

云南大学

数学与统计学院本科生

学年论文

**题 目：基于时间序列模型探讨云南经济增长潜力**

**系 别： 统计系**

**专 业： 统计学**

**学 号： 20121910043**

**姓 名： 蓝欣欣**

**指导教师： 唐年胜（教授）**

**中文摘要**

English Abstract

1. **绪论** 
   1. **研究背景**

改革开放后,云南省经济已经保持了三十多年的高速增长。但增长方式仍然粗放，结构还比较单一，区域之间和城乡之间发展很不平衡，并且其间差距呈现出不断拉大的趋势，经济发展中还存在着诸多急需要解决的困难和问题，加快发展的任务仍然十分艰巨。探寻云南经济增长潜力

* 1. **研究意义**

1. **时间序列分析模型**
   1. **指数平滑法**
      1. **方法介绍**

指数平滑法是[布朗](http://baike.baidu.com/view/53555.htm" \t "_blank)所提出，布朗认为时间序列的态势具有稳定性或规则性，所以时间序列可被合理地顺势推延；他认为最近的过去态势，在某种程度上会持续的未来，所以将较大的权数放在最近的资料。

指数平滑法是生产预测中常用的一种方法。也用于中短期经济发展趋势预测，所有预测方法中，指数平滑是用得最多的一种。简单的全期[平均法](http://baike.baidu.com/view/1590324.htm" \t "_blank)是对时间数列的过去数据一个不漏地全部加以同等利用；移动平均法则不考虑较远期的数据，并在[加权移动平均法](http://baike.baidu.com/view/1184751.htm" \t "_blank)中给予近期资料更大的权重；而指数平滑法则兼容了全期平均和移动平均所长，不舍弃过去的数据，但是仅给予逐渐减弱的影响程度，即随着数据的远离，赋予逐渐收敛为零的权数。

也就是说指数平滑法是在[移动平均法](http://baike.baidu.com/view/627827.htm" \t "_blank)基础上发展起来的一种时间序列分析预测法，它是通过计算指数平滑值，配合一定的时间序列[预测模型](http://baike.baidu.com/view/1590251.htm" \t "_blank)对现象的未来进行预测。其原理是任一期的指数平滑值都是本期实际观察值与前一期指数平滑值的加权平均。具体表达式为：

（2.1.1）

其中，表示第t期的指数平滑值，为平滑系数。

* + 1. **模型确立**

本文研究的是1978～2013年的云南省生产总值问题，通过对数据的简单观察，不难发现该组数据具有长期趋势，所以我们可以采用分析趋势问题的指数模型。

布朗二次指数平滑法和Holt指数平滑法都是能够分析趋势问题的指数方法。前者只含有一个参数，又称单参数指数平滑法；后者含两个参数，亦称双参数指数平滑法。考虑到双参数法比单参数法增加一个参数，在处理数据变动和消除随机干扰方面应比单参数法有更大的优势性，所以本文选择Holt指数平滑法。其基本公式如下：

（2.1.2）

（2.1.3）

（2.1.4）

其中，为经过趋势调整后的指数平滑值，为趋势增量，L为预测步长，为t+L期的预测值，和为两个相互独立的平滑参数。

* + 1. **模型预测**

将1978～2013年云南省经济总量的数据录入Eviews，并选择指数平滑法中的Holt-winters无季节影响的双参数方法，令平滑参数和自动取值，通过建立的指数平滑模型得到2013～2015年的预测结果如下表1。

**表1 指数平滑法的预测值**

|  |  |  |
| --- | --- | --- |
| 年份 | GRP | 95%的置信区间 |
| 2013 | 11748.14 |  |
| 2014 | 13136.43 |  |
| 2015 | 14551.96 |  |

* 1. **ARIMA（p,d,q）模型**
     1. **理论基础**

**Ⅰ 纯AR(p)过程**

Xt=ϕ1Xt-1+ ϕ2Xt-2 + … + ϕpXt-p + μt (\*)

