ARIMA models

1. Definitions

1.1. AR models. An AR(p) (AutoRegressive of order p) model is a discrete time linear equations with noise, of the form

$$X_t = \alpha_1 X_{t-1} + ... + \alpha_p X_{t-p} + \varepsilon_t.$$

Here p is the order, $\alpha_1, ..., \alpha_p$ are the parameters or coefficients (real numbers), ε_t is an error term, usually a white noise with intensity σ^2. The model is considered either on integers $t \in \mathbb{Z}$, thus without initial conditions, or on the non-negative integers $t \in \mathbb{N}$. In this case, the relation above starts from $t = p$ and some initial condition $X_0, ..., X_{p-1}$ must be specified.

Example 1. *We have seen above the simplest case of an AR(1) model*

$$X_t = \alpha X_{t-1} + \varepsilon_t.$$

With $|\alpha| < 1$ and $Var[X_t] = \frac{\sigma^2}{1-\alpha^2}$, it is a wide sense stationary process (in fact strict sense since it is gaussian, see also below). The autocorrelation coefficient decays exponentially:

$$\rho(n) = \alpha^n.$$

Even if the general formula is not so simple, one can prove a similar result for any AR model.

In order to model more general situations, it may be convenient to introduce models with non-zero average, namely of the form

$$(X_t - \mu) = \alpha_1 (X_{t-1} - \mu) + ... + \alpha_p (X_{t-p} - \mu) + \varepsilon_t.$$

When $\mu = 0$, if we take an initial condition having zero average (this is needed if we want stationarity), then $E[X_t] = 0$ for all t. We may escape this restriction by taking $\mu \neq 0$. The new process $Z_t = X_t - \mu$ has zero average and satisfies the usual equation

$$Z_t = \alpha_1 Z_{t-1} + ... + \alpha_p Z_{t-p} + \varepsilon_t.$$

But X_t satisfies

$$\begin{aligned} X_t &= \alpha_1 X_{t-1} + ... + \alpha_p X_{t-p} + \varepsilon_t + (\mu - \alpha_1\mu - ... - \alpha_p\mu) \\ &= \alpha_1 X_{t-1} + ... + \alpha_p X_{t-p} + \varepsilon_t + \widetilde{\mu}. \end{aligned}$$

1.2. Time lag operator. Let S be the space of all sequences $(x_t)_{t\in\mathbb{Z}}$ of real numbers. Let us define an *operator* $L : S \to S$, a map which transform sequences in sequences. It is defined as

$$Lx_t = x_{t-1}, \qquad \text{for all } t \in \mathbb{Z}.$$

We should write $(Lx)_t = x_{t-1}$, with the meaning that, given a sequence $x = (x_t)_{t\in\mathbb{Z}} \in S$, we introduce a new sequence $Lx \in S$, that at time t is equal to the original sequence at time $t-1$, hence the notation $(Lx)_t = x_{t-1}$. For shortness, we drop the bracket and write $Lx_t = x_{t-1}$, but it is clear that L operates on the full sequence x, not on the single value x_t.

The map L is called *time lag operator*, or *backward shift*, because the result of L is a shift, a translation, of the sequence, backwards (in the sense that we observe the same sequence but from one position shifted on the left).

If we work on the space S^+ of sequences $(x_t)_{t\in\mathbb{N}}$ defined only for non-negative times, we cannot define this operator since, given $(x_t)_{t\in\mathbb{N}}$, its first value is x_0, while the first value of Lx should be x_{-1} which does not exist. Nevertheless, if we forget the first value, the operation of backward shift is meaningful also here. Hence the notation $Lx_t = x_{t-1}$ is used also for sequences $(x_t)_{t\in\mathbb{N}}$, with the understanding that one cannot take $t = 0$ in it.

