Quantitative Analysis of Finance I

ECON90033

WEEK 10

STRUCTURAL VECTOR AUTOREGRESSION

IMPULSE RESPONSE ANALYSIS

FORECAST ERROR VARIANCE DECOMPOSITION

COINTEGRATION

→ SVAR is for r/ship
↳ VAR for f/cast

<u>Reference</u>:
HMPY: § 4.5-4.6

Dr László Kónya
January 2023

STRUCTURAL VECTOR AUTOREGRESSION (*SVAR*)

start w/ reduced form, then SVAR if possible

- Recall that a structural vector autoregressive (*SVAR*) system is a set of autoregressive structural equations, where the time path of each endogenous variable is determined by its own history and by current and past realizations of the other endogenous variables.

b/c we can observe all RHS

can estimate

- Like structural simultaneous equation models, a *SVAR* model cannot be estimated directly, it must be recovered from the corresponding standard *VAR* model, if possible.
- When is *SVAR* identifiable, i.e., under what conditions can we identify a *SVAR* model from a *VAR* model?
- To keep it simple, consider a pair of two-variable *VAR* and *SVAR* models with a single lag (see slides #6-8 of week 9): → need to estimate ε variance

$$\begin{aligned} y_t &= a_{10} + a_{11} y_{t-1} + a_{12} z_{t-1} + u_{1t} \\ z_t &= a_{20} + a_{21} y_{t-1} + a_{22} z_{t-1} + u_{2t} \end{aligned} \quad \xrightarrow{?} \quad \begin{aligned} y_t &= b_{10} - b_{12} z_t + \gamma_{11} y_{t-1} + \gamma_{12} z_{t-1} + \varepsilon_{yt} \\ z_t &= b_{20} - b_{21} y_t + \gamma_{21} y_{t-1} + \gamma_{22} z_{t-1} + \varepsilon_{zt} \end{aligned}$$

VAR has six regression coefficients, while *SVAR* has eight.

2 intercepts & 4 slopes $\neq$

∴ need assumptions to reduce # of params

On top of these coefficients, the OLS estimation of *VAR* yields estimates of the variances of u_{1t} and u_{2t}, and also of the covariance between them, while in *SVAR* the error terms are uncorrelated, so we must estimate only the variances of ε_{yt} and ε_{zt}.

→ if you "over-identify" you don't get a unique solution

→ All in all, *SVAR* contains 10 unknown parameters, whereas the estimation of *VAR* yields only 9 estimates.

→ *SVAR* is underidentified and without some restriction(s) on the structural parameters it is not possible to recover them from the estimated *VAR* coefficients.

What restriction(s) can be imposed on the structural parameters?

Sometimes it is reasonable to assume that the relationship between y_t and z_t is asymmetric because y_t has a contemporaneous effect on z_t, but z_t does not affect y_t, implying that $b_{12} = 0$.

i.e. assume causality directions & timing of impacts

→ Need to assume the 1st & then use this to plug into the 2nd

$$\begin{aligned} y_t &= b_{10} + \gamma_{11} y_{t-1} + \gamma_{12} z_{t-1} + \varepsilon_{yt} \\ z_t &= b_{20} - b_{21} y_t + \gamma_{21} y_{t-1} + \gamma_{22} z_{t-1} + \varepsilon_{zt} \end{aligned}$$

This *SVAR* is a recursive system, i.e., the endogenous variables are determined one at a time in sequence (y_t first and then z_t).

This recursive *SVAR* in matrix form

$$\mathbf{B}\mathbf{x}_t = \boldsymbol{\Gamma}_0 + \boldsymbol{\Gamma}_1\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t \quad \text{where} \quad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ b_{21} & 1 \end{bmatrix} \longrightarrow \mathbf{B}^{-1} = \begin{bmatrix} 1 & 0 \\ -b_{21} & 1 \end{bmatrix}$$

… and the relationship between the *VAR* error terms ($\mathbf{u}_t$) and the *SVAR* error terms ($\boldsymbol{\varepsilon}_t$) is

$$\mathbf{u}_t = \mathbf{B}^{-1}\boldsymbol{\varepsilon}_t = \begin{bmatrix} 1 & 0 \\ -b_{21} & 1 \end{bmatrix}\begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} = \begin{bmatrix} \varepsilon_{yt} \\ -b_{21}\varepsilon_{yt} + \varepsilon_{zt} \end{bmatrix}$$

How to calculate error terms!!

$\longrightarrow$ Given a recursive *SVAR*, in the corresponding *VAR* only the ε_{yt} (= u_{2t}) shock has a direct effect on y_t, while both the ε_{yt} and ε_{zt} shocks have instantaneous effect on z_t via u_{2t}.

Note that although an ε_{zt} shock has no direct effect on y_t, it does so on y_{t+1}.

The decomposition of the *VAR* error terms (u_{1t}, u_{2t}) in this asymmetrical way is called Cholesky decomposition (named after the French military officer and mathematician, André-Louis Cholesky, 1875-1918).

Due to the $b_{12} = 0$ restriction imposed on the *SVAR* system, there are only 9 unknown parameters left, just like in the corresponding *VAR* system, thus it is possible to obtain estimates of the remaining *SVAR* parameters from the OLS estimates of the *VAR* parameters.

