MSDS 413

Time Series Analysis & Forecasting

Module 1: Exploratory Data Analysis (EDA)

Canvas - TS1 - Gaussian-Based Time Series Analysis

1. Time Series Definition

1. Ordered sequence indexed by time
2. Constant intervals between all time periods
3. Variable is stochastic

2. Gauss-Markov Theory (and Assumptions)

OLS estimation is the BLUE of the general model $Y = Bx + \epsilon$ if:

1. Linearity of the parameters being estimated
2. Random sampling
3. Non-Collinearity of independent variables
4. Exogeneity: independent variables are uncorrelated with error term
5. Homoskedastic errors

OLS = BLUE when $\hat{\varepsilon}_i \overset{iid}{\sim} N(0, \sigma_i^2)$

3. Hypothesis Testing for Normality

H_{10}: Homoskedasticity (constant variance, σ_i^2)

H_{20}: ε_i independent of ε_j for all $i \neq j$

H_{30}: Identically normal, $\epsilon_i \overset{iid}{\sim} N(0, \sigma_i^2)$

4. Identifying a Time Series

1. Data must be a time-ordered sequence
2. (For this class) Data must have constant time intervals between periods
3. The variable must be stochastic

5. Tests for Gaussian Data

1. Skew
   H_{10}: skew = 0
2. Kurtosis
   H_{20}: excess kurtosis = 0

- Histogram
- QQ Plot

6. Obtaining a Linear Additive Model

Skew in a time series likely indicates increasing or decreasing variance,

Transformation via natural logarithm can stabilize variance by compressing skewed tail, resulting in a more normal probability distribution function.

Natural log transformation also makes a model additive, rather than multiplicative

Example: $X_t = c\exp(\Phi_1 X_{t-1})\exp(\Phi_2 X_{t-2})$ ← a multiplicable model

$$\log(X_t) = \log(c\exp(\Phi_1 X_{t-1})\exp(\Phi_2 X_{t-2}))$$
$$= \log(c) + \log(\exp(\Phi_1 X_{t-1})\exp(\Phi_2 X_{t-2}))$$
$$= \log(c) + (\Phi_1 X_{t-1} + \Phi_2 X_{t-2}) \leftarrow \text{an additive model}$$

Log transformation expresses change in percent, rather than normal units. To obtain a prediction in normal units, the anti-log must be taken.

7. Decomposition of a Time Series

Time Series data can be composed of up to three parts:

1. Trend
2. Seasonality
3. Stochasticity

~

Forecasting - Hyndman and Athanasopoulos

Chapter 6 - Time Series Decomposition

Section 6.6 - Time Series Components

Time series decomposition can be done using many kinds of models.

Two distinct types of models are additive models and multiplicative models.

Each describe how the season, trend/cycle, and stochastic components are combined.

Additive: $Y_t = T_t + S_t + R_t$

Multiplicative: $Y_t = T_t \times S_t \times R_t$

- Additive models are appropriate when the magnitude of the seasonal component stays constant. If it varies with the level of the time-series, use a multiplicative model.

TS1 Introduction

1. Exploratory Data Analysis

1.1 EDA Techniques

- Timeplot
- Summary Statistics
- Histogram
- QQ Plot
- Skew
- Kurtosis
- Constant variance
- Augmented Dickey-Fuller Test

1.2 Probability Distributions

A time series' probability distribution determines which kind of forecasting model is most appropriate.

Normal - tractable statistical qualities, characteristics rare, no lower bound

Lognormal - transformation of observations forces normality, has a lower bound

Stable -

Scale Mixture -

Discrete - Binomial (two-state) and Multi-nomial (multi-state) are often used.

1.3 Using EDA

Log transformation results in an additive model. To test if this is appropriate, verify that the data has additive symmetry. Mean = 0.

2. Time Series Decomposition: Trend and Seasonal Components

A common method for adjusting time series for trend is to difference the data. This can remove a linear trend. Differencing twice can remove a non-linear trend.

