Introduction into Vector models for multivariate problems

Motivation: "...Economists quickly discovered that one or two variables are not enough..."

VAR(1) (Vector autoregression model in so called reduced form)

$$\mathbf{y}_t = \phi_0 + \mathbf{\Phi}\, \underbrace{\mathbf{y}_{t-1}}_{lag\ 1} + \varepsilon_t$$

matrix of dimension m

$$
\begin{pmatrix} y_{1,t} \\ y_{2,t} \\ . \\ . \\ \dots \\ y_{m,t} \end{pmatrix}
=
\begin{pmatrix} \phi_{10} \\ \phi_{20} \\ . \\ . \\ . \\ \phi_{m0} \end{pmatrix}
+
\begin{pmatrix} \Phi_{11} & \Phi_{12} & \cdot & \Phi_{1m} \\ \Phi_{21} & \Phi_{22} & \cdot & \Phi_{2m} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ \Phi_{m1} & \Phi_{m2} & \cdot & \Phi_{mm} \end{pmatrix}
\begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ . \\ . \\ . \\ y_{m,t-1} \end{pmatrix}
+
\begin{pmatrix} \varepsilon_{1,t} \\ \epsilon_{2,t} \\ . \\ . \\ . \\ \varepsilon_{m,t} \end{pmatrix}
$$

$$\boldsymbol{\varepsilon}_t \equiv \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_{m,t} \end{pmatrix}, \qquad \boldsymbol{\varepsilon}_s^{\mathrm{T}} \equiv (\varepsilon_{1,s}, \varepsilon_{2,s}, \cdot, \cdot, \cdot, \varepsilon_{m,s})$$

$$\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s^T \equiv \begin{pmatrix} \varepsilon_{1,t}\varepsilon_{1,s} & \varepsilon_{1,t}\varepsilon_{2,s} & \cdot & \cdot & \epsilon_{1,t}\epsilon_{m,s} \\ \varepsilon_{2,t}\epsilon_{1,s} & \epsilon_{2,t}\epsilon_{2,s} & \cdot & \cdot & \varepsilon_{2,t}\epsilon_{m,s} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \varepsilon_{m,t}\epsilon_{1,s} & \epsilon_{m,t}\epsilon_{2,s} & \cdot & \cdot & \varepsilon_{m,t}\varepsilon_{m,s} \end{pmatrix}$$

$$\mathsf{E}[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s^T] \equiv \begin{pmatrix} \mathsf{E}[\varepsilon_{1,t}\varepsilon_{1,s}] & \mathsf{E}[\varepsilon_{1,t}\varepsilon_{2,s}] & \cdot & \cdot & \mathsf{E}[\varepsilon_{1,t}\varepsilon_{m,s}] \\ \mathsf{E}[\varepsilon_{2,t}\varepsilon_{1,s}] & \mathsf{E}[\varepsilon_{2,t}\varepsilon_{2,s}] & \cdot & \cdot & \mathsf{E}[\varepsilon_{2,t}\varepsilon_{m,s}] \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \mathsf{E}[\varepsilon_{m,t}\varepsilon_{1,s}] & \mathsf{E}[\varepsilon_{m,t}\varepsilon_{2,s}] & \cdot & \cdot & \mathsf{E}[\varepsilon_{m,t}\varepsilon_{m,s}] \end{pmatrix}$$

Assumption: matrix $\mathbf{\Sigma}$ is positively definite (constant in time)

$$\mathrm{E}[\varepsilon_t \varepsilon_s^T] \equiv \delta_{ts} \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \cdot & \cdot & \Sigma_{1m} \\ \Sigma_{21} & \Sigma_{22} & \cdot & \cdot & \Sigma_{2m} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \Sigma_{m1} & \Sigma_{m2} & \cdot & \cdot & \Sigma_{mm} \end{pmatrix} = \delta_{ts} \mathbf{\Sigma}$$

Kronecker symbol - white noise

$$\delta_{ts} = \begin{cases} 1 & t = s \\ 0 & t \neq s \end{cases}$$

Model $VAR(1)$ variants: case $m = 2$, diagonal $\mathbf{\Sigma}$:

$$y_{1,t} = \phi_{10} + \Phi_{11} y_{1,t-1} + \Phi_{12} y_{2,t-1} + \varepsilon_{1,t}$$
$$y_{2,t} = \phi_{20} + \Phi_{21} y_{1,t-1} + \Phi_{22} y_{2,t-1} + \varepsilon_{2,t}$$

- $\Phi_{12} = \Phi_{21} = 0$, $\{y_{1,t}\}$ and $\{y_{2,t}\}$ are not correlated

$$y_{1,t} = \phi_{10} + \Phi_{11} y_{1,t-1} + \varepsilon_{1,t}$$
$$y_{2,t} = \phi_{20} + \Phi_{22} y_{2,t-1} + \varepsilon_{2,t}$$

