Søren Johansen

Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models

Preprint
January
1989
3

Institute of Mathematical Statistics
University of Copenhagen

ISSN 0902-8846

Søren Johansen

ESTIMATION AND HYPOTHESIS TESTING OF COINTEGRATION VECTORS IN GAUSSIAN VECTOR AUTOREGRESSIVE MODELS

Preprint 1989 No. 3

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

January 1989

ABSTRACT

The purpose of this paper is to present the maximum likelihood estimators and likelihood ratio tests for a series of hypotheses about cointegration vectors and their loadings in a Gaussian vector autoregressive model, which includes seasonal dummies and a constant term. We find the asymptotic distribution of the likelihood ratio test for the hypothesis of r cointegration vectors, and it turns out to depend on the relation between the constant term and the loadings to the cointegration vectors. We then show that asymptotic inference about the cointegration vectors and their loadings can be performed by the usual χ^2 methods. We find an asymptotic representation of the maximum likelihood estimator, which is used to derive its asymptotic distribution and the distribution of some simple Wald tests.

Keywords: Cointegration, error correction models, maximum likelihood estimation, likelihood ratio test, Gaussian vector autoregressive processes

1. *Introduction and summary*

A large number of papers are devoted to the analysis of the concept of *cointegration* defined first by Granger (1981), Granger and Weiss (1983), and studied further by Engle and Granger (1987). Under this heading the topic has been also been studied by Stock (1987), Phillips and Ouliaris (1986), (1987), Phillips (1988), Johansen (1988b), Johansen and Juselius (1988), (1989). The main statistical technique that has been applied is *Regression with integrated regressors*, which has been studied by Phillips (1987), Phillips and Park (1986a), (1986b), (1987) and Sims, Stock and Watson (1986). Similar problems have been studied under the name *commom trends* see Stock and Watson (1987).

The purpose of this paper is to present some new results on the maximum likelihood estimators and likelihood ratio tests for cointegration in Gaussian vector autoregressive models which allows for a constant term and seasonal dummies. This brings in the technique of *reduced rank regression*, see Velu, Reinsel and Wichern (1986), and Ahn and Reinsel (1987) as well as the notion of *canonical analysis* Box and Tiao (1981) Velu, Wichern and Reinsel (1987), Pena and Box (1987), and the very elegant paper by Tso (1981). In Johansen (1988b) the likelihood based theory was presented for such a model without the constant term and the seasonal dummies, but it turns out that this term plays a crucial role for the interpretation of the model, as well as for the statistical and the probabilistic analysis.

A detailed statistical analysis illustrating the techniques by data on money demand from Denmark and Finland is given in Johansen and Juselius (1989), and the present paper deals with the underlying probability theory that allows one to make asymptotic inference.

The structure of the paper is the following: The next Section describes very briefly the estimators and test statistics studied in the subsequent Sections. A more detailed account can be found in Johansen and Juselius (1989) together with some illustartive examples. Section 3 gives a simple proof of Granger's representation theorem which clarifies the role of the constant term and gives a condition for the process to be integrated of order 1. We also state in rather condensed form the basic results on the processes as can be derived by the results of Phillips and Durlauf (1986) by applying the methods in Johansen (1988b). In Section 4 the asymptotic distribution of the likelihood ratio test statistic for the hypothesis of r cointegration vectors is derived. It turns out that the presence of the trend gives rise to some new limit distributions. Section 5 gives an asymptotic representation of the maximum likelihood estimator suitably normalized, and the results are then applied in Section 6 to show that asympotic inference about linear restrictions on the cointegration vectors and loadings can be performed using the χ^2 distribution, and in Section 7 we apply the asymptotic distribution of the maximum likelihood estimators to derive some very simple Wald tests.

2. *The statistical analysis of cointegration*

Consider a general VAR model with Gaussian errors written on the form

(2.1) $$\Delta X_t = \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} - \Pi X_{t-k} + \Phi D_t + \mu c_t + \epsilon_t, \quad (t = 1,\ldots,T),$$

where $c_t = 1$ and D_t are seasonal dummies orthogonal to the constant term, such that they sum to zero over a year, say. Further ϵ_t, $t = 1,\ldots,T$ are independent p-dimensional Gaussian variables with mean zero and variance matrix Λ. The values $X_{1-k},\ldots,X_0$ are considered fixed and the likelihood function is calculated for given values of these.

The model (2.1) is denoted by H_1 and we formulate the hypothesis of (at most) r cointegration vectors as

(2.2) $$H_2: \Pi = \alpha\beta',$$

where α and β are p×r matrices. Sometimes we shall compare models with different number of cointegration vectors, and we shall then use the notation $H_2(r)$. We shall also investigate a series of models expressed in terms of the cointegration vectors β and the loadings α.

Linear restrictions on β are expressed as

$$H_3: \beta = H\varphi,$$

where H(p×s) is known and φ(s×r) is the parameter to be estimated. Similarly linear restrictions on α are expressed as

$$H_4: \alpha = A\psi,$$

were A(p×m) is known and ψ(m×r) is to be estimated. It is convenient to formulate a general model

(2.3) $$H_5: \beta = H\varphi \text{ and } \alpha = A\psi,$$

which contains the previous as special cases.

It turns out that the role of the constant term is crucial for the statistical analysis as well as for the probabilistic analysis. It is proved in Theorem 3.1 that under certain conditions the process given by (2.1) is integrated of order 1 with a linear trend which is essentially determined by $\alpha_{\perp}'\mu$, where $\alpha_{\perp}$ is a p×(p−r) matrix of rank p−r consisting of vectors orthogonal to the vectors in α. The presence of the linear trend changes the analysis and it is therefore convenient to define a series of models H_i^* where the $*$ indicates that apart from the restrictions imposed under H_i we also impose the restriction $\mu = \alpha\beta_0'$, where the parameter β_0 has the interpretation as an intercept in the cointegration relation. In this case clearly $\alpha_{\perp}'\mu = 0$, and the linear trend is absent.

We shall treat the models H_i in detail and mention the results for the models H_i^*, and sometimes comment on the proofs when they require special attention.

In order to formulate the main result of this section we shall introduce some notation. In Johansen (1988b) it was shown how one can estimate the parameters in the model H_2, see (2.2).

We define $Z_{0t} = \Delta X_t$, $Z_{1t} = (\Delta X_{t-1}',\ldots,\Delta X_{t-k+1}',D_t',c_t)'$ and $Z_{kt} = X_{t-k}$, and we let Γ consist of the parameters $(\Gamma_1,\ldots,\Gamma_{k-1},\Phi,\mu)$. Then the model becomes

(2.4) $$Z_{0t} = \Gamma Z_{1t} - \Pi Z_{kt} + \epsilon_t.$$

With this notation define the product moment matrices

(2.5) $$M_{ij} = T^{-1}\sum_{t=1}^{T} Z_{it}Z_{jt}', \qquad (i,j = 0,1,k),$$

the residuals

$$R_{it} = Z_{it} - M_{i1}M_{11}^{-1}Z_{1t}, \qquad (i = 0,k),$$

and the residual sums of squares

$$S_{ij} = M_{ij} - M_{i1}M_{11}^{-1}M_{1j}, \qquad (i,j = 0,k). \tag{2.6}$$

Then the estimate of Γ for fixed values of α, β and Λ is found to be

$$\hat{\Gamma} = (M_{01} + \Pi M_{k1})M_{11}^{-1}. \tag{2.7}$$

Thus the residuals are found by regressing ΔX_t and X_{t-k} on the lagged differences, the dummies and the constant. This gives the concentrated likelihood function with respect to the parameters $\Gamma_1,\ldots,\Gamma_{k-1},\Phi$, and μ:

$$L_{max}^{-2/T}(\alpha,\beta,\Lambda) = |\Lambda|\exp\{T^{-1}\sum_{t=1}^{T}(R_{0t} + \alpha\beta' R_{kt})\, \Lambda^{-1}(R_{0t} + \alpha\beta' R_{kt})\}. \tag{2.8}$$

This function is easily minimized for fixed β to give

$$\hat{\alpha}(\beta) = -\, S_{0k}\beta(\beta' S_{kk}\beta)^{-1}, \tag{2.9}$$

$$\hat{\Lambda} = S_{00} - S_{0k}\beta(\beta' S_{kk}\beta)^{-1}\beta' S_{k0}, \tag{2.10}$$

together with

$$L_{max}^{-2/T}(\beta) = |S_{00}|\,|\beta'(S_{kk} - S_{k0}S_{00}^{-1}S_{0k})\beta|/|\beta' S_{kk}\beta|. \tag{2.11}$$

This again is minimized by the choice $\hat{\beta} = (\hat{v}_1,\ldots,\hat{v}_r)$, where $\hat{v}_1,\ldots,\hat{v}_p$ are the eigenvectors of the equation

$$|\lambda S_{kk} - S_{k0}S_{00}^{-1}S_{0k}| = 0, \tag{2.12}$$

normed by $\hat{v}' S_{kk}\hat{v} = I$. The maximized likelihood function is found from

$$L_{max}^{-2/T} = |S_{00}|\prod_{i=1}^{r}(1-\hat{\lambda}_i). \tag{2.13}$$

This procedure is given in Johansen (1988b), and consists of well known multivariate techniques from the theory of partial canonical correlations, see Anderson (1984) and Tso (1981).

Under the hypothesis H_5 we shall transform the matrices S_{ij} some more. Together with A we shall consider $B(p\times(p-r)) = A_{\perp}$, and introduce the notation

$$S_{hh.b} = H'S_{kk}H - H'S_{k0}B(B'S_{00}B)^{-1}B'S_{0k}H,$$
$$S_{aa.b} = A'S_{00}A - A'S_{00}B(B'S_{00}B)^{-1}B'S_{00}A,$$

and similarly for $S_{ha.b}$, $S_{ah.b}$, V_{ab}, S_{bb} etc.

