ARIMA MODELS

Auto-Regressive (AR) Model

- Auto-regressive(AR) and moving average(MA) models are popular models that are frequently used for forecasting
- AR and MA models are combined to create models such as auto-regressive moving average (ARMA) and auto-regressive integrated moving average(ARIMA) models.
- ARMA models are basically regression models
- Auto-regression simply means regression of a variable on itself measured at different time periods
- One of the fundamental assumptions of AR model is that the time series is assumed to be a stationary process
- When the time series data is not stationary, then we have to convert the non-stationary time-series data to stationary time-series data before applying AR models

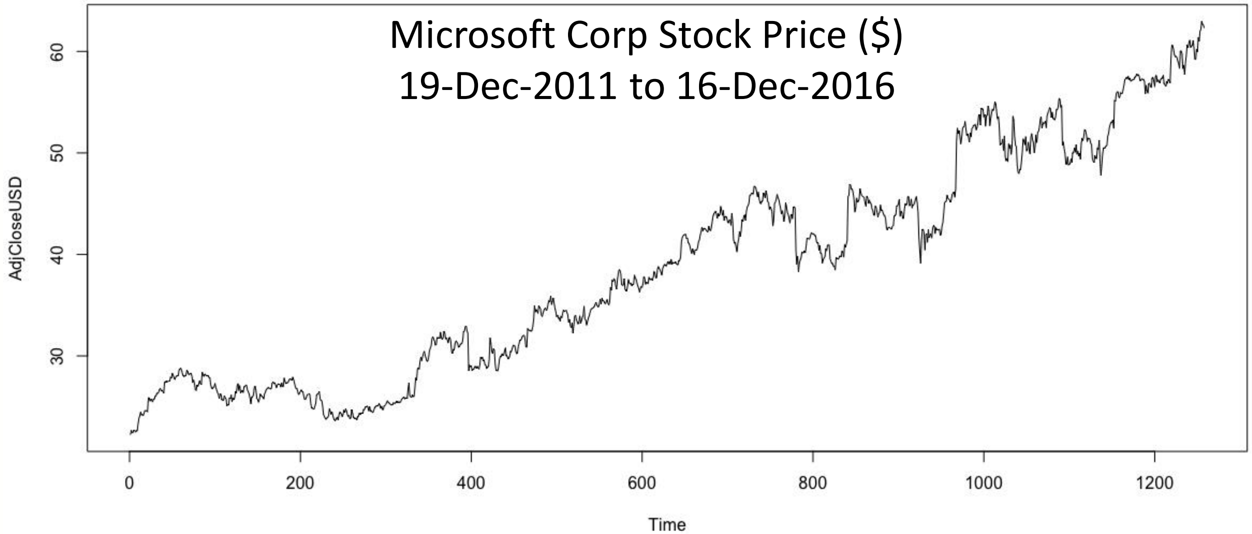
Stationary and Non-Stationary

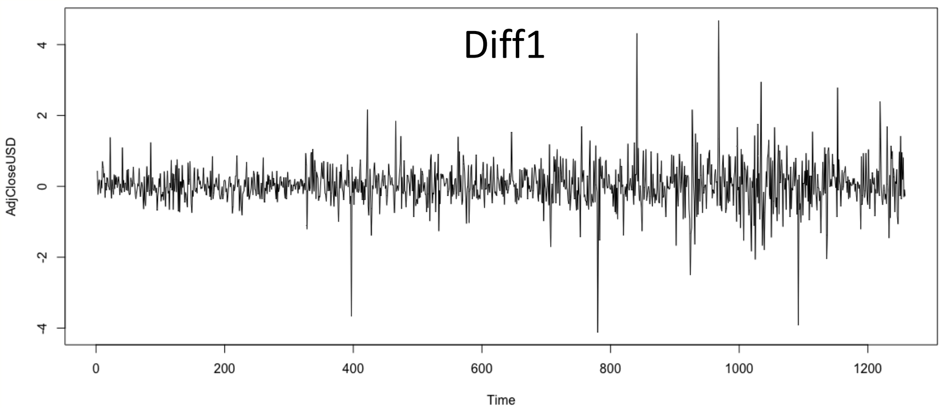
Stationary data has constant statistical properties – mean, variance, autocorrelation, etc. – over time

