Recap

- AR(p) and MA(q)
- ARMA(p, q)
  - The Ψ-weights of an ARMA Model : $\phi(z)\psi(z) = \theta(z)$
- The ACF of an ARMA
- Partial ACF: order of dependence of ARMA or AR.

Table: Behavior of the ACF and PACF for 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

- To estimate the partial ACF, Yule-Walker equation
- For a given lag k, it can be shown that the ϕ_{kk} satisfy the Yule-Walker equations:

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \cdots + \phi_{kk}\rho_{j-k} \text{ for } \ j = 1, 2, \cdots, k$$

- Given ρ's, solve Yule-Walker to obtain ϕ_{kk}.
- If the process is AR(p), then $\phi_{pp} = \phi_p$.
- Given r', solve Yule-Walker to obtain $\hat{\phi}_{kk}$.

Levinson (1947) and Durbin (1960) gave an efficient method for obtaining the solutions to the Yule-Walker equation for either theoretical or sample partial autocorrelations. It is shown that the equations can be solved recursively as follows:

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_j}$$

where

$$\phi_{k,j} = \phi_{k-1,j} - \phi_{k,k}\phi_{k-1,k-j}$$

Replace ρ's with r's, obtain sample PACF.

Observations

- any time series without a constant mean over time is non-stationary.
- Consider the following model:

$$Y_t = \mu_t + X_t$$

  where μ_t is a nonconstant mean function and X_t is a zero-mean, stationary series
- Such model is reasonable only if there are good reasons for believing that the deterministic trend is appropriate 'forever'.
- Just because a segment of the time series looks like it is increasing (or decreasing) approximately linearly, do we believe that the linearity is intrinsic to the process and will persist in the future?

- Frequently in applications, particularly in business and economics, we cannot legitimately assume a deterministic trend.
- Recall the random walk time series, it appears s strong upward trend, but it has a constant, zero-mean, and contains no deterministic trend at all.

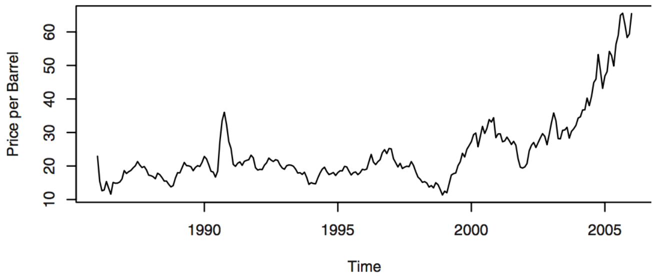


Figure: The monthly price of a barrel of crude oil from January 1986 through January 2006

- The series displays considerable variation, especially since 2001, and a stationary model does not seem to be reasonable.
- We will discover in later classes that no deterministic trend model works well for this series,
- but one of the nonstationary models that have been described as containing stochastic trends does seem reasonable.

Stationarity Through Differencing

Consider again the AR(1) model:

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

We have seen that assuming e_t is a true ‘innovation’ (that is, e_t is uncorrelated with $Y_{t-1}, Y_{t-2}, \cdots$), we must have $|\phi| < 1$.

Stationarity Through Differencing

Consider $|\phi| \geq 1$:

$$Y_t = 3Y_{t-1} + e_t$$

Iterating into the past we have:

$$Y_t = e_t + 3e_{t-1} + 3^2 e_{t-2} + \cdots + 3^{t-1} e_1 + 3^t Y_0$$

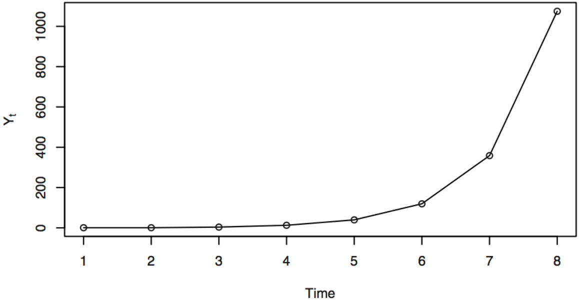
The influence of the distant past values of Y_t and e_t does not die out, rather the weights applied to Y_0 and e_1 grow exponentially large.