Remark 1. *The time lag operator is a linear operator.*

The powers, positive and negative, of the lag operator are denoted by L^k:

$$L^k x_t = x_{t-k}, \qquad \text{for } t \in \mathbb{Z}$$

(or, for $t \geq \max(k, 0)$, for sequences $(x_t)_{t\in\mathbb{N}}$).

With this notation, the AR model reads

$$\left(1 - \sum_{k=1}^{p} \alpha_k L^k\right) X_t = \varepsilon_t.$$

1.3. MA models. A MA(q) (Moving Average with orders p and q) model is an explicit formula for X_t in terms of noise of the form

$$X_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + ... + \beta_q \varepsilon_{t-q}.$$

The process is given by a (weighted) average of the noise, but not an average from time zero to the present time t; instead, an average moving with t is taken, using only the last $q+1$ times.

Using time lags we can write

$$X_t = \left(1 + \sum_{k=1}^{q} \beta_k L^k\right) \varepsilon_t.$$

1.4. ARMA models. An ARMA(p, q) (AutoRegressive Moving Average with orders p and q) model is a discrete time linear equations with noise, of the form

$$\left(1 - \sum_{k=1}^{p} \alpha_k L^k\right) X_t = \left(1 + \sum_{k=1}^{q} \beta_k L^k\right) \varepsilon_t$$

or explicitly

$$X_t = \alpha_1 X_{t-1} + ... + \alpha_p X_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + ... + \beta_q \varepsilon_{t-q}.$$

We may incorporate a non-zero average in this model. If we want that X_t has average μ, the natural procedure is to have a zero-average solution Z_t of

$$Z_t = \alpha_1 Z_{t-1} + ... + \alpha_p Z_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + ... + \beta_q \varepsilon_{t-q}$$

and take $X_t = Z_t + \mu$, hence solution of

$$X_t = \alpha_1 X_{t-1} + ... + \alpha_p X_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + ... + \beta_q \varepsilon_{t-q} + \widetilde{\mu}$$

with

$$\widetilde{\mu} = \mu - \alpha_1 \mu - ... - \alpha_p \mu.$$

1.5. Difference operator. Integration. The *first difference operator*, Δ, is defined as

$$\Delta X_t = X_t - X_{t-1} = (1 - L) X_t.$$

If we call

$$Y_t = (1 - L) X_t$$

then we may reconstruct X_t from Y_t by *integration*:

$$X_t = Y_t + X_{t-1} = Y_t + Y_{t-1} + X_{t-2} = ... = Y_t + ... + Y_1 + X_0$$

having the initial condition X_0.

The *second difference operator*, Δ^2, is defined as

$$\Delta^2 X_t = (1 - L)^2 X_t.$$

Assume we have

$$Y_t = (1 - L)^2 X_t.$$

Then

$$Y_t = (1 - L) Z_t$$
$$Z_t = (1 - L) X_t$$

so we may first reconstruct Z_t from Y_t:

$$Z_t = Y_t + ... + Y_2 + Z_1$$

where

$$Z_1 = (1 - L) X_1 = X_1 - X_0$$

(thus we need $X_1$and X_0); then we reconstruct X_t from Z_t:

$$X_t = Z_t + ... + Z_1 + X_0.$$

All this can be generalized to Δ^d, for any positive integer d.

1.6. ARIMA models. An ARIMA(p, d, q) (AutoRegressive Integrated Moving Average with orders p, d, q) model is a discrete time linear equations with noise, of the form

$$\left(1 - \sum_{k=1}^{p} \alpha_k L^k\right)(1 - L)^d X_t = \left(1 + \sum_{k=1}^{q} \beta_k L^k\right)\varepsilon_t.$$

It is a particular case of ARMA models, but with a special structure. Set $Y_t := (1 - L)^d X_t$. Then Y_t is an ARMA(p, q) model

$$\left(1 - \sum_{k=1}^{p} \alpha_k L^k\right) Y_t = \left(1 + \sum_{k=1}^{q} \beta_k L^k\right)\varepsilon_t$$

and X_t is obtained from Y_t by d successive *integrations*. The number d is thus the order of integration.

