Cointegration Lecture II: Estimating the CVAR

Julia Giese
Nuffield College
julia.giese@economics.ox.ac.uk

Hilary Term 2008

Recap

Last week:

- Looked at *estimation* and *specification* of *unrestricted* model.
- Discussed properties of Π when data is *non-stationary*.
- Learned about how *deterministic components* may enter the VAR.

This week:

- Discuss *estimation* under the *reduced rank* hypothesis.
- Consider tests for *rank determination*.
- Briefly introduce tests to check for *parameter constancy*.

Outline

Likelihood analysis of rank hypothesis

Model:

$$\Delta \mathbf{x}_t = \Pi \mathbf{x}_{t-1} + \Gamma_1 \Delta \mathbf{x}_{t-1} + \ldots + \Gamma_{k-1} \Delta \mathbf{x}_{t-k+1} + \phi \mathbf{D}_t + \varepsilon_t$$

Usual multivariate regression in *unrestricted* estimation. But what if hypothesis $\mathcal{H}_r$: $\text{rank}(\Pi) \leq r \leq p$ and hence $\Pi = \alpha\beta'$?

$\Rightarrow$ Regression problem with *non-linear restriction* on parameters.

Maximise likelihood by

- Numerical procedure: Newton-Raphson.
- Analytic procedure: use linear algebra of symmetric matrices.

Derivation of ML estimator[1]

For simplicity, consider a VAR(1) with no deterministic terms:

$$\underbrace{\Delta \mathbf{x}_t}_{\mathbf{R}_{0t}} = \alpha\beta' \underbrace{\mathbf{x}_{t-1}}_{\mathbf{R}_{1t}} + \varepsilon_t \quad \text{with} \quad \varepsilon_t \sim \text{iid } \mathrm{N}_p(\mathbf{0}, \Omega)$$

For fixed β, obtaining estimates for α and Ω is a regression problem:

$$\begin{aligned} \hat{\alpha}(\beta) &= \mathbf{S}_{01}\beta(\beta'\mathbf{S}_{11}\beta)^{-1} \\ \hat{\Omega}(\beta) &= \mathbf{S}_{00} - \mathbf{S}_{01}\beta(\beta'\mathbf{S}_{11}\beta)^{-1}\beta'\mathbf{S}_{10} \end{aligned}$$

where $\mathbf{S}_{ij} = T^{-1}\sum_t \mathbf{R}_{it}\mathbf{R}'_{jt}$.

Note that the estimator for $\Omega(\beta)$ is found by ML maximisation (and substituting for α).

[1] See Juselius (2006), Ch. 7.2.

Substituting both estimators into the log-likelihood function gives

$$\ln L(\beta) = -\frac{T}{2} \ln |\hat{\Omega}(\beta)| \; - \text{constant terms.} \tag{1}$$

Maximising (1) implies minimising $|\hat{\Omega}(\beta)|$.

Since it can be shown that

$$\begin{vmatrix} A & B \\ B' & C \end{vmatrix} = |A| \cdot |C - B'A^{-1}B| = |C| \cdot |A - BC^{-1}B'|$$

we can express the determinant of $\hat{\Omega}(\beta)$ as

$$\begin{aligned} |\hat{\Omega}(\beta)| &= |\mathbf{S}_{00} - \mathbf{S}_{01}\beta(\beta'\mathbf{S}_{11}\beta)^{-1}\beta'\mathbf{S}_{10}| \\ &= |\mathbf{S}_{00}| \cdot \frac{|\beta'(\mathbf{S}_{11} - \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01})\beta|}{|\beta'\mathbf{S}_{11}\beta|} \end{aligned}$$

Another result states that the function

$$f(x) = \frac{|Y'MY|}{|Y'NY|} = \frac{|\beta'(\mathbf{S}_{11} - \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01})\beta|}{|\beta'\mathbf{S}_{11}\beta|}$$

is minimised by solving the eigenvalue problem

$$|\rho N - M| = |\rho\mathbf{S}_{11} - \mathbf{S}_{11} + \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01}| = |\lambda\mathbf{S}_{11} - \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01}| = 0$$

where $\lambda = 1 - \rho$ for eigenvalues $1 \geq \hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_p \geq 0$ and corresponding eigenvectors $\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_p$, such that

$$\mathbf{S}_{11}\hat{\mathbf{v}}_i\hat{\lambda}_i = \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01}\hat{\mathbf{v}}_i$$

and

$$\hat{\mathbf{v}}_i'\mathbf{S}_{11}\hat{\mathbf{v}}_i = 1 \quad \text{and} \quad \hat{\mathbf{v}}_i'\mathbf{S}_{11}\hat{\mathbf{v}}_j = 0 \text{ for } i \neq j.$$