In addition, it is also possible to estimate the $\{\varepsilon_{yt}\}$ and $\{\varepsilon_{zt}\}$ series from the OLS residuals of the *VAR* system (u_{1t}-hat and u_{2t}-hat):

$$\hat{\varepsilon}_{yt} = \hat{u}_{1t} \quad , \quad \hat{\varepsilon}_{zt} = \hat{b}_{21}\hat{\varepsilon}_{yt} + \hat{u}_{2t} = \hat{b}_{21}\hat{u}_{1t} + \hat{u}_{2t}$$

A recursive *SVAR* like this one can be recovered from *VAR*, or it can be estimated directly with the OLS method applied to each equation one at a time in the given order.

Note:

a) The previous Cholesky decomposition assumes that y_t ‘precedes’ z_t in the sense that y_t does have a contemporaneous effect on z_t but z_t does not have a contemporaneous effect on y_t. These assumptions imply an ordering of the variables. Namely, y_t is supposed to be ‘prior’ to z_t.

   This is, however, an arbitrary assumption. Instead of $b_{12} = 0$, one could also assume that $b_{21} = 0$, implying that z_t is ‘prior’ to y_t.

→ Unless there is some theoretical reason in favour of a particular ordering of the variables, in practice it is important to check how robust the results are to alternative ordering.

b) In general, the importance of the ordering depends on the correlation between the error terms in the standard *VAR* system. Namely, the weaker the correlation between $\{u_{1t}\}$ and $\{u_{2t}\}$, the less important the ordering is; and ordering makes no difference at all when $\{u_{1t}\}$ and $\{u_{2t}\}$ are uncorrelated.

However, if the analysis involves several endogenous variables, it is very unlikely that all pair-wise correlations are small, and it would be impractical to experiment with all possible orderings.

c) Recursive *SVAR* models can be too restrictive in terms of economic theory.

An alternative and more general approach is to consider non-recursive *SVAR* specifications based on $m(m-1)/2$ restrictions consistent with economic theory, where m is the number of endogenous variables.

With this theory-based approach it is possible to implement short-run restrictions, long-run restrictions, or sign restrictions.

These restrictions, however, can be just as arbitrary, heroic and incredible as the ones imposed on simultaneous equation systems – the very reason why Sims advocated standard *VAR*s the first place.

A more recent approach is based on statistical identification methods which aim to exploit the information content of the data, like heteroskedasticity of structural shocks or changes in volatility, etc. (*R svars* package).

We do not discuss the identification and estimation of *SVAR* models in this course. We shall just make use of recursive *SVAR*s indirectly for impulse response analysis and forecast error variance decomposition.

IMPULSE RESPONSE ANALYSIS

- The aim of impulse response analysis is to trace the effects of the structural innovations or shocks ($\varepsilon_{.t}$) on the entire time paths of the left-hand side variables.

  Recall that a *VAR*(*p*) system can be written as (slide #18, week 9)

$$\mathbf{x}_t = \mathbf{A}_0 + \mathbf{A}_1\mathbf{x}_{t-1} + \mathbf{A}_2\mathbf{x}_{t-2} + \ldots + \mathbf{A}_p\mathbf{x}_{t-p} + \mathbf{u}_t$$

$$\longrightarrow \quad \mathbf{A}(L)\mathbf{x}_t = \mathbf{A}_0 + \mathbf{u}_t$$ where $$\mathbf{A}(L) = \mathbf{I} - \mathbf{A}_1 L - \mathbf{A}_2 L^2 - \ldots - \mathbf{A}_p L^p$$

  If the **A**(*L*) matrix polynomial has all its roots outside the unit circle, *VAR*(*p*) is stable,

$$\mathbf{A}^{-1}(L) = \left[\mathbf{I} - \mathbf{A}_1 L\right]^{-1} = \sum_{i=0}^{\infty}\left(\mathbf{A}_1 L\right)^i$$

  and the *VAR* system has a *VMA* representation:

$$\mathbf{x}_t = \mathbf{A}^{-1}(L)\times\left(\mathbf{A}_0 + \mathbf{u}_t\right) = \sum_{i=0}^{\infty}\left(\mathbf{A}_1 L\right)^i \times\left(\mathbf{A}_0 + \mathbf{u}_t\right)$$

<u>Note</u>: The ‘original’ *VAR* is mainly used to generate forecasts, while the *VMA* representation is used to calculate the forecast errors and to study the dynamic properties of the system.

→ VAR = forecast
→ VMA = forecast errors & study dyn. properties.

