MA Models as an Infinite-Lag AR Models

There are two major approaches to MA models:

- An extension of a white-noise series.
- An Infinite order AR model with some restrictions.

The former is pretty straight-forward, since one way to think about a MA average model is that it is a sum of weighted white-noise. Of course, this only holds true when the residuals resemble white-noise.

In his short lecture, we are going to focus on the links between AR and MA models and how one can be expressed in terms of the other. In particular, we are going to focus on how a special case of the Autoregressive model can be expressed as a simple moving-average model.

$$AR(\infty) = MA(1)$$

Infinite-Lag AR Model
(with certain restrictions)

Simple MA Model

365√DataScience

$AR(\infty)$ Model

Let's start by elaborating what we mean by "infinite"-lag AR model: $AR(\infty)$

By "infinite"-lag AR model, we mean an autoregressive model, which takes the values of *infinitely* many previous periods. Of course, this is just a **theoretical** concept.

We cannot have an **infinite** data set, even when working with big data. Thus, no data set would ever satisfy the information criteria required for this kind of model. Hence, an infinite lag autoregressive model is not applicable in practice.

Despite this, let us write out what a modelling such a variable would look like. Since will be modelling returns in the start of this section, suppose that "r" is a variable, which follows this model. Then, the $AR(\infty)$ looks like so:

Parameter Coefficients

Residuals

$$r_t = c + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \cdots + \epsilon_t$$

Value of Interest

Constant

Past Values

365√DataScience

Parameter Restrictions

Having absolute freedom on infinitely many parameter coefficients is unreasonable, since we would have no idea of how the data behaves. Instead, we are going to try and make the model practical and assume all the ϕ coefficients obey certain restrictions.

A special case of this idea is when all the coefficients are negative and progressing geometrically. In other words, the model looks like so:

$$r_t = c - \theta_1 r_{t-1} - \theta_1^2 r_{t-2} - \theta_1^3 r_{t-3} - \cdots + \epsilon_t$$

Here, the coefficient θ_1(theta one) is essentially $-\phi_1$. In fact, a more general representation would suggest that $\boldsymbol{\phi_i = -\theta_1^i}$.

Since all models include a constant and a residual, even if they end up being insignificant, we can rewrite the equation on top. If we move all the AR parts of the model on the left, we are left with a different representation of the same model.

$$r_t + \theta_1 r_{t-1} + \theta_1^2 r_{t-2} + \theta_1^3 r_{t-3} + \cdots = c + \epsilon_t$$

Past and Present Values

By the definition of AR Models, we can express r_{t-1} in a similar way.

More precisely, if the model suggests that $r_t + \theta_1 r_{t-1} + \theta_1^2 r_{t-2} + \theta_1^3 r_{t-3} + \cdots = c + \epsilon_t$, then clearly we can conclude the following.

$$r_{t-1} + \theta_1 r_{t-2} + \theta_1^2 r_{t-3} + \theta_1^3 r_{t-4} + \cdots = c + \epsilon_{t-1}$$

Remember that, unlike the AR model, the MA one does include any values of past period, but only past errors. Thus, we need to get rid of all the r_{t-i} components. We are aware this sounds completely out of the blue, but please be patient a bit more since everything will fall into place in just a bit.

We can multiply both sides of the r_{t-1} equation to get this new one:

$$\theta_1 r_{t-1} + \theta_1^2 r_{t-2} + \theta_1^3 r_{t-3} + \theta_1^4 r_{t-4} + \cdots = \theta_1 c + \theta_1 \epsilon_{t-1}$$

Combining the Equations

Now, let's observe the r_t and the new r_{t-1} equations.

$$r_t + \theta_1 r_{t-1} + \ \theta_1^2 r_{t-2} + \theta_1^3 r_{t-3} + \cdots = c + \epsilon_t$$

$$\theta_1 r_{t-1} + \theta_1^2 r_{t-2} + \ \theta_1^3 r_{t-3} + \theta_1^4 r_{t-4} + \cdots = \theta_1 c + \theta_1 \epsilon_{t-1}$$

If you look closely, all the AR components (like $\theta_1 r_{t-1}$ and $\theta_1^2 r_{t-2}$) for the two equations are identical. Hence, we can subtract the bottom one from the top one.

The resulting equation leaves r_t only on the left and looks like so:

$$r_t = c + \epsilon_t - \ \theta_1 c - \ \theta_1 \epsilon_{t-1}$$

Rearranging the Model

We can obviously transform the equation to account for the parts using the constant c. Then, the model becomes:

$$r_t = c(1-\theta_1) + \epsilon_t - \theta_1\epsilon_{t-1}$$

Additionally, we can also rearrange the model so that we get the familiar look of a constant, some parameters and their coefficients and a residual.

$$r_t = c(1-\theta_1) - \theta_1\epsilon_{t-1} + \epsilon_t$$

Notice how the first part of the model ($c(1-\theta_1)$) is time-invariant. What we mean is that it does not rely on "t", so the value is constant in every time period. Hence, we can substitute it with some other constant variable. Thus, the new equation looks like:

$$r_t = c_0 - \theta_1\epsilon_{t-1} + \epsilon_t$$

MA(1) Model

Now let's quickly examine this new representation of the model.

$$r_t = c_0 - \theta_1 \epsilon_{t-1} + \epsilon_t$$

Value of Interest

Constant

Past Errors

Residuals

As you can see, this current model predicts current values solely based on the mistakes in our predictions one period ago. Hence, it contains a Moving-Average aspect for precisely 1 lag.

Additionally, since ϵ_t and ϵ_{t-1} are the residuals from an AR model, then we expect them to be white noise. Hence, why we stated that an MA model is also known as an extension of white noise. Since adding two WN series together $-\theta_1\epsilon_{t-1}$and ϵ_t, results in a WN series (mean and variance remain the same), then an MA model is essentially predicting what is known as white noise, with a drift, with the "drift" coming from the c_0 factor.

365√DataScience