquad (1.66)$$

где последовательность $\delta_1, \delta_2, \ldots$ образует белый шум.

Условия стационарности ряда (1.66) (необходимые и достаточные) определяются как

$$\begin{cases} |\alpha_1| < 2, \\ \alpha_2 < 1 - |\alpha_1|. \end{cases} \qquad (1.67)$$

В рамках общей теории моделей те же самые условия стационарности получаются из требования, чтобы все корни соответствующего характеристического уравнения лежали бы вне единичного круга. Характеристическое уравнение для модели авторегрессии 2-го порядка имеет вид: $1 - \alpha_1 z - \alpha_2 z^2 = 0$.

Автокорреляционная функция процесса Юла подсчитывается следующим образом. Два первых значения r(1) и r(2) определены соотношениями

$$r(1) = \frac{\alpha_1}{1-\alpha_2},$$
$$r(2) = \alpha_2 + \frac{\alpha_1^2}{1-\alpha_2}, \qquad (1.68)$$

а значения для $r(\tau)$, $\tau = 3, 4,\ldots$ вычисляются с помощью рекуррентного соотношения

$$r(\tau) = \alpha_1 r(\tau - 1) + \alpha_2 r(\tau - 2). \qquad (1.69)$$

Модели авторегрессии p-го порядка – AR(p) ($p \geq 3$). Эти модели, образуя подмножество в классе общих линейных моделей, сами составляют достаточно широкий класс моделей. Если в общей линейной модели (1.63) полагать все параметры π_j, кроме первых p коэффициентов, равными нулю, то мы приходим к определению AR(p)-модели

$$\varepsilon_t = \sum_{j=1}^{p} \alpha_j \varepsilon_{t-j} + \delta_t, \qquad (1.70)$$

где последовательность случайных величин $\delta_1, \delta_2,\ldots$ образует белый шум.

Условия стационарности процесса, генерируемого моделью (1.70), также формулируются в терминах корней его характеристического уравнения

$$1 - \alpha_1 z - \alpha_2 z^2 - \ldots - \alpha_p z^p = 0. \qquad (1.71)$$

Для стационарности процесса необходимо и достаточно, чтобы все корни характеристического уравнения лежали бы вне единичного круга, т.е. превосходили бы по модулю единицу.

1.4.2. Модели скользящего среднего порядка q (MA(q)-модели)

Рассмотрим частный случай общего линейного процесса (1.62), когда только первые q из весовых коэффициентов β_j ненулевые. В это случае процесс имеет вид

$$\varepsilon_t = \delta_t - \theta_1\delta_{t-1} - \theta_2\delta_{t-2} - \ldots - \theta_q\delta_{t-q}, \qquad (1.72)$$

где символы $-\theta_1,\ldots, \theta_q$ используются для обозначения конечного набора параметров β, участвующих в (1.62). Процесс (1.72) называется моделью скользящего среднего порядка q (MA(q)).

Двойственность в представлении AR- и MA-моделей и понятие обратимости MA-модели. Из (1.62) и (1.63) видно, что один и тот же общий линейный процесс может быть представлен либо в виде AR-модели бесконечного порядка, либо в виде MA-модели бесконечного порядка.

Соотношение (1.72) может быть переписано в виде

$$\delta_t = \varepsilon_t + \theta_1\delta_{t-1} + \theta_2\delta_{t-2} + \ldots + \theta_q\delta_{t-q}. \qquad (1.73)$$

Откуда

$$\delta_t = \varepsilon_t - \pi_1\varepsilon_{t-1} - \pi_2\varepsilon_{t-2} - \ldots, \qquad (1.74)$$

где коэффициенты π_j (j = 1, 2,…) определенным образом выражаются через параметры $\theta_1,\ldots, \theta_q$. Соотношение (1.74) может быть записано в виде модели авторегрессии бесконечного порядка (т.е. в виде обращенного разложения)

$$\varepsilon_t = \sum_{j=1}^{\infty} \pi_j \varepsilon_{t-j} + \delta_t. \qquad (1.75)$$

Известно, что условие обратимости MA(q)-модели (т.е. условие сходимости ряда $\sum_{j=1}^{\infty} \pi_j$) формулируется в терминах характеристического уравнения модели (1.75) следующим образом.

Все корни характеристического уравнения $1 - \theta_1 z - \theta_2 z^2 - \ldots - \theta_q z^q = 0$ должны лежать вне единичного круга, т.е. $|z_j| > 1$ для всех j = 1, 2,…, q.

Взаимосвязь процессов AR(q) и MA(q). Сделаем ряд замечаний о взаимосвязях между процессами авторегрессии и скользящего среднего.

Для конечного процесса авторегрессии порядка p δ_t может быть представлено как конечная взвешенная сумма предшествующих ε, или ε_t может быть представлено как бесконечная сумма предшест-

вующих δ. В то же время, в конечном процессе скользящего среднего порядка q ε_t может быть представлено как конечная взвешенная сумма предшествующих δ или δ_t – как бесконечная взвешенная сумма предшествующих ε.

Конечный процесс MA имеет автокорреляционную функцию, обращающуюся в нуль после некоторой точки, но так как он эквивалентен бесконечному процессу AR, его частная автокорреляционная функция бесконечно протяженная. Главную роль в ней играют затухающие экспоненты и (или) затухающие синусоиды. И наоборот, процесс AR имеет частную автокорреляционную функцию, обращающуюся в нуль после некоторой точки, но его автокорреляционная функция имеет бесконечную протяженность и состоит из совокупности затухающих экспонент и или затухающих синусоид.

Параметры процесса авторегрессии конечного порядка не должны удовлетворять каким-нибудь условиям для того, чтобы процесс был стационарным. Однако для того чтобы процесс MA был обратимым, корни его характеристического уравнения должны лежать вне единичного круга.

Спектр процесса скользящего среднего является обратным к спектру соответствующего процесса авторегрессии.

1.4.3. Авторегрессионные модели со скользящими средними в остатках (ARMA(p, q)-модели)

Представление процесса типа MA в виде процесса авторегрессии неэкономично с точки зрения его параметризации. Аналогично процесс AR не может быть экономично представлен с помощью модели скользящего среднего. Поэтому для получения экономичной параметризации иногда бывает целесообразно включить в модель как члены, описывающие авторегрессию, так и члены, моделирующие остаток в виде скользящего среднего. Такие линейные процессы имеют вид

$$\varepsilon_t = \alpha_1\varepsilon_{t-1} + \ldots + \alpha_p\varepsilon_{t-p} + \delta_t - \theta_1\delta_{t-1} - \ldots - \theta_q\delta_{t-q} \qquad (1.76)$$

и называются процессами авторегрессии – скользящего среднего порядка (p, q)(ARMA(p, q)).

Стационарность и обратимость ARMA(p, q)-процессов. Записывая процесс (1.76) в виде

$$\varepsilon_t = \sum_{j=1}^{p} \alpha_j \varepsilon_{t-j} + \overline{\delta}_{qt}, \qquad (1.77)$$

где $\overline{\delta}_{qt} = \delta_t - \theta_1 \delta_{t-1} - \ldots - \theta_q \delta_{t-q}$, можно провести анализ стационарности (8) по той же схеме, что и для AR(p)-процессов. При этом различие “остатков” $\overline{\delta}_{qt}$ и δ_e никак не повлияет на выводы, определяющие условия стационарности процесса авторегрессии. Поэтому процесс (1.74) является стационарным тогда и только тогда, когда все корни характеристического уравнения AR(p)-процесса лежат вне единичного круга.

1. *Что такое «белый шум»?*
2. *Дайте определение Марковским процессам.*
3. *Охарактеризуйте стационарный процесс.*

1.5. Модели нестационарных временных рядов и их идентификация

1.5.1. Модель авторегрессии-проинтегрированного скользящего среднего (ARIMA(p, k, q)-модель)

Эта модель предложена Дж. Боксом и Г. Дженкинсом [15]. Она предназначена для описания нестационарных временных рядов x_t, обладающих следующими свойствами:

- анализируемый временной ряд аддитивно включает в себя составляющую f(t), имеющую вид алгебраического полинома (от параметра времени t) некоторой степени k > 1; при этом коэффициенты этого полинома могут быть как стохастической, так и нестохастической природы;

- ряд $x_t^k, t = 1,\ldots,T-k,$ получившийся из x_t после применения к нему k-кратной процедуры метода последовательных разностей, может быть описан моделью ARMA(p, q).

Это означает, что ARIMA(p, k, p)-модель анализируемого процесса x_t, может быть записана в виде

$$x_t^k = \alpha_1 x_{t-1}^k + \alpha_2 x_{t-2}^k + \ldots + \alpha_p x_{t-p}^k + \delta - \theta_1 \delta_{t-1} - \ldots - \theta_q \delta_{t-q}, \qquad (1.78)$$

где

$$x_t^k = \Delta^k x_t = x_t - C_k^1 x_{t-1} + C_k^2 x_{t-2} - \ldots + (-1)^k x_{t-k}.$$

Заметим, что классу моделей ARIMA принадлежит и простейшая модель стохастического тренда – процесс случайного блуждания (или просто случайное блуждание). Случайное блуждание определяется аналогично процессу авторегрессии первого порядка (1.63), но только у случайного блуждания $\alpha = 1$, так что

$$\varepsilon_t = \varepsilon_{t-1} + \delta_t. \qquad (1.79)$$

Ряд первых разностей случайного блуждания δ_t представляет собой белый шум, т.е. процесс ARMA(0, 0). Поэтому само случайное блуждание входит в класс моделей ARIMA как модель ARIMA(0, 1, 0).

Идентификация ARIMA-моделей. В первую очередь, следует подобрать порядок k модели. Первый тип критерия подбора основан на отслеживании поведения величины $\hat{\sigma}^2(k)$ в зависимости от k: в качестве верхней оценки для порядка k определяется то значение k_0, начиная с которого тенденция к убыванию $\hat{\sigma}^2(k)$ гасится и само значение $\hat{\sigma}^2(k)$ относительно стабилизируется. Второй тип критерия подбора порядка k ARIMA-модели основан на анализе поведения автокорреляционных функций процессов Δx_t, $\Delta^2 x_t$,…. последовательные преобразования анализируемого процесса x_t с помощью операторов Δ, Δ^2,… нацелены на устранение его нестационарности. Поэтому до тех пор, пока $l < k$ процессы $\Delta^l x_t$ будут оставаться нестационарными, что будет выражаться в отсутствии быстрого спада в поведении их выборочной автокорреляционной функции. Поэтому предполагается, что необходимая для получения стационарности степень k разности Δ достигнута, если автокорреляционная функция ряда $x_t^k = \Delta^k x_t$ быстро затухает.

После подбора порядка k анализируется уже не сам ряд x_t, а его k-е разности. Идентификация этого ряда сводится к идентификации ARMA(p, q)моделей.

1.5.2. Модели рядов, содержащих сезонную компоненту

Под временными рядами, содержащими сезонную компоненту, понимаются процессы, при формировании значений которых обязательно присутствовали сезонные и/или циклические факторы.

Один из распространенных подходов к прогнозированию состоит в следующем: ряд раскладывается на долговременную, сезонную (в том числе, циклическую) и случайную составляющие; затем долговременную составляющую подгоняют полиномом, сезонную – рядом Фурье, после чего прогноз осуществляется экстраполяцией этих подогнанных значений в будущее. Однако этот подход может приводить к серьезным ошибкам. Во-первых, короткие участки стационарного ряда (а в экономических приложениях редко бывают достаточно длинные временные ряды) могут выглядеть похожими на фрагменты полиномиальных или гармонических функций, что приведет к их неправомерной аппроксимации и представлению в качестве неслучайной составляющей. Во-вторых, даже если ряд действительно включает неслучайные полиномиальные и гармонические компоненты, их формальная аппроксимация может потребовать слишком большого числа параметров, т.е. получающаяся параметризация модели оказывается неэкономичной.

Принципиально другой подход основан на модификации ARIMA-моделей с помощью «упрощающих операторов». Схематично процедура построения сезонных моделей, основанных на ARIMA-конструкциях, модифицированных с помощью упрощающих операторов $\nabla_T = 1 - F^T_{_}$, может быть описана следующим образом (детальное описание соответствующих процедур см., например, в [15]:

- применяем к наблюдаемому ряду x_t операторы Δ и ∇_T для достижения стационарности;

- по виду автокорреляционной функции преобразованного ряда $x_{k,K}^{(T)}(t)$ подбираем пробную модель в классе ARMA- или модифицированных (в правой части) ARMA-моделей;

- по значениям соответствующих автоковариаций ряда $x_{k,K}^{(T)}(t)$ получаем (методом моментов) оценки параметров пробной модели;

Диагностическая проверка полученной модели (анализ остатков в описании реального ряда x_t с помощью построенной модели) может либо подтвердить правильность модели, либо указать пути ее

улучшения, что приводит к новой подгонке и повторению всей процедуры.

1.5.3. Прогнозирование на базе ARIMA-моделей

ARIMA-модели охватывают достаточно широкий спектр временных рядов, а небольшие модификации этих моделей позволяют весьма точно описывать и временные ряды с сезонностью. Начнем обсуждение проблемы прогнозирования временных рядов с методов, основанных на использовании ARIMA-моделей. Мы говорим об ARIMA-моделях, имея в виду, что сюда входят как частные случаи AR-, MA- и ARMA-модели. Кроме того, будем исходить из того, что уже осуществлен подбор подходящей модели для анализируемого временного ряда, включая идентификацию этой модели. Поэтому в дальнейшем предполагается, что все параметры модели уже оценены.

Будем прогнозировать неизвестное значение x_{t+l}, $l \geq 1$ полагая, что x_t – последнее по времени наблюдение анализируемого временного ряда, имеющееся в нашем распоряжении. Обозначим такой прогноз $\hat{x}_t^l$.

Заметим, что хотя $\hat{x}_t^l$ и $\hat{x}_{t-1}^{l+1}$ обозначают прогноз одного и того же неизвестного значения x_{t+l}, но вычисляются они по-разному, т.к. являются решениями разных задач.

Ряд x_τ, анализируемый в рамках ARIMA(p, k, q)-модели, представим (при любом $\tau > k$) в виде

$$\left(1-\alpha_1 L-\ldots-\alpha_p L^p\right)\sum_{j=0}^{k}(-1)^k C_k^j x_{\tau-j}=\delta_\tau-\theta_1\delta_{\tau-1}-\ldots-\theta_q\delta_{\tau-q}, \qquad (1.80)$$

где L – оператор сдвига функции времени на один временной такт назад.

Из соотношения (1.80) можно выразить x_τ для любого $\tau = t - q,\ldots, t-1, t, t+1,\ldots, t++l$. Получаем

$$x_\tau=\left(\sum_{j=1}^{p}\alpha_j L^j\right)x_\tau+\left(1-\sum_{j=1}^{p}\alpha_j L^j\right)\left(\sum_{i=0}^{k}(-1)^i C_k^i x_{\tau-i}\right)+\left(1-\sum_{j=1}^{q}\theta_j L^j\right)\delta_\tau. \qquad (1.81)$$

Правые части этих соотношений представляют собой линейные комбинации p + k предшествующих (по отношению к левой части) значений анализируемого процесса x_τ, дополненные линейными комбинациями текущего и q предшествующих значений случайных остатков δ_τ. Причем коэффициенты, с помощью которых эти линейные комбинации подсчитываются, известны, т.к. выражаются в терминах уже оцененных параметров модели.

Этот факт и дает возможность использовать соотношения (1.81) для построения прогнозных значений анализируемого временного ряда на l тактов времени вперед. Теоретическую базу такого подхода к прогнозированию обеспечивает известный результат, в соответствии с которым наилучшим (в смысле среднеквадратической ошибки) линейным прогнозом в момент времени t с упреждением l является условное математическое ожидание случайной величины x_{t+l}, вычисленное при условии, что все значения x_τ до момента времени t. Этот результат является частным случаем общей теории прогнозирования (см. [237, 198, 235]).

Условное математическое ожидание $E(x_{t+l} \mid x_1,\ldots, x_t)$ получается применением операции усреднения к обеим частям (10) при $\tau = t + l$ с учетом следующих соотношений:

$$E(x_{t-j} \mid x_1,\ldots, x_t) = x_{t-j} \quad \text{при всех } j = 0, 1, 2,\ldots, t-1 \qquad (1.82)$$

$$E(x_{t+j} \mid x_1,\ldots, x_t) = \hat{x}_t^j \quad \text{при всех } j = 1, 2,\ldots; \qquad (1.83)$$

$$E(x_{t+j} \mid x_1,\ldots, x_t) = 0 \quad \text{при всех } j = 1, 2,\ldots; \qquad (1.84)$$

$$E(x_{t-j} \mid x_1,\ldots, x_t) = x_{t-j} - \hat{x}_{t-j-1}^1 \quad \text{при } j = 0, 1, 2,\ldots, t-1. \qquad (1.85)$$

Таким образом, определяется следующая процедура построения прогноза по известной до момента траектории временного ряда: по формулам (1.81) вычисляются ретроспективные прогнозы $\hat{x}_{t-q-1}^1$, $\hat{x}_{t-q}^1,\ldots, \hat{x}_{t-1}^1$ по предыдущим значениям временного ряда; при этом при вычислении начальных прогнозных значений $\hat{x}_{t-q+m-1}$ для x_{t-q+m} (m = 0, 1,…) по формулам (1.81) вместо условных средних $E(\delta_{t-q+m-j} \mid x_1,\ldots, x_{t-q+m})$, которые в общем случае следовало бы вычислять по формулам (1.85), подставляются их безусловные значения, равные нулю; используя формулы для $\tau > t$ и правила (1.82)–(1.85) подсчитываются условные математические ожидания для вычисления прогнозных значений.

Описанная процедура выглядит достаточно сложной. Однако при реалистичных значениях параметров p, q и k эта процедура в действительности оказывается весьма простой.

1. *В чем состоят основные отличия стационарных временных рядов от нестационарных?*
2. *В чем состоит идентификация моделей ARIMA?*
3. *Какова последовательность процесса идентификации моделей прогнозирования, содержащих сезонную компоненту?*
4. *Каков алгоритм (процедура) построения прогнозов на базе модели ARIMA?*

1.6. Адаптивные методы прогнозирования

Считается, что характерной чертой адаптивных методов прогнозирования является их способность непрерывно учитывать эволюцию динамических характеристик изучаемых процессов, «подстраиваться» под эту эволюцию, придавая, в частности, тем больший вес и тем более высокую информационную ценность имеющимся наблюдениям, чем ближе они к текущему моменту прогнозирования. Однако деление методов и моделей на «адаптивные» и «неадаптивные» достаточно условно. В известном смысле любой метод прогнозирования адаптивный, т.к. все они учитывают вновь поступающую информацию, в том числе наблюдения, сделанные с момента последнего прогноза. Общее значение термина заключается, по-видимому, в том, что «адаптивное» прогнозирование позволяет обновлять прогнозы с минимальной задержкой и с помощью относительно несложных математических процедур. Однако это не означает, что в любой ситуации адаптивные методы эффективнее тех, которые традиционно не относятся к таковым.

Простейший вариант метода (метод экспоненциального сглаживания [151]) уже рассматривался в связи с задачей выявления неслучайной составляющей анализируемого временного ряда. Постановка задачи прогнозирования с использованием простейшего варианта метода экспоненциального сглаживания формулируется следующим образом.

Пусть анализируемый временной ряд x_τ, $\tau = 1, 2,\ldots, t$ представлен в виде

$$x_\tau = a_0 + \varepsilon_\tau, \qquad (1.86)$$

где a_0 – неизвестный параметр, не зависящий от времени, а ε_τ – случайный остаток со средним значением, равным нулю, и конечной дисперсией. Как известно, экспоненциально взвешенная скользящая средняя ряда x_τ в точке t $\bar{x}_t(\lambda)$ с параметром сглаживания (параметром адаптации) λ ($0 < \lambda < 1$) определяется формулой

$$\bar{x}_t(\lambda) = \frac{1-\lambda}{1-\lambda^t}\sum_{j=0}^{t-1}\lambda^j x_{t-j}, \tag{1.87}$$

которая дает решение задачи: $\bar{x}_t(\lambda) = \arg\min_a \sum_{j=0}^{t-1}\lambda^j\left(x_{t-j}-a\right)^2$.

Коэффициент сглаживания λ можно интерпретировать также как коэффициент дисконтирования, характеризующий меру обесценения наблюдения за единицу времени.

Для рядов с «бесконечным прошлым» формула (1.87) сводится к виду

$$\bar{x}_t(\lambda) = (1-\lambda)\sum_{j=0}^{\infty}\lambda^j x_{t-j}. \tag{1.88}$$

В соответствии с простейшим вариантом метода экспоненциального сглаживания прогноз $\hat{x}_t^1$ для неизвестного значения x_{t+1} по известной до момента времени t траектории ряда x_t строится по формуле

$$\hat{x}_t^1 = \bar{x}_t(\lambda), \tag{1.89}$$

где значение $\bar{x}_t(\lambda)$ определено формулой (1.87) или (1.88), соответственно для короткого или длинного временного ряда.

Формула (1.89) удобна, в частности, тем, что при появлении следующего (t + 1)-го наблюдения x_{t+1} пересчет прогнозирующей функции $\hat{x}_{t+1}^1 = \bar{x}_{t+1}(\lambda)$ производится с помощью простого соотношения $\bar{x}_{t+1}(\lambda) = \lambda\bar{x}_t(\lambda) + (1-\lambda)x_{t+1}$.

Метод экспоненциального сглаживания можно обобщить на случай полиномиальной неслучайной составляющей анализируемого временного ряда, т.е. на ситуации, когда вместо (1.86) постулируется

$$x_{t+\tau} = a_0 + a_1\tau + \ldots + a_k\tau^k + \varepsilon_\tau, \qquad (1.90)$$

где $k \geq 1$. В соотношении (1.90) начальная точка отсчета времени сдвинута в текущий момент времени t, что облегчает дальнейшие вычисления. Соответственно, в схеме простейшего варианта метода прогноз $\hat{x}_t^1$ значения x_{t+1} будет определяться соотношениями (1.90) при $\tau = 1$ и (4)

$$\hat{x}_t^1 = \hat{x}_{t+1} = \hat{a}_0^{(k)}(t,\lambda) + \hat{a}_1^{(k)}(t,\lambda) + \ldots + \hat{a}_k^{(k)}(t,\lambda), \qquad (1.91)$$

где оценки $\hat{a}_j(t,\lambda), j = 0,1,\ldots,k$ получаются как решение оптимизационной задачи

$$\sum_{j=0}^{\infty} \lambda^j \left(x_{t-j} - a_0 - a_1 j - \ldots - a_k j^k\right)^2 \to \min_{a_0, a_1, \ldots, a_k}. \qquad (1.92)$$

Решение задачи (1.92) не представляет принципиальных трудностей.

Рассмотрим еще несколько методов, использующих идеологию экспоненциального сглаживания, которые развивают метод Брауна в различных направлениях.

Метод Хольта. Хольт [195] ослабил ограничения метода Брауна, связанные с его однопараметричностью, введением двух параметров сглаживания λ_1 и λ_2 ($0 < \lambda_1, \lambda_2 < 1$). В его модели прогноз $\hat{x}_t^l$ на l тактов времени в текущий момент t также определяется линейным трендом вида

$$\hat{x}_t^l = \hat{a}_0(t,\lambda_1,\lambda_2) + l\hat{a}_1(t,\lambda_1,\lambda_2), \qquad (1.93)$$

где обновление прогнозирующих коэффициентов производится по формулам

$$\begin{aligned} \hat{a}_0(t+1,\lambda_1,\lambda_2) &= \lambda_1 x_t + (1-\lambda_1)\left(\hat{a}_0(t,\lambda_1,\lambda_2) + \hat{a}_1(t,\lambda_1,\lambda_2)\right), \qquad (1.94) \\ \hat{a}_1(t+1,\lambda_1,\lambda_2) &= \lambda_2\left(\hat{a}_0(t+1,\lambda_1,\lambda_2) - \hat{a}_1(t,\lambda_1,\lambda_2)\right) + \\ &+ (1-\lambda_2)\hat{a}_1(t,\lambda_1,\lambda_2). \end{aligned}$$

Таким образом, прогноз по данному методу является функцией прошлых и текущих данных, параметров λ_1 и λ_2, а также начальных значений $\hat{a}_0(0,\lambda_1,\lambda_2)$ и $\hat{a}_1(0,\lambda_1,\lambda_2)$.

Метод Хольта–Уинтерса. Уинтерс [236] развил метод Хольта так, чтобы он охватывал еще и сезонные эффекты. Прогноз, сделанный в момент t на l тактов времени вперед, равен

$$\hat{x}_t^l = [\hat{a}_0(t) + l\hat{a}_1(t)]\omega_{t+l-N}, \tag{1.95}$$

где ω_τ – коэффициент сезонности, а N – число временных тактов, содержащихся в полном сезонном цикле. Сезонность в этой формуле представлена мультипликативно. Метод использует три параметра сглаживания λ_1, λ_2, λ_3 ($0 < \lambda_j < 1$, j = 1, 2, 3), а его формулы обновления имеют вид

$$\begin{aligned}
&\hat{a}_0(t+1) = \lambda_1 \frac{x_{t+1}}{\omega_{t+1-N}} + (1-\lambda_1)[\hat{a}_0(t) + \hat{a}_1(t)],\\
&\omega_{t+1} = \lambda_2 \frac{x_{t+1}}{\hat{a}_0(t+1)} + (1-\lambda_2)\omega_{t+1-N},\\
&\hat{a}_1(t+1) = \lambda_3[\hat{a}_0(t+1) - \hat{a}_0(t)] + (1-\lambda_3)\hat{a}_1(t).
\end{aligned} \tag{1.96}$$

Как и в предыдущем случае, прогноз строится на основании прошлых и текущих значений временного ряда, параметров адаптации λ_1, λ_2 и λ_3, а также начальных значений $\hat{a}_0(0), \hat{a}_1(0)$ и ω_0.

Аддитивная модель сезонности Тейла–Вейджа. В экономической практике чаще встречаются экспоненциальные тенденции с мультипликативно наложенной сезонностью. Поэтому перед использованием аддитивной модели члены анализируемого временного ряда обычно заменяют их логарифмами, преобразуя экспоненциальную тенденцию в линейную, а мультипликативную сезонность – в аддитивную. Преимущество аддитивной модели заключается в относительной простоте ее вычислительной реализации. Рассмотрим модель вида (в предположении, что исходные данные прологарифмированы)

$$\begin{aligned}
&x_\tau = a_0(\tau) + \omega_\tau + \delta_\tau,\\
&a_0(\tau) = a_0(\tau-1) + a_1(\tau),
\end{aligned} \tag{1.97}$$

где $a_0(\tau)$ – уровень процесса после элиминирования сезонных колебаний, $a_1(\tau)$ – аддитивный коэффициент роста, ω_t – аддитивный коэффициент сезонности, δ_t – белый шум.

Прогноз, сделанный в момент t на l временных тактов вперед, подсчитывается по формуле

$$\hat{x}_t^l = \hat{a}_0(t) + l\hat{a}_1(t) + \hat{\omega}_{t-N+l}, \qquad (1.98)$$

где коэффициенты $\hat{a}_0, \hat{a}_1$ и ω вычисляются рекуррентным образом с помощью следующих формул обновления

$$\begin{aligned} \hat{a}_0(\tau) &= \hat{a}_0(\tau-1) + \hat{a}_1(\tau-1) + \lambda_1\left[x_\tau - \hat{x}_{\tau-1}^1\right]; \\ \hat{a}_1(\tau) &= \hat{a}_1(\tau-1) + \lambda_1\lambda_2\left[x_\tau - \hat{x}_{\tau-1}^1\right]; \\ \hat{\omega}_t &= \hat{\omega}_{\tau-N} + (1-\lambda_1)\lambda_3\left[x_\tau - \hat{x}_{\tau-1}^1\right] \end{aligned} \qquad (1.99)$$

В этих соотношениях, как и прежде, N – число временных тактов, содержащихся в полном сезонном цикле, а λ_1, λ_2 и λ_3 – параметры адаптации.

1. *В чем состоит отличительная особенность адаптивных моделей прогнозирования?*
2. *Какова методологическая особенность метода Хольта?*
3. *В чем состоит усовершенствование метода Хольта в методе Хольта-Уинтерса?*

1.7. Метод группового учета аргументов

В настоящее время большую популярность для конкретных задач прогнозирования приобретает так называемый метод группового учета аргументов (МГУА), представляющий собой дальнейшее развитие метода регрессионного анализа. Он основан на некоторых принципах теории обучения и самоорганизации, в частности на принципе «селекции», или направленного отбора [52,51].

Метод реализует задачи синтеза оптимальных моделей высокой сложности, адекватной сложности исследуемого объекта (здесь под моделями понимается система регрессионных уравнений). Так, алгоритмы МГУА, построенные по схеме массовой селекции, осуще-

ствляют перебор возможных функциональных описаний объекта. При этом полное описание объекта [51]

$$y = f(x_1, \quad x_2, \quad ..., \quad x_m), \tag{1.100}$$

где f – некоторая функция, например степенной полином, заменяется рядами частных описаний:

1-й ряд селекции -

$$y_1 = f(x_1, x_2), \quad y_2 = f(x_1, x_3), \quad ..., \quad y_S = f(x_{m-1} x_m);$$

2-й ряд селекции -

$$z_1 = f(y_1, y_2), \quad z_2 = f(y_1, y_3), \quad ..., \quad z_k = f(y_{n-1} y_n).$$

и т. д.

Рассматриваются различные сочетания входных и промежуточных переменных, и для каждого сочетания строится модель, причем при построении рядов селекции используются самые регулярные переменные. Понятие регулярности является одним из основных в методе МГУА. Регулярность определяется минимумом среднеквадратической ошибки переменных на отдельной проверочной последовательности данных (исходный ряд делится на обучающую и проверочную последовательности). В некоторых случаях в качестве показателя регулярности используется коэффициент корреляции. Ряды строятся до тех пор, пока регулярность повышается, т. е. снижается ошибка или увеличивается коэффициент корреляции. Таким образом, из всей совокупности моделей выбирается такая, которая является оптимальной с точки зрения выбранного критерия.

Рассмотрим некоторые алгоритмы МГУА [51].

В алгоритмах с линейными полиномами в качестве частных описаний используются соотношения вида

$$y_k = a_0 + a_1 x_i + a_2 y_i, \quad 0 < i \le m. \tag{1.101}$$

Алгоритм синтезирует модели с последовательно увеличивающимся числом учитываемых аргументов. Так, модели первого селекционного ряда включают по два аргумента, модели второго ряда – три-четыре и т. д.

Алгоритмы с ковариациями и квадратичными описаниями оперируют с частными описаниями вида

$$y_i = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j; \tag{1.102}$$

$$y_k = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2 .$$

В данном случае модели усложняются не только за счет увеличения числа учитываемых аргументов, но и за счет роста степени описания.

В алгоритмах с последовательным выделением трендов в качестве таковых рассматриваются уравнения регрессии по одному аргументу, включая время: $f(x_1)$, $f(x_2)$, ..., $f(x_m)$. Для построения моделей используются частные описания вида

$$y = a_0 + a_1 f(x_1) + a_2 f(x_2) + ... + a_m f(x_m) . \quad (1.103)$$

Алгоритм работает таким образом, что вначале выделяется первый тренд и рассчитывается соответствующее отклонение (первый остаток) истинных значений функции от тренда. После чего это отклонение аппроксимируется вторым трендом и определяется второй остаток и т. д. На практике выделяют до пяти-шести трендов.

Среди основных алгоритмов МГУА наибольший интерес представляет обобщенный алгоритм, обеспечивающий получение наиболее точных моделей благодаря использованию в качестве опорной функции аддитивной и мультипликативной моделей трендов [51].

С целью сокращения числа входных аргументов в обобщенном алгоритме используется рассмотренный выше алгоритм последовательного выделения трендов для выбора оптимальной опорной функции, после чего осуществляется перебор всех возможных комбинаций выделенных трендов, либо в классе сумм, либо в классе произведений. Пусть, например, выбрана зависимость

$$\varphi_{+1} = f(t) + f_1(x_1) + f_2(x_2) + f_3(x_3) + f_4(x_4) . \quad (1.104)$$

Обобщенный алгоритм МГУА предусматривает перебор двенадцати комбинаций [51]

$$
\begin{aligned}
&\varphi_{+1} = f(t) + f_1(x_1) + f_2(x_2) + f_3(x_3) + f_4(x_4);\\
&\varphi_{+1} = f(t) + f_1(x_1) + f_2(x_2, x_3, x_4);\\
&\varphi_{+1} = f(t) + f_1(x_2) + f_2(x_2, x_3, x_4);\\
&\varphi_{+1} = f(t) + f_1(x_3) + f_2(x_2, x_3, x_4);\\
&\varphi_{+1} = f(t) + f_1(x_4) + f_2(x_2, x_3, x_4);\\
&\varphi_{+1} = f(t) + f_1(x_1) + f_2(x_2) + f_3(x_3, x_4);\\
&\varphi_{+1} = f(t) + f_1(x_1) + f_2(x_3) + f_3(x_2, x_4);\\
&\varphi_{+1} = f(t) + f_1(x_1) + f_2(x_4) + f_3(x_2, x_3);\\
&\varphi_{+1} = f(t) + f_1(x_2) + f_2(x_3) + f_3(x_1, x_4);\\
&\varphi_{+1} = f(t) + f_1(x_2) + f_2(x_4) + f_3(x_1, x_2);\\
&\varphi_{+1} = f(t) + f_1(x_3) + f_2(x_4) + f_3(x_1, x_2);\\
&\varphi_{+1} = f(t) + f_1(x_1, x_2, x_3, x_4).
\end{aligned}
\tag{1.105}
$$

В результате перебора определяется комбинация, дающая наиболее регулярное решение.

В случаях, когда процесс описывается большим числом переменных, использование обобщенного алгоритма затруднено и для сокращения перебора рекомендуется применять алгоритм с многофазной селекцией проекторов (операторов ортогонального проектирования). Понятие проектора введено в МГУА по аналогии с графическим представлением метода наименьших квадратов, согласно которому вектор выходной величины проектируется на плоскость аргументов. В соответствии с этим все алгоритмы МГУА рассматриваются как варианты последовательного проектирования выходной величины на плоскости переменных на каждом ряду селекции.

Выбирая определенный вид оператора проектирования, можно получить тот или иной алгоритм МГУА. Например, в алгоритме МГУА с последовательным выделением трендов выходная величина y проектируется только на первом ряду, далее проектируется уже остаток $\Delta r = y - f_r$ на оси наиболее эффективной переменной. Аппроксимационная функция имеет вид

$$
f_{n+1} = \overline{P}_{j0}\varphi + \overline{P}_{j1}\Delta_1 + \ldots + \overline{P}_{jn}\Delta_n, \tag{1.106}
$$

где $\overline{P}_{jr}\Delta_r$ – частное описание в виде полинома степени I по одной из наиболее эффективных переменных x_j отбираемых по коэффициенту корреляции: $0 \le l \le m$; P_{ij} – оператор ортогонального проектирования на подпространство $L_j = \{x_i^0, \quad x_i^1, \quad ..., \quad x_i^m\}$, $1 \le j \le (m+1)S, \quad j = \overline{1,l}$. Оператор P_{ij} можно представить в виде суперпозиции проекторов, соответствующих различным алгоритмам МГУА. Таким образом, выражение для f_{n+1} задает множество алгоритмов МГУА, отличающихся по способу проектирования остатка.

В ряде случаев для упрощения вида частных описаний и простоты определения оценок их коэффициентов используют прием ортогонализации переменных.

Непосредственное использование МГУА для целей прогнозирования основывается на теоремах, изложенных в [52, 51].

1. При любом разделении полного полинома заданной степени на частные полиномы критерий минимума среднеквадратической ошибки, определяемой на обучающей последовательности (первый критерий), позволяет однозначно определить оптимальные оценки всех коэффициентов, если число точек в обучающей последовательности больше числа членов каждого из частных полиномов по крайней мере на единицу.

2. При заданной степени полного полинома имеется много вариантов разбиения его на частные полиномы. Полный перебор всех комбинаций по критерию среднеквадратической ошибки, измеряемой на отдельной проверочной последовательности данных, позволяет найти единственное наилучшее разделение.

3. При постепенном нарастании степени полного полинома до некоторого ограниченного значения ошибка на проверочной последовательности либо непрерывно падает, либо имеет минимум по крайней мере при одном значении степени.

4. Если точки ранжированы по величине дисперсии, то имеется единственное значение отношения числа точек проверочной последовательности к числу точек обучающей последовательности, при котором достигается минимум числа рядов селекции и степени полного полинома.

5. В многорядном процессе алгоритмов МГУА среднеквадратическая ошибка от ряда к ряду не может возрастать независимо от пути, по которому идет селекция.

В качестве критериев получения оптимальной модели по МГУА или критерия регуляризации (точности) используются критерии [51]: $\Delta^2(2) \to \min$ и $\Delta(1) \to \min$, где $\Delta(1)$ – среднеквадратическая ошибка на проверочной последовательности данных (первая разность реальных и прогнозных значений); $\Delta^2(2)$ - среднеквадратическая ошибка приращений (вторая разность этих значений). Так, по критерию $\Delta(1) \to \min$ расчет проводится следующим образом [51]:

1. По заданному уравнению регрессии для периода упреждения прогноза $T_y = 1$ и периода предыстории $T_n = 2$ строится зависимость x(t).

2. Находятся рассогласования [x*(t)-x(t)] для всех точек проверочной последовательности ($N_{пр}$), x*(t) - реализации процесса, x(t) – прогнозные значения.

3. Рассчитывается величина ошибки

$$\delta(1) = \frac{1}{N_{пр}} \sqrt{\sum_{j=1}^{N_{пр}} \left[x_j^*(t) - x_j(t)\right]^2}; \quad \Delta(1) = \frac{\delta(1) \cdot 100\%}{\frac{1}{N_{пр}} \sqrt{\sum_{j=1}^{N_{пр}} \left[x_j^*(t)\right]^2}}. \tag{1.107}$$

Расчет по критерию $\Delta^2(2)$ проводится по схеме [51].

1. При $T_y = 1$, $T_n = 2$ при помощи прямого обучения по алгоритму МГУА находится уравнение регрессии для ошибки, например: $\Delta x_0 = f(\Delta x_{-1}, \Delta x_{-2}, x_{-1}, x_{-2})$. Затем строится кривая $\Delta x(t)$ для каждого частного уравнения регрессии.

2. Находятся разности $\left[\Delta x^*(t) - \Delta x(t)\right]$ для всех экспериментальных точек.

3. Рассчитывается величина ошибок

$$\delta^2(2) = \frac{1}{N - T_y} \sqrt{\sum_{j=1}^{N-T_y} \left[\Delta x_j^*(t) - \Delta x_j(t)\right]^2} \tag{1.108}$$

$$\Delta^2(2) = \frac{\delta^2(2) \cdot 100\%}{\frac{1}{N - T_y} \sum_{j=1}^{N-1} \left[\Delta x_j^*(t)\right]^2}. \tag{1.109}$$

Использование критерия $\Delta^2(2)$, как показывает анализ, повышает точность прогноза. Метод МГУА может быть эффективно использован для получения так называемых «системных многократных дифференциальных прогнозов». Под системой в данном случае можно понимать группу определенным образом связанных между собой «входных» и «выходных» показателей с заданным описанием связей, элементов, процессов, структуры.

Для получения прогнозов поведения сложных систем предполагается выполнение определенных условий [52, 51]:

1. Границы системы выбираются таким образом, чтобы можно было исключить лишь наименее важные связи системы с внешней средой.

2. Система включает определенное число переменных M и обратных связей f. Для получения надежного прогноза при построении модели достаточно использовать любые $m \geq M - f$ переменные.

3. Выбранные переменные не должны повторять друг друга.

4. Плохо прогнозируемые переменные следует исключить из модели.

Прогноз называется системным, если одновременно прогнозируются не менее T характеристических переменных системы. Переменные прогнозируются одновременно, шаг за шагом. При этом устраняется один из основных недостатков однократного прогноза – аргументы уравнений прогнозирования «не стареют» (носят последние по времени отсчета индексы).

Многократный прогноз можно вести на основе как алгебраических, так и дифференциальных или интегральных уравнений.

При получении долгосрочных дифференциальных прогнозов важным является установление устойчивости поведения системы. Наиболее распространенным способом установления области устойчивости (для линейных моделей) являются методы Ляпунова, критерии Гурвица – Рауса.

При реализации прогнозов важно установить критерий качества полученных прогнозных результатов. В [53] устанавливается своеобразная иерархия критериев прогноза в зависимости от глубины прогнозирования. Так, для краткосрочного прогноза в качестве критерия селекции предлагается использовать критерий регулярности – величину среднеквадратической ошибки, определяемой на точках проверочной последовательности, не участвующей в получении оценок коэффициентов. Для среднесрочных прогнозов предлагается использовать критерий несмещенности как более эффективный. При

наличии информации об изменении взаимосвязанных переменных появляется возможность использовать критерий, который является одним из наиболее эффективных при долгосрочном прогнозировании, именно критерий баланса переменных, т. е. минимизации суммы квадратов рассогласований самих значений промежуточных переменных и их модельных представлений. Данный критерий определяет «жесткость», неизменность структуры исследуемого объекта.

1. *В чем состоит суть метода МГУА?*
2. *Дайте определение понятия «регулярности»?*
3. *Опишите алгоритмическую последовательность применения метода МГУА.*

1.8. Теория распознавания образов

Весьма перспективным в настоящее время является использование для прогнозирования методов теории распознавания образов. Непосредственно с использованием этой теории решается комплекс задач, имеющих важное значение в прогнозировании [4, 24, 25, 67, 75, 105, 126].

Уже использование экстраполяционных методов для прогнозирования временных рядов предполагает однородность динамики. Действительно, исходный временной ряд, являющийся основой прогнозирования какого-либо процесса, может содержать в себе интервалы, внутри которых динамика характеризуется определенными отличными от других интервалов условиями. Естественно, эти интервалы на перспективу искажают полученный прогнозный результат. В этой связи возникает необходимость четкого выделения тех интервалов, для которых характерна однородная динамика. Решение этого вопроса эффективно реализуется с помощью методов теории распознавания образов [126].

Другим важным приложением теории распознавания образов для получения прогнозов является описание и прогнозирование поведения какого-либо объекта по набору признаков, определяющих поведение этого объекта.

Процедура прогнозирования на основе методов распознавания образов состоит в том, что выбираются классы состояний исследуемых объектов, которые могут быть заданы как диапазонами изменения некоторых параметров, так и определенными качественными характеристиками. По совокупности признаков, определяющих со-

стояние объектов, находится соответствие принадлежности каждого нового объекта (или объекта в будущем понятии времени) к определенному классу. Это позволяет дать прогноз состояния объекта или указать диапазон изменения параметров, характеризующих его на прогнозируемый период.

Одной из важнейших проблем, возникающих при получении конкретных прогнозов, является оценка исходной информации. При прогнозировании развития сложной системы возникает ситуация, когда поведение системы может быть описано с помощью многих различных показателей. Реализация прогнозов по всей совокупности этих показателей приводит к необходимости учитывать и взаимосвязи между ними, что подчас бывает весьма затруднительно. Ситуация облегчается, когда для реализации прогнозов используется аппарат распознавания образов и прогнозируются возможные варианты развития сложной системы. В этой связи важной является задача определения качества исходной информации, т. е. рассматриваемых показателей, для возможного описания исследуемой системы.

Интересным является при построении прогнозных моделей использование сочетаний методов, например, регрессионного анализа и распознавания образов.

1. *Охарактеризуйте основные постулаты теории распознавания образов.*
2. *В чем состоит процедура прогнозирования на основе методов распознавания образов?*
3. *Какие проблемы возникают при получении прогнозов на основе рассмотренного метода?*

1.9. Прогнозирование с использованием нейронных сетей, искусственного интеллекта и генетических алгоритмов

Жесткие статистические предложения о свойствах временных рядов ограничивают возможности методов математической статистики, теории распознавания образов, теории случайных процессов и т.п. Дело в том, что многие реальные процессы не могут адекватно быть описаны с помощью традиционных статистических моделей, поскольку, по сути, являются существенно нелинейными, и имеют

либо хаотическую, либо квазипериодическую, либо смешанную (стохастика + хаос-динамика+детерминизм) основу [14 ,20, 21].

В данной ситуации адекватным аппаратом для решения задач диагностики и прогнозирования могут служить специальные искусственные сети [29, 36, 104, 109] реализующие идеи предсказания и классификации при наличии обучающих последовательностей, причем, как весьма перспективные, следует отметить радиально-базисные структуры, отличающиеся высокой скоростью обучения и универсальными аппроксимирующими возможностями [114].

В его основе нейроинтеллекта лежит нейронная организация искусственных систем, которая имеет биологические предпосылки. Способность биологических систем к обучению, самоорганизации и адаптации обладает большим преимуществом по сравнению с современными вычислительными системами. Первые шаги в области искусственных нейронных сетей сделали в 1943 г. В.Мак-Калох и В.Питс. Они показали, что при помощи пороговых нейронных элементов можно реализовать исчисление любых логических функций [36]. В 1949 г. Хебб предложил правило обучения, которое стало математической основой для обучения ряда нейронных сетей [29]. В 1957-1962 гг. Ф. Розенблатт предложил и исследовал модель нейронной сети, которую он назвал персептроном [104]. В 1959 г. В. Видроу и М. Хофф предложили процедуру обучения для линейного адаптивного элемента – ADALINE. Процедура обучения получила название "дельта правило" [36]. В 1969 г. М. Минский и С. Пайперт опубликовали монографию "Персептроны", в которой был дан математический анализ персептрона, и показаны ограничения, присущие ему. В 80-е годы значительно расширяются исследования в области нейронных сетей. Д. Хопфилд в 1982 г. дал анализ устойчивости нейронных сетей с обратными связями и предложил использовать их для решений задач оптимизации. Т.Кохонен разработал и исследовал самоорганизующиеся нейронные сети. Ряд авторов предложил алгоритм обратного распространение ошибки, который стал мощным средством для обучения многослойных нейронных сетей [29, 36, 104]. В настоящее время разработано большое число нейросистем, применяемых в различных областях: прогнозировании, управлении, диагностике в медицине и технике, распознавании образов и т.д [1, 4, 28, 47, 51, 52].

Нейронная сеть – совокупность нейронных элементов и связей между ними. Основной элемент нейронной сети - это *формальный*

нейрон, осуществляющий операцию нелинейного преобразования суммы произведений входных сигналов на весовые коэффициенты

$$y = F\left(\sum_{i=1}^{n} w_i x_i\right) F(WX), \tag{1.110}$$

где $X = (x_1, x_2, ..., x_n)^T$ - вектор входного сигнала; $W = (w_1, w_2, ..., w_n)$ весовой вектор; F - оператор нелинейного преобразования.

Для обучения сети используются различные алгоритмы обучения и их модификации [9, 11, 22, 42, 70, 139]. Очень трудно определить, какой обучающий алгоритм будет самым быстрым при решении той или иной задачи. Наибольший интерес для нас представляет *алгоритм обратного распространения ошибки,* так как является эффективным средством для обучения многослойных нейронных сетей прямого распространения [85, 127]. Алгоритм минимизирует среднеквадратичную ошибку нейронной сети. Для этого с целью настройки синаптических связей используется метод градиентного спуска в пространстве весовых коэффициентов и порогов нейронной сети. Следует отметить, что для настройки синаптических связей сети используется не только метод градиентного спуска, но и методы сопряженных градиентов, Ньютона, квазиньютоновский метод [94]. Для ускорения процедуры обучения вместо постоянного шага обучения предложено использовать адаптивный шаг обучения $\alpha(t)$. Алгоритм с адаптивным шагом обучения работает в 4 раза быстрее. На каждом этапе обучения сети он выбирается таким, чтобы минимизировать среднеквадратическую ошибку сети [29, 36].

Для прогнозирующих систем на базе НС наилучшие качества показывает гетерогенная сеть, состоящая из скрытых слоев с нелинейной функцией активации нейронных элементов и выходного линейного нейрона. Недостатком большинства рассмотренных нелинейных функций активации является то, что область выходных значений их ограничена отрезком [0,1] или [-1,1]. Это приводит к необходимости масштабирования данных, если они не принадлежат указанным выше диапазонам значений. В работе предложено использовать логарифмическую функцию активации для решения задач прогнозирования, которая позволяет получить прогноз значительно точнее, чем при использовании сигмоидной функции.

Анализ различных типов НС показал, что НС может решать задачи сложения, вычитания десятичных чисел, задачи линейного ав-

торегрессионного анализа и прогнозирования временных рядов с использованием метода «скользящего окна» [70].

Проведенный анализ многослойных нейронных сетей и алгоритмов их обучения позволил выявить ряд недостатков и возникающих проблем:

1. Неопределенность в выборе числа слоев и количества нейронных элементов в слое;
2. Медленная сходимость градиентного метода с постоянным шагом обучения;
3. Сложность выбора подходящей скорости обучения α. Так как маленькая скорость обучения приводит к скатыванию НС в локальный минимум, а большая скорость обучения может привести к пропуску глобального минимума и сделать процесс обучения расходящимся;
4. Невозможность определения точек локального и глобального минимума, так как градиентный метод их не различает;
5. Влияние случайной инициализации весовых коэффициентов НС на поиск минимума функции среднеквадратической ошибки.

Большую роль для эффективности обучения сети играет архитектура НС [114]. При помощи трехслойной НС можно аппроксимировать любую функцию со сколь угодно заданной точностью [14, 110]. Точность определяется числом нейронов в скрытом слое, но при слишком большой размерности скрытого слоя может наступить явление, называемое перетренировкой сети. Для устранения этого недостатка необходимо, чтобы число нейронов в промежуточном слое было значительно меньше, чем число тренировочных образов. С другой стороны, при слишком маленькой размерности скрытого слоя можно попасть в нежелательный локальный минимум. Для нейтрализации этого недостатка можно применять ряд методов описанных в [94, 127].

Прогнозирование с использованием теории генетических алгоритмов. Впервые идея использования генетических алгоритмов для обучения (machine learning) была предложена в 1970-е годы [241, 245, 244, 246]. Во второй половине 1980-х к этой идее вернулись в связи с обучением нейронных сетей. Они позволяют решать задачи прогнозирования (в последнее время наиболее широко генетические алгоритмы обучения используются для банковских прогнозов), классификации, поиска оптимальных вариантов, и совершенно незаменимы в тех случаях, когда в обычных условиях решение задачи основано на интуиции или опыте, а не на строгом (в ма-

тематическом смысле) ее описании. Использование механизмов генетической эволюции для обучения нейронных сетей кажется естественным, поскольку модели нейронных сетей разрабатываются по аналогии с мозгом и реализуют некоторые его особенности, появившиеся в результате биологической эволюции [10, 28, 100].

