Another simple time series model is the first order autoregression, denoted by AR(1). The series \( \{x_t\} \) is AR(1) if it satisfies the iterative equation (called a difference equation)

\[
x_t = \alpha x_{t-1} + \epsilon_t ,
\]

where \( \{\epsilon_t\} \) is a zero-mean white noise. We use the term autoregression since (1) is actually a linear regression model for \( x_t \) in terms of the explanatory variable \( x_{t-1} \). That is, \( x_t \) is being modeled as a regression on its own past. We will see that \( \epsilon_t \) is uncorrelated with past values of the AR series \( x_t \). Thus, \( \epsilon_t \) represents the new contribution to \( x_t \), and we can think of \( \{\epsilon_t\} \) as a series of random shocks, or innovations.

The definition (1) is implicit, since \( x_t \) is defined in terms of its own past. It is useful to try to write the AR(1) process explicitly in terms of present and past innovations. Substituting for \( x_{t-1} \) in (1) gives

\[
x_t = \alpha [\alpha x_{t-2} + \epsilon_{t-1}] + \epsilon_t = \epsilon_t + \alpha \epsilon_{t-1} + \alpha^2 x_{t-2} .
\]

Substituting for \( x_{t-2} \) in (2) gives

\[
x_t = \epsilon_t + \alpha \epsilon_{t-1} + \alpha^2 [\alpha x_{t-3} + \epsilon_{t-2}] = \epsilon_t + \alpha \epsilon_{t-1} + \alpha^2 \epsilon_{t-2} + \alpha^3 x_{t-3} .
\]

Continuing this process, we see that for any \( N \),

\[
x_t = \sum_{j=0}^{N-1} \alpha^j \epsilon_{t-j} + \alpha^N x_{t-N} .
\]

The value of the parameter \( \alpha \) strongly affects the behavior of the AR(1) process. Suppose that \( N \) in equation (3) is large, so that the underlying series \( \{x_t\} \) started a long time before our observations \( x_1, \ldots, x_n \) were formed. Then if we have \(-1 < \alpha < 1\), the last term in (3) will be negligible and the weight given to shocks which occurred a long time ago will also be extremely small. The resulting series \( \{x_t\} \) will be stationary. If, on the other hand, we have \(|\alpha| > 1\), the last term in (3) will be large in magnitude and the weights given to distant shocks will be much greater than those given to more recent ones. The model is then said to be explosive, since the series mean and variance both explode as \( t \) grows. The explosive model is not considered useful for economic time series. Finally, if \( \alpha = 1 \), we
obtain a useful model called the \textit{random walk} which is neither explosive nor stationary. It will be discussed later.

Here, we assume that \(-1<\alpha<1\), and that \(N\) is very large, so that (3) is effectively

\[
x_t = \sum_{j=0}^{\infty} \alpha^j \varepsilon_{t-j}.
\]

Equation (4) is called an MA\((\infty)\) representation for the AR\((1)\) process, since it expresses \(x_t\) as a moving average of an infinite number of present and past shocks. It follows from (4) that \(E[x_t \varepsilon_{t+1}] = 0\). Thus the future shock is uncorrelated with the present data. More generally, we have

\[
E[x_t \varepsilon_{t+k}] = 0
\]

for all \(t\) and all positive \(k\). Thus, for example, the present shock is uncorrelated with all past series values. Also, all future shocks are uncorrelated with all present and past series values.

Another consequence of (4) is that

\[
var(x_t) = E[x_t^2] = (1 + \alpha^2 + \alpha^4 + \alpha^6 + \cdots) \text{var} \varepsilon = \frac{\text{var} \varepsilon}{1 - \alpha^2}.
\]

The covariance between \(x_t\) and \(x_{t-1}\) is

\[
cov(x_t, x_{t-1}) = E[x_t x_{t-1}] = E[(\alpha x_{t-1} + \varepsilon_t)(x_{t-1})]
\]

\[
= \alpha E[x_{t-1}^2] + E[\varepsilon_t x_{t-1}] = \alpha \text{var}(x_t).
\]

Thus, the correlation between \(x_t\) and \(x_{t-1}\) is

\[
corr(x_t, x_{t-1}) = \frac{cov(x_t, x_{t-1})}{\text{var} x_t} = \alpha.
\]

so the degree of smoothness of \(\{x_t\}\) is determined by \(\alpha\) : \(\{x_t\}\) is most smooth for \(\alpha\) near +1, very unsmooth for \(\alpha\) near −1.