如果随机扰动项是一个白噪声(μt=εt)，则称(\*)式为一纯AR(p)过程（pure AR(p) process），记为：

Xt=ϕ1Xt-1+ ϕ2Xt-2 + … + ϕpXt-p +εt

**Ⅱ 纯MA(q)过程**

如果随机扰动项不是一个白噪声，通常认为它是一个q阶的移动平均（moving average）过程MA(q)：

μt=εt - θ1εt-1 - θ2εt-2 - … - θqεt-q

该式给出了一个纯MA(q)过程（pure MA(p) process）。

**Ⅲ 一般的自回归移动平均（autoregressive moving average）过程ARMA（p，q）**

将纯AR(p)与纯MA(q)结合，得到一个一般的自回归移动平均（autoregressive moving average）过程ARMA（p，q）：

Xt=ϕ1Xt-1+ ϕ2Xt-2 + … + ϕpXt-p + εt - θ1εt-1 - θ2εt-2 - … - θqεt-q

该式表明：

（1）一个随机时间序列可以通过一个自回归移动平均过程生成，即该序列可以由其自身的过去或滞后值以及随机扰动项来解释。

（2）如果该序列是平稳的，即它的行为并不会随着时间的推移而变化，那么我们就可以通过该序列过去的行为来预测未来。这也正是随机时间序列分析模型的优势所在。

**Ⅳ 自回归单整移动平均时间序列ARIMA(p，d，q)**

ARIMA模型全称为自回归移动平均模型(Autoregressive Integrated Moving Average Model，简记ARIMA)，是由[博克思](http://wiki.mbalib.com/w/index.php?title=%E5%8D%9A%E5%85%8B%E6%80%9D&action=edit)([Box](http://wiki.mbalib.com/w/index.php?title=Box&action=edit))和[詹金斯](http://wiki.mbalib.com/w/index.php?title=%E8%A9%B9%E9%87%91%E6%96%AF&action=edit)([Jenkins](http://wiki.mbalib.com/w/index.php?title=Jenkins&action=edit))于70年代初提出的一著名时间序列预测方法，所以又称为box-jenkins模型、博克思-詹金斯法。其中ARIMA（p，d，q）称为差分自回归移动平均模型，AR是自回归， p为自回归项; MA为移动平均，q为移动平均项数，d为时间序列成为平稳时所做的差分次数。

如果我们将一个非平稳时间序列通过d次差分，将它变为平稳的，然后用一个平稳的ARMA(p，q)模型作为它的生成模型，则我们就说该原始时间序列是一个自回归单整移动平均（autoregressive integrated moving average）时间序列，记为ARIMA(p，d，q)。

ARIMA模型的基本思想是：将预测对象随时间推移而形成的数据序列视为一个随机序列，用一定的数学模型来近似描述这个序列。这个模型一旦被识别后就可以从时间序列的过去值及现在值来预测未来值。现代统计方法、[计量经济模型](http://wiki.mbalib.com/wiki/%E8%AE%A1%E9%87%8F%E7%BB%8F%E6%B5%8E%E6%A8%A1%E5%9E%8B)在某种程度上已经能够帮助企业对未来进行预测。

