Time series analysis

Lecture 7. Unit root and structure changes

Dr. Khamidov Obidjon

Backshift notation

- Backward shift operator, B, is defined as

$$BY_t = Y_{t-1}$$

- Two applications of B to Y_t, shifts the data back two periods:

$$B(BY_t) = B^2 Y_t = Y_{t-2}$$

- A shift to the same quarter last year will use B^4 which is

$$B^4 Y_t = Y_{t-4}$$

Backshift notation

- The backward shift operator can be used to describe the differencing process. A first difference can be written as

$$Y_t' = Y_t - Y_{t-1} = Y_t - BY_t = (1-B)Y_t$$

- The second order differences as

$$\begin{aligned} Y_t'' &= (Y_t' - Y_{t-1}') \\ &= (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) \\ &= Y_t - 2Y_{t-1} + Y_{t-2} \\ &= (1 - 2B + B^2)Y_t \\ &= (1-B)^2 Y_t \end{aligned}$$

Backshift notation

- Example;
  - ARMA(1,1) or ARIMA(1,0,1) model

$$Y_t = c + \phi_1 Y_{t-1} + e_t - \theta_1 e_{t-1}$$

$$(1 - \phi_1 B) Y_t = c + (1 - \theta_1 B) e_t$$

  - ARMA(p,q) or ARIMA(p,0,q) model

$$Y_t = c + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}$$

$$(1 - \phi_1 B - \cdots - \phi_p B^p) Y_t = c + (1 - \theta_1 B - \cdots - \theta_q B^q) e_t$$

Backshift notation

– ARIMA(1,1,1)

$$Y_t - Y_{t-1} = c + \phi_1 (Y_{t-1} - Y_{t-2}) + e_t - \theta_1 e_{t-1}$$

$$(1 - \phi_1 B)(1 - B) Y_t = c + (1 - \theta_1 B) e_t$$

Estimating the parameters

- Once a tentative model has been selected, the parameters for the model must be estimated.
- The method of least squares can be used for RIMA model.
- However, for models with an MA components, there is no simple formula that can be used to estimate the parameters.
- Instead, an iterative method is used. This involves starting with a preliminary estimate, and refining the estimate iteratively until the sum of the squared errors is minimized.

Estimating the parameters

- Another method of estimating the parameters is the maximum likelihood procedure.
- Like least squares methods, these estimates must be found iteratively.
- Maximum likelihood estimation is usually favored because it has some desiable statistical properties.

Estimating the parameters

- After the estimates and their standard errors are determined, t values can be constructed and interpreted in the usual way.
- Parameters that are judged significantly different from zero are retained in the fitted model; parameters that are not significantly different from zero are dropped from the model.

Estimating the parameters

- There may have been more than one plausible model identified, and we need a method to determine which of them is preferred.
- Akaike's Information Criterion (AIC)

$$AIC = -2\log L + 2m$$

  - L denotes the likelihood
  - m is the number of parameters estimated in the model: m = p+q+P+Q

Estimating the parameters

- Because not all computer programs produce the AIC or the likelihood L, it is not always possible to find the AIC for a given model.
- A useful approximation to the AIC is:

$$AIC = n(1+\log(2\pi)) + n\log\sigma^2 + 2m$$

Diagnostic Checking

- Before using the model for forecasting, it must be checked for adequacy.
- A model is adequate if the residuals left over after fitting the model is simply white noise.
- The pattern of ACF and PACF of the residuals may suggest how the model can be improved.

Diagnostic Checking

- For example
  - Significant spikes at the seasonal lags suggests adding seasonal component to the chosen model
  - Significant spikes at small lags suggest increasing the non-seasonal AR or MA components of the model.

Diagnostic Checking

- A portmanteau test can also be applied to the residuals as an additional test of fit.
- If the portmanteau test is significant, then the model is inadequate.
- In this case we need to go back and consider other ARIMA models.
- Any new model will need their parameters estimated and their AIC values computed and compared with other models.

