ARIMA MODELLING FOR ORISSA GSDP

[1]Bhavna Seth
Assistant Professor
Department of Economics
Dyal Singh College, University of Delhi, New Delhi, India.

Abstract : Orissa is one of the fastest growing state economies in India with a growth rate of 8.48% in 2014-2015. GSDP of Orissa for the years 1980 to 2009 at 2004 2005 constant prices is aimed to be forecasted for 2009 in this paper. To model the GSDP series and to forecast it , we are using **Box Jenkins Methodology.** The paper discusses step wise Box Jenkins Methodology, including Identification, Estimation of the model, Diagnostic Checking, and Forecasting. Using Eviews software, we find our model forecasts very close to the actual model. Hence, we conclude that our model fits the data very well and can be used for further forecasting.

IndexTerms : Forecasting, GSDP, Box Jenkins Methodology

I. INTRODUCTION

Orissa is one of the fastest growing state economies in India with a growth rate of 8.48% in 2014-2015. GSDP of Orissa for the years 1980 to 2009 at 2004 2005 constant prices is aimed to be forecasted for 2009 in this paper. To model the GSDP series and to forecast it , we are using **Box Jenkins Methodology.** The paper discusses step wise Box Jenking Methodology, including Identification, Estimation of the model, Diagnostic Checking, and Forecasting.

II. DATA AND METHODOLGY

Box Jenking Methodoly

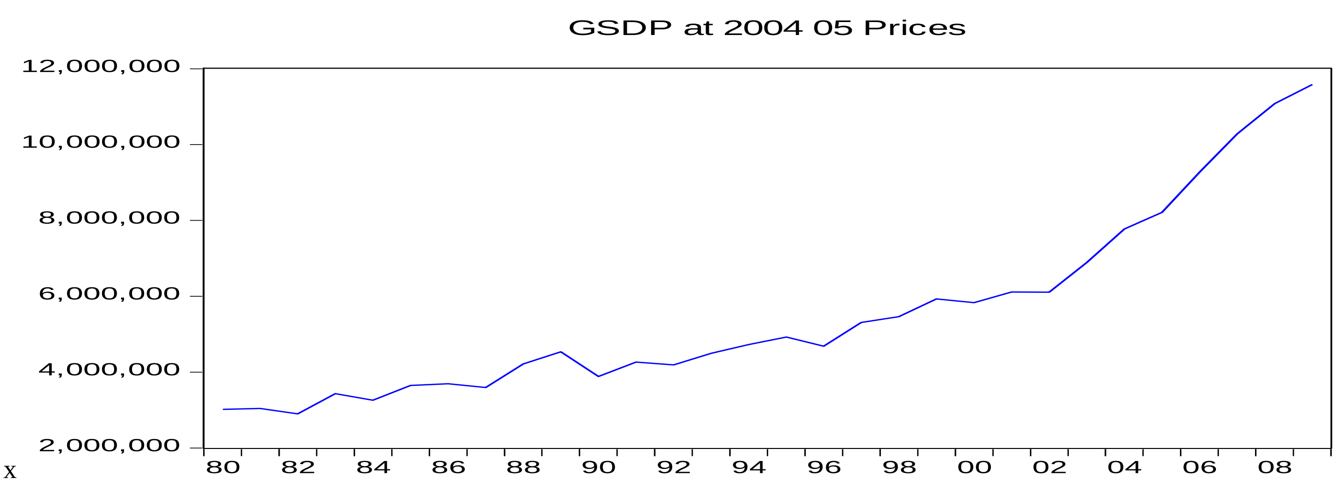
Box Jenking Methodology comes in handy in knowing whether the series follows a purely AR process (and the lag p) , MA process (and the lag q) or an ARMA process (p, q). The methods consists of these steps :

Step 1 : Identification

This step involves finding out the appropriate p (AR lag), q(MA lag) and d(number of differencing reqired to make the series stationary). The chief tools in this step are autocorrelation function(ACF) , partial autocorrelation function (PACF), and the correlograms, which plots ACF and PACF against each lag.

