Interpretating Macroeconomic Time Series

Sciences-Po M2 2009-2010

Benoit Mojon (Banque de France)

List of lectures
Lecture 1 and 2: Overview and Filtering
Lecture 2 and 3: RBC calibration
Lecture 3 and 4: Toward the standard DSGE model
Lecture 4 and 5 : VAR estimation
Lecture 5 : VAR identification
Lecture 6 : Maximum likelihood estimation of a DSGE

1 Estimation and identification of VAR models

Outline

1. **Introduction**
2. **Estimation of VAR models**
3. The MA representation of a VAR
4. **Identification**

Reading list

Hamilton chapter 9, 10 and 12 (ask Aurélien Poissonier where available on the net)
Leeper, Sims and Zha (Brookings-1996)
Gali (AER-1999)

Exercise for December 14

Compare the IRFs of a monetary policy shock in a VAR of the US with various identification schemes:
Choleski, CEE recursive identification
Sims information based identification
Blanchard and Quah Long term restrictions

1.1 Introduction

1.1.1 Shortcomings of structural models

- What if the theory is wrong?
- Difficulty of estimation

1.1.2 Alternative: reduced form models (including VARs)

Remember the general form of time series models

$$Y_t = f(Y_t, Y_{t-1}, ..., X_t, X_{t-1}, ..., \Theta, \varepsilon_t)$$

VARs can be described as

$$Y_t = c + \Phi(L)Y_t + \varepsilon_t$$

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ... + \Phi_p Y_{t-p} + \varepsilon_t$$

with

$$\begin{aligned} \varepsilon_t &\sim N(0, \Omega) \\ E_t(\varepsilon_{t+k}) &= 0, k > 0 \end{aligned}$$

Main advantages of VARs

1. simplicity
2. consistency with rational expectations!
3. they have a probabilistic structure
4. they are widely used to forecast, simulate the impact of shocks

Main drawbacks of VARs

1. black box
2. idendification "as" problematic as standard models
3. degrees of freedom

Examples:

- CEE (JPE-2005): description of "conditional forecast/moments". What are the "objective/true" effects of monetary policy according to a "reduced form theory free" dynamic model?

$$Y_t = (GDP_t, Cons_t, Inv_t, Hours_t, Money_t, R_t, ...)$$

However, in CEE, the identification restrictions require sticky prices; Uhlig (and others) show that prices may very well adjust rapidly to monetary policy shocks

- Blanchard-Quah (AER-1989) : describe the effects of demand and supply shocks, using "long-term" identification restrictions

$$Y_t = (GDP_t, Unemployment_t)$$

- Gali (AER-1999): show that hours do not increase following a productivity shock: this turns RBC theory on its head

1.1.3 Estimation of VARs

Hamilton chapter 11, section 1

Assume the vector of variables

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ... + \Phi_p Y_{t-p} + \varepsilon_t$$

with

$$\begin{aligned} \varepsilon_t &\sim N(0, \Omega) \\ E_t(\varepsilon_{t+k}) &= 0, k > 0 \end{aligned}$$

How can we define the likelihood of this process for a sample of T observations?

Let's take the first observation of Y_t, i.e. $Y_{-p}, Y_{-p+1}, ..., Y_0$ as given and define the joint density of the sample of observations Y_1 to Y_T.

$$f_{Y_T,...,Y_1|Y_{-p},Y_{-p+1},....,Y_0}(Y_T, ..., Y_1 \mid Y_{-p}, Y_{-p+1}, ..., Y_{-p+p-1}, Y_0; \Theta) \quad (1)$$

what are the value of the models parameters $\Theta \equiv \Phi_1, \Phi_2, ..., \Phi_p$ that maximise this likelihood function?

$$\begin{aligned} \varepsilon_t &\sim N(0, \Omega) \text{ hence} \quad (2) \\ Y_t &\mid Y_{t-p}, Y_{t-p+1}, ..., Y_{t-1} \sim N(c + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ... + \Phi_p Y_{t-p}, \Omega) \end{aligned}$$

Let's use the following notation:

$$X_t = \begin{bmatrix} 1 \\ Y_{t-1} \\ Y_{t-2} \\ : \\ Y_{t-p} \end{bmatrix}$$

X_t is a $(np + 1) \times 1$ matrix

$$\Pi' = \left[\begin{array}{ccccc} c & \Phi_1 & \Phi_2 & ... & \Phi_p \end{array} \right]$$