Equation of first differencing: $Y_t = \nabla X_t = X_t - X_{t-1}$, where ∇ is the difference operator.

A common method for adjusting for seasonality is differencing at the seasonality's frequency.

Equation: $Y_t = \nabla^4 X_t = X_t - X_{t-4}$ for quarterly seasonality.

Lecture Notes

- Include plots in report and include description of how to read it
- Use LaTex for writing equations when requested
- When making and deciding between multiple models, only the final model needs complete diagnostics, but all need enough to make comparisons valid
- One third of a page of text, include diagrams of forecasts
- Time must be a relevant component. Sounds obvious, but doesn't.
- Time Series - observations of a random variable through time and through a probability space
- Innovation - observations of residuals
- Residuals: $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$, $Cov(\epsilon_i, \epsilon_j) = 0$, $i \neq j$ → OLS = BLUE!
  iid - identically individually distributed: a collection of random variables with identical statistical distributions
  - For time-series, aim for identically distributed, not independent but are statistically independent
- F-statistic: indication of model predictive ability relative to mean model, which just takes and predicts the mean of the variable of interest
- Histogram: geom_histogram(fill=, color=)
- Normality Tests: ① Shapiro ② Lilliefors ③ Anderson-Darling } null hypothesis: normality
- Sample size and normality: central limit theorem?
- Confidence interval with zero: value can be considered 'statistically' zero, indistinguishable statistically
- Constant Variance Test: levene test (categorical variables)
- Skew: if mean = median, no skew, does median fall within mean CI?
- Stochasticity: variance CI, doesn't include 0
- Obtaining a Linear Additive Model - log transformation
- STL Decomposition: patterns in residuals indicate insufficient model

MSDS 413

Time Series Analysis & Forecasting

Module 2: Introduction to Gaussian Based Models

Canvas - TS2 - Gaussian-Based Time Series Models

Accounting for a Linear Trend in Time Series

Testing for a linear trend in time series data can be done by testing parameter values of an SLR of the time series. $X_t = B_0 + B_1 t + \varepsilon$ implies that x increases by $B_0 + B_1$ for each successive period. If $B_1 = 0$, then there is no linear trend. This assumes mean 0, so $B_0 = 0$.

To correct for a linear trend, take the first difference of the time-series.

If X_t is a time-series of n+1 observations, and $x_t = m_t + y_t$ where m is the slope of the trend line and y is the stochastic component, then

$d_t = X_t - X_{t-1}$

if $x_t = m_t + y_t$, then $x_t - x_{t-1} = (m_t + y_t) - (m_{t-1} + y_{t-1})$

$= y_t - y_{t-1}$

Identifying Auto-Correlation and Stationarity

ADF Test (Augmented Dickey-Fuller) - test for random walk stationarity. If unit roots are present, the observations are interdependent, and the data is not stationary.

KPSS Test - is a test for linear-trend stationarity. If no unit roots are present, differencing is required to account for random-walk drift.

Reading ACF and PACF plots:
- Quick return to zero implies stationarity.
- line outside of CI represent correlation.

White Noise

A time series of mean zero and finite, constant variance is called white noise.

$Z_t = \text{iid } WN(0, \sigma^2)$ where $\sigma^2 < \infty$ (also indicates observations are independent and identically distributed.)

McLeod Li Test - tests for the null hypothesis of homoskedasticity, constant variance is assumed if all test statistics are above 0.05.

Auto-Regressive (AR) Models

An AR(p) model has the general form:

$Z_t = X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p}$

where X is a mean zero, stationary time-series

$Z \sim WN(0, \sigma^2)$

Lecture Notes (43:56)

Review

- Additive models: if data is right skewed, model is likely multiplicative
  removed skew means model is additive (symmetrical?)
  equivalent percentiles on either side of the mean have the same magnitude.

