

90-84

Faculteit der Economische Wetenschappen en Econometrie

ET

05348

Serie Research Memoranda

The program "Pjohcoin"

for applications of the co-integration tests of Johansen

Ben Vogelvang

Research Memorandum 1990-84
December 1990





THE PROGRAM "PJOHCOIN",
for applications of the co-integration tests of Johansen

Ben Vogelvang
Vakgroep Econometrie
FEWEC-VU

Comments welcome
November 1990

Summary

This is the description of the Fortran program "Pjohcoin", which can be used on the IBM-3090 mainframe of SARA in a simple and interactive way. The program can be applied to test for the number of co-integration relationships which may exist between a number of economic variables, according to the the maximum likelihood approach of Johansen (and Juselius). If these relationships do exist, then restrictions on the parameters of these relationships can be tested as well with this program.

Contents

1. Introduction
2. Summary of the maximum-likelihood approach of Johansen
3. Description of the program PJOHCOIN and how to use it
 - a. The use of the program
 - b. An example of the output of the program
- References

1. INTRODUCTION

In this manual the applications of the tests belonging to the multivariate approach of Johansen (1988), (1989), with help of the program "Pjohcoin", has been described. That program is available on the IBM-3090, from now on. The program is available in translated form on a minidisk of the author which may be linked by any user of the VM-machine, who knows the password, during his or her own session. The procedure of Johansen and Juselius to test for co-integration relationships is not yet common available in software packages, which is a reason to describe and promote this program in this manner, as these methods are beginning to become rather popular. Besides, the program can only improve if is often used.

The procedures of Johansen have been programmed to use for VAR-models without constant terms, with constant terms (linear trend in the data) or with a restriction on the constant terms that eliminates the trend in the long-run relationships. The asymptotical distribution of the tests is dependent on the assumption that has been made concerning the constant term in the VAR-model. The correct quantiles of this distribution are given in the output depending on the assumptions which have been made. So far, seasonal dummies have not been included in the program. The program gives the output on the screen and writes the results to an output file. The output file contains more results than have been written to the screen, like parameter estimates which have not to be judged immediately. An option is available to write all or only the most important results to the output file.

In Section 2 the procedure of Johansen will be summarized in condensed form, for reasons of completeness, but it is also necessary for the introduction of the usual notation, which is used both in the program description and in the output. Next, in Section 3 the program is discussed and its use is explained, followed by some examples of the output.

Anyone who wishes to use the program, can receive the password, necessary to read the minidisk, from the author. Comments on the program or particular ideas concerning the output are very welcome, as certain imperfections in the program can certainly be expected.

2. SUMMARY OF THE MAXIMUM LIKELIHOOD APPROACH OF JOHANSEN

This section gives a condensed résumé of the theoretical and applied work of Johansen (1988), (1989) and Johansen and Juselius (1989) to the problem of testing for and estimating of co-integration vectors in a multivariate context. The procedure consists of the following stages. The first step is to test for the number of co-integrating vectors in a multivariate system. This is done with a likelihood-ratio test, that has an asymptotic distribution which can be approximated by a χ^2 distribution. Johansen computed fractiles of the distribution by simulation, up to five co-integrating relationships, depending on the assumption concerning the constant term. In following steps the co-integrating space is estimated, which dimension equals the number of co-integrating vectors, after which hypotheses about possible restrictions on the co-integrating vectors are tested concerning economic interpretations of the results.

We consider first the unrestricted vector-autoregressive (VAR) process, which is integrated of order 1, of p variables and a constant term.

$$X_t = \Pi_1 X_{t-1} + \dots + \Pi_k X_{t-k} + \mu + \varepsilon_t, \quad (2.1)$$

with ε_t being a sequence of i.i.d. p -dimensional Gaussian random vectors, distributed as $N(0, A)$. This specification can be rewritten as:

$$\nabla X_t = \Gamma_1 \nabla X_{t-1} + \dots + \Gamma_{k-1} \nabla X_{t-k+1} - \Pi X_{t-k} + \mu + \varepsilon_t, \quad (2.2)$$

where $\Gamma_i = -I + \Pi_1 + \dots + \Pi_i$,

and $\Pi = I - \Pi_1 - \dots - \Pi_k$.

Then the research concentrates on the matrix Π , because of the long-run information that can be obtained from knowledge of this matrix, which is often called the impact matrix. The following situations may occur concerning the matrix Π :

$\text{rk}(\Pi) = p$, The matrix has full rank, implying a stationary process X_t ;

$\text{rk}(\Pi) = 0$, The matrix Π is zero, implying an integrated vector process X_t ;

$\text{rk}(\Pi) = r$, $0 < r < p$, implying the existence of $p \times r$ matrices α and β of rank r , giving a non-linear constraint on the coefficients Π_1, \dots, Π_k : $\Pi = \alpha \beta'$.

with β being the matrix with co-integrating vectors; so $\beta' X_t$ is stationary.

With reference to Johansen we memorize that the parameters α and β cannot be estimated, as they are not uniquely determined, but the space spanned by β can be estimated. First two matrices of residuals are computed, originating from the regression of ∇X_t on $\nabla X_{t-1}, \dots, \nabla X_{t-k+1}, 1$, and X_{t-k} on the same set of regressors $\nabla X_{t-1}, \dots, \nabla X_{t-k+1}, 1$. Denote these residuals by R_{0t} and R_{kt} . These residuals can also be computed according their definitions:

$$\begin{aligned} R_{0t} &= Z_{0t} - M_{01}M_{11}^{-1}Z_{1t}, \\ R_{kt} &= Z_{kt} - M_{k1}M_{11}^{-1}Z_{1t}, \end{aligned} \quad (2.3)$$

where $Z_{0t} = \nabla X_t$, $Z_{1t} = (\nabla X_{t-1}, \dots, \nabla X_{t-k+1}, 1)'$, and $Z_{kt} = X_{t-k}$. Z_{0t} is a $(p \times 1)$, Z_{1t} is a $\{[p(k-1)+1] \times 1\}$ and Z_{kt} is a $(p \times 1)$ vector. The matrices M_{ij} are defined as:

$$M_{ij} = T^{-1} \sum_{t=1}^T Z_{it}Z_{jt}'. \quad (2.4)$$

M_{01} is a $[p \times \{p(k-1)+1\}]$, M_{11} is a $\{[p(k-1)+1] \times [p(k-1)+1]\}$, and M_{k1} is a $[p \times \{p(k-1)+1\}]$ matrix. Then the $(p \times p)$ moment matrices S_{00} , S_{kk} and S_{k0} are computed:

$$S_{ij} = T^{-1} \sum_{t=1}^T R_{it}R_{jt}'. \quad (2.5)$$

Johansen proposes the next procedure to test for the number of co-integration relationships. Solve the equation

$$|\lambda S_{kk} - S_{k0}S_{00}^{-1}S_{0k}| = 0,$$

giving the p eigenvalues λ_i and determine the corresponding eigenvectors. Let E be the matrix with eigenvectors, then E is normalized such that $E'S_{kk}E = I$.

Three possible assumptions can be made with respect to the constant term μ .

- First one may assume $\mu = 0$, implying that Z_{1t} is a $[p(k-1) \times 1]$ vector, resulting in lower dimensions of the matrices M too.
- Secondly $\mu \neq 0$, than the non-stationary process X_t contains a linear trend, which is the situation as described above.
- Lastly $\mu \neq 0$ and the trend is absent in the process; this implies that the restriction $\mu = \alpha\beta_0'$ holds. In this case the vectors Z_{1t} and Z_{kt} are

redefined: $Z_{1t}^* = (\nabla X_{t-1}, \dots, \nabla X_{t-k+1})'$ and $Z_{kt}^* = (X_{t-k}, 1)'$, which implies that the dimensions of M^* and S_{ij}^* are also different. There is no difference in applying the procedure to compute the likelihood ratio tests in these different situations.

The number of co-integrating vectors r is determined by means of the likelihood-ratio test statistic (2.6) for H_0 : there are at most r co-integration vectors.

$$-2\ln(Q) = -T \sum_{i=r+1}^p \ln(1 - \hat{\lambda}_i), \quad (2.6)$$

where $\hat{\lambda}_{r+1}, \dots, \hat{\lambda}_p$ are the $p - r$ smallest eigenvalues. So the matrix β which columns span the co-integration space, consists of the r eigenvectors belonging to the r largest eigenvalues. Then the matrix α and the restricted impact matrix $-\Pi_r$ are computed as: $\alpha = -S_{0k}\beta$ and $\Pi_r = \alpha\beta'$.

The next step concerns the economic interpretation of the matrices α and β , which is not straightforward in this multivariate analysis. If $r = 1$, it is probably possible to normalize with respect to one of the variables, and to test for any exclusion restriction. But with $r > 1$ we cannot arbitrarily normalize the co-integration equations. In fact, only the hypotheses which co-integrating vectors exist can be tested. With the matrix β , we have computed a basis of the co-integration space. Therefore hypotheses can be tested, stemming from economic knowledge, which we have of the problem that we are investigating, concerning restrictions on the co-integrating vectors: $H_0: \beta = H\phi$, with H a known $(p \times s)$ matrix of constants and ϕ a $(s \times r)$ matrix of unknown parameters, ($r \leq s \leq p$). If $s = p$, then no restrictions are placed upon the choice of the co-integration vectors, and if $s = r$, then the co-integration space is fully specified. These restrictions are imposed on all the co-integration vectors, as otherwise no meaningful conclusions can be drawn. Johansen proves that the next procedure can be used to test the hypothesis H_0 . First solve the equation:

$$|\lambda H' S_{kk} H - H' S_{k0} S_{00}^{-1} S_{0k} H| = 0,$$

which gives the s eigenvalues λ_i^* . Let E^* be the matrix with eigenvectors, then E^* is normalized such that $E^{*'} H' S_{kk} H E^* = I$. Further the null is tested

with the likelihood-ratio test:

$$-2\ln(Q) = T \sum_{i=1}^r \ln[(1 - \lambda_i^*) / (1 - \hat{\lambda}_i)], \quad (2.7)$$

with λ_i^* and $\hat{\lambda}_i$ the r largest eigenvalues. This test statistic is asymptotically distributed as $\chi^2(r(p-s))$. Some indications concerning the formulation of H_0 can be obtained from the eigenvectors in β , where e.g. opposite signs can indicate that the difference of two variables has to enter a co-integration relationship or one large coefficient may indicate that the corresponding variable is stationary.

3. DESCRIPTION OF THE PROGRAM PJOHCOIN AND HOW TO USE IT

a. The use of the program

The program "Pjohcoin" has been written in Fortran for the VM machine of SARA, using real arrays in double precision (which is really a necessity on this machine), and the library IMSL for subroutines to invert matrices (DLINRG) and to compute eigenvalues and eigenvectors (DGVCSF, DLFTDS).

A data file has to be prepared before running the program, according to the sequence:

```
name series 1
..data series 1..
.....
name series p
..data series p..
```

The program reads the file COIN DATA A; so the data file has to be copied to a file with this name. All the output is written on the file COIN OUT A, while a part of the results, like the test statistics is also written on the screen.

The following commands have to be given to start the program:

```
LINK V67U073A 193 555 <password>      [links the minidisk 193 of the author,
                                         555 is an arbitrary number]
ACCESS 555 <fm>                        [e.g. when only the filemode A is in use, than
                                         file mode B can be taken]
ADD FORTV2                              [adding the Fortran compiler]
ADD IMSL                                [adding the IMSL library]
LOAD PJOHCOIN(START                     [starting the program]
```

From then on the computer asks you to give the relevant information to perform the tests. Summarized, this concerns the following data.

- *Information about the constant term.*

Typing a 0 means: no constant term, the table of Johansen (1988) is used,
typing a 1 means: the inclusion of a constant term (trend in the process),
typing a 2 means: the inclusion of a constant term with the restriction $\mu = \alpha\beta_0'$, implying the absence of the trend. In the last two situations the

tables of Johansen (1989) are used.

- *Number of observations (T) and variables (p).*

At this moment the following restrictions are effective: $p \leq 10$ (inclusive of a possible constant term) and $T \leq 200$.

- *An option is available to test only for the number of co-integrating vectors r , or to compute also the restricted parameters after which "hypotheses H " can be tested and restricted parameters are computed.*

- *An option is available to transform the variables into their logs.*

- *The order k of the VAR process.*

- *An option is available to choose between the most relevant output or all the output of the program (like the residual matrices R (eq. 2.3) and the matrices M (eq. 2.4).*

- *The value of r and s to test "hypotheses H ".*

- *An option to stop or to continue for one more "hypothesis H ".*

b. An example of the output of the program

The output of the program, used to compute the results in Vogelvang (1990) is given here. This concerns four agricultural commodity prices: coffee, cocoa, tea, sugar, and an world inflation indicator.

The file COIN OUT A, with some comments:

This is the program PJOHCOIN, which can be used to test for the number of co-integrating relationships among max. 10 variables with max. 200 observations, and to test for hypotheses according to Johansen (1988), or (1989).

(c) Ben Vogelvang, Department of Econometrics VU Amsterdam.

Comments welcome.

The series have to be stored in the file COIN DATA A, according to: name1, series1, name2, series2, etc.

output is written on the screen (partial) and in the file COIN OUT A (all).

Trend is assumed in the series, restriction is not effective.

Number of observations: 132, number of variables: 5

Both the test on the number (r) of co-integrating relationships is computed and hypotheses are tested.