Stationarity Through Differencing

Exhibit 5.2 **Simulation of the Explosive "AR(1) Model"** $Y_t = 3Y_{t-1} + e_t$

t	1	2	3	4	5	6	7	8
e_t	0.63	−1.25	1.80	1.51	1.56	0.62	0.64	−0.98
Y_t	0.63	0.64	3.72	12.67	39.57	119.33	358.63	1074.91

(a) Simulation of the Explosive AR(1) model: $Y_t = 3Y_{t-1} + e_t$



(b) An Explosive AR(1) Series

Stationarity Through Differencing

The explosive behaviour of such a model is reflected in the model's variance and covariance functions.

$$
\begin{aligned}
Var(Y_t) &= \frac{1}{8}(9^t - 1)\sigma_e^2 \\
Cov(Y_t, Y_{t-k}) &= \frac{3^k}{8}(9^{t-k} - 1)\sigma_e^2
\end{aligned}
$$

Notice that

$$
Corr(Y_t, Y_{t-k}) = 3^k \sqrt{\frac{9^{t-k} - 1}{9^t - 1}} \approx 1
$$

for large t and moderate k

Stationarity Through Differencing

The same general exponential growth or explosive behaviour will occur for any $|\phi| > 1$. If $\phi = 1$, the AR(1) model equation is

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

This the relationship satisfied by the random walk. Alternatively, we can write this as

$$\nabla Y_t = e_t$$

where $\nabla Y_t = Y_t - Y_{t-1}$ is the **first difference** of Y_t.

Stationarity Through Differencing

The random walk is easily extended to a more general model whose first difference is some stationary process, not just white noise. Several somewhat different sets of assumptions can lead to models whose first difference is a stationary process.

Stationarity Through Differencing

Suppose

$$Y_t = M_t + X_t$$

where M_t is a series that is changing only slowly over time. Here M_t could be either deterministic or stochastic. If we assume that M_t is approximately constant over every two consecutive time points, we might estimate (predict) M_t at t by choosing β_0 so that

$$\sum_{j=0}^{1} (Y_{t-j} - \beta_{0,t})^2$$

is minimized.

Stationarity Through Differencing

This clearly leads to

$$\hat{M}_t = \frac{1}{2}(Y_t + Y_{t-1})$$

and the 'detrended' series at time t is then

$$Y_t - \hat{M}_t = Y_t - \frac{1}{2}(Y_t + Y_{t-1}) = \frac{1}{2}(Y_t - Y_{t-1}) = \frac{1}{2}\nabla Y_t$$

Stationarity Through Differencing

A second set of assumptions might be that M_t is stochastic and changes slowly over time governed by a random walk model. Suppose, for example, that

$$Y_t = M_t + e_t \qquad \text{with} \qquad M_t = M_{t-1} + \epsilon_t$$

where e_t and ϵ_t are independent white noise series. Then

$$\begin{aligned} \nabla Y_t &= \nabla M_t + \nabla e_t \\ &= \epsilon_t + e_t - e_{t-1} \end{aligned}$$

which would have the autocorrelation function of MA(1) series with

$$\rho_1 = -\frac{1}{2 + (\sigma_\epsilon^2/\sigma_e^2)}$$

In either of these situtations, we see ∇Y_t is a stationary process.

Stationarity Through Differencing

See the oil price time series again.

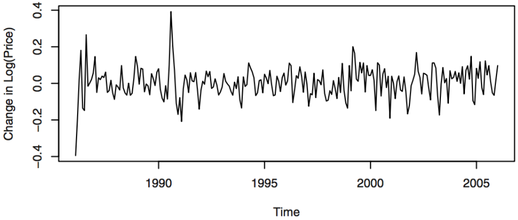


Figure: The difference Series of the Logs of the Oil Price Time series

- This series is more like a stationary process than the original one.
- There are outliers that should be considered.