Основные компоненты генетических алгоритмов – это стратегии репродукций, мутаций и отбор "индивидуальных" нейронных сетей (по аналогии с отбором индивидуальных особей) [28].

Важным недостатком генетических алгоритмов является сложность для понимания и программной реализации. Однако преимуществом является эффективность в поиске глобальных минимумов адаптивных рельефов, так как в них исследуются большие области допустимых значений параметров нейронных сетей. Другая причина того, что генетические алгоритмы не застревают в локальных минимумах – случайные мутации, которые аналогичны температурным флуктуациям метода имитации отжига.

В [28, 88, 100] есть указания на достаточно высокую скорость обучения при использовании генетических алгоритмов. Хотя скорость сходимости градиентных алгоритмов в среднем выше, чем генетических алгоритмов.

Генетические алгоритмы дают возможность оперировать дискретными значениями параметров нейронных сетей. Это упрощает разработку цифровых аппаратных реализаций нейронных сетей. При обучении на компьютере нейронных сетей, не ориентированных на аппаратную реализацию, возможность использования дискретных значений параметров в некоторых случаях может приводить к сокращению общего времени обучения.

В рамках «генетического» подхода в последнее время разработаны многочисленные алгоритмы обучения нейронных сетей, различающиеся способами представления данных нейронной сети в "хромосомах", стратегиями репродукции, мутаций, отбора [241, 244, 245].

1. *Каковы основные предпосылки применения нейронных сетей в прогнозировании?*
2. *В чем состоит принципиальная концепция построения нейронных сетей?*
3. *Какие типы нейросетевых структур используются в прогнозировании?*
4. *Для чего используются генетические алгоритмы в процессах обучения нейронных сетей?*

Выводы к теоретической главе

В этой главе были рассмотрены и проанализированы некоторые методы и алгоритмы прогнозирования, имеющие четкую математическую формализацию и позволяющие нам работать с временными рядами. Отметим, что на практике, кроме рассмотренных методов, для прогнозирования широко используются методы экспертных оценок, теория межотраслевого баланса, методы, основанные на теории игр, вариационного исчисления, спектрального анализа и др. [31, 32, 35, 41, 46, 49, 56, 96]. В последнее время все большее внимание уделяется исследованию и прогнозированию финансовых временных рядов с использованием теории динамических систем, теория хаоса. Это достаточно новая область, которая представляет собой популярный и активно развивающийся раздел математических методов экономики [12, 33, 50, 58, 59, 78, 82, 101, 140].

Рассмотренные в данной главе методы, помимо очевидных преимуществ и плюсов, имеют ряд существенных недостатков.

- Недостатки метода наименьших квадратов (МНК). Использование процедуры оценки, основанной на методе наименьших квадратов, предполагает обязательное удовлетворение целого ряда предпосылок, невыполнение которых может привести к значительным ошибкам: 1. Случайные ошибки имеют нулевую среднюю, конечные дисперсии и ковариации; 2. Каждое измерение случайной ошибки характеризуется нулевым средним, не зависящим от значений наблюдаемых переменных; 3. Дисперсии каждой случайной ошибки одинаковы, их величины независимы от значений наблюдаемых переменных (гомоскедастичность); 4. Отсутствие автокорреляции ошибок, т. е. значения ошибок различных наблюдений независимы друг от друга; 5. Нормальность. Случайные ошибки имеют нормальное распределение; 6. Значения эндогенной переменной x свободны от ошибок измерения и имеют конечные средние значения и дисперсии.
- Проблемы выбора адекватной модели. Выбор модели в каждом конкретном случае осуществляется по целому ряду статистических критериев, например, по дисперсии, корреляционному отношению и др.
- Дисконтирование. Классический метод наименьших квадратов предполагает равноценность исходной информации в модели. В реальной же практике будущее поведение процесса значительно в большей степени определяется поздними наблюдениями, чем

ранними. Это обстоятельство породило так называемое дисконтирование, т. е. уменьшение ценности более ранней информации.

- Проблема оценки достоверности прогнозов. Важным моментом получения прогноза с помощью МНК является оценка достоверности полученного результата. Для этой цели используется целый ряд статистических характеристик: 1. Оценка стандартной ошибки; 2. Средняя относительная ошибка оценки; 3. Среднее линейное отклонение; 4. Корреляционное отношение для оценки надежности модели; 5. Оценка достоверности выбранной модели через значимость индекса корреляции по Z-критерию Фишера; 6. Оценка достоверности модели по F-критерий Фишера; 7. Наличие автокорреляций (критерий Дарбина – Уотсона).
- Недостатки, обусловленные жесткой фиксацией тренда. Жесткие статистические предложения о свойствах временных рядов ограничивают возможности методов математической статистики, теории распознавания образов, теории случайных процессов и т.п., так как многие реальные процессы не могут адекватно быть описаны с помощью традиционных статистических моделей, поскольку по сути являются существенно нелинейными и имеют либо хаотическую, либо квазипериодическую, либо смешанную основу.
- Проблемы и недостатки метода экспоненциального сглаживания. Для метода экспоненциального сглаживания основным и наиболее трудным моментом является выбор параметра сглаживания α, начальных условий и степени прогнозирующего полинома. Кроме того, для определения начальных параметров модели остаются актуальными перечисленные недостатки МНК и проблема автокорреляций.
- Проблемы и недостатки метода вероятностного моделирования. Недостатком модели является требование большого количества наблюдений и незнание начального распределения, что может привести к неправильным оценкам.
- Проблемы и недостатки метода адаптивного сглаживания. При наличии достаточной информации можно получить надежный прогноз на интервал больший, чем при обычном экспоненциальном сглаживании. Но это лишь при очень длинных рядах. К сожалению, для данного метода нет строгой процедуры оценки необходимой или достаточной длины исходной информации, для конечных рядов нет конкретных условий оценки точности про-

гноза. Поэтому для конечных рядов существует риск получить весьма приблизительный прогноз, тем более что в большинстве случаев в реальной практике встречаются ряды, содержащие не более 20 – 30 точек.

- Проблемы и недостатки метода Бокса – Дженкинса (модели авторегрессии – скользящего среднего). Проблемы связанны, прежде всего, с неоднородностью временных рядов и практической реализации метода из-за своей сложности.
- Проблемы и недостатки методов, реализованных на базе нейронных сетей. Проблемы неопределенности в выборе числа слоев и количества нейронных элементов в слое, медленная сходимость градиентного метода с постоянным шагом обучения, сложность выбора оптимальной скорости обучения α, влияние случайной инициализации весовых коэффициентов НС на поиск минимума функции среднеквадратической ошибки. Одна из наиболее серьезных трудностей при обучении – это явление переобучения. Кроме того, использование немасштабированных данных может привести к «параличу» сети.
- Проблемы и недостатки методов, реализованных на базе генетических алгоритмов. Это проблема кодировки информации, содержащейся в модели нейронной сети, а также сложность понимания и программной реализации.

При построении реальных прогнозов всегда необходимо учитывать не только специфические особенности применяемых методов, но и их ограничения и недостатки.

ГЛАВА 2. ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ АДАПТИВНЫХ МЕТОДОВ ПРОГНОЗИРОВАНИЯ

2.1. Общие положения

Наиболее часто при практическом построении прогнозов экономических показателей приходиться учитывать их сезонность и цикличность. Для прогнозирования несезонных и сезонных процессов используется различный математический аппарат. Динамика многих финансово-экономических показателей имеет устойчивую колебательную составляющую. При исследовании месячных и квартальных данных часто наблюдаются внутри годичные сезонные колебания соответственно с периодом 12 и 4 месяцев. При использовании дневных наблюдений часто наблюдаются колебания с недельным (пятидневным) циклом. В этом случае для получения более точных прогнозных оценок необходимо правильно отобразить не только тренд, но и колебательную компоненту. Решение этой задачи возможно только при использовании специального класса методов и моделей [6, 62, 79].

В основе сезонных моделей лежат их несезонные аналоги, которые дополнены средствами отражения сезонных колебаний. Сезонные модели способны отражать как относительно постоянную сезонную волну, так и волну, динамически изменяющуюся в зависимости от тренда. Первая форма относится к классу аддитивных, а вторая – к классу мультипликативных моделей. Большинство моделей имеет обе эти формы. Наиболее широко в практике используются модели Хольта-Уинтерса, авторегрессии, модели Бокса-Дженкинса [15, 34].

При краткосрочном прогнозировании обычно более важна динамика развития исследуемого показателя на конце периода наблюдений, а не тенденция его развития, сложившаяся в среднем на всем периоде предыстории. Свойство динамичности развития финансово-экономических процессов часто преобладает над свойством инерционности, поэтому более эффективными являются адаптивные методы, учитывающие информационную неравнозначность данных.

Адаптивные модели и методы имеют механизм автоматической настройки на изменение исследуемого показателя. Инструментом прогноза является модель, первоначальная оценка параметров кото-

рой производится по нескольким первым наблюдениям. На ее основе делается прогноз, который сравнивается с фактическими наблюдениями. Далее модель корректируется в соответствии с величиной ошибки прогноза и вновь используется для прогнозирования следующего уровня, вплоть до исчерпания всех наблюдений. Таким образом, она постоянно «впитывает» новую информацию, приспосабливается к ней, и к концу периода наблюдения отображает тенденцию, сложившуюся на текущий момент. Прогноз получается как экстраполяция последней тенденции. В различных методах прогнозирования процесс настройки (адаптации) модели осуществляется по-разному. Базовыми адаптивными моделями являются:

- модель Брауна;
- модель Хольта;
- модель авторегрессии.

Первые две модели относятся к схеме скользящего среднего, последняя – к схеме авторегрессии. Многочисленные адаптивные методы, базирующиеся на этих моделях, различаются между собой способом числовой оценки параметров, определения параметров адаптации и компоновкой.

Согласно схеме **скользящего среднего**, оценкой текущего уровня является взвешенное среднее всех предшествующих уровней, причем веса при наблюдениях убывают по мере удаления от последнего (текущего) уровня, т. е. информационная ценность наблюдений тем больше, чем ближе они к концу периода наблюдений.

Согласно схеме **авторегрессии**, оценкой текущего уровня является взвешенная сумма порядков моделей «*p*» предшествующих уровней. Информационная ценность наблюдений определяется не их близостью к моделируемому уровню, а теснотой связи между ними.

Обе эти схемы имеют механизм отображения колебательного (сезонного или циклического) развития исследуемого процесса.

2.2. Полиномиальные модели временных рядов. Метод экспоненциальной средней

Пусть дан временной ряд x_t некоторого экономического показателя (например, динамика акций некоторой российской компании за 25 дней), включающий n = 25 наблюдений. Приняв первоначально коэффициент адаптации $\alpha = 0{,}5$ и период упреждения $\tau = 1$, требуется

аппроксимировать ряд с помощью адаптивной полиномиальной модели:

1. Нулевого порядка (p=0);
2. Первого порядка (p=1);
3. Второго порядка (p=2);
4. Оценить точность и качество прогнозов;
5. Сделать прогноз.

2.2.1. Адаптивная полиномиальная модель нулевого порядка (p=0)

Экспоненциальная средняя имеет вид

$$S_t = \alpha x_t + \beta S_{t-1}, \quad \beta = 1 - \alpha \tag{2.1}$$

Начальное условие: $S_0 = \hat{a}_{1,0}$, где $\hat{a}_{1,0}$ примем как среднее значение, например, первых пяти наблюдений.

$$\hat{a}_{1,0} = \frac{1}{5}\sum_{t=1}^{5} x_t = 511. \tag{2.2}$$

Расчетное (прогнозное) модельное значение с периодом упреждения τ будем определять из соотношения

$$\hat{x}_t^{\bullet} = S_{t-\tau} = 511. \tag{2.3}$$

Ошибку определим по формуле 2.4.

$$E = \frac{(x_t - \hat{x}^*)^2}{x_t}. \tag{2.4}$$

В Приложении 1 приведены методики оценки адекватности и точности прогноза.

Используя формулу 2.1 и принятое значение α = 0,5, рассчитаем.

При $t = 1$

$$S_1 = \alpha x_i + (1-\alpha)S_0 = 0{,}5 \cdot 520 + 0{,}5 \cdot 511 = 515{,}5\,.$$

$$\hat{x}_1^{\bullet} = S_0 = 511$$

При t =2

$$S_2 = 0{,}5 \cdot 497 + 0{,}5 \cdot 515{,}5 = 506{,}25$$

$$\hat{x}_2^{\bullet} = S_1 = 515{,}5$$

При t = 3

$$S_3 = 0{,}5 \cdot 504 + 0{,}5 \cdot 506{,}25 = 505{,}125$$

$$\hat{x}_3^{\bullet} = S_2 = 506{,}25$$

	A	B	C	D	E	F	G
1	τ	t	x_t	**p=0**			
2				**St**	$\hat{x}^*$	$(x_t - \hat{x}^*)^2$	ошибка
3	1	0		511			
4	1	1	520	515,5	511	81,00	0,16
5	1	2	497	506,25	515,5	342,25	0,69
6	1	3	504	505,125	506,25	5,06	0,01
7	1	4	525	515,063	505,125	395,02	0,75
. . .							
27	1	24	560	541,884	523,769	1312,69	2,34
28	1	25	529	535,442	541,884	166,01	0,31
29	1	26			535,442		
	α	**0,5**					
	β	**0,5**					

Рисунок 2.1. Прогнозирование временного ряда x_t на шаг вперед (адаптивная полиномиальная модель нулевого (p = 0) порядка)

Нами сделан прогноз на один шаг вперед, однако его нельзя считать оптимальным. Для получения адекватного прогноза необходимо подобрать такое значение α, чтобы сумма квадратов отклонений и ошибка прогноза была минимальной.

Для определения оптимального значения α протабулируем его от 0,1 до 0,9 с шагом 0,1. Затем каждый раз подставим его в расчетную модель для получения прогноза и величины ошибки. Таким об-

разом подбирается такое значение α, при котором ошибка также будет минимальна.

Распределение ошибки прогнозирования относительно параметра α показано на рисунке 2.2.

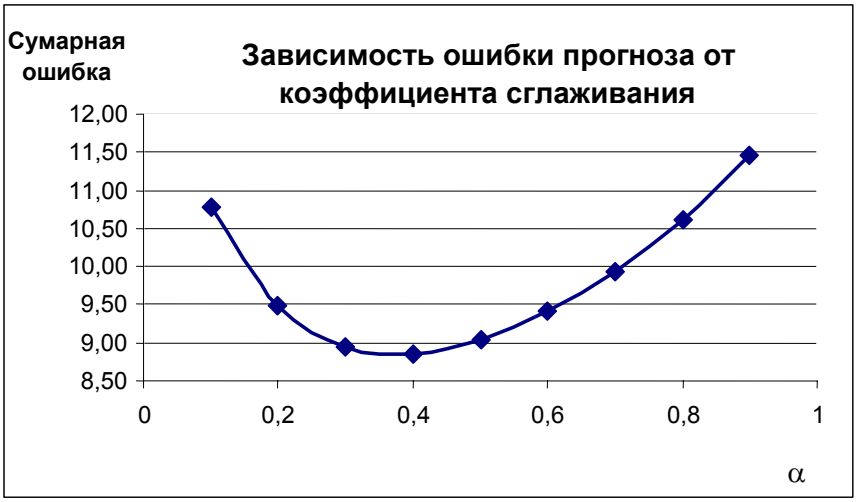


Рисунок 2.2. Зависимость ошибки прогнозирования от α

Рисунок 2.3. Результаты прогноза с оптимальным значением α

По рисунку 2.2 видно, что для модели нулевого порядка оптимальным является значение α = 0,4, которое определяется на основании минимума суммарной ошибки E = 8,85.

Результаты полученного прогноза показаны на рисунке 2.3.

Численные значения прогнозирования показаны на рисунке 2.4.

	A	B	C	D	E	F	G
1	τ	**t**	x_t	**p=0**			
2				**St**	$\hat{x}^*$	$(x_t - \hat{x}^*)^2$	ошибка
3	1	0		511			
4	1	1	520	514,6	511	511	81,00
5	1	2	497	507,56	514,6	515,5	309,76
6	1	3	504	506,136	507,56	506,25	12,67
7	1	4	525	513,682	506,136	505,125	355,85
. . .							
27	1	24	560	537,895	523,159	523,769	1357,27
28	1	25	529	534,337	537,895	541,884	79,13
29	1	26			534,337	535,442	
	α	**0,4**					
	β	**0,6**					

Рисунок 2.4. Результаты прогноза при α = 0,4

Из рисунка 2.4 видно, что в результате подбора оптимального значения α был получен прогноз на 26 шаг равный 534,3. Данное значение прогноза можно считать более точным.

2.2.2. Адаптивная полиномиальная модель первого порядка (p=1)

Первоначально по данным временного ряда x_t, находим МНК-оценку линейного тренда

$$\hat{x}_t = \hat{a}_1 + \hat{a}_2 t \qquad (2.5)$$

и принимаем $\hat{a}_{1,0} = \hat{a}_1$ и $\hat{a}_{2,0} = \hat{a}_2$.

Для нахождения коэффициентов $\hat{a}_{1,0}$ и $\hat{a}_{2,0}$ на графике временного ряда x_t, добавим линию тренда (рисунок 2.5). В нашем случае уравнение тренда имеет вид

$$\hat{x}_t = 498 + 1{,}2t,$$

откуда $\hat{a}_{1,0} = \hat{a}_1 = 498$ и $\hat{a}_{2,0} = \hat{a}_2 = 1{,}2$.

Рисунок 2.5. Оценка линии регрессии МНК

Экспоненциальные средние 1-го и 2-го порядка определяются как

$$S_t = \alpha x_t + \beta S_{t-1} \tag{2.6}$$

$$S_t^{[2]} = \alpha S_t + \beta S_{t-1}^{[2]} \tag{2.7}$$

где β=1-α.

Отсюда начальные условия

$$S_0 = \hat{a}_{1,0} - \frac{\beta}{\alpha}\hat{a}_{2,0} \tag{2.8}$$

$$S_0^{[2]} = \hat{a}_{1,0} - \frac{2\beta}{\alpha}\hat{a}_{2,0} \qquad (2.9)$$

Оценка модельного (прогнозируемого) значения ряда с периодом упреждения τ равна

$$\hat{x}_t^{\bullet} = \left(2 + \frac{\alpha}{\beta}\tau\right)S_{t-\tau} - \left(1 + \frac{\alpha}{\beta}\tau\right)S_{t-\tau}^{[2]} \qquad (2.10)$$

$$S_0 = \hat{a}_{1,0} - \frac{\beta}{\alpha}\hat{a}_{2,0} = 498 - 0{,}5/0{,}5 \cdot 1{,}2 = 496{,}8$$

$$S_0^{[2]} = \hat{a}_{1,0} - \frac{2\beta}{\alpha}\hat{a}_{2,0} = 498 - 2 \cdot 1{,}2 = 495{,}6$$

По формуле (2.10) найдем

$$\hat{x}_t^{\bullet} = \left(2 + \frac{\alpha}{\beta}\tau\right)S_{t-\tau} - \left(1 + \frac{\alpha}{\beta}\tau\right)S_{t-\tau}^{[2]} = \left(2 + \frac{0{,}5}{0{,}5} \cdot 1\right) \cdot 496{,}8 - \left(1 + \frac{0{,}5}{0{,}5} \cdot 1\right) \cdot 495{,}6 = 499{,}2$$

	A	B	C	D	E	F	G	H
1	τ	t	x_t	**p=1**				
2				St	St[2]	$\hat{x}^*$	$(x_t - \hat{x}^*)^2$	ошибка
3	1	0		496,80	495,60			
4	1	1	520	508,40	502,00	499,20	432,64	0,83
5	1	2	497	502,70	502,35	521,20	585,64	1,18
6	1	3	504	503,35	502,85	503,40	0,36	0,00
7	1	4	525	514,18	508,51	504,35	426,42	0,81
. . .								
27	1	24	560	541,88	532,37	525,61	1182,83	2,11
28	1	25	529	535,44	533,90	560,92	1018,85	1,93
29	1	26				538,52		
	α	**0,5**						
	β	**0,5**						

Рисунок 2.6. Результаты расчетов прогнозной модели при α = 0,5

При $t = 1$ экспоненциальные средние равны

$$S_1 = \alpha x_1 + \beta S_0 = 0{,}5 \cdot 520 + 0{,}5 \cdot 496{,}8 = 508{,}4$$
$$S_1^{[2]} = \alpha S_1 + \beta S_0^{[2]} = 0{,}5 \cdot 508{,}4 + 0{,}5 \cdot 495{,}6 = 502{,}00$$

По формуле (2.10) найдем

$$\hat{x}_t^{\bullet} = \left(2 + \frac{\alpha}{\beta}\tau\right) S_{t-\tau} - \left(1 + \frac{\alpha}{\beta}\tau\right) S_{t-\tau}^{[2]} = \left(2 + \frac{0{,}5}{0{,}5} \cdot 1\right) \cdot 508{,}4 - \left(1 + \frac{0{,}5}{0{,}5} \cdot 1\right) \cdot 502{,}0 = 521{,}2$$

Результаты расчетов $\hat{x}_t^{\bullet}$ приведены на рисунке 2.6. Для нашего примера первоначально рассчитываются прогнозные значения при $\alpha = 0{,}5$ и $\tau = 1$. Далее необходимо определить оптимальное значение α, исходя из соображения минимума суммарной ошибки. Для этого, как и в первой модели, подбираем такое значение α, при котором суммарная ошибка будет минимальной.

На рисунке 2.7 показаны результаты определения оптимального параметра сглаживания.

Рисунок 2.7. Определение оптимального значения α

На рисунке видно, что минимум ошибки работы прогнозной мо-

дели будет при $\alpha = 0{,}1$ (суммарная ошибка $E = 9{,}06$).

Результаты прогнозирования при выбранном оптимальном значении α покажем на рисунке 2.8.

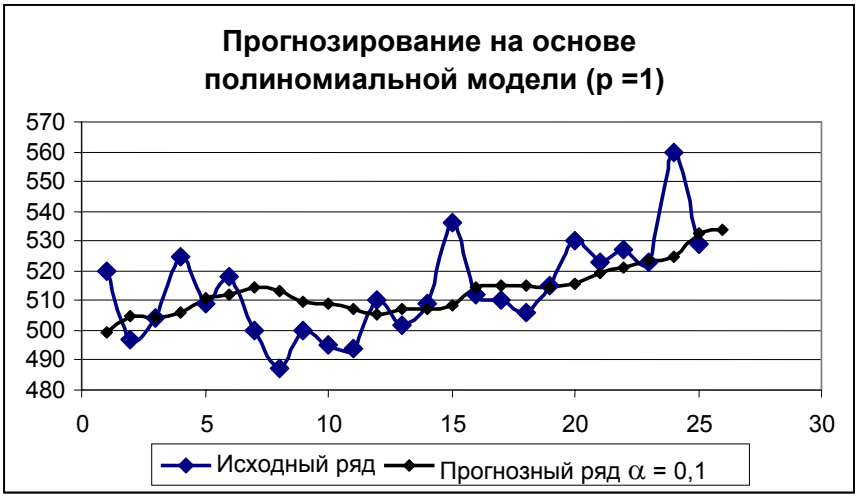


Рисунок 2.8. Результаты прогноза

2.2.3. Адаптивная полиномиальная модель второго порядка (р=2)

По данным временного ряда x_t находим МНК-оценку параболического тренда

$$\hat{x}_t = \hat{a}_1 + \hat{a}_2 t + \hat{a}_3 t^2 \qquad (2.11)$$

Для модели второго порядка уравнение параболического тренда имеет вид (см. рисунок 2.9):

$$\hat{x}_t = 515{,}96 - 2{,}79t + 0{,}15t^2$$

Отсюда

$$\hat{a}_{1,0} = \hat{a}_1 = 515{,}96; \quad \hat{a}_{2,0} = \hat{a}_2 = -2{,}79; \text{ и } \hat{a}_{3,0} = \hat{a}_3 = 0{,}15$$

Рисунок 2.9. Нахождение МНК-оценки параболического тренда по данным временного ряда x_t

Экспоненциальные средние 1-го, 2-го и 3-го порядка

$$\begin{aligned} S_t &= \alpha x_t + \beta S_{t-1}; \\ S_t^{[2]} &= \alpha S_t + \beta S_{t-1}^{[2]}; \\ S_t^{[3]} &= \alpha S_t^{[2]} + \beta S_{t-1}^{[3]} \end{aligned} \qquad (2.12)$$

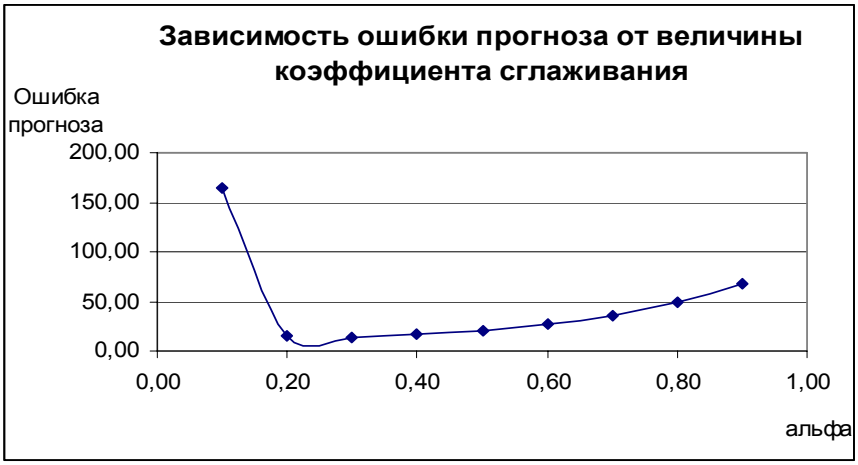


Рисунок 2.10. Определение оптимального значения α

Начальные условия определяются по следующим формулам:

$$S_0 = \hat{a}_{1,0} - \frac{\beta}{\alpha}\hat{a}_{2,0} + \frac{\beta(2-\alpha)}{2\alpha^2}\hat{a}_{3,0}; \qquad (2.13)$$

$$S_0^{[2]} = \hat{a}_{1,0} - \frac{2\beta}{\alpha}\hat{a}_{2,0} + \frac{\beta(3-2\alpha)}{\alpha^2}\hat{a}_{3,0}; \qquad (2.14)$$

$$S_0^{[3]} = \hat{a}_{1,0} - \frac{3\beta}{\alpha}\hat{a}_{2,0} + \frac{3\beta(4-3\alpha)}{2\alpha^2}\hat{a}_{3,0} \qquad (2.15)$$

Оценка модельного (прогнозируемого) значения с периодом упреждения τ находим из выражения

$$\hat{x}_t^{\bullet} = \left[6\beta^2 + (6-5\alpha)\alpha\cdot\tau + \alpha^2\tau^2\right]\frac{S_{t-\tau}}{2\beta^2} - \left[\begin{matrix}6\beta^2 + 2(5-4\alpha)\alpha\tau + \\ 2\alpha^2\tau^2\end{matrix}\right]\frac{S_{t-\tau}^{[2]}}{2\beta^2} + \\ + \left[2\beta^2 + (4-3\alpha)\alpha\tau + \alpha^2\tau^2\right]\frac{S_{t-\tau}^{[3]}}{2\beta^2}. \qquad (2.16)$$

Как и в предыдущих примерах, определяем оптимальное значение коэффициента сглаживания (см. рисунок 2.10). С учетом полученного оптимально значения α = 0,1 (E = 9,06) построим прогноз (см. рисунок 2.11).

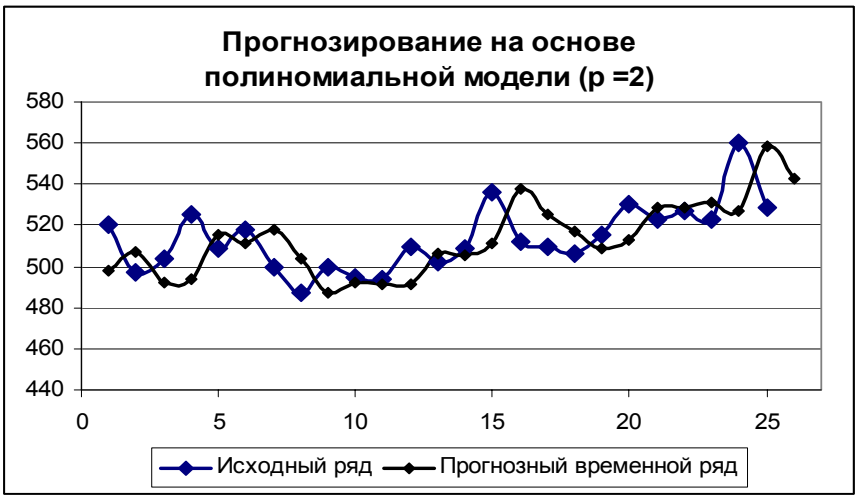


Рисунок 2.11. Прогноз на основе модели (р = 2)

Задания для самостоятельного выполнения.

1. *Используя данные таблиц Приложения Б, построить прогноз для* адаптивной полиномиальной модели нулевого (p=0), первого (p=1), второго (p=2) порядков.
2. *Подобрать оптимальные параметры прогнозных моделей.*
3. *Рассчитать ошибку прогнозирования, дополнительно руководствуясь теоретическими положениями, приведенными в Приложении А.*

2.3. Прогнозирование с использованием модели Уинтерса (экспоненциального сглаживания с мультипликативной сезонностью и линейным ростом)

Данную модель удобно использовать при небольшом объеме исходных данных. Допустим, мы имеем количество родившихся в каждом квартале (тыс. чел.) за два года (n = 8). Данные об этом представлены в таблице 2.1.

Таблица 2.1. Исходные x_t и расчетные значения количества родившихся по кварталам (2003-2005 гг.).

t	Год	Цикл k_t	Квартал (фаза цикла) ν_t	x_t	$\hat{x}_t$	$\frac{x_t}{\hat{x}_t}$	$\hat{x}_t^*$	Ошибка	$(x_t - \hat{x}_t^*)$
1	2003	1	1	499	483,91	1,031	508,28	1,9%	9,28
2			2	475	475,36	0,999	486,55	2,4%	4,55
3			3	452	466,82	0,968	452,32	0,1%	0,32
4			4	415	458,27	0,906	422,58	1,8%	7,58
5	2004	2	1	481	449,72	1,070	457,84	4,8%	23,16
6			2	467	441,17	1,059	443,15	5,1%	23,85
7			3	431	432,63	0,996	422,95	1,9%	8,05
8			4	412	424,08	0,972	397,88	3,4%	14,12
9	2005	3	Прогноз				448,16	Дисперсия:	56,04
							σ:		7,49

Требуется определить расчетные значения и прогноз (при t = 9) количества родившихся, воспользовавшись моделью экспоненци-

ального сглаживания с мультипликативной сезонностью и линейным ростом (модель Уинтерса) при периоде упреждения $\tau = 1$ и параметрах адаптации $\alpha_1 = 0{,}2$; $\alpha_2 = 0{,}3$ и $\alpha_3 = 0{,}4$.

Сезонная модель Уинтерса с линейным ростом имеет вид

$$x_t = a_{1,t} f_{v_t k_t} + \varepsilon_t, \tag{2.17}$$

где x_t – исходный временной ряд $t = 1,2,...,n$;

$a_{1,t}$ – параметр характеризующий линейную тенденцию развития процесса, т.е. средние значения уровня исследуемого временного ряда x_t в момент t;

f_{vtkt} – коэффициент сезонности для v_t фазы k_t-го цикла; v_t=1,2,…,l где $v_t=t-l(k_t-1)$;

l – число фаз в полном цикле (в месячных временных рядах l =12, в квартальных $l=4$ и т.д.);

ε_t – случайная ошибка. Обычно предполагается, что вектор $\varepsilon \in N_n\left(0, \sigma^2 I_n\right)$, где $\varepsilon = \left(\varepsilon_1, ..., \varepsilon_t, ..., \varepsilon_n\right)^T$

I_n – единичная матрица размерности $(n \times n)$.

Адаптивные параметры модели оцениваются с помощью рекуррентной экспоненциальной схемы по данным временного ряда x_t, состоящего из n наблюдений

$$\begin{cases} \hat{a}_{1,t} = \alpha_1 \dfrac{x_t}{\hat{f}_{v_t, k_{t-1}}} + (1-\alpha)\left(\hat{a}_{1,t-1} + \hat{a}_{2,t-1}\right) \\ \hat{f}_{v_t . k_t} = \alpha_2 \dfrac{x_t}{\hat{a}_{1,t}} + (1-\alpha_2)\hat{f}_{v_t, k_{t-1}} \\ \hat{a}_{2,t} = \alpha_3\left(\hat{a}_{1,t} - \hat{a}_{1,t-1}\right) + (1-\alpha_3)\hat{a}_{2,t-1} \\ \hat{x}_t^{\bullet} = \left(\hat{a}_{1,t-\tau} + \tau \hat{a}_{2,t-\tau}\right)\hat{f}_{v_t, k_t - 1} \end{cases} \tag{2.18}$$

где $a_{2,t}$ – прирост среднего уровня ряда от момента t - 1 к моменту t;

$\hat{x}_t^{\bullet} = \hat{x}_\tau(t)$ – расчетное значение временного ряда, определяемое для момента времени t с периодом упреждения τ, т.е. по данным момента (t -τ);

α_1, α_2, α_3 – параметры адаптации экспоненциального сглаживания, причем ($0 < \alpha_1, \alpha_2\ \alpha_3 < 1$).

При этом увеличение α_j (j = 1,2,3) ведет к увеличению веса более

поздних наблюдений, а уменьшение α_i – к улучшению сглаживания случайных отклонений. Эти два требования находятся в противоречии, и поиск компромиссного сочетания значений составляет задачу оптимизации модели.

Экспоненциальное выравнивание всегда требует предыдущей оценки сглаживаемой величины. Когда процесс адаптации только начинается, должны быть начальные значения, предшествующие первому наблюдению. В нашей задаче предстоит определить начальные условия: $\hat{a}_{1,0}$; $\hat{a}_{2,0}$; $\hat{f}_{v_t,0}$, где $v_t = 1,2,...,l$.

Таким образом, расчетные значения $\hat{x}_t^{\bullet}$ являются функцией всех прошлых значений исходного временного ряда x_t, параметров α_1, α_2 и α_3 и начальных условий. Влияние начальных условий на расчетное значение зависит от величины весов α_j и длины ряда, предшествующего моменту t. Влияние $\hat{a}_{1,0}$; $\hat{a}_{2,0}$ обычно уменьшается быстрее, чем $\hat{f}_{v_t,0}$, $\hat{a}_{1,t}$ *и* $\hat{a}_{2,t}$ пересматриваются на каждом шаге, а $\hat{f}_{v_t,k_t}$ только один раз за цикл.

Решение.

Первоначально по $n = 8$ наблюдениям временного ряда x_t найдем МНК-оценку линейного тренда $\hat{x}_t = a_0 + a_1 t$. В результате расчета имеем

$$\hat{x}_t = 492{,}46 - 8{,}5476 \cdot t$$

Определим начальные условия

$$\hat{a}_{1,0} = \hat{a}_0 = 492{,}46; \quad \hat{a}_{2,0} = \hat{a}_1 = -8{,}5476$$

Мультипликативные коэффициенты сезонности нулевого цикла.

$\hat{f}_{v_t,0}$ определим как среднюю арифметическую индексов сезонности $x_t / \hat{x}_t$ для v_t-й фазы в исходном временном ряду

$$\hat{f}_{1,0} = \frac{1{,}031 + 1{,}070}{2} = 1{,}050; \quad \hat{f}_{2,0} = \frac{0{,}999 + 1{,}059}{2} = 1{,}029;$$

$$\hat{f}_{3,0} = \frac{0{,}968 + 0{,}996}{2} = 0{,}982; \quad \hat{f}_{4,0} = \frac{0{,}906 + 0{,}972}{2} = 0{,}939.$$

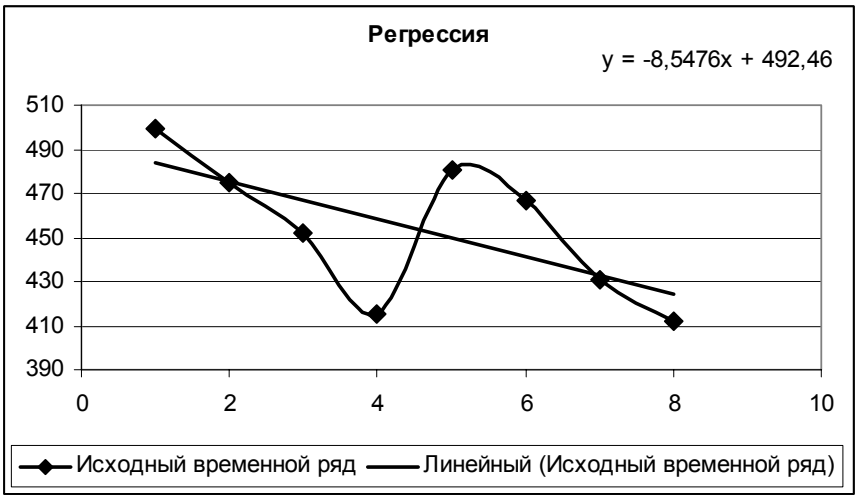


Рисунок 2.12. МНК-оценка линейного тренда

Расчеты будем проводить при параметрах адаптации $\alpha_1 = 0{,}2$; $\alpha_2 = 0{,}3$; $\alpha_3 = 0{,}4$ и периоде упреждения $\tau = 1$.

Расчетные значения для 1-го цикла ($k_t = 1, v_t = t$).
Согласно (18) при $t = 1$ имеем

$$\hat{x}_1^{\bullet} = \left(\hat{a}_{1,0} + \hat{a}_{2,0}\right) \cdot \hat{f}_{1,0} = \left(492{,}46 - 8{,}5476\right) \cdot 1.050 = 508{,}28$$

$$\hat{a}_{1,1} = \alpha_1 \cdot \frac{x_1}{\hat{f}_{1,0}} + \left(1 - \alpha_1\right)\left(\hat{a}_{1,0} + \hat{a}_{2,0}\right) = 0.2\frac{499}{1{,}050} + \left(1 - 0{,}2\right) \times$$
$$\times \left(492{,}46 - 8{,}5476\right) = 482{,}14$$

$$\hat{f}_{1,1} = \alpha_2 \frac{x_1}{\hat{a}_{1,1}} + \left(1 - \alpha_2\right) \cdot \hat{f}_{1,0} = 0.3\frac{499}{482{,}14} + \left(1 - 0{,}3\right) \cdot 1{,}050 = 1{,}046$$

$$\hat{a}_{2,1} = \alpha_3\left(\hat{a}_{1,1} - a_{1,0}\right) + \left(1 - \alpha_3\right) \cdot \hat{a}_{2,0} = 0.4\left(482{,}14 - 492{,}46\right) + 0.6 \times$$
$$\times \left(-8{,}5476\right) = -9{,}255;$$

при $t = 2$

$$\hat{x}_2^{\bullet} = \left(\hat{a}_{1,1} + \hat{a}_{2,1}\right) \cdot \hat{f}_{2,0} = \left(482{,}14 - 9{,}255\right) \cdot 1{,}029 = 486{,}55$$

$$\hat{a}_{1,2} = \alpha_1 \cdot \frac{x_2}{\hat{f}_{2,0}} + \left(1 - \alpha_1\right)\left(\hat{a}_{1,1} + \hat{a}_{2,1}\right) = 0{,}2\frac{475}{1{,}029} + 0.8 \times$$

$$\times \left(482{,}14 - 9{,}255\right) = 470{,}64$$

$$\hat{f}_{2,1} = \alpha_2 \frac{x_2}{\hat{a}_{1,2}} + \left(1 - \alpha_2\right) \cdot \hat{f}_{2,0} = 0{,}3\frac{475}{470{,}64} + 0{,}7 \cdot 1{,}029 = 1{,}023$$

$$\hat{a}_{2,2} = \alpha_3\left(\hat{a}_{1,2} - a_{1,1}\right) + \left(1 - \alpha_3\right) \cdot \hat{a}_{2,1} = 0{,}4\left(470{,}64 - 482{,}14\right) + 0{,}6 \times$$

$$\times \left(-9{,}255\right) = -10{,}153$$

при $t = 3$

$$\hat{x}_3^{\bullet} = \left(\hat{a}_{1,2} + \hat{a}_{2,2}\right) \cdot \hat{f}_{3,0} = \left(470{,}64 - 10{,}153\right) \cdot 0{,}982 = 452{,}32$$

$$\hat{a}_{1,3} = \alpha_1 \cdot \frac{x_3}{\hat{f}_{3,0}} + \left(1 - \alpha_1\right)\left(\hat{a}_{1,2} + \hat{a}_{2,2}\right) = 0{,}2\frac{452}{0.982} + 0.8 \times$$

$$\left(470{,}64 - 10{,}153\right) = 460{,}43$$

$$\hat{f}_{3,1} = \alpha_2 \frac{x_3}{\hat{a}_{1,3}} + \left(1 - \alpha_2\right) \cdot \hat{f}_{3,0} = 0{,}3\frac{452}{460{,}43} + 0.7 \cdot 0.982 = 0{,}982$$

$$\hat{a}_{2,3} = \alpha_3\left(\hat{a}_{1,3} - a_{1,2}\right) + \left(1 - \alpha_3\right) \cdot \hat{a}_{2,2} = 0{,}4\left(460{,}43 - 470{,}64\right) + 0{,}6 \times$$

$$\left(-10{,}153\right) = -10{,}179$$

при $t = 4$

$$\hat{x}_4^{\bullet} = \left(\hat{a}_{1,3} + \hat{a}_{2,3}\right) \cdot \hat{f}_{4,0} = \left(460{,}43 - 10{,}179\right) \cdot 0{,}939 = 422{,}58$$

$$\hat{a}_{1,4} = 0{,}2\frac{415}{0{,}939} + 0{,}8\left(460{,}43 - 10{,}179\right) = 448{,}63$$

$$\hat{f}_{4,1} = 0{,}3\frac{415}{448{,}63} + 0{,}7 \cdot 0{,}939 = 0{,}934$$

$$\hat{a}_{2,4} = 0{,}4\left(448{,}63 - 460{,}43\right) + 0{,}6 \cdot \left(-10{,}179\right) = -10{,}825$$

Расчетные значения для 2-го цикла(k_t = 2,v_t = t-4). Здесь нам понадобятся коэффициенты сезонности, найденные для 1-го цикла

$$\hat{f}_{1,1} = 1{,}046; \quad \hat{f}_{2,1} = 1{,}023; \quad \hat{f}_{3,1} = 0{,}982 \text{ è } \hat{f}_{4,1} = 0{,}934$$

при t = 5

$$\hat{x}_5^{\bullet} = \left(\hat{a}_{1,4} + \hat{a}_{2,4}\right) \cdot \hat{f}_{1,1} = \left(448{,}63 - 10{,}825\right) \cdot 1{,}046 = 457{,}84$$

Т.к. $\hat{x}_5^{\bullet}$ относится ко 2-му циклу(k_t = 2), при выборе $\hat{f}_{v_t, k_t - 1}$ исходили, что v_t = 5-4 = 1

$$\hat{a}_{1,5} = \alpha_1 \cdot \frac{x_5}{\hat{f}_{1,1}} + \left(1 - \alpha_1\right)\left(\hat{a}_{1,4} + \hat{a}_{2,4}\right) = 0{,}2\frac{481}{1{,}046} + 0{,}8 \times$$
$$\left(448{,}63 - 10{,}825\right) = 442{,}24$$
$$\hat{f}_{1,2} = \alpha_2 \frac{x_5}{\hat{a}_{1,5}} + \left(1 - \alpha_2\right) \cdot \hat{f}_{1,1} = 0{,}3\frac{481}{442{,}24} + 0{,}7 \cdot 1{,}046 = 1{,}058$$
$$\hat{a}_{2,5} = \alpha_3\left(\hat{a}_{1,5} - a_{1,4}\right) + \left(1 - \alpha_3\right) \cdot \hat{a}_{2,4} = 0{,}4\left(442{,}24 - 448{,}63\right) + 0{,}6 \times$$
$$\times\left(-10{,}825\right) = -9{,}053$$

при t = 6

$$\hat{x}_6^{\bullet} = \left(\hat{a}_{1,5} + \hat{a}_{2,5}\right) \cdot \hat{f}_{2,1} = \left(442{,}24 - 9{,}053\right) \cdot 1{,}023 = 443{,}15$$
$$\hat{a}_{1,6} = \alpha_1 \cdot \frac{x_6}{\hat{f}_{2,1}} + \left(1 - \alpha_1\right)\left(\hat{a}_{1,5} + \hat{a}_{2,5}\right) = 0{,}2\frac{467}{1{,}023} + 0{,}8 \times$$
$$\times\left(442{,}24 - 9{,}053\right) = 437{,}85$$
$$\hat{f}_{2,2} = \alpha_2 \frac{x_6}{\hat{a}_{1,6}} + \left(1 - \alpha_2\right) \cdot \hat{f}_{2,1} = 0{,}3\frac{467}{437{,}85} + 0{,}7 \cdot 1{,}023 = 1{,}036$$
$$\hat{a}_{2,2} = 0{,}4\left(437{,}85 - 442{,}24\right) + 0{,}6 \cdot \left(-9{,}053\right) = -7{,}187$$

при $t = 7$

$$\hat{x}_7^{\bullet} = (437{,}85 - 7{,}187) \cdot 0{,}982 = 422{,}95$$

$$\hat{a}_{1,7} = 0{,}2\frac{431}{0{,}982} + 0{,}8(437{,}85 - 7{,}187) = 432{,}30$$

$$\hat{f}_{3,2} = 0{,}3\frac{431}{432{,}30} + 0{,}7 \cdot 0{,}982 = 0{,}987$$

$$\hat{a}_{2,7} = 0{,}4(432{,}30437{,}85) + 0{,}6 \cdot (-7{,}187) = -6{,}531$$

при $t = 8$

$$\hat{x}_8^{\bullet} = (432{,}30 - 6{,}531) \cdot 0{,}934 = 397{,}88$$

$$\hat{a}_{1,8} = 0{,}2\frac{412}{0{,}934} + 0{,}8(432{,}30 - 6{,}531) = 428{,}79$$

$$\hat{f}_{4,2} = 0{,}3\frac{412}{428{,}79} + 0{,}7 \cdot 0{,}934 = 0{,}942$$

$$\hat{a}_{2,8} = 0{,}4(428{,}79 - 432{,}30) + 0{,}6 \cdot (-6{,}531) = -5{,}323$$

при $t = 9$ (прогноз)

$$\hat{x}_9^{\bullet} = (\hat{a}_{1,8} + \hat{a}_{2,8}) \cdot \hat{f}_{1,2} = (428{,}79 - 5{,}323) \cdot 1{,}058 = 448{,}16$$

Расчетные значения и прогноз $\hat{x}_t^{\bullet}$, полученный по временному ряду x_t, представлены в таблице 2.1 и на рисунке 2.13. Из представленного графика можно сделать вывод, что модель экспоненциального сглаживания с мультипликативной сезонностью Уинтерса более предпочтительна, нежели регрессионная модель.

Результаты прогноза можно улучшить, подобрав оптимальные значения α.

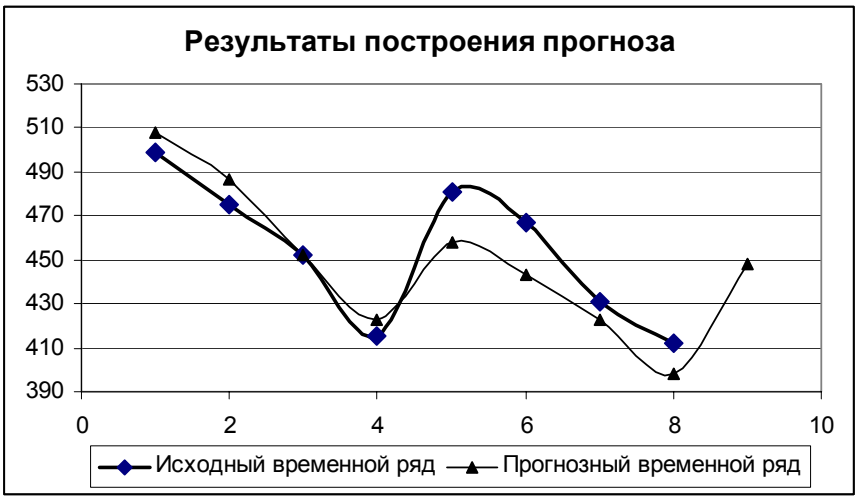


Рисунок 2.13. Результаты прогноза

Задания для самостоятельного выполнения.

1. *Используя данные таблиц Приложения Б, построить прогноз* с использованием модели Уинтерса (экспоненциального сглаживания с мультипликативной сезонностью и линейным ростом).
2. *Подобрать оптимальные параметры прогнозной модели.*
3. *Рассчитать ошибку прогнозирования, дополнительно руководствуясь теоретическими положениями, приведенными в Приложении А.*

2.4. Прогнозирование объема производства по модели Тейла-Вейджа

В таблице 2.2 представлены данные (x_t) об объеме производства по кварталам за 2003 и 2005 гг. (млн. куб. м).

Таблица 2.2. Исходные и расчетные значения объема производства по кварталам

t	год	Цикл k_t	Квартал (фаза цикла) v_t	x_t	$\hat{x}_t$	$\Delta t = x_t - \hat{x}_t$	$\hat{x}_t^*$	ошибка	$(x_t - \hat{x}_t^*)$
1	2003	1	1	7,2	6,82	0,38	6,93	3,7%	0,27
2			2	6,5	6,63	-0,13	6,52	0,2%	-0,02
3			3	6,1	6,44	-0,34	6,47	6,1%	-0,33
4			4	6,3	6,25	0,05	6,28	0,3%	0,07
5	2004	2	1	5,9	6,05	-0,15	6,26	6,2%	-0,31
6			2	5,7	5,86	-0,16	5,66	0,7%	0,08
7			3	6	5,67	0,33	5,47	8,8%	0,56
8			4	5,5	5,48	0,02	5,52	0,4%	0,02
9	2005	3	1	Прогноз			5,37	Дисперсия: 0,07	
								σ	0,27

Требуется по модели экспоненциального сглаживания с аддитивной сезонностью и линейным ростом (модель Тейла-Вейджа) определить расчетные значения $\hat{x}_t^*$ и прогноз при $t = 9$, приняв период упреждения $\tau = 1$, а параметры адаптации равными $\alpha_1 = 0{,}1$; $\alpha_2 = 0{,}4$; $\alpha_3 = 0{,}3$.

