

Forecast and Analysis of Fixed Assets Investment in Henan Province Based on ARIMA Model

Mao Yang, Changchun Wang

Henan University of Technology, Zhengzhou Henan
Email: yangmao20022002@163.com

Received: Dec. 8th, 2017; accepted: Dec. 22nd, 2017; published: Dec. 29th, 2017

Abstract

ARIMA model can solve the modeling problem of non-stationary time series well, and it has a good application time in the short-term prediction of time series model. By collecting the data from 1989 to 2014, combined with EViews software, finally, the ARIMA model is applied to the analysis and forecast of the total investment of fixed assets in Henan province. The results show that the ARIMA model can be used to analyze the total assets of Henan province. The results show that the model can solve the problem of estimating and forecasting the investment of fixed assets in Henan province, and the prediction accuracy is high so as to provide the government with a fixed asset investment ratio and the amount of investment to provide a reliable reference to promote the healthy development of the economy.

Keywords

Henan Province, Whole Society Fixed Asset Investment, Time Series Analysis, ARIMA Forecast

基于ARIMA模型的河南省全社会固定资产投资预测分析

杨 茂, 王长春

河南工业大学, 河南 郑州
Email: yangmao20022002@163.com

收稿日期: 2017年12月8日; 录用日期: 2017年12月22日; 发布日期: 2017年12月29日

摘要

ARIMA模型能较好的解决非平稳时间序列的建模问题,并且在时间序列模型的短期预测方面有很好的应用时间,文章通过搜集从1989年~2014年的数据,结合EViews软件,建立河南省全社会固定资产投资总额自回归移动平均模型ARIMA(p, d, q)的时间序列模型并进行检验,最后将ARIMA模型应用于河南省全社会固定资产投资总额分析及预测。结果表明,该模型能较好的解决河南省全社会固定资产投资的估计和预测问题,预测精度较高。从而为政府提供固定资产投资比例和投资金额提供可靠的参考依据,促进经济的健康发展。

关键词

河南省, 全社会固定资产投资, 时间序列分析, ARIMA预测

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

全社会固定资产投资是指某个国家或者地区内固定资产投资的数量和范围,建造和购置固定资产的经济活动,即固定资产再生产活动,是反映建设规模的综合性指标[1]。在国民经济发展过程中,社会固定资产投资是影响经济增长的一个重要因素,对经济增长具有直接的拉动作用。历来,许多学者从各方面证实了这一点,全社会固定资产投资与经济增长的彼此拉动也是实际数据的彰显。早期,学者胡春和仲继银就利用1981~1999年的国内固定资产投资的数据,在《我国固定资产投资与经济增长周期关系的实证分析》中,通过建立固定资产投资的增长率与国内生产总值增长率间相关关系的回归模型,论证了固定资产投资对经济增长的推动作用。学者熊兴在《关于甘肃省固定资产投资与经济增长关系的实证研究》文中,就利用甘肃省1978~2011年的数据,从三次产业的角度分别论证了固定资产投资对经济增长的不同拉动作用。在《中国固定资产投资与经济增长相互动态决定的实证研究》中,作者龚曙明就运用回归模型研究了固定资产投资对经济增长的动态推动效应。大量论证表明了固定资产投资与经济增长的拉动作用,因此利用最新的数据研究全社会固定资产投资状况对经济发展具有非常重要的指导意义。全社会固定资产投资的实证分析大多运用时间序列分析方法,文章运用ARIMA模型进行河南省全社会固定资产投资的时间序列分析,从统计分析角度研究河南省历年固定资产投资的发展状况。

2. 序列的平稳性检验

2.1. 原序列的平稳性检验

文章数据来源于《2015河南省统计年鉴》,选取1989年~2014年河南省全社会固定资产投资总额为指标 Y 。运用EViews软件绘制原始数据的时间序列图,如图1所示,从图中可观察到该序列呈指数分布,初步判断 Y 是一个非平稳的时间序列。

文章运用ADF检验来检验序列的平稳性,ADF检验法的基本原理是通过 n 次差分的方法将原始的非平稳序列转化为平稳序列,对ADF检验,检验统计量是检验回归滞后因变量的 t 统计量,由于是单边

