Time Series Analysis – Lecture 4

Models for Stationary Time Series

Dr. Kevin Burke

1 Preliminary Note

We use $\{e_t\}$ to denote a **white noise** process. This is a process with the following properties:

$$E(e_t) = 0$$
$$\text{Cov}(e_t, e_{t-k}) = \begin{cases} \sigma_e^2 & \text{for } k = 0 \\ 0 & \text{otherwise} \end{cases}$$

2 General Linear Process

The **general linear process** is a model for a stationary time series in terms of a weighted combination of the present and past white noise terms:

$$Y_t = e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \cdots$$

which can be compactly written as

$$Y_t = \sum_{i=0}^{\infty} \psi_i \, e_{t-i}$$

where $\psi_0 = 1$.

The series can also be written in terms of the **backshift operator** as follows:

$$\begin{aligned} Y_t &= e_t + \psi_1 B^1 e_t + \psi_2 B^2 e_t + \cdots \\ &= (1 + \psi_1 B^1 + \psi_2 B^2 + \cdots)\, e_t \\ &= \psi(B)\, e_t \end{aligned}$$

where $\psi(B) = (1 + \psi_1 B^1 + \psi_2 B^2 + \cdots)$.

The mean function is given by

$$\begin{aligned} \mu_t = E(Y_t) &= E\left(\sum_{i=0}^{\infty} \psi_i \, e_{t-i}\right) \\ &= \sum_{i=0}^{\infty} \psi_i \, E(e_{t-i}) \\ &= \sum_{i=0}^{\infty} \psi_i \,(0) = 0. \end{aligned}$$

If we require a non-zero mean function then of course we can simply redefine the process as $Y_t = \mu + \sum_{i=0}^{\infty} \psi_i \, e_{t-i}$.

The autocovariance function is

$$\begin{aligned} \gamma_{t,t-k} &= \text{Cov}(Y_t, Y_{t-k}) \\ &= \text{Cov}\left(\sum_{i=0}^{\infty} \psi_i \, e_{t-i}, \sum_{j=0}^{\infty} \psi_j \, e_{t-k-j}\right) \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \psi_i \psi_j \text{Cov}\left(e_{t-i},\, e_{t-k-j}\right) \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \psi_i \psi_j \text{Cov}\left(e_{t-i},\, e_{t-(j+k)}\right) \end{aligned}$$

Note that $\text{Cov}\left(e_{t-i},\, e_{t-(j+k)}\right) = \sigma_e^2$ when $i = j + k$ and zero otherwise.

Therefore, the autocovariance function for the general linear process is

$$\begin{aligned} \gamma_{t,t-k} &= \sum_{j=0}^{\infty} \psi_{j+k} \psi_j \sigma_e^2 \\ &= \sigma_e^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} = \gamma_k. \end{aligned}$$

Clearly the mean function is constant and the acvf function depends only on k. These are requirements for stationarity but note that since the acvf is an infinite sum, we must constrain the weights ψ_i in some way to ensure the sum is *convergent* so that $|\gamma_k| < \infty$ (also required for stationarity).

First recall that the Cauchy-Schwarz inequality is $|\text{Cov}(X, Y)| \leq \sqrt{\text{Var}(X)\text{Var}(Y)}$ which, in the case of a stationary time series becomes

$$|\gamma_k| \leq \sqrt{\gamma_0 \gamma_0} = \gamma_0$$

For the general linear process we have that

$$\gamma_0 = \sigma_e^2 \sum_{j=0}^{\infty} \psi_j^2.$$

Hence, if $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ then $\gamma_0 < \infty$ and, by the Cauchy-Schwarz inequality, $|\gamma_k| \leq \gamma_0 < \infty$. So $\boldsymbol{\sum_{j=0}^{\infty} \psi_j^2 < \infty}$ **is required for stationarity**.

Often we wish to write a time series model in the following form:

$$Y_t = \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \cdots + e_t$$

so that the current value in the series is a weighted combination of the previous values plus a purely random white noise term, e_t. A series that can be written in this form is known as **invertible**. This is an important property for the purpose of forecasting because it **implies that past values can predict future values**.

We can also write this in terms of the backshift operator

$$\begin{aligned} Y_t - \pi_1 Y_{t-1} - \pi_2 Y_{t-2} - \cdots &= e_t \\ (1 - \pi_1 B^1 - \pi_2 B^2 - \cdots) Y_t &= e_t \\ \pi(B) Y_t &= e_t. \end{aligned}$$

If Y_t is a general linear process, i.e., $Y_t = \psi(B)e_t$, then the above equation becomes

$$\pi(B)\psi(B)e_t = e_t$$

and hence

$$\pi(B)\psi(B) = 1.$$

This relationship can be used to determine the π weights for a given set of ψ weights and vice versa, i.e., we can use this to convert the model from one form to the other.

3 Moving Average Process: MA(q)

A general linear process with a fixed number of ψ weights is called a **moving average (MA) process**.

3.1 MA(1)

The MA(1) process is given by

$$Y_t = e_t - \theta e_{t-1}$$

which is a general linear process with $\psi_1 = -\theta$ and all other $\psi_i = 0$ for $i > 1$. Hence, this is a stationary process.

From the theory developed in Section 2 we have that $E(Y_t) = 0$ and the acvf is

$$\gamma_k = \sigma_e^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}$$

$$= \begin{cases} \sigma_e^2(\psi_0^2 + \psi_1^2 + \cdots) & = \sigma_e^2(1+\theta^2) & \text{for } k = 0 \\ \sigma_e^2(\psi_0\psi_1 + \psi_1\psi_2 + \cdots) & = \sigma_e^2(-\theta) & \text{for } k = 1 \\ & = 0 & \text{for } k \geq 2 \end{cases}$$

It is instructive to work out the mean and acvf directly (already covered in Lecture 2 for MA(1) with $\theta = -\frac{1}{2}$):

$$\begin{aligned} E(Y_t) &= E(e_t) - \theta E(e_{t-1}) \\ &= 0 - \theta(0) \\ &= 0. \end{aligned}$$

$$\begin{aligned} \text{Cov}(Y_t, Y_{t-k}) &= \text{Cov}(e_t - \theta e_{t-1}, e_{t-k} - \theta e_{t-k-1}) \\ &= \begin{cases} (1+\theta^2)\sigma_e^2 & \text{for } k = 0 \\ -\theta\sigma_e^2 & \text{for } k = 1 \\ 0 & \text{for } k \geq 2 \end{cases} \end{aligned}$$

As found using general linear process theory. The autocorrelation function is therefore

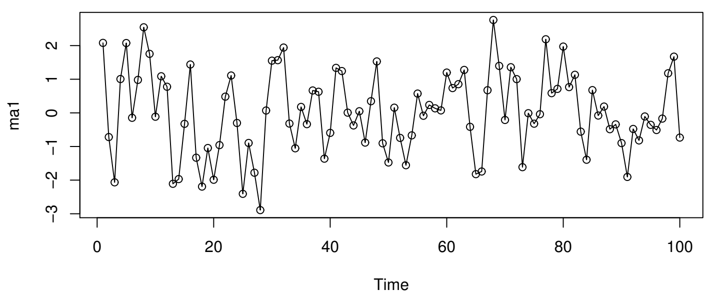
$$\begin{aligned} \rho_k &= \frac{\gamma_k}{\gamma_0} \\ &= \begin{cases} 1 & \text{for } k = 0 \\ \dfrac{-\theta}{1+\theta^2} & \text{for } k = 1 \\ 0 & \text{for } k \geq 2 \end{cases} \end{aligned}$$

Note that the sign of ρ_1 is the opposite of the sign of θ. It is easy to show that for the MA(1) process $|\boldsymbol{\rho_1}| \leq \mathbf{0.5}$.

