

时间序列分析在山东省GDP分析预测中的应用

刘子瑜, 吴清

上海工程技术大学, 上海

收稿日期: 2022年7月1日; 录用日期: 2022年7月27日; 发布日期: 2022年8月4日

摘要

GDP可以反映一个地区经济的发展变化情况, 本文利用山东省1978~2020年生产总值的数据, 通过图检验和假设检验两种方式确定平稳性, 通过二阶差分转换为平稳序列, 拟合出ARIMA(1,2,0)模型, ARIMA被广泛应用于单变量时间序列数据预测, 便利性十分突出, 仅使用内生变量就足够达到预测效果。借助模型预测未来几年山东省GDP仍将处于快速发展的水平, 对未来山东省的发展布局起到了一定的现实意义。

关键词

时间序列模型, 山东GDP, 预测

Application of Time Series Analysis in GDP Analysis and Forecasting in Shandong Province

Ziyu Liu, Qing Wu

Shanghai University of Engineering Science, Shanghai

Received: Jul. 1st, 2022; accepted: Jul. 27th, 2022; published: Aug. 4th, 2022

Abstract

GDP can reflect the development and change of a region's economy. In this paper, we use the data of GDP of Shandong Province from 1978 to 2020 to determine the smoothness by both graph test and hypothesis test, and fit the ARIMA(1,2,0) model by converting it into a smooth series through second-order difference. ARIMA is widely used for univariate time series data forecasting, the convenience is outstanding and the use of endogenous variables alone is sufficient to achieve the forecasting effect. With the help of the model, it is predicted that the GDP of Shandong Province

will remain at a rapid level in the coming years, which plays a realistic role in the future layout of Shandong Province.

Keywords

Time Series Models, Shandong GDP, Forecast

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

1. 引言

通过度量核算期内, 经济社会运用生产要素最终生产出的全部产品和服务的价值总量, GDP 在反映国民经济发展状况、正确制定战略和政策以及检测政策科学性和有效性等诸方面均起着不可替代的作用。

1970 年, George 和 Gwilym [1]-[6]对 ARIMA 模型主要步骤包括识别、预估、检验和预测进行了理论和时间的综合性阐述。近些年来, 国内学者利用时间序列模型也对大量地区的 GDP 做了预测。何新易 [7]在 2010 年利用 ARIMA 模型预测中国在未来五年里, 经济仍保持较高水平的增长; 魏宁[8]等人利用 ARIMA 模型对中西部某省份 GDP 进行预测, 确定 ARIMA(1,2,1)为最优拟合模型, 并对未来 6 年做出了预测; 瞿海情[9]等人利用 ARIMA 模型预测湖北省 GDP, 得到最优模型 ARIMA(0,2,3), 认为 ARIMA 模型能较好地反映其发展趋势并进行短期预测。

山东省地理位置佳, 人口众多, 有着丰厚的文化底蕴, 其经济发展水平在国内一直处于较高位置, 重要特点是“稳”。长期占据全国省份 GDP 排名榜单的第三位, 仅次于广东和江苏, 对全国经济总量的影响可谓举足轻重。然而, 自 2017 年开始, 山东省的 GDP 增长遇到不小的挑战, 经济下行压力加大, 省内经济增长缺乏活力带来的问题进一步显现, 具体来说就是实体经济发展困难, 内需增长放缓, 新旧动能转换空档期等问题。在此背景下, 研究和预测山东省 GDP 具有重要意义。与此同时, 鉴于山东省与国家宏观经济政策的契合程度高, 研究也可以反映出我国 GDP 总量的发展趋势, 预测今后国家整体的经济动向。

2. 数据来源与研究方法

2.1. 数据来源

中国统计局官网。

2.2. 研究方法

时间序列分析, 应用于解决具有随机性、季节性以及平稳性等特点的时间序列问题, 根据对历史资料的变动趋势的分析, 对未来数据评估预测。主要集中在经济、金融、商业数据的分析领域。本文采用的模型是自回归滑动平均求和模型, 又称为 ARIMA(p,d,q), AR 是“自回归”, p 为自回归项数; MA 为“滑动平均”, q 为滑动平均项数, d 是使其成为平稳序列所做的差分次数(阶数)。

