**ГЛАВА 3. Метод прогнозирования временных рядов по выборке максимального подобия с разными уравнениями аппроксимации для положительных и отрицательных значений.**

Метод прогнозирования временных рядов по выборке максимального подобия с разными уравнениями аппроксимации для положительных и отрицательных значений основывается на следующих алгоритмах:

1. Алгоритм прогнозирования по выборке максимального подобия без учета внешних факторов.
2. Алгоритм прогнозирования по выборке максимального подобия с учетом значимых внешних факторов.

**3.1. Алгоритм прогнозирования по выборке максимального подобия без учета внешних факторов**

При прогнозировании временных рядов часто используют интеграцию первого порядка, то есть прогнозирование не самого ряда или процесса, а его изменения или прироста.

Перейдем к формальному описанию постановки задачи.

Пусть дан временной ряд ,  – уровни временного ряда,  – временные метки (моменты или интервалы наблюдений), n – количество уровней временного ряда. Тогда ряд отклонений  получается путем вычитания из уровня ряда  предыдущего уровня ряда: .

Алгоритм прогнозирования по выборке максимального подобия с разными уравнениями аппроксимации для положительных и отрицательных значений без учета внешних факторов состоит из следующих шагов:

*Шаг 1*. Определение выборки предшествующей прогнозу.

*Шаг 2.* Разбиение выборки предшествующей прогнозу на две подвыборки положительных и отрицательных значений.

*Шаг 3.* Определение выборки аппроксимации.

*Шаг 4*. Определение выборки прогноза, т.е. отрезка временного ряда следующего после выборки аппроксимации длиной прогноза.

*Шаг 5.* Разбиение выборки аппроксимации. Выборка аппроксимации разбивается на две выборки отрицательных и положительных значений.

*Шаг 6*. Разбиение выборки прогноза. Отрезок временного ряда следующий после выборки аппроксимации длиной прогноза на положительные и отрицательные значения.

*Шаг 7*. Построение уравнений аппроксимации для положительных и отрицательных значений выборки аппроксимации.

*Шаг 8.* Получение прогноза. Непосредственная аппроксимация положительных значений выборки прогноза с помощью уравнения для положительных значений, отрицательных значений – с помощью уравнения для отрицательных элементов выборки.

Перейдем к формальному описанию каждого шага алгоритма:

*Шаг 1. Определение выборки предшествующей прогнозу.*

Во временном ряду ,  прогнозному значению предшествует выборка длиной *М*, причем .

*Шаг 2. Разбиение выборки предшествующей прогнозу.*

Выборка, предшествующая прогнозу , разбивается на две подвыборки , длиной  – для положительных значений и , длиной  – для отрицательных.

*Шаг 3. Определение выборки аппроксимации.*

Моделирование временных рядов при помощи выборок основано на предположении, что временной ряд представляет собой последовательность выборок [Singh, S. Pattern Modelling in Time-Series Forecasting / S. Singh // Cybernetics and Systems-AnInternational Journal. 2000, Vol. 31, No. 1. P. 49-65.]. Основываясь на данном предположении, допустим, что в исследуемом временном ряду  существует выборка  с началом отсчета в точке  длиной *М*.

Определение выборки  осуществляется по следующему алгоритму:

1. Значения выборки  последовательно сравниваются со значениями временного ряда  с лагом в единицу.
2. Для каждого сдвига считается значение коэффициента корреляции:



1. Из полученных значений коэффициентов корреляции выбирается наибольшее и определяется, какому значению лага  оно соответствует.

*Шаг 4. Определение выборки прогноза.*

Если необходимо спрогнозировать *р* уровней временного ряда, то выборка прогноза является отрезком временного ряда, следующим после  длиной *р*, т.е. .

*Шаг 5. Разбиение выборки аппроксимации*.

Так как во временном ряду значения могут быть и положительными и отрицательными, то вычисление аппроксимации по классической линейной модели не подходит. Полученная выборка  разбивается на две:  – для положительных значений выборки и  – для отрицательных значений выборки. Выборка  длиной , выборка  длиной .

*Шаг 6. Разбиение выборки прогноза.*

Выборка прогноза  разбивается на две подвыборки , длиной  для положительных значений выборки и , длиной  для отрицательных.

*Шаг 7. Построение уравнений аппроксимации для положительных и отрицательных значений выборки аппроксимации*.

