A Correlogram tale

In data analysis, we usually start with the descriptive statistical properties of the sample data (e.g. mean, standard deviation, skew, kurtosis, empirical distribution, etc.). These calculations are certainly useful, but they do not account for the order of the observations in the sample data.

Time series analysis demands that we pay attention to order, and thus requires a different type of descriptive statistics: time series descriptive statistics, or simply correlogram analysis. The correlogram analysis examines the time-spatial dependency within the sample data, and focuses on the empirical auto-covariance, auto-correlation, and related statistical tests. Finally, the correlogram is a cornerstone for identifying the model and model order(s).

What does a plot for auto correlation (AFC) and/or partial auto-correlation (PACF) tell us about the underlying process dynamics?

This tutorial is a bit more theoretical than prior tutorials in the same series, but we will do our best to drive the intuitions home for you.

Background

First, we'll start with a definition for the auto-correlation function, simplify it, and investigate the theoretical ACF for an ARMA-type of process.

Auto-correlation function (ACF)

By definition, the auto correlation for lag k is expressed as follows:

$$ACF(k) = \rho_k = \frac{\gamma_k}{\gamma_o}$$

$$\gamma_k = E\left[(x_{t-k} - \mu_{t-k})(x_t - \mu_t)\right] = E\left[x_{t-k}x_t\right] - \mu^2$$

$$\gamma_o = E\left[(x_t - \mu)^2\right] = E\left[x_t^2\right] - \mu^2 = \sigma^2$$

Where:

ρ_k = auto-correlation function for lag k

γ_k = auto-covariance for lag k

$\gamma_o = \sigma^2$ = time series variance **(un-conditional)**

$\mu_{t-k} = \mu_t = \mu$ = stationary time series unconditional mean

SPIDERFINANCIAL
www.spiderfinancial.com

Phone: 1-888-427-9486
1-312-324-0367
Fax: 1-312-238-9092
info@spiderfinancial.com

Furthermore, let's assume that $\{x_t\}$ is generated from a weak-stationary process with a zero mean (i.e. $\mu = 0$).

$$\rho_k = \rho_{-k} = \frac{E[x_t x_{t-k}]}{E[x_t^2]} = \frac{E[x_t x_{t-k}]}{\sigma^2}$$

For finite sample data, the empirical auto-correlation is expressed as follows:

$$\hat{\rho}_k = \frac{\sum_{t=k}^{N}(x_{t-k} - \overline{x})(x_t - \overline{x})}{\sum_{t=k}^{N}(x_t - \overline{x})^2} = \frac{\sum_{t=k}^{N}x_{t-k}x_t - (N-k)\overline{x}^2}{\sum_{t=k}^{N}x_t^2 - (N-k)\overline{x}^2}$$

And

$$\hat{\rho}_k \sim N(0, \sigma_{\rho_k}^2)$$

$$\sigma_{\rho_k}^2 = \frac{1 + \sum_{l=1}^{k-1}\hat{\rho}_l^2}{N}$$

Where:

N =number of non-missing observations in the time series

$\overline{x}$ =the time series sample average

NOTE: Using the sample auto-correlation estimate ($\hat{\rho}_k$) and error ($\sigma_{\rho_k}^2$), we can easily perform a one-sample mean test to examine its statistical significance, but what about a joint test for a set of auto-correlation factors? For that, we use the Ljung-Box Test (white-noise test).

$$H_o : \rho_1 = \rho_2 = \rho_3 = ... = \rho_m = 0$$
$$H_1 : \exists \rho_{1 \le k \le m} \ne 0$$

The Ljung-Box test is discussed in great details as part of our "White-noise tutorial." Please refer to that paper for more details.