A similar argument shows that

\[
corr(x_t, x_{t-k}) = \alpha^k
\]

for all \(k\). Thus, there is always some correlation between present and future values, but this correlation
dies down as we look further into the future. The implication is that future values are always forecastable, but forecasting becomes more difficult (inaccurate) as the lead time increases.

Forecasting for AR models is achieved by the same strategy used earlier for MA models: obtain an expression for the desired future value in terms of the \( \{x_t\} \) and the \( \{\varepsilon_t\} \), and then replace all unknown terms by their optimal forecasts (which may be zero). Specifically, for one-step prediction in the AR(1) model, we have

\[
x_{n+1} = \alpha x_n + \varepsilon_{n+1}
\]

The optimal forecast of \( \varepsilon_{n+1} \) is zero since \( \varepsilon_{n+1} \) is uncorrelated with all present and past values of \( \{x_t\} \). Thus, the optimum forecast of \( x_{n+1} \) is

\[
f_{n+1} = \alpha x_n
\]

The one-step forecast error is

\[
e_{n+1} = x_{n+1} - f_{n+1} = \varepsilon_{n+1}
\]

so the sequence of one-step errors is a white noise. For two-step prediction, use the relation

\[
x_{n+2} = \alpha x_{n+1} + \varepsilon_{n+2}
\]

replace \( \varepsilon_{n+2} \) by 0, and replace \( x_{n+1} \) by its optimal forecast \( f_{n+1} \) to obtain

\[
f_{n+2} = \alpha f_{n+1} = \alpha^2 x_n
\]

In general, the optimal \( h \)-step forecast is

\[
f_{n+h} = \alpha^h x_n
\]

As the lead time \( h \) increases, the forecast approaches zero (i.e., the series mean).

A generalization of the AR(1) model is the \( p \)'th order autoregression AR(p), generated by

\[
x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \cdots + \alpha_p x_{t-p} + \varepsilon_t
\]

The solution to this difference equation depends on the \( p \) starting values of \( \{x_t\} \) and on the \( \{\varepsilon_t\} \) series. The AR(p) series will be stationary if the largest root \( \theta \) of the equation (in the complex variable \( z \))
\[ z^p = \alpha_1 z^{p-1} + \alpha_2 z^{p-2} + \cdots + \alpha_{p-1} z + \alpha_p \]

satisfies \(|\theta| < 1\). In this case, the correlation function \( corr(x_t, x_{t-k}) \) (which is, because of stationarity, a function of \( k \) alone) will approximately lie in the region \( \pm A \pi^k \). if \( k \) is not too small. This helps determine the shape of the plot of the correlation function against \( k \). (More about this later.)

To forecast an AR(p) model with known parameters, use the usual strategy. For one-step forecasts, use

\[ x_{n+1} = (\alpha_1 x_n + \cdots + \alpha_p x_{n-p+1}) + \epsilon_{n+1}, \]

and replace \( \epsilon_{n+1} \) by zero to obtain the forecast

\[ f_{n,1} = \alpha_1 x_n + \cdots + \alpha_p x_{n-p+1}. \]

For two-step forecasts, use

\[ x_{n+2} = \alpha_1 x_{n+1} + (\alpha_2 x_n + \cdots + \alpha_p x_{n-p+2}) + \epsilon_{n+2}, \]

replace \( \epsilon_{n+2} \) by zero and replace \( x_{n+1} \) by its optimal forecast \( f_{n,1} \) to obtain

\[ f_{n,2} = \alpha_1 f_{n,1} + (\alpha_2 x_n + \cdots + \alpha_p x_{n-p+2}). \]

h-step forecasts are obtained similarly.
Yen Per U.S. Dollar
Jan 1985 to July 1992

Mean Adjusted Data and
Random Walk Prediction

Adjusted Data and
AR(1) Prediction, .5 x(t-1)

This Month’s vs Last
Month’s Exchange Rate

Exchange Rate


120 160 200 240


-20 0 20 40 60 80


-20 0 20 40 60 80


120 160 200 240

120 160 200 240

Last Month