$\Pi'$is a $n \times (np+1)$ matrix
we can rewrite (2) as

$$Y_t \mid Y_{t-p}, Y_{t-p+1}, ..., Y_{t-1} \sim N(\Pi' X_t, \Omega)$$

the conditional density of observation Y_t can be written as

$$\begin{aligned} f_{Y_t|Y_{t-1},Y_{t-2},....,Y_{t-p}}(Y_t \quad | \quad & Y_{t-1}, Y_{t-2}, ..., Y_{t-p}; \Theta) = \\ & (2\pi)^{-n/2} \left|\Omega^{-1}\right|^{1/2} \exp\left[-\frac{1}{2}(Y_t - \Pi' X_t)' \Omega^{-1} (Y_t - \Pi' X_t)\right] \end{aligned}$$

considering now the likelihood of the full sample:
(1) can be written as

$$\begin{aligned} f_{Y_T,...,Y_1|Y_{t-p},Y_{t-p+1},....,Y_0}(Y_T, ..., Y_1 \quad & | \quad Y_{-p}, Y_{-p+1}, ..., Y_0; \Theta) = \\ f_{Y_{T-1},...,Y_1|Y_{t-p},Y_{t-p+1},....,Y_0}(Y_{T-1}, ..., Y_1 \quad & | \quad Y_{-p}, Y_{-p+1}, ..., Y_0; \Theta) \times \\ f_{Y_t|Y_{t-1},Y_{t-2},....,Y_{t-p}}(Y_t \quad & | \quad Y_{t-1}, Y_{t-2}, ..., Y_{t-p}; \Theta) \end{aligned}$$

Applying this decomposition of the likelihood recursively, we get

$$\begin{aligned} f_{Y_T,...,Y_1|Y_{t-p},Y_{t-p+1},....,Y_0}(Y_T, ..., Y_1 \quad & | \quad Y_{-p}, Y_{-p+1}, ..., Y_0; \Theta) = \\ \prod_{t=1}^{T} f_{Y_t|Y_{t-1},Y_{t-2},....,Y_{t-p}}(Y_t \quad & | \quad Y_{t-1}, Y_{t-2}, ..., Y_{t-p}; \Theta) \end{aligned}$$

in logs (because sums are much easier to handle than products)

$$\begin{aligned} L(\Theta) \quad = \quad & \sum_{t=1}^{T} \log(f_{Y_t|Y_{t-1},Y_{t-2},....,Y_{t-p}}(Y_t \mid Y_{t-1}, Y_{t-2}, ..., Y_{t-p}; \Theta)) = \\ & -Tn/2 \log(2\pi) + \frac{T}{2} \log\left|\Omega^{-1}\right| - \frac{1}{2} \sum_{t=1}^{T} \left[(Y_t - \Pi' X_t)' \Omega^{-1} (Y_t - \Pi' X_t)\right] \end{aligned}$$

We can now maximise this (log) likelihood with respect to the parameter vector Θ (that are all contained in Π') in the above expression.

One result that makes VAR so attractive (and so much used) is that the MLE (max likelihood estimator of Π' is the OLS estimator of individual equations of the model!

$$\Pi' = \left[\sum_{t=1}^{T} Y_t X_t'\right] \left[\sum_{t=1}^{T} X_t X_t'\right]^{-1}$$

$$\hat{\pi}_j' = \left[\sum_{t=1}^{T} y_{j,t}X_t'\right]\left[\sum_{t=1}^{T} X_tX_t'\right]^{-1}$$

that is the coefficients of equation j are given by the OLS regression of variable y_j on X.

