Let x be a zero-mean stochastic process. A roughly linear relationship between successive observations may be described by a simple **autoregressive model** of the form

$$x_t = \phi x_{t-1} + u_t$$

where u is zero-mean white noise.

A stationary process x satisfying this **first-order difference equation** is called a **first-order autoregressive process** (AR(1) process).

Substituting $\phi x_{t-2}+u_{t-1}$ for x_{t-1}, $\phi x_{t-3}+u_{t-2}$ for x_{t-2}, … gives

$$\begin{aligned} x_t &= \phi x_{t-1}+u_t \\ &= \phi(\phi x_{t-2}+u_{t-1})+u_t \\ &= \phi^2 x_{t-2}+\phi u_{t-1}+u_t \\ &= \phi^2(\phi x_{t-3}+u_{t-2})+\phi u_{t-1}+u_t \\ &= \phi^3 x_{t-3}+\phi^2 u_{t-2}+\phi u_{t-1}+u_t \\ &\;\;\vdots \\ &= \phi^k x_{t-k}+\sum_{j=0}^{k-1}\phi^j u_{t-j} . \end{aligned}$$

If k is large and $|\phi|<1$, then the first part of this expression is negligible, hence

$$x_t = \sum_{j=0}^{\infty}\phi^j u_{t-j} .$$

IR

Suppose that $|\phi|<1$. Then the solution

$$x_t = \sum_{j=0}^{\infty} \phi^j u_{t-j}$$

of the difference equation

$$x_t = \phi x_{t-1} + u_t$$

is stationary because

$$E(x_t) = E\sum_{j=0}^{\infty} \phi^j u_{t-j} = \sum_{j=0}^{\infty} \phi^j E(u_{t-j}) = \sum_{j=0}^{\infty} \phi^j 0 = 0,$$

$$\text{Var}(x_t) = \sum_{j=0}^{\infty} (\phi^j)^2 \,\text{Var}(u_{t-j}) = \sum_{j=0}^{\infty} (\phi^2)^j \sigma^2 = \sigma^2 \sum_{j=0}^{\infty} (\phi^2)^j = \frac{\sigma^2}{1-\phi^2},$$

and

$$\text{Cov}(x_t, x_{t-k}) = \ldots = \frac{\phi^{|k|}\sigma^2}{1-\phi^2}$$

do not depend on t.

For example, the autocovariance at lag $k=1$ is given by

$$\text{Cov}(x_t, x_{t-1})$$

$$= E(u_t + \phi u_{t-1} + \phi^2 u_{t-2} + \ldots)(u_{t-1} + \phi u_{t-2} + \phi^2 u_{t-3} + \ldots)$$

$$= E(u_t + \phi u_{t-1} + \phi^2 u_{t-2} + \ldots) \sum_{j=1}^{\infty} \phi^{j-1} u_{t-j}$$

$$= E(u_t \sum_{j=1}^{\infty} \phi^{j-1} u_{t-j} + \phi u_{t-1} \sum_{j=1}^{\infty} \phi^{j-1} u_{t-j} + \phi^2 u_{t-2} \sum_{j=1}^{\infty} \phi^{j-1} u_{t-j} + \ldots)$$

$$= \sum_{j=1}^{\infty} \phi^{j-1}\, E u_t u_{t-j} + \sum_{j=1}^{\infty} \phi^{j}\, E u_{t-1} u_{t-j} + \sum_{j=1}^{\infty} \phi^{j+1}\, E u_{t-2} u_{t-j} + \ldots$$

$$= (\phi^0 0 + \phi^1 0 + \ldots) + (\phi^1 \sigma^2 + \phi^2 0 + \ldots) + (\phi^2 0 + \phi^3 \sigma^2 + \ldots) + \ldots$$