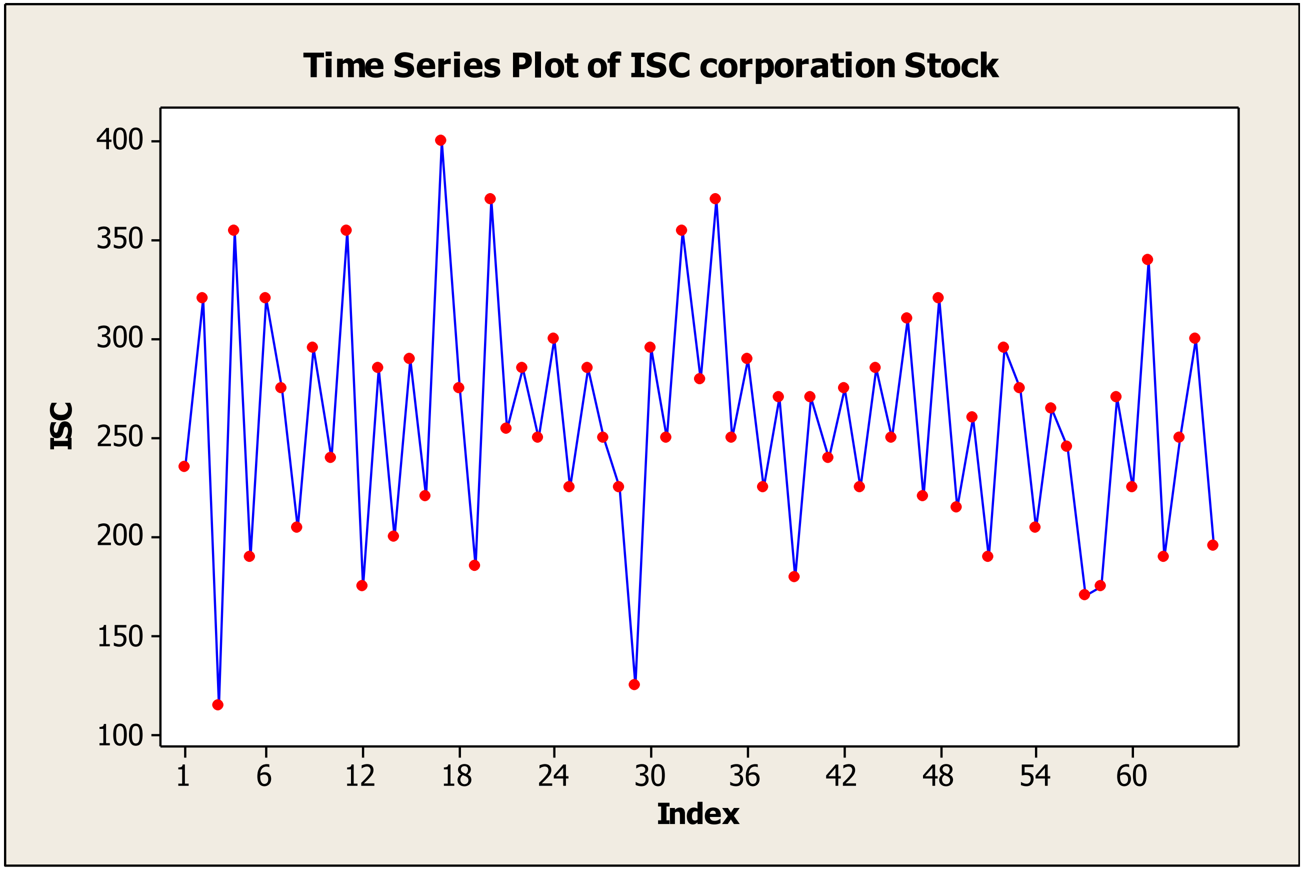
Diagnostic Checking

- Usually, the the model with the smallest AIC will have residuals which resemble white noise.
- Occasionally, it might be necessary to adopt a model with not quite the smallest AIC value, but with better behaved residuals.

