Content

- Cross-correlations
- VAR model in standard/reduced form
- Properties of VAR(1), VAR($p$)
- Structural VAR, identification
- Estimation, model specification, forecasting
- Impulse responses

Everybody should recapitulate the properties of univariate AR(1) models!
Notation

Here we use the multiple equations notation. Vectors are of the form: $y_t = (y_{1t}, \ldots, y_{kt})'$ a column vector of length $k$.

$y_t$ contains the values of $k$ variables at time $t$.

$k$ … the number of variables (i.e. number of equations)

$T$ … the number of observations

We do not distinguish in notation between data and random variables.
Motivation
From an empirical/atheoretical point of view observed time series movements are often related with each another. E.g.

- GDP growth and unemployment rate show an inverse pattern,
- oil prices might be a leading indicator for other energy prices, which on the other hand have an effect on oil.
Weak stationarity: multivariate

We assume our series $y_t$ are **weakly stationary**. I.e. for $k = 2$

$$E(y_t) = \begin{bmatrix} E(y_{1t}) \\ E(y_{2t}) \end{bmatrix} = \mu$$

and

$$\text{Cov}(y_t, y_{t-\ell}) = E(y_t - \mu)(y_{t-\ell} - \mu)'$$

$$\text{Cov}(y_t, y_{t-\ell}) = \begin{bmatrix} \text{Cov}(y_{1t}, y_{1, t-\ell}) & \text{Cov}(y_{1t}, y_{2, t-\ell}) \\ \text{Cov}(y_{2t}, y_{1, t-\ell}) & \text{Cov}(y_{2t}, y_{2, t-\ell}) \end{bmatrix} = \Gamma_{\ell}$$

are independent of time $t$. $\ell = \ldots, -1, 0, 1, \ldots$

- $\mu$ is the mean vector, $\mu = (\mu_1, \ldots, \mu_k)'$.
- $\Gamma_0$ the contemporaneous/concurrent covariance matrix
- $\Gamma_0 = [\gamma_{ij}(0)]$ is a $(k \times k)$ matrix.
- $\Gamma_{\ell}$ the **cross-covariance matrix** of order $\ell$, $\Gamma_{\ell} = [\gamma_{ij}(\ell)]$
Cross-correlations
Let $D = \text{diag}(\sigma_1, \ldots, \sigma_k)$ with $\sigma_i = \sqrt{\gamma_{ii}(0)}$, the diagonal matrix of the standard deviations of $y_i$’s. Then

$$\rho_\ell = D^{-1} \Gamma_\ell D^{-1}$$

with $\rho_\ell = [\rho_{ij}(\ell)]$, is the **cross-correlation matrix** of order $\ell$.

$$\rho_{ij}(\ell) = \frac{\gamma_{ij}(\ell)}{\sqrt{\gamma_{ii}(0)} \sqrt{\gamma_{jj}(0)}} = \frac{\text{Cov}(y_{it}, y_{j,t-\ell})}{\sigma_i \sigma_j}$$

$\rho_{ij}(\ell)$ is the correlation coefficient between $y_{it}$ and $y_{j,t-\ell}$. 

Cross-correlation matrices, CCMs
Cross-correlation matrices: properties

We say with $\ell > 0$, if with $\gamma_{ij}(\ell) = \text{Cov}(y_{it}, y_{j,t-\ell})$
\[
\rho_{ij}(\ell) \neq 0: \quad y_j \text{ leads } y_i.
\]
\[
\rho_{ji}(\ell) \neq 0: \quad y_i \text{ leads } y_j.
\]
$\rho_{ii}(\ell)$ is the autocorrelation coefficient of order $\ell$ of variable $y_i$.