The following graph of the GSDP series for the years 1980 to 2009 shows an upward trend , which means the data series is non stationary. Also, the ACF declines over time and are statistically significant for 16 lags.

Since the Orissa GSDP series is nonstationary, we need to make it stationary so that we can apply Box Jenking Methodology to it.



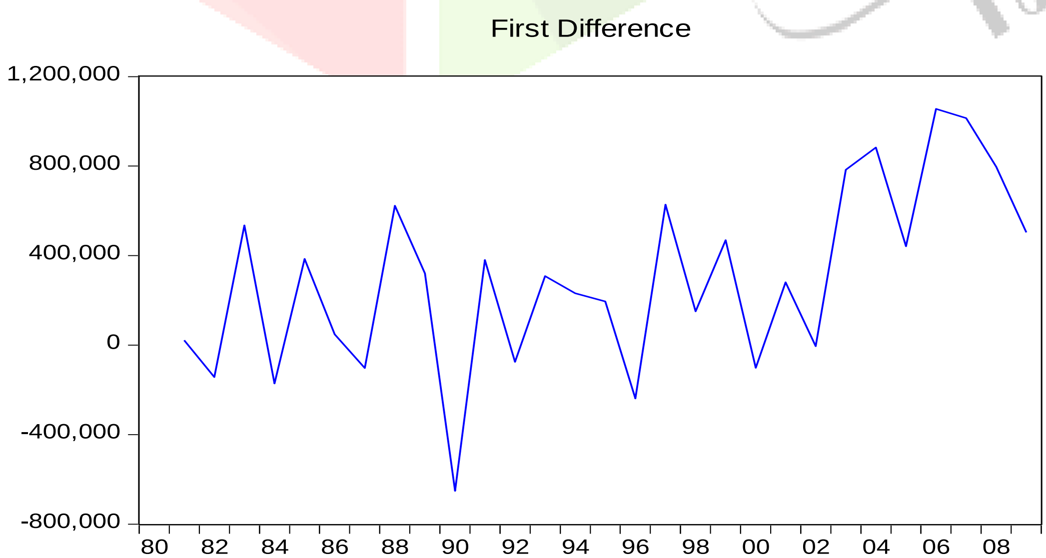
x

1.1 Nonstationarity of GSDP series

<table>
<tr><th>Autocorrelation</th><th>Partial Correlation</th><th></th><th>AC</th><th>PAC</th><th>Q-Stat</th><th>Prob</th></tr>
<tr><td>. |******|</td><td>. |******|</td><td>1</td><td>0.853</td><td>0.853</td><td>24.100</td><td>0.000</td></tr>
<tr><td>. |***** |</td><td>. *| . |</td><td>2</td><td>0.706</td><td>-0.082</td><td>41.173</td><td>0.000</td></tr>
<tr><td>. |**** |</td><td>. *| . |</td><td>3</td><td>0.562</td><td>-0.074</td><td>52.387</td><td>0.000</td></tr>
<tr><td>. |*** |</td><td>. | . |</td><td>4</td><td>0.443</td><td>0.003</td><td>59.636</td><td>0.000</td></tr>
<tr><td>. |**. |</td><td>. | . |</td><td>5</td><td>0.345</td><td>-0.010</td><td>64.196</td><td>0.000</td></tr>
<tr><td>. |**. |</td><td>. | . |</td><td>6</td><td>0.256</td><td>-0.040</td><td>66.814</td><td>0.000</td></tr>
<tr><td>. |* . |</td><td>. | . |</td><td>7</td><td>0.185</td><td>-0.004</td><td>68.237</td><td>0.000</td></tr>
<tr><td>. |* . |</td><td>. | . |</td><td>8</td><td>0.138</td><td>0.034</td><td>69.070</td><td>0.000</td></tr>
<tr><td>. |* . |</td><td>. | . |</td><td>9</td><td>0.092</td><td>-0.048</td><td>69.453</td><td>0.000</td></tr>
<tr><td>. | . |</td><td>. | . |</td><td>10</td><td>0.051</td><td>-0.020</td><td>69.580</td><td>0.000</td></tr>
<tr><td>. | . |</td><td>. *| . |</td><td>11</td><td>-0.013</td><td>-0.123</td><td>69.588</td><td>0.000</td></tr>
<tr><td>. | . |</td><td>. | . |</td><td>12</td><td>-0.060</td><td>0.008</td><td>69.781</td><td>0.000</td></tr>
<tr><td>. *| . |</td><td>. | . |</td><td>13</td><td>-0.111</td><td>-0.065</td><td>70.474</td><td>0.000</td></tr>
<tr><td>. *| . |</td><td>. | . |</td><td>14</td><td>-0.137</td><td>0.027</td><td>71.598</td><td>0.000</td></tr>
<tr><td>. *| . |</td><td>. *| . |</td><td>15</td><td>-0.173</td><td>-0.087</td><td>73.512</td><td>0.000</td></tr>
<tr><td>.**| . |</td><td>. | . |</td><td>16</td><td>-0.205</td><td>-0.043</td><td>76.406</td><td>0.000</td></tr>
</table>

