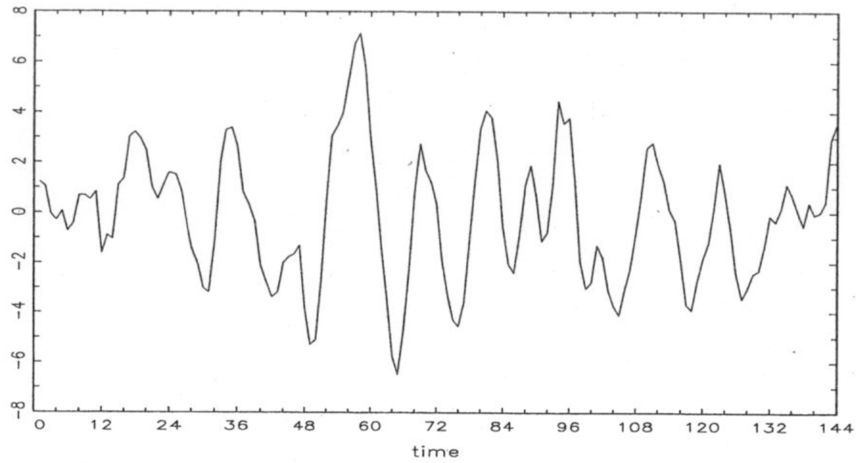

Figure 2.3: Simulated AR(2) model, $n = 144$ with $\phi_1 = 1.5$ and $\phi_2 = -0.75$.

The associated polynomial is

$$\alpha(z) = 1 - \alpha_1 z - \cdots - \alpha_p z^p.$$

Suppose $\alpha(z)$ has r distinct roots, z_1 with multiplicity m_1, z_2 with multiplicity m_2, ..., and z_r with multiplicity m_r, such that $m_1 + m_2 + \cdots m_r = p$. The general solution to the difference equation (2.35) is

$$u_n = z_1^{-n} P_1(n) + z_2^{-n} P_2(n) + \cdots + z_r^{-n} P_r(n), \qquad (2.36)$$

where $P_j(n)$, for $j = 1, 2, ..., r$, is a polynomial in n, of degree $m_j - 1$. Given p initial conditions $u_0, ..., u_{p-1}$, we can solve for the $P_j(n)$ explicitly.

Example 2.10 Determining the ψ-weights for a Causal ARMA(p,q)

For a causal ARMA(p,q) model, $\phi(B)x_t = \theta(B)w_t$, where the zeros of $\phi(z)$ are outside the unit circle, recall that we may write

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j},$$

where the ψ-weights are determined using Property P2.1.

For the pure MA(q) model, $\psi_0 = 1$, $\psi_j = \theta_j$, for $j = 1, ..., q$, and $\psi_j = 0$, otherwise. For the general case of ARMA(p,q) models, the task of solving

for the ψ-weights is much more complicated, as was demonstrated in Example 2.6. The use of the theory of homogeneous difference equations can help here. To solve for the ψ-weights in general, we must match the coefficients in $\psi(z)\phi(z) = \theta(z)$:

$$(\psi_0 + \psi_1 z + \psi_2 z^2 + \cdots)(1 - \phi_1 z - \phi_2 z^2 + \cdots) = (1 + \theta_1 z + \theta_2 z^2 + \cdots).$$

The first few values are

$$\begin{aligned} \psi_0 &= 1 \\ \psi_1 - \phi_1\psi_0 &= \theta_1 \\ \psi_2 - \phi_1\psi_1 - \phi_2\psi_0 &= \theta_2 \\ \psi_3 - \phi_1\psi_2 - \phi_2\psi_1 - \phi_3\psi_0 &= \theta_3 \\ &\vdots \end{aligned}$$

where we would take $\phi_j = 0$ for $j > p$, and $\theta_j = 0$ for $j > q$. The ψ-weights satisfy the homogeneous difference equation given by

$$\psi_j - \sum_{k=1}^{p} \phi_k \psi_{j-k} = 0, \quad j \geq \max(p, q+1), \qquad (2.37)$$

with initial conditions

$$\psi_j - \sum_{k=1}^{j} \phi_k \psi_{j-k} = \theta_j, \quad 0 \leq j \leq \max(p, q+1). \qquad (2.38)$$

The general solution depends on the roots of the AR polynomial $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$, as seen from (2.37). The specific solution will, of course, depend on the initial conditions.

