25 Spring 439/639 TSA: Lecture 21

Dr Sergey Kushnarev

Table of contents

Multiplicative seasonal ARIMA model

Similar to seasonal ARMA$(p,q)\times(P,Q)_s$, we can also combine a nonseasonal ARIMA(p,d,q) and a seasonal ARIMA$(P,D,Q)_s$.

First, we need to define seasonal ARIMA$(P,D,Q)_s$.

Recall that in the nonseasonal version, we say $Y_t \sim$ ARIMA(p,d,q) if $\nabla^d Y_t \sim$ ARMA(p,q). Where the differencing operator ∇ is

$$\nabla Y_t = Y_t - Y_{t-1} = (1-B)\ Y_t, \quad \text{so } \nabla^d Y_t = (1-B)^d\ Y_t.$$

We need a seasonal analogue for this. The seasonal differencing operator of period s, is defined as

$$\nabla_s Y_t = Y_t - Y_{t-s} = (1-B^s)\ Y_t.$$

We say $Y_t \sim$ ARIMA$(P,D,Q)_s$ if $\nabla_s^D Y_t \sim$ ARMA$(P,Q)_s$. This is called seasonal ARIMA (SARIMA).

Similarly, for multiplicative seasonal ARIMA: we say $Y_t \sim$ ARIMA$(p,d,q)\times$ARIMA$(P,D,Q)_s$ if $\nabla^d\nabla_s^D Y_t \sim$ ARMA$(p,q)\times$ ARMA$(P,Q)_s$. In other words,

$$\text{if } \nabla^d\nabla_s^D Y_t \sim \text{ARMA}(p,q)\times\text{ARMA}(P,Q)_s, \quad \text{then } Y_t \sim \text{ARIMA}(p,d,q)\times\text{ARIMA}(P,D,Q)_s.$$

Using the AR/MA polynomial, $Y_t \sim$ ARIMA$(p,d,q)\times$ ARIMA$(P,D,Q)_s$ can be written as

$$\phi(B)\ \Phi(B)\ (1-B)^d\ (1-B^s)^D\ Y_t = \theta(B)\ \Theta(B)\ e_t,$$

where $\phi(x), \Phi(x), \theta(x), \Theta(x)$ are the AR from nonseasonal, AR from seasonal, MA from nonseasonal, MA from seasonal respectively.

Example. Consider this model

$$Y_t = 0.5\,Y_{t-1} + Y_{t-4} - 0.5\,Y_{t-5} + e_t - 0.3\,e_{t-1}.$$

It can be written as

$$\begin{aligned}(1-0.5B-B^4+0.5B^5)\ Y_t &= (1-0.3B)\ e_t,\\ (1-0.5B)(1-B^4)\ Y_t &= (1-0.3B)\ e_t.\end{aligned}$$

This is an ARIMA$(1,0,1)\times$ARIMA$(0,1,0)_4$: $(1-0.5B)$ and $(1-0.3B)$ are nonseasonal AR/MA polynomial, with orders $p=q=1$. $(1-B^4)^1$ is a seasonal differencing operator of period $s=4$, with order $D=1$.

Example. Consider this model

$$Y_t = Y_{t-4} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}.$$

Rewrite it as

$$(1 - B^4)\ Y_t = (1 - \theta_1 B - \theta_2 B^2)\ e_t.$$

This is an $\text{ARIMA}(0,0,2) \times \text{ARIMA}(0,1,0)_4$.

Example. Consider this model

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + e_t - 0.1\, e_{t-1} - 0.1\, e_{t-12} + 0.01\, e_{t-13}.$$

Rewrite it as

$$\begin{aligned}(1 - B - B^{12} + B^{13})\ Y_t &= (1 - 0.1B - 0.1B^{12} + 0.01B^{13})\ e_t,\\ (1 - B)(1 - B^{12})\ Y_t &= (1 - 0.1B)(1 - 0.1B^{12})\ e_t.\end{aligned}$$

This is an $\text{ARIMA}(0,1,1) \times \text{ARIMA}(0,1,1)_{12}$: $(1-0.1B)$ and $(1-0.1B^{12})$ are the nonseasonal and seasonal (period $s = 12$) MA parts, with orders $q = Q = 1$. $(1 - B)^1$ and $(1 - B^{12})^1$ are nonseasonal and seasonal (period $s = 12$) differencing operators, with orders $d = D = 1$.

Cross-covariance, cross-correlation function

Previously, we studied forecasting, i.e. use the past values of (Y_t) to predict future Y_t. Now we consider a different setting. We may use the past values of another time series (X_t) to help predict Y_t.

