

# 基于ARIMA模型的股票价格变动规律和预测的研究

徐新新

(哈尔滨商业大学, 哈尔滨 150028)

**摘要:**随着股市的不断发展和完善,越来越多的炒股爱好者进入了股市,但是经验的缺乏和交易系统的欠缺使很多入市者的资金不断缩水,从而演绎了一场场悲剧。因此,投资分析是股票投资的关键,能够为投资者提供良好的分析工具。以分析沪市股票价格指数为例,运用ARIMA(p,d,q)模型对股票价格指数进行回归建模分析,观察股票价格的变化规律,进而运用模型对未来一定时期内股票价格走势进行预测,全方位地让投资者了解股票分析有助于提高决策分析的科学性,减少盲目性;降低投资风险、提高投资收益。

**关键词:**股票价格;变动规律;ARIMA(p,d,q)模型;预测

**中图分类号:**F830.91 **文献标志码:**A **文章编号:**1673-291X(2016)19-0077-02

## 引言

股票投资作为一种风险和收益都相对较高的金融产品一直是人们进行投资和理财的重要途径。正是基于其风险和收益都较高的特点,所以,广大股票市场参与者均要在投资时做好充分的分析,结合自己的理性预期做出合理判断,进而选择合适的入市时机和资金量的投入,理性理财和投资。本文以分析沪市股票价格指数为例,运用ARIMA(p,d,q)模型对股票价格指数进行回归建模分析,观察股票价格的变化规律,进而运用模型对未来一定时期内股票价格走势进行预测,让投资者对股票价格的变动有一定程度上的了解,帮助投资者进行投资分析,进而做出科学的决策。

## 一、数据及研究分析

### (一)数据描述

以上海证券交易所的收盘指数为样本,搜集了沪市股票价格收盘指数s从1995年1月至2011年8月的数据,总共200组月度数据,并对数据进行初步整理。

### (二)数据处理

首先利用Eviews软件对这两百组原始数据描述其时间序列图,观察序列s存在时间趋势,因此在进行平稳性检验时选择含有趋势项和截距项的,对序列s进行单位根检验的结果:ADF检验的P值为1.0000大于给定检验水平0.05,因此接受原假设,说明序列s存在单位根,即序列s为非平稳时间序列。ARIMA(p,d,q)模型建立的前提是保证序列的平稳性<sup>[1]</sup>,为了消除序列s的时间趋势,对序列进行平稳化处理,再取序列的一阶自然对数差分序列,检验结果可以看出,ADF检验的P值为0.0000小于给定检验水平0.05。因此,拒绝原

假设,说明新序列不存在单位根,即新序列为平稳的时间序列,可对新序列建立ARIMA(p,d,q)模型。

### (三)模型的建立和识别

本文对序列s建立如式(1)所示的ARIMA(p,d,q)模型。

$$\Phi(L)\omega = c + \Theta(L)\varepsilon_t \quad (1)$$

式中, $\omega = (1-L)^d y_t$ ;  $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$ ;  $\Theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q$

ARIMA(p,d,q)模型的识别就是指具体的三个参数的确定。

1.d参数的识别。对于ARIMA(p,d,q)模型的d系数的识别是最简单的,对分析的原序列进行单位根检验,如果是有单位根的,则对其差分后的序列进行判断,如果d阶差分后的序列为平稳的,则称序列为d阶单整序列,对这个差分后序列建立ARMA模型即可<sup>[2]</sup>。

2.p,q参数的识别。应使用自相关函数图ACF和偏自相关函数图PACF图对p,q做出判断。我们运用自相关系数和偏自相关系数这两个统计量来识别ARIMA(p,d,q)模型的系数特点和模型的阶数,不断反复地进行试验及验证,选出各项都符合实际统计要求的具体形式。

### (四)模型的估计

本文利用Eviews软件对ARIMA(p,d,q)模型进行估计。利用偏自相关系数和相关系数进行选择 and 判断,不断重复验证,对模型形式进行初步估计,初步筛选处拟合优度较高,回归系数均显著的模型。

### (五)模型的诊断

ARIMA建模需要对比各种可能模型的优劣,取舍得到较优的模型。对模型的估计结果进行比较,挑选出各项统计指标均符合要求的模型形式,并结合对回归模型的残差序列

收稿日期:2016-05-28

作者简介:徐新新(1992-),女,黑龙江青冈人,硕士研究生,从事金融学研究。

相关性的检验,最后确定较优的、合理的回归模型。调整  $R^2$  表示模型的整体拟合优度,该值介于 0 和 1 之间,拟合的程度越好那么此数值则越大。AIC 和 SC 都是信息准则的表现形式。数值越小精度越高,结果越可信。如果残差序列相关性检验显著拒绝原假设,这也是模型挑选的最重要条件。综上所述,估计结果如下:

$$\begin{aligned} \hat{ls}_t &= \hat{\phi}_9 \hat{u}_{t-9} + \hat{\varepsilon}_t + \hat{\theta}_9 \hat{\varepsilon}_{t-9} \\ &= 0.7597 \hat{ls}_{t-9} + \hat{\varepsilon}_t - 0.9216 \hat{\varepsilon}_{t-9} \end{aligned} \quad (2)$$