Consider the ARMA process given in (2.27), $x_t = 0.9x_{t-1} + 0.5w_{t-1} + w_t$. Because $\max(p, q+1) = 2$, using (2.38), we have $\psi_0 = 1$ and $\psi_1 = 0.9 + 0.5 = 1.4$. By (2.37), for $j = 2, 3, ...$, the ψ-weights satisfy $\psi_j - 0.9\psi_{j-1} = 0$. The general solution is $\psi_j = c\,0.9^j$. To find the specific solution, use the initial condition $\psi_1 = 1.4$, so $1.4 = 0.9c$ or $c = 1.4/0.9$. Finally, $\psi_j = 1.4(0.9)^{j-1}$, for $j \geq 1$, as we saw in Example 2.6.

2.4 Autocorrelation and Partial Autocorrelation Functions

We begin by exhibiting the ACF of an MA(q) process, $x_t = \theta(B)w_t$, where $\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$. Because x_t is a finite linear combination of white noise terms, the process is stationary with mean

$$E(x_t) = \sum_{j=0}^{q} \theta_j E(w_{t-j}) = 0,$$

where we have written $\theta_0 = 1$, and with autocovariance function

$$\begin{aligned}\gamma(h) = \text{cov}(x_{t+h}, x_t) &= E\left[\left(\sum_{j=0}^{q}\theta_j w_{t+h-j}\right)\left(\sum_{k=0}^{q}\theta_j w_{t-k}\right)\right] \\ &= \begin{cases}\sigma_w^2 \sum_{j=0}^{q-h}\theta_j\theta_{j+h}, & 0 \le h \le q \\ 0, & h > q.\end{cases}\end{aligned} \tag{2.39}$$

Recall that $\gamma(h) = \gamma(-h)$, so we will only display the values for $h \ge 0$. The cutting off of $\gamma(h)$ after q lags is the signature of the MA(q) model. Dividing (2.39) by $\gamma(0)$ yields the ***ACF of an MA(q)***:

$$\rho(h) = \begin{cases}\frac{\sum_{j=0}^{q-h}\theta_j\theta_{j+h}}{1+\theta_1^2+\cdots+\theta_q^2}, & 1 \le h \le q \\ 0, & h > q.\end{cases} \tag{2.40}$$

For a causal ARMA(p,q) model, $\phi(B)x_t = \theta(B)w_t$, where the zeros of $\phi(z)$ are outside the unit circle, write

$$x_t = \sum_{j=0}^{\infty}\psi_j w_{t-j}.$$

It follows immediately that $E(x_t) = 0$. Also, the autocovariance function of x_t can be written as:

$$\gamma(h) = \text{cov}(x_{t+h}, x_t) = \sigma_w^2 \sum_{j=0}^{\infty}\psi_j\psi_{j+h}, \quad h \ge 0. \tag{2.41}$$

We could then use (2.37) and (2.38) to solve for the ψ-weights. In turn, we could solve for $\gamma(h)$, and the ACF $\rho(h) = \gamma(h)/\gamma(0)$. As in Example 2.8, it is also possible to obtain a homogeneous difference equation directly in terms of $\gamma(h)$. First, we write

$$\begin{aligned}\gamma(h) &= \text{cov}(x_{t+h}, x_t) = E\left[\left(\sum_{j=1}^{p}\phi_j x_{t+h-j} + \sum_{j=0}^{q}\theta_j w_{t+h-j}\right)x_t\right] \\ &= \sum_{j=1}^{p}\phi_j\gamma(h-j) + \sigma_w^2\sum_{j=h}^{q}\theta_j\psi_{j-h}, \quad h \ge 0,\end{aligned} \tag{2.42}$$

where we have used the fact that for $h \ge 0$,

$$E(x_{t+h}w_t) = E\left[\left(\sum_{j=0}^{\infty}\psi_j w_{t+h-j}\right)w_t\right] = \psi_h\sigma_w^2.$$