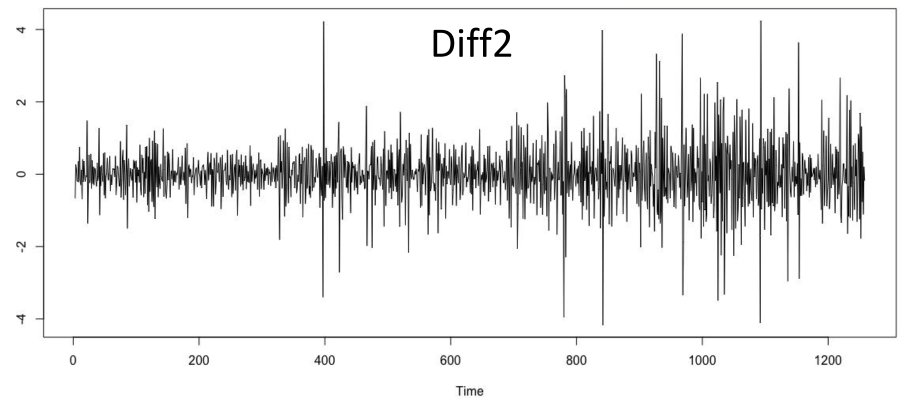
If the data is stationary, forecasting is easier!

Differencing to convert non-stationary to stationary

Converting non-stationary to stationary data







AutoRegressive (AR) Models

GREAT LAKES INSTITUTE OF MANAGEMENT, CHENNAI

$\hat{y}_t$ **depends only on its own past values** $y_{t-1}, y_{t-2}, y_{t-3}$, **etc. Thus,**

$\hat{y}_t$**=f(**$y_{t-1}, y_{t-2}, y_{t-3}, \ldots, \varepsilon_t$**)**

Autoregressive model where it depends on p of its past values ***(lagged values having significant relationship with the most recent value)*** **is an AR(p) model and is represented as:**

$$\hat{y}_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + e_t$$

AutoRegressive (AR) Models

- AR model with lag 1, AR(1) can be written as

$$y_{t+1} = \beta_0 + \beta_1 y_t + e_{t+1}$$

- Observe that, we need to have $|\beta_1| < 1$
- One of the important tasks in using auto-regressive model in forecasting is the model identification, which is, identifying the value of p(the number of lags)
- Standard approaches used for model identification: Auto-Correlation Function(ACF) and Partial Auto-Correlation Function(PACF)
- Auto-Correlation is the correlation between Y_t measured at different time periods
- Auto-correlation indicates the memory of a process, i.e., how far in time can it remember what has happened before.
- A plot of auto-correlation for different values of k is called auto-correlation function(ACF) or correlogram

AutoRegressive (AR) Models

- **Partial auto-correlation of lag k, ρ_{pk}, is the correlation between Y_t and Y_{t-k} when the influence of all intermediate values $(Y_{t-1}, Y_{t-2}, \cdots, Y_{t-k+1})$ is removed from both Y_t *and* Y_{t-k} . i.e., the additional prediction Y_{t-k} adds to AR(k-1) model.**
- **A plot of partial auto-correlation for different values of k is called partial auto-correlation function(PACF)**
- **Hypothesis tests can be carried out to check whether the auto-correlation and partial auto-correlation values are different from zero**
- **Null and alternative hypotheses:**
  - $H_0: \rho_k = 0$ and $H_A: \rho_k \neq 0$ where ρ_k is the auto-correlation of order k
  - $H_0: \rho_{pk} = 0$ and $H_A: \rho_{pk} \neq 0$ where ρ_{pk} is the partial auto-correlation of order k