Stationary Second-difference Models

Again assume $Y_t = M_t + X_t$, but now assume that M_t is linear in time over three consecutive time points. We can now estimate (predict) M_t at the middle time point t by choosing $\beta_{0,t}$ and $\beta_{1,t}$ to minimize

$$\sum_{j=-1}^{1} (Y_{t-j} - (\beta_{0,t} + j\beta_{1,t}))^2$$

The solution yields

$$\hat{M}_t = \frac{1}{3}(Y_{t+1} + Y_t + Y_{t-1})$$

Stationary Second-difference Models

Thus the detrended series is

$$
\begin{aligned}
Y_t - \hat{M}_t &= Y_t - \frac{1}{3}(Y_{t+1} + Y_t + Y_{t-1}) \\
&= \left(-\frac{1}{3}\right)(Y_{t+1} - 2Y_t + Y_{t-1}) \\
&= \left(-\frac{1}{3}\right)\nabla(\nabla Y_{t+1}) \\
&= \left(-\frac{1}{3}\right)\nabla^2 Y_{t+1}
\end{aligned}
$$

a constant multiple of the centered **second difference** of Y_t. Notice that we have differenced twice, but both differences at lag 1.

Stationary Second-difference Models

Alternatively, we might assume that

$$Y_t = M_t + e_t \text{ where } M_t = M_{t-1} + W_t \text{ and } W_t = W_{t-1} + \epsilon_t$$

where e_t and ϵ_t independent white noise time series. Here the stochastic trend M_t is such that its **rate of change**, ∇M_t, is changing slowly over time. Then

$$\nabla Y_t = \nabla M_t + \nabla e_t = W_t + \nabla e_t$$

and

$$\begin{aligned} \nabla^2 Y_t &= \nabla W_t + \nabla^2 e_t \\ &= \epsilon_t + (e_t - e_{t-1}) - (e_{t-1} - e_{t-2}) \\ &= \epsilon_t + e_t - 2e_{t-1} + e_{t-2} \end{aligned}$$

which has the autocorrelation function of an MA(2) process.

Stationary Second-difference Models

The important point is that the second difference of the non-stationary process Y_t is stationary.

Definition

A time series $\{Y_t\}$ is said to follow an **integrated autoregressive moving average** model if the dth difference $W_t = \nabla^d Y_t$ is a stationary ARMA process. If $\{W_t\}$ follows an ARMA(p, q) model, we say that $\{Y_t\}$ is an ARIMA(p, d, q) process.

For practical purposes, we can usually take $d = 1$ or at most 2.

ARIMA Models

Consider than an ARIMA$(p, 1, q)$ process. With $W_t = Y_t - Y_{t-1}$, we have

$$W_t = \phi_1 W_{t-1} + \cdots + \phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q}$$

or in terms of the observed series

$$\begin{aligned} Y_t - Y_{t-1} &= \phi_1(Y_{t-1} - Y_{t-2}) + \phi_2(Y_{t-2} - Y_{t-3}) + \cdots \\ &\quad \phi_p(Y_{t-p} - Y_{t-p-1}) + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \end{aligned}$$

ARIMA Models

which we may rewrite as

$$\begin{aligned} Y_t &= (1+\phi_1)Y_{t-1} + (\phi_2 - \phi_1)Y_{t-2} + \cdots \\ &\quad + (\phi_p - \phi_{p-1})Y_{t-p} - \phi_p Y_{t-p-1} + e_t - \theta_1 e_{t-1} - \cdots - \theta_q \theta_{t-q} \end{aligned}$$

We call this the **difference equation form** of the model.

ARIMA Models

Notice that it appears to be an ARMA$(p+1, q)$ process. However, the characteristic polynomial satisfies

$$1-(1+\phi_1)x-(\phi_2-\phi_1)x^2-(\phi_3-\phi_2)x^3-\cdots-(\phi_p-\phi_{p-1})x^p+\phi_p x^{p+1}$$
$$= (1-\phi_1 x-\phi_2 x^2-\cdots-\phi_p x^p)(1-x)$$

This factorization clearly shows the root at $x = 1$, which implies nonstationarity.
The remaining roots, however, are the roots of the characteristic polynomial of the *stationary process* ∇Y_t.

ARIMA Models

- Explicit representations of the observed series in terms of either W_t or the white noise series underlying W_t are more difficult than in the stationary case.
- Since nonstationary processes are not in statistical equilibrium, we cannot assume that they go infinitely into the past or that they start at $t = -\infty$.
- However, we can and shall assume that they start at some time point $t = -m$, say, where $-m$ is earlier than time $t = 1$, at which point we first observed the series.

For convenience, we take $Y_t = 0$ for $t < -m$. The difference equation $Y_t - Y_{t-1} = W_t$ can be solved by summing both sides from $t = -m$ to $t = t$ to get the representation

$$Y_t = \sum_{j=-m}^{t} W_j$$

for the ARIMA$(p, 1, q)$ process.