- $\Phi_{12} = 0$, $\{y_{1,t}\}$ and $\{y_{2,t}\}$ there is unidirectional dependence

$$\begin{aligned} y_{1,t} &= \phi_{10} + \Phi_{11} y_{1,t-1} + \varepsilon_{1,t} \\ y_{2,t} &= \phi_{20} + \Phi_{21} y_{1,t-1} + \Phi_{22} y_{2,t-1} + \varepsilon_{2,t} \end{aligned}$$

Cholesky decomposition Hermitian, positive-definite matrix into the product of a lower triangular matrix ($\mathbf{L}$) and its conjugate transpose ($\mathbf{L}^T$ as it follows

$$\boldsymbol{\Sigma} = \mathbf{LDL}^T \quad = \begin{pmatrix} \boxed{1} k & 0 & 0 & \cdot & 0 \\ L_{21} & \boxed{1} & 0 & \cdot & 0 \\ L_{31} & L_{32} & \boxed{1} & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & 0 \\ L_{m1} & L_{m2} & L_{m3} & \cdot & \boxed{1} \end{pmatrix}$$

$$\times \begin{pmatrix} D_1 & 0 & 0 & \cdot & 0 \\ 0 & D_2 & 0 & \cdot & 0 \\ 0 & 0 & D_3 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & D_m \end{pmatrix}$$

$$\times \begin{pmatrix} \boxed{1} & L_{21} & L_{31} & \cdot & L_{m1} \\ 0 & \boxed{1} & L_{32} & \cdot & L_{m2} \\ 0 & 0 & \boxed{1} & \cdot & L_{m3} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \boxed{1} \end{pmatrix}$$

Structural form of VAR(1) [SVAR(1)]

$$
\begin{aligned}
\mathbf{L}^{-1}\mathbf{y}_t &= \mathbf{L}^{-1}\phi_0 + \mathbf{L}^{-1}\boldsymbol{\Phi}\mathbf{y}_{t-1} + \mathbf{L}^{-1}\varepsilon_t \\
\mathbf{L}^{-1}\mathbf{y}_t &= \mathbf{L}^{-1}\phi_0 + \mathbf{DL}\mathbf{y}_{t-1} + \mathbf{L}^{-1}\varepsilon_t \\
\mathbf{L}^{-1}\mathbf{y}_t &= \tilde{\phi}_0 + \tilde{\boldsymbol{\Phi}}\mathbf{y}_{t-1} + \mathbf{u}_t
\end{aligned}
$$

Important: The inverse of an invertible lower triangular matrix will be a lower triangular matrix.

Motivation: Structural shocs produced by the random component are independent and thus diagonal

$$
\begin{aligned}
\tilde{\phi}_0 &= \mathbf{L}^{-1}\phi_0\,, \qquad \tilde{\boldsymbol{\Phi}} = \mathbf{DL}\mathbf{y}_{t-1} \\
\mathbf{u}_t &= \mathbf{L}^{-1}\varepsilon_t \\
\mathrm{E}[\mathbf{u}_t] &= \mathbf{L}^{-1}\mathrm{E}[\varepsilon_t] = 0 \\
\mathrm{E}[\mathbf{u}_t\mathbf{u}_t^T] &= \mathrm{E}[\mathbf{L}^{-1}\varepsilon_t(\mathbf{L}^{-1}\varepsilon_t)^T] \\
&= \mathbf{L}^{-1}\mathrm{E}[\varepsilon_t\varepsilon_t^T](\mathbf{L}^{-1})^T = \mathbf{L}^{-1}\boldsymbol{\Sigma}(\mathbf{L}^{\mathrm{T}})^{-1} \\
&= \mathbf{L}^{-1}\mathbf{LDL}^T(\mathbf{L}^{\mathrm{T}})^{-1} = \mathbf{D}
\end{aligned}
$$

Structure of the left hand side $\mathbf{L}^{-1}\mathbf{y}_t$: (System of simultaneous equations, problem of forecast)

$$\mathbf{L}^{-1} \equiv \begin{pmatrix} \lambda_{11} & 0 & 0 & \cdot & 0 \\ \lambda_{21} & \lambda_{22} & 0 & \cdot & 0 \\ \lambda_{31} & \lambda_{32} & \lambda_{33} & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \lambda_{m1} & \lambda_{m2} & \lambda_{m3} & \cdot & \lambda_{mm} \end{pmatrix}$$

$$\mathbf{L}^{-1}\mathbf{y}_t = \begin{pmatrix} \lambda_{11} & 0 & 0 & \cdot & 0 \\ \lambda_{21} & \lambda_{22} & 0 & \cdot & 0 \\ \lambda_{31} & \lambda_{32} & \lambda_{33} & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \lambda_{m1} & \lambda_{m2} & \lambda_{m3} & \cdot & \lambda_{mm} \end{pmatrix} \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ \ldots \\ \\ y_{m,t} \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_{11}y_{1,t} \\ \lambda_{11}y_{1,t} + \lambda_{22}y_{2,t} \\ \ldots \\ \\ \lambda_{m1}y_{1,t} + \lambda_{m2}y_{2,t} + \ldots \lambda_{mm}y_{m,t} \end{pmatrix}$$