Example

- The analyst for the ISC Corporation was asked to develop forecasts for the closing prices of ISC stock. The stock has been languishing for some time with little growth, and senior management wanted some projections to discuss with the board of directors. The ISC stock prices are plotted in the following slide.

Example

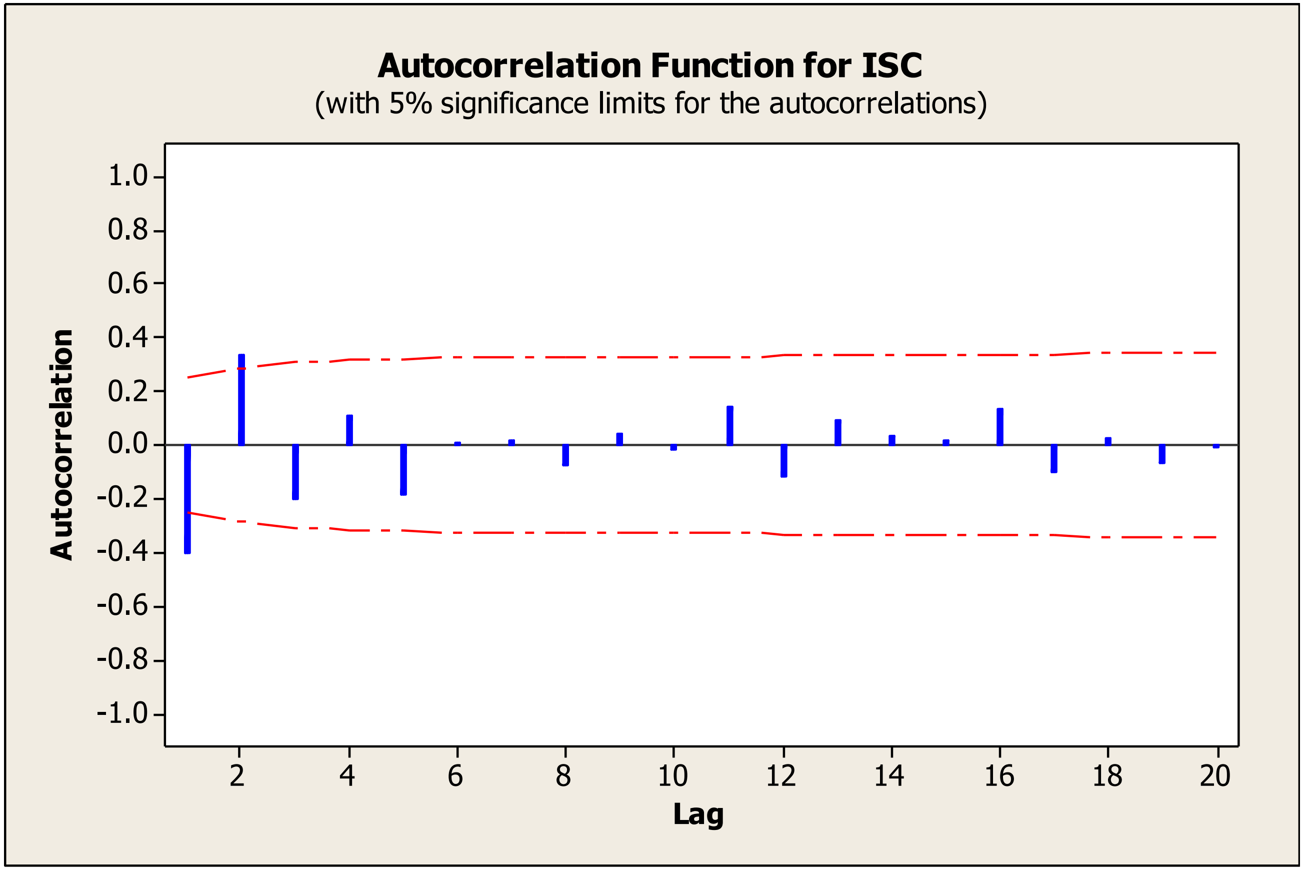


Time Series Plot of ISC corporation Stock
ISC
400
350
300
250
200
150
100
1
6
12
18
24
30
36
42
48
54
60
Index