THEOREM 2.1: *Under the hypothesis H_5: $\beta = H\varphi$ and $\alpha = A\varphi$ where H is p×s and α is p×m, the maximum likelihood estimators are found as follows: First solve*

(2.14) $$|\lambda S_{hh.b} - S_{ha.b}S_{aa.b}^{-1}S_{ah.b}| = 0,$$

to give eigenvalues $\hat{\lambda}_1 > \ldots > \hat{\lambda}_s$, and eigenvectors $\hat{v}_1, \ldots, \hat{v}_s$. Then

(2.15) $$\hat{\beta} = H(\hat{v}_1, \ldots, \hat{v}_r)$$

(2.16) $$\hat{\alpha} = -(A'A)^{-1}S_{ak.b}\hat{\beta}.$$

The estimate of Λ is found from

(2.17) $$\hat{\Lambda}_{bb} = S_{bb},$$

(2.18) $$\hat{\Lambda}_{ab} = S_{ab} + A'\hat{\alpha}\hat{\beta}'S_{kb},$$

(2.19) $$\hat{\Lambda}_{aa.b} = S_{aa} - A'\hat{\alpha}\hat{\alpha}'A$$

The estimate for Γ is found from (2.7) and the maximized likelihood function is

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

PROOF: We insert the value $\beta = H\varphi$ and $\alpha = A\psi$ in the concentrated likelihood function (2.8), and find

$$L_{max}^{-2/T}(\psi,\varphi,\Lambda) =$$

$$|\Lambda|\exp\{T^{-1}\sum_{t=1}^{T}(R_{0t} + A\psi\varphi'H'R_{kt})\ \Lambda^{-1}(R_{0t} + A\psi\varphi'H'R_{kt})\}.$$

We shall use the properties of the density of a multivariate Gaussian distribution to decompose this likelihood function into the marginal density of $R_{bt} = B'R_{0t}$, which does not contain the parameters of interest φ and ψ, and the conditional density given R_{bt}. The first factor gives the contribution

$$L_{max}^{-2/T}(\Lambda_{bb}) = |\Lambda_{bb}|\ \exp\{\ T^{-1}\sum_{t=1}^{T} R_{bt}'\Lambda_{bb}^{-1}\ R_{bt}\}/|B'B|,$$

where $\Lambda_{bb} = B'\Lambda B$. Maximizing we find

$$\hat{\Lambda}_{bb} = S_{bb} = B'S_{00}B,$$

which proves (2.17). The relevant part of the maximized likelihood function is

$$L_{max}^{-2/T} = |S_{bb}|/|B'B|.$$

The conditional likelihood function becomes

$$L_{max}^{-2/T}(\psi,\varphi,\Lambda_{aa.b},\Lambda_{ab}\Lambda_{bb}^{-1}) = |\Lambda_{aa.b}|\exp\{T^{-1}\sum_{t=1}^{T} R_t'\Lambda_{aa.b}^{-1}R_t\}/|A'A|,$$

where

$$R_t = R_{at} - \Lambda_{ab}\Lambda_{bb}^{-1}R_{bt} + A'A\psi\varphi'R_{ht}.$$

Minimizing with respect the parameter $\Lambda_{ab}\Lambda_{bb}^{-1}$ gives rise to yet another regression of R_{at} and R_{ht} on R_{bt}, which gives the estimate

$$\hat{\Lambda}_{ab}\hat{\Lambda}_{bb}^{-1}(\varphi,\psi) = (S_{ab} + A'A\psi\varphi'S_{hb})S_{bb}^{-1}.$$

This proves (2.18) and gives the new residuals

$$R_{a.bt} = R_{at} - S_{ab}S_{bb}^{-1}R_{bt},$$

and

$$R_{h.bt} = R_{ht} - S_{hb}S_{bb}^{-1}R_{bt}.$$

If we now define the new parameter $\tilde{\psi} = A'A\psi$, then the likelihood function is reduced to the form (2.8) in terms of $R_{a.bt}$, $R_{h.bt}$, φ, $\tilde{\psi}$ and $\Lambda_{aa.b}$. Hence the solution can be found from the relations (2.9),(2.10),(2.13) and (2.7).

If the equation (2.1) is multiplied by A' and B' respectively the second equation does not contain the parameters of interest, since $B'\alpha = B'A\psi = 0$. Hence it appears natural to estimate these equations first and then analyse the conditional model for $A'\Delta X_t$ given $B'\Delta X_t$ and the past values of the process. Since also $\beta = H\varphi$, then $\alpha\beta'X_{t-k} = \alpha\varphi'H'X_{t-k}$, which shows that the levels of the process appear only through the transformation H. It is these two operations of conditioning and transformation that form the basis for the proof of Theorem 2.1. We can now easily derive the various likelihood ratio tests using the relation (2.20) for the maximized likelihood function, since for r = p we have $H_2(p) = H_1$, and since the factor $|S_{00}|$ cancels in all the ratios.

COROLLARY 2.2: *The likelihood ratio test statistic for the hypothesis H_2 versus H_1 is given by*

$$(2.21) \qquad -2ln(Q;H_2|H_1) = -T\sum_{i=r+1}^{p} ln(1-\hat{\lambda}_i),$$

whereas the likelihood ratio test statistic for $H_2(r)$ versus $H_2(r+1)$ is given by

$$(2.22) \qquad -2ln(Q;r|r+1) = -T ln(1-\hat{\lambda}_{r+1}).$$

In order to express the likelihood ratio test statistics of the various hypotheses about restrictions on β and α, we shall indicate by a subscript the model under which the eigenvalues are calculated. Thus $\lambda_{3.1}$ is the largest eigenvalue calculated from (2.14) under the model H_3, i.e. under the assumption that $\alpha = A\psi$ and $H = I$. We can then formulate the results about the test statistics for the hyptheses about restrictions on β and α.

COROLLARY 2.3: *The likelihood ratio test statistic of the restriction $\beta = H\varphi$ under the assumption that $\alpha = A\psi$ is given by*

$$(2.23) \qquad -2ln(Q;H_5|H_4) = -T\sum_{t=1}^{T} ln\{(1-\hat{\lambda}_{5.i})/(1-\hat{\lambda}_{4.i})\},$$

and the likleihood ratio test of the hypothesis $\alpha = A\psi$ under the assumption that $\beta = H\varphi$ is given by

$$(2.24) \qquad -2ln(Q;H_5|H_3) = -T\sum_{t=1}^{T} ln\{(1-\hat{\lambda}_{5.i})/(1-\hat{\lambda}_{3.i})\}.$$

We shall conclude this section by pointing out how one can analyse the models H_i^*. First we note that if $\mu = \alpha\beta_0'$, then

$$-\alpha\beta'X_{t-k} + \mu c_t = -\alpha\beta'X_{t-k} + \alpha\beta_0'c_t = -\alpha\beta^{*\prime}X_{t-k}^*,$$

for $\beta^* = (\beta', -\beta_0')'$ and $X_{t-k}^* = (X_{t-k}', c_t)'$. In these new variables the model looks like

$$\Delta X_t = \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} - \alpha\beta^{*\prime}X_{t-k}^* + \Phi D_t + \epsilon_t, \quad (t = 1,\dots,T).$$

The analysis is now easily performed as in the Theorem 2.1 by defining $Z_{0t}^* = \Delta X_t$, $Z_{1t}^* = (\Delta X_{t-1}', \dots, \Delta X_{t-k+1}', D_t')'$, and $Z_{kt}^* = X_{t-k}^*$, as well as the moment matrices M_{ij}^* and S_{ij}^*, see (2.5) and (2.6). In order to keep the same meaning of the matrix H in these hypotheses we extend H to the

$(p+1)\times(s+1)$ matrix H^* by defining

$$H^* = \begin{bmatrix} H & 0 \\ 0 & 1 \end{bmatrix},$$

and define $S^*_{hh} = H^{*\prime}S^*_{kk}H^*$ etc.

With this notation exactly the same results hold for the models H^*_i as given by Theorem 2.1, Corollary 2.2 and Corollary 2.3, with proper change of notation to allow for the *.

3. *Grangers representation theorem and the limiting behaviour of the process*

When we want to investigate the distributional properties of the estimates and the test statistics we have to make more assumptions about the process. The basic assumption is that for the characteristic polynomial

(3.1) $\quad \Pi(z) = I - \Pi_1 z - \ldots - \Pi_k z^k,$

we have that $|\Pi(z)| = 0$ implies that either $|z| > 1$ or $z = 1$, which guarantees that the non-stationarity of X_t can be removed by differencing.

Now write the model defined by (2.1), as

(3.2) $\quad \Pi X_t + \Pi_1(L)\Delta X_t = \epsilon_t + \mu c_t + \Phi D_t,$

where $\Pi = I - \Pi_1 - \ldots - \Pi_k = \Pi(1)$.

The first result that we want to prove is the fundamental result about error correction models of order 1 and their structure. The basic result is due to Granger (1981), see Engle and Granger (1987) or Johansen

(1988a), but we shall give a very simple proof here, which contains a condition for the process to be integrated of order 1 and also clarifies the role of the constant term. We shall define the $p\times(p-r)$ matrices $\alpha_\perp$ and $\beta_\perp$ such that $\beta_\perp'\beta = 0$ and such that

(3.3) $$\alpha_\perp\alpha_\perp' = \Lambda^{-1}(I - \alpha(\alpha'\Lambda^{-1}\alpha)^{-1}\alpha'\Lambda^{-1}).$$

Note that $\alpha_\perp'\alpha = 0$ and that $\alpha_\perp'\Lambda\alpha_\perp = I$. This choice of $\alpha_\perp$ will simplify the calculations later in the proofs.

THEOREM 3.1: *(Grangers representation theorem). If*

(3.4) $$\Pi = \alpha\beta',$$

for α *and* β *of dimension* $p\times r$ *and rank* r *and if*

(3.5) $$\alpha_\perp'\Pi_1(1)\beta_\perp,$$

has full rank $p - r$, *then we can write the process in the moving average form* $\Delta X_t = C(L)(\epsilon_t - \mu c_t - \Phi D_t)$ *and the following representation holds:*

(3.6) $$C = C(1) = \beta_\perp(\alpha_\perp'\Pi_1(1)\beta_\perp)^{-1}\alpha_\perp'.$$

It follows that

(3.7) ΔX_t *is stationary,*

(3.8) X_t *is non-stationary, with linear trend* $\psi t = C\mu t$,

(3.9) $\beta' X_t$ *is stationary,*

and hence (3.2) can be interpreted as an error correction model. If in particular $\alpha_\perp'\mu = 0$ *the linear trend disappears and* $\beta' X_t$ *has mean* $(\alpha'\alpha)^{-1}\alpha'\mu$ *and* ΔX_t *has mean zero, apart from terms involving the seasonal dummies.*

Strictly speaking the processes ΔX_t and $\beta' X_{t-k}$ equal a stationary process plus the term involving the seasonal dummies, but we shall call such a process stationary. Note that the relation between (3.4) and (3.6) shows that for this type of process there is a very nice symmetry between the singularity of the "impact" matrix Π for the autoregressive representation and the singularity of the "impact" matrix for the moving average representation, in the sense that what is the null space for C' is the range space for Π and what is the range space for Π' is the null space for C. It is this symmetry that allows the results for this type of process to be exceptionally simple.