* + 1. **建立模型**

**Ⅰ 平稳性判断**

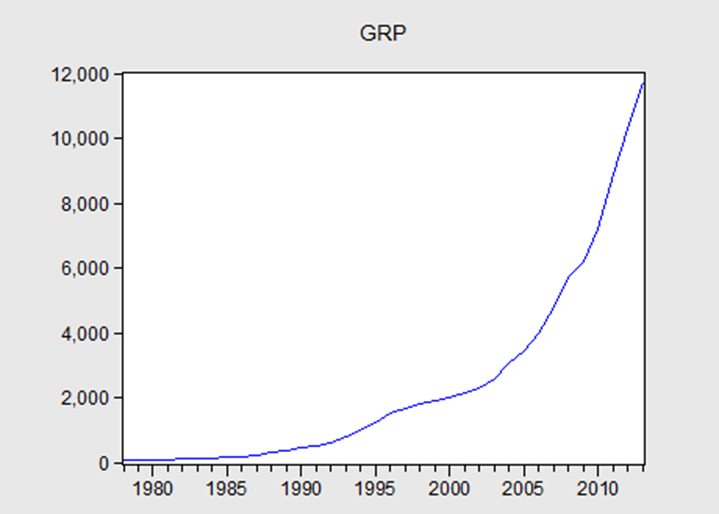
****

图1 序列云南省生产总值的曲线图

从图中可以看出，中国从1978年到2013年的云南省生产总值具有明显的上升趋势，显现出指数增长的趋势，初步识别为一个非平稳序列，接下来进行ADF检验来判断该序列的平稳性。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Null Hypothesis: GRP has a unit root | | | |  |
| Exogenous: Constant | | |  |  |
| Lag Length: 6 (Automatic based on SIC, MAXLAG=9) | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  | t-Statistic | Prob.\* |
|  |  |  |  |  |
|  |  |  |  |  |
| Augmented Dickey-Fuller test statistic | | | 0.842979 | 0.9931 |
| Test critical values: | 1% level |  | -3.679322 |  |
|  | 5% level |  | -2.967767 |  |
|  | 10% level |  | -2.622989 |  |
|  |  |  |  |  |
|  |  |  |  |  |
| \*MacKinnon (1996) one-sided p-values. | | | |  |
|  |  |  |  |  |

图2 序列云南省生产总值的ADF检验结果

由上图可看出，云南省生产总值序列以较大的P值，即99.31％的概率接受原假设，即存在单位根，序列非平稳。因此应该先对其做平稳化处理。

**Ⅱ 序列平稳化**

首先考虑取对数，做出云南省生产总值的曲线图及ADF检验，发现其仍然呈现非平稳趋势。因此考虑取对数后再进行一阶差分，记为G。

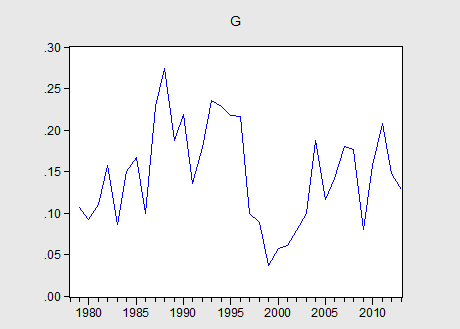


图3序列G的曲线图

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Null Hypothesis: G has a unit root | | | |  |
| Exogenous: Constant | | |  |  |
| Lag Length: 0 (Automatic based on SIC, MAXLAG=8) | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  | t-Statistic | Prob.\* |
|  |  |  |  |  |
|  |  |  |  |  |
| Augmented Dickey-Fuller test statistic | | | -3.126231 | 0.0340 |
| Test critical values: | 1% level |  | -3.639407 |  |
|  | 5% level |  | -2.951125 |  |
|  | 10% level |  | -2.614300 |  |
|  |  |  |  |  |
|  |  |  |  |  |
| \*MacKinnon (1996) one-sided p-values. | | | |  |

图4 序列G的ADF检验结果

通过看图，可初步识别序列已平稳。ADF的值为-3.126231，通过置信度为5%及10%的检验，因此我们可以看作它通过了ADF检验，为一平稳序列。在这里应该注意的是要防止过度差分。一般来说平稳序列差分得到的仍然是平稳序列，但当差分次数过多时存在两个缺点，（1）序列的样本容量减小；（2）方差变大；所以建模过程中要防止差分过度。对于一个序列，差分后若数据的极差变大，说明差分过度。此处，我们认为一阶差分已可以消除序列的非平稳性。

