АНАЛИЗ, МОДЕЛИРОВАНИЕ И ПРОГНОЗИРОВАНИЕ ОСТАТКОВ ПРОДУКЦИИ НА СКЛАДЕ

Бозняков Р.В.
Научный руководитель: Семенов М.Е.
Томский политехнический университет
634050,Россия, г. Томск, пр-т Ленина, 30
E-mail: romario1131@bk.ru

Введение

В современной экономике математическое моделирование и компьютерный анализ данных применяются как методы неценовой конкуренции. Использование моделей управления складом с возможностью контроля и прогнозирования текущей ликвидности позволяет более эффективно управлять предприятием.

Цель данной работы – построение математической модели, описывающей временной ряд остатков продукции на складе, с возможностью прогноза значений ряда.

Под моделью временного ряда будем понимать уравнение, которое связывает наблюдение, полученное в определенный момент времени, с наблюдениями, полученными ранее по той же и/или другим характеристикам изучаемой переменной. Существует различные модели, которые используются для описания временных рядов. Среди таких моделей можно выделить: модели авторегрессии, скользящего среднего, а также комбинации на их основе.

Авторегрессионным процессом AR(p) порядка p называют стохастический процесс X_t, который определяется следующим соотношением [1, 2]

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + a_2 X_{t-2} + \ldots + \alpha_p X_{t-p} + \varepsilon_t,$$

где ε_t – процесс типа «белый шум» с нулевым математическим ожиданием $\mu_\varepsilon = 0$. Данная модель временного ряда основана на предположении, что поведение исследуемого явления в будущем определяется только его текущими и предыдущими состояниями.

Процессом скользящего среднего MA(q) порядка q называют стохастический процесс X_t, который определяется соотношением:

$$X_t = \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \ldots - \beta_q \varepsilon_{t-q}$$

где ε_t – процесс типа «белый шум» с нулевым математическим ожиданием $\mu_\varepsilon = 0$ и дисперсией $\sigma_\varepsilon^2 = \sigma^2$. В моделях скользящего среднего MA(q) среднее текущее значение стационарного стохастического процесса представляется в виде линейной комбинации текущего и прошедших значений ошибки, обладающей свойствами «белого шума» [1, 2].

Комбинация процессов AR(p) и MA(q) называется авторегрессионным процессом скользящего среднего, обозначается ARMA(p, q). Модель ARMA(p, q) имеет следующий вид [1, 3]:

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \alpha_p X_{t-p} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \ldots - \beta_q \varepsilon_{t-q}.$$

Модель ARIMA(p, d, q) – интегрированная модель авторегрессии – скользящего среднего. Эта модель является расширением моделей ARMA для нестационарных временных рядов. Значения параметров p и q соответствуют количество компонент от моделей авторегрессии AR(p) и скользящего среднего MA(q) соответственно, а параметр d – определяет порядок интегрирования.

В данной работе для анализа временного ряда выбрана модель ARIMA(p, d, q).

Построение математической модели

Для построения математической модели временного ряда необходимо последовательно определить значения параметров модели – d, p и q.

Исходные данные, для которых необходимо построить математическую модель и сделать прогноз, остатки продукции на складе в период с 01.02.2008 г. по 21.05.2014 г. с интервалом 10 дней (228 значений).

Для определения значения параметра d проведена проверка исследуемого ряда на стационарность (проверка гипотезы о наличии во временном ряду авторегрессии более чем первого порядка). Были выбраны тест Дики-Фуллера и параметрические тесты – критерии Стьюдента и Фишера. Во всех расчетах уровень значимости выбран 0,05.

В итоге исследуемый ряд оказался нестационарным интегрируемым первого порядка рядом, следовательно, для исследуемого временного ряда выбрана модель ARIMA(p, 1, q).

Рис. 1. Коррелограмма автокорреляционной функции

Для определения значения параметра p были построены автокорреляционная (АКФ) и частная автокорреляционная (ЧАКФ) функции временного ряда, а также коррелограммы указанных функций (рис. 1 и 2).

АКФ функция экспоненциально убывает, не меняя знак (рис. 1), ЧАКФ имеет существенные выбросы на лагах 1 и 2, при этом остальные значения не значимы (рис. 2).

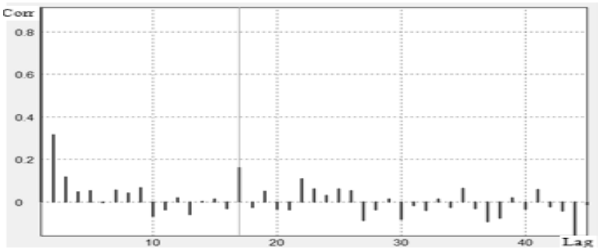

Рис. 2. Коррелограмма частной автокорреляционной функции

Из анализа коррелограмм можно сделать вывод, что модель ARIMA для исследуемого ряда содержит два параметра от модели AR, т.е. p=2. На данном этапе предположим, что моделью для нашего ряда будет модель ARIMA(2, 1, q).