Example 2. *The random walk is ARIMA*$(0, 1, 0)$.

We may incorporate a non-zero average in the auxiliary process Y_t and consider the equation

$$\left(1 - \sum_{k=1}^{p} \alpha_k L^k\right)(1 - L)^d X_t = \left(1 + \sum_{k=1}^{q} \beta_k L^k\right)\varepsilon_t + \widetilde{\mu}$$

with

$$\widetilde{\mu} = \mu - \alpha_1 \mu - ... - \alpha_p \mu.$$

2. Stationarity, ARMA and ARIMA processes

Under suitable conditions on the parameters, there are stationary solutions to ARMA models, called *ARMA processes*.

In the simplest case of AR(1) models, we have proved the stationarity (with suitable variance of the initial condition) when the parameter α satisfies $|\alpha| < 1$. In general, there are conditions but they are quite technical and we address the interested reader to the specialized literature. In the sequel we shall always use sentences of the form: "consider a stationary solution of the following ARMA model", meaning implicitly that it exists, namely that we are in the framework of such conditions. Our statements will therefore hold only in such case, otherwise are just empty statements.

Integration brakes stationarity. Solutions to ARIMA models are always non-stationary if we take Y_t stationary (in this case the corresponding X_t is called *ARIMA process*). For instance, the random walk is not stationary. The kind of growth of such processes is not always trivial. But if we include a non-zero average, namely we consider the case

$$\left(1 - \sum_{k=1}^{p} \alpha_k L^k\right)(1 - L)^d X_t = \left(1 + \sum_{k=1}^{q} \beta_k L^k\right)\varepsilon_t + \widetilde{\mu}$$

then we have the following: if $d = 1$, X_t has a *linear trend*; if $d = 2$, a *quadratic trend*, and so on. Indeed, at a very intuitive level, if Y_t is a stationary solution of the associated ARMA model, with mean μ, then its integration produces a trend: a single step integration gives us

$$X_t = Y_t + ... + Y_1 + X_0$$

so the stationary values of Y accumulate linearly; a two step integration produces a quadratic accumulation, and so on. When $\mu = 0$, the sum $Y_t + ... + Y_1$ has a lot of cancellations, so the trend is sublinear (roughly it behaves as a square root). But the cancellations become statistically not significant when $\mu \neq 0$. If $\mu > 0$ and $d = 1$ we observe an average linear growth; if If $\mu < 0$ and $d = 1$ we observe an average linear decay. This is also related to the ergodic theorem: since Y_t is stationary and its autocorrelation decays at infinity, we may apply the ergodic theorem and have that

$$\frac{Y_t + ... + Y_1}{t} \to E[Y_1] = \mu$$

(in mean square). Hence

$$Y_t + ... + Y_1 \sim \mu \cdot t.$$

There are fluctuations, roughly of the order of a square root, around this linear trend.

3. Correlation function

3.1. First results. Assume we have a stationary, mean zero, ARMA(p, q) process. Set $\gamma_n := R(n) = E[X_n X_0]$.

PROPOSITION 1. *For every $j > q$,*

$$\gamma_j = \sum_{k=1}^{p} \alpha_k \gamma_{j-k}.$$

PROOF. Recall that $R(-n) = R(n)$. Recall also that

$$\left(1 - \sum_{k=1}^{p} \alpha_k L^k\right) X_t = \left(1 + \sum_{k=1}^{q} \beta_k L^k\right) \varepsilon_t$$