Unit Root related to stationarity

$$
\begin{aligned}
\mathbf{y}_t &= z^t \mathbf{y}_0 \\
\mathbf{y}_t &= \mathbf{\Phi} \mathbf{y}_{t-1} \\
\mathbf{y}_0 z^t &= \mathbf{\Phi} \mathbf{y}_0 \frac{z^t}{z} \\
\left(\mathbf{I} - \mathbf{\Phi} \frac{1}{z} \right) \mathbf{y}_0 &= 0 \\
\left(\mathbf{I} z - \mathbf{\Phi} \right) \mathbf{y}_0 &= 0
\end{aligned}
$$

$$
\det(\mathbf{I} z - \mathbf{\Phi}) = 0 \quad \text{(roots within the unit cicle)}
$$

$$
\det(\mathbf{I} - \mathbf{\Phi} z_1) = 0, \quad \text{(out of the unit cycle if} z_1 = 1/z)
$$

When **stationary**, the matrix $\mathbf{I} - \boldsymbol{\Phi}$ is regular

$$\mathbf{y}_t - \boldsymbol{\mu} = \boldsymbol{\Phi}(\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \varepsilon_t$$

$$\mathbf{y}_t = \phi_0 + \Phi \mathbf{y}_{t-1} + \varepsilon_t$$

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\mu} - \boldsymbol{\Phi}\boldsymbol{\mu} + \boldsymbol{\Phi}\mathbf{y}_{t-1} + \varepsilon_t \\ &= \underbrace{\boldsymbol{\mu}(\mathbf{I} - \boldsymbol{\Phi})}_{\phi_0} + \boldsymbol{\Phi}\mathbf{y}_{t-1} + \varepsilon_t \end{aligned}$$

VAR(1) as a linear process

$$\mathbf{y}_t = \boldsymbol{\mu} + \varepsilon_t + \boldsymbol{\Phi}\varepsilon_{t-1} + \boldsymbol{\Phi}^2\varepsilon_{t-2} + \boldsymbol{\Phi}^3\varepsilon_{t-3} + \dots$$

$$\begin{aligned} \mathbf{y}_t - \boldsymbol{\mu} &= \varepsilon_t + \boldsymbol{\Phi}\varepsilon_{t-1} + \boldsymbol{\Phi}^2\varepsilon_{t-2} + \boldsymbol{\Phi}^3\varepsilon_{t-3} + \dots \\ \boldsymbol{\Phi}(\mathbf{y}_{t-1} - \boldsymbol{\mu}) &= \boldsymbol{\Phi}\varepsilon_{t-1} + \boldsymbol{\Phi}^2\varepsilon_{t-2} + \boldsymbol{\Phi}^3\varepsilon_{t-3} + \dots \\ \mathbf{y}_t - \boldsymbol{\mu} - \boldsymbol{\Phi}(\mathbf{y}_{t-1} - \boldsymbol{\mu}) &= \varepsilon_t \end{aligned}$$

$$
\begin{aligned}
(\mathbf{y}_t - \boldsymbol{\mu}) &= \left[\varepsilon_t + \mathbf{\Phi}\varepsilon_{t-1} + \mathbf{\Phi}^2\varepsilon_{t-2} + \mathbf{\Phi}^3\varepsilon_{t-3} + \ldots\right] \\
(\mathbf{y}_t - \boldsymbol{\mu})^T &= \left[\varepsilon_t + \mathbf{\Phi}\varepsilon_{t-1} + \mathbf{\Phi}^2\varepsilon_{t-2} + \mathbf{\Phi}^3\varepsilon_{t-3} + \ldots\right]^T \\
(\mathbf{y}_t - \boldsymbol{\mu})^T &= \varepsilon_t^T + \varepsilon_{t-1}^T\mathbf{\Phi}^T + \varepsilon_{t-2}^T[\mathbf{\Phi}^2]^T + \varepsilon_{t-3}^T[\mathbf{\Phi}^3]^T + \ldots
\end{aligned}
$$

When stationary $\rightarrow$ variance covariance matrix exists

$$
\mathrm{Var}(\mathbf{y}_t) = \mathrm{E}\left[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})^T\right]
$$

$$
\mathrm{Var}(\mathbf{y}_t) = \left(\begin{array}{lllll}
\mathrm{E}[(y_{1,t}-\mu_1)^2] & \mathrm{E}[(y_{1,t}-\mu_1)(y_{2,t}-\mu_2)] & \cdot & \cdot & \mathrm{E}[(y_{1,t}-\mu_1)(y_{m,} \\
\mathrm{E}[(y_{2,t}-\mu_2)(y_{1,t}-\mu_1)] & \mathrm{E}[(y_{2,t}-\mu_2)^2] & \cdot & \cdot & \mathrm{E}[(y_{2,t}-\mu_2)(y_{m,} \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\mathrm{E}[(y_{m,t}-\mu_m)(y_{1,t}-\mu_1)] & \mathrm{E}[(y_{m,t}-\mu_m)(y_{2,t}-\mu_2)] & \cdot & \cdot & \mathrm{E}[(y_{m,t}-\mu_m)^2]
\end{array}\right.
$$