Trends

- Auto-correlation cannot properly be assessed with a linear trend in the data, must first remove linear trend. Differencing!
- Test linear trend by assessing B_1 in SLR as described above
- Differenced mean zero indicates linear trend is gone

Stationarity

- Gauss-Markov Assumptions for regression require independance of residuals, stationary requires constant covariance

Auto-Regressive Models

$\Phi(B) X_t = Z_t, \quad Z_t \sim WN(0, \sigma^2)$

General Form of AR(p): $X_t - \phi_1 X_{t-1} - \dots - \phi_p X_{t-p} = Z_t$

observation - (coefficients · observations) = white noise residuals

- For AR models, use the plot of the PACF, p is the number of consecutive significant lags starting at 0
- Information Criterion are valid model comparison measures when using identical kinds of models

TS2 Introduction - Linear Trends, Stationarity, and Auto-Regressive Models

2. Stationarity and Invertability Conditions for a Linear Process

2.2 Stationarity

Stationarity - Fixed (no trend) mean and constant variance.

Weak Stationarity - Fixed mean, finite variance, and covariance only dependant only on lag

Strict Stationarity - see definition of stationarity above.
- When a time-series follows a normal distribution, it is strictly stationary.
- Why can strict stationary time series have dependant observations?

3. Auto-Regressive Models (AR)

3.3 AutoCorrelation Functions and Partial AutoCorrelation Functions

Autocorrelation measures the dependency of preceding events on successive events. Assessment of autocorrelation attempts to find a linear combination of realizations that can predict a variable within error. The coefficients represent to correlation between lags. The sign of the coefficient indicates which side of the mean the observations are relative to that lag. ACF can be indicative of stationarity. PACF represents the correlation between lags after adjusting for the observations in between. Spikes above critical line indicates the order of a fitting AR model.

4. Unit Root Non-Stationarity

Time series is stationary if all unit roots' absolute values are greater than 1.

If there are unit roots with imaginary numbers, the data may have cycles.

- Detection of unit roots in residuals indicates the residuals are not stationary

6. Model Comparison

6.1 Akaike Information Criterion

Lower AIC = better model

6.3 Out-of-Sample Comparison

Mean Squared Forecasting Error (MSFE) is a measurement used in model comparison for out-of-sample testing. Lower MFSE = better

Lecture Code Section

- Seasonality in residuals indicates that decomposition didn't extract all of the seasonality
- Regression for slope: `lm(Value ~ Date, data = y)`, readout has coef t-test
- Find frequency function finds periodicity (peak to peak in ACF that shows sin)
- Breusch-Pagan Test: Constant variance
- McLeod-Li Test: Constant variance, use `which()` to return int of sig lags
- Law of Large Numbers: as sample size increases, sample mean approaches population mean
- Central Limit Theorem: 'random sample estimates of population parameters are normally distributed even if the population distribution ~~isn't~~ when n > 50
- Cycles: results are period lengths, drop 'business cycle' terminology, cycles

MSDS 413

Time-Series Analysis & Forecasting

Module 3: ARMA Models

Lecture Notes

Recap

- Median falling within mean CI range is indicative of symmetry
- Models we have discussed so far are mean models, finding expected value

Moving Average Models

General Form: $X_t = Z_t - \theta_1 Z_{t-1} - \dots - \theta_q Z_{t-q}$

- The coefficients in a moving average model are applied to innovations of a white noise process
- Use ACF to find order of a moving average model
- AR models use realizations of a process, MA models use the difference of these realizations from the mean
- AR → PACF → order = significant lags
- MA → ACF → order = significant lags - 1 (starts at zero)
- ARMA assumes no linear trend (hence no differencing)
- ACF values indicate the correlation of all previous lags (t, t_{-1}, t_{-2}, etc...) joint distribution
- PACF values indicate the correlation of only one lag (t, $t-s$) conditional distribution
- Identify period of seasons via patterns in ACF plot
- Autocorrelation in the residuals indicates that the model didn't capture all autocorrelation

Module 5: Nonstationary Time Series & Modelling Covariates

Lecture Notes

Nonstationary Time Series

- Nonstationary indicates trend and/or seasonality
- Differencing can remove trend → integrated time series
- Order of trend pattern indicates order of differencing
- ARIMA object in R has lambda, 0 = log transformation, 0.5 = square root
  - Lambda = box-cox transformation