Notice that for every n and m we have

$$E[X_{t-n} L^m X_t] = E[X_{t-n} X_{t-m}] = \gamma_{m-n}.$$

Then

$$\begin{aligned} \gamma_j - \sum_{k=1}^{p} \alpha_k \gamma_{j-k} &= E[X_t X_{t-j}] - \sum_{k=1}^{p} \alpha_k E[X_{t-k} X_{t-j}] \\ &= E\left[X_{t-j}\left(X_t - \sum_{k=1}^{p} \alpha_k X_{t-k}\right)\right] \\ &= E\left[X_{t-j}\left(1 + \sum_{k=1}^{q} \beta_k L^k\right) \varepsilon_t\right]. \end{aligned}$$

In the case $j > q$ we have $L^k \varepsilon_t$ independent of X_{t-j}, for $k \leq q$, hence

$$\gamma_j - \sum_{k=1}^{p} \alpha_k \gamma_{j-k} = 0.$$

The proof is complete. □

This formula allows us to compute the full autocorrelation function when $q = 0$ (AR processes).

EXAMPLE 3. *Consider the simple case*

$$X_t = aX_{t-1} + \varepsilon_t.$$

We get

$$\gamma_j - \alpha\gamma_{j-1} = 0$$

for every $j > 0$*, namely*

$$\gamma_1 = \alpha\gamma_0$$
$$\gamma_2 = \alpha\gamma_1$$
$$\ldots$$

where $\gamma_0 = E\left[X_0^2\right]$*. Hence*

$$\gamma_j = \alpha^j\gamma_0.$$

On its own,

$$Var\left[X_t\right] = a^2 Var\left[X_{t-1}\right] + Var\left[\varepsilon_t\right]$$

hence

$$\gamma_0 = a^2\gamma_0 + \sigma^2$$

which gives us $\gamma_0 = \frac{\sigma^2}{1-a^2}$*. This is the same result found in Chapter 2.*

EXAMPLE 4. *Consider the next case,*

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \varepsilon_t.$$

We get

$$\gamma_j = \alpha_1\gamma_{j-1} + \alpha_2\gamma_{j-2}$$

for every $j > 0$*, namely*

$$\gamma_1 = \alpha_1\gamma_0 + \alpha_2\gamma_{-1}$$
$$\gamma_2 = \alpha_1\gamma_1 + \alpha_2\gamma_0$$
$$\ldots$$

The first equation, in view of $\gamma_{-1} = \gamma_1$*, gives us*

$$\gamma_1 = \frac{\alpha_1}{1-\alpha_2}\gamma_0.$$

Hence again we just need to find γ_0*. We have*

$$Var\left[X_t\right] = a_1 Var\left[X_{t-1}\right] + a_2 Var\left[X_{t-2}\right] + \sigma^2 + 2Cov\left(X_{t-1}, X_{t-2}\right)$$

hence

$$\gamma_0 = a_1\gamma_0 + a_2\gamma_0 + \sigma^2 + 2\gamma_1.$$

This is a second relation between γ_0 *and* γ_1*; together, they will gives us both quantities.*

3.2. ARMA as infinite MA. Assume always that we have a stationary, mean zero, ARMA process. Assume it is defined for all integers (in particular we use the noise for negative integers). From

$$\left(1-\sum_{k=1}^{p}\alpha_k L^k\right)X_t=\left(1+\sum_{k=1}^{q}\beta_k L^k\right)\varepsilon_t$$

we get, when proper convergence conditions are satisfied,

$$X_t=\sum_{j=0}^{\infty}\varphi_j L^j\varepsilon_t$$

where $\sum_{j=0}^{\infty}\varphi_j x^j$ is the Taylor expansion of the function

$$g(x)=\frac{1+\sum_{k=1}^{q}\beta_k x^k}{1-\sum_{k=1}^{p}\alpha_k x^k}.$$

Indeed, assume this function g has the Taylor development $g(x)=\sum_{j=0}^{\infty}\varphi_j x^j$ in a neighborhood U of the origin. Then, for each $x\in U$,

$$\left(1-\sum_{k=1}^{p}\alpha_k x^k\right)\sum_{j=0}^{\infty}\varphi_j x^j=\left(1+\sum_{k=1}^{q}\beta_k x^k\right).$$