... still for VAR(1)

$$\begin{aligned} \text{Var}(\mathbf{y}_t) &= \text{E}[\varepsilon_t \varepsilon_t^T] + \mathbf{\Phi}\text{E}[\varepsilon_{t-1}\varepsilon_{t-1}^T]\mathbf{\Phi}^T \\ &+ \mathbf{\Phi}^2\text{E}[\varepsilon_{t-2}\varepsilon_{t-2}^T][\mathbf{\Phi}^2]^T + \dots \end{aligned}$$

$$\text{Var}(\mathbf{y}_t) = \mathbf{\Sigma} + \mathbf{\Phi}\mathbf{\Sigma}\mathbf{\Phi}^T + \mathbf{\Phi}^2\mathbf{\Sigma}[\mathbf{\Phi}^2]^T + \dots$$

VAR(p)

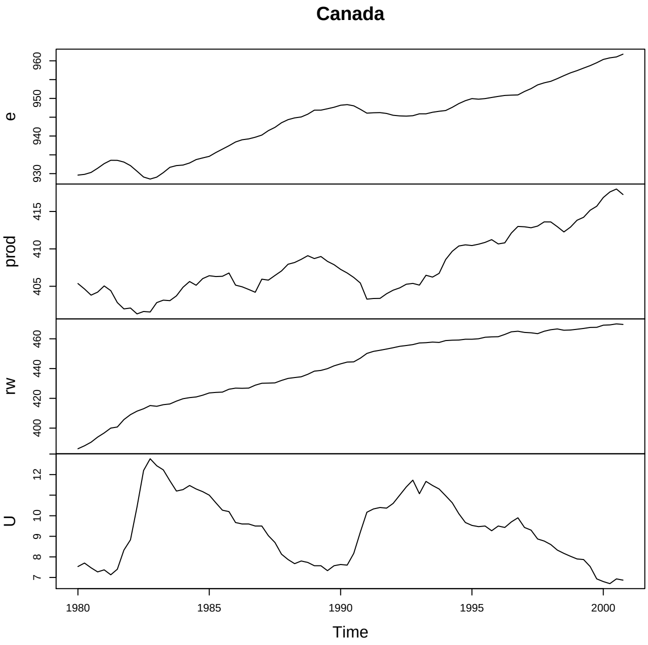
$$\mathbf{y}_t = \phi_0 + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{\Phi}_2 \mathbf{y}_{t-2} + \mathbf{\Phi}_3 \mathbf{y}_{t-3} \dots \mathbf{\Phi}_p \mathbf{y}_{t-p} + \varepsilon_t$$

$$\mathbf{y}_{t-k}^T = (y_{1,t-k}, y_{2,t-k}, \dots, y_{m,t-k})$$

$$\mathbf{\Phi}_k = \begin{pmatrix} \Phi_{k;11} & \Phi_{k;12} & \cdot & \Phi_{k;1m} \\ \Phi_{k;21} & \Phi_{k;22} & \cdot & \Phi_{k;2m} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \Phi_{k;m1} & \Phi_{k;m2} & \cdot & \Phi_{k;mm} \end{pmatrix}, \qquad k = 1, 2, \dots p$$

```
>var.2c<-VAR(Canada,p=2,type="const",ic=AIC)
>summary(var.2c)
```

```
VAR Estimation Results:
=========================
Endogenous variables: e, prod, rw, U
---------------------------------------
"e"    employment
"prod" labor productivity
"U"    unemployment rate
"rw"   real wage
----------------------------------------
Deterministic variables: const
Sample size: 82
Log Likelihood: -175.819
```



Canada
e
960
950
940
930
prod
415
410
405
rw
460
440
420
400
U
12
10
9
8
7
1980
1985
1990
1995
2000
Time

```
VAR(y = Canada, p = 2, type = "const", ic = AIC)
Estimation results for equation e:
==================================
e =  e.l1 + prod.l1 + rw.l1 + U.l1 + e.l2 +
     prod.l2 + rw.l2 + U.l2 + const

          Estimate Std. Error t value Pr(>|t|)
e.l1     1.638e+00  1.500e-01  10.918  < 2e-16 ***
prod.l1  1.673e-01  6.114e-02   2.736  0.00780 **
rw.l1   -6.312e-02  5.524e-02  -1.143  0.25692
U.l1     2.656e-01  2.028e-01   1.310  0.19444
e.l2    -4.971e-01  1.595e-01  -3.116  0.00262 **
prod.l2 -1.017e-01  6.607e-02  -1.539  0.12824
rw.l2    3.844e-03  5.552e-02   0.069  0.94499
U.l2     1.327e-01  2.073e-01   0.640  0.52418
const   -1.370e+02  5.585e+01  -2.453  0.01655 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

roots(var.2c) # of the characteristic polynomial
[1] 0.9950338 0.9081062 0.9081062 0.7380565 0.7380565 0.1856381
[8] 0.1428889 # stable model has all roots less than 1
```

Order of integration of the variables

All the variables **y** used **have to be of the same order of integration** $I(d)$.
Possible Cases:

◯ All the variables are $I(0)$ **stationary** one is in the standard case, i.e. a VAR in level (see Dickey Fuller test).