From (2.42), we can write a general homogeneous equation:

$$\gamma(h) - \phi_1\gamma(h-1) - \cdots - \phi_p\gamma(h-p) = 0, \quad h \ge \max(p, q+1), \tag{2.43}$$

with initial conditions

$$\gamma(h) - \sum_{j=1}^{p}\phi_j\gamma(h-j) = \sigma_w^2\sum_{j=h}^{q}\theta_j\psi_{j-h}, \quad 0 \le h < \max(p, q+1). \tag{2.44}$$

Dividing (2.43) and (2.44) through by $\gamma(0)$ will allow us to solve for the ACF, $\rho(h) = \gamma(h)/\gamma(0)$.

Example 2.11 The ACF of an ARMA(1,1)

Consider the causal ARMA(1,1) process $x_t = \phi x_{t-1} + \theta w_{t-1} + w_t$, where $|\phi| < 1$. Based on (2.43), the autocovariance function satisfies

$$\gamma(h) - \phi\gamma(h-1) = 0, \quad h = 2, 3, ...,$$

so the general solution is $\gamma(h) = c\phi^h$, for $h = 1, 2, ...$. To solve for c, we use (2.44):

$$\begin{aligned}\gamma(0) &= \phi\gamma(1) + \sigma_w^2[1 + \theta\phi + \theta^2] \\ \gamma(1) &= \phi\gamma(0) + \sigma_w^2\theta.\end{aligned}$$

Solving for $\gamma(0)$ and $\gamma(1)$, we obtain:

$$\begin{aligned}\gamma(0) &= \sigma_w^2\frac{1 + 2\theta\phi + \theta^2}{1 - \phi^2} \\ \gamma(1) &= \sigma_w^2\frac{(1 + \theta\phi)(\phi + \theta)}{1 - \phi^2}.\end{aligned}$$

Because $\gamma(1) = c\phi$, we have $c = \gamma(1)/\phi$, so the general solution is

$$\gamma(h) = \sigma_w^2\frac{(1 + \theta\phi)(\phi + \theta)}{1 - \phi^2}\phi^{h-1}.$$

Finally, dividing through by $\gamma(0)$ yields the ACF

$$\rho(h) = \frac{(1 + \theta\phi)(\phi + \theta)}{1 + 2\theta\phi + \theta^2}\phi^{h-1}, \quad h \ge 1. \tag{2.45}$$

Example 2.12 The ACF of an AR(p)

For a causal AR(p), it follows immediately from (2.43) that

$$\rho(h) - \phi_1\rho(h-1) - \cdots - \phi_p\rho(h-p) = 0, \quad h \ge p. \tag{2.46}$$

Let $z_1, ..., z_r$ denote the roots of $\phi(z)$, each with multiplicity $m_1, ..., m_r$, respectively, where $m_1 + \cdots + m_r = p$. Then, from (2.37), the general solution is

$$\rho(h) = z_1^{-h}P_1(h) + z_2^{-h}P_2(h) + \cdots + z_r^{-h}P_r(h), \quad h \geq p, \tag{2.47}$$

where $P_j(h)$ is a polynomial in h of degree $m_j - 1$.

Recall that for a causal model, all of the roots are outside the unit circle, $|z_i| > 1$, for $i = 1, ..., r$. If all the roots are real, then $\rho(h)$ dampens exponentially fast to zero as $h \to \infty$. If some of the roots are complex, then they will be in conjugate pairs and $\rho(h)$ will dampen, in a sinusoidal fashion, exponentially fast to zero as $h \to \infty$. In the case of complex roots, the time series will appear to be cyclic in nature. This, of course, is also true for ARMA models in which the AR part has complex roots.

The Partial Autocorrelation Function (PACF)

We have seen in (2.40) that, for MA(q) models, the ACF will be zero for lags greater than q. Moreover, because $\theta_q \neq 0$, the ACF will not be zero at lag q. Thus, the ACF provides a considerable amount of information about the order of the dependence when the process is a moving average process. If the process, however, is ARMA or AR, the ACF alone tells us little about the orders of dependence. Hence, it is worthwhile pursuing a function that will behave like the ACF for AR models, namely, the ***partial autocorrelation function (PACF)***.

To motivate the idea, consider a causal AR(1) model, $x_t = \phi x_{t-1} + w_t$. Then,

$$\begin{aligned}\gamma(2) = \text{cov}(x_t, x_{t-2}) &= \text{cov}(\phi x_{t-1} + w_t, x_{t-2}) \\ &= \text{cov}(\phi^2 x_{t-2} + \phi w_{t-1} + w_t, x_{t-2}) = \phi^2\gamma(0).\end{aligned}$$

This result follows from causality because x_{t-2} involves $\{w_{t-2}, w_{t-3}, ...\}$, which are all uncorrelated with w_t and w_{t-1}. The correlation between x_t and x_{t-2} is not zero, as it would be for an MA(1), because x_t is dependent on x_{t-2} through x_{t-1}. Suppose we break this chain of dependence by removing (*partial out*) x_{t-1}. That is, we consider the correlation between $x_t - \phi x_{t-1}$ and $x_{t-2} - \phi x_{t-1}$, because it is the correlation between x_t and x_{t-2} with the linear dependence of each on x_{t-1} removed. In this way, we have broken the dependence chain between x_t and x_{t-2}. In fact,

$$\text{cov}(x_t - \phi x_{t-1}, x_{t-2} - \phi x_{t-1}) = \text{cov}(w_t, x_{t-2} - \phi x_{t-1}) = 0.$$

To formally define the PACF for mean-zero stationary time series, let x_h^{h-1} denote the best linear predictor of x_h based on $\{x_1, x_2, ..., x_{h-1}\}$. We will discuss prediction in detail in the next section, but for now, we note x_h^{h-1} has the form:

$$x_h^{h-1} = \beta_1 x_{h-1} + \beta_2 x_{h-2} + \cdots + \beta_{h-1}x_1, \tag{2.48}$$

where the β's are chosen to minimize the mean square linear prediction error, $E(x_h - x_h^{h-1})^2$. In addition, let x_0^{h-1} denote the minimum mean square linear predictor of x_0 based on $\{x_1, x_2, ..., x_{h-1}\}$. As will be seen in the next section, x_0^{h-1} can be written as

$$x_0^{h-1} = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{h-1}x_{h-1}. \tag{2.49}$$

Equation (2.48) can be thought of as the linear regression of x_h on the past, $x_{h-1}, ..., x_1$, and (2.49) can be thought of as the linear regression of x_0 on the future, $x_1, ..., x_{h-1}$. The coefficients, $\beta_1, ..., \beta_{h-1}$ are the same in (2.48) and (2.49), which means that, for stationary processes, linear prediction backward in time is equivalent to linear prediction forward in time. We will discuss this result further in the next section.

Formally, for a stationary time series, x_t, we define the ***partial autocorrelation function (PACF)***, ϕ_{hh}, $h = 1, 2, ...$, by

$$\phi_{11} = \text{corr}(x_1, x_0) = \rho(1) \tag{2.50}$$

and

$$\phi_{hh} = \text{corr}(x_h - x_h^{h-1}, x_0 - x_0^{h-1}), \quad h \geq 2. \tag{2.51}$$

Both $(x_h - x_h^{h-1})$ and $(x_0 - x_0^{h-1})$ are uncorrelated with $\{x_1, x_2, ..., x_{h-1}\}$. By stationarity, the PACF, ϕ_{hh}, is the correlation between x_t and x_{t-h} with the linear effect of $\{x_{t-1}, ..., x_{t-(h-1)}\}$, on each, removed. If the process x_t is Gaussian, then $\phi_{hh} = \text{corr}(x_t, x_{t-h} \mid x_{t-1}, ..., x_{t-(h-1)})$. That is, ϕ_{hh} is the correlation coefficient between x_t and x_{t-h} in the bivariate distribution of (x_t, x_{t-h}) conditional on $\{x_{t-1}, ..., x_{t-(h-1)}\}$.

Example 2.13 The PACF of a Causal AR(1)

Consider the PACF of the AR(1) process given by $x_t = \phi x_{t-1} + w_t$, with $|\phi| < 1$. By definition, $\phi_{11} = \rho(1) = \phi$. To calculate ϕ_{22}, consider the prediction of x_2 based on a linear function of x_1, say, $x_2^1 = \alpha x_1$. We choose α to minimize

$$E(x_2 - \alpha x_1)^2 = \gamma(0) - 2\alpha\gamma(1) + \alpha^2\gamma(0).$$

Taking derivatives and setting the result equal to zero, we have $\alpha = \gamma(1)/\gamma(0) = \rho(1) = \phi$. Thus, $x_2^1 = \phi x_1$. Next, consider the prediction of x_0 based on a linear function of x_1: $x_0^1 = \alpha x_1$. We choose α to minimize

$$E(x_0 - \alpha x_1)^2 = \gamma(0) - 2\alpha\gamma(1) + \alpha^2\gamma(0).$$

Analogously, we have $x_0^1 = \phi x_1$, which agrees with the claim that prediction forward in time is equivalent to prediction backward in time. Hence, $\phi_{22} = \text{corr}(x_2 - \phi x_1, x_0 - \phi x_1)$. But, note

$$\text{cov}(x_2 - \phi x_1, x_0 - \phi x_1) = \gamma(2) - 2\phi\gamma(1) + \phi^2\gamma(0) = 0$$

since $\gamma(h) = \gamma(0)\phi^h$. Thus, $\phi_{22} = 0$. In the next example, we will see that in this case $\phi_{hh} = 0$, for all $h > 1$.

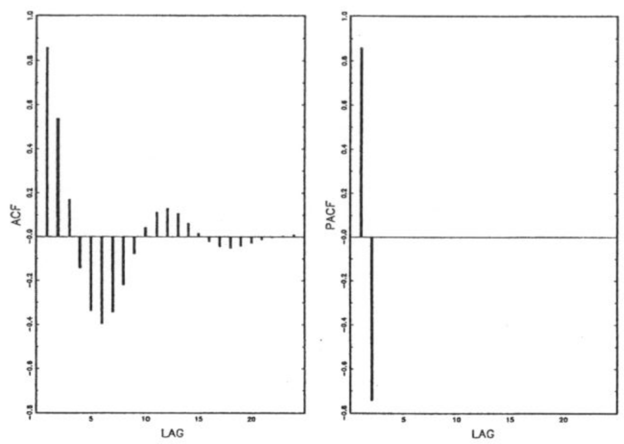

Figure 2.4: The ACF and PACF, to lag 24, of an AR(2) model, with $\phi_1 = 1.5$ and $\phi_2 = -0.75$.

Example 2.14 The PACF of a Causal AR(p)

Let $x_t = \sum_{j=1}^{p} \phi_j x_{t-j} + w_t$, where the roots of $\phi(z)$ are outside the unit circle. When $h > p$, then

$$x_h^{h-1} = \sum_{j=1}^{p} \phi_j x_{h-j}.$$

We have not proven this obvious result yet, but we will prove it in the next section. Thus, when $h > p$,

$$\begin{aligned} \phi_{hh} &= \text{corr}(x_h - x_h^{h-1}, x_0 - x_0^{h-1}) \\ &= \text{corr}(w_h, x_0 - x_0^{h-1}) = 0, \end{aligned}$$

since, by causality, $x_0 - x_0^{h-1}$ depends only on $\{w_{h-1}, w_{h-2}, ...\}$; recall equation (2.49). When $h \leq p$, ϕ_{pp} is not zero, and $\phi_{11}, ..., \phi_{p-1,p-1}$ are not necessarily zero. Figure 2.4 shows the ACF and the PACF of the AR(2) model presented in Example 2.9.

Example 2.15 The PACF of an invertible MA(q)

For an invertible MA(q), we can write $x_t = \sum_{j=1}^{\infty} \pi_j x_{t-j} + w_t$. Moreover, no finite representation exists. From this result, it should be apparent that the PACF will never cut off, as in the case of an AR(p).

Table 2.1 Behavior of the ACF and PACF for Causal and Invertible ARMA Models

	AR(p)	MA(q)	ARMA(p, q)
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

For an MA(1), $x_t = w_t + \theta w_{t-1}$, with $|\theta| < 1$, calculations similar to Example 2.13 will yield $\phi_{22} = (-\theta)^2/(1 + \theta^2 + \theta^4)$. For the MA(1) in general, we can show that

$$\phi_{hh} = \frac{(-\theta)^h(1-\theta^2)}{1-\theta^{2(h+1)}}, \quad h \geq 1.$$

In the next section, we will discuss methods of calculating the PACF. The PACF for MA models behaves much like the ACF for AR models. Also, the PACF for AR models behaves much like the ACF for MA models. Because an invertible ARMA model has an infinite AR representation, the PACF will not cut off. We may summarize these results in Table 2.1.

Example 2.16 Preliminary Analysis of the Recruitment Series

We consider the problem of modeling the Recruitment series (number of new fish) shown in Figure 1.4. There are 453 months of observed recruitment ranging over the years 1950-1987. The ACF and the PACF given in Figure 2.5 are consistent with the behavior of an AR(2). The ACF has cycles corresponding roughly to a 12-month period, and the PACF has large values for $h = 1, 2$ and then is essentially zero for higher order lags. Based on Table 2.1, these results suggest that a second-order ($p = 2$) autoregressive model might provide a good fit. Although we will discuss estimation in detail in Section 2.6, we ran a regression (see Section 1.8) using the data triplets $\{(x_3, x_2, x_1), (x_4, x_3, x_2), ..., (x_{453}, x_{452}, x_{451})\}$ to fit a model of the form

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + w_t$$

for $t = 3, 4, \ldots, 453$. The values of the estimates were $\widehat{\phi}_0 = 6.74(1.11)$, $\widehat{\phi}_1 = 1.35(.04), \widehat{\phi}_2 = -.46(.04)$, and $\widehat{\sigma}_w^2 = 90.31$, where the estimated standard errors are in parentheses.

2.5 Forecasting

In ***forecasting***, the goal is to predict future values of a time series, x_{n+m}, $m = 1, 2, ...$, based on the data collected to the present, $\boldsymbol{x} = \{x_n, x_{n-1}, ..., x_1\}$. Throughout this section, we will assume x_t is stationary and the model parameters are known. The problem of forecasting when the model parameters are unknown will be discussed in the next section; also, see Problem 2.25. The ***minimum mean square error predictor*** of x_{n+m} is $x^n_{n+m} = E(x_{n+m}|x_n, x_{n-1}, ..., x_1)$ because the conditional expectation minimizes the mean square error

$$E\{x_{n+m} - g(\boldsymbol{x})\}^2, \tag{2.52}$$

where $g(\boldsymbol{x})$ is a (measurable) function of the observations $\boldsymbol{x}$. This result follows by iterating the expectation, $E\{x_{n+m} - g(\boldsymbol{x})\}^2 = E[E\{[x_{n+m} - g(\boldsymbol{x})]^2 \mid \boldsymbol{x}\}]$, and then observing that the inner expectation is minimized when $x^n_{n+m} = E(x_{n+m} \mid \boldsymbol{x})$; see Problem 2.13.

First, we will restrict attention to predictors that are linear functions of the data, that is, predictors of the form

$$x^n_{n+m} = \alpha_0 + \sum_{k=1}^{n} \alpha_k x_k, \tag{2.53}$$

where $\alpha_0, \alpha_1, ..., \alpha_n$ are real numbers. Linear predictors of the form (2.53) that minimize the mean square prediction error (2.52) are called ***best linear predictors***.