3. 模型建立

3.1. 原始序列平稳性检验

由图 1 时序图看出, 山东省 GDP 从 1980 至今尽管速度上不一致, 但趋势始终保持稳步上升, 初步认为是非平稳序列。

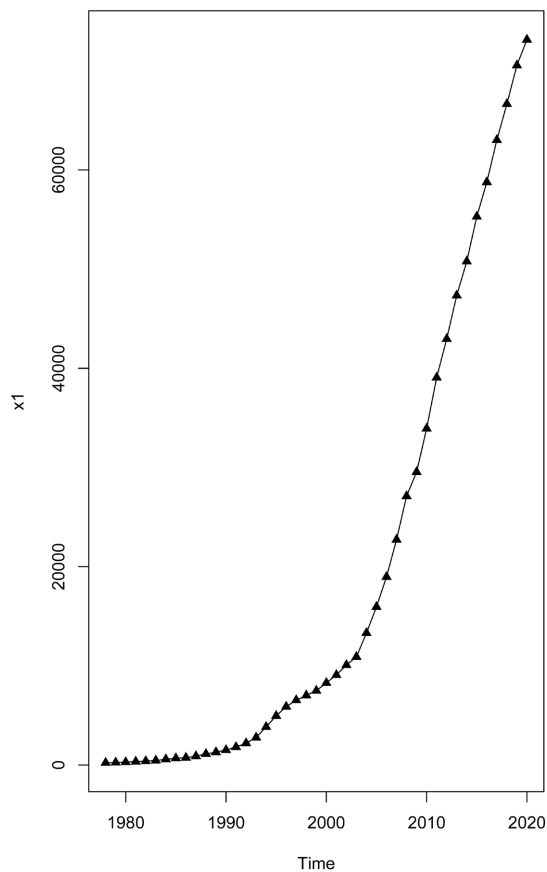


Figure 1. Time-series diagram

图 1. 时序图

由图 2 结果显示, 上述三种类型各个模型 p 值均 >0.05 , 最终确定山东省 GDP 序列为非平稳序列。

```
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
lag ADF p.value
[1,] 0 13.049 0.990
[2,] 1 0.343 0.737
Type 2: with drift no trend
lag ADF p.value
[1,] 0 9.109 0.990
[2,] 1 0.443 0.981
Type 3: with drift and trend
lag ADF p.value
[1,] 0 1.069 0.990
[2,] 1 -0.808 0.954
----
Note: in fact, p.value = 0.01 means p.value <= 0.01
```

Figure 2. Unit root test

图 2. 单位根检验

3.2. 差分处理

图 3, 1 阶差分时序图观察可得, 原序列中部分存在长期趋势, 但仍可观察出有长期递增的趋势, 而图 4, 2 阶差分时序图可以看出较完整的长期趋势, 并且没有了明显递增趋势或周期。即原序列在经过二阶差分处理后变成平稳序列。