Example

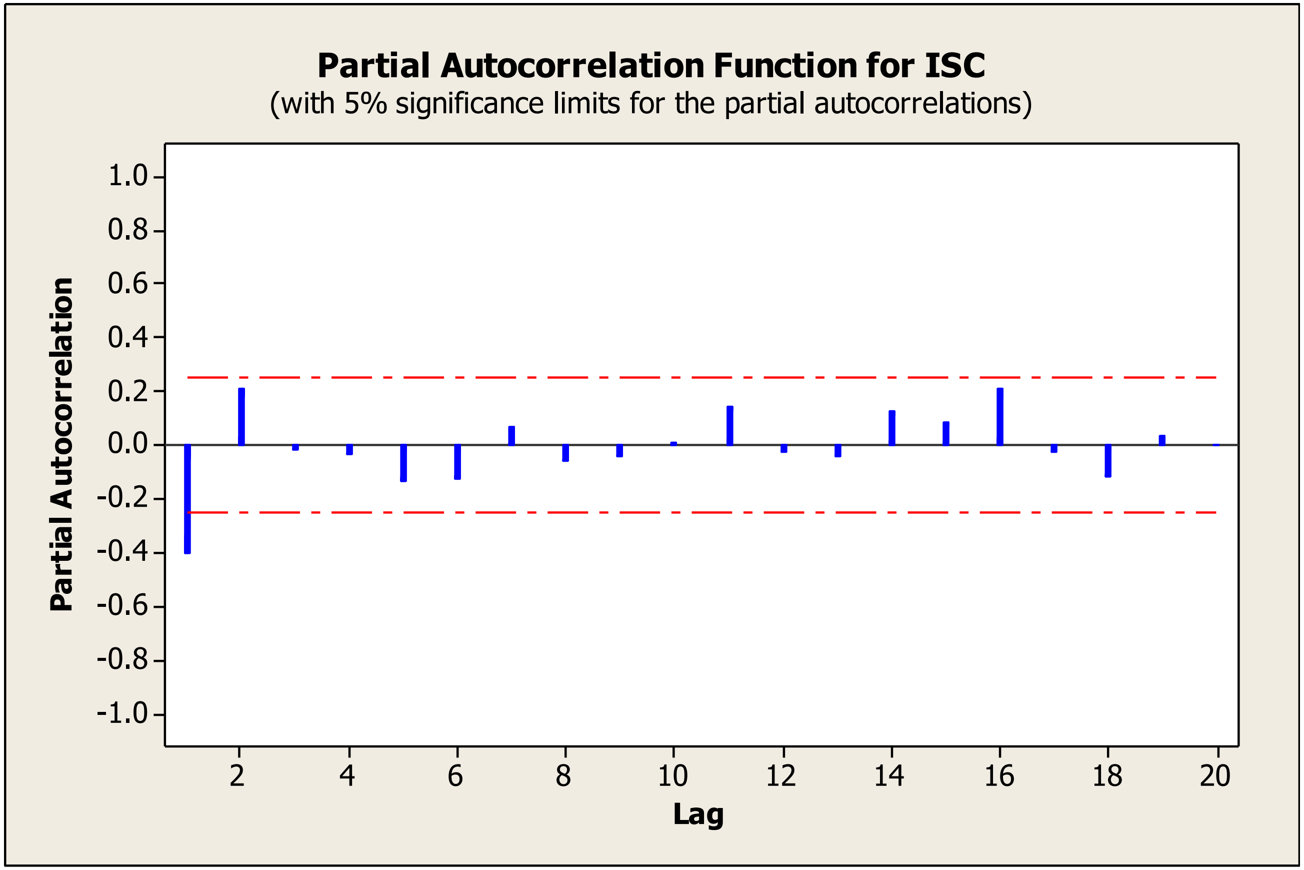
- The plot of the stock prices suggests the series is stationary.
- The stock prices vary about a fixed level of approximately 250.
- Is the Box-Jenkins methodology appropriate for this data series?
- The ACF and PACF for the stock price series are reported in the following two slides.