Differencing to make the series stationary

Since we find our GSDP series to be non stationary, we need to make it stationary for further process. To make this series stationar , we difference it once. As it can be see that the first differenced GSDP series does not show a particular trend and the ACF and PACF for all the lags are insignificant. This shows the first difference of GSDP is series is stationary. Therefore, we can go ahead with Box Jenkins Methodology.



Date: 10/03/16 Time: 22:46
Sample: 1980 2009

Included observations: 29

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. |* . |	. |* . |	1	0.167	0.167	0.8983	0.343
. |**. |	. |**. |	2	0.324	0.304	4.3850	0.112
. |**. |	. |* . |	3	0.247	0.180	6.5007	0.090
. |* . |	. | . |	4	0.154	0.020	7.3513	0.118
. | . |	. *| . |	5	0.071	-0.079	7.5400	0.183
. |**. |	. |* . |	6	0.217	0.146	9.3812	0.153
.**| . |	***| . |	7	-0.238	-0.352	11.697	0.111
. | . |	. | . |	8	0.067	0.030	11.888	0.156
. | . |	. |* . |	9	0.054	0.186	12.021	0.212
. |* . |	. |* . |	10	0.077	0.169	12.303	0.265
. *| . |	. *| . |	11	-0.090	-0.202	12.709	0.313
. | . |	. *| . |	12	-0.020	-0.176	12.730	0.389

Step 2. Estimation of the Model

Since we have a stationary series with us, we can go ahead with estimation of the model.The following models were tested for significance and

Model	AIC	SIC	Significance
ARMA(1,0)	28.76476	28.85992	Not significant
ARMA(1,1)	28.56931	28.71205	Intercept not Significant, AR(1) MA(1) Significant
ARMA(1,2)	28.61988	28.81020	MA(2) Not significant
ARMA(1,3)	28.72512	28.96301	MA(2) MA(3) not significant
ARMA(2,0)	27.167	28.86557	Not significant
ARMA(2,1)	27.98860	28.18058	AR(2) not significant
ARMA(2,2)	28.441177	28.68114	Not Significant
ARMA(2,3)	28.74117	29.02913	Not Significant
ARMA(3,0)	28.61317	28.90565	Not Significant
ARMA(3,1)	28.66371	28.90565	AR(2) AR(3) not significant
ARMA(3,2)	28.67200	28.96233	Not Significant
ARMA(3,3)	28.09018	28.42890	Not Significant
ARMA(0,1)	28.75280	28.84710	Not Significant
ARMA(0,2)	28.72001	28.86145	Not Significant
ARMA(0,3)	28.76922	28.95781	Not Significant

