Introduction

In economics, most cases we want to model relationships *between* variables, and often simultaneously. That means we need to move from univariate time series to multivariate. We do it in two steps. First we have only one dependent variable, in the ADL model, before we go to models with several dependent variables (VARs).

Autoregressive distributed lag models

AR(p) specifications

$$y_t = \mu + \sum_{j=1}^{p} \phi_j y_{t-j} + u_t$$

Such models are easily expanded by adding additional predictors x_t to the autoregressive compoent, creating what is called the *autoregressive distributed lag model*

$$y_t = \mu + \sum_{j=1}^{p} \phi_j y_{t-j} + \sum_{j=1}^{q} b_j x_{t-j} + u_t$$

We term this the $ADL(p, q)$ model. Note that x_t can be a vector of variables observed at time t.
This model relies on stationarity of the series.

Example

Let us model changes in US inflation as a function of past values of inflation, using data from 1962–1999.

The estimated AR(1) relathionsip is

$$\widehat{\Delta Inf_t} = \underset{(0.14)}{0.02} - \underset{(0.106)}{0.211} \Delta Inf_{t-1}$$

$\bar{R}^2 = 0.04$ (Numbers in parenthesis are standard errors)

Why not add a few lags?

The estimated AR(4) relationship is

$$\widehat{\Delta Inf_t} = \underset{(0.12)}{0.02} - \underset{(0.10)}{0.21} \Delta Inf_{t-1} - \underset{(0.09)}{0.32} \Delta Inf_{t-2} + \underset{(0.09)}{0.19} \Delta Inf_{t-3} - \underset{(0.10)}{0.04} \Delta Inf$$

$\bar{R}^2 = 0.21$

Note that several of these coefficients are highly significant.

In macroeconomics, there is a relationship between unemployment and inflation such that if unemployment is currently high, inflation tends to fall. (the short-run Phillips curve).
Let us therefore add past values of unemployment, first just one lag

$$\begin{aligned}\widehat{\Delta Inf_t} &= \underset{(0.55)}{1.42} - \underset{(0.26)}{0.09}\,\Delta Inf_{t-1} - \underset{(0.10)}{0.40}\,\Delta Inf_{t-2} + \underset{(0.08)}{0.11}\,\Delta Inf_{t-3} - \underset{(0.}{0.0} \\ &\quad - \underset{(0.10)}{0.23}\,Unemp_{t-1}\end{aligned}$$

$\bar{R}^2 = 0.22$
t-stat Unemployment one lag is -2.33, which is significant at the 5% level.

And then three more lags

$$\widehat{\Delta Inf_t} = \underset{(0.55)}{1.42} - \underset{(0.26)}{0.09}\,\Delta Inf_{t-1} - \underset{(0.10)}{0.40}\,\Delta Inf_{t-2} + \underset{(0.08)}{0.11}\,\Delta Inf_{t-3} - 0.0$$
$$-\underset{(0.47)}{2.68}\,Unemp_{t-1} + \underset{(0.89)}{3.43}\,Unemp_{t-2} - \underset{(0.89)}{1.04}\,Unemp_{t-3} -$$

$\bar{R}^2 = 0.35$
F-stat joint test of second through fourth lag 4.93 (p-value 0.003), hence jointly significant

Granger Causality

This setup allows us to define the concept of *Granger Causality*. A variable x_t is said to *Granger Cause* another variable y_t if past values of x_t are significant determinants of current values of y_t.
Given an ADL(p,q) model

$$y_t = \mu + \sum_{j=1}^{p} \phi_j y_{t-j} + \sum_{j=1}^{q} b_j x_{t-j} + u_t$$

x is said to Granger cause y if the hypothesis that all coefficients b_j are jointly zero is rejected.
The null hypothesis in a Granger causailty test is that one of the regressors in x has no predictive content for y beyind that contained in the other regressors.
In the inflation example above, Unemployment is said to Granger cause inflation, since the coefficients on the unemployment variable is significant.
Important qualification: The Granger Causaility concept is something of a misnomer. It really means that x is an useful predictor of y, not that x *causes* y in the usual sense of the word,

Many will claim that the Norwegian stock market is driven by oil prices, that what happens at the Oslo Stock Exchange is related to what happens to the oil price. The question is whether the oil price has any predictive content for the Norwegian market. Let us look at monthly data, and estimate the model

$$R_{m,t} = \mu + \phi R_{m,t-1} + b \Delta OP_{t-1}$$

where R_m is a stock market index and ΔOP the change in the oil price (in USD).

- Estimate the model.
- Does Oil Granger cause the Norwegian stock market?