Consider, for example, a stable two-variable *VAR*(1) system:

$$y_t = a_{10} + a_{11} y_{t-1} + a_{12} z_{t-1} + u_{1t}$$
$$z_t = a_{20} + a_{21} y_{t-1} + a_{22} z_{t-1} + u_{2t}$$

$$\longrightarrow \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \underbrace{\sum_{i=0}^{\infty} (\mathbf{A}_1 L)^i \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix}}_{\boldsymbol{\mu}} + \sum_{i=0}^{\infty} (\mathbf{A}_1 L)^i \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

with $\boldsymbol{\mu} = \begin{bmatrix} \mu_y \\ \mu_z \end{bmatrix}$

and

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \frac{1}{1 - b_{12} b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \mu_y \\ \mu_z \end{bmatrix} + \frac{1}{1-b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{y,t-i} \\ \varepsilon_{z,t-i} \end{bmatrix}$$

Introducing

$$\boldsymbol{\varphi}_i = \mathbf{A}_1^i \mathbf{B}^{-1} = \frac{1}{1-b_{12}b_{21}} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} = \begin{bmatrix} \varphi_{11}(i) & \varphi_{12}(i) \\ \varphi_{21}(i) & \varphi_{22}(i) \end{bmatrix}$$

we get the final form of the *VMA*(∞) representation of the *VAR*(1) system:

$$\mathbf{x}_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \mu_y \\ \mu_z \end{bmatrix} + \sum_{i=0}^{\infty} \underbrace{\begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix}}_{\boldsymbol{\varphi}_i} \begin{bmatrix} \varepsilon_{y,t-i} \\ \varepsilon_{z,t-i} \end{bmatrix} = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \boldsymbol{\varphi}_i \, \boldsymbol{\varepsilon}_{t-i}$$

The elements of the $\boldsymbol{\varphi}_i$ (i = 0, 1, 2, ...) matrices measure the effects of the ε_{yt} and ε_{zt} shocks on current and future values of the $\{y_t\}$ and $\{z_t\}$ sequences,

and the $\{\varphi_{11}(i)\}$, $\{\varphi_{12}(i)\}$, $\{\varphi_{21}(i)\}$, $\{\varphi_{22}(i)\}$ (i = 0, 1, 2, ...) sets of coefficients are called impulse response functions.

The first four coefficients, i.e., the elements of $\boldsymbol{\varphi}_0 = \mathbf{B}^{-1}$, are called impact multipliers; and the accumulated effects, i.e., intermediate multipliers, are

$\varepsilon_{yt} \rightarrow \{y_{t+i}\}$	$\varepsilon_{zt} \rightarrow \{y_{t+i}\}$	$\varepsilon_{yt} \rightarrow \{z_{t+i}\}$	$\varepsilon_{zt} \rightarrow \{z_{t+i}\}$
$\sum_{i=0}^{n} \varphi_{11}(i)$	$\sum_{i=0}^{n} \varphi_{12}(i)$	$\sum_{i=0}^{n} \varphi_{21}(i)$	$\sum_{i=0}^{n} \varphi_{22}(i)$

Letting n approach infinity, we obtain the long-run multipliers, i.e., the overall impacts of one-unit changes in ε_{yt} and ε_{zt} on the $\{y_t\}$, $\{z_t\}$ sequences.

Given that $\{y_t\}$, $\{z_t\}$ are stationary, these infinite sums are convergent, and the long-run multipliers are finite.

Ex 1: Consider the following $VAR(1)$ system

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \underbrace{\begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.1 \end{bmatrix}}_{\mathbf{A}_1} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \longrightarrow \mathbf{I} - \mathbf{A}_1 L = \begin{bmatrix} 1-0.8L & -0.2L \\ -0.4L & 1-0.1L \end{bmatrix}$$

↳ Does the system have an inverse?

a) Determine whether the $\mathbf{x}_t = [y_t, z_t]'$ vector process is stationary.

det = 10/9 > 1
∴ Stable VAR system

The inverse characteristic equation is

$$|\mathbf{I} - \mathbf{A}_1 L| = (1 - 0.8L)(1 - 0.1L) - (-0.02L)(-0.4L) = 1 - 0.9L = 0$$

$\longrightarrow$ There is a single root, L = 10/9. It is outside the unit circle, so this *VAR*(1) system is stable and the $\{y_t\}$, $\{z_t\}$ processes are stationary.

b) Derive the impulse response functions assuming that $u_{1t} = \varepsilon_{yt} + 0.5\,\varepsilon_{zt}$ and $u_{2t} = \varepsilon_{zt}$.

$$\mathbf{u}_t = \mathbf{B}^{-1}\boldsymbol{\varepsilon}_t = \underbrace{\begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}}\begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \longrightarrow \boldsymbol{\varphi}_i = \mathbf{A}_1^i \mathbf{B}^{-1} = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.1 \end{bmatrix}^i \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}$$

Note process in a 1-pager

Therefore,

$$\boldsymbol{\varphi}_1 = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.1 \end{bmatrix}\begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.8 & 0.6 \\ 0.4 & 0.3 \end{bmatrix}$$

$$\boldsymbol{\varphi}_2 = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.1 \end{bmatrix}^2 \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.72 & 0.18 \\ 0.36 & 0.09 \end{bmatrix} \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.72 & 0.54 \\ 0.36 & 0.27 \end{bmatrix}$$

$$\boldsymbol{\varphi}_3 = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.1 \end{bmatrix}^3 \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.648 & 0.162 \\ 0.324 & 0.081 \end{bmatrix} \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.648 & 0.486 \\ 0.324 & 0.243 \end{bmatrix}$$