As it can be seen from the above table , the only model with significant results is ARIMA(1,1,1) without intercept . We are dropping the intercept because it comes out to be insignificant in the results. Also , it has the lowed SIC and AIC. The result of the model is shown below:

Major Findings:

- Intercept is insignificant
- AR(1) and MA(1) are significant
- AIC SIC are low
- R^2 is 0.25 , showing 25% of the variation is explained by AR(1) and MA(1)

Model : ARIMA(1,1,1)

$$\Delta Y_t = 0.9619\ \Delta Y_{t-1} - 0.9999\ u_{t-1}$$

$$Y_t - Y_{t-1} = 0.9619\ (Y_{t-1} - Y_{t-2}) - 0.9999\ u_{t-1}$$

Se = (0.068249) (0.165148)

T = (14.09421) (14.09421)

Where ΔY_t = first differences of Orissa GSDP

ARMA(1,1)

Dependent Variable: DY
Method: Least Squares
Date: 10/03/16 Time: 23:32
Sample (adjusted): 1982 2009
Included observations: 28 after adjustments
Failure to improve SSR after 39 iterations
MA Backcast: 1981

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	818439.0	741446.6	1.103841	0.2802
AR(1)	0.961918	0.068249	14.09421	0.0000
MA(1)	-0.999901	0.165148	c	0.0000

R-squared	0.256351	Mean dependent var	305138.4
Adjusted R-squared	0.196859	S.D. dependent var	410385.4
S.E. of regression	367779.7	Akaike info criterion	28.56931
Sum squared resid	3.38E+12	Schwarz criterion	28.71205
Log likelihood	-396.9704	Hannan-Quinn criter.	28.61295
F-statistic	4.309010	Durbin-Watson stat	2.145701
Prob(F-statistic)	0.024666		

Inverted AR Roots	.96
Inverted MA Roots	1.00

Date: 10/04/16 Time: 00:07
Sample: 1982 2009
Included observations: 28
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
.**| . |	.**| . |	1	-0.285	-0.285	2.5335	
. |* . |	. | . |	2	0.106	0.027	2.8984	
.**| . |	.**| . |	3	-0.320	-0.308	6.3422	0.012
. | . |	.**| . |	4	-0.027	-0.235	6.3678	0.041
. *| . |	.**| . |	5	-0.122	-0.232	6.9145	0.075
. |* . |	. | . |	6	0.184	-0.035	8.2056	0.084
. |* . |	. | . |	7	0.074	0.041	8.4235	0.134
. *| . |	. *| . |	8	-0.091	-0.203	8.7689	0.187
.**| . |	***| . |	9	-0.210	-0.389	10.710	0.152
. |* . |	. | . |	10	0.201	0.067	12.597	0.126
. *| . |	. | . |	11	-0.067	-0.024	12.818	0.171
. |* . |	. *| . |	12	0.105	-0.185	13.393	0.203

Step 3: Diagnostic Checking

One simple diagnostic is to obtain residuals of the model estimated and obtain the ACF and PACF of these residuals. The ACF and PACF are shown below :

Date: 10/15/16 Time: 13:37
Sample: 1982 2009
Included observations: 28
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
. *| . |	. *| . |	1	-0.075	-0.075	0.1763	
. |* . |	. |* . |	2	0.172	0.168	1.1362	
. |* . |	. |* . |	3	0.090	0.117	1.4090	0.235
. | . |	. | . |	4	0.037	0.024	1.4563	0.483
. *| . |	. *| . |	5	-0.076	-0.113	1.6690	0.644
. |* . |	. |* . |	6	0.173	0.145	2.8136	0.589
***| . |	***| . |	7	-0.374	-0.352	8.4098	0.135
. | . |	. *| . |	8	-0.022	-0.112	8.4308	0.208
. | . |	. | . |	9	-0.025	0.073	8.4585	0.294
. | . |	. |* . |	10	0.056	0.178	8.6059	0.377
. *| . |	. *| . |	11	-0.148	-0.120	9.6818	0.377
. | . |	. *| . |	12	-0.027	-0.182	9.7203	0.465

As we can see from the above, none of the ACF or PACFs are significant . Therefore, there is not be any need to look for another ARIMA model.