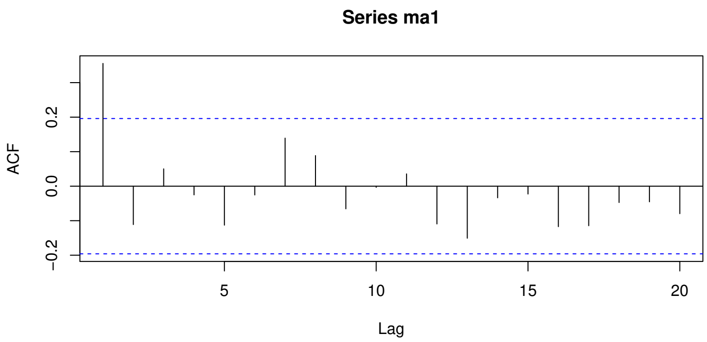
Example 3.1. *MA(1) with* $\theta = -0.9$
With $\theta = -0.9$ the lag-1 autocorrelation is

$$\rho_1 = \frac{-\theta}{1+\theta^2} = \frac{0.9}{1+0.81} = 0.4972$$

which means that neighbouring values of the time series will hang together due to positive correlation and this is seen in the time series plot below.



The acf for this series shows that only the lag-1 autocorrelation is significantly different to zero.



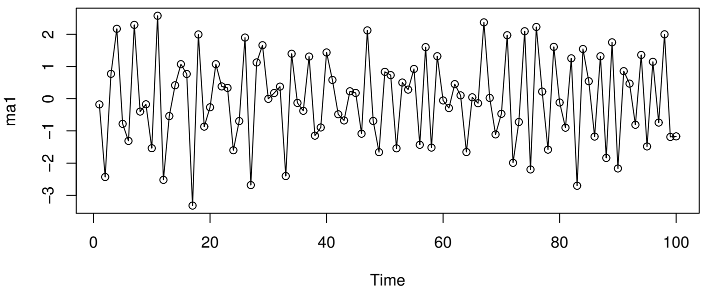

```
set.seed(196311)
ma1 <- arima.sim(n=100, model=list(ma=c(0.9)) )
dev.new(width=8, height=4)
plot(ma1, type="o")
(acf(ma1))
```

Note: In R, the process is defined $\boldsymbol{Y_t = e_t + \theta e_{t-1}}$; hence, the $\boldsymbol{\theta}$ parameter has the opposite sign.

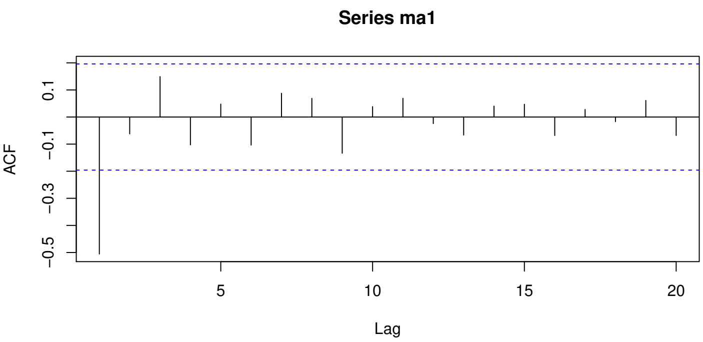
Example 3.2. *MA(1) with* $\theta = 0.9$
With $\theta = 0.9$ the lag-1 autocorrelation is

$$\rho_1 = \frac{-\theta}{1+\theta^2} = \frac{-0.9}{1+0.81} = -0.4972$$

which means that neighbouring values of the time series will tend to oscillate around the mean and this is seen in the time series plot below.



The acf for this series shows that only the lag-1 autocorrelation is significantly different to zero.



3.2 MA(2)

The MA(2) process is given by

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}$$

which is a general linear process with $\psi_1 = -\theta_1$, $\psi_2 = -\theta_2$ and all other $\psi_i = 0$ for $i > 2$. Hence, this is a stationary process.

Of course we could use the theory from Section 2 to work out the acvf but we will do it directly:

$$\begin{aligned} \mathrm{Cov}(Y_t, Y_{t-k}) &= \mathrm{Cov}(e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}, \\ &\qquad e_{t-k} - \theta_1 e_{t-k-1} - \theta_2 e_{t-k-2}) \\ &= \begin{cases} (1+\theta_1^2+\theta_2^2)\sigma_e^2 & \text{for } k=0 \\ (-\theta_1+\theta_1\theta_2)\sigma_e^2 & \text{for } k=1 \\ -\theta_2\,\sigma_e^2 & \text{for } k=2 \\ 0 & \text{for } k \geq 3 \end{cases} \end{aligned}$$

and, hence, the acf is

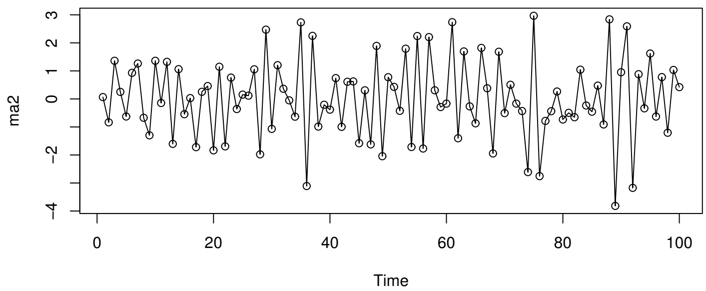
$$\rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} 1 & \text{for } k=0 \\ \dfrac{-\theta_1+\theta_1\theta_2}{1+\theta_1^2+\theta_2^2} & \text{for } k=1 \\ \dfrac{-\theta_2}{1+\theta_1^2+\theta_2^2} & \text{for } k=2 \\ 0 & \text{for } k \geq 3 \end{cases}$$

Example 3.3. *MA(2) with* $\theta_1 = 1.0$ *and* $\theta_2 = -0.6$
With $\theta_1 = 1.0$ and $\theta_2 = -0.6$ the lag-1 and lag-3 autocorrelations are

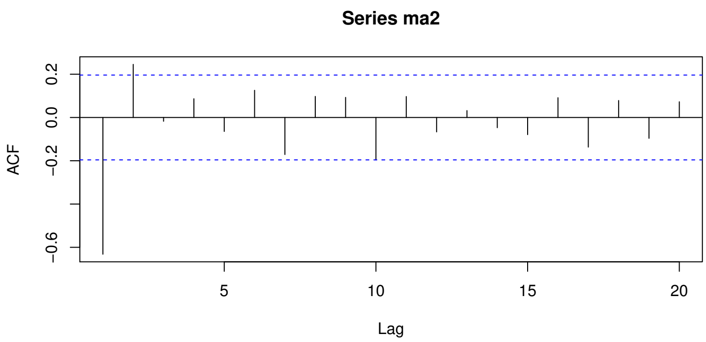
$$\rho_1 = \frac{-1.0 + 1.0(-0.6)}{1+(1.0)^2+(-0.6)^2} = -0.677,$$

$$\rho_2 = \frac{-(-0.6)}{1+(1.0)^2+(-0.6)^2} = 0.254.$$

The relatively strong negative lag-1 autocorrelation will cause the series to oscillate around the mean and this is seen in the time series plot below.



