A Quick Overview of Time Series Analysis and Forecasting

Dr. Goutam Chakraborty

Lecture

Outline of Session

- Forecasting
  - Introduction of forecasting for decision making
  - Discuss different types of forecasting approaches
  - Discuss methodological issues in forecasting
  - Demonstrate building different time series models using a data set
  - Selecting the best forecasting model

Forecasting Introduction

- Why is it needed?
- What is a time series database?
- Some background on forecasting
- Limitations of classical univariate forecasting
- What is data mining for forecasting?
- A quick overview of forecasting process

What is a Forecast?

- Forecast – a statement about the value of a variable of interest at a *future time*.
  - We make forecasts about such things as sales (demand), weather, resource availability, …
  - Forecasts are an important element in making informed decisions
  - Planning starts with forecasting…

Why do we need forecasting?

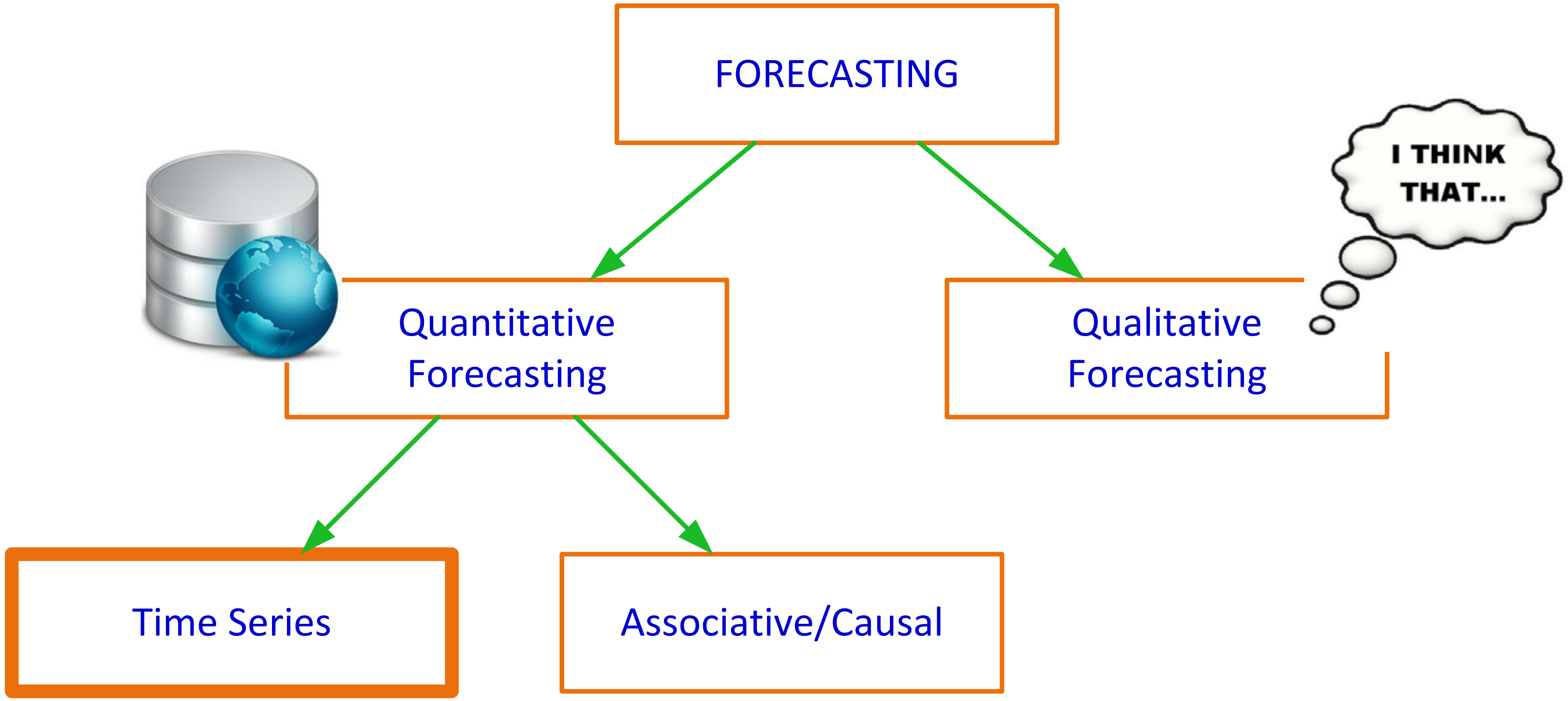
- Almost every aspects of business can leverage good forecasts for business gain. Examples include:
  - Marketing: new product forecasts
  - Sales: planning salesforce deployment and optimization
  - Operations: resource and asset planning
  - Supply chain: right product at right time at right place..
  - Finance: company’s financial health for reporting…
  - Strategy: strategic planning
  - Others

