ПРОГНОЗИРОВАНИЕ НА ОСНОВЕ ВРЕМЕННЫХ РЯДОВ С УЧЕТОМ ДИСБАЛАНСОВ ПРЕДЫДУЩИХ ПЕРИОДОВ

В. Л. Литвинов, К. В. Литвинова

Самарский государственный технический университет, Россия
Московский государственный университет им. М. В. Ломоносова, Россия
E-mail: vladlitvinov@rambler.ru, kristinalitvinova900@rambler.ru

Используя метод наименьших квадратов разработана математическая модель прогнозирования временных рядов, позволяющая учитывать суммарные дисбалансы предыдущих периодов. Определены параметры модели, при которых дисперсия прогнозирования будет наименьшей. В среде Matlab разработана компьютерная программа, предназначенная для математического моделирования и получения численных характеристик прогнозных значений. Для оценки достоверности модели использован критерий Фишера.

FORECASTING ON THE BASIS OF TIME SERIES INCLUDING PREVIOUS PERIOD IMBALANCES

V. L. Litvinov, K. V. Litvinova

Using the least squares method, a mathematical model for predicting time series has been developed, which allows taking into account the total imbalances of previous periods. The model parameters are determined for which the forecasting variance will be the smallest. In the Matlab environment, a computer program has been developed for mathematical modeling and obtaining numerical characteristics of predictive values. Fisher's criterion was used to assess the reliability of the model.

Во все времена особую актуальность имело прогнозирование социально–экономических и природных явлений. Правильный прогноз особенно важен, т.к. он позволяет предугадать возможную динамику прогнозируемых величин, проследить их изменение в будущем.

В настоящее время для прогнозирования применяются различные методы, начиная от простой экстраполяции до разработки и применения математических моделей. Одним из возможных подходов к решению подобного рода задач является применение метода множественной регрессии [2]. Суть этого подхода состоит в том, что на основании многолетних данных о величинах и ряда показателей строится уравнение множественной регрессии, связывающее значения с уровнями перечисленных факторов.

Для прогнозирования могут применяться самые разные математические функции. Наиболее часто, используются линейная, экспоненциальная и логистическая. При этом прогнозирование, основанное на применении линейной и экспоненциальной функций, иногда чисто условно называют экстраполяционным методом, а прогнозирование, основанное на применении логистической функции, – аналитическим методом.

Разработанная в настоящей статье методика прогнозирования, основанная на методе наименьших квадратов, позволяет учитывать суммарные дисбалансы предыдущих периодов. В большинстве случаев процессы, протекающие в природе и обществе, не предполагают резких изменений. Резкие изменения, происходящие в какой-то период, должны впоследствии сгладиться. Чем больше накапливается отклонений от среднего значения, тем сильнее действуют выравнивающие факторы. Если в предыдущие периоды исследуемая величина была больше нормы, то в последующие следует ожидать её снижение и наоборот. В связи с высокой трудоёмкостью расчётов была составлена и использована компьютерная программа в среде Matlab.

Пусть дан некоторый временной ряд:

Временной ряд

i	1	2	3	…	n
x_i	x_1	x_2	x_3	…	x_n

Здесь i – номер периода времени; n – общее число периодов; x_i – числовое значение некоторого показателя.

Первоначально необходимо определить общую тенденцию изменения временного ряда. В данной статье такая зависимость взята в линейном виде:

$$\overline{x}_i = P_0 + P_1 \cdot i, \tag{1}$$

где параметры P_0 и P_1 находятся по методу наименьших квадратов (МНК). Для нахождения $\overline{x}_i$ может быть использована любая функциональная зависимость, например, экспоненциальная.

Зависимость (1) определяет некоторую среднюю тенденцию изменения x_i. В отдельные периоды времени x_i может отклоняться от среднего значения $\overline{x}_i$. Для балансировки этих отклонений предлагается следующая формула прогнозирования на один период вперёд:

$$\hat{x}_{n+1} = \overline{x}_{n+1} + P_2 \sum_{i=n-k+1}^{n} (x_i - \overline{x}_i) \tag{2}$$

Здесь сигма представляет собой сумму дисбалансов, накопленных за k предыдущих периодов. Параметр P_2 находится по методу наименьших квадратов с использованием значений временного ряда. Значение k подбирается таким образом, чтобы дисперсия прогнозирования была наименьшей. Сначала принимается $k=1$. С увеличением k дисперсия сначала уменьшается, а затем начинает увеличиваться. В качестве k принимается значение, соответствующее минимуму дисперсии.

