Vector Auto-Regressive Models

Laurent Ferrara[1]

[1]University of Paris Nanterre

M2 Oct. 2018

Overview of the presentation

1. Vector Auto-Regressions
    - Definition
    - Estimation
    - Testing
2. Impulse responses functions (IRF)
    - Concept
    - General IRF
3. Forecasting
    - Concept
    - Variance decomposition
4. Extensions
5. Applications
    - US-EA GDP relationships

Vector Auto-Regressions: Short introduction

- The VAR models are widely used in economic analysis.
- While simple and easy to estimate, they make it possible to conveniently capture the dynamics of multivariate systems.
- VAR popularity is mainly due to Sims (1980) influential work.
- Reference text books : Hamilton (1994), Tsay (2014)

Vector Auto-Regressions: Notations

- Let y_t denote an $(n \times 1)$ vector of random variables. y_t follows a p^{th} order Gaussian VAR if, for all t, we have

$$y_t = c + \Phi_1 y_{t-1} + \ldots \Phi_p y_{t-p} + \varepsilon_t$$

  where c is n-vector, Φ_i are $n \times n$ matrices, $\varepsilon_t \sim N(0, \Omega)$ with Ω is a positive-definite covariance matrix.

- Consequently, the conditional distribution :

$$y_t \mid y_{t-1}, y_{t-2}, \ldots, y_{-p+1} \sim N(c + \Phi_1 y_{t-1} + \ldots \Phi_p y_{t-p}, \Omega).$$

Vector Auto-Regressions: Exemple $n = 2$

VAR(1)for $y_t = (y_{1,t}, y_{2,t})$:

$$\begin{aligned} y_{1,t} &= c_1 + \phi_{11} y_{1,t-1} + \phi_{12} y_{2,t-1} + \varepsilon_{1,t} \\ y_{2,t} &= c_2 + \phi_{21} y_{1,t-1} + \phi_{22} y_{2,t-1} + \varepsilon_{2,t}. \end{aligned}$$

where $\varepsilon_{1,t} \sim GWN(\sigma^2_{\varepsilon_1})$, $\varepsilon_{2,t} \sim GWN(\sigma^2_{\varepsilon_2})$ and $\rho(\varepsilon_{1,t}, \varepsilon_{2,t}) = 0$
ϕ_{11} and ϕ_{22} are autoregressive coefficients,
ϕ_{21} and ϕ_{12} are coefficients measuring the linear dependence between y_1 and y_2.

Vector Auto-Regressions: Notations

- In lag operator notation:

$$\Phi(B)y_t = c + \varepsilon_t$$

where $\Phi(B) = I_n - \Phi_1 B - \ldots - \Phi_p B^p$, B the backwards operator

- The VAR(p) is stationary if all the roots of

$$\det(I_n - \Phi_1 z - \ldots - \Phi_p z^p) = 0$$

lie outside the unit circle (ie: have modulus greater than 1)

- If y_t is stationary, the unconditional mean vector is:

$$\mu = (I_n - \Phi_1 - \ldots - \Phi_p)^{-1} c = (\Phi(1))^{-1} c$$

Rewritting a VAR(p) as a VAR(1)

Let define $Y_t = (y'_t, y'_{t-1}, \ldots, y'_{t-p+1})$ the stacked pn-vector, then the VARp process can be rewritten

$$Y_t = \mathbf{\Phi} Y_{t-1} + \varepsilon^*_t$$

with $\varepsilon^*_t = (\varepsilon'_t, 0')$ with 0 being a $k(p-1)$-vector of zeros such that

$$Y_t \quad = \quad \begin{bmatrix} c \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_p \\ 1 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} Y_{t-1} + \begin{bmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Rewritting a VAR(p) as a MA process

For sake of simplicity, assume y_t is VAR(1) with $\mu = 0$.

$$\begin{aligned} y_t &= \varepsilon_t + \phi_1 y_{t-1} \\ y_t &= \varepsilon_t + \phi_1(\varepsilon_{t-1} + \phi_1 y_{t-1}) \\ y_t &= \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2(+\varepsilon_{t-2} + \phi_1 y_{t-2}) \\ y_t &= \dots \end{aligned}$$

$$y_t = \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \phi_1^3 \varepsilon_{t-3} + \dots$$

ie general form :

$$y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \psi_3 \varepsilon_{t-3} + \dots$$

Rewritting a VAR(p) as a MA process

If y_t is a general stationary VAR(p) process, it has a Wold representation of the following form :

$$y_t = \varepsilon_t + \sum_{j=1}^{\infty} \psi_j \varepsilon_{t-j}$$

with

$$\psi_i = \sum_{j=1}^{\min(i,p)} \phi_j \psi_{i-j}$$

Vector Auto-Regressions: Unconditional variance

- The unconditional matrix of variance-covariance of y_t is

$$Var(y) = \lim_{t\to\infty} E_0((y_t - \bar{y}_t)(y_t - \bar{y}_t)')$$

where $\bar{y}_t$ denotes the unconditional mean of y.