Example 1 - MA(q) model
Let's start with a simple moving average model with order q.

$$x_t - \mu = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + ... + \theta_q a_{t-q}$$
$$x_t - \mu = (1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q)a_t$$

Where

$$a_t \sim i.i.d \sim N(0, \sigma^2)$$

$$E\left[a_t \times a_{t-j}\right] = \begin{cases} 0 & j \neq 0 \\ \sigma^2 & j = 0 \end{cases}$$

Now, let's compute the ACF for different lags.

$$\gamma_o = E\left[x_t x_t\right] = E\left[a_t^2 + \theta_1^2 a_{t-1}^2 + \theta_2^2 a_{t-2}^2 + \theta_3^2 a_{t-3}^2 + \ldots + \theta_q^2 a_{t-q}^2\right] = \sigma^2\left(1 + \sum_{j=1}^{q} \theta_j^2\right)$$

$$\gamma_1 = E\left[x_{t-1} x_t\right] = E\left[\theta_1 a_{t-1}^2 + \theta_1 \theta_2 a_{t-2}^2 + \theta_2 \theta_3 a_{t-3}^2 + \ldots + \theta_{q-1} \theta_q a_{t-q-1}^2\right] = \sigma^2\left(\theta_1 + \sum_{j=1}^{q-1} \theta_j \theta_{j+1}\right)$$

$$\gamma_2 = E\left[x_{t-2} x_t\right] = E\left[\theta_2 a_{t-2}^2 + \theta_1 \theta_3 a_{t-3}^2 + \theta_2 \theta_4 a_{t-4}^2 + \ldots + \theta_{q-2} \theta_q a_{t-q-2}^2\right] = \sigma^2\left(\theta_2 + \sum_{j=1}^{q-2} \theta_j \theta_{j+2}\right)$$

$$\gamma_{k<q} = E\left[x_{t-k} x_t\right] = E\left[\theta_k a_{t-k}^2 + \theta_1 \theta_{k+1} a_{t-k-1}^2 + \theta_2 \theta_{k+2} a_{t-k-2}^2 + \ldots + \theta_{q-k} \theta_q a_{t-q-k}^2\right] = \sigma^2\left(\theta_3 + \sum_{j=1}^{q-k} \theta_j \theta_{j+k}\right)$$

$$\gamma_{k>q} = E\left[x_{t-k} x_t\right] = 0$$

$$\rho_k = \begin{cases} \dfrac{\theta_3 + \sum_{j=1}^{q-k} \theta_j \theta_{j+k}}{1 + \sum_{j=1}^{q} \theta_j^2} & k \leq q \\ 0 & k \succ q \end{cases}$$

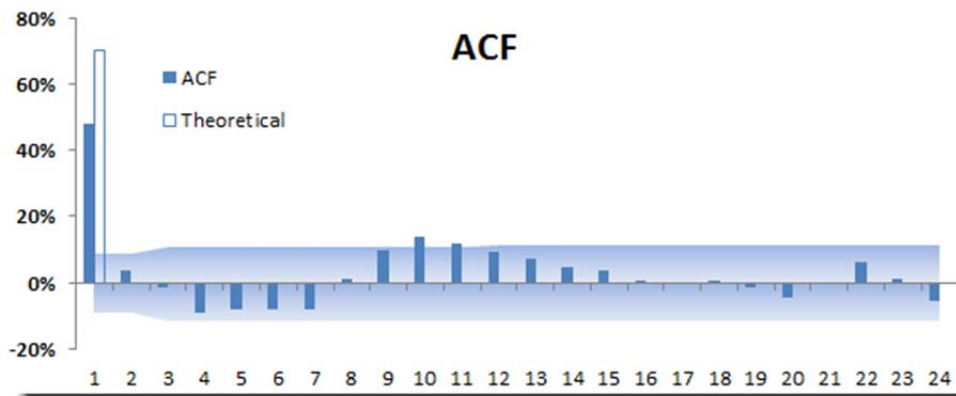The ACF plot for an MA(q) process is non-zero for the first q lags.
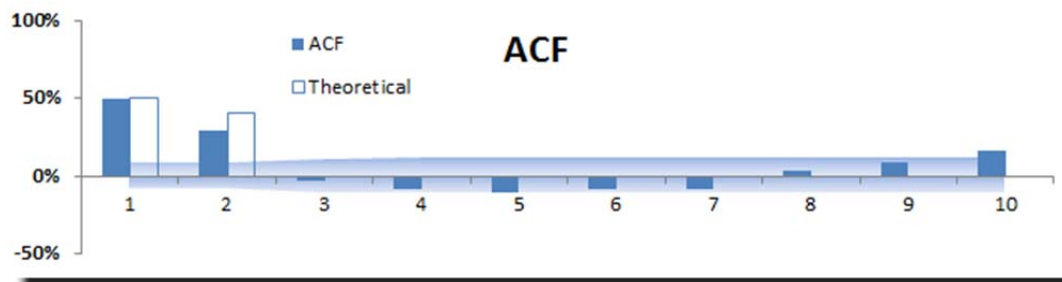

