EC408 Topics in Applied Econometrics

B Fingleton, Dept of Economics, Strathclyde University

Applied Econometrics

- What is spurious regression?
- How do we check for stochastic trends?
- Cointegration and Error Correction Models
- Autoregressive distributed lag (ADL) models
- VAR models

Applied Econometrics

VAR models

- Vector error correction models
- Multiple cointegrating vectors
- Johansen’s procedure

Multiple cointegrating vectors

- g variables, it is convenient to collect these together and represent them as the g x 1 vector
- Given that we have g variables, we may be able to discover <u>more than one linear combination</u> of the g variables in Y that is stationary, with each linear combination, or cointegrating relationship, being uncorrelated with, or orthogonal to, the others
- Recognising this will give a better specified model
- First we write out the VECM in terms of differences and levels
- This can be shown to be mathematically equivalent to models in which the error correction term (the lagged residuals) is explicit

the vector error correction representation

$$\Delta Y_t = \phi_1 \Delta Y_{t-1} + \phi_2 \Delta Y_{t-2} + \ldots + \phi_p \Delta Y_{t-p} + P_0 Y_{t-1} + U_t \qquad (72)$$

Where $\Delta Y_t = Y_t - Y_{t\text{-}1}$ is a g by 1 vector of differences at time t of g endogenous variables. The terms $\Delta Y_{t\text{-}1} = Y_{t\text{-}1} - Y_{t\text{-}2}, \Delta Y_{t\text{-}2} = Y_{t\text{-}2} - Y_{t\text{-}3}, \ldots, \Delta Y_{t\text{-}p} = Y_{t\text{-}p} - Y_{t\text{-}p-1}$ are the lagged differences. There are p lags, so ϕ_{j} applies to the j'th lag. It is a g x g matrix of coefficients to be estimated. Also U_{t} is a g x 1 vector of error terms. P_0 is a g x g matrix which is referred to as the (restricted) long-run matrix.

the vector error correction representation

let the number of lags p = 1, then equation (72) becomes

$$\Delta \mathrm{Y}_t = \phi_1 \Delta \mathrm{Y}_{t-1} + P_0 \mathrm{Y}_{t-1} + \mathrm{U}_t$$

for three variables, say Y_{1t}, Y_{2t}, Y_{3t}

$$\begin{bmatrix} \Delta Y_{1t} \\ \Delta Y_{2t} \\ \Delta Y_{3t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix} \begin{bmatrix} \Delta Y_{1t-1} \\ \Delta Y_{2t-1} \\ \Delta Y_{3t-1} \end{bmatrix} + \begin{bmatrix} P_{011} & P_{012} & P_{013} \\ P_{021} & P_{022} & P_{023} \\ P_{031} & P_{032} & P_{033} \end{bmatrix} \begin{bmatrix} Y_{1t-1} \\ Y_{2t-1} \\ Y_{3t-1} \end{bmatrix} + \begin{bmatrix} U_{1t} \\ U_{2t} \\ U_{3t} \end{bmatrix}$$

so that for variable Y_1

$$\Delta Y_{1t} = \phi_{11} \Delta Y_{1t-1} + \phi_{12} \Delta Y_{2t-1} + \phi_{13} \Delta Y_{3t-1} + P_{011} Y_{1t-1} + P_{012} Y_{2t-1} + P_{013} Y_{3t-1} + U_{1t}$$

and so on for Y_2, Y_3

the vector error correction representation

A mathematically equivalent way to write out this model is in terms of levels and lagged levels,

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \\ Y_{3t} \end{bmatrix} = \begin{bmatrix} 1+P_{011} & P_{012} & P_{013} \\ P_{021} & 1+P_{022} & P_{023} \\ P_{031} & P_{032} & 1+P_{033} \end{bmatrix} \begin{bmatrix} Y_{1t-1} \\ Y_{2t-1} \\ Y_{3t-1} \end{bmatrix} + \begin{bmatrix} U_{1t} \\ U_{2t} \\ U_{3t} \end{bmatrix} = \begin{bmatrix} \pi_{111} & \pi_{112} & \pi_{113} \\ \pi_{121} & \pi_{122} & \pi_{123} \\ \pi_{131} & \pi_{132} & \pi_{133} \end{bmatrix} \begin{bmatrix} Y_{1t-1} \\ Y_{2t-1} \\ Y_{3t-1} \end{bmatrix} + \begin{bmatrix} U_{1t} \\ U_{2t} \\ U_{3t} \end{bmatrix}$$