Reading the data

```
library(zoo)
OPm<- read.zoo("../../../../data/commodities/oil_prices/bre
                format="%Y%m%d",header=FALSE,skip=1)
Rm <- read.zoo("../../../../data/norway/stock_market_indice
                header=TRUE,sep=";",format="%Y%m%d")
ew <- Rm$EW
OPmdiff <- diff(log(OPm))
```

Running the regression on lagged data

```
> Oil1 = lag(OPmdiff,-1)
> ew1 = lag(ew,-1)
> data <- merge(ew,ew1,Oil1,all=FALSE)
> reg <- lm( data[,1]~data[,2]+data[,3])
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.010903   0.003396   3.210  0.00148 **
data[, 2]    0.321901   0.057939   5.556 6.45e-08 ***
data[, 3]   -0.034944   0.030659  -1.140  0.25536

Residual standard error: 0.05509 on 279 degrees of freedom
Multiple R-squared: 0.09965,Adjusted R-squared: 0.0932
F-statistic: 15.44 on 2 and 279 DF,  p-value: 4.368e-07
```

The coefficient on the change in the oil price the previous is negative, but not significant. (With a slightly longer time series I have actually found a negative and significant relation.) The negative sign is counterintuive. Why is an increase in the oil price bad news for the norwegian stock market?
Check what happens the *same*month, the correlation matrix

```
> cor(data)
          OPmdiff        ew
OPmdiff 1.0000000 0.1876173
ew      0.1876173 1.0000000
```

Lag length selection

Several ways of selecing lags

- Testing of significance using t-tests and F-tests.
- Formal information criteria, for example Akaike's information criterion.

$$AIC = \ln\left(\hat{\sigma}^2\right) + \frac{2k}{T}$$

where k is the lag length.

We look at the link between oil price and the Norwegian stock market. The question is whether the oil price has any predictive content for the Norwegian market.

We ask what the correct lag length is for modelling this, and compare the two model formulations

$$R_{m,t} = \mu + \phi R_{m,t-1} + b\Delta OP_{t-1} + e_t$$

and

$$R_{m,t} = \mu + \phi_1 R_{m,t-1} + \phi_2 R_{m,t-2} + b_1 \Delta OP_{t-1} + b_2 \Delta OP_{t-2} + e_t$$

where R_m is a market index and ΔOP the change in the oil price (in NOK).

Use monthly data prices for Brent oil, and an equally weighted index for the Norwegian stock market.

- A heuristic way of choosing lag length is to start at the max lag length, and see if the last lag is significant, and remove one lag at a time. Do this for this case.
- Another way to do the choice is to do use a formal information criterion. Choose one of these two model formulations using an information criterion.

Reading the data

```
library(zoo)
OPm<- read.zoo("../../../../data/commodities/oil_prices/bre
                format="%Y%m%d",header=FALSE,skip=1)
Rm <- read.zoo("../../../../data/norway/stock_market_indice
                header=TRUE,sep=";",format="%Y%m%d")
ew <- Rm$EW
OPmdiff <- diff(log(OPm))
```

Running the regression on lagged data

```
> Oil1 = lag(OPmdiff,-1)
> ew1 = lag(ew,-1)
> Oil2 = lag(OPmdiff,-2)
> ew2 = lag(ew,-2)
> data <- merge(ew,ew1,Oil1,ew2,Oil2,all=FALSE)
> reg2 <- lm( data[,1]~data[,2]+data[,3]+data[,4]+data[,5])

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.010418   0.003475   2.998  0.00297 **
data[, 2]    0.314773   0.060931   5.166 4.59e-07 ***
data[, 3]   -0.037554   0.030981  -1.212  0.22648
data[, 4]    0.027192   0.061573   0.442  0.65911
data[, 5]    0.014061   0.030891   0.455  0.64934

Residual standard error: 0.0553 on 276 degrees of freedom
Multiple R-squared: 0.1015,Adjusted R-squared: 0.0885
F-statistic: 7.796 on 4 and 276 DF,  p-value: 5.764e-06
```

Then do the one period model

```
> data <- merge(ew,ew1,Oil1,all=FALSE)
> reg1 <- lm( data[,1]~data[,2]+data[,3])

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.010903   0.003396   3.210  0.00148 **
data[, 2]    0.321901   0.057939   5.556 6.45e-08 ***
data[, 3]   -0.034944   0.030659  -1.140  0.25536

Residual standard error: 0.05509 on 279 degrees of freedom
Multiple R-squared: 0.09965,Adjusted R-squared: 0.0932
F-statistic: 15.44 on 2 and 279 DF,  p-value: 4.368e-07
```

The model comparison is by calculating the AIC criterion

```
> AIC(reg1)
[1] -829.6623
> AIC(reg2)
[1] -822.5492
```

The one lag model has lower AIC, it is preferred

Forecasting

What we have seen before in MA/AR/ARMA applies as before, the additional problem will be in forecasting the future values of x_{t+i}.

Vector Auto Regressions

Vector Auto Regressions (VARs) are the natural next step following autoregressive distributed lag models. The important feature that these models add is the possibility of *feedback* between variables, x can influence y, but y can also influence x.