Figure 1: ACF Plot for a simulated MA(1) process

SPIDERFINANCIAL
www.spiderfinancial.com

Phone: 1-888-427-9486
1-312-324-0367
Fax: 1-312-238-9092
info@spiderfinancial.com

Figure 2: ACF plot for a simulated MA(2) process

INTUITION ONE:

1. The MA has a **finite memory** of size q
2. The ACF plot **shows the memory size requirement** of the model
3. An **ARMA model with finite memory** can be fully described using an MA type of model. The coefficients' values of the MA model are the values of the auto-correlation function.

Example 2 - AR(1) model

Next, let's look at a simple auto-regressive (AR) model of order 1.

$$x_t = \phi_o + \phi_1 x_{t-1} + a_t$$
$$(1 - \phi_1 L)x_t = \phi_o + a_t$$
$$(1 - \phi_1 L)(x_t - \mu) = a_t$$

Where

$$\mu = \frac{\phi_o}{1 - \phi_1} = \text{long-run process (un-conditional) average}$$

$$a_t \sim i.i.d \sim N(0, \sigma^2)$$

Let's compute the auto-correlation function of an AR(1) process:

$$(x_t - \mu) = y_t = \frac{a_t}{(1 - \phi_1 L)}$$

Assuming $\|\phi_1\| < 1$, then AR(1) can be represented as an **infinite** MA model as below:

SPIDERFINANCIAL
www.spiderfinancial.com

Phone: 1-888-427-9486
1-312-324-0367
Fax: 1-312-238-9092
info@spiderfinancial.com

$$(x_t - \mu) = y_t = \frac{a_t}{(1 - \phi_1 L)} = (1 + \sum_{k=1}^{\infty} \phi_1^k L^k) a_t$$

$$\gamma_o = E\left[x_t^2 \right] = \left(1 + \sum_{k=1}^{\infty} \phi_1^{2k} \right) \sigma_a^2 = \frac{\sigma_a^2}{1 - \phi_1^2}$$

$$\rho_k = \frac{\phi_1^k \left(1 + \sum_{j=1}^{\infty} \phi_1^{2j} \right)}{1 + \sum_{j=1}^{\infty} \phi_1^{2j}} = \phi_1^k$$

The ACF plot for an AR type of process is infinite, but the ρ_k decays exponentially.
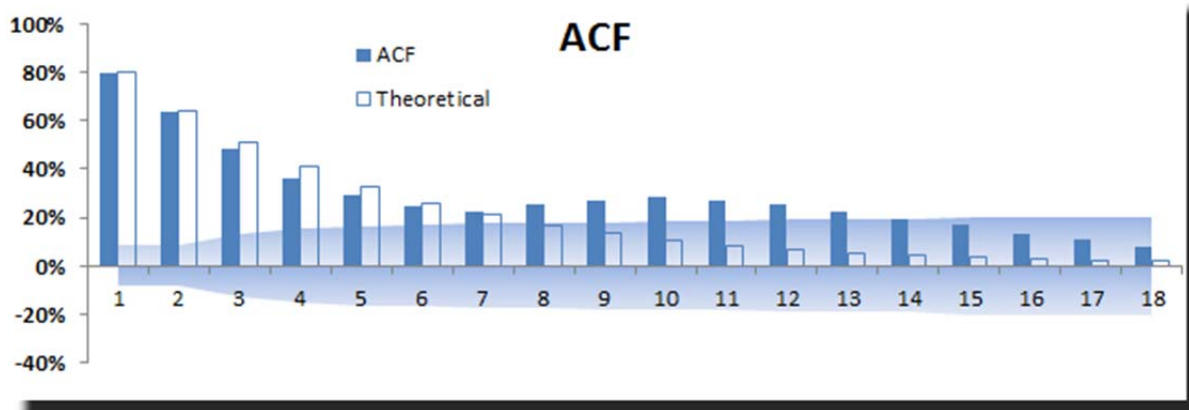

