1. Introduction

Data of time series are prevalent. Modeling time series is therefore an enterprise with many benefits.

We start with, surprisingly, noise, which in principle and in practice, cannot be modeled except for its statistical distribution.

ARMA models are generally made up of an autoregressive component with p lags, a moving average component with q past values of noise, as well the current noise u_t. Mathematically, ARMA(p, q) is expressed as

$$Y_t = \theta + \lambda_1 Y_{t-1} + \cdots + \lambda_p Y_{t-p} + u_t + \alpha_1 u_{t-1} + \cdots + \alpha_q u_{t-q}.$$

The back-shift operator, also known as the lag operator, is to change the discrete time back by one step

$$Bu_t := u_{t-1}.$$

2. AR(1) and Stationarity

To explain the essence of a time series, we set $\theta = 0$ for convenience. For a start, we look at AR(1):

$$Y_t = \lambda Y_{t-1} + u_t,$$

which can also be expressed as

$$(1 - \lambda B)Y_t = u_t,$$

Interestingly, we can write it as

$$Y_t = (1 - \lambda B)^{-1} u_t.$$

Knowing that if $|x| < 1$, then we can have $(1 - x)^{-1} = 1 + x + x^2 + \cdots$. It follows that if $|\lambda| < 1$,

$$Y_t = \big(1 + \lambda B + \lambda^2 B^2 + \cdots\big) u_t = \sum_{j=0}^{\infty} \lambda^j u_{t-j}.$$

Since Y_t is a sum of noise at different time $t - j \leqq t$, which is stationary, and therefore Y_t is also stationary. That' why $|\lambda| < 1$ is the criterion for AR(1) to be stationary.

We may craft the determination of the stationarity of AR(1) as a characteristic equation.

$$1 - \lambda x^\star = 0.$$

The root of this linear equation is $x^\star = \dfrac{1}{\lambda}$. By insisting that the solution $|x^\star| > 1$, we will ensure $|\lambda| < 1$.

3. AR(2) and Stationarity

In the same vein, AR(2) can be written as

$$(1-\lambda_1 B-\lambda_2 B^2)Y_t=u_t$$

Being quadratic, we can factorize it as

$$1-\lambda_1 B-\lambda_2 B^2=\big(1-\phi_a B\big)\big(1-\phi_b B\big).$$

To the right-hand side, we need to set ϕ_a and ϕ_b in such a way that

$$\phi_a+\phi_b=\lambda_1, \quad \text{and} \quad \phi_a\phi_b=-\lambda_2 \tag{1}$$

By inverting, we obtain

$$Y_t=\left(\frac{1}{\big(1-\phi_a B\big)\big(1-\phi_b B\big)}\right)u_t.$$

Now, by applying the elementary technique of common fraction, we obtain

$$\frac{1}{\big(1-\phi_a B\big)\big(1-\phi_b B\big)}=\left(\frac{\phi_a}{\phi_a-\phi_b}\right)\frac{1}{1-\phi_a B}-\left(\frac{\phi_b}{\phi_a-\phi_b}\right)\frac{1}{1-\phi_b B}.$$

It follows that

$$\begin{aligned}Y_t&=\left(\frac{\phi_a}{\phi_a-\phi_b}\right)\sum_{j=0}^{\infty}\phi_a^j u_{t-j}-\left(\frac{\phi_b}{\phi_a-\phi_b}\right)\sum_{j=0}^{\infty}\phi_b^j u_{t-j}\\&=\sum_{j=0}^{\infty}\left(\left(\frac{\phi_a}{\phi_a-\phi_b}\right)\phi_a^j-\left(\frac{\phi_b}{\phi_a-\phi_b}\right)\phi_b^j\right)u_{t-j}.\end{aligned}$$

Again, we find that Y_t composes purely of noise, which indicates that Y_t is stationary, if

$$|\phi_a|<1 \qquad \text{and} \qquad |\phi_b|<1.$$

This is because, as before, for $\frac{1}{(1-\phi_i B)}$, $i=a,b$, to be expendable as the infinite series $1+\phi_i B+\phi_i^2B^2+\cdots$, $|\phi_i|$ must be less than 1.