so that for variable Y_1

$$Y_{1t} = (1+P_{011})Y_{1t-1} + P_{012}Y_{2t-1} + P_{013}Y_{3t-1} + U_{1t} = \pi_{111}Y_{1t-1} + \pi_{112}Y_{2t-1} + \pi_{113}Y_{3t-1} + U_{1t}$$

$$Y_{2t} = P_{021}Y_{1t-1} + (1+P_{022})Y_{2t-1} + P_{023}Y_{3t-1} + U_{2t} = \pi_{121}Y_{1t-1} + \pi_{122}Y_{2t-1} + \pi_{123}Y_{3t-1} + U_{2t}$$

$$Y_{3t} = P_{031}Y_{1t-1} + P_{032}Y_{2t-1} + (1+P_{033})Y_{3t-1} + U_{3t} = \pi_{131}Y_{1t-1} + \pi_{132}Y_{2t-1} + \pi_{133}Y_{3t-1} + U_{3t}$$

and π_1 is a g x g matrix of coefficients specific to lag 1

Additional variables in the ADL

we do reject Ho: $\delta = 0$ in favour of $\delta < 0$ for INFLAT, since ADF-INFLAT = -6.197** Critical values used in ADF test: 5%=-3.439, 1%=-4.019.

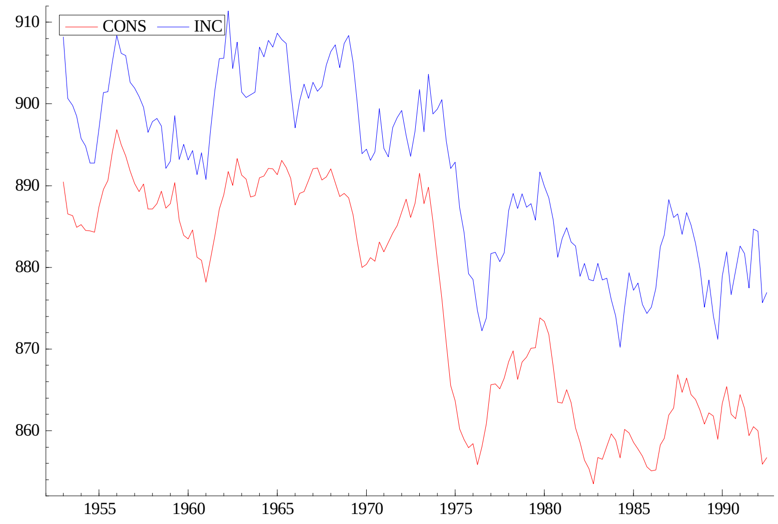

Fig 10 variables CONS and INC, both ~ I(1)

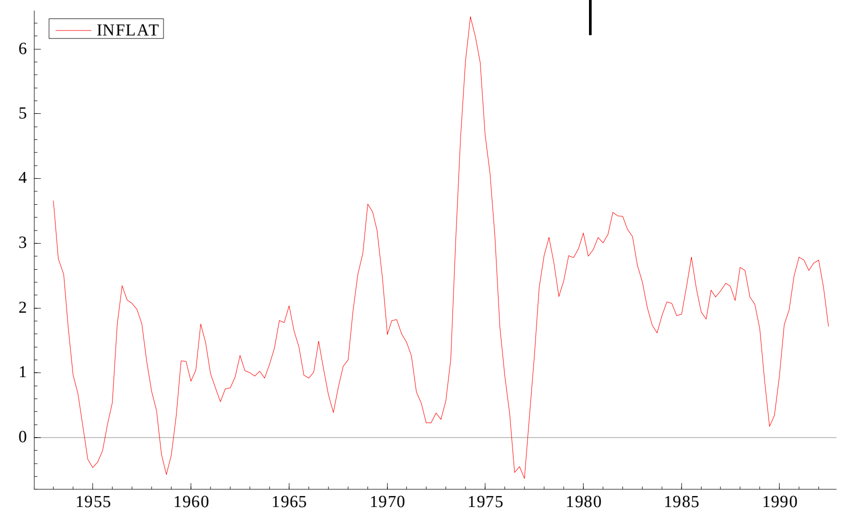

Fig 11 Variable INFLAT ~ I(0)