PROOF: If we multiply the equation (3.2) by α' and $\alpha_\perp'$ we get the equations

$$\alpha'\alpha\beta' X_t + \alpha'\Pi_1(L)\Delta X_t = \alpha'(\epsilon_t + \mu c_t + \Phi D_t),$$

$$\alpha_\perp'\Pi_1(L)\Delta X_t = \alpha_\perp'(\epsilon_t + \mu c_t + \Phi D_t).$$

To discuss the properties of the process X_t we shall solve the equations for X_t and express it in terms of the ϵ_t's. We therefore introduce the variables $Z_t = (\beta'\beta)^{-1}\beta' X_t$ and $Y_t = (\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\Delta X_t$ as new variables, from which ΔX_t can be recovered:

$$\Delta X_t = \beta_\perp Y_t + \beta\Delta Z_t.$$

This gives the equations

$$\alpha'\alpha\beta'\beta Z_t + \alpha'\Pi_1(L)\beta\Delta Z_t + \alpha'\Pi_1(L)\beta_\perp Y_t = \alpha'(\epsilon_t + \mu c_t + \Phi D_t), \tag{3.10}$$

$$\alpha_\perp'\Pi_1(L)\beta\Delta Z_t + \alpha_\perp'\Pi_1(L)\beta_\perp Y_t = \alpha_\perp'(\epsilon_t + \mu c_t + \Phi D_t). \tag{3.11}$$

The matrix function defining this new system, consisting of Z_t and Y_t, takes the form:

$$\tilde{A}(z) = \begin{bmatrix} \alpha'\alpha\beta'\beta + \alpha'\Pi_1(z)\beta(1-z) & \alpha'\Pi_1(z)\beta_\perp \\ \alpha_\perp'\Pi_1(z)\beta(1-z) & \alpha_\perp'\Pi_1(z)\beta_\perp \end{bmatrix}.$$

For z = 1 this has determinant

$$|\alpha'\alpha|\,|\beta'\beta|\,|\alpha_\perp'\Pi_1(1)\beta_\perp|,$$

which is non-zero by assumption (3.4) and (3.5), hence z = 1 is not a root. For $z \neq 1$ we use the representation

$$\tilde{A}(z) = (\alpha,\alpha_\perp)'\Pi(z)(\beta,\beta_\perp(1-z)^{-1}),$$

which gives the determinant as

$$|\tilde{A}(z)| = |(\alpha,\alpha_\perp)|\,|\Pi(z)|\,|(\beta,\beta_\perp)|(1-z)^{-(p-r)},$$

which shows that all roots of $|\tilde{A}(z)| = 0$ are outside the unit disk, by the assumption about $\Pi(z)$, see (3.1).

This shows that the system defined by (3.10) and (3.11) is invertible and that Y_t and Z_t are stationary processes, and hence that ΔX_t is stationary apart from the contribution from the centered dummies. This proves (3.7) and (3.9). From the representation of the processes Z_t and Y_t we can get a representation of ΔX_t by multiplying by the matrix $(\beta\Lambda,\beta_\perp)$. Hence

$$C(L) = (\beta\Lambda,\beta_\perp)\tilde{A}(L)^{-1}(\alpha,\alpha_\perp)'$$

For L = 1 we get (3.6). By summation of ΔX_t we find that X_t contains the non-stationary component $\beta_\perp' \sum_{s=0}^{t} Y_s$ together with a linear trend $\psi t = C\mu t$, which proves (3.8).

Note that μ enters the linear trend only through $\alpha_\perp'\mu$, and that the linear trend ψ is contained in the span of $\beta_\perp$, and hence cancels if we consider the components $\beta' X_t$. The seasonal dummies are so constructed that they remain bounded even after summation over t and hence do not contribute to the linear trend.

Since we have proved that ΔX_t and $\beta' X_t$ are stationary the stochastic components of $Z_{1t}' = (\Delta X_{t-1}', \ldots, \Delta X_{t-k+1}', D_t', c_t)$ are stationary. We define

$$\mathrm{Var}\begin{bmatrix} \Delta X_t \\ \beta' X_{t-k} \end{bmatrix} | Z_{1t} = \begin{bmatrix} \Sigma_{00} & \Sigma_{0k}\beta \\ \Sigma_{0k}\beta & \beta'\Sigma_{kk}\beta \end{bmatrix}.$$

From the equation (2.7)

(3.12) $$\Delta X_t = \Gamma Z_{1t} - \Pi Z_{kt} + \epsilon_t,$$

one finds immidiately the results of the next Lemma.

LEMMA 3.2: *The following relations hold*

(3.13) $$\Sigma_{00} = -\alpha\beta'\Sigma_{k0} + \Lambda,$$

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

and hence

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

These relations imply that

(3.16) $$(\alpha'\Sigma_{00}^{-1}\alpha)^{-1}\alpha'\Sigma_{00}^{-1} = (\alpha'\Lambda^{-1}\alpha)^{-1}\alpha'\Lambda^{-1},$$

and

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

The following technical results are essentially given in Johansen (1988b) based on the results by Phillips and Durlauf (1986). The results have been modified to allow for the linear trend in the process. We let W be a Brownian motion in p dimensions with covariance matrix Λ and define $\overline{W} = \int W(u)du$, where all such integrals here and in the following are from 0 to 1. One can then prove that $\psi' X_t$ is approximately linearly increasing, whereas if γ is chosen such that $\gamma'\psi = 0$ and $sp(\gamma,\psi) = sp(\beta_\perp)$, then $\gamma' X_t$ is non-stationary but with a constant mean. The results of the asymptotic behaviour of the process are summarized in

LEMMA 3.3: *As* $T \to \infty$ *we have*

$$T^{-1/2}\gamma' X_{[Tt]} \overset{w}{\to} \gamma' CW(t), \tag{3.18}$$

$$T^{-1}\psi' X_{[Tt]} \overset{P}{\to} \psi'\psi t. \tag{3.19}$$

Furthermore $\beta' X_t$ *is stationary.*

Using these results one can describe the asymptotic properties of the various moment matrices M_{ij} and S_{ij} which are basic for the properties of the estimators and tests. We shall not give these in detail here since the proofs mimic the proofs in Johansen (1988b) but we shall rather summarize the results in two Lemmas.

LEMMA 3.4: *Let* $A_T = (\beta, T^{-1/2}\gamma, T^{-1}\psi)$, *then if* $\psi = C\mu \neq 0$

$$A_T' S_{kk} A_T \overset{w}{\to} \tag{3.20}$$

$$\begin{bmatrix} \beta'\Sigma_{kk}\beta & 0 & 0 \\ 0 & \gamma' C\int(W-\bar{W})(W-\bar{W})'du C'\gamma & \gamma' C\int(u-1/2)(W-\bar{W})du\psi'\psi \\ 0 & \psi'\psi\int(u-1/2)(W-\bar{W})'du C'\gamma & (\psi'\psi)^2/12 \end{bmatrix}$$

Further

$$(3.21)\quad T^{1/2}A_T'(S_{k0} + S_{kk}\beta\alpha') =$$

$$T^{1/2}A_T'[T^{-1}\sum_{t=1}^{T}(X_{t-k} - M_{k1}M_{11}^{-1}Z_{1t})\epsilon_t'] \xrightarrow{w} \begin{bmatrix} N(0,\beta'\Sigma_{kk}\beta\otimes\Lambda) \\ \gamma'C\int(W-\overline{W})dW' \\ \psi'\psi\int(u-1/2)dW' \end{bmatrix}.$$

Finally $\beta'S_{k0} \overset{a.s.}{\rightarrow} \beta'\Sigma_{k0}$ *and* $S_{00} \overset{a.s.}{\rightarrow} \Sigma_{00}$. *If* $\psi = 0$, *then the results (3.20) and (3.21) should be modified by choosing* $\gamma = \beta_\perp$ *and deleting the terms involving* ψ.

Since $c_t = 1$ is included in the regressors the process X_t is corrected for its mean in the preliminary regressions. This is seen to be reflected in the asymptotics by the subtraction of $\overline{W}$ and 1/2. Since the process contains a linear trend only p–r–1 components $(\gamma'C)$ of the process W enter the result. The trend is described by replacing the last component of W by u.

If we want to discuss the hypotheses H_i^* which also restricts μ to have no trend component, i.e. $\psi = 0$ or $\mu = \alpha\beta_0'$, then we get different asymptotic results. We shall calculate the estimates using the matrices S_{ij}^*, see Section 2, and define $\beta^{*\prime} = (\beta', -\beta_0')$, and choose $\xi' = (0,1)$ and $\gamma^{*\prime} = (\beta_\perp', 0)$. Then if β has full rank the vectors (β^*, γ^*, ξ) are r+(p–r)+1 = p+1 linearly independent vectors spanning R^{p+1}. We can then prove

LEMMA 3.5: *Let* $B_T = (\beta^*, T^{-1/2}\gamma^*, \xi)$, *then* $T^{-1}\sum_{t=1}^{T} \beta' X_{t-k} \overset{a.s.}{\rightarrow} \beta_0'$, *and*

$$(3.22) \quad B_T' S_{kk}^* B_T \overset{w}{\rightarrow} \begin{bmatrix} \beta'\Sigma_{kk}\beta & 0 & 0 \\ 0 & \beta_\perp' C\int WW'du C'\beta_\perp & \beta_\perp' C\int W du \\ 0 & \int W' du C'\beta_\perp & 1 \end{bmatrix}.$$

Further

$$(3.23) \quad T^{1/2} B_t'(S_{k0}^* + S_{kk}^* \beta^* \alpha') =$$

$$T^{1/2} B_T'[T^{-1}\sum_{t=1}^{T}(X_{t-k}^* - M_{k1}^* M_{11}^{*-1} Z_{1t}^*)\epsilon_t'] \overset{w}{\rightarrow} \begin{bmatrix} N(0, \beta'\Sigma_{kk}\beta \otimes \Lambda) \\ \beta_\perp' C\int W dW' \\ \int dW' \end{bmatrix}$$

Finally $\beta^{*\prime} S_{k0}^* \overset{a.s.}{\rightarrow} \beta'\Sigma_{k0}$ *and* $S_{00}^* \overset{a.s.}{\rightarrow} \Sigma_{00}$.

By moving the constant term to the vector X_{t-k}, we no longer correct for the mean in the process W and the added 1 gives an extra dimension to the matrix S_{kk}^*. It is seen that the constant term plays an important role for the formulation of the limiting results, either because it implies a linear trend for the non-stationary part of the process, or because it enters the cointegration vector. The two cases require a different normalization. The seasonal dummies do not play an equally important role once they have be orthogonalized to the constant term. The reason for the is that quantities like $T^{-1}\sum_{t=1}^{T} D_t \Delta X_t'$ and $T^{-1}\sum_{t=1}^{T} D_t X_{t-k}'$ remain bounded as $T \rightarrow \infty$. The crucial property, which is applied to see this, is that the partial sums of D_t remain bounded.

4. *Asymptotic properties of the Likelihood ratio tests for cointegration*

In Johansen (1988b) the likelihood ratio test of H_2, i.e. of r cointegrating vectors, was discussed in the model with no constant term. It turns out, as will be shown below, that when a constant is included in the model not only the test statistics are changed, but also their asymptotic distributions are changed. Furthermore the analysis also depends on whether or not the processes are allowed to contain a linear trend in the non-stationary part, that is whether or not $\alpha_\perp'\mu = 0$.