◯ All the variables are $I(d)$ **non-stationary** with $d > 0$;

• variables are **cointegrated**: i.e. the **error correction term** has to be included into the VAR. The model becomes a **Vector error correction model (VECM)** which can be seen as a restricted VAR.

Note: two or more series are individually integrated (in the time series sense) but some linear combination of them has a lower order of integration, then the series are said to be cointegrated. A common example is where the individual series are first-order integrated (I(1)) but some (cointegrating) vector of coefficients exists to form a stationary linear combination of them.

• The variables are **not cointegrated**: the variables have first to be **differenced** d times and one has a VAR in difference.

Johansen test

Johansen test, is a procedure for testing **cointegration** of several $I(1)$ time series. This test permits **more than one cointegrating relationship**

Note: is more generally applicable than the EngleGranger test which is based on the DickeyFuller (or the augmented) test for unit roots in the residuals from a single estimated cointegrating relationship.

```
ca.jo  {urca}   R Documentation
Johansen Procedure for VAR

Description

Conducts the Johansen procedure on a given data set. Th
or "eigen" statistics are reported and the matrix of ei
as well as the loading matrix.
```

Usage

```
ca.jo(x, type = c("eigen", "trace"), constant = FALSE, K = 2,
         spec=c("longrun", "transitory"), season = NULL,
         dumvar = NULL, ctable = c("A1", "A2", "A3"))
```

```
>Can.vecm<-ca.jo(Canada,ecdet="const",
                 type="eigen",
                 K=2,
                 spec="longrun",
                 season=4)
```

```
summary(Can.vecm)
######################
# Johansen-Procedure #
######################

Test type: maximal eigenvalue statistic (lambda max) , without linear trend
and constant in cointegration

Eigenvalues (lambda):
[1]  5.386738e-01  2.538453e-01  1.170751e-01  8.234188e-02 -2.170923e-15
```

```
Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 3 |  7.05  7.52  9.24 12.97
r <= 2 | 10.21 13.75 15.67 20.20
r <= 1 | 24.01 19.77 22.00 26.81
r = 0  | 63.44 25.56 28.14 33.24
```

What do we conclude by looking at the output? We start going up and compare the **test** statistic with the critical values.

- 63.44 is more than the critical values 25.56, 28.14, 33.24 and thus we reject null hypothesis about zero vector ($r = 0$) cointegration;
- 24.01 is less than the critical vale 26.77 and thus we fail in the cointegration hypothesis about the significance $r <= 1$ at the level 10 percent;

```
Eigenvectors, normalised to first column:
(These are the cointegration relations)

                e.l2      prod.l2        rw.l2          U.l2       constant
e.l2       1.0000000    1.0000000    1.0000000      1.000000     1.00000000
prod.l2    0.1225630    0.1857997   -0.3412087      4.129626    -0.04190071
rw.l2     -0.1404706   -0.3748493   -0.3407458     -1.025258    -0.32552059
U.l2       4.3248159    2.6283235    1.6230256      4.805482     1.56637295
constant -984.2496385 -878.4818889 -670.4038503 -2222.196123 -797.96458695

Weights W:
(This is the loading matrix)
```

	e.l2	prod.l2	rw.l2	U.l2	constant
e.d	-0.006598744	0.162109462	-0.02095845	0.007065933	1.725162e-11
prod.d	0.008061043	0.218797421	-0.01833601	-0.016798969	1.508248e-11
rw.d	-0.106917498	-0.003616112	0.14001491	-0.005394187	-3.263382e-12
U.d	-0.021157219	-0.098175004	-0.04241109	-0.004763840	-1.682723e-11

Eigenvectors, normalised to first column:
(These are the cointegration relations)

	LRM.l2	LRY.l2	IBO.l2	IDE.l2	constant
LRM.l2	1.000000	1.0000000	1.0000000	1.000000	1.0000000
LRY.l2	-1.032949	-1.3681031	-3.2266580	-1.883625	-0.6336946
IBO.l2	5.206919	0.2429825	0.5382847	24.399487	1.6965828
IDE.l2	-4.215879	6.8411103	-5.6473903	-14.298037	-1.8951589
constant	-6.059932	-4.2708474	7.8963696	-2.263224	-8.0330127

Weights W:
(This is the loading matrix)

	LRM.l2	LRY.l2	IBO.l2	IDE.l2	constant
LRM.d	-0.21295494	-0.00481498	0.035011128	2.028908e-03	1.739631e-13
LRY.d	0.11502204	0.01975028	0.049938460	1.108654e-03	7.149427e-14
IBO.d	0.02317724	-0.01059605	0.003480357	-1.573742e-03	2.186400e-14
IDE.d	0.02941109	-0.03022917	-0.002811506	-4.767627e-05	3.091445e-14