Properties of $\Gamma_\ell$ and $\rho_\ell$:

- $\Gamma_0$, the covariance matrix, is symmetric (and pos def).
- $\Gamma_\ell$ with $\ell \neq 0$, is i.g. not symmetric.
  $\rho_{ij}(\ell) \neq \rho_{ji}(\ell)$ in general
- $\gamma_{ij}(\ell) = \gamma_{ji}(-\ell)$ and so $\Gamma_\ell = \Gamma'_{(-\ell)}$:
  \[
  \text{Cov}(y_{it}, y_{j,t-\ell}) = \text{Cov}(y_{j,t-\ell}, y_{it}) = \text{Cov}(y_{jt}, y_{i,t+\ell}) = \text{Cov}(y_{jt}, y_{i,t-(-\ell)})
  \]

Therefore we consider only CCMs with $\ell \geq 0$. 
Types of relationships

We distinguish several types of relationships between 2 series \( i \) and \( j \) \((i \neq j)\):

- **no linear relationship:**\( \rho_{ij}(\ell) = \rho_{ji}(\ell) = 0 \) for all \( \ell \geq 0 \)
- **concurrently correlated:**\( \rho_{ij}(0) \neq 0 \)
  Some cross-correlations of higher order may be \( \neq 0 \).
- **uncoupled:**\( \rho_{ij}(\ell) = \rho_{ji}(\ell) = 0 \) for all \( \ell > 0. \)
  There is no lead-lag but possibly a concurrent relationship.
- **unidirectional** relationship (wrt time) from \( i \) to \( j \):
  \( \rho_{ij}(\ell) = 0 \) for all \( \ell > 0 \) and \( \rho_{ji}(\nu) \neq 0 \) for some \( \nu > 0 \)
- **feedback** relationship (wrt time):
  \( \rho_{ij}(\ell) \neq 0 \) for some \( \ell > 0 \) and \( \rho_{ji}(\nu) \neq 0 \) for some \( \nu > 0 \)
Sample cross-correlations, properties

The sample values are straightforward as in the univariate case.

\[ \hat{\gamma}_{ij}(\ell) = \frac{1}{T} \sum_{t=\ell+1}^{T} (y_{it} - \bar{y}_i) (y_{jt} - \bar{y}_j) \]

- The estimates of \( \gamma \)'s and \( \rho \)'s are consistent (but biased in small samples).
- Under the hypothesis of multivariate white noise, the cross-correlations may be tested individually with the \((\pm 1.96/\sqrt{T})\)-rule.
- The distribution in small samples may deviate considerably from the expected, e.g. for stock returns due to heteroscedasticity and fat tails. Then bootstrap resampling methods are recommended.
Cross-correlations and autocorrelated $y_i$’s

- Cross-correlations can be interpreted in a nice way only, if *at least one* of both series is *white noise*.
- Otherwise the autocorrelation structure of one series may interact with that of the other and ‘spurious’ cross effects may be observed.
- One way to resolve this problem is to use a *prewhitening* technique.
  - Find a univariate model for one of the two series.
  - Apply this *estimated* model also to the other.
  - Interpret the cross-correlations between the residuals of the 1st model and the ‘residuals’ of the 2nd model, instead.
Multivariate portmanteau test

The univariate **Ljung-Box statistic** $Q(m)$ has been generalized to $(k \times k)$-dimensional CCMs. The null hypothesis

$$H_0 : \rho_1 = \ldots = \rho_m = 0$$

is tested against the alternative

$$H_A : \rho_i \neq 0 \quad \text{for some } 1 \leq i \leq m$$

with

$$Q_k(m) = T^2 \sum_{\ell=1}^{m} \frac{1}{T - \ell} \text{tr}(\hat{\Gamma}_\ell \hat{\Gamma}_0^{-1} \hat{\Gamma}_\ell \hat{\Gamma}_0^{-1}) \sim \chi^2(k^2 m)$$

which is $\chi^2$ distributed with $(k^2 m)$ degrees of freedom.
This test is suitable for the case where the interdependence of essentially (univariate) white noise returns is in question.

If only a single series is not white noise, i.e. \( \rho_{ii}(\ell) \neq 0 \) for some \( \ell > 0 \), the test will reject as the univariate tests will do.
The VAR model
The VAR in standard form

A model taking into account/approximating multivariate dynamic relationships is the VAR($p$), **vector autoregression** of order $p$.

$$y_t = \phi_0 + \Phi_1 y_{t-1} + \ldots + \Phi_p y_{t-p} + \epsilon_t$$

- $y_t$ is a vector of length $k$. There are $k$ equations.
- $p$ is the order of the VAR.
- $\{\epsilon_t\}$ is a sequence of serially uncorrelated random vectors with concurrent full rank covariance matrix $\Sigma$ (not diagonal i.g.). They are coupled. 
  \((\Sigma = \Gamma_0^{(\epsilon)}, \Gamma_\ell^{(\epsilon)} = 0, \ell \neq 0)\)
- $\phi_0$ is a $(k \times 1)$ vector of constants.
- $\Phi$’s are $(k \times k)$ coefficient matrices.