This gives $\hat{\beta} = (\hat{\mathbf{v}}_1, ...\hat{\mathbf{v}}_p)$, and

$$\hat{\lambda}_i = \hat{\beta}_i' \mathbf{S}_{10} \mathbf{S}_{00}^{-1} \mathbf{S}_{01} \hat{\beta}_i.$$

Now

$$\begin{aligned} |\hat{\Omega}(\hat{\beta})| &= |\mathbf{S}_{00}| \cdot \frac{|\hat{\beta}'(\mathbf{S}_{11} - \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01})\hat{\beta}|}{|\hat{\beta}'\mathbf{S}_{11}\hat{\beta}|} \\ &= |\mathbf{S}_{00}| \prod_{i=1}^{p} (1 - \hat{\lambda}_i) \end{aligned}$$

and

$$-2 \ln L_{\max} = T \ln \left(|\mathbf{S}_{00}| \prod_{i=1}^{p} (1 - \hat{\lambda}_i) \right).$$

Outline

Different ways of choosing the rank

- Most formal: rank test, but can have low power in small samples (Bartlett factor used to correct for this).[2]

Ch. 8.5 and 8.6 in Juselius (2006) discuss how to take additional information into account when choosing a rank:

- Recursive version of rank test.
- Absolute size of roots.
- Stationarity properties of unrestricted cointegrating relations.
- Information in columns of unrestricted α.
- Economic interpretability of results.

[2] See Johansen (2002).

The rank test[3]

Want to test a model with rank $r \leq p$ against alternative of full rank:

$\mathcal{H}_p$: rank $= p$, i.e. no unit roots; $\mathbf{x}_t$ is stationary.
$\mathcal{H}_r$: rank $= r$, i.e. $p - r$ unit roots, r cointegrating relations; $\mathbf{x}_t$ is non-stationary. Assumptions from before remain valid.

The LR test statistic for the rank (also called *trace* test) is found as:

$$\begin{aligned} LR(\mathcal{H}_r|\mathcal{H}_p) &= -2\ln Q(\mathcal{H}_r/\mathcal{H}_p) \\ &= T\ln\left\{\frac{|\mathbf{S}_{00}|(1-\hat{\lambda}_1)(1-\hat{\lambda}_2)\ldots(1-\hat{\lambda}_r)}{|\mathbf{S}_{00}|(1-\hat{\lambda}_1)(1-\hat{\lambda}_2)\ldots(1-\hat{\lambda}_r)\ldots(1-\hat{\lambda}_p)}\right\} \\ &= -T\sum_{j=r+1}^{p}\ln(1-\hat{\lambda}_j) \end{aligned}$$

[3] See Juselius (2006), Ch. 8.1.

Asymptotic distribution of rank test

Theorem

Johansen (1996), Theorem 6.1. *Assumptions A, B. Then*

$$LR\left(\mathcal{H}_r|\mathcal{H}_p\right) \xrightarrow{\mathsf{D}} \mathsf{DF}_z\left(p-r\right),$$

where $\mathsf{DF}_z\left(n\right)$ *is a Dickey-Fuller type distribution, see Johansen (1996), Table 15.1. Assumption C not necessary, see Nielsen (2000).*

The $\mathsf{DF}_z\left(n\right)$ distribution can be represented as

$$\operatorname{tr}\left\{\int_0^1 (dB)B'\left[\int_0^1 BB'du\right]^{-1}\int_0^1 B(dB)'\right\}$$

where B is an n-dimensional standard Brownian motion, $B=\Omega^{-\frac{1}{2}}W$.[4]

[4] See Johansen (1996), Ch. 11.

‘Top-to-bottom’ rank testing

$$\underbrace{(\text{rank}\Pi \leq 0)}_{\mathcal{H}_0} \subset \cdots \subset \underbrace{(\text{rank}\Pi \leq r)}_{\mathcal{H}_r} \subset \cdots \subset \underbrace{(\text{rank}\Pi \leq p)}_{\mathcal{H}_p}$$

Estimate r by first test which is accepted (‘*top-to-bottom*’ procedure):

$$\mathcal{H}_0|\mathcal{H}_p, \quad \mathcal{H}_1|\mathcal{H}_p, \quad \mathcal{H}_2|\mathcal{H}_p, \ldots$$

Estimator $\hat{r}$ is consistent in the sense:

$$\mathsf{P}_r\left(\hat{r} = i\right) \to \begin{cases} 0 & \text{if } i < r, \\ 95\% & \text{if } i = r, \\ \leq 5\% & \text{if } i \geq r. \end{cases}$$

Some output from CATS

Rank trace test for 5-dimensional VAR(2) of global money market:

p-r	r	Eig.Value	Trace	Trace*	Frac95	P-Value	P-Value*
5	0	0.383	121.752	105.255	88.554	0.000	0.002
4	1	0.339	76.792	66.591	63.659	0.002	0.027
3	2	0.225	38.343	33.547	42.770	0.133	0.315
2	3	0.123	14.688	11.760	25.731	0.607	0.826
1	4	0.026	2.479	1.838	12.448	0.918	0.964

Based on the results above, one would either choose a rank of 2. Columns with a * denote test statistics and p-values corrected by the Bartlett factor for small sample size. To determine the rank conclusively, one also needs to check other indicators of rank.