ARIMA Models

The ARIMA$(p, 2, q)$ process can be dealt with similarly by summing twice to get the representations

$$
\begin{aligned}
Y_t &= \sum_{j=-m}^{t} \sum_{i=-m}^{j} W_i \\
&= \sum_{j=0}^{t+m} (j+1) W_{t-j}
\end{aligned}
$$

These representations have limited use but can be used to investigate the covariance properties of ARIMA models and also to express Y_t in terms of the white noise series $\{e_t\}$

ARIMA Models

- If the process contains no autoregressive terms, we call it an integrated moving average and abbreviate the name to $\text{IMA}(d, q)$.
- If no moving average terms are present, we denote the model as $\text{ARI}(p, d)$.

The IMA(1,1) Model

The simple IMA(1,1) model satisfactorily represents numerous time series, especially those arising in economics and business. In difference equation form, the model is

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

The IMA(1,1) Model

To write Y_t explicitly as a function of present and past noise values, we use the form

$$Y_t = \sum_{j=-m}^{t} W_j$$

and the fact that $W_t = e_t - \theta e_{t-1}$ in this case. After a little rearrangement, we can write

$$Y_t = e_t + (1-\theta)e_{t-1} + (1-\theta)e_{t-2} + \cdots + (1-\theta)e_{-m} - \theta e_{-m-1}$$

Notice that in contrast to our stationary ARMA models, the weights on the white noise terms *do not die out* as we go into the past. Since we are assuming that $-m < 1$ and $0 < t$, we may usefully think of Y_t as mostly an equally weighted accumulation of a large number of white noise values.

The IMA(1,1) Model

We can derive the variances and correlations of IMA(1,1) as follows:

$$\begin{aligned} Var(Y_t) &= [1+\theta^2+(1-\theta)^2(t+m)]\sigma_e^2 \\ Corr(Y_t, Y_{t-k}) &= \frac{1-\theta+\theta^2+(1-\theta)^2(t+m-k)}{[Var(Y_t)Var(Y_{t-k})]^{1/2}} \\ &\approx \sqrt{\frac{t+m-k}{t+m}} \\ &\approx 1 \text{ for large } m \text{ and moderate } k \end{aligned}$$

as t increases, $Var(Y_t)$ increases and could be quite large. Also, the correlation between Y_t and Y_{t-k} will be strongly positive for many lags $k = 1, 2, \cdots$.

The IMA(2,2) Model

Consider

$$Y_t = M_t + e_t \text{ where } M_t = M_{t-1} + W_t \text{ and } W_t = W_{t-1} + \epsilon_t$$

This leads to an IMA(2,2) model. Write it in Difference Equation form:

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

or

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

The IMA(2,2) Model

The representation

$$Y_t = \sum_{j=0}^{t+m} (j+1) W_{t-j}$$

may be used to express Y_t in terms of $e_t, e_{t-1}, \cdots$. After some algebra, we find that

$$\begin{aligned} Y_t &= e_t + \sum_{j=1}^{t+m} \Psi_j e_{t-j} - [(t+m+1)\theta_1 + (t+m)\theta_2] e_{-m-1} \\ &\quad -(t+m+1)\theta_2 e_{-m-2} \end{aligned}$$

where $\Psi_j = 1 + \theta_2 + (1 - \theta_1 - \theta_2) j$ for $j = 1, 2, 3, \cdots, t+m$. Once more we see that the Ψ-weights do not die out but form a linear function of j.

The IMA(2,2) Model

- Variances and correlations for Y_t can be obtained from the representation given in previous equation, but the calculations are tedious.
- We shall simply note that the variance of Y_t increases rapidly with t and again $Corr(Y_t, Y_{t-k})$ is nearly 1 for all moderate k.

The IMA(2,2) Model

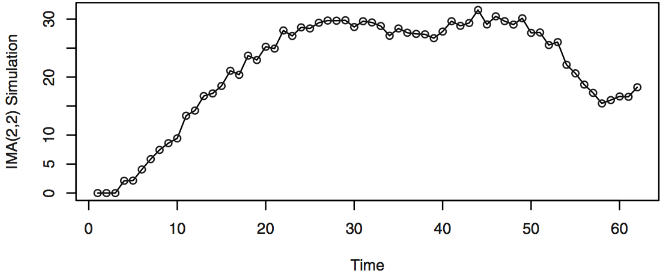


Figure: Simulation of an IMA(2,2) Series with $\theta_1 = 1$ and $\theta_2 = -0.6$.

