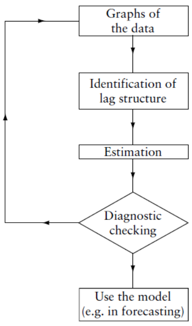
ECONOMETRICS

M.A. Program in Economics
YTU
Department of Economics

Time Series Modeling for Stationary Data and Box-Jenkins Modeling Methodology

1

Steps in Time-series Modeling



Data needs to be stationary. In practice, usually an appropriate transformation of data is used, e.g. taking first differences

2

Mixed Autoregressive Moving Average ARMA(p,q) Models

ARMA(1,1) Model

$$y_t = \alpha + \rho_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}, \quad \varepsilon_t \sim iidN(0, \sigma^2)$$

ARMA(p,q) Model

$$y_t = \alpha + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \ldots + \rho_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}$$

3

Box-Jenkins Methodology

- The orders of an ARMA process, p and q are unknown in practice
- Box - Jenkins Modeling philosophy uses sample and partial autocorrelation functions to determine the lag orders
- They advocate the benefits of parsimony, thus suggested using as few parameters as possible
- The time series under investigation should be covariance-stationary

4

Box-Jenkins Methodology

- Transform the data if necessary so that the assumption of covariance-stationarity is a reasonable one
- Make an initial guess of small values for p and q for an ARMA(p,q) model that might describe the transformed series
- Estimate the AR and MA parameters
- Perform diagnostic analysis to confirm that the model is consistent with the observed features of the data

5

How to determine ARMA orders, p and q?

- Analysis of Correlogram, e.g. for an AR(1) Model SACF will die out to zero as the lag order increases but the sample PACF will cut off at lag 1
- For an MA(1) model SACF will cut off at lag 1 but sample PACF will decrease to zero exponentially
- For an AR(p) model SACF decreases exponentially and SPACF cuts off at lag p.
- For an MA(q) model SACF cuts off at lag q but SPACF decays to zero exponentially.
- In practice the correlogram analysis may not be that clear-cut. In this case try low-order ARMA(p,q) models and choose the one with lowest Information Criteria (e.g. AIC or SBC)

6

ARIMA(p,d,q) Modeling

- If the null hypothesis cannot be rejected then the series said to have a unit root, denoted $I(1)$ series. This means that series are integrated of order 1.
- If they are $I(1)$ then they should be converted in some way to become $I(0)$ - series are integrated of order 0.
- The term d stands for the the degree of integration in a univariate time series, i.e., whether they have a unit root or not
- The usual practice is to take first differences of the series
- Then we would have an $ARIMA(p,1,q)$ model, where the lag orders for AR and MA terms are to be chosen using data based selection procedures.

7

ARIMA(p,d,q) Modling

- Identification: specify p, d and q
  - test for a unit root to determine d
  - analyze correlogram (sample ACF) and partial correlogram (sample PACF) to set initial p and q
  - take seasonal differences or deseasonalize the series using seasonal dummies or other methods
- Estimation: requires solving a set of nonlinear equations - most computer packages can handle this
- Diagnostic Checking: apply a variety of tests to see whether the estimated model fits the data adequately. if the model is not adequate repeat these steps.

8

Sample Autocorrelation Function

- The parameter γ_j is called the jth order or lag j autocovariance of y_t and a plot of γ_j against j is called the autocovariance function.
- The autocorrelations of y_t are defined by

$$\rho_j = \frac{Cov(y_t, y_{t-j})}{\sqrt{Var(y_t)Var(y_{t-j})}} = \frac{\gamma_j}{\gamma_0}$$

- Based on a sample of T observtions, ρ_j can be estimated using the following **sample autocorrelation function**

$$\hat{\gamma}_j = \frac{1}{T}\sum_{t=j+1}^{T}(y_t - \bar{y})(y_{t-j} - \bar{y}) \qquad (1)$$

$$\hat{\rho}_j = \frac{\hat{\gamma}_j}{\hat{\gamma}_0} \qquad (2)$$

where $\bar{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$ is the sample mean.