Модель экспоненциального сглаживания с аддитивной сезонностью и линейным ростом (модель Г. Теша и С. Вейджа).

Аддитивная модель, имеющая самостоятельную ценность в экономических исследованиях, интересна еще и тем, что позволяет строить модель с мультипликативной сезонностью и экспоненциальной тенденцией. Для этого необходима замена значений первоначального временного ряда их логарифмами, что преобразует экспоненциальную тенденцию в линейную и одновременно мультипликативную сезонную модель в аддитивную.

Пусть наблюдение x_t относится к v_t-й фазе k_t-го цикла, где $v_t = t - l(k_t - l)$, l - число фаз в цикле (для квартального временного ряда $l = 4$, а для месячного $l = 12$).

Модель с аддитивной сезонностью и линейным ростом можно представить в виде

$$\begin{aligned} x_t &= a_{1,t} + g_{v_t k_t} + \varepsilon_t \\ a_{1,t} &= a_{1,t-1} + a_{2,t} \end{aligned} \qquad (2.19)$$

где x_i – среднее значение уровня временного ряда в момент времени t после исключения сезонных колебаний;

$a_{2,t}$ – аддитивный коэффициент роста от момента t-1 к моменту t;

$g_{v_t k_t}$ – аддитивный коэффициент сезонности для v_t-й фазы k_t-го цикла;

ε_t – белый шум.

Оценки параметров модели будем искать при коэффициентах сглаживания α_1, α_2, α_3, где ($0< \alpha_1, \alpha_2, \alpha_3<1$) по следующим процедурам адаптации

$$\begin{aligned} \hat{a}_{1,t} &= \alpha_1\left(x_t - \hat{g}_{v_t,k_{t-1}}\right) + \left(1-\alpha_1\right)\left(\hat{a}_{1,t-1} + \hat{a}_{2,t-1}\right); \\ \hat{g}_{v_t k_t} &= \alpha_2\left(x_t - \hat{a}_{1,t}\right) + \left(1-\alpha_2\right)\hat{g}_{v_t k_{t-1}}; \\ \hat{a}_{2,t} &= \alpha_3\left(\hat{a}_{1,t} - \hat{a}_{1,t-1}\right) + \left(1-\alpha_3\right)\hat{a}_{2,t-1}; \\ \hat{x}_t^* &= \hat{a}_{1,t-\tau} + \tau \cdot a_{2,t-\tau} + \hat{g}_{v_t,k_{t-1}}. \end{aligned} \qquad (2.20)$$

Начальные условия экспоненциального сглаживания определяют по исходному временному ряду x_t ($t = 1,2,\ldots,n$).

Первоначально по временному ряду x_t, содержащему $n = 8$ наблюдений, находим МНК - оценку линейного уравнения регрессии

$$\hat{x}_t = \hat{\theta} + \hat{\theta}_1 t = 7{,}0071 - 0{,}1905t$$

Откуда $\hat{a}_{1,0} = \hat{\theta}_0 = 7{,}0071$; $\hat{a}_{2,0} = \hat{\theta}_1 = -0{,}1905$.

В таблице 2.2 представлены расчетные значения $\hat{x}_t$ и отклонения $\Delta_t = x_t - \hat{x}_t$. Тогда начальные значения аддитивных коэффициентов сезонности равны

$$\hat{g}_{1,0} = \frac{0{,}38 - 0{,}15}{2} = 0{,}1144;$$
$$\hat{g}_{2,0} = \frac{-0{,}13 - 0{,}16}{2} = -0{,}1451;$$
$$\hat{g}_{3,0} = \frac{-0{,}34 + 0{,}33}{2} = -0{,}0046;$$
$$\hat{g}_{4,0} = \frac{0{,}05 + 0{,}02}{2} = 0{,}0359.$$

Расчеты проведем при параметрах адаптации $\alpha_1 = 0{,}1;\quad \alpha_2 = 0{,}4;\quad \alpha_3 = 0{,}3$ и периоде упреждения $\tau = 1$.

Расчет модельных значений на 2003 г. (Первый цикл: $v_t = t$; $k_t = 1$; $\tau = 1$) Исходные данные для расчета

$$\hat{g}_{1,0} = 0{,}1144;$$
$$\hat{g}_{2,0} = -0{,}1451;$$
$$\hat{g}_{3,0} = -0{,}0046;$$
$$\hat{g}_{4,0} = 0{,}0359.$$

Согласно (20) при $t = 1$ имеем

$$\hat{x}_1^* = \hat{a}_{1,0} + \hat{a}_{2,0} + \hat{g}_{1,0} = 7{,}0071 - 0{,}1905 + 0{,}1144 = 6{,}93$$
$$\hat{a}_{1,1} = 0{,}1 \cdot (7{,}2 - 0{,}1144) + (1 - 0{,}1) \cdot (7{,}0071 - 0{,}1905) = 6{,}844$$
$$\hat{g}_{1,1} = 0{,}4 \cdot (7{,}2 - 6{,}844) + 0{,}6 \cdot 0{,}1144 = 0{,}211$$
$$a_{2,1} = 0{,}3 \cdot (6{,}844 - 7{,}0071) + 0{,}7 \cdot (-0{,}1905) = -0{,}182$$

при $t = 2$

$$\hat{x}_2^* = 6{,}844 - 0{,}182 - 0{,}1451 = 6{,}52$$
$$\hat{a}_{1,2} = 0{,}1 \cdot (6{,}5 + 0{,}1451) + 0{,}9 \cdot (6{,}844 - 0{,}182) = 6{,}6595$$
$$\hat{g}_{2,1} = 0{,}4 \cdot (6{,}5 - 6{,}6595) + 0{,}6 \cdot (-0{,}1451) = -0{,}1508$$
$$\hat{a}_{2,2} = 0{,}3 \cdot (6{,}6595 - 6{,}844) + 0{,}7 \cdot (-0{,}182) = -0{,}183$$

при $t = 3$
$$\hat{x}_3^* = 6{,}6595 - 0{,}183 - 0{,}0046 = 6{,}472$$
$$\hat{a}_{1,3} = 0{,}1 \cdot (6{,}1 + 0{,}0046) + 0{,}9 \cdot (6{,}6595 - 0{,}183) = 6{,}4394$$
$$\hat{g}_{3,1} = 0{,}4 \cdot (6{,}1 - 6{,}4394) + 0{,}6 \cdot (-0{,}0046) = -0{,}1385$$
$$\hat{a}_{2,3} = 0{,}3 \cdot (6{,}4394 - 6{,}6595) + 0{,}7 \cdot (-0{,}183) = -0{,}194$$

при $t = 4$
$$\hat{x}_4^* = 6{,}4394 - 0{,}194 + 0{,}0359 = 6{,}281$$
$$\hat{a}_{1,4} = 0{,}1 \cdot (6{,}3 - 0{,}0359) + 0{,}9 \cdot (6{,}4394 - 0{,}194) = 6{,}2472$$
$$\hat{g}_{4,1} = 0{,}4 \cdot (6{,}3 - 6{,}2472) + 0{,}6 \cdot 0{,}0359 = 0{,}0427$$
$$\hat{a}_{2,4} = 0{,}3 \cdot (6{,}2472 - 6{,}4394) + 0{,}7 \cdot (-0{,}194) = -0{,}194$$

Расчет значений за 2004 г. (Второй цикл: $v_t = t$-4; $k_t = 2$).
Исходные данные для расчета

$$\hat{g}_{1,1} = 0{,}211; \quad \hat{g}_{2,1} = -0{,}1508; \quad \hat{g}_{3,1} = -0{,}1385; \quad \hat{g}_{4,1} = 0{,}0427$$

При расчете $\hat{x}_5^*$ учитывается, что 5-я точка относится ко 2-му циклу (k_5 = 2), поэтому $v_5 = t - l$ (k_5 - 1) = 5-4(2-1) = 1 и $\hat{g}_{\upsilon_5 k_{5-1}} = \hat{g}_{1,1} = 0{,}211$. Тогда
$$\hat{x}_5^* = 6{,}2472 - 0{,}194 + 0{,}211 = 6{,}265$$
$$\hat{a}_{1,5} = 0{,}1 \cdot (5{,}9 - 0{,}211) + 0{,}9 \cdot (6{,}2472 - 0{,}194) = 6{,}0172$$
$$\hat{g}_{1,2} = 0{,}4 \cdot (5{,}9 - 6{,}0172) + 0{,}6 \cdot 0{,}211 = 0{,}0799$$
$$\hat{a}_{2,5} = 0{,}3 \cdot (6{,}0172 - 6{,}2472) + 0{,}7 \cdot (-0{,}194) = -0{,}204$$

при $t = 6$

$$\hat{x}_6^* = 6{,}0172 - 0{,}204 - 0{,}1508 = 5{,}662$$
$$\hat{a}_{1,6} = 0{,}1 \cdot (5{,}7 + 0{,}1508) + 0{,}9 \cdot (6{,}0172 - 0{,}204) = 5{,}8165$$
$$\hat{g}_{2,2} = 0{,}4 \cdot (5{,}7 - 5{,}8165) + 0{,}6 \cdot (-0{,}1508) = -0{,}1371$$
$$\hat{a}_{2,6} = 0{,}3 \cdot (5{,}8165 - 6{,}0172) + 0{,}7 \cdot (-0{,}204) = -0{,}203$$

при $t = 7$

$$\hat{x}_7^* = 5{,}8165 - 0{,}203 - 0{,}1385 = 5{,}475$$
$$\hat{a}_{1,7} = 0{,}1 \cdot (6 + 0{,}1385) + 0{,}9 \cdot (5{,}8165 - 0{,}203) = 5{,}6658$$
$$\hat{g}_{3,2} = 0{,}4 \cdot (6 - 5{,}6658) + 0{,}6 \cdot (-0{,}1385) = 0{,}0506$$
$$\hat{a}_{2,7} = 0{,}3 \cdot (5{,}6658 - 5{,}8165) + 0{,}7 \cdot (-0{,}203) = -0{,}188$$

при $t = 8$

$$\hat{x}_8^* = 5{,}6658 - 0{,}188 + 0{,}0427 = 5{,}521$$
$$\hat{a}_{1,8} = 0{,}1 \cdot (5{,}5 + 0{,}0427) + 0{,}9 \cdot (5{,}6658 - 0{,}188) = 5{,}4761$$
$$\hat{g}_{4,2} = 0{,}4 \cdot (5{,}5 - 5{,}4761) + 0{,}6 \cdot 0{,}0427 = 0{,}0352$$
$$\hat{a}_{2,8} = 0{,}3 \cdot (5{,}4761 - 5{,}6658) + 0{,}7 \cdot (-0{,}188) = -0{,}188$$

При расчете прогнозного значения $\hat{x}_9^*$ учитывалось, что момент $t = 9$ принадлежит 3-му циклу, поэтому k_9-1 = 2 и v_9 = 9-4·2 = 1, а $\hat{g}_{v_9,k_{9-1}} = \hat{g}_{1,2}$=0,0799. Тогда

$$\hat{x}_9^* = \hat{a}_{1,8} + a_{2,8} + g_{1,2} = 5{,}4761 - 0{,}188 + 0{,}0799 = 5{,}368$$

В качестве оценок $\hat{g}_{v_t,0}$ принимают средние значения отклонений $\Delta_t = x_t - \hat{x}_t$, соответствующих v_t-й фазе исходного временного ряда, где $v_t = 1, 2, \ldots, l$.

Рассчитанные по модели Тейла-Вейджа значения временного ряда $\hat{x}_t^*$ представлены в таблице 2.2 и на рисунке 2.14, где они сопоставляются с исходным временным рядом x_t.

Рисунок 2.14. Результаты прогноза

Прогноз можно сделать более точным при определении оптимальных параметров сглаживания.

Задания для самостоятельного выполнения.

1. *Используя данные таблиц Приложения Б, построить прогноз с использованием модели Тейла-Вейджа.*
2. *Подобрать оптимальные параметры прогнозной модели.*
3. *Рассчитать ошибку прогнозирования, дополнительно руководствуясь теоретическими положениями, приведенными в Приложении А.*

ГЛАВА 3. ПРАКТИЧЕСКАЯ РЕАЛИЗАЦИЯ МНОГОФАКТОРНЫХ МОДЕЛЕЙ ПРОГНОЗИРОВАНИЯ

Перед подробным рассмотрением собственно методики и практики применения ее в прогнозирования следует отметить, что в большинстве учебных пособий рассматриваются только линейные модели. Объясняется это, прежде всего, тем, что их достаточно просто исследовать, используя ручной счет или табличные процессоры. На основе простых моделей строится дальнейший прогноз. Однако, вручную рассчитать и провести исследование нелинейных многофакторных моделей, а тем более сделать на их основе прогноз почти невозможно, особенно если в модели учитываются более трех факторов.

Для решения данной проблемы использован статистический пакет Statistica. Данному программному продукту посвящен ряд работ и учебных пособий известных авторов, однако в них почти не рассмотрены практические вопросы построения нелинейных многофакторных моделей.

Построение прогнозов на основе многофакторных моделей рассмотрим на примере прогнозирования убытков ОАО «Невинномысский хлебокомбинат». Убытки выбраны потому, что перед нами стоит задача не только определения их возможной величины, т.е. определение величины риска, но и выявление факторов производства, которые этому способствуют.

Задание:

В последние три года на некотором предприятии является убыточным одно из производств – производство халвы, которое снижает общую прибыль предприятия.

Поэтому необходимо провести исследование причин и выявить факторы, оказывающие наибольшее влияние на снижение прибыли.

Данные за последние 30 месяцев берутся из бухгалтерского баланса предприятия. Соответственно, оптимальное число факторных признаков равно пяти, несущественные факторные признаки будут исключены обратным методом пошаговой регрессии.

Введем следующие переменные:

X1 – процент реализации халвы (за месяц);
X2 – стоимость 1 тонны сырья, в частности, семечек (в рублях);
X3 – затраты на один рубль произведенной халвы (в рублях);
X4 – расход сырья (семечек) на одну тонну халвы (в долях);
X5 – стоимость электроэнергии (в рублях);
У – соответственно убытки предприятия от данного производства.

Таблица 3.1. Исходные данные

7	Y	X1	X2	X3	X4	X5
1	50000	90,3	2600	0,77	0.689	0,58
2	58000	86.4	3000	0,873	0,699	0,58
3	65000	73,2	3900	0,898	0,71	0,58
4	66000	72,1	3900	0,89	0,7	0,58
5	67000	71,3	3900	0,89	0,71	0,59
6	68000	70,9	3900	0,89	0,72	0,6
7	69000	68,7	4000	0,9	0,73	0,6
8	70000	67	4100	0,9	0,74	0,6
9	71000	66	4150	0,9	0,75	0,6
10	72000	65	4160	0,9	0,76	0,6
11	73000	64	4170	0,89	0,77	0,61
12	75000	50,4	4180	0,89	0,78	0,62
13	80000	51,2	4200	0,9	0,79	0,77
14	79000	53,8	4300	0,89	0,8	0,78
15	83000	50,3	4400	0,88	0,81	0,79
16	85000	50	4500	0,9	0,82	0,8
17	79000	55,6	4600	0,88	0,83	0,8
18	83000	53,1	4700	0,88	0,84	0,81
19	86000	50,1	4800	0,89	0,85	0,81
20	89000	49,9	4900	1	0,86	0,83
21	90000	50,8	4900	1	0,87	0,84
22	91000	50,6	4900	1	0,88	0,86
23	90000	51,4	4900	1	0,89	0,9
24	91000	52,3	5000	1	0,9	1
25	93000	50	5000	1	0,9	1,1
26	100000	50	5100	1	0,9	1,1
27	100000	51,3	5200	1	0,9	1,1
28	100000	50	5300	1	0,9	1,1
29	110000	50	5400	1	0,9	1,1
30	110000	50	5500	1	0,9	1,1

3.1. Линейные многофакторные модели

В пакете при запуске приложения появляется *Переключатель модулей*, представленный на рисунке 3.1.

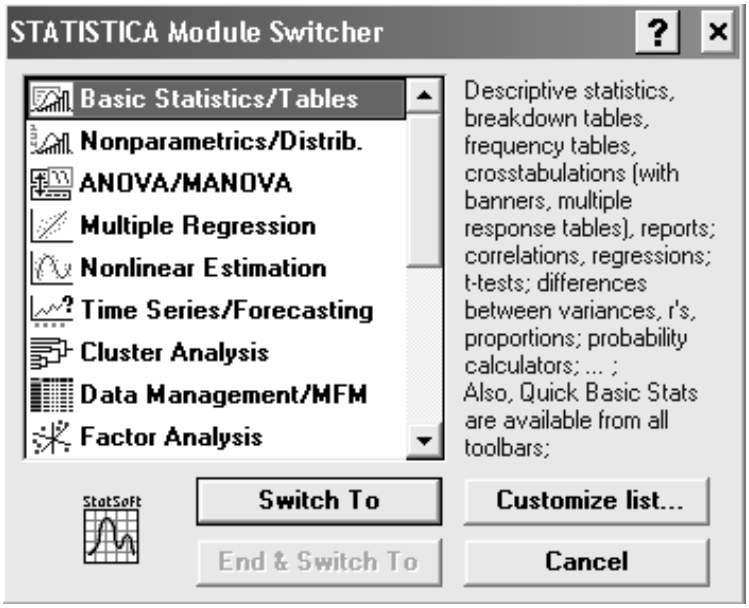


Рисунок 3.1. Переключатель модулей. Основное меню системы STATISTICA

Для решения примера необходим модуль ***Multiple Regression*** множественной регрессии и выполнить команду ***Switch To***.

Замечание: решение данной задачи состоит из двух этапов:
- *рассмотрение линейных уравнений регрессии;*
- *рассмотрение нелинейных уравнений регрессии;*

После нажатия кнопки Switch to появится чистая таблица. Для заполнения ее имеющимися данными, увеличим число строк с 10 до 30. Для этого необходимо на панели инструментов выбрать **Cases** → *Add*:

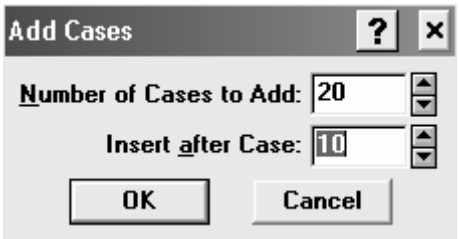


Рисунок 3.2. Добавление данных

Возможно также уменьшение числа столбцов (нам необходимо 6) командой Vars→ *Delete*. После того как таблица заполнена, выполняем следующую последовательность действий.

Меню *Analysis* → *Resume Analysis*→ *появится диалоговое окно Multiple Regression* (рисунок 3.3).

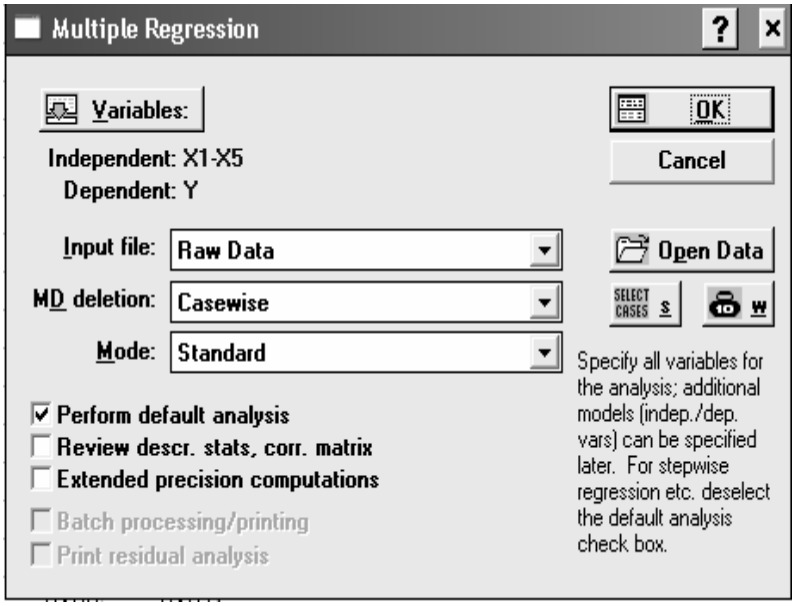


Рисунок 3.3. Multiple Regression (Множественная регрессия)

Variables – переменные.
Independent:X1-X5 – независимые переменные.
Dependent :Y – зависимая переменная.
Input file: входной файл.

Выбираем стандартный метод оценивания *Mode- Standard*.

В левом верхнем углу расположена кнопка *Variables* (переменные), она необходима для того, чтобы определить переменные, которые будут участвовать в процессе построения модели. После нажатия на нее появляется следующее диалоговое окно, представленное на рисунке 3.4.

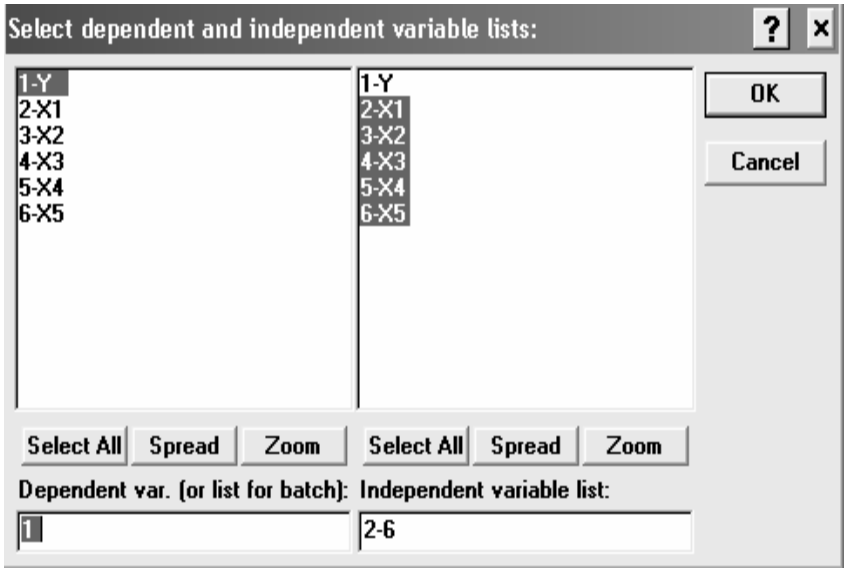


Рисунок 3.4. Выбор зависимых и независимых переменных

Среди зависимых переменных (dependent) выделяете только *Y* (в левом столбце), среди независимых (independent) выделяете все, кроме *Y* (в правом столбце). Нажимайте кнопку ОК.

Программа произведет оценивание параметров, в результате чего появится следующее диалоговое окно *«Результаты множественной регрессии»*, представленное на рисунке 3.5.

Как можно видеть, в программе красным цветом в нижней части окна выделены те факторы, которые являются значимыми, и *синим – незначимые.* Это не значения коэффициентов модели, а лишь бета-коэффициенты. Отмеченные факторы необходимо исключать, они не значимы.

Рассмотрим информационную часть этого окна. В нем содержатся краткие сведения о результатах анализа.

Dep. var. (имя зависимой переменной) – *Y*.

No of Cases (число наблюдений, по которым построена регрессия). В нашей задаче это число равно 30.

Multiple R **–** (коэффициент множественной корреляции).

R^2 – наиболее важный показатель множественный коэффициент детерминации (он показывает долю общего разброса относительно выборочного среднего зависимой переменной).

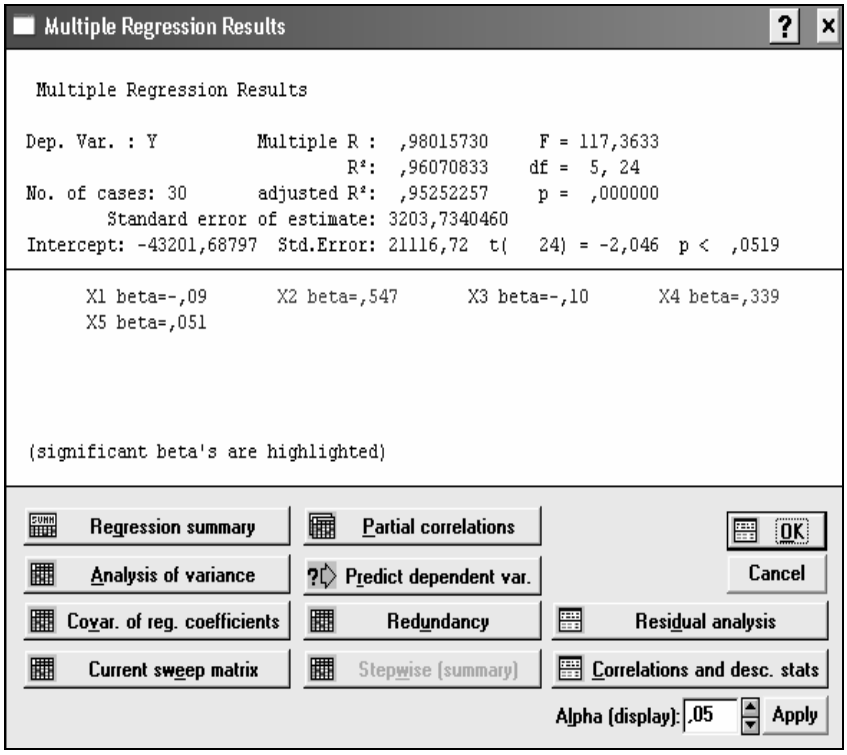


Рисунок 3.5. Результаты множественной регрессии

Важен этот показатель и потому, что по нему можно найти число неучтенных факторов в модели. Для этого вычисляется величина σ так

$$\sigma = (1-R^2)*100\%.$$

Эта величина не должна превышать значения σ = 15%.

Важно отметить, что предварительно в *MS Excel* необходимо было просчитать парные коэффициенты множественной корреляции. Но это возможно осуществить и в данной программе. Для этого необходимо выполнить следующую последовательность действий.

В диалоговом окне (см. рисунок 3.5) в правом нижнем углу необходимо нажать на кнопку *Correlations and desc.stats → Correlations*. После этого появится диалоговое окно «Корреляции», представленное на рисунке 3.6.

Correlations (ybitki.sta)

Continue...

	X1	X2	X3	X4	X5	Y
X1	1,00	-,23	,16	-,16	-,24	-,30
X2	-,23	1,00	-,75	,93	,08	,96
X3	,16	-,75	1,00	-,82	,02	-,80
X4	-,16	,93	-,82	1,00	,03	,95
X5	-,24	,08	,02	,03	1,00	,12
Y	-,30	,96	-,80	,95	,12	1,00

Рисунок 3.6. Матрица парных коэффициентов корреляции

Анализируя матрицу парных коэффициентов, можно сделать вывод, что скорее всего в модель будут входить элементы: x_2, x_3, x_4. Это связано с сильной зависимостью между этими показателями и y (соответственно 0,96; -0,8; 0,95), то есть значение парного коэффициента корреляции между переменной и y должно быть максимально приближено к 1.

Как видно (см. рисунок 3.7), фактор x_3 будет входить в модель со знаком минус, то есть, соответственно, уменьшать регрессионную модель убытков.

Иная картина складывается с парным коэффициентом корреляции между переменными. Как раз этот показатель должен быть как можно меньше. Если же коэффициент парной корреляции между признаками более 0,7, то это говорит о наличии мультиколлинеарности, которая искажает модель в целом, и от нее необходимо избавляться путем последовательного исключения факторов, между которыми она определена.

Для того, чтобы получить непосредственно коэффициенты уравнения регрессии, необходимо в диалоговом окне, представленном на рисунке 3.5, нажать кнопку *Regression Summary,* и, соответственно, появится диалоговое окно, представленное на рисунке 3.8.

Regression Summary for Dependent Variable: Y

Continue...

R= ,98015730 RI= ,96070833 Adjusted RI= ,95252257
F(5,24)=117,36 p<,00000 Std.Error of estimate: 3203,7

N=30	BETA	St. Err. of BETA	B	St. Err. of B	t(24)	p-level
Intercpt			-43201,7	21116,72	-2,04585	,051879
X1	-,087291	,043217	-48,6	24,08	-2,01983	,054700
X2	,547315	,117052	12,1	2,59	4,67584	,000095
X3	-,099849	,071159	-29,5	21,05	-1,40318	,173369
X4	,339036	,133063	649,0	254,73	2,54793	,017658
X5	,050598	,042014	37,2	30,86	1,20432	,240210

Рисунок 3.8. Суммарная регрессия

В верхней части диалогового окна на рисунке 3.8 представлена практически та же информация, что и в верхней части окна на рисунке 3.5. RI – коэффициент множественной детерминации. Добавляется F(5,24) = 117,36 – коэффициент Фишера-Снедекора (где 5 и 24 – степени свободы $v_1 = 5$ и $v_2 = 24$). Что касается **нижней части** диалогового окна (таблица), то:

- в *первом столбце* перечислены переменные;
- *во втором (BETA)* – бета-коэффициенты;
- *в третьем (St.Err of BETA)* – ошибка для бета-коэффициентов;
- *в четвертом (B)* – коэффициенты, которые стоят перед факторными признаками уравнения регрессии;
- *в пятом (St. Err of B*) – стандартная ошибка для коэффициентов при факторных признаках уравнения регрессии;
- *в шестом (t(24))* – коэффициент Стьюдента, в соответствии с которым осуществляется пошаговое исключение незначимых факторов.

Важно также отметить, что верхняя строка таблицы рассчитана для свободного члена, который входит в уравнение регрессии в данном случае со знаком «минус» и равен -43201.

Исходя из содержимого диалогового окна на рисунке 3.8, получим следующую модель

$$\mathbf{y} = -43201 - 48{,}6\mathbf{x}_1 + 12{,}1\mathbf{x}_2 - 29{,}5\mathbf{x}_3 + 649\mathbf{x}_4 + 37{,}2\mathbf{x}_5 \qquad (3.1)$$

В столбце «*B*» представлены коэффициенты регрессии при соответствующих признаках (факторах).

Коэффициент B_1 (при независимой переменной x_1) =-48,6
Коэффициент B_2 (при независимой переменной x_2) =12,1
Коэффициент B_3 (при независимой переменной x_3) =-29,5
Коэффициент B_4 (при независимой переменной x_4) =649
Коэффициент B_5 (при независимой переменной x_5) =37,2

Если экономически интерпретировать представленную модель, то можно сказать, что на уменьшение убытков влияют такие факторы, как процент реализации халвы, то есть чем больше реализуется данный продукт, тем меньше убытки, и затраты на один рубль произведенной продукции.

Если рассматривать экономический смысл последнего фактора, то он должен входить в уравнение регрессии со знаком «плюс», так как увеличение затрат приводит к увеличению убытков, и, соответственно, уменьшению прибыли. Но в нашем примере этот фактор входит в уравнение регрессии со знаком «минус», что может свидетельствовать лишь о том, что в этот период работы предприятия необходимо увеличить затраты на производство, рекламу, повысить качество продукции, персонала, внедрить новые технологии.

На величину результата влияют и такие показатели, как стоимость тонны сырья (чем она выше, тем больше величина убытков) и расход сырья (чем он выше, тем больше убытков) и стоимость электроэнергии (чем она быстрее растет и чем она выше, тем, соответственно, больше убытков).

Если рассмотреть коэффициент множественной детерминации, то важно отметить, что R^2 = 0,96 (чем он ближе к 1, тем лучше и сильнее связь). Этот показатель является практически самым высоким.

На основе данного показателя определяется число неучтенных факторов, в данной модели эта величина составляет 4%.

Важным элементом анализа является оценка адекватности модели. Для этого необходимо проанализировать критерий ***Фишера-Снедекора (F)***, который также представлен в диалоговом окне на рисунке 3.8. В нашем примере $F(5,24) = 117,36$. Расчетное значение F_P необходимо сравнить с табличным, которое при данных степенях свободы 5 v_1=5 и v_2 = 24 будет равно F_T = 2,62.

Условие адекватности модели $\boldsymbol{F_P > F_T}$.

В нашем примере данное условие выполняется (117,36 > 2,62). При определении адекватности модели возможно выполнение одного из трех условий.

1. Если построенная модель на основе ее проверки по F-критерию Фишера в целом адекватна и все коэффициенты регрессии значимы, то такая модель может быть использована для принятия решений и прогноза..

2. Модель по F-критерию Фишера адекватна, но часть коэффициентов регрессии незначима. В этом случае модель пригодна для принятия некоторых решений, но не для производства прогнозов.

3. Модель по F-критерию Фишера адекватна, но все коэффициенты регрессии незначимы. В этом на ее основе решения прогнозы не принимаются и не осуществляются.

Представленная первая модель, являясь адекватной, включает в себя ряд незначимых факторов, поэтому на ее основе нельзя осуществлять прогнозы, но можно принимать некоторые решения.

Для того чтобы на основе данной модели можно было осуществлять прогнозы, необходимо исключить незначимые факторы.

Как определить, какой фактор наименее значим?

Для этого необходимо сравнить факторы по критерию Стьюдента, представленному на рисунке 3.8, в столбце *б* (t(24)) и проанализировать значения этого критерия. Из всех значений выбирается наименьшее (в данном случае это 1,2 для x_5).

Таким образом, наименее значимым является фактор x_5, то есть стоимость электроэнергии, который исключается из числа независимых факторов.

Необходимо учитывать, что коэффициент Стьюдента для различных факторов при сравнении берется **по модулю**.

После этого возвращаемся к диалоговому окну (см. рисунок 3.9) «Выбор зависимых и независимых переменных» и выделяем в правом столбце x_1,x_2,x_3,x_4 (в левом так и остается зависимая переменная у).

Как можно видеть из рисунка 3.10, из незначимых факторов остался только x_3, то есть затраты на один рубль произведенной продукции, фактор x_1 стал значимым.

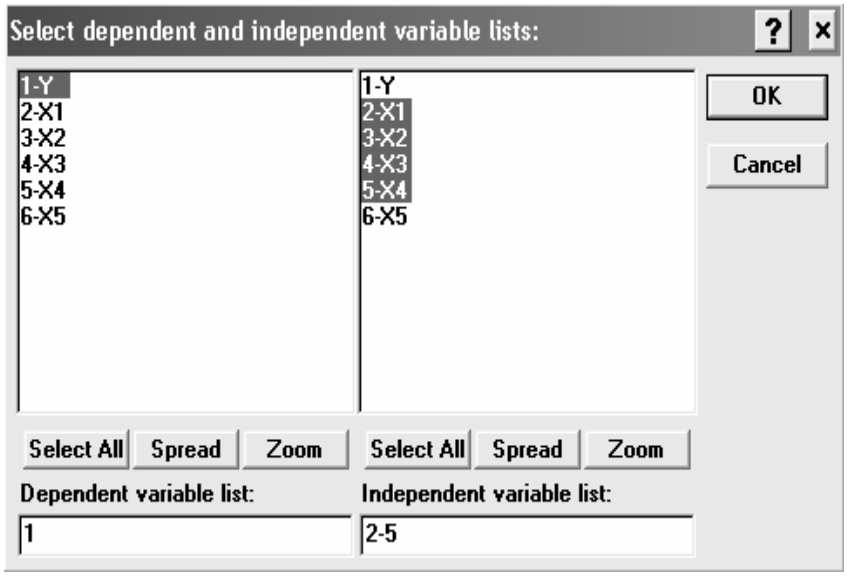


Рисунок 3.9. Выбор зависимых и независимых переменных

Нажимаем ОК. и получим следующее диалоговое окно «Результат множественной регрессии»

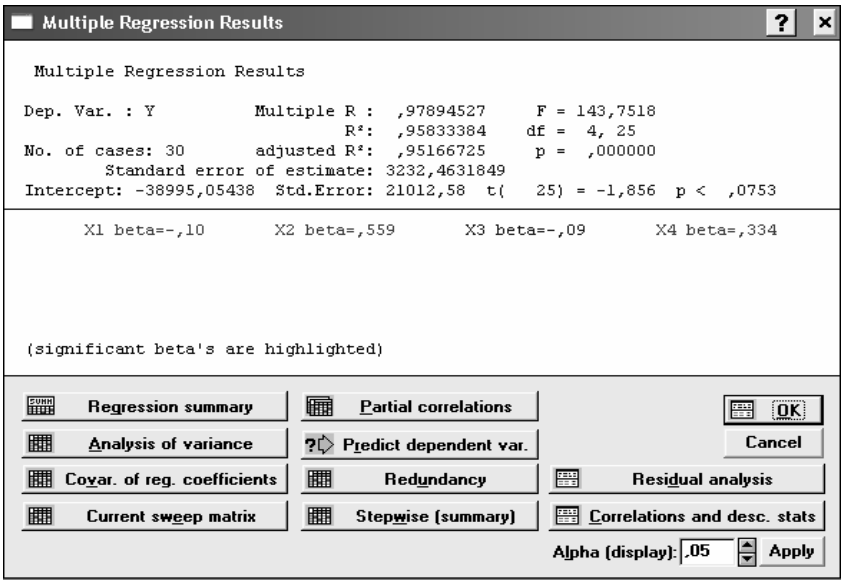


Рисунок 3.10. Результат множественной регрессии

Нажимая кнопку *Regression Summary*,чтобы определить коэффициенты при факторных признаках и ряд важных оценочных показателей, получим результаты, представленные на рисунке 3.11.

Regression Summary for Dependent Variable: Y

Continue...

R= ,97894527 RI= ,95833384 Adjusted RI= ,95166725
F(4,25)=143,75 p<,00000 Std.Error of estimate: 3232,5

N=30	BETA	St. Err. of BETA	B	St. Err. of B	t(25)	p-level
Intercpt			-38995,1	21012,58	-1,85580	,075309
X1	-,098560	,042570	-54,9	23,72	-2,31526	,029094
X2	,559417	,117665	12,4	2,60	4,75430	,000070
X3	-,092172	,071508	-27,3	21,15	-1,28896	,209211
X4	,333816	,134185	639,0	256,88	2,48772	,019883

Рисунок 3.11. Суммарная регрессия

Уравнение регрессии выглядит следующим образом (по аналогии с предыдущей моделью)

$$\mathbf{y} = -38995 - 54{,}9\mathbf{x}_1 + 12{,}4\mathbf{x}_2 - 27{,}3\mathbf{x}_3 + 639\mathbf{x}_4 . \qquad (3.2)$$

Если анализировать данную модель, то можно сказать, что она незначительно отличается от предыдущей модели, то есть коэффициенты, стоящие перед факторными признаками уравнения регрессии, изменились мало.

Проанализируем остальные показатели. В первую очередь, необходимо сказать, что коэффициент множественной детерминации *RI* (R^2) остался практически неизменным и составляет 0,96. Отсюда и не изменилось число неучтенных факторов, оно по-прежнему составляет 4%. Еще раз подчеркивается незначимость фактора x_5 (то есть стоимости электроэнергии, возможно, из-за того, что ее величина уже включается в величину производимых затрат на 1 рубль произведенного продукта).

Можно опять утверждать, что модель адекватна, так как при степенях свободы $v_1 = 4$ и $v_2 = 25$ имеем $F_P > F_T$ (143,75 > 2,76). Но она включает в себя незначимый фактор x_3, который необходимо исключить.

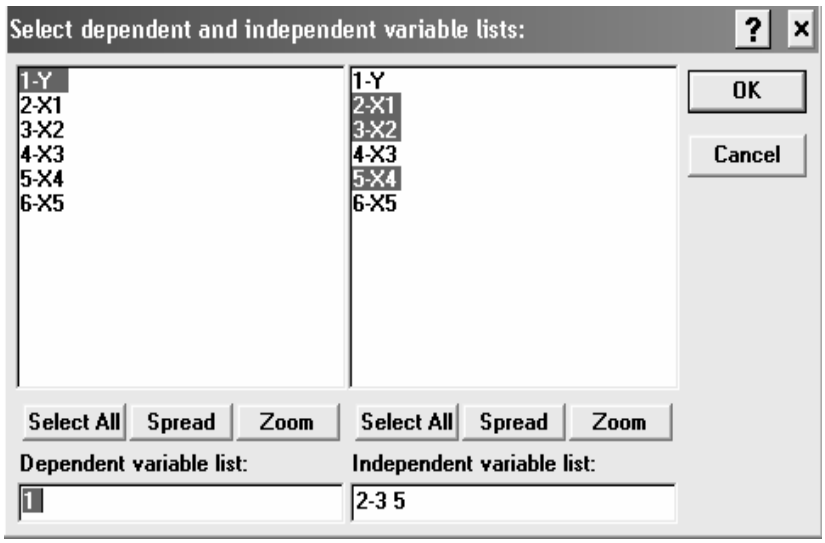


Рисунок 3.12. Суммарная регрессия

Для этого возвращаемся опять к диалоговому окну *«Зависимые и независимые переменные»* и по тем же правилам исключаем фактор x_3 – затраты на один рубль произведенной продукции.

Полученные следующие результаты показаны на рисунке 3.13.

Как видно из него, все перечисленные факторы являются значимыми:

- x_1 – процент реализации халвы в месяц;
- x_2 – стоимость 1 тонны сырья (семечек) в рублях;
- x_4 – расход сырья на одну тонну халвы.

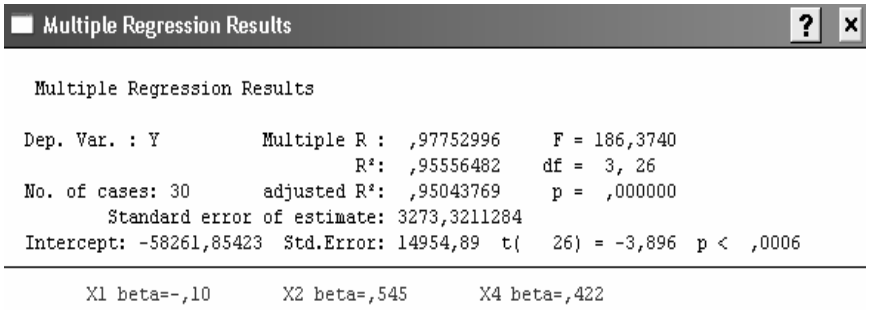


Рисунок 3.13. Результаты множественной регрессии

На основе следующего диалогового окна составим уравнение множественной регрессии.

Regression Summary for Dependent Variable: Y

Continue...

R= ,97752996 RI= ,95556482 Adjusted RI= ,95043769
F(3,26)=186,37 p<,00000 Std.Error of estimate: 3273,3

N=30	BETA	St. Err. of BETA	B	St. Err. of B	t(26)	p-level
Intercpt			-58261,9	14954,89	-3,89584	,000613
X1	-,101745	,043035	-56,7	23,98	-2,36421	,025819
X2	,545389	,118642	12,1	2,62	4,59693	,000097
X4	,421881	,116949	807,6	223,88	3,60739	,001290

Рисунок 3.14. Суммарная регрессия

Уравнение регрессии будет иметь следующий вид

$$\mathbf{y} = -58261{,}9 - 56{,}7\mathbf{x}_1 + 12{,}1\mathbf{x}_2 + 807\mathbf{x}_4 \qquad (3.3)$$

Анализируя данное уравнение, можно сказать, что коэффициенты при первом и втором факторных показателях изменились незначительно, при третьем, последнем факторном признаке коэффициент в уравнении значительно возрос. Поэтому можно сказать, что в первую очередь на уровень убытков влияет последний фактор – расход сырья на одну тонну халвы, причем при увеличении этого показателя увеличивается и уровень убытков. В меньшей степени на результативный признак влияет процент реализации халвы, и еще меньше влияет стоимость 1 тонны сырья.

Анализируя остальные показатели, можно сказать, что множественный коэффициент детерминации *RI* =0,96, что соответствует предыдущим значениям. Даже притом, что исключались незначимые факторы, число неучтенных факторов так и остается 4%. Важно также отметить, что модель является адекватной, так как $F_P > F_T$, (при степенях свободы v_1=3 и v_2 = 26 имеем 143,75 > 2,98). На ее основе можно принимать решения и строить прогнозы.

Таким образом, наиболее оптимальным является линейное уравнение регрессии следующего вида

$$\mathbf{y} = -58261{,}9 - 56{,}7\mathbf{x}_1 + 12{,}1\mathbf{x}_2 + 807\mathbf{x}_4 \qquad (3.4)$$

3.2. Нелинейные многофакторные модели

Вернемся к диалоговому окну *«Множественной регрессии»*. В опции *Mode* – выбираем не стандартный метод – *Fixed non linear* (нелинейные модели). Нажимаем ОК.

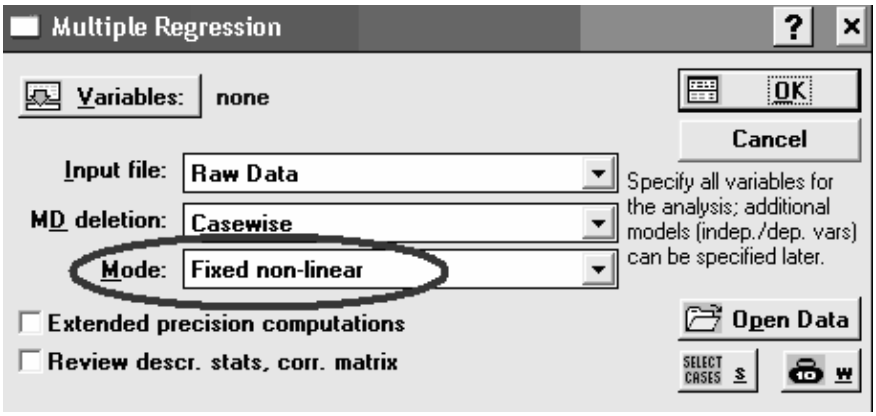


Рисунок 3.15. Множественная регрессия

После того, как нажмем кнопку ОК, появится следующее диалоговое окно, где необходимо выбрать переменные для нелинейного анализа. В построении нелинейных моделей может участвовать максимум четыре фактора (переменные).

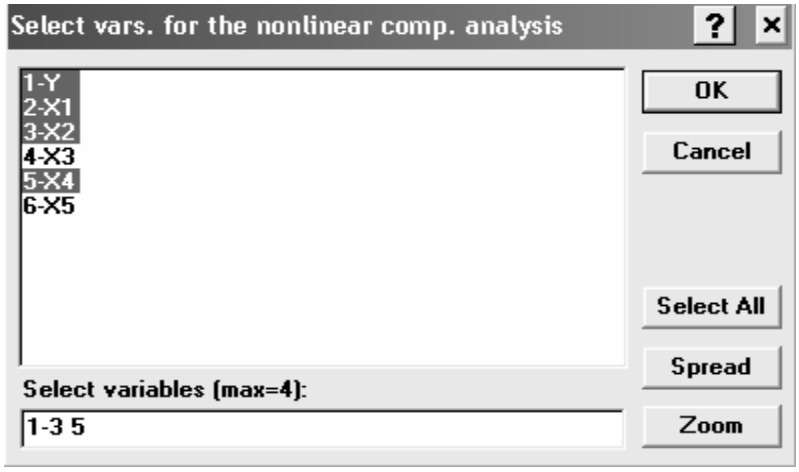


Рисунок 3.16. Выбор переменных для нелинейного анализа

С учетом уже построенных линейных моделей выбираем наиболее значимые факторы x_1, x_2, x_4, и, конечно, y. Нажимаем кнопку ОК. Появится следующее диалоговое окно *«Нелинейные компоненты регрессии»*, где можно выбрать порядок факторов, что определит порядок уравнения регрессии. Вначале построим уравнение регрессии второго порядка, для этого ставим соответствующий флажок.

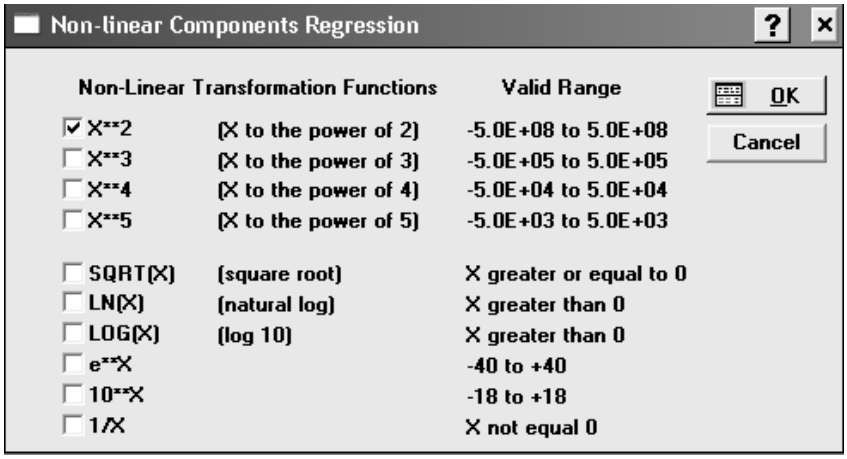


Рисунок 3.17. Нелинейные компоненты регрессии

Нажимаем ОК. После этого появится следующее диалоговое окно *«Определение модели;* метод (*Method*) – Standart и нажимаем кнопку *Variables*.

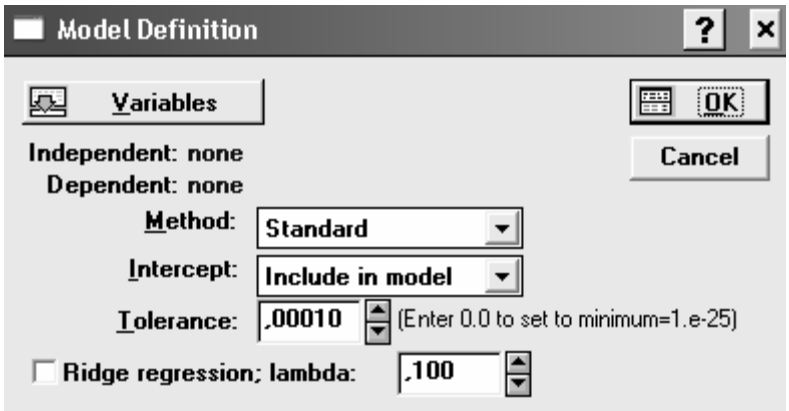


Рисунок 3.18. Определение модели

Нажимаем кнопку *Variables*. После чего появится следующее диалоговое окно «*Зависимые и независимые переменные*».