- Let denote with y_t^* the vector $\left[\begin{array}{cccc} y_t' & y_{t-1}' & \dots & y_{t-p}' \end{array}\right]'$, we have

$$\begin{aligned}
y_t^* &= \begin{bmatrix} c \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_p \\ 1 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} y_{t-1}^* + \begin{bmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\
y_t^* &= c^* + \Phi y_{t-1}^* + \varepsilon_t^*
\end{aligned}$$

Vector Auto-Regressions: Unconditional variance

- It is then easy to get the Wold's decomposition of y_t^*:

$$\begin{aligned} y_t^* &= c^* + \Phi\left(c^* + \Phi y_{t-2}^* + \varepsilon_{t-1}^*\right) + \varepsilon_t^* \\ &= c^* + \varepsilon_t^* + \Phi(c^* + \varepsilon_{t-1}^*) + \ldots + \Phi^k(c^* + \varepsilon_{t-k}^*) + \ldots \end{aligned}$$

- The ε_t^*'s being iid, we have

$$Var(y) = \Omega + \Phi\Omega\Phi' + \ldots + \Phi^k\Omega\Phi'^k + \ldots$$

Parameter estimation methods

3 main types of techniques for VAR models

1. Least-Squares (LS)
2. Maximum Likelihood (ML)
3. Bayesian methods

Under the multivariate Normality assumption for the error terms, ML estimates are asymptotically equivalent to the LS estimates (see Hamilton, 94, for a proof)

VAR estimation: LS

Let's rewrite y_t supposed to follow a VAR(p) model for $t = p+1, \dots, T$:

$$y_t = c + \Phi_1 y_{t-1} + \dots + \Phi_p y_{t-p} + \varepsilon_t$$

as:

$$y_t' = x_t' \Pi + \varepsilon_t'$$

where $x_t = (1, y_{t-1}', \dots, y_{t-p}')'$ is a $np+1$-vector and $\Pi' = \begin{bmatrix} c & \Phi_1 & \Phi_2 & \dots & \Phi_p \end{bmatrix}$ is a $n \times (np+1)$ matrix.
Matrix notation:

$$\mathbf{Z} = \mathbf{X}\Pi + \varepsilon$$

where $\mathbf{Z}$ is a $(T-p) \times n$ matrix ith *ith* row being y_{p+i}' and $\mathbf{X}$ is a $(T-p) \times (np+1)$ design matrix with ith row being x_{p+i}' and ε is a $(T-p) \times n$ matrix with ith row being ε_{p+i}'

VAR estimation: LS

The LSE of Π, denoted with $\hat{\Pi}_{LS}$ is given by

$$\hat{\Pi}_{LS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \tag{1}$$

$$\hat{\Pi}_{LS} = \left[\sum_{t=1}^{T} x_t x_t'\right]^{-1} \left[\sum_{t=1}^{T} y_t x_t'\right] \tag{2}$$

Vector Auto-Regressions: MLE

- Denoting with Π the matrix $\begin{bmatrix} c & \Phi_1 & \Phi_2 & \dots & \Phi_p \end{bmatrix}'$ and with x_t the vector $\begin{bmatrix} 1 & y_{t-1}' & y_{t-2}' & \dots & y_{t-p}' \end{bmatrix}'$, the log-likelihood is given by

$$\begin{aligned} L(Y_T;\theta) &= -(Tn/2)\log(2\pi) + (T/2)\log\left|\Omega^{-1}\right| \\ &\quad -\frac{1}{2}\sum_{t=1}^{T}\left[\left(y_t - \Pi' x_t\right)' \Omega^{-1} \left(y_t - \Pi' x_t\right)\right]. \end{aligned}$$