9

Correlogram

- can use sample autocorrelation function or sample correlogram to see the degree of linear dependence of a time series with its lags.
- $\hat{\rho}_j$ is normally distributed if $y_t \sim iid(0, \sigma^2)$, i.e,.

$$\hat{\rho}_j \sim N\left(0, \frac{1}{T}\right)$$

- This is based on the central limit theorem: $\sqrt{T}\hat{\rho}_j \sim N(0,1)$
- Hence the %95 confidence limit around zero is $\pm\frac{1.96}{\sqrt{T}}$

10

Partial Autocorrelation Function (PACF)

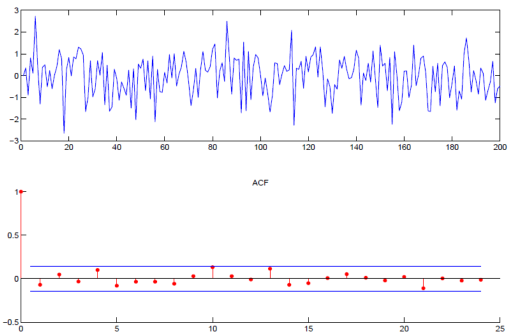
- PACF is useful tool for determining the order p of an AR model
- To find the sample PACF the following AR models are estimated in consecutive orders:

$$\begin{aligned}
y_t &= \mu_1 + \rho_{1,1}y_{t-1} + \varepsilon_{1t} \\
y_t &= \mu_2 + \rho_{2,1}y_{t-1} + \rho_{2,2}y_{t-2} + \varepsilon_{2t} \\
y_t &= \mu_3 + \rho_{3,1}y_{t-1} + \rho_{3,2}y_{t-2} + \rho_{3,3}y_{t-3} + \varepsilon_{3t} \\
y_t &= \mu_4 + \rho_{4,1}y_{t-1} + \rho_{4,2}y_{t-2} + \rho_{4,3}y_{t-3} + \rho_{4,4}y_{t-4} + \varepsilon_{4t}
\end{aligned}$$

- The estimate $\hat{\rho}_{1,1}$ from the first equation is called the lag-1 sample PACF of y_t. The estimate $\hat{\rho}_{2,2}$ from the second equation is called the lag-2 sample PACF of y_t, etc.
- For an AR(p) series, the sample PACF cuts off at lag p.
- The plot of sample PACFs against lag order is a useful tool to

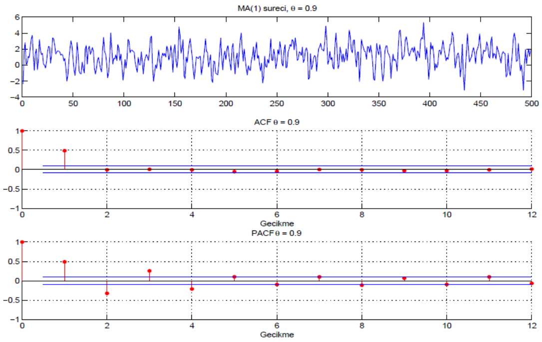
11

A Simulated White Noise Process and its Sample ACF

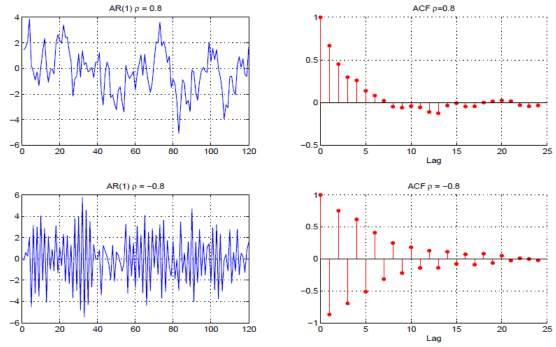


12

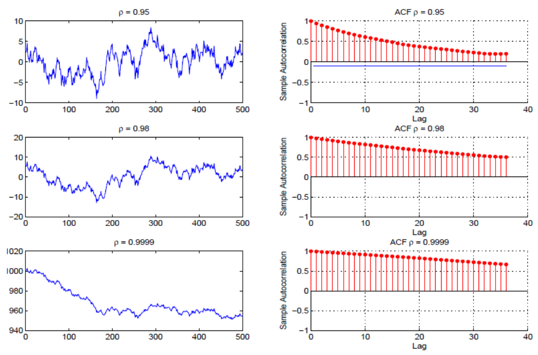
Sample ACF and PACF for a Simulated MA(1) Process



Simulated AR(1) Processes



Simulated AR(1) Processes



Diagnostic Checking

- If the model fits the data well then the residuals from the estimated model should be white noise. For this, we can use Box-Pierce statistic:

$$Q = n \sum_{k=1}^{K} r_k^2$$

where r_k is the kth order autocorrelation of residuals, n is the number of observations used in the estimation and K is a preselected number of autocorrelations (say 20 or more)

- Under the null hypothesis of white noise residuals Q will have a chi-square distribution with $K - p - q$ degrees of freedom.
- If Q is larger than the critical value than the residuals is not white noise.