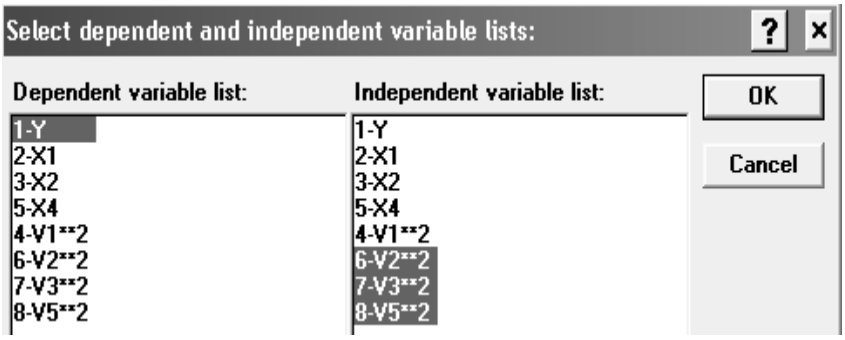


Рисунок 3.19. Зависимые и независимые переменные

В диалоговом окне, представленном на рисунке 3.19, в левом столбце (*Dependent*) выбираем также зависимую переменную – *y*, в правом столбце выбираем независимые переменные *V2**2,V3**2,V5**2*, что означает, что в нелинейное уравнение регрессии будут входить x_1^2, x_2^2, x_4^2. То есть,

- *V2**2* - это x_1^2;
- *V3**2* - это x_2^2;
- *V5**2* - это x_4^2.

Нажимаем ОК, и программа произведет оценивание параметров модели, по завершении которого появится следующее диалоговое окно (см. рисунок 3.20).

Очевидно, что в представленной модели выделяются один незначимый фактор и два значимых фактора.

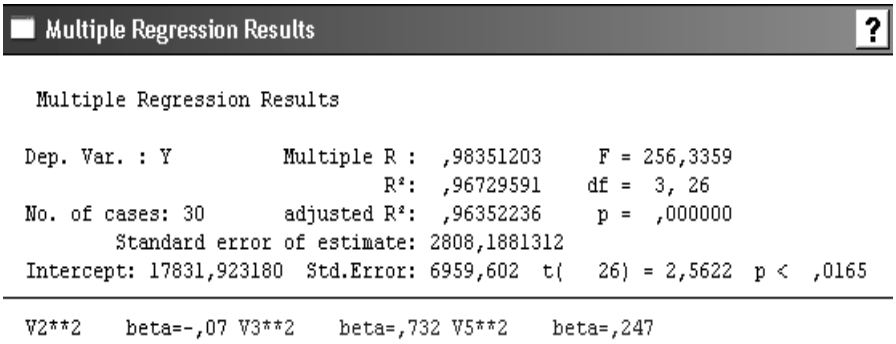


Рисунок 3.20. Результаты множественной регрессии

Нажимаем кнопку *Regression summary* и получим следующее диалоговое окно *«Суммарная регрессия для зависимой переменной Y»*

Regression Summary for Dependent Variable: Y

Continue...

R= ,98351203 RI= ,96729591 Adjusted RI= ,96352236
F(3,26)=256,34 p<,00000 Std.Error of estimate: 2808,2

N=30	BETA	St. Err. of BETA	B	St. Err. of B	t(26)	p-level
Intercpt			17831,92	6959,602	2,56220	,016543
V2**2	-,069414	,037480	-,23	,126	-1,85204	,075408
V3**2	,732104	,118900	,00	,000	6,15732	,000002
V5**2	,246957	,117335	2,11	1,001	2,10472	,045134

Рисунок 3.21. Суммарная регрессия
для зависимой переменной Y

Очевидно, что данная модель не является оптимальной и наилучшей, но если рассмотреть коэффициент множественной детерминации $RI = 0{,}97$, то можно сказать, что он выше, чем у наилучшей линейной модели. Можно сделать вывод о том, что неучтенных факторов всего 3%. Что касается адекватности модели, то модель является адекватной, так как $F_P > F_T$ (при степенях свободы $v_1 = 3$ и $v_2 = 26$ имеем 256,34 > 2,98), но на ее основе нельзя осуществлять прогнозы, так как она включает в себя незначимый фактор – x_1^2. Уравнение регрессии будет иметь следующий вид

$$\mathbf{y} = 17831{,}92 - 0{,}23\mathbf{x}_1^2 + 0{,}002\mathbf{x}_2^2 + 2{,}11\mathbf{x}_4^2 \qquad (3.5)$$

Если проанализировать его, то можно сказать, что знаки коэффициентов, стоящих перед факторными признаками, экономически верны: чем выше процент реализации халвы, тем ниже убытки, и чем больше стоимость сырья и его расход, тем выше убытки предприятия.

Для того, чтобы оптимизировать и улучшить модель, необходимо исключить незначимый фактор x_1^2.

После его исключения уравнение регрессии примет следующий вид

$$\mathbf{y} = 18480{,}4 + 0{,}002\mathbf{x}_2^2 + 1{,}61\mathbf{x}_4^2 \qquad (3.6)$$

Multiple Regression Results

```
Multiple Regression Results

Dep. Var. : Y          Multiple R :  ,98131617     F = 351,1817
                            R²:  ,96298142    df =  2, 27
No. of cases: 30       adjusted R²:  ,96023930     p =  ,000000
        Standard error of estimate: 2931,8371650
Intercept: 18480,379589  Std.Error: 7256,843  t(   27) = 2,5466  p <  ,0169

V3**2    beta=,801 V5**2    beta=,189
```

Рисунок 3.22. Результаты множественной регрессии

Regression Summary for Dependent Variable: Y

Continue...

R= ,98131617 RI= ,96298142 Adjusted RI= ,96023930
F(2,27)=351,18 p<,00000 Std.Error of estimate: 2931,8

N=30	BETA	St. Err. of BETA	B	St. Err. of B
Intercpt			18480,3795886	7256,84294883
V3**2	,80052963855	,11799000615	,0020917	,00030830
V5**2	,18852536076	,11799000615	1,6090754	1,00705188

Рисунок 3.23 Суммарная регрессия
для зависимой переменной Y

Причем коэффициент множественной детерминации снизился (RI = 0,96), и число неучтенных факторов стало опять 4%. Модель адекватна, так как $F_P > F_T$ (при степенях свободы v_1=2 и v_2 = 27 имеем 351,18 > 3,35), но на ее основе нельзя осуществлять прогнозы, так как она включает в себя незначимый фактор – x_4^2. Исключая его, получим следующую модель (см. рисунок 3.24).

Тогда уравнение регрессии примет следующий вид

$$y = 29612{,}17 + 0{,}003 \cdot x_2^2 \qquad (3.7)$$

Multiple Regression Results

```
Multiple Regression Results

Dep. Var. : Y          Multiple R :  ,97953107     F = 663,0358
                                R²:  ,95948112    df =  1, 28
No. of cases: 30       adjusted R²:  ,95803402     p =  ,000000
        Standard error of estimate: 3012,0456076
Intercept: 29612,167822  Std.Error: 2086,294  t(   28) = 14,194  p <  ,0000

V3**2    beta=,980
```

Рисунок 3.24 Результаты множественной регрессии

Regression Summary for Dependent Variable: Y

Continue...

R= ,97953107 RI= ,95948112 Adjusted RI= ,95803402
F(1,28)=663,04 p<,00000 Std.Error of estimate: 3012,0

N=30	BETA	St. Err. of BETA	B	St. Err. of B	t(28)	p-level
Intercpt			29612,168	2086,2939	14,193671	,0000000
V3**2	,9795311	,0380408	,003	,0001	25,749483	,0000000

Рисунок 3.25 Суммарная регрессия для зависимой переменной Y

Важно отметить, что при увеличении порядка уравнения регрессии значения параметров становятся хуже. Тем не менее, рассмотрим нелинейные **модели третьего порядка,** которые стоятся аналогично (см. рисунки 3.26 и 3.27).

Как видно из рисунка 3.24, в данной модели остается лишь один значимый фактор. А коэффициент множественной детерминации остается прежним – *RI* = 0,96, что говорит о 4% неучтенных факторов. Можно сказать, что она не является оптимальной, так как коэффициент множественной детерминации не является самым высоким. Кроме того, модель включает в себя всего один фактор x_2 (стоимость одной тонны сырья в рублях).

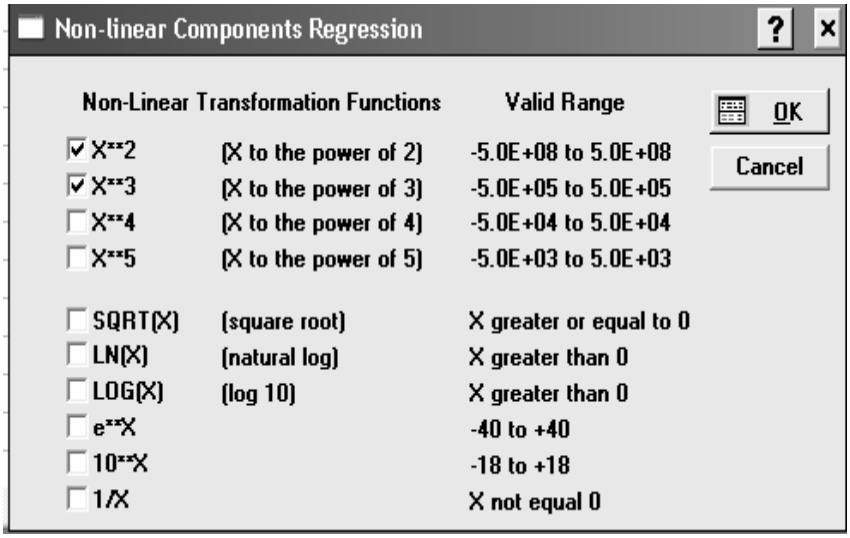


Рисунок 3.26. Нелинейные компоненты регрессии

Нажимаем ОК.

Выбираем переменные только третьего порядка.

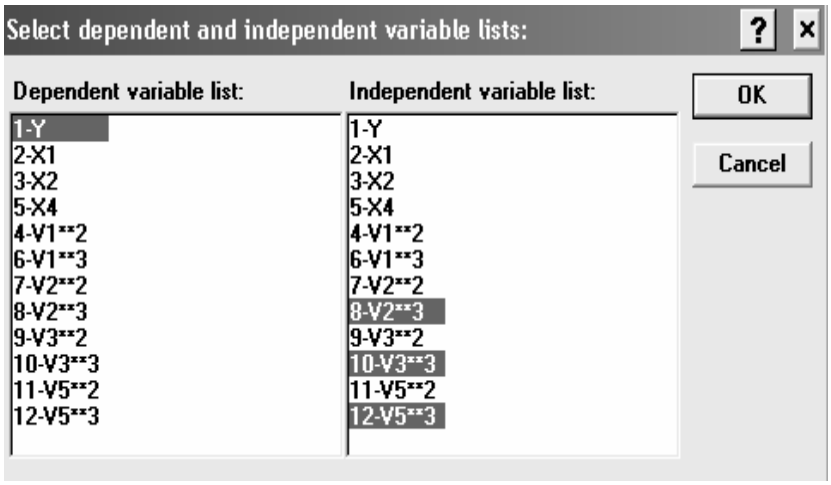


Рисунок 3.27. Выбор зависимых и независимых переменных

Multiple Regression Results

```
Multiple Regression Results

Dep. Var. : Y          Multiple R :  ,98553004     F = 292,9868
                              R²:  ,97126946    df =  3, 26
No. of cases: 30       adjusted R²:  ,96795440     p =  ,000000
        Standard error of estimate: 2632,0676635
Intercept: 39778,207350  Std.Error: 4471,492  t(   26) = 8,8960  p <  ,0000

V2**3    beta=-,04 V3**3    beta=,821 V5**3    beta=,165
```

Рисунок 3.28. Результаты множественной регрессии

Regression Summary for Dependent Variable: Y

Continue...

R= ,98553004 RI= ,97126946 Adjusted RI= ,96795440
F(3,26)=292,99 p<,00000 Std.Error of estimate: 2632,1

N=30	BETA	St. Err. of BETA	B	St. Err. of B	t(26)	p-level
Intercpt			39778,21	4471,492	8,89596	,000000
V2**3	-,044242	,035927	-,00	,001	-1,23143	,229186
V3**3	,820548	,119539	,00	,000	6,86426	,000000
V5**3	,165085	,118439	,01	,006	1,39384	,175166

Рисунок 3.29 Суммарная регрессия для зависимой переменной Y

Как можно видеть из рисунка 3.29, значимым в модели третьего порядка является только один фактор x_2^2, остальные факторы не значимы. Коэффициент множественной детерминации RI = 0,97 (число неучтенных факторов 3%), что выше, чем в предыдущей модели второго порядка.

Regression Summary for Dependent Variable: Y

Continue...

R= ,98409372 RI= ,96844046 Adjusted RI= ,96731333
F(1,28)=859,21 p<,00000 Std.Error of estimate: 2658,3

N=30	BETA	St. Err. of BETA	B	St. Err. of B
Intercpt			44794,8921336	1340,85017325
V3**3	,98409372280	,03357270615	,0000004	,00000001

Рисунок 3.30. Суммарная регрессия
для зависимой переменной Y

Модель является адекватной, так как $F_P > F_T$ (при степенях свободы $v_1 = 3$ и $v_2 = 26$ имеем 292,99 > 2,98), но на ее основе нельзя осуществлять прогнозы, так как она включает в себя незначимые факторы - x_1^3 и x_4^3. Исключая незначимые факторные признаки (процент реализации халвы и расход сырья) получим следующую модель

$$y = 44794{,}89 + 0{,}0000004 \cdot x_2^3 \qquad (3.8)$$

Если анализировать уравнение (3.8), то можно сказать, что по показателям, показанным на рисунке 3.30, оно лучше, чем уравнение (3.7), хотя и то и другое включают всего один фактор (стоимость 1 тонны сырья). Но коэффициент множественной детерминации модели третьего порядка выше (RI = 0.97, число неучтенных факторов 3 %) Данная модель является, как и все предыдущие, адекватной, так как $F_P > F_T$ (при степенях свободы v_1=1 и v_2=28 имеем 859,21 > 4,20), и на ее основе можно принимать решения и осуществлять прогнозы.

Смешанные модели, в которых используются факторы второго, третьего и выше порядков исключены из рассмотрения, так как при их построении ухудшается значение коэффициента множественной детерминации.

Таким образом, на данный момент наилучшей является модель третьего порядка, с одним факторным признаком.

Рассмотрим логарифмические модели.

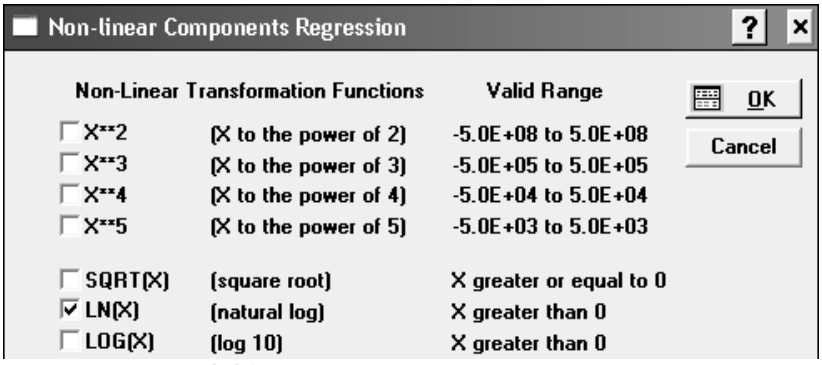


Рисунок 3.31. Нелинейные компоненты регрессии

Выбираем следующие факторные признаки.

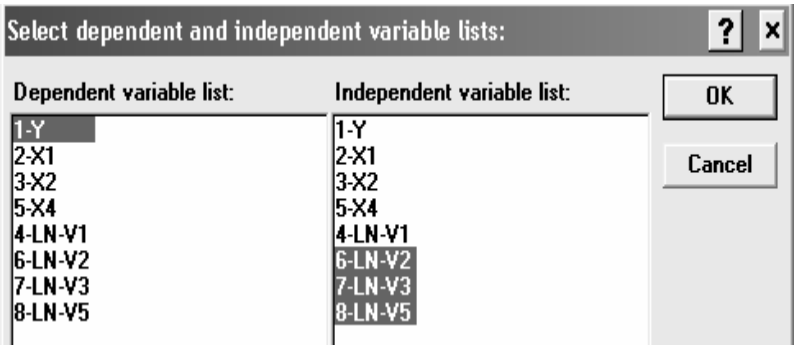


Рисунок 3.32. Выбор зависимых и независимых переменных

Где LN V2 – натуральный логарифм x_1, LN V3 –натуральный логарифм x_2, LN V5 – натуральный логарифм x_4.

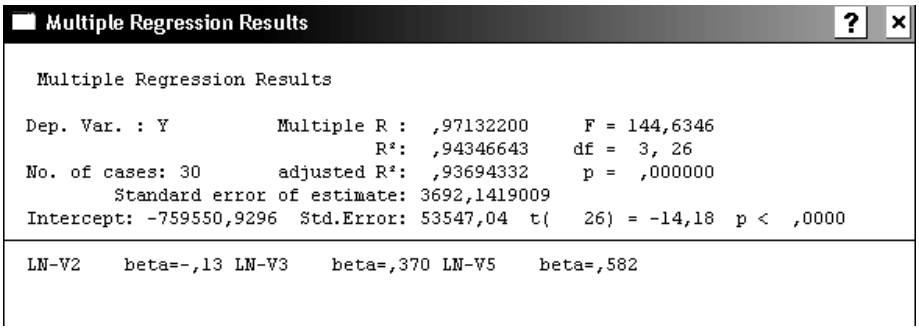
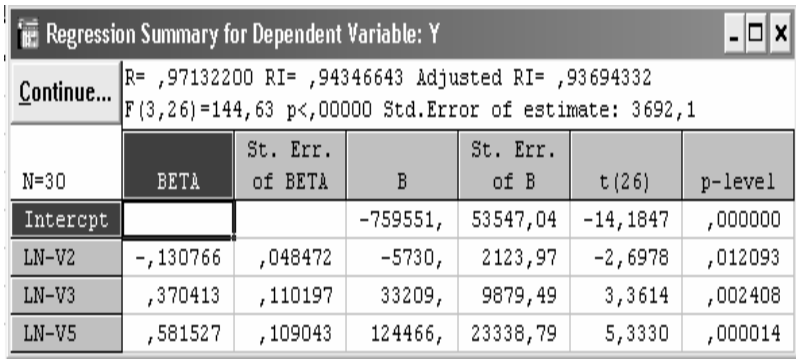


Рисунок 3.33. Результаты множественной регрессии



N=30	BETA	St. Err. of BETA	B	St. Err. of B	t(26)	p-level
Intercpt			-759551,	53547,04	-14,1847	,000000
LN-V2	-,130766	,048472	-5730,	2123,97	-2,6978	,012093
LN-V3	,370413	,110197	33209,	9879,49	3,3614	,002408
LN-V5	,581527	,109043	124466,	23338,79	5,3330	,000014

Рисунок 3.34. Суммарная регрессия
для зависимой переменной Y

На основе данных из рисунков 3.33-3.34 можно получить следующее уравнение регрессии

$$y = -759551 - 5730\log x_1 + 33209\log x_2 + 124466\log x_4 \qquad (3.9)$$

Если анализировать его, то можно сказать определенно, что модель является адекватной, на ее основе можно принимать решения и осуществлять прогнозы. Но определяющим здесь является коэффициент множественной детерминации. Для данной модели RI = 0,94, что говорит о 6% неучтенных факторов.

Если рассматривать смешанные модели, в которые входят как факторы высших порядков, так и логарифмические, то определенно можно сказать, что коэффициент множественной детерминации не изменится и не превысит значения 0,97. Если рассматривать модели с натуральным и десятичным логарифмом, то коэффициент множественной детерминации опять-таки равен 0,94, что не является лучшим показателем.

Выводы.

Можно выбрать две лучшие модели (уравнения (3.8) и (3.9)).

На их основе можно принимать решения и осуществлять прогнозы.

Если сравнивать эти две модели по коэффициенту множественной детерминации и числу неучтенных факторов, то логарифмическая модель уступает модели третьего порядка. Но если рассматривать их по числу входящих факторов, то оптимальной можно считать именно логарифмическую модель, включающую в себя больше значимых факторов.

Задания для самостоятельного выполнения.

1. *Используя данные таблицы Приложения В, построить прогноз с использованием многофакторных моделей прогнозирования.*
2. *Подобрать оптимальные параметры прогнозной модели.*
3. *Рассчитать ошибку прогнозирования, дополнительно руководствуясь теоретическими положениями, приведенными в Приложении А.*

ГЛАВА 4. НЕЙРОСЕТЕВОЕ ПРОГНОЗИРОВАНИЕ ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ В ПАКЕТЕ STATISTICA NEURAL NETWORKS

Нейронным сетям и прогнозированию на их основе посвящено множество статей, научных работ, монографий. При необходимости детального разбора теории нейронных сетей можно обратиться к следующим источникам [29, 36, 37, 127, 57, 92]. Поэтому теоретические основы построения нейронных сетей не рассматриваются, а основное внимание сосредоточено на практической реализации нейросетевого прогнозирования. Пакет **Statistica Neural Networks** выбран не случайно, так как он позволяет моделировать различные типы сетей, использовать несколько алгоритмов их обучения, а также многое другое.

4.1. Основные возможности пакета Statistica Neural Networks

4.1.1. Создание набора данных

Запустите *ST Neural Networks.* (Меню Пуск → Программы→ STATISTICA Neural Networks)

В пакете *ST Neural Network,* обучающие данные хранятся в виде набора *(Data Set),* содержащих некоторое количество наблюдений, для каждого из которых заданы значения нескольких входных и выходных переменных. Как правило, данные берутся из какого-то внешнего источника (например, системы *STATISTICA* или электронной таблицы). Однако новый набор данных можно создать прямо в пакете *ST Neural Networks*. Для этого нужно проделать следующие действия.

1. Войти в диалоговое окно *Создать набор данных - Create Data Set* с помощью команды *Набор данных - Data Set...* из меню *Файл - Создать - File-New.*

2. Ввести значения числа *входных (Inputs)* и *выходных (Outputs)* переменных. Например, две входные переменные и одна выходная.

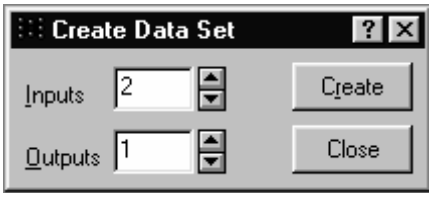


Рисунок 4.1. Создать набор данных

3. Нажать кнопку *Создать – Create.*

Для нашего примера мы используем уже имеющиеся данные, хранящиеся в файле *Series_g.sta* . Программа *ST Neural Networks* автоматически открывает окно *Редактор данных - Data Set Editor.*

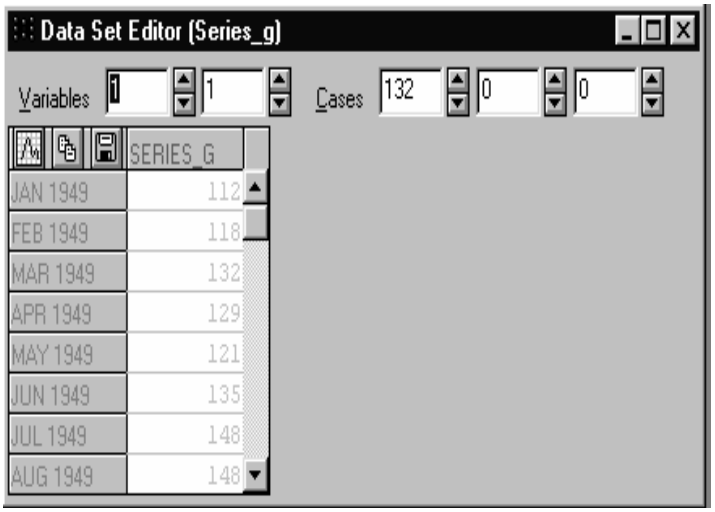


Рисунок 4.2. Редактор данных

Основной элемент окна *Редактор данных - Data Set Editor* - это таблица, содержащая все записи набора данных. Каждому наблюдению соответствует одна строка таблицы. У входных переменных заголовок столбца черного цвета, у выходных - голубого; входы от выходов отделяются темной вертикальной линией. Добавление новых наблюдений и редактирование уже имеющихся данных осуществляется обычным для таблицы способом.

4.1.2. Добавление наблюдений

1. Выберите ячейку таблицы, щелкнув по ней мышью.

2. Нажмите клавишу СТРЕЛКА ВНИЗ. Всякий раз при попытке выйти вниз за границы таблицы программа *ST Neural Networks* создает новое наблюдение.

4.1.3. Удаление лишних наблюдений

Если вы случайно создали лишнее наблюдение, его можно удалить следующим образом.

1. Щелкните в средней части метки строки, соответствующей лишнему наблюдению, метки расположены в левой части таблицы. Вся строка станет выделенной.

2. Нажмите клавиши CTRL+X. Наблюдение будет удалено.

4.1.4. Изменение переменных и наблюдений

В пакете *ST Neural Networks* имеется возможность присваивать имена отдельным наблюдениям и переменным. Чтобы присвоить наблюдению имя, сделайте следующее.

1. Дважды щелкните в средней части метки строки этого наблюдения. Метки строк расположены в левой части таблицы. Появится текстовый курсор (серая вертикальная полоса).

2. Введите имя. В качестве метки строки по умолчанию берется ее номер. Не обращайте на него внимания - он отображается только в строках, которым не присвоено имя, и исчезнет сразу же как только вы начнете вводить символы.

3. С помощью клавиш СТРЕЛКА ВЛЕВО и СТРЕЛКА ВПРАВО курсор может передвигать по буквам имени, клавишами DELETE и BACKSPACE - удалять лишние символы, с помощью клавиш СТРЕЛКА ВВЕРХ И СТРЕЛКА ВНИЗ можно перейти к именам других наблюдений, клавиша ESCAPE прерывает редактирование.

Аналогично присваиваются имена переменным - для этого нужно отредактировать метки столбцов.

4.1.5. Другие возможности редактирования данных

Таблицы пакета *ST Neural Networks* предлагают большой набор средств, облегчающих создание наборов данных и последующую

работу с ними.

- **Перемещение активной ячейки.** Осуществляется клавишами СТРЕЛКА ВЛЕВО, СТРЕЛКА ВПРАВО, СТРЕЛКА ВВЕРХ, СТРЕЛКА ВНИЗ, HOME, END, PAGE UP, PAGE DOWN.

- **Выделение диапазона ячеек.** Производится перетаскиванием указателя мыши или клавишами курсора при нажатой клавише SHIFT.

- **Копирование и вставка.** Чтобы скопировать выделенный диапазон ячеек в буфер обмена, нажмите CTRL+C. Чтобы вставить содержимое буфера обмена в таблицу – нажмите CTRL+V. Можно копировать и вставлять целые строки и столбцы целиком. Возможен также обмен данными между пакетом *ST Neural Networks* и другими приложениями.

- **Вставка.** В любом месте таблицы можно вставить новую строку или столбец. Поместите курсор мыши на линию, разделяющую метки двух соседних строк или столбцов (при этом курсор превратится в двухстороннюю стрелку), и щелкните кнопкой - откроется полоса вставки. После нажатия клавиши INSERT будет вставлена новая строка столбец.

- Чтобы **назначить тип переменной** - *Входная - Input, Выходная -Output, Входная/Выходная - Input/Output* или *Неучитываемая - ignored,* выберите переменную, щелкнув на метке соответствующего столбца, затем нажмите правую кнопку мыши и выберите нужный тип из контекстного меню.

- Чтобы задать номинальную переменную (например *Пол - {Муж, Жен}),* выберите переменную, щелкнув на метке соответствующего столбца, затем нажмите правую кнопку мыши и выберите команду *Определение-Definition* из контекстного меню.

- Чтобы задать тип подмножества *Обучающее – Training, Контрольное - Verification, Тестовое - Test* или *Неучитываемое - Ignored,* выбирайте наблюдения, щелкая на метках их строк, нажимайте правую кнопку мыши и выбирайте нужный тип из контекстного меню.

- Все перечисленные возможности доступны также через команды *Наблюдения – Cases…* и *Переменные – Variables…* меню Правка – *Edit.*

4.1.6. Создание новой сети

Создать новую сеть в пакете *ST Neural Networks* можно либо средствами диалогового окна *Создать сеть – Create Network,* доступ к которому осуществляется через команду *File → New → Network*…. Кроме того, можно создать сеть, пользуясь автоматическим конструктором сети (кнопка).

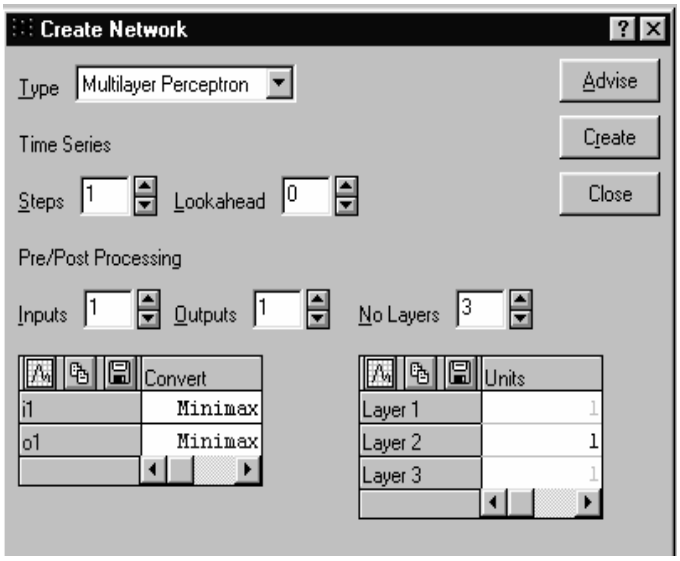


Рисунок 4.3. Создать сеть

4.1.7. Создание сети

1. Выберите тип сети из выпадающего списка *Тип – Type.* Тип *Многослойный персептрон - Multilayer Perceptron* предлагается по умолчанию.

2. Нажмите кнопку *Совет – Advise.* Программа *ST Neural Networks* установит параметры по умолчанию для пре/пост-процессирования и конфигурации сети исходя из типа переменных, составляющих исходные данные. В этом диалоговом окне можно задать и некоторые другие параметры, в том числе параметры временного ряда *(Time Series) Временное окно - Steps* и *Горизонт - Lookahead* параметры преобразования и подстановки пропущенных

значений при пре/пост - процессировании, ширину слоев сети.

3. Нажмите кнопку *Создать - Create,* в результате будет создана новая сеть.

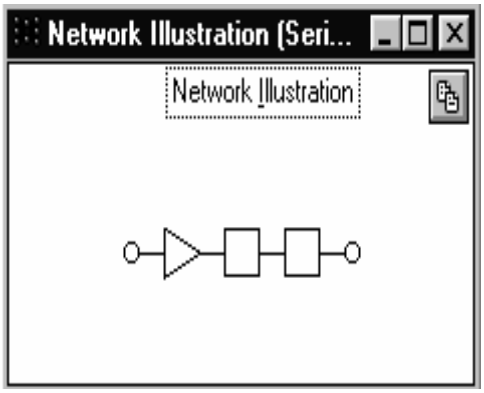


Рисунок 4.4. Иллюстрация сети

4.1.8. Сохранение набора данных и сети

1. Откройте диалоговое окно *Сохранить набор данных - Save Data Set* с помощью команды *Набор данных - Data Set* из меню *Файл - Сохранить как – File - Save as.*

2. Введите имя файла данных в поле *Имя файла - File Name.*

3. Нажмите кнопку *Сохранить – Save.*

Сеть сохраняется аналогичным образом с помощью окна *Сохранить сеть - Save Networks,* в качестве стандартного расширения имени файла сети используются «*.*net*» или «*.*bnt*».

4.1.9. Обучение сети

Следующий шаг после задания набора данных и построения подходящей сети - это обучение.

В пакете *ST Neural Networks* реализованы основные алгоритмы обучения многослойных персептронов: методы обратного распространения, сопряженных градиентов и Левенберга-Маркара.

Суть метода обратного распространения

1. Алгоритм обратного распространения последовательно обучает сеть на данных из обучающего множества. На каждой итерации все наблюдения из обучающего множества (в данном случае оно совпадает со всем набором данных) по очереди подаются на вход сети. Сеть обрабатывает их и выдает выходные значения.

2. Эти выходные значения сравниваются с целевыми выходными значениями, которые также содержатся в наборе исходных данных, и ошибка, то есть разность между желаемым и реальным выходом, используется для корректировки весов сети так, чтобы уменьшить эту ошибку.

3. Алгоритм должен находить компромисс между различными наблюдениями и менять веса таким образом, чтобы уменьшить суммарную ошибку на всем обучающем множестве, поскольку алгоритм обрабатывает наблюдения по одному, общая ошибка на отдельных шагах обязательно будет убывать.

В пакете *ST Neural Networks* отслеживается общая ошибка сети - на графике, а также ее ошибки на отдельных наблюдениях на гистограмме.

Обучение методом обратного распространения

1. Откроите окно *График ошибки обучения – Training Error Graph* с помощью команды *График обучения - Training Graph* меню *Статистики -Statistics* или кнопки (см. рис 4.5).

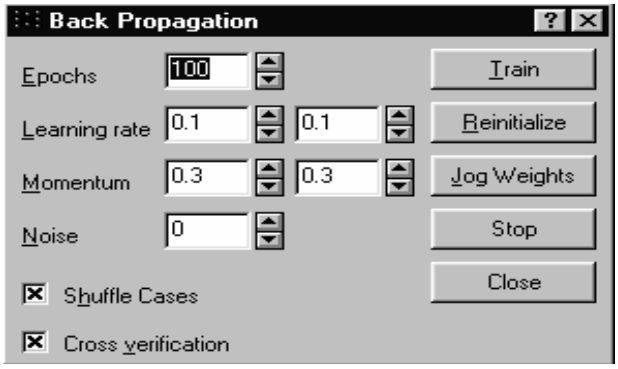


Рисунок 4.5. Обратное распространение

2. Откройте диалоговое окно *Обратное распространение - Back Propagation* с помощью команды *Обратное распространение – Back Propagation…* меню Обучение многослойного персептрона – *Train-Multilayer Perceptron* или кнопки .

3. Нажмите кнопку *Обучить – Train,* в диалоговом окне *Обратное распространение – Back Propagation* - будет запущен алгоритм обучения. При этом на график будет выводиться ошибка.

4. Повторно нажимайте кнопку *Обучить – Train,* чтобы алгоритм переходил к очередным эпохам.

4.1.10. Оптимизация обучения

Режим работы алгоритма обратного распространения зависит от ряда параметров, большинство из которых собрано в диалоговом окне *Обратное распространение – Back Propagation* (рисунок 4.5).

- *Эпохи – Epochs.* Задает число эпох обучения, которые проходят при одном нажатии клавиши *Обучить –Train.* Значение по умолчанию 100 вполне приемлемо.

- *Скорость обучения – Learning rate.* При увеличении скорости обучения алгоритм работает быстрее, но в некоторых задачах это может привести к неустойчивости.

- *Инерция – Momentum.* Этот параметр улучшает (ускоряет) обучение в ситуациях, когда ошибка мало меняется, а также придает алгоритму дополнительную устойчивость. Значение этого параметра всегда должно лежать в интервале [0;1). Часто рекомендуется использовать высокую скорость обучения в сочетании с небольшим коэффициентом инерции и наоборот.

- *Перемешивать наблюдения - Shuffle Cases.* При использовании этой функции порядок, в котором наблюдения подаются на вход сети, меняется в каждой новой эпохе. Это добавляет в обучение некоторый шум, так что ошибка может испытывать небольшие колебания. Однако при этом меньше вероятность того, что алгоритм «застрянет», и общие показатели его работы обычно улучшаются.

4.1.11. Выполнение повторных прогонов

Если вы хотите сравнить результаты работы алгоритма в разных

вариантах, воспользуйтесь кнопкой *Переустановить - Reinitialize* диалогового окна *Обратное распространение - Back Propagatio* (рисунок 4.5). В результате веса сети вновь будут установлены случайным образом для начала следующего сеанса обучения. Если теперь после кнопки *Переустановить -Reinitialize* нажать кнопку *Обучить - Train,* на графике начнет рисоваться новая линия.

Совет. Чтобы сделать сравнение более наглядным, можно рисовать линии разными цветами. Введите значение в поле *Метка - Label* в окне *график обучения - Training Graph,* (рисунок 4.6) и тогда следующая линия будет парирована другим цветом, а указанная метка будет выведена справа от графика как условное обозначение.

4.1.12. Ошибки для отдельных наблюдений

В окне *График обучения - Trainig Graph* выводится суммарная ошибка сети. Но иногда бывает полезно проследить за тем, как алгоритм обучения воспринимает отдельные наблюдения.

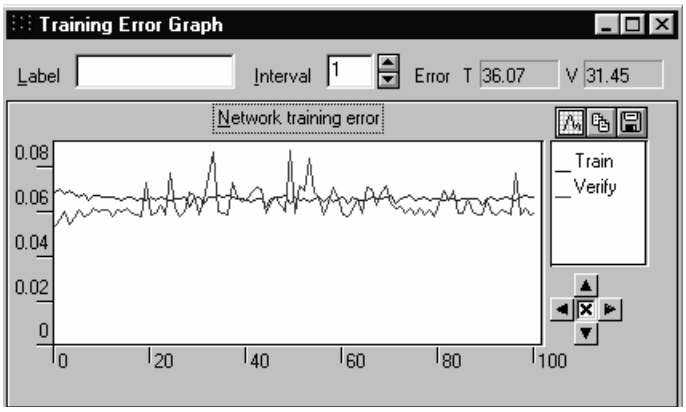


Рисунок 4.6. График обучения

В пакете *ST Neural Networks* это делается в окне *Ошибки наблюдений… - Case Errors,* которое открывается командой *Ошибки наблюдений - Case Errors…* меню *Статистики - Statistics* или кнопкой .

Ошибки на отдельных наблюдениях выводятся в виде гистограммы (рисунок 4.7).

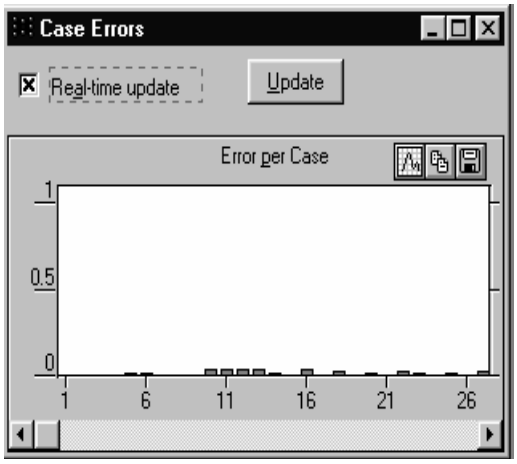


Рисунок 4.7. Ошибки наблюдений

В конце сеанса обучения ошибки пересчитываются. Имеется также возможность следить за тем, как они меняются в процессе обучения - для этого служит функция *Пересчитывать по ходу - Real-time update* окна *Ошибки наблюдений - Case Errors;* ее нужно активизировать перед запуском алгоритма обратного распространения. Сделав это, вы сможете наблюдать, как алгоритм пытается искать компромисс между обучающими и мешающими наблюдениями.

4.2. Запуск нейронной сети

После того, как сеть обучена, ее можно запустить на исполнение. В пакете *ST Neural Networks* это можно сделать в нескольких вариантах:

- на текущем наборе данных - в целом или на отдельных наблюдениях;
- на другом наборе данных - в целом или на отдельных наблюдениях (такой набор данных уже может не содержать выходных значений и предназначаться исключительно для тестирования);
- на одном конкретном наблюдении, для которого значения переменных введены пользователем, а не взяты из какого-то файла данных, из другого приложения с помощью интерфейса прикладного программирования SNN API.

4.2.1. Обработка наблюдений по одному

Для обработки отдельных наблюдений из набора данных служит окно *Прогнать одно наблюдение - Run Single Case,* доступ к которому осуществляется *Run → Single Case ...* или кнопкой на панели инструментов (рисунок 4.8).

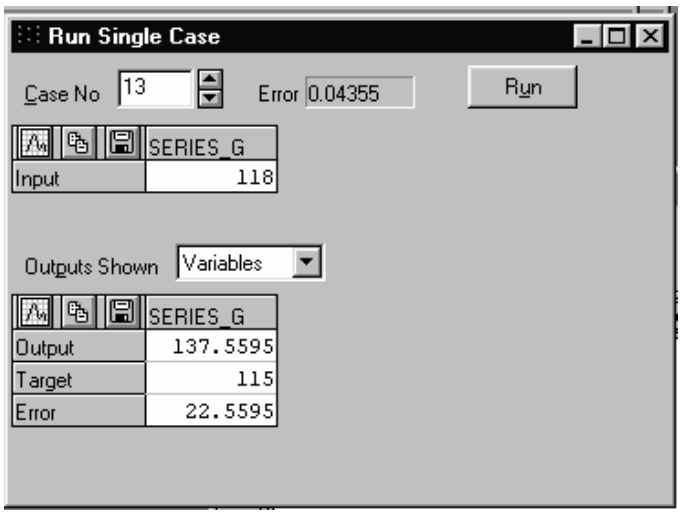


Рисунок 4.8. Прогнать одно наблюдение

В поле *Номер наблюдения* - Case *No* задается номер наблюдения, подлежащего обработке. Чтобы обработать текущее наблюдение, нажмите кнопку *Запуск – Run,* а для обработки какого-либо другого наблюдения введите соответствующий номер в поле *Номер наблюдения – Case No* и нажмите клавишу ВВОД (рисунок 4.8).

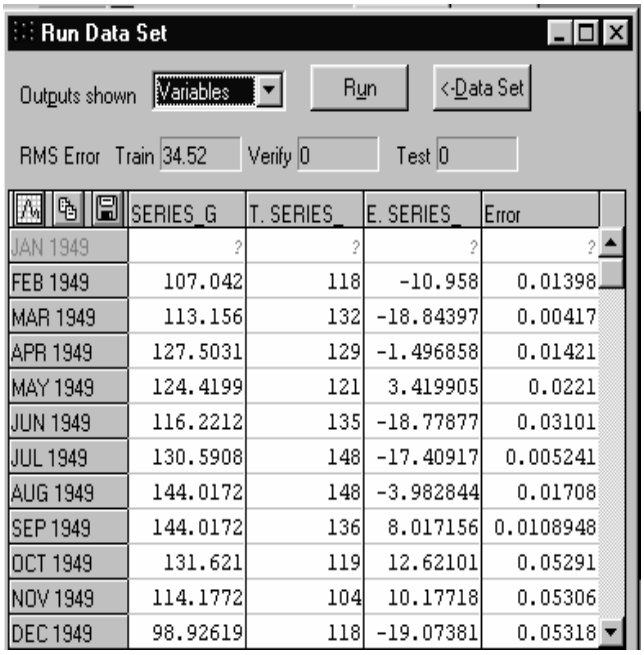
Значения входных переменных для текущего наблюдения отображаются в таблице, расположенной в верхней части окна, а выходные значения в нижней таблице.

Помимо фактического выходного значения, которое выдает сеть, выводится также целевое значение и ошибка.

В пакете *ST Neural Networks* предусмотрены различные форматы вывода (для этого служит выпадающий список *Показывать при выводе - Outputs Shown).* Сейчас мы использовали установленный по умолчанию вариант *Переменные – Variables.*

4.2.2. Прогон всего набора данных

Для тестирования сети на всем наборе данных служит окно *Прогнать набор данных - Run Data Set,* доступ к которому осуществляется через пункт *Набор данных Data* Set… меню *Запуск - Run* или кнопкой Нажмите КНОПКУ *Запуск – Run,* чтобы протестировать сеть, и результаты будут выведены в таблицу в нижней части окна (рисунок 4.9).



	SERIES_G	T. SERIES_	E. SERIES_	Error
JAN 1949	?	?	?	?
FEB 1949	107.042	118	-10.958	0.01398
MAR 1949	113.156	132	-18.84397	0.00417
APR 1949	127.5031	129	-1.496858	0.01421
MAY 1949	124.4199	121	3.419905	0.0221
JUN 1949	116.2212	135	-18.77877	0.03101
JUL 1949	130.5908	148	-17.40917	0.005241
AUG 1949	144.0172	148	-3.982844	0.01708
SEP 1949	144.0172	136	8.017156	0.0108948
OCT 1949	131.621	119	12.62101	0.05291
NOV 1949	114.1772	104	10.17718	0.05306
DEC 1949	98.92619	118	-19.07381	0.05318

Рисунок 4.9. Прогнать набор данных

В таблице окна *Прогнать набор данных - Run Data Set* содержатся следующие значения (перечисленные слева направо): фактические выходы сети, целевые выходные, значения ошибки и суммарная ошибка по каждому наблюдению. Над таблицей выдается *ОКОНЧАТЕЛЬНАЯ СРЕДНЕКВАДРАТИЧЕСКАЯ ОШИБКА* (СКО) – RMS ERROR сети на этом наборе данных.

4.2.3. Тестирование на отдельном наблюдении

Иногда необходимо протестировать сеть на отдельном наблюдении, не принадлежащем никакому набору данных. Причины для этого могут быть такие.

- Обученная сеть используется для построения прогнозов на новых данных с неизвестными выходными значениями.

- Вы хотите поэкспериментировать с сетью, например, проверить чувствительность результата к малым изменениям в данных.

Тестирование заданных пользователем наблюдений проводится из окна *Прогнать отдельное наблюдение - Run One-off...,* доступ к которому осуществляется через *Run* → *One-off...*(рисунок 4.10).

Рисунок 4.10 Прогнать отдельное наблюдение

Для этого нужно ввести входные значения в таблицу, расположенную в верхней части окна, и нажать кнопку *Запуск – Run,* результаты будут выведены в нижнюю таблицу.

4.3. Создание сети типа Многослойный персептрон

Для создания сети типа многослойный персептрон в первую очередь необходимо выполнить следующие действия.

Пуск→Программы → *STATISTICA Neural Networks*→ *STATISTICA Neural Networks*→ в появившемся окне *Open Data Set* выбираете файл *SERIES_G.*

Возможны два варианта создания сети: с помощью мастера создания сети и самостоятельно.

Рассмотрим вначале вариант создания сети типа многослойный персептрон с помощью мастера.

4.3.1. Создание сети типа многослойный персептрон с помощью мастера

После того как вы запустили программу, перед вами появятся следующие диалоговые окна.

Data Set Editor (Редактор набора данных) (рисунок 4.11), где будут представлены все имеющиеся значения нашей единственной переменной и диалоговое окно мастера - *Intelligent Problem Solver-Basic or Advanced* (рисунок 4.12) - окно помощника решения задач.

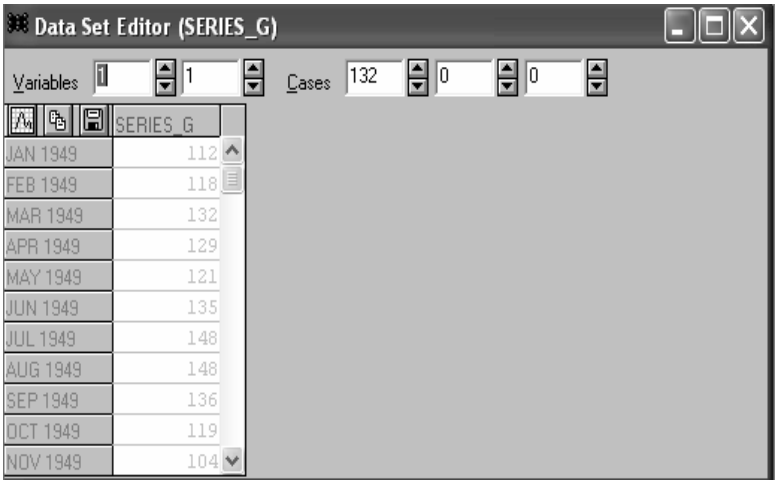
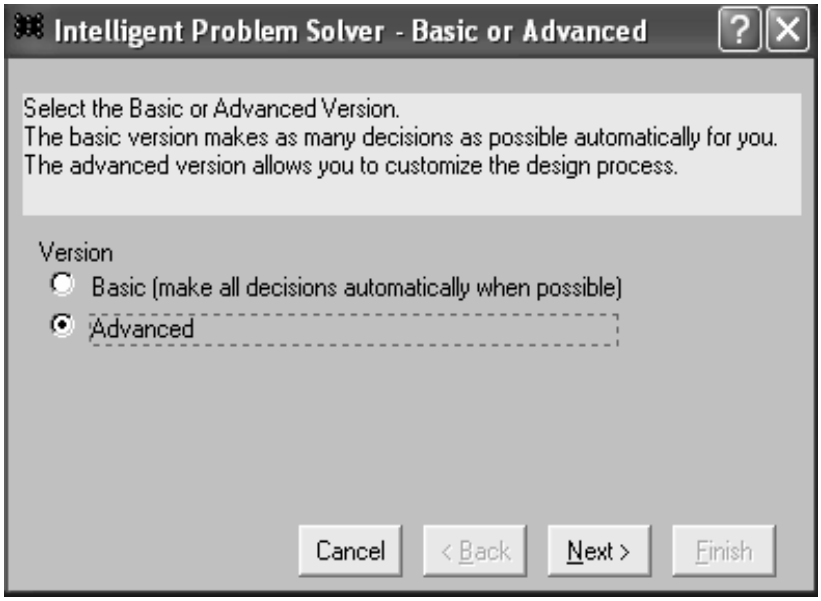


Рисунок 4.11. Редактор набора данных

В появившемся окне *Intelligent Problem Solver-Basic* or *Advanced* (рисунок 4.12) выбираем второй вариант *Advanced* (продвинутая версия, с помощью которой можно осуществить обычный процесс)→ нажимаем *Next*.