Assume

$$\sum_{j=0}^{\infty}\varphi_j^2<\infty.$$

One can prove that the series $X_t:=\sum_{j=0}^{\infty}\varphi_j L^j\varepsilon_t$ converges in mean square; and also $\sum_{j=0}^{\infty}\varphi_j L^{j-k}\varepsilon_t$ for every k, and thus we have

$$\left(1-\sum_{k=1}^{p}\alpha_k L^k\right)\sum_{j=0}^{\infty}\varphi_j L^j\varepsilon_t=\left(1+\sum_{k=1}^{q}\beta_k L^k\right)\varepsilon_t.$$

Therefore $X_t:=\sum_{j=0}^{\infty}\varphi_j L^j\varepsilon_t$ solves the equation which defines the ARMA process. This is always a solution; if we started with a given solution X_t, only under a uniqueness condition in law (that we omit to discuss), we are sure that such solution can be represented in the form $X_t=\sum_{j=0}^{\infty}\varphi_j L^j\varepsilon_t$.

EXAMPLE 5. *Consider the simple case*

$$X_t=aX_{t-1}+\varepsilon_t.$$

We have

$$g(x)=\frac{1}{1-\alpha x}$$

hence

$$g(x)=\sum_{j=0}^{\infty}(\alpha x)^j$$

(recall the geometric series). The series converges for $|\alpha x| < 1$*. We need* $|\alpha| < 1$ *to have* $\sum_{j=0}^{\infty} \varphi_j^2 < \infty$*. Therefore the ARMA process is*

$$X_t = \sum_{j=0}^{\infty} \alpha^j L^j \varepsilon_t.$$

3.3. Correlation function, part 2. Assume always that we have a stationary, mean zero, ARMA process and set $\gamma_n := R(n) = E[X_n X_0]$. Under proper conditions,

$$X_t = \sum_{i=0}^{\infty} \varphi_i L^i \varepsilon_t.$$

Hence we may compute $E\left[X_{t-j}\left(1 + \sum_{k=1}^{q} \beta_k L^k\right) \varepsilon_t\right]$ also for $j \leq q$, the case which was left aside above, in the proof of Proposition 1.

PROPOSITION 2. *Under the previous assumptions, for all* $j = 0, ..., q$ *we have*

$$\gamma_j - \sum_{k=1}^{p} \alpha_k \gamma_{j-k} = \sum_{i=0}^{q-j} \varphi_i \beta_{i+j} \sigma^2.$$

Thus, for every $j \geq 0$ *we may write*

$$\gamma_j - \sum_{k=1}^{p} \alpha_k \gamma_{j-k} = \sum_{i=0}^{\infty} \varphi_i \beta_{i+j} \sigma^2 1_{i+j \in \{0,...,q\}}.$$

PROOF. We have (we set $\beta_0 = 1$), from the proof of Proposition 1 and the identity $X_{t-j} = \sum_{i=0}^{\infty} \varphi_i L^i \varepsilon_{t-j}$,

$$\begin{aligned}
\gamma_j - \sum_{k=1}^{p} \alpha_k \gamma_{j-k} &= E\left[X_{t-j} \sum_{k=0}^{q} \beta_k L^k \varepsilon_t\right] \\
&= \sum_{i=0}^{\infty} \sum_{k=0}^{q} \varphi_i \beta_k E\left[L^i \varepsilon_{t-j} L^k \varepsilon_t\right] \\
&= \sum_{i=0}^{\infty} \sum_{k=0}^{q} \varphi_i \beta_k \delta_{i+j,k} \sigma^2 \\
&= \sum_{i=0}^{\infty} \varphi_i \beta_{i+j} \sigma^2 1_{i+j \in \{0,...,q\}}.
\end{aligned}$$