$$= \phi^1 \sigma^2 + \phi^3 \sigma^2 + \phi^5 \sigma^2 + \ldots$$

$$= \phi \sigma^2 \left((\phi^2)^0 + (\phi^2)^1 + (\phi^2)^2 + \ldots \right)$$

$$= \frac{\phi \sigma^2}{1-\phi^2}.$$

IM

Using the lag operator L we can write the equation

$$x_t = \phi x_{t-1} + u_t$$

as

$$u_t = x_t - \phi x_{t-1} = L^0(x_t) - \phi L\,(x_t)$$
$$= (L^0 - \phi L)(x_t)$$
$$= (1 - \phi L)(x_t)$$

and the equation

$$x_t = \sum_{j=0}^{\infty} \phi^j u_{t-j}$$

as

$$x_t = \sum_{j=0}^{\infty} \phi^j L^j (u_t)$$

$$= (\sum_{j=0}^{\infty} \phi^j L^j)(u_t)$$

$$= (\sum_{j=0}^{\infty} (\phi L)^j)(u_t).$$

A comparison of the equations

$$(1 - \phi L)(x_t) = u_t$$

and

$$x_t = (\sum_{j=0}^{\infty} (\phi L)^j)(u_t)$$

suggests that the operator

$$\sum_{j=0}^{\infty} (\phi L)^j$$

is the inverse operator of

$$1 - \phi L$$

and, more generally, that the lag operator follows the usual algebraic rules. We may therefore write

$$(1 - \phi L)^{-1} = \sum_{j=0}^{\infty} (\phi L)^j$$

or

$$\frac{1}{1 - \phi L} = \sum_{j=0}^{\infty} (\phi L)^j \,.$$

A stationary process x satisfying the p'th-order difference equation

$$x_t=\phi_1 x_{t-1}+\ldots+\phi_p x_{t-p}+u_t,$$

where u is zero-mean white noise and $\phi_p\neq 0$, is called an **autoregressive process of order *p*** (AR(p) process).

A stationary process X satisfying

$$x_t=\phi_1 x_{t-1}+\ldots+\phi_p x_{t-p}+u_t+\theta_1 u_{t-1}+\ldots+\theta_q u_{t-q},$$

where u is zero-mean white noise, $\phi_p\neq 0$, and $\theta_q\neq 0$, is called an **autoregressive moving average process of order (*p*,*q*)** (ARMA(p,q) process).

The equation

$$x_t=\phi_1 x_{t-1}+\ldots+\phi_p x_{t-p}+u_t+\theta_1 u_{t-1}+\ldots+\theta_q u_{t-q}$$

can also be written as

$$x_t-\phi_1 x_{t-1}-\ldots-\phi_p x_{t-p}=u_t+\theta_1 u_{t-1}+\ldots+\theta_q u_{t-q}$$

or as $(1-\phi_1 L-\ldots-\phi_p L^p)(x_t)=(1+\theta_1 L+\ldots+\theta_q L^q)(u_t)$.

An ARMA(p,0) process is an AR(p) process.
An ARMA(0,q) is called a **moving average process of order *q*** (MA(q) process).

A process x is called an **autoregressive integrated moving average process of order (*p*,*d*,*q*)** (ARIMA(p,d,q) process) if its d'th difference is an ARMA(p,q) process, i.e.,

$$(1-\phi_1 L-\ldots-\phi_p L^p)(\Delta^d x_t)=(1+\theta_1 L+\ldots+\theta_q L^q)(u_t).$$

ARIMA(p,d,q) processes may be generalized by permitting the degree of differencing, d, to take fractional values.

The **fractional differencing operator** $\Delta^d=(1-L)^d$ is defined as a power series expansion in integer powers of L:

$$(1-L)^d=1-dL+d(d-1)\frac{L^2}{2}-d(d-1)(d-2)\frac{L^3}{3!}+\ldots$$

A process is called an **autoregressive fractionally integrated moving average process** (ARFIMA(p,d,q) process) if the fractionally differenced process is an ARMA process.