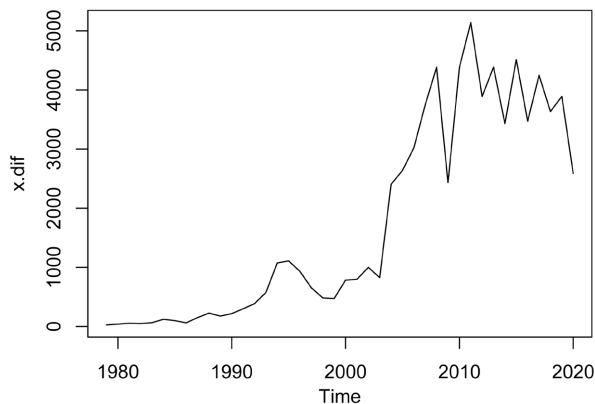


Figure 3. 1st order differential timing diagram

图 3.1 阶差分时序图

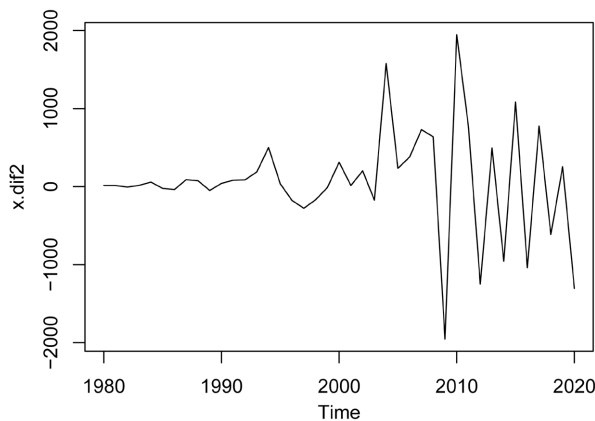


Figure 4. 2nd order differential timing diagram

图 4.2 阶差分时序图

3.3. 模型识别

观察图 5, 2 阶差分序列自相关图和图 6, 2 阶差分序列偏自相关图, 分别选择 ARIMA(2,2,0), ARIMA(1,2,1), ARIMA(1,2,0), 通过它们的 AIC 值得出最优拟合模型为 ARIMA(1,2,0), 如表 1 模型选择。

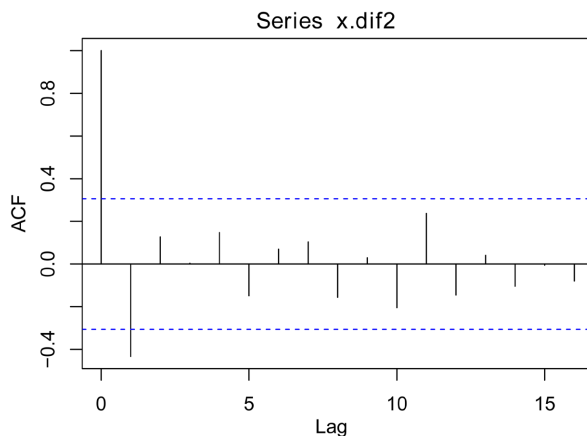


Figure 5. 2nd order differential serial autocorrelation plots

图 5.2 阶差分序列自相关图

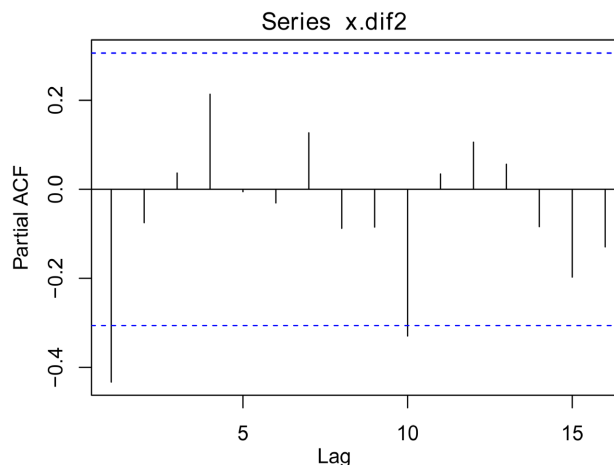


Figure 6. 2nd order differential serial partial autocorrelation plots
图 6. 2 阶差分序列偏自相关图

Table 1. Model selection
表 1. 模型选择

模型	AIC
ARIMA(2,2,0)	651.13
ARIMA(1,2,1)	651.16
ARIMA(1,2,0)	649.27

3.4. 模型检验

对白噪声检验来说, 确定一个模型的拟合效果要看它所提取的信息的能力。即残差项中不能存在关联性, 表 2 白噪声检验显示, 延迟阶数在 6 阶还是 12 阶的 P 值均大于 0.05, 认为该序列是白噪声序列。ARIMA(1,2,0)模型对原数列的相关信息提取充分, 模型的残差满足方差齐性假设, 模型残差序列为白噪声序列。

