ECON 4160, Spring term 2013. Lecture 4

Estimation theory for VARs and derived models. Multipliers. Granger causality (concept). Typology.

Ragnar Nymoen

University of Oslo

12 September 2013

Some references to Lecture 4

- Ch. 13.4-13.5 and 13.7 in DM,
- Lecture note 3 and 4 about VARs and ADL models.

Theoretical moments of AR(1) I

- The simplest VAR is the univariate and stationary $AR(1)$ model (here in the Gaussian version with normal disturbances)

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t, \quad |\phi_1| < 1,\ \varepsilon_t \sim IN\left(0, \sigma_\varepsilon^2\right)\ \forall t \quad (1)$$

- Obtain the solution conditional on Y_0 by repeated backward solution

$$Y_t = \phi_0 \sum_{i=0}^{t-1} \phi_1^i + \phi_1^t Y_0 + \sum_{i=0}^{t-1} \phi_1^i \varepsilon_{t-i} \quad (2)$$

Theoretical moments of AR(1) II

- The **conditional** expectation and variance become

$$E(Y_t \mid Y_0) = E(\phi_0 \sum_{i=0}^{t-1} \phi_1^i + \phi_1^t Y_0)$$
$$= \phi_0 \frac{1-\phi_1^t}{1-\phi_1} + \phi_1^t Y_0 \quad (3)$$

$$Var(Y_t \mid Y_0) = \sigma_\varepsilon^2 \frac{1-\phi_1^{2t}}{1-\phi_1^2} \quad (4)$$

The unconditional expectation and variance can be found by setting $t \longrightarrow \infty$ in these two expressions, or by calculating $E(Y_t)$ and $Var(Y_t)$ from the unconditional solution:

$$Y_t = \phi_0 \sum_{i=0}^{\infty} \phi_1^i + \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i} \quad (5)$$

Theoretical moments of AR(1) III

$$E(Y_t) = \frac{\phi_0}{1-\phi_1} = \mu \qquad (6)$$

$$Var(Y_t) = \frac{\sigma_\varepsilon^2}{1-\phi_1^2} = \tau_0 \qquad (7)$$

- What about the autocovariance function? Note that we can re-write

  $$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t, \quad |\phi_1| < 1$$

  as

  $$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \varepsilon_t, \quad |\phi_1| < 1 \qquad (8)$$

  by adding and subtracting $(\mu - \phi_1\mu)$.

Theoretical moments of AR(1) IV

- Then the first autocovariance

$$\tau_1 = E[(Y_t - \mu)(Y_{t-1} - \mu)]$$

is found by multiplying (8) by $(Y_{t-1} - \mu)$ and then taking the expectation:

$$\tau_1 = E[(Y_t - \mu)(Y_{t-1} - \mu)] = E\left[\phi_1(Y_{t-1} - \mu)^2 + \varepsilon_t(Y_{t-1} - \mu)\right]$$
$$= \phi_1 \tau_0$$

- The second autocovariance τ_2 is

$$\begin{aligned}\tau_2 &= E\left[(Y_t - \mu)(Y_{t-2} - \mu)\right] \\ &= E\left[\phi_1(Y_{t-1} - \mu)(Y_{t-2} - \mu) + \varepsilon_t(Y_{t-2} - \mu)\right] \\ &= \phi_1 \tau_1 = \phi_1^2 \tau_0\end{aligned}$$

Theoretical moments of AR(1) V

- This shows that the autocovariance function τ_j $(j = 0, 1, 2, \ldots)$ follows the **same dynamics as the variable** Y_t itself.
- The same is true for the autocorrelation function (ACF)

$$\zeta_j = \frac{\tau_j}{\tau_0} = \phi_1^j \text{ for } j = 0, 1, 2, ... \quad (9)$$

- This generalizes to $AR(2)$ and $AR(p)$, but we skip the exact expressions here. ECON 5101 stuff.

ML estimation of AR(1) I

- The following is a self-contained argument for why OLS estimation of AR(1) results in conditional ML estimators that are good approximations to exact MLEs if the time series is long enough.
- In particular we avoid the references to the *Kalman filter*, mentioned by DM, which is important, but beyond the scope of this course, and it is not need to understand the relationship between MLE and OLS estimation

ML estimation of AR(1) II

- For the model AR(1) (1) the unconditional pdf for Y_1 is:

$$f(Y_1) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_\varepsilon^2/(1-\phi_1^2)}} exp\left(\frac{-[Y_1 - \phi_0/(1-\phi_1)]^2}{2\sigma_\varepsilon^2/(1-\phi_1^2)}\right) \quad (10)$$