THEOREM 4.1: *Under the hypothesis* $H_2 \; : \; \Pi = \alpha\beta'$ *the statistic* $-2ln(Q;H_2|H_1)$ *has a limit distribution which, if* $\alpha_\perp'\mu \neq 0$, *can be expressed in terms of a p-r dimensional Brownian motion B with i.i.d components as*

(4.1) $$tr\{\int dBF'[\int FF'du]^{-1}\int FdB'\},$$

where

(4.2) $$F_i(u) = B_i(u) - \int B_i(u)du, \qquad (i = 1,\ldots,p-r-1),$$

and

(4.3) $$F_i(u) = u - 1/2, \qquad (i = p-r).$$

If in fact $\alpha_\perp'\mu = 0$, *then the asymptotic distribution* $-2ln(Q;H_2|H_1)$ *is given by (4.1) with* $F = B - \overline{B}$.

PROOF: The likelihood ratio test statistic of H_2 in H_1 is given in the form (2.21)

$$-2\ln(Q;H_2|H_1) = -T\sum_{i=r+1}^{p}\ln(1-\hat{\lambda}_i),$$

where the eigenvalues $\hat{\lambda}_{r+1},\ldots,\hat{\lambda}_p$ are the smallest eigenvalues in the equation

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

see (2.12), or (2.14) with H = A = I. We now multiply (4.4) by the matrix A_T' and A_T, see Lemma 3.4 and let $T \to \infty$. We then use the result that the ordered eigenvalues are continuous functions of the coefficient matrices and find that they converge to the ordered eigenvalues of the equation

$$\lambda^{p-r}|\Phi|\,|\lambda\beta'\Sigma_{kk}\beta - \beta'\Sigma_{k0}\Sigma_{00}^{-1}\Sigma_{0k}\beta| = 0,$$

where Φ is a notation for

(4.5) $$\Phi = \begin{bmatrix} \gamma'C\int(W-\bar{W})(W-\bar{W})'duC'\gamma & \gamma'C\int(u-1/2)(W-\bar{W})du\psi'\psi \\ \psi'\psi\int(u-1/2)(W-\bar{W})'duC'\gamma & (\psi'\psi)^2/12 \end{bmatrix}.$$

This shows that the r largest solutions of (4.4) converge to the solutions of

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

and that the rest converge to zero. Now define $\rho = T\lambda$ and let $\lambda \to 0$, then we multiply by $(\beta,\gamma,T^{-1/2}\psi)$ and its transposed, see Lemma 3.4, and find that the p-r smallest solutions of the equation (4.4), multiplied by T, will converge to the solutions of the equation

$$\left|\rho\begin{bmatrix} 0 & 0 \\ 0 & \Phi \end{bmatrix} - (\beta,\gamma,T^{-1/2}\psi)'S_{k0}S_{00}^{-1}S_{0k}(\beta,\gamma,T^{-1/2}\psi)\right| = 0,$$

and hence to the solutions of

(4.7) $$|\beta'\Sigma_{k0}\Sigma_{00}^{-1}\Sigma_{0k}\beta|\,|\rho\Phi - (\gamma,T^{-1/2}\psi)'S_{k0}NS_{0k}(\gamma,T^{-1/2}\psi)| = 0,$$

where N is a notation for

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

This converges by Lemma 3.4 to the same expression with S replaced by Σ. It follows from Lemma 3.2 and the choice of $\alpha_\perp$, see (3.3), that this limit equals $\alpha_\perp\alpha_\perp'$. Hence, from (4.7) it follows that in the limit ρ must be a solution to

$$(4.8)\qquad \left|\rho\begin{bmatrix}\gamma' C\int(W-\bar{W})(W-\bar{W})'duC'\gamma & \gamma' C\int(u-1/2)(W-\bar{W})du\psi'\psi \\ \psi'\psi\int(u-1/2)(W-\bar{W})'duC'\gamma & (\psi'\psi)^2/12\end{bmatrix}\right.$$

$$\left. - \begin{bmatrix}\gamma' C\int(W-\bar{W})dW'\alpha_\perp \\ \psi'\psi\int(u-1/2)dW'\alpha_\perp\end{bmatrix}\begin{bmatrix}\gamma' C\int(W-\bar{W})dW'\alpha_\perp \\ \psi'\psi\int(u-1/2)dW'\alpha_\perp\end{bmatrix}'\right| = 0.$$

Now $C = \beta_\perp\tau\alpha_\perp'$ for a non-singular $(p-r)\times(p-r)$ matrix τ, see (3.6), such that we can introduce the $(p-r)$-dimensional process $U = \alpha_\perp' W$ with variance matrix $\alpha_\perp'\Lambda\alpha_\perp = I$, and the $(p-r+1)$-dimensional process $\tilde{F}$ with the first p-r components equal to those of $U - \bar{U}$ and the last component equal to u-1/2. We can then write (4.8) as

$$|M(\rho\int\tilde{F}\tilde{F}'du - \int\tilde{F}dU'\int dU\tilde{F}')M'| = 0,$$

where the $(p-r)\times(p-r+1)$ matrix M has the form

$$M = \begin{bmatrix}\gamma'\beta_\perp\tau & 0 \\ 0 & \psi'\psi\end{bmatrix},$$

The expression can be simplified by noting that $\psi'\psi$ cancels. Still the process U enters into the integrals with the factor $\gamma'\beta_\perp\tau$ which are p-r-1 linearly independent combinations of the components of U. By multiplying by $(\gamma'\beta_\perp\tau\tau'\beta_\perp'\gamma)^{-1/2}$ we can turn these into orthonormal components and by supplementing these vectors with an extra orthonormal vector, which is $(\mu'\alpha_\perp\alpha_\perp'\mu)^{-1/2}\alpha_\perp'\mu$, we can transform the process U by an orthogonal matrix O to the process $B = OU$. Then the equation can be written as

$$(4.9)\qquad |\rho\int FF'du - \int FdB'\int dBF'| = 0,$$

where F is given by (4.2) and (4.3). This equation has p-r roots. Thus we have seen that the p-r smallest roots of (4.4) decrease to zero at the rate T^{-1} and that $T\lambda$ converge to the roots of (4.9). From the expression for the likelihood ratio test statistic we find that

$$- 2\ln(Q;H_2|H_1)) \simeq T \sum_{i=r+1}^{p} \hat{\lambda}_i \xrightarrow{w}$$

$$\sum_{i=r+1}^{p} \hat{\rho}_i = tr\{\int dBF'[\int FF'du]^{-1}\int FdB'\}.$$

Note that if $\psi = 0$, i.e. the linear trend is missing, then again applying Lemma 3.4 we can choose $\gamma = \beta_\perp$, and the results have to be modified by leaving out the terms containing ψ. This means that if $\alpha_\perp'\mu = 0$ holds, then the test of H_2 in H_1 is distributed as

(4.10) $$T_2 = tr\{\int dU(U-\bar{U})'[\int(U-\bar{U})(U-\bar{U})'du]^{-1}\int(U-\bar{U})dU'\}.$$

This completes the proof of Theorem 4.1.

COROLLARY 4.2: *The likelihood ratio test statistic* $- 2\ln(Q;r|r+1)$ *of the hypothesis* $H_2(r)$, *of r or less cointegrating vectors, in* $H_2(r+1)$, *see (2.22), is asympotically distributed as the maximum eigenvalue of (4.9), where F is given by (4.2) and (4.3) if* $\alpha_\perp'\mu \neq 0$, *and* $F = B - \bar{B}$ *if* $\alpha_\perp'\mu = 0$.

Next we shall investigate the test of H_2^* in H_1.

THEOREM 4.3: *Under the hypothesis* H_2^* : $\Pi = \alpha\beta'$ *and* $\mu = \alpha\beta_0'$ *the likelihood ratio statistic* $- 2\ln(Q;H_2^*|H_1)$ *is distributed as (4.1), but with the process F given by*

(4.11) $$F_i(u) = B_i(u) \qquad (i = 1,\ldots,p-r),$$

(4.12) $$F_i(u) = 1 \qquad (i = p-r+1).$$

Similarly the asymptotic distribution of $- 2\ln(Q^*;r|r+1)$ *of* $H_2^*(r)$ *in* $H_2^*(r+1)$ *is given by the maximum eigenvalue of (4.9) with F as in (4.11) and (4.12).*

PROOF: The estimation under H_2^* involved the solution of the equation

$$(4.13)\qquad |\lambda S_{kk}^* - S_{k0}^* S_{00}^{*-1} S_{0k}^*| = 0,$$

see (2.12) with the S_{ij} replaced by S_{ij}^*. Now multiply by $B_T' = (\beta^*, T^{-1/2}\gamma^*, \xi)'$ and B_T, see Lemma 3.5, and let $T \to \infty$. The roots of (4.13) converge to the roots of the equation

$$\left|\begin{bmatrix} \lambda\beta'\Sigma_{kk}\beta - \beta'\Sigma_{k0}\Sigma_{00}^{-1}\Sigma_{0k}\beta & 0 & 0 \\ 0 & \lambda\beta_\perp' C\int WW'duC'\beta_\perp & \lambda\beta_\perp' C\int Wdu \\ 0 & \lambda\int W'duC'\beta_\perp & \lambda \end{bmatrix}\right| = 0.$$

which shows that the r largest solutions of (4.13) converge to the roots of the same limiting equation as before, see (4.6). Now multiply instead by $(\beta^*, \gamma^*, T^{1/2}\xi)'$ and its transposed and let $\rho = T\lambda$ and $\lambda \to 0$, then we obtain, by an argument similar to that given in the proof of Theorem 4.1, that in the limit the p-r smallest roots normalized by T will converge in distribution to the roots of the equation

$$\left|\rho\begin{bmatrix} \beta_\perp' C\int WW'duC'\beta_\perp & \beta_\perp' C\int Wdu \\ \int W'duC'\beta_\perp & 1 \end{bmatrix} - \begin{bmatrix} \beta_\perp' C\int WdW'\alpha_\perp \\ \int dW'\alpha_\perp \end{bmatrix}\begin{bmatrix} \beta_\perp' C\int WdW'\alpha_\perp \\ \int dW'\alpha_\perp \end{bmatrix}'\right| = 0.$$

Again we can introduce the p-r dimensional process $U = \alpha_\perp' W$ and cancel the matrices $\beta_\perp'\beta_\perp\tau$ to see that the test statistic has a limit distribution which is given by

$$(4.14)\qquad T_2^* = tr\{\int dU\begin{bmatrix} U \\ 1 \end{bmatrix}'[\int \begin{bmatrix} U \\ 1 \end{bmatrix}\begin{bmatrix} U \\ 1 \end{bmatrix}'du]^{-1}\int \begin{bmatrix} U \\ 1 \end{bmatrix}dU'\},$$

which completes the proof of Theorem 4.3. The result for the maximal eigenvalue follows similarly.