We say that an ARMA(p,q) representation

$$(1-\phi_1 L-\ldots-\phi_p L^p)(x_t)=(1+\theta_1 L+\ldots+\theta_q L^q)(u_t),$$

is **causal** if x_t can be expressed in terms of present and past shocks, i.e.,

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

It is called **invertible** if u_t has a representation of the form

$$u_t=\sum_{j=0}^{\infty}\xi_j x_{t-j}\ .$$

Causality is equivalent to

$$\Phi(z)=1-\phi_1 z-\ldots-\phi_p z^p\neq 0\ \forall |z|\leq 1$$

and invertibility is equivalent to

$$\Theta(z)=1+\theta_1 z+\ldots+\theta_q z^q\neq 0\ \forall |z|\leq 1.$$

For example, if the only zero $1/\phi_1$ of the polynomial

$$\Phi(z)=1-\phi_1 z$$

satisfies

$$|1/\phi_1|>1,$$

then

$$|\phi_1|<1$$

and

$$x_t=\sum_{j=0}^{\infty}\phi^j u_{t-j}\ ,$$

hence the AR(1) representation

$$(1-\phi_1 L)(x_t)=u_t$$

is causal.

Exercise: Determine which of the following ARMA representations are causal and which are invertible.

(i)	$x_t - 0.7x_{t-1} = u_t$	M1
(ii)	$x_t + 0.4x_{t-1} - 1.5x_{t-2} = u_t$	
(iii)	$x_t = u_t - 0.2u_{t-1}$	
(iv)	$x_t = u_t + 0.8u_{t-1} + 2.6u_{t-2}$	M2
(v)	$x_t = u_t - \sqrt{2}\, u_{t-1} + u_{t-2}$	
(vi)	$x_t = u_t + u_{t-2}$	
(vii)	$x_t - x_{t-1} = u_t - 0.3u_{t-1}$	M3
(viii)	$x_t - 0.2x_{t-1} = u_t + 0.4u_{t-1} - 1.6u_{t-2}$	
(ix)	$x_t - 1.2x_{t-1} + 0.1x_{t-2} = u_t + 0.3u_{t-1} - 1.2u_{t-2}$	

Exercise: Fit an ARMA(1,2) model to the growth rates of the quarterly GDP.

```
r.dm <- r-mean(r) # r is demeaned
p <- 1; q <- 2
arima(r.dm,order=c(p,0,q),include.mean=FALSE,
        transform.pars=TRUE)
```

If **transform.pars=TRUE**, the AR parameters are checked and, if necessary, transformed to ensure causality.

We obtain the following estimates of ϕ_1, θ_1, θ_2, and σ^2:

```
Coefficients:
       ar1     ma1     ma2
    0.2844  0.0547  0.1557

sigma^2 estimated as 7.734e-05
```

Exercise: Show that the MA(∞) representation of the causal ARMA(1,1) process

$$x_t - \phi x_{t-1} = u_t + \theta u_{t-1}$$

is given by

$$x_t = u_t + (\theta+\phi)\sum_{j=1}^{\infty}\phi^{j-1}u_{t-j}\ .$$

Solution 1: We have

$$\tfrac{1+\theta z}{1-\phi z} = \sum_{j=0}^{\infty}\psi_j z^j$$

or, equivalently,

$$(1+\theta z) = (1-\phi z)(\ \psi_0 + \psi_1 z + \psi_2 z^2 + \ldots).$$

Equating coefficients of z^j we obtain:

$j=0$: $1=\psi_0$,

$j=1$: $\theta = \psi_1 - \phi\psi_0 \Rightarrow \psi_1 = \theta + \phi\psi_0 = \theta + \phi$,

⋮

Solution 2: We have

$$x_t = \tfrac{1+\theta L}{1-\phi L}u_t = \tfrac{1}{1-\phi L}u_t + \theta L\tfrac{1}{1-\phi L}u_t$$

$$= \sum_{j=0}^{\infty}\phi^j u_{t-j} + \theta L\sum_{j=0}^{\infty}\phi^j u_{t-j}$$

$$= u_t + \phi\sum_{j=1}^{\infty}\phi^{j-1}u_{t-j} + \theta\sum_{j=1}^{\infty}\phi^{j-1}u_{t-j}\ .$$