Для полученных выборок аппроксимации  и  строятся уравнения аппроксимации:





где  и  – единичные векторы, а , , ,  – коэффициенты аппроксимации, их значение должно быть таким, чтобы

,



то есть квадрат отклонений модельных значений от реальных должен быть минимален.

Значение коэффициентов уравнений находятся с помощью метода наименьших квадратов:

,

где , , 

,

где , , 

*Шаг 8. Получение прогноза*.

Для получения прогноза каждое значение подвыборки  аппроксимируется с помощью уравнения , а значения подвыборки  с помощью уравнения . Полученные прогнозные значения выстраиваются в том же порядке, в котором были соответствующие значения в выборке .

Значение параметра М определяется по алгоритму идентификации модели, предложенному И.А. Чучуевой [Чучуева И. А. Модель экстраполяции временных рядов по выборке максимального подобия // Информационные технологии. 2010. №12. С. 43 – 47.]:

1) Исходный временной ряд  разделяется на три части в пропорции 2:2:1.

Полученные части:

* *базовый период* (40%),
* *тестовый период*(40%),
* *контрольный период* (20%) временного ряда соответственно.

2) Исходя из решаемой задачи прогнозирования, определяется значение *P*, а также диапазон возможных значений параметра *M.* Первоначально рекомендуется брать широкий диапазон возможных значений *M*, например, *M* ∈[*P*,30 ⋅ *P*], а затем его последовательно уточнять.

3) Для каждого значения параметра *M* из установленного диапазона, прогнозируется значения временного ряда внутри *тестового периода*.

4) По результатам прогнозирования определяем среднюю абсолютную ошибку прогноза (mean absolute error, *MAE*)



где K – количество значений временного ряда, попавших в *тестовый период*.

5) Строится график зависимости *MAE* от *M* для *тестового периода* и определяется диапазон значений *M*, соответствующий устойчивому минимуму *MAE*.

6) На последнем шаге экспертом выбирается параметр модели *М* внутри диапазона устойчивого минимума.

**3.2. Алгоритм прогнозирования по выборке максимального подобия с учетом значимых внешних факторов.**

Методы для определения значимых внешних факторов подробно описаны в главе 2.

При введение в модель внешних факторов формальная постановка задачи принимает вид:

Пусть дан временной ряд ,  – уровни временного ряда,  – временные метки (моменты или интервалы наблюдений), n – количество уровней временного ряда. Каждому значению уровня временного ряда  соответствуют значения внешних факторов . Исходные данные удобнее представлять в виде таблицы, пример которой приведен в таблице 3.1.

Таблица 3.1 – Исходные данные при прогнозировании временных рядов с учетом внешних факторов.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *i* |  |  |  |  | … |  |
| 1 |  |  |  |  | … |  |
| 2 |  |  |  |  | … |  |
| … | … | … | … | … | … | … |
| n |  |  |  |  | … |  |

Каждый из факторов представляет собой временной ряд.

Так как в рассматриваемом алгоритме прогнозируются не сами уровни ряда, а отклонения, то исходный зависимый временной ряд принимает вид: . Каждый объясняющий фактор так же представляется в виде временного ряда отклонений: .

Алгоритм прогнозирования по выборке максимального подобия с разными уравнениями аппроксимации для положительных и отрицательных значений с учетом внешних факторов состоит из следующих шагов:

*Шаг 1*. Определение прогнозного значения каждого фактора.

*Шаг 2*. Определение выборок предшествующих прогнозу.

*Шаг 3.* Разбиение выборок предшествующих прогнозу на подвыборки положительных и отрицательных значений.

*Шаг 4.* Определение выборок аппроксимации. (в Y и Х)

*Шаг 5*. Определение выборок прогноза, т.е. отрезка временного ряда следующего после выборки аппроксимации длиной прогноза в каждом временном ряду показателя и факторов.

*Шаг 6.* Разбиение выборок аппроксимации. Выборки аппроксимации разбиваются на подвыборки отрицательных и положительных значений.

*Шаг 7*. Разбиение выборок прогноза. Отрезки временных рядов следующие после выборок аппроксимации длиной прогноза разбиваются на положительные и отрицательные значения.

*Шаг 8*. Построение уравнений аппроксимации для положительных и отрицательных значений выборки аппроксимации с учетом внешних факторов.