Time Series Data

- These are set of values for variables that are recorded for specific points in time.
  - Best practice : Have the values for variables recorded in time intervals which are equally spaced such as weekly, monthly, quarterly etc.
- For a company, the time series data are generated from internal operations (sales, production, ..)
- Some examples of publicly available (free) data include for research:
  - All kinds of economic time series data from US government (https://api.census.gov/data/timeseries.html)

	A	B	C
1	**Year**	**Month**	**Revenue**
2	2013	October	$ 97,766
3	2013	November	$ 98,685
4	2013	December	$ 99,677
5	2014	January	$ 101,943
6	2014	February	$ 100,499
7	2014	March	$ 111,142
8	2014	April	$ 98,389
9	2014	May	$ 113,633
10	2014	June	$ 100,946
11	2014	July	$ 102,125
12	2014	August	$ 109,513
13	2014	September	$ 99,813
14	2014	October	$ 108,047
15	2014	November	$ 107,090
16	2014	December	$ 109,764

A Simple Forecasting Taxonomy



Forecasting Approaches

- Qualitative/Judgmental Forecasting (opinion based)
  - These use subjective inputs such as opinions from managers, executives, and domain experts
    - Delphi method (consensus forecast)
      - Surveys/questionnaires are filled-out individually; then consolidated
- Quantitative Forecasting (data driven)
  - Time-series forecast (univariate) - projection of historical data
  - Associative models (multiple time series) - development of associative methods that attempt to use causal variables to make a forecast

Two Approaches of Quantitative Time Series Forecasting

- Univariate Time Series Forecast
  - Only one target (Y) variable over time is modeled
  - Model the **historical pattern in Y over time and extend it to future time**
  - Generally works well in *short term* forecasts
  - Examples of models : Naïve, moving average, smoothing, ARMA,…
  - Do not provide "drivers" of forecasts

- Multiple Time Series Forecast
  - Y (target) variable over time is modeled along with many X (independent) variables over time
  - Model the **relationship between X's and Y over time and extend it to future time**
  - Generally works better for *medium/long* forecasts
  - Examples of models: VARMAX, UCM, and others
  - Provide insights as to what drives Y

Features Common to All Forecasts

1. Forecasts are not perfect!!
2. Techniques assume some underlying patterns that *existed in the past will persist* into the future
3. Forecasts for groups of items are generally more accurate than those for individual items
   - Aggregate forecasts are usually more accurate
4. Forecast accuracy decreases as the forecasting horizon increases
   - It is easier to forecast the near future