A VAR model of Inflation and Unemployment

For example, if we have two time series variables

$$y_t = \beta_{10} + \beta_{11} y_{t-1} + \beta_{12} y_{t-2} + \cdots + \beta_{1p} y_{t-p} + \gamma_{11} x_{t-1} + \gamma_{12} x_{t-2} + \cdots \gamma_{1p}$$

$$x_t = \beta_{20} + \beta_{21} y_{t-1} + \beta_{22} y_{t-2} + \cdots + \beta_{2p} y_{t-p} + \gamma_{21} x_{t-1} + \gamma_{22} x_{t-2} + \cdots \gamma_{2p}$$

$$\begin{aligned}
\widehat{\Delta Inf_t} &= \underset{(0.47)}{1.32} - \underset{(0.09)}{0.36}\,\Delta Inf_{t-1} - \underset{(0.10)}{0.34}\,\Delta Inf_{t-2} + \underset{(0.08)}{0.07}\,\Delta Inf_{t-3} - \\
&\quad - \underset{(0.47)}{2.68}\,Unemp_{t-1} + \underset{(0.89)}{3.43}\,Unemp_{t-2} - \underset{(0.89)}{1.04}\,Unemp_{t-3} - \\
\widehat{Unemp_t} &= \underset{(0.09)}{0.12} + \underset{(0.02)}{0.043}\,\Delta Inf_{t-1} + \underset{(0.015)}{0.000}\,\Delta Inf_{t-2} + \underset{(0.16)}{0.021}\,\Delta Inf_{t-3} \\
&\quad + \underset{(0.12)}{1.68}\,Unemp_{t-1} + \underset{(0.20)}{0.70}\,Unemp_{t-2} - \underset{(0.20)}{0.03}\,Unemp_{t-3} -
\end{aligned}$$

For the first equation $\bar{R}^2 = 0.35$, and for the second $\bar{R}^2 = 0.975$

Setup in matrix notation

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \Phi_1 \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \Phi_2 \begin{bmatrix} y_{t-2} \\ x_{t-2} \end{bmatrix} + \cdots$$

Using the previous example, we would observe that

$$\Phi_1 = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \gamma_{11} & \gamma_{12} \end{bmatrix}$$

We will often collapse the variables into vectors too,

$$\mathbf{y}_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix}$$

$$\mathbf{y}_t = \mathbf{c} + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \cdots + \Phi_p \mathbf{y}_{t-p} + \varepsilon_t$$

Note that the number of parameters grow very fast.

Lag length selection

We want to construct an information criterion we can apply to the choice of lag length. Remember in the univariate case we use a version of the AIC criterion

$$AIC = \ln\left(\frac{\sum_{t=1}^{T} u_t^2}{T}\right) + \frac{2K}{T}$$

where u_t is the residual, T the number of observations, K the number of parameters.
Here we have a situation where we have several equations, we can not apply such a criterion directly, we need to aggregate somehow

$$y_t = \beta_{10} + \beta_{11} y_{t-1} + \cdots + \beta_{1p} y_{t-p} + \gamma_{11} x_{t-1} + \cdots \gamma_{1p} x_{t-p} + u_{1t}$$

$$x_t = \beta_{20} + \beta_{21} y_{t-1} + \cdots + \beta_{2p} y_{t-p} + \gamma_{21} x_{t-1} + \cdots \gamma_{2p} x_{t-p} + u_{2t}$$

Define

$$\hat{\mathbf{\Sigma}}_u = \begin{bmatrix} \sum_t u_{1t}^2 & \sum_t u_{1t}u_{2t} \\ \sum_t u_{1t}u_{2t} & \sum_t u_{2t}^2 \end{bmatrix}$$

The information criterion is calculated as

$$AIC = \ln\left(\frac{\det \mathbf{\Sigma}_u}{T}\right) + \frac{2K}{T}$$

Granger causality

Most important use of VAR's: Testing for Granger causality. Similar method to before. Test for whether the coefficent on a predictor x are zero or not. If we reject a null of all coefficients on an explanatory variable x being zero, the variable x is said to Granger cause y, if y is the relevant dependent variable.

Caveat

VAR estimation is “model-free” in the sense that we are not estimating an explicit economic model, we are exploring possible linear relationships between economic variables. A prime use of these investigations will be forecasting. It is also useful to generate “stylized facts” that our economic theories should be able to explain. But it should be clear that a VAR or some other time series model does not *explain* any causality or other regularities in the data.