*Шаг 9.* Получение прогноза. Непосредственная аппроксимация положительных значений выборки прогноза с помощью уравнения для положительных значений, отрицательных значений – с помощью уравнения для отрицательных элементов выборки.

Перейдем к формальному описанию каждого шага алгоритма:

*Шаг 1. Определение прогнозного значения каждого фактора.*

Так как каждый из факторов в отдельности является временным рядом, то получение прогнозных значений возможно с помощью алгоритма прогнозирования по выборке максимального подобия с разными уравнениями аппроксимации для положительных и отрицательных значений. Если фактор представлен в виде фиктивной переменной, то его значение находится эмпирически.

*Шаг 2. Определение выборки предшествующей прогнозу.*

Выборка, предшествующая прогнозу определяется в каждом из временных рядов: зависимом и факторов. Длина выборки M определяется по алгоритму идентификации модели по зависимому временному ряду . Определяются следующие выборки:



*Шаг 3. Разбиение выборки предшествующей прогнозу.*

Каждая из полученных на шаге 2 выборка разбивается на две подвыборки для положительных и отрицательных значений.

Предшествующая прогнозу , разбивается на две подвыборки , для положительных значений и  – для отрицательных.  разбивается на  и  для положительных и отрицательных значений соответственно.

Длина всех положительных выборок: , отрицательных: .

*Шаг 4. Определение выборок аппроксимации.*

Во временном ряду показателя  находится выборка  с началом отсчета в точке  длиной *М*.

Определение выборки  осуществляется по следующему алгоритму:

1. Значения выборки  последовательно сравниваются со значениями временного ряда  с лагом в единицу.
2. Для каждого сдвига считается значение коэффициента корреляции:



Из полученных значений коэффициентов корреляции выбирается наибольшее и определяется, какому значению лага  оно соответствует.

Так как длины временных рядов факторов и показателя одинаковы, то в каждом из рядов факторов находится выборка длиной *М* и началом отсчета в точке *k*: .

*Шаг 5. Определение выборок прогноза.*

Если необходимо спрогнозировать *р* уровней временного ряда, то выборкой прогноза является отрезок временного ряда, следующий после выборки аппроксимации длиной *р*. Так для ряда показателя, выборка аппроксимации , а выборка прогноза . Для временных рядов факторов выборка прогноза: 

*Шаг 6. Разбиение выборок аппроксимации.*

Выборка аппроксимации показателя  разбивается на две:  – для положительных значений выборки и  – для отрицательных значений выборки. Выборки аппроксимации факторов  разбиваются на  и .

Длины выборок для каждого временного ряда  и  для положительных и отрицательных значений соответственно.

*Шаг 7. Разбиение выборок прогноза.*

Выборка прогноза  разбивается на две подвыборки  для положительных значений и  для отрицательных. Аналогично разбиваются выборки прогноза факторов  на  и .

Длины выборок для каждого временного ряда  и  для положительных м отрицательных значений соответственно.

*Шаг 8. Построение уравнений аппроксимации для положительных и отрицательных значений выборки аппроксимации с учетом внешних факторов.*

На основе полученных выборок аппроксимации , , ,  строятся уравнения аппроксимации:





где  и  – единичные векторы, а ,  – коэффициенты аппроксимации, где . Значение коэффициентов аппроксимации должно быть таким, чтобы:

,



то есть квадрат отклонений модельных значений от реальных должен быть минимален.

Значение коэффициентов аппроксимации находятся с помощью метода наименьших квадратов:

,

где , 



,

где , 



*Шаг 9. Получение прогноза.*

Для получения прогноза каждое значение подвыборки  аппроксимируется с помощью уравнения:



а значения подвыборки  с помощью уравнения:



Полученные прогнозные значения выстраиваются в том же порядке, в котором были соответствующие значения в выборке .

**3.3. ВЫВОДЫ**

1. Разработан метод прогнозирования отклонений временного ряда по выборке максимального подобия с разными уравнениями аппроксимации для положительных и отрицательных значений.
2. Построена математическая модель прогнозирования по указанному методу.
3. Разработаны алгоритмы прогнозирования отклонений временного ряда с учетом и без учета значимых внешних факторов. Временные ряды факторов для повышения точности прогнозирования представляются в виде рядов отклонений.
4. Модель позволяет учитывать неограниченное количество внешних факторов, что является неоспоримым преимуществом при прогнозировании экономических процессов.

.