Granger Causality analysis

$$\text{MSE}(\underline{\textit{with}}\ \textit{variable}\ y_l) < \text{MSE}(\underline{\textit{without}}\ \textit{variable}\ y_l)$$

In general: in the VAR(p) model $\underline{l\text{th ordered variable}}$ does not Granger-cause the $\underline{j-\text{th ordered variable}}$ if and only if the joint null hypothesis:

$$\boxed{H_0: \quad \Phi_{k;jl} = 0 \text{ for all } k = 1, 2, \ldots, p}$$

```
> data(Canada)
> var.2c<-VAR(Canada,p=2,type="const")
> causality(var.2c,cause="e")
\$ Granger

Granger causality H0: e do not Granger-cause prod rw U
data:  VAR object var.2c
F-Test = 6.2768, df1 = 6, df2 = 292, p-value = 3.206e-06

\$Instant
       H0: No instantaneous causality between: e and prod rw U
data:  VAR object var.2c
Chi-squared = 26.0685, df = 3, p-value = 9.228e-06
```

(1) Null: employment "e" **does not Granger cause** remaining three variables: prod, rw, and U
(2) Null: employment "e" **does not Granger cause** remaining three variables: prod, rw, and U, The Wald test statistics is asymptotically Chi-squared.
Due to low $p-$values in both types of the causality tests **we reject noncausality**, supporting the **possible causation**.
There are macroeconomic reasons to believe in the causal relation is at all plausible.

VECM(1) from VAR(1)

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{\Phi}\mathbf{y}_{t-1} + \epsilon_t \\
\mathbf{y}_t &= \mathbf{y}_{t-1} + \Delta\mathbf{y}_t \\
\Delta\mathbf{y}_t &= (\mathbf{\Phi} - \mathbf{I})\mathbf{y}_{t-1} + \epsilon_t
\end{aligned}$$

VECM(2) from VAR(2)

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{\Phi}_1\mathbf{y}_{t-1} + \mathbf{\Phi}_2\mathbf{y}_{t-2} + \epsilon_t \\
\mathbf{y}_t &= \mathbf{y}_{t-1} + \Delta\mathbf{y}_t \\
\mathbf{y}_{t-2} &= \mathbf{y}_{t-1} - \Delta\mathbf{y}_{t-1} \\
\Delta\mathbf{y}_t &= (\mathbf{\Phi}_1 - \mathbf{\Phi}_2 - \mathbf{I})\mathbf{y}_{t-1} - \mathbf{\Phi}_2\Delta\mathbf{y}_{t-1} + \epsilon_t
\end{aligned}$$

VECM(3) from VAR(3)

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{\Phi}_1\mathbf{y}_{t-1} + \mathbf{\Phi}_2\mathbf{y}_{t-2} + \mathbf{\Phi}_3\mathbf{y}_{t-3} + \epsilon_t \\
\mathbf{y}_t &= \mathbf{y}_{t-1} + \Delta\mathbf{y}_t \\
\mathbf{y}_{t-2} &= \mathbf{y}_{t-1} - \Delta\mathbf{y}_{t-1} \\
\mathbf{y}_{t-3} &= \mathbf{y}_{t-2} - \Delta\mathbf{y}_{t-2} \\
\Delta\mathbf{y}_t &= (\mathbf{\Phi}_1 + \mathbf{\Phi}_2 - \mathbf{I})\mathbf{y}_{t-1} + \mathbf{\Phi}_3\mathbf{y}_{t-2} - \mathbf{\Phi}_2\Delta\mathbf{y}_{t-1} - \mathbf{\Phi}_3\Delta\mathbf{y}_{t-2} + \epsilon_t \\
\Delta\mathbf{y}_t &= (\mathbf{\Phi}_1 + \mathbf{\Phi}_2 + \mathbf{\Phi}_3 - \mathbf{I})\mathbf{y}_{t-1} - (\mathbf{\Phi}_2 + \mathbf{\Phi}_3)\Delta\mathbf{y}_{t-1} - \mathbf{\Phi}_3\Delta\mathbf{y}_{t-2} + \epsilon_t
\end{aligned}$$

$$\Delta \mathbf{y}_t = \underbrace{(\mathbf{\Phi}_1 + \mathbf{\Phi}_2 + \mathbf{\Phi}_3 - \mathbf{I})}_{\mathbf{\Pi}} \mathbf{y}_{t-1} + \underbrace{-(\mathbf{\Phi}_2 + \mathbf{\Phi}_3)}_{\mathbf{\Gamma}_1} \Delta \mathbf{y}_{t-1} + \underbrace{-\mathbf{\Phi}_3}_{\mathbf{\Gamma}_2} \Delta \mathbf{y}_{t-2} + \epsilon_t$$

VECM(3)

$$\Delta \mathbf{y}_t = \mathbf{\Pi} \mathbf{y}_{t-1} + \mathbf{\Gamma}_1 \Delta \mathbf{y}_{t-1} + \mathbf{\Gamma}_2 \Delta \mathbf{y}_{t-2} + \epsilon_t$$

