STAT 153 & 248 - Time Series
Lecture Twenty Three

Fall 2025, UC Berkeley

Aditya Guntuboyina

November 20, 2025

1 The Box-Jenkins Time Series Modeling Strategy

Box and Jenkins popularized the following strategy for modeling an observed time series $y_1, \ldots, y_n$:

1. Generally $y_1, \ldots, y_n$ will exhibit various kinds of trends. Preprocess the data to transform it to another series x_t which does not have any discernible trends.
2. Fit an ARMA(p, q) model for appropriate p and q to the transformed data x_t.

The preprocessing in the first step above is usually done in one of the following two ways:

1. **Differencing**. The first difference of $\{y_t\}$ is given by $\nabla y_t := y_t - y_{t-1}$ for $t = 2, \ldots, n$. The second difference is given by

$$\begin{aligned}\nabla^2 y_t &= \nabla(\nabla y_t) \\ &= \nabla(y_t - y_{t-1}) = \nabla y_t - \nabla y_{t-1} = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}.\end{aligned}$$

   Higher order differences $\nabla^k y_t$ are defined recursively. Note that the length of the time series comes down after each successive differencing. For example, ∇y_t has length $n-1$, $\nabla^2 y_t$ has length $n-2$ and so on. Differencing usually eliminates increasing/decreasing trends. Usually one or two orders of differencing is enough to take care of increasing/decreasing trends.

2. **Seasonal Differencing**. Seasonal differencing is used to eliminate seasonal trends. Suppose we have a dataset having seasonal trends with period s (for example, for monthly datasets, $s = 12$). The seasonal first difference of y_t with period s is defined as

$$\nabla_s y_t := y_t - y_{t-s}$$

   Note that $\nabla_s y_t$ is a time series of length $n - s$. The second order seasonal difference is

$$\nabla_s^2 y_t = \nabla_s(\nabla_s y_t) = y_t - 2y_{t-s} + y_{t-2s}$$

   and higher order seasonal differences are defined recursively. Seasonal differences eliminate seasonal trends. Usually, in datasets having seasonal and increasing/decreasing trends, one first takes a seasonal difference. This often eliminates seasonality and might also eliminate the linear trend. If a linear trend still persists, one takes a regular difference of the seasonal differenced series. This will often give a series with no trend and seasonality.

To the transformed data x_t, one fits an ARMA(p, q) model which can be done via the `ARIMA` function from the statsmodels library. The order p and q can be determined via a model selection criterion such as AIC or BIC.

2 ARIMA models

ARIMA stands for AutoRegressive Integrated Moving Average. ARIMA is essentially differencing plus ARMA.

Definition 2.1 (ARIMA). *A time series model y_t is said to be ARIMA(p, d, q) if*

$$\phi(B)((\nabla^d y_t) - \mu) = \theta(B)\epsilon_t,$$

where $\epsilon_t \overset{texti.i.d}{\sim} N(0, \sigma^2)$.

ARIMA models are fit by the function `ARIMA()` in statismodels. The mean μ above is taken to be zero by default when the order parameter d in `ARIMA` is strictly larger than zero.

3 Seasonal ARMA Models

Seasonal ARMA models are often useful while modeling datasets having seasonal features (e.g., monthly datasets). We say that $\{y_t\}$ is a seasonal ARMA$(P,\ Q)$ process with period s if it satisfies the difference equation $\Phi(B^s)(y_t - \mu) = \Theta(B^s)\epsilon_t$ where $\epsilon_t \overset{\text{i.i.d}}{\sim} N(0, \sigma^2)$ and

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

and

$$\Theta(B^s) = 1 + \Theta_1 B^2 + \Theta_2 B^{2s} + \cdots + \Theta_Q B^{Qs}.$$

The seasonal ARMA$(P,\ Q)$ model with period s is a special case of an ARMA$(Ps,\ Qs)$ model. However the seasonal model has $P + Q + 1$ (the 1 is for σ^2) parameters while a general ARMA$(Ps,\ Qs)$ model will have $Ps + Qs + 1$ parameters. So the seasonal models are much sparser.

Causal stationary solution exists when every root of $\Phi(z^s)$ (equivalently, $\Phi(z)$) has modulus strictly larger than one.

