Time Series Analysis

Seasonal Models

Hossein Moradi Rekabdarkolaee

South Dakota State University

email: *hossein.moradirekabdarkolaee@sdstate.edu*.

Office: CAME Building, *Room: 240*.

Introduction

- So far, we have learned
  - Different components of time series
    - Trend,
    - Seasonality,
    - Error.
  - Different type of time series
    - Stationary (weak and strong)
    - Non-stationary.
  - Different models for time series
    - Autoregressive,
    - Moving Average,
    - ARMA,
    - ARIMA.

Introduction

- Different way to find the order of time series
  - ACF,
  - PACF,
  - EACF,
  - Information Criteria-based method (AIC, BIC, AICc, etc.)
- Different approaches to estimate the parameters of the model we choose to fit the time series.
  - Method of Moment,
  - Least Squares,
  - Maximum Likelihood.

Introduction

- After fitting the model, we did
  - Model diagnostics or model criticism,
  - What does it mean
  - How to do it.
- We studied the prediction at a future time for different time series models.
- In this section we will study seasonal models.

Seasonal Time Series

- Seasonality is periodic pattern in time series.
- Examples of the seasonal data can be :
  - Average Monthly Temperature Data.
  - Levels of carbon dioxide (CO_2) in the atmosphere.
  - Electricity consumption of a house.
  - Quarterly sales of a given store or company.
- There is two ways to model seasonality, depending on its type:
  - **Deterministic Seasonality**
  - **Stochastic Seasonality**

Deterministic Seasonality

- Deterministic seasonal models such as
  - Seasonal means plus linear time trend,
  - Sums of cosine curves at various frequencies plus linear time trends.
- Before starting this, note that we will denote the Period of seasonal pattern denoted by s
  - Monthly data with an annual pattern, then $s = 12$,
  - Daily data with a weekly pattern, then $s = 7$,
  - Quarterly data are seasonal with $s = 4$, etc.

Deterministic Seasonality

- Consider series,
$$Y_t = S(t) + X_t$$
- where
  - X_t is stationary series,
  - $S(t)$ is a deterministic seasonal component, i.e., it is a periodic function with period s, that is $S(t) = S(t+s), \forall t$.
- To estimate deterministic function $S(t)$, we fit s number of separate means (one for each time within the period).

$$\widehat{S}(t) = \begin{cases} \widehat{\mu}_1 & \text{for } t = 1, 1+s, 1+2s, \ldots \\ \widehat{\mu}_2 & \text{for } t = 2, 2+s, 2+2s, \ldots \\ \vdots & \qquad \vdots \\ \widehat{\mu}_s & \text{for } t = s, 2s, 3s, \ldots \end{cases}$$

Stochastic Seasonality

- More common situation is when seasonality is stochastic,

$$Y_t = S_t + X_t,$$

  - S_t is a random variable, typically dependent on past values of Y_t
  - For example, monthly sales with an annual pattern, where this year's Jan sales depend on last year's Jan sales,
  - or quarterly ice cream sales are seasonal with period 4.

- For stochastic process Y_t, it is seasonal with periodicity s, if Y_t and Y_{t+s} have the same distribution.

- A simple way to model this behavior is to look at auto-regression at multiple lags of period s.

Seasonal MA Models

- Consider the time series generated according to

$$Y_t = e_t - \Theta e_{t-12},$$

- Then we can show that

$$Cov(Y_t, Y_{t-1}) = Cov(e_t - \Theta e_{t-12}, e_{t-1} - \Theta e_{t-13}) = 0,$$

$$Cov(Y_t, Y_{t-12}) = Cov(e_t - \Theta e_{t-12}, e_{t-12} - \Theta e_{t-24}) = -\Theta \sigma^2.$$

- Generalizing this idea, a seasonal MA(Q) model of order Q with seasonal period s is defined as

$$Y_t = e_t - \Theta_1 e_{t-s} - \Theta_2 e_{t-2s} - \ldots - \Theta_Q e_{t-Qs},$$