- The MLE of Π, denoted with $\hat{\Pi}$ is given by

$$\hat{\Pi}' = \left[\sum_{t=1}^{T} y_t x_t'\right] \left[\sum_{t=1}^{T} x_t x_t'\right]^{-1}. \tag{3}$$

Vector Auto-Regressions: MLE

Proof of equation (3)

Let's rewrite the last term of the log-likelihood

$$\sum_{t=1}^{T}\left[\left(y_t-\Pi' x_t\right)'\Omega^{-1}\left(y_t-\Pi' x_t\right)\right] =$$

$$\sum_{t=1}^{T}\left[\left(y_t-\hat{\Pi}' x_t+\hat{\Pi}' x_t-\Pi' x_t\right)'\Omega^{-1}\left(y_t-\hat{\Pi}' x_t+\hat{\Pi}' x_t-\Pi' x_t\right)\right] =$$

$$\sum_{t=1}^{T}\left[\left(\hat{\varepsilon}_t+(\hat{\Pi}-\Pi)' x_t-\right)'\Omega^{-1}\left(\hat{\varepsilon}_t+(\hat{\Pi}-\Pi)' x_t\right)\right]$$

where the j^{th} element of the $(n \times 1)$ vector $\hat{\varepsilon}_t$ is the sample residual for observation t from an OLS regression of y_{jt} on x_t.

Vector Auto-Regressions: MLE

$$\sum_{t=1}^{T}\left[\left(y_t-\Pi' x_t\right)' \Omega^{-1}\left(y_t-\Pi' x_t\right)\right] \quad =$$

$$\sum_{t=1}^{T} \hat{\varepsilon}'_t \Omega^{-1} \hat{\varepsilon}_t + 2\sum_{t=1}^{T} \hat{\varepsilon}'_t \Omega^{-1}(\hat{\Pi}-\Pi)' x_t$$

$$+\sum_{t=1}^{T} x_t'(\hat{\Pi}-\Pi)\Omega^{-1}(\hat{\Pi}-\Pi)' x_t$$

Vector Auto-Regressions: MLE

Let's apply the trace operator on the second term (that is a scalar):

$$\begin{aligned}
\sum_{t=1}^{T} \hat{\varepsilon}'_t \Omega^{-1} (\hat{\Pi} - \Pi)' x_t &= trace\left(\sum_{t=1}^{T} \hat{\varepsilon}'_t \Omega^{-1} (\hat{\Pi} - \Pi)' x_t \right) \\
&= trace\left(\sum_{t=1}^{T} \Omega^{-1} (\hat{\Pi} - \Pi)' x_t \hat{\varepsilon}'_t \right) \\
&= trace\left(\Omega^{-1} (\hat{\Pi} - \Pi)' \sum_{t=1}^{T} x_t \hat{\varepsilon}'_t \right)
\end{aligned}$$

Vector Auto-Regressions: MLE

Given that, by construction, the sample residuals are orthogonal to the explanatory variables, this term is equal to zero.
If $\tilde{x}_t = (\hat{\Pi} - \Pi)' x_t$, we have

$$\sum_{t=1}^{T} \left[\left(y_t - \Pi' x_t \right)' \Omega^{-1} \left(y_t - \Pi' x_t \right) \right] =$$

$$\sum_{t=1}^{T} \hat{\varepsilon}'_t \Omega^{-1} \hat{\varepsilon}_t + \sum_{t=1}^{T} \tilde{x}'_t \Omega^{-1} \tilde{x}_t$$

Since Ω is a positive definite matrix, Ω^{-1} is as well. Consequently, the smallest value that the last term can take is obtained when $x_t^* = 0$,ie when $\Pi = \hat{\Pi}$.

Vector Auto-Regressions: MLE

- Assume that we have computed $\hat{\Pi}$, the MLE of is the matrix $\hat{\Omega}$ that maximizes $\Omega \xrightarrow{\ell} L(Y_T; \hat{\Pi}, \Omega)$.
- Denoting with $\hat{\varepsilon}_t$ the estimated residual $y_t - \hat{\Pi} x_t$, we have

$$\begin{aligned} L(Y_T; \hat{\Pi}, \Omega) &= -(Tn/2)\log(2\pi) + (T/2)\log\left|\Omega^{-1}\right| \\ &\quad -\frac{1}{2}\sum_{t=1}^{T}\left[\hat{\varepsilon}_t'\Omega^{-1}\hat{\varepsilon}_t\right]. \end{aligned}$$