COROLLARY 4.4: *The asymptotic distribution of the likelihood ratio test* $-2ln(Q;H_2^*|H_2)$ *for the hypothesis* H_2^* *given the hypothesis* H_2, *i.e.* $\alpha_\perp'\mu = 0$, *when there are r cointegration vectors, is asymptotically distributed as* χ^2 *with p-r degrees of freedom.*

PROOF: From the relation

$$\begin{pmatrix}U\\1\end{pmatrix}'[\int\begin{pmatrix}U\\1\end{pmatrix}\begin{pmatrix}U\\1\end{pmatrix}'du]^{-1}\begin{pmatrix}U\\1\end{pmatrix} = 1 + (U-\bar{U})'[\int(U-\bar{U})(U-\bar{U})'du]^{-1}(U-\bar{U})$$

it follows that $T_2^* = U(1)'U(1) + T_2$, see (4.10) and (4.14). The likelihood ratio test statistic of H_2^* in H_2 is the difference of the two test statistics considered in Theorem 4.1 and Theorem 4.3. They furthermore have the same variables entering the asymptotic expansions and hence the distribution can be found by subtracting the above random variables T_2 and T_2^*. The test statistics T_2 and T_2^* can be considered multivariate versions of the Dickey-Fuller test statistics, see Fuller (1976).

5. *Asymptotic properties of the estimators under the assumption of cointegration and linear restrictions on α and β*

We shall choose to give the asymptotic properties of the estimators under the hypothesis H_5 where restrictions are imposed on both α and β, since the other hypotheses are special cases. We shall in this section denote the maximum likelihood estimators under H_5 by ^. Under the hypothesis H_5 : $\alpha = A\psi$ and $\beta = H\varphi$, where A is p×m and H is p×s of full rank the estimates are found as follows:

For fixed β we find, see (2.9) and (2.16)

(5.1) $$\hat{\alpha}(\beta) = -A(A'A)^{-1}A'S_{0k.b}\beta(\beta'S_{kk.b}\beta)^{-1},$$

where $B = A_{\perp}$, whereas $\hat{\beta}$ is found from (2.15)

The distribution of $\hat{\beta}$ is clearly concentrated on the sp(H), and we shall in the proof use the natural coordinate system in that space. We choose, apart from β, the projection of ψ onto sp(H), $\eta = H(H'H)^{-1}H'\psi$, which is orthogonal to β, since $\psi = \beta_{\perp}\tau\alpha_{\perp}'\mu$, and supplement with the s-r-1 vectors $\gamma \in$ sp(H), such that (β,γ,η) consists of s mutually orthogonal vectors which span sp(H). We can then decompose the estimator $\hat{\beta}_i = H\hat{v}_i$, see (2.15), as follows

$$\hat{\beta}_i = \beta b_i + \gamma c_i + \eta f_i,$$

where, for instance, $b_i = (\beta'\beta)^{-1}\beta'\hat{\beta}_i$. Since the parameter β is not identified we can not hope to find a reasonable estimator without some arbitrary normalization. A general type of normalization can be found as follows: Take a p×r matrix c such that $c'\beta$ has full rank r, and define the normalized estimator

$$\hat{\beta}_c = \hat{\beta}(c'\hat{\beta})^{-1}.$$

It is seen that this estimator is independent of the choice of maximum likelihood estimator. If we choose $c = (I,0)'$, where I is r×r then $\hat{\beta}$ is normalized such that the first r rows constitute an identity matrix. We shall in the following choose $c = (\beta'\beta)^{-1}\beta$, which gives a convenient formulation of the theory. The results for other normalizations can then easily be deduced. With this choice of normalization we find that we shall divide $\hat{\beta}$ by $b = (\beta'\beta)^{-1}(\beta'\hat{\beta})$, and we find the following representation for $\hat{\beta}b^{-1}$:

PROPOSITION 5.1: *If $\eta \neq 0$, then it holds that under the hypothesis H_5 : $\alpha = A\psi$ and $\beta = H\varphi$ the maximum likelihood estimator $\hat{\beta}$ has the representation*

$$(5.2) \qquad \begin{bmatrix} T(\gamma'\gamma)^{-1}\gamma' \\ T^{3/2}\eta' \end{bmatrix}(\hat{\beta}b^{-1} - \beta) \simeq$$

$$\begin{bmatrix} T^{-1}\gamma'S_{kk}\gamma & T^{-3/2}\gamma'S_{kk}\eta \\ T^{-3/2}\eta'S_{kk}\gamma & T^{-2}\ \eta'S_{kk}\eta \end{bmatrix}^{-1} \begin{bmatrix} \gamma' \\ T^{-1/2}\eta' \end{bmatrix} T^{-1}\sum_{t=1}^{T}(X_{t-k}-\bar{X}_k)\epsilon_t'\Lambda^{-1}\alpha(\alpha'\Lambda^{-1}\alpha)^{-1},$$

where $\bar{X}_k = T^{-1}\sum_{t=1}^{T} X_{t-k}$. The right hand side converges in distribution to

$$(5.3) \qquad (\int FF'du)^{-1}\int FdV',$$

where F and V are independent processes defined by

$$(5.4) \qquad F_i(u) = \gamma_i'C(W-\bar{W}), \qquad (i = 1,\ldots,s-r-1),$$

$$(5.5) \qquad F_i(u) = u - 1/2, \qquad (i = s-r),$$

and

$$V = (\alpha'\Lambda^{-1}\alpha)^{-1}\alpha'\Lambda^{-1}W,$$

with variance given by $var(V) = (\alpha'\Lambda^{-1}\alpha)^{-1}$.

If $\eta = 0$, *then the above results have to be modified by deleting the terms contaning* η *and by choosing* γ *such that* (β,γ) *span* $sp(H)$ *hence* (5.3) *holds with*

(5.6) $$F = \gamma' C(W - \overline{W}).$$

PROOF: With the notation $S(\lambda) = \lambda S_{kk.b} - S_{ka.b}S_{aa.b}^{-1}S_{ak.b}$ we have that the vector $\hat{\beta}_i$ is the solution to the equation $S(\hat{\lambda}_i)\hat{\beta}_i = 0$. In the new coordinates (β,γ,η) this implies that

$$\gamma' S(\hat{\lambda}_i)\beta b_i + \gamma' S(\hat{\lambda}_i)\gamma c_i + \gamma' S(\hat{\lambda}_i)\eta f_i = 0,$$
$$\eta' S(\hat{\lambda}_i)\beta b_i + \eta' S(\hat{\lambda}_i)\gamma c_i + \eta' S(\hat{\lambda}_i)\eta f_i = 0,$$

and hence

$$\hat{\beta}_i - \beta b_i = (\gamma,\eta)(c_i',f_i')' =$$
$$- (\gamma,\eta)\begin{bmatrix}\gamma' S(\hat{\lambda}_i)\gamma & \gamma' S(\hat{\lambda}_i)\eta \\ \eta' S(\hat{\lambda}_i)\gamma & \eta' S(\hat{\lambda}_i)\eta\end{bmatrix}^{-1}\begin{bmatrix}\gamma' S(\hat{\lambda}_i)\beta \\ \eta' S(\hat{\lambda}_i)\beta\end{bmatrix} b_i.$$

The normalization is different for $\hat{\beta}$ in the direction η and in the direction γ due to the trend in the process. Hence we shall multiply both sides of the equation by $(T\gamma, T^{3/2}\eta)'$. Now note for instance that by Lemma 3.4 the term

$$T^{-1}\gamma' S(\hat{\lambda}_i)\gamma = T^{-1}(\hat{\lambda}_i\gamma'(S_{kk} - S_{kb}S_{bb}^{-1}S_{bk})\gamma - \gamma' S_{ka.b}S_{aa.b}^{-1}S_{ak.b}\gamma)$$

$$\simeq T^{-1}\hat{\lambda}_i\gamma' S_{kk}\gamma \simeq \hat{\lambda}_i\gamma' C\int(W-\overline{W})(W-\overline{W})'du C'\gamma.$$

Thus the conditioning on $B'R_{0t}$ is not relevant for this part of the asymptotics. We can then apply Lemma 3.4 and find

$$\begin{bmatrix} T\gamma' \\ T^{3/2}\eta' \end{bmatrix}(\hat{\beta}_i - \beta b_i) \simeq$$
$$-\begin{bmatrix} \gamma'\gamma & 0 \\ 0 & \eta'\eta \end{bmatrix} \begin{bmatrix} T^{-1}\gamma' S_{kk}\gamma & T^{-3/2}\gamma' S_{kk}\eta \\ T^{-3/2}\eta' S_{kk}\gamma & T^{-2}\ \eta' S_{kk}\eta \end{bmatrix}^{-1} \begin{bmatrix} \gamma' S(\hat{\lambda}_i)\beta \\ T^{-1/2}\eta' S(\hat{\lambda}_i)\beta \end{bmatrix} b_i\hat{\lambda}_i .$$

The second factor converges in distribution to Φ, see (4.5), and we shall evaluate the third factor. We begin by

(5.7) $$\gamma' S(\hat{\lambda}_i)\beta b_i = (\hat{\lambda}_i\gamma' S_{kk.b}\beta - \gamma' S_{ka.b}S_{aa.b}^{-1}S_{ak.b}\beta)b_i .$$

We shall first evaluate

(5.8) $$\begin{aligned} \gamma' S_{ka.b} &= \gamma'(S_{ka.b} + S_{kk}\beta\alpha' A) - \gamma' S_{kk}\beta\alpha' A \\ &\simeq \gamma'(S_{ka.b} + S_{kk}\beta\alpha' A) + \gamma' S_{kk}\beta(\beta' S_{kk.b}\beta)^{-1}\beta' S_{ka.b} \\ &\simeq \gamma'(S_{ka.b} + S_{kk}\beta\alpha' A) + \gamma' S_{kk.b}\beta(\beta' S_{kk.b}\beta)^{-1}\beta' S_{ka.b}, \end{aligned}$$

where we have applied the expression (5.1) for α and replaced $S_{kk.b}\beta$ by $S_{kk}\beta$, since $S_{bk}\beta \simeq B'\Sigma_{0k} = 0$, when $B'\alpha = 0$, see (3.14). Now insert the last term of (5.8) into (5.7) and we find

$$\gamma' S_{kk.b}\beta(\beta' S_{kk.b}\beta)^{-1}\beta' S(\hat{\lambda}_i)\beta b_i \simeq 0.$$

The first term of (5.8) inserted into (5.7) gives

$$\gamma' S(\hat{\lambda}_i)\beta b_i \simeq -\gamma'(S_{ka.b} + S_{kk}\beta\alpha' A)S_{aa.b}^{-1}S_{ak.b}\beta b_i .$$

From the relation (2.4) it follows that

$$S_{k0} + S_{kk}\beta\alpha' = T^{-1}\sum_{t=1}^{T}(X_{t-k} - M_{k1}M_{11}^{-1}Z_{1t})\epsilon_t' .$$

Hence

(5.9) $$\gamma'(S_{ka.b} + S_{kk}\beta\alpha' A) =$$
$$\gamma'\{T^{-1}\sum_{t=1}^{T}(X_{t-k}-M_{k1}M_{11}^{-1}Z_{1t})\epsilon_t'\}(A-BS_{bb}^{-1}S_{ba}) .$$

The limit of this is found from Lemma 3.4 and summing up we find that $\gamma' S(\hat{\lambda}_i)\beta b_i$ has the same limit distribution as

$$- \gamma' C\int(W-\bar{W})dW'(A-B\Sigma_{bb}^{-1}\Sigma_{ba})\Sigma_{aa.b}^{-1}\Sigma_{ak.b}\beta b_i .$$