Seasonal MA Models

- Therefore the seasonal MA(Q) model of order Q is defined as

$$Y_t = e_t - \Theta_1 e_{t-s} - \Theta_2 e_{t-2s} - \ldots - \Theta_Q e_{t-Qs},$$

- with seasonal MA characteristic polynomial

$$\Theta(x) = 1 - \Theta_1 x^s - \Theta_2 x^{2s} - \ldots - \Theta_Q x^{Qs},$$

- It is well-known that such a series is always stationary and that the autocorrelation function will be nonzero only at the seasonal lags of $s, 2s, 3s, \ldots, Qs$.
- In particular, $k = 1, 2, \ldots, Q$,

$$\rho_{ks} = \frac{-\Theta_k + \Theta_1\Theta_{k+1} + \Theta_2\Theta_{k+2} + \ldots, -\Theta_{Q-k}\Theta_Q}{1 + -\Theta_1^2 + \Theta_2^2 + \ldots + \Theta_Q^2}$$

Seasonal AR Models

- Seasonal autoregressive models can also be defined as

$$Y_t = \Phi Y_{t-12} + e_t,$$

- where e_t is independent of $Y_{t-1}, Y_{t-2}, \ldots$.
- Stationarity is implied by $|\Phi| < 1$ and $E[Y_t] = 0$.
- Similar to non-seasonal AR models, the Yule-Walker equation can be obtained

$$\rho_k = \Phi \rho_{k-12}, \quad \text{for} \quad k \geq 1$$

- Clearly, $\rho_{12} = \Phi \rho_0$ and $\rho_{24} = \Phi^2$.
- More generally, we have

$$\rho_k = \begin{cases} \Phi^{\frac{k}{12}} & \text{for } k = 12, 24, 36, \ldots \\ 0 & \text{o.w.} \end{cases}$$

Seasonal AR Models

- A seasonal AR(P) model of order P and seasonal period s is defined as

$$Y_t = \Phi_1 Y_{t-s} + \Phi_2 Y_{t-2s} + \ldots + \Phi_P Y_{t-Ps} + e_t,$$

- where e_t is independent of $Y_{t-1}, Y_{t-2}, \ldots$.

- The seasonal AR characteristic function

$$\Phi(x) = 1 - \Phi_1 x^s - \Phi_2 x^{2s} - \ldots - \Phi_P x^{Ps}$$

- For stationarity the roots of $\Phi(x) = 0$ need to be greater than one in absolute value.

- It can be shown that the autocorrelation function is nonzero only at lags $s, 2s, 3s, \ldots$, where it behaves like a combination of decaying exponentials and damped sine functions.

Seasonal ARMA Model

- A seasonal ARMA model of order (P, Q) and period s, denoted by SARMA$(P, Q)_s$, is

$$Y_t = \Phi_1 Y_{t-s} + \Phi_2 Y_{t-2s} + \ldots + \Phi_P Y_{t-Ps} \\ + e_t - \Theta_1 e_{t-s} - \Theta_2 e_{t-2s} - \Theta_Q e_{t-Qs}$$

- The SARMA$(P, Q)_s$ model is causal (then stationary) if all the roots of $\Phi(x^s)$ be greater than one in absolute value.

- The SARMA$(P, Q)_s$ model is invertible if and only if the roots of $\Theta(x^s)$ be greater than one in absolute value.

- ACF and PACF of SARMA$(P, Q)_s$ models behaves similarly to ACF and PACF of usual ARMA(p, q) at multiple lags of s, and is zero elsewhere

Multiplicative Seasonal ARMA Models

- We can combine seasonal SARMA$(P, Q)_s$ models with simple nonseasonal ARMA(p, q) models.
- These types of models would develop parsimonious models that contain autocorrelation for the seasonal lags but also for low lags of neighboring series values.
- This model is called multiplicative SARMA$(p, q) \times (P, Q)_s$ model with seasonal period s.