Steps in the Forecasting Process

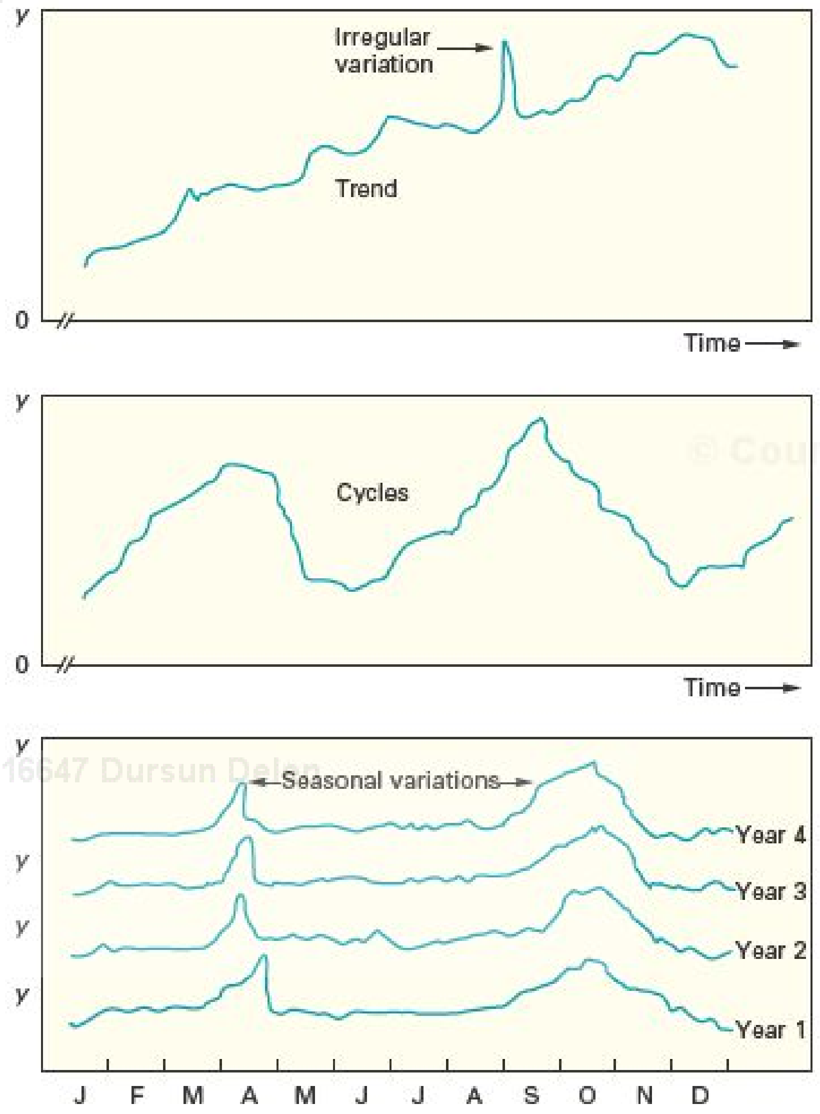
1. Determine the purpose of the forecast
2. Establish a time horizon
3. Select a forecasting technique
4. Obtain, clean, and analyze appropriate data
5. Make the forecast
6. Monitor the forecast (for accuracy and reliability…)

Terms and Metrics Used in Time-Series

- Terms:
  - *Equally spaced* data over time
  - Components of a time series data
    - Trend, seasonality, cyclicality, random, irregular, error..
- Forecast error indices (metrics)
    - MAD (or, MAE), MSE, RMSE, MAPE, …

Time-Series Components

- Trend
  - A long-term up or down movement in data
- Seasonality
  - Regular variation that repeats itself at the same time
- Cycles
  - Wavelike variations lasting *long time*
- Irregular variation
  - Due to unusual circumstances that do not repeat regularly
- Random variation
  - What’s left after the four mentioned above



Forecast Accuracy and Control

- Forecasters want to minimize forecast errors
  - It is nearly impossible to correctly forecast real-world variable values on a regular basis
  - So, it is important to provide an indication of the extent to which the forecast might deviate from the value of the variable that actually occurs
- Forecast accuracy is an important criterion for selection of a forecasting technique
- Forecast errors should be monitored and (if unacceptable) underlying models should be changed (adopted or recreated)
  - **Error = Actual – Forecast**

Forecast Accuracy Metrics

$$\text{MAD or MAE} = \frac{\sum \left|\text{Actual}_t - \text{Forecast}_t\right|}{n}$$

Mean Absolute Deviation (MAD) or, Mean Absolute Error (MAE) weights all errors evenly

$$\text{MSE} = \frac{\sum \left(\text{Actual}_t - \text{Forecast}_t\right)^2}{n-1}$$