Figure 3: ACF plot for a simulated AR(1) process

Example 3 - AR(p) model

Now, let's get a bit more ambitious and look for an AR model with order p.

$$x_t = \phi_o + \phi_1 x_{t-1} + \phi_2 x_{t-2} + ... + \phi_p x_{t-p} + a_t$$

$$(1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p) x_t = \phi_o + a_t$$

$$(1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p)(x_t - \mu) = a_t$$

$$(1 - r_1 L)(1 - r_2 L)...(1 - r_p L)(x_t - \mu) = a_t$$

Where

$$\mu = \frac{\phi_o}{1 - \phi_1 - \phi_2 - ... - \phi_p} = \text{long-run unconditional process mean}$$

$r_1, r_2, ..., r_p$ = characteristic roots of the AR(p) model

SPIDERFINANCIAL
www.spiderfinancial.com

Phone: 1-888-427-9486
       1-312-324-0367
Fax: 1-312-238-9092
info@spiderfinancial.com

Let's compute the auto-correlation function of an AR(p) process. Using partial-fraction decomposition, we break an AR(p) process into a set of p AR(1) process.

$$(x_t - \mu) = y_t = \frac{a_t}{(1 - r_1 L)(1 - r_2 L)...(1 - r_p L)} = (\frac{\alpha_1}{(1 - r_1 L)} + \frac{\alpha_2}{(1 - r_2 L)} + ... + \frac{\alpha_p}{(1 - r_p L)})a_t$$

Let's assume all characteristic roots fall outside the unit circle, and therefore the $\|r_k\| \prec 1$

$$(x_t - \mu) = y_t = \left(1 + \alpha_1 \sum_{k=1}^{\infty} r_1^k L^k + \alpha_2 \sum_{k=1}^{\infty} r_2^k L^k + ... + \alpha_p \sum_{k=1}^{\infty} r_p^k L^k\right)a_t$$

$$(x_t - \mu) = y_t = \left(1 + \sum_{k=1}^{\infty} \psi_k L^k\right)a_t$$

$$\psi_k = \sum_{j=1}^{p} \alpha_j r_j^k$$

$$\gamma_o = \left(1 + \sum_{j=1}^{\infty} \psi_j^2\right)\sigma_a^2$$

$$\rho_k = \frac{\gamma_k}{\gamma_o} = \frac{\psi_k + \sum_{j=1}^{\infty} \psi_j \psi_{j+k}}{1 + \sum_{j=1}^{\infty} \psi_j^2}$$

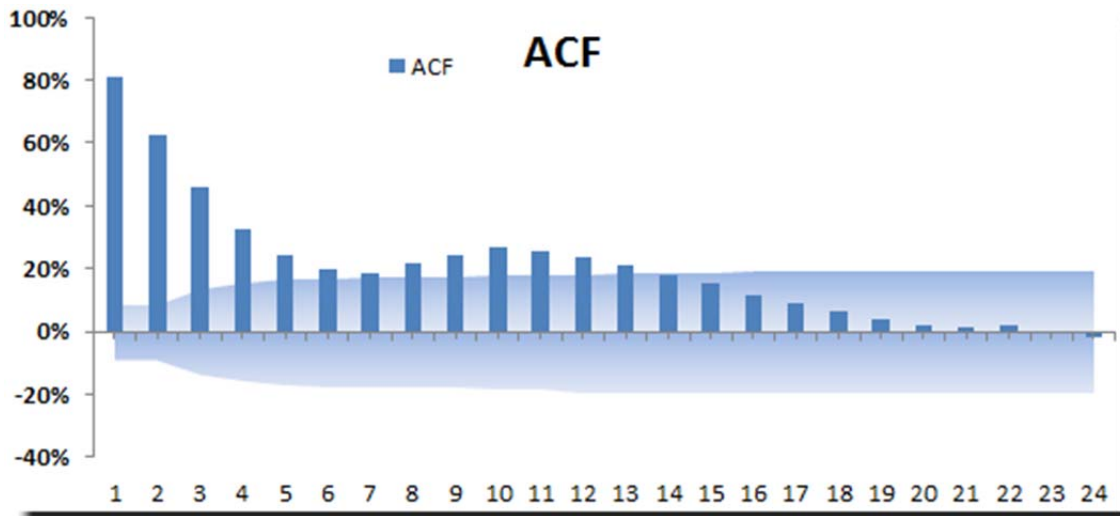This ACF plot is also infinite, but the actual shape can follow different patterns.

