ARMA Model Derivations

Random Walk Model

Consider a time series variable y_t that follows the process: $y_t = y_{t-1} + \varepsilon_t$, where ε_t is a White Noise error term:

$$y_t = y_{t-1} + \varepsilon_t$$

$$E(\varepsilon_t) = 0$$

$$Var(\varepsilon_t) = \sigma^2$$

$$Cov(\varepsilon_t, \varepsilon_{t-s}) = 0 \; for \; all \; s \neq 0$$

If we backdate the original equation one period, $y_{t-1} = y_{t-2} + \varepsilon_{t-1}$, and substitute into the original equation for y_{t-1}:

$$y_t = y_{t-2} + \varepsilon_t + \varepsilon_{t-1}$$

Repeated substitution gives:

$$y_t = \sum_{i=1}^{t} \varepsilon_i$$

And

$$E(y_t) = \sum_{i=1}^{t} E(\varepsilon_i) = 0$$

$$Var(y_t) = \sum_{i=1}^{t} Var(\varepsilon_i) = t\sigma^2$$

Because ε_t is independent White Noise.

$$Cov(y_t, y_{t-s}) = E(\varepsilon_t + \varepsilon_{t-1} + \cdots + \varepsilon_{t-s} + \cdots \varepsilon_1)(\varepsilon_{t-s} + \cdots + \varepsilon_1)$$

When we multiply the two parts of this expression, any term where the time periods do not match, e.g., $E(\varepsilon_{t-i}\varepsilon_{t-j})$ with $i \neq j$, goes to zero, again because ε is White Noise. The equation then reduces to:

$$Cov(y_t, y_{t-s}) = E(\varepsilon_{t-s}^2) + \cdots + E(\varepsilon_1^2) = (t-s)\sigma^2$$

Because the variance and autocovariances of y_t depend on t, this type of Random Walk variable is not stationary.

AR(1) Model

Now consider a simple AR(1) model: $y_t = \phi_1 y_{t-1} + \varepsilon_t$, where again, ε_t is a White Noise error term and ϕ_1 is the (constant) AR parameter.

Backdate this equation one period: $y_{t-1} = \phi_1 y_{t-2} + \varepsilon_{t-1}$ and substitute into the original equation:

$$y_t = \phi_1(\phi_1 y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$$

Expanding and rearranging:

$$y_t = \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 y_{t-2}$$

By repeated substitution we can *invert* the AR(1) form into an infinite MA form to give:

$$y_t = \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \phi_1^3 \varepsilon_{t-3} + \cdots$$

And:

$$E(y_t) = 0$$

$$Var(y_t) = \sigma^2 \sum_{i=0}^{\infty} \phi_1^{2i}$$

For y_t to be stationary, we require $Var(y_t) < \infty$. If $|\phi_1| < 1$, or equivalently, $\phi_1^2 < 1$, the infinite sum in the variance equation reduces to $\frac{1}{1-\phi_1^2}$, so:

$$Var(y_t) = \frac{\sigma^2}{1 - \phi_1^2}$$

Note that because $E(y_t) = 0, Var(y_t) = E(y_t^2)$.

For y_t to be stationary, we also require it autocovariances, $Cov(y_t, y_{t-s})$ to be independent of t. Because $E(y_t) = 0, Cov(y_t, y_{t-s}) = E(y_t y_{t-s})$. For the first autocovariance,

$$E(y_t, y_{t-1}) = E[(\phi_1 y_{t-1} + \varepsilon_t) y_{t-1}] = \phi_1 E(y_{t-1}^2) + E(y_{t-1}\varepsilon_t)$$

In the first portion: $\phi_1 E(y_{t-1}^2), E(y_{t-1}^2) = Var(y_t)$.

In the last portion:

$$E(y_{t-1}\varepsilon_t) = E[(\varepsilon_{t-1} + \phi_1 \varepsilon_{t-2} + \phi_1^2 \varepsilon_{t-3} + \phi_1^3 \varepsilon_{t-4} + \cdots)\varepsilon_t] = 0$$

Because all the $\varepsilon's$ are from different time periods. The first autocovariance then reduces to:

$$E(y_t, y_{t-1}) = \frac{\phi_1 \sigma^2}{1 - \phi_1^2}$$

For the second autocovariance, write $y_t = \phi_1^2 y_{t-2} + \phi_1 \varepsilon_{t-1} + \varepsilon_t$, then:

$$E(y_t y_{t-2}) = E[(\phi_1^2 y_{t-2} + \phi_1 \varepsilon_{t-1} + \varepsilon_t) y_{t-2}]$$