The conditional distribution if Y_2 given Y_1 follows directly from (1)

$$(Y_2 \mid Y_1) \sim N(\phi_0 + \phi_1 Y_1, \sigma_\varepsilon^2),$$

with pdf

$$f(Y_2 \mid Y_1) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_\varepsilon^2}} exp\left(\frac{-[Y_2 - \phi_0 - \phi_1 Y_1]^2}{2\sigma_\varepsilon^2}\right). \quad (11)$$

ML estimation of AR(1) III

The simultaneous pdf of Y_2 and Y_1 is

$$f(Y_{2,} Y_1) = f(Y_2 \mid Y_1) \cdot f(Y_1)$$

$f(Y_3 \mid Y_{2,} Y_1)$ will be similar to (11) (Y_3 replaces Y_2; Y_2 replaces Y_1) and $f(Y_3, Y_2, Y_1)$ becomes

$$\begin{aligned} f(Y_3, Y_2, Y_1) &= f(Y_3 \mid Y_2) \cdot f(Y_2, Y_1) \\ &= f(Y_3 \mid Y_2) \cdot f(Y_2 \mid Y_1) \cdot f(Y_1). \end{aligned}$$

ML estimation of AR(1) IV

By induction the likelihood for the whole sample can be written as

$$f(Y_{T,} Y_{T-1}, \ldots, Y_1) = f(Y_1) \cdot \prod_{t=2}^{T} f(Y_t \mid Y_{t-1}). \tag{12}$$

MLE is found by maximisation of the log-likelihood function

$$L = \ln(f(Y_1)) - \frac{T}{2}(\ln(2\pi/\sigma_\varepsilon^2)) - \sum_{t=2}^{T} \frac{[Y_t - \phi_0 - \phi_1 Y_{t-1}]^2}{2\sigma_\varepsilon^2}. \tag{13}$$

ML estimation of AR(1) V

- If we consider the first term as known, we see that the ML estimators of ϕ_0 and ϕ_1 are found as the OLS estimators

$$\min_{\hat{\phi}_0,\hat{\phi}_1}\{\sum_{t=2}^{T}[Y_t-\hat{\phi}_0-\hat{\phi}_1 Y_{t-1}]^2\} \tag{14}$$

in the same way as for static regression model, see for Lecture 2.
- Conditioning on known Y_1 means that the OLS estimators of ϕ_0 and ϕ_1 in the AR(1) model are conditional MLE.
- Since Y_1 is without importance for the likelihood function when $T \longrightarrow \infty$, the practical interpretation is that OLS estimation gives a good approximation to the exact ML estimators when T is sufficiently large.

ML estimation of AR(1) VI

- The **OLS estimators are biased**, since Y_{t-1} is only pre-determined, not strictly exogenous (write (2) for Y_{t-1} if you need a reminder about this).
- The **OLS estimators are consistent,** because asymptotically, Y_{t-1} is uncorrelated with the sequence of future distubances. For $\phi_0 = 0$, we can write this as

$$\text{plim}\left(\widehat{\phi}_1 - \phi_1\right) = \frac{\text{plim}\,\frac{1}{T}\sum_{t=2}^{T} Y_{t-1}\varepsilon_t}{\text{plim}\,\frac{1}{T}\sum_{t=2}^{T} Y_{t-1}^2} = \frac{0}{\frac{\sigma_\varepsilon^2}{1-\phi_1^2}} = 0.$$

if $E\left(Y_{t-1}\varepsilon_t\right) = 0$ and $|\phi_1| < 1$ (**stationarity**), and with reference to the Law of large numbers and Slutsky's theorem

How large is "large"? I

- As usual, we should ask how relevant the asymptotic estimation theory is.
- By experimentation with Monte Carlo analysis of the AR(1) model in PcGive, you will get an impression about the importance of the values of ϕ_1 and σ_ε^2 in the DGP
- In particular, find that for $\phi_1 = 0.99$ ("almost non-stationary") the bias declines very slowly, but much faster already with $\phi_1 = 0.95$ and $\phi_1 = 0.80$ for example.
- This shows that with time series, the answer to "How large is large?" (or "How small is large?"); depends no only on T but also on ϕ_1 in particular.

Estimation of ARMA models I

- First: The condition about **invertibility** of MA processes has nothing to with the question about stationarity.
- In is a more technical requirement about the roots or the polynomial associated the MA part of the process: Invertible MA processes can be represented as an infinite AR process
- By a similar argument as we have given for AR(1), approximate MLE of ARMA(1,1) models are obtained by NLS estimation.
- In this course, focus on AR(p) and VAR(p).

The VAR estimation theorem I

- The results about OLS estimators of AR(1) being approximately MLE generalizes directly to AR(p) and to VAR(p).
- The condition is that the roots of the associated polynomials are on the stationary side of the unit-circle, and that the disturbances ("input series") are white–noise, as secured by a univariate (AR(p) case) or multivariate ($VAR(p)$) normal distribution.
- (The following states the same as on page 595-597 in DM.)