Suppose (X_t, Y_t) is a vector time series,

$$(X_1, Y_1), (X_2, Y_2), ..., (X_t, Y_t), ...$$

The **Cross-covariance function (CCVF)** is defined as

$$\gamma_{t,s}(X, Y) \overset{\text{def}}{=} \text{Cov}(X_t, Y_s).$$

We can also define the **joint (weak) stationarity** for the vector time series (X_t, Y_t) (which is a generalization of the weak stationarity of a single time series (Y_t)). A vector time series (X_t, Y_t) is (weakly) stationary if it satisfies

- $\mathbb{E}[X_t]$ is a constant μ_X for all t, $\mathbb{E}[Y_t]$ is a constant μ_Y for all t.
- $\text{Var}(X_t)$ is a constant for all t, $\text{Var}(Y_t)$ is a constant for all t.
- ACVF $\gamma_{t,s}(X) = \text{Cov}(X_t, X_s)$ only depends on the lag difference $t - s$, $\gamma_{t,s}(Y) = \text{Cov}(Y_t, Y_s)$ only depends on the lag difference $t - s$.
- CCVF $\gamma_{t,s}(X, Y) = \text{Cov}(X_t, Y_s)$ only depends on the lag difference $t - s$.

So the first three conditions are just saying (X_t) and (Y_t) are both stationary. The only new requirement is the last condition on CCVF.

If the joint stationarity holds, then we can replace the notation $\gamma_{t,s}(X, Y)$ by $\gamma_{t-s}(X, Y)$, since it only depends on the lag difference $t - s$. For example, assuming joint stationarity,

$$\begin{aligned}\gamma_0(X, Y) &= \gamma_{t,t}(X, Y) = \text{Cov}(X_t, Y_t), \text{ for any } t\\ \gamma_1(X, Y) &= \gamma_{t+1,\,t}(X, Y) = \text{Cov}(X_{t+1},\ Y_t), \text{ for any } t\\ \gamma_{-1}(X, Y) &= \gamma_{t-1,\,t}(X, Y) = \text{Cov}(X_{t-1},\ Y_t), \text{ for any } t\end{aligned}$$

Note: For a single stationary time series (Y_t), the ACVF has the property that $\gamma_k(Y) = \gamma_{-k}(Y)$ by the symmetry of covariance. But for a joint stationary vector time series (X_t, Y_t), in general, $\gamma_k(X, Y) \neq \gamma_{-k}(X, Y)$.

Similarly, we can define **cross-correlation function (CCF)**. For simplicity, assume the vector time series (X_t, Y_t) is jointly stationary. The CCF is

$$\rho_k(X,Y) \stackrel{\text{def}}{=} \text{corr}(X_t, Y_{t-k}) = \frac{\gamma_k(X,Y)}{\sqrt{\gamma_0(X)\cdot\gamma_0(Y)}}.$$

Example. Consider (X_t, Y_t), where $X_t \sim \text{iid}(0, \sigma_x^2)$, and

$$Y_t = \beta_0 + \beta_1 X_{t-d} + e_t, \quad e_t \sim \text{iid}(0, \sigma_e^2),$$

and $(X_t), (e_t)$ are independent. For this vector time series (X_t, Y_t), we can show that the CCF is

$$\begin{cases} \rho_{-d}(X,Y) = \text{corr}(X_t, Y_{t+d}) = \frac{\beta_1\sigma_x}{\sqrt{\beta_1^2\sigma_x^2+\sigma_e^2}} \\ \rho_k(X,Y) = 0, \quad \text{if } k \neq -d. \end{cases}$$

Exercise: verify this CCF.

Bartlett's theorem on sample CCF

Given the observed samples from a vector time series (X_t, Y_t), we can also obtain sample CCF $r_m(X,Y)$. (Similar to the way we constructed sample ACF.)

We also have another version of Bartlett's theorem for sample CCF: when sample size n is large, the sampling distribution of the sample CCF $r_m(X,Y)$ is approximately

$$r_m(X,Y) \sim \mathcal{N}\left(\rho_m(X,Y),\ \frac{1}{n}\left(1 + 2\sum_{k=1}^{\infty}\rho_k(X)\,\rho_k(Y)\right)\right).$$

This may lead to **"spurious correlation"**: even the theoretical CCF $\rho_m(X,Y)$ is small (or zero), the sample CCF $r_m(X,Y)$ may still be "large" which seemingly implies correlation at lag m. *Note:* here "large" is in the sense of comparing to the standard "$\frac{2}{\sqrt{n}}$ rule" used in the software.

Example. Suppose $X_t \sim \text{AR}(1)$, $Y_t \sim \text{AR}(1)$, and $(X_t), (Y_t)$ are independent. So the theoretical CCF $\rho_m(X,Y) = 0$ for any m, and the variance term in Bartlett theorem above is

$$\text{Var}(r_m(X,Y)) = \frac{1}{n}\left(1 + 2\sum_{k=1}^{\infty}\phi_X^k\phi_Y^k\right) = \frac{1}{n}\left(\frac{1+\phi_X\phi_Y}{1-\phi_X\phi_Y}\right).$$

For example, if $\phi_X = \phi_Y = \frac{1}{2}$, then $\text{Var}(r_m(X,Y)) \approx \frac{1.67}{n}$. So the sampling distribution has larger variance than the standard $\frac{1}{n}$, which makes the standard "$\frac{2}{\sqrt{n}}$ rule" not reliable here.