$\longrightarrow$ $\varphi_{11}(i) : 0.8, 0.72, 0.648, \ldots$ $\longrightarrow$ $\varphi_{11}(i) = (0.8)(0.9)^{i-1} \underset{i=1,2,\ldots}{\longrightarrow} 0$

$\varphi_{12}(i) : 0.6, 0.54, 0.486, \ldots$ $\longrightarrow$ $\varphi_{12}(i) = (0.6)(0.9)^{i-1} \underset{i=1,2,\ldots}{\longrightarrow} 0$

Similarly,

$\varphi_{21}(i) = (0.4)(0.9)^{i-1} \underset{i=1,2,\ldots}{\longrightarrow} 0$ $\qquad$ $\varphi_{22}(i) = (0.3)(0.9)^{i-1} \underset{i=1,2,\ldots}{\longrightarrow} 0$

$\rightarrow$ Note values decay, $\therefore$ stable

Ex 1:

In Ex 1 of the week 9 lectures, we estimated a *VAR*(4) model of the first difference of the unemployment rate (*DUNR*), the first difference of the long-term interest rate (*DLIR*) and the rate of inflation (*INF*) in Australia using quarterly data from 1969 Q4 to 2022 Q3.

a) Obtain the impulse response functions for 12 quarters ahead using the Cholesky decomposition and assuming that contemporaneously only shocks to *DLIR* can shift *DLIR*, only shocks to *DLIR* and *DUNR* can shift *DUNR*, but all shocks can affect *INF*.

$\longrightarrow$ *DLIR* is 'prior' to *DUNR*, and *DUNR* is 'prior' to *INF*.

This is a different ordering than the one we used when we estimated our *VAR*. Namely,

$$DLIR \rightarrow DUNR \rightarrow INF$$

Ordering has no impact on *VAR*, but it does on *SVAR*, impulse responses and variance decomposition (see later).

```
data_v2 = ts(data.frame(DLIR, DUNR, INF), frequency = 4,
             start = c(1969,4), end = c(2022,3))
var4_v2 = VAR(data_v2, p = 4, type = "both")
```

Estimated impulse responses with confidence interval are provided by the *irf()* and *cirf()* functions.

cum. impact over N periods

irf.DLIR_all = irf(var4_v2, impulse = "DLIR", response = c("DLIR","DUNR", "INF"), n.ahead = 12, boot = TRUE, ci = 0.95)
plot(irf.DLIR_all, main = "Responses to shock to DLIR")

cirf.DLIR_all = irf(var4_v2, impulse = "DLIR", response = c("DLIR","DUNR", "INF"), n.ahead = 12, boot = TRUE, ci = 0.95, cumulative = TRUE)
plot(cirf.DLIR_all, main = "Cumulative responses to shock to DLIR")

What's impact in N periods time

→ error term in eqn of DLIR has been increased by 1 unit.

These commands return two plots.

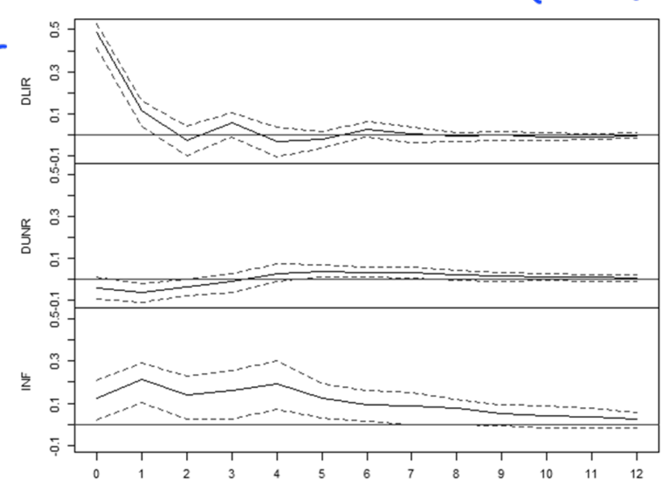
impact on system eqn

↑ 1 std dev of the error term!

The first shows the impact multipliers of the three endogenous variables following a one standard deviation shock to *DLIR* 1-12 quarters later, and the corresponding approximate 95% confidence intervals.

Apparently, all three variables have some significant immediate and delayed responds to shocks to *DLIR*.

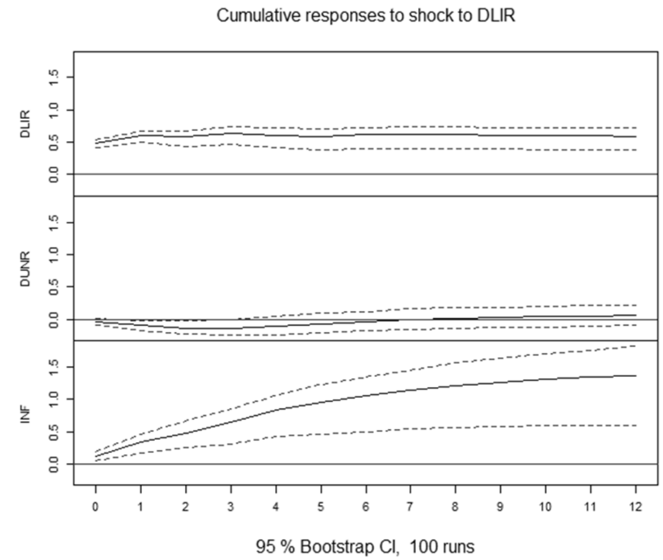
Responses to shock to DLIR



95 % Bootstrap CI, 100 runs

If the bands cross 0 then the point est. is insignif.