ADF-CONS = -3.034 which indicates that Ho: $\delta = 0$ should not be rejected in favour of $\delta < 0$ using critical values : 5%=-3.439, 1%=-4.019, We also find that $\delta = 0$ is not rejected for INC, since ADF-INC = -3.14 Critical values used in ADF test: 5%=-3.439, 1%=-4.019.

the vector error correction representation

As a numerical example, consider P_0 to be as follows,

Restricted long-run matrix, rank 2

	CONS	INC	INFLAT
CONS	-0.15042	0.14999	-1.2089
INC	0.072210	-0.070129	-0.49009
INFLAT	0.019763	-0.019382	-0.026605

$$\pi_1 = P_0 + \mathrm{I}$$

Where I is a g x g identity matrix

$$\mathrm{I} = \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$$

π_1 is the following set of numerical coefficients

```
CONS =  + 0.8496*CONS_1 + 0.15*INC_1 - 1.209*INFLAT_1

 INC =  + 0.07221*CONS_1 + 0.9299*INC_1 - 0.4901*INFLAT_1

INFLAT =  + 0.01976*CONS_1 - 0.01938*INC_1 + 0.9734*INFLAT_1
```

the vector error correction representation

More generally, with p > 1

$$\mathrm{Y}_t = \sum_{i=1}^{p} \pi_i \mathrm{Y}_{t-i} + \mathrm{U}_t$$

in this π_i is a g by g matrix of coefficients specific to lag *i* (73)

And it follows that

$$P_0 = \sum_{i=1}^{p} \pi_i - \mathrm{I} \qquad (74)$$

I is a g x g identity matrix, and P_0 is the g x g (restricted) long-run matrix

the vector error correction representation

As a numerical example with p = 2, we have P_0 as follows

Restricted long-run matrix, rank 2

	CONS	INC	INFLAT
CONS	-0.19280	0.19211	-1.4366
INC	-0.048286	0.049558	-1.1167
INFLAT	-0.0011498	0.0013482	-0.11461

Then π_1 is the set of numerical coefficients attached to the lag 1 terms, and π_2 is the set of numerical coefficients attached to the lag 2 terms

```
CONS =  + 0.6958*CONS_1 + 0.1454*INC_1 - 1.182*INFLAT_1 + 0.1114*CONS_2

        + 0.04668*INC_2 - 0.2543*INFLAT_2

 INC =  - 0.3309*CONS_1 + 0.98*INC_1 + 0.2796*INFLAT_1 + 0.2826*CONS_2

        + 0.0696*INC_2 - 1.396*INFLAT_2

INFLAT =  - 0.001926*CONS_1 + 0.01442*INC_1 + 1.535*INFLAT_1 + 0.0007761*CONS_2

          - 0.01307*INC_2 - 0.65*INFLAT_2
```

the vector error correction representation

$$P_{011} = \pi_{111} + \pi_{211} - 1$$

$$P_{011} = 0.6958 + 0.1114 - 1 = -0.19280$$

$$P_{012} = \pi_{112} + \pi_{212}$$

$$P_{012} = 0.1454 + 0.04668 = 0.19211$$

Hence and so on.... for example

$$P_{033} = \pi_{133} + \pi_{233} - 1$$

$$P_{033} = 1.535 + 0.65 - 1 = -0.11461$$

the vector error correction representation

We need $P_0 Y_{t-1}$ to be I(0) to balance the fact that $\Delta Y_t \sim I(0)$ in equation (72).

Given that we have g variables, we may be able to discover more than one linear combination of the g variables in Y that is stationary, with each linear combination, or cointegrating relationship, being uncorrelated with, or orthogonal to, the others.

Although P_0 is a g x g matrix, it could be the result of r < g cointegrating vectors.