Similarly we find

$$T^{-1/2}\eta' S(\hat{\lambda}_i)\beta b_i \simeq$$

$$- \eta'\eta\int(u-1/2)dW'(A-B\Sigma_{bb}^{-1}\Sigma_{ba})\Sigma_{aa.b}^{-1}\Sigma_{ak.b}\beta b_i .$$

Combining these results we obtain the asymptotic representation and the limiting expression

$$\begin{bmatrix} T\gamma' \\ T^{3/2}\eta' \end{bmatrix}(\hat{\beta b}^{-1} - \beta) \simeq$$

$$-\begin{bmatrix} \gamma'\gamma & 0 \\ 0 & \eta'\eta \end{bmatrix}\Phi^{-1}\begin{bmatrix} \gamma' C\int(W-\bar{W})\, dW' \\ \eta'\eta\int(u-1/2)dW' \end{bmatrix}(A - B\Sigma_{bb}^{-1}\Sigma_{ba})\Sigma_{aa.b}^{-1}\Sigma_{ak.b}\beta b D(\hat{\lambda})b^{-1},$$

where $D(\hat{\lambda}) = \mathrm{diag}(\hat{\lambda}_1,\ldots,\hat{\lambda}_r)$, and Φ is given by (4.5). Now

$$\beta' S_{kk.b}\beta b D(\hat{\lambda}) - \beta' S_{ka.a}S_{aa.b}^{-1}S_{ak.b}\beta b \overset{P}{\to} 0$$

which shows that

$$(A-B\Sigma_{bb}^{-1}\Sigma_{ba})\Sigma_{aa.b}^{-1}\Sigma_{ak.b}\beta b D(\hat{\lambda})b^{-1} \overset{P}{\to}$$

$$(A-B\Sigma_{bb}^{-1}\Sigma_{ba})\Sigma_{aa.b}^{-1}\Sigma_{ak.b}\beta(\beta'\Sigma_{ka.b}\Sigma_{aa.b}^{-1}\Sigma_{ak.b}\beta)^{-1}\beta'\Sigma_{kk.b}\beta = \kappa,$$

say. We shall now reduce this expression. From (3.14) we find that since $B'\alpha = 0$ it follows that $\Sigma_{bk}\beta = 0$, which shows that $\Sigma_{kk.b}\beta = \Sigma_{kk}\beta - \Sigma_{kb}\Sigma_{bb}^{-1}\Sigma_{bk}\beta = \Sigma_{kk}\beta$, and that $\Sigma_{ak.b}\beta = \Sigma_{ak}\beta$. With these simplifications we now have

$$\beta'\Sigma_{k0}\Sigma_{00}^{-1}\Sigma_{0k}\beta = (\beta'\Sigma_{ka},\beta'\Sigma_{kb})\begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{bmatrix}^{-1}\begin{bmatrix} \Sigma_{ak}\beta \\ \Sigma_{bk}\beta \end{bmatrix}$$

$$= \beta'\Sigma_{ka}\Sigma^{aa}\Sigma_{ak} = \beta'\Sigma_{ka}\Sigma_{aa.b}^{-1}\Sigma_{ak}\beta.$$

Finally we have that

$$(A - B\Sigma_{bb}\Sigma_{ba})\Sigma_{aa.b}^{-1}\Sigma_{ak.b}\beta = \Sigma_{00}^{-1}\Sigma_{0k}\beta.$$

This relation is checked by first multiplying from the left by $A'\Sigma_{00}$ to get an identity and then by $B'\Sigma_{00}$ to get zero. Thus we have the expression for κ given by

$$\kappa = \Sigma_{00}^{-1}\Sigma_{0k}\beta(\beta'\Sigma_{k0}\Sigma_{00}^{-1}\Sigma_{0k}\beta)^{-1}\beta'\Sigma_{kk}\beta.$$

Now apply (3.14) to replace $\beta'\Sigma_{k0}$ by an expression in α and we find that it equals $\Sigma_{00}^{-1}\alpha(\alpha'\Sigma_{00}^{-1}\alpha)^{-1}$, which by (3.16) equals $\Lambda^{-1}\alpha(\alpha'\Lambda^{-1}\alpha)^{-1}$ from which the expression for the variance follows. Notice that it is at this point that the condition (3.5) becomes crucial, and that the definition of $\alpha_{\perp}$ turns out to be convenient.

Note that the limiting distribution for fixed F is Gaussian with mean zero and variance

$$(\int FF'du)^{-1} \otimes (\alpha'\Lambda^{-1}\alpha)^{-1}.$$

Thus the limiting distribution of $\hat{\beta}b^{-1} - \beta$, which lies in sp(H) and is orthogonal to β, is a mixture of Gaussian distributions, see also Jeganathan (1988) for a general theory of locally asymptotically mixed normal models. Note also that the restrictions on α, as expressed by $\alpha = A\psi$, do not enter into the asymptotic representation for the estimate of β. This is the key to the asymptotic results for the test statistics in the next section.

We shall now find the distribution of $\hat{\beta}^*$ in the model H_5^*, i.e. when $\mu = \alpha\beta_0'$, $\beta = H\varphi$ and $\alpha = A\psi$. We introduce γ such that β and γ span H and define $\gamma^{*\prime} = (\gamma',0)$ and $\xi' = (0,1)$, which together with β^* gives a convenient coordinate system to work with in the description of the limiting distribution.

PROPOSITION 5.2: *Under the hypothesis* H_5^* : $\alpha = A\psi$, $\beta = H\varphi$, *and* $\mu = \alpha\beta_0'$ *the maximum likelihood estimator* $\hat{\beta}^*$ *has the representation*

$$(5.10) \qquad \begin{bmatrix} T(\gamma^{*\prime}\gamma^*)^{-1}\gamma^{*\prime} \\ T^{1/2}\ \xi' \end{bmatrix} (\hat{\beta}^* b^{*-1} - \beta^*) \simeq$$

$$\begin{bmatrix} T^{-1}\gamma^{*\prime}S_{kk}\gamma^* & T^{-1/2}\gamma^{*\prime}S_{kk}\xi \\ T^{-1/2}\ \xi' S_{kk}\gamma^* & \xi' S_{kk}\xi \end{bmatrix}^{-1} \begin{bmatrix} \gamma^{*\prime} \\ T^{-1/2}\xi' \end{bmatrix} T^{-1} \sum_{t=1}^{T} X_{t-k}^* \epsilon_t' \Lambda^{-1}\alpha(\alpha'\Lambda^{-1}\alpha)^{-1}.$$

The right hand side converges in distribution to (5.3) *with*

$F_i = U_i$, $\qquad (i = 1,\dots,p-r)$,

$F_i = 1$, $\qquad (i = p-r+1)$.

The estimate of β is superconsistent, see Stock (1987), whereas the estimate of β_0' is consistent with the usual rate. If we estimate β under the further restriction that $\beta_0 = 0$, i.e. the coefficients to the constant term are restricted to zero, the result (5.2) holds with the term involving ξ deleted and the limiting distribution is given by (5.3) with F = U, see Lemma 8 in Johansen (1988b).

PROOF: The proof of this result follows the lines of the proof of Proposition 5.1 applying the results in Lemma 3.5.

Next we shall find an asymptotic representation and the asymptotic distribution of the estimate of α suitably normalized.

PROPOSITION 5.3: *Under the hypothesis* H_5: $\alpha = A\psi$ *and* $\beta = H\varphi$ *the estimator* $\hat{\alpha}$ *has the representation*

$$T^{1/2}A'(\hat{\alpha}b' - \alpha) \simeq$$

$$- T^{-1/2}(A - B\Lambda_{bb}^{-1}\Lambda_{ba})'\{\sum_{t=1}^{T} \epsilon_t(X'_{t-k} - Z'_{1t}M_{11}^{-1}M_{1k})\}\beta(\beta'\Sigma_{kk}\beta)^{-1}.$$

which converges weakly to a Gaussian distribution of dimension $p \times r$ *with mean zero and variance matrix*

$$\Lambda_{aa.b} \otimes (\beta'\Sigma_{kk}\beta)^{-1}.$$

PROOF: In the definition (5.1) of $\hat{\alpha} = \hat{\alpha}(\hat{\beta})$ we can, by (5.2), replace $\hat{\beta}$ by βb^{-1}, which shows

$$A'\hat{\alpha} = - S_{ak.b}\beta(\beta'S_{kk.b}\beta)^{-1}b'^{-1} + O_P(T^{-1}),$$

and hence

$$T^{1/2}A'(\hat{\alpha}b' - \alpha) \simeq - T^{1/2}(S_{ak.b}\beta + A'\alpha\beta'S_{kk.b}\beta)(\beta'\Sigma_{kk}\beta)^{-1}$$

$$\simeq - T^{1/2}(S_{ak.b}\beta + A'\alpha\beta'S_{kk}\beta)(\beta'\Sigma_{kk}\beta)^{-1}.$$

Similarly to (5.9) we can prove

$$\beta'(S_{ka.b} + S_{kk}\beta\alpha'A\ ) = \beta'\{T^{-1}\sum_{t=1}^{T}(X_{t-k} - M_{k1}M_{11}^{-1}Z_{1t})\}\epsilon'_t(A-BS_{bb}^{-1}S_{ba}).$$

The asymptotic distribution now follows from Lemma 3.4.

PROPOSITION 5.4: *Under the hypothesis* H_5^*: $\alpha = A\psi$, $\beta = H\varphi$ *and* $\alpha'_\perp\mu = 0$ *the estimator* $\hat{\alpha}^*$ *has the representation*

$$T^{1/2}A'(\hat{\alpha}^*b^{*\prime} - \alpha) \simeq$$

$$- T^{1/2}(A-B\Lambda_{bb}^{-1}\Lambda_{ba})'\sum_{t=1}^{T}\epsilon_t(X^{*\prime}_{t-k} - Z^{*\prime}_{1t}M^{*-1}_{11}M^{*}_{1k})\beta^{*}(\beta'\Sigma_{kk}\beta)^{-1},$$

which converges weakly to a Gaussian distribution with mean zero and variance matrix

$$\Lambda_{aa.b} \otimes (\beta'\Sigma_{kk}\beta)^{-1}.$$

PROOF: We have

$$A'\hat{\alpha}^{*} = - S_{ak.b}\hat{\beta}^{*}(\hat{\beta}^{*\prime}S_{kk.b}\hat{\beta}^{*})^{-1},$$

and again we want to replace $\hat{\beta}^{*}$ by $\beta^{*}b^{*-1}$. The expansion

$$\hat{\beta}^{*} = \beta^{*}b^{*} + \gamma^{*}g^{*} + \xi f^{*},$$

together with Lemma 3.5 shows that $g^{*} \in O_p(T^{-1})$ while $f^{*} \in O_p(T^{-1/2})$. Thus dropping the term involving g^{*} involves an error of the order of magnitude T^{-1}, which disappers even when we multiply by $T^{1/2}$. It is seen that

$$\hat{\beta}^{*\prime}S^{*}_{kk.b}\hat{\beta}^{*} \simeq b^{*\prime}\beta^{*\prime}S^{*}_{kk.b}\beta^{*}b^{*}.$$

What remains to investigate is the term

$$S^{*}_{ak.b}\xi = S^{*}_{ak}\xi - S^{*}_{ab}S^{*-1}_{bb}S^{*}_{bk}\xi \ ,$$

where for instance $S^{*}_{ak}\xi = A'M^{*}_{0k}\xi - A'M^{*}_{01}M^{*-1}_{11}M^{*}_{1k}\xi$. Under the hypothesis H^{*}_{2} , where ΔX_t has a mean zero, the terms $S^{*}_{ak}\xi$ and $S^{*}_{bk}\xi$ converge to zero in probability of the order $T^{-1/2}$, hence

$$S^{*}_{ak.b}\xi f^{*} \in O_p(T^{-1}).$$

The rest of the proof now follows as in Proposition 5.3. The results about $\hat{\alpha}$ and $\hat{\beta}$ are mostly of a technical nature, since there does not seem to be a natural normalization. The one chosen is convenient for expressing the results. The more useful results on the estimation of the parameters are collected in the next Theorem.