ME

Exercise: Show that the autocovariances

$$\gamma(k)=\mathrm{Cov}(x_t,x_{t-k})$$

of a causal ARMA(1,1) process

$$x_t-\phi x_{t-1}=u_t+\theta u_{t-1}$$

satisfy

$$\gamma(k)=\phi\gamma(k-1), \text{ if } k>1.$$

Solution: Multiplying the difference equation by x_{t-k} and taking expectations we obtain for $k>1$

$$x_t\, x_{t-k} - \phi x_{t-1}\, x_{t-k} = u_t\, x_{t-k} + \theta u_{t-1}\, x_{t-k} ,$$

$$E x_t\, x_{t-k} - \phi E x_{t-1}\, x_{t-k} = E u_t\, x_{t-k} + \theta\, E u_{t-1}\, x_{t-k} ,$$

$$\gamma(k)-\phi\gamma(k-1)=0.$$

Exercise: Find the stationary solution of the noncausal univariate AR(1) equation

$$x_t=\phi x_{t-1}+u_t,\ |\phi|>1.$$

Solution: $x_t=\phi x_{t-1}+u_t \Rightarrow x_{t-1}=\frac{1}{\phi}x_t-\frac{1}{\phi}u_t$

$$\Rightarrow x_t=\frac{1}{\phi}x_{t+1}-\frac{1}{\phi}u_{t+1}$$

$$=\frac{1}{\phi}\left(\frac{1}{\phi}x_{t+2}-\frac{1}{\phi}u_{t+2}\right)-\frac{1}{\phi}u_{t+1}$$

$$=\frac{1}{\phi^2}x_{t+2}-\frac{1}{\phi^2}u_{t+2}-\frac{1}{\phi}u_{t+1}$$

$$\vdots$$

$$=-\sum_{j=1}^{\infty}\frac{1}{\varphi^j}u_{t+j}$$

MN

Exercise: Show that the stationary solution

$$x_t = -\sum_{j=1}^{\infty} \frac{1}{\varphi^j} u_{t+j}$$

of the noncausal univariate AR(1) equation

$$x_t = \phi x_{t-1} + u_t, \ |\phi| > 1, \ \text{var}(u_t) = \sigma^2$$

also satisfies the causal univariate equation

$$x_t = \frac{1}{\phi} x_{t-1} + v_t$$

for a suitable white noise v.

Solution: If $v_t = x_t - \frac{1}{\phi} x_{t-1}$

$$= -\frac{1}{\phi} x_{t-1} + x_t$$

$$= \frac{1}{\phi} \sum_{j=1}^{\infty} \frac{1}{\varphi^j} u_{t-1+j} - \sum_{j=1}^{\infty} \frac{1}{\varphi^j} u_{t+j}$$

$$= \frac{1}{\phi^2} u_t + \left(\frac{1}{\phi^3} - \frac{1}{\phi}\right) u_{t+1} + \left(\frac{1}{\phi^4} - \frac{1}{\phi^2}\right) u_{t+2} + \ldots,$$

then the mean of v_t is zero 0 and its variance is given by

$$\sigma^2\left(\frac{1}{\phi^4} + \left(\frac{1}{\phi^3} - \frac{1}{\phi}\right)^2 + \left(\frac{1}{\phi^4} - \frac{1}{\phi^2}\right)^2 + \ldots\right)$$

$$= \sigma^2\left(\frac{1}{\phi^4} + \left(\frac{1}{\phi^6} + \frac{1}{\phi^8} + \ldots\right) - 2\left(\frac{1}{\phi^4} + \frac{1}{\phi^6} + \ldots\right) + \left(\frac{1}{\phi^2} + \frac{1}{\phi^4} + \ldots\right)\right)$$

$$= \frac{\sigma^2}{\phi^2}.$$

Moreover, if $h > 0$, then

$$\text{E}v_t v_{t+h} = \sigma^2\left(\left(\frac{1}{\phi^{h+2}} - \frac{1}{\phi^h}\right)\frac{1}{\phi^2} + \left(\frac{1}{\phi^{h+3}} - \frac{1}{\phi^{h+1}}\right)\left(\frac{1}{\phi^3} - \frac{1}{\phi}\right) + \ldots\right)$$