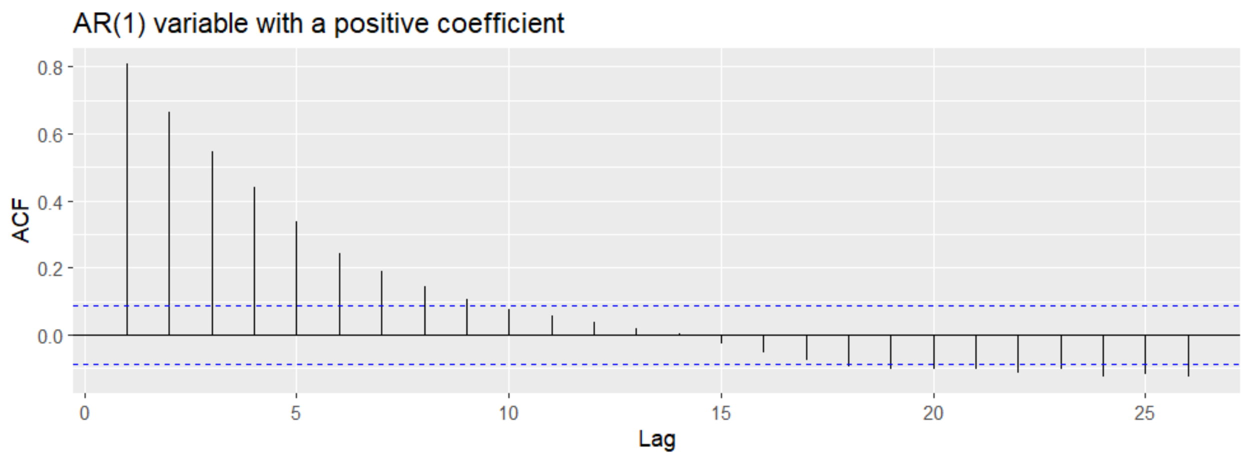
Which reduces to:

$$E(y_t, y_{t-2}) = \frac{\phi_1^2 \sigma^2}{1 - \phi_1^2}$$

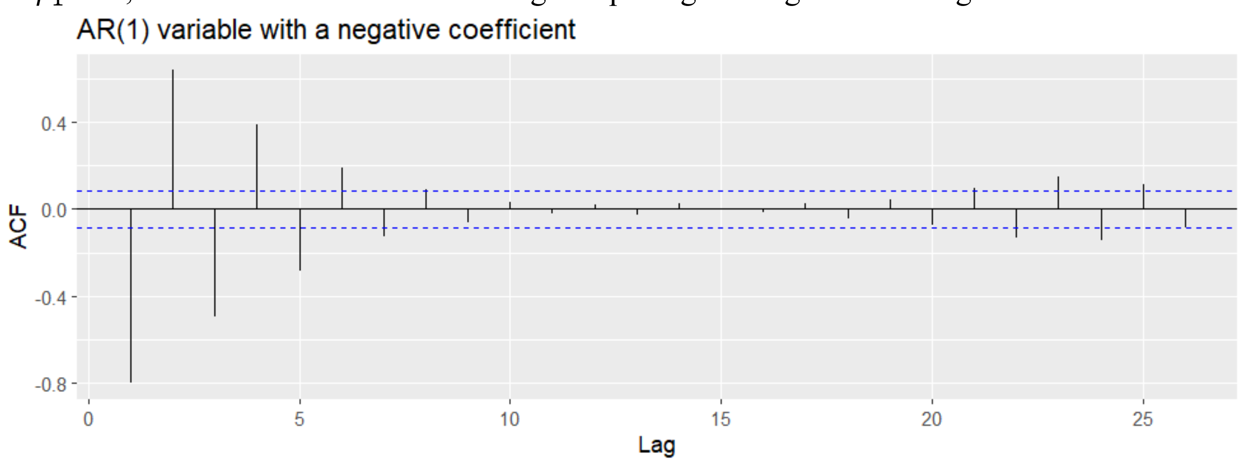
Generally,

$$E(y_t, y_{t-s}) = \frac{\phi_1^s \sigma^2}{1 - \phi_1^2}$$

Thus, the autocovariance function (ACF) for the simple AR(1) model does not depend on t, only on the *level of displacement*, s. Because we required $|\phi_1| < 1$, the ACF will *dampen* as we raise ϕ_1 to higher and higher powers. If $\phi_1 > 0$, the ACF will exhibit one-sided dampening from the positive side:



If $\phi_1 < 0$, the ACF will exhibit oscillating dampening starting from the negative side.