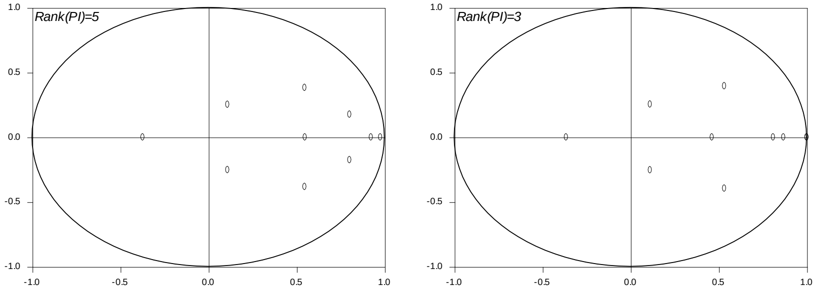

Figure: The roots of the companion matrix for the global money market model, $r = 5$ and $r = 3$.

Largest 3 roots:

- For $r = 5$: 0.976, 0.923, 0.820.
- For $r = 4$: 1.000, 0.949, 0.807.
- For $r = 3$: 1.000, 1.000, 0.868.

Rank test: Including constant level[5]

$$\text{Model} \quad : \quad \underbrace{\Delta \mathbf{x}_t}_{\mathbf{R}_{0t}} = (\Pi, \Pi_c) \underbrace{\begin{pmatrix} \mathbf{x}_{t-1} \\ 1 \end{pmatrix}}_{\mathbf{R}_{1t}} + \varepsilon_t,$$

$$\text{Hypothesis} \quad : \quad \mathcal{H}_r^c : \text{rank}\,(\Pi, \Pi_c) \leq r \quad \text{or} \quad (\Pi, \Pi_c) = \alpha(\beta', \beta_0).$$

Theorem

Johansen (1996), Theorem 6.2. *Assumptions A and B. Then*

$$LR(\mathcal{H}_r^c | \mathcal{H}_p^c) \xrightarrow{\mathsf{D}} \mathsf{DF}_c\,(p-r)\,,$$

where $\mathsf{DF}_c\,(n)$ *tabulated by Johansen (1996), Table 15.2.*

Estimate r by first test which is accepted:

$$(\mathcal{H}_0^c | \mathcal{H}_p^c), \quad (\mathcal{H}_1^c | \mathcal{H}_p^c), \quad (\mathcal{H}_2^c | \mathcal{H}_p^c), \ldots$$

[5] This choice of hypothesis is advocated by Nielsen and Rahbek (2000).

Rank test: Including linear trend

$$\begin{aligned}
\text{Model} &: \quad \Delta \mathbf{x}_t = (\Pi, \Pi_l) \begin{pmatrix} \mathbf{x}_{t-1} \\ t \end{pmatrix} + \mu_c + \varepsilon_t, \\
\text{Hypothesis} &: \quad \mathcal{H}_r^l : \text{rank}\,(\Pi, \Pi_l) \leq r \quad \text{or} \quad (\Pi, \Pi_l) = \alpha(\beta', \beta_1), \\
\text{“Residuals”} &: \quad \mathbf{R}_{0t} = (\Delta X_t |\, 1)\,, \mathbf{R}_{1t} = \left(\begin{matrix} \mathbf{x}_{t-1} \\ t \end{matrix} \middle| \, 1 \right).
\end{aligned}$$

Theorem

Johansen (1996), Theorem 6.3. *Assumptions A and B. Then*

$$LR(\mathcal{H}_r^l | \mathcal{H}_p^l) \xrightarrow{\mathsf{D}} \mathsf{DF}_l\,(p-r)\,,$$

where $\mathsf{DF}_l\,(n)$ *tabulated by Johansen (1996), Table 15.4.*

Estimate r by first test which is accepted:

$$(\mathcal{H}_0^l | \mathcal{H}_p^l), \quad (\mathcal{H}_1^l | \mathcal{H}_p^l), \quad (\mathcal{H}_2^l | \mathcal{H}_p^l), \ldots$$

Outline

Estimation in the CVAR

Determination of rank

Recursive tests

Appendix: Algebra of symmetric matrices

Recursive tests[6]

Once we have specified a model and settled on a rank, we need to check that the coefficients are constant over the sample:

- Use recursive tests where the sample size is gradually expanded with each estimation.
- Tests can start at the beginning (forward recursive tests) or at the end (backward recursive tests) of the sample.
- Tests are easiest to evaluate in graphical form, where the x-axis is a timeline.