$$= \frac{\sigma^2}{\phi^h}\left(\left(\frac{1}{\phi^2} - \frac{1}{\phi^0}\right)\frac{1}{\phi^2} + \left(\frac{1}{\phi^3} - \frac{1}{\phi^1}\right)\left(\frac{1}{\phi^3} - \frac{1}{\phi}\right) + \ldots\right)$$

$$= \frac{\sigma^2}{\phi^h}\left(\left(\frac{1}{\phi^4} - \frac{1}{\phi^2}\right) + \left(\frac{1}{\phi^6} - 2\frac{1}{\phi^4} + \frac{1}{\phi^2}\right) + \left(\frac{1}{\phi^8} - 2\frac{1}{\phi^6} + \frac{1}{\phi^4}\right) + \ldots\right)$$

$$= 0.$$

Thus v is white noise. MC

This exercise shows that the stationary solution of the noncausal equation also satisfies a causal equation. Nothing will therefore be lost if we consider only causal equations.

Exercise: Show that the invertible MA(1) model

$$x_t = u_t + \tfrac{1}{2} u_{t-1}, \ \text{var}(u_t) = 4$$

implies the same autocovariance function as the non-invertible MA(1) model

$$x_t = u_t + 2u_{t-1}, \ \text{var}(u_t) = 1.$$

MO

Remark: Since we can only observe x_t and not u_t, we cannot distinguish between the two models. They are therefore called **observationally equivalent**.

Exercise: Use the model-selection criteria AIC and BIC to find the "best" ARMA(p,q) model with $p\leq3$, $q\leq3$ for the (demeaned) growth rates of the quarterly GDP.

Hint: Trying to strike a balance between the goodness of fit (measured by the log likelihood L) and the complexity of the model (measured by the total number $p+q+1$ of model parameters $\phi_1,\ldots,\phi_p,\theta_1,\ldots,\theta_q,\sigma^2$), these criteria select that model which minimizes

$$\mathrm{AIC}(p,q) = -2\,L + 2(p+q+1)$$

and

$$\mathrm{BIC}(p,q) = -2\,L + (p+q+1)\,\log(n),$$

respectively.

```
AIC <- BIC <- matrix(nrow=4,ncol=4)
for (p in 0:3) for (q in 0:3) {
  h <- arima(r.dm,order=c(p,0,q),include.mean=F)
  AIC[p+1,q+1] <- -2*h$loglik+2*(p+q+1)
  BIC[p+1,q+1] <- -2*h$loglik+(p+q+1)*log(n) }
0:3; cbind(0:3,AIC); 0:3; cbind(0:3,BIC)
```

	0	1	2	3
0	-1755.977	-1784.283	-1795.342	-1795.452
1	-1794.456	-1793.917	-1795.024	-1793.477
2	-1794.800	-1794.174	-1796.062	-1792.485
3	-1795.332	-1794.940	-1801.370	-1800.740

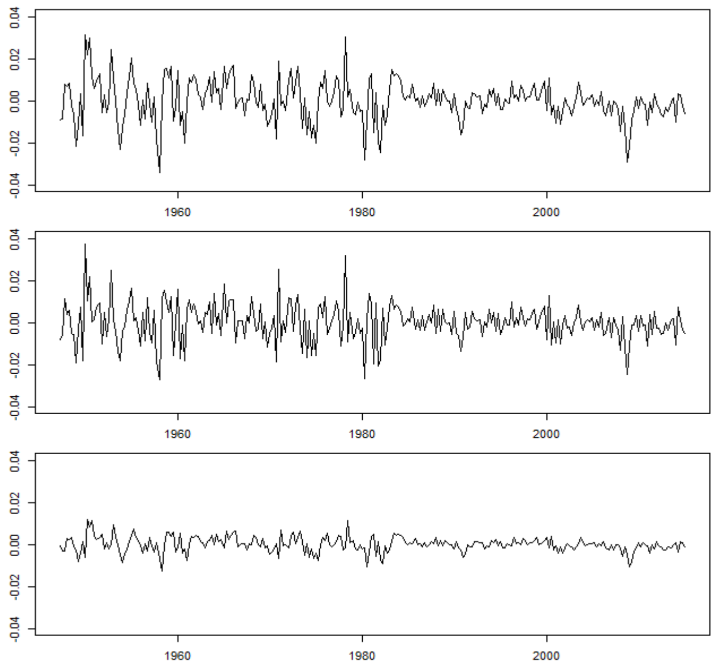
	0	1	2	3
0	-1752.368	-1777.064	-1784.514	-1781.014
1	-1787.237	-1783.089	-1780.586	-1775.430
2	-1783.971	-1779.737	-1778.014	-1770.828
3	-1780.894	-1776.893	-1779.713	-1775.473