THEOREM 5.5: *Under the hypothesis* H_5 : $\alpha = A\psi$ *and* $\beta = H\varphi$ *the asymptotic distribution of* $T^{1/2}(\hat{\Pi} - \Pi)$ *is Gaussian in* $p \times r$ *dimensions with mean 0 and a variance matrix given by*

$$\{A(A'A)^{-1}\Lambda_{aa.b}(A'A)^{-1}A'\} \otimes \{\beta(\beta'\Sigma_{kk}\beta)^{-1}\beta'\},$$

which is consistently estimated by

$$\{A(A'A)^{-1}A(S_{00} - \hat{\alpha}\hat{\alpha}')A(A'A)^{-1}A'\} \otimes \{\hat{\beta}\hat{\beta}'\}.$$

The parameters $\Lambda, \Gamma_1, \ldots, \Gamma_{k-1}, \Phi$, *and* μ *are consistently estimated. A similar result holds for the estimate of the parameters under the model* H_5^*.

PROOF. The results about $\hat{\Pi}$ follow directly from Proposition 5.3 by noting that $\hat{\Pi} = \hat{\alpha}\hat{\beta}' = \hat{\alpha}b'(\hat{\beta}b^{-1})'$, where we can replace $\hat{\beta}b^{-1}$ by β without changing the asymptotic distribution. This implies that the asymptotic distributions of $\hat{\Pi}$ is not influenced by the restrictions on β, and that the asymptotic variance comes from $\hat{\alpha}$ exclusively. The consistency of the parameters in Γ come from (2.7) together with the properties of the moment matrices M_{ij}. The consistency of $\hat{\Lambda}$ can be derived from (2.17),(2.18) and (2.19) by noting that, for instance,

$$\hat{\Lambda}_{ab} = S_{aa} + A'\hat{\alpha}\hat{\beta}'S_{kb} \overset{a.s.}{\rightarrow} \Sigma_{aa} + A'\alpha\beta'\Sigma_{kb},$$

but (3.14) implies that $\beta'\Sigma_{k0}B = 0$.

6. *The asymptotic distribution of the likelihood ratio test statistics of restrictions on* α *and* β.

We shall now find the limiting distributions of the likelihood ratio test statistics for the various hypotheses discussed in Section 2.

We shall first give a useful approximation to the test statistic of a simple hypothesis for β.

LEMMA 6.1: *The likelihood ratio test of a simple hypothesis about* β *in* H_5 *has the representation*

$$- 2\ln(Q;\beta|H_5) \simeq T tr\{var(V)^{-1}(\hat{\beta}b^{-1}-\beta)'S_{kk}(\hat{\beta}b^{-1}-\beta)\}, \tag{6.1}$$

which is and has asymptotically distributed as

$$tr\{var(V)^{-1}\int dV\{\int FF'du\}^{-1}\int FdV'\}, \tag{6.2}$$

where F *is given by* (5.4) *and* (5.5) *if* $\eta = H(H'H)^{-1}H'\psi \neq 0$ *and by* (5.6) *if* $\eta = 0$. *The distribution of* (6.2) *is* $\chi^2(r(s-r))$. *A similar representation and the same limit result holds for* $- 2\ln(Q;\beta|H_5^*)$.

PROOF: By a Taylors expansion of the partially maximized likelihood function we obtain (6.1) and Proposition 5.1 now implies the result about the limit distribution.

We shall give the asymptotic distribution of the test statistics for testing H_5 versus either H_3 or H_4, and H_5^* versus either H_3^* or H_4^*.

THEOREM 6.2: *The likelihood ratio test* $-2ln(Q;H_5|H_4)$ *of the hypothesis* $\beta = H\varphi$ *in the model* $H_4 : \alpha = A\psi$, *where* A *is* $p\times m$ *and* H *is* $p\times s$, *is asymptotically distributed as* $\chi^2(r(p-s))$. *The same result holds for* $-2ln(Q;H_5^*|H_4^*)$.

PROOF: If $L(\beta)$ denotes the concentrated likelihood function after maximizing with respect to all other parameters, then

$$Q(H_5|H_4) = L(\hat{\beta}_5)/L(\hat{\beta}_4) =$$
$$\{L(\hat{\beta}_5)/L(\beta)\}/\{L(\hat{\beta}_4)/L(\beta)\} = Q_5/Q_4,$$

where Q_4 and Q_5 are the tests of a simple hypothesis about β in H_4 and H_5 respectively. We can now apply Lemma 6.1, and find that $-2ln(Q;H_4|H_5)$ is the difference between two expressions like (6.1) with different estimates $\hat{\beta}_4$ and $\hat{\beta}_5$ respectively. Let the γ chosen under H_5, see (5.2), be denoted γ_H, then γ_H is chosen orthogonal to both $\beta \in sp(H)$ and to $\eta = H(H'H)^{-1}H'\psi \in sp(H)$. This implies that there is a matrix h such that $\gamma_H = \gamma h$. The rest of the proof now follows the proof of Theorem 4 in Johansen (1988b). The proof assumes that in fact $\eta \neq 0$, but the same technique works if $\eta = 0$, since then one should only work with the alternative form of the limiting distribution as discussed in Proposition 4.1. The result for H_5^* is proved similarly.

THEOREM 6.3: *The asymptotic distribution of the likelihood ratio test statistic* $-2ln(Q;H_5|H_3)$ *of the restriction* $\alpha = A\psi$ *in the model* H_3: $\beta = H\varphi$ *is asymptotically distributed as* $\chi^2(r(p-m))$. *The same result holds for* $-2ln(Q;H_5^*|H_3^*)$.

PROOF: Let us introduce the notation $\hat{\alpha}_3(\beta) = -S_{0k}\beta(\beta' S_{kk}\beta)$ and $\hat{\alpha}_5(\beta) = -A(A'A)^{-1}S_{ak.b}\beta(\beta' S_{kk.b}\beta)^{-1}$, see (2.9) and (5.1), and let $L(\alpha,\beta)$ denote the concentrated likelihood function. As before consider the decomposition

$$Q(H_5|H_3) = L(\hat{\alpha}_5(\hat{\beta}_5),\hat{\beta}_5)/L(\hat{\alpha}_3(\hat{\beta}_3),\hat{\beta}_3) = Q_1 Q_2/Q_3,$$

where $Q_1 = L(\hat{\alpha}_5(\beta),\beta)/L(\hat{\alpha}_3(\beta),\beta)$, $Q_3 = L(\hat{\alpha}_5(\beta),\beta)/L(\hat{\alpha}_5(\hat{\beta}_5),\hat{\beta}_5)$, and $Q_2 = L(\hat{\alpha}_3(\beta),\beta)/L(\hat{\alpha}_3(\hat{\beta}_3),\hat{\beta}_3)$.

Now the first factor, Q_1, is just the test of the restrictions $\alpha = A\psi$ for a fixed value of β. Such a hypothesis is just a hypothesis on the parameters of the stationary process $(\Delta X_t, \beta' X_t)$ and is therefore asymptotically distributed as $\chi^2(r(p-m))$. It is shown in Lemma 6.1 that the two other test statistics Q_2 and Q_3, which correspond to simple hypotheses about β and which only differ in the restrictions on α, have the same asymptotic representation. Hence their ratio tends to 1 and the result follows.

7. *Wald tests for hypotheses about α and β*

The asymptotic distribution of the estimators are expressed in an arbitrary normalization. The proper way to exploit such a result is to ask questions about the parameters which are invariant under such normalizations. One way of doing this is to ask for the distribution of the Wald test statistics.

We shall consider Wald tests which are very easy to calculate once the eigenvectors and eigenvalues have been calculated under the hypothesis H_2. Let us first consider a test for the hypothesis concerning α

and let us express it as

$$H_4 : B'\alpha = 0.$$

Note that the hypothesis is invariant to different normalizations of α. A Wald test can be constructed from the statistic $B'\hat{\alpha}$.

THEOREM 7.1: *Under the hypothesis H_4: $B'\alpha = 0$ where B is $p\times(p-m)$ of full rank the asymptotic distribution of*

$$(7.1) \qquad Ttr\{(B'(S_{00} - \hat{\alpha}\hat{\alpha}')B)^{-1}(B'\hat{\alpha}\hat{\alpha}'B)\}$$

is χ^2 with $(p-m)r$ degrees of freedom, where $\hat{\alpha}$ is the maximum likelihood estimator of α under H_2.

PROOF: In view of the results about the asymptotic distribution of $\hat{\alpha}$, see Proposition 5.3, we can consider the statistic

$$Ttr\{\Lambda_{bb}^{-1}B'\hat{\alpha}\hat{b}'(\beta'\Sigma_{kk}\beta)\hat{b}\hat{\alpha}'B\},$$

which is asymptotically distributed as χ^2 with (p-m)r degrees of freedom. To apply the test we need consistent estimates for the variance matrices, and we thus insert the consistent estimates for $\beta'\Sigma_{kk}\beta$

$$b'^{-1}\hat{\beta}'S_{kk}\hat{\beta}b^{-1} = b'^{-1}b^{-1},$$

and for Λ_{bb}

$$\hat{\Lambda}_{bb} = B'(S_{00} - \hat{\alpha}\hat{\alpha}')B,$$

and the result follows.

Let us next consider the hypothesis H_4 but expressed as

$$H_4 : K'\beta = 0.$$

where K is $p\times(p-s)$ of full rank. This suggests a Wald test on the statistic $K'\hat{\beta}$ and the problem is again how to normalize it.

We shall first prove an intermediate result. Let (β,γ,ψ) be as in Lemma 3.4 and let $\hat{v}$ be the eigenvectors of (4.4) corresponding to the eigenvalues $\hat{\lambda}_{r+1},\ldots,\hat{\lambda}_p$.