The VAR estimation theorem II

- Assume the $VAR(p)$ for the vector time series $\mathbf{y}_t = (Y_{1t}, Y_{1t}, \ldots, Y_{kt})' \quad t = 1, 2, \ldots, T$:

$$\mathbf{y}_t = \sum_{i=1}^{p} \mathbf{\Pi}_i \mathbf{y}_{t-i} + \mathbf{YD}_t + \boldsymbol{\epsilon}_t \tag{15}$$

  where $\mathbf{\Pi}_i$ $(i = 1, 2, \ldots, p)$ contains the AR coefficients, $\mathbf{YD}_t$ contains constant terms and other deterministic term, and $\boldsymbol{\epsilon}_t$ is multivariate Gaussian $\boldsymbol{\epsilon}_t \sim IN(\mathbf{0}, \mathbf{\Sigma})$.
- The following important results about estimation and inference in the VAR are true:

1. If y_t is stationary, the OLS estimators for $\mathbf{\Pi}_i$ $(i = 1, 2, \ldots, p)$, and $\mathbf{Y}$ are conditional MLEs.

The VAR estimation theorem III

2. These OLS estimators are consistent and asymptotically normally distributed.
3. If $\mathbf{\Sigma}$ is consistently estimated, hypotheses about single parameters can be tested by OLS t-ratios that have asymptotical standard normal distributions.
4. OLS estimated coefficients standard errors can be used to construct confidence intervals that are accurate in large samples
5. An LR tests for a joint hypothesis with r restrictions is $-2(L_R^* - L_U^*)$ and is asymptotically $\chi^2(r)$ under the null hypothesis the the r restrictions are valid.

- Remark to 3 and 4: Nothing wrong in using the t-distribution rather than $N(0, 1)$!

The VAR estimation theorem IV

- Remark to 5: As in standard regression: L_R^* and L_U^* are one-to-one with RSS_R and RSS_U, meaning that they can be used to construct F-versions of the joint test.

Consequence for ADL estimation I

Since an ADL model is a conditional model based on a VAR, all the results 1.- 5. from the VAR estimation theorem carry over to estimation of the coefficients of single equation ADL models and to hypothesis testing and confidence intervals.

- The delta method can be used to make inference about non-linear derived parameters.
- Since there are many models that are special cases of the ADL model, the estimation theorems also for these models

Multipliers and Granger Causality I

- In line with Lecture note 4 we can write the ADL model with k regressors as

$$\phi(L)Y_t = \phi_0 + \sum_{j=1}^{k} \beta_j(L)X_{jt} + \varepsilon_t \tag{16}$$

where

$$\phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i \tag{17}$$

$$\beta_j(L) = \sum_{i=0}^{p} \beta_{ji} L^i,\ j = 1, 2, \ldots k \tag{18}$$

and $\varepsilon_t \sim IN(0, \sigma_\varepsilon^2 M)$.

Multipliers and Granger Causality II

- The dynamic response of Y_t to temporal and/or permanent shifts in one of the X variables are called dynamic multipliers.
- The period t responses to shocks in period t are simply the regression coefficients β_{j0} $(j = 1, 2, \ldots, k)$ and are often called called the **impact multipliers**
- The long-run multipliers are the partial derivatives of the static solution for $Y_t = Y^*$ and $X_t = X^*$

$$\phi(1)Y = \phi_0 + \sum_{j=0}^{k} \beta_j(1)X_{jt}$$

(note that the disturbance is omitted by convention), i.e.

$$\frac{\partial Y^*}{\partial X_j^*} = \frac{\sum_{i=0}^{p} \beta_{ji}}{1 - \sum_{i=1}^{p} \phi_i} \tag{19}$$

Multipliers and Granger Causality III

- After estimation in PcGive, (19) is obtained from **Test-Dynamic Analysis-Static long-run solution**
- The long-run multipliers are given with delta-method type standard errors.
- Clearly, (19) only has good meaning if

$$X_{t-i} \rightarrow Y_t \text{ for } i = 0, 1, 2, \ldots \text{and } Y_{t-j} \nrightarrow X_t \text{ for } j = 1, 2, \ldots \tag{20}$$

which is called one-way **Granger-causality** from X to Y.

Multipliers and Granger Causality IV

- We can also be interested in the dynamic multipliers:

$$\frac{\partial Y_t}{\partial X_{t-j}} \text{ for } j = 1, 2,$$

  and the **cumulation** of these, which show how the effects of a permanent change in X build up from the impact multiplier to the long-run multiplier.
- In PcGive they are given in normalized form, divided by the corresponding long-run multiplier, under the name Cumulative normalized lag-weights.

ADL based typology I

- Many relevant models are either *re-parameterizations* or special cases of (16).
- We briefly review the most useful in class