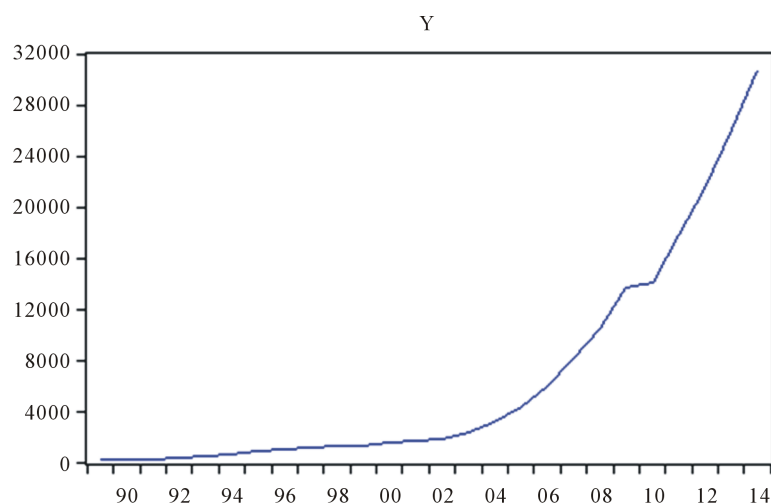


Figure 1. Timing chart of total fixed assets investment in Henan from 1989 to 2014

图 1. 中国 1989~2014 年河南省全社会固定资产投资总额时序图

检验, 当计算得到的 t 统计量的值小于临界值时拒绝原假设(即否定存在单位根) [2], ADF 检验结果见表 1。从表 1 中, ADF 检验中得到的 t 值为 10.55245, 大于三个检验水平的临界值, 表明该时间序列存在单位根, 为非平稳时间序列。非平稳的数据是不能用来建立时间序列模型的, 我们需要对数据进行平稳化处理。

2.2. 对数序列的平稳性检验

对于含有指数趋势的时间序列, 可以通过对序列取自然对数来消除指数趋势。由于这一序列有着非常明显的指数趋势, 因此我们对它进行取自然对数的运算来消除指数趋势的影响, 将取对数后的序列命名为 $\{y_t\}$, 即 $y_t = \log(Y)$ 。作出序列的自相关图分别如图 2 所示。

对 $\{y_t\}$ 序列作单位根检验, 原假设: $H_0: |\lambda| \geq 1$; 备择假设: $H_1: |\lambda| < 1$, 检验结果如表 2 所示。

根据这一检验结果, 我们看到对原序列取对数以后得序列 $\{y_t\}$ 依然没有平稳, 根据自相关-偏自相关图可以看出, 序列中仍有着明显的增长趋势, 采用 ADF 检验方法对序列进行单位根检验, 进一步验证序列的非平稳性, 检验结果如表 2 所示, t 值为 2.866057, 大于临界值, 接受原假设即 $H_0: |\lambda| \geq 1$; 表明序列 $\{y_t\}$ 存在单位根, 为非平稳序列。根据规律, 一阶差分可以消除线性规律, 所以需要先对序列 $\{y_t\}$ 进行差分处理来消除趋势的影响。

2.3. 序列 $\{y_t\}$ 的差分处理

我们将序列 $\{y_t\}$ 进行一阶差分处理, 得到一个新序列 $\{x_t\}$, 即 $x_t = (1-B)y_t$ 。

画出序列 $\{x_t\}$ 的时序图, 并进行相应的单位根检验, 如图 3、表 3 所示。

图 3 为一阶差分后的时序图, 从图中可以看出随着时间的推移没有明显的波动变化趋势, 初步判断其是一个平稳的时间序列。采用 ADF 检验方法对序列 $\{x_t\}$ 进行单位根检验, 进一步判断该序列的平稳性, ADF 检验结果见表 3。从表 3 可以看出, ADF 检验中得到的 t 值为 -3.910395, 小于 5% 及 10% 显著性水平下的临界值, 应当拒绝原假设 $H_0: |\lambda| \geq 1$; 表明该时间序列所有的根在单位圆内, 为平稳时间序列。因此, 河南省全社会固定资产投资数据是一阶单整序列, 记为 $x_t \sim I(1)$ 。接下来可以针对该序列作进一步的模型拟合。