**Ⅲ 模型识别**

利用Eviews6做出二阶对数差分的自相关以及偏相关函数图，以判断模型阶数。

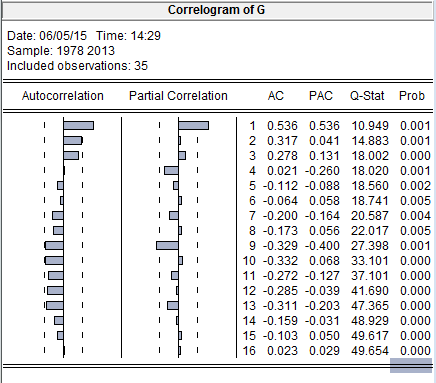


图5 序列G的自相关及偏相关函数图

可以看出，自相关系数在3阶后结尾，偏相关系数在1阶后结尾，初步取识别模型结果为p=3，q=1，建立ARIMA（1，1，3）模型。

* + 1. **选择最优的ARIMA(p,d,q)模型**

利用Eviews6对模型进行估计。我们可以对同一个平稳序列建立多个适合模型，但比较AIC和SC的值，以及综合考虑其他检验统计量，考虑模型的简约原则，我们认为ARMA（1，1，3）模型是较优选择。 如下图所示：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent Variable: G | | |  |  |
| Method: Least Squares | | |  |  |
| Date: 06/07/15 Time: 11:27 | | |  |  |
| Sample (adjusted): 1980 2013 | | |  |  |
| Included observations: 34 after adjustments | | | |  |
| Convergence achieved after 26 iterations | | | |  |
| MA Backcast: 1977 1979 | | |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  | Coefficient | Std. Error | t-Statistic | Prob. |
|  |  |  |  |  |
|  |  |  |  |  |
| C | 0.146866 | 0.021969 | 6.685038 | 0.0000 |
| AR(1) | 0.268725 | 0.422466 | 0.636088 | 0.5297 |
| MA(1) | 0.205752 | 0.394387 | 0.521702 | 0.6058 |
| MA(2) | 0.266760 | 0.221642 | 1.203563 | 0.2385 |
| MA(3) | 0.366705 | 0.211745 | 1.731822 | 0.0939 |
|  |  |  |  |  |
|  |  |  |  |  |
| R-squared | 0.344934 | Mean dependent var | | 0.147869 |
| Adjusted R-squared | 0.254580 | S.D. dependent var | | 0.059950 |
| S.E. of regression | 0.051760 | Akaike info criterion | | -2.949356 |
| Sum squared resid | 0.077693 | Schwarz criterion | | -2.724891 |
| Log likelihood | 55.13905 | Hannan-Quinn criter. | | -2.872807 |
| F-statistic | 3.817591 | Durbin-Watson stat | | 1.927446 |
| Prob(F-statistic) | 0.013014 |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Inverted AR Roots | .27 | |  |  |
| Inverted MA Roots | .22+.71i | .22-.71i | -.65 | |
|  |  |  |  |  |
|  |  |  |  |  |

图6 ARIMA（1，1，3）模型

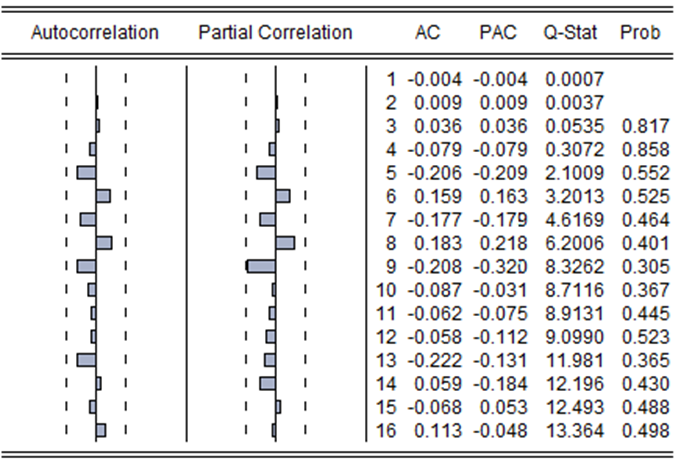


图7 ARIMA（1，1，3）模型的残差序列的相关图

对ARIMA（1，1，3）模型的残差序列进行白噪声检验，可以看出ACF和PACF都没有显著异于零，Q统计量的P值都远远大于0.05，因此我们可以认为该残差序列为白噪声序列，模型信息提取比较充分。

* + 1. **ARIMA(1,1,3)模型**

grp＝0.146866＋0.268725grp+ 0.268725+0.2667600.366705

* + 1. **模型预测**

采用ARIMA(1,1,3)模型预测2012~2015年云南省生产总值数据如下：

|  |  |  |
| --- | --- | --- |
| 年份 | GRP | 95%的置信区间 |
| 2013 | 11940.41 |  |
| 2014 | 13575.15 |  |
| 2015 | 15722.72 |  |

1. **模型预测比较**
2. **总结**