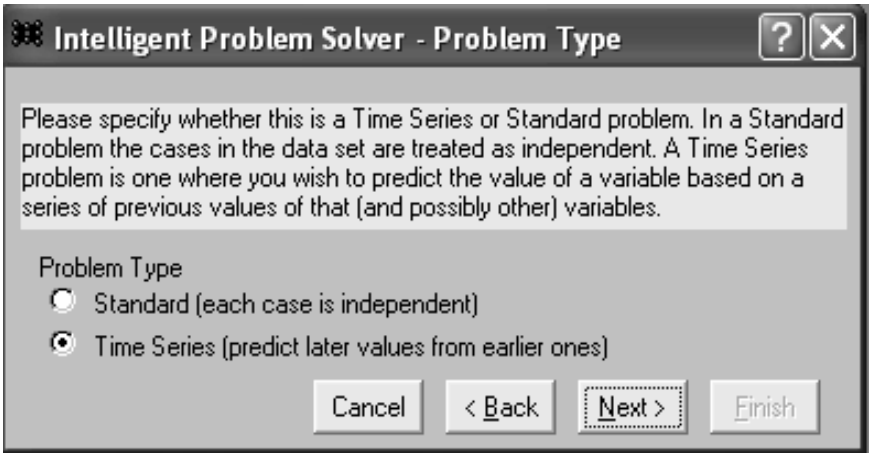
Выберите основную или расширенную версию. Основная версия делает максимальное количество решений за вас. Расширенная версия позволяет контролировать процесс создания нейронных сетей

Рисунок 4.12. Окно помощника решения задач - Основная или продвинутая версия

После нажатия клавиши *Next* появится следующее диалоговое окно - Intelligent Problem Solver - Problem Type (рисунок 4.13), где снова выбираем второй вариант - *Time Series* (predict later values from earlier ones) → *Next*.

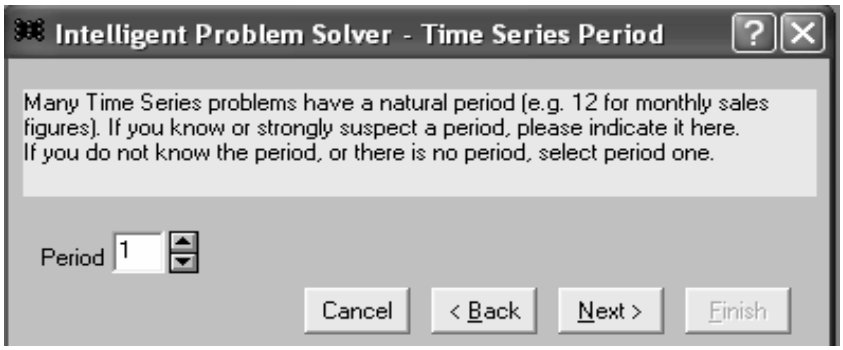
Опять нажимаем клавишу продолжения *Next,* и на экран выводится следующее диалоговое окно мастера - *Intelligent Problem Solver-Time Series Period* (рисунок 4.14), где устанавливаем период равный 1.

Установленный период означает, на сколько шагов вперед осуществляется предсказание (если он неизвестен, обычно вводится значение 1, иначе по умолчанию задается – 0). Предполагается, что мастер создания сети будет определять оптимальный размер шагов автоматически.



Укажите, является исходные данные временным рядом или нет. В обычном ряду набор данных считается независимым, а во временном ряду каждое последующее значение зависит от предыдущих или от других переменных.

Рисунок 4.13. Окно помощника решения задач - Тип задачи



Многие временные ряды имеют натуральный период (лаг). Если вы четко знаете период, то укажите его. Если не знаете или периода нет, выберите значение равным 1.

Рисунок 4.14. Окно помощника решения задач.
Период временного ряда

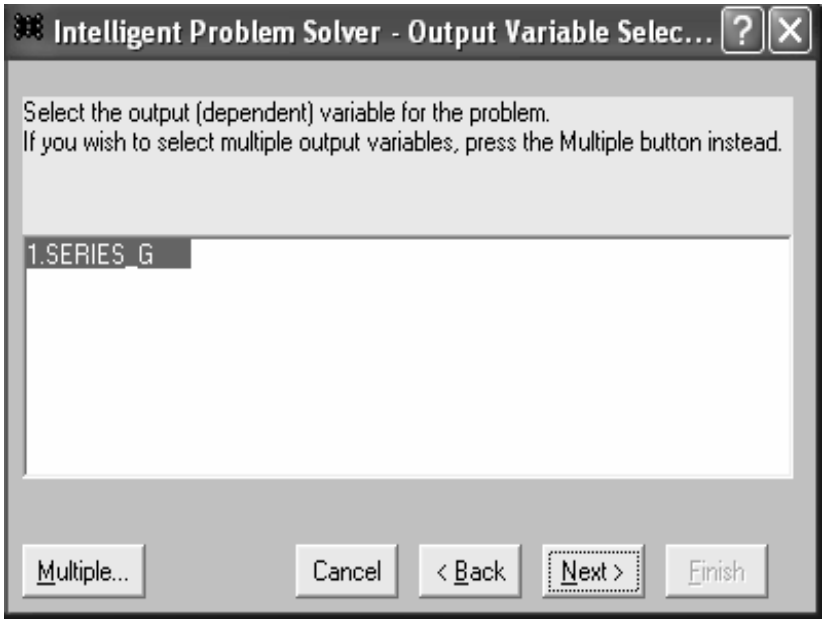
Замечание. Обычно очередное значение временного ряда прогнозируется по некоторому числу его предыдущих значений (прогноз на один шаг вперед во времени). В представленном пакете можно выполнять прогноз на любое число шагов. После того, как вычислено очередное предполагаемое значение, оно подставляется

обратно в ряд, и с его помощью получается следующий прогноз – это называется проекцией временного ряда.

После нажатия клавиши *Next* появится следующее окно (рисунок 4.15), где будет указана автоматически, в нашем случае, выходная переменная *SERIES_G*, которая является также и входной переменной (так как в выбранном файле *SERIES_G*, имеется всего один ряд значений). Поэтому после нажатия клавиши *Next* в появившемся диалоговом окне, показанном на рисунке 4.16, будет также автоматически выбрана переменная SERIES_G.

Причем данная переменная является входной и выходной переменной.

Если Вы желаете создать сеть с множественными выводами, нажмите кнопку -*Multiple*, где Вы можете выбрать множественные переменные вывода. Далее нажмите *Next*.



Выберите выходные (зависимые) переменные. Если необходимо, выберите несколько, нажимая на кнопку Multiple.

Рисунок 4.15. Окно помощника решения задач.
Выбор выходной переменной

Замечание. Данные, подаваемые на вход и снимаемые с выхода, должны быть правильно подготовлены. Один из способов – масштабирование

$$x = (x' - m) \cdot c \qquad 4.1$$

где x' – исходный вектор;
x – масштабированный;
c – масштабный коэффициент.

Масштабирование желательно, чтобы привести данные в допустимый диапазон. Если этого не сделать, то нейроны входного слоя окажутся в постоянном насыщении или будут все время заторможенными.

Простейшей из масштабируемых функций пакета *STATISTICA Neural Networks* является минимаксная функция: она находит минимальное и максимальное значение переменной по обучающему множеству и выполняет линейное преобразование (с применением коэффициента масштаба и смещения).

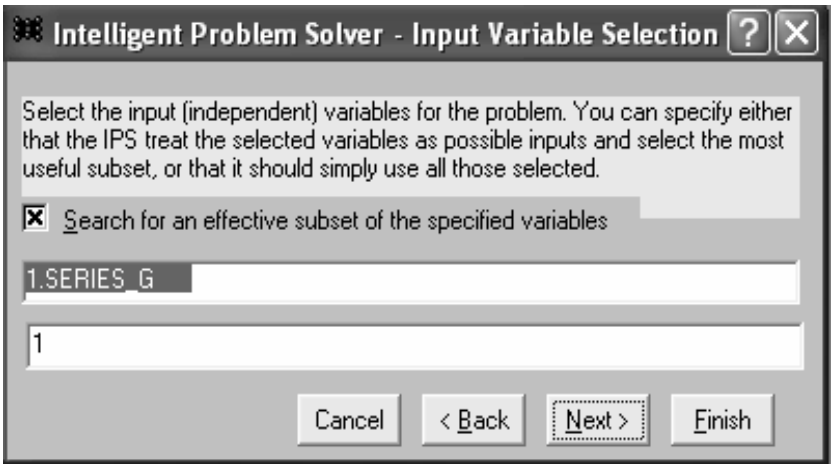


Рисунок 4.16. Окно помощника решения задач.
Выбор входной переменной

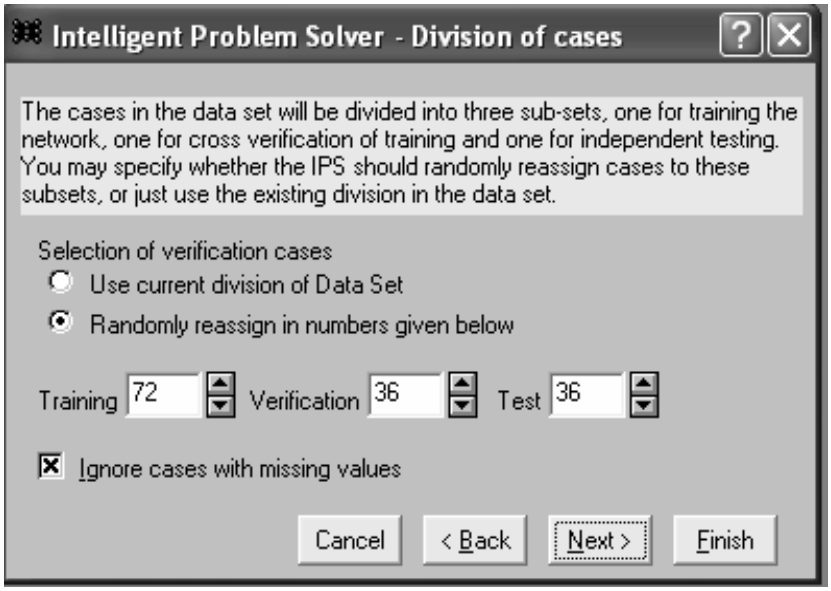
В следующем окне - *Intelligent Problem Solver - Division of cases* (рисунок 4.17), выбираем вариант автоматического распределения

значений для обучающей (72) значения, верификационной (36) значений и тестовой (36) значений выборки-*Randomly reassign in numbers given below*. А также ставим флажок в окне – *Ignore cases with missing values*) (игнорирование неизвестных величин).

Случаи, в наборе данных, будут разделены на три поднабора, один - чтобы обучить сеть, другой - для проверки обучения и третий – для независимого испытания.

Вы можете определить выбор случаев для проверки.

Как Вы можете заметить, наибольшей является обучаемая выборка, и равными – выборка для проверки обучения и выборка для независимого испытания. Причем необходимо учитывать, что ошибка на выборке, используемой для проверки обучения, не должна превышать ошибку на независимой выборке.



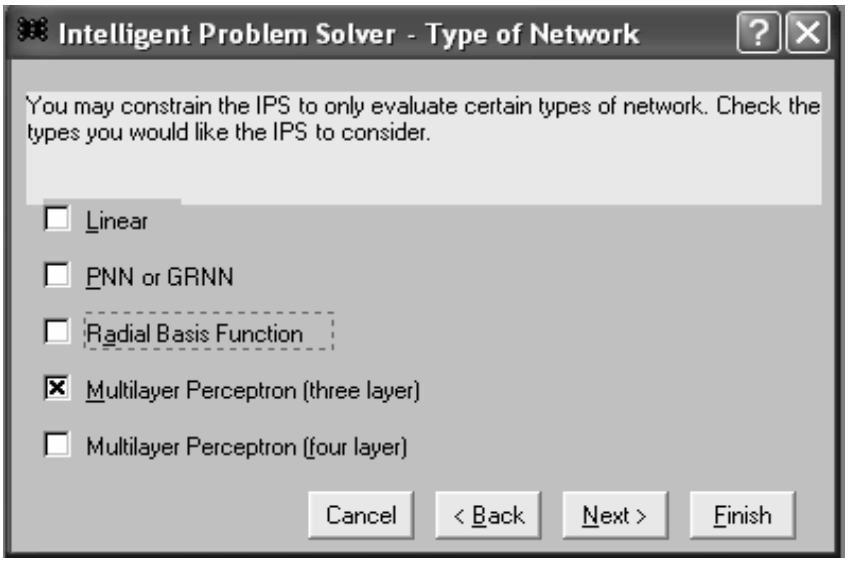
Варианты в множестве данных будут поделены на три подмножества. Одно для обучения сети, другое для верификации, третье для независимого тестирования. Укажите, должен ли IPS случайным образом перемешивать данные в этих подмножествах или использовать без изменений.

Рисунок 4.17. Окно помощника решения задач.
Распределение случаев (значений переменной)

В следующем диалоговом окне, показанном на рисунке 4.18 - *Intelligent Problem Solver – Type of Network*, выбираем тип строящейся сети – *Multiplayer Perceptron* (*three layer*).

Замечание. Нет строго определенной процедуры для выбора количества нейронов и количества слоев сети. Чем больше количество нейронов и слоев, тем шире возможности сети, тем медленнее она обучается и работает и тем более нелинейной может быть зависимость вход-выход.

Существует возможность указания мастеру создания сети создать и дать оценку только определенному типу сети. При решении нашей задачи выбираем многослойный персептрон (3 слоя), так как это цель нашего построения.



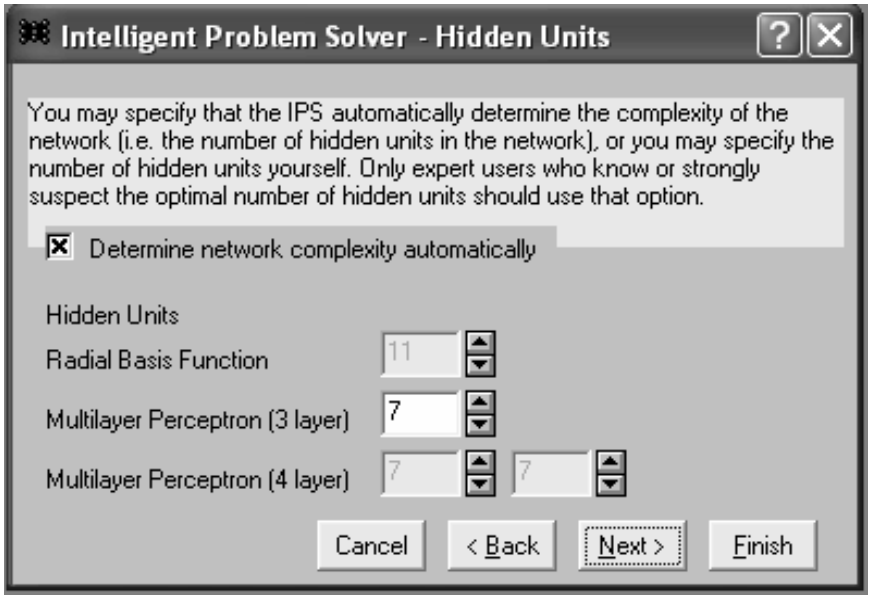
Вы можете указать, какой тип сети необходимо создать

Рисунок 4.18. Окно помощника решения задач. Тип сети

Как Вы можете видеть из диалогового окна, представленного на рисунке 4.18, типы сетей упорядочены в порядке возрастания сложности обучения. Линейные сети почти не требуют обучения и включены потому, что дают хорошую точку отсчета для сравнения эффективности различных методов. Сети PNN и GRNN также довольно просты, радиальные базисные функции устроены несколько сложнее, а трех- и четырех – многослойные персептроны – это очень сложные конструкции.

Для трехслойного MLP (многослойного персептрона) поиск сводится к выбору числа элементов в скрытом слое. По существу, это линейный поиск для функции, содержащей помехи (при разных прогонах обучения сети, даже с одним и тем же числом элементов, могут получаться немного различные результаты).

После нажатия клавиши продолжения получим следующее диалоговое окно - *Intelligent Problem Solver – Hidden Units* (рисунок 4.19), где ставим флажок автоматического определения.



Вы можете указать количество нейронов в скрытом слое самостоятельно. Однако вы можете указать, чтобы IPS автоматически определял сложность нейронной сети.

Рисунок 4.19. Окно помощника решения задач. Скрытые элементы слоя

Затем появляется (после нажатия Next) следующее диалоговое окно - *Intelligent Problem Solver – Duration of Design P…*(рисунок 4.20), где выбираем длительность процесса создания сети по времени – *Quick*.

Длительность процесса проектирования:

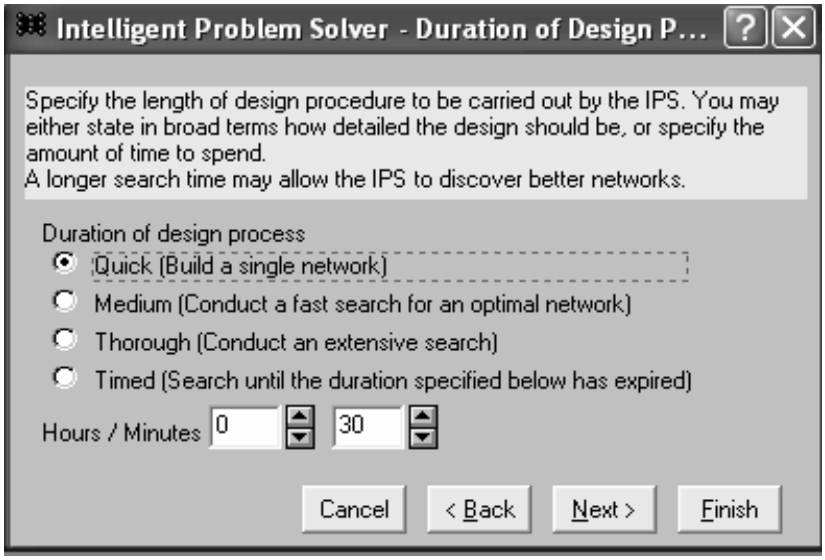
Быстрый - минимальный поиск. Обнаруженные сети будут подоптимальными.

Средний по продолжительности – достаточен, чтобы определить местонахождение почти оптимального решения.

Полный, всесторонний, тщательный – рекомендуется для проблем большего масштаба.

Последний вариант предполагает установку определенного времени построения в часах и минутах.

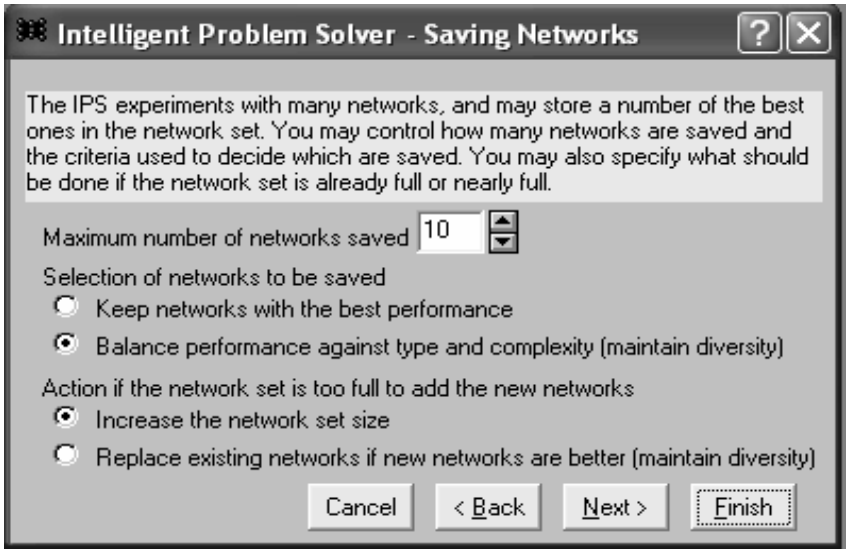
После нажатия *Next* получаем следующее диалоговое окно - Intelligent *Problem Solver – Saving Networks* (рисунок 4.21). Это диалоговое окно позволит Вам определить, сколько сетей должно быть сохранено, и определить критерии, в соответствии с которыми мастер принимает решение сохранить сеть.



Укажите длительность процесса создания сети.

Рисунок 4.20. Окно помощника решения задач.
Продолжительность проектирования сети

Замечание. Основным ограничением при построении сетей является время, уходящее на обучение тестирование сетей. Даже учитывая то, что MLP небольшие и быстро работают. Но важным преимуществом является то, что алгоритм осуществляет перебор гораздо большего числа вариантов, чем способен проделать человек, и соответственно результаты автоматического конструктора оказываются лучше.



IPS проводит вычислительные эксперименты с различными типами сетей и сохраняет наилучший вариант. Вы можете указать количество сетей для сохранения, а также указать критерии сохранения.

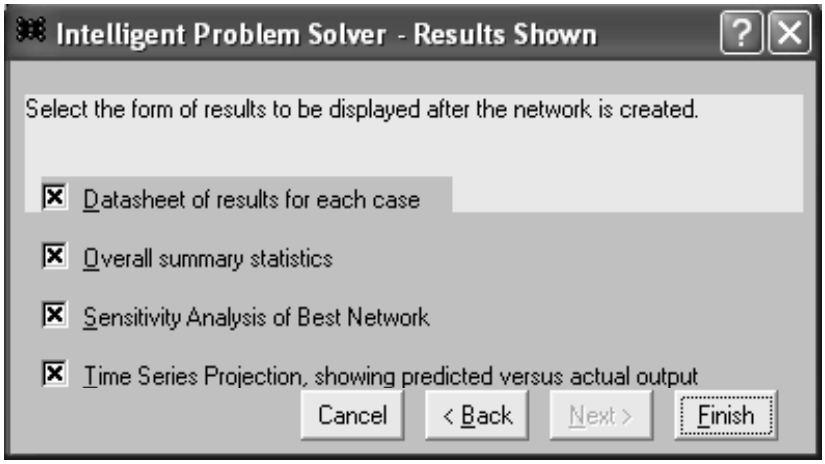
Рисунок 4.21. Окно помощника решения задач.
Сохранение сети

В появившемся диалоговом окне (после нажатия клавиши *Next*) – Intelligent *Problem Solver – Results Shown*, представленном на рисунке 4.22, выбираем все варианты:

- таблица данных по каждому случаю;
- общая полная суммарная статистика;
- анализ лучшей (адекватной) сети;
- проектирование временных рядов.

Возможность выбора представления данных.

А затем→ *Finish.*



Выбрать форму представления данных.
-таблица данных;
- общая, суммарная статистика;
- анализ лучшей сети;
- проектирование временных рядов.
Выбираем все варианты.

Рисунок 4.22. Окно помощника решения задач.
Показать данные

В итоге на экране появятся четыре окна: рисунки 4.23 - 4.26.

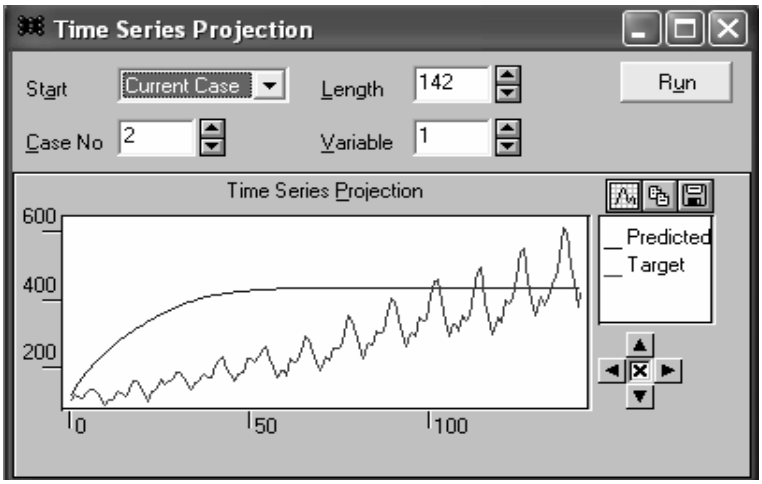


Рисунок 4.23. Прогноз временного ряда

На рисунке 4.24 представлена вся информация по построению сети, где оговорено, что было построено и рассмотрено 10 сетей и представлены характеристики наилучшей сети.

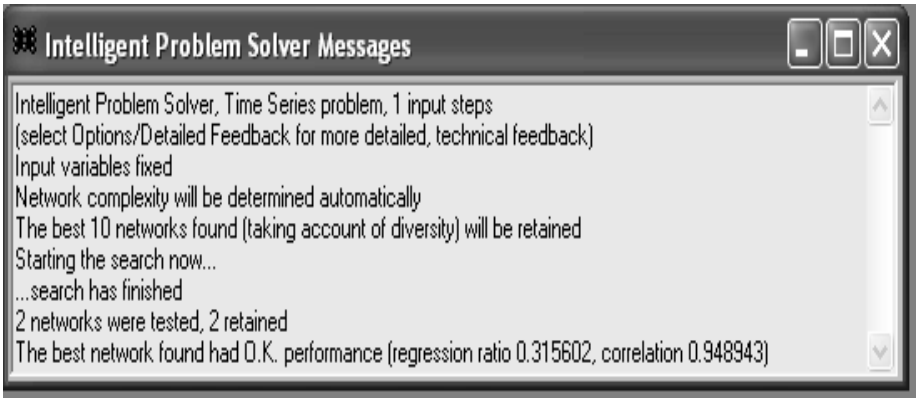
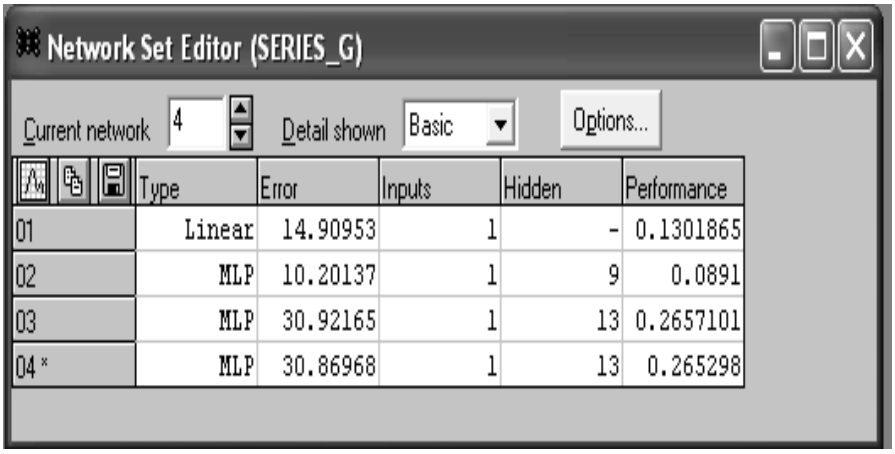


Рисунок 2.24. Сообщение помощника решения задач

Из всех сетей было выбрано четыре (рисунок 4.25) оптимальные, и наилучшей из них была признана четвертая (04*). Причем, как видно из представленной таблицы (рисунок 4.25), рассматривались и линейные сети, а не только множественный персептрон. Кроме типа сети указывается также и ошибка для сетей, причем, заметим, что у наилучшей сети ошибка не самая минимальная, и это обусловлено тем, что в представленном диалоговом окне (рисунок 4.25) был выбран баланс между качеством и параметрами сети.



	Type	Error	Inputs	Hidden	Performance
01	Linear	14.90953	1	-	0.1301865
02	MLP	10.20137	1	9	0.0891
03	MLP	30.92165	1	13	0.2657101
04 *	MLP	30.86968	1	13	0.265298

Рисунок 4.25. Редактор установки сети

На рисунке 4.26 – Network Illustration, графически представляется построенная итоговая нейронная сеть.

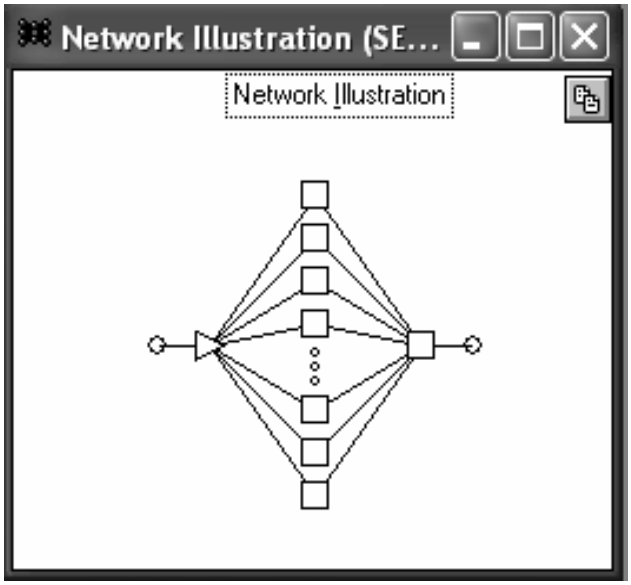


Рисунок 4.26. Иллюстрация сети.

4.4. Построение нейронной сети без мастера

Второй вариант построения нейронной сети заключается в самостоятельном построении, без мастера. Для этого необходимо выполнить следующее.

Пуск→Программы → *STATISTICA Neural Networks*→ *STATISTICA Neural Networks*→ в появившемся окне *Open Data Set* выбираете файл *SERIES_G.*→*File*→*New*→*Network.*

Перед вами появится представленное на рисунке 4.27 диалоговое окно – *Create Network*.

Где выбираем тип сети, которую строим -*Multilayer Perceptron*. Затем указываем параметры временного ряда: *Step -1 Lookahead* (горизонт) - 0. Входную переменную - *Inputs* 1, выходную переменную, которая совпадает с входной - *Outputs* 1. Всего сеть состоит из трех слоев - No Layers 3. После этого нажимаем клавиши - *Advise* (совет), а затем кнопку- *Create* (создать).

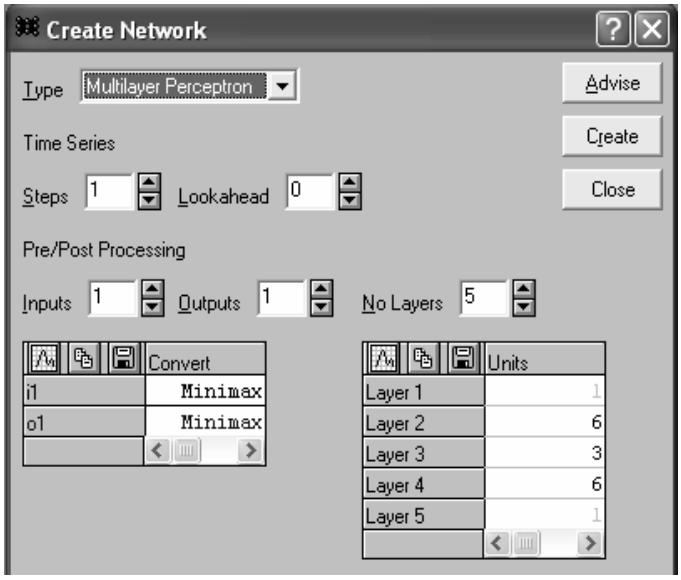


Рисунок 4.27. Создание сети

В результате будет создана новая нейронная сеть, представленная графически на рисунке 4.28.

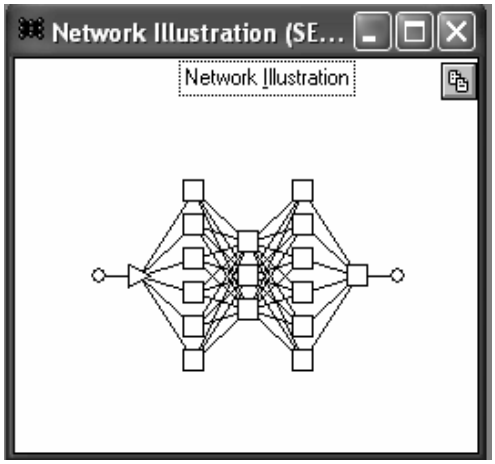


Рисунок 4.28. Иллюстрация сети

Таким образом, были рассмотрены два варианта построения нейронных сетей типа многослойный персептрон: возможность са-

мостоятельного построения нейронной сети и возможность построения нейронной сети с помощью автоматического конструктора.

4.5. Обучение сети

4.5.1. Обучение методом обратного распространения ошибки

Суть метода обратного распространения заключается в следующем.

1. Алгоритм обратного распространения последовательно обучает сеть на данных из обучающего множества. На каждой итерации (они называются эпохами) все наблюдения из обучающего множества по очереди попадаются на вход сети. Сеть обрабатывает их и выдает выходные значения.

2. Эти выходные значения сравниваются с целевыми выходными значениями, которые также содержатся в наборе исходных данных, и ошибка, то есть разность между желаемым и реальным выходом, используется для корректировки весов сети так, чтобы уменьшить эту ошибку.

3. Алгоритм должен находить компромисс между различными наблюдениями и менять веса таким образом, чтобы уменьшить суммарную ошибку на всем обучающем множестве; поскольку алгоритм обрабатывает наблюдения по одному, общая ошибка на отдельных шагах не обязательно будет убывать.

Алгоритм обучения сети методом обратного распространения ошибки

1. Создаем нейронную сеть либо с помощью мастера *File→ Intelligent Problem Solver*, либо самостоятельно *File →New*.

2. Вторым этапом является непосредственно обучение сети.

2.1. Открываем окно *График ошибки*: *Statistics → Training error graph* (рисунок 4.29).

В представленном диалоговом окне на графике изображается среднеквадратическая ошибка на обучающем множестве на данной эпохе. С помощью расположенных под графиком элементов управления можно менять масштаб изображения, а если график целиком

не помещается в окне, под ним появляются линейки прокрутки. Флажок, расположенный между стрелками масштаба, включает режим автоматического масштабирования (при котором размер графика выбирается так, чтобы он точно помещался в окне).

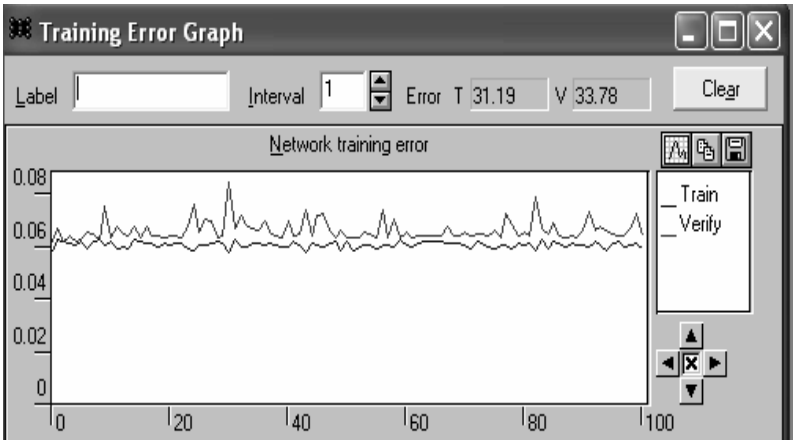


Рисунок 4.29. График ошибки обучения

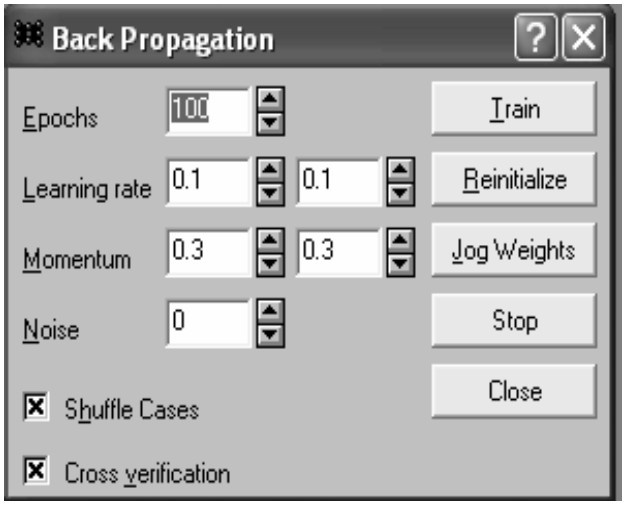


Рисунок 4.30. Обратное распространение

Чтобы сделать сравнение более наглядным, можно рисовать линии разными цветами. Введите значение в поле *Метка* - ***Label*** в окне - *Training Graph,* и тогда следующая линия будет парирована другим цветом.

2.2. Открываем диалоговое окно *Обратное распространение*: *Train→Multilayer Perceptron→Back Propagation* (рисунок 4.30).

2.3. Открываем диалоговое окно ошибки: *Statistics→ Case Errors* (рисунок 4.31).

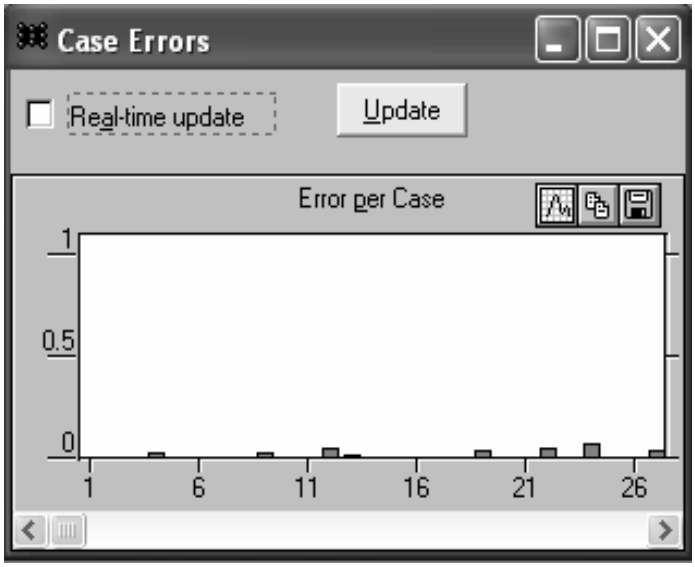


Рисунок 4.31. Ошибки наблюдений

Представленное на рисунке 4.31 диалоговое окно позволяет определить ошибки сети (во время и по результатам обучения).

Замечание. Чтобы обучить нейронную сеть решению какой-либо задачи, необходимо таким образом подправлять веса каждого элемента, чтобы уменьшалась ошибка-расхождение между действительным и желаемым выходом. Для этого необходимо, чтобы нейронная сеть вычисляла производную ошибки по весам. Таким образом, она должна вычислять, как изменяется ошибка при небольшом увеличении или уменьшении каждого веса.

Также имеется возможность следить за тем, как ошибки меняются в процессе обучения - для этого служит функция *Пересчитывать по ходу - Real-time update;* ее нужно активизировать перед запуском алгоритма обратного распространения. Сделав это, вы сможете наблюдать, как алгоритм пытается искать компромисс между обучающими и мешающими наблюдениями.

Все окна открываются и располагаются таким образом, чтобы они не пересекались.

3. Нажимаем кнопку *Обучить – Train,* в диалоговом окне *Обратное распространение - Back Propagation* будет запущен алгоритм обучения. При этом на график (рисунок 4.31) будет выводится ошибка.

4. Повторно нажимаем кнопку *Обучить - Train*, чтобы алгоритм переходил к очередным эпохам.

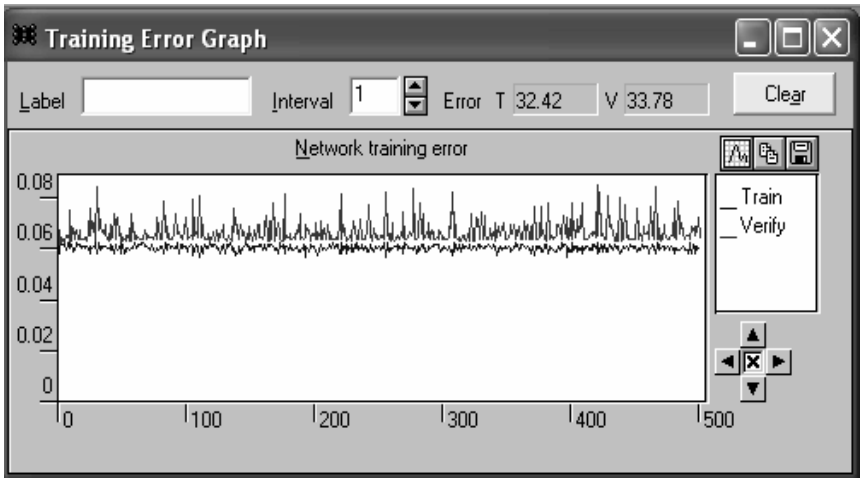


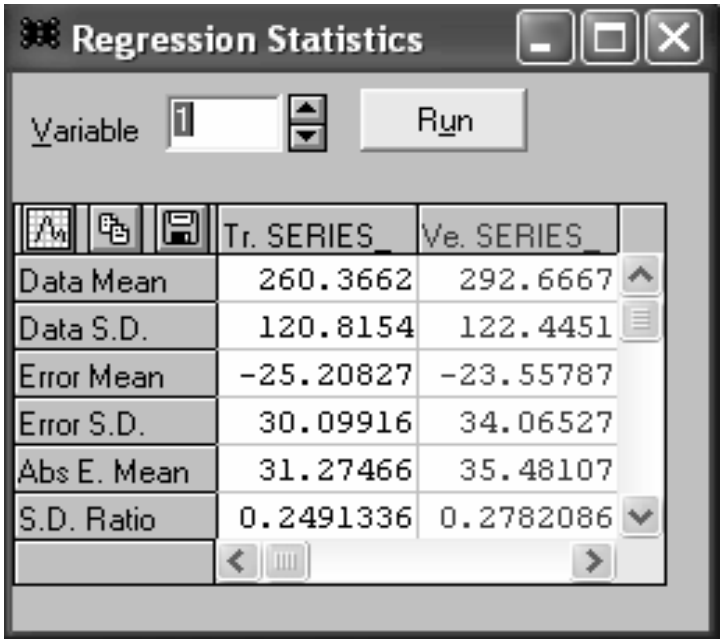
Рисунок 4.32. График ошибки обучения

Причем обращаем внимание на величину ошибки (рисунок 4.31), чем больше число итераций, тем больше величина ошибки. Наилучшую оценку обученной сети можно получить из диалогового окна *Regression Statistics* (рисунок 4.33), здесь анализируется показатель *S.D.Ratio* (отношение стандартного отклонения ошибки к стандартному отклонению данных), если он меньше 0.1, это означает прекрасное качество регрессии. В нашем примере он превышает значение 0.7, что также является достаточно неплохим результатом обучения.

ST Neural Networks автоматически вычисляет среднее и стандартное отклонение обучения и поднаборов проверки, а также вычисляются средние и стандартные отклонения ошибок предсказания.

Кроме того, *ST Neural Networks* показывает стандартный Pearson-R коэффициент корреляции между фактическими и предсказанными значениями. Совершенное предсказание будет иметь коэффи-

циент корреляции 1,0. Корреляция 1,0 не обязательно указывает совершенное предсказание (только предсказание, которое является совершенно линейно коррелированным с фактическими данными), хотя практически коэффициент корреляции - хороший индикатор работы. Это обеспечивает простой способ сравнения работы (выполнения) нашей нейронной сети со стандартным методом наименьших квадратов.



Средние данные – среднее значение переменных.

Data S.D. - стандартное отклонение переменной.

Abs. E. Mean - абсолютное значение ошибки (разница между целевыми и фактическими значениями переменной).

Error S.D. - стандартное отклонение ошибок для переменной .

S.D. Ration - стандартное отношение отклонения.

Корреляция - Pearson-R коэффициент корреляции между целевыми и фактическими значениями.

Рисунок 4.33. Статистика регрессии

5. Если модель не достаточно «хорошо обучена», можно оптимизировать некоторые параметры, от которых зависит работа алгоритма обратного распространения.

5.1. *Эпохи-Epochs.* Задает число эпох обучения, которые проходятся при одном нажатии клавиши Обучить - Train. Значение по умолчанию установлено 100.

5.2. *Скорость обучения-Learning rate.* При увеличении скорости алгоритм работает быстрее, но в некоторых задачах это может привести к неустойчивости. Установим в начале максимальное значение 0,9, но убедившись, что оно не подходит, так как это приводит к неустойчивости (скорее всего из-за того, что данные зашумлены), подбором определяем значение 0,1, наилучшее для нашей сети и процесса обучения.

5.3. *Инерция-Momentum*. Этот параметр ускоряет обучение в ситуациях, когда ошибка мало меняется, а также придает алгоритму дополнительную устойчивость. Вообще этот показатель выбирают небольшим, при высокой скорости. Но для данного алгоритма применяется величина равная 0,9. Для рассматриваемой сети наилучшей была признана величина 0,1.

5.4. Перемешивать наблюдения - *Shuffle Cases*. Данный параметр не позволит алгоритму «застрять», улучшая показатели его работы, так как порядок, в котором наблюдения подаются на вход сети, меняется в каждой эпохе.

6. Если после оптимизации параметров алгоритма обратного распространения ошибки меняются незначительно, то возможно изменить веса, которые меняются автоматически, если нажать клавишу *Reinitialize*, предварительно очистив окно *Training error graph (Clear)*. А затем опять начать процесс обучения, нажав клавишу *Train* (обучение). Стоит также проанализировать и изменившееся диалоговое окно *Regression Statistics* (рисунок 4.35).

7. Последним моментом является запуск нейронной обученной сети, который может осуществляться тремя способами.

7.1. Обработка наблюдений по одному: *Run* → *Run Single Case* → *Case No* (введите номер наблюдения, подлежащего обработке) → нажмите клавишу *Run* и в поле *Output* проследите выходное значение.

7.2. Прогон всего набора данных: *Run* → *Run Data Set* → нажмите клавишу *Run,* и результаты будут выведены в нижней части окна. Наибольший интерес представляет среднеквадратическая ошибка – *RMS error* сети на данном наборе данных.

7.3. Тестирование на отдельном наблюдении: *Run* → *Run One-off Case* →введите входное значение и нажмите клавишу Run, и результаты будут выведены в поле Output.

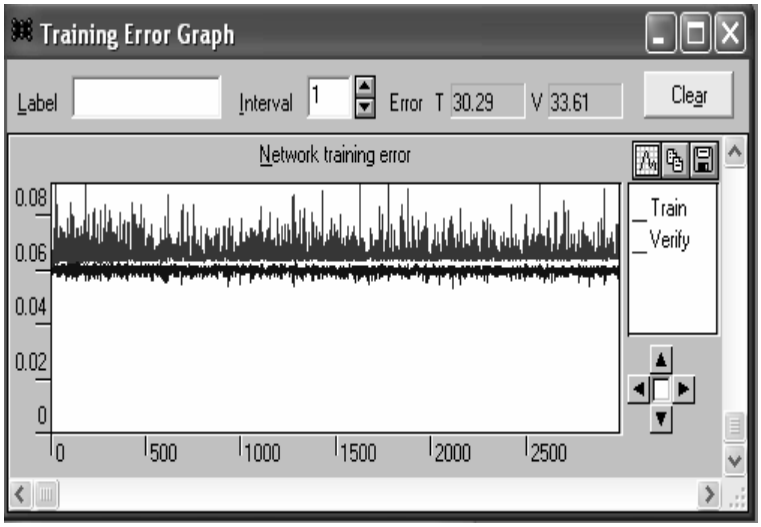
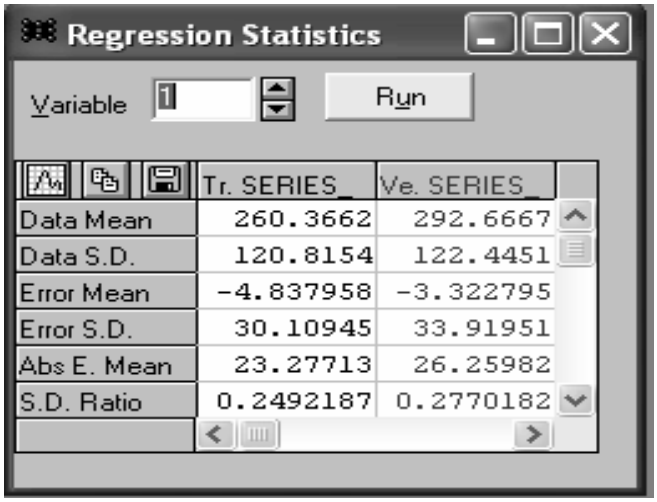


Рисунок 4.34. График ошибки обучения



	Tr. SERIES_	Ve. SERIES_
Data Mean	260.3662	292.6667
Data S.D.	120.8154	122.4451
Error Mean	-4.837958	-3.322795
Error S.D.	30.10945	33.91951
Abs E. Mean	23.27713	26.25982
S.D. Ratio	0.2492187	0.2770182

Рисунок 4.35. Статистика регрессии

4.5.2. Обучение с помощью метода Левенберга-Маркара

Метод Левенберга-Маркара считается одним из лучших алгоритмов нелинейной оптимизации, известных на сегодняшний день, и

это один из самых известных алгоритмов обучения нейронных сетей. Он имеет два существенных ограничения.

1. Алгоритм можно применять только для относительно небольших сетей (в пределах нескольких сотен нейронов);

2. Алгоритм годится только для сетей с одним выходом.

Замечание. Применение алгоритма Левенберга-Маркара связано с одной особенностью. Оценивая очередной вариант сети, алгоритм отвергает его, если при этом увеличилась ошибка. Представленная программа изображает на графике ошибки рассмотренных ею вариантов сетей, при этом график обучения может оказаться очень «зазубренным». Это, однако, не означает плохой работы алгоритма, поскольку предыдущий вариант сети хранится в памяти до тех пор, пока он не будет превзойден.

Вначале создадим с помощью советчика - *Intelligent Problem Solver* многослойный персептрон с тремя слоями и в промежуточном слое возьмем шесть элементов. Затем командами: *File → New → Network*, видоизменяем сеть, создав в промежуточном слое шесть элементов (установив в поле *Layer* 2 цифру 6).

В итоге получим следующую сеть, представленную на рисунке 4.36.

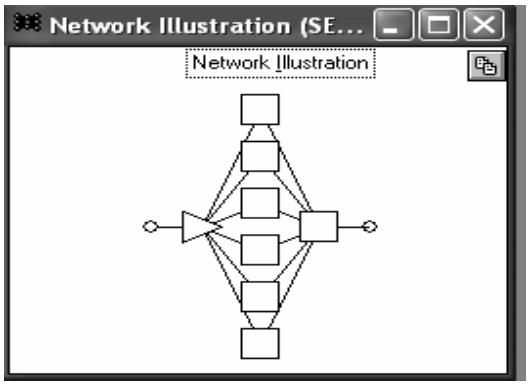


Рисунок 4.36. Иллюстрация сети

Для обучения сети с помощью метода Левенберга-Маркара необходимо выполнить следующие действия:

- *Statistics → Training error graph;*
- *Statistics → Case Errors;*
- *Statistics → Regression Statistics;*
- *Train → Multilayer Perceptron → Levenberg-Marquardt.*

В результате будут открыты четыре окна, с соответствующими названиями.