Multiplicative Seasonal ARMA Models

- We can obtain multiplicative SARMA$(p, q) \times (P, Q)_s$ models by multiplying the corresponding polynomials as a model with AR characteristic polynomial $\phi(x)\Phi(x)$ and MA characteristic polynomial $\theta(x)\Theta(x)$ where

$$\begin{cases} \phi(x) & = 1 - \phi_1 x - \phi_2 x^2 - \ldots - \phi_p x^p \\ \Phi(x) & = 1 - \Phi_1 x^s - \Phi_2 x^{2s} - \ldots - \Phi_P x^{Ps} \end{cases}$$

- and

$$\begin{cases} \theta(x) & = 1 - \theta_1 x - \theta_2 x^2 - \ldots - \theta_q x^q \\ \Theta(x) & = 1 - \Theta_1 x^s - \Theta_2 x^{2s} - \ldots - \Theta_Q x^{Qs} \end{cases}$$

Multiplicative SARMA$(p, q) \times (P, Q)_s$ Model

- Consider SARMA$(0, 1) \times (0, 1)_{12}$ model, then we have

$$Y_t = e_t - \theta e_{t-1} - \Theta e_{t-12} + \theta\Theta e_{t-13},$$

- Then, we can show that

$$\begin{aligned}
\gamma_0 &= (1 + \theta^2)(1 + \Theta^2)\sigma^2, \\
\gamma_1 &= Cov(Y_t, Y_{t-1}) = -\theta(1 + \Theta^2), \\
\gamma_{11} &= Cov(Y_t, Y_{t-11}) = \theta\Theta = \gamma_{13}, \\
\gamma_{12} &= Cov(Y_t, Y_{t-12}) = -\Theta, \\
\rho_1 &= \frac{-\theta}{1 + \theta^2}, \\
\rho_{11} &= \rho_{13} = \frac{\theta\Theta}{(1 + \theta^2)(1 + \Theta^2)}, \\
\rho_{12} &= \frac{-\Theta}{1 + \Theta^2}.
\end{aligned}$$

Multiplicative SARMA$(p, q) \times (P, Q)_s$ Model

- Consider SARMA$(0, 1) \times (0, 1)_{12}$ model, then we have

$$Y_t = e_t - \theta e_{t-1} - \Theta e_{t-12} + \theta\Theta e_{t-13}$$

- Suppose $\theta = -0.5$ and $\Theta = -0.8$, then we have

$$\rho_1 = \frac{-\theta}{1+\theta^2} = 0.4,$$

$$\rho_{11} = \rho_{13} = \frac{\theta\Theta}{(1+\theta^2)(1+\Theta^2)} = 0.19,$$

$$\rho_{12} = \frac{-\Theta}{1+\Theta^2} = 0.49$$

- Consider SARMA$(0, 1) \times (1, 0)_{12}$ model, then we have

$$Y_t = \Phi Y_{t-12} + e_t - \theta e_{t-1}$$

- Then, we can show that

$$\begin{aligned}
\gamma_1 &= \Phi\gamma_{11} - \theta\sigma^2, \\
\gamma_k &= \Phi\gamma_{11}, \qquad \text{for} \quad k \geq 2 \\
\gamma_0 &= \left[\frac{1+\theta^2}{1-\Phi^2}\right]\sigma^2, \\
\rho_{12k} &= \Phi^k, \qquad \text{for} \quad k \geq 1 \\
\rho_{12k-1} &= \rho_{12k+1} = \left(\frac{-\theta}{1+\theta^2}\right)\Phi^k, \quad \text{for} k = 0, 1, 2, \ldots
\end{aligned}$$

Nonstationary Seasonal ARIMA Models

- Often, the seasonal component is not stationary, that is we have

$$Y_t = S_t + e_t, \quad \text{with} \quad S_t = S_{t-s} + \varepsilon_t,$$

- where e_t and ε_t are independent white noises.
- In this formulation, the $\{S_t\}$ is a seasonal random walk.
- In such cases, ACF at lags that are multiples of s does not tail off exponentially fast.
- To make series stationary, we can use seasonal differencing, i.e., differencing at lag s.