Table 1. ADF test results of the sequence Y
表 1. Y 序列的 ADF 检验结果

		t-Statistic	Prob.
Augmented Dickey-Fuller test statistic		10.55245	1.0000
Test critical values:	1% level	-3.724070	
	5% level	-2.986225	
	10% level	-2.632604	

Table 2. ADF test results
表 2. ADF 检验结果

		t-Statistic	Prob.
Augmented Dickey-Fuller test statistic		2.866057	1.0000
Test critical values:	1% level	-3.808546	
	5% level	-3.020686	
	10% level	-2.650413	

Table 3. ADF test results
表 3. ADF 检验结果

		t-Statistic	Prob.
Augmented Dickey-Fuller test statistic		-3.910395	0.0081
Test critical values:	1% level	-3.808546	
	5% level	-3.020686	
	10% level	-2.650413	

Sample: 1989 2014
 Included observations: 26

Auto Correlation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.885	0.885	22.805	0.000	
2	0.762	-0.097	40.422	0.000	
3	0.639	-0.069	53.354	0.000	
4	0.518	-0.068	62.227	0.000	
5	0.407	-0.031	67.970	0.000	
6	0.299	-0.057	71.233	0.000	
7	0.202	-0.034	72.802	0.000	

Figure 2. Autocorrelation-partial autocorrelation diagram
图 2. 自相关 - 偏自相关图

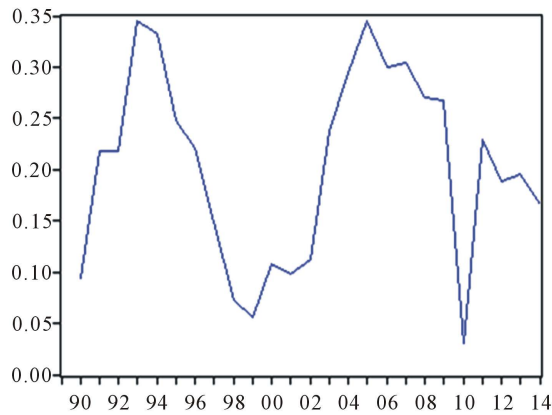


Figure 3. x_t timing diagram
图 3. x_t 时序图

3. 模型建立

ARIMA(p, d, q)模型的识别与定阶可以通过样本的自相关与偏自相关函数的观察获得[3]。经过平稳性检验, 我们建立ARIMA(p, d, q), 模型经过一阶差分后平稳, 所以 $d = 1$, 为了获得自回归和移动平均回归中的 p 和 q , 可以通过自相关和偏自相关图来观察。图4为 $\{x_t\}$ 序列的自相关图和偏自相关图。

从图4中可以看出, 自相关系数在滞后阶数1、5、6和7时显著不为0, 因此 $q = 1$ 、 $q = 5$ 、 $q = 6$ 或者 $q = 7$ 。偏自相关系数在滞后阶数为1时显著不为0, 因此 $p = 1$ 。为了确定模型的最终结构, 需要结合AIC和SC准则定阶, 最终来确定模型的阶数。文章建立4个模型, 运用Eviews软件根据AIC和SC值来定阶, 结果如表4。

从表4中观察到ARMA(p, q)取(1,5)时, AIC和SIC的值最小。综合比较DW统计量以及 R^2 , 一般DW统计量越接近2, R^2 越接近1说明模型拟合的越好, 最后选定的ARIMA(p, d, q)模型取(1, 1, 5)时, 模型拟合的较好。为了进一步确定选定的模型是否能够充分的提取数据列的信息, 对ARIMA(1,1,5)进行检验。我们将该模型中的自回归和移动平均中部分不显著的系数放弃, 最终的估计结果如图5所示。