Figure 4: ACF plot for a simulated AR(2) process

INTUITION TWO:

SPIDERFINANCIAL
www.spiderfinancial.com

Phone: 1-888-427-9486
1-312-324-0367
Fax: 1-312-238-9092
info@spiderfinancial.com

1. An AR process can be represented by an infinite MA process
2. The AR has **infinite memory,** but the effect diminishes over time
3. Exponential smoothing functions are special cases of an AR process, and they also possess infinite memory

Example 4 - ARMA (p,q) model

By now, we see what the ACF plot of a pure MA and AR process looks like, but what about a mixture of the two models?

Question: why do we need to consider a mixture model like ARMA, since we can represent any model as an MA or an AR model? **Answer:** we are trying to reduce the memory requirement and the complexity of the process by super-imposing the two models.

$$x_t = \phi_o + \phi_1 x_{t-1} + \phi_2 x_{t-2} + ... + \phi_p x_{t-p} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + ... + \theta_q a_{t-q}$$

$$(1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p) x_t = \phi_o + (1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q) a_t$$

$$(1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p)(x_t - \mu) = (1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q) a_t$$

$$(1 - r_1 L)(1 - r_2 L)...(1 - r_p L)(x_t - \mu) = (1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q) a_t$$

$$(x_t - \mu) = y_t = \frac{(1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q)}{(1 - r_1 L)(1 - r_2 L)...(1 - r_p L)} a_t$$

$$y_t = \left(\frac{\alpha_1}{(1 - r_1 L)} + \frac{\alpha_2}{(1 - r_2 L)} + ... + \frac{\alpha_p}{(1 - r_p L)} \right) \times (1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q) a_t$$

$$y_t = \left(1 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + ... \right) \times (1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q) a_t$$

$$y_t = (1 + \omega_1 L + \omega_2 L^2 + ... + \omega_q L^q + ...) a_t$$

Where

$$\mu = \frac{\phi_o}{1 - \phi_1 - \phi_2 - ... - \phi_p} = \text{long-run unconditional process mean (from the AR component)}$$

$r_1, r_2, ..., r_p$ = inverse of the characteristic roots of the AR(p) model

$$\psi_k = \sum_{j=1}^{p} \alpha_j r_j^k$$

SPIDERFINANCIAL
www.spiderfinancial.com

Phone: 1-888-427-9486
1-312-324-0367
Fax: 1-312-238-9092
info@spiderfinancial.com

$$\omega_1 = \theta_1 + \psi_1$$

$$\omega_2 = \theta_2 + \theta_1 \times \psi_1 + \psi_2$$

$$\omega_3 = \theta_3 + \theta_2 \times \psi_1 + \theta_1 \times \psi_2 + \psi_3$$

$$\omega_q = \theta_q + \sum_{j=1}^{q-1} \theta_{q-j} \times \psi_j + \psi_q$$

$$\omega_{q+k} = \sum_{j=0}^{q-1} \theta_{q-j} \times \psi_{j+k} + \psi_{q+k}$$

$$\gamma_o = \left(1 + \sum_{j=1}^{\infty} \omega_j^2\right)\sigma_a^2$$

Using the MA(q) auto-correlation formula, we can compute the ARMA(p,q) auto-correlation functions for their MA representation.