The second plot shows the intermediate multipliers of the three endogenous variables following a one standard deviation shock to *DLIR* 1-12 quarters later, and the corresponding approximate 95% confidence intervals.

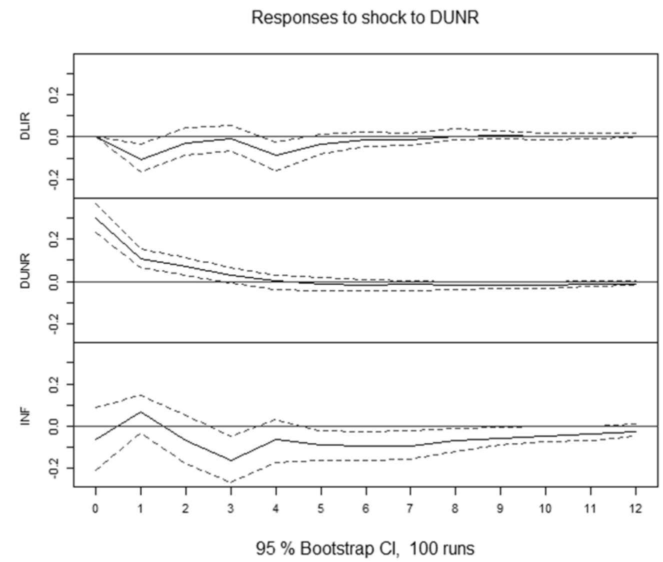


As we saw, this *VAR* is stable, and in the case of stable *VAR*s the responses to shocks are either insignificant, or all die out to zero very quickly.

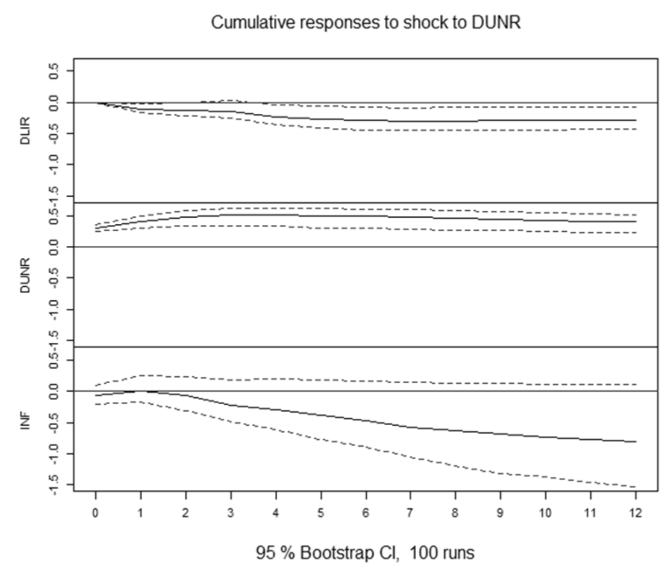
Accordingly, the intermediate multipliers (i.e., the cumulative responses) stabilize within a few quarters.

The impact and intermediate multipliers for a shock to *DUNR* and to *INF* can be obtained similarly.

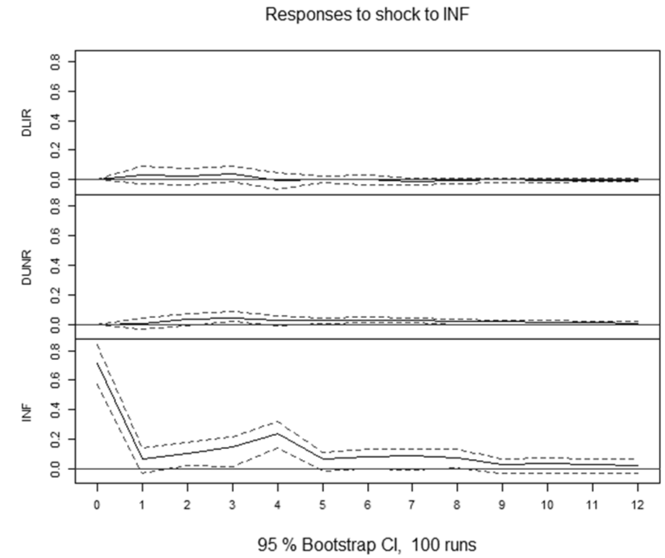
```
irf.DUNR_all = irf(var4_v2, impulse = "DUNR",
              response = c("DUNR", "DLIR", "INF"),
              n.ahead = 12, boot = TRUE, ci = 0.95)
plot(irf.DUNR_all,
     main = "Responses to shock to DUNR")
```