Outliers

- In the event of outliers, create a vector that is a dummy variable indicating the presence of outlier
- ARIMAX model : ARMA model with indicator vector, x → exogenous

R Code Walkthrough

- Create regression of value ~ time
- Use polynomial of order of trend
- Check residuals as time series, trend removed?
- ACF/PACF analysis of residuals indicates model order
- xreg component in ARIMA

Module 7: Multivariate Time Series Models

Recap

- Assessing auto-correlation and cross-correlation between time series
- Time-series represented as vectors can be combined to analyze time series as matrices
- Additional tests to assess EDA/correlation structure (check EDA list)
- Multivariate gaussian residual models:
  1) Vector Autoregressive Models (VAR)
  2) Vector ARMA models
  3) Cointegration

Vector Auto-Regressive (VAR models)

- General Form: $y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + \underbrace{DC_t}_{\text{multi-variate x-reg}} + u_t$
- Ideal is to have stationarity in both auto and cross correlations
- If unit roots are present in the time series, integration of at least order one
- If cross-correlation exists over time, it is possible that one is a leading indicator of the other

Module 4 - ARMA Models with Seasonality

Lecture Notes

Recap

- KPSS test looks for random walk
  - ↳ Random Walk: a random process describing a path along a mathematical space
- ADF test looks for drift
  - ↳ Stochastic Drift: change in the expected value of a stochastic process
- Cycling creates unit roots, unit roots → not stationary
- Levine's Test - constant variance of categorical variables
- Autocorrelation in residuals → autocorrelation that model didn't capture
- ARMA Models: Causal - infinite series of white noise innovations
  Invertible - infinite series of time series observations

SARMA Models

General Form: SARMA(Q,P)$_m$

where m = frequency of seasonal component

$$X_t = \mu + \sum_{i=1}^{p} \phi_i X_{t-i} + Z_t + \sum_{i=1}^{q} \theta Z_{t-i} + \sum_{i=1}^{P} \Phi X_{t-i} + \sum_{i=1}^{Q} \Theta Z_{t-i}$$

- Assume linear trend stationarity
- Find m, seasonal frequency, by using the find frequency function or peak-peak in ACF
- Auto.arima (seasonal = T) to find suggested orders
- Find P and Q :
  - Find frequency of time series
  - ACF and PACF of diff(data, period)

ARMA x SARMA model

- Orders p,q are identified in the same method that we have been using
- Orders P,Q are identified by using ACF and PACF of data differenced at frequency
- Differencing in ARMA component does not need to be repeated in SARMA component

Module 6 : Volatility Models

Lecture Notes

Volatility

- To understand volatility, look at the residuals of expected value (mean) models
- Expected value models assume constant variance
- Time series with changing variance suggest residuals have time-conditional variance
- Auto-Regressive Conditional Heteroskedasticity
- ARCH modeling is similar to AR modeling in the sense of time dependency
- GARCH ≅ ARMA , useful for volatility clusters
- IGARCH ≅ ARIMA, I → integrated → differencing
- ARMA + GARCH
- TGARCH, GJR-GARCH, EGARCH model asymmetric data (skew)
- APARCH ? -

Example

- ARCH test on ARIMA residuals
  - lags in arch test ≅ p? pace of mean reversion
- If alpha+beta = 1, IGARCH, $\alpha + \beta < 1$, stable
- KPSS Unit-Root Drift tests for random-walk
- ADF tests for drift
- Unit root in residuals → IGARCH

- Model diagnostics of ARCH models
  - ACF/PACF → auto-correlation
  - Skew/Kurtosis → normality
  - KPSS/ADF → stationarity
  - t-test → mean zero
  - Box-Ljung → auto-correlation (use residuals^2 to better model variance)