The model is called **VARX**, if additional explanatories are included.
Example of VAR(1), $k = 2$

$$y_t = \phi_0 + \Phi_1 y_{t-1} + \epsilon_t$$

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \phi^{(0)}_1 \\ \phi^{(0)}_2 \end{bmatrix} + \begin{bmatrix} \phi^{(1)}_{11} & \phi^{(1)}_{12} \\ \phi^{(1)}_{21} & \phi^{(1)}_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix}$$

or equation by equation

$$y_{1t} = \phi^{(0)}_1 + \phi^{(1)}_{11} y_{1,t-1} + \phi^{(1)}_{12} y_{2,t-1} + \epsilon_{1t}$$

$$y_{2t} = \phi^{(0)}_2 + \phi^{(1)}_{21} y_{1,t-1} + \phi^{(1)}_{22} y_{2,t-1} + \epsilon_{2t}$$

The concurrent relationship between $y_1$ and $y_2$ is measured by the off-diagonal elements of $\Sigma$.

If $\sigma_{12} = 0$ there is no concurrent relationship.

$$\sigma_{12} = \text{Cov}(\epsilon_{1t}, \epsilon_{2t}) = \text{Cov}(y_{1t}(\epsilon_{1t}), y_{2t}(\epsilon_{2t})|y_{t-1})$$
Example of VAR(1), $k = 2$

$\Phi_1$ measures the dynamic dependence in $y$.

$\phi_{12}$:
$\phi_{12}$ measures the linear dependence of $y_{1t}$ on $y_{2,t-1}$ in the presence of $y_{1,t-1}$.
If $\phi_{12} = 0$, $y_{1t}$ does not depend on $y_{2,t-1}$. Then, $y_{1t}$ depends only on its own past.

Analogous for equation 2 and $\phi_{21}$.

- If $\phi_{12} = 0$ and $\phi_{21} \neq 0$:
  There is a unidirectional relationship from $y_1$ to $y_2$.

- If $\phi_{12} = 0$ and $\phi_{21} = 0$:
  $y_1$ and $y_2$ are coupled (in the sense that possibly $\sigma_{12} \neq 0$).

- If $\phi_{12} \neq 0$ and $\phi_{21} \neq 0$:
  There is a feedback relationship between both series.
Properties of the VAR
Properties of the VAR(1)

We investigate the statistical properties of a VAR(1).

\[ y_t = \phi_0 + \Phi y_{t-1} + \epsilon_t \]

Taking expectations we get

\[ \mu = E(y_t) = \phi_0 + \Phi E(y_{t-1}) \quad \text{or} \quad \mu = E(y_t) = (I - \Phi)^{-1} \phi_0 \]

if \((I - \Phi)^{-1}\) exists.

We demean the series as in the univariate case and denote

\[ \tilde{y}_t = y_t - \mu \]

\[ \tilde{y}_t = \Phi \tilde{y}_{t-1} + \epsilon_t \]

Repeated substitutions gives the MA(\(\infty\)) representation

\[ \tilde{y}_t = \epsilon_t + \Phi \epsilon_{t-1} + \Phi^2 \epsilon_{t-2} + \Phi^3 \epsilon_{t-3} + \ldots \]
Properties of the VAR(1)

We see that $y_{t-\ell}$ is predetermined wrt $\epsilon_t$.

- $\text{Cov}(\epsilon_t, y_{t-\ell}) = 0$, $\ell > 0$. $\epsilon_t$ is uncorrelated with all past $y$'s.

- $\epsilon_t$ may be interpreted as a shock or innovation at time $t$. It has possibly an effect on the future but not on the past.

- $\text{Cov}(y_t, \epsilon_t) = \Sigma$. (Multiply by $\epsilon_t'$ and take expectations.)

- $y_t$ depends on the innovation $\epsilon_{t-j}$ with the coeff matrix $\Phi^j$. This implies that $\Phi^j$ has to vanish for stationarity when $j \to \infty$.