This is getting intense! Some of you might be wondering why we haven't used VAR or a state space representation to simplify the notations. I made a point to stay in the time domain, and avoided any new ideas or math tricks as they would not serve our intentions here: Implying the exact AR/MA order using the ACF values by themselves, which is anything but precise.

> **(1) Intuition:** The ACF values can be thought of as the coefficient values of the equivalent MA model.
>
> **(2) Intuition:** The conditional variance has no barrier (effect) on the auto-correlation calculations.
>
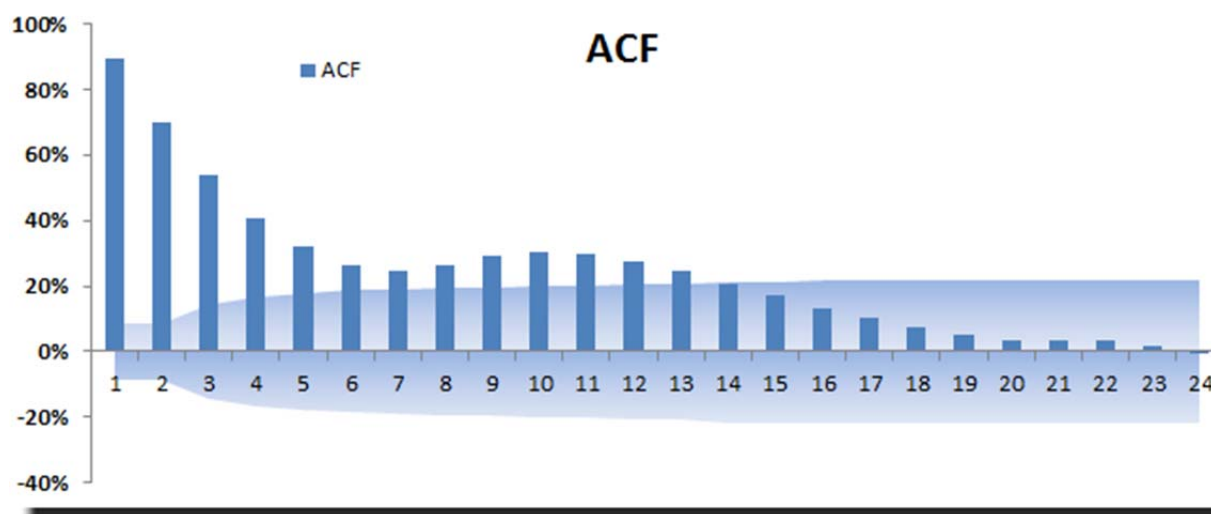> **(3) Intuition** : The long-run mean (μ) also does not have any barrier (effect) on the auto-correlations.

Figure 5: ACF plot for simulated ARMA(1,1) process

SPIDERFINANCIAL
www.spiderfinancial.com

Phone: 1-888-427-9486
1-312-324-0367
Fax: 1-312-238-9092
info@spiderfinancial.com

Partial Auto-correlation function (PACF)

By now, we have seen that identifying the model order (MA or the AR) is non-trivial for non-simple cases, so we need another tool – partial auto-correlation function (PACF).

The partial autocorrelation function (PACF) plays an important role in data analysis aimed at identifying the extent of the lag in an autoregressive model. The use of this function was introduced as part of the Box-Jenkins approach to time series modeling, whereby one could determine the appropriate lags p in an AR(p) model or in an extended ARIMA(p,d,q) model by plotting the partial auto correlation functions.

Simply put, the PACF for lag k is the regression coefficient for the kth term, as shown below:

$$x_t = b_{1,1}x_{t-1} + \varepsilon_{1,t}$$

$$x_t = b_{1,2}x_{t-1} + b_{2,2}x_{t-2} + \varepsilon_{2,t}$$

$$x_t = b_{1,3}x_{t-1} + b_{2,3}x_{t-2} + b_{3,3}x_{t-3} + \varepsilon_{3,t}$$

$$...$$

$$x_t = b_{1,k}x_{t-1} + b_{2,k}x_{t-2} + ... + b_{k,k}x_{t-k} + \varepsilon_{k,t}$$

The PACF assumes the underlying model is an AR(k) and uses multiple regressions to compute the last regression coefficient.