Nonstationary Seasonal ARIMA Models

- Let's Define the seasonal backshift operator ∇_s^D to be the D^{th}-order difference operator at lag s.
- Then by applying an appropriate order seasonal differencing, $\nabla_s^D Y_t$ will be stationary.
- For instance, the first seasonal differencing is

$$\nabla_s Y_t = S_t - S_{t-s} + e_t - e_{t-s} = \varepsilon_t + e_t - e_{t-s}$$

- This model can be generalized for cases that has a nonseasonal stochastic trend.

Integrated Seasonal ARMA Models

- The multiplicative seasonal autoregressive integrated moving average model, or SARIMA$(p, d, q) \times (P, D, Q)_s$, is given as follows

$$W_t = \nabla^d \nabla_s^D Y_t$$

- The d is the order of ordinary differencing, and D is the order of seasonal differencing
- Clearly, such models represent a broad, flexible class from which to select an appropriate model for a particular time series.
- We can still use AIC and/or BIC to select model orders and utilize the usual diagnostics to check model fit

Model Specification, Fitting, and Checking

- Model specification, fitting, and diagnostic checking for seasonal models follow the same general techniques developed for nonseasonal models.
- For model specification, we would start looking at the time series plot, ACF, and PACF.
- We want to detect if there are certain lags and their multiplication that are significant.
- We also want to make sure that we de-trand the data.
- This could be done using differencing or fitting a model as trend.

Model Specification, Fitting, and Checking

- Model specification criteria are similar to those in the ARIMA model.
- The goal is to follow Box-Jenkins' "principle of parsimony" for model selection.
- That is, we increase the order of the multiplicative SARIMA model with caution.
- After determining the order of the SARIMA model, we have to run the model diagnostic.
- If we are satisfied with the model, it can be used for forecasting.

Forecasting Seasonal Models

- We can use the similar idea of forecasting function for nonseasonal ARMA, given in Ch 9, for SARMA models:

$$\widehat{Y}_t(\ell) = E[Y_{t+\ell}|I_t] = E[Y_{t+\ell}|Y_1, Y_2, \ldots, Y_{t-1}, Y_t]$$

- For example, consider the model $\text{ARIMA}(0,1,1) \times (1,0,1)_{12}$

$$Y_t = Y_{t-1} + \Phi Y_{t-12} - \Phi Y_{t-13} + e_t - \theta e_{t-1} - \Theta e_{t-12} + \theta\Theta e_{t-13}$$

- The one and two-step-ahead forecast is

$$\widehat{Y}_t(1) = Y_t + \Phi Y_{t-11} - \Phi Y_{t-12} - \theta e_t - \Theta e_{t-11} + \theta\Theta e_{t-12}$$
$$\widehat{Y}_t(2) = \widehat{Y}_t(1) + \Phi Y_{t-10} - \Phi Y_{t-11} - \Theta e_{t-10} + \theta\Theta e_{t-11}$$

- Then for $\ell > 13$, we have

$$\widehat{Y}_t(\ell) = \widehat{Y}_t(\ell-1) + \Phi\widehat{Y}_t(\ell-12) - \Phi\widehat{Y}_t(\ell-13)$$

Forecasting Seasonal Models

- For a seasonal model $AR(1)_{12}$

$$Y_t = \phi Y_{t-12} + e_t,$$

- The forecast is

$$\hat{Y}_t(\ell) = \phi \hat{Y}_t(\ell - 12)$$

- Iterating back on ℓ, we have

$$\hat{Y}_t(\ell) = \phi^{k-1} \hat{Y}_{t+r-11}$$

- where $\ell = 12k + r + 1$ with $0 \leq r < 12$ and $k = 0, 1, 2, \dots$ i.e. k is the integer part of $(\ell - 1)/12$ and $r/12$ is the fractional part.
- The weights ψ are

$$\phi = \begin{cases} \phi^{j/12} & \text{for } \; j = 0, 12, 24, \dots \\ 0 & O.W. \end{cases} \quad \text{and} \;\; var(e_t(\ell)) = \left[\frac{1 - \phi^{2k+2}}{1 - \phi^2} \right] \sigma^2$$