Diagnostic Checking

- A more recent test is the Ljung-Box test:

$$LJB = n'(n'+2) \sum_{k=1}^{K} \left(\frac{r_k^2}{n'-k} \right)$$

where $n' = n - d$, the number of observations used after the series has been differenced d times

- LJB statistic also has a chi-square distribution with $K - p - q$ degrees of freedom.

Jarque-Bera Normality Test

- Based on the property that a normal distribution has skewness 0 and kurtosis 3.
- can be calculated as

$$JB = \frac{T}{6} \left(\hat{s}^2 + \frac{(\hat{k}-3)^2}{4} \right)$$

where $\hat{s}$ and $\hat{k}$ are sample skewness and kurtosis, respectively

- JB has a chi-squared distribution with 2 degrees of freedom, i.e.,

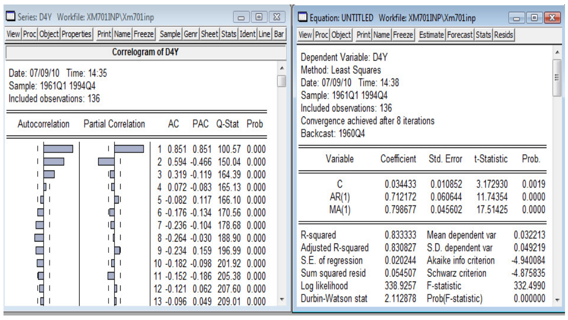
$$JB \sim \chi^2(2)$$

- the null hypothesis is that the series under study is normally distributed against the alternative that it is not.
- H_0 is rejected for large values of JB

Eviews Implementation

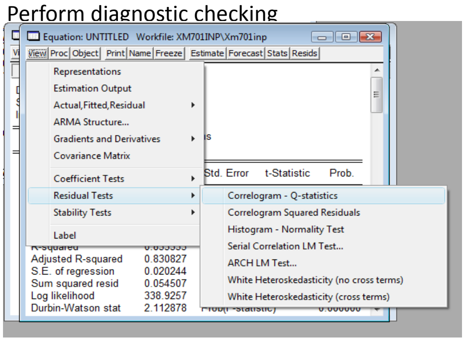
- Open the series under investigation
- Check for stationarity: View>Unit Root Test>ADF
- Or use any other unit root test, eg Philips-Perron
- Inspect SACF and SPACF, and LBQ tests
- View>Correlogram
- Transform the series if nonstationary, what is d in ARIMA(p,d,q)?
- d=1 => take first differences
- Choose lag orders p and q
- Estimate ARMA(p,q) model: quick>estimate equation

Eviews Implementation



Eviews Implementation

- Perform diagnostic checking



Read Eviews Manual Ch. 17

Vector Autoregressive Models (VAR)

- We now consider a column vector of k different variables: $\mathbf{y}_t = (\begin{array}{cccc} y_{1t} & y_{2t} & \cdots & y_{kt} \end{array})'$
- A pth order vector autoregression or VAR(p) is written as follows

$$\mathbf{y}_t = \mathbf{c} + \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{A}_2\mathbf{y}_{t-2} + \ldots + \mathbf{A}_p\mathbf{y}_{t-p} + \epsilon_t$$

where $\mathbf{c}$ is a $k \times 1$ vector of constants, $\mathbf{A}_i$, $i = 1, 2, ..., k$ is a $k \times k$ matrices of coefficients and ϵ_t is a vector of white noise process with the following properties:

$$E(\epsilon_t) = \mathbf{0} \quad \text{for all } t$$

and

$$E(\epsilon_t \epsilon_s') = \begin{cases} \mathbf{\Omega}, & s = t; \\ \mathbf{0}, & s \neq t. \end{cases}$$