The proof is complete. □

This is a general formula. A more direct approach to compute $E\left[X_{t-j}\left(1 + \sum_{k=1}^{q} \beta_k L^k\right) \varepsilon_t\right]$ even if $j \leq q$ consists in the substitution of the equation for X_{t-j}:

$$E\left[X_{t-j}\left(1 + \sum_{k=1}^{q} \beta_k L^k\right) \varepsilon_t\right] = E\left[\left(\sum_{k=1}^{p} \alpha_k L^k X_{t-j} + \left(1 + \sum_{k=1}^{q} \beta_k L^k\right) \varepsilon_{t-j}\right)\left(1 + \sum_{k=1}^{q} \beta_k L^k\right) \varepsilon_t\right].$$

The products involving $L^k \varepsilon_{t-j}$ and $L^{k'} \varepsilon_t$ are easily computed. Then we have products of the form

$$E\left[L^k X_{t-j} L^{k'} \varepsilon_t\right]$$

the worse of which is

$$E\left[L^1 X_{t-j} L^q \varepsilon_t\right].$$

If $j \geq q$, it is zero, otherwise not, but we may repeat the trick and go backward step by step. In simple examples we may compute all γ_j by this strategy.

Example 6. *Consider the simple case*

$$X_t = \alpha X_{t-1} + \varepsilon_t + \beta \varepsilon_{t-1}.$$

We get

$$\gamma_j - \alpha \gamma_{j-1} = 0$$

for every $j > 1$, namely

$$\gamma_2 = \alpha \gamma_1$$
$$\gamma_3 = \alpha \gamma_2$$
$$...$$

but from this relations we miss γ_1 and γ_0. About γ_1, we have

$$\begin{aligned} \gamma_1 - \alpha \gamma_0 &= E\left[X_{t-1}\left(1 + \beta L\right) \varepsilon_t\right] = \beta E\left[X_{t-1} \varepsilon_{t-1}\right] \\ &= \beta E\left[\left(\alpha X_{t-2} + \varepsilon_{t-1} + \beta \varepsilon_{t-2}\right) \varepsilon_{t-1}\right] = \beta \sigma^2. \end{aligned}$$

Therefore γ_1 is given in terms of γ_0 (and then iteratively also γ_2, γ_3 and so on). On its own,

$$Var\left[X_t\right] = a^2 Var\left[X_{t-1}\right] + \sigma^2 + \beta^2 \sigma^2 + 2\alpha\beta Cov\left(X_{t-1}, \varepsilon_{t-1}\right)$$

hence

$$\gamma_0 = a^2 \gamma_0 + \sigma^2 + \beta^2 \sigma^2 + 2\alpha\beta Cov\left(X_{t-1}, \varepsilon_{t-1}\right).$$

Moreover,

$$Cov\left(X_{t-1}, \varepsilon_{t-1}\right) = Cov\left(\alpha X_{t-2} + \varepsilon_{t-1} + \beta \varepsilon_{t-2}, \varepsilon_{t-1}\right) = \sigma^2$$

hence

$$\gamma_0 = a^2 \gamma_0 + \sigma^2 + \beta^2 \sigma^2 + 2\alpha\beta\sigma^2.$$

This gives us γ_0.

4. Power spectral density

We work under the assumptions of the previous sections, in particular that X is a stationary ARMA process and $g(x) = \frac{1+\sum_{k=1}^{q} \beta_k x^k}{1-\sum_{k=1}^{p} \alpha_k x^k}$ has the Taylor development $g(x) = \sum_{j=0}^{\infty} \varphi_j x^j$ in a complex neighborhood U of the origin which includes the closed ball of radius 1. Moreover, we assume the assumptions of Wiener-Khinchin theorem of Chapter 3.