AutoRegressive (AR) Models

- **The null hypothesis is rejected when $|\boldsymbol{\rho_k}| > \frac{1.96}{\sqrt{n}}$ and $|\boldsymbol{\rho_{pk}}| > \frac{1.96}{\sqrt{n}}$**
- **To select the appropriate p in the auto-regressive model, the following thumb rule may be used**
- **The number of lags is p when**
  - The partial auto-correlation, $|\boldsymbol{\rho_{pk}}| > \frac{1.96}{\sqrt{n}}$ for first p values (first p lags) and cuts off to zero
  - The auto-correlation function(ACF), ρ_k, decreases exponentially
- **Note that the model identification is an iterative process and may require additional inputs**
- **The model identification using ACF and PACF can not be taken as conclusive evidence for the number of lags in AR process**

AutoRegressive (AR) Models: Examples

- **AR(1):** $\boldsymbol{y_t = \beta_0 + \beta_1 y_{t-1} + \epsilon_t}$

- **AR(3):** $\boldsymbol{y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \epsilon_t}$

Moving Average (MA) Models

GREAT LAKES INSTITUTE OF MANAGEMENT, CHENNAI

A moving average model is one where $\hat{y}_t$ depends only on the random error term, i.e.,

$$\hat{y}_t = f(\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3}, \ldots..)$$

Moving average model where it depends on q of its past values (*residuals or lagged errors from earlier estimates*) is an MA(q) model and is represented as

$$\hat{y}_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots. + \theta_q e_{t-q}$$

Moving Average (MA) Models

- **The value of q (number of lags) in the moving average process can be identified using the following rule**
- **Auto-correlation value, $|\rho_p| > \frac{1.96}{\sqrt{n}}$ for first q values (first q lags) and cuts of to zero**
- **The partial auto-correlation function decreases exponentially**

Moving Average (MA) Models: Examples

- **MA(1):** $y_t = \theta_0 + \theta_1\epsilon_{t-1} + \epsilon_t$

- **MA(3):** $y_t = \theta_0 + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \theta_3\epsilon_{t-3} + \epsilon_t$

AutoRegressive Moving Average (ARMA) Models

ARMA(p,q)

Mix of both AR and MA

General form of such a time-series model

$$\hat{y}_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots + \theta_q e_{t-q}$$

ARMA Models: Examples

- **ARMA(1, 1):** $y_t = \beta_0 + \beta_1 y_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t$
- **ARMA(2,1):** $y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \theta_1 \epsilon_{t-1} + \epsilon_t$

Integration of order d

Non-stationary series can be made stationary series after differencing.

Integration of order d

A series that is **differentiated d times**

Denoted as **I(d)**

If the series is differentiated **once** then it is **I(1)**.

I(0) means series is **stationary** without **differencing**

ARIMA Models: Examples

- **ARIMA(1, 0, 1):** $y_t = \beta_0 + \beta_1 y_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t$

- **ARIMA(2, 1, 1):** $\Delta y_t = \beta_0 + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + \theta_1 \epsilon_{t-1} + \epsilon_t$

  where $\Delta \mathrm{y}_t = y_t - y_{t-1}$

Box–Jenkins Methodology

Model identification and model selection

Make sure variables are stationary. Difference as necessary to get a constant mean and transformations to get constant variance.

Check for seasonality: Decays and spikes at regular intervals in ACF plots.

Parameter estimation

Compute coefficients that best fit the selected model.

