ARIMA

ARIMA, 'AutoRegressive Integrated Moving Average', is a forecasting algorithm based on the idea that the information in the past values of the time series can alone be used to predict the future values.

Any 'non-seasonal' time series that exhibits patterns and is not a random white noise can be modeled with ARIMA models.

An ARIMA model is characterized by 3 terms: p, d, q

where,

p is the order of the AR term

p, q, d

q is the order of the MA term

d is the number of differencing required to make the time series stationary

If a time series, has seasonal patterns, then you need to add seasonal terms and it becomes SARIMA, short for 'Seasonal ARIMA'.

Auto Regressive' in ARIMA means it is a linear regressin model that uses its own lags as predictors. Linear regression models, as you know, work best when the predictors are not correlated and are independent of each other

The most common approach is to difference it. That is, subtract the previous value from the current value. Sometimes, depending on the complexity of the series, more than one differencing may be needed.

The value of d, therefore, is the minimum number of differencing needed to make the series stationary. And if the time series is already stationary, then d = 0.

Note:

A model with (only) two AR terms would be specified as an ARIMA of order (2,0,0).

A model with one AR term, a first difference, and one MA term would have order (1,1,1).

For the last model, ARIMA (1,1,1), a model with one AR term and one MA term is being applied to the variable Zt=Xt−Xt−1Zt=Xt−Xt−1. A first difference might be used to account for a linear trend in the data.

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

The differencing order refers to successive first differences. For example, for a difference order = 2 the variable analyzed is zt=(xt−xt−1)−(xt−1−xt−2)zt=(xt−xt−1)−(xt−1−xt−2), the first difference of first differences. This type of difference might account for a quadratic trend in the data.

$Z_t = (x_t - x_{t-1}) - (x_{t-1} - x_{t-2})$

A pure **Auto Regressive (AR only) model** is one where Yt depends only on its own lags. That is, Yt is a function of the 'lags of Yt'.

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + .. + \beta_p Y_{t-p} + \epsilon_1$$

linear
quadratic
expon

β

where y(t-1) is the lag of the series and beta is the coefficient of lag that the model estimates, alpha is the intercept terms, also estimated by the model.

α

Moving Average (MA only) model is one where Yt depends only on the lagged forecast errors.

$$Y_t = \alpha + \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + .. + \phi_q \epsilon_{t-q}$$

where the error terms are the errors of the autoregressive models of the respective lags. The errors Et and E(t-1) are the errors from the following equations :

$$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + .. + \beta_0 Y_0 + \epsilon_t$$

$$Y_{t-1} = \beta_1 Y_{t-2} + \beta_2 Y_{t-3} + .. + \beta_0 Y_0 + \epsilon_{t-1}$$

An ARIMA model is one where the time series was differenced at least once to make it stationary and you combine the AR and the MA terms. So the equation becomes:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + .. + \beta_p Y_{t-p} \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + .. + \phi_q \epsilon_{t-q}$$

The forecasting equation is constructed as follows. First, let y denote the d^{th} difference of Y, which means:

If d=0: $y_t = Y_t$

If d=1: $y_t = Y_t - Y_{t-1}$

If d=2: $y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2}$

Note that the second difference of Y (the d=2 case) is not the difference from 2 periods ago. Rather, it is the *first-difference-of-the-first difference*, which is the discrete analog of a second derivative, i.e., the local acceleration of the series rather than its local trend.

In terms of y, the general forecasting equation is:

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \ldots - \theta_q e_{t-q}$$

Here the moving average parameters (θ's) are defined so that their signs are negative in the equation, following the convention introduced by Box and Jenkins.

Types of *nonseasonal* ARIMA models that are commonly used:

ARIMA(1,0,0) = first-order autoregressive model AR(1) AR(1)

If the series is stationary and autocorrelated, perhaps it can be predicted as a multiple of its own previous value, plus a constant. The forecasting equation in this case is

$$\hat{Y}_t = \mu + \phi_1 Y_{t-1}$$

which is Y regressed on itself lagged by one period. This is an "ARIMA(1,0,0)+constant" model. If the mean of Y is zero, then the constant term would not be included.