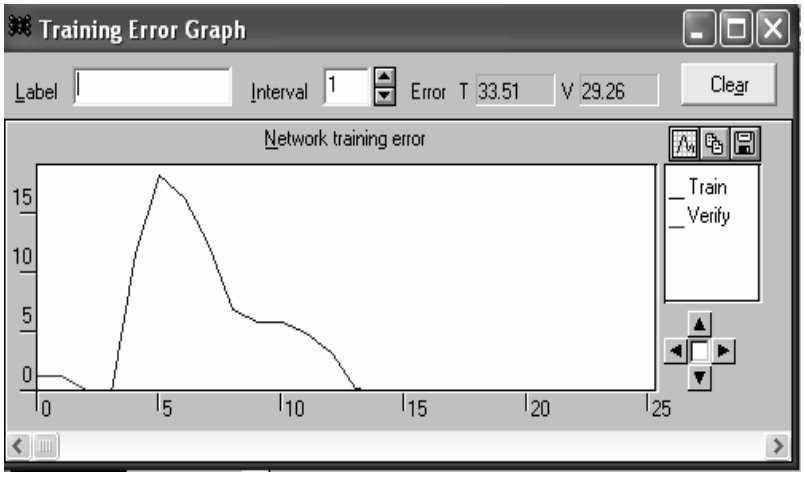


Рисунок 4.37. График ошибки обучения

Рисунок 4.38. Обучение методом Левенберга-Маркара

Вначале окно *Tranding error graph* будет пустым, то есть на нем не будет отражен график изменения ошибки. Обучение заключается в том, что в окне *Levenberg-Marquardt* необходимо нажать клавишу *Train*, после этого лишь начнется обучение сети и в окне *Tranding error graph* появится график изменения ошибки, причем в верхней части этого окна указывается ошибка на обучающей и верификационной части выборки (их значения не должны сильно различаться).

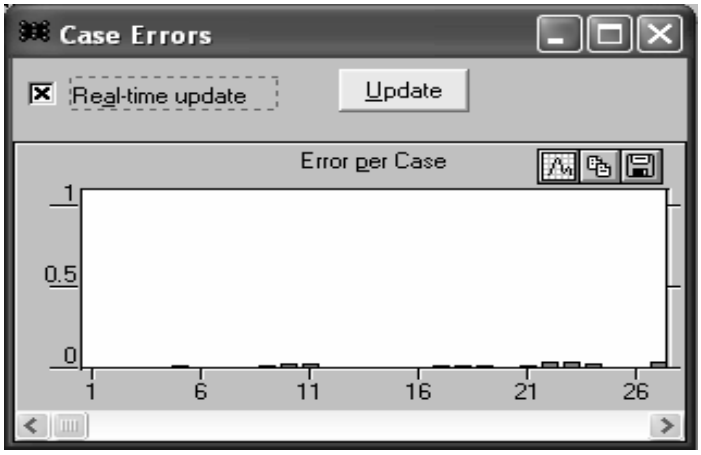
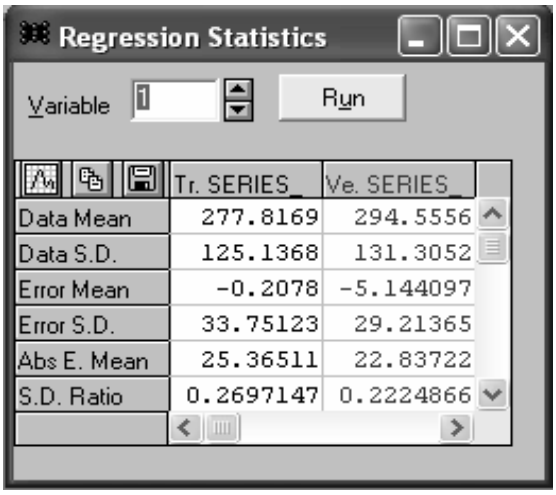


Рисунок 4.39. Ошибки наблюдений



	Tr. SERIES_	Ve. SERIES_
Data Mean	277.8169	294.5556
Data S.D.	125.1368	131.3052
Error Mean	-0.2078	-5.144097
Error S.D.	33.75123	29.21365
Abs E. Mean	25.36511	22.83722
S.D. Ratio	0.2697147	0.2224866

Рисунок 4.40. Статистика регрессии

Кроме того, динамику изменения ошибки можно проследить в окне *Case Errors*. Но более точную оценку обученной сети и оценку результата работы сети в задаче регрессии можно получить из диалогового окна *Regression Statistics*. Здесь подсчитывают среднее и стандартное отклонение для выходных переменных и ошибки сети, а также отношение стандартного отклонения ошибки к стандартному отклонению данных.

Причем важно заметить, если величина Отношение ст.откл. – *S.D.Ratio* меньше 0,1, это означает прекрасное качество регрессии. В

нашем случае эта величина превышает 0,2, что также является хорошим показателем обучения нейронной сети.

4.5.3. Алгоритм выполнения обучения сети с помощью метода Левенберга-Маркара

1. Создаем нейронную сеть. Вначале создадим с помощью мастера - *Intelligent Problem Solver* многослойный персептрон с тремя слоями и в промежуточном слое возьмем шесть элементов. С помощью мастера, строим основную нейронную сеть, а затем командами: *File-New-Network*, видоизменяем сеть, создав в промежуточном слое шесть элементов (установив в Layer 2 цифру 6).

2. Затем начнем обучение с помощью метода Левенберга-Маркара. Для этого выполняем следующие команды:

– *Statistics→Tranding error graph*;
– *Statistics→ Case Errors*;
– *Statistics→ Regression Statistics*;
– *Train→Multilayer Perceptrons→ Levenberg-Marquardt*.

3. В диалоговом окне *Levenberg-Marquardt* нажимаем клавишу *Train* (обучение) несколько раз (то есть обучение происходит на различных эпохах, если же нажать только один раз, то получим обучение лишь на одной эпохе).

4. Если же ошибка на обучаемой и верификационной выборке очень велика, то стоит изменить весовые коэффициенты. Они автоматически поменяются, если нажать клавишу *Reinitialize* , предварительно очистив окно *Training error graph* (*Clear*). А затем опять начинаем процесс обучения, нажав клавишу *Train* (обучение).

5. Анализируем полученные данные по выведенным диалоговым окнам. Что касается диалогового окна *Regression Statistics*, то анализируется показатель *S.D.Ratio*, если он меньше 0,1, это означает прекрасное качество регрессии.

4.6. Генетические алгоритмы отбора входных данных

Один из самых трудных вопросов, который приходится решать разработчику нейросетевых приложений, - это вопрос о том, какие данные взять в качестве входных для нейронной сети. Этот вопрос сложен в силу сразу нескольких причин.

- Чаще всего при применении нейронных сетей в реальных задачах заранее не бывает точно известно, как прогнозируемый показатель связан с имеющимися данными. Поэтому собирают больше разнообразных данных, среди которых предположительно есть и важные, и такие, чья ценность неизвестна или сомнительна.

- В задачах нелинейной природы среди параметров могут быть взаимозависимые и избыточные. Например, может случиться так, что каждый из двух параметров сам по себе ничего не значит, но оба они вместе несут чрезвычайно полезную информацию. Это же может относиться к совокупности из нескольких параметров. Сказанное означает, что попытки ранжировать параметры по степени важности могут быть неправильными в принципе.

- Из-за «проклятия размерности» иногда лучше просто убрать некоторые переменные, в том числе и несущие значимую информацию, чтобы хоть как-то уменьшить число входных переменных, а значит и сложность задачи, и размеры сети. Вопреки здравому смыслу, такой прием иногда действительно улучшает способность сети к обобщению (Bishop, 1995) [10].

Единственный способ получить полную гарантию того, что входные данные выбраны наилучшим образом, состоит в том, чтобы перепробовать все возможные варианты входных наборов данных и архитектур сетей и выбрать из них наилучший. На практике это сделать невозможно из-за огромного количества вариантов.

Можно попытаться поэкспериментировать в среде пакета *ST Neural Networks* - последовательно строить сети с различными наборами входных переменных, чтобы постепенно составить себе картину того, какие же входные переменные действительно нужны. Можно воспользоваться методом регуляризации весов по Вигенду *(Weigend Weight Regiilariation)* и в окне *Редактор сети - Network Editor* посмотреть, у каких входных переменных выходящие веса сделаны нулевыми (это говорит о том, что данная переменная игнорируется).

Самое действенное средство решения данного вопроса в пакете *ST Neural Networks* -- это *Генетический алгоритм отбора входных данных - Genetic*, *Algorithm Input Selection.* Этот алгоритм выполняет большое число экспериментов с различными комбинациями входных данных, строит для каждой из них вероятностную либо обобщенно-регрессионную сеть, оценивает ее результаты и использует их в дальнейшем поиске наилучшего варианта.

Генетические алгоритмы являются очень эффективным инструментом поиска в комбинаторных задачах как раз такого типа (где

требуется принимать ряд взаимосвязанных решений «да/нет»). Этот метод требует большого времени счета (обычно приходится строить и проверять многие тысячи сетей), однако реализованная в пакете *ST Neural Networks* его комбинация с быстро обучающимися сетями типа PNN/GRNN позволяет ускорить его работу настолько, насколько это вообще возможно [244].

Выполнение алгоритма

Прогонять этот алгоритм мы будем на модифицированном варианте усеченного набора данных про ирисы, но прежде сделаем хотя бы одну переменную *Выходной - Output* через диалоговое окно Выбор типов переменных– *Set Variable Types* → *Типы –Type*→ *Переменные – Variables*→ меню Правка – *Edit.*

В состав программы входит файл данных *Iris.sta* или *IrisDat.sta,* относящийся к задаче Фишера о классификации ирисов.

Рассматривается три вида цветков ириса: *Setosa, Versicolor* и *Virginica.* Всего имеется по 50 экземпляров каждого вида, и для каждого из них измерены четыре величины: длина и ширина чашелистика, длина и ширина лепестка. Цель состоит в том, чтобы научиться предсказывать тип нового (пока не известного) цветка по данным этих измерений.

Эта задача интересна сразу но нескольким причинам. Во-первых, один из классов *(Setosa)* линейно отделим от двух других. Оставшиеся два разделить гораздо труднее, и кроме того, имеется несколько экземпляров, выбивающихся из общей картины, которые легко отнести к другому классу. Такие выбросы являются хорошей проверкой надежности работы сети на сложных данных.

Представление нескольких классов

Загрузите файл *Iris.sta* или *IrisDat.sta* в пакет *ST Neural Networks* и взгляните на данные. Здесь имеется четыре входных переменных и одна выходная. Выходная переменная — номинальная с тремя возможными значениями: *Iris = {Setosa, Versicol, Virginic}.*

Формирование уменьшенного набора данных

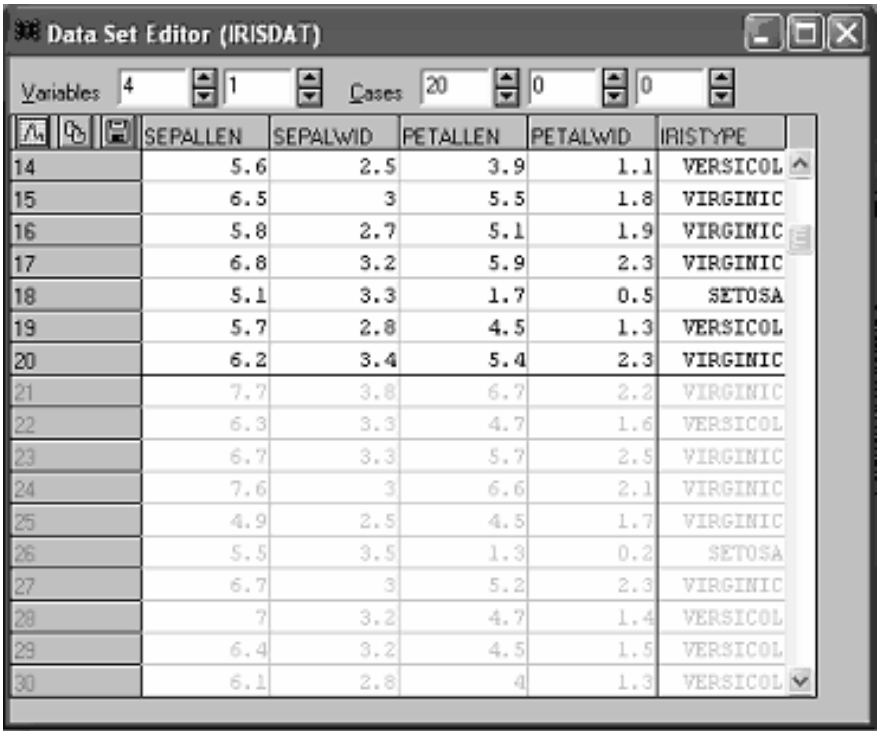
Данные про ирисы содержа 150 наблюдений - это довольно много, и сеть будет обучаться медленно. Для учебных целей давайте сократим объем данных до 60 наблюдений. При этом мы заодно по-

знакомимся с некоторыми возможностями *Редактора данных - Data Set Editor.*

Кроме того, мы разобьем данные на две группы: обучающее множество *(Training Set)* будет использоваться для обучения сети, а контрольное множество *(Verification Set)* - для проверки качества ее работы.

Чтобы сформировать сокращенный набор данных, сделайте следующее:

1. Откройте набор данных *Iris.sta* или *IrisDat.sta* и вызовите *Редактор данных - Data Editor.*



	SEPALLEN	SEPALWID	PETALLEN	PETALWID	IRISTYPE
14	5.6	2.5	3.9	1.1	VERSICOL
15	6.5	3	5.5	1.8	VIRGINIC
16	5.8	2.7	5.1	1.9	VIRGINIC
17	6.8	3.2	5.9	2.3	VIRGINIC
18	5.1	3.3	1.7	0.5	SETOSA
19	5.7	2.8	4.5	1.3	VERSICOL
20	6.2	3.4	5.4	2.3	VIRGINIC
21	7.7	3.8	6.7	2.2	VIRGINIC
22	6.3	3.3	4.7	1.6	VERSICOL
23	6.7	3.3	5.7	2.5	VIRGINIC
24	7.6	3	6.6	2.1	VIRGINIC
25	4.9	2.5	4.5	1.7	VIRGINIC
26	5.5	3.5	1.3	0.2	SETOSA
27	6.7	3	5.2	2.3	VIRGINIC
28	7	3.2	4.7	1.4	VERSICOL
29	6.4	3.2	4.5	1.5	VERSICOL
30	6.1	2.8	4	1.3	VERSICOL

Рисунок 4.41. Сокращение объема данных

2. Выделите наблюдения 21-50. Для этого прокрутите таблицу так, чтобы стало видно наблюдение номер 20, щелкните на метке его строки, и либо протащите мышь до нижней границы диапазона, либо выделите его клавишами СТРЕЛКА вниз и PAGE DOWN при нажатой клавише SHIFT.

3. Нажмите правую кнопку мыши и выберите из контекстного меню пункт *Не учитывать - Ignore.* Отмеченные наблюдения будут выделены серым цветом, и программа *ST Neural Networks* не будет использовать их при обучении.

4. Проделайте все то же самое для наблюдений 71-100 и 121-150 (рисунок 4.41).

5. Поменяйте число наблюдений в обучающем множестве - (оно указано в поле *Обучающее - Training)* с 60 на 30. Программа *ST Neural Networks* автоматически отнесет оставшиеся наблюдения к контрольному множеству, так что теперь у нас будет 30 обучающих и 30 контрольных наблюдений.

6. Нажмите кнопку *Перемешать - Shuffle* — обучающие и контрольные наблюдения будут взяты случайным образом среди всех 60 наблюдений. Вы заметите это, взглянув на таблицу. Обучающие наблюдения показаны черным цветом, а контрольные - красным.

7. Затем переместите указатель мыши на линию, разделяющую заголовки столбцов переменных *SLENGTH* и *SWIDTH.* Указатель превратится в двустороннюю стрелку. Щелкните мышью - при этом появится голубая полоса вставки, затем нажмите три раза клавишу ENTER - в файл будут добавлены три новые переменные. Так как они очевидным образом не несут в себе никакой информации, следует ожидать, что *Генетический алгоритм отбора входных данных - GA Input Selection* удалит их.

Откроем диалоговое окно *Input Feature Selection* с помощью меню *Обучение-Дополнительно - Train-Auxiliary* (рисунок 4.42).

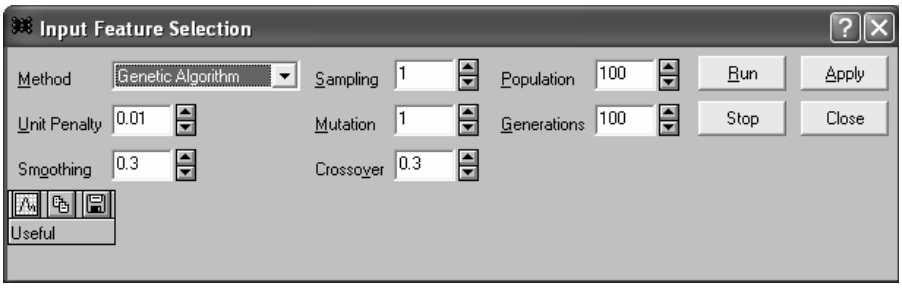


Рисунок 4.42. Диалоговое окно *Input Feature Selection*

Это окно содержит много управляющих параметров. Среди них - *Популяция - Population* и *Поколения - Generations, Выборка – Sampling, Скорость мутации - Mutation Rate* и *Скорость скрещивания - Crossover Rate.* Если вы не знакомы с генетическими алгоритмами,

то не стоит менять два последних параметра (рекомендуется взять значения по умолчанию). Параметры *Популяция - Population* и *Поколения - Generations* определяют, как много усилий алгоритм затрачивает на поиск.

Схема работы генетического алгоритма такова. Берется случайный набор, популяция битовых строк (в нашем случае отдельный бит, соответствующий каждому входу, показывает, учитывать или нет соответствующую входную переменную) и оценивается степень их пригодности (т.е. качество получаемых решений). Затем плохие строки исключаются из рассмотрения, а из оставшихся порождаются новые строки с помощью искусственных генетических операций мутации и скрещивания. Таким образом, возникает новая популяция, и весь процесс повторяется, порождая все новые поколения, а в конце его отбирается наилучший экземпляр. Параметр *Популяция - Population* задает объем популяции индивидуумов, а параметр *Поколения - Generations* определяет, сколько раз будет повторен цикл отбора порождения-оценки. Произведение этих двух чисел равно общему числу операций оценивания, которые алгоритм должен будет выполнить, и каждое оценивание включает построение PNN или GRNN сети и ее тестирование на контрольном множестве.

При построении PNN или GRNN сети необходимо выбрать коэффициент сглаживания *(Smoothing)*. В общем случае следует самостоятельно провести ряд экспериментов со всей совокупностью входных переменных, строя PNN или GRNN сети с различными коэффициентами сглаживания, и выбрать подходящее значение. К счастью, сети PNN и GRNN не слишком чувствительны к точному выбору коэффициента сглаживания, и в нашем случае вполне подойдет значение по умолчанию.

Как уже говорилось, иногда бывает полезно уменьшить число входов даже ценой некоторой потери точности, поскольку это улучшает способности сети к обобщению и уменьшает размер сети и время счета. Можно создать дополнительный стимул к исключению лишних переменных, назначив «штраф» за элемент *(Unit Penalty)*. Это число будет умножаться на количество элементов, и результат будет прибавляться к уровню ошибки при оценке качества сети. Таким образом, будут штрафоваться большие по размеру сети. Обычно значения этого параметра (если он используется) берутся в интервале 0,01-0,001. В нашей задаче дополнительные переменные не несут никакой информации и действительно будут ухудшать качество се-

ти, поэтому нет необходимости специально задавать еще и штраф за элемент.

Нажмите кнопку *Запуск - Run.* При выполнении генетического алгоритма с параметрами по умолчанию он проделает 10000 оцениваний (популяция объемом 100 и 100 поколений). Однако в нашей задаче имеется всего семь кандидатов во входные переменные (четыре настоящих и три добавленные), поэтому число всевозможных комбинаций равно всего 128 (два в седьмой степени). Программа *ST Neural Networks* сама обнаружит это обстоятельство и вместо описанных действий выполнит оценивание полным перебором вариантов (соответствующая информация будет выдана в строке сообщений).

По окончании работы алгоритма откроется окно с таблицей, в которой будет указано, какие переменные были признаны полезными, а какие нет (соответственно *Да - Yes* или *Нет - No*) (рисунок 4.43). Переменные, которые не рассматривались как кандидаты во входной набор (в данной задаче - выходная переменная) будут помечены как неучитываемые. Если вы все сделали правильно, алгоритм отберет настоящие переменные задачи и отбросит вновь добавленные.

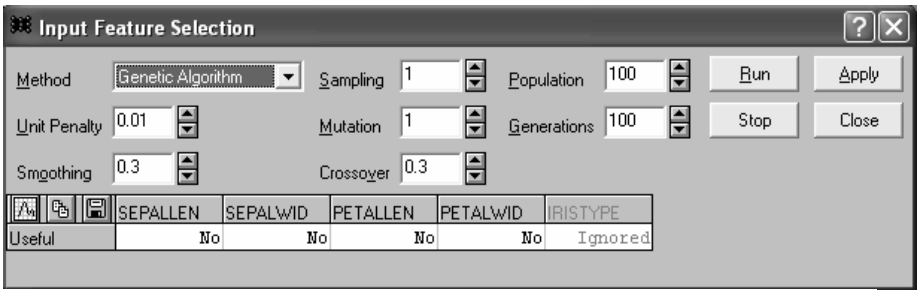


Рисунок 4.43. Окончание работы алгоритма

Нажмите кнопку *Применить - Apply.* Программа *ST Neural Networks* применит найденный шаблон к исходным данным, и у тех переменных, которые были отвергнуты алгоритмом, изменит тип на *Неучитываемая - Ignore.*

Теперь сделаем параметр *Штраф за элемент - Unit Penalty* равным 0,01 и снова нажмем кнопку *Запуск - Run.* На этот раз мы побуждаем алгоритм уменьшать число входов, даже ценой некоторого увеличения ошибки. Конкретный результат будет зависеть от того, какие наблюдения были взяты в обучающее и контрольное множе-

ства. Скорее всего окажется, что переменные *PLENGTH* и *PWIDTH* будут отобраны, а переменные *SLENGTH* и *SWIDTH* - отвергнуты. Экспериментируя с различными значениями штрафа за элемент, вы сможете приблизительно упорядочить входные переменные по степени важности.

4.7. Пример применения нейронных сетей в задачах прогнозирования и проблемы идентификации моделей прогнозирования на нейронных сетях

Жесткие статистические предположения о свойствах временных рядов ограничивают возможности классических методов прогнозирования. В данной ситуации адекватным аппаратом для решения задач прогнозирования могут служить специальные искусственные нейронные сети (НС) [22, 116].

В этой связи возникает особо важная задача определения структуры и типа прогнозирующей нейронной сети. На сегодняшний день нет алгоритма или метода, позволяющего дать однозначный ответ на этот вопрос. Однако предложены способы настройки числа нейронов в процессе обучения, которые обеспечивают построение нейронной сети для решения задачи и дают возможность избежать избыточности. Эти способы настройки можно разделить на две группы: конструктивные алгоритмы и алгоритмы сокращения.

В основе алгоритмов сокращения лежит принцип постепенного удаления из нейронной сети синапсов и нейронов. В начале работы алгоритма обучения с сокращением число нейронов скрытых слоях сети заведомо избыточно.

Алгоритмы сокращения можно рассматривать как частный случай алгоритмов контрастирования. Существуют два подхода к реализации алгоритмов сокращения метод штрафных функций и метод проекций. Для реализации первого в целевую функцию алгоритма обучения вводится штрафы за то, что значения синаптических весов отличны от нуля; пример штрафа

$$C = \sum_{i=1}^{N} \sum_{j=1}^{M} w_{ij}^2 \qquad (4.2)$$

где w_{ij}^2 – синаптический вес, i – номер нейрона, j – номер входа, N – число нейронов, M – размерность входного сигнала нейронов;

Метод проекций реализуется следующим образом. Синаптический вес обнуляется, если его значение попало в заданный диапазон

$$w_{ij} = \begin{cases} 0, -\varepsilon \le w_{ij} \le \varepsilon \\ w_{ij}, w_{ij} < -\varepsilon, w_{ij} > \varepsilon \end{cases} \tag{4.3}$$

где ε – некоторая константа.

Алгоритмы сокращения имеют по крайней мере два недостатка:

1. Нет методики определения числа нейронов скрытых слоев, которое является избыточным, поэтому перед началом работы алгоритма нужно угадать это число.

2. В процессе работы алгоритма сеть содержит избыточное число нейронов, поэтому обучение идет медленно.

Предшественником конструктивных алгоритмов можно считать методику обучения многослойных сетей, включающую в себя следующие шаги:

1. Выбор начального числа нейронов в скрытых слоях.

2. Инициализация сети, то есть присваивание синаптическим весами смещениям сети случайных значений из заданного диапазона.

3. Обучение сети по заданной выборке.

4. Завершение в случае успешного обучения; если сеть обучить не удалось, то число нейронов увеличивается, и повторяются шаги со второго по четвертый.

В конструктивных алгоритмах число нейронов в скрытых слоях также изначально мало и постепенно увеличивается. В отличие от описанной методики, в конструктивных алгоритмах сохраняются навыки, приобретенные сетью до увеличения числа нейронов.

Конструктивные алгоритмы различаются правилами задания значений параметров в новых – добавленных в сеть – нейронах:

1. Значения параметров – случайные числа из заданного диапазона;

2. Значения синаптических весов нового нейрона определяются путем расщепления (splitting) одного из старых нейронов.

Первое правило не требует значительных вычислений, однако его использование приводит к некоторому увеличению значения функции ошибки после каждого добавления нового нейрона. В результате случайного задания значений параметров новых нейронов

может появиться избыточность в числе нейронов скрытого слоя. Расщепление нейронов лишено двух указанных недостатков.

Самым большим недостатком алгоритма является экспоненциальный рост времени вычислений при увеличении размерности сети. Для преодоления указанного недостатка предлагается упрощенный алгоритм расщепления, который не требует значительных вычислений.

Кроме описанных способов выбора нейронов для расщепления, может быть использован анализ чувствительности, в процессе которого строятся матрицы Гессе - матрицы вторых производных функции ошибки по параметрам сети. По величине модуля второй производной судят о важности значения данного параметра для решения задачи. Параметры с малыми значениями вторых производных обнуляют. Анализ чувствительности имеет большую вычислительную сложность и требует много дополнительной памяти.

4.7.1. Сравнительный анализ нейронных сетей

Актуальность данной тематики продиктована поиском адекватных моделей нейронных сетей (НС), определяемые типом и структурой НС, для задач прогнозирования. В ходе исследования установлено, что радиальные базисные сети (RBF) обладают рядом преимуществ перед сетями типа многослойных персептрон (MLP) [121, 123]. Во-первых, они моделируют произвольную нелинейную функцию с помощью одного промежуточного слоя. Тем самым отпадает вопрос о числе слоев. Во-вторых, параметры линейной комбинации в выходном слое можно полностью оптимизировать с помощью известных методов моделирования, которые не испытывают трудностей с локальными минимумами, мешающими при обучении MLP. Поэтому сеть RBF обучается очень быстро (на порядок быстрее MLP) [85,107].

С другой стороны, до того как применять линейную оптимизацию в выходном слое сети RBF, необходимо определить число радиальных элементов, положение их центров и величины отклонений. Для устранения этой проблемы предлагается использовать автоматизированный конструктор сети, который выполняет за пользователя основные эксперименты с сетью.

Другие отличия работы RBF от MLP связаны с различным представлением пространства модели: «групповым» в RBF и «плоскост-

ным» в MLP. Опыт показывает, что для правильного моделирования типичной функции сеть RBF, с ее более эксцентричной поверхностью отклика, требует несколько большего числа элементов. Следовательно, модель, основанная на RBF, будет работать медленнее и потребует больше памяти, чем соответствующий MLP (однако она гораздо быстрее обучается, а в некоторых случаях это важнее).

С «групповым» подходом связано и неумение сетей RBF экстраполировать свои выводы за область известных данных. При удалении от обучающего множества значение функции отклика быстро спадает до нуля. Напротив, сеть MLP выдает более определенные решения при обработке сильно отклоняющихся данных, однако, в целом, склонность MLP к некритическому экстраполированию результата считается его слабостью. Сети RBF более чувствительны к «проклятию размерности» и испытывают значительные трудности, когда число входов велико.

Для оценки точности и адекватности результатов прогнозирования, а также структуры нейронной сети использовались следующие статистические показатели:

1. Data Mean. – среднее значение целевой выходной переменной;
2. Data S.D. – среднеквадратическое отклонение целевой выходной переменной;
3. Error Mean – средняя ошибка выходной переменной (остаток между целевой и реальной переменной);
4. Abs. E. Mean – средняя абсолютная ошибка (разница между целевой и реальной выходной переменной);
5. Error S.D. – стандартное отклонение ошибки выходной переменной;
6. S.D. Ratio – среднеквадратическое отклонение ошибок выходной переменной;
7. Correlation – коэффициент корреляции Спирмена, вычисленный между целевым вектором и реальным выходным вектором.

Исследования проводились в пакете STATISTICA Neural Networks 4.0. Используются данные биржи «ММВБ» в период с 29.05.1997 по 24.06.2003.

На рисунке 4.44 показана динамика курса акций российской компании ОАО «РАО ЕЭС».

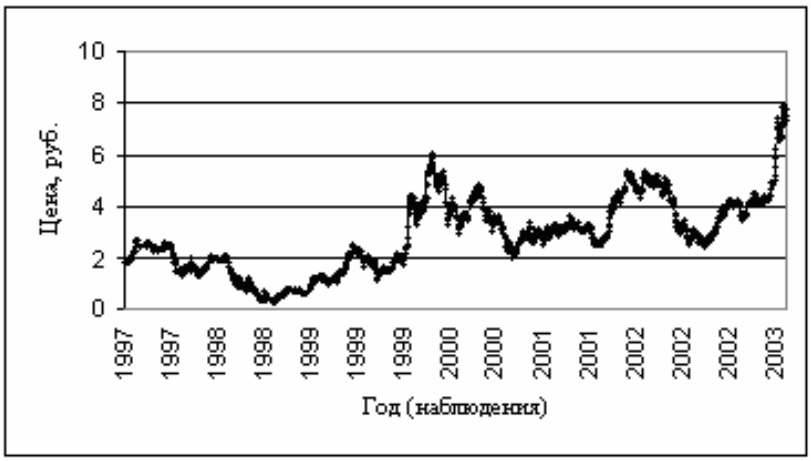


Рисунок 4.44. Динамика курса акций российской компании ОАО «РАО ЕЭС» в период с 29.05.1997 по 24.06.2003 гг.

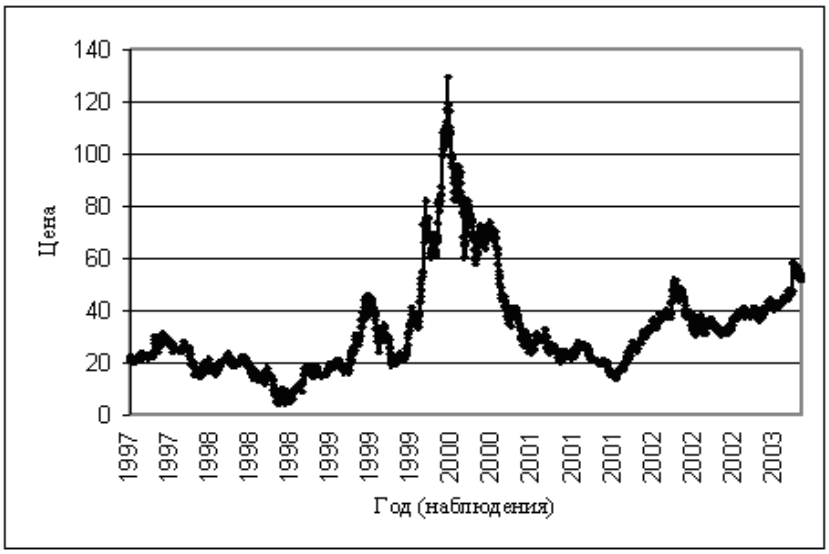


Рисунок 4.45. Динамика курса акций российской компании ОАО «Ростелеком» (1536 наблюдений) в период с 29.05.1997 по 24.06.2003 гг.

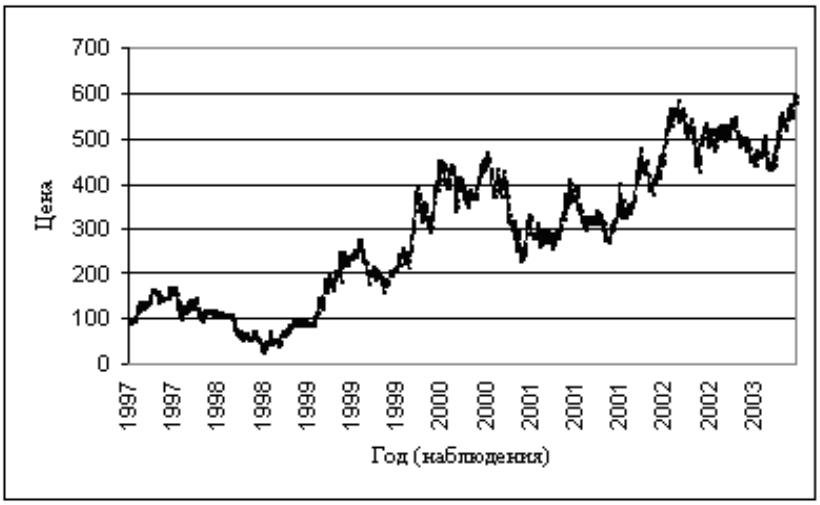


Рисунок 4.46. Динамика курса акций российской компании ОАО «Лукоил» (1511 наблюдений) в период с 29.05.1997 по 24.06.2003 гг.

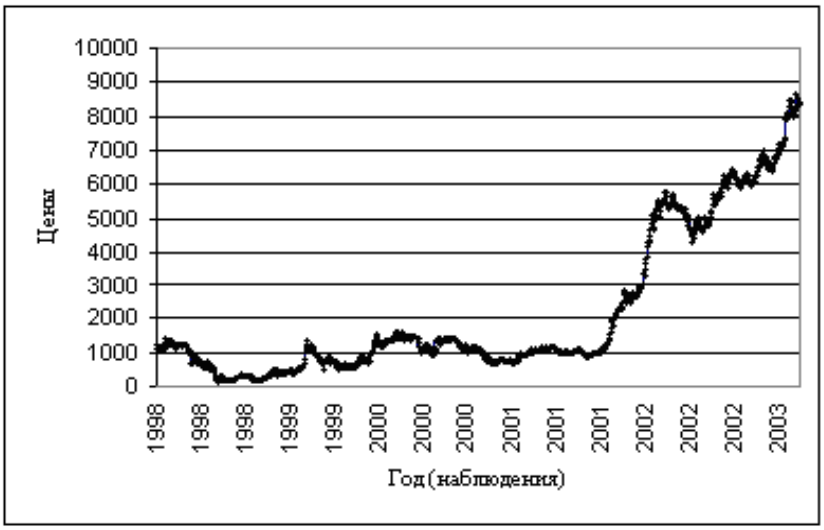


Рисунок 4.47. Динамика курса акций Сбербанка (1304 наблюдений) в период с 29.05.1997 по 24.06.2003 гг.

Сначала проведем исследования для курса акций ОАО «РАО ЕЭС». Каждая таблица показывает найденные типы нейронных

структур для исследуемого временного ряда. В первом столбце таблиц стоит значение лага, с которым данные подаются на вход НС. Во втором столбце указано количество проведенных испытаний, следствием которых стал выбор наилучшей, по всем характеристикам, НС.

Таблица 4.1. Нейронные сети типа RBF

Лаг	Количество испытаний	Параметры отобранной НС Тип НС Inputs: Hidden:Outputs	Вид лучшей НС	Характеристики
1	16	RBF 1:7:1 TPerf, VPerf, TePerf 0,1198; 0,1174; 0,1321		Найден адекватный тип сети (регрессионное отношение 0,08168, корреляция 0.995359, среднеквадратическая ошибка предсказания 0,1172432)
3	24	RBF 3:14:1 TPerf, VPerf, TePerf 0,1234; 0,1215; 0,1315		Найден адекватный тип сети (регрессионное отношение 0.089525, корреляция 0.995992, среднеквадратическая ошибка предсказания 0.1199467)
5	21	RBF 5:14:1 TPerf, VPerf, TePerf 0,1241; 0,195; 0,1354		Найден адекватный тип сети (регрессионное отношение 0.095182, корреляция 0.995465, среднеквадратическая ошибка предсказания 0.1403166)

В третьем столбце указаны основные характеристики НС: Тип сети RBF или MPL; внутренняя структура (например 1:7:1 означает, что сеть имеет один входной нейрон, 7 нейронов скрытого слоя и один выходной нейрон), среднеквадратическая ошибка предсказания (TPerf, VPerf, TePerf) для обучающей, волидационной и тестовой выборок.

В четвертом столбце таблиц показан внешний вид НС. В пятом столбце указаны дополнительные характеристики НС, такие как регрессионное отношение, корреляция, среднеквадратическая ошибка предсказания НС.

Таблица 4.2. Нейронные сети типа MLP (трехслойная сеть)

Лаг	Количество испытаний	Параметры отобранной НС Тип НС Inputs: Hidden:Outputs	Вид лучшей НС	Характеристики
1	16	MLP 1:10:1 TPerf, VPerf, TePerf 0,1221; 0,1193; 0,1328		Найден адекватный тип сети (регрессионное отношение 0.083055, корреляция 0.996550, среднеквадратическая ошибка предсказания 0.1191367)
3	20	MLP 3:12:1 TPerf, VPerf, TePerf 0,127; 0,125; 0,134		Найден адекватный тип сети (регрессионное отношение 0.085872, корреляция 0.996326, среднеквадратическая ошибка предсказания 0.1253478)
5	25	MLP 5:7:1 TPerf, VPerf, TePerf 0,1227; 0,1203; 0,1313		Найден адекватный тип сети (регрессионное отношение 0.095182, корреляция 0.995465, среднеквадратическая ошибка предсказания 0.1403166)

Можно сделать следующие выводы. Каждый из двух описанных подходов имеет свои достоинства и недостатки. Действие радиальных функций очень локально, в то время как при линейном подходе охватывается все пространство входов. Поэтому, как правило, RBF-сети имеют больше элементов, чем MLP-сети, однако MLP может делать необоснованные обобщения в ситуациях, когда ему попадается набор данных, непохожий ни на какие наборы из обучающего множества, в то время как RBF в таком случае всегда будет выдавать почти нулевой отклик [91, 92].

Таблица 4.3. Нейронные сети типа MLP (четырехслойная сеть)

Лаг	Количество испытаний	Параметры отобранной НС Тип НС Inputs: Hidden:Outputs	Вид лучшей НС	Характеристики
1	16	MLP 1:13:13:1 TPerf, VPerf, TePerf 0,1227; 0,1197; 0,133		Найден адекватный тип сети (регрессионное отношение 0.083461, корреляция 0.996511, среднеквадратическая ошибка предсказания 0.1180339)
3	16	MLP 3:13:8:1 TPerf, VPerf, TePerf 0,1233; 0,1207; 0,1338		Найден адекватный тип сети (регрессионное отношение 0.084064, корреляция 0.996471, среднеквадратическая ошибка предсказания 0.1207132)
5	16	MLP 5:4:2:1 TPerf, VPerf, TePerf 0,1227; 0,1194; 0,1325		Найден адекватный тип сети (регрессионное отношение 0.083177, корреляция 0.996543, среднеквадратическая ошибка предсказания 0,1194306)

Анализируя результаты предварительных исследований таблиц 4.1 – 4.3 можно заключить, что сеть типа RBF (1:7:1) таблица 4.1 имеет большее предпочтение. Об этом можно судить по регрессионному отношению и среднеквадратической ошибке предсказания. TPerf, VPerf, TePerf показывает ошибку, которая вычисляется как корень квадратный из среднего значения ошибки на каждом шаге обучения для обучающей, верификационной и тестовой выборок.

4.7.2. Исследование нейросетевых структур для курсов акций ОАО «Ростелеком», ОАО «Лукойл», «Сбербанк»

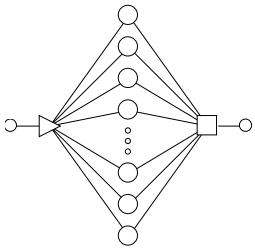
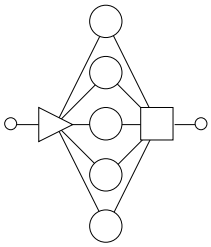
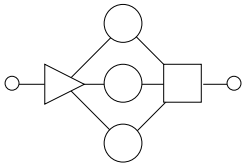
Исследования проводились для рассмотренных временных рядов. Все результаты исследований также сведены в таблицы. В таблице 4.4 показаны результаты поиска оптимальных нейросетевых

структур для курсов акций ОАО «Ростелеком», ОАО «Лукоил», «Сбербанк». Исследования проводились для сетей типа RBF.

В таблицах 4.5 – 4.6 показаны результаты поиска оптимальных нейросетевых структур для тех же временных рядов. Исследования проводились для сетей типа MLP (трех и четырехслойная сеть). Известно, что поиск типа нейронной сети и структуры достаточно трудоемкая процедура, поэтому решалась промежуточная задача определения начального «прототипа», после определения которого, велось экспериментальное уточнение дальнейшей структуры НС.

Выводы по результатам вычислительных экспериментов: Анализируя таблицы 4.4 – 4.6, для курса акций ОАО «Ростелеком» выбрана оптимальная структура нейронной сети со следующими показателями: тип – MLP; трехслойная структура: 3 входных нейронов – 6 нейронов скрытого слоя – 1 выходной нейрон (таблица 4.5). Однако возможно использование сети типа – MLP; четырехслойная структура: 1 входных нейрона – 10 нейронов скрытого слоя – 1 выходной нейрон (таблица 4.5).

Таблица 4.4. Нейронные сети типа RBF

Лаг	Параметры отобранной НС Для акций ОАО «Ростелеком»	Параметры отобранной НС для акции ОАО «Лукоил»	Параметры отобранной НС для акций «Сбербанк» (для последних 405 знач.)
1	2	3	4
1	8 тестов  RBF 1:11:1 TPerf, VPerf, TePerf 2,054; 1,074; 1,266 (регрессионное отношение 0.114370, корреляция 0.993442, среднеквадратическая ошибка предсказания 1.073946)	15 тестов  RBF 1:5:1 TPerf, VPerf, TePerf 9.223; 13.890; 12.320 (регрессионное отношение 0.181738, корреляция 0.983840, среднеквадратическая ошибка предсказания 13.88914) Сеть претерпела переобучение. Нужно увеличить обучающую выборку.	13 теста  RBF 1:3:1 TPerf, VPerf, TePerf 110.8; 103.7; 152.0 (регрессионное отношение 0.190076, корреляция 0.981947, среднеквадратическая ошибка предсказания 103.6565)

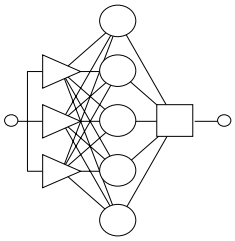
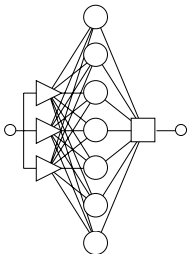
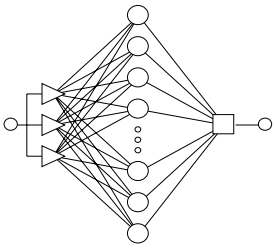
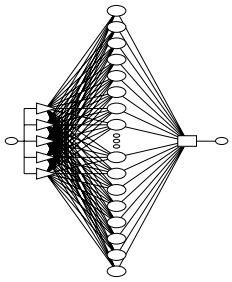
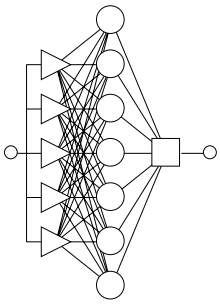
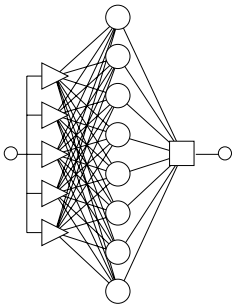
1	2	3	4
3	8 тестов  RBF 3:5:1 TPerf, VPerf, TePerf 2,483; 1,263; 1,35 (регрессионное отношение 0.134435, корреляция 0.990930, среднеквадратическая ошибка предсказания 1.262944)	8 тестов  RBF 3:7:1 TPerf, VPerf, TePerf 11,06; 12,95; 13,49 (регрессионное отношение 0.157377, корреляция 0.987700, среднеквадратическая ошибка предсказания 14.08784) Сеть претерпела переобучение. Нужно увеличить обучающую выборку.	24 тестов  RBF 3:22:1 TPerf, VPerf, TePerf 120.7; 124.3; 2076.0 (регрессионное отношение 0.223885v, корреляция 0.974682, среднеквадратическая ошибка предсказания 124.2555)
5	27 тестов  RBF 5:44:1 TPerf, VPerf, TePerf 2,484; 1,858; 2,916 (регрессионное отношение 0.098558, корреляция 0.995316, среднеквадратическая ошибка предсказания 1,857529)	16 тестов  RBF 5:7:1 TPerf, VPerf, TePerf 12.10; 14.57; 303.80 (регрессионное отношение 0.265997, корреляция 0.964260, среднеквадратическая ошибка предсказания 14.57255)	15 тестов  RBF 5:8:1 TPerf, VPerf, TePerf 120.6; 135.6; 5162.0 (регрессионное отношение 0.251100, корреляция 0.968655, среднеквадратическая ошибка предсказания 135.5618)

Таблица 4.5. Нейронные сети типа MLP (трехслойная сеть)

Лаг	Параметры отобранной НС для акций ОАО «Ростелеком»	Параметры отобранной НС для акций ОАО «Лукоил»	Параметры отобранной НС для акций «Сбербанк» (для последних 405 знач.)
1	2	3	4
1	16 тестов MLP 1:10:1 TPerf, VPerf, TePerf 2,128; 1,07; 1,201 (регрессионное отношение 0.113987, корреляция 0.993490, среднеквадратическая ошибка предсказания 1.070176)	16 тестов MLP 1:11:1 TPerf, VPerf, TePerf 9.217; 10.580; 12.610 (регрессионное отношение 0.195551, корреляция 0.980696, среднеквадратическая ошибка предсказания 10.58064)	69 тестов MLP 1:47:1 TPerf, VPerf, TePerf 106,7; 97.19; 188,5 (регрессионное отношение 0.179922, корреляция 0.983726, среднеквадратическая ошибка предсказания 97.18809)
3	16 тестов MLP 3:6:1 TPerf, VPerf, TePerf 2.104; 1.7; 1.529 (регрессионное отношение 0.085128, корреляция 0.996371, среднеквадратическая ошибка предсказания 1.699702)	16 тестов MLP 3:8:1 TPerf, VPerf, TePerf 9,259; 10,570; 11,840 (регрессионное отношение 0.195361, корреляция 0.980733, среднеквадратическая ошибка предсказания 10.5714)	16 тестов MLP 3:13:1 TPerf, VPerf, TePerf 108.9; 98.74; 248.4 (регрессионное отношение 0.179700, корреляция 0.983764, среднеквадратическая ошибка предсказания 98.74388)

Продолжение таблицы 4.5

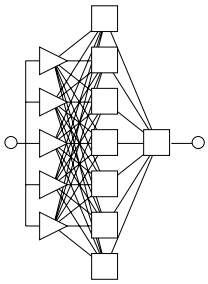
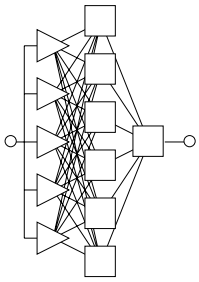
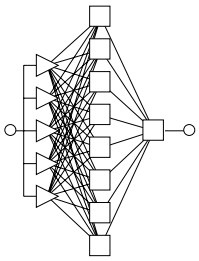
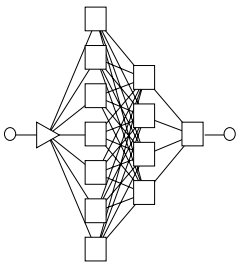
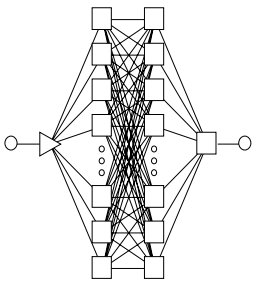
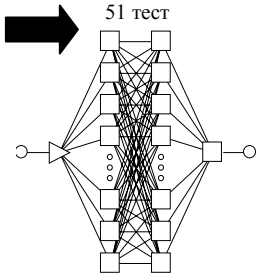
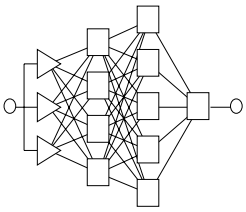
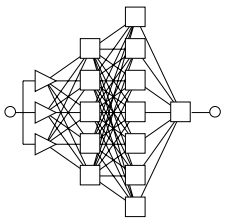
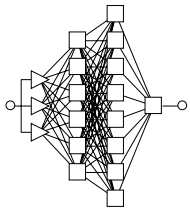
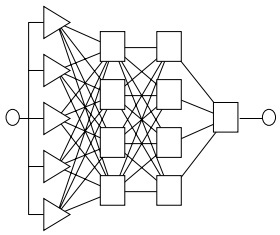
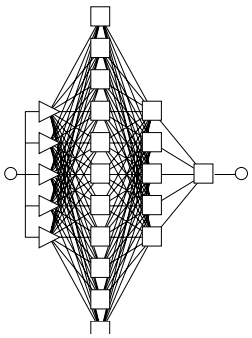
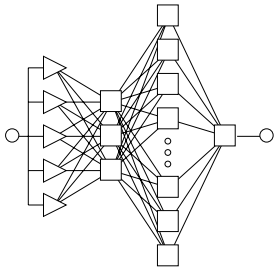
1	2	3	4
5	16 тестов  MLP 5:7:1 TPerf, VPerf, TePerf 2.022; 1.233; 1.200 (регрессионное отношение 0.187880, корреляция 0.982552, среднеквадратическая ошибка предсказания 1.232779)	16 тестов  MLP 5:6:1 TPerf, VPerf, TePerf 9,248; 10,510; 13,850 (регрессионное отношение 0.194242, корреляция 0.981016, среднеквадратическая ошибка предсказания 10.50778)	16 тестов  MLP 5:8:1 TPerf, VPerf, TePerf 117.1; 103.3; 295.6 (регрессионное отношение 0.187530, корреляция 0.982329, среднеквадратическая ошибка предсказания 103.2507)

Таблица 4.6. Нейронные сети типа MLP (четырехслойная сеть)

Лаг	Параметры отобранной НС для акций ОАО «Ростелеком»	Параметры отобранной НС для акций ОАО «Лукоил»	Параметры отобранной НС для акций «Сбербанк» (для последних 405 знач.)
1	2	3	4
1	16 тестов  MLP 1:7:4:1 TPerf, VPerf, TePerf 2.051; 1.248; 1.202 (регрессионное отношение 0.190008, корреляция 0.982021, среднеквадратическая ошибка предсказания 1.247681)	16 тестов  MLP 1:13:13:1 TPerf, VPerf, TePerf 9,337; 10,600; 16,120 (регрессионное отношение 0.195906, корреляция 0.980633, среднеквадратическая ошибка предсказания 10.60465)	51 тест  MLP 1:15:13:1 TPerf, VPerf, TePerf 106,6; 98,1; 144,2 (регрессионное отношение 0.181767, корреляция 0.983668, среднеквадратическая ошибка предсказания 98.09992)

1	2	3	4
3	16 тестов  MLP 3:4:5:1 TPerf, VPerf, TePerf 2.029; 1.255; 1.200 (регрессионное отношение 0.190985, корреляция 0.982249, среднеквадратическая ошибка предсказания 1.254811)	16 тестов  MLP 3:5:7:1 TPerf, VPerf, TePerf 9,382; 10,570; 14,370 (регрессионное отношение 0.195302, корреляция 0.980751, среднеквадратическая ошибка предсказания 10.56612)	32 теста  MLP 3:6:8:1 TPerf, VPerf, TePerf 112.3; 99.66; 240.7 (регрессионное отношение 0.181885, корреляция 0.983326, среднеквадратическая ошибка предсказания 99.6591)
5	16 тестов  MLP 5:4:4:1 TPerf, VPerf, TePerf 2.105; 1.231; 1.243 (регрессионное отношение 0.187464, корреляция 0.982542, среднеквадратическая ошибка предсказания 1.230627)	16 тестов  MLP 5:11:5:1 TPerf, VPerf, TePerf 9,256; 10,580; 17,210 (регрессионное отношение 0.195390, корреляция 0.980889, среднеквадратическая ошибка предсказания 10.57533)	16 тестов  MLP 5:3:9:1 TPerf, VPerf, TePerf 126.6; 102.7; 263.4 (регрессионное отношение 0.189757, корреляция 0.982160, среднеквадратическая ошибка предсказания 102.6798)

На рисунке 4.48 показано изменение ошибки прогнозирования в процессе обучения нейронной сети для курса акций ОАО «Ростелеком», которая составляет в среднем 1,5-2%. Выброс ошибок на интервале 600-800 вызван значительным ростом цен в период 2000 года.