Forecasting Seasonal Models

- For seasonal $MA(1)_{12}$,

$$Y_t = e_t - \Theta e_{t-12} + \theta_0.$$

- The forecast is

$$\begin{cases} \hat{Y}_t(1) = -\Theta e_{t-11} + \theta_0, \\ \hat{Y}_t(2) = -\Theta e_{t-10} + \theta_0, \\ \vdots \\ \hat{Y}_t(12) = -\Theta e_t + \theta_0, \end{cases}$$

- and $\hat{Y}_t(\ell) = \theta_0$ for $\ell > 12$.
- For this model $\psi_0 = 1$, $\psi_{12} = -\Theta$ and $\psi_j = 0$ and the variance is

$$\begin{cases} \sigma^2 & 1 \leq \ell \leq 12 \\ (1+\Theta^2)\sigma^2 & 12 \leq \ell \end{cases}$$

Forecasting Seasonal Models

- For an $ARIMA(0, 0, 0) \times (0, 1, 1)_{12}$

$$Y_t - Y_{t-12} = e_t - \Theta e_{t-12}$$

- or

$$Y_{t+\ell} = Y_{t+\ell-12} + e_{t+\ell} - \Theta e_{t+\ell-12}$$

- The forecast is

$$\begin{cases} \hat{Y}_t(1) = Y_{t-11} - \Theta e_{t-11}, \\ \hat{Y}_t(2) = Y_{t-10} - \Theta e_{t-10}, \\ \vdots \\ \hat{Y}_t(12) = Y_t - \Theta e_t, \end{cases}$$

- and $\hat{Y}_t(\ell) = \hat{Y}_t(\ell - 12)$ for $\ell > 12$.

Forecasting Seasonal Models

- If we invert this model, we have

$$Y_t = (1 - \Theta)(Y_{t-12} + \Theta Y_{t-24} + \Theta^2 Y_{t-36} + \dots) + e_t$$

- Thus, we have

$$\begin{cases} \hat{Y}_t(1) = (1 - \Theta) \sum_{j=0}^{\infty} \Theta^j Y_{t-11-12j}, \\ \hat{Y}_t(2) = (1 - \Theta) \sum_{j=0}^{\infty} \Theta^j Y_{t-10-12j}, \\ \vdots \\ \hat{Y}_t(12) = (1 - \Theta) \sum_{j=0}^{\infty} \Theta^j Y_{t-12j}. \end{cases}$$

- Thus, we have $\psi_j = 1 - \Theta$ for $j = 12, 24, \dots$ and 0 otherwise.

- The forecast error variance is

$$var(e_t(\ell)) = [1 + k(1 - \Theta)^2]\sigma^2$$

Forecasting Seasonal Models

- For the $ARIMA(0,1,1)\times(0,1,1)_{12}$ model

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + e_t - \theta e_{t-1} - \Theta e_{t-12} + \theta\Theta e_{t-13},$$

- the forecast is

$$\begin{cases} \hat{Y}_t(1) = Y_t + Y_{t-11} - Y_{t-12} - \theta e_t - \Theta e_{t-11} + \theta\Theta e_{t-12}, \\ \hat{Y}_t(2) = \hat{Y}_t(1) + Y_{t-10} - Y_{t-11} - \Theta e_{t-10} + \theta\Theta e_{t-11}, \\ \vdots \\ \hat{Y}_t(12) = \hat{Y}_t(11) - Y_t + Y_{t-1} - \Theta e_t + \theta\Theta e_{t-1}, \\ \hat{Y}_t(13) = \hat{Y}_t(12) + \hat{Y}_t(1) - Y_t + \theta\Theta e_t \end{cases}$$

- and for $\ell > 13$

$$\hat{Y}_t(\ell) = \hat{Y}_t(\ell-1) + \hat{Y}_t(\ell-12) - \hat{Y}_t(\ell-13).$$

SARIMA Model Selection

- SARIMA offers a flexible framework for time series modeling and forecasting
- It is appropriate for time series with geometrically decaying ACF.
- However, model selection can be complicated since we need to compare AIC/BIC of all possible SARIMA$(p, d, q) \times (P, D, Q)_s$ combinations
- We can check ACF/PACF of (possibly differenced) data for guidance, and use a stepwise selection method