- Notice the smooth change in the process values (and the unimportance of the zero-mean function).
- The increasing variance and the strong, positive neighboring correlations dominate the appearance of the time series plot.

The IMA(2,2) Model

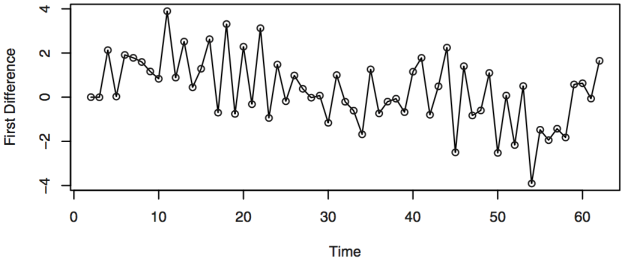


Figure: First Difference of the Simulated IMA(2,2) Series.

- This series is also nonstationary, as it is governed by an IMA(1,2) model.

The IMA(2,2) Model

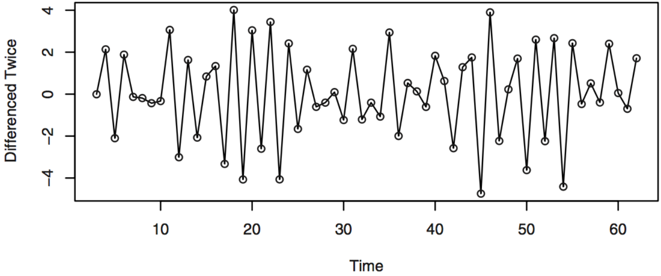


Figure: Second Difference of the Simulated IMA(2,2) Series.

- These values arise from a stationary MA(2) model with $\theta_1 = 1$ and $\theta_2 = -0.6$.
- The theoretical autocorrelations for this model are $\rho_1 = -0.678$ and $\rho_2 = 0.254$. These correlation values seem to be reflected in the appearance of the time series plot.

The ARI(1,1) Model

The ARI(1,1) process will satisfy:

$$Y_t - Y_{t-1} = \phi(Y_{t-1} - Y_{t-2}) + e_t$$

or

$$Y_t = (1 + \phi)Y_{t-1} - \phi Y_{t-2} + e_t$$

where $|\phi| < 1$.
Notice that this looks like a special AR(2) model. However, one of the roots of the corresponding AR(2) characteristic polynomial is 1, and this is not allowed in stationary models.

The ARI(1,1) Model

To find the Ψ-weights in this case, we shall use a technique that will generalize to arbitrary ARIMA models.
It can be shown that the Ψ-weights can be obtained by equating like powers of x in the identity:

$$(1-\phi_1 x-\phi_2 x^2-\cdots-\phi_p x^p)(1-x)^d(1+\Psi_1 x+\Psi_2 x^2+\cdots)$$
$$=(1-\theta_1 x-\theta_2 x^2-\cdots-\theta_q x^q)$$

The ARI(1,1) Model

In our case, this relationship reduces to

$$(1 - \phi x)(1 - x)(1 + \Psi_1 x + \Psi_2^x + \cdots) = 1$$

or

$$[1 - (1 + \phi)x + \phi x^2](1 + \Psi_1 x + \Psi_2 x^2 + \cdots) = 1$$

Equating like powers of x on both sides, we obtain

The ARI(1,1) Model

$$\begin{aligned} -(1+\phi) + \Psi_1 &= 0 \\ \phi - (1+\phi)\Psi_1 + \Psi_2 &= 0 \end{aligned}$$

or in general,

$$\Psi_k = (1+\phi)\Psi_{k-1} - \phi\Psi_{k-2}$$

with $\Psi_0 = 1$ and $\Psi_1 = 1 + \phi$.

This recursion with starting values allows us to compute as many Ψ-weights as necessary.
It can also be shown that in this case an explicit solution to the recursion is given as

$$\Psi_k = \frac{1-\phi^{k+1}}{1-\phi} \text{ for } \ k \geq 1$$

Questions?