The ACF and PACF of seasonal ARMA models are **non-zero** only at the seasonal lags $h = 0, s, 2s, 3s, \ldots$. At these seasonal lags, the ACF and PACF of these models behave just as the case of the unseasonal ARMA model: $\Phi(B)X_t = \Theta(B)\epsilon_t$.

4 Multiplicative Seasonal ARMA Models

For the `co2` dataset (from the time series analysis textbook by Cryer and Chan), for the first and seasonal differenced data, we saw that the sample autocorrelations seem nonnegligible at lags 0, 1, 11, 12, 13 and those at all other lags seem negligible. This behaviour can be produced in a MA(13) model but that model will have 14 parameters possibly leading to overfitting.

We can get a much more parsimonious model for this dataset by *combining* the MA(1) model with a seasonal MA(1) model of period 12. Specifically, consider the model

$$y_t = (1 + \Theta B^{12})(1 + \theta B)\epsilon_t = \left(1 + \theta B + \Theta B^{12} + \theta\Theta B^{13}\right)\epsilon_t = \epsilon_t + \theta\epsilon_{t-1} + \Theta\epsilon_{t-12} + \theta\Theta\epsilon_{t-13}.$$

It is easy to check that model has the autocorrelation function:

$$\rho(1) = \frac{\theta}{1+\theta^2} \quad \text{and} \quad \rho(12) = \frac{\Theta}{1+\Theta^2}$$

and

$$\rho(11) = \rho(13) = \frac{\theta\Theta}{(1+\theta^2)(1+\Theta^2)}.$$

At every other lag $h > 0$, the autocorrelation $\rho_X(h)$ equals zero. Based on this ACF (and the sample ACF calculated from the data), this model can be suitable for the first and seasonal differenced data in the co2 dataset.

More generally, we can combine, by multiplication, ARMA and seasonal ARMA models to obtain models which have special autocorrelation properties with respect to seasonal lags. The **Multiplicative Seasonal Autoregressive Moving Average Model** ARMA(p, q) $\times$ (P, Q)$_s$ is defined via the difference equation:

$$\Phi(B^s)\phi(B)(y_t - \mu) = \Theta(B^s)\theta(B)\epsilon_t.$$

The model we looked at above for the co2 dataset is ARMA(0, 1) $\times$ (0, 1)$_{12}$.

Another example of a multiplicative seasonal ARMA model is ARMA(0, 1) $\times$ (1, 0)$_{12}$ (this is same as $MA(1) \times AR(1)_{12}$)

$$(y_t - \mu) - \Phi(y_{t-12} - \mu) = \epsilon_t + \theta\epsilon_{t-1}.$$

The autocorrelation function of this model can be checked to be $\rho(12h) = \Phi^h$ for $h \geq 0$ and

$$\rho(12h-1) = \rho(12h+1) = \frac{\theta}{1+\theta^2}\Phi^h \qquad \text{for } h = 0, 1, 2, \ldots$$

and $\rho(h) = 0$ at all other lags.

When we have a dataset whose ACF and PACF show interesting patterns at seasonal lags, consider using a multiplicative seasonal ARMA model. You may use the `Statsmodels` functions `arma_acf` and `arma_pacf` to understand the autocorrelation and partial autocorrelation functions of these models.

5 SARIMA Models

These models are obtained by combining differencing with multiplicative seasonal ARMA models. These models are denoted by ARIMA(p, d, q) $\times$ (P, D, Q)$_s$. This means that after differencing d times and seasonal differencing D times (with period s), we get a multiplicative seasonal ARMA model. In other words, $\{y_t\}$ is ARIMA(p, d, q) $\times$ (P, D, Q)$_s$ if it satisfies the difference equation:

$$\Phi(B^s)\phi(B)\nabla_s^D\nabla^d(y_t - \mu) = \delta + \Theta(B^s)\theta(B)\epsilon_t.$$

Recall that $\nabla_s^d = (1 - B^s)^d$ and $\nabla^d = (1 - B)^d$ denote the differencing operators.

In the co2 example, we wanted to use the model ARMA$(0, 1) \times (0, 1)_{12}$ to the seasonal and first differenced data: $\nabla\nabla_{12}X_t$. In other words, we want to fit the SARIMA model with nonseasonal orders $0, 1, 1$ and seasonal orders $0, 1, 1$ with seasonal period 12 to the original co2 dataset. This model can be fit to the data using the function `ARIMA` with the `seasonal_order` argument.