- GARCH model for skewed data → student's t-distribution (cond.dist = 'sstd')
- GARCHM : the mean is a function of variance (allowmean = T)
- TGARCH : stepwise shift in the variance
- $\alpha + \beta$ = persistance

MSDS - 413

Time-Series Analysis & Forecasting

Module 7: Multivariate Time Series Models

Recap

- Assessing auto-correlation and cross-correlation between time series
- Time-series represented as vectors can be combined to analyze time series as matrices
- Additional tests to assess EDA/correlation structure (check EDA list)
- Multivariate gaussian residual models:
  1) Vector Autoregressive Models (VAR)
  2) Vector ARMA models
  3) Cointegration

Vector Auto-Regressive (VAR models)

- General Form: $y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + \underbrace{DC_t}_{\text{multi-variate x-reg}} + u_t$

- Ideal is to have stationarity in both auto and cross correlations
- If unit roots are present in the time series, integration of at least order one
- If cross-correlation exists over time, it is possible that one is a leading indicator of the other

Vector ARMA Models

- Regular EDA needs to be applied to both time series independent of each other

Example: S Persei & X-Ray Data

EDA

- Single and double line plots
- Scatter plot: x and y axes each represent a time series, non-time-dependent plot, add a regression line of best fit, look for slope and R^2 (see code)
- Fit a regression on the two timeseries to look at linear interdependance $(corr)^2 = R^2$
- Multi-variate Normal Test, require(MVN), mardia test
- Li-Mcleod Test: Multivariate test for presence of ARCH effects
- Cross-Correlation Function - output: $PACF_{x_1}$, CCF_{x_1,x_2}, CCF_{x_2,x_1}, $PACF_{x_2}$ $x_2 x_1$ / $x_1 x_2$ is this correct?
- MQ test for cross-correlation over lags (similar visual to mcleod li)
- CCM 2x2, lag 1

ccv x_1 -   +   ·   ccov x_1, x_2 (lead)   lead and lag will have opposite signs

ccov $x_1 x_2$ (lag)   ·   - · cor x_2 + = leading, - = lagging

Modeling & Model Diagnostics

VARselect provides suggested VAR parameters

  - Set lag.max based on EDA, significant parameters → p

- Portmanteau test - independence in lags
- VAR forecast produces two plots, one for each timeseries based on cross-correlation

Module 8: Multivariate Volatility Models

Questions and Recap

- $(\alpha + \beta) \leq 1$, closer to 1, persistent → not good for model diagnostics, indicates model didn't explain variance well

Multi-Variate Volatility Models

- Possible cross-correlations in the volatility of variables in a multi-variate system
- Can contain combinations of two or more univariate time series, mean, volatility or combination mean/volatility models.
- Similar to uni-variate volatility models, multivariate volatility models are fitted on the residuals of multivariate mean models.
- Multiple univariate volatility vectors are combined into a matrix.
- Three aspects of multivariate volatility timeseries to test:
  1) Volatility of each series
  2) Rolling correlation
  3) DCC volatility models (Dynamic cross-correlation)

Rolling Correlation

- Focus is on whether variables have a lead/lag relationship
- Changes in cross-correlation structure of volatility over time can be analyzed
- 'Sliding window' is the number of observations used to calculate correlation at one interval of the cross-correlation structure (matrix).

Vectorization (VEC) Model

- Similar in concept to a VAR model, but for volatility modeling instead of mean modeling
- Model does not account for if one time series affects the volatility of others
  ↳ does not allow for 'dynamic dependence'
- Diagonal Vectorization (VEC) matrices must be positive definite (PD)

Dynamic Conditional Correlation (DCC Model)

- Variation of a multivariate GARCH model
  Variable in multi-variate timeseries are assumed to follow Students t-distribution
- Something about α and β values and significance
  - provide α and β for univariate and multivariate persistence

Example

1) EDA
   - Line plots
   - Take first difference of timeseries (no justification), line plot
   - ARCH Effects test (univariate)
2) GARCH Model

   diagnostics:
   - Look at parameter significance in summary(model) output
   - Sum coefficients to assess persistence
   - Histograms of variable & residuals
   - QQ plot of variable & residuals
   - ACF Plots of variable^2 & residuals^2 (look for quick drop ⇒ stationary)
   - MVN test on residuals?
   - ADF/KPSS tests on residuals?
   - Model fit output contains two diagnostic tests: Jarque Bera Test, Box-Ljung test
3) fgarch::garchFit() (these are univariate)

   model diagnostics:
   - model output contains the results of statistical tests on models' residuals
   - Standardized-Residuals Tests
   - Jarque-Bera test - Ho: normal
   - Shapiro-Wilkins test - Ho: normal
   - Ljung-Box test - Ho: non-auto-correlated, independant
   - LM arch test - Ho: no arch, independant

   * volatility = resi^2

   * plot(m) allows to print model diagrams automatically
   - ACF plot
4) Rolling Correlations
   - Plot
   - Lead/lag relationship: positive and negative values of rolling-correlation indicate direction of lead/lag relationship between variables.
     - Which sign a variables is mapped to is determined by model/function input
   - 'Calculate Exponentially-Weighted Moving Average covariances and correlations
     - set lambda to 0.94
     - transformation results in new visualization of cov & corr
     - Half-life

Module 8: Multivariate Volatility Models

Questions and Recap

- $(\alpha+\beta) \leq 1$, closer to 1, persistent → not good for model diagnostics, indicates model didn't explain variance well

Multi-Variate Volatility Models

- Possible cross-correlations in the volatility of variables in a multi-variate system
- Can contain combinations of two or more univariate time series, mean, volatility or combination mean/volatility models,
- Similar to uni-variate volatility models, multivariate volatility models are fitted on the residuals of multivariate mean models.
- Multiple univariate volatility vectors are combined into a matrix.
- Three aspects of multivariate volatility timeseries to test:
  1) Volatility of each series
  2) Rolling correlation
  3) DCC volatility models (Dynamic cross-correlation)

Rolling Correlation

- Focus is on whether variables have a lead/lag relationship
- Changes in cross-correlation structure of volatility over time can be analyzed
- 'Sliding window' is the number of observations used to calculate correlation at one interval of the cross-correlation structure (matrix).

Vectorization (VEC) Model

- Similar in concept to a VAR model, but for volatility modeling instead of mean models
- Model does not account for if one time series affects the volatility of others
  - ↳ does not allow for 'dynamic dependence'
- Diagonal Vectorization (VEC) matrices must be positive definite (PD)

Dynamic Conditional Correlation (DCC Model)

- Variation of a multivariate GARCH model
- Variable in multi-variate time series are assumed to follow Students t-distribution
- Something about α and β values and significance
  - provide α and β for univariate and multivariate persistence

Example

1) EDA
   - Line plots
   - Take first difference of time series (no justification), line plot
   - ARCH Effects test (univariate)
2) GARCH Model

   diagnostics
   - Look at parameter significance in summary(model) output
   - Sum coefficients to assess persistence
   - Histograms of variable & residuals
   - QQ plot of variable & residuals
   - ACF Plots of variable^2 & residuals^2 (look for quick drop ⇒ stationary)
   - MVN test on residuals?
   - ADF/KPSS tests on residuals?
   - Model fit output contains two diagnostic tests: Jarque Bera Test, Box-Ljung test
3) fgarch::garchFit() (these are univariate)

   model diagnostics
   - model output contains the results of statistical tests on models' residuals
   - Standardized-Residuals Tests
   - Jarque-Bera test · Ho: normal
   - Shapiro-Wilkins test - Ho: normal
   - Ljung-Box test - Ho: non-auto-correlated, independant
   - LM arch test - Ho: no arch, independant

   * volatility = resi^2

   * plot(m) allows to print model diagrams automatically
   - ACF plot
4) Rolling Correlations
   - Plot
   - Lead/lag relationship: positive and negative values of rolling-correlation indicate direction of lead/lag relationship between variables.
     - Which sign a variables is mapped to is determined by model/function input
   - 'Calculate Exponentially-Weighted Moving Average covariances and correlations
     - set lambda to 0.94
     - transformation results in new visualization of cov & corr
     - Half-life: (Tsay) -