Step 4: Forecasting

On the basis of model , we want to forecaste GSDP for the year 2010 . But the series we have is the differenced GSDP, So we will have to integrate the first differenced series. Thus to obtain the forecaste value of GSDP, we rewrite the model as :

ARIMA (1,1,1)

$$Y_t - Y_{t-1} = 0.9619 (Y_{t-1} - Y_{t-2}) - 0.9999 u_{t-1}$$

$$Y_t = Y_{t-1} + 0.9619 Y_{t-1} + 0.9619Y_{t-2} – 0.9999 u_{t-1}$$

$$Y_{2010} = 1.9619 Y_{2009} – 0.9619 Y_{2008} - 0.9999 u_{2009}$$

$$Y_{2010} = 1.9619 (11585113) – 0.9619(11081178) – 0.9999(-16613.27)$$

$$Y_{2010} = 120886459.685$$

Similarly, we can forecast the values of further years also by putting the values in the equaltion. The following table shows the forecasted values as per our model:

Year	GSDP at 2004 05 Prices	First Difference (ΔYt)	ut	Forecated Y
1980	3019766			
1981	3041237	21471		
1982	2898300	-142937	-111555	
1983	3433353	535053	534071.9	
1984	3261593	-171760	-203871	3414002.01
1985	3646900	385307	323250.8	3300228.07

1986	3694728	47828	-43032.6	3694308.3
1987	3592204	-102524	-221092	3783762.05
1988	4214647	622443	477222.5	3714656.16
1989	4535081	320434	149576.2	4336200.1
1990	3883162	-651919	-847438	4693745.25
1991	4263824	380662	161421.2	4103434.26
1992	4188390	-75434	-317493	4468577.74
1993	4496557	308167	44158.05	4433291.65
1994	4728488	231931	-53191.6	4748829.2
1995	4923531	195043	-110389	5004768.75
1996	4684672	-238859	-563828	5221521.12
1997	5311965	627293	283532.2	5018684.66
1998	5462975	151010	-210827	5631854.26
1999	5932446	469471	90245.41	5819037.81
2000	5830376	-102070	-498022	6293793.77
2001	6110766	280390	-131651	6230166.69
2002	6105838	-4928	-432445	6512110.69
2003	6889860	784022	341617.9	6533499.6
2004	7772943	883083	426358.8	7302427.03
2005	8214472	441529	-28969.9	8196064.37
2006	9270083	1055611	571861.9	8668145.79
2007	10284562	1014479	517984.3	9713670.54
2008	11081178	796616	287861.1	10742456.9
2009	11585113	503935	-16613.3	11559610.7
Forecasting (2010)				12086459.7

III. CONCLUSION

As we can see, the forecasted values are very close to the actual values of GSDP. Therefore, We can conclude that our model is good enough to forecast the future values of the GSDP.

REFERENCES

[1]Andersen, A., and Weiss, A., (1984) "Forecasting: The Box-Jenkins Approach" in The Forecasting Accuracy of Major Time Series Methods, Makridakis et. al, John Wiley & Sons, Chichester, England.

[2]Box, G.E.P., and Jenkins, G., (1970) Time Series Analysis, Forecasting and Control, HoldenDay, San Francisco.

[3]Box, G.E.P., and Pierce, D.A., (1970) "Distribution of the Residual Autocorrelations in Autoregressive-Integrated Moving-Average Time Series Models", Journal of the American Statistical Association, 65, 1509-1526.

[4]Markridakis S. and M. Hibon, “ARMA Models and the Box Jenkins Methodoly”, Working Paper, INSTEAD