<u>The approach developed by Johansen(1988) is designed to seek the actual number r of cointegrating relationships</u>

the vector error correction representation

There may also be some deterministic variables in the VAR, such as a time trend or a constant term, or some other non-modelled exogenous variables. We denote these by X. We also need these to be I(0) for our equation to balance, with I(0) variable throughout. Now the specification becomes

$$Y_t = \sum_{i=1}^{p} \pi_i Y_{t-1} + \sum_{j=0}^{r} \Gamma_j X_{t\text{-}j} + U_t \qquad (75)$$

the vector error correction representation

P_0 is g x g matrix of long-run responses

For the model to 'work' the left hand side and the right hand side must be I(0)

We know, since it assumed that Y is I(1), that ΔY_t is a set of I(0) variables

The rank of P_0 is the number of linearly independent rows of the matrix, and is given by the number of non-zero eigenvalues (characteristic roots).

Mathematically, since we assume Y is I(1), P_0 cannot be full rank (equal to g) and $P_0 Y_{t-1}$ ~I(0).

We have to place restrictions on the rank of P_0 so that the rank r < g.

If the rank of P_0 is r, this equals the number of independent cointegrating relationships between the g variables.

the vector error correction representation

The long run matrix P_0 can be decomposed

, it is the product of the g x r matrix of cointegrating vectors (β) and another g x r matrix (α)

the vector error correction representation

α **(3 x 1)**

```
   -0.19545
  -0.090935
 -0.0063286
```

α_{ij} is responsiveness of i'th variable to disequilibrium at t-1 given by j'th cointegrating vector

here j = 1

β **(3 x 1)**

```
     1.0000
   -0.99413
     6.3027
```

Cointegrating vector

$P_0 = \alpha\beta'$ **(3 x 3)**

```
   -0.19545      0.19431      -1.2319
  -0.090935     0.090402     -0.57314
 -0.0063286    0.0062915    -0.039887
```

Rank = 1

Equation 1: d_CONS

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE	
EC1	-0.195452	0.0247129	-7.909	<0.00001	***

Equation 2: d_INC

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE	
EC1	-0.0909355	0.0436303	-2.084	0.03876	**

Equation 3: d_INFLAT

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE
EC1	-0.00632863	0.00648877	-0.975	0.33090

Johansen's procedure : an informal introduction

designed to estimate the actual number of cointegrating linear combinations r

tests how many of the eigenvalues are significantly different from zero

The rank r will range from 0 to g

If r is zero, then that indicates there are no stationary linear combinations of the levels of the variables in Y~I(1)

If r = 1, then that means there is just one cointegrating vector

if r = g, then every linear combination of the variables in Y is stationary

this implies that all the series in Y ~ I(0)

this contradicts the assumption that Y ~ I(1) giving the left hand side variables ΔY_t ~I(0)

we want P_0 to be less than full rank, or the number of columns in β to be less than g

Johansen's procedure : an informal introduction

first we fit unrestricted reduced forms (URFs) for each of the endogenous variables CONS, INC, INFLAT

We then fit cointegrating equations, reducing the rank on P_0 from r = g = 3 to r = 0

With r = 3 we see that the matrix β has 3 columns and these are the 3 separate cointegrating vectors

α **(3 x 3)**

-0.15038	0.14992	-1.2053
0.072278	-0.070260	-0.48341
0.019754	-0.019364	-0.027522

β **(3 x 3)**

1	0	0
0	1	0
0	0	1

$P_0 = \alpha\beta'$ **(3 x 3)**

-0.15038	0.14992	-1.2053
0.072278	-0.070260	-0.48341
0.019754	-0.019364	-0.027522

log-likelihood -748.758136 This is exactly the same as for the unrestricted URF

Thus the r = 3 specification entails no simplification of the URF model. They are identical

Johansen's procedure : an informal introduction

when we fit the model with r = 2, giving only 2 cointegrating vectors

α (3 x 2)

-0.15042	0.14999
0.072210	-0.070129
0.019763	-0.019382

β (3 x 2)

1.0000	0.0000
0.0000	1.0000
-560.33	-569.98

$P_0 = \alpha\beta'$ (3 x 3)