下面进一步对模型的残差序列进行检验,其 ADF 检验结果可以看出,在 5% 的显著水平下拒绝原假设,说明残差序列为平稳的。

模型的拟合效果及残差序列(如图 1 所示)。

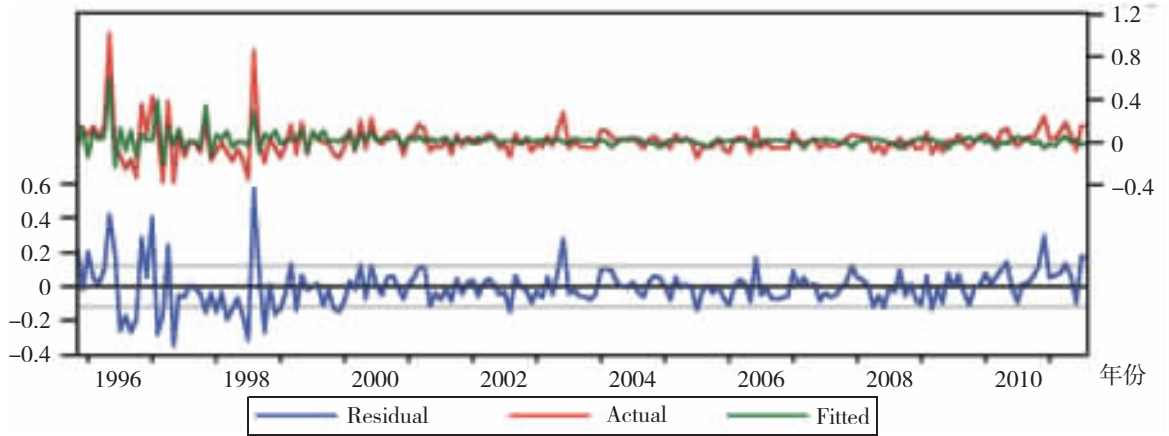


图 1 序列  $ls$  的实际值和拟合值及残差序列图

由图 1 可以观察到该模型比较好的拟合了  $ls$  序列,回归方程的残差序列基本上也是一个零均值的平稳序列。以上的检验进一步说明了该模型的合理性。

## 二、模型的预测

模型的预测,实际上就是利用序列已观测到的样本值对

序列在未来某个时刻的取值进行估计。本文利用建立好的模型进行执行样本外的动态预测,对未来一定时期内的股票价格指数进行预测,预测结果(如图 2 所示)。

图 2 显示的红线部分代表预测区间,随着时间推移证明预测效果还是不错的;根据图中右侧评价指标可以看出,预测误差的均方根为 0.1646,绝对预测误差均值为 0.1646,相对

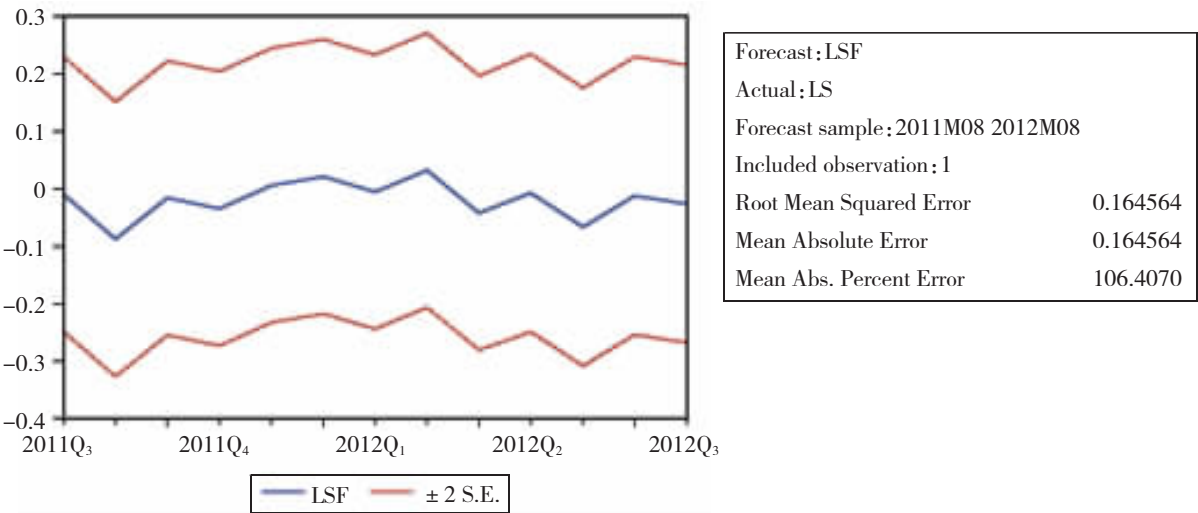


图 2 预测值折线图和评价指标

误差绝对值平均为 106.407,因此本次预测精度是很好的,也说明了建立的模型是非常适合的。还可以看出  $ls$  序列的预测几乎是平稳的,则说明原序列是呈波动上升趋势的,即股票

价格收盘指数在 2011 年 8 月的未来一定时期内将呈波动上升趋势且事实证明 2011 年 8 月后的上证收盘价确实呈波动上升趋势,也说明了模型的合理性。

## 参考文献:

- [1] 徐耀东.ARIMA 模型在安徽省 GDP 的研究与应用[J].铜陵学院学报,2009,(6):19-21.
- [2] 时曦.期货对现货价格预测的作用——基于 ARIMA 模型的实证探讨[J].商业时代,2012,(14):10-15.

[责任编辑 陈丽敏]