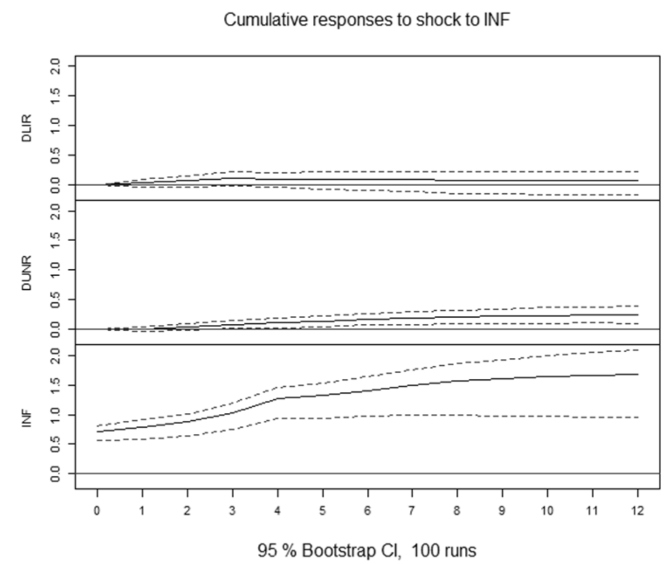

```
cirf.DUNR.all_v2 = irf(var4_v2, impulse = "DUNR",
              response = c("DUNR", "DLIR", "INF"),
              n.ahead = 12, boot = TRUE, ci = 0.95,
              cumulative = TRUE)
plot(cirf.DUNR_all_v2,
     main = "Cumulative responses to shock to DUNR")
```



```
irf.INF_all = irf(var4_v2, impulse = "INF",
        response = c("DUNR", "DLIR", "INF"),
        n.ahead = 12, boot = TRUE, ci = 0.95)
plot(irf.INF_all,
     main = "Responses to shock to INF")
```




```
cirf.INF_all = irf(var4_v2, impulse = "INF",
          response = c("DUNR", "DLIR", "INF"),
          n.ahead = 12, boot = TRUE, ci = 0.95,
          cumulative = TRUE)
plot(cirf.INF_all,
     main = "Cumulative responses to shock to INF")
```



FORECAST ERROR VARIANCE DECOMPOSITION

- Forecast error variance decomposition is a potentially useful tool for calculating the proportions of the movements in a sequence due to its ‘own’ shocks versus to shocks to other variables.

  Recall the *VMA*(∞) representation of the two-variable *VAR*(1) model (see slides #8-10):

$$\mathbf{x}_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \mu_y \\ \mu_z \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \varphi_{11}(i) & \varphi_{12}(i) \\ \varphi_{21}(i) & \varphi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{y,t-i} \\ \varepsilon_{z,t-i} \end{bmatrix} = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \boldsymbol{\varphi}_i \, \boldsymbol{\varepsilon}_{t-i}$$

$$\mathbf{x}_{t+h} = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \boldsymbol{\varphi}_i \boldsymbol{\varepsilon}_{t+h-i} = \boldsymbol{\mu} + \sum_{i=0}^{h-1} \boldsymbol{\varphi}_i \boldsymbol{\varepsilon}_{t+h-i} + \sum_{i=h}^{\infty} \boldsymbol{\varphi}_i \boldsymbol{\varepsilon}_{t+h-i}$$

$$E_t(\mathbf{x}_{t+h}) = \boldsymbol{\mu} + \sum_{i=h}^{\infty} \boldsymbol{\varphi}_i \boldsymbol{\varepsilon}_{t+h-i}$$

⟶ The *h*-period ahead forecast error is

$$\mathbf{x}_{t+h} - E_t(\mathbf{x}_{t+h}) = \sum_{i=0}^{h-1} \boldsymbol{\varphi}_i \boldsymbol{\varepsilon}_{t+h-i}$$

The first element of this vector, i.e., the h-step ahead forecast error for y_t, is

$$y_{t+h} - E_t(y_{t+h}) = \sum_{i=0}^{h-1} (\varphi_{11}(i)\varepsilon_{y,t+h-i} + \varphi_{12}(i)\varepsilon_{z,t+h-i})$$

Denoting the variance of this forecast error as $\sigma_y^2(h)$ and the constant variances of Y and Z as σ_y^2 and σ_z^2, we obtain

$$\sigma_y^2(h) = \sigma_y^2 \sum_{i=0}^{h-1} \varphi_{11}^2(i) + \sigma_z^2 \sum_{i=0}^{h-1} \varphi_{12}^2(i)$$

$\longrightarrow$ Since on the right side every term is non-negative, $\sigma_y^2(h)$ is a non-decreasing function of h.

$$\longrightarrow \quad 1 = \underbrace{\frac{\sigma_y^2}{\sigma_y^2(h)} \sum_{i=0}^{h-1} \varphi_{11}^2(i)} + \underbrace{\frac{\sigma_z^2}{\sigma_y^2(h)} \sum_{i=0}^{h-1} \varphi_{12}^2(i)}$$

The proportion of $\sigma_y^2(h)$ that is due to shocks in the $\{\varepsilon_{yt}\}$ sequence.

The proportion of $\sigma_y^2(h)$ that is due to shocks in the $\{\varepsilon_{zt}\}$ sequence.