-0.15042	0.14999	-1.2089
0.072210	-0.070129	-0.49009
0.019763	-0.019382	-0.026605

the log-likelihood is log L_R **= -748.797631,** only marginally less than for r = 3

Johansen's procedure : an informal introduction

With r = 1, there is only 1 cointegrating vector

α **(3 x 1)**

```
 -0.19545
 -0.090935
 -0.0063286
```

β **(3 x 1)**

```
 1.0000
 -0.99413
 6.3027
```

$P_0 = \alpha\beta'$ **(3 x 3)**

-0.19545	0.19431	-1.2319
-0.090935	0.090402	-0.57314
-0.0063286	0.0062915	-0.039887

The log-likelihood is -756.56227

Johansen's procedure : an informal introduction

The results for r = 1 show a much bigger drop in log-likelihood , which is down to log L_R **= -756.562268**

This suggests that the simplification involved in just having one cointegrating vector is too great

Overall, the assumption that r = 2 appears to be the most acceptable, given the larger falls in the likelihood with lower r. However, we cannot really be sure about this until the significance of the changes in likelihood have been formally tested, which is the function of Johansen's procedure for establishing the rank r of P_0.

Model	T	p		log-likelihood	SC	HQ	AIC
SYS(51)	158	5	COINT	-756.56227	9.7369	9.6794	9.6400
SYS(50)	158	8	COINT	-748.79763	9.7348	9.6427	9.5797
SYS(49)	158	9	COINT	-748.75814	9.7663	9.6627	9.5919
SYS(48)	158	9	OLS	-748.75814	9.7663	9.6627	9.5919

Johansen test

- **<u>Trace test</u>**
- Compares likelihoods for rank r model and var model (full rank)
- If the difference is significant, we <u>cannot</u> assume rank is r and eliminate higher ranks
- If the difference is not significant, we can assume rank is r

- Null hypothesis : rank $\leq$ r
- Alternative hypothesis : r $<$ rank $\leq$ full rank

Johansen test

- **Maximum eigenvalue test**
- Compares likelihoods for rank r model and rank r+1 model
- If the difference is significant, rank r + 1 improves likelihood
- and we assume rank is r + 1

- If the difference is not significant, we can assume rank is r

- Null hypothesis : rank at most = r
- Alternative hypothesis : rank = r + 1

Johansen's procedure

```
Johansen test:
Number of equations = 3
Lag order = 1
Estimation period: 1953:2 - 1992:3 (T = 158)

Case 1: No constant
Rank Eigenvalue Trace test p-value   Lmax test  p-value
   0    0.44574     108.85 [0.0000]     93.239 [0.0000]
   1   0.093611     15.608 [0.0127]     15.529 [0.0071]
   2 0.00049982   0.078991 [0.8414]   0.078991 [0.8327]
```

Johansen's procedure

In terms of likelihood, this shows exactly the same results

The null hypothesis in this case is that r <= 1, in other words columns 2 and 3 of β are null (there is at most one cointegrating vector)

comparing the log-likelihoods for r = 3 versus r = 1, we obtain a test statistic 15.608 = $2\left\{\log L_U - \log L_R\right\}$ = 2(-748.75814 + 756.56227)

$$prob\left\{\chi_4^2 \geq 15.608\right\} = 0.0036$$

for reasons similar to those that produce the non-standard distributions for the Dickey-Fuller test statistic, χ_4^2 is <u>not the correct reference distribution</u>

Both gretl and PcGIVE provide the appropriate p-value, equal to 0.013

Johansen's procedure

Ho: r <= 2

$$2\{\log L_U - \log L_R\} = 2(-748.75814 + 748.79763) = 0.078991$$

p-value equal 0.841 (nb compare this with the p-value of 0.7787 given by the theoretical χ^2_1 distribution)

Johansen's procedure

Given that we have established the rank of P_0 and hence the number of cointegrating vectors, we can then move forward in the knowledge that we have a balanced model with stationary variables.

we can obtain estimates of the dependencies within the data that are not spurious, and we will ultimately be able to produce more credible forecasts and a richer and more informative picture of the interrelationships between the variables.