- $\hat{\Omega}$ is a symmetric positive definite matrix. Fortunately, it turns out that that the unrestricted matrix that maximizes the latter expression is a symmetric postive definite matrix. Indeed,

$$\frac{\partial \ell(\Omega)}{\partial \Omega} = \frac{T}{2}\Omega' - \frac{1}{2}\sum_{t=1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t' \Longrightarrow \hat{\Omega}' = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t'.$$

Vector Auto-Regressions: Likelihood-Ratio test

- The simplicity of the VAR framework and the tractability of its MLE contribute to convenience of various econometric tests. We illustrate this here with the likelihhod ratio test.
- The maximum value achieved by the MLE is

$$\begin{aligned} L(Y_T; \hat{\Pi}, \hat{\Omega}) &= -(Tn/2)\log(2\pi) + (T/2)\log\left|\hat{\Omega}^{-1}\right| \\ &\quad -\frac{1}{2}\sum_{t=1}^{T}\left[\hat{\varepsilon}_t'\hat{\Omega}^{-1}\hat{\varepsilon}_t\right]. \end{aligned}$$

Vector Auto-Regressions: Likelihood-Ratio test

- The last term is

$$\begin{aligned}\sum_{t=1}^{T}\hat{\varepsilon}_t'\hat{\Omega}^{-1}\hat{\varepsilon}_t &= \text{trace}\left[\textstyle\sum_{t=1}^{T}\hat{\varepsilon}_t'\hat{\Omega}^{-1}\hat{\varepsilon}_t\right] \\ &= \text{trace}\left[\sum_{t=1}^{T}\hat{\Omega}^{-1}\hat{\varepsilon}_t\hat{\varepsilon}_t'\right] = \text{trace}\left[\hat{\Omega}^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t'\right] \\ &= \text{trace}\left[\hat{\Omega}^{-1}\left(T\hat{\Omega}\right)\right] = Tn.\end{aligned}$$

- Therefore

$$L(Y_T;\hat{\Pi},\hat{\Omega}) = -(Tn/2)\log(2\pi) + (T/2)\log\left|\hat{\Omega}^{-1}\right| - Tn/2.$$

which is easy to calculate.

Vector Auto-Regressions: Likelihood-Ratio test

- For instance, assume that we want to test the null hypothesis that a set of variable follows a VAR(p_0) against the alternative specification of p_1 lags (with $p_1 > p_0$).
- Let us respectively denote with $\hat{L}_0$ and $\hat{L}_1$ the maximum log-likelihoods obtained withp_0 and p_1 lags. Under the null hypothesis, we have

$$2\left(\hat{L}_1 - \hat{L}_0\right) = T\left(\log\left|\hat{\Omega}_1^{-1}\right| - \log\left|\hat{\Omega}_0^{-1}\right|\right)$$

  which asymptotically has a χ^2 distribution with degrees of freedom equal to the number of restrictions imposed under H_0 (compared with H_1), ie $n^2(p_1 - p_0)$.

Vector Auto-Regressions: Criteria

- Adding lags quickly consume degrees of freedom. If lag length is p, each of the n equations contains $n \times p$ coefficients plus the intercept term.
- Adding lengths improve in-sample fit, but is likely to result in over-parameterization and affect the out-of-sample prediction performance.
- To select appropriate lag length, some criteria can be used (they have to be minimized)

$$
\begin{aligned}
AIC(p) &= \log\left|\hat{\Omega}\right| + \frac{2}{T}N \\
SBIC(p) &= \log\left|\hat{\Omega}\right| + \frac{\log T}{T}N \\
HQ(p) &= \log\left|\hat{\Omega}\right| + \frac{2\log(\log T)}{T}N
\end{aligned}
$$

where $N = n \times p^2 + p$.

Vector Auto-Regressions: Granger Causality

- Granger (1969) developed a method to analyze the causal relationship among variables systematically.
- The approach consists in determining whether the past values of $y_{1,t}$ can help to explain the current $y_{2,t}$.
- Let us denote three information sets

$$\begin{aligned} I_{1,t} &= \{y_{1,t}, y_{1,t-1}, \ldots\} \\ I_{2,t} &= \{y_{2,t}, y_{2,t-1}, \ldots\} \\ I_t &= \{y_{1,t}, y_{1,t-1}, \ldots y_{2,t}, y_{2,t-1}, \ldots\}. \end{aligned}$$