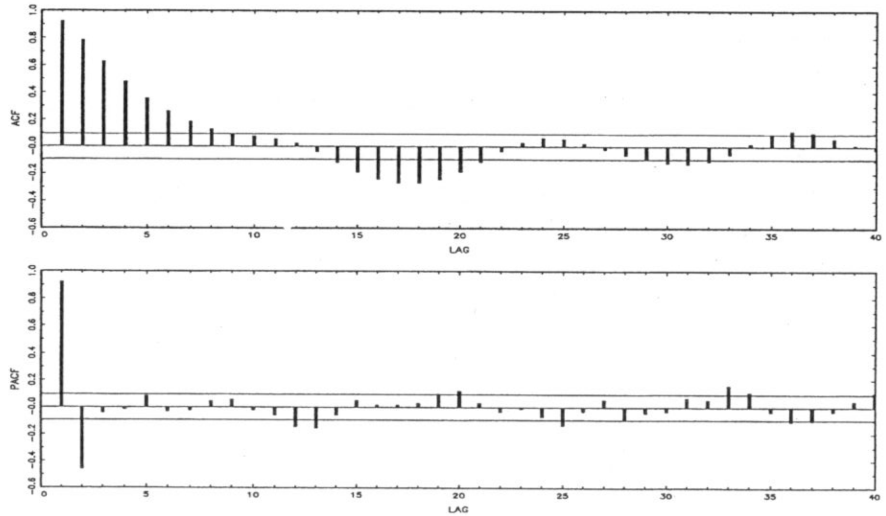

Figure 2.5: ACF and PACF of the Recruitment series.

As we shall see, linear prediction depends only on the second-order moments of the process, which are easy to estimate from the data. Much of the material in this section is enhanced by the theoretical material presented in Section T2.15. For example, Theorem 2.3 states that if the process is Gaussian, minimum mean square error predictors and best linear predictors are the same. The following property, which is based on the projection theorem, Theorem 2.1 of Section T2.15, is a key result.

Property P2.3: Best Linear Prediction for Stationary Processes
Given data $x_1, ..., x_n$, the best linear predictor, $x^n_{n+m} = \alpha_0 + \sum_{k=1}^{n} \alpha_k x_k$, of x_{n+m}, for $m \geq 1$, is found by solving

$$E\left[x_{n+m} - x^n_{n+m}\right] = 0;$$
$$E\left[\left(x_{n+m} - x^n_{n+m}\right) x_k\right] = 0, \quad k = 1, ..., n. \tag{2.54}$$

The equations specified in (2.54) are called the ***prediction equations***, and they are used to solve for the coefficients $\{\alpha_0, \alpha_1, ..., \alpha_n\}$. If $E(x_t) = \mu$, the first equation of (2.54) is $E(x^n_{n+m}) = E(x_{n+m}) = \mu$, which implies $\alpha_0 = \mu(1 - \sum_{k=1}^{n} \alpha_k)$. Hence, the form of the BLP is $x^n_{n+m} = \mu + \sum_{k=1}^{n} \alpha_k(x_k - \mu)$. Thus, until we discuss estimation, there is no loss of generality in considering the case that $\mu = 0$, in which case, $\alpha_0 = 0$.

Consider, first, ***one-step-ahead prediction***. That is, given $\{x_1, ..., x_n\}$, we wish to forecast the value of the time series at the next time point, x_{n+1}. The BLP of x_{n+1} is

$$x^n_{n+1} = \phi_{n1}x_n + \phi_{n2}x_{n-1} + \cdots + \phi_{nn}x_1, \tag{2.55}$$

where, for purposes that will become clear shortly, we have written α_k in (2.54), as $\phi_{n,n+1-k}$ in (2.55), for $k = 1, ..., n$. Using Property P2.3, the coefficients $\{\phi_{n1}, \phi_{n2}, ..., \phi_{nn}\}$ satisfy

$$E\left[\left(x_{n+1} - \sum_{j=1}^{n} \phi_{nj} x_{n+1-j}\right) x_{n+1-k}\right] = 0, \quad k = 1, ..., n,$$

or

$$\sum_{j=1}^{n} \phi_{nj}\gamma(k-j) = \gamma(k), \quad k = 1, ..., n. \tag{2.56}$$

The prediction equations (2.56) can be written in matrix notation as

$$\Gamma_n \boldsymbol{\phi}_n = \boldsymbol{\gamma}_n, \tag{2.57}$$

where $\Gamma_n = \{\gamma(k-j)\}_{j,k=1}^n$ is an $n \times n$ matrix, $\boldsymbol{\phi}_n = (\phi_{n1}, ..., \phi_{nn})'$ is an $n \times 1$ vector, and $\boldsymbol{\gamma}_n = (\gamma(1), ..., \gamma(n))'$ is an $n \times 1$ vector.