模型的具体形式为:

$$x_t = 0.199550 + 0.507968x_{t-1} - 0.082020\varepsilon_{t-1} - 0.887127\varepsilon_{t-5} + \varepsilon_t$$

$$t = (15.76010) \quad (2.452251) \quad (-0.939909) \quad (-14.49139)$$

$$S.E. = 0.055682, R^2 = 0.688939, \bar{R}^2 = 0.642279, F = 14.76533$$

$$AIC = -2.787294, SC = -2.590951, DW = 2.089939$$

我们可以推出时间序列 $\{y_t\}$ 的预测公式为:

$$y_t = y_{t-1} + 0.199550 + 0.507968x_{t-1} - 0.082020\varepsilon_{t-1} + \varepsilon_t$$

进而推出时间序列 $\{Y\}$ 的预测公式为:

$$Y = e^{y_{t-1} + 0.199550 + 0.507968x_{t-1} - 0.082020\varepsilon_{t-1} + \varepsilon_t}$$

根据图5中得到的数据: R^2 接近于1, 说明模型拟合的很好; DW值接近2, 说明不存在自相关关系; AIC和SC值非常小, 说明模型拟合的相对合理。

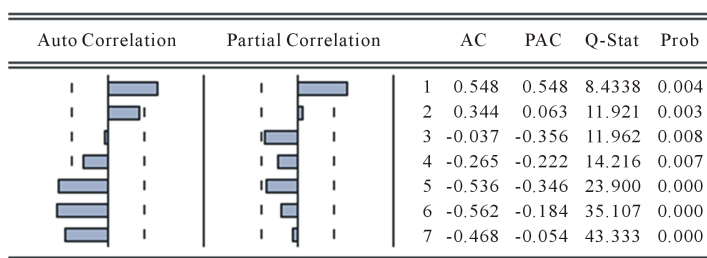


Figure 4. x_t autocorrelation-partial autocorrelation of sequences

图4. x_t 序列的自相关 - 偏自相关图

Table 4. Comparison of different ARMA(p, q) model AIC and SC values

表4. 不同ARMA(p, q)模型AIC和SC值比较

ARMA(p, q)	(1,1)	(1,5)	(1,6)	(1,7)
AIC	-1.957741	-2.568044	-2.320666	-2.356181
SC	-1.859570	-2.224445	-1.927981	-1.914411

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.199547	0.012658	15.76436	0.0000
AR(1)	0.507911	0.207138	2.452040	0.0235
MA(1)	-0.082020	0.087301	-0.939515	0.3587
MA(5)	-0.887101	0.061216	-14.49138	0.0000
R-squared	0.688939	Mean dependent var		0.208579
Adjusted R-squared	0.642279	S.D. dependent var		0.093099
S.E. of regression	0.055682	Akaike info criterion		-2.787294
Sum squared resid	0.062011	Schwarz criterion		-2.590951
Log likelihood	37.44752	Hannan-Quinn criter.		-2.735204
F-statistic	14.76533	Durbin-Watson stat		2.089860
Prob(F-statistic)	0.000027			

Figure 5. First-order difference model ARIMA(1,1,5) of the total social investment in fixed assets

图 5. 对数全社会固定资产投资总额一阶差分 ARIMA(1,1,5) 模型

4. 模型的检验

文章对模型的残差序列进行独立性检验, 这里采用拉格朗日乘数检验法(LM 检验)对残差序列的样本自相关系数进行检验, 原假设是残差序列相互独立。拉格朗日乘数检验的思想是: 在约束条件下, 可以用拉格朗日方法构造目标函数。如果约束有效, 则最大化拉格朗日函数所得估计量应位于最大化无约束所得参数估计值附近。这里也是构造一个 LM 统计量, 该统计量服从卡方分布[4]。表 5 为 LM 检验结果, 可以看出统计量值的 p 值均大于 0.05, 应当接受原假设, 所以残差序列相互独立, 是白噪声序列。

另外从残差序列的自相关图(图 6)可以看出, 残差序列是纯随机序列, ARIMA(1,1,5)模型的建立较为合理, 充分提取了数据中的信息。残差通过了白噪声检验, 因而可以判断该模型是可行的, 可用于预测分析。