Estimation methodology of VAR's

When we estimate VAR's, it is usually done in vector form:

$$\mathbf{y}_t = \mathbf{c} + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \cdots + \Phi_p \mathbf{y}_{t-p} + \varepsilon_t$$

If y_t is an n-vector, each of Φ_i is an $n \times n$ matrix. This is one of the characteristics of a VAR, the large number of parameters to estimate:

$$\begin{array}{lcl} \mathbf{c} & : & 1 \times n \\ \Phi_1 & : & n \times n \\ & \vdots & \\ \Phi_p & : & n \times n \end{array}$$

or $n + p \cdot (n \times n) = n + pn^2 = n(1 + pn)$ parameters to estimate. So if we use 3 lags and 4 variables, we get $(1 + 3 \cdot 4) \cdot 4 = 13 \cdot 4 = 52$ parameters to estimate.

Estimating VAR's thus involves a lot of parameters, which is one reason for why it is usually estimated by Maximum Likelihood, we need to use the information in the sample as efficient as possible. Another way to make the estimation more feasible is to impose restrictions on some of the parameters, which will make the estimation of the remaining parameters more precise.
VAR's are usually estimated by assuming the error vector ϵ_t is multivariate normal.

$$\epsilon_t \sim \mathcal{N}(\mathbf{0}, \Sigma)$$

Let us go through the steps involved in this estimation. The distribution function for a multivariate normal random variable x is

$$f(x) = \frac{1}{(2\pi)^2} \frac{1}{|\Sigma|^{\frac{n}{2}}} e^{-\frac{1}{2}x'\Sigma^{-1}x}$$

where n is the length of the vector x.
We want to estimate the VAR

$$\mathbf{y}_t = \mathbf{c} + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \cdots + \Phi_p \mathbf{y}_{t-p} + \varepsilon_t$$

with

$$\epsilon_t \sim \mathcal{N}(\mathbf{0}, \Sigma)$$

which implies

$$\mathbf{y}_t \sim \mathcal{N}(\mathbf{c} + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \cdots + \Phi_p \mathbf{y}_{t-p}, \Sigma)$$

We use this to specify the conditional likelihood for one observation.

$$\begin{aligned} & f(\mathbf{y}_t | \mathbf{y}_{t-1} \cdots \mathbf{y}_{t-2}) \\ = \quad & \frac{1}{(2\pi)^2} \frac{1}{|\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\left(\mathbf{y}_t - \mathbf{c} - \Phi_1 \mathbf{y}_{t-1} - \cdots \Phi_p \mathbf{y}_{t-p}\right)' \Sigma^{-1} \left(\mathbf{y}_t - \mathbf{c} - \right.\right. \end{aligned}$$

The conditional likelihood function for a set of $T + p$ observations will be a product of T of these individual likelihood functions. Estimation will consist of maximizing

$$\log \Pi_{t=p}^{T+p} f(y_t | y_{t-1} \cdots \mathbf{y}_{t-2})$$

$$= \log \left(\frac{1}{(2\pi)^2} \right) + \log \left(\frac{1}{|\Sigma|^{\frac{n}{2}}} \right)$$

$$-\frac{1}{2} \sum_{t=p}^{T+p} \left(\mathbf{y}_t - \mathbf{c} - \Phi_1 \mathbf{y}_{t-1} - \cdots \Phi_p \mathbf{y}_{t-p} \right)' \Sigma^{-1} \left(\mathbf{y}_t - \mathbf{c} - \Phi_1 \mathbf{y}_{t-1} - \cdots \right.$$

with respect to the parameters

$$\mathbf{c}, \Phi_1 \cdots \Phi_p, \Sigma$$

Since we are doing Maximum Likelihood, the usual test statistics are available, and are used to test hypotheses. In particular, Granger type tests of causality can be performed by looking at the off–diagonal elements of Φ_i.

Identification issues in VARs

In system of equations we had the important question of identification, is a property of the model, and we either had to impose inclusion restrictions, or use instruments.

The good news with VAR is that as long as all explanatory variables are prior to time t,

$$\mathbf{y}_t = \mathbf{c} + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \cdots + \Phi_p \mathbf{y}_{t-p} + \varepsilon_t$$

there is no identification issue, all parameters are identified.

The problem of identificaiton rears its ugly head again, if we try to impose restrictions between *contemporary* variables, observed at time t, which will generally be expressed as

$$\Phi_0 \mathbf{y}_t = \mathbf{c} + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \cdots + \Phi_p \mathbf{y}_{t-p} + \varepsilon_t$$

where Φ_0 is a constant matrix.