Example

5. Univariate Normal GARCH(1,1) using ugarchfit()
   - Large output for model diagnostics and comparison
   - sum(coef(m))
6. DCC Estimation (~Multivariate GARCH)
   a) DCC Model Building and coefficient analysis
   - Sum α and β for individual variable models to assess persistence
   - Sum joint? values? for joint persistence
     - ↳ low joint = one doesnt strongly influence the other, closer to one, tied together

   sum(coef(dcc.fit)[3:4]) → MSFT

   [7:8] → GSPC

   [9:10] → joint

   b) DCC Rolling Covariance, Correlation
   - Strong influence of one variable on another is extreme values, direction by sign
   - Joint sum of $\alpha+\beta$ calculated above is indicative of strength of influence, closer to one, stronger, more persistent
   - Red line is the correlation of both time series assuming all observations are iid
7. Forecast
   - Forecasts are in the rolling covariance and correlation plots
   - Expected rolling correlation between variables based on 'trend' of rolling correlation
   - 'Interdependent volatility/variance'

Module 9 - Non-Gaussian Time-Series Models

Canvas Notes - Exponential, Smoothing, and Fourier Analysis

1) Exponential Smoothing

Holt-Winters Method

- Two variations, additive and multiplicative
- Appropriate variation determined by characteristics of seasonal component
- Additive appropriate when seasonal variations are constant,
  multiplicative appropriate when seasonal variations are changing in proportion to the level of the series
- Dampening parameter can be added to the model, $0 \le \phi \le 1$, usually $\phi \ge 0.8$

2) Prophet

- Time series forecasting system from Meta that can handle multiple trends and seasonal components, and can model holidays
- Value = trend + seasonal value + holiday + error

3) Fourier Analysis

Frequency Domain Spectral Analysis - decomposition of time series into component sine and cosine functions with different frequencies

DFT Model - a regression-like model

- Assumptions of Fourier transformation:
  - Stationarity
  - Infinite duration
  - Constant time intervals
  - Sinusoidal over white noise

Example

EDA

- Time plot
- Analyze seasonal component to determine multiplicative or additive
- Fixed magnitude of seasonal component ⇒ additive

Model Fitting & Comparison

- Compare reasonability of model forecasts
  - Trend
  - Magnitude of seasonal activity
  - Range of confidence interval
- Compare error measurements (RMSE)
- Compare distribution of residuals
  - Time plot
  - Normality (histogram, QQ plot)
  - Stationarity (time plot, ADF/KPSS, McLeod-Li)
  - Auto-correlation (ACF plot)

Prophet

- Plot of seasonal pattern
- Plot of forecast
- Plots of components (trend, seasonality)
- Cross-validation feature
- Plots of MAPE's

Lecture Notes

Parametric Models: models with estimated parameters, assumptions have been normality, stationarity, distribution of residuals describes fit

Non-Gaussian Models

- Exponential smoothing
- Generalized Additive Models (GAM)
- Spectral Decomposition
- Random Forests
- Recurrent Neural Networks (RNN)
- Hidden Markov Models (HMM)

Exponential Smoothing

- Value at time t is a weighted average of previous values
- Weights decline exponentially
- Trend and seasonality can be added
- Two variations: Additive and Multiplicative
- Model configuration: ES(e, t, s)

  where: e = error type   Types: A = additive   N = None
  t = trend type   M = multiplicative
  s = seasonal type   Z = auto
- Damping ~ Regularization

Prophet

- Generalized Additive Model (GAM)
- Linear model where value is the sum of smoothed functions
- Gaps in data can exist and time intervals can be non-constant

Fourier Analysis

- Discrete Fourier Transformation (DFT)
- A time series function is of period p when $f(t) = f(t \pm p)$ assuming stationarity
- P = fundamental period of time series