Please note that $PACF(k) = b_{k,k}$, and that PACF $b_{k,k} \neq b_{k,k+1} \neq b_{k,k+2} \neq b_{k,k+2} \neq b_{k,k+2}...$

> **Quick intuition:** the PACF values can be thought of (roughly speaking) as the coefficient values of the equivalent AR model.

How is the PACF helpful to us? Assuming we have an AR(p) process, then the PACF will have significant values for <u>the first p lags, and will drop to zero afterwards.</u>

What about the MA process? The MA process has non-zero PACF values for a (theoretically) infinite number of lags.

Example 4: MA (1)

$$x_t - \mu = y_t = (1 + \theta L)a_t$$

$$\frac{y_t}{(1 + \theta L)} = a_t$$

Assuming $\|\theta\| \prec 1$ (invertible), the MA process can be represented as infinite AR:

SPIDERFINANCIAL
www.spiderfinancial.com

Phone: 1-888-427-9486
          1-312-324-0367
Fax: 1-312-238-9092
info@spiderfinancial.com

$$(1 - \theta L + \theta^2 L^2 - \ldots)\, y_t = a_t$$

$$y_t = a_t + \theta\, y_{t-1} - \theta^2 y_{t-2} + \theta^3 y_{t-3} + \ldots + (-1)^{n+1} \theta^n y_{t-n} + \ldots$$

The PACF of an MA is expected to have significant values to a large number of lags.

How about MA(q)? The same conclusion follows through here, so that the MA(q) can be inverted to an infinite AR sequence.
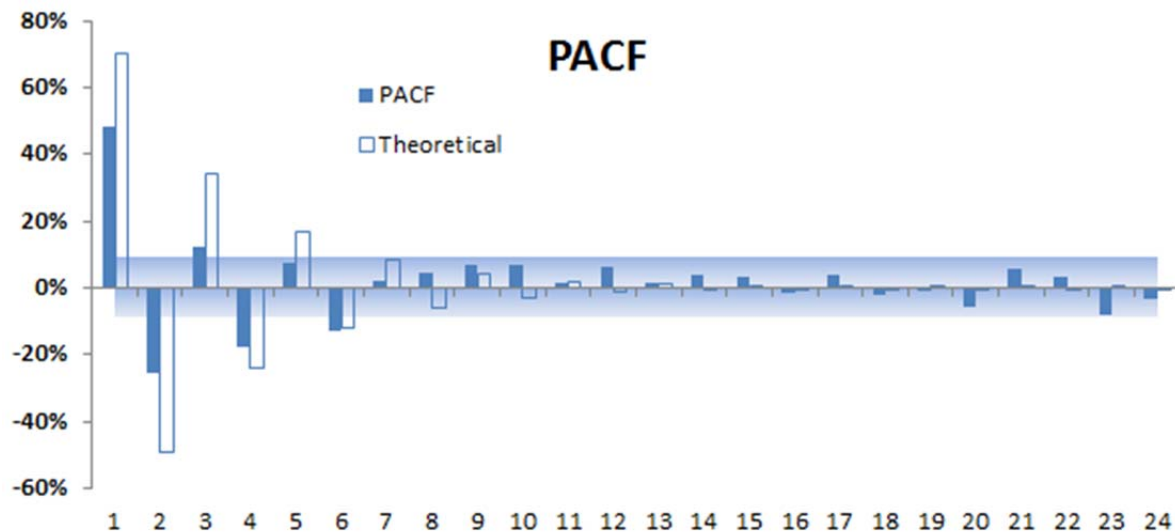

Figure 6: PACF Plot for simulated MA(1) process

Conclusion

In this tutorial, we discussed the auto and partial auto correlation functions and their role in identifying the order of the underlying ARMA process: ACF for the MA order and the PACF for the AR order.

Furthermore, we showed how more than one model can be used to generate the same ACF (and PACF) plots (i.e. correlogram).

With the exception of trivial cases, the process of identifying the proper model order is never clear-cut. The process demands that we entertain several candidate models, fit their parameters, validate the assumptions, compare their fit and finally select the best one.