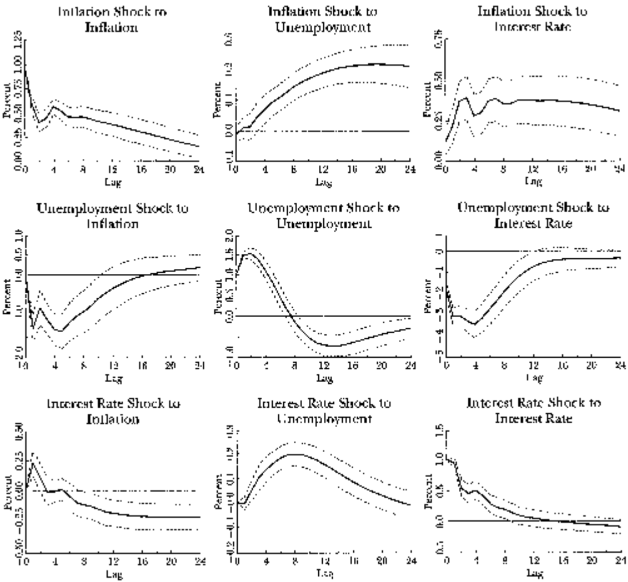
Such a system is in general not identified, and will need to be fixed in the "usual" way, either by theoretical restrictions coming from a model, or extra information.

Impulse response functions

Based on a MA representation of a VAR or a distributed lag model, it is possible to construct forecasts which tells us the time $t + p$ period response to a unit shock at time t. This is called the *impulse response function*.
Consider an example taken from Stock and Watson [2001].

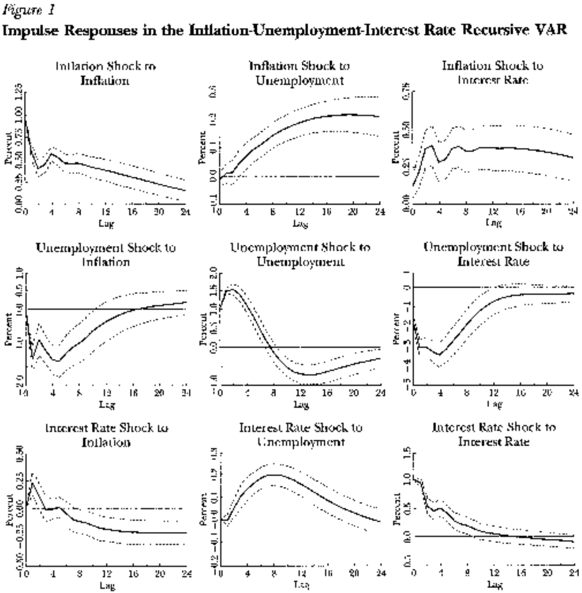
Figure 1

Impulse Responses in the Inflation-Unemployment-Interest Rate Recursive VAR



In their Journal of Economic Perspectives survey of VAR methodology Stock and Watson [2001] model a VAR of three variables: Inflation, Unemployment, and Interest rates.
You want to replicate their picture showing the Impulse Response relationships Using US data 1960–2000, estimate their model as a simple VAR, and plot impulse response diagrams for the model.

The original picture:



Solution

Reading the data

```
library(xts)
library(vars)
                                        # fed funds rate
FFRate <- read.zoo("../data/FRB_H15_fed_funds_daily.csv",
                     skip=6,header=FALSE,na.strings="ND",se
                     format="%Y-%m-%d")
FFRate <- na.locf(FFRate)
qff <- as.xts(FFRate[endpoints(FFRate,on="quarters")])
names(qff) <- "R"
                                        # cpi and then infl
CPI <- read.table("../data/cpi.csv",skip=7,header=TRUE,sep=
cpi <- ts(CPI$Value,frequency=12,start=c(1913,1))
cpi <- as.xts(cpi)
cpi <- cpi[endpoints(cpi,on="quarters")]
names(cpi) <- "CPI"
infl <- 400*diff(log(cpi))
names(infl) <-"Infl"
```

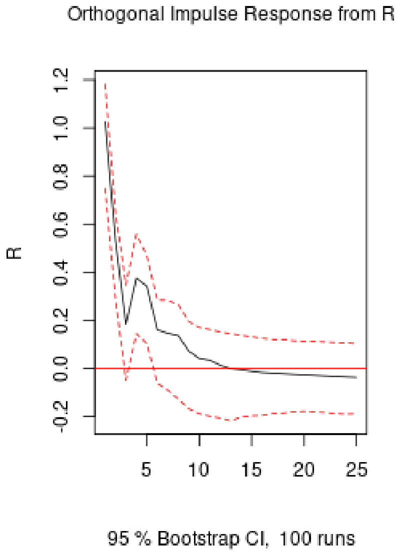
```
# unemployment
Unemployment <- read.table("../data/unemployment.csv",sep='
unemployment <- as.xts(ts(Unemployment$Value,frequency=12,s
unemployment <- na.locf(unemployment)
unemployment <- unemployment[endpoints(unemployment,on="qua
names(unemployment) <- "Unempl"

infl    <- infl["1960/2000"]
unempl <- unemployment["1960/2000"]
ff     <- qff["1960/2000"]
length(infl)
length(unempl)
length(ff)
                                          # can not use merge
data <- cbind(as.matrix(infl),as.matrix(unempl),as.matrix(f
```

