Søren Johansen

Statistical Analysis of Cointegration Vectors

Preprint
October
1987

7

Institute of Mathematical Statistics
University of Copenhagen

ISSN 0902-8846

Søren Johansen

STATISTICAL ANALYSIS OF COINTEGRATION VECTORS

Preprint 1987 No. 7

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

October 1987

STATISTICAL ANALYSIS OF COINTEGRATION VECTORS

by

SØREN JOHANSEN

INSTITUTE OF MATHEMATICAL STATISTICS

UNIVERSITY OF COPENHAGEN

12. October 1987

ABSTRACT.

We consider a non stationary vector autoregressive process which is integrated of order 1, and generated by i.i.d Gaussian errors. We then derive the maximum likelihood estimator of the space of cointegration vectors and the likelihood ratio test of the hypothesis that it has a given number of dimensions. Further we test linear hypotheses about the cointegration vectors.

The asymptotic distribution of these test statistics are found and one is described by a natural multivariate version of the usual test for a unit root in an autoregressive process, and the other by a χ^2 test.

1. INTRODUCTION

The idea of using cointegration vectors in the study of non stationary time series comes from the work of Granger (1981), Granger & Weiss (1983), Granger & Engle (1985), and Engle & Granger (1987). The connection with error correcting models has been investigated by a number of authors, see Davidson (1986), Stock (1985) and Johansen (1985) among others.

Granger & Engle (1987) suggest estimating the cointegration relations using regression, and these estimators have been investigated by Stock (1985), Phillips (1985) and Phillips & Durlauf (1985), Phillips & Park (1986^a),(1986^b) and (1987), Phillips & Ouliaris (1986), Stock & Watson (1987), and Sims,Stock & Watson (1986). The purpose of this paper is to derive maximum likelihood estimators of the cointegration vectors for an autoregressive process with independent Gaussian errors, and to derive a likelihood ratio test for the hypothesis that there is a given number of these.

This programme will not only give good estimates and test statistics in the Gaussian case, but will also yield estimators and tests, the properties of which can be investigated under various assumptions about the underlying data generating process. The reason for expecting the estimators to behave better than the regression estimates is that they take into account the error structure of the underlying process, which the regression estimates do not.

The processes we shall consider are defined from a sequence $\{\epsilon_t\}$ of i.i.d. p-dimensional Gaussian random variables with mean zero and variance matrix Λ. We shall define the process X_t by

(1.1) $$X_t = \Pi_1 X_{t-1} + \ldots + \Pi_k X_{t-k} + \epsilon_t, \quad t = 1,2,\ldots$$

for given values of $X_{-k+1},\ldots,X_0$. We shall work in the conditional distribution given the starting values, since we shall allow the process X_t to be non stationary. We define the matrix polynomium

$$A(z) = I - \Pi_1 z - \ldots - \Pi_k z^k$$

and we shall be concerned with the situation where the determinant $|A(z)|$ has roots at $z = 1$. The general structure of such processes and the relation to error correction models was studied in the above references.

We shall in this paper mainly consider a very simple case where X_t is integrated of order 1, such that ΔX_t is stationary, and where the impact matrix

$$A(z)|_{z=1} = \Pi = I - \Pi_1 - \ldots - \Pi_k$$

has rank $r < p$. If we express this as

(1.2) $$\Pi = \alpha\beta'$$

for suitable $p \times r$ matrices α and β, then we shall assume that although ΔX_t is stationary and X_t is non stationary as a vector process, still the linear combinations given by $\beta' X_t$ are stationary. In the terminology of Granger this means that the vector process X_t is cointegrated with cointegration vectors β. The space spanned by β is the space spanned by the rows of the matrix Π, which we shall call the cointegration space.

In this paper we shall derive the likelihood ratio test for the hypothesis given by (1.2), and derive the maximum likelihood estimator of the cointegration space. Then we shall find the likelihood ratio test of the hyptothesis that the cointegration space is restricted to lie in a certain subspace, representing the linear restrictions that one may want to impose on the cointegration vectors.

The results we obtain can briefly be described as follows: the estimation of β is performed by first regressing ΔX_t and X_{t-k} on the lagged differences. From the residuals of these regressions we calculate a $2p \times 2p$ matrix of product moments. We can now show that the estimate of β is the empirical canonical variates of X_{t-k} with respect to ΔX_t corrected for the lagged differences.

The likelihood ratio test is now a function of certain eigenvalues of the product moment matrix corresponding to the smallest squared canonical correlations. The test of the linear restrictions involve yet another set of eigenvalues of a reduced product moment matrix. The asymptotic distributions of the first test statistic involve an integral of a multivariate Brownian motion with respect to itself, and turns out to depend on just one parameter, namely the dimension of the process, and can hence be tabulated by simulationor approximated by a χ^2 distribution. The second test statistic is asymptotically distributed as χ^2 with the proper degrees of freedom.

2. MAXIMUM LIKELIHOOD ESTIMATION OF COINTEGRATION VECTORS AND LIKELIHOOD RATIO TESTS OF HYPOTHESES ABOUT COINTEGRATION VECTORS.

We want to estimate the space spanned by β from observations X_t, $t = -k+1,\ldots,T$. For any $r \leq p$ we formulate the model as the hypothesis

$$H_0 : \text{rank}(\Pi) \leq r \quad \text{or } \Pi = \alpha\,\beta' \tag{2.1}$$

where α and β are $p\times r$ matrices.

Note that there are no other constraints on $\Pi_1,\ldots,\Pi_k$ than (2.1). Hence a wide class containing stationary as well as non stationary processes is considered.

The parameters α and β can not be estimated since they form an overparametrisation of the model, but one can estimate the space spanned by β which is the range space of Π. If we choose a suitable base in this space then we can also estimate the individual cointegration vectors.

We can now formulate the main result about the estimation of $sp(\beta)$ and the test of the hypothesis (2.1).

THEOREM 1. The maximum likelihood estimator of the space spanned by β is the space spanned by the r canonical variates corresponding to the r largest squared canonical correlations between the residuals of X_{t-k} and ΔX_t corrected for the effect of the lagged differences of the X process.

The likelihood ratio test statistic for the hypothesis that there are at most r cointegration vectors is

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

where $\hat{\lambda}_{r+1},\ldots,\hat{\lambda}_p$ are the p-r smallest squared canonical correlations.

Next we shall investigate the test of linear hypotheses on β. In the case we have $r = 1$, i.e. only one cointegration vector, it seems natural to test that certain variables do not enter into the cointegration vector, or that certain linear constraints are satisfied, for instance that the variables X_{1t} and X_{2t} only enter through their difference $X_{1t} - X_{2t}$. If $r \geq 2$ then a hypothesis of interest could be that the variables X_{1t} and X_{2t} enter through their difference only in all the cointegration vectors, since if two different linear combinations would occur then any coefficients to X_{1t} and X_{2t} would be possible. Thus it seems that some natural hypotheses on β can be formulated as

$$H_1 \; : \beta = H\varphi \tag{2.2}$$

where $H(p{\times}s)$ is a known matrix of full rank s, and $\varphi(s{\times}r)$ is a matrix of unknown parameters. We assume that $p \geq s \geq r$. If $s = p$ then no restrictions are placed upon the choice of cointegration vectors, and if $s = r$ then the cointegration space is fully specified.

THEOREM 2. The maximum likelihood estimator of the cointegration space under the assumption that it is restricted to sp(H) is given as the space spanned by the canonical variates corresponding to the r largest squared canonical correlations between the residuals of $H'X_{t-k}$ and ΔX_t corrected for the lagged differences of X_t.

The likelihood ratio test now becomes

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

where $\lambda_1^*,\ldots,\lambda_r^*$ are the r largest squared canonical correlations.

Proof. We shall here give the proofs of both theorems. Before studying the likelihood function it is convenient to reparametrise the model (1.1) such that the parameter of interest Π enters explicitly. We write

(2.3) $$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \dots + \Gamma_{k-1}\Delta X_{t-k+1} + \Gamma_k X_{t-k} + \epsilon_t$$

where

$$\Gamma_i = -I + \Pi_1 + \dots + \Pi_i \ , \ i = 1,\dots,k.$$

Note that (2.1) gives a non linear constraint on the coefficients $\Pi_1,\dots,\Pi_k$, but that the parameters $(\Gamma_1,\dots,\Gamma_{k-1},\alpha,\beta,\Lambda)$ have no constraints imposed. In this way the impact matrix $\Pi = -\Gamma_k$ is found as the coefficient of the lagged levels in a non linear least squares regression of ΔX_t on lagged differences and lagged levels. Under the constraint (2.1) we shall maximise the likelihood function with respect to the parameters

$$\alpha,\beta,\Gamma_1,\dots,\Gamma_{k-1},\Lambda.$$

The maximisation over the parameters $\Gamma_1,\dots,\Gamma_{k-1}$ is easy since it just leads to an ordinary least squares regression of $\Delta X_t + \alpha\ \beta' X_{t-k}$ on the lagged differences. Let us do this by first regressing ΔX_t on the lagged differences giving the residuals R_{0t} and then regressing X_{t-k} on the lagged differences giving the residuals R_{kt}. After having performed these regressions the partially maximised likelihood function or likelihood profile becomes proportional to

$$L(\alpha,\beta,\Lambda) = |\Lambda|^{-T/2}\exp\{-1/2 \sum_{t=1}^{T} (R_{0t} + \alpha\beta' R_{kt})'\Lambda^{-1}(R_{0t} + \alpha\beta' R_{kt})\}.$$

For fixed β we can maximise over α and Λ by a usual regression of R_{0t} on $-\beta' R_{kt}$ which gives the well known result

(2.4) $$\hat{\alpha}(\beta) = - S_{0k}\beta(\beta' S_{kk}\beta)^{-1},$$

and

(2.5) $$\hat{\Lambda}(\beta) = S_{00} - S_{0k}\beta(\beta' S_{kk}\beta)^{-1}\beta' S_{k0},$$

where we have defined product moment matrices of the residuals as

(2.6) $$S_{ij} = T^{-1} \sum_{t=1}^{T} R_{it} R_{jt}' \quad i,j = 0,k.$$

The likelihood profile now becomes proportional to

$$|\hat{\Lambda}(\beta)|^{-T/2}$$

and it remains to solve the minimisation problem

$$\min |S_{00} - S_{0k}\beta(\beta' S_{kk}\beta)^{-1}\beta' S_{k0}| \ ,$$

where the minimisation is over all p×r matrices β. The well known matrix relation, see Rao (1973),

$$|S_{00}| \, |\beta' S_{kk}\beta - \beta' S_{k0} S_{00}^{-1} S_{0k}\beta| =$$
$$|\beta' S_{kk}\beta| \, |S_{00} - S_{0k}\beta(\beta' S_{kk}\beta)^{-1}\beta' S_{k0}|$$

shows that we shall minimise

$$|\beta' S_{kk}\beta - \beta' S_{k0} S_{00}^{-1} S_{0k}\beta| / |\beta' S_{kk}\beta|$$

with respect to the matrix β.

We now let D denote the diagonal matrix of ordered eigenvalues $\hat{\lambda}_1 > \ldots > \hat{\lambda}_p$ of $S_{k0}S_{00}^{-1}S_{0k}$ with respect to S_{kk}, i.e. the solutions to the equation

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

and E the matrix of the corresponding eigenvectors, then

$$S_{kk} E D = S_{k0} S_{00}^{-1} S_{0k} E$$

where E is normalised such that

$$E' S_{kk} E = I.$$

Now choose $\beta = E\ \varphi$ where φ is p×r, then we shall minimise

$$|\varphi'\varphi - \varphi' D\varphi| / |\varphi'\varphi|.$$

This can be accomplished by choosing φ to be the first r unit vectors or by choosing $\hat{\beta}$ to be the first r eigenvectors of $S_{k0}S_{00}^{-1}S_{0k}$ with respect to S_{kk}, that is the first r columns of E. These are called the canonical variates and the eigenvalues are the squared canonical correlations of R_k with respect to R_0. For the details of these calculations the reader is referred to Anderson (1984) chapter 12. Note that all possible choices of the optimal β can be found from $\hat{\beta}$ by $\beta = \hat{\beta}\rho$ for ρ an r×r matrix of full rank. The estimators derived here are related to the NLS estimators given by Stock (1985). Note that $\hat{\beta}' S_{kk} \hat{\beta} = I$ such that the estimate of the other parameters are given by

$$\hat{\alpha} = -S_{0k}\hat{\beta}(\hat{\beta}' S_{kk}\hat{\beta})^{-1} = -S_{0k}\hat{\beta} \tag{2.8}$$

which clearly depends on the choice of the optimising β, whereas

$$\hat{\Pi} = -S_{0k}\hat{\beta}\hat{\beta}' \tag{2.9}$$

and

$$\hat{\Lambda} = S_{00} - S_{0k}\hat{\beta}\hat{\beta}' S_{k0} = S_{00} - \hat{\alpha}\hat{\alpha}' \tag{2.10}$$

and the maximised likelihood as given by

$$L_{max}^{-2/T} = |S_{00}| \prod_{i=1}^{r} (1-\hat{\lambda}_i) \tag{2.11}$$

do not depend on the choice of $\hat{\beta}$.

With this notation it is easy to express the estimates of Π and Λ without the constraint (2.1). These follow from (2.4) and (2.5) for β = I and give

$$\hat{\Pi} = -S_{0k}S_{kk}^{-1}$$

and

$$\hat{\Lambda} = S_{00} - S_{0k}S_{kk}^{-1}S_{k0}$$

as well as the expression for the determinant

$$|\hat{\Lambda}| = |S_{00}| \prod_{i=1}^{p}(1-\hat{\lambda}_i). \tag{2.12}$$

If we now want a test that there are at most r cointegrating vectors then the likelihood ratio test statistic is the ratio of (2.11) and (2.12) and can be expressed as

$$-2\ln Q = - T \sum_{i=r+1}^{p} \ln(1-\hat{\lambda}_i) \tag{2.13}$$

where $\hat{\lambda}_{r+1} > \ldots > \hat{\lambda}_p$ are the p-r smallest eigenvalues. This completes the proof of Theorem 1.

Notice how this analysis allows one to calculate all p eigenvalues and eigenvectors at once, and then make inference about the number of important cointegration relations, by testing how many of the λ's that are zero.

Next consider Theorem 2. It is apparent from the derivation of $\hat{\beta}$ that if $\beta = H\varphi$ is fixed, then regression of R_{0t} on $- \varphi'H'R_{kt}$ is still a simple linear regression and the analysis is as before with R_{kt} replaced by $H'R_{kt}$. Thus the matrix φ can be estimated as the eigenvectors corresponding to the r largest eigenvalues of $H'S_{k0}'S_{00}^{-1}S_{0k}H$ with respect to $H'S_{kk}H$, i.e. the solution to

$$(2.14) \qquad |\lambda H' S_{kk} H - H' S_{k0} S_{00}^{-1} S_{0k} H| = 0.$$

Let the s eigenvalues be denoted by λ_i^* , i=1,...,s. Then the likelihood ratio test of H_1 in H_0 can be found from two expressions like (2.11) and is given by

$$(2.15) \qquad -2\ln Q = T \sum_{i=1}^{r} \ln\{(1-\lambda_i^*)/(1-\hat{\lambda}_i)\},$$

which completes the proof of Theorem 2.

In the next section we shall find the asymptotic distribution of the test statistics (2.13) and (2.15) and show that the cointegration space, the impact matrix Π and the variance matrix Λ are estimated consistently.

3. ASYMPTOTIC PROPERTIES OF THE ESTIMATORS AND THE TEST STATISTICS.

In order to derive properties of the estimators we need to impose more precise conditions on the parameters of the model, such that they correspond to the situation we have in mind, namely of a process that is integrated of order 1, but still has r cointegration vectors β.

First of all we want all roots of $|A(z)| = 0$ to satisfy $|z| > 1$ or possibly $z = 1$. This implies that the non stationarity of the process can be removed by differencing. Next we shall assume that X_t is integrated of order 1, i.e. that ΔX_t is stationary and that the hypothesis (2.1) is satisfied by some non singular α and β. Correspondingly we can express ΔX_t in terms of the ϵ's by its moving average representation

$$\Delta X_t = \sum_{j=0}^{\infty} C_j \epsilon_{t-j}$$

for some exponentially decreasing coefficients C_j. Under suitable conditions on these coefficients it is known that this equation determines an error correction model of the form (2.3), where $\Gamma_k X_{t-k} = -\Pi X_{t-k}$ represents the error correction term containing the stationary components of X_{t-k}, i.e. $\beta' X_{t-k}$. Moreover the null space for $C = \sum_{j=0}^{\infty} C_j$ given by $\{\xi | \xi' C = 0\}$ is exactly the range space of Γ_k, i.e. the space spanned by the columns in β and vice versa. We thus have the following representations

$$\Pi = \alpha\beta' \text{ and } C = \gamma\varphi\delta'$$

where φ is $(p-r)x(p-r)$, γ and δ are $px(p-r)$ and all three are non singular, and $\gamma'\beta = \delta'\alpha = 0$. We shall later choose δ and γ in a convenient way, see Johansen (1985) or the references to Granger (1981) and Granger & Engle (1985), and Engle & Granger (1987) for the details of these results.

Let us now formulate

THEOREM 3 Under the hypothesis that there are r cointegrating vectors the estimate of the cointegration space as well as Π and Λ are consistent, and the likelihood ratio test statistic of this hypothesis is asymptotically distributed as

$$\mathrm{tr}\{\int_0^1 B dB' [\int_0^1 BB' du]^{-1} \int_0^1 dB B'\}$$

where B is a p-r dimensional Brownian motion with covariance matrix I.

In order to understand the structure of this limit distribution one should notice that if B is a Brownian motion with I as the covariance matrix, then the stochastic integral $\int_0^t B dB'$ is a matrix valued martingale, with quadratic variation process

$$\int_0^t \mathrm{Var}(B dB') = \int_0^t BB' du \otimes I$$

where the integral $\int_0^t BB' du$ is an ordinary integral of the continuous matrix valued process BB'. With this notation the limit distribution in Theorem 3 can be considered as a multivariate version of the square of a martingale $\int B dB'$ divided by its variance process $\int BB' du$. Notice that for $r = p-1$ i.e. for testing $p-1$ cointegration relations one obtains the limit distribution with a 1 dimensional Brownian motion, i.e.

$$\left(\int_0^1 B dB\right)^2 / \int_0^1 B^2 du = ((B(1)^2 - 1)/2)^2 / \int_0^1 B^2 du$$

which is the square of the usual "unit root" distribution see Dickey & Fuller (1976).

Table 1

A surprisingly accurate description of the results in Table 1 is obtained by approximating the distributions by $c\chi^2(f)$ for suitable values of c and f. By equating the mean of the distributions based on 10000 observations to those of a $c\chi^2$ with $f = 2m^2$ degrees of freedom we obtain values of c, and it turns out that we can use the empirical relation

$$c = .85 - .58/f.$$

Notice that the hypothesis of r cointegrating relations reduces the number of parameters in the Π matrix from p^2 to $rp + r(p-r)$, thus one could expect $(p - r)^2$ degrees of freedom if the usual asymptotics would hold. In the case of non stationary processes it is known that this does not hold but a very good approximation is given by the above choice of

$2(p - r)^2$ degrees of freedom.

THEOREM 4 The likelihood ratio test of the hypothesis

$$H_0 : \beta = H\varphi$$

of restricting the r dimensional cointegration space to an s dimensional subspace of R^p is asymptotically distributed as χ^2 with $r(p - s)$ degrees of freedom.

We shall now give the proof of these Theorems, through a series of intermediate results. We shall first give some expressions for variances and their limits, then show how the algorithm for deriving the maximum likelihood estimator can be followed by a probabilistic analysis ending up with the asymptotic properties of the estimator and the test statistics.

We can represent X_t as $X_t = \sum_{j=1}^{t} \Delta X_j$, where X_0 is a constant which we shall take to be zero to simplify the notation. We shall describe the stationary process ΔX_t by its covariance function

$$\psi(i) = \mathrm{Var}(\Delta X_t, \Delta X_{t+i})$$

and we define the matrices

$$\mu_{ij} = \psi(i-j) = E(\Delta X_{t-i}, \Delta X'_{t-j}) \ , \ i,j = 0,\dots,k-1$$

$$\mu_{ki} = \sum_{j=k-i}^{\infty} \psi(j) \quad i = 0,\dots,k-1$$

and

$$\mu_{kk} = - \sum_{j=-\infty}^{\infty} |j| \psi(j).$$

Finally define

$$\psi = \sum_{j=-\infty}^{\infty} \psi(j).$$

Note the following relations

$$\psi(i) = \sum_{j=0}^{\infty} C_j \Lambda C'_{j+i},$$

$$\psi = \sum_{j=0}^{\infty} C_j \Lambda \sum_{j=0}^{\infty} C'_j = C\Lambda C',$$

$$\mathrm{Var}(X_{t-k}) = \sum_{j=-t+k}^{t-k} (t-k-|j|)\Psi(j),$$

$$\mathrm{Cov}(X_{t-k}, \Delta X_{t-i}) = \sum_{j=k-i}^{t-i} \Psi(j)$$

which show that

$$\mathrm{Var}(X_T/T^{1/2}) \to \sum_{i=-\infty}^{\infty} \Psi(i) = \Psi,$$

and

$$\mathrm{Cov}(X_{T-k}, \Delta X_{T-i}) \to \sum_{j=k-i}^{\infty} \Psi(j) = \mu_{ki}$$

whereas the relation

$$\mathrm{Var}(\beta' X_{T-k}) = (T-k) \sum_{j=-T+k}^{T-k} \beta'\Psi(j)\beta - \sum_{j=-T+k}^{T-k} |j| \beta'\Psi(j)\beta.$$

shows that

$$\mathrm{Var}(\beta' X_{T-k}) \to \beta' \mu_{kk} \beta,$$

since $\beta' C = 0$ implies that $\beta'\psi = 0$, such that the first term vanishes in the limit. Note that the non stationary part of X_t makes the variance matrix tend to infinity, except for the directions given by the vectors in β, since $\beta' X_t$ is stationary.

The calculations involved in the maximum likelihood estimation all center around the product moment matrices

$$M_{ij} = T^{-1}\sum_{t=1}^{T} \Delta X_{t-i}\Delta X_{t-j}' \quad ,i,j = 0,\dots,k-1,$$

$$M_{ki} = T^{-1}\sum_{t=1}^{T} X_{t-k}\Delta X_{t-i}' \quad ,i = 0,\dots,k-1,$$

and

$$M_{kk} = T^{-1}\sum_{t=1}^{T} X_{t-k}\, X_{t-k}'.$$

We shall first give the asymptotic behaviour of these matrices, then find the asymptotic properties of S_{ij} and finally apply these results to the estimators and the test statistic. The methods are inspired by Phillips (1985) even though I shall stick to the Gaussian case, which make the results somewhat simpler.

In order to formulate the results we need a Brownian motion W in p dimensions with covariance function $t\Lambda$.

LEMMA 1. As $T \to \infty$ we have

(3.1) $$T^{-1/2}X_{[Tt]} \xrightarrow{w} CW(t)$$

(3.2) $$M_{ij} \to \mu_{ij}, \quad i,j = 0,\dots,k-1$$

(3.3) $$M_{ki} \to C\int_0^1 WdW'C' + \mu_{ki} \quad i= 0,\dots,k-1$$

(3.4) $$\beta' M_{kk}\beta \to \beta'\mu_{kk}\beta$$

(3.5) $$T^{-1}M_{kk} \to C\int_0^1 W(u)W'(u)du\ C'.$$

Note that for any $\xi \in R^p$, $\xi' M_{kk}\xi$ is of the order of T unless ξ is in the space spanned by β, in which case it is convergent. Note also that the stochastic integrals enter as limits of the non stationary part of the process X_t, and that they disappear when multiplied by β, since $\beta' C = 0$.

Proof. We shall use the fact that

$$T^{-1/2}\sum_{j=0}^{[Tt]} \epsilon_j \xrightarrow{w} W(t) \text{ as } T \to \infty.$$

From the representation

$$X_t = \sum_{j=0}^{t} \Delta X_j = \sum_{j=0}^{t} \sum_{i=0}^{\infty} C_i \epsilon_{j-i} =$$

$$(\sum_{i=0}^{\infty} C_i)(\sum_{s=0}^{t} \epsilon_s) + \sum_{i=0}^{\infty} C_i \sum_{s=-i}^{-1} \epsilon_s - \sum_{i=0}^{\infty} C_i \sum_{s=t-i+1}^{t} \epsilon_s.$$

We find with t replaced by [Tt] and by dividing by $T^{1/2}$ that the first term on the right hand side converges to CW and that the last two terms tend to zero. This proves (3.1).

The result (3.2) follows by noting that since $\{\epsilon_t\}$ are i.i.d., then $\{\Delta X_t\}$ is ergodic and hence

$$M_{ij} = T^{-1}\sum_{t=1}^{T} \Delta X_{t-i} \Delta X'_{t-j} \to E(\Delta X_{t-i} \Delta X'_{t-j}) = \mu_{ij}, \quad i,j = 0,\dots,k-1.$$

To prove (3.3) we need the following representation

$$M_{ki} = T^{-1}\sum_{t=1}^{T} X_{t-k} \Delta X'_{t-i} = T^{-1}\sum_{t=1}^{T} \sum_{j=1}^{t-k} \Delta X_j \Delta X'_{t-i}$$

$$= T^{-1}\sum_{t=1}^{T} \sum_{j=1}^{t-k} \sum_{\upsilon=0}^{\infty} \sum_{\mu=0}^{\infty} C_\upsilon \epsilon_{j-\upsilon} \epsilon'_{t-i-\mu} C'_\mu.$$

Now consider the term for each value of υ and μ without the coefficients C_υ and C_μ. We then get if $t-k-\upsilon \geq t-i-\mu$ (or $k+\upsilon \leq i+\mu$)

$$T^{-1}\sum_{t=1}^{T} \sum_{s=-\upsilon}^{t-i-\mu-1} \epsilon_s \epsilon'_{t-i-\mu} + T^{-1}\sum_{t=1}^{T} \epsilon_{t-i-\mu} \epsilon'_{t-i-\mu} + T^{-1}\sum_{t=1}^{T} \sum_{s=k+\upsilon}^{i+\mu-1} \epsilon_{t-s} \epsilon'_{t-i-\mu}$$

which converges to

$$\int_0^1 W dW' + \Lambda + 0.$$

If $t-k-\upsilon < t-i-\mu$ then the first term is the same and the remaining terms now become

$$-T^{-1}\sum_{t=1}^{T}\sum_{s=i+\mu+1}^{k+\upsilon}\epsilon_{t-s}\epsilon'_{t-i-\mu}$$

which converges to 0 as $T\to\infty$. Collecting the terms we get that

$$M_{ki}\to\sum_{\upsilon=0}^{\infty}C_\upsilon\int_0^1 WdW'\sum_{\mu=0}^{\infty}C'_\mu+\sum_{\upsilon=0}^{\infty}\sum_{\mu=\upsilon+k-i}^{\infty}C_\upsilon\Lambda C'_\mu.$$

The first term is just $C\int_0^1 WdW'C'$ and the last term is $\sum_{\mu=0}^{\infty}\Psi(\mu+k-i)=\mu_{ki}$.

The relation (3.4) follows since $\beta'X_t$ is stationary.

Finally we shall show (3.5). From the weak convergence of $T^{-1/2}X_{[Tt]}$ to CW it follows by the continuous mapping theorem that

$$\int_0^1 T^{-1/2}X_{[Tt]}\,T^{-1/2}X'_{[Tt]}\,dt=T^{-1}\sum_{t=1}^{T}(X_{t-k}/T^{1/2})(X'_{t-k}/T^{1/2})=T^{-1}M_{k,k}$$

converges to $C\int_0^1 W(u)W'(u)duC'$.

This completes the proof of Lemma 1. We shall now apply the results to find the asymptotic properties of S_{ij}, $i,j=0,k$, see (2.6). These can be expressed in terms of the M_{ij}'s as follows:

$$S_{ij}=M_{ij}-M_{i*}M_{**}^{-1}M_{*j}\quad i,j=0,k$$

where

$$M_{**}=\{M_{ij},i,j=1,\dots,k-1\}$$

$$M_{k*}=\{M_{ki},i=1,\dots,k-1\}$$

and

$$M_{0*}=\{M_{0i},i=1,\dots,k-1\},$$

A similar notation is introduced for the μ_{ij}'s. It is convenient to have the notation

$$\Sigma_{ij} = \mu_{ij} - \mu_{i*}\mu_{**}^{-1}\mu_{*j} \quad i,j = 0,k.$$

We now get

LEMMA 2. The following relations hold

$$\Sigma_{00} = \Gamma_k\Sigma_{k0} + \Lambda \tag{3.6}$$

$$\Sigma_{0k}\Gamma_k' = \Gamma_k\Sigma_{kk}\Gamma_k' \tag{3.7}$$

and hence since $\Gamma_k = -\alpha\beta'$

$$\Sigma_{00} = \alpha(\beta'\Sigma_{kk}\beta)\alpha' + \Lambda. \tag{3.8}$$

Proof. From the defining equation for the process X_t we find the equations

$$M_{0i} = \Gamma_1 M_{1i} + \dots + \Gamma_{k-1}M_{k-1,i} + \Gamma_k M_{ki} + T^{-1}\sum_{t=1}^{T}\epsilon_t \Delta X_{t-i}' \tag{3.9}$$

$i = 0,1,\dots,k-1$

$$M_{0k} = \Gamma_1 M_{1k} + \dots + \Gamma_{k-1}M_{k-1,k} + \Gamma_k M_{kk} + T^{-1}\sum_{t=1}^{T}\epsilon_t X_{t-k}'. \tag{3.10}$$

Now let $T \to \infty$, then we get the equations

$$\mu_{00} = \Gamma_1\mu_{10} + \dots + \Gamma_{k-1}\mu_{k-1,0} + \Gamma_k\mu_{k0} + \Lambda \tag{3.11}$$

$$\mu_{0i} = \Gamma_1\mu_{1i} + \dots + \Gamma_{k-1}\mu_{k-1,i} + \Gamma_k\mu_{ki}, \quad i = 1,\dots,k-1 \tag{3.12}$$

$$\mu_{0k}\beta = \Gamma_1\mu_{1k}\beta + \dots + \Gamma_{k-1}\mu_{k-1,k}\beta + \Gamma_k\mu_{kk}\beta \tag{3.13}$$

If we solve the equations (3.12) for the matrices Γ_* and insert into (3.11) and (3.13) we get (3.6) and (3.7).

We shall now find the asymptotic properties of S_{ij}.

LEMMA 3 For $T \to \infty$ it holds, that if δ is chosen such that $\delta'\alpha = 0$, then

(3.14) $$S_{00} \to \Sigma_{00}$$

(3.15) $$\delta' S_{0k} \to \delta' \int_0^1 dW W' C'$$

(3.16) $$\beta' S_{k0} \to \beta' \Sigma_{k0}$$

(3.17) $$T^{-1} S_{kk} \to C \int_0^1 W(u) W'(u) du C'$$

(3.18) $$\beta' S_{kk} \beta \to \beta' \Sigma_{kk} \beta.$$

Proof. All relations follow from Lemma 1 except the second. If we solve for Γ_* in the equations (3.9) and insert the solution into (3.10) and use the definition of S_{ij} in terms of the M's, then we get

(3.19) $$S_{0k} = T^{-1}\sum_{t=1}^{T} \epsilon_t X_{t-k}' + \Gamma_k S_{kk} - \sum_{i=1}^{k-1}\sum_{j=1}^{k-1} T^{-1} \sum_{t=1}^{T} \epsilon_t \Delta X_{t-i}' M^{ij} M_{jk}.$$

The last term goes to zero as $T \to \infty$, since ϵ_t and $\Delta X_{t-i}'$ are stationary and uncorrelated. The second term vanishes when multiplied by δ', since $\delta'\Gamma_k = -\delta'\alpha\beta' = 0$, and the first term converges to the integral as stated.

We shall now turn to the proof of Theorem 3.

We let $P_\alpha(\Lambda)$ denote the projection of R^p onto the column space spanned α with respect to the matrix Λ^{-1}, i.e.

$$P_\alpha(\Lambda) = \alpha(\alpha'\Lambda^{-1}\alpha)^{-1}\alpha'\Lambda^{-1}$$

We then choose δ $(p\times(p-r))$ of full rank to satify

$$\delta\delta' = \Lambda^{-1}(I - P_\alpha(\Lambda)).$$

Note that $\delta'\alpha = 0$, and that $\delta'\Lambda\delta = I$ of dimension $(p-r)\times(p-r)$. Note also that $P_\alpha(\Lambda) = P_\alpha(\Sigma_{00})$ since Σ_{00} is given by (3.8). This relation is well known from the theory of random coefficient regression, see Rao (1965) or Johansen (1984). Similarly we choose γ $(p\times(p-r))$ of full rank to satisfy

$$\gamma\gamma' = \Lambda^{-1}(I - P_\beta(\Lambda))$$

such that $\gamma'\beta = 0$. Note that the matrices (γ,β) and (δ,α) have full rank p.

We want to express the estimation problem in the coordinates given by the p vectors in β and γ. This can be done as follows:

The maximum likelihood estimation involves finding λ as solution to the equation

$$\left|\lambda \begin{bmatrix} \beta'S_{kk}\beta & \beta'S_{kk}\gamma \\ \gamma'S_{kk}\beta & \gamma'S_{kk}\gamma \end{bmatrix} - \begin{bmatrix} \beta'S_{k0}S_{00}^{-1}S_{0k}\beta & \beta'S_{k0}S_{00}^{-1}S_{0k}\gamma \\ \gamma'S_{k0}S_{00}^{-1}S_{0k}\beta & \gamma'S_{k0}S_{00}^{-1}S_{0k}\gamma \end{bmatrix}\right| = 0.$$

We shall first discuss the eigenvalues. The eigenvalues are bounded between 0 and 1, and for $T \to \infty$ we can find which limit points are possible. Let $S = S(\lambda) = \lambda S_{kk} - S_{k0}S_{00}^{-1}S_{0k}$, where λ has been chosen as an eigenvalue, so that $|S| = 0$, then

$$(3.20) \qquad 0 = \left|\binom{\beta'}{\gamma'}S(\beta\ \gamma)\right| = |\gamma'S\gamma|\,|\beta'S\beta - \beta'S\gamma(\gamma'S\gamma)^{-1}\gamma'S\beta|$$
$$= |\beta'S\beta|\,|\gamma'S\gamma - \gamma'S\beta(\beta'S\beta)^{-1}\beta'S\gamma|.$$

As $T \to \infty$ the term $\gamma'S_{kk}\gamma \to \infty$, see (3.17). Now take a subsequence T' such that $\lambda = \lambda(T') \to \upsilon > 0$, then $|\gamma'S\gamma| \to \infty$ and we get from the first decomposition in (3.20) that for T' sufficiently large the second factor must be zero, i.e.

$$(3.21)\qquad |\beta'S\beta - \beta'S\gamma(\gamma'S\gamma)^{-1}\gamma'S\beta| = 0.$$

Using the results (3.14),(3.16) and (3.18) from Lemma 3, we find that in the limit υ must satisfy the equation

$$(3.22)\qquad |\upsilon\beta'\Sigma_{kk}\beta - \beta'\Sigma_{k0}\Sigma_{00}^{-1}\Sigma_{0k}\beta| = 0.$$

If on the other hand $\lambda(T') \to 0$, then it is seen that $T'\lambda(T')$ tends to some constant υ say. The second decomposition in (3.20) shows by using (3.14),(3.16) and (3.18) that since the first factor converges to $|\beta'\Sigma_{k0}\Sigma_{00}^{-1}\Sigma_{0k}\beta|$, then for T' sufficiently large we have that the second factor is zero:

$$(3.23)\qquad |\gamma'S\gamma - \gamma'S\beta(\beta'S\beta)^{-1}\beta'S\gamma| = 0.$$

Using the results from Lemma 3 again we find that in the limit υ must satisfy the equation

$$(3.24)\qquad |\upsilon\gamma'C\int_0^1 W(u)W'(u)du\gamma C' - \gamma'N\gamma| = 0$$

where

$$N=\lim\{S_{k0}S_{00}^{-1}S_{0k} - S_{k0}S_{00}^{-1}S_{0k}\beta(\beta'S_{k0}S_{00}^{-1}S_{0k}\beta)^{-1}\beta'S_{k0}S_{00}^{-1}S_{0k}\}.$$

We can now apply Lemma 3 and find that

$$S_{00}^{-1} - S_{00}^{-1}S_{0k}\beta(\beta'S_{k0}S_{00}^{-1}S_{0k}\beta)^{-1}\beta'S_{k0}S_{00}^{-1}$$

converges to the same expression with S replaced by Σ. Now apply (3.6) and (3.7) to show that the limit equals

$$\Sigma_{00}^{-1}(I - P_{\alpha}(\Sigma_{00})) = \Lambda^{-1}(I - P_{\alpha}(\Lambda)) = \delta\delta'.$$

Finally from Lemma 3 we also find the limit of $\delta'S_{0k}$ and hence that the limit υ must satisfy the equation

$$|\upsilon\gamma'C\int_0^1 W(u)W'(u)duC'\gamma - \gamma'C\int_0^1 WdW'\delta\delta'\int_0^1 dWW'C'\gamma| = 0.$$

The representation

$$C = \gamma\varphi\delta'$$

for some non singular matrix φ now implies since $|\gamma'\gamma| \neq 0$ and $|\varphi| \neq 0$ that

$$\left|\nu\delta'\int_0^1 W(u)W(u)'du\delta - \delta'\int_0^1 WdW'\delta\delta'\int_0^1 dWW'\delta\right| = 0$$

Now $B = \delta'W$ is a Brownian motion with variance $\delta'\Lambda\delta = I$, which shows that

$$\left|\nu\int_0^1 B(u)B(u)'du - \int_0^1 BdB'\int_0^1 dBB'\right| = 0 \tag{3.25}$$

We have now seen that the possible limit points of $\lambda(T')$ are the eigenvalues of (3.22) or 0, and that if 0 then the limits $T'\lambda(T')$ must satisfy the equation (3.25).

Now let T' be chosen such that $\hat{\lambda}_1(T'),\ldots,\hat{\lambda}_p(T')$ all converge and if to zero then also $T'\lambda(T')$ converges. The limiting values have to be eigenvalues in the above matrices (3.22) and (3.25), and there are a total of $r+(p-r) = p$ such eigenvalues. This shows that the limit points are uniquely defined, and hence that $\lambda_i(T)$ is convergent and if to zero then so is $T\lambda_i(T)$, and that the r largest eigenvalues converge to those determined by (3.22).

The test statistic

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

will converge to the sum of the eigenvalues given by (3.25), which gives the second statement of Theorem 2. Next consider the eigenvectors.

An eigenvector $\hat{v}' = (\hat{x}', \hat{y}')$ satisfies the equations

$$\hat{\lambda}\beta' S_{kk}\beta\hat{x} + \hat{\lambda}\beta' S_{kk}\gamma\hat{y} = \beta' S_{k0} S_{00}^{-1} S_{0k}\beta\hat{x} + \beta' S_{k0} S_{00}^{-1} S_{0k}\gamma\hat{y}$$
$$\hat{\lambda}\gamma' S_{kk}\beta\hat{x} + \hat{\lambda}\gamma' S_{kk}\gamma\hat{y} = \gamma' S_{k0} S_{00}^{-1} S_{0k}\beta\hat{x} + \gamma' S_{k0} S_{00}^{-1} S_{0k}\gamma\hat{y}.$$

Let now $\hat{\lambda}$ equal one of the r largest eigenvalues. We see that since $\hat{\lambda}\gamma' S_{kk}\gamma \to \infty$ we have that $\hat{y}$ must go to zero like T^{-1}. Thus the component of the eigenvector which does not belong to the cointegration space must go to zero like T^{-1}. In this sense the cointegration space is estimated consistently.

In the limit $\hat{x}$ must satisfy the equation

$$\lambda\beta'\Sigma_{kk}\beta x = \beta'\Sigma_{k0}\Sigma_{00}^{-1}\Sigma_{0k}\beta x,$$

i.e. be an eigenvector of the equation (3.22). Let now $x = (x_1, \ldots, x_r)$ denote all the eigenvectors in (3.22), then we have seen that $\hat{\beta} = \beta\hat{x} + \gamma\hat{y} \to \beta x$. With this result it follows from (2.4) that

$$\hat{\alpha} \to -\Sigma_{0k}\beta x$$

whereas (2.8) gives

$$\hat{\Pi} \to -\Sigma_{0k}\beta x x'\beta' = -\Sigma_{0k}\beta(\beta'\Sigma_{kk}\beta)^{-1}\beta' = \alpha\beta' = \Pi.$$

We also find

$$\hat{\Lambda} \to \Sigma_{00} - \Sigma_{0k}\beta x x'\beta'\Sigma_{k0} = \Sigma_{00} - \alpha(\beta'\Sigma_{kk}\beta)\alpha' = \Lambda.$$

This completes the proof of Theorem 3, and we shall therefore turn to the proof of Theorem 4.

The proof of consistency of the eigenvalues and the eigenspace as well as Π and Λ is the same as before. Recall that for Theorem 4 we assume that $\beta = H\varphi$. We let λ^* denote any of the r largest eigenvalues given by (2.13) and $\hat{\lambda}$ and $\hat{e}$ denotes the corresponding eigenvalue and eigenvector from (2.7).

From (3.21) we find with $S(\hat{\lambda}) = \hat{\lambda}S_{kk} - S_{k0}S_{00}^{-1}S_{0k}$ that

$$|\beta'S(\hat{\lambda})\beta - \beta'S(\hat{\lambda})\gamma(\gamma'S(\hat{\lambda})\gamma)^{-1}\gamma'S(\hat{\lambda})\beta| = 0 \tag{3.26}$$

where γ is chosen as above.

Similarly λ^* has to satisfy the equation

$$|\varphi'H'S(\lambda^*)H\psi - \varphi'H'S(\lambda^*)H\psi(\psi'H'S(\lambda^*)H\psi)^{-1}H'\psi'S(\lambda^*)H\varphi| = 0, \tag{3.27}$$

where $\psi(s\times(s-r))$ is so chosen that $(\varphi,\psi)(s\times s)$ is of full rank, i.e. such that $H(\varphi,\psi)$ spans the space sp(H). Note that $H\varphi = \beta$, which simplifies the expression (3.27). Note also that we can in fact choose ψ such that $H\psi \in sp(H)\cap sp(\gamma)$, i.e. of the form $H\psi = \gamma\eta$. This representation will be useful later. We can now write the equations as follows

$$|\beta'S(\hat{\lambda})\beta - \beta'S(\hat{\lambda})P_{\gamma}(S(\hat{\lambda})^{-1})\beta|, \tag{3.28}$$

and

$$|\beta'S(\lambda^*)\beta - \beta'S(\lambda^*)P_{H\psi}(S(\lambda^*)^{-1})\beta| = 0. \tag{3.29}$$

The first term in the two expressions is O(1) and the second expression is in both cases $O(T^{-1})$, since $T^{-1}\gamma'S_{kk}\gamma$ is convergent. Thus it is to be expected that the difference in the eigenvalues is of the order T^{-1}. Thus we put $\hat{\lambda} = \lambda^* + \rho/T$, and we now want to expand (3.28) around the point λ^*. For this we need the following lemma

LEMMA 4 Let A be a p×p symmetric matrix with eigenvalues $\lambda_1 > \dots > \lambda_{p-1} > \lambda_p = 0$. and corresponding eigenvectors e_i $i = 1,\dots,p$. Let B denote a (p×p) matrix, then

$$|A + tB| = t \prod_{i=1}^{p-1} \lambda_i \; e_p'Be_p + o(t). \tag{3.30}$$

Proof. This follows by diagonalising the matrix A, and expanding the determinant, starting with the terms in the diagonal, since all other terms will be of lower order.

We now write (3.28) as

$$|\beta' S(\lambda^*)\beta - \beta' S(\lambda^*)P_{H\psi}(S(\lambda^*)^{-1})\beta +$$

$$T^{-1}\beta'\{\rho S_{kk} + TS(\lambda^*)P_{H\psi}(S(\lambda^*)^{-1}) - TS(\hat{\lambda})P_{\gamma}(S(\hat{\lambda})^{-1})\}\beta| = 0.$$

Now expand using Lemma 4 with $t = T^{-1}$. Since the other eigenvalues of the equation (3.29) are different from λ^* we find that the first term in the expansion (3.30) can not be zero, hence the second must be zero, which gives the equation for ρ:

$$\hat{\rho}\hat{e}'\beta' S_{kk}\beta\hat{e} + T\hat{e}'\beta'\{S(\lambda^*)P_{H\psi}(S(\lambda^*)^{-1}) - S(\hat{\lambda})P_{\gamma}(S(\hat{\lambda})^{-1})\}\beta\hat{e} \simeq 0.$$

The coefficient of ρ will converge to 1, because of the normalisation of the eigenvectors. Thus we have the following equation for the determination of the limiting distribution of ρ

$$(3.31)\quad T(\hat{\lambda} - \hat{\lambda}^{*}) = \rho \simeq T\hat{e}'\beta'\{S(\hat{\lambda})P_{\gamma}(S(\hat{\lambda})^{-1}) - S(\lambda^{*})P_{H\psi}(S(\lambda^{*})^{-1})\}\beta\hat{e}$$

Note that since $sp(H\psi) \in sp(\gamma)$, the expression in { } is approximately a projection onto a complement of $sp(H\psi)$ in $sp(\gamma)$ which has dimension p - s. Hence in the limit $\rho \geq 0$, corresponding to the idea that by choosing the eigenvectors in R^{p} one can achieve larger eigenvalues than if we restrict the eigenvectors to sp(H).

We shall use the above representation (3.31) of ρ to find its limiting distribution. Let us first consider the expression

$$(3.32)\qquad T\hat{e}'\beta'S(\hat{\lambda})P_{\gamma}(S(\hat{\lambda})^{-1})\beta\hat{e} = \hat{e}'\beta'S(\hat{\lambda})\gamma(\gamma'S(\hat{\lambda})\gamma/T)^{-1}\gamma'S(\hat{\lambda})\beta\hat{e}.$$

It follows from (3.17) that the middle factor in (3.32) can be evaluated as

$$T^{-1}\gamma'S(\hat{\lambda})\gamma = \hat{\lambda}T^{-1}\gamma'S_{kk}\gamma - T^{-1}\gamma'S_{k0}S_{00}^{-1}S_{0k}\gamma \rightarrow \lambda\gamma'C\int_0^1 W(u)W'(u)du C'\gamma$$

From (3.19) we find that

$$S_{0k}\gamma = T^{-1}\sum_{t=1}^{T}\epsilon_t X'_{t-k}\gamma + \Gamma_k S_{kk}\gamma + o_P(1) = \int_0^1 dWW'C'\gamma - \alpha\beta'S_{kk}\gamma + o_P(1).$$

Then we get the first factor in (3.32)

$$(3.33)\quad \hat{e}'\beta'S(\hat{\lambda})\gamma = \hat{\lambda}\hat{e}'\beta'S_{kk}\gamma - \hat{e}'\beta'S_{k0}S_{00}^{-1}\{-\alpha\beta'S_{kk}\gamma + \int_0^1 dWW'C'\gamma + o_P(1)\}$$

$$= \{\hat{\lambda}e' + e'\beta'S_{k0}S_{00}^{-1}\alpha\}\beta'S_{kk}\gamma - e'\beta'S_{k0}S_{00}^{-1}\int_0^1 dWW'C'\gamma + o_P(1).$$

In the limit the first term has coefficient

$$\lambda e' + e'\beta'\Sigma_{k0}\Sigma_{00}^{-1}\alpha.$$

We shall use the fact that e is an eigenvector together with (3.7) to show that this is zero.

We have in fact

$$\lambda e'\beta\Sigma_{kk}\beta = e'\beta'\Sigma_{k0}\Sigma_{00}^{-1}\Sigma_{0k}\beta = -e'\beta'\Sigma_{k0}\Sigma_{00}^{-1}\alpha(\beta'\Sigma_{kk}\beta),$$

which shows that the coefficient to $\beta'S_{kk}\gamma$ tends to zero. Thus we find by collecting the results (3.32) and (3.33) and inserting them into (3.31) that

$$T\hat{e}'\beta'S(\hat{\lambda})P_{\gamma}(S(\hat{\lambda})^{-1})\beta\hat{e}$$

is asymptotically distributed as

$$(3.34)\quad \lambda^{-1}e'\beta'\Sigma_{k0}\Sigma_{00}^{-1}\int_0^1 dWW'C'\gamma(\gamma'C\int_0^1 WW'duC'\gamma)^{-1}\gamma'C\int_0^1 WdW'\Sigma_{00}^{-1}\Sigma_{0k}\beta e$$

and a similar expression is valid for the other component of ρ in (3.31) only γ is replaced by $\gamma\eta$, where $\eta(p-r,s-r)$ is chosen of full rank such that $\gamma\eta \in sp(H)$.

Now define

$$B_i = (\lambda_i(1-\lambda_i))^{-1/2} e_i'\beta'\Sigma_{k0}\Sigma_{00}^{-1}W \in R^1 \quad i= 1,\ldots,r.$$

and

$$A = \gamma'CW \in R^{p-r},$$

then A and B are multivariate Brownian motions. We then calculate

$$\mathrm{Cov}(B_i,A) = (\lambda_i(1-\lambda_i))^{-1/2} e_i'\beta'\Sigma_{k0}\Sigma_{00}^{-1}\Lambda C'\gamma.$$

From the relations (3.7) and (3.8) , i.e.

$$\Sigma_{0k}\beta = \alpha(\beta'\Sigma_{kk}\beta)\alpha'$$

and

$$\Lambda = \Sigma_{00} + \alpha(\beta'\Sigma_{kk}\beta)\alpha'$$

one finds that $\alpha'\delta = 0$ implies that the covariance is zero and hence that A and B are independent. Using the same relation and the eigenvalue properties one can show that $V(B_i) = 1$, $i = 1,\ldots,r$.

The expression for the likelihood ratio test is

$$-2\ln Q = T \sum_{i=1}^{r} \ln\{(1 - \lambda_i^*)/(1 - \hat{\lambda}_i)\} \simeq \sum_{i=1}^{r} T(\hat{\lambda}_i - \lambda_i^*)/(1 - \hat{\lambda}_i)$$

and from (3.31) we then find

$$-2\ln Q \simeq \sum_{i=1}^{r} \int_0^1 A'dB_i\{\int_0^1 AA'du\}^{-1}\int_0^1 AdB_i - \int_0^1 A'dB_i\eta\{\eta'\int_0^1 AA'du\eta\}^{-1}\eta'\int_0^1 AdB_i .$$

We shall first find the distribution of this approximating statistic for given value of A. Now notice that the terms in the sum are independent and that for fixed A, $Y_i = \int_0^1 AdB_i$ is p - r dimensional Gaussian with mean 0 and variance $V = \int_0^1 AA'du$. Hence

$$Y'V^{-1}Y = Y'(V^{-1} - \eta(\eta'V\eta)^{-1}\eta')Y + Y'\eta(\eta'V\eta)^{-1}\eta'Y$$

is a decomposition of the χ^2 distribution with p - r degrees of freedom on the left into two independent χ^2 distributions with degrees of freedom p - s and s - r respectively. Thus the distribution of the approximation to the likelihood ratio test statistic is, for fixed A, given by χ^2 with r(p - s) degrees of freedom. Since this distribution does not involve A it follows that the limiting distribution of the likelihood ratio test statistic is the same. This completes the proof of Theorem 4.

ACKNOWLEDGEMENT

The simulations were carefully performed by Marc Andersen with the support of the Danish Social Science Research Council.

REFERENCES.

-Anderson, T.W. (1984): An introduction to multivariate statistical analysis. Wiley, New York.

-Davidson, J. (1986): Cointegration in linear dynamic systems. Mimeo LSE 46.pp

-Dickey, D.A. & Fuller, W.A. (1976): Distribution of the estimator for autoregressive time series with a unit root. J. Amer. Statist. Assoc. 74, 427-431.

-Engle, R.F. & Granger, C.W.J. (1987): Co-integration and error correction: representation, estimation and testing. Econometrica 55,251-276.

-Granger, C.J. (1981): Some properties of time series data and their use in econometric model specification. Journal of Econometrics 16 121 -130.

-Granger, C.W.J. & Engle, R.F. (1985): Dynamic model specification with equlibrium constraints. Mimeo UCSD.

-Granger, C.W.J. & Weiss, A.A. (1983) Time series analysis of error correction models. In Studies in economic time series and multivariate statistics. eds S.Karlin,T.Amemiya & L.A.Goodman, Academic Press, New York.

-Johansen, S. (1984): Functional relations, random coefficients, and nonlinear regression with application to kinetic data. Lecture Notes in Statistics. Springer, New York.

-Johansen, S. (1985): The mathematical structure of error correction models.Preprint KUIMS 37 pp.

-Phillips, P.C.B. & Durlauf, S.N. (1985): Multiple time series regression with integrated processes. Cowles Foundation discussion Paper No. 768.

-Phillips, P.C.B. (1985): Understanding spurious regression in econometrics. Cowles Foundation discussion paper No. 757.

-Phillips, P.C.B. & Park, J.Y. (1986): Asymptotic equivalence of OLS and GLS in regression with integrated regressorsCowles Foundation discussion paper No.802

-Phillips, P.C.B. & Ouliaris, S. (1986) : Testing for cointegration. Cowles Foundation discussion paper No 809.

-Phillips, P.C.B. & Park, J.Y. (1986): Statistical inference in regressions with integrated processes : Part 1. Cowles Foundation discussion paper No. 811.

-Phillips, P.C.B. & Park, J.Y. (1987): Statistical inference in regressions with integrated processes : Part 2. Cowles Foundation discussion paper No. 819.

-Rao, C.R. (1965):The theory of least squares when the parameters are stochastic and its applications to the analysis of growth curves. Biometrika 52,447-458.

-Rao, C.R. (1973): Linear statistical inference and its applications, 2. Wiley, New York.

-Sims, A., Stock, J.H. & Watson, M.W. (1986): Inference in linear time series models with some unit roots. Preprint.

-Stock, J.H. & Watson, M.W. (1987): Testing for common trends. Working paper in Econometrics, the Hoover Institution, Stanford.

Table 1

The quantiles in the distribution of the test statistic

$$tr\{\int_0^1 dBB'(\int_0^1 B(u)B(u)'du)^{-1}\int_0^1 BdB'\}$$

where B is an m-dimensional Brownian motion with covariance matrix I. The Table is constructed from 10.000 simulations using the random number generator in Poly Pascal 8087. The uncertainty is about 0.1.

m	2.5%	5%	10%	50%	90%	95%	97.5%
1	0.0	0.0	0.0	0.6	2.9	4.2	5.3
2	1.6	1.9	2.5	5.4	10.3	12.0	13.9
3	7.0	7.8	8.8	14.0	21.2	23.8	26.1
4	16.0	17.4	19.2	26.3	35.6	38.6	41.2
5	28.3	30.4	32.8	42.1	53.6	57.2	60.3

PREPRINTS 1986

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK, TELEPHONE + 45 1 35 31 33.

No. 1 Jespersen, N.C.B.: On the Structure of Simple Transformation Models.

No. 2 Dalgaard, Peter and Johansen, Søren: The Asymptotic Properties of the Cornish-Bowden-Eisenthal Median Estimator.

No. 3 Jespersen, N.C.B.: Dichotomizing a Continuous Covariate in the Cox Regression Model.

No. 4 Asmussen, Søren: On Ruin Problems and Queues of Markov-Modulated M/G/1 Type.

No. 5 Asmussen, Søren: The Heavy Traffic Limit of a Class of Markovian Queueing Models.

No. 6 Jacobsen, Martin: Right Censoring and the Kaplan - Meier and Nelson - Aalen Estimators.

PREPRINTS 1987

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK, TELEPHONE + 45 1 35 31 33.

No. 1 Jensen, Søren Tolver and Johansen, Søren: Estimation of Proportional Covariances.

No. 2 Rootzén, Holger: Extremes, Loads, and Strengths.

No. 3 Bertelsen, Aksel: On the Problem of Testing Reality of a Complex Multivariate Normal Distribution.

No. 4 Gill, Richard D. and Johansen, Søren: Product-Integrals and Counting Processes.

No. 5 Leadbetter, M.R. and Rootzén, Holger: Extremal Theory for Stochastic Processes.

No. 6 Tjur, Tue: Block Designs and Electrical Networks.

No. 7 Johansen, Søren: Statistical Analysis of Cointegration Vectors.