Model checking

Check if residuals are independent of each other and constant in mean and variance over time (white noise).

http://www.ncss.com/wp-content/themes/ncss/pdf/Procedures/NCSS/The_Box-Jenkins_Method.pdf

ARIMA(p,d,q) Model

p **is the number of autoregressive [AR(p)] terms (a linear regression of the current value of the series against one or more prior values of the series)**

Maximum lag beyond which PACF is 0

d **is the number of non-seasonal differences (order of the differencing) used to make the time series stationary [I(d)]**

q **is the order of the moving average [MA(q)] model**

Maximum lag beyond which the ACF is 0

Non-seasonal ARIMA models are denoted ARIMA(p,d,q)
Seasonal ARIMA (SARIMA) models are denoted ARIMA(p,d,q)(P,D,Q)$_m$, where m refers to the number of periods in each season and (P,D,Q) refer to the autoregressive, differencing and moving average terms of the seasonal part of the ARIMA model.

SARIMA Models: Examples

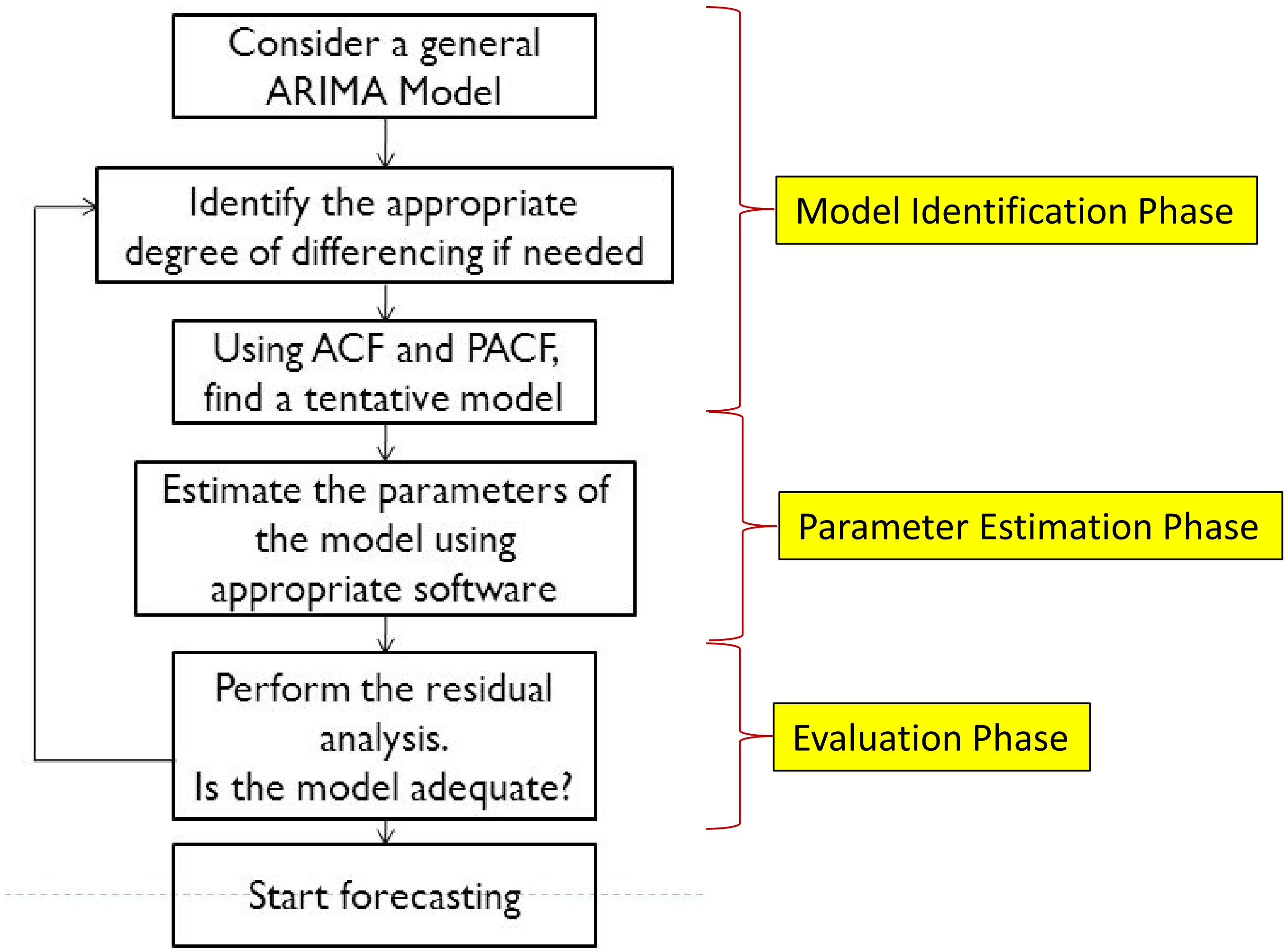
- **ARIMA(0, 0, 1)(0,0,1)_12:** $\boldsymbol{y_t = \beta_0 + \theta_1 \epsilon_{t-1} + \Phi_1\, \epsilon_{t-12} + \theta_1 \Phi_1\, \epsilon_{t-13} + \epsilon_t}$

- **ARIMA(1, 0, 0)(1, 0, 0)_12:** $\boldsymbol{y_t = \beta_0 + \beta_1 y_{t-1} + \Theta_1 y_{t-12} + \beta_1 \Theta_1 y_{t-13} + \epsilon_t}$

Selection of p and q

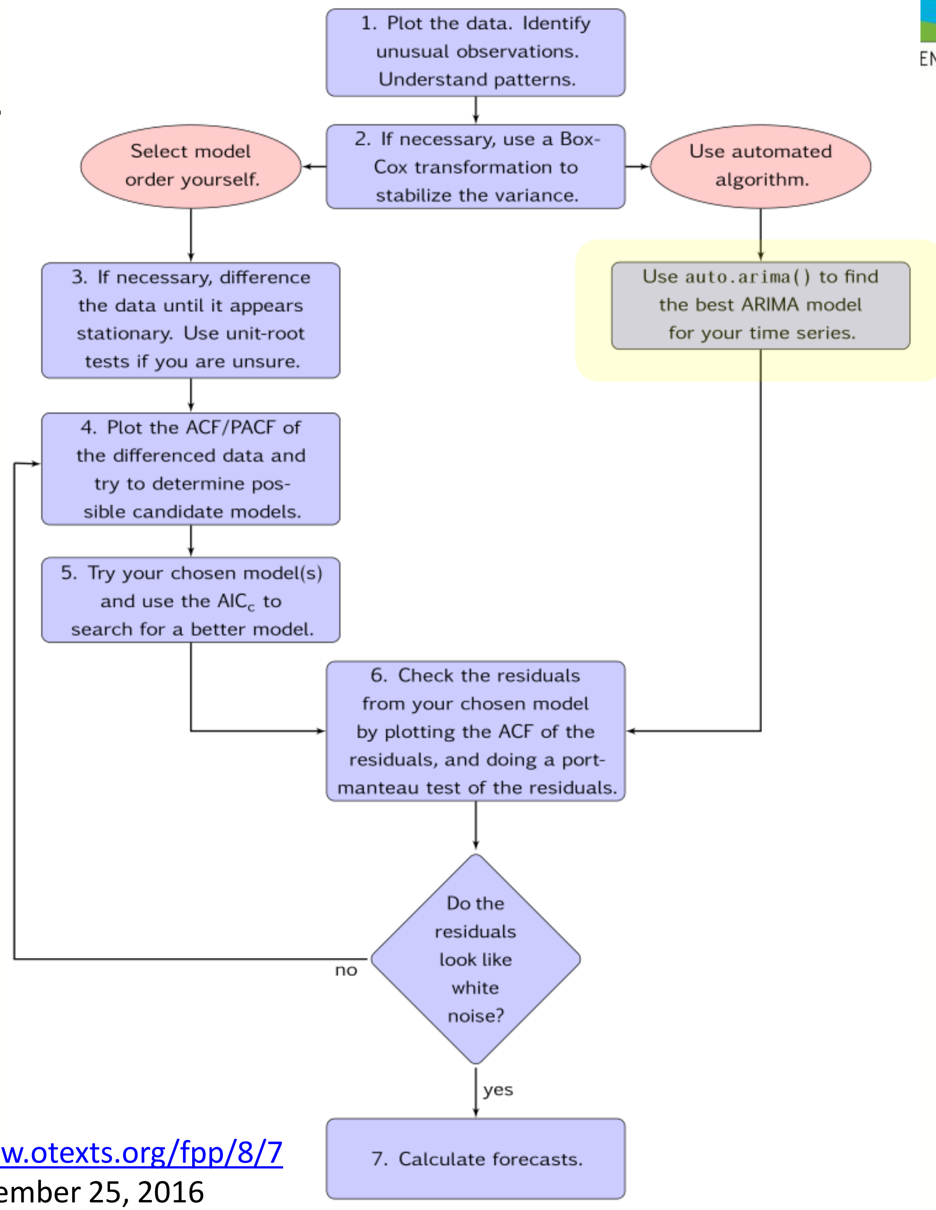
Process	MA(q)	AR(p)	ARMA(p,q)
Auto-correlation function (ACF)	Cuts of at lag q	Infinite. Tails off. Damped exponentials and/or cosine waves	Infinite. Tails off. Damped exponentials and/or after q-p
Partial auto-correlation function (PACF)	Infinite. Tails off. Damped exponentials and/or cosine waves	Cuts of at lag p	Infinite. Tails off. Damped exponentials and / or after p-q