From (1), simple algebraic operations lead to a quadratic equation for $i=a,b$:

$$\phi_i^2-\lambda_1\phi_i-\lambda_2=0.$$

The solution is

$$\phi_i=\frac{\lambda_1\pm\sqrt{\lambda_1^2+4\lambda_2}}{2}.$$

Notice that (1) is symmetric between a and b. Therefore, without loss of generality, we let

$$\phi_a=\frac{\lambda_1+\sqrt{\lambda_1^2+4\lambda_2}}{2},$$

and

$$\phi_b=\frac{\lambda_1-\sqrt{\lambda_1^2+4\lambda_2}}{2}.$$

Now, we re-express these relations between ϕ_i $(i = a, b)$ and $\lambda_j, j = 1, 2$ as

$$2\phi_a + \lambda_1 = \sqrt{\lambda_1^2 + 4\lambda_2}$$
$$2\phi_b + \lambda_1 = -\sqrt{\lambda_1^2 + 4\lambda_2}.$$

Upon subtraction, we obtain

$$\phi_a - \phi_b = \sqrt{\lambda_1^2 + 4\lambda_2}. \tag{2}$$

Now, the characteristic equation for AR(2) is

$$\lambda_2 x^2 + \lambda_1 x - 1 = 0. \tag{3}$$

The solutions of this characteristic equation (3) are

$$x_a = \frac{-\lambda_1 + \sqrt{\lambda_1^2 + 4\lambda_2}}{2\lambda_2} \quad \text{and} \quad x_b = \frac{-\lambda_1 - \sqrt{\lambda_1^2 + 4\lambda_2}}{2\lambda_2}.$$

We see that because of (1) and (2), it follows that

$$x_a = \frac{-\phi_a - \phi_b + (\phi_a - \phi_b)}{2(-\phi_a\phi_b)} = \frac{1}{\phi_a}.$$

Likewise

$$x_b = \frac{-\phi_a - \phi_b - (\phi_a - \phi_b)}{2(-\phi_a\phi_b)} = \frac{1}{\phi_b}.$$

Since $|\phi_i| < 1$, $i = a, b$, the solutions x_i of the characteristic equation must be strictly larger than 1.

The same technique of common fraction applies for AR(p). That is, the characteristic equation for AR(p) is the polynomial equation

$$\lambda_p x^p + \lambda_{p-1} x^{p-1} + \cdots + \lambda_1 x - 1 = 0.$$

In principle, there are p roots, some or all of which can be complex numbers. For AR(p) to be stationary, the absolute value (or length) of every root must be larger than 1.

4. Duality of ACF and PACF Diagrams

AR and MA are dual to each other with respect to the ACF vis-à-vis PACF diagrams. We simulate the AR(2) process with $\lambda_1 = 0.5$ for Y_{t-1} and $\lambda_2 = 0.3$ for Y_{t-2}, with σ_u being 10%. The ACF and PACF of AR(2) are plotted in Figure 1. Clearly, the ACF diagram has exponentially decaying autocorrelation coefficients. By contrast, in the PACF diagram, there are two distinctive PACF coefficients that are statistically significant.

Next, we simulate an MA(2) process with $\alpha_1 = -0.7$ for u_{t-1} and $\alpha_2 = -0.25$ for u_{t-2}, again with σ_u being 10%. This time round, the ACF has two distinct coefficients that are statistically significant. On the other hand, in the PACF diagram, this invertible MA(2) process displays an exponentially decaying characteristics.

Figure 1: AR(2)

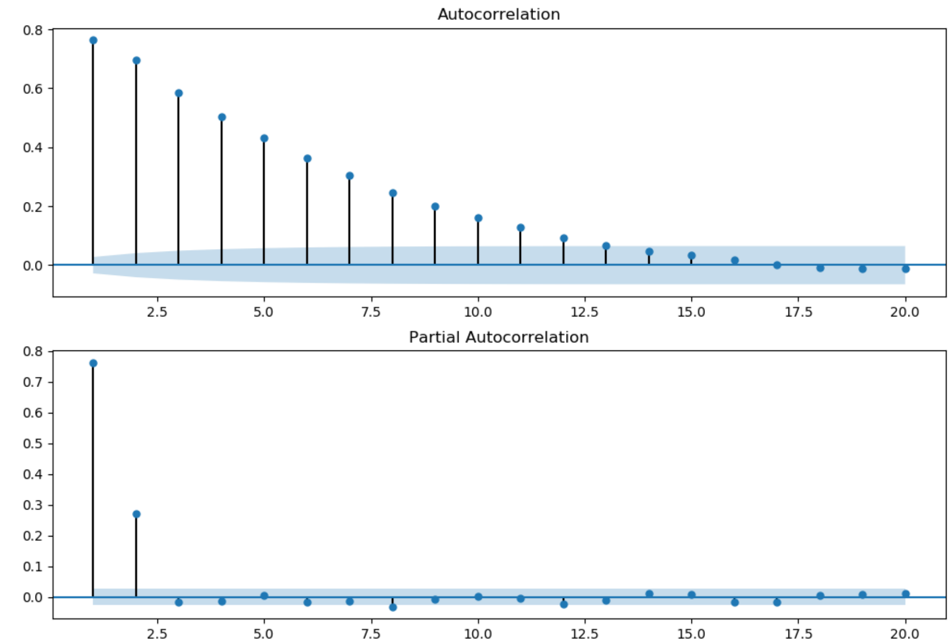

Figure 2: MA(2)