Similar decomposition exists for the $\sigma_z^2(h)$ forecast error variance too.

If Y is an exogenous variable within this system, implying that Z does not Granger cause it, then shocks in $\{\varepsilon_{zt}\}$ do not have any impact on the σ_y^2 forecast error variance, so the second term on the right side of the previous formula is zero and the first is equal to one. In the other extreme, if the first term is zero and thus the second is one for all h, the $\{\varepsilon_{zt}\}$ shocks account for the whole forecast error variance of $\{y_t\}$, so the $\{y_t\}$ sequence is entirely endogenous.

Note:

a) In practice, the forecast error variance of a variable is typically explained by its 'own' shocks at short horizons and by shocks to other variables at longer horizons.

b) Just like impulse response function analysis, variance decomposition requires the estimation of matrix **B**, and thus the restriction of some of its elements. We can again rely on the Cholesky decomposition and attribute all the one-period ahead forecast error variance of y_t to ε_{yt}, or all the one-period ahead forecast error variance of z_t to ε_{zt}.

c) It is recommended to study the variance decompositions at various forecast horizons. If the *VAR* system is stable, as h is getting bigger, the variance decompositions should approach some constants.

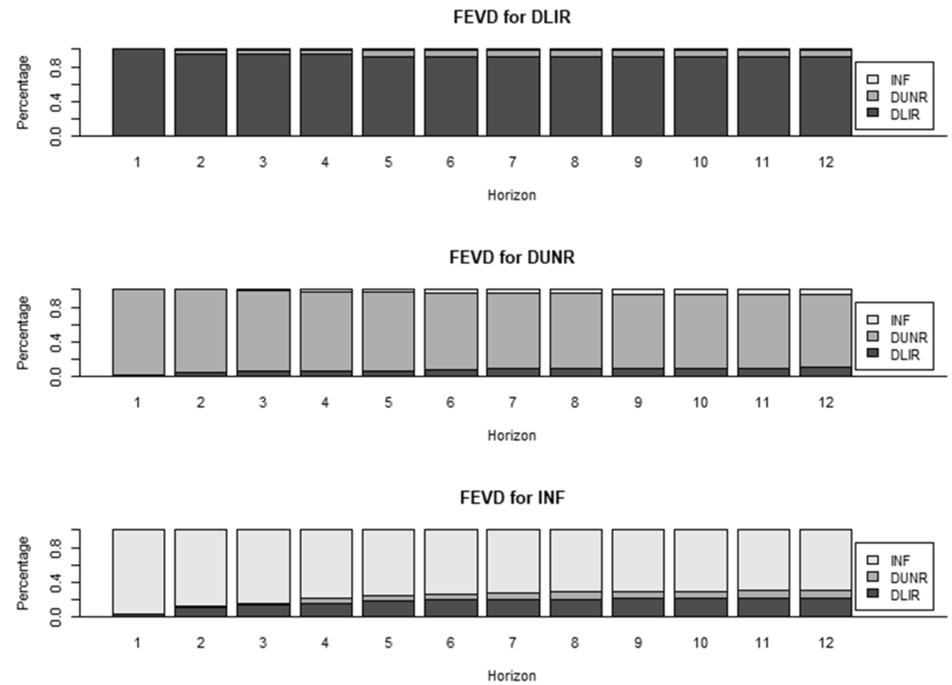
(Ex 1)

g) Obtain forecast error variance decompositions for 1-12 quatres ahead using Cholesky decomposition with the same ordering as in part (a), i.e.

$$DLIR \rightarrow DUNR \rightarrow INF$$

After having estimated the *VAR*(4) model, we can use the *fevd()* function of the *vars* package.

fevd.var4_v2 = fevd(var4_v2, n.ahead = 12)
plot(fevd.var4_v2)



In the first quarter, the forecast error variances of all variables are mainly driven by their own shocks.

In the long run, each sequence seems to converge, and the forecast error variances are mainly (about 90%, 85%, 70%) driven by own shocks.

It was mentioned earlier that the importance of the ordering in the Cholesky decomposition, and hence also in impulse response analysis and forecast error variance decomposition, depends on the correlations between the error terms in the standard *VAR* system.

These correlations are shown in the third part of the *VAR* printout:

```
Correlation matrix of residuals:
          DLIR      DUNR       INF
DLIR    1.0000   -0.1260    0.1683
DUNR   -0.1260    1.0000   -0.1087
INF     0.1683   -0.1087    1.0000
```

None of these correlation coefficients is large enough (in absolute value) to worry about ordering of the variables this time.

→ Note importance of correlations!!

COINTEGRATION

- Traditionally all non-stationary variables used in regression analysis were to be differenced in order to remove the stochastic trends from the data series.

  ← Otherwise, regressions could be unbalanced and spurious, rendering the standard estimation and inferential techniques likely misleading.

  In a multivariate context, however, it is quite possible that integrated variables have a stationary linear combination, a so-called cointegrated relationship, that makes differencing unnecessary.

  The idea of cointegration is due to Granger (1981, 1986), and it was further developed by Engle and Granger (1987), Stock (1987) and Johansen (1988).

  This concept and the related methodology became popular very quickly and nowadays they are essential tools in time-series econometrics.