Table 2. White noise test
表 2. 白噪声检验

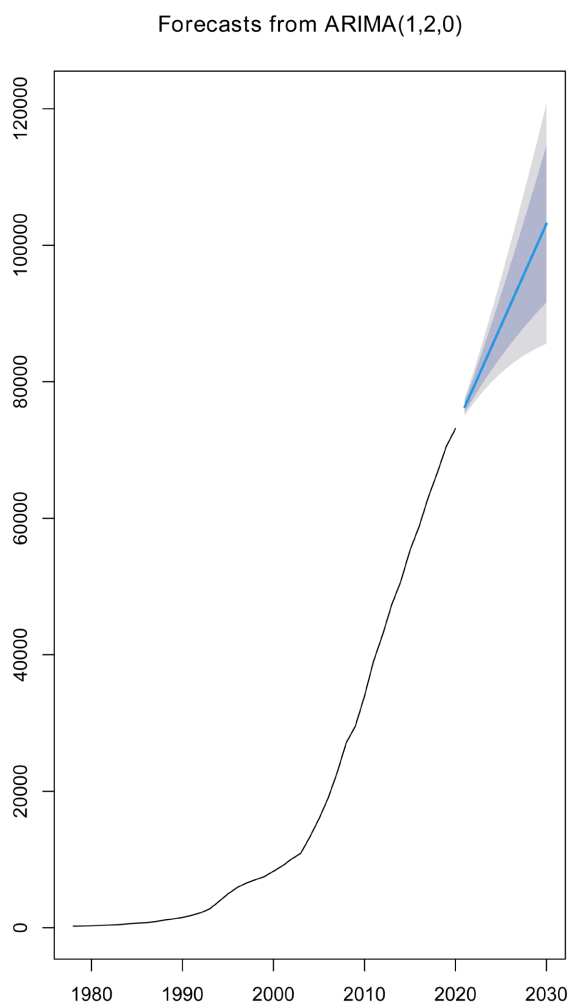
延迟阶数	p 值
6 阶	0.7669
12 阶	0.6683

3.5. 模型预测

利用通过检验的 ARIMA(1,2,0)模型, 预测 2021 年的山东省 GDP, 通过表 3 实际值与预测值对比可观察, 两者相对误差在 10%以内, 数据上看, 该模型短期内预测的值较准确, 图 7 模型预测图中可看出, 预测数与山东省 GDP 实际数据的误差区域不大, 也就是预测数据与实际发生情况基本吻合, 由此说明 ARIMA(1,2,0)模型拟合效果良好, 具备短期内达到预测目的的功能, 具有一定现实参考意义。

Table 3. Actual versus predicted values**表 3.** 实际值与预测值对比

年份	预测值	实际值	相对误差
2021	76295.99	83095.9	-8.91%

**Figure 7.** Model prediction chart**图 7.** 模型预测图

4. 结论

利用时间序列分析的方法对改革开放以来山东省 GDP 数据建模, 得出 ARIMA(1,2,0)模型, 经白噪声检验后说明该模型拟合效果具有说服力。随后又利用模型对 2021 年山东省 GDP 进行预测, 相对误差控制在 10%以内, 是一个较理想的结果。从预测结果来看山东省 GDP 在 2022~2030 年内仍保持较高增长速度。但需注意的是, 该结果只是一个基于拟合模型的预测值, 始终会和实际数据有所出入, 现实情况会受到国家宏观政策、经济环境、突发灾害等多种因素的影响。本文对山东省未来十年 GDP 进行预测可以提供一个参考, 有利于政府有关部门做好风险调控, 做好对经济运行中蕴藏风险的防范, 同时, 根据 GDP 的发展统筹全省整体发展布局, 拉动内需, 从根源拉动经济增长, 提升居民购买力。

参考文献

- [1] Box, G.E.P. (1976) Science and Statistics. *Journal of the American Statistical Association*, **71**, 791-799.
<https://doi.org/10.1080/01621459.1976.10480949>
- [2] Box, G.E.P. (1987) *Empirical Model-Building and Response Surfaces*. John Wiley & Sons, Inc., State of New Jersey.
- [3] Box, G.E.P. (2005) *Statistics for Experimenters*. 2nd Edition, John Wiley & Sons, Inc., State of New Jersey.
- [4] Jenkins, G.M. (1983) *Case Studies in Time Series Analysis*. John Wiley & Sons, Inc., State of New Jersey.
- [5] Jenkins, G.M. (1979) *Practical Experience with Modelling and Forecasting Time Series*. John Wiley & Sons, Inc., State of New Jersey.
- [6] Box, G.E.P. and Jenkins, G.M. (1970) *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, Inc., State of New Jersey.
- [7] 何新易. 基于时间序列模型的中国 GDP 增长预测分析[J]. 财经理论与实践, 2012, 33(4): 96-99.
- [8] 魏宁, 边宽江, 袁志发. 基于 ARIMA 模型的陕西省 GDP 分析与预测[J]. 安徽农业科学, 2010, 38(9): 4933-4935.
<https://doi.org/10.13989/j.cnki.0517-6611.2010.09.168>
- [9] 瞿海情, 何先平. 基于时间序列分析的湖北省 GDP 预测模型研究[J]. 湖北经济学院学报(人文社会科学版), 2021, 18(9): 37-39.