5. 模型的预测

首先进行样本内静态预测, 图 7 为 1989~2014 年全社会固定资产投资总额预测值与真实值的趋势图, 从图中可以看出样本期内河南省社会固定资产投资总额的真实值和预测值两条曲线的拟合情况非常好, 走势基本一致, 说明模型拟合的很好, 非常合理。通过计算样本内各年份河南省全社会固定资产投资总额的真实值和预测值, 1989~2014 年内预测的平均误差为 2.0%, 一般来说预测的误差越小, 预测的精度就越高, 说明模型拟合的就越合理, 由于样本期内预测的平均误差非常小, 所以预测的精度非常高, 这对于未来几年河南省全社会固定资产投资总额的预测具有重要的参考意义。

为了进一步分析河南省全社会固定资产投资总额, 利用建立的 ARIMA 模型进行样本外动态预测, 由于该时间序列模型经过取对数再做一阶差分才基本平稳, 且模型数据有限, 所以所获得的模型反映的是短期变化关系, 而非长期变化关系, 因此只对未来四年 2015 年~2018 年的全社会固定资产投资总额进行预测, 表 6 为预测值数据。

从预测的结果来看, 未来4年的河南省全社会固定资产投资总额仍将保持较高速增长, 从模型的预测公式我们可以看到, 河南省固定资产投资总额与其第一期的滞后值、第五期的随机扰动项密切相关, 从参数估计值来看, 与第一期的滞后值正相关, 与第五期的随机扰动项负相关。因此, 政府在引导投资时要考虑到这点。近几年全社会固定资产投资总额与资本形成总额出现严重背离且差额逐渐拉大, 目前差额已超过30%, 也就是说超过30%的固定资产投资不能形成资本, 这值得我们关注和警惕。近几年房地产成为热门, 土地交易金在全社会固定资产投资总额中所占比重逐渐提高, 到2012年已经超过30%, 金额上达

Table 5. Residual LM test results

表 5. 残差的 LM 检验结果

F 统计量	0.951271	P 值	0.5024
$T \times R^2$ 统计量	8.084326	Probability	0.3252

Table 6. Estimated total investment in fixed assets of Henan province from 2015 to 2018 (Unit: 100 million Yuan)

表 6. 2015 年~2018 年河南省全社会固定资产投资总额预测值(单位: 亿元)

2015 年	40371.04534 亿元
2016 年	47207.65016 亿元
2017 年	57916.4796 亿元
2018 年	69188.68803 亿元

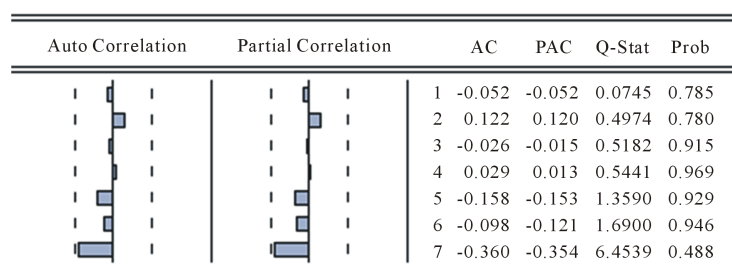


Figure 6. Residual sequence autocorrelation-partial autocorrelation analysis

图 6. 残差序列自相关 - 偏自相关分析图

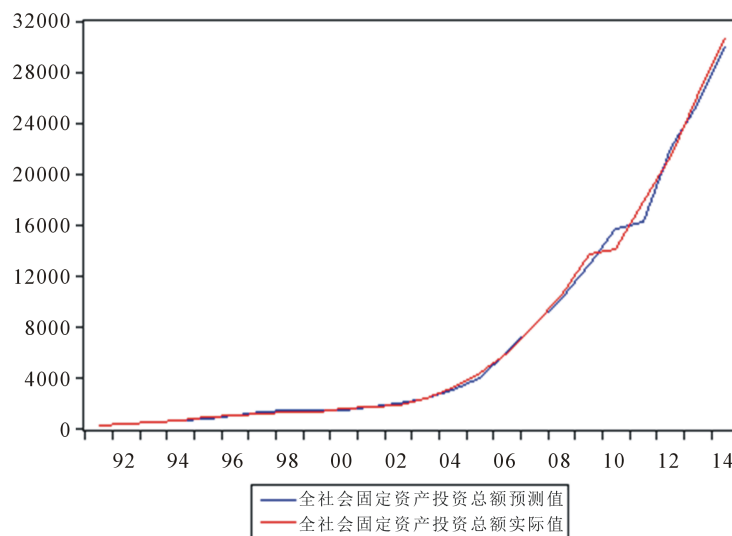


Figure 7. Real and static estimates of total fixed assets investment in Henan province