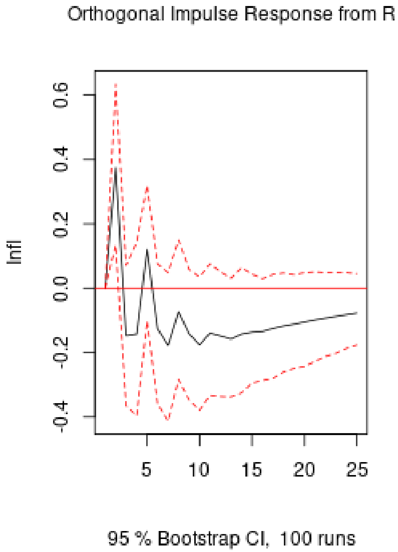
Performing the analysis

```
var <- VAR(data,p=4)
summary(var)
causality(var,cause="R")
causality(var,cause="Infl")
causality(var,cause="Unempl")
ir <- irf(var, impulse="R",response="R",n.ahead=24)
...
```

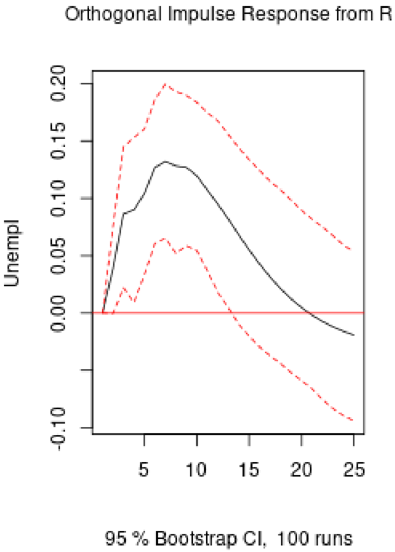
```
ir <- irf(var, impulse="R",response="R",n.ahead=24)
```



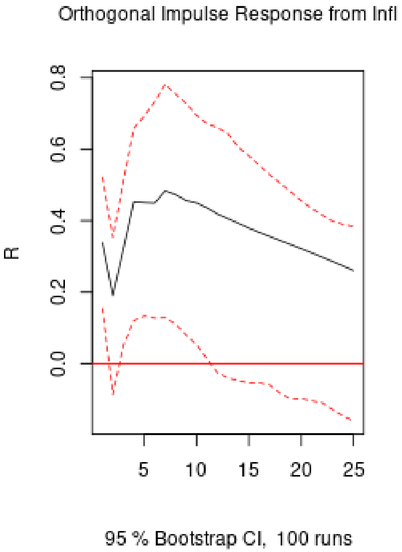
```
ir <- irf(var, impulse="R",response="Infl",n.ahead=24)
```



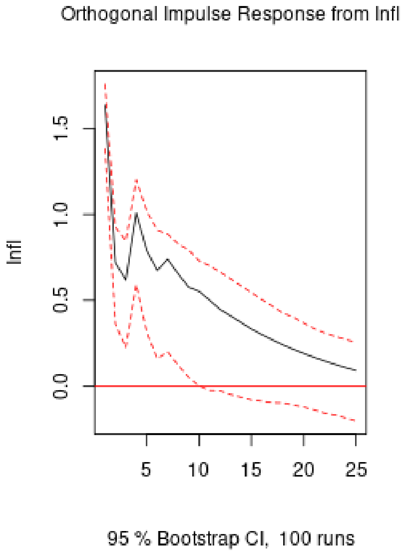
```
ir <- irf(var, impulse="R",response="Unempl",n.ahead=24)
```



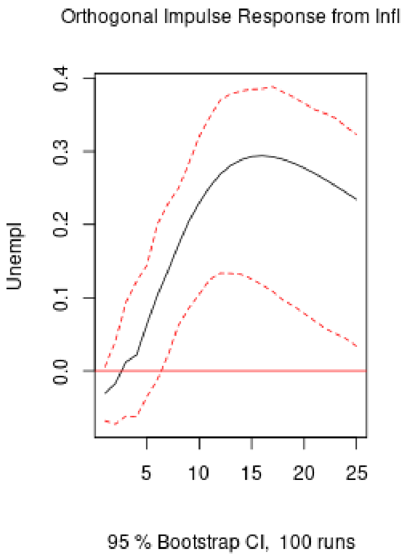
```
ir <- irf(var, impulse="Infl",response="R",n.ahead=24)
```



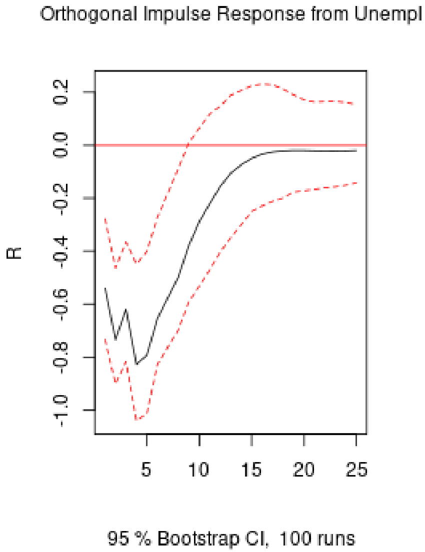
```
ir <- irf(var, impulse="Infl",response="Infl",n.ahead=24)
```