[6] See Juselius (2006), Ch. 9.

The recursive tests can be grouped into four broad categories:

1. Recursive tests of the *full model*: For example, tests for the constancy of the *likelihood function*.
2. Recursive tests based on *eigenvalues*: For example, recursively calculated trace tests, eigenvalues, and fluctuation tests of eigenvalues. These may indicate constancy problems around α and β.
3. Recursive tests of the constancy of the *cointegration space*: For example, "max test of a constant β" and test of "β_t equal to known β".
4. Recursive tests of *predictive failure* both for full system and individual series.

Outline

Diagonalising symmetric matrices[7]

A matrix M is *symmetric* if $M = M'$.
A matrix Λ is *diagonal* if off diagonal elements are zero.
A matrix V is *orthogonal* if $V'V = I_p$ or equivalently $V^{-1} = V'$.

Examples: $M = \begin{pmatrix} a & c \\ c & b \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$

$$V = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \text{ with } V^{-1} = V' = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

Lemma

Johansen (1996), Lemma A.3

M symmetric $\Rightarrow$ there exists orthogonal matrix V and diagonal Λ so

$$\underbrace{V'}_{p\times p}\underbrace{M}_{p\times p}\underbrace{V}_{p\times p} = \underbrace{\Lambda}_{p\times p} \qquad \textit{or equivalently} \qquad MV = V\Lambda.$$

[7] Johansen (1996), Ch. A.1, Simon and Blume (1994), Mirsky (1961).

Eigenvalues and eigenvectors for symmetric matrices

If $\mathbf{v}$ is a column vector of V and λ is a diagonal element of Λ then $M\mathbf{v} = \mathbf{v}\lambda$.
If $M\mathbf{v} = \mathbf{v}\lambda$ then $\mathbf{v}$ is an *eigenvector* of M and λ is an *eigenvalue* of M.
Eigenvalues also solve

$$|\lambda I - M| = 0.$$

The proof uses that $|AB| = |A||B|$ so

$$\begin{aligned} |\lambda I - M| &= |\lambda VV' - M| \\ &= |VV'||\lambda I - V'MV| = |\lambda I - \Lambda| = 0. \end{aligned}$$

Positive definiteness
A symmetric matrix is *positive definite* if all eigenvalue are positive.
A symmetric matrix is *positive semi-definite* if all eigenvalue are non-negative.

Joint diagonalisation of two symmetric matrices

Lemma

Johansen (1996), Lemma A.5

M, N symmetric so M positive semi-definite, N positive definite
$\Rightarrow$ there exists matrix V and diagonal, positive semidefinite Λ so

$$V'NV = I_p, \qquad V'MV = \Lambda, \qquad \textit{and} \qquad NV\Lambda = MV.$$

Λ has elements $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ solving $0 = |\lambda_j N - M|$.
V has columns $\mathbf{v}_1, \ldots, \mathbf{v}_p$ solving $N\mathbf{v}_j\lambda_j = M\mathbf{v}_j$.

Maximising ratio of determinants

Lemma

Johansen (1996), Lemma A.8

M, N symmetric and positive definite.

Suppose $\lambda_1 > \cdots > \lambda_p > 0$ and $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are those given in Lemma A.5.

Then

$$f(x) = \frac{|x'Mx|}{|x'Nx|}$$

is maximised among $x \in \mathbf{R}^{p \times r}$ with full column rank by

$$\hat{x} = (\mathbf{v}_1, \ldots, \mathbf{v}_r) \qquad \text{and} \qquad f(\hat{x}) = \Pi_{j=1}^{r} \lambda_j.$$

Exercises

1. Exam 2006, Question 5 (iv)-(vi).
2. Exam 2003, Question 2.
3. Exam 2001, Question 5 (i) , (ii) .
4. Exercise 3.1, Johansen (1996).
5. Exercise 4.11, Johansen (1996).

References

- Doornik, J.A., Nielsen, B. and Hendry, D.F. (1998), Inference in cointegrating models: UK M1 revisited, *Journal of Economic Surveys*, 12, 533–572.
- Johansen, S. (2002), A small sample correction of the test for cointegrating rank in the vector autoregressive model, *Econometrica*, 70, 1929–1961.
- Mirsky, L. (1961) *An introduction to linear algebra.* Corrected print Oxford University Press, reprinted by Dover.
- Nielsen, B. (2000). The asymptotic distribution of likelihood ratio test statistics for cointegration in unstable vector autoregressive processes. Nuffield College working paper no. 24.
- Nielsen, B. and Rahbek, A. (2000) *Oxford Bulletin of Economics and Statistics.*
- Simon, C.P. and Blume, L.E. (1994) *Mathematics for Economists*, Ch. 23.7-23.8.