图 7. 河南省全社会固定资产投资总额的真实值与静态预测值

到12万亿人民币以上, 名副其实地出现经济虚拟化、经济地产业化, 使得全社会固定资产投资总额逐渐大于资本形成总额, 两者之间差距越来越大[5]。同时经济的虚拟化、经济地产业化也是我国近几年用扩大投资的方式拉动经济日益困难, 拉动力日益减弱经济增长率降低的一个重要原因[6]。我国经济增长的70%以上要靠投资拉动, 而投资的30%以上不能对经济增长有所贡献, 那么至少降低了20%的经济增长率。如

果一个国家或地区固定资产投资有较大增长不能称之为坏事, 但是每个地区经济结构不同, 总有一些问题会影响自身的均衡发展, 制约经济增长。投资过热有时候也会成为自身的累赘, 不一定是好的, 要看经济结构能不能承受, 所以说, 要做到经济结构全面和相对合理是非常困难的。所以政府应当注意: 1) 政府应当充分贯彻中央各项宏观经济政策, 引导并调整和优化产业投资结构; 2) 政府应当合理引导投资比例, 提升投资对经济的拉动力, 多渠道筹措建设资金, 积极争取中央投资和信贷基金支持[7]; 3) 提高投资收益, 大力扶持一些影响力大、有长远发展的投资项目, 着力推进产业集聚区建设, 积极推进一批重大项目尽快开工建设, 切实加快在建项目建设进度, 加快保障性安居工程建设, 努力实现投资平稳较快增长, 为河南省社会发展提供有力支撑[8]; 4) 坚持把扩大投资需求、调整投资结构作为“转方式、调结构”最直接最有效的方式, 激发民众的热情, 拓展民间投资领域。

综上所述, ARIMA模型较好的解决了非平稳序列的建模问题, 在短期预测方面有着重要的意义。文章借助Eviews软件, 将ARIMA模型应用于河南省全社会固定资产投资总额的预测, 预测值与真实值非常接近, 预测精度较高, 说明模型拟合的较好, 相对合理, 为河南省全社会固定资产投资的制定提供了一个客观和科学的根据。

基金项目

融资租赁与“百城建设提质工程”融资模式研究(河南省政府决策研究招标课题, 立项编号: 2017B276)。

参考文献 (References)

- [1] 陈悦华, 廖造壮. 基于灰色理论的武汉市全社会固定资产投资预测[J]. 科学技术与工程, 2012(14): 3535-3538.
- [2] 郑博儒, 李雪莉. 基于 X-12-ARIMA 的全社会固定资产投资预测[J]. 特区经济, 2012(11): 237-238.
- [3] 王丽丽, 何强, 李芬. 影响社会固定资产投资因素的实证分析[J]. 当代经济, 2011(18): 142-143.
- [4] 吴海军. ARIMA 模型在北京市全社会固定资产投资预测中的应用[J]. 经济研究导刊, 2007(2): 131-133.
- [5] 褚晓飞, 赵冉. 中国如何走出储备“水槽”——基于 ARIMA 模型的外汇储备适度性分析[J]. 河南科技大学学报(社会科学版), 2015(4): 77-81.
- [6] 桑秀丽, 李哲, 肖汉杰, 王天朝. 基于 ARIMA-GARCH 与抛物线的 CPI 组合预测模型[J]. 统计与决策, 2015(15): 23-25.
- [7] 吴新南, 王志伟. 基于 ARIMA 模型的浙江省农业总产值预测[J]. 现代商业, 2015(28): 54-55.
- [8] 陈蔚. 基于线性 ARIMA 与非线性 BP 神经网络组合模型的进出口贸易预测[J]. 统计与决策, 2015(22): 47-49.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-2251, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: sa@hanspub.org