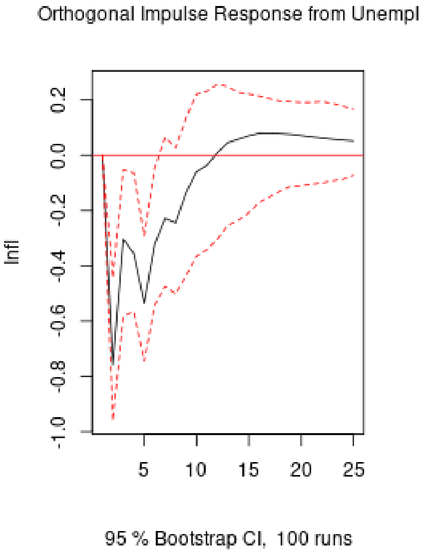
```
ir <- irf(var, impulse="Infl",response="Unempl",n.ahead=24)
```



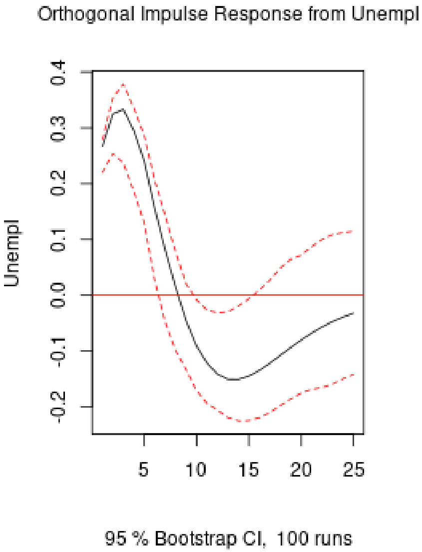
```
ir <- irf(var, impulse="Unempl",response="R",n.ahead=24)
```



```
ir <- irf(var, impulse="Unempl",response="Infl",n.ahead=24)
```



```
ir <- irf(var, impulse="Unempl",response="Unempl",n.ahead=2
```



Exercise: Oil and The Oslo Stock Exchange in a VAR

We look at the link between oil price and the Norwegian stock market. The question is whether the oil price has any predictive content for the Norwegian market. Use monthly data prices for Brent oil, and an equally weighted index for the Norwegian stock market.

Consider modelling this as a VAR.

- Compare model formulations with one or two lags.
- Do oil Granger Cause the Norwegian Stock market?
- Plot the impulse response functions following a shock to oil prices.

Exercise Solution

Reading the data

```
> library(zoo)
> OPm<- read.zoo("../data/brent_monthly.txt",
+                    format="%Y%m%d",header=FALSE,skip=1)
> Rm <- read.zoo("../data/market_portfolio_returns_monthly
+                    header=TRUE,sep=";",format="%Y%m%d")
> ew <- Rm$EW
> data <- merge(OPmdiff,ew,all=FALSE)
```

Note the last command, it takes intersection to align the data.

Exercise Solution ctd

Now, running VAR analysis, need to load library vars

```
> library(vars)
```

Exercise Solution ctd

A one period VAR:

```
> onelag=VAR(data,p=1)
> summary(onelag)

VAR Estimation Results:
=========================
Endogenous variables: OPmdiff, ew
Deterministic variables: const
Sample size: 294
Log Likelihood: 685.569
Roots of the characteristic polynomial:
0.2691 0.0834
Call:
VAR(y = data, p = 1)
```

Exercise Solution ctd

```
Estimation results for equation OPmdiff:
========================================
OPmdiff = OPmdiff.l1 + ew.l1 + const

           Estimate Std. Error t value Pr(>|t|)
OPmdiff.l1 0.037395   0.058993   0.634   0.5267
ew.l1      0.279053   0.111097   2.512   0.0126 *
const      0.001593   0.006420   0.248   0.8042

Residual standard error: 0.1067 on 291 degrees of freedom
Multiple R-Squared: 0.02549,Adjusted R-squared: 0.01879
F-statistic: 3.806 on 2 and 291 DF,  p-value: 0.02335
```

Exercise Solution ctd

```
Estimation results for equation ew:
===================================
ew = OPmdiff.l1 + ew.l1 + const

            Estimate Std. Error t value Pr(>|t|)
OPmdiff.l1 -0.038203   0.030175  -1.266  0.20650
ew.l1       0.315122   0.056825   5.545 6.58e-08 ***
const       0.010161   0.003284   3.094  0.00216 **

Residual standard error: 0.05456 on 291 degrees of freedom
Multiple R-Squared: 0.09569,Adjusted R-squared: 0.08947
F-statistic:  15.4 on 2 and 291 DF,  p-value: 4.411e-07
```