Для определения оптимального значения параметра q были построены следующие модели – ARIMA(2, 1, 0), ARIMA(2, 1, 1), ARIMA(2, 1, 2) соответственно:

$$Y_t = 2154{,}4 + 0{,}61Y_{t-1} + 0{,}38Y_{t-2},$$

$$Y_t = 1{,}15Y_{t-1} - 0{,}158Y_{t-2} - 34{,}46 - 0{,}63\varepsilon_{t-1},$$

$$Y_t = 1{,}825Y_{t-1} - 0{,}825Y_{t-2} - 14{,}94 - 1{,}32\varepsilon_{t-1} + 0{,}36\varepsilon_{t-2}$$

Для нахождения коэффициентов моделей использован метод наименьших квадратов.

В качестве критериев оптимальности значения параметра q использованы (табл. 1): максимальное значение коэффициента детерминации (R^2), минимальное значение информационного критерия Акаике (AIC), минимальное значение стандартной ошибки (E).

	ARIMA (2, 1, 0)	ARIMA (2, 1, 1)	**ARIMA (2, 1, 2)**
α_0	2154,5	-34,45	-14,94
α_1	0,60	1,15	1,82
α_2	0,38	-0,15	-0,82
β_1	–	-0,63	-1,31
β_2	–	–	0,36
R^2	86,45%	87,06%	**87,22%**
E	525 408,20	514 551,00	**512 561,00**
AIC	29,20	29,16	**29,16**

Таблица 1. Показатели построенных моделей

Сравнение критериев показало, что наилучшей моделью для исследуемого временного ряда, является модель ARIMA(2, 1, 2). Эта модель имеет наименьшее значение ошибки E и критерия AIC, наибольшее значение коэффициента R^2 (в табл. 1 выделено жирным шрифтом).

Все коэффициенты α_0, α_1, α_2, β_1, β_2 построенной модели ARIMA(2, 1, 2) статистически значимы. Исходный временной ряд и предложенная модель ARIMA(2, 1, 2) изображены на рис. 3.

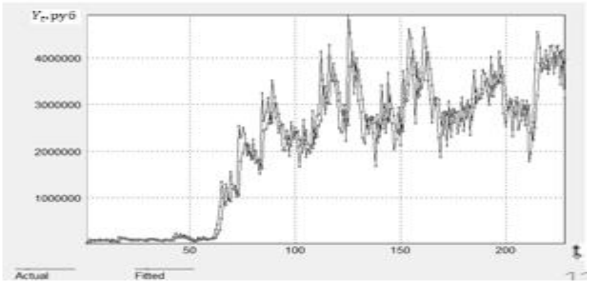

Рис. 3. Временной ряд и модель ARIMA(2, 1, 2)

Построение прогноза

На основе полученной модели ARIMA(2, 1, 2) построен прогноз на три декады вперед – 01.06.2014, 11.06.2014, 21.06.2014. В результате показано, что на 01.06.2014 на складе останется продукции на сумму 3,54 млн. руб., на 11.06.2014 – 3,57 млн. руб., на 21.06.2014 – 3,59 млн. руб., доверительный интервал равен 0,512 млн. руб., доверительные границы интервала выбраны равными $\pm 2E$.

Заключение

В работе сделан анализ исходных данных и построена интегрированная модель авторегрессии – скользящего среднего ARIMA(2, 1, 2). На основании предложенной модели сделан прогноз на три периода. Прогноз показал, что остатки товарной продукции будут увеличиваться каждые 10 дней на 0,02 млн руб.

В предположении, что случайные изменения остатков товара на складе на каждом временном интервале не зависят друг от друга, в дальнейших исследованиях планируется использовать нелинейные модели.

Литература

1. Эконометрика: лабораторный практикум: учебное пособие / Н.И. Шанченко. – Ульяновск: УлГТУ, 2011. – 117 с.

2. Трегуб А.В., Трегуб И.В. Методика построения модели ARIMA для прогнозирования динамики временных рядов // Вестник Московского государственного университета леса – Лесной вестник. – 2011. – № 5. – С. 179-183.

3. Гребенников А.В., Крюков Ю.А., Чернягин Д.В. Моделирование сетевого трафика и прогнозирование с помощью модели ARIMA // Системный анализ в науке и образовании, 2011. – Вып. 1. – www.sanse.ru/download/79.