Mean Squared Error (MSE) weights errors according to their squared values, RMSE is square root of MSE

$$\text{MAPE} = \frac{\sum \frac{\left|\text{Actual}_t - \text{Forecast}_t\right|}{\text{Actual}_t} \times 100}{n}$$

Mean Absolute Percent Error (MAPE) weights errors according to relative error

An Example of Forecast Error Calculation

Period	Actual (A)	Forecast (F)	(A-F) Error	|Error|	Error^2	[|Error|/Actual]x100
1	107	110	-3	3	9	2.80%
2	125	121	4	4	16	3.20%
3	115	112	3	3	9	2.61%
4	118	120	-2	2	4	1.69%
5	108	109	1	1	1	0.93%
			Sum	13	39	11.23%
				n = 5	*n-1* = 4	*n* = 5
				MAD	**MSE**	**MAPE**
				= 2.6	= 9.75	= 2.25%

RMSE = 3.12

Basic Models in Univariate Time Series Forecasts

- Naïve forecasts
- Averaging forecasts
  - Simple average
  - Moving average
  - Weighted moving average
  - Exponential smoothing
- Adjusting for:
  - Trend
  - Seasonality

Naïve Forecast

- Naïve Forecast
  - Uses a single previous value of a time series as the basis for a forecast
    - The forecast for a time period is equal to the previous time period's value
  - Can be used when
    - The time series is fairly stable
    - No time or expertise to build other models

Averaging

- These techniques work best when a series tends to vary about an average
  - Averaging techniques smooth out variations in the data
  - They can handle *step or gradual* changes in the level of a series
  - Techniques include
    - Simple average
    - Moving average
    - Weighted moving average
    - Exponential smoothing

Moving Average

- Technique that averages a number of the most recent actual values in generating a forecast

$$F_t = \mathrm{MA}_t = \frac{\sum_{i=1}^{n} A_{t-i}}{n}$$

where

F_t = Forecast for time period t

MA_t = n period moving average

A_{t-1} = Actual value in period $t-1$

n = Number of periods in the moving average

Moving Average – Example

Compute a three-period moving average forecast given demand for shopping carts for the last five periods.

Period	Demand	
1	42	
2	40	
3	43	
4	40	the 3 most recent demands
5	41	

$$F_6 = \frac{43 + 40 + 41}{3} = 41.33$$

If actual demand in period 6 turns out to be 38, the moving average forecast for period 7 would be

$$F_7 = \frac{40 + 41 + 38}{3} = 39.67$$

Moving Average

- As new data become available, the forecast is updated by adding the newest value and dropping the oldest and then re-computing the average
- The number of data points included in the average determines the model's sensitivity to changes in data pattern
  - Fewer data points used– model is more responsive
  - More data points used– model is less responsive

Weighted Moving Average

- The most recent values in a time series are given more weight in computing a forecast
  - The choice of weights, w, is somewhat arbitrary (often using expert judgement) and involves some trial and error

$$F_t = w_n A_{t-n} + w_{n-1} A_{t-(n-1)} + \ldots + w_1 A_{t-1}$$

where

w_t = weight for period t, w_{t-1} = weight for period $t-1$, etc.

A_t = the actual value for period t, A_{t-1} = the actual value for period $t-1$, etc.

Weighted Moving Average - Example

Compute a weighted average forecast using a weight of .40 for the most recent period, .30 for the next most recent, .20 for the next, and .10 for the next.

If the actual demand for period 6 is 39, forecast demand for period 7 using the same weights as in part *a*.

Period	Demand
1	42
2	40
3	43
4	40
5	41

$$F_6 = .10(40) + .20(43) + .30(40) + .40(41) = 41.0$$

$$F_7 = .10(43) + .20(40) + .30(41) + .40(39) = 40.2$$

Exponential Smoothing

- A weighted averaging method that is based on the previous forecast *plus a percentage of the forecast error*