Exercise Solution ctd

```
Covariance matrix of residuals:
          OPmdiff       ew
OPmdiff 0.011376 0.000927
ew      0.000927 0.002976

Correlation matrix of residuals:
         OPmdiff     ew
OPmdiff  1.0000 0.1593
ew       0.1593 1.0000
```

Exercise Solution ctd

Adding one period is merely a case of specifying the number of lags as 2:

```
> twolags=VAR(data,p=2)
> summary(twolags)

VAR Estimation Results:
=========================
Endogenous variables: OPmdiff, ew
Deterministic variables: const
Sample size: 293
Log Likelihood: 684.365
Roots of the characteristic polynomial:
0.3923 0.2218 0.2218 0.195
Call:
VAR(y = data, p = 2)
```

Exercise Solution ctd

```
Estimation results for equation OPmdiff:
========================================
OPmdiff = OPmdiff.l1 + ew.l1 + OPmdiff.l2 + ew.l2 + const

              Estimate Std. Error t value Pr(>|t|)
OPmdiff.l1  0.0272818  0.0593455   0.460    0.646
ew.l1       0.2190597  0.1160100   1.888    0.060 .
OPmdiff.l2 -0.0195816  0.0591992  -0.331    0.741
ew.l2       0.2111329  0.1173484   1.799    0.073 .
const      -0.0006059  0.0065351  -0.093    0.926

Residual standard error: 0.1066 on 288 degrees of freedom
Multiple R-Squared: 0.03635,Adjusted R-squared: 0.02296
F-statistic: 2.716 on 4 and 288 DF,  p-value: 0.03016
```

Exercise Solution ctd

```
Estimation results for equation ew:
===================================
ew = OPmdiff.l1 + ew.l1 + OPmdiff.l2 + ew.l2 + const

           Estimate Std. Error t value Pr(>|t|)
OPmdiff.l1 -0.041439   0.030474  -1.360  0.17495
ew.l1       0.305441   0.059571   5.127  5.4e-07 ***
OPmdiff.l2  0.014343   0.030399   0.472  0.63740
ew.l2       0.037540   0.060258   0.623  0.53379
const       0.009569   0.003356   2.852  0.00467 **

Residual standard error: 0.05473 on 288 degrees of freedom
Multiple R-Squared: 0.0983,Adjusted R-squared: 0.08577
F-statistic: 7.849 on 4 and 288 DF,  p-value: 5.126e-06
```

Exercise Solution ctd

```
Covariance matrix of residuals:
          OPmdiff        ew
OPmdiff 0.0113607 0.0009085
ew      0.0009085 0.0029956

Correlation matrix of residuals:
        OPmdiff     ew
OPmdiff  1.0000 0.1557
ew       0.1557 1.0000
```

Exercise Solution ctd

Causality tests, first for one lag, then for two lags

```
> causality(onelag,cause="OPmdiff")
$Granger
Granger causality H0: OPmdiff do not Granger-cause ew

data:  VAR object onelag
F-Test = 1.6029, df1 = 1, df2 = 582, p-value = 0.206

$Instant
H0: No instantaneous causality between: OPmdiff and ew

data:  VAR object onelag
Chi-squared = 7.2762, df = 1, p-value = 0.006987
```

Can not reject the null of no causality for the one-lag case.

Exercise Solution ctd

```
> causality(twolags,cause="OPmdiff")
$Granger
Granger causality H0: OPmdiff do not Granger-cause ew

data:  VAR object twolags
F-Test = 1.0065, df1 = 2, df2 = 576, p-value = 0.3662

$Instant
H0: No instantaneous causality between: OPmdiff and ew

data:  VAR object twolags
Chi-squared = 6.9384, df = 1, p-value = 0.008437
```

Summary

Autoregressive distributed lag models $ADL(p, q)$:

$$y_t = \mu + \sum_{j=1}^{p} \phi_j y_{t-j} + \sum_{j=1}^{q} b_j x_{t-j} + u_t$$

Granger Causality
Lag length selection

- Hypothesis tests
- Information criteria

Summary

Vector Auto Regressions (VARs) Allow for feedback
Stack several distributed lag models

$$\mathbf{y}_t = \mathbf{c} + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \cdots + \Phi_p \mathbf{y}_{t-p} + \varepsilon_t$$

where for example

$$\mathbf{y}_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix}$$

Lag length selection

- Hypothesis tests
- Information criteria

Granger causality
Impulse response functions:
What is the future effect of a unit shock to one variable?

James H Stock and Mark W Watson. Vector autoregressions. *Journal of Economic Perspectives*, 15(4):101–115, Autumn 2001.