- Thus, the eigenvalues of $\Phi$ have to be smaller 1 in modulus. This is necessary and sufficient for weak stationarity. Then also $(I - \Phi)^{-1}$ exists.

(Cp. AR(1) process: $y_t = \alpha y_{t-1} + \epsilon_t$ with $|\alpha| < 1$.)

- $\Gamma_\ell = \Phi \Gamma_{\ell-1}$, $\ell > 0$.

Many properties generalize from the univariate AR to the VAR($p$).
We define the (matrix) polynomial (for a VAR($p$))

$$\Phi(L) = I - \Phi_1 L - \ldots - \Phi_p L^p$$

For stability it is required that the roots of the characteristic equation

$$|\Phi(z)| = 0, \quad |z| > 1$$

are outside the unit circle. E.g. VAR(1): $\Phi(L) = I - \Phi L$.

The VAR in standard form is well defined and can be used to approximate any weakly stationary process arbitrarily well by choosing a suitable order $p$. 
Representation of a VAR($p$) as a VAR(1)

\[
\text{VAR}(p): \quad y_t = \phi_0 + \Phi_1 y_{t-1} + \ldots + \Phi_p y_{t-p} + \epsilon_t
\]

Every VAR($p$) can be written as a ($k \cdot p$)-dimensional VAR(1),

\[
x_t = \Phi^* x_{t-1} + \xi_t
\]

\[x_t = (\tilde{y}_{t-p+1}', \ldots, \tilde{y}_t')' \quad \text{and} \quad \xi_t = (0, \ldots, 0, \epsilon_t')' \quad \text{both} \quad (k \cdot p \times 1).
\]

\[
\Phi^* = \begin{bmatrix}
0 & I & 0 & \ldots & 0 \\
0 & 0 & I & \ldots & 0 \\
0 & 0 & \ldots & \ddots & \ldots \\
0 & 0 & \ldots & 0 & I \\
\Phi_p & \Phi_{p-1} & \Phi_{p-2} & \ldots & \Phi_1
\end{bmatrix}
\]

The ($k \cdot p \times k \cdot p$) matrix $\Phi^*$ is called \textbf{companion matrix} to the matrix polynomial \(\Phi(L) = I - \Phi_1 L - \ldots - \Phi_p L^p\).
The last component of $x_t$ is the mean corrected $y_t$, $\tilde{y}_t$.

The last row of $\Phi^*$ is essentially the VAR($p$) recursion in reverse order.
Structural VAR and identification
Structural VAR

The VAR in standard form is also called VAR in **reduced form**, as it does not contain the concurrent relationships in \( y \) explicitly. A VAR in **structural form** is

\[
\Theta y_t = \theta_0 + \Theta_1 y_{t-1} + \ldots + \Theta_p y_{t-p} + \eta_t
\]

\( \Theta \) is the coefficient matrix of the \( y_t \)'s. Its diagonal elements are all 1. \( \eta_t \) is serially uncorrelated as \( \epsilon_t \), but its concurrent covariance matrix, \( \Omega \), is diagonal. (The concurrent eff are captured in a possibly non diagonal \( \Theta \).)

If \( \Theta \) is invertible, multiplication with \( \Theta^{-1} \) yields the VAR in standard form with

\[
\Phi_j = \Theta^{-1} \Theta_j
\]

\[
\epsilon_t = \Theta^{-1} \eta_t \quad \text{and} \quad \Sigma = \Theta^{-1} \Omega (\Theta^{-1})'
\]
Comparing the number of coefficients including the error covariances, we find

- Standard form: \( 0 + k + pk^2 + [(k^2 - k)/2 + k] \)
- Structural form: \( (k^2 - k) + k + pk^2 + k \)

The number of parameters of a generally specified structural model is always (except \( k = 1 \)) larger than that of the reduced form.
Identification

- There is always (at least) one structural form corresponding to a standard form (e.g. via the Cholesky decomposition of $\Sigma$, see e.g. Tsay p.350)
- However, the representation in a structural form is not unique without putting the required number of restrictions on the parameters in the $\Theta$ matrices. From the point of view of the structural form this is called the **identification problem**.
- A ‘disadvantage’ of the structural framework is that the concurrent relationships of the $y$’s can be interpreted only together with some economic theory. Sims’(1980) conjecture was that the dynamic relationships can be interpreted well, even without economic theory. Unfortunately, this does not hold. (See the discussion of orthogonalized shocks $\eta_t$ below.)
Estimation of a VAR in standard form

We assume the true relationship is a VAR($p$), the model is correctly specified, and the error is normal or close to normal.