$$F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1})$$

where

F_t = Forecast for period t

F_{t-1} = Forecast for the previous period

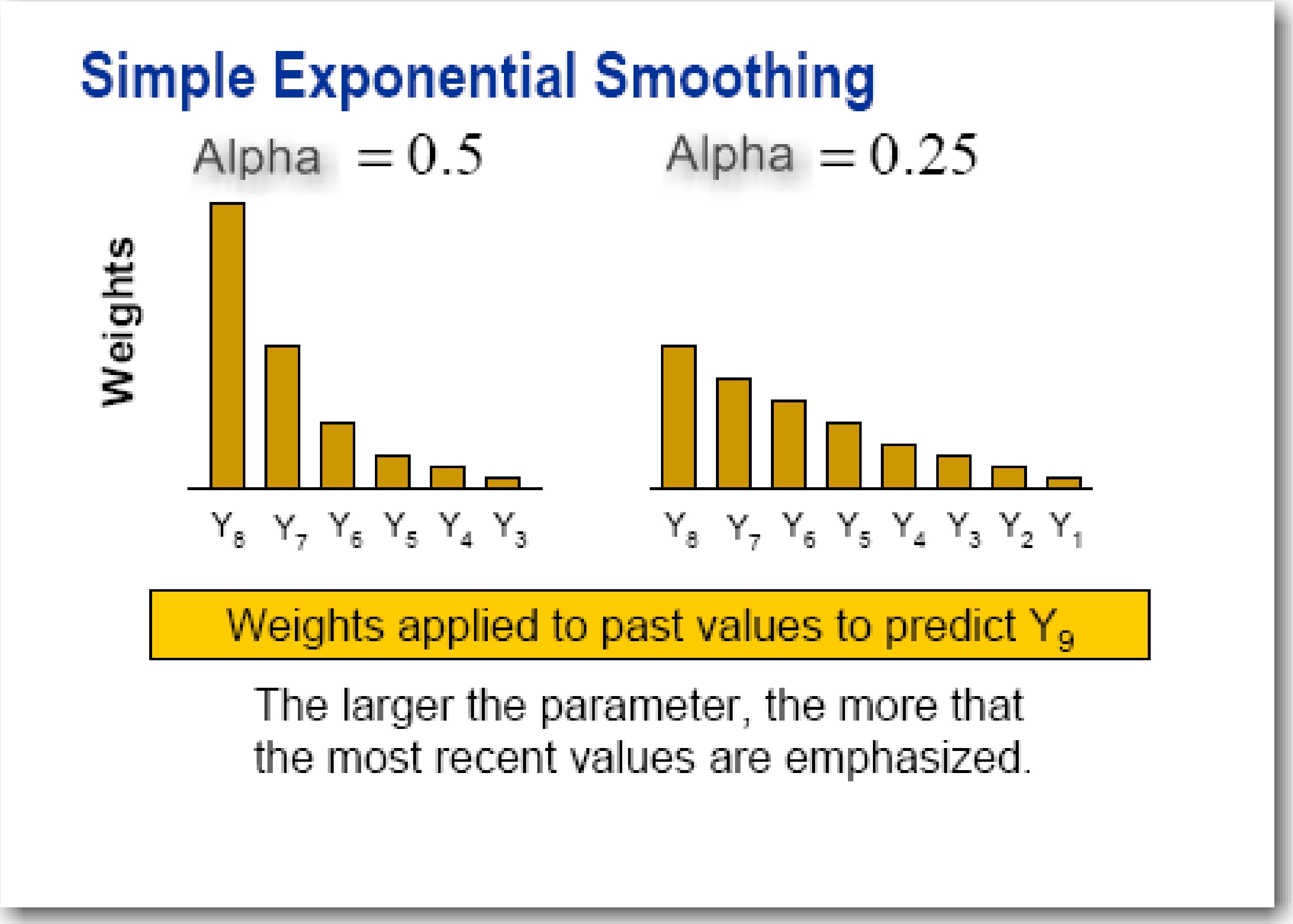
α = Smoothing constant

A_{t-1} = Actual demand or sales from the previous period

Exponential Smoothing (contd.)