Example



Autocorrelation Function for ISC
(with 5% significance limits for the autocorrelations)
Autocorrelation
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
2
4
6
8
10
12
14
16
18
20
Lag

Example


Partial Autocorrelation Function for ISC
(with 5% significance limits for the partial autocorrelations)
Partial Autocorrelation
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
2
4
6
8
10
12
14
16
18
20
Lag

Example

- The sample ACF alternate in sign and decline to zero after lag 2.
- The sample PACF are similar are close to zero after time lag 2.
- These are consistent with an AR(2) or ARIMA(2,0,0) model
- Using MINITAB an AR(2) model is fit to the data.
- WE include a constant term to allow for a nonzero level.

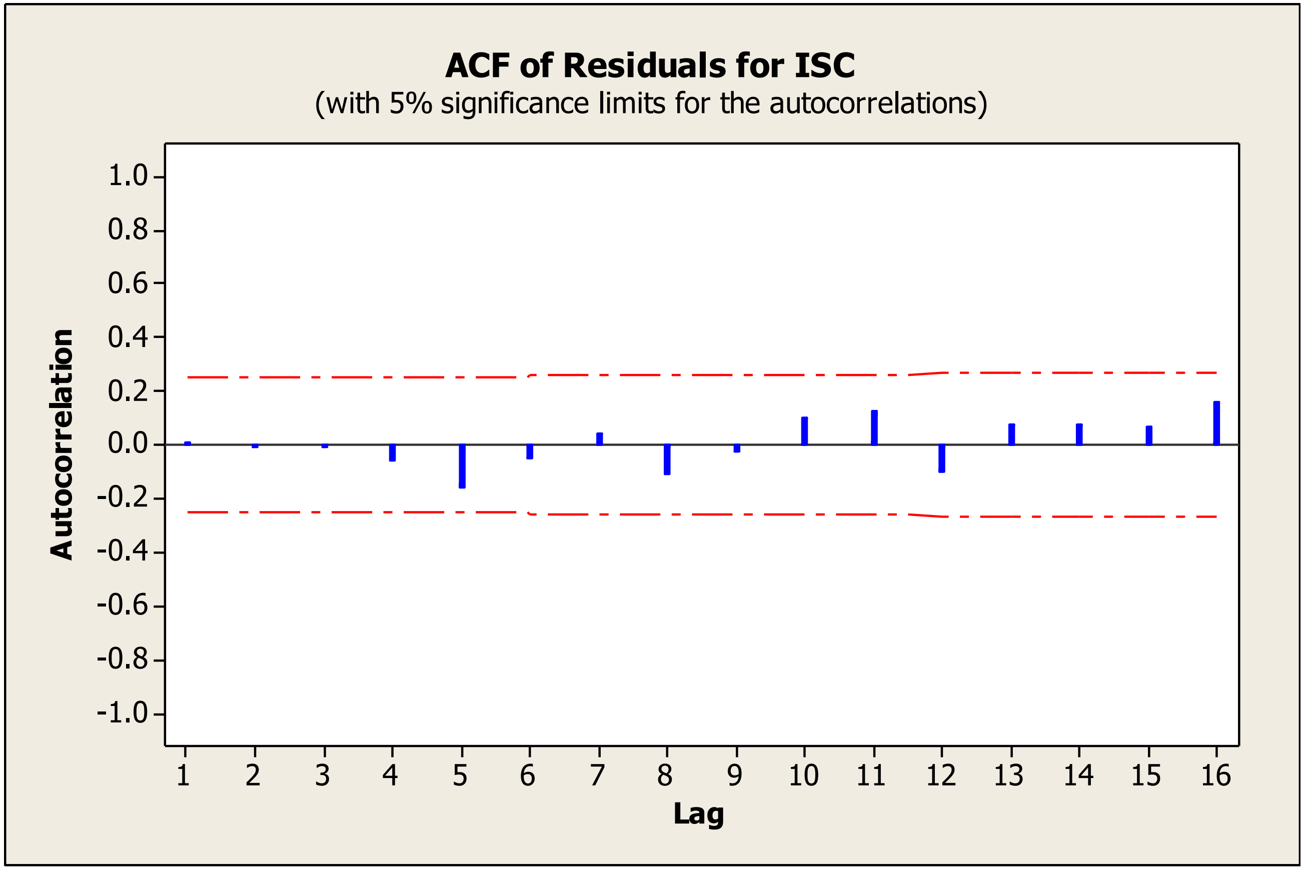
Example

- The estimated coefficient ϕ_2 is not significant (t=1.75) at 5% level but is significant at the 10% level.
- The residual ACF and PACF are given in the following two slides.
- The ACF and PACF are well within their two standard error limits.

Final Estimates of Parameters

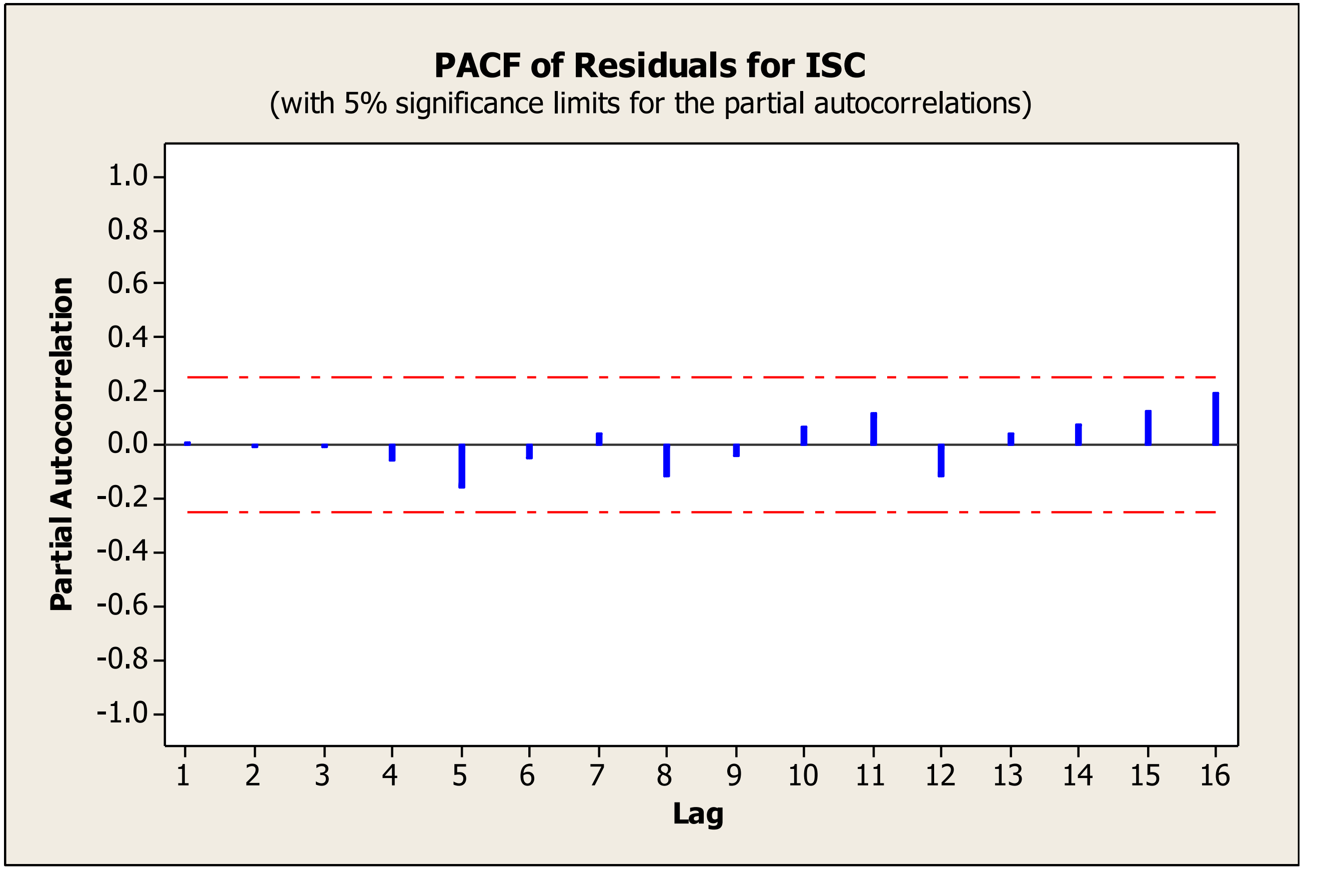
Type	Coef	SE Coef	T	P
AR 1	-0.3243	0.1246	-2.60	0.012
AR 2	0.2192	0.1251	1.75	0.085
Constant	284.903	6.573	43.34	0.000