▶ LS: If all variables are included in each equation (no restrictions on the $\Phi_j$-parameters, $j > 0$) simple single equation LS is consistent and asy normal. Clearly, single equation LS cannot be efficient, since the cross-covariances in $\Sigma$ are ignored.

▶ GLS: Using the estimated covariances of $\epsilon_t$ improves efficiency.

Interpretation of the $t$-statistics of the single equation estimates is straight forward.
Curse of dimensionality

A VAR($p$) in standard form has

$$p k^2$$

parameters not counting the constant terms and the error variance-covariances.

The number of the parameters is

- linear in the order of the VAR, and
- increases quadratically with the dimension $k$.

The number of observations $(k T)$ increases only linearly with $k$.

Remark: Empirically, it is difficult to interpret e.g. a significant $\hat{\Phi}^{(9)}_{7,23}$. 
Model selection

Like in the univariate ARMA modeling, single coefficients in $\Phi_j$, $j > 0$, are not set to zero a priori. The model selection procedure starts with a maximal plausible order $p_{\text{max}}$. All models with $p = 0, 1, \ldots, p_{\text{max}}$ are estimated. The models can be ranked according to an information criterion.

The (conditional) ML estimate of the concurrent error covariance matrix of the model with order $p$ is

$$\hat{\Sigma}(p) = \frac{1}{T} \sum_{t=p+1}^{T} \hat{\epsilon}_t(p) \hat{\epsilon}_t'(p)$$

The model selection criteria are

$$\text{AIC}(p) = \log(|\hat{\Sigma}(p)|) + 2p k^2 / T$$

$$\text{SBC}(p) = \log(|\hat{\Sigma}(p)|) + \log(T) p k^2 / T$$
As in the univariate case the residuals of the chosen model should be (multivariate) white noise. Here the $\rho^{(\epsilon)}_\ell$ for $\ell > 0$ should be zero.

The multivariate Ljung-Box test may be applied with number of freedoms as the number of cross-correlation coefficients under test reduced for the number of estimated parameters:

$$k^2(m - p)$$
Forecasting
Forecasting

Existing concurrent structural relationships between the endogenous variables are demanding to interpret. However, due to the unique representation of the reduced/standard form the model is suitable for forecasting.

The 1-step ahead forecast (conditional on data up to $T$) is

$$y_T(1) = \phi_0 + \Phi_1 y_T + \ldots + \Phi_p y_{T-p+1}$$

The 1-step ahead forecast error, $e_T(1) = y_{T+1} - y_T(1)$, is

$$e_T(1) = \epsilon_{T+1}$$
Forecasting

The 2-step ahead forecast is

\[ y_T(2) = \phi_0 + \Phi_1 y_T(1) + \Phi_2 y_T + \ldots + \Phi_p y_{T-p+2} \]

The 2-step ahead forecast error is

\[ e_T(2) = y_{T+2} - y_T(2) = \epsilon_{T+2} + \Phi_1 [y_{T+1} - y_T(1)] = \epsilon_{T+2} + \Phi_1 \epsilon_{T+1} \]

The \( \ell \)-step ahead forecast \( y_T(\ell) \) converges to the mean vector \( \mu \) as the forecast horizon \( \ell \) increases.

The covariance matrices are

- for \( e_T(1) \) \( \Sigma \),
- for \( e_T(2) \) \( \Sigma + \Phi_1 \Sigma \Phi_1' \).

The covariance matrix of the \( \ell \)-step ahead forecast error converges to the covariance matrix of \( y_t \) as \( \ell \) increases.
Impulse responses
Impulse responses

Analogous to the Wold representation of univariate weakly stationary process the VAR($p$) can be written in terms of concurrent and past innovations.

$$y_t = \mu + \epsilon_t + \Psi_1 \epsilon_{t-1} + \Psi_2 \epsilon_{t-2} + \ldots$$

This moving average representation makes the impact of an innovation $\epsilon_{t-i}$ at $(t - i)$ on $y_t$ at $t$ explicit. It is $\Psi_i$.

