Seasonal ARIMA

Dr. Natalia Khorunzhina

CBS

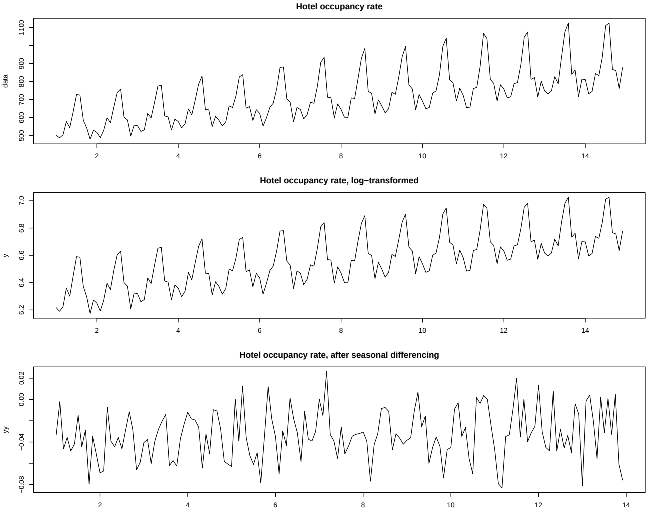
Seasonal differencing

- If the data is not stationary, taking differences can help making it stationary.
- Seasonal data is not stationary. Taking seasonal differences can fix the non-stationarity of this kind.
- A seasonal difference is the difference between an observation and the previous observation from the same season. So, if s is the number of seasons, then

$$z_t = y_t - y_{t-s}$$

- To distinguish seasonal differences from ordinary differences, we sometimes refer to ordinary differences as “first differences”, meaning differences at lag 1.

Hotel occupancy rate



Seasonal differencing

- Sometimes it is necessary to take both a seasonal difference and a first difference to obtain stationary data
- After a round of seasonal differences, test for Unit Root. If the data were not sufficiently stationary, then take an extra round of differencing. This time it will be first differences, since the seasonality is removed in the first round of seasonal differences.
- The twice-differenced series is

$$\begin{aligned} z_t &= (y_t - y_{t-s}) - (y_{t-1} - y_{t-s-1}) \\ &= y_t - y_{t-1} - y_{t-s} + y_{t-s-1} \end{aligned}$$

Interpretation of seasonal differencing

- First differences are the change between one observation and the next - a growth over consecutive periods.
- Seasonal differences are the change between one year to the next - a growth between consecutive years.

Seasonal ARIMA models

- A seasonal ARIMA model is formed by including additional seasonal terms

$$ARIMA(p, d, q)(P, D, Q)[s]$$

  where s is the number of observations per year, (p, d, q) is the order of non-seasonal part of the model, and (P, D, Q) is the order of the seasonal part of the model.

- The seasonal part of an AR or MA model will be seen in the seasonal lags of the PACF and ACF.

Seasonal ARIMA models

In considering the appropriate seasonal orders for a seasonal ARIMA model, restrict attention to the seasonal lags.

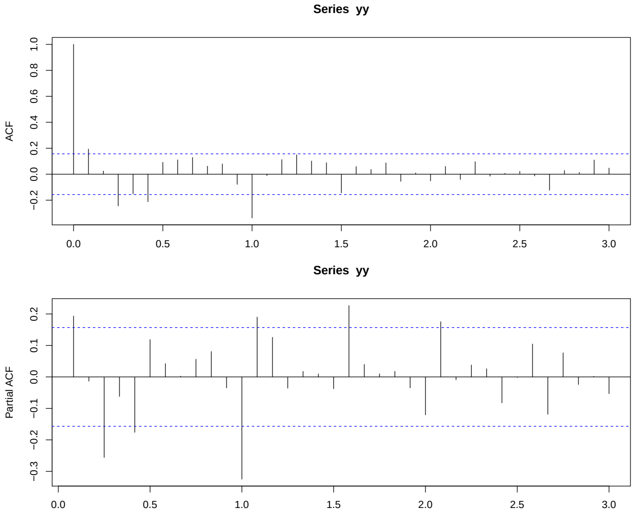
For example, an ARIMA$(0,0,0)(0,1,1)_{12}$ model will show:

- a spike at lag 12 in the ACF but no other significant spikes
- exponential decay in the seasonal lags of the PACF (i.e., at lags 12, 24, 36 etc.)

Similarly, an ARIMA$(0,0,0)(1,1,0)_{12}$ model will show:

- exponential decay in the seasonal lags of the ACF
- a single significant spike at lag 12 in the PACF

ACF and PACF after taking seasonal differences: Hotel example



ACF and PACF after taking seasonal differences: Hotel example

- The significant spike at lag 1 in the ACF suggests a non-seasonal MA(1) component, whereas the significant spike at lag 12 in the ACF suggests a seasonal MA(1) component.
- Consequently, we can estimate an ARIMA$(0, 0, 1)(0, 1, 1)_{12}$ model, indicating seasonal difference only, and non-seasonal and seasonal MA(1) components.
- By analogous logic applied to the PACF, we could also have estimated an ARIMA($(1, 0, 0)(1, 1, 0)_{12}$ model.
- The ACF and PACF give a mixed message, so alternatively we could find an ARIMA$(p, 0, q)(P, 1, Q)_{12}$ by minimizing AIC.

Forecast with Seasonal ARIMA: Hotel example