The acf for this series shows that only the lag-1 and lag-2 autocorrelations are significantly different to zero.




```
set.seed(5928371)
ma2 <- arima.sim(n=100, model=list(ma=c(-1.0, 0.6)) )
dev.new(width=8, height=4)
plot(ma2, type="o")
(acf(ma2))
```

3.3 MA(q)

The general MA(q) process is given by

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q}$$

and is simply a general linear process with a finite number of non-zero ψ weights, i.e., $\psi_i = -\theta_i$ for $i = 1, \ldots, q$ and $\psi_i = 0$ for $i > q$. As it is a general linear process with a finite number of terms, it is clearly convergent. Thus, **all MA processes are stationary.**

The MA(q) process has the property that **only the first q autocorrelations are non-zero.**

This process can be written in terms of the backshift operator

$$\begin{aligned} Y_t &= (1 - \theta_1 B^1 - \theta_2 B^2 - \ldots - \theta_q B^q)\, e_t \\ &= \theta(B)\, e_t \end{aligned}$$

where $\theta(B) = 1 - \theta_1 B^1 - \theta_2 B^2 - \ldots - \theta_q B^q$ is known as the **MA characteristic polynomial**.

Note: `R` defines the MA process using plus signs before the θ coefficients.

3.4 Invertibility

Consider the MA(1) process

$$Y_t = e_t - \theta e_{t-1} = (1 - \theta B) e_t.$$

Although this series is stationary, we will need to constrain θ so that it is **invertible**. Recall that if a series is invertible it can be written as

$$Y_t = \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \cdots + e_t$$

$$\Rightarrow (1 - \pi_1 B^1 - \pi_2 B^2 - \cdots) Y_t = e_t.$$

Note that the MA(1) process can be written

$$e_t = Y_t + \theta e_{t-1}$$

and by successively back-substituting we get

$$\begin{aligned} e_t &= Y_t + \theta e_{t-1} \\ &= Y_t + \theta(Y_{t-1} + \theta e_{t-2}) \\ &= Y_t + \theta Y_{t-1} + \theta^2 e_{t-2} \\ &= Y_t + \theta Y_{t-1} + \theta^2 (Y_{t-2} + \theta e_{t-3}) \\ &= Y_t + \theta Y_{t-1} + \theta^2 Y_{t-2} + \theta^3 Y_{t-3} + \cdots \\ &= (1 + \theta B^1 + \theta^2 B^2 + \cdots)\, Y_t \\ &= \pi(B)\, Y_t \end{aligned}$$

which is clearly in the invertible form with $\pi_i = -\theta^i$.

Note we could have also found $\pi_i = -\theta^i$ as follows:

$$\begin{aligned} \pi(B)\, Y_t &= e_t \\ \pi(B)(1 - \theta B)\, e_t &= e_t \\ \pi(B)(1 - \theta B) &= 1. \\ \pi(B)(1 - \theta B) &= 1 + 0B^1 + 0B^2 + \cdots \end{aligned}$$

We then multiply out $\pi(B)(1 - \theta B)$ and match the coefficients of B^i with those on the right hand side.

In order for the inversion to be meaningful, the sum $1 + \theta B^1 + \theta^2 B^2 + \cdots$ must *converge* which clearly requires $|\theta| < 1$.

It is easy to show that if $|\theta| > 1$, then Y_t can only be expressed in terms of the future values $(Y_{t+1}, Y_{t+2}, \ldots)$. However, this is not useful since we then need to know the future to predict the future. In contrast, an invertible time series can be used to **predict the future using the past** which is why this property is important.

General Condition for Invertibility

We have only considered the specific case of the MA(1) process above. However, in general, the condition for invertibility is that **the roots of the MA characteristic equation**

$$1 - \theta_1 B^1 - \theta_2 B^2 - \ldots - \theta_q B^q = 0$$

must have modulus greater than one. The proof of this result is beyond the scope of this course.

As a specific example, consider again the MA(1) process whose characteristic equation is

$$\begin{aligned} 1 - \theta B &= 0 \\ -\theta B &= -1 \\ B &= \frac{1}{\theta} \end{aligned}$$

For invertibility we require

$$\left|\frac{1}{\theta}\right| > 1$$

$$\Rightarrow |\theta| < 1$$

which is the same as what we found by inverting the series manually.

4 Autoregressive Process: AR(p)

An autoregressive process is defined by the relationship that the current value of the series, Y_t, depends on past values $(Y_{t-1}, Y_{t-2}, \ldots)$ plus a white noise term, e_t. The white noise term encapsulates what cannot be explained by the past values, i.e., **e_t is *independent* of the past which implies $\mathrm{Cov}(e_t, Y_{t-k}) = 0$ for $k \geq 1$**.

4.1 AR(1)

The AR(1) process is given by

$$Y_t = \phi Y_{t-1} + e_t$$

which is clearly in the invertible form with $\pi_1 = \phi$ and all other $\pi_i = 0$ for $i > 1$.

It can be shown (see Section 4.4) that the AR(1) process can be written as

$$Y_t = \sum_{i=0}^{\infty} \phi^i e_{t-i}$$

which is a general linear process with $\psi_i = \phi^i$ and, hence, this is a stationary process (as long as $|\boldsymbol{\phi}| < \mathbf{1}$ so that the sum converges).

From the theory developed in Section 2 we have that $E(Y_t) = 0$ and the acvf is

$$\begin{aligned}
\gamma_k &= \sigma_e^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} \\
&= \sigma_e^2 \sum_{j=0}^{\infty} \phi^j \phi^{j+k} \\
&= \sigma_e^2 \phi^k \sum_{j=0}^{\infty} \phi^{2j} \\
&= \sigma_e^2 \phi^k \sum_{j=0}^{\infty} (\phi^2)^j \\
&= \sigma_e^2 \phi^k \frac{1}{1-\phi^2} \\
&\quad \text{(geometric sum with } a = 1 \text{ and } r = \phi^2) \\
&= \phi^k \frac{\sigma_e^2}{1-\phi^2}
\end{aligned}$$

where we have used the fact that $|\phi^2| < 1$ as a result of $|\phi| < 1$.

We will now work out the acvf directly (without converting to general linear process form. This is important because AR models of order greater than 1 can be difficult to convert to general linear process form.

$$\begin{aligned}
\mathrm{Cov}(Y_t, Y_{t-k}) &= \mathrm{Cov}(\phi Y_{t-1} + e_t, Y_{t-k}) \\
&= \phi\, \mathrm{Cov}(Y_{t-1}, Y_{t-k}) + \mathrm{Cov}(e_t, Y_{t-k}).
\end{aligned}$$

Important: Note that we leave Y_{t-k} rather than filling in $Y_{t-k} = \phi Y_{t-k-1} + e_{t-k}$. This is an important trick for deriving the acvf for an AR process. The next trick is to assume that the series is stationary (we will worry about the conditions for stationarity in Section 4.4).

Assuming stationarity means that the covariance depends only the time lag, i.e., $\mathrm{Cov}(Y_t, Y_{t-k}) = \gamma_{t-(t-k)} = \gamma_k$ and $\mathrm{Cov}(Y_{t-1}, Y_{t-k}) = \gamma_{t-1-(t-k)} = \gamma_{k-1}$, which gives

$$\gamma_k = \phi\, \gamma_{k-1} + \underbrace{\mathrm{Cov}(e_t, Y_{t-k})}_{=\,?}.$$

Note that for $k \geq 1$, $\mathrm{Cov}(e_t, Y_{t-k}) = 0$ as e_t is independent of the past whereas, for $k = 0$,

$$\begin{aligned}
\mathrm{Cov}(e_t, Y_{t-k}) &= \mathrm{Cov}(e_t, Y_t) \\
&= \phi\, \mathrm{Cov}(e_t, Y_{t-1}) + \mathrm{Cov}(e_t, e_t) \\
&= \phi\,(0) + \mathrm{Var}(e_t) \\
&= \sigma_e^2.
\end{aligned}$$

From this we get

$$\begin{aligned}
\gamma_k &= \phi\, \gamma_{k-1} + \mathrm{Cov}(e_t, Y_{t-k}) \\
&= \begin{cases} \phi\, \gamma_{-1} + \sigma_e^2 = \quad \phi\, \gamma_1 + \sigma_e^2 & k = 0 \\ \phi\, \gamma_{k-1} + 0 = \quad \phi\, \gamma_{k-1} & k \geq 1 \end{cases}
\end{aligned}$$

where, for $k = 0$ we have used the fact that $\gamma_{-k} = \gamma_k$ by definition (see Lecture 2).

Putting the above together we get

$$\begin{aligned}
\gamma_0 &= \phi\, \gamma_1 + \sigma_e^2 \\
\gamma_0 &= \phi\,(\phi\, \gamma_0) + \sigma_e^2 \\
&\quad \text{(since } \gamma_k = \phi\, \gamma_{k-1} \text{ for } k \geq 1) \\
&= \phi^2 \gamma_0 + \sigma_e^2 \\
\Rightarrow \gamma_0(1-\phi^2) &= \sigma_e^2 \\
\gamma_0 &= \frac{\sigma_e^2}{1-\phi^2}.
\end{aligned}$$

Note: We could also work out γ_0 by evaluating $\mathrm{Var}(Y_t)$.

Applying $\gamma_k = \phi\, \gamma_{k-1}$ then gives

$$\begin{aligned}
\gamma_1 &= \phi\, \gamma_0 = \phi\, \frac{\sigma_e^2}{1-\phi^2} \\
\gamma_2 &= \phi\, \gamma_1 = \phi^2\, \frac{\sigma_e^2}{1-\phi^2} \\
&\vdots \\
\gamma_k &= \phi\, \gamma_{k-1} = \phi^k\, \frac{\sigma_e^2}{1-\phi^2}
\end{aligned}$$

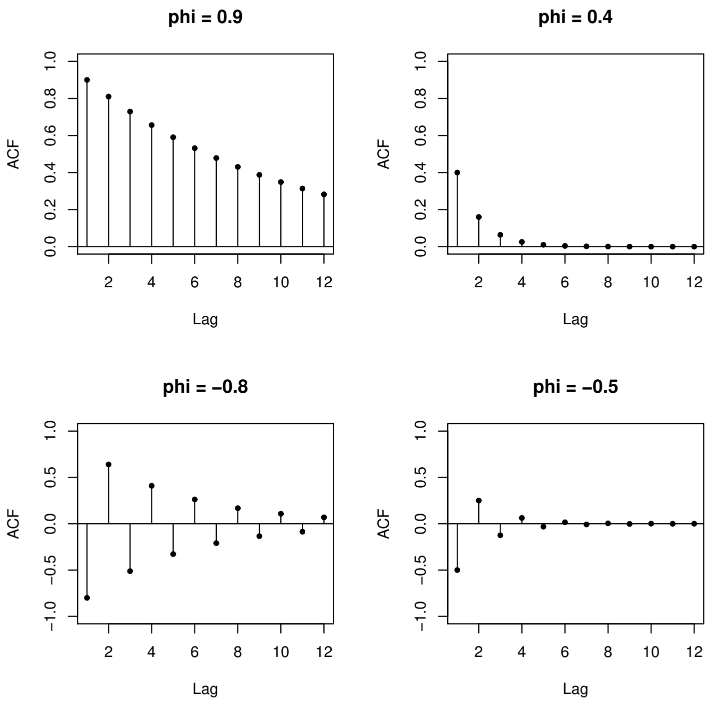
which is the same as what we found using general linear process theory.

The autocorrelation function is therefore

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\phi^k \frac{\sigma_e^2}{1-\phi^2}}{\frac{\sigma_e^2}{1-\phi^2}}$$
$$= \phi^k.$$

Again it is clear that we require $|\phi| < 1$ so that the autocorrelations do not exceed 1 in magnitude.

Since $|\phi| < 1$, the autocorrelations are decaying with k (starting from γ_0). If ϕ is positive, then all autocorrelations will be positive whereas, if ϕ is negative, the first autocorrelation will be negative and the signs of successive autocorrelations will alternate sign thereafter.



The above plots have been created using the `ARMAacf` function:

```
ACF <- ARMAacf(ar=c(0.9),lag.max=12)
plot(x=1:12, y=ACF[-1], xlab="Lag", ylab="ACF", type="h")
abline(h=0)
```

It is worth noting here that we can work out the autocorrelation function simply using $\gamma_k = \phi\, \gamma_{k-1}$ (for $k \geq 1$) without deriving an expression for γ_k itself:

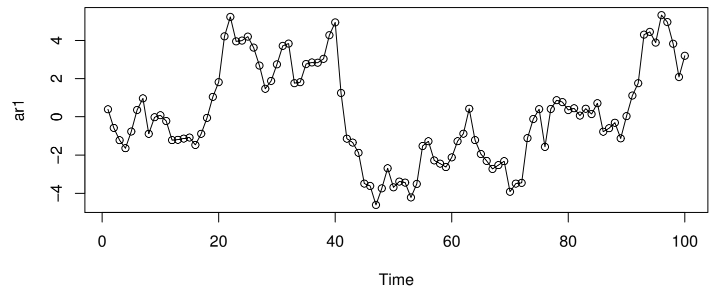
$$\gamma_k = \phi\, \gamma_{k-1}$$
$$\frac{\gamma_k}{\gamma_0} = \phi\, \frac{\gamma_{k-1}}{\gamma_0}$$
$$\Rightarrow \rho_k = \phi\, \rho_{k-1}.$$

Using this relationship we get

$$\rho_1 = \phi\, \rho_0 = \phi$$
$$\text{(since by definition } \rho_0 = 1 \text{ [Lecture 2])}$$
$$\rho_2 = \phi\, \rho_1 = \phi^2$$
$$\rho_3 = \phi\, \rho_2 = \phi^3$$
$$\vdots$$
$$\rho_k = \phi\, \rho_{k-1} = \phi^k.$$

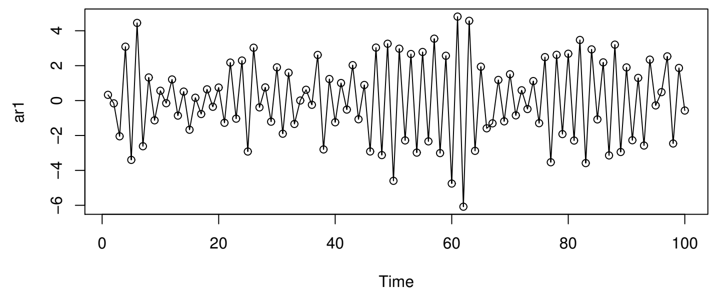
This is the best approach if we do not require an expression for γ_k.

Example 4.1. *AR(1) with* $\phi = 0.9$
With $\phi = 0.9$, the autocorrelations decay from $\rho_1 = 0.9$, i.e., $\rho_2 = 0.81$, $\rho_3 = 0.729$, etc. Thus neighbouring values will have a strong tendency to hang together leading to a very smooth time series plot:




```
set.seed(1211391)
ar1 <- arima.sim(n=100, model=list(ar=c(0.9)) )
dev.new(width=8, height=4)
plot(ar1, type="o")
```

Example 4.2. *AR(1) with* $\phi = -0.9$
With $\phi = -0.9$, the autocorrelations decay from $\rho_1 = -0.9$ but alternate sign, i.e., $\rho_2 = 0.81$, $\rho_3 = -0.729$, etc. This causes a strong oscillation effect in the series:



In contrast to the earlier MA(1) examples, we see that the AR(1) process can produce both smoother and wilder time series. This is to be expected since ρ_1 is restricted for the MA(1) process ($|\rho_1| < 0.5$) with no autocorrelations after lag 1. On the other hand ρ_1 is not restricted for the AR(1) process (apart from the usual $|\rho_1| < 1$) with autocorrelations persisting beyond lag 1.

4.2 AR(2)

The AR(2) process is given by

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t$$

which is clearly in the invertible form with $\pi_1 = \phi_1$, $\pi_2 = \phi_2$ and all other $\pi_i = 0$ for $i > 2$.

The AR(2) process can be written in general linear process form with constraints on the ϕ weights to ensure stationarity (see Section 4.4). In particular it can be shown (details omitted) that for stationarity we require: • $\phi_1 + \phi_2 < 1$ • $\phi_2 - \phi_1 < 1$ and • $|\phi_2| < 1$.

Although it is possible to convert the AR(2) process to general linear form, and then use the theory from Section 2 to derive the autocovariance function, it is much easier to do this directly:

$$\begin{aligned}\mathrm{Cov}(Y_t, Y_{t-k}) &= \mathrm{Cov}(\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t, Y_{t-k}) \\ &= \phi_1\,\mathrm{Cov}(Y_{t-1}, Y_{t-k}) + \phi_2\,\mathrm{Cov}(Y_{t-2}, Y_{t-k}) \\ &\qquad + \mathrm{Cov}(e_t, Y_{t-k}) \\ \Rightarrow \gamma_k &= \phi_1\,\gamma_{k-1} + \phi_2\,\gamma_{k-2} + \mathrm{Cov}(e_t, Y_{t-k})\end{aligned}$$

where, following the same line as in the AR(1) process, we have assumed stationarity.

By definition, e_t is independent of the past and, hence, $\mathrm{Cov}(e_t, Y_{t-k}) = 0$ for $k \geq 1$. Also, as in Section 4.1, it is easy to show that $\mathrm{Cov}(e_t, Y_t) = \sigma_e^2$. Hence,

$$\begin{aligned}\gamma_k &= \phi_1\,\gamma_{k-1} + \phi_2\,\gamma_{k-2} + \mathrm{Cov}(e_t, Y_{t-k}) \\ &= \begin{cases} \phi_1\,\gamma_1 + \phi_2\,\gamma_2 + \sigma_e^2 & k = 0 \\ \phi_1\,\gamma_{k-1} + \phi_2\,\gamma_{k-2} & k \geq 1 \end{cases}\end{aligned}$$

where, for $k = 0$ we have used the fact that $\gamma_{-1} = \gamma_1$ and $\gamma_{-2} = \gamma_2$.

From the above we have

$$\begin{aligned}\gamma_1 &= \phi_1\,\gamma_0 + \phi_2\,\gamma_{-1} \\ &= \phi_1\,\gamma_0 + \phi_2\,\gamma_1 \\ (1-\phi_2)\gamma_1 &= \phi_1\,\gamma_0 \\ \gamma_1 &= \frac{\phi_1}{1-\phi_2}\,\gamma_0\end{aligned}$$

and

$$\begin{aligned}\gamma_2 &= \phi_1\,\gamma_1 + \phi_2\,\gamma_0 \\ &= \frac{\phi_1^2}{1-\phi_2}\,\gamma_0 + \phi_2\,\gamma_0 \\ &= \left(\frac{\phi_1^2}{1-\phi_2} + \phi_2\right)\gamma_0 \\ &= \frac{\phi_1^2 + \phi_2 - \phi_2^2}{1-\phi_2}\,\gamma_0.\end{aligned}$$

Plugging these into the equation $\gamma_0 = \phi_1\,\gamma_1 + \phi_2\,\gamma_2 + \sigma_e^2$ allows us to solve for γ_0 yielding:

$$\gamma_0 = \frac{(1-\phi_2)\sigma_e^2}{1 - \phi_2 - \phi_1^2 - \phi_1^2\phi_2 - \phi_2^2 + \phi_2^3}$$

which can be substituted back into the expressions for γ_1 and γ_2. From here we can derive all γ_k values using $\gamma_k = \phi_1\,\gamma_{k-1} + \phi_2\,\gamma_{k-2}$ and, hence, all values of $\rho_k = \frac{\gamma_k}{\gamma_0}$.

If we are only interested in ρ_k (and not γ_k) then we have (for $k \geq 1$):

$$\begin{aligned}\gamma_k &= \phi_1\,\gamma_{k-1} + \phi_2\,\gamma_{k-2} \\ \frac{\gamma_k}{\gamma_0} &= \phi_1\,\frac{\gamma_{k-1}}{\gamma_0} + \phi_2\,\frac{\gamma_{k-2}}{\gamma_0} \\ \Rightarrow \rho_k &= \phi_1\,\rho_{k-1} + \phi_2\,\rho_{k-2}.\end{aligned}$$

From this we get

$$\begin{aligned}\rho_1 &= \phi_1\,\rho_0 + \phi_2\,\rho_{-1} \\ &= \phi_1\,(1) + \phi_2\,\rho_1 \\ (1-\phi_2)\,\rho_1 &= \phi_1 \\ \rho_1 &= \frac{\phi_1}{1-\phi_2}\end{aligned}$$

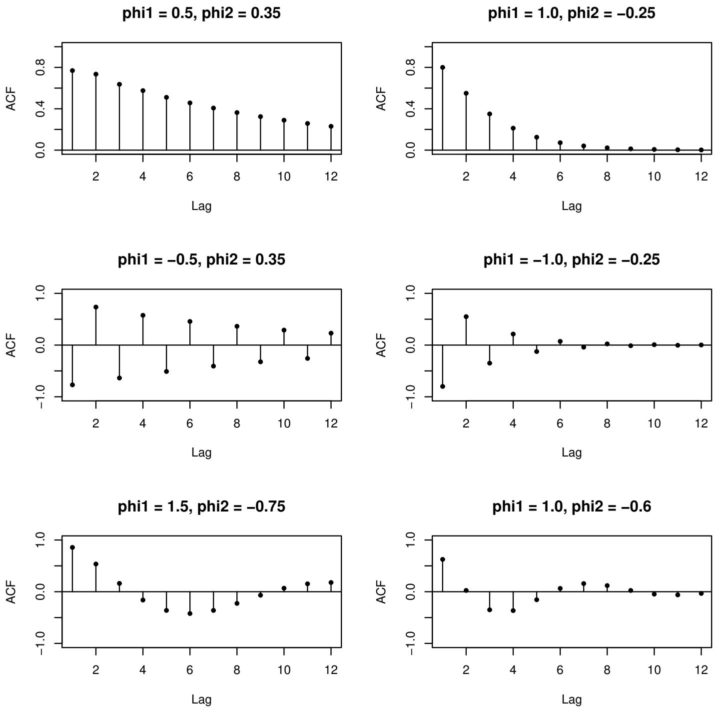
and, hence,

$$\begin{aligned}\rho_2 &= \phi_1\,\rho_1 + \phi_2\,\rho_0 \\ &= \frac{\phi_1^2}{1-\phi_2} + \phi_2.\end{aligned}$$

Continuing in this way, all ρ_k values can be determined. In fact, an explicit solution for ρ_k can be derived by solving the *recurrence relation* $\rho_k = \phi_1\,\rho_{k-1} + \phi_2\,\rho_{k-2}$ with the above expressions for ρ_1 and ρ_2 (we will not pursue this).

The autocorrelation function can take a variety of shapes which are similar to (but more general than) those from the AR(1) case. In particular, the autocorrelations decay over time (and may oscillate).

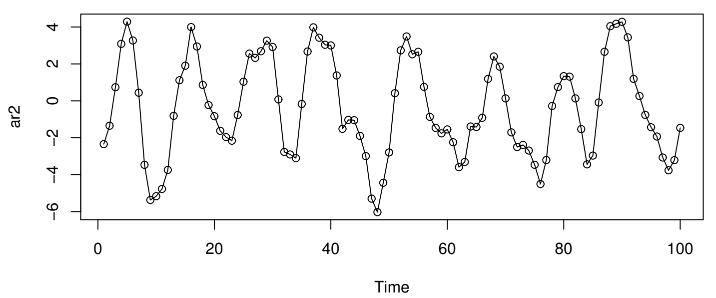
In addition to this, the autocorrelation function for the AR(2) process can also behave as a damped sine wave (when the roots of the *AR characteristic equation* are complex). This produces cyclic behaviour in the series which may look somewhat similar to seasonality (even though no seasonality exists). This is apparent in the next example.



The above plots have been created using the `ARMAacf` function:

```
ACF <- ARMAacf(ar=c(0.5,0.35), lag.max=12)
plot(x=1:12, y=ACF[-1], xlab="Lag", ylab="ACF", type="h")
abline(h=0)
```

Example 4.3. *AR(2) with* $\phi_1 = 1.5$ *and* $\phi_2 = -0.75$
It is clear that the autocorrelation structure for this process produces cyclic behaviour in the series:




```
set.seed(9271231)
ar2 <- arima.sim(n=100, model=list(ar=c(1.5,-0.75)) )
dev.new(width=8, height=4)
plot(ar2, type="o")
```

4.3 AR(p)

The general AR(p) process is given by

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t$$

which is in the invertible form with $\pi_i = \phi_i$ for $i = 1, \ldots, p$. Thus, **all AR processes are invertible**.

This process can be written in terms of the backshift operator

$$\begin{aligned} Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \cdots - \phi_p Y_{t-p} &= e_t \\ (1 - \phi_1 B^1 - \phi_2 B^2 - \cdots - \phi_p B^p)\, Y_t &= e_t \\ \phi(B) Y_t &= e_t \end{aligned}$$

where $\phi(B) = 1 - \phi_1 B^1 - \phi_2 B^2 - \cdots - \phi_p B^p$ is known as the **AR characteristic polynomial**.

Note that for $k \geq 1$ we have (assuming stationarity as before):

$$\begin{aligned} \gamma_k &= \mathrm{Cov}(Y_t, Y_{t-k}) \\ &= \mathrm{Cov}(\phi_1 Y_{t-1} + \phi_1 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t, Y_{t-k}) \\ &= \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p} \end{aligned}$$

where $\mathrm{Cov}(e_t, Y_{t-k}) = 0$ since $k \geq 1$.

Dividing through by γ_0 gives

$$\rho_k = \phi_1 \rho_{k-1} + \phi_1 \rho_{k-2} + \cdots + \phi_p \rho_{k-p}$$

which is an important *recurrence relation* for calculating ρ_k values.

First, setting $k = 1, 2, \ldots, p$ gives a system of equations called the **Yule-Walker equations**:

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2\, \rho_1 + \cdots + \phi_p\, \rho_{p-1} \\ \rho_2 &= \phi_1\, \rho_1 + \phi_2 + \cdots + \phi_p\, \rho_{p-2} \\ &\vdots \\ \rho_p &= \phi_1\, \rho_{p-1} + \phi_2\, \rho_{p-2} + \cdots + \phi_p \end{aligned}$$

For a given set of ϕ values, these equations can be solved to produce $\rho_1, \rho_2, \ldots, \rho_p$ (in fact we have done this for the AR(1) and AR(2) cases). Once we have these, we can compute further ρ_k values using the recurrence relation.

Note that when $k = 0$ we get

$$\begin{aligned} \gamma_0 &= \mathrm{Cov}(\phi_1 Y_{t-1} + \phi_1 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t, Y_t) \\ &= \phi_1 \gamma_1 + \phi_2 \gamma_2 + \cdots + \phi_p \gamma_p + \sigma_e^2. \end{aligned}$$

where it is easy to show (as in Section 4.1) that $\mathrm{Cov}(e_t, Y_t) = \sigma_e^2$. Dividing by γ_0 gives

$$1 = \phi_1 \rho_1 + \phi_1 \rho_2 + \cdots + \phi_p \rho_p + \frac{\sigma_e^2}{\gamma_0}$$

$$\Rightarrow \gamma_0 = \frac{\sigma_e^2}{1 - \phi_1 \rho_1 - \phi_1 \rho_2 - \cdots - \phi_p \rho_p}$$

which is useful if we require the series variance $\gamma_0 = \mathrm{Var}(Y_t)$.

4.4 Stationarity

Consider the AR(1) process

$$Y_t = \phi Y_{t-1} + e_t$$

We need to constrain ϕ to ensure that the series is stationary. For example, $\phi = 1$ leads to a random walk which we already know is non-stationary (see Lecture 2).

We aim to write the AR(1) process as a general linear process, i.e.,

$$\begin{aligned} Y_t &= (1 + \psi_1 B^1 + \psi_2 B^2 + \cdots)\, e_t \\ &= \psi(B)\, e_t. \end{aligned}$$

This can be done by successively back-substituting as follows:

$$\begin{aligned} Y_t &= e_t + \phi Y_{t-1} \\ &= e_t + \phi\,(e_{t-1} + \phi Y_{t-2}) \\ &= e_t + \phi e_{t-1} + \phi^2 Y_{t-2} \\ &= e_t + \phi e_{t-1} + \phi^2 (e_{t-2} + \phi Y_{t-3}) \\ &= e_t + \phi e_{t-1} + \phi^2 e_{t-2} + \cdots \\ &= (1 + \phi B^1 + \phi^2 B^2 + \cdots)\, e_t \\ &= \psi(B)\, Y_t \end{aligned}$$

where $\psi_i = \phi^i$. In order for the above to be meaningful, the sum $1 + \phi B^1 + \phi^2 B^2 + \cdots$ must *converge* which clearly requires $|\phi| < 1$.

Alternatively we could have found $\psi_i = \phi^i$ by noting that the AR(1) process can be written as $(1 - \phi B) Y_t = e_t$ and proceeding as follows:

$$\begin{aligned} Y_t &= \psi(B) e_t \\ Y_t &= \psi(B)(1 - \phi B) Y_t \\ 1 &= \psi(B)(1 - \phi B). \\ 1 + 0B^1 + 0B^2 + \cdots &= \psi(B)(1 - \phi B) \end{aligned}$$

We then multiply out $\psi(B)(1 - \phi B)$ and match the coefficients of B^i with those on the left hand side.

General Condition for Stationarity

We have only considered the specific case of the AR(1) process above. However, in general, the condition for stationarity is that **the roots of the AR characteristic equation**

$$1 - \phi_1 B^1 - \phi_2 B^2 - \ldots - \phi_q B^q = 0$$

must have modulus greater than one. The proof of this result is beyond the scope of this course. Note that this is the same as the condition for invertibility of the MA process.

Consider again the AR(1) process whose characteristic equation is

$$\begin{aligned} 1 - \phi B &= 0 \\ -\phi B &= -1 \\ B &= \frac{1}{\phi} \end{aligned}$$

For stationarity we require

$$\left| \frac{1}{\phi} \right| > 1$$

$$\Rightarrow |\phi| < 1$$

which is the same as what we found by manually writing the process as a general linear process.

In the AR(2) case, the characteristic equation is

$$1 - \phi_1 B^1 - \phi_2 B^2 = 0$$

which is a quadratic in B and, hence, the roots are

$$\frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}.$$

It can be shown that the requirement that these roots have modulus greater than one is equivalent to requiring that $\phi_1 + \phi_2 < 1, \quad \phi_2 - \phi_1 < 1 \quad \text{and} \quad |\phi_2| < 1$.

5 Autoregressive Moving Average Process: ARMA(p, q)

The ARMA model contains both AR and MA components and can, therefore, handle a wide variety of different correlation structures.

5.1 ARMA(1,1)

The ARMA(1,1) process is given by

$$Y_t = \phi Y_{t-1} + e_t - \theta e_{t-1}.$$

We can write this as a general linear process by successively back-substituting as follows:

$$\begin{aligned}
Y_t &= e_t - \theta e_{t-1} + \phi Y_{t-1} \\
&= e_t - \theta e_{t-1} + \phi(e_{t-1} - \theta e_{t-2} + \phi Y_{t-2}) \\
&= e_t + (\phi - \theta)e_{t-1} - \phi\theta e_{t-2} + \phi^2 Y_{t-2} \\
&= e_t + (\phi - \theta)e_{t-1} - \phi\theta e_{t-2} + \phi^2(e_{t-2} - \theta e_{t-3} + \phi Y_{t-3}) \\
&= e_t + (\phi - \theta)e_{t-1} + \phi(\phi - \theta)e_{t-2} - \phi^2\theta e_{t-3} + \phi^3 Y_{t-3} \\
&= e_t + (\phi - \theta)e_{t-1} + \phi(\phi - \theta)e_{t-2} - \phi^2\theta e_{t-3} \\
&\quad + \phi^3(e_{t-3} - \theta e_{t-4} + \phi Y_{t-4}) \\
&= e_t + (\phi - \theta)e_{t-1} + \phi(\phi - \theta)e_{t-2} + \phi^2(\phi - \theta)e_{t-3} + \cdots
\end{aligned}$$

which converges if $|\phi| < 1$ (note: same as the AR(1) case). Therefore, $\psi_0 = 1$ and $\psi_j = \phi^{j-1}(\phi - \theta)$ for $j \geq 1$.

Using the results from Section 2, we can work out the autocovariance

$$\begin{aligned}
\gamma_k = \sigma_e^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k} &= \sigma_e^2 \left(\psi_0 \psi_k + \sum_{j=1}^{\infty} \psi_j \psi_{j+k} \right) \\
&= \sigma_e^2 \left((1)\psi_k + \sum_{j=1}^{\infty} \phi^{j-1}(\phi - \theta)\phi^{j+k-1}(\phi - \theta) \right) \\
&= \sigma_e^2 \left(\psi_k + (\phi - \theta)^2 \phi^{k-2} \sum_{j=1}^{\infty} (\phi^2)^j \right) \\
&= \sigma_e^2 \left(\psi_k + (\phi - \theta)^2 \phi^{k-2} \frac{\phi^2}{1 - \phi^2} \right) \\
&\qquad \text{(geometric sum with } a = \phi^2 \text{ and } r = \phi^2\text{)} \\
&= \sigma_e^2 \left(\psi_k + \frac{(\phi - \theta)^2 \phi^k}{1 - \phi^2} \right) \\
&= \sigma_e^2 \frac{\psi_k(1 - \phi^2) + (\phi - \theta)^2 \phi^k}{1 - \phi^2}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\gamma_0 &= \sigma_e^2 \frac{(1)(1 - \phi^2) + (\phi - \theta)^2}{1 - \phi^2} \\
&= \sigma_e^2 \frac{1 - \phi^2 + \phi^2 - 2\phi\theta + \theta^2}{1 - \phi^2} \\
&= \sigma_e^2 \frac{1 - 2\phi\theta + \theta^2}{1 - \phi^2}
\end{aligned}$$

and, for $k \geq 1$,

$$\begin{aligned}
\gamma_k &= \sigma_e^2 \frac{\phi^{k-1}(\phi - \theta)(1 - \phi^2) + (\phi - \theta)^2 \phi^k}{1 - \phi^2} \\
&= \sigma_e^2 \frac{\phi^{k-1}(\phi - \theta)[1 - \phi^2 + (\phi - \theta)\phi]}{1 - \phi^2} \\
&= \phi^{k-1} \sigma_e^2 \frac{(\phi - \theta)[1 - \phi^2 + \phi^2 - \theta\phi]}{1 - \phi^2} \\
&= \phi^{k-1} \sigma_e^2 \frac{(\phi - \theta)(1 - \theta\phi)}{1 - \phi^2}
\end{aligned}$$

Alternatively, we can work out the acvf directly (following a similar route to Section 4.1):

$$\begin{aligned}
\mathrm{Cov}(Y_t, Y_{t-k}) &= \mathrm{Cov}(\phi Y_{t-1} + e_t - \theta e_{t-1}, Y_{t-k}) \\
&= \phi\, \mathrm{Cov}(Y_{t-1}, Y_{t-k}) + \mathrm{Cov}(e_t, Y_{t-k}) \\
&\quad - \theta \mathrm{Cov}(e_{t-1}, Y_{t-k})
\end{aligned}$$

$$\Rightarrow \gamma_k = \phi\, \gamma_{k-1} + \mathrm{Cov}(e_t, Y_{t-k}) - \theta \mathrm{Cov}(e_{t-1}, Y_{t-k})$$

where, as before, we have assumed stationarity to simplify $\mathrm{Cov}(Y_t, Y_{t-k})$ and $\mathrm{Cov}(Y_{t-1}, Y_{t-k})$.

Consider the term $\mathrm{Cov}(e_t, Y_{t-k})$. This equals zero for $k \geq 1$ as e_t is independent of the past. For $k = 0$

$$\begin{aligned}
&\mathrm{Cov}(e_t, Y_t) \\
&= \phi\, \mathrm{Cov}(e_t, Y_{t-1}) + \mathrm{Cov}(e_t, e_t) - \theta \mathrm{Cov}(e_t, e_{t-1}) \\
&= \phi\,(0) + \mathrm{Var}(e_t) - \theta(0) \\
&= \sigma_e^2.
\end{aligned}$$

Now consider the term $\mathrm{Cov}(e_{t-1}, Y_{t-k})$. This equals zero for $k \geq 2$ as e_{t-1} is independent of the past. For $k = 1$

$$\begin{aligned}
&\mathrm{Cov}(e_{t-1}, Y_{t-1}) \\
&= \phi\, \mathrm{Cov}(e_{t-1}, Y_{t-2}) + \mathrm{Cov}(e_{t-1}, e_{t-1}) - \theta \mathrm{Cov}(e_{t-1}, e_{t-2}) \\
&= \phi\,(0) + \mathrm{Var}(e_{t-1}) - \theta(0) \\
&= \sigma_e^2
\end{aligned}$$

and for $k = 0$ we have

$$\begin{aligned}
&\mathrm{Cov}(e_{t-1}, Y_t) \\
&= \phi\, \mathrm{Cov}(e_{t-1}, Y_{t-1}) + \mathrm{Cov}(e_{t-1}, e_t) - \theta \mathrm{Cov}(e_{t-1}, e_{t-1}) \\
&= \phi\, \sigma_e^2 + 0 - \theta\sigma_e^2 \\
&= \sigma_e^2(\phi - \theta)
\end{aligned}$$

Thus, we have the following

$$\begin{aligned}
\gamma_0 &= \phi\, \gamma_1 + \sigma_e^2[1 - \theta(\phi - \theta)] \\
\gamma_1 &= \phi\, \gamma_0 - \theta\sigma_e^2 \\
\gamma_k &= \phi\, \gamma_{k-1} \qquad\qquad k \geq 2
\end{aligned}$$

Using the above we have

$$\begin{aligned}
\gamma_0 &= \phi\, \gamma_1 + \sigma_e^2[1 - \theta(\phi - \theta)] \\
\Rightarrow \gamma_0 &= \phi\,(\phi\, \gamma_0 - \theta\sigma_e^2) + \sigma_e^2[1 - \theta(\phi - \theta)] \\
(1 - \phi^2)\gamma_0 &= \sigma_e^2[1 - \phi\theta - \theta(\phi - \theta)] \\
(1 - \phi^2)\gamma_0 &= \sigma_e^2[1 - 2\phi\theta + \theta^2] \\
\gamma_0 &= \sigma_e^2 \frac{1 - 2\phi\theta + \theta^2}{1 - \phi^2}
\end{aligned}$$

$$
\begin{aligned}
\gamma_1 &= \phi\,\gamma_0 - \theta\sigma_e^2 \\
\Rightarrow \gamma_1 &= \phi\,\sigma_e^2 \frac{1-2\phi\theta+\theta^2}{1-\phi^2} - \theta\sigma_e^2 \\
&= \sigma_e^2 \frac{\phi - 2\phi^2\theta + \phi\theta^2 - \theta + \phi^2\theta}{1-\phi^2} \\
&= \sigma_e^2 \frac{\phi - \phi^2\theta - \theta + \phi\theta^2}{1-\phi^2} \\
&= \sigma_e^2 \frac{\phi(1-\phi\theta) - \theta(1-\phi\theta)}{1-\phi^2} \\
&= \sigma_e^2 \frac{(\phi-\theta)(1-\phi\theta)}{1-\phi^2}
\end{aligned}
$$

$$
\gamma_k = \phi\,\gamma_{k-1} \qquad k \geq 2
$$

$$
\Rightarrow \gamma_2 = \phi\,\gamma_1 = \phi\sigma_e^2 \frac{(\phi-\theta)(1-\phi\theta)}{1-\phi^2}
$$

$$
\gamma_3 = \phi\,\gamma_2 = \phi^2\sigma_e^2 \frac{(\phi-\theta)(1-\phi\theta)}{1-\phi^2}
$$

$$
\vdots
$$

$$
\gamma_k = \phi\,\gamma_{k-1} = \phi^{k-1}\sigma_e^2 \frac{(\phi-\theta)(1-\phi\theta)}{1-\phi^2}
$$

Notice this formula works for γ_1 too. This result is the same as what we found using general linear process theory.

Dividing by γ_0 gives the autocorrelation function (for $k \geq 1$)

$$
\begin{aligned}
\rho_k = \frac{\gamma_k}{\gamma_0} &= \frac{\phi^{k-1}\sigma_e^2 \frac{(\phi-\theta)(1-\phi\theta)}{1-\phi^2}}{\sigma_e^2 \frac{1-2\phi\theta+\theta^2}{1-\phi^2}} \\
&= \frac{\phi^{k-1}(\phi-\theta)(1-\phi\theta)}{1-2\phi\theta+\theta^2}.
\end{aligned}
$$

5.2 ARMA(p, q)

The general ARMA(p,q) process is given by

$$
\begin{aligned}
Y_t = {} & \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} \\
& + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q}
\end{aligned}
$$

i.e., it has **p AR terms and q MA terms**.

Note that the process can be written in terms of the backshift operator as follows:

$$
\begin{aligned}
& Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \cdots - \phi_p Y_{t-p} \\
& \quad = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q} \\
& (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)\, Y_t \\
& \quad = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)\, e_t
\end{aligned}
$$

$$
\Rightarrow \phi(B)\,Y_t = \theta(B)\,e_t
$$

where

$$
\phi(B) = 1 - \phi_1 B^1 - \phi_2 B^2 - \cdots - \phi_p B^p
$$

is the AR characteristic polynomial and

$$
\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q
$$

is the MA characteristic polynomial.

The ARMA(p,q) is a flexible model which can handle a wide variety of autocorrelation structures.

5.3 Invertibility, Stationarity and Parameter Redundancy

AR processes are always invertible. Thus, for the ARMA process, *invertibility depends only on the invertibility of the MA component*:

- **The ARMA process is invertible if the roots of $\theta(B)$ have modulus greater than one.**

On the other hand, MA processes are always stationary. Thus, for the ARMA process, *stationarity depends only on the stationarity of the AR component*:

- **The ARMA process is stationary if the roots of $\phi(B)$ have modulus greater than one.**

We must also be aware of parameter redundancy (i.e., more parameters than needed) in ARMA models. Consider the white noise process $Y_t = e_t$. Applying $1 - 0.7B$ to both sides gives:

$$
\begin{aligned}
Y_t - 0.7Y_{t-1} &= e_t - 0.7e_{t-1} \\
Y_t &= 0.7Y_{t-1} + e_t - 0.7e_{t-1}
\end{aligned}
$$

which looks like an ARMA(1,1) model. However, we know that it is a white noise process. If we did not know this, note that when we write the model in the form $\phi(B)\,Y_t = \theta(B)\,e_t$ we get

$$
\begin{aligned}
(1 - 0.7B)Y_t &= (1 - 0.7B)e_t \\
Y_t &= e_t
\end{aligned}
$$

by dividing both sides by $1 - 0.7B$.

- **The polynomials $\phi(B)$ and $\theta(B)$ must have no common factors to ensure there is no parameter redundancy.**