THEOREM 1.

$$S\left(\omega\right)=\frac{\sigma^2}{2\pi}\left|\frac{1+\sum_{k=1}^{q}\beta_k e^{-ik\omega}}{1-\sum_{k=1}^{p}\alpha_k e^{-ik\omega}}\right|^2.$$

PROOF. We have ($\mathbb{Z}_T$ denotes the set of all $n\in\mathbb{Z}$ such that $|n|\leq T/2$)

$$\widehat{X}_T\left(\omega\right)=\frac{1}{\sqrt{2\pi}}\sum_{n\in\mathbb{Z}_T}e^{-i\omega n}X_n=\frac{1}{\sqrt{2\pi}}\sum_{n\in\mathbb{Z}_T}\sum_{j=0}^{\infty}\varphi_j e^{-i\omega n}\varepsilon_{n-j}$$

$$\widehat{X}_T^*\left(\omega\right)=\frac{1}{\sqrt{2\pi}}\sum_{n'\in\mathbb{Z}_T}\sum_{j'=0}^{\infty}\varphi_{j'}e^{i\omega n'}\varepsilon_{n'-j'}$$

$$E\left[\widehat{X}_T\left(\omega\right)\widehat{X}_T^*\left(\omega\right)\right]=\frac{1}{2\pi}E\left[\sum_{n\in\mathbb{Z}_T}\sum_{n'\in\mathbb{Z}_T}\sum_{j=0}^{\infty}\sum_{j'=0}^{\infty}\varphi_j\varphi_{j'}e^{-i\omega n}e^{i\omega n'}E\left[\varepsilon_{n-j}\varepsilon_{n'-j'}\right]\right]$$

$$=\frac{\sigma^2}{2\pi}\sum_{n\in\mathbb{Z}_T}\sum_{j=0}^{\infty}\sum_{j'=0}^{\infty}\varphi_j\varphi_{j'}e^{-i\omega n}e^{i\omega(n-j+j')}=|\mathbb{Z}_T|\frac{\sigma^2}{2\pi}\sum_{j=0}^{\infty}\sum_{j'=0}^{\infty}\varphi_j e^{-i\omega j}\varphi_{j'}e^{i\omega j'}=|\mathbb{Z}_T|\frac{\sigma^2}{2\pi}\left|\sum_{n=0}^{\infty}\varphi_n e^{-i\omega n}\right|^2.$$

The cardinality $|\mathbb{Z}_T|$ of $\mathbb{Z}_T$ has the property $\lim_{T\to\infty}|\mathbb{Z}_T|/T=1$, hence we get

$$S\left(\omega\right)=\frac{\sigma^2}{2\pi}\left|\sum_{n=0}^{\infty}\varphi_n e^{-i\omega n}\right|^2.$$

Now, it is sufficient to use the the relation $\frac{1+\sum_{k=1}^{q}\beta_k x^k}{1-\sum_{k=1}^{p}\alpha_k x^k}=\sum_{j=0}^{\infty}\varphi_j x^j$ for $x=e^{-i\omega}$. The proof is complete.

REMARK 2. *Consider the case $q=0$ and write the formula with $\omega=2\pi f$*

$$S\left(f\right)=\frac{\sigma^2}{2\pi}\frac{1}{\left|1-\sum_{k=1}^{p}\alpha_k e^{-2\pi ikf}\right|^2}.$$

Assume there is only $k=p$:

$$S\left(f\right)=\frac{\sigma^2}{2\pi}\frac{1}{\left|1-\alpha_p e^{-2\pi ipf}\right|^2}.$$

The maxima are at $pf\in\mathbb{Z}$, namely for $f=\frac{1}{p}$. The function $S\left(f\right)$ immediately shows the periodicity in the recursion

$$X_t=a_pX_{t-p}+\varepsilon_t.$$

□

4.1. Example.

$$X_t = 0.8 \cdot X_{t-12} + \varepsilon_t$$

$$S(f) = \frac{\sigma^2}{2\pi} \left| \frac{1}{1 - 0.8 \cdot e^{-2\pi i \cdot 12 \cdot f}} \right|^2$$