- $F_{t+1} = \alpha A_t + (1-\alpha)F_t$
- Similarly, $F_t = \alpha A_{t-1} + (1-\alpha)F_{t-1}$ substituting this value in the above equation, we get:
- $F_{t+1} = \alpha A_t + (1-\alpha)*\{\alpha A_{t-1} + (1-\alpha)F_{t-1}\}$
- Or, $F_{t+1} = \alpha A_t + (1-\alpha)\alpha A_{t-1} + (1-\alpha)^2 F_{t-1}$
- Again, $F_{t-1} = \alpha A_{t-2} + (1-\alpha)F_{t-2}$ Substituting this value in the above equation, we get:
- $F_{t+1} = \alpha A_t + (1-\alpha)\alpha A_{t-1} + (1-\alpha)^2 *\{\alpha A_{t-2} + (1-\alpha)F_{t-2}\}$
- Or, $F_{t+1} = \alpha A_t + (1-\alpha)\alpha A_{t-1} + (1-\alpha)^2 *\alpha A_{t-2} + (1-\alpha)^3 F_{t-2}$
- Continuing in this way, we can show that F_{t+1} is a weighted average of **ALL** past values:
- $F_{t+1} = \alpha A_t + (1-\alpha)\alpha A_{t-1} + (1-\alpha)^2 *\alpha A_{t-2} + (1-\alpha)^3 \alpha A_{t-3} + \ldots.$

SES (contd.)



Simple Exponential Smoothing
Alpha = 0.5
Alpha = 0.25
Weights
Y8 Y7 Y6 Y5 Y4 Y3
Y8 Y7 Y6 Y5 Y4 Y3 Y2 Y1
Weights applied to past values to predict Y9
The larger the parameter, the more that the most recent values are emphasized.

How do we choose α ?

- The value of α is between 0 and 1
- A value closer to 0 indicates series is very random, a value closer to 1 indicates forecast depends heavily on changes in recent values
- In practice, α values of 0.05-0.40 works well for most simple smoothing models
- Choose α such that the model minimizes some criterion such as the RMSE, MAPE
  - We will rely on the software to do it for us!
- Advantage of SES - requires few data points and simple to implement
- Disadvantage - forecast lags original and has **no** ability to model trend/seasonality

Techniques for Trend

- Linear trend equation
  - Similar to simple linear regression
- Non-linear trends
  - Parabolic trend equation
  - Exponential trend equation
  - Growth curve trend equation

Linear Trend

- A simple data plot can reveal the existence and nature of a trend
- Linear trend equation given below
- Slope and intercept may be estimated from historical data

$$F_t = a + bt$$

where

F_t = Forecast for period t

a = Value of F_t at $t = 0$

b = Slope of the line

t = Specified number of time periods from $t = 0$

Holt's Exponential Smoothing for Trend

- The model will have two parameters .
  - α and β are smoothing constants
  - α same as in simple exponential smoothing model (averaging past observations)
  - β parameter captures the trend component
- Choose α, β such that the model minimizes some criterion such as the RMSE, MAPE
  - Rely on the software to do it for us!

Techniques for Seasonality

- Seasonality is expressed in terms of the amount that actual values deviate from the average value of a series
- Models of seasonality
  - **Additive**
    - Seasonality is expressed as a quantity that gets added or subtracted from the time-series average in order to incorporate seasonality
  - **Multiplicative**
    - Seasonality is expressed as a percentage of the average (or trend) amount which is then used to multiply the value of a series in order to incorporate seasonality

Winter's Smoothing Model

- More complicated (3 parameters) model
  - α β γ are smoothing constants for data stationarity, trend and seasonality
- But, it can handle both **seasonality** and **linear trend** in the data
- Need more data points to get estimates for parameters
- Software will choose α β γ so as to minimize error indices such as RMSE, MAPE etc.

Techniques for Cycles

- Cycles are similar to seasonal variations but are of longer duration
- They are generally ignored in univariate TS forecasting models
- If needed, *the explanatory (multivariate) approach* is often used
  - Search for another variable that relates to, and may be leads, the variable of interest
    - Housing starts precede demand for products and services directly related to construction of new homes
    - If a high correlation can be established with a leading variable, it can develop an equation that describes the relationship, enabling forecasts to be made