This impact is the same as of $\epsilon_t$ on $y_{t+i}$.

${\Psi}_i$ is called **impulse response function**, IRF.
Impulse responses

Since the components of $\epsilon_t$ are correlated the separating out of the effects of single innovations is difficult. Thus we consider

$$y_t = \mu + \eta_t + \Psi^*_1 \eta_{t-1} + \Psi^*_2 \eta_{t-2} + \ldots$$

where the components of $\eta_t$ are uncorrelated.

$$\Psi^*_j \eta_{t-j} = [\Psi_j L] \eta_{t-j} = \Psi_j [L \eta_{t-j}] = \Psi_j \epsilon_{t-j}$$

$L$ is the Cholesky factor of $\Sigma$: $\Sigma = LDL'$ and $\epsilon_t = L\eta_t$

The diagonal matrix $D$ is the covariance matrix of $\eta_t$. Cov($\eta_t, \eta_t$) = $D$ and Cov($\epsilon_t, \epsilon_t$) = $\Sigma$.

Then the covariances of $L\eta_{t-j}$ are those of $\epsilon_{t-J}$.

$$E(\epsilon_t \epsilon_t') = E(L\eta_t)(L\eta_t)' = E(L(\eta_t \eta_t')L') = LDL' = \Sigma$$
Impulse responses

Element $\psi^*_ij(\ell)$ of $\Psi^*_\ell$ is the impact of innovation $\eta_{j,t}$ of size 1 on $y_{i,t+\ell}$.

In practice instead of size 1, the standard deviation of the shock is used. Then the diagonal elements of $D$ are transformed to 1:

$$LDL' = \tilde{L}\tilde{L}$$

and $\tilde{\Psi} = \Psi\tilde{L}$ is used instead of $\Psi^*$.

Transformation from the representation of $y$ in $\epsilon$ to the representation in $\eta$ is a trial to infer from the errors of the reduced form to the structural errors. However, they are not unique in general.

The impulse responses will change when the Cholesky algorithm does not start with element $(1,1)$ of $\Sigma$ - as is commonly done - but e.g. with $(k, k)$. 
Exercise and References
Exercises

Choose at least 2 of (1, 2, 3, 4) and 2 of (5G, 6G, 7G, 8G).

1 The univariate AR(2) process
   (a) $y_t = 0.1y_{t-1} + 0.2y_{t-2} + \epsilon_t$
   (b) $y_t = 1 + 0.1y_{t-1} + 0.2y_{t-2} + \epsilon_t$
   is given.
   Write the process as a VAR(1) using the companion matrix. Show that both
   univariate and the multivariate representations are equivalent.

2 Series of length $T = 100$ are available.
   (a) $k = 5$, (b) $k = 20$
   variables should be modeled as VAR($p$).
   What is the maximal order $p$, which can be used in this VAR?
Exercises

3. Give
   (a) the 5-step ahead forecast of a VAR(1), \( y_T(5) \), and
   (b) its covariance matrix.

4. Show that the roots of \(|I - \Phi z| = 0\) are \((1/\lambda_\Phi)\), the inverse eigenvalues of \(\Phi\),
   for \(\lambda \neq 0\).

5G. Choose a set of returns and test for
   (a) univariate, (b) multivariate
   dynamic dependencies with the Ljung-Box test. Use
   Ex2_5_LjungBox_R.txt.
Exercises

6G Consider returns of various stock indices before and after 9/11 2001 and interpret the dynamic dependencies in low order VARs in standard form. Choose 2 stock indices first and use $p = 1, 2$. Then extend the analysis to three dimensions. Interpret the results wrt unidirectional, feedback and coupled relationships. Use Ex2_6_DynDep_R.txt.
7G Find the best fitting model for the dynamic dependencies of a set of stock returns wrt the AIC or SBC. Test the residuals for dynamic dependencies. See 6G and Ex2_7_VARselection_R.txt.

8G Choose a data set and a suitable VAR. Calculate
(a) 1- to 10-step ahead forecasts with forecast intervals,
(b) try to interpret the impulse response function,
(c) interpret the composition of the forecast error variances.
Use Ex2_8_VARfc_R.txt.
References

Tsay 8.1-2
Christopher A. Sims(1980), Macroeconomics and Reality, *Econometrica* 48