- We say that $y_{1,t}$ Granger-causes $y_{2,t}$ if

$$E\left[y_{2,t} \mid I_{2,t-1}\right] \neq E\left[y_{2,t} \mid I_{t-1}\right].$$

Vector Auto-Regressions: Granger Causality

- To get the intuition behind the testing procedure, consider the following bivariate VAR(p) process:

$$
\begin{aligned}
y_{1,t} &= \Phi_{10} + \Sigma_{i=1}^{p} \Phi_{11}(i) y_{1,t-i} + \Sigma_{i=1}^{p} \Phi_{12}(i) y_{2,t-i} + u_{1,t} \\
y_{2,t} &= \Phi_{20} + \Sigma_{i=1}^{p} \Phi_{21}(i) y_{1,t-i} + \Sigma_{i=1}^{p} \Phi_{22}(i) y_{2,t-i} + u_{2,t}.
\end{aligned}
$$

- Then,$y_{1,t}$ does not Granger-cause $y_{2,t}$ if

$$
\Phi_{21}(1) = \Phi_{21}(2) = \ldots = \Phi_{21}(p) = 0.
$$

- Therefore the hypothesis testing is

$$
\begin{cases}
H_0 : & \Phi_{21}(1) = \Phi_{21}(2) = \ldots = \Phi_{21}(p) = 0 \\
H_A : & \Phi_{21}(1) \neq 0 \text{ or } \Phi_{21}(1) \neq 0 \text{ or} \ldots \Phi_{21}(p) \neq 0.
\end{cases}
$$

Vector Auto-Regressions: Granger Causality

- Rejection of H_0 implies that some of the coefficients on the lagged $y_{1,t}$'s are statistically significant.
- This can be tested using the F-test or asymptotic chi-square test.
  - The F-statistic is $F = \frac{(RSS-USS)/p}{USS/(T-2p-1)}$ (where RSS: restricted residual sum of squares, USS: unrestriced residual sum of squares)
  - Under H_0, the F-statistic is distributed as $F(p, T-2p-1)$
  - In addition, $pF \to \chi^2(p)$.

Vector Auto-Regressions: Granger Causality

See *RATS* example on the US-EA GDP growth relationships

Vector Auto-Regressions: Impulse responses

- Objective: analyzing the effect of a given shock on the endogenous variables.
- Let the stationary VAR(p) system:

$$y_t = c + \Phi_1 y_{t-1} + \dots \Phi_p y_{t-p} + \varepsilon_t$$

- Assume the system receives a shock at t: $\varepsilon_t = \delta$
- bf Definition
  A standard $IRF(h, \delta)$ describes the effects of the shock at date $t + h$ compared to a zero-shock $\varepsilon_t = 0$, assuming that $\varepsilon_{t+h} = 0$ for all $h > 0$.
- The Generalized IRF by Koop, Pesaran and Potter (1996):

$$\begin{aligned} & GIRF(h, \delta, F_{t-1}) \\ = \quad & E\{y_{t+h} | \varepsilon_t = \delta; \varepsilon_{t+h} = 0, h > 0; F_{t-1}\} \\ - \quad & E\{y_{t+h} | \varepsilon_{t+h} = 0, h \geq ; F_{t-1}\} \end{aligned}$$

Vector Auto-Regressions: Impulse responses

Exemple of a centered univariate AR(1) : $x_t = \phi x_{t-1} + \varepsilon_t$.
Assume $x_{t-1} = 0$, thus $x_t = \varepsilon_t = \delta$.

$$IRF(1,\delta) = E(x_{t+1}|\varepsilon_t = \delta, \varepsilon_{t+1} = 0, F_{t-1}) - E(x_{t+1}|\varepsilon_t = \varepsilon_{t+1} = 0, F_{t-1})$$

$$IRF(1,\delta) = \phi\delta$$

$$IRF(2,\delta) = \phi^2\delta....$$

$$IRF(h,\delta) = \phi^h\delta$$

Remark : IRF is proportional to the size of the shock and independent of past history

Vector Auto-Regressions: Impulse responses

- Let us consider a stationary vector random variable y_t that presents the following Wold's decomposition:

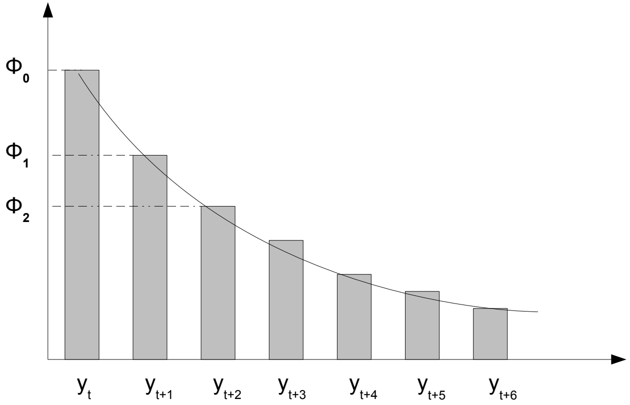
$$y_t = \varepsilon_t + \sum_{j=1}^{\infty} \Psi_j \varepsilon_{t-j}.$$

- The hth impulse response function of the shock ε_t on $y_t, y_{t+1}, \ldots$ is given by $\Psi_h \delta$ and vanishes as $h \to \infty$
- Formally, the impulse response of the shock ε_t on the variable y is defined as

$$\frac{\partial y_{t+h}}{\partial \varepsilon_t} = \Psi_h.$$

Vector Auto-Regressions: Impulse responses

Dynamics of $y_t, y_{t+1}, y_{t+2}, \ldots$ when $\varepsilon_t = 1, \varepsilon_{t+1} = 0, \varepsilon_{t+2} = 0, \ldots$



Vector Auto-Regressions: Forecasting

- The best linear 1-step-ahead forecast that minimizes the MSE if parameters are known is the conditional expectation :

$$\hat{y}_t(1) = c + \Phi_1 y_t + \ldots + \Phi_p y_{t-p+1}$$

- for $h > 1$, the iterations lead to

$$\hat{y}_t(h) = c + \Phi_1 \hat{y}_t(h-1) + \ldots + \Phi_p \hat{y}_t(h-p)$$

where $\hat{y}_t(h-p) = y_{t-p}$ if $h < p$

Forecast variance decomposition

- The h-step-ahead forecast error in y_t is defined as the difference between the actual value of y_t and its VAR-based forecast $\hat{y}_t(h)$, expressed according to their infinite MA representation:

$$y_t = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}$$

$$y_{t+h} - \hat{y}_t(h) = \sum_{j=0}^{h-1} \Psi_j \varepsilon_{t+h-j}$$

Forecast variance decomposition

- It is straigthforward to compute the forecast error variance of a variable in y_t for the h-step forecast horizon as well as the corresponding shares of individual innovations (Lutkepohl, 2015)
- In applications h corresponds to the business cycle horizon (between 1,5y and 8y)

$$V(y_{t+h} - \hat{y}_t(h)) = \sum_{j=0}^{h-1} \Psi_j \Omega \Psi_j'$$

- The contribution of the i^{th} shock to the forecast of k^{th} variable is

$$\sum_{j=0}^{h-1} \{e_k' \Psi_j e_i\}^2$$

where e_i is the i^{th} column of Ω

Vector Auto-Regressions: Applications

See *RATS* example on the US-EA GDP growth relationships

Extension: VAR-X

- Let y_t denote an $(n \times 1)$ vector of random variables. y_t is Gaussian VAR(p) with exogeneous variables $x_t = (x_t^1, \ldots, x_t^m)$ of dimension m, for all t

$$y_t = \Phi_1 y_{t-1} + \ldots \Phi_p y_{t-p} + C x_t + \varepsilon_t$$

where $\varepsilon_t \sim N(0, \Omega)$ and

$$C = \begin{pmatrix} c_{11} & \cdots & c_{1m} \\ \vdots & c_{ij} & \vdots \\ c_{n1} & \cdots & c_{nm} \end{pmatrix}$$

Vector Auto-Regressions: Exemple $n = 2$

VAR(1)for $y_t = (y_{1,t}, y_{2,t})$:

$$
\begin{aligned}
y_{1,t} &= \phi_{11} y_{1,t-1} + \phi_{12} y_{2,t-1} + c_{11} x_{t-1}^1 + c_{12} x_{t-1}^2 + \varepsilon_{1,t} \\
y_{2,t} &= \phi_{21} y_{1,t-1} + \phi_{22} y_{2,t-1} + c_{21} x_{t-1}^1 + c_{22} x_{t-1}^2 + \varepsilon_{2,t}.
\end{aligned}
$$

Vector Auto-Regressions: Extensions

- VAR in levels
- Bayesian VAR
- Non-linear VAR (Smooth-Transition VAR, Markov-Switching VAR)
- Factor-Augmented VAR