Vector Autoregressive Models (VAR)

- For example for $k = 2$ and $p = 1$ we have the following system of simultaneous equations

$$\begin{aligned} \mathbf{y}_t &= \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} \\ &= \mathbf{c} + \mathbf{A}_1\mathbf{y}_{t-1} + \epsilon_t \\ &= \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix} \end{aligned}$$

- Or written out explicitly:

$$\begin{aligned} y_{1t} &= c_1 + a_{11}y_{1,t-1} + a_{12}y_{2,t-1} + \epsilon_{1t} \\ y_{2t} &= c_2 + a_{21}y_{1,t-1} + a_{22}y_{2,t-1} + \epsilon_{2t} \end{aligned}$$

Vector Autoregressive Models (VAR)

- In a VAR, each variable is expressed as a linear combination of lagged values of itself and lagged values of all other variables.
- In practice, a VAR may include deterministic terms such as time trends and/or seasonal effects, and exogenous variables
- Recall that in a univariate AR(p) process the behavior of y_t depends on the AR parameters
- Similarly for VAR(p)s the behavior of $\mathbf{y}_t$s will depend on the properties of the coefficient matrices $\mathbf{A}_i$s

Vector Autoregressive Models (VAR)

- All variables in a VAR system must be stationary
- For example, for the first order VAR model the eigenvalues of the coefficient matrix $\mathbf{A}$ must be less than one
- If they are not stationary then we should look for a cointegration relationship between variables
- If there exists a cointegration relationship then we should estimate an Error Correction Model (ECM)

25

Vector Autoregressive Models (VAR)

- In practice we first need to determine if the series under investigation are stationary or not
- If they are found to be $I(0)$ i.e., stationary then the estimation can be carried out using the formulation written above with OLS (conditional on the first p observations)
- The order of a VAR process, p can be chosen using either Likelihood Ratio test or using information criteria such as AIC or SIC

26

Testing for Granger-causality in a VAR(p)

- Suppose that a 2-variable VAR is specified as

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix}$$

- Here the lagged value of $\mathbf{y_2}$ plays no role in the determination of y_1
- Thus, $\mathbf{y_2}$ is said to not Granger-cause y_1
- This hypothesis can be tested using Likelihood Ratio test

$$LR = -2[l_r - l_u] \sim \chi^2(q)$$

where l_r is the value of log-likelihood from the restricted model and l_u is the value of log-likelihood from the unrestricted model and q is the number of restrictions.

27

Impulse Response Functions

- Consider the following 2-variable VAR(1) model:

$$\begin{aligned} y_{1t} &= a_{11}y_{1,t-1} + a_{12}y_{2,t-1} + \epsilon_{1t} \\ y_{2t} &= a_{21}y_{1,t-1} + a_{22}y_{2,t-1} + \epsilon_{2t} \end{aligned}$$

- A perturbation in ϵ_{1t} has an immediate and one-for-one effect in y_{1t}
- In period $t+1$, that perturbation in y_{1t} affects $y_{1,t+1}$ through the first equation and also affects $y_{2,t+1}$ through the second equation.
- These effects work through to period $t+2$ and so on
- Thus, a perturbation in one innovation in the VAR system sets up a chain reaction over time in all variables in the VAR.
- Impulse response functions calculate these chains reactions

28

Vector Error Correction Models (ECM)

- If the variables in a VAR system are integrated of order one or more individually then we cannot use the unrestricted VAR model discussed so far
- The presence of nonstationary variables raises the possibility of cointegrating relations
- Therefore we must determine the number of cointegrating relations among the variables
- Then we can estimate the VAR, incorporating the cointegrating relations from the previous step
- We can use Johansen's methodology to determine the cointegrating rank.