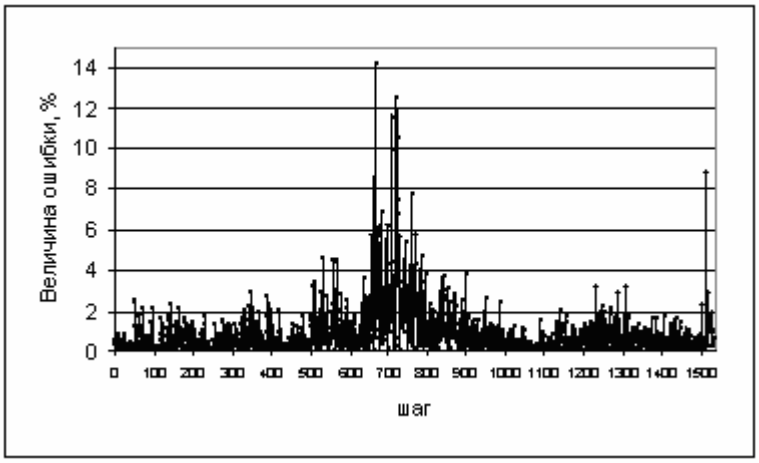


Рисунок 4.48. Изменение ошибки прогнозирования в процессе обучения нейронной сети для курса акций ОАО «Ростелеком»

Для курса акций ОАО «Лукоил» выбрана оптимальная структура нейронной сети со следующими показателями: тип – MLP; трехслойная структура: 3 входных нейронов – 8 нейронов скрытого слоя – 1 выходной нейрон (таблица 4.5). Однако возможно использование сети типа – MLP; трехслойная структура: 1 входных нейрона – 11 нейронов скрытого слоя – 1 выходной нейрон (таблица 4.5).

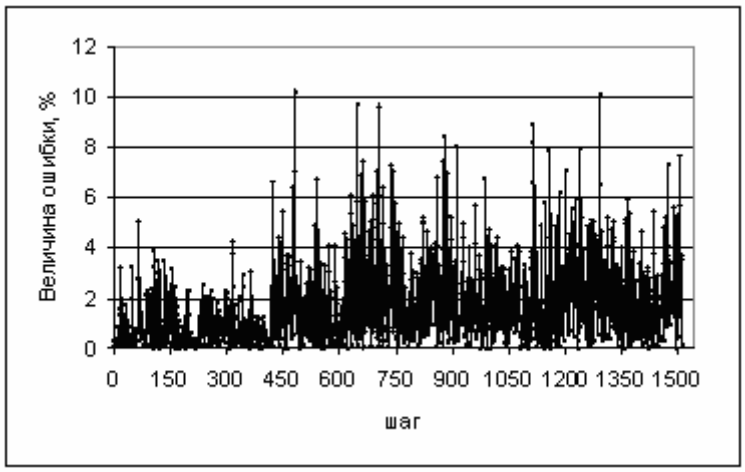


Рис. 4.49. Изменение ошибки прогнозирования в процессе обучения нейронной сети для курса акций ОАО «Лукоил»

На рисунке 4.49 показано изменение ошибки прогнозирования, которая составляет в среднем 3%.

Для эмитента – «Сбербанк» проанализировав общую тенденцию развития можно выделить два временных интервала 1998-2001 и конец 2001 – 2003. Поэтому в целях получения адекватных прогнозов рассмотрен второй период (примерно 405 значений временного ряда).

Выбрана оптимальная структура нейронной сети со следующими показателями: тип – MLP; трехслойная структура: 1 входных нейронов – 13 нейронов скрытого слоя – 1 выходной нейрон (таблица 4.5).

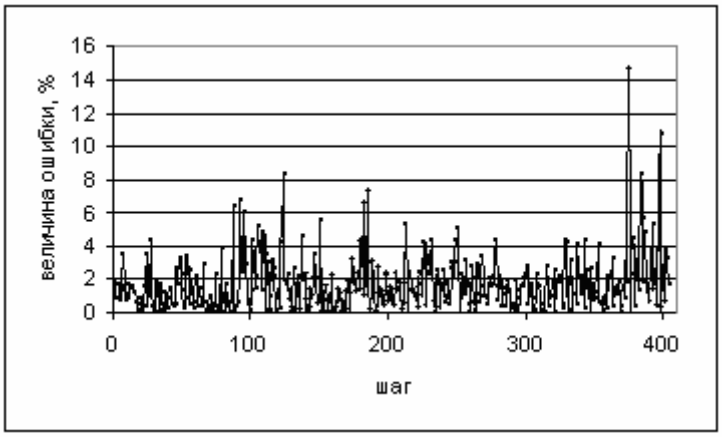


Рисунок 4.50. Изменение ошибки прогнозирования в процессе обучения нейронной сети для курса акций «Сбербанк»

Однако возможно использование сети типа – MLP; четырехслойная структура: 1 входных нейрона – 15 нейронов 1-го скрытого слоя – 13 нейронов 2-го скрытого слоя – 1 выходной нейрон (таблица 4.6). На рисунке 4.50 показано изменение ошибки прогнозирования, которая составляет в среднем 2%.

Можно видеть, что наблюдается рост величины ошибки для последних тестовых значений временного ряда. Это свидетельствует о том, что величина обучающей выборки имеет не достаточное количество данных, однако такой рост ошибки не приведет к росту ошибок прогнозирования НС и составит в среднем 3-7%.

4.8. Сравнительная оценка классических и нейросетевых методов прогнозирования

Сравнительная оценка прогнозов проведена для исследуемых временных рядов динамики курсов акций для последних 200 значений, так как именно эта часть выборок считается тестовой для всех исследуемых моделей прогнозирования. Исследование проведено для моделей экспоненциальной средней по следующей схеме.

Адаптивная полиномиальная модель первого порядка (p = 1).

Экспоненциальные средние 1-го и 2-го порядков определяются как

$$\begin{aligned} S_t &= \alpha x_t + \beta S_{t-1} \\ S_t^{[2]} &= \alpha S_t + \beta S_{t-1}^{[2]}, \end{aligned} \tag{4.4}$$

где β = 1-α.

Отсюда начальные условия

$$\begin{aligned} S_0 &= \hat{a}_{1,0} - \frac{\beta}{\alpha}\hat{a}_{2,0} \\ S_0^{[2]} &= \hat{a}_{1,0} - \frac{2\beta}{\alpha}\hat{a}_{2,0} \end{aligned} \tag{4.5}$$

Оценка модельного значения ряда с периодом упреждения τ равна

$$\hat{x}_t^{\bullet} = \left(2 + \frac{\alpha}{\beta}\tau\right)S_{t-\tau} - \left(1 + \frac{\alpha}{\beta}\tau\right)S_{t-\tau}^{[2]} \tag{4.6}$$

Адаптивная полиномиальная модель второго порядка (p=2). Экспоненциальные средние 1-го, 2-го и 3-го порядков

$$\begin{aligned} S_t &= \alpha x_t + \beta S_{t-1}; \\ S_t^{[2]} &= \alpha S_t + \beta S_{t-1}^{[2]}; \\ S_t^{[3]} &= \alpha S_t^{[2]} + \beta S_{t-1}^{[3]}. \end{aligned} \tag{4.7}$$

Начальные условия определим как

$$
\begin{aligned}
S_0 &= \hat{a}_{1,0} - \frac{\beta}{\alpha}\hat{a}_{2,0} + \frac{\beta(2-\alpha)}{2\alpha^2}\hat{a}_{3,0};\\
S_0^{[2]} &= \hat{a}_{1,0} - \frac{2\beta}{\alpha}\hat{a}_{2,0} + \frac{\beta(3-2\alpha)}{\alpha^2}\hat{a}_{3,0};\\
S_0^{[3]} &= \hat{a}_{1,0} - \frac{3\beta}{\alpha} + \frac{3\beta(4-3\alpha)}{2\alpha^2}\hat{a}_{3,0}.
\end{aligned}
\tag{4.8}
$$

где $\hat{a}_{1,0} = \hat{a}_1; \quad \hat{a}_{2,0} = \hat{a}_2; и\ \hat{a}_{3,0} = \hat{a}_3.$

Оценка модельного значения с периодом упреждения τ находим из выражения

$$
\begin{aligned}
\hat{x}_t^{\bullet} &= \left[6\beta^2 + (6-5\alpha)\alpha\cdot\tau + \alpha^2\tau^2\right]\frac{S_{t-\tau}}{2\beta^2} - \left[6\beta^2 + 2(5-4\alpha)\alpha\tau + 2\alpha^2\tau^2\right]\frac{S_{t-\tau}^{[2]}}{2\beta^2} +\\
&+ \left[2\beta^2 + (4-3\alpha)\alpha\tau + \alpha^2\tau^2\right]\frac{S_{t-\tau}^{[3]}}{2\beta^2}.
\end{aligned}
\tag{4.9}
$$

Весьма эффективным и надежным методом прогнозирования является экспоненциальное сглаживание. Основные достоинства метода состоят в возможности учета весов исходной информации, в простоте вычислительных операций, в гибкости описания различных динамик процессов. Метод экспоненциального сглаживания дает возможность получить оценку параметров тренда, характеризующих не средний уровень процесса, а тенденцию, сложившуюся к моменту последнего наблюдения. Наибольшее применение метод нашел для реализации среднесрочных прогнозов. Для метода экспоненциального сглаживания основным и наиболее трудным моментом является выбор параметра сглаживания α, начальных условий и степени прогнозирующего полинома. На рисунке 4.51 показана зависимость качества прогнозов от изменения параметра сглаживания α для адаптивной модели первого порядка, построенной для курса акций ОАО «РАО ЕЭС». Из рисунка хорошо видно, как влияет на качество прогнозов изменение параметра сглаживания α.

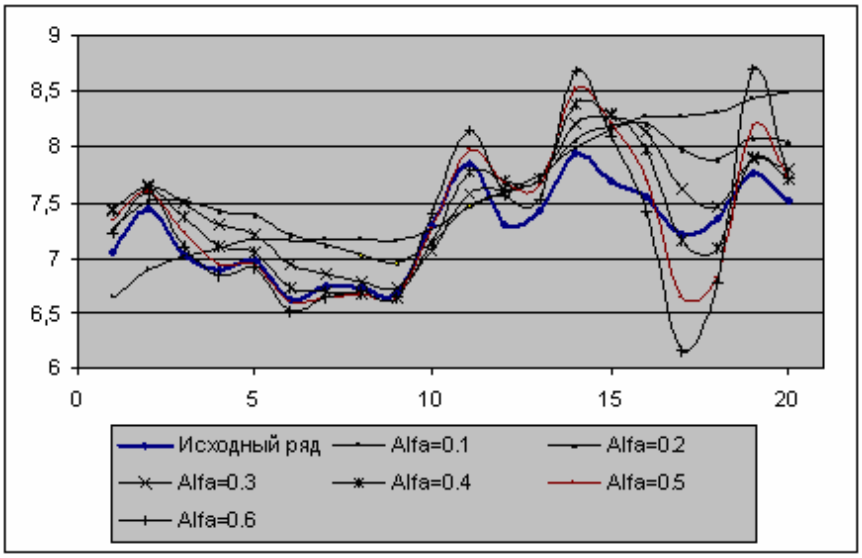


Рисунок 4.51. Зависимость качества прогнозов от изменения параметра сглаживания α на примере курса акций ОАО «РАО ЕЭС» для последних 20 значений ряда

В зависимости от величины параметра прогнозные оценки поразному учитывают влияние исходного ряда наблюдений: чем больше α, тем больше вклад последних наблюдений в формирование тренда, а влияние начальных условий быстро убывает. При малом α прогнозные оценки учитывают все наблюдения, при этом уменьшение влияния более «старой» информации происходит медленно. На рисунке 4.52 показана зависимость изменения ошибки прогноза в зависимости от изменения α для модели первого порядка.

В результате численных экспериментов удалось найти оптимальные значения параметра α для моделей первого и второго порядка, которые составили 0,38 и 0,2 соответственно. Однако полученные оптимальные модели не показали лучших прогнозов, чем нейронные сети. Так, на рисунках 4.53 – 4.56 показаны результаты прогнозов исследованных моделей и полученных нейронных сетей для исследуемых временных рядов.

Так, для акций ОАО «РАО ЕЭС» средняя относительная ошибка прогноза для модели p = 1 составила E = 3,81%, для модели p = 2, E = 4,02%, для НС (структуры 1:10:1) E = 3,90%; для НС (структуры 1:13:13:1) E = 3,88%.

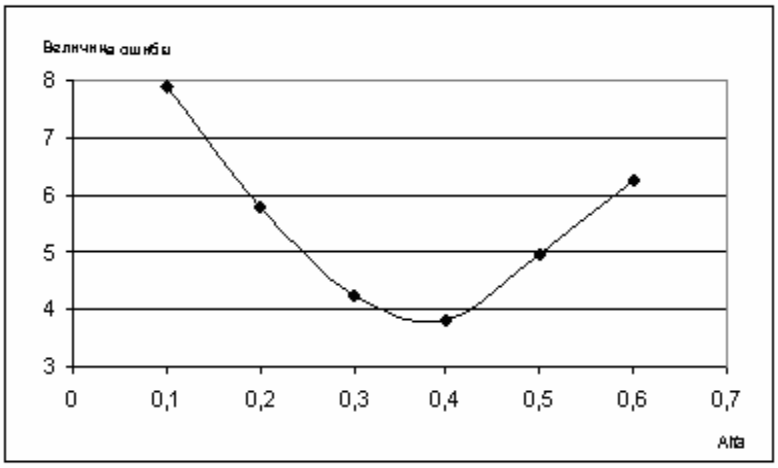


Рисунок 4.52. Зависимость изменения ошибки прогноза в зависимости от изменения α

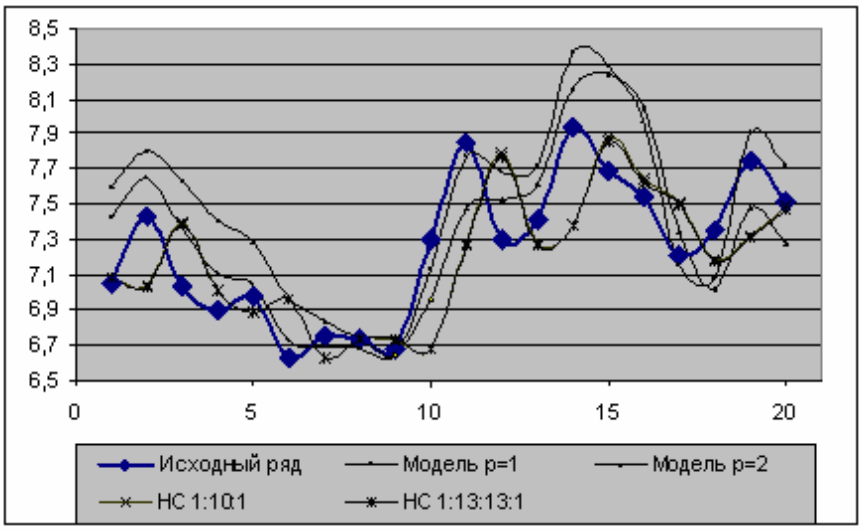


Рисунок 4.53. Результаты прогнозов адаптивных моделей и полученных нейронных сетей для динамики курса акций ОАО «РАО ЕЭС» для последних 20 значений ряда

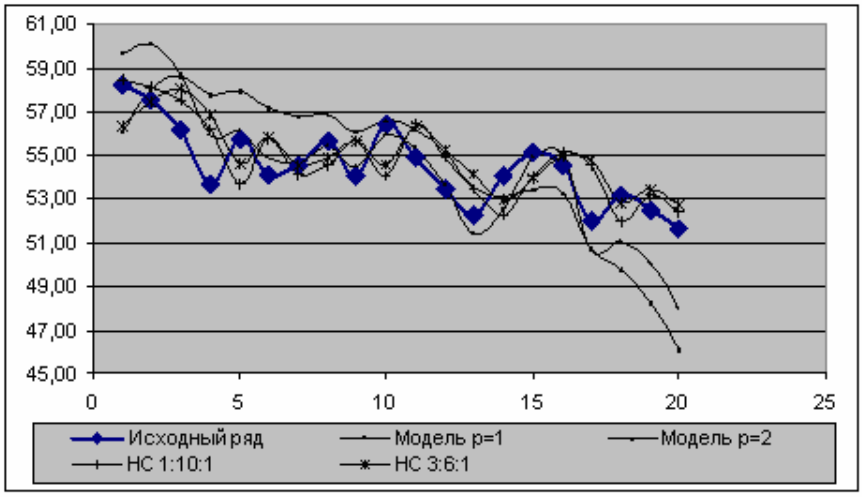


Рисунок 4.54. Результаты прогнозов адаптивных моделей и полученных нейронных сетей для динамики курса акций ОАО «Ростелеком» для последних 20 значений ряда

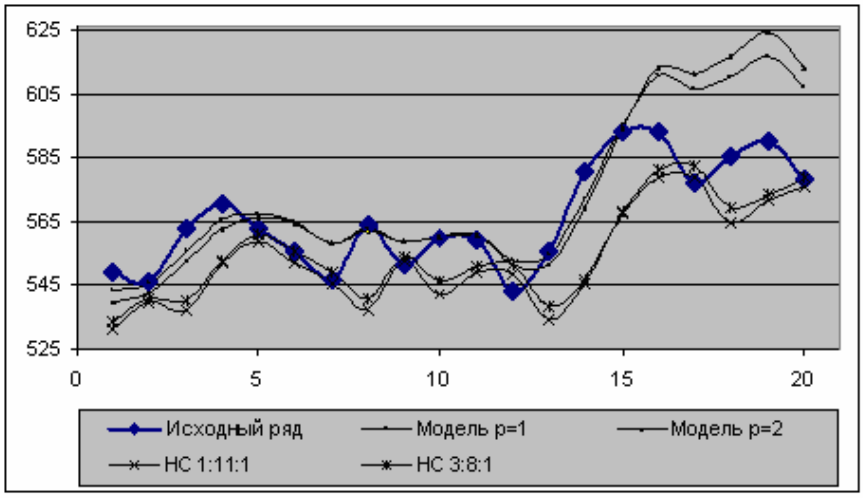


Рисунок 4.55. Результаты прогнозов адаптивных моделей и полученных нейронных сетей для динамики курса акций ОАО «Лукоил» для последних 20 значений ряда

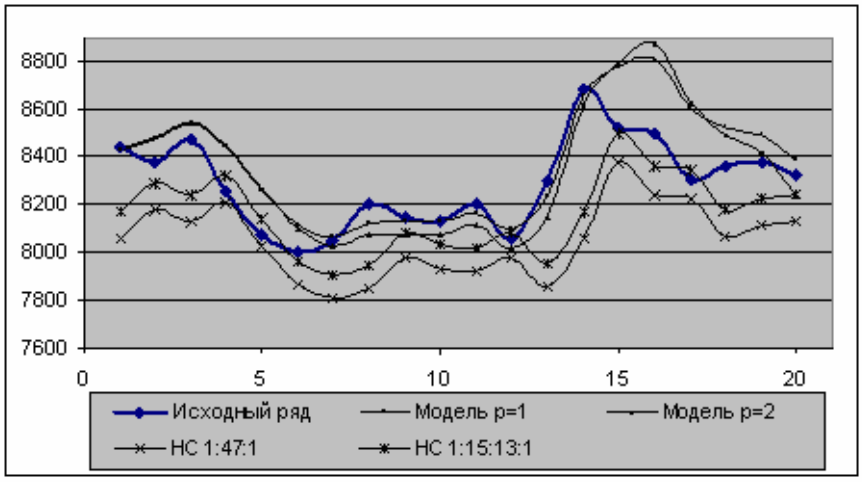


Рисунок 4.56. Результаты прогнозов адаптивных моделей и полученных нейронных сетей для динамики курса акций «Сбербанк» для последних 20 значений ряда

Для акций ОАО «Ростелеком» средняя относительная ошибка прогноза для модели p = 1, составила E = 2,56%, для модели p = 2, E = 4,26%, для подобранных НС (структуры 1:10:1) E = 2,41%, для НС (структуры 3:6:1) E = 2,46%.

Для акций ОАО «Лукойла» средняя относительная ошибка прогноза для модели p = 1, составила E = 2,60%, для модели p = 2, E = 3,11%, для подобранных НС (структуры) E = 2,67%, для НС (структуры) E = 2,46%.

Для акций «Сбербанка» средняя относительная ошибка прогноза для модели p = 1, составила E = 1,61%, для модели p = 2, E = 1,87%, для подобранных НС (структуры 1:47:1) E = 3,18%, для НС (структуры 1:15:13:1) E = 1,98%. Высокая ошибка прогнозирования для НС вызвана тем, что рассматриваемый ряд имеет не достаточное количество данных, что привело к недообученности НС. Но в целом нейронные сети показали наилучшие результаты по сравнению с классическими моделями прогнозирования.

Литература

1. Абовский Н.П. и др. Разработка практического метода нейросетевого прогнозирования. //Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение » Сб.докл., 2002. – С. 1089 – 1097.
2. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрии. – М.: ЮНИТИ, 1998. – 465 с.
3. Акушский И.Я. Машинная арифметика в остаточных классах. – М: Советское радио, 1968. – 440 с.
4. Алексеев В.И., Максимов А.В. Использование нейронных сетей с двухмерными слоями для распознавания образов//Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение »: Сб. докл., 2002. – С. 69–72.
5. Амербаев В.М. Теоретические основы машинной арифметики. – Алма-Ата: Наука КазССР, 1976. – 324 с.
6. Андерсон Т. Статистический анализ временных рядов. – М.: Наука, 1976. – 343 с.
7. Аркин В. И, Евстигнеев И.В. Вероятностные модели управления экономической динамики. – М.: Наука, 1979. – 176с.
8. Ашманов С.А. Математические модели и методы в экономике. М.: Изд. МГУ, 1981. – 158 с.
9. Барский А.Б. Обучение нейросети методом трассировки //Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение»: Сб. докл., 2002. – С. 862 – 898.
10. Батищев Д.И. Генетические алгоритмы решения экстремальных задач. – Воронеж: ВГУ, 1994. – 135 с.
11. Белим С.В. Математическое моделирование квантового нейрона//Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение»: Сб.докл., 2002. – С. 899 – 900.
12. Берже П., Помо И., Видаль К. Порядок в хаосе. О детерминированном подходе к турбулентности: Пер. с франц. – М.: Мир, 1991. – 368 с.
13. Бирман Э.Г. Сравнительный анализ методов прогнозирования //НТИ. Сер.2 – 1986. – №1. – С. 11–16.
14. Бодянский Е.В., Кучеренко Е.И. Диагностика и прогнозирование временных рядов многослойной радиально-базисной нейронной сети //Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение»: Сб. докл., 2002. – С. 69–72.

15. Бокс Дж., Дженкинс Г. (1974) Анализ временных рядов. Прогноз и управление. – М.: Мир, 1974. – Вып. 1, 2.
16. Болн Б., Хуань К. Дж. Многомерные статистические методы для экономики. М.: Наука, 1979. – 348 с.
17. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.: Наука, 1965. – 35 с.
18. Боярский А.Я., и др. Математическая статистика для экономистов. – М.: Статистика, 1979. – 253 с.
19. Бурдо А.И., Тихонов Э.Е. К вопросу систематизации методов и алгоритмов прогнозирования//Материалы межрегиональной конференции "Студенческая наука – экономике научно-технического прогресса". Ставрополь: СевКав ГТУ, 2001. – С. 33 – 34.
20. Бурдо А.И., Тихонов Э.Е. К вопросу совершенствования систем прогнозирования//Материалы XXXVIII юбилейной отчетной научной конференции за 1999 год: В 3 ч./Воронеж. гос. технол. акад. Воронеж, 2000.– Ч.2. – С. 211–215.
21. Бурдо А.И., Тихонов Э.Е. Об одном подходе к проблеме прогнозирования количественных характеристик производственных систем//Материалы XXX НТК профессорко-преподавательского состава. Ставрополь: СевКав ГТУ, 2000.– С.225–226.
22. Бутенко А.А. и др. Обучение нейронной сети при помощи алгоритма фильтра Калмана. //Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение »: Сб. докл., 2002. – С. 1120 – 1125.
23. Бухштаб А.А. Теория чисел. – М.: Государственное учебно-методическое издательство мин. Просвящения РСФСР, 1960. – 375 с.
24. Бытачевский Е.А., Козуб В.В. Использование нейронных сетей для распознавания визуальных образов//Материалы IV РНТ конференции «Вузовская наука – Северо-Кавказскому региону» Ставрополь, 2000. – С. 52–54.
25. Васильев В.И. Распознающие системы. – Киев: Навукова думка, – 1969. – 354 с.
26. Виноградов И.М. Основы теории чисел. – М.: Издательство «Наука», Гл. ред. физ.-мат.лит., 1965.– 173 с.
27. Владимирова Л.П. Прогнозирование и планирование в условиях рынка: Учебное пособие. – М.: Издательский дом «Дашков и К», 2000. – 308 с.

28. Вороновский Г.К., и др. Генетические алгоритмы, нейронные сети и проблемы виртуальной реальности. – Х.: ОСНОВА, 1997. – 112 с.
29. Галушкин А.И. Теория нейронных сетей. Кн. 1: Учеб. Пособие для вузов. – М.: ИПРЖР, 2001. – 385 с.:ил.
30. Гвишиани Д.М., Лисичкин В.А. Прогностика. М.: «Знание», 1968. – 421 с.
31. Гельфан И.М., Фомин С.В. Вариационное исчисление. – М.: Мир, 1961. – 321 с.
32. Гладыщевский А.И. Методы и модели отраслевого экономического прогнозирования. – М.: «Экономика», 1997. – 143 с.
33. Гласс Л., Мэки М. От часов к хаосу: ритмы жизни. – М.: Мир, 1991. – 153с.
34. Глущенко В.В. Прогнозирование. 3-е издание. – М.: Вузовская книга, 2000. – 208 с.
35. Голованова Н.Б., Кривов Ю.Г. Методические вопросы использования межотраслевого баланса в прогнозных расчетах//Взаимосвязи НТП и экономического развития: Сб.науч.тр./АН СССР. СО, ИЭиОПП. – Новосибирск, 1987. – С. 62–77.
36. Головко В.А. Нейронные сети: обучение, организация и применение. Кн.4:Учеб.пособие для вузов/Общая ред. А.И. Галушкина. – М.: ИПРЖР, 2001. – 256 с.
37. Горбань А.Н.Обучение нейронных сетей.–М.:СП“ПараГраф”, 1990. – 159с.
38. Горелик Е.С. и др. Об одном подходе к задаче формализации процесса прогнозирования //Автоматика и телемеханика. – 1987. – №2. – С.129-136.
39. Гренандер У. Случайные процессы и статистические выводы. (пер. с нем.) ИЛ. 1961. – 167с.
40. Гренджер К., Хатанака М. Спектральный анализ временных рядов в экономике. Пер.с англ. – М.: Статистика., 1972. – 312 с.
41. Грень Е. Статистические игры и их применение. – М.: Наука, 1975. – 243 с.
42. Гусак А.Н. и др. Подход к послойному обучению нейронной сети прямого распространения//Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение » Сб.докл., 2002. – С. 931 – 933.
43. Давидович Б.Я. и др. Методы прогнозирования спроса. М., 1972. –193с.

44. Дженкинс Г., Ватс Д. (1971, 1972) Спектральный анализ и его применения. – М.: Мир, 1971, 1972. – Вып. 1,2.
45. Джонстон Дж. (1980) Эконометрические методы. – М.: Статистика, 1980. – 431 с.
46. Добров Г.М., Ершов Ю.В. и др. Экспертные оценки в научно-техническом прогнозировании – Киев: Наукова Думка, 1974. – 159 с.
47. Еремин Д.М. Система управления с применением нейронных сетей//Приборы и системы. Управление, контроль, диагностика. – 2001. –№9 – С. 8–11.
48. Заенцев И.В. Нейронные сети: основные модели/Учебное пособие к курсу «Нейронные сети» – Воронеж: ВГУ, 1999. – 76 с.
49. Зайкин В.С. Применение простых цепей Маркова для прогнозирования расходов населения//Проблемы моделирования народного хозяйства, 4IV. Новосибирск, 1973. – С. 45 –47.
50. Занг В.Б. Синергитическая экономика. Время и перемены в нелинейной экономической теории. – М.: Мир, 1999. – 216с.
51. Ибираимова Т.Б. Прогнозирование тенденций финансовых рынков с помощью нейронных сетей//Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение » Сб.докл., 2002 г. – С. 745 – 755.
52. Иванов М.Н. Анализ роста курса акций с применением нейронных сетей. //Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение» Сб.докл., 2002 г. – С. 756 – 772.
53. Ивахненко А.Г. Долгосрочное прогнозирование и управление сложными системами. – Киев: Наукова думка, 1975. – 340 с.
54. Ивахненко А.Г., Лапа Р.Г., Предсказание случайных процессов. – Киев: Наукова думка, 1971. – 416с.
55. Ивахненко А.Г., Степаненко В.С. Особенности применения метода группового учета аргументов в задачах прогнозирования случайных процессов//Автоматика. –1986. – №5. – С. 3-14.
56. Ивахненко А.Г., Юрачков Ю.П. Моделирование сложных систем по экспертным данным. – М.: Радио и связь, 1987. – 119 с.
57. Калан, Роберт. Основные концепции нейронных сетей.: Пер. с англ. – М.: Издательский дом «Вильямс», 2001.– 288 с.
58. Капица С. П., Курдюмов С. П., Малинецкий Г. Г. Синергетика и прогнозы будущего. М.: Наука, 1997. – 236с.
59. Касти Дж. Связность, сложность и катастрофы: Пер. с англ. – М.: Мир, 1982, – 216 с.

60. Кейн Э. Экономическая статистика и эконометрия. М.: Наука, 1977, Вып. 1, 2.
61. Кемени Дж., Снелл Дж. Конечные цепи Маркова. М.: Наука, 1970. – 136 с.
62. Кендэл М. Временные ряды. Пер. с англ. Ю.П. Лукашина. – М.: «Финансы и статистика», 1979. – 198 с.
63. Кильдинов Г.С., Френкель А.А. Анализ временных рядов и прогнозирование. М. Статистика. 1973. – 432 с.
64. Клеопатров Д.И., Френкель А.А. Прогнозирование экономических показателей с помощью метода простого экспоненциального сглаживания. – Статистический анализ экономических временных рядов и прогнозирование. – М.: Наука, 1973. – 298с.
65. Кобринский Н.Е., и др. Экономическая кибернетика: Учебник для студентов вузов и фак., обучающихся по спец. «Экономическая кибернетика». – М.: Экономика, 1982. – 408 с.
66. Колмогоров А.Н. Об энтропии на единицу времени как метрическом инварианте автоморфизмов. ДАН СССР, 1959. – Т.124 – С.754-755
67. Кондратьев А.И. Теоретико-игровые распознающие алгоритмы. – М.: Наука, 1990. – 272 с.
68. Кульбак С. Теория информации и статистика. – М.:, Наука, 1967. – 408с.
69. Кучин Б.Л., Якушева Е.В. Управление развитием экономических систем: технический прогресс, устойчивость. – М.: Экономика, 1990. – 156с.
70. Лащев А.Я., Глушич Д.В. Синтез алгоритмов обучения нейронных сетей. //Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение » Сб.докл., 2002 г. – С. 997 – 999.
71. Левин В.Л. Выпуклый анализ в пространстве измеримых функций и его применение в математике и экономике. М.: Наука, 1985. – 352с.
72. Легостаева И.Л., Ширяев А.Н. Минимальные веса в задаче выделения тренда случайного процесса. – «Теория вероятностей и ее применение», 1971, – Т. XVI, – №2.
73. Лизер С. Эконометрические методы и задачи. – М.: Статистика, 1971. – 141с.
74. Лисичкин В.А. Теория и практика прогностики. – М.: Наука, 1972. – 223с.

75. Литовченко Ц.Г. Нейрокомпьютер для обнаружения и распознавания сложных динамических образов//Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение » Сб.докл., 2002 г. – С. 69–72.
76. Ллойд Э., Ледерман У. (1990) (ред.) Справочник по прикладной статистике. – М.: Финансы и статистика, 1990. – Том 2.
77. Лопатников Л.И. Экономико-математический словарь /АНССР. ЦЭМИ, – М.: Наука, 1987. – 506 с.
78. Лоскутов А.Ю., Михайлов А.С. Введение в синергетику: Учеб. Руководство. –М.: Наука. гл. ред. физ.-мат.лит., 1990.– 272 с.
79. Льюис К.Д. Методы прогнозирования экономических показателей./Пер. с англ. Демиденко Е.З. – М.: Финансы и статистика, 1986 г. – 132 с.
80. Ляпунов А.М. Собр. соч. Т.1,2. –М.:Изд-во АН СССР, 1954–1956.
81. Максимов В.А. Прогнозирование доходности инвестиций на фондовом рынке//Экономика и математические методы, 2001. Т. 37–№1. – С. 37 – 46.
82. Малинецкий Г.Г., Потапов А.Б. Современные проблемы нелинейной динамики – М.: Эдиториал УРСС, 2000.– 336с.
83. Математическая энциклопедия: Гл.ред. И.М. Виноградов, т.3 Коо-Од – М.: Советская энциклопедия, 1982. – 1184 с.
84. Махортых С.А., Сычев В.В. Алгоритм вычисления размерности стохастического аттрактора и его применение к анализу электрофизиологических данных.– Пущино, 1998. – 34с.
85. Медведев В.С., Потемкин В.Г. Нейронные сети. MATLAB 6 /Под общ.ред. к.т.н. В.Г. Потемкина. – М.: ДИАЛОГ-МИФИ, 2002. – 496 с.
86. Минский М., Пайперт С. Персептроны. – М.: Мир, 1971. – 365с.
87. Михайлов Ю.Б. Алгоритм выбора прогнозирующей зависимости, обеспечивающей наилучшую точность прогноза//Приборы и системы. Управление, контроль, диагностика., 2000. – №12. – С. 11 – 19.
88. Моделирование функционирования развивающихся систем с изменяющейся структурой. Сб. науч. тр./АН УССР. Ин-т кибернетики им. В.М. Глушакова. – Киев: 1989. – 140 с.
89. Моришма М. Равновесие, устойчивость, рост. – М.: Наука, 1972. 314 с.

90. Морозова Т.Г., Пикулькин А.В., Тихонов В.Ф., и др. Прогнозирование и планирование в условиях рынка. Учеб. Пособие для вузов. Под ред. Т.Г. Морозовой, А.В. Пикулькина. – М.: ЮНИТИ-ДАНА, 1999. – 318 с.
91. Нейроинформатика и ее приложения //Материалы 3 Всероссийского семинара, 6-8 октября 1995 г. Ч. 1 /Под редакцией А.Н.Горбаня. – Красноярск: изд. КГТУ, 1995. – 229 с.
92. Нейронные сети. STSTISTICA Neural Networks: Пер. с англ. – М.: Горячая линия – Телеком. 2001. – 182 с.
93. Никайдо Х. Выпуклые структуры и математическая экономика. – М.: Мир. 1972. – 127с.
94. Новиков А.В., и др. Метод поиска экстремума функционала оптимизации для нейронной сети с полными последовательными связями //Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение» Сб.докл., 2002 г. – С. 1000 – 1006.
95. Осуга М. Обработка знаний. – М.: Мир, 1989. – 239 с.
96. Оуэн Г. Теория игр. – М.: Наука, 1971. – 359 с.
97. Песин Я.Б. Характеристические показатели Ляпунова и гладкая эргодическая теория. УМН, 1977.– Т.32. – С.55–112.
98. Петерс Э. Хаос и порядок на рынке капитала. – М.: Мир, 2000. – 333с.
99. Половников В.А., и др. Оценивание точности и адекватности моделей экономического прогнозирования // Математическое моделирование экономических процессов: Сб. науч. тр./МЭСИ – М., 1986. – С. 37–47.
100. Пятецкий В.Е., Бурдо А.И. Имитационное моделирование процесса создания обучающихся систем. – В сб.: Имитационное моделирование производственных процессов. Под. ред. Мироносецкого Н.Б., – Новосибирск. 1979. – 68 с.
101. Пятецкий В.Е., Бурдо А.И., Литвяк В.С. Построение стохастических моделей прогнозирования параметров производственных систем//Рук. деп. в Черметинформации, 1987.– № 4161.
102. Развитие российского финансового рынка и новые инструменты привлечения инвестиций. – М., 1998. – 233 с.
103. Рожков Л.Н., Френкель А.А. Выбор оптимального параметра сглаживания в методе экспоненциального сглаживания. – Основные проблемы и задачи научного прогнозирования. – М.: Наука, 1972.- 154 с.

104. Розенблат Ф. Принципы нейродинамики: Персептрон и теория механизмов мозга. Пер. с англ. – М.: Мир, 1965. - 175 с.
105. Розин Б.Б. Распознавание образов в экономических исследованиях. – М.: Статистика, 1973. – 198 с.
106. Романов А.Н., Одинцов Б.Е. Советующие системы в экономике: Учебное пособие для вузов.– М.: ЮНИТИ-ДАНА, 2000. – 487 с.
107. Сергеев А.В. Прогнозирование временных рядов с помощью нейронных сетей с радиальными базисными функциями//Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение » Сб.докл, 2002 г. – С. 1187 – 1191.
108. Серебрянников М.Г., Первозванский А.А. Выявление скрытых периодичностей. М.: Наука, 1965. – 244с.
109. Сигеру О., и др. Нейроуправление и его приложения. Пер. с англ. под ред. А.И. Галушкина. – М.: ИПРЖР, 2001. – 321 с.
110. Статевич В.П., Шумков Е.А. Новый принцип построения самообучаемых нейросетей//Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение » Сб.докл., 2002. – С. 1037 – 1040.
111. Стратонович Р.Л. Теория информации. –М.: Сов. Радио. – 1975. – 424с.
112. Теория прогнозирования и принятия решений. Учеб. пособие. Под. ред. С.А. Саркисяна – М.: Высш. Школа, 1977. – 351 с.
113. Тихонов Э.Е., Бурдо А.И. К вопросу совершенствования автоматизированных систем прогноза//Материалы межрегиональной конференции "Студенческая наука – экономике научно-технического прогресса". Ставрополь: СевКав ГТУ, 2000. – С.30–31.
114. Тихонов Э.Е. Об одном подходе к прогнозированию с помощью нейронных сетей//Материалы третьей МНК "Студенческая наука – экономике России". Ставрополь: СевКав ГТУ, 2002.– С. 69 – 70.
115. Тихонов Э.Е. Об одном подходе к вопросу повышения надежности работы нейронных сетей с использованием непозиционной системы остаточных классов//Материалы XXXII НТК профессорко-преподавательского состава, аспирантов и студентов СевКав ГТУ за 2002 год. Т 2: Технические и прикладные науки г. Ставрополь: СевКав ГТУ, 2003. – С. 29 – 30.
116. Тихонов Э.Е. Об одном подходе к разработке системы прогнозирования с использованием модулярных вычислений на ней-

ронных сетях//Материалы РНК «Теоретические и прикладные проблемы современной физики», Ставрополь: СГУ, 2002. – С. 449 – 453. Тихонов Э.Е. Проблемы идентификации моделей прогнозирования на нейронных сетях//Компьютерная техника и технологии: Сб. трудов регион. НТК. Ставрополь: СевКав ГТУ, 2003. – С.187–191.

117. Тихонов Э.Е. Сравнительный анализ радиально базисной нейронной сети и сети типа – многослойный персептрон на примере прогнозирования объема экспорта//Компьютерная техника и технологии: Сб. трудов регион. НТК. Ставрополь: СевКав ГТУ, 2003. – С.184–186.
118. Тихонов Э.Е., Зайцева И.В. Прогнозирование на основе рядов Фурье//Компьютерная техника и технологии: Сб. трудов регион. НТК. Ставрополь: СевКав ГТУ, 2003. – С.173–175.
119. Тихонов Э.Е., Кучаев А.Г. Об одном подходе к вопросу разработки информационной системы прогнозирования для задач поддержки принятия решений// Компьютерная техника и технологии: Сб. трудов регион. НТК. Ставрополь: СевКав ГТУ, 2003.– С.191–193.
120. Тихонов Э.Е. Об одном подходе к разработке алгоритмов модулярных вычислений на нейронных сетях//Материалы РНК «Теоретические и прикладные проблемы современной физики», Ставрополь: СГУ, 2002. – С.443 – 448.
121. Тихонов Э.Е. Сравнительный анализ традиционных методов прогнозирования с методами прогнозирования на нейронных сетях//Компьютерная техника и технологии: Сб. трудов регион. НТК. Ставрополь: СевКав ГТУ, 2003. – С.179–183.
122. Тихонов Э.Е. Проблемы идентификации моделей прогнозирования на нейронных сетях//Компьютерная техника и технологии: Сб. трудов регион. НТК. Ставрополь: СевКав ГТУ, 2003. – С.187–191.
123. Тихонов Э.Е. Методы и алгоритмы прогнозирования экономических показателей на базе нейронных сетей и модулярной арифметики. Дисс. …канд. тех. наук. – Ставрополь, 2003.– 139с.
124. Томпсон Дж.М.Т. Неустойчивости и катастрофы в науке и технике: Пер. с англ. – М.: Мир, 1985. – 254 с.
125. Трисеев Ю.П. Долгосрочное прогнозирование экономических процессов (системные методы). – Киев: Наукова Думка, 1987. – 132с.

126. Ту Дж., Гонсалес Р. Принципы распознавания образов. – М.: Наука, 1978. – 139 с.
127. Уоссермен Ф. Нейрокомпьютерная техника: теория и практика. – М.: ЮНИТИ, 1992. – 240 с.
128. Федер Е. Фракталы. - М.: Мир, 1991. – 143с.
129. Френкель А.А. Математические методы анализа динамики и прогнозирования производительности труда. – М.: Наука, 1972.
130. Хенан Э. Анализ временных рядов. – М.: Статистика, 1964. – 215с.
131. Хенан Э.Дж. Многомерные временные ряды. – М.: Мир, 1986. – 346с.
132. Ховард Р.А. Динамическое прогнозирование и марковские процессы. – М.: сов. Радио, 1964. – 365с.
133. Червяков Н.И., Тихонов Э.Е. Применение нейронных сетей для задач прогнозирования и проблемы идентификации моделей прогнозирования//Нейрокомпьютеры: разработка, применение. – М.: Радиотехника, – 2003. – № 10. – С.25-31.
134. Червяков Н.И., Тихонов Э.Е. Предсказание фрактальных временных рядов с помощью нейронных сетей // Нейрокомпьютеры: разработка, применение. – М.: Радиотехника, – 2003. – № 10. – С.19-24.
135. Червяков Н.И., Тихонов Э.Е. Совершенствование методов прогнозирования на базе нейронных сетей с использованием непозиционной системы остаточных классов//Нейрокомпьютеры: разработка, применение. – М.: Радиотехника, – 2003. – № 10. – С.13-18.
136. Червяков Н.И., Сахнюк П.А. Применение нейроматематики для реализации вычислений в конечных кольцах //Нейрокомпьютеры: разработка, применение, 1999. – № 1.– С. 75–84
137. Червяков Н.И., Шапошников А.В., Сахнюк П.А., Калмыков И.А. Применение модулярных вычислений для нейрообработки//Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение» НКП, 2002. Под ред. проф. А. Галушкина – М., 2002. – С. 1053–1056.
138. Чуев Ю.В., Михайлов Ю.Б., Кузьмин В.И. Прогнозирование количественных характеристик процессов. М., «Сов. радио», 1975. – 400 с.

139. Шибхузов З.М. Конструктивный TOWER алгоритм для обучения нейронных сетей из ΣΠ – нейронов//Труды VIII Всероссийской конференции «Нейрокомпьютеры и их применение » Сб.докл., 2002. – С. 1066 – 1072.
140. Шустер Г. Детерминированный хаос: Введение: Пер. с англ. – М.: Мир, 1988. – 240 с.
141. Экономика переходного периода. Очерки экономической политики посткоммунистической России 1991 – 1997. – М., 1998.
142. Экономические межотраслевые модели целевого прогнозирования экономики/Б.В. Седелев, и др.; ВНИИСИ. – Препр. – М., 1987. – 59с.
143. Ямпольский С.М., Хилюк Ф.М., Лисичкин В.А. Проблемы научно-технического прогнозирования. М.: Экономика, 1969. – 189 с.
144. Almon S. (1960) “The Distributed Lag between Capital Appropriations and Expenditures”, Econometrica, 30, 178-196.
145. Andrews D.W.K. (1991) “Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation,” Econometrica, 59, 817–858.
146. Ardeni P.G., D. Lubian (1991) “Is There Trend Reversion in Purchaising Power Parity”, European Economic Review, 35, 1035-1055.
147. Banerjee A., R.L. Lumsdaine, J.H. Stock (1992) “Recursive and Sequential Tests of the Unit Root and Trend Break Hypotheses: Theory and International Evidence”, Journal of Business and Economic Statistics, 10, 271-287.
148. Bierens H.J. (1997) “Testing the Unit Root with Drift Hypothesis Against Nonlinear Trend Stationarity, with an Application to the US Price Level and Interest Rate”, Journal of Econometrics, 81, 29-64.
149. Bollerslev, Tim (1986) “Generalized Autoregressive Conditional Heteroskedasticity,” Journal of Econometrics, 31, 307–327.
150. Brown R.G. (1962) “Smoothing, Forecasting and Prediction of Discrete Time-Series”. Prentice-Hall, New Jersey.
151. Brown R.G. (1963) “Smoothing, Forecasting and Prediction”. Prentice-Hall, Englewood Cliffs, N.Y.
152. Cagan P. (1956) “The Monetary Dynamics of Hyperinflation”, in: “Studies in the Quantity Theory of Maney”. Chicago, University of Chicago Press.

153. Chan K.H, J.C.Hayya, J.K.Ord (1977) “A Note on Trend Removal Methods: The Case of polynomial versus vatiate differencing”, Econometrica, 45, 737-744.
154. Cheung Y.-W., K.S.Lay (1995) “Lag Order and Critical Values of a Modified Dickey-Fuller Test”, Oxford Bulletin of Economics and Statistics, 57, №3, 411-419.
155. Cheung Y.-W., M.D. Chinn (1996) “Deterministic, Stochastic, and Segmented Trends in Aggregate Output: a Cross-country Analysis”, Oxford Economic Papers, 48, №1, 134-162.
156. Cheung Y.-W., M.D. Chinn (1997) “Further Investigation of the Uncertain Unit Root in GDP”, Journal of Business and Economic Statistics, 15, 68-73.
157. Christiano L.J., M. Eichenbaum (1990) “Unit Roots in Real GDP: Do We Know, and Do We Care?”, Carnegie-Rochester Conference Series on Public Policy, 32, 7-62.
158. Clark P.K. (1989) “Trend Reversion in Real Output and Unemployment”, Journal of Econometrics, 40,15-32.
159. Cochrane J.H. (1998) “How Big is the Random Walk in GNP?”, Journal of Political Economy, 96, 893-920.
160. Cogley T. (1990) “International Evidence on the Size of the Random Walk in Output”, Journal of Political Economy, 98, 501-518.
161. Copeland L.S. (1991) “Cointegration Tests with Daily Exchange Rate Data”, Oxford Bulletin of Economics and Statistics, 53, 185-198.
162. Davidson R., J.G. MacKinnon (1993) Estimation and Inference in Econometrics, Oxford University Press
163. den Haan W.J. (2000) “The Comovement Between Output and Prices”, Journal of Monetary Economics, 46, №1, 3-30.
164. Dickey D.A. (1976) “Estimation and Hypothesis Testing for Nonstationary Time Series”, Ph.D. dissertation, Iowa State University.
165. Dickey D.A., S. Pantula (1987) “Determining the Order of Differencing in Autoregressive Processes”, Journal of Business and Economic Statistics, 15, 455-461.
166. Dickey D.A., W.A. Fuller (1979) “Distribution of the Estimators for Autoregressive Time Series with a Unit Root”, Journal of the American Statistical Association, 74, 427–431.
167. Dickey D.A., W.R.Bell, R.B. Miller (1986) “Unit Roots in Time Series Models: Tests and Implications”, American Statistican, 40, 12-26.