AIC selects the ARMA(3,2) model and BIC selects the ARMA(1,0) model.

Unfortunately, AIC and BIC rarely select the same model. And to make matters worse, there are a lot more model-selection criteria that could be used. So it seems that automatic model selection with a model-selection criterion still contains a subjective element because we have to select the model-selection criterion first.

Exercise: Plot the (demeaned) growth rates of the quarterly GDP, the residuals from the ARMA(1,0) model, which is the best model for this time series according to BIC, and the fitted values.

```
par(mfrow=c(3,1),mar=c(2,2,1,1))
p <- 1; q <- 0
h <- arima(r.dm,order=c(p,0,q),include.mean=FALSE,
        transform.pars=TRUE)
plot(d[2:n],r.dm,type="l",ylim=c(-0.04,0.04))
plot(d[2:n],h$residuals,type="l",ylim=c(-0.04,0.04))
plot(d[2:n],r.dm-h$residuals,type="l",ylim=c(-0.04,0.04))
```



This ARMA model explains very little. The residuals are almost of the same size as the data.

The joint probability density of a sample $x_1,\ldots,x_n$ from a Gaussian AR(1) process

$$x_t = \phi x_{t-1} + u_t$$

with i.i.d. $N(0,\sigma^2)$ innovations u_t is given by

$$\begin{aligned}
&f(x_1,\ldots,x_n)\\
&= f(x_1)f(x_2|x_1)f(x_3|x_1,x_2)\ldots f(x_n|x_1,\ldots,x_{n-1})\\
&= f(x_1)f(x_2|x_1)f(x_3|x_2)\ldots f(x_n|x_{n-1})\\
&= f(x_1)\prod_{t=2}^{n} f(x_t|x_{t-1})\\
&= \frac{1}{\sqrt{2\pi\frac{\sigma^2}{1-\rho^2}}}\exp\Big(-\frac{x_1^2}{2\frac{\sigma^2}{1-\rho^2}}\Big)\prod_{t=2}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\Big(-\frac{(x_t-\phi x_{t-1})^2}{2\sigma^2}\Big).
\end{aligned}$$

The maximum likelihood (ML) estimates and the conditional ML estimates of the parameters ϕ and σ^2 are obtained by maximization of the log likelihood function and the log likelihood function conditioned on the first observation, respectively, with respect to ϕ and σ^2.

We have

$$\begin{aligned}
&\log f(x_1,\ldots,x_n;\phi,\sigma^2)\\
&= -\frac{1}{2}\log\Big(\frac{2\pi\sigma^2}{1-\rho^2}\Big) - \frac{x_1^2(1-\rho^2)}{2\sigma^2}\\
&\quad -\frac{n-1}{2}\log(2\pi\sigma^2) - \sum_{t=2}^{n}\frac{(x_t-\phi x_{t-1})^2}{2\sigma^2}.
\end{aligned}$$

and

$$\begin{aligned}
&\log f(x_2,\ldots,x_n|x_1)\\
&= -\frac{n-1}{2}\log(2\pi\sigma^2) - \sum_{t=2}^{n}\frac{(x_t-\phi x_{t-1})^2}{2\sigma^2}.
\end{aligned}$$

ML

Exercise: Find the conditional ML estimates of the parameters ϕ and σ^2.

MM