In a *second-order* autoregressive model (ARIMA(2,0,0)), there would be a Y_{t-2} term on the right as well, and so on. Depending on the signs and magnitudes of the coefficients, an ARIMA(2,0,0) model

ARIMA(0,1,0) = random walk: If the series Y is not stationary, the simplest possible model for it is a random walk model, which can be considered as a limiting case of an AR(1) model in which the autoregressive coefficient is equal to 1, i.e., a series with infinitely slow mean reversion. The prediction equation for this model can be written as:

$$\hat{Y}_t - Y_{t-1} = \mu$$

or equivalently

$$\hat{Y}_t = \mu + Y_{t-1}$$

where the constant term is the average period-to-period change (i.e. the long-term drift) in Y. This model could be fitted as a *no-intercept regression model* in which the first difference of Y is the dependent variable. Since it includes (only) a nonseasonal difference and a constant term, it is classified as an "ARIMA(0,1,0) model with constant." The random-walk-*without*-drift model would be an ARIMA(0,1,0) model *without* constant

ARIMA(1,1,0) = differenced first-order autoregressive model: If the errors of a random walk model are autocorrelated, perhaps the problem can be fixed by adding one lag of the dependent variable to the prediction equation--i.e., by regressing *the first difference of Y* on itself lagged by one period. This would yield the following prediction equation:

$$\hat{Y}_t - Y_{t-1} = \mu + \phi_1(Y_{t-1} - Y_{t-2})$$

$$\hat{Y}_t - Y_{t-1} = \mu$$

which can be rearranged to

$$\boxed{\hat{Y}_t = \mu + Y_{t-1} + \phi_1 (Y_{t-1} - Y_{t-2})}$$

This is a first-order autoregressive model with one order of nonseasonal differencing and a constant term--i.e., an ARIMA(1,1,0) model.

PROBLEMS

Suppose that an AR(1) model is estimated to be $x_t=40+0.6x_{t-1}+w_t$. with the other details as,

Suppose that we have n = 100 observations, $\hat{\sigma}_w^2 = 4$ and $x_{100} = 80$. We wish to forecast the values at times 101 and 102, and create prediction intervals for both forecasts.

First we forecast time 101.

$$\begin{aligned} x_{101} &= 40 + 0.6x_{100} + w_{101} \\ x_{101}^{100} &= 40 + 0.6(80) + 0 = 88 \end{aligned}$$

The standard error of the forecast error at time 101 is

$$\sqrt{\hat{\sigma}_w^2 \sum_{j=0}^{1-1} \psi_j^2} = \sqrt{4(1)} = 2.$$

The 95% prediction interval for the value at time 101 is 88 ± 2(1.96), which is 84.08 to 91.96. We are therefore 95% confident that the observation at time 101 will be between 84.08 and 91.96. If we repeated this exact process many times, then 95% of the computed prediction intervals would contain the true value of x at time 101.

The forecast for time 102 is

$$x_{102}^{100} = 40 + 0.6(88) + 0 = 92.8$$

The relevant standard error is

$$\sqrt{\hat{\sigma}_w^2 \sum_{j=0}^{2-1} \psi_j^2} = \sqrt{4(1 + 0.6^2)} = 2.332$$

A 95% prediction interval for the value at time 102 is 92.8 ± (1.96)(2.332).

Υ 5%

95%

$x_t = 40 + 0.6\, x_{t-1} + w_t$

If $t = 101$

$x_{101} = 40 + 0.6\, x_{100} + w_{101}$

$88 \pm 2(1.96)$

$x_{102} = 40 + 0.6\, x_{101} + w_{102}$

$92.8 \pm (1.96)(2.332)$