Proof. rewrite the last term of the log likelihood as

$$\begin{aligned}
& \sum_{t=1}^{T}\left[(Y_t - \Pi' X_t)'\Omega^{-1}(Y_t - \Pi' X_t)\right] \\
= & \sum_{t=1}^{T}\left[(Y_t - \hat{\Pi}' X_t + \hat{\Pi}' X_t - \Pi' X_t)'\Omega^{-1}(Y_t - \hat{\Pi}' X_t + \hat{\Pi}' X_t - \Pi' X_t)\right] \\
= & \sum_{t=1}^{T}\left[(\hat{\varepsilon}_t + (\hat{\Pi} - \Pi)X_t)'\Omega^{-1}(\hat{\varepsilon}_t + (\hat{\Pi} - \Pi)'X_t)\right] \\
= & \sum_{t=1}^{T}\left[\hat{\varepsilon}_t'\Omega^{-1}\hat{\varepsilon}_t\right] + 2\sum_{t=1}^{T}\hat{\varepsilon}_t'\Omega^{-1}(\hat{\Pi} - \Pi)X_t + \sum_{t=1}^{T} X_t'(\hat{\Pi} - \Pi)\Omega^{-1}(\hat{\Pi} - \Pi)'X_t
\end{aligned}$$

where $\hat{\varepsilon}_t$ are the sample residuals of the OLS regressions.

$$\sum_{t=1}^{T}\hat{\varepsilon}_t'\Omega^{-1}(\hat{\Pi} - \Pi)X_t = trace\left[\Omega^{-1}(\hat{\Pi} - \Pi)\sum_{t=1}^{T} X_t\hat{\varepsilon}_t'\right] = 0$$

because OLS residuals are, by construction, orthogonal to regressors X_t.

We are now aiming at minimising the following quadratic form

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

which, beause Ω (and hence Ω^{-1}) are positive semi-definite matrices, admits a minimum when $\hat{\Pi} = \Pi$.

CQFD ■

Also note that the MLE of Ω is $\hat{\Omega} = 1/T\sum_{t=1}^{T}\left[\hat{\varepsilon}_t\hat{\varepsilon}_t'\right]$. See Hamilton page 295 and 296.

1.1.4 The MA representation of a VARs

Let's consider a simple NK inspired 3 variable economy VAR of order 2

$$\begin{pmatrix} dy_t \\ dp_t \\ r_t \end{pmatrix} = \begin{bmatrix} a & b & c \\ d & f & g \\ h & i & j \end{bmatrix}\begin{pmatrix} dy_{t-1} \\ dp_{t-1} \\ r_{t-1} \end{pmatrix} + \begin{bmatrix} k & l & m \\ n & o & p \\ q & s & w \end{bmatrix}\begin{pmatrix} dy_{t-2} \\ dp_{t-2} \\ r_{t-2} \end{pmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{bmatrix}$$

in matrix form we have

$$Y_t = A_1 Y_t + A_2 Y_t + \varepsilon_t$$

we can actually rewrite this VAR of order 2 as a VAR of order one as follows:

$$Z_t = \begin{bmatrix} Y_t \\ Y_{t-1} \end{bmatrix} = \begin{bmatrix} dy_t \\ dp_t \\ r_t \\ dy_{t-1} \\ dp_{t-1} \\ r_{t-1} \end{bmatrix} \text{ and } \zeta_t = \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix} = \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$Z_t = BZ_{t-1} + \zeta_t$$

with

$$B = \begin{bmatrix} a & b & c & k & l & m \\ d & f & g & n & o & p \\ h & i & j & q & s & w \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

We can now more easily invert the "AR" form of the cannonic form of the VAR:

$$\begin{aligned} Z_t &= BZ_{t-1} + \zeta_t \\ Z_t &= (I - B(L))^{-1} \zeta_t \\ Z_t &= \sum_{j=0}^{\infty} C_j \zeta_{t-j} \end{aligned}$$

Hence, the realisations of the variables can be represented as the dynamic effects of past realisation of the innovations ε_t. This will be directly used to represent estimates of the impulse response functions of a given variable (say inflation or the interest rate) to an identified shock, such as a monetary policy shock, a "demand" or a "supply" shock.

The **IRFs** of variable m to shock k be given by the sequences of a element m,k of the C_j matrices:

$$c_{m,k,j} \text{ for } j = 0, ..., h$$

Likewise, provided that we can transform the innovations ε_t into economically meaningfull "shocks" u_t that we construct as orthogonal to one another, we can use the MA representation to decompose the variance of economic variables the various economic shocks that we have included in the model.

$$
\begin{aligned}
Z_t &= \sum_{j=0}^{\infty} C_j P \zeta_{t-j} \\
P\zeta_{t-j} &\sim N(0, D), \text{ with } D \text{ a diagonal matrix}
\end{aligned}
$$