VECM(p)

$$\begin{aligned}
\Delta \mathbf{y}_t &= \mathbf{\Pi} \mathbf{y}_{t-1} + \mathbf{\Gamma}_1 \Delta \mathbf{y}_{t-1} + \ldots + \mathbf{\Gamma}_{p-1} \Delta \mathbf{y}_{t-p+1} + \epsilon_t \\
\mathbf{\Pi} &= \mathbf{\Phi}_1 + \mathbf{\Phi}_2 + \ldots + \mathbf{\Phi}_p - \mathbf{I} \\
\mathbf{\Gamma}_1 &= -\mathbf{\Phi}_2 - \mathbf{\Phi}_3 + \ldots - \mathbf{\Phi}_p \\
\mathbf{\Gamma}_2 &= -\mathbf{\Phi}_3 - \mathbf{\Phi}_4 + \ldots - \mathbf{\Phi}_p \\
\ldots\ldots & \\
\mathbf{\Gamma}_{p-2} &= -\mathbf{\Phi}_{p-1} - \mathbf{\Phi}_p \\
\mathbf{\Gamma}_{p-1} &= -\mathbf{\Phi}_p
\end{aligned}$$

Theorem (Granger): Let $\mathbf{y}_t \sim I(1)$ then the matrices $\boldsymbol{\alpha}$, $\boldsymbol{\beta}^{\mathrm{T}}$ of the order $m \times r$ exist such that

$$\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}}, \qquad \mathrm{Rank}(\mathbf{\Pi}) = r < m$$

and r cointegration relations $\boldsymbol{\beta}^T \mathbf{y}_t$ are of the order $I(0)$ (i.e. they are stationary)

Note: the summation of elements:
$\Pi_{ik} = \sum_{j=1}^{r} \alpha_{ij}\beta_{jk}^{\mathrm{T}} = \sum_{j=1}^{r} \alpha_{ij}\beta_{kj}$, $i = 1, 2, \ldots, m$, $k = 1, 2, \ldots m$
$[\boldsymbol{\beta}^T \mathbf{y}_t]_j = \sum_{k=1}^{m} \beta_{jk}^{T} y_{k,t} = \sum_{k=1}^{m} \beta_{kj} y_{k,t}$, $j = 1, 2, \ldots r$

Special cases:

- If the $\mathbf{\Pi} = 0$ there is no cointegration, the nonstationarity vanishes by taking the differences $\Delta \mathbf{y}_t$ of the original data;
- For $\mathrm{Rank}(\mathbf{\Pi}) = m$ there is no $\mathbf{y}_t \sim I(1)$, and thus instead stationarity of the processes exists (they are $\mathbf{y}_t \sim I(0)$)

Long run - model of equilibrium (fixed point)

$$
\begin{aligned}
\Delta \mathbf{y}_t &= \Delta \mathbf{y}^* = 0\,, \qquad \mathbf{y}_t = \mathbf{y}^*\,, \\
\mathbf{\Pi} \mathbf{y}^* &= 0 \\
\boldsymbol{\alpha}\boldsymbol{\beta}^T \mathbf{y}^* &= 0 \\
\boldsymbol{\beta}^T \mathbf{y}^* &= 0\,, \qquad \mathbf{y}^* \perp \boldsymbol{\beta}^T
\end{aligned}
$$

Stability of the equlibrium, (of the fixed point)

$$
\begin{aligned}
\Delta \mathbf{y}_t &= \Delta \mathbf{y}_0 \exp(\lambda t) = \Delta \mathbf{y}_0 z^t, \qquad z = \exp(\lambda) \\
\Delta \mathbf{y}_{t-1} &= \Delta \mathbf{y}_0 z^{t-1} = \Delta \mathbf{y}_0 \frac{z^t}{z} \\
\Delta \mathbf{y}_{t-2} &= \Delta \mathbf{y}_0 z^{t-2} = \Delta \mathbf{y}_0 \frac{z^t}{z^2} \\
\Delta \mathbf{y}_{t-p+1} &= \Delta \mathbf{y}_0 z^{t-p+1} = \Delta \mathbf{y}_0 \frac{z^t}{z^{p-1}}
\end{aligned}
$$

After the substitution into homogeneous VECM we obtained

$$
\left[-\mathbf{I} - (\mathbf{\Pi} + \mathbf{\Gamma}_1) z^{-1} + \mathbf{\Gamma}_2 z^{-2} + \ldots + \mathbf{\Gamma}_{p-1} z^{1-p}\right] \Delta \mathbf{y}_0 = 0
$$
$$
\left[-\mathbf{I}\, z^{p-1} - (\mathbf{\Pi} + \mathbf{\Gamma}_1) z^{p-2} + \mathbf{\Gamma}_2 z^{p-3} + \ldots + \mathbf{\Gamma}_{p-1}\right] \Delta \mathbf{y}_0 = 0
$$