29

Multivariate Cointegration

- Recall, that cointegration refers to a linear combination of nonstationary variables.
- Cointegrating vector is not unique. If $(\beta_1, \beta_2 \ldots, \beta_k)$ is a cointegrating vector, then for any nonzero value of λ, $(\lambda\beta_1, \lambda\beta_2, \ldots, \lambda\beta_k)$ is also a cointegrating vector.
- Normalize the cointegrating vector using of the variables. e.g. to normalize with respect to y_{1t} select $\lambda = 1/\beta_1$
- If y_t has k nonstationary components there may be as many as $k-1$ linearly independent cointegrating vectors.
- The number of cointegrating vectors is called the cointegrating rank of y_t

30

Johansen's Test for Cointegration

- Consider the following VAR system

$$y_t = A_1 y_{t-1} + , \ldots, A_p y_{t-p} + B x_t + \epsilon_t$$

or

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + B x_t + \epsilon_t$$

where $\Pi = \sum_{i=1}^{p} A_i - I$, and $\Gamma_i = -\sum_{j=i+1}^{p} A_j$

- If the coefficient matrix Π has reduced rank $r < k$ then there exists $k \times r$ matrices α and β (each with rank r) such that $\Pi = \alpha\beta'$ and $\beta' y_t$ is stationary, $I(0)$.
- Johansen's procedure: estimate the full VAR(p) and test the reduced rank restriction (check the rank of Π).
- λ_{trace} and λ_{max} statistics.
- can be implemented in Eviews.

31

Johansen's Test for Cointegration

Testing for the number of cointegration relations

- *Step 1: Test* $H_0: r = 0$ *against* $H_1: r \geq 1$. First test the null hypothesis that there is no cointegration and that there are m stochastic trends. This corresponds to the hypothesis that $\lambda_1 = \cdots = \lambda_m = 0$, and the relevant test statistic is (7.40) with $r = 0$. If H_0 is not rejected, then there is no cointegration. If H_0 is rejected, continue with step 2.
- *Step 2: Test* $H_0: r = 1$ *against* $H_1: r \geq 2$. In a similar way as in step 1, apply the test (7.40) with $r = 1$. If H_0 is not rejected then there is a single cointegration relation and there are $(m - 1)$ common trends. If H_0 is rejected, continue with step 3.
- *Step 3: Iteratively test* $H_0: rank(\Pi) = r$ *against* $H_1: rank(\Pi) \geq r + 1$. Repeat the test (7.40) iteratively, increasing the value of r by one in each step. Continue until the first time that H_0 is not rejected. Then the number of cointegration relations is equal to r and the number of (common) trends is $(m - r)$.

32

Forecasting

- Once we've run a time-series regression we can use it for forecasting into the future
- Can calculate a point forecast and forecast interval in the same way we got a prediction and prediction interval with a cross-section
- Rather than use in-sample criteria like adjusted R^2, often want to use out-of-sample criteria to judge how good the forecast is

33

Out-of-Sample Criteria

- Idea is to note use all of the data in estimating the equation, but to save some for evaluating how well the model forecasts
- Let total number of observations be $n + m$ and use n of them for estimating the model
- Use the model to predict the next m observations, and calculate the difference between your prediction and the truth

34

Out-of-Sample Criteria (cont)

- Call this difference the forecast error, which is $\hat{e}_{n+h+1}$ for $h = 0, 1, \ldots, m$
- Calculate the root mean square error (RMSE)

35

Out-of-Sample Criteria (cont)

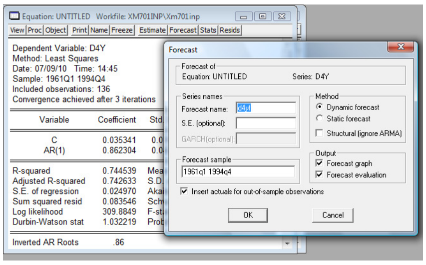
- Call this difference the forecast error, which is $\hat{e}_{n+h+1}$ for $h = 0, 1, \ldots, m$
- Calculate the root mean square error and see which model has the smallest, where

$$RMSE = \left(m^{-1} \sum_{h=0}^{m-1} \hat{e}_{n+h+1}^2 \right)^{1/2}$$

36

Forecasting in Eviews

- Estimate an appropriate model that passes diagnostic tests
- Open forecast button in equation view



Forecasting in Eviews

- Specify:
- Forecast name
- Forecasting method: dynamic or static
- Dynamic: calculates multi-step forecast starting with the first period in the forecast sample
- Static: calculates a sequence of one-step-ahead forecasts. Uses actual values for the lagged dependent variables
- Sample range: let n=n1+n2, use n1 observations for estimation and use the rest to form out-of-sample forecasts
- Read Eviews Manual Ch. 18