168. Dickey, D.A., W.A. Fuller (1981) “Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root”, Econometrica, 49, 1057-1072.
169. Dolado H., T. Jenkinson, S. Sosvilla-Rivero (1990) “Cointegration and Unit Roots”, Journal of Economic Surveys, 4, 243-273.
170. Dutt S.D. (1998) “Purchasing Power Parity Revisited: Null of Cointegration Approach”, Applied Economic Letters, 5, 573-576.
171. Dutt S.D., D. Ghosh (1999) “An Empirical Examination of Exchange Market Efficiency”, Applied Economic Letters, 6, №2, 89-91.
172. Dweyer G.P., Wallace M.S. (1992) “Cointegration and Market Efficiency”, Journal of International Money and Finance, 11 318-327.
173. Elliott G., T.J. Rothenberg, J.H. Stock (1996) “Efficient Tests for an Autoregressive Unit Root”, Econometrica, 64, 813-836.
174. Enders W. (1995) “Applied Econometric Time Series”, Wiley, New York
175. Engle R., Kraft D. (1983) B Applied Time Series Analysis of Economic Data, Washington D.C.: Bureau of the Gensus.
176. Engle R.F., C.W.J. Granger (1991) “Cointegrated Economic Time Series: An Overview with New Results”, in R.F. Engle and C.W.J. (eds.), Long-Run Economic Relationships, Readings in Cointegration, Oxford University Press, 237-266.
177. Engle, R. (1983) “Estimates of the Variance of U.S. Inflation Based on the ARCH Model,” Journal of Money, Credit and Banking, 15, 286–301.
178. Engle, R. F. (1982) “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation,” Econometrica, 50, 987–1008.
179. Fama E.F., French K.R. (1988) “Permanent and Temporary Components of Stock Prices”, Journal of Political Economy, 96, 246-273.
180. Friedman M. (1957) “Theory of the Consumption Function”. Princeton, N.J.: Princeton University Press.
181. Fuller W.A. (1976) Introduction to Statistical Time Series, Wiley, New York.
182. Fuller W.A. (1996) Introduction to Statistical Time Series, 2nd Ed, Wiley, New York
183. Funke N., J. Thornton (1999) “The demand for money in Italy, 1861-1988”, Applied Economic Letters, 6, №5, 299-301.

184. Ghysels E., H.S. Lee, J. Noh (1994) “Testing for Unit Roots in Seasonal Time Series: Some Theoretical Extensions and a Monte Carlo Investigation”, Journal of Econometrics, 62, 415-442.
185. Ghysels E., Perron P. (1992) “The Effect of Seasonal Adjustment Filters on Tests for a Unit Root”, Journal of Econometrics, 55, 57-98.
186. Gragg J. (1983) “More Efficient Estimation in the Presence of Heteroscedasticity of Unknown Form”, Econometrica, 51, 751-763.
187. Granger C.W.J. (1963) “The Effect of Varjing Month-Length in the Analisis of Economic Time Series”, L’Industria, 1, 3, Milano.
188. Green W.H. (1997) “Econometric Analysis”. 3rd edition, Prentice-Hall.
189. Hafer R.W., D.W. Jansen (1991) “The Demand for Money in the United States: Evidence from Cointegration Tests”, Journal of Money, Credit, and Banking, 23 (1991), 155-168.
190. Hall A. (1994) “Testing for a Unit Root in Time Series with Pretest Data-Based Model Selection”, Journal of Business and Economic Statistics, 12, 451-470.
191. Hamilton, James D. (1994) Time Series Analysis, Princeton University Press, Princeton.
192. Hasan M.S. (1998) “The Choice of Appropriate Monetary Aggregate in the United Kingdom”, Applied Economic Letters, 5, №9, 563-568.
193. Hatanaka M. (1996) Time Series-Based Econometrics: Unit Roots and Cointegration, Oxford University Press.
194. Holden D., Perman R. (1994) “Unit Roots and Cointegration for Economist”, в сборнике Cointegration for the Applied Economists (редактор Rao B.B.), Macmillan.
195. Holt C.C. (1957) “Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages”, Carnegie Inst. Tech. Res. Mem., 52.
196. Johansen S., K. Juselius (1990) “Maximum Likelihood Estimation and Inferences on Cointegration–with applications to the demand for money,” Oxford Bulletin of Economics and Statistics, 52, 169–210.
197. Kim B. J.C., Mo Soowon (1995) “Cointegration and the long-run forecast of exchange rates“, Economics Letters, 48, №№ 3-4, 353-359.
198. Kolmogoroff A. (1939) “Sur L’interpolation et L’extrapolation des Suites Stationaires”, Compt. Rend., 208, 2043.

199. Koyck L.M. (1954) “Distributed Lags and Investment Analysis”. North-Holland Publishing Company, Amsterdam.
200. Kwiatkowski D., P.C.B. Phillips, P. Schmidt, Y. Shin (1992) “Testing of the Null Hypothesis of Stationary against the Alternative of a Unit Root”, Journal of Econometrics, 54, 159-178.
201. Leybourne S., T. Mills, P. Newbold (1998) “Spurious Rejections by Dickey-Fuller Tests in the Presence of a Break Under Null”, Journal of Econometrics, 87, 191-203.
202. Leybourne S.J. (1995) “Testing for Unit Roots Using Forward and Reverse Dickey-Fuller Regressions”, Oxford Bulletin of Economics and Statistics, 57, 559-571.
203. Lumsdaine R.L., Kim I.M. (1997) “Structural Change and Unit Roots”, The Review of Economics and Statistics, 79, 212-218.
204. MacKinnon, J.G. (1991) “Critical Values for Cointegration Tests,” Глава 13 в Long-run Economic Relationships: Readings in Cointegration, edited by R.F.Engle and C.W.J. Granger, Oxford University Press.
205. Maddala G.S., In-Moo Kim (1998) Unit Roots, Cointegration, and Structural Change. Cambridge University Press, Cambridge.
206. Metin K. (1995) “An Integrated Analysis of Turkish Inflation”, Oxford Bulletin of Economics and Statistics, 57, №4, 513-532.
207. Milas C. (1998) “Demand for Greek Imports Using Multivariate Cointegration Technique”, Applied Economics, 30, №11, 1483-1492.
208. Mills T.C. (1993) The Econometric Modeling of Financial Time Series. Cambridge University Press, Cambridge.
209. Molana H. (1994) “Consumption and Fiscal Theory. UK Evidence from a Cointegration Approach”, Dundee Discussion Papers, University of Dundey, Dundey, Scotland.
210. Murray C.J., C.R. Nelson (2000) “The Uncertain Trend in U.S. GDP”, Journal of Monetary Economics, 46, 79-95.
211. Nadal-De Simone F., W.A. Razzak (1999) “Nominal Exchange Rates and Nominal Interest Rate Differentials”, IMF Working Paper WP/99/141.
212. Nelson C.R., C.I. Plosser (1982) “Trends and Random Walks in Macroeconomic Time Series”, Journal of Monetary Economics, 10, 139-162.
213. Nelson C.R., H. Kang (1981) “Spurious Periodicity in Inappropriately Detrended Time Series”, Journal of Monetary Economics, 10, 139-162.

214. Nerlove M. (1956) “Estimates of the Elasticities of Supply of Selected Agricultural Commodities”, Jorn. Farm Econ., 38, 496-509.
215. Nerlove M. (1958) “The Dynamics of Supply: Estimation of Farmers Response to Price”. The Johns Hopkins Press. Baltimore.
216. Newey W., K. West (1987) “A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” Econometrica, 55, 703–708
217. Newey W., K. West (1994) “Automatic Lag Selection in Covariance Matrix Estimation,” Review of Economic Studies, 61, 631–653.
218. Ng S., P. Perron (1995) “Unit Root Tests in ARMA models With Data-Dependent Methods for the Selection of the Truncation Lag”, Journal of American Statistical Association, 90, 268-281.
219. Nunes L.S., Newbold P., C.-M. Kuan (1997) “Testing for Unit Roots With Breaks. Evidence on the Great Crash and the Unit Root Hypothesis Reconsidered”, Oxford Bulletin of Economics and Statistics, 59, №4, 435-448.
220. Perron P. (1988) “ Trends and Random Walks in Macroeconomic Time Series: Furter Evidence from a New Approach”, Journal of Economic Dynamic and Control, 12, 297-332.
221. Perron P. (1989a) “The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis, Econometrica, 577, 1361-1401.
222. Perron P. (1989b) “Testing for a Random Walk: A Simulation Experiment When the Sampling Interval Is Varied” – в сборнике Advances in Econometrics and Modeling (редактор B.Ray), Kluwer Academic Publishers, Dordrecht and Boston.
223. Perron P. (1997) "Further evidence on breaking trend functions in macroeconomic variables, Journal of Econometrics, 80, №2, 355-385.
224. Perron P., Vogelsang T.J. (1993) “Erratum”, Econometrica, 61, №1, 248-249.
225. Phillips P.C.B. (1987) “Time Series Regression with a Unit Root”, Econometrica, 55, 277-301.
226. Phillips P.C.B., P. Perron (1988) “Testing for a Unit Root in Time Series Regression,” Biometrika, 75, 335–346.
227. Said E., D.A. Dickey (1984) “Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order,” Biometrika, 71, 599–607.

228. Schmidt P., Phillips P.C.B. (1992) “LM Tests for a Unit Root in the Presence of Deterministic Trends”, Oxford Bulletin of Economics and Statistics, 54, 257-287.
229. Schwert G.W. (1989) “Tests for Unit Roots: A Monte Carlo Investigation”, Journal of Business and Economic Statistics, 7, 147-159.
230. Shiller R.J., Perron P. (1985) “Testing the Random Walk Hypothesis: Power versus Frequency of Observation”, Economic Letters, 18, 381-386.
231. Solow R.M. (1960) “On a Family of Lag Distributions”, Econometrica, 28, 393-406.
232. Taylor A.M.R. (2000) “The Finite Sample Effects of Deterministic Variables on Conventional Methods of Lag-Selection in Unit-Root Tests”, Oxford Bulletin of Economics and Statistics, 62, 293-304.
233. Walker G. (1931) “On Periodicity in Series of Related Terms”, Proc. Royal Soc., A131, 518.
234. White H., I. Domovitz (1984) “Nonlinear Regression with Dependent Observations”, Econometrica, 52, 143-162.
235. Wiener N. (1949) “Extrapolation, Interpolation and Smoothing of Stationary Time Series”. John Wiley, New York.
236. Winters P.R. (1960) “Forecasting Sales by Exponentially Weighted Moving Averages”, Mgmt. Sci., 6, 324.
237. Wold H.O. (1932) “A Study in the Analysis of Stationary Time Series”. Almquist and Wieksell, Uppsala.
238. Woodward G., R. Pillarisetti (1999) “Empirical Evidence on Alternative Theories of Inflation and Unemployment: a Re-Evaluation for the Scandinavian Countries”, Applied Economic Letters, 6, №1, 55-58.
239. Yule G.U. (1927) “On a Method of Investigating Periodicities in Disturbed Series”, Phil. Trans., A226, 227.
240. Zivot E., Andrews D.W.K. (1992) “Further Evidence on the Great Crash, the Oil Price Shock and the Unit Root Hypothesis”, Journal of Business and Economic Statistics, 10, 251-270.
241. (EHIPS) Генетические алгоритмы. Режим доступа [http://www.iki.rssi.ru/ehips/genetic.htm 29.08.2002]
242. Аргуткина Н.Л. О совершенствовании методов прогнозирования, основанных на экспоненциальном сглаживании. Конф. Маркетологов ВНПК «Прогнозирование» Режим доступа [http://www.marketing.spb.ru/conf]

243. Билл Вильямс Торговый Хаос – М.: ИК Аналитика, 2000. – 328 с. Режим доступа [http://yamdex.narod.ru/books/Williams/Tradxaos/Ссылки.htm]
244. Генетические алгоритмы и машинное обучение. Режим доступа [http://www.math.tsu.ru/russian/center/ai_group/ai_collection/docs/faqs/ai/part5/faq3.html]
245. Генетические алгоритмы обучения. Режим доступа [http://www.hamovniki.net/~alchemist/NN/DATA/Gonchar/Main.htm]
246. Лекции по нейронным сетям и генетическим алгоритмам. Режим доступа [http://infoart.baku.az/inews/30000007.htm]
247. Прогностика. Терминология, вып. 92. – М.: «Наука», 1978. Режим доступа [http://www.icc.jamal.ru/library/koi/POLITOLOG/bunchuk.txt]
248. Яковлев В.Л., Яковлева Г.Л., Лисицкий Л.А. Применение нейросетевых алгоритмов к анализу финансовых рынков. Режим доступа [http://neurnews.iu4.bmstu.ru/neurnews.html]
249. Яковлев В.Л., Яковлева Г.Л., Лисицкий Л.А. Создание математических моделей прогнозирования тенденций финансовых рынков, реализуемых при помощи нейросетевых алгоритмов. Режим доступа [http://neurnews.iu4.bmstu.ru/neurnews.html]

Приложение А

Методики оценки адекватности и точности прогноза

Для сравнения различных альтернативных прогнозов необходим критерий оценки качества прогноза. Используются следующие критерии.

1. Коэффициент несовпадения ретроспективного предсказания P_i с наблюдавшимися значениями A_i, предложенный Тейлом

$$L = \frac{\left[\frac{1}{n}\sum_{i=1}^{n}(P_i - A_i)^2\right]^{\frac{1}{2}}}{\left[\frac{1}{n}\sum_{i=1}^{n}P_i^2\right]^{\frac{1}{2}} + \left[\frac{1}{n}\sum_{i=1}^{n}A_i^{\,2}\right]^{\frac{1}{2}}}. \qquad (1)$$

Коэффициент легко вычисляется, и его значения принадлежат отрезку [0,1], причем на концах отрезка он имеет следующую содержательную интерпретацию: при $L = 0$ – отличное качество прогноза; при $L = 1$ – плохое качество прогноза. Может применяться как в случае статистического, так и в случае качественного прогнозирования.

Стандартное отклонение. Пусть $e_t = y_t - \widehat{y}_t$ – ошибка прогноза, причем $y_t \neq 0$, $\widehat{y}$ – прогноз. Определим среднее абсолютное отклонение ошибки – MAD_t сгладив экспоненциально ошибки e_t

$$MAD_t = \alpha\left|e_t\right| + (1-\alpha)MAD_{t-1.} \qquad (2)$$

Очевидно, что MAD_t неотрицательно. Оказывается, что для довольно большого класса распределений значение стандартного отклонения несколько больше значения среднего абсолютного отклонения и пропорционально ему. Константа пропорциональности для различных распределений колеблется между 1,2 и 1,3 (для нормального распределения – $\sqrt{\pi/2}$). То есть

$$\sigma_t \approx 1{,}25MAD_t. \qquad (3)$$

Имеем следующую процедуру оценки качества прогноза:

1. Вычислим ошибку прогноза e_t.
2. Вычислим значение MAD_t по формуле (2). Для прогноза экономических показателей удовлетворительной является $MAD_0 = 0{,}1S_0$, где S_0 – начальное значение экспоненциальной средней ряда.
3. Вычислим стандартное отклонение по формуле (3).
4. При относительно малом горизонте прогнозирования с достаточной степенью уверенности можно утверждать, что будущее значение прогнозируемого значения.

3. Среднеабсолютная процентная ошибка (Mean Absolute Percentage Error)

$$MAPE = \frac{1}{n}\sum_{t=1}^{n-1}\frac{|e_t|}{y_t}\cdot 100\% . \qquad (4)$$

Показатель *MAPE*, как правило, используется для сравнения точности прогнозов разнородных объектов прогнозирования, поскольку он характеризует относительную точность прогноза.

Для прогнозов высокой точности $MAPE < 10\%$, хорошей – $10\% < MAPE < 20\%$, удовлетворительной – $MAPE > 50\%$.

Целесообразно пропускать значения ряда, для которых $y_t = 0$.

4.Средняя процентная ошибка (Mean Percentage Error) и средняя ошибка (Mean Error).

MPE характеризует относительную степень смещенности прогноза. При условии, что потери при прогнозировании, связанные с завышением фактического будущего значения, уравновешиваются занижением, идеальный прогноз должен быть несмещенным, и обе меры должны стремиться к нулю. Средняя процентная ошибка не определена при нулевых данных и не должна превышать 5%

$$MPE = \frac{1}{n}\sum_{t=1}^{n-1}\frac{e_t}{y_t}100\% . \qquad (5)$$

Абсолютное смещение характеризует средняя ошибка

$$MPE = \frac{1}{n}\sum_{t=1}^{n-1}e_t . \qquad (6)$$

5.Средний квадрат ошибки (Mean Square Error) и сумма квадратов (Sum Square Error).

Соответственно

$$MSE = \frac{1}{n}\sum_{t=1}^{n-1} e_t^2 \ . \qquad (7)$$

и

$$SSE = \sum_{t=1}^{n-1} e_t^2 . \qquad (8)$$

Чаще всего *MSE* и *SSE* используется при выборе оптимального метода прогнозирования. В большинстве пакетов программ по прогнозированию именно эти два показателя принимаются в качестве критериев при оптимальном выборе параметров моделей.

Приложение Б

Варианты заданий для выполнения практических работ с использованием адаптивных моделей прогнозирования

Таблица Б1. Темпы инфляции. Месячные данные с 1996 г. по 2005 г. (%).

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Январь	6,2	245,0	25,8	17,9	17,8	4,1	2,3	1,4	8,4	2,3
Февраль	4,8	38,3	24,7	10,8	11,0	2,8	1,5	0,9	4,1	1,0
Март	6,3	29,9	20,1	7,4	8,9	2,8	1,4	0,6	2,8	0,6
Апрель	63,5	21,7	18,8	8,5	8,5	2,2	1,0	0,4	3,0	0,9
Май	3,0	12,0	18,1	6,9	7,9	1,6	0,9	0,5	2,2	1,8
Июнь	1,2	18,6	19,9	6,0	6,7	1,2	1,1	0,1	1,9	2,6
Июль	0,6	11,0	22,4	5,3	5,4	0,7	0,9	0,2	2,8	1,8
Август	0,5	8,6	25,8	4,6	4,6	-0,2	-0,1	3,7	1,2	1,0
Сентябрь	1,1	11,5	23,1	7,7	4,5	0,3	-0,3	38,4	1,5	
Октябрь	3,5	22,9	19,5	15,0	4,7	1,2	0,2	4,5	1,4	
Ноябрь	8,9	26,1	16,4	15,0	4,5	1,9	0,6	5,7	1,2	
Декабрь	12,1	25,4	12,5	16,4	3,2	1,4	1,0	11,6	1,3	

Таблица Б2. Экспорт. Месячные данные с 1999 г. по 2005 г.

	1999	2000	2001	2002	2003	2004	2005
Январь	4,09	5,64	5,80	7,00	6,00	4,60	6,80
Февраль	4,46	6,15	6,80	6,70	5,90	5,00	7,90
Март	4,76	6,76	7,80	7,40	6,80	5,90	8,70
Апрель	4,74	6,75	7,00	6,80	6,20	6,50	8,20
Май	5,83	6,86	7,50	6,70	6,10	5,10	
Июнь	6,34	7,02	6,90	6,90	6,60	5,30	
Июль	5,77	6,31	7,30	7,50	6,30	6,30	
Август	6,11	6,34	7,00	7,00	5,80	6,10	
Сентябрь	6,55	6,73	7,10	7,10	6,00	6,30	
Октябрь	6,01	7,10	8,60	7,90	6,10	6,80	
Ноябрь	6,31	7,73	8,10	8,30	5,90	7,40	
Декабрь	6,59	7,80	8,70	8,90	7,30	9,40	

Таблица Б3. Импорт. Месячные данные с 1999 г. по 2005 г.

	1999	2000	2001	2002	2003	2004	2005
Январь	3,55	3,70	4,80	4,80	5,90	2,80	2,60
Февраль	3,85	4,41	5,80	5,10	6,10	3,10	3,40
Март	4,14	4,86	6,00	5,70	6,60	3,60	3,70
Апрель	3,64	4,28	6,10	6,20	6,30	3,40	3,50
Май	4,06	4,72	5,70	5,50	5,90	3,00	
Июнь	4,35	5,22	5,50	5,80	5,90	3,40	
Июль	3,76	5,21	6,10	6,50	5,60	3,40	
Август	4,09	5,00	5,80	6,10	5,20	3,20	
Сентябрь	4,46	5,07	5,30	6,20	3,10	3,30	
Октябрь	4,33	5,46	5,70	6,90	3,10	3,50	
Ноябрь	4,77	6,30	5,60	6,50	3,10	3,60	
Декабрь	5,48	6,60	6,40	8,40	3,60	4,10	

Таблица Б4. Валовой внутренний продукт. Квартальные данные с 1999 г. по 2005 г. (млрд. руб.).

	1999	2000	2001	2002	2003	2004	2005
I квартал	88 000	253 000	456 000	539 000	551 000	837 000	1 389 100
II квартал	130 000	353 000	509 000	594 000	626 000	1 042 000	1 557 300
III квартал	168 000	443 000	570 000	679 000	694 000	1 276 000	
IV квартал	225 000	491 000	611 000	667 000	825 000	1 391 000	

Таблица Б5. Доходы федерального бюджета. Месячные данные с 1997 г. по 2005 г. (млрд. руб.).

	1997	1998	1999	2000	2001	2002	2003	2004	2005
Январь	40	756	2378	9327	13091	14660	18902	27758	64913
Февраль	74	779	2285	10953	14244	19021	19034	26925	73387
Март	88	1006	5402	13450	22778	22112	22832	34403	83524
Апрель	153	1310	3735	15855	14644	26927	22237	44830	92223
Май	112	1001	10604	15491	22533	26583	23331	39746	101450
Июнь	98	1531	5174	15037	27781	16821	21698	52921	
Июль	193	1626	4758	19422	22510	21979	22248	55544	
Август	199	1753	5903	25616	19978	34725	21525	52214	
Сентябрь	290	2145	7108	23767	20858	22175	20248	53011	
Октябрь	612	2066	8944	24424	22265	26005	23690	58280	
Ноябрь	444	3020	9253	32810	26409	30472	27591	69257	
Декабрь	666	4635	13308	20725	54909	61352	59052	96820	

Таблица Б6. Налоговые доходы федерального бюджета. Месячные данные с 1997 г. по 2005 г. (млрд. руб.).

	1997	1998	1999	2000	2001	2002	2003	2004	2005
Январь	32	729	2316	7798	11252	11460	15793	24579	56841
Февраль	69	729	2099	7698	10360	14667	15412	24060	65851
Март	76	914	3421	10062	16172	19487	18745	31444	73470
Апрель	150	1382	3824	13581	14118	24375	18882	39246	80890
Май	101	946	8580	14009	13037	23683	19049	33585	88130
Июнь	100	1135	4017	13254	17473	13522	17423	42333	
Июль	176	1438	5192	16216	18214	16505	18379	47658	
Август	183	1320	5373	15710	17241	15935	15520	42869	
Сентябрь	287	1407	6232	16936	17617	17409	15421	40239	
Октябрь	631	1977	8245	21579	18189	20352	19285	49602	
Ноябрь	405	2174	8655	18289	22756	20811	23928	57409	
Декабрь	648	2599	12376	15333	42294	45341	38147	72001	

Таблица Б7. Интенсивность промышленного производства. Сезонно скорректированные месячные данные с 1990 г. по 2005 г.

	1990	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Январь		94,42	80,53	69,57	50,42	47,26	42,48	40,34	41,90	40,96	46,26
Февраль		93,03	79,78	69,78	48,95	46,61	42,27	40,55	41,49	41,41	46,48
Март		91,35	79,10	69,42	48,06	45,89	42,08	40,77	41,17	41,96	46,48
Апрель		89,58	78,02	68,64	47,32	45,46	41,85	40,97	40,80	42,71	46,49
Май		88,24	76,27	67,50	46,60	45,48	41,61	41,24	40,18	43,59	46,66
Июнь		87,59	74,03	65,92	46,13	45,79	41,42	41,67	39,32	44,34	47,06
Июль		87,50	71,92	64,01	45,97	45,97	41,29	42,19	38,48	44,81	47,68
Август		87,40	70,34	62,06	46,05	45,73	41,21	42,62	38,00	44,94	
Сентябрь		86,74	69,31	60,20	46,31	45,05	41,08	42,82	38,10	44,87	
Октябрь		85,36	68,71	58,12	46,75	44,16	40,83	42,82	38,73	44,85	
Ноябрь		83,52	68,61	55,54	47,23	43,35	40,51	42,67	39,60	45,16	
Декабрь	95,44	81,77	69,00	52,75	47,48	42,80	40,30	42,35	40,38	45,73	

Приложение В

Статистика для выполнения практических работ с использованием многофакторных моделей прогнозирования

Социально - демографические показатели регионов России

где:

x1, x2 – рождаемость, смертность населения (на 1000 человек);

x3, x4 – браки и разводы на 1000 населения;

x5 – миграционный прирост населения на 10000 населения;

x6 – коэффициенты младенческой смертности (число детей, умерших в возрасте до 1 года, на 1000 родившихся);

x7 – удельный вес безработных в численности экономически активного населения (%), классифицируемых в соответствии с методологией МОТ;

x8 – денежные доходы с учетом индекса цен;

x9 – соотношение денежного дохода и прожиточного минимума (%);

x10 – соотношение средней оплаты труда с учетом выплат социального характера и прожиточного минимума трудоспособного населения (%);

x11 – численность населения с денежными доходами ниже прожиточного минимума в % от численности населения региона;

Таблица В1. Социально - демографические показатели регионов России

		x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11
1	2	3	4	5	6	7	8	9	10	11	12	13
	Российская Федерация	9,3	15	7,3	4,5	34	18,1	8,8	85,6	202	179	24,7
	Северный район	8,7	14,2	6,8	5	-43	18,5	9,5	84,1	175,6	168,2	23,5
1	Республика Карелия	8,5	16,3	6,8	5,6	23	17,4	11,6	84,7	163	151	23,6
2	Республика Коми	9,3	12,6	7,2	5,5	-101	25,3	11,7	74,6	194	239	19,2
3	Архангельская область	8,7	14,6	6,5	4,2	-31	16,2	11,3	71,4	152	192	26,9
4	Вологодская область	8,6	16,2	6,1	4	41	17,4	8,1	79,9	190	205	20,1
5	Мурманская область	8,1	11,4	7,7	6,4	-148	15,9	12,9	79,9	183	198	22
	Северо - западный район	7,2	17,3	7,7	5,3	50	14,9	11,2	85,5	196,5	186,2	21,2
6	г. Санкт - Петербург	7	15,9	8,2	5,1	13	13,8	9,8	101,1	229	172	20

1	2	3	4	5	6	7	8	9	10	11	12	13
7	Ленинградская область	7,2	18,2	7,4	6,1	120	14,3	11	76,6	146	167	29,1
8	Новгородская область	7,9	19,7	6,4	4,7	79	19,8	9,3	8,5	174	144	22,8
9	Псковская область	7,7	20,8	6,9	5,2	95	17,1	11,7	72,5	128	111	42,7
	Центральный район	7,7	17,3	7,6	4,8	56	16,6	11,5	75,5	134,1	121,2	35,2
10	Брянская область	9,2	15,9	7,8	5,3	68	16,7	9,3	68,9	169	148	22,7
11	Владимировская область	7,6	16,4	6,7	4,7	70	15,5	12,3	64,3	144	150	27,9
12	Ивановская область	7,3	18,3	6,3	4,9	46	19,6	14,9	68,1	138	133	33,7
13	Калужская область	7,9	16,4	6,8	5	109	17,6	8,3	59,7	197	155	26,6
14	Костромская область	7,9	17	6,3	4,4	52	20,1	8,7	72,7	182	159	30,5
15	г. Москва	8	16,9	8,2	4,6	28	15,5	5,2	98,7	520	197	19,1
16	Московская область	7,2	17,6	8,1	5,2	60	16,1	9,5	104,1	143	165	31,2
17	Орловская область	8,7	16	7,6	4,4	52	18,9	7,2	69,3	214	161	22,7
18	Рязанская область	7,8	17,9	7,2	4,3	47	15,7	6,4	60,7	158	163	24,4
19	Смоленская область	8	16,9	6,9	4,7	84	16,8	9,6	79	185	146	19,8
20	Тверская область	7,5	19,4	6,7	4,6	105	19,3	8	64	153	165	28,6
21	Тульская область	7,3	19,4	7,4	5	59	20,1	5,9	75,5	200	175	16,2
22	Ярославская область	7,6	17,3	7,1	5,3	65	12	11,5	66,8	180	154	21,3
	Волго - Вятский район	8,6	15,8	6,6	3,7	37	16,4	10,3	55,3	178,5	165,2	32,1
23	Республика Марий Эл	9,6	13	6,4	3,5	35	16,8	11,2	67,9	120	117	43,2
24	Республика Мордовия	9	14,1	7	3,3	11	15,2	10,3	75,2	132	126	34,7
25	Чувашская Республика	10,2	13	7,1	3,2	28	16,1	9,6	74,6	145	121	27,3
26	Кировская область	8,1	16,3	6,2	3,9	18	17,1	9,2	77,8	137	121	32
27	Нижегородская область	8	17,5	6,7	4	56	16,4	7,8	81,7	181	182	22
	Центральное - Черноземн	8,5	16,3	7,8	4,4	79	16,4	6,6	83,4	165	182	23,1
28	Белгородская область	9,4	14,8	8,1	5	129	14,7	5,6	85,3	200	195	19,9
29	Воронежская область	8,3	16,6	7,7	4,4	71	15,4	7,4	71,3	182	157	23,1
30	Курская область	8,5	16,7	8	4,1	68	17,1	5,9	68,1	179	177	20,2
31	Липецкая область	8,4	16,1	7,6	4,6	78	16,7	6,3	74,6	181	191	18,6
32	Тамбовская область	8,4	17,3	7,3	4,1	53	19,4	10	52,2	183	170	22
	Поволжский район	9,3	14,1	7,1	4,4	62	18,5	12,5	78,4	164	165	35,1
33	Республика Калмыкия	13,5	10,5	7,1	3,4	-69	15,8	19,7	58	100	120	60,3
34	Республика Татарстан	10,4	12,9	7	3,9	40	18,5	6,4	89,4	194	225	22,1
35	Астраханская область	10,1	13,5	7,1	4,6	79	18,6	13,1	62,8	143	137	32,3
36	Волгоградская область	9,1	14,6	7,5	5,1	90	19,1	10,3	71	141	160	33,2
37	Пензенская область	8,2	15	7,1	4	44	14,7	12,5	71	148	121	30,2
38	Самарская область	8,6	14,8	7,3	4,8	82	14	7,3	82,6	188	207	21,2

1	2	3	4	5	6	7	8	9	10	11	12	13
39	Саратовская область	8,9	14,5	6,9	4,2	57	23,6	9,6	72,3	138	117	35,3
40	Ульяновская область	8,9	13,4	6,7	3,9	65	21,8	7,8	68,4	198	206	16,3
	Северо - Кавказский район	12	13,6	7,9	4,1	49	19	8,9	53,8	165	138	26
41	Республика Адыгея	10,7	14,4	8	4	40	18,7	9,9	71,8	129	130	46,4
42	Республика Дагестан	21,8	7,5	6,9	1,3	20	17,6	22,3	64,2	86	79	43
43	Кабардино – Балкарская Республика	13,7	10,4	7,1	3,4	-33	14,5	14,7	72,3	128	102	42,5
44	Карачаево - Черкесская Республика	12,9	10,3	7	3,3	-11	16,3	24	56,1	123	107	45,7
45	Республика Северная Осетия	13,3	13	6,6	2,6	62	17,8	24	65,3	128	101	42,8
46	Краснодарский край	10	15,3	8,8	5	132	19,2	8,8	77,1	175	160	32,4
47	Ставропольский край	10,7	13,5	8,1	4,5	91	21,7	9,2	77	151	154	39,6
48	Ростовская область	9,2	15,8	8	4,8	58	18,7	8,2	75,6	146	140	33,4
	Уральский район	9,5	14,5	6,9	4,2	36	18,3	7,5	65,2	156,2	133	27,5
49	Республика Башкортостан	11,2	12,7	7,3	3,7	55	18,3	7,3	79,2	158	191	32,4
50	Удмуртская Республика	9,4	13,7	6,7	3,4	32	18,4	11,2	68,9	158	144	26,1
51	Курганская область	9	14,6	7,3	4,3	15	22,6	8,5	58,9	113	132	50,4
52	Оренбургская область	10,3	13,5	7,5	4	56	19,7	6,9	70,8	115	145	49,3
53	Пермская область	9,2	15,8	5,9	3,8	18	18,9	8,6	83,3	184	175	25,7
54	Свердловская область	8,5	15,6	6,7	4,7	36	17,5	8,5	87,9	163	169	29,5
55	Челябинская область	9	14,8	7	4,7	28	16,6	8,3	85,2	171	182	27,9
	Западно - Сибирский	9,4	13,5	7,3	4,9	33	19,3	9,8	79,8	175,4	164	24,3
56	Республика Алтай	14,2	13,1	7,3	3,8	70	27,9	11,3	66,8	188	148	26,2
57	Алтайский край	8,7	14,7	7,3	4,4	34	20,8	10,8	78,1	158	146	33,7
58	Кемеровская область	8,9	16,6	7	4,9	30	19,6	6,6	84	254	260	16,1
59	Новосибирская область	8,5	14,1	7	4,5	56	15,9	9,5	101,4	136	156	39,8
60	Омская область	10,2	12,3	7,3	4,6	6	16,3	5,2	92,8	157	170	29,7
61	Томская область	9,1	13	7	5,3	23	21,3	8,5	83,2	173	190	30,6
62	Тюменская область	10,6	9,8	7,9	5,7	34	21,3	6,1	75,6	290	293	19,2
	Восточно – Сибирский	11	13,7	6,8	4	4	19,6	11,4	36	165,4	214	18,2
63	Республика Бурятия	11,7	12	6,5	3,5	0	15,2	13,7	83,5	122	155	55,2
64	Республика Тыва	20	13	5,9	1,9	-16	28	14,7	53,8	84	101	73,2
65	Республика Хакасия	9,9	14	7,1	4,4	62	24,6	9,6	75,1	161	201	25,3
66	Красноярский край	9,8	14	7,2	4,8	7	19,8	9	69,7	246	296	24,2
67	Иркутская область	10,6	14,6	6,3	3,3	7	18,1	9,2	69,2	170	215	32,3
68	Читинская область	12,2	12,8	6,9	4	-25	20,8	10,2	80,6	99	112	66,5
	Дальневосточный	10,2	12,6	7,1	5,3	-136	20,5	7,6	74,3	186,4	165	63

1	2	3	4	5	6	7	8	9	10	11	12	13
69	Республика Саха (Якутия)	15,3	9,8	8	4,7	-182	19,5	6,4	74,2	170	201	29,2
70	Еврейская автономная область	10,9	13,6	7,3	5,2	-66	26,4	15,9	76,4	171	185,9	29,4
71	Чукотский автономный округ	9,8	8,6	7,3	8,9	-978	34	5,2	64,6	162	195,2	28,7
72	Приморский край	9,4	13,1	6,6	4,7	-42	21,5	10,7	77,1	144	170	31,8
73	Хабаровский край	9,3	13,1	6,6	5,7	-69	17,8	11,6	82,4	153	171	29,4
74	Амурская область	10,1	12	7,2	4,9	-11	23,6	12,5	83,5	175	187	37,9
75	Камчатская область	9,1	11,2	7,9	6,7	-280	15,4	8,5	88	211	228	22,7
76	Маганданская область	8,3	10,9	7,2	7,1	-759	14,2	10,4	64,4	202	187	24,6
77	Сахалинская область	8,9	17	7,2	5,7	-301	22,7	12,7	58,2	145	169	24,6
78	Калининградская область	8,6	13,6	7,8	6	113	15,4	9,4	75,7	155	145	26,6

Приложение С

Статистика по Ставропольскому краю

Сельское хозяйство Ставрополья: Статистический сборник / Территориальный орган Федеральной службы государственной статистики по Ставропольскому краю - 2005г. – 103с.

Таблица С1. Реализация основной продукции сельского хозяйства сельхозпредприятиями Ставропольского края (тысяч тонн).

	1991	1992	1993	1994	1995	1996	1997
1	2	3	4	5	6	7	8
Зерновые культуры-всего	2777,2	2604,6	2440,0	2193,7	1956,8	1888,9	2244,6
из них:							
пшеница		2172,0	2056,2	1897,3	1619,3	1487,8	1854,4
рожь		0,1	0,2	0,2	0,7	0,9	0,5
просо		92,6	55,9	12,7	13,3	16,1	22,5
гречиха		6,9	5,7	5,7	3,6	6,0	5,6
кукуруза		147,6	93,6	66,0	40,0	67,0	103,2
ячмень		141,5	155,5	125,1	173,5	194,7	154,1
овес		8,0	13,4	8,5	10,1	8,1	10,4
зернобобовые		11,6	13,5	19,0	17,3	25,7	17,2
в т.ч.горох		11,3	13,4	17,6	17,0	25,6	17,0
прочие зерновые		24,3	46,0	59,2	79,0	82,6	76,7
Масличные культуры	191,8	123,6	165,7	182,5	166,5	260,4	186,0
в т.ч. подсолнечник	163,2	113,3	162,1	179,8	146,5	249,7	175,9
Сахарная свекла		218,6	247,8	14,3	91,4	221,7	228,5
Картофель	50,4	43,4	32,0	20,9	12,3	7,7	11,3
Овощи -всего	168,2	113,7	97,8	76,6	60,0	43,8	38,7
из них:							
помидоры	47,6	21,6	25,4	30,7	18,4	5,8	5,6
огурцы	10,9	7,8	5,3	5,6	4,4	4,0	4,0
лук репчатый	25,2	20,9	13,1	12,2	13,3	12,1	11,0
капуста	40,4	34,9	26,6	14,9	12,7	12,4	10,7
морковь	6,6	4,5	4,0	1,9	2,3	2,0	1,7
свекла столовая		6,7	8,2	4,4	3,1	2,9	3,0
Бахчевые культуры	38,6	17,2	10,8	8,3	14,7	6,5	5,0
Плоды и ягоды- всего	66,7	54,1	39,0	25,0	18,5	18,2	22,3
в т.ч.косточковые	10,5	18,1	7,6	2,8	2,9	3,6	4,1
семечковые	55,4	35,4	30,7	21,9	15,4	14,4	17,9
ягоды	0,8	0,6	0,7	0,3	0,2	0,2	0,3
Виноград	28,9	33,5	21,1	13,3	30,7	23,8	18,2
Скот и птица(в живом весе) - всего	284,1	236,2	185,6	153,5	91,9	98,3	80,1
из них:							
крупный рогатый скот	100,8	94,9	67,4	66,2	42,6	45,7	39,5
овцы и козы	60,3	52,1	52,3	39,3	21,0	22,6	18,5
свиньи	57,9	50,6	28,5	19,2	13,8	14,9	11,0

1	2	3	4	5	6	7	8
птица	64,3	37,8	31,9	23,0	10,4	10,4	9,9
прочие виды скота	0,8	0,8	5,5	5,9	4,1	4,7	1,2
Молоко	712,1	515,9	436,7	364,8	289,8	243,1	189,3
Яйца, млн.штук	450,3	347,4	296,7	274,5	238,5	260,4	216,9
Шерсть (в физическом весе), тонн	27954	13307	18285	19195	10046	8542	6809
Мед,тонн	130	85	111	74	57	62	65

Продолжение

	1998	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Зерновые культуры-всего	2336,3	1806,8	2284,8	2762,8	3873,9	2565,7	3436,4
из них:							
пшеница	1957,0	1424,0	1749,9	2041,2	2982,2	2045,5	2790,6
рожь	0,3	0,2	1,2	1,4	1,4	0,8	1,4
просо	33,3	16,0	12,4	22,6	28,2	18,5	53,7
гречиха	3,2	0,8	2,0	1,8	5,8	2,2	4,4
кукуруза	86,2	32,8	40,1	32,0	54,1	113,1	129,8
ячмень	141,7	232,5	354,5	507,5	598,7	247,1	306,8
овес	9,2	9,2	28,2	38,6	19,9	17,0	10,3
зернобобовые	15,3	8,8	10,9	23,5	56,6	18,3	40,7
в т.ч.горох	15,1	8,7	10,8	21,1	54,9	17,5	40,0
прочие зерновые	90,1	71,7	85,6	94,2	126,8	102,5	98,6
Масличные культуры	197,9	132,0	153,6	128,9	142,2	190,6	259,8
в т.ч. подсолнечник	176,9	121,5	128,5	108,5	126,6	156,20	183,1
Сахарная свекла	181,9	156,3	197,4	205,5	382,8	301,0	780,5
Картофель	6,7	3,8	6,8	3,9	4,6	2,6	5,9
Овощи -всего	34,1	34,8	27,1	31,7	30,4	19,6	19,4
из них:							
помидоры	8,2	5,6	3,7	6,2	4,5	4,3	4,3
огурцы	4,1	5,3	5,2	4,8	4,2	4,1	4,4
лук репчатый	9,2	10,8	8,8	9,5	10,0	4,9	6,5
капуста	6,7	8,4	5,2	6,2	5,0	2,8	2,1
морковь	1,0	1,0	1,0	0,9	1,7	0,9	0,9
свекла столовая	2,9	2,0	1,5	2,0	1,7	1,1	0,9
Бахчевые культуры	4,4	2,8	3,0	1,5	2,6	1,8	1,3
Плоды и ягоды- всего	11,9	9,4	11,2	11,8	11,6	9,8	6,8
в т.ч.косточковые	0,6	2,4	3,2	1,9	1,3	1,1	0,7
семечковые	11,1	6,8	7,7	9,4	9,9	8,5	5,9
ягоды	0,2	0,2	0,3	0,5	0,4	0,2	0,2
Виноград	9,2	16,4	21,7	7,2	12,7	6,9	16,4
Скот и птица(в живом весе)- всего	64,9	55,8	68,1	68,2	70,1	87,2	83,2
из них:							
крупный рогатый скот	28,2	20,3	23,5	21,5	20,4	24,5	21,5
овцы и козы	14,6	9,0	9,7	10,2	9,5	10,8	10,7
свиньи	9,2	9,1	11,4	8,5	11,1	16,3	14,0
птица	12,0	17,1	23,1	27,7	28,8	35,2	36,5
прочие виды скота	0,9	0,3	0,4	0,3	0,3	0,4	0,5
Молоко	165,8	154,4	145,3	140,6	145,1	128,3	110,4
Яйца, млн.штук	171,9	136	174,8	225,6	389,2	267,9	205,8

1	2	3	4	5	6	7	8
Шерсть (в физическом весе),тонн	4804	3796	4057	3692	3439	3758	3248
Мед,тонн	70	20	28	20	20	18	17

Таблица С2. Производство основных продуктов животноводства в Ставропольском крае в 1940 - 2004г.г. (в хозяйствах всех категорий; тысяч тонн)

Годы	**Скот и птица на убой в живом весе**	**Молоко**	**Яйца млн.штук**	**Шерсть**
1	**2**	**3**	**4**	**5**
1940	128,5	333,1	315,8	11,1
1950	87,7	292,8	189,5	8,4
1955	148,5	439,6	435,2	19,0
1960	198,6	643,3	581,5	25,2
1965	285,5	855,9	807,1	29,5
1970	270,2	829,5	985,2	31,2
1971	338,0	846,8	1078,6	32,3
1972	334,2	844,3	1100,8	30,6
1973	334,6	940,0	1168,8	29,0
1974	350,9	953,0	1279,6	29,8
1975	337,7	936,4	1194,1	30,5
1976	246,3	719,7	865,7	27,7
1977	273,4	771,0	1009,9	29,1
1978	305,5	820,4	1105,5	30,0
1979	340,7	825,0	1109,2	33,7
1980	311,6	806,4	1123,0	28,3
1981	344,2	844,5	1226,7	30,0
1982	350,4	843,3	1248,7	29,4
1983	363,6	867,3	1293,0	30,4
1984	366,1	853,8	1316,4	30,3
1985	359,3	871,6	1334,6	30,7
1986	400,8	919,6	1377,3	31,7
1987	424,5	955,7	1414,2	32,9
1988	439,7	1016,3	1462,4	32,9
1989	467,5	1054,8	1415,0	34,0
1990	477,0	1066,1	1306,6	33,3
1991	431,9	1014,2	1247,6	30,2
1992	396,8	856,0	1099,9	28,6
1993	367,9	811,0	1004,7	20,2
1994	319,2	786,5	989,3	16,7
1995	267,5	732,1	909,8	14,0
1996	262,0	654,1	920,2	13,4
1997	236,7	572,9	792,1	10,1
1998	207,3	524,9	707,5	7,1
1999	193,8	526,7	660,0	6,0
2000	194,6	542,8	699,3	6,2
2001	197,1	544,6	747,3	5,8
2002	204,0	553,4	903,7	5,5
2003	224,0	568,9	782,2	6,0
2004	231,5	544,2	706,4	6,0

Таблица С3. Структура инвестиций в основной капитал агропромышленного комплекса Ставропольского края по источникам финансирования (миллионов рублей)

Наименование	1998	1999	2000	2001	2002	2003	2004
1	2	3	4	5	6	7	8
Всего	689,3	1089	1361,8	3467,5	3969,9	4834,1	5454,6
из них:							
по предприятиям (без малого предпринимательства)	408,8	718,6	888,0	1448,6	1727,9	2256,2	3035,4
в том числе:							
собственные средства	324,8	542,9	767,2	1066,9	1226,8	1324,5	1887,9
из них: прибыль	95,8	244,3	407,7	476,5	642,6	709,1	997,0
амортизация			129,5	300,5	354,4	488,9	511,8
привлеченные средства	84,0	175,7	120,8	381,7	501,1	931,7	1147,5
из них:							
кредиты банков	16,7	0,4	21,3	212,2	358,2	698,2	601,4
заемные средства других организаций	13,2	10,9	15,3	58,2	55,6	60,2	202,5
бюджетные средства	36,7	51,1	72,1	103,3	73,5	130,3	182,5
в том числе за счет:							
федерального бюджета	33,5	37,5	58,5	58,0	34,8	63,6	95,4
и местных бюджетов	3,2	13,6	13,6	45,3	38,7	66,7	87,1
средства внебюджетных фондов	0,6	14,1	10,1	1,8	2,0	1,3	1,4
прочие	16,8	99,2	2,0	6,3	11,8	41,7	159,7
из них:							
средства вышестоящих организаций					1,2	16,7	28,8
средства от эмиссии акций		63,3		5,0		4,7	

Таблица С4. Продукция сельского хозяйства Ставропольского края (в фактических действовавших ценах, миллионов рублей)

	1990	1991	1992	1993	1994	1995	1996	1997
Хозяйства всех категорий								
Всего	3,5	7,8	83,6	564,3	1656,1	5179,4	7280,8	8778,9
Растениеводство	1,6	3,1	45,6	294,3	875,2	3060,7	3573,5	4042,5
Животноводство	1,9	4,7	38,0	270,0	780,9	2118,7	3707,3	4736,4
Сельхозпредприятия								
Всего	2,8	5,6	59,4	374,2	1088,7	3297,5	3849,5	3806,3
Растениеводство	1,5	2,5	40,0	223,6	635,4	2392,1	2554,8	2881,2
Животноводство	1,3	3,2	19,4	150,7	453,3	905,4	1294,7	925,1
Хозяйства населения								
Всего	0,7	2,1	23,6	151,9	501,7	1657,6	3107,8	4596,5
Растениеводство	0,1	0,6	5,5	37,9	191,4	483,3	787,8	881,7
Животноводство	0,6	1,5	18,1	114,0	310,3	1174,3	2320,0	3714,8

Крестьянские (фермерские) хозяйства

Всего		0,012	0,6	38,2	65,7	224,3	323,5	376,1
Растениеводство		0,009	0,14	32,8	48,4	185,3	230,9	279,6
Животноводство		0,003	0,5	5,4	17,3	39,0	92,6	96,5

Продолжение

	1998	1999	2000	2001	2002	2003	2004(пред.)
Хозяйства всех категорий							
Всего	9174,2	15605	19636	24825	27068	30174,8	46280,1
Растениеводство	3368,2	7707,5	10553	14477	15790	18051,1	32601,1
Животноводство	5806,0	7897,8	9082,8	10349	11278,0	12123,7	13679,0
Сельхозпредприятия							
Всего	3351,5	7474,7	10311	14003	14539,0	15059,9	27109,8
Растениеводство	2471,9	5589,4	7908,1	10993	10865	11551,1	22731,8
Животноводство	879,6	1885,3	2403,1	3009,4	3674,3	3508,8	4378,0
Хозяйства населения							
Всего	5480,2	7457,7	8222,4	9126,1	10219,1	11756,9	13803,7
Растениеводство	684,4	1610,8	1772,0	2066,9	2921,9	3485,6	4927,6
Животноводство	4795,8	5846,9	6450,4	7059,2	7297,2	8271,3	8876,1
Крестьянские (фермерские) хозяйства							
Всего	342,5	672,9	1101,9	1696,5	2309,4	3357,9	5366,6
Растениеводство	211,9	507,3	872,6	1416,5	2002,9	3014,4	4941,8
Животноводство	130,6	165,6	229,3	280,0	306,5	343,5	424,8

Научное издание

Тихонов Э.Е.

Практика прогнозирования финансовых рынков

Учебное пособие

Редактор Тихонов Э.Е.
Корректура автора