For the Partial Autocorrelation Function (PACF), the first autocorrelation is simply ϕ_1. After controlling for the effect of y_{t-1}, all other partial autocorrelations (e.g. t-2, t-3, etc.) should drop to effectively zero.

MA(1) Model

Now consider a simple MA(1) model: $y_t = \theta_1 \varepsilon_{t-1} + \varepsilon_t$, giving

$$E(y_t) = 0$$

$$Var(y_t) = E(y_t^2) = \sigma^2(1 + \theta^2)$$

To get the ACF, we need $E(y_t y_{t-s})$. For the first autocovariance:

$$E(y_t y_{t-1}) = E[(\theta_1 \varepsilon_{t-1} + \varepsilon_t)(\theta_1 \varepsilon_{t-2} + \varepsilon_{t-1})] = \sigma^2 \theta_1$$

For the second autocovariance (and all others):

$$E(y_t y_{t-2}) = E[(\theta_1 \varepsilon_{t-1} + \varepsilon_t)(\theta_1 \varepsilon_{t-3} + \varepsilon_{t-2})] = 0$$

Generally,

$$E(y_t, y_{t-s}) = \begin{Bmatrix} \sigma^2 \theta_1 & s = 1 \\ 0 & s > 1 \end{Bmatrix}$$

So, provided θ_1 is finite, the MA process is stationary and the ACF drops off abruptly.

To get the PACF, backdate and rearrange the original equation to get:

$$\varepsilon_{t-1} = y_{t-1} - \theta_1 \varepsilon_{t-2}$$

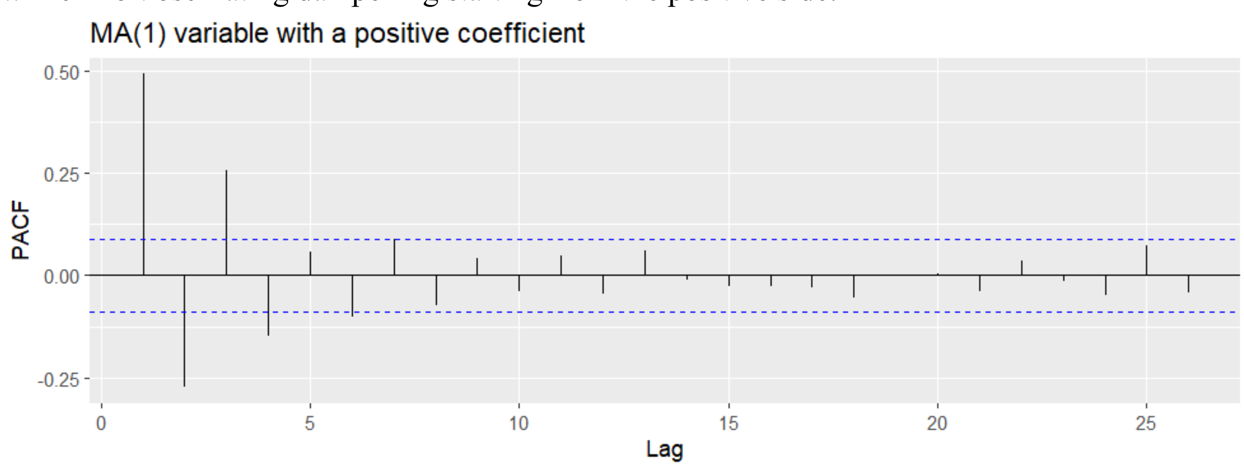
Substituting for ε_{t-1} in the original equation:

$$y_t = \theta_1 y_{t-1} - \theta_1^2 \varepsilon_{t-2} + \varepsilon_t$$

By repeated substitution to eliminate all the lagged ε terms, we can invert the MA(1) form into an infinite AR form to give:

$$y_t = \theta_1 y_{t-1} - \theta_1^2 y_{t-2} + \theta_1^3 y_{t-3} - \theta_1^4 y_{t-4} \dots + \varepsilon_t$$

In order for this to be stable (whereby events in the distant past don't have an infinite impact on what we observe today), we require $|\theta_1| < 1$. Thus, if $\theta_1 > 0$, the PACF for an MA(1) process will exhibit oscillating dampening starting from the positive side:



And if $\theta_1 < 0$, the PACF will exhibit one sided dampening from the negative side: