AR, MA and ARMA models

- The autoregressive process of order p or $AR(p)$ is defined by the equation

$$X_t = \sum_{j=1}^{p} \phi_j X_{t-j} + \omega_t$$

  where $\omega_t \sim N(0, \sigma^2)$

- $\phi = (\phi_1, \phi_2, \ldots, \phi_p)$ is the vector of model coefficients and p is a non-negative integer.

- The AR model establishes that a realization at time t is a linear combination of the p previous realization plus some noise term.

- For $p = 0$, $X_t = \omega_t$ and there is no autoregression term.

- The lag operator is denoted by B and used to express lagged values of the process so $BX_t = X_{t-1}$, $B^2X_t = X_{t-2}$, $B^3X_t = X_{t-3}$,..., B^dX_{t-d}.
- If we define

$$\Phi(B) = 1 - \sum_{j=1}^{p} \phi_j B^j = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$$

  the AR(p) process is given by the equation $\Phi(B)X_t = \omega_t; t = 1, \ldots, n.$
- $\Phi(B)$ is known as the *characteristic* polynomial of the process and its roots determine when the process is stationary or not.
- The moving average process of order q or $MA(q)$ is

defined as

$$X_t = \omega_t + \sum_{j=1}^{q} \theta_j \omega_{t-j}$$

- Under this model, the observed process depends on previous $\omega_t' s$
- $MA(q)$ can define correlated noise structure in our data and goes beyond the traditional assumption where errors are iid.
- In lag operator notation, the $MA(q)$ process is given by the equation $X_t = \Theta(B)\omega_t$ where $\Theta(B) = 1 + \sum_{j=1}^{q} \theta_j B^j$.
- The general autoregressive moving average process of orders p and q or $ARMA(p, q)$ combines both AR and MA models into a unique representation.

- The ARMA process of orders p and q is defined as

$$X_t = \sum_{j=1}^{p} \phi_j X_{t-j} + \sum_{j=1}^{q} \theta_j \omega_{t-j} + \omega_t$$

- In lag operator notation, the $ARMA(p, q)$ process is given by $\Phi(B)X_t = \Theta(B)\omega_t, t = 1, \ldots, n$
- Lets focus on the AR process and its characteristic polynomial.
- The characteristic polynomial can be expressed as:

$$\Phi(B) = \prod_{i=1}^{p} (1 - \alpha_i B)$$

where the $\alpha' s$ are the reciprocal roots.

- If $\beta_1, \beta_2, \ldots, \beta_p$ are such that $\Phi(\beta_i) = 0$ (roots of the polynomial) then $\beta_1 = 1/\alpha_1, \beta_2 = 1/\alpha_2, \ldots, \beta_p = 1/\alpha_p$

- *Theorem:* If $X_t \sim AR(p)$, X_t is a stationary process if and only if the modulus of all the roots of the characteristic polynomial are greater than one, i.e. if $||\beta_i|| > 1$ for all $i = 1, 2, \ldots, p$ or equivalently if $||\alpha_i|| < 1, i = 1, 2, \ldots p$.
- The $\alpha_i' s$ are also known as the *poles* of the AR process.
- This theorem follows from the *general linear process* theory.
- Some of the poles or reciprocal roots can be real number and some can be complex numbers and we need to distinguish between the 2 cases.

- For the complex case, we will use the representation

$$\alpha_i = r_i exp(\pm\omega_i i), i = 1, \ldots, C$$

so C is the total number of conjugate pairs and $2C$ is the total number of complex poles.

- r_i is the modulus of α_i and ω_i its frequency.
- The real reciprocal roots are denoted as

$$\alpha_i = r_i, i = 1, 2, \ldots R$$

- Example. Consider the AR(1) process $X_t = \phi X_{t-1} + \omega_t$.
- In lag-operator notation this process is $(1 - \phi B)X_t = \omega_t$ and the characteristic polynomial is $\Phi(B) = (1 - \phi B)$.
- If $\Phi(B) = (1 - \phi B) = 0$, the only characteristic root is

$\beta = 1/\phi$ (assuming $\phi \neq 0$).

- The AR(1) process is stationary if only if $|\phi| < 1$ or $-1 < \phi < 1$.
- The case where $\phi = 1$ corresponds to a Random Walk process with a zero drift, $X_t = X_{t-1} + \omega_t$
- This is a non-stationary explosive process.
- If we recursive apply the AR(1) equation, the Random Walk process can be expressed as $X_t = \omega_t + \omega_{t-1} + \omega_{t-2} + \dots$. Then, $Var(X_t) = \sum_{t=0}^{\infty} \sigma^2 = \infty$.
- Example. AR(2) process $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \omega_t$
- The characteristic polynomial is now

$\Phi(B) = (1 - \phi_1 B - \phi_2 B^2)$

- The solutions to $\Phi(B) = 0$ are

$$\beta_1 = \frac{-\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2}; \beta_2 = \frac{-\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2}$$

- The reciprocal roots are

$$\alpha_1 = \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2}; \alpha_2 = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

- The AR(2) is stationary if and only if $||\alpha_1|| < 1$ and $||\alpha_2|| < 1$
- These two conditions imply that $||\alpha_1 \alpha_2|| = |\phi_2| < 1$ and $||\alpha_1 + \alpha_2|| = |\phi_1| < 2$ which means $-1 < \phi_2 < 1$ and $-2 < \phi_1 < 2$.

- For α_1 and α_2 real numbers, $\phi_1^2 + 4\phi_2 \geq 0$ which implies $-1 < \alpha_2 \leq \alpha_1 < 1$ and after some algebra $\phi_1 + \phi_2 < 1$; $\phi_2 - \phi_1 < 1$
- In the complex case $\phi_1^2 + 4\phi_2 < 0$ or $\frac{\phi_1^2}{-4} > \phi_2$
- If we combine all the inequalities we obtain a region bounded by the lines $\phi_2 = 1 + \phi_1$; $\phi_2 = 1 - \phi_1$; $\phi_2 = -1$.
- This is the region where the AR(2) process is stationary.
- For an AR(p) where $p \geq 3$, the region where the process is stationary is quite abstract.
- For the stationarity condition of the MA(q) process, we need to rely on the *general linear process.*
- A general linear process is a random sequence X_t of the

form,

$$X_t = \sum_{j=0}^{\infty} a_j \omega_{t-j}$$

where ω_t is a white noise sequence with variance σ^2.

- In lag operator notation, the general linear is given by the expression $X_t = \Phi(B)^{-1}\omega_t$ where $\Phi(B)^{-1} = \sum_{j=0}^{\infty} a_j B^j$.
- Note firstly that by the definition of the linear process, $E(X_t) = 0$.
- Then, the covariance between X_t and X_s is

$$E[X_t X_s] = \sum_{j=0}^{\infty}\sum_{l=0}^{\infty} a_j a_l E[X_{t-j} X_{s-l}]$$

$$= \sigma^2 \sum_{j=0}^{\infty} a_j a_{j+s-t} \quad ; (s \geq t)$$

- The last expression depends on t and s only through the difference $s - t$. Therefore, the process is stationary if $\sum_{j=0}^{\infty} a_j a_{j+k}$ is finite for all non-negative integers k
- Setting $k = 0$ we require that $\sum_{j=0}^{\infty} a_j^2 < \infty$
- Given that a correlation is always between -1 and 1,

  $$|\gamma_k| \leq \gamma_0$$

  so if $\gamma_0 < \infty$ then $\sum_{j=0}^{\infty} a_j a_{j+k} < \infty$.
- Then X_t is stationary if and only if $\sum_{j=0}^{\infty} a_j^2 < \infty$
- The MA(q) process can be written as a **general linear**

process of the form $X_t = \sum_{j=0}^{\infty} a_j \omega_{t-j}$ where

$$a_j = \begin{cases} \theta_j & \text{j=0,}\ldots\text{,q} \\ 0 & \text{j=q+1,q+2,}\ldots \end{cases}$$

with $\theta_0 = 1$.

- For the MA(q) process $\sum_{j=0}^{\infty} a_j^2 = \sum_{j=0}^{q} \theta_j^2 < \infty$ so a moving average process is always stationary.
- For the ARMA(p,q) process given by $\Phi(B)X_t = \Theta(B)\omega_t$ X_t is stationary if only if the roots of $\Phi(B) = 0$ have all modulus greater than 1 or all the reciprocal roots have a modulus less than one.
- A related concept to stationary linear process is invertible process.

- *Definition:* A process X_t is invertible if

$$X_t = \sum_{j=1}^{\infty} a_j X_{t-j} + \omega_t$$

with the restriction that $\sum_{j=1}^{\infty} a_j^2 < \infty$

- Basically, an invertible process is an infinite autoregression.

- By definition the AR(p) is invertible. We can set $a_1 = \phi_1, a_2 = \phi_2, \ldots a_p = \phi_p$ and $a_j = 0, j > p$. Then $\sum_{j=1}^{\infty} a_j^2 = \sum_{=1}^{p} \phi_j^2$ which is finite.

- For an MA(q) process we have $X_t = \Theta(B)\omega_t$. If we find a polynomial $\Theta(B)^{-1}$ such that $\Theta(B)\Theta(B)^{-1} = 1$ then we can invert the process since $\Theta(B)^{-1}X_t = \omega_t$

- The MA(q) process is invertible if and only if the roots of $\Theta(B)$ have all modulus greater than one.
- To illustrate this last point consider the MA(1) process $X_t = (1 - \theta B)\omega_t$
- If $|\theta| < 1$ then

$$\Theta(B)^{-1} = \frac{1}{(1 - \theta B)} = \sum_{j=0}^{\infty} \theta^j B^j$$

- Since $|\theta| < 1$ then $\sum_{j=0}^{\infty} \theta^j < \infty$ and so the process is invertible and has the representation

$$X_t = \sum_{j=1}^{\infty} \theta^j X_{t-j} + \omega_t$$

- The ARMA(p,q) process is invertible whenever the MA part of the process is invertible, i.e. when $\Theta(B)$ has reciprocal roots with modulus less than one.

Autocorrelation and Partial Autocorrelation

- The partial autocorrelation function (PACF) of a process Z_t is defined as $P_k = Corr(Z_t, Z_{t+k} | Z_{t+1}, \ldots, Z_{t+k-1}); k = 0, 1, 2, 3, \ldots$
- This PACF is equal to the ordinary correlation between $Z_t - \hat{Z}_t$ and $Z_{t+k} - \hat{Z_{t+k}}$ where $\hat{Z}_t$ and $\hat{Z_{t+k}}$ are the "best" linear estimators for Z_t and Z_{t+k} respectively.
- This PACF can also be derived through an autoregressive model of order k

$$Z_{t+k} = \phi_{k1} Z_{t+k-1} + \phi_{k2} Z_{t+k-2} + \ldots + \phi_{kk} Z_t + \omega_{t+k}$$

- The coefficients $\phi_{k1}, \phi_{k2}, \ldots, \phi_{kk}$ define the PACF.
- We have a set of linear equations for which the solution can be obtained via Cramer's Rule.
- R/S-plus include an option to compute the PACF.

```
> acf(x,type=''partial'')
```

- The partial autocorrelation can be derived as follows. Suppose that Z_t is zero mean stationary process.
- Consider a regression model where Z_{t+k} is regressed on k lagged variables $Z_{t+k-1}, Z_{t+k-2}, \ldots, Z_t$, i.e.,

$$Z_{t+k} = \phi_{k1} Z_{t+k-1} + \phi_{k2} Z_{t+k-2} + \ldots + \phi_{kk} Z_t + \omega_{t+k}$$

- ϕ_{ki} denotes the i-th regression parameter and ω_{t+k} is a normal error term uncorrelated with Z_{t+k-j} for $j \geq 1$.
- Multiplying Z_{t+k-j} on both sides of the above regression equation and taking the expectation, we get

$$\gamma_j = \phi_{k1} \gamma_{j-1} + \phi_{k2} \gamma_{j-2} + \ldots + \phi_{kk} \gamma_{j-k}$$

- If we divide by γ_0 we get,

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \ldots + \phi_{kk}\rho_{j-k}$$

- For $j = 1, 2, \ldots, k$, we have the following system of equations:

$$\begin{aligned}
\rho_1 &= \phi_{k1}\rho_0 + \phi_{k2}\rho_1 + \ldots + \phi_{kk}\rho_{k-1} \\
\rho_2 &= \phi_{k1}\rho_1 + \phi_{k2}\rho_0 + \ldots + \phi_{kk}\rho_{k-2} \\
&\vdots \\
\rho_k &= \phi_{k1}\rho_{k-1} + \phi_{k2}\rho_{k-2} + \ldots + \phi_{kk}\rho_0
\end{aligned}$$

- Using Cramer's rule successively for $k = 1, 2, \ldots$, we have

$$\phi_{11} = \rho_1$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_1 \end{vmatrix}}$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}$$

$$
\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}}
$$

- As a function of k, ϕ_{kk} is usually referred to as the partial autocorrelation function (PACF).
- A computer package will produce an estimate of $\phi_{k,k}$ using $\hat{\rho}_k$

- **Example**. Consider the stationary AR(1) process , $X_t = \alpha X_{t-1} + \omega_t$ with $-1 < \alpha < 1$. (change ϕ to α)
- Previously, we establish that the autocorrelation function for an AR(1) is $\rho_k = \alpha^k$.
- If we apply Cramer's rule $\phi_{11} = \rho_1 = \alpha$.
- Also

$$\phi_{22} \quad = \quad \frac{\begin{vmatrix} 1 & \alpha \\ \alpha & \alpha^2 \end{vmatrix}}{\begin{vmatrix} 1 & \alpha \\ \alpha & 1 \end{vmatrix}} = 0$$

- In fact, it can be checked that $\phi_{kk} = 0$ for any $k \geq 2$.

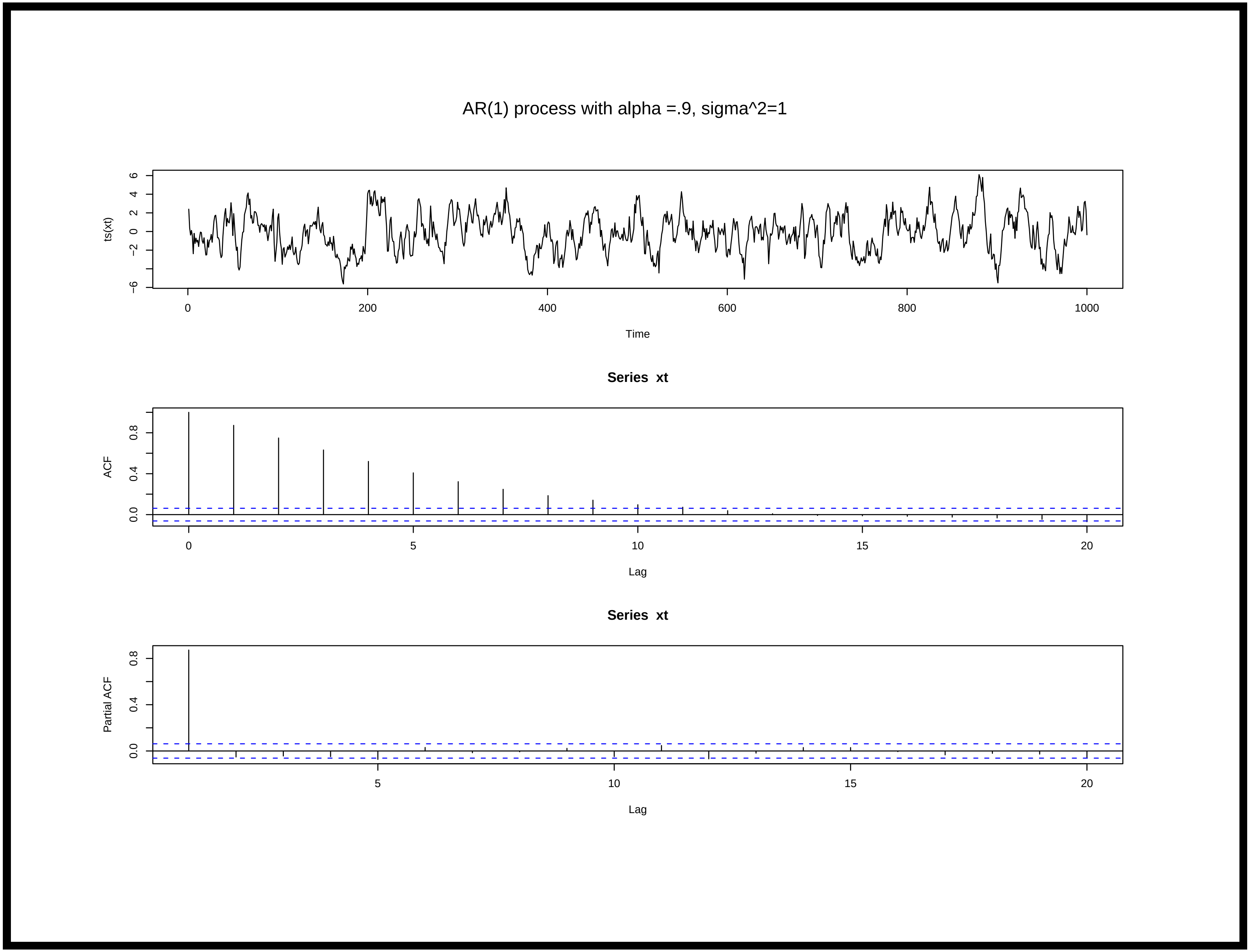
- The result is that for the AR(1)

$$\phi_{kk} = \begin{cases} \alpha & \text{k=1} \\ 0 & k \geq 2 \end{cases}$$

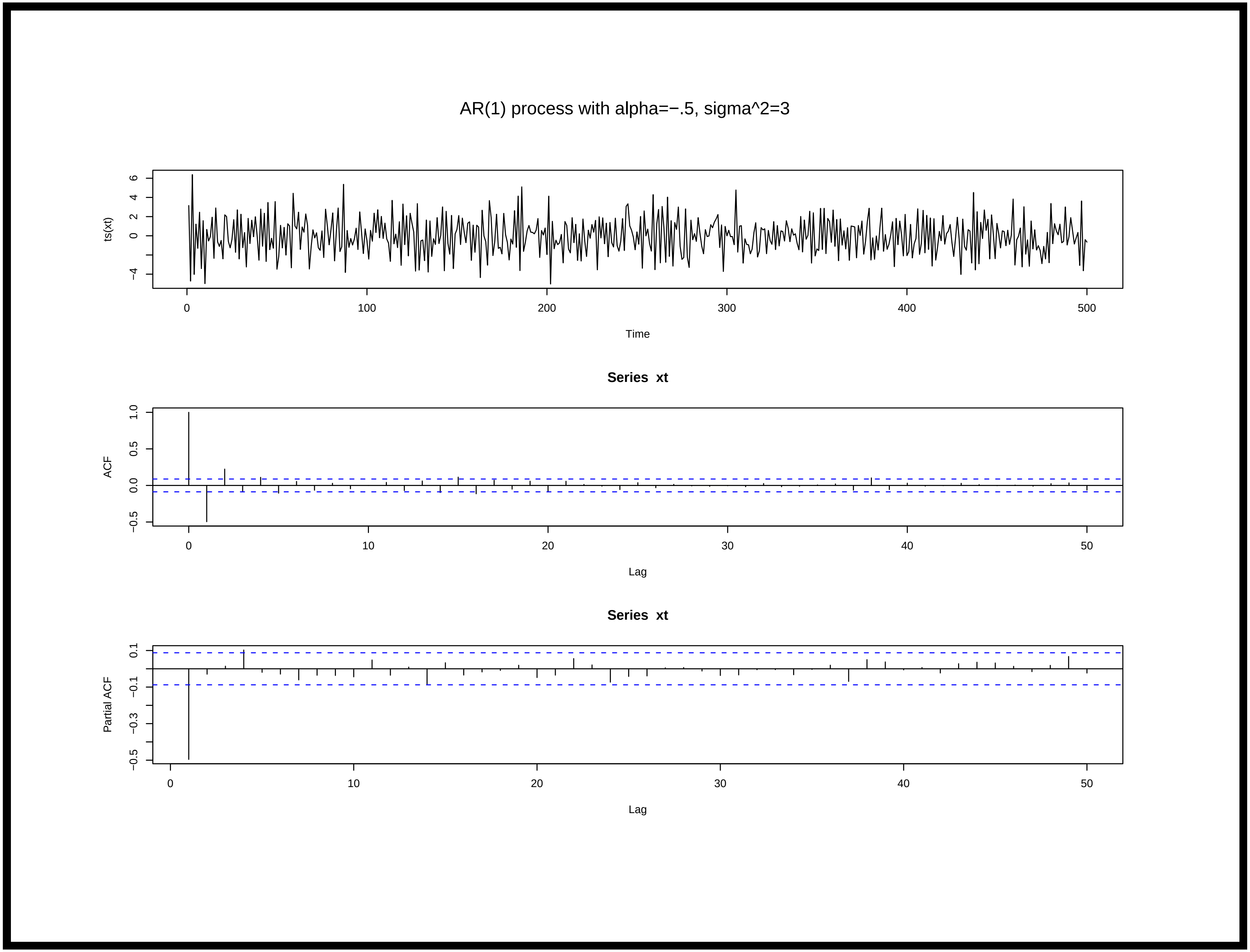
  so the partial autocorrelation of an AR(1) cuts down to zero after lag 1.

- Examples. Simulation of an AR(1) process with coefficient $\alpha = 0.9$ and $\alpha = -0.5$. 1000 observations in each case. In the second case $\sigma^2 = 9$.

```
alpha=.9
xt=arima.sim(1000,model=list(ar=alpha))
par(mfrow=c(3,1),oma=c(2,2,2,2))
ts.plot(ts(xt))
acf(xt,lag=20)
acf(xt,type="partial",lag=20)
mtext("AR(1) process with alpha =.9,
sigma^2=1",outer=T,cex=1.1)
```


AR(1) process with alpha =.9, sigma^2=1
ts(xt)
Time
Series xt
ACF
Lag
Series xt
Partial ACF
Lag

```
# Now with a variance different to one
epsilon=rnorm(500,mean=0,sd=sqrt(3))
alpha=-.5
xt=arima.sim(500,model=list(ar=alpha),innov=epsilon)
par(mfrow=c(3,1),oma=c(2,2,2,2))
ts.plot(ts(xt))
acf(xt,lag=50)
acf(xt,type="partial",lag=50)
mtext("AR(1) process with alpha=-.5,
sigma^2=3",outer=T,cex=1.1)
```


AR(1) process with alpha=−.5, sigma^2=3
ts(xt)
Time
Series xt
ACF
Lag
Series xt
Partial ACF
Lag

- **Example.** The AR(2) model

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \omega_t$$

- First we need to find the autocorrelation function of the process. That will allow us to find the PACF.
- By multiplying the AR(2) equation by X_{t-k} (both sides) and taking the expected value, we get

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}, k = 1, 2, \ldots$$

- Dividing by γ_0 gives,

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, k = 1, 2, \ldots$$

  which defines a *linear difference equation* for ρ_k.
- These difference equations can be difficult to solve.

- For this type of equations, it is recommended (Diggle) to explore a solution of the form $\rho_k = \lambda^k$ and try to determine the value of λ.
- If we substitute λ^k in the difference equation, we get

$$\lambda^k = \phi_1 \lambda^{k-1} + \phi_2 \lambda^{k-2}$$

- Which gives the equation

$$\lambda^2 - \phi_1 \lambda - \phi_2 = 0$$

- The solution to this equation is

$$\lambda = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

- This expression gives α_1 and α_2, the reciprocal roots of the characteristic polynomial.

- Since α_1 and α_2 are solutions of the equation, then a linear combination is also a solution.
- Then, the general solution to the difference equation takes the form,

$$\rho_k = a\alpha_1^k + b\alpha_2^k, k = 0, 1, 2, \ldots$$

where a and b are constants to be determined.

- Given this form, the $\rho_k' s$ will have an exponential behavior.
- To find values for a and b, note that for $k = 0, 1$, the difference equations are:

$$\rho_0 \quad = \quad 1 = a + b$$

$$\rho_1 = \phi_1/(1-\phi_2) = a\alpha_1 + b\alpha_2$$

- Assuming the AR(2) satisfies the stationarity conditions, we can find the values of ρ_k recursively.
- For the PACF, following Cramer's rule (HW exercise) it can be shown that the first two partial correlations are:

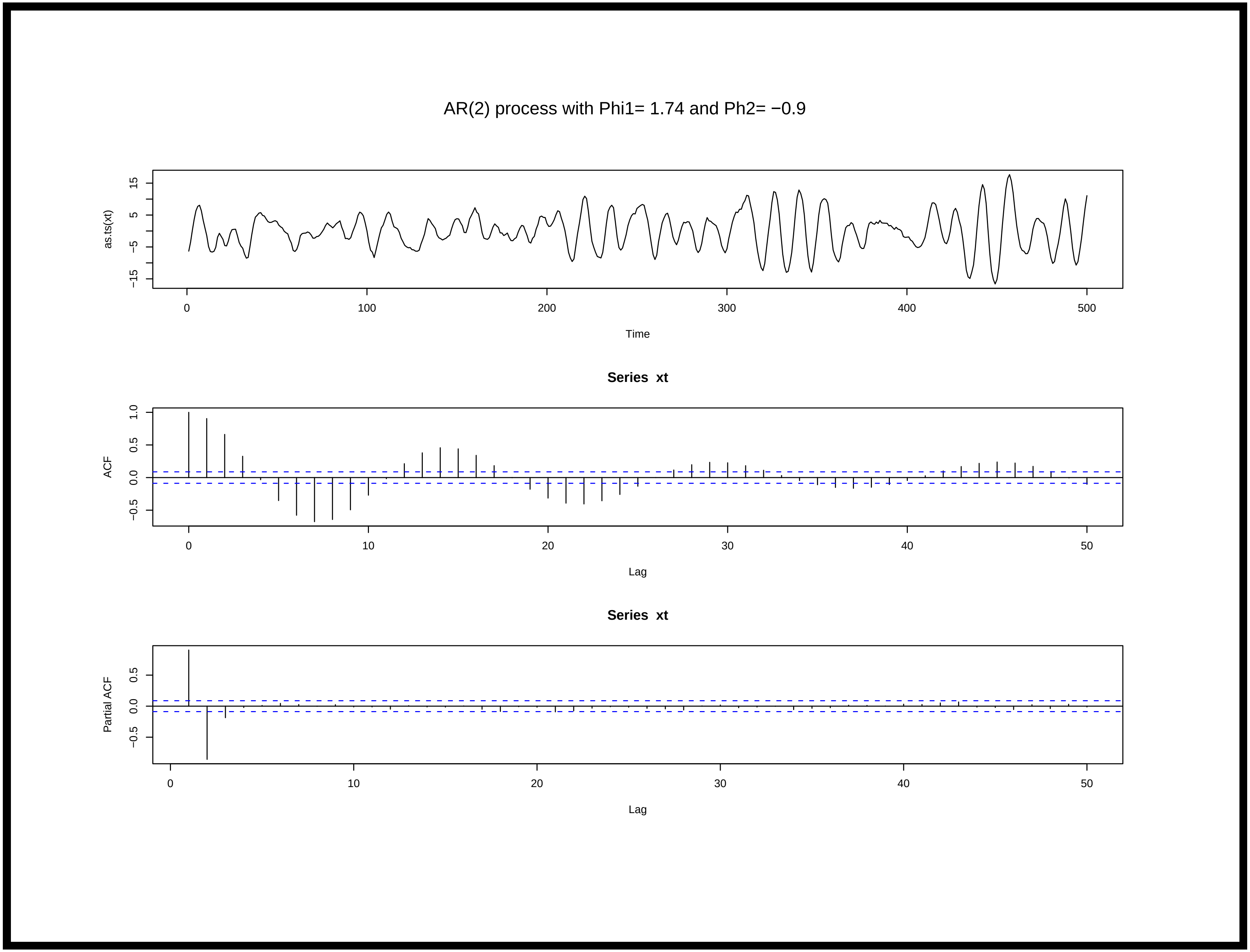
$$P_1 = \frac{\phi_1}{1-\phi_2}$$
$$P_2 = \phi_2$$

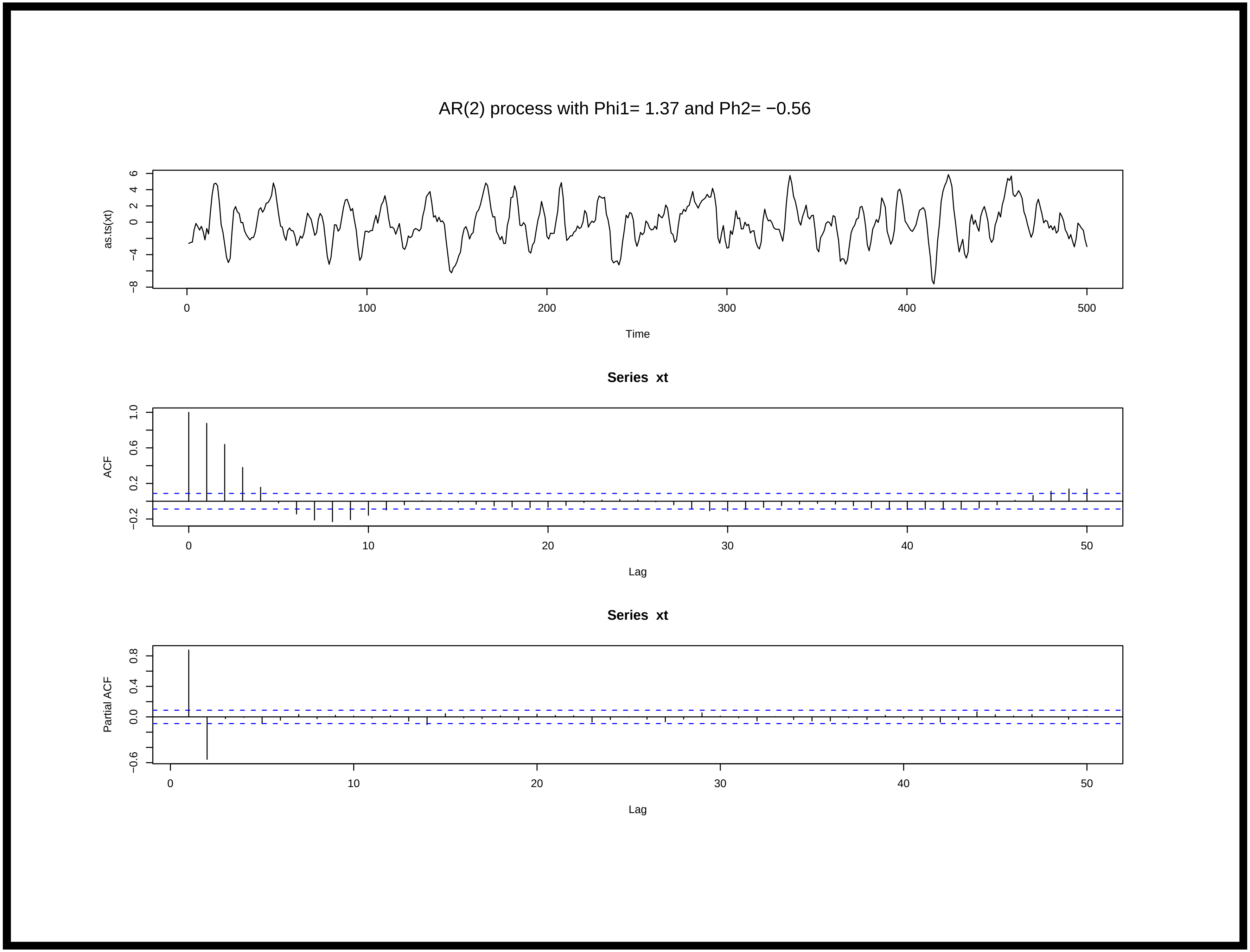
  and the remaining P_k's are zero.
- Consider an AR(2) with complex reciprocal roots $\alpha_{1,2} = rexp(\pm\omega i)$
- The characteristic polynomial is

$$\Phi(B) = (1 - re^{-i\omega}B)(1 - re^{i\omega}B) = (1 - 2rcos(\omega)B + r^2B^2)$$

- The AR coefficients are given by $\phi_1 = 2rcos(\omega)$, $\phi_2 = -r^2$
- Lets look at 500 simulated observations of an AR(2) process with $r = 0.95$ ($r = 0.75$) and $\omega = 2\pi/15$. Here is the R code to obtain this.

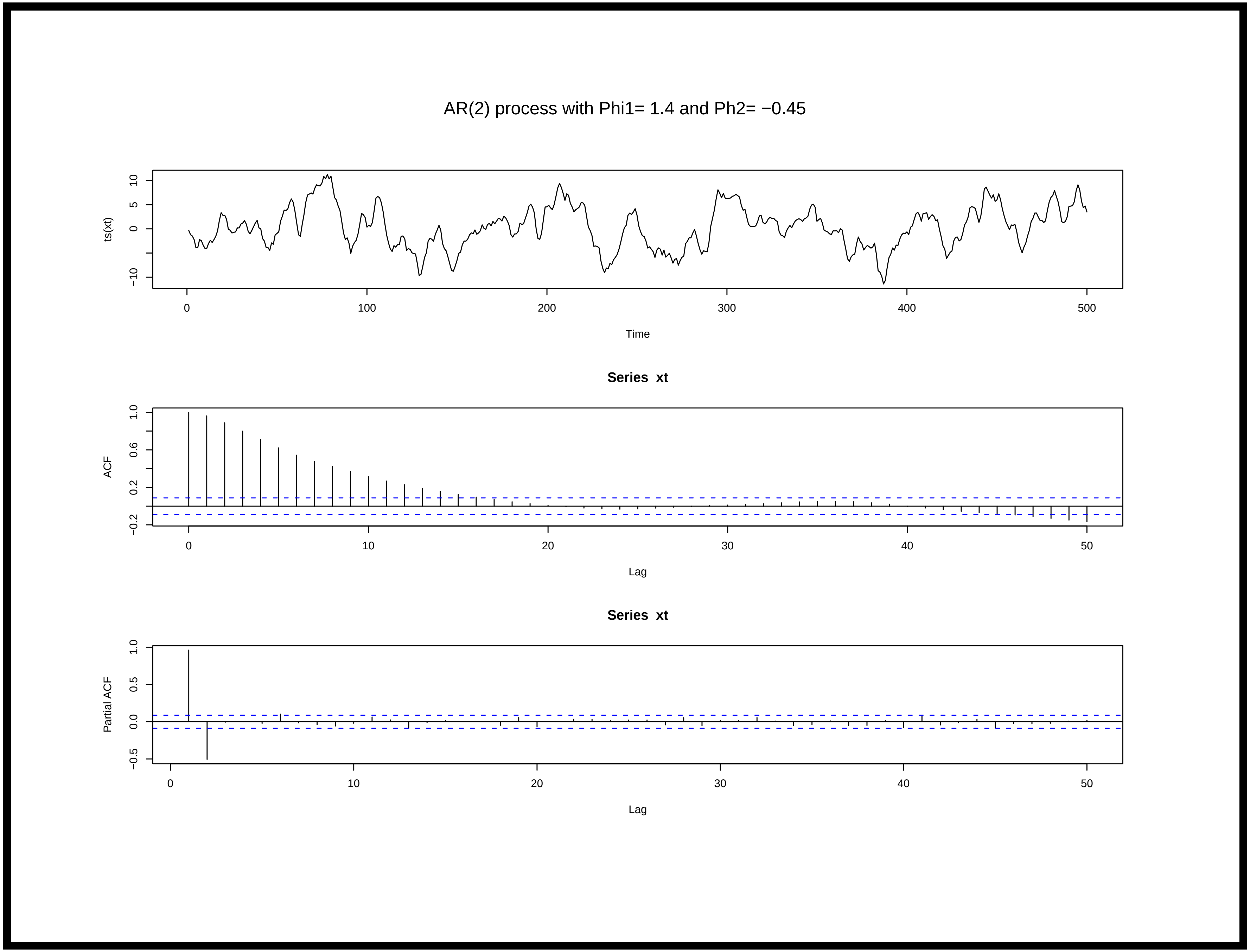
```
r=0.95
w=2*pi/15
phi1=2*r*cos(w)
phi2=-r^2
xt=arima.sim(500,model=list(ar=c(phi1,phi2)))
par(mfrow=c(3,1),oma=c(2,2,2,2))
ts.plot(as.ts(xt))
acf(xt,lag=50)
acf(xt,lag=50,type="partial")
```


AR(2) process with Phi1= 1.74 and Ph2= −0.9
as.ts(xt)
Time
Series xt
ACF
Lag
Series xt
Partial ACF
Lag

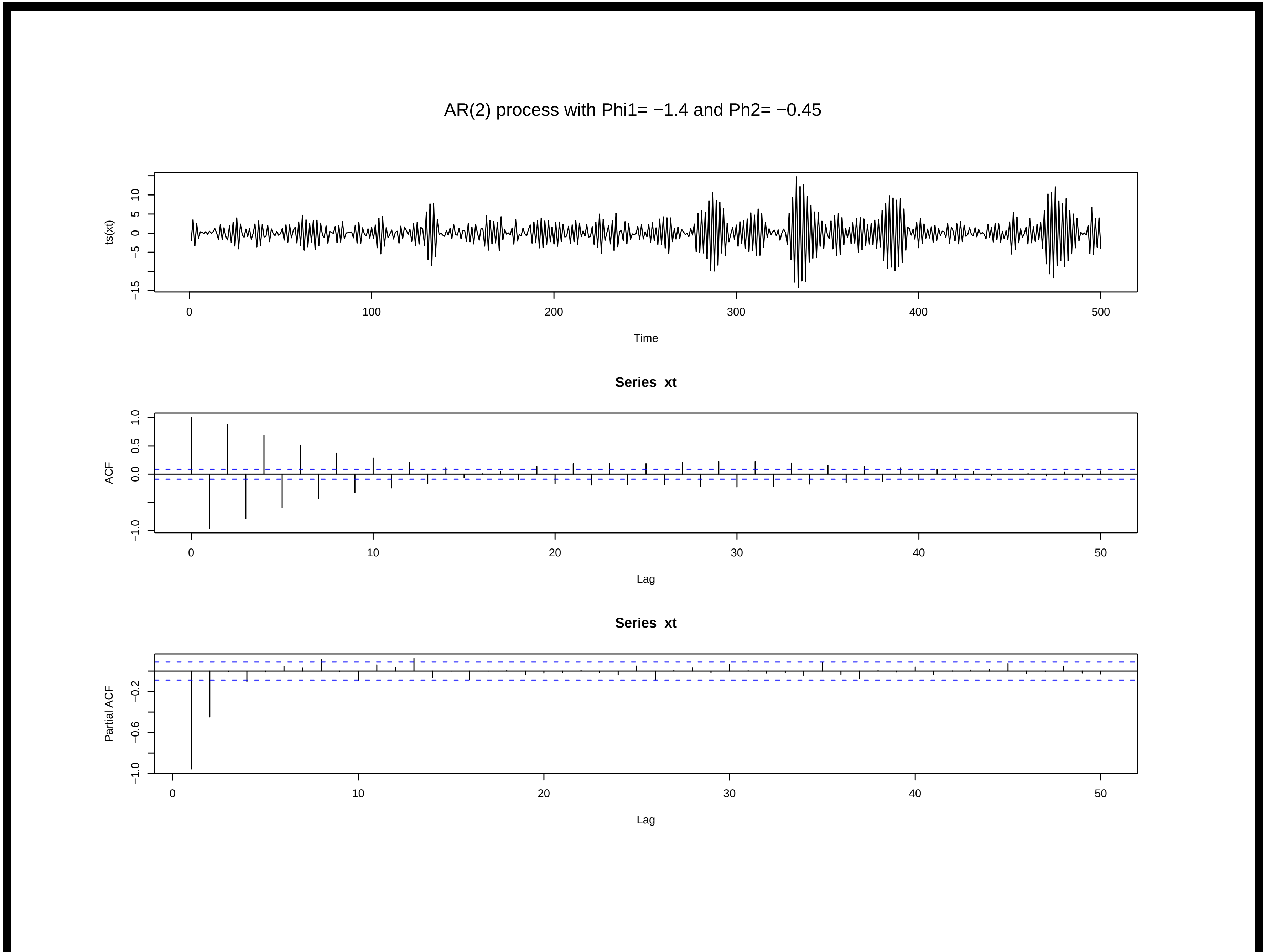

AR(2) process with Phi1= 1.37 and Ph2= −0.56
as.ts(xt)
Time
Series xt
ACF
Lag
Series xt
Partial ACF
Lag

- For the case of an AR(2) with two real roots r_1 and r_2, the characteristic polynomial is $\Phi(B) = (1 - r_1 B)(1 - r_2 B) = (1 - (r_1 + r_2)B + r_1 r_2 B^2)$
- The AR(2) coefficients are $\phi_1 = r_1 + r_2$ and $\phi_2 = -r_1 r_2$
- Lets look at a simulated process with $r_1 = 0.9$ and $r_2 = 0.5$.
- Then we will consider the case $r_1 = -0.9$ and $r_2 = -.0.5$

```
r1=0.9
r2=0.5
phi1=r1+r2
phi2=-r1*r2
xt=arima.sim(500,model=list(ar=c(phi1,phi2)))
```



AR(2) process with Phi1= 1.4 and Ph2= −0.45
ts(xt)
−10
0
5
10
0
100
200
300
400
500
Time
Series xt
ACF
−0.2
0.2
0.6
1.0
0
10
20
30
40
50
Lag
Series xt
Partial ACF
−0.5
0.0
0.5
1.0
0
10
20
30
40
50
Lag



AR(2) process with Phi1= −1.4 and Ph2= −0.45
ts(xt)
Time
Series xt
ACF
Lag
Series xt
Partial ACF
Lag