Time Series Model Building Using ARIMA



Model Selection in Practice

There are techniques that automate model selection



Source: https://www.otexts.org/fpp/8/7
Last accessed: November 25, 2016

Model Selection in Practice

auto.arima() in R / Python use Hyndman-Khandakar algorithm.

1. **Find *d* for stationarity.**
2. ***p* and *q* are selected using a stepwise search to minimize AICc on the differenced data.**
   a. The best model among the following is selected as the current model: ARIMA(2,*d*,2), ARIMA(0,*d*,0), ARIMA(1,*d*,0), ARIMA(0,*d*,1).
   b. *p* and *q* are varied by ± 1.
   c. 2(b) is repeated till AICc doesn't reduce further.

Time Series Model Building Using ARIMA

A nice summary of rules for identifying ARIMA models

http://people.duke.edu/~rnau/arimrule.htm

Summary

- Forecasting is one of the important tasks carried out using analytics by many organizations since accurate forecasting is important for taking several decisions such as man-power planning, materials requirement planning, budgeting and supply chain related issues
- Forecasting is carried out on a time-series data in which the dependent variable Y_t is observed at different time periods t
- Several techniques such as moving average, exponential smoothing, and auto-regressive models are used for forecasting future value of Y_t
- The forecasting models are validated using accuracy measures such as RMSE, MAPE, AIC and BIC
- Simple techniques such as moving average and exponential smoothing may outperform complex regression based models in certain scenarios. Thus, it is important to develop forecasting models using several techniques before selecting the final model
- Regression model in the presence of independent variables may outperform other techniques
- AR models are regression based models in which dependent variable is Y_t and the independent variables are Y_{t-1}, Y_{t-2}, etc

Summary

- AR models can be used only when the data is stationary
- MA models are regression models in which the independent variables are past error values
- ARIMA has 3 components: AR with p lags, MA with q lags and integration which is differencing the original data to make it stationary
- One of the necessary conditions of acceptance of ARIMA model is that the residuals should follow white noise
- In ARIMA, the model identification, i.e., identifying the value of p in AR and q in MA, is achieved through ACF and PACF
- The stationarity of time-series data is usually checked using Dickey-Fuller and Augmented Dickey Fuller test
- The overall model accuracy of forecasting model is tested using Ljung-Box test