Example



ACF of Residuals for ISC
(with 5% significance limits for the autocorrelations)
Autocorrelation
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Lag

Example


PACF of Residuals for ISC
(with 5% significance limits for the partial autocorrelations)
Partial Autocorrelation
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Lag

Example

- The p-value for the Ljung-Box statistics for m = 12, 24, 36, and 48 are all large (> 5%) indicating an adequate model.
- We use the model to generate forecasts for

MS = 2808 DF = 62

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	6.3	13.3	18.2	29.1
DF	9	21	33	45
P-Value	0.707	0.899	0.983	0.969

Example

- The forecasts are generated by the following equation.

$$\hat{Y}_t = c + \hat{\phi}_1 Y_{t-1} + \hat{\phi}_2 Y_{t-2}$$

$$\hat{Y}_{66} = 284.9 + (-.324)Y_{65} + .219Y_{64}$$

$$= 284.9 - .324(195) + .219(300) = 287.4$$

$$\hat{Y}_{67} = 284.9 + (-.324)\hat{Y}_{66} + .219Y_{65}$$

$$= 284.9 - .324(287.4) + .219(195) = 234.5$$

Example

- The 95% prediction limits are approximately

$$\hat{Y} \pm 2s$$

- The 95% prediction limits for period 66 are

$$287.4 \pm 2\sqrt{2808}$$

$$287.4 \pm 106$$

$$(181.4,\ 393.4)$$

Final comments

- In ARIMA modeling, it is not good practice to include AR and MA parameters to “cover all possibilities” suggested by the sample ACF and Sample PACF.
- This means, when in doubt, start with a model containing few parameters rather than many parameters.The need for additional parameters will be evident from the residual ACF and PACF.

Final comments

- Least square estimates of AR and MA parameters in ARIMA models tend to be highly correlated. When there are more parameters than necessary, this leads to unstable models that can produce poor forecasts.

Final comments

- To summarize, start with a small number of clearly justifiable parameters and add one parameter at a time as needed.

Final comments

- If parameters in a fitted ARIMA model are not significant, delete one parameter at a time and refit the model. Because of high correlation among estimated parameters, it may be the case that a previously non-significant parameter becomes significant.

Reference and sourse:

1. Multivariate Time Series Analysis: With R and Financial Applications by Ruey S. Tsay
2. Time Series Analysis by James Douglas Hamilton
3. The Analysis of Time Series: An Introduction with R Chapman & Hall/CRC Texts in Statistical Science)
4. Machine Learning for Time Series Forecasting with Python by Francesca Lazzeri
5. Time Series Analysis for the Social Sciences (Analytical Methods for Social Research) Part of: Analytical Methods for Social Research (14 Books)
6. Introduction to Probability, Statistics, and Random Processes by Hossein Pishro-Nik
7. Introduction to Time Series and Forecasting (Springer Texts in Statistics) Part of: Springer Texts in Statistics