Статистические характеристики вычисляются по следующим формулам:

$\overline{X} = \dfrac{\sum_{i=1}^{n} x_i}{n}$ – среднее значение экспериментальных данных;

$D_Э = \dfrac{\sum_{i=1}^{n} (x_i - \overline{X})^2}{(n-1)}$ – экспериментальная дисперсия;

$\overline{x}_i = P_0 + P_1 \cdot i$ – среднее значение признака, определенное по МНК;

$y_i = x_i - \overline{x_i}$ – отклонение от среднего значения признака (дисбаланс);

$t_i = \sum\limits_{j=i-k+1}^{i} y_i$ – накопленные дисбалансы за k предыдущих периодов.

Первоначальный этап прогнозирования связан с анализом временного ряда, который позволяет охарактеризовать закономерность изменения показателя во времени.

Линейный тренд выглядит следующим образом:

$$\overline{x}_i = P_0 + P_1 \cdot i\,.$$

Здесь $\overline{x}_i$ – среднее значение в i–й период; i – порядковый номер.

Сущность метода наименьших квадратов заключается в нахождении параметров модели, при которых минимизируется сумма квадратов отклонений эмпирических значений от теоретических. Расчёт параметров модели произведен с помощью разработанной компьютерной программы в среде Matlab.

Прогноз предлагается находить по формуле:

$$\hat{x}_{i+1} = P_0 + P_1 \cdot (i+1) + P_2 t_i, \qquad (3)$$

где $\hat{x}_{i+1}$ – прогнозируемое значение в i+1 периоде.

Используя МНК, получена формула для нахождения P_2:

$$P_2 = \frac{\sum\limits_{i=k}^{n-1} (y_{i+1} \cdot t_i)}{\sum\limits_{i=k}^{n-1} t_i^2}.$$

Теоретическая дисперсия находится по формуле:

$$D_T = \frac{\sum\limits_{i=k+1}^{n} (x_i - \hat{x}_i)^2}{n-m-k},$$

где m – число параметров, используемых для прогнозирования (в данном случае их три (P_0, P_1, P_2)).

В данной статье разработана математическая модель прогнозирования временных рядов на основе метода наименьших квадратов, учитывающая суммарные дисбалансы предыдущих периодов.

Для получения численных и графических характеристик прогнозных значений разработано специальное программное обеспечение. Достоверность указанной методики проверена при помощи критерия Фишера, получены положительные результаты.

Разработанная методика удобна тем, что для прогнозирования необходимы только статистические данные. Тем самым, она может получить широкое применение в различных отраслях, требующих количественного прогноза. С помощью разработанной модели возможно также прогнозирование объёма продаж, курсов валют, акций, среднегодовых температур, рождаемости и т.д.

СПИСОК ЛИТЕРАТУРЫ

1. *Гмурман В. Е.* Теория вероятности и математическая статистика. М. : Высш. шк., 2003. 479 с.

2. Теория статистики: Учебник / Под ред. Проф. *Р.А. Шмойловой*. 3-е изд., перераб. М. : Финансы и статистика, 2000. 560 с.

3. *Баврин И. И.* Теория вероятностей и математическая статистика: Учебник. М. : Высш. шк., 2005. 160 с.

4. *Баврин И. И., Матросов В. Л.* Общий курс высшей математики: Учеб. для студентов физ.-мат. спец. пед. вузов. М. : Просвещение, 1995. 464 с.

5. *Мышкис А. Д.* Элементы теории математических моделей. Изд. 5-е. М. : Книжный дом «ЛИБРОКОМ», 2011. 192 с.

6. *Зельдович Я. Б., Мышкис А. Д.* Элементы прикладной математики. М. : Наука, 1972. 592 с.

7. *Просветов Г. И.* Прогнозирование и планирование. М. : РДЛ, 2005.