The following variables have been read:

coffee cocoa tea sugar inflation

Variables have been transformed to logs

SOO:

0.0043	0.0003	0.0016	0.0000	0.0001
0.0003	0.0044	0.0005	-0.0013	0.0001
0.0016	0.0005	0.0072	0.0020	0.0001
0.0000	-0.0013	0.0020	0.0135	0.0002
0.0001	0.0001	0.0001	0.0002	0.0002

SKK:

0.0411	0.0250	0.0045	0.0066	-0.0085
0.0250	0.0591	0.0155	-0.0003	-0.0124
0.0045	0.0155	0.0546	-0.0401	0.0033
0.0066	-0.0003	-0.0401	0.2583	0.0205
-0.0085	-0.0124	0.0033	0.0205	0.0157

(eq. 2.5)

SKO:

-0.0041	-0.0014	-0.0031	-0.0001	-0.0004
-0.0005	-0.0033	-0.0021	-0.0032	-0.0024
-0.0006	-0.0002	-0.0057	0.0027	-0.0002
0.0008	0.0032	-0.0005	-0.0050	0.0020
0.0004	0.0010	-0.0005	-0.0009	-0.0005

PI-unrestricted:

0.1248	-0.0297	0.0522	0.0008	0.0579
0.0001	0.1164	-0.0186	-0.0114	0.2665
0.0130	-0.0515	0.1326	0.0152	-0.0705
-0.0293	-0.1594	0.1339	0.0718	-0.3942
0.0010	-0.0220	0.0209	0.0066	0.0014

GAMMA-unrestricted:

0.1049	0.1782	-0.0007	0.0315	0.2712	0.0448	0.0574
-0.1182	0.0042	-0.0304	0.0597	0.0366	-0.0585	-0.3083
0.1612	0.1993	0.1443	-0.0044	-0.4499	-0.0595	-0.0176
-0.1955	0.3286	-0.0558	0.0979	0.1293	0.4909	0.1656
-0.0138	-0.0182	-0.0122	-0.0013	-0.0538	0.0341	-0.0033

To save space, these are not all the parameter estimates. GAMMA has $p(k-1)+1$ [or $p(k-1)$] columns. The first p columns belong to lag 1 (F_1), the second p columns belong to lag 2 (F_2), etc.. The last column [$p(k-1)+1$] contains the possible constant terms.

Eigenvalues:

0.0258	0.0421	0.1237	0.2024	0.2920	$\hat{\lambda}_i$
--------	--------	--------	--------	--------	-------------------

Eigenvectors:

0.2285	0.8083	1.0000	0.5059	0.1022
-0.3668	-0.5636	-0.2229	-0.5147	0.5153
-0.2511	-0.7571	0.0849	1.0000	-0.2815
-0.6462	0.0228	-0.0604	0.1755	-0.1242
1.0000	1.0000	-0.5961	0.7373	1.0000

New normalized eigenvectors:

0.7326	2.9403	4.5824	1.8677	0.8774
-1.1761	-2.0501	-1.0214	-1.9001	4.4242
-0.8051	-2.7541	0.3891	3.6917	-2.4166
-2.0718	0.0831	-0.2767	0.6478	-1.0662
3.2064	3.6375	-2.7318	2.7219	8.5857

Johansen's LR-test:

3.3778	8.9282	25.9657	55.1413	99.6792	(eq. 2.6)
--------	--------	---------	---------	---------	-----------

90% Quantiles:

6.6910	15.5830	28.4360	45.2480	65.9560
--------	---------	---------	---------	---------

95% Quantiles:

8.0830	17.8440	31.2560	48.4190	69.9770
--------	---------	---------	---------	---------

97.5% Quantiles:

9.6580	19.6110	34.0620	51.8010	73.0310
--------	---------	---------	---------	---------

The matrix alfa:

0.0059	0.0177	0.0176
0.0273	0.0113	-0.0079
-0.0078	0.0202	0.0048
-0.0389	0.0042	0.0042
-0.0027	0.0051	-0.0021

PI restricted:

0.1186	-0.0256	0.0579	0.0003	0.0506
0.0090	0.1075	-0.0275	-0.0197	0.2868
0.0529	-0.0780	0.0955	0.0201	-0.0254
-0.0069	-0.1844	0.1112	0.0430	-0.3341
-0.0027	-0.0194	0.0244	0.0067	-0.0035

De matrix H:

0.0000	-1.0000	1.0000	1.0000
1.0000	0.0000	-1.0000	-1.0000
-1.0000	1.0000	1.0000	1.0000
0.0000	0.0000	0.0000	1.0000
1.0000	1.0000	0.0000	1.0000

Eigenvalues:

0.0320	0.1163	0.2012	0.2919
--------	--------	--------	--------

(λ_i^*)