  ← The estimated cointegrating relationships can often be interpreted as equilibrium relations.

Ex 2:

To illustrate the concept of cointegration, let us consider

a) A simple model of money demand.

Economic theory suggests that

- Individuals want to hold money balances of a certain *real* value, so the demand for nominal money holdings is proportional to the price level.
- As real income and the number of transactions increase, individuals want to hold larger real money balances.
- Since interest rate is the opportunity cost of holding money, it is negatively related to money demand.

These assumptions suggest the following model:

$$\ln M_t = \beta_0 + \beta_1 \ln P_t + \beta_2 \ln Y_t + \beta_3 r_t + \varepsilon_t$$

where M_t: long-run money demand; P_t: price level;
Y_t: real total income; r_t: interest rate.

The first assumption implies that $\beta_1 = 1$, while upon the second and third assumptions it is expected that $\beta_2 > 1$ and $\beta_3 < 0$.

As usual, the error term, ε_t, is assumed to be a white noise.

⟵ If the theory is correct, any deviation from the mean, i.e., from the long-run equilibrium, must be temporary in nature.
By contrast, if ε_t were a random walk, it would have a stochastic trend and there would be no tendency for ln M_t to return to E (ln M_t).

→ If correct

⟶ $$\boxed{\varepsilon_t = \ln M_t - \beta_0 - \beta_1 \ln P_t - \beta_2 \ln Y_t - \beta_3 r_t \quad : I(0)}$$

ε should be stationary!!

Yet, similarly to many other macroeconomic variables, lnM_t, lnP_t, ln Y_t, and maybe even r_t, are most likely $I(1)$.

⟶ Since an $I(0)$ variable cannot be equal to an $I(1)$ variable (unbalanced equation), ε_t : $I(0)$ if and only if the linear combination of lnM_t, lnP_t, ln Y_t and r_t, defined by the right side of this equality, is stationary.

In other words, economic theory suggests that the time paths of the four most likely non-stationary variables must be linked.

This simple model of money demand is just one example, but it illustrates that equilibrium theories involving variables that are likely random walks require that these variables have a stationary linear combination.

all of this stochastic trends neutralise eachother

b) As a second example, consider a small open economy which produces a single good and can borrow and lend in international markets.

$$\longrightarrow \boxed{X_t - M_t = Y_t - C_t - I_t = -B_t + (1+r_t)B_{t-1}}$$

where X_t : exports; M_t : imports;
Y_t : national output; C_t : consumption; I_t : investment;
B_t : net international borrowing (borrowing – lending);
r_t : world interest rate, assumed to be stationary with mean r.

$$\boxed{\underbrace{M_t + (r_t - r)B_{t-1}}_{Z_t} + (1+r)B_{t-1} = X_t + B_t}$$

It can be shown (e.g., Kónya, 2009, pp. 370-373) that

$$\boxed{Z_t + rB_{t-1} = X_t + \sum_{j=1}^{\infty} \lambda^j (\Delta X_{t+j} - \Delta Z_{t+j}) + r \lim_{n\to\infty} \lambda^{n+1} B_{t+n} \quad ; \quad \lambda = \frac{1}{1+r}}$$

Suppose that $\{X_t\}$ and $\{Z_t\}$ are random walks with drift, i.e.

$$\boxed{X_t = \alpha_1 + X_{t-1} + \varepsilon_{1t}} \quad \text{and} \quad \boxed{Z_t = \alpha_2 + Z_{t-1} + \varepsilon_{2t}}$$

$$\longrightarrow \boxed{Z_t + rB_{t-1} = X_t + \sum_{j=1}^{\infty} \lambda^j (\alpha_1 - \alpha_2 + \varepsilon_{1,t+j} - \varepsilon_{2,t+j}) + r \lim_{n \to \infty} \lambda^{n+1} B_{t+n}}$$

… and, since $0 < \lambda < 1$,

$$\underbrace{Z_t + rB_{t-1}}_{\substack{MM_t \\ \text{Imports plus interest} \\ \text{payments on net debt}}} = X_t + \underbrace{\frac{\alpha_1 - \alpha_2}{1 - \lambda}}_{-\alpha} + \underbrace{\sum_{j=1}^{\infty} \lambda^j (\varepsilon_{1,t+j} - \varepsilon_{2,t+j})}_{\substack{-\varepsilon_t \\ \text{Stationary } MA(\infty) \\ \text{error term}}} + \underbrace{r \lim_{n \to \infty} \lambda^{n+1} B_{t+n}}_{0}$$

$$\longrightarrow \boxed{X_t = \alpha + MM_t + \varepsilon_t}$$

$\longrightarrow$ $\{X_t\}$ and $\{MM_t\}$ are $I(1)$ but have a stationary linear combination.

WHAT SHOULD YOU KNOW?

- Structural vector autoregressive (*SVAR*) models
- Cholesky decomposition and ordering of the variables
- Impulse response analysis
- Forecast error variance decomposition
- Cointegration

BOARD OF FAME

André-Louis Cholesky (1875-1918):

French military officer, geodesist, and mathematician

Linear algebra, Cholesky decomposition