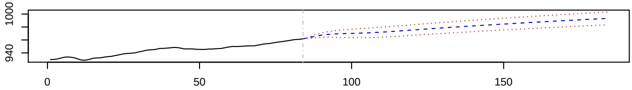
VAR - out of sample forecast

- "predict" function from "vars"

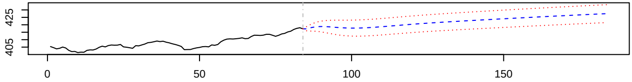
- h-step ahead forecast with flexible choice
$h = 1, 2, \ldots n$ where
n is chosen under the name "n.ahead"

```
> var.2c<-VAR(Canada,p=2,type="const")
> library(vars)
> var.f100<-predict(var.2c,n.ahead=10,ci=0.95)
> names(var.f100)
  [1] "fcst"  "endog" "model"  "exo.fcst"
> layout(matrix(1:3, nrow=3, ncol=1))
> var.f100<-predict(var.2c,n.ahead=100,ci=0.95)
```

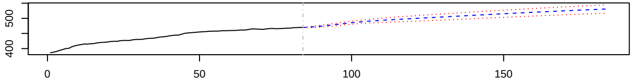
Forecast of series e



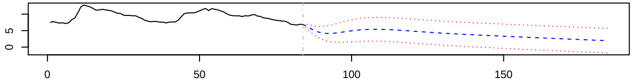
Forecast of series prod



Forecast of series rw



Forecast of series U



Impulse-Response:

Motivation: when the Fed changes the federal fund (interst) rate by 50 basis points it is akin to the administered impulse

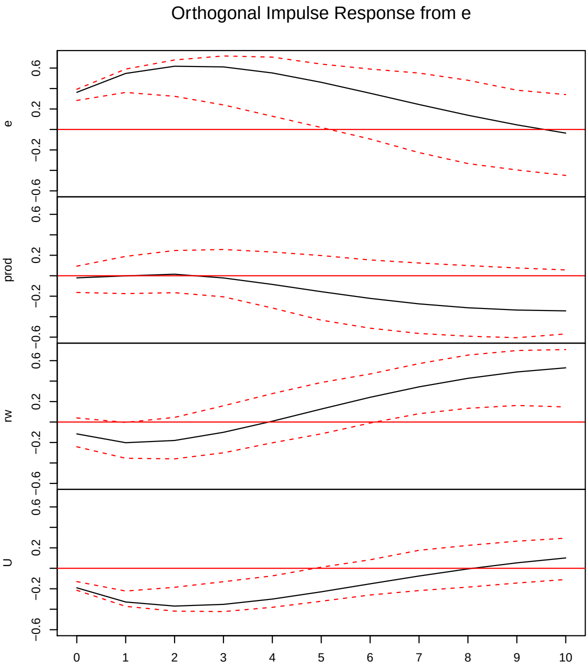
$\mathbf{u} = \mathbf{L}\mathbf{u}_t$, $E[u_t^T u_t] = \mathbf{D}$ (Cholesky decomposition)

$$
\begin{aligned}
\mathbf{L}\mathbf{u}_t &= \begin{pmatrix} 1 & 0 & 0 & \cdot & 0 \\ L_{21} & 1 & 0 & \cdot & 0 \\ L_{31} & L_{32} & 1 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & 0 \\ L_{m1} & L_{m2} & L_{m3} & \cdot & 1 \end{pmatrix} \begin{pmatrix} \boxed{u_{1,t}} \\ u_{2,t} \\ u_{3,1} \\ \dots \\ u_{m,t} \end{pmatrix} \\
&= \begin{pmatrix} \boxed{u_{1,t}} \\ L_{21}\boxed{u_{1,t}} + u_{2,t} \\ L_{31}\boxed{u_{1,t}} + L_{32}u_{2,t} + u_{3,t} \\ \dots \\ L_{m1}\boxed{u_{1t}} + L_{m2}u_{2,t} + \dots L_{mm}u_{m,t} \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{y}_t &= \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t + \boldsymbol{\Phi}\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\Phi}^2\boldsymbol{\varepsilon}_{t-2} + \boldsymbol{\Phi}^3\boldsymbol{\varepsilon}_{t-3} + \\
\mathbf{y}_t &= \boldsymbol{\mu} + \mathbf{L}\mathbf{u}_t + \boldsymbol{\Phi}\mathbf{L}\mathbf{u}_{t-1} + \boldsymbol{\Phi}^2\mathbf{L}\mathbf{u}_{t-2} + \boldsymbol{\Phi}^3\mathbf{L}\mathbf{u}_{t-3} +
\end{aligned}
$$

R implementation of impulse response

```
> library("vars")
> data(Canada)
> var.2c<-VAR(Canada,p=2,type="const")
> library(vars)
> irf.e=irf(var.2c, impulse="e")
> pdf("fig_irfe.pdf")
> plot(irf.e)
> dev.off()
```


Orthogonal Impulse Response from e
e
prod
rw
U
0.6
0.2
−0.2
−0.6
0 1 2 3 4 5 6 7 8 9 10