Eigenvectors:

0.2862	0.3910	0.4539	1.0000
0.6967	-0.5505	0.5258	0.2833
1.0000	1.0000	1.0000	0.5415
-0.8611	-0.1479	0.2134	-0.1405

New normalized eigenvectors:

0.5620	1.4683	1.2183	7.4255
1.3684	-2.0670	1.4114	2.1034
1.9640	3.7550	2.6841	4.0209
-1.6912	-0.5554	0.5727	-1.0436

Test value to test the restrictions H: 1.30558 (eq. 2.7)

Number of co-integrating relationships, $r = 3$

Number of columns of H, $s = 4$

Number of degrees of freedom: 3

Chi-square value at the 5%-level: 7.82

[$\chi^2_{.05}(1), \dots, \chi^2_{.05}(10)$ are included in the program]

The restricted matrix beta:

0.8739	1.8454	5.2666
4.4482	-2.0384	-1.7313
-2.3449	3.4499	-0.3357
-1.0436	0.5727	-0.5554
8.4853	3.2024	-1.1540

The matrix alfa:

0.0061	0.0168	0.0177
0.0274	0.0112	-0.0085
-0.0075	0.0191	0.0020
-0.0386	0.0032	0.0010
-0.0027	0.0052	-0.0016

PI restricted:

0.1297	-0.0379	0.0379	-0.0065	0.0851
0.0001	0.1136	-0.0226	-0.0175	0.2781
0.0390	-0.0757	0.0828	0.0177	-0.0048
-0.0226	-0.1800	0.1012	0.0416	-0.3186
-0.0010	-0.0200	0.0250	0.0067	-0.0043

REFERENCES

- Johansen, S. (1988): 'Statistical analysis of cointegration vectors', *Journal of Economic Dynamics and Control*, 12, 231-254.
- Johansen, S. (1989): 'Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models', Preprint University of Copenhagen.
- Johansen, S., and K. Juselius (1989): 'The full information maximum likelihood procedure for inference on cointegration - with applications', Preprint University of Copenhagen.
- Vogelvang, E. (1990): 'Testing for co-integration with spot prices of some related agricultural commodities, Free University Amsterdam, Dept. of Economics and Econometrics, *Research Memorandum* 1990-1, 35p.

1990-60	P. Nijkamp H.J.Scholten	New Information Systems: The use of retail information systems in the Netherlands
1990-61	P. Nijkamp J. Poot	Endogenous Technological Progress and Spatial Interdependence
1990-62	P. Nijkamp A. Reggiani	Strategic Long-Term Transportation Planning and Dynamic Transportation Models: A Life Cycle Interpretation
1990-63	P. Nijkamp A. Reggiani	Spatio-Temporal Processes in Dynamic Logit Models
1990-64	P. Nijkamp P. Rietveld I. Salomon	Barriers in Spatial Interactions and Communications: A Conceptual Exploration
1990-65	P. Nijkamp	Measurement in Conservation Planning
1990-66	C.P.M. Wilderom P. Nijkamp A.M.C. Vollering	Services for the Elderly in Europe: A Cross-National Comparative Analysis
1990-67	J.M. Vleugel P. Nijkamp P. Rietveld	Network Infrastructure and Regional Development; A Case Study for North-Holland
1990-68	H.J. Blommestein P. Nijkamp	A Comparison of Some Analytical Procedures for Nonmetric Information in Conflict Analysis
1990-69	F. Brouwer P. Nijkamp	Modelling Interactions between Economic Development and Environmental Change: A Policy Life-Cycle Interpretation
1990-70	P. Nijkamp	Spatial Developments in The United States of Europe: Glorious Victories or Great Defeats
1990-71	M. van Geenhuizen P. Nijkamp P. Townroe	Company Life History Analysis and Technogenesis: A Spatial View
1990-72	J.C.J.M. van den Bergh	A General Dynamic Economic-Ecological Models for Regional Sustainable Development
1990-73	J.C.J.M. van den Bergh P. Nijkamp	Aggregate Dynamic Economic-Ecological Models for Sustainable Development
1990-74	P. Nijkamp A. Perrels L. Schippers	Strategic Evaluation of New Infrastructure in Europe
1990-75	P. Nijkamp J. Vleugel	Critical Success Factors for New European Network Infrastructure
1990-76	P. Nijkamp	Regional Economic Growth and Regional Policy: A European Perspective
1990-77	R.J. Boucherie	Aggregation of Stochastic Processes
1990-78	H.J. Bierens L. Broersma	The Relation between Unemployment and Interest Rate: Some Empirical Evidence