LEMMA 7.2: *If $K'\beta = 0$, $K'\gamma \neq 0$, and $\psi \neq 0$, we have*

$$TK'\hat{v}\hat{v}'K \simeq TK'(\gamma,\psi)\{(\gamma,\psi)'S_{kk}(\gamma,\psi)\}^{-1}(\gamma,\psi)'K, \tag{7.2}$$

which converges in distribution to

$$K'\gamma[\gamma'C\int GG'duC'\gamma]^{-1}\gamma'K, \tag{7.3}$$

where $G = W-\overline{W} - a(t-1/2)$ and $a = \int W(u)(u-1/2)du/\int(u-1/2)^2du$, i.e. G is W corrected for constant and trend. If $\psi = 0$ then (7.2) and (7.3) hold with $\gamma = \beta_\perp$ and $G = W - \overline{W}$. If, however, $K'\gamma = 0$ and $\psi \neq 0$, then

$$T^2K'\hat{v}\hat{v}'K \to K'\psi[\psi'\psi\int(u-1/2-a(W-\overline{W}))^2\psi'\psi]^{-1}\psi'K,$$

where $a = \int(u-1/2)W(u)du/\int(W-\overline{W})^2du$, i.e. the process that appears is the trend corrected for the mean and the Brownian motion.

PROOF: We first expand

$$\hat{v} = \beta e + \gamma g + \psi f \tag{7.4}$$

and then note that from

$$\hat{v}'S_{k0}S_{00}^{-1}S_{0k}\hat{v} = \mathrm{diag}(\hat{\lambda}_{r+1},\ldots,\hat{\lambda}_p) \in O_P(T^{-1}),$$

it follows that $\hat{v}$ and hence also the coordinates (e,g,f) are $O_P(T^{-1/2})$. From the normalization $\hat{v}'S_{kk}\hat{v} = I$, it even follows that $f \in O_P(T^{-1})$ and that, since $e \xrightarrow{P} 0$, we have

$$(\gamma g+\psi f)'S_{kk}(\gamma g+\psi f) \xrightarrow{P} I. \tag{7.5}$$

Note finally that from (7.4) we have, since $K'\beta = 0$, that

$$K'\hat{v} = K'(\gamma g+\psi f) = K'(\gamma,\psi)(g',f')'. \tag{7.6}$$

Now insert (7.5) and (7.6) into $TK'\hat{v}\hat{v}'K$ and we get

$$(7.7)\qquad TK'(\gamma g+\psi f)\{(\gamma g+\psi f)'S_{kk}(\gamma g+\psi f)\}^{-1}(\gamma g+\psi f)'K$$
$$= TK'(\gamma,\psi)\{(\gamma,\psi)'S_{kk}(\gamma,\psi)\}^{-1}(\gamma,\psi)'K$$
$$= K'(\gamma,T^{-1/2}\psi)\begin{bmatrix} T^{-1}\gamma'S_{kk}\gamma & T^{-3/2}\gamma'S_{kk}\psi \\ T^{-3/2}\psi'S_{kk}\gamma & T^{-2}\psi'S_{kk}\psi \end{bmatrix}^{-1}(\gamma,T^{-1/2}\psi)'K.$$

If $K'\gamma \neq 0$ then the term involving $T^{-1/2}\psi$ are of smaller order of magnitude and what remains is

$$T^{-1}K'\gamma[\gamma'S_{kk}\gamma - \gamma'S_{kk}\psi(\psi'S_{kk}\psi)^{-1}\psi'S_{kk}\gamma]^{-1}\gamma'K,$$

which converges as stated by application of Lemma 3.4. If $\psi = 0$ we can drop the terms involving ψ and choose $\gamma = \beta_\perp$ and apply Lemma 3.4 again. If instead $K'\gamma = 0$, then we must normalize by T^2, and then the result follows from (7.7).

We shall also need a result for the model H_2^* where restrictions have been placed on both Π and μ. Let $\hat{v}^*$ denote the eigenvectors of (4.13) corresponding to $\hat{\lambda}^*_{r+1},\dots,\hat{\lambda}^*_{p+1}$.

LEMMA 7.3: *If $K'\beta = 0$, then we let $K^*=(K',0)'$, such that $K^{*\prime}\xi = 0$, then it holds that*

$$(7.8)\qquad TK^{*\prime}\hat{v}^*\hat{v}^{*\prime}K^* \simeq TK'\beta_\perp(\beta_\perp'S_{kk}\beta_\perp)^{-1}\beta_\perp'K$$

which converges in distribution to

$$(7.9)\qquad K'\beta_\perp(\beta_\perp'C\textstyle\int FF'du C'\beta_\perp)^{-1}\beta_\perp'K,$$

where $F = W-\overline{W}$.

PROOF: The proof is almost the same as for Lemma 7.2 but differs since the smallest eigenvalue of (4.13) is always equal to 0. We apply the coordinate system given by $\beta^* = (\beta',-\beta_0')'$, $\gamma^* = (\beta_\perp',0)'$ and $\xi = (0,1)'$ and find

$$\hat{v}^* = \beta^* e^* + \gamma g^* + \xi f^*.$$

We now use the equation $\hat{v}^{*\prime} S^{*}_{kk} \hat{v}^{*} = I$ to note that $\hat{v}^{*}$ and also (e^{*}, g^{*}, f^{*}) are bounded. Further

$$\hat{v}^{*\prime}_i (S^{*}_{k0} S^{*-1}_{00} S^{*}_{0k}) \hat{v}^{*}_i = \hat{\lambda}^{*}_i \in o_P(T^{-1}),$$

implies that the coordinates e^{*} and g^{*} are $O_P(T^{-1/2})$ and that

$$(\gamma^{*} g^{*} + \xi f^{*})' S^{*}_{kk} (\gamma^{*} g^{*} + \xi f^{*}) \overset{P}{\to} I.$$

By an argument similar to (7.7) we find

$$TK^{*\prime}(\gamma^{*}, \xi)\{(\gamma^{*}, \xi)' S^{*}_{kk} (\gamma^{*}, \xi)\}^{-1} (\gamma^{*}, \xi)' K^{*} =$$

$$TK' \beta_{\perp} (\gamma^{*\prime} S^{*}_{kk} \gamma^{*} - \gamma^{*\prime} S^{*}_{kk} \xi (\xi' S^{*}_{kk} \xi)^{-1} \xi' S^{*}_{kk} \gamma^{*})^{-1} \beta_{\perp} K,$$

since $K^{*\prime}\xi = 0$ and $K^{*\prime}\gamma^{*} = \beta_{\perp}$. Now the proof is as before using that

$$\gamma^{*\prime} S^{*}_{kk} \gamma^{*} - \gamma^{*\prime} S^{*}_{kk} \xi (\xi' S^{*}_{kk} \xi)^{-1} \xi' S^{*}_{kk} \gamma^{*} = \beta_{\perp}' S_{kk} \beta_{\perp}.$$

THEOREM 7.4: *Under the hypothesis $H_3 : K'\beta = 0$, where K is $p \times (p-s)$ of full rank, the asymptotic distribution of*

(7.9) $$T\,tr\{(K'\hat{\Pi}'(S_{00} - \hat{\alpha}\hat{\alpha}')^{-1}\hat{\Pi}K)(K'\hat{\upsilon}\hat{\upsilon}'K)^{-1}\}$$

or

(7.10) $$T\,tr\{(K'\hat{\beta}(\hat{D}^{-1} - I)^{-1}\hat{\beta}'K)(K'\hat{\upsilon}\hat{\upsilon}'K)^{-1}\},$$

where $\hat{D} = diag(\hat{\lambda}_1, \ldots, \hat{\lambda}_r)$, is χ^2 with $(p-s)r$ degrees of freedom.

PROOF: It follows from Proposition 5.3 that if $K'\gamma \neq 0$, and $\psi \neq 0$ the asymptotic distribution of $TK'(\hat{\beta}b^{-1} - \beta) = TK'\hat{\beta}b^{-1}$ for given F, see (5.4) and (5.5) is Gaussian with mean zero and covariance matrix $K'\gamma(\gamma' C \int GG' du C'\gamma)^{-1}\gamma' K \otimes (\alpha'\Lambda^{-1}\alpha)^{-1}$, where G is W corrected for mean and trend. By Lemma 7.2 this can be estimated by

$$TK'\hat{v}\hat{v}'K \otimes (b\hat{\alpha}'(S_{00} - \hat{\alpha}\hat{\alpha}')^{-1}\hat{\alpha}b')^{-1}.$$

Hence the asymptotic distribution for given G of

$$T\,tr\{(K'\hat{\beta}b^{-1}(b\hat{\alpha}'(S_{00} - \hat{\alpha}\hat{\alpha}')^{-1}\hat{\alpha}b')b'^{-1}\hat{\beta}'K(K'\hat{v}\hat{v}'K)^{-1}\}$$

is $\chi^2((p-s)r))$. Since this result does not depend on G, the result holds unconditionally.

The alternative form of the test statistic, which is very convenient to calculate once the eigenvalues and the eigenvectors from (4.4) have been found, is derived as follows: We insert the estimates into the identity

$$\beta'\Sigma_{kk}\beta(\beta'\Sigma_{k0}\Sigma_{00}^{-1}\Sigma_{0k}\beta)^{-1}\beta'\Sigma_{kk}\beta - \beta'\Sigma_{kk}\beta = (\alpha'\Lambda^{-1}\alpha)^{-1}.$$

If $\psi = 0$ the same result holds, by applying the limit results that hold under this assumption, see (5.2).

Finally if $K'\gamma = 0$ then $K'\psi \neq 0$. In this case Proposition 5.1 implies that $T^{3/2}K'\hat{\beta}b^{-1}$ has an asymptotic distribution, which for given F is Gaussian with mean zero and variance matrix

$$K'\psi(\psi'\psi\int(u-1/2-a(W-\overline{W}))^2du\psi'\psi)^{-1}\psi'K \otimes (\alpha'\Lambda^{-1}\alpha)^{-1},$$

where $a = \int(u-1/2)W(u)du/\int(W-\overline{W})^2du$. This, however, can be estimated by

$$T^2K'\hat{v}\hat{v}'K \otimes (b\hat{\alpha}'(S_{00} - \alpha\alpha')^{-1}\hat{\alpha}'b),$$

see Lemma 7.2, and then the proof is as before.

Next we shall give the result for testing the restrictions on β when there is no trend.

THEOREM 7.5: *Under the hypothesis* H_3^* : $K'\beta = 0$ *and* $\alpha_\perp'\mu = 0$, *where* K *is* $p\times(p-s)$ *of full rank, the asymptotic distribution of*

$$(7.9) \qquad Ttr\{(K^{*\prime}\hat{\Pi}^{*\prime}(S_{00} - \hat{\alpha}^*\hat{\alpha}^{*\prime})^{-1}\hat{\Pi}^*K^*)(K^{*\prime}\hat{v}^*\hat{v}^{*\prime}K^*)^{-1}\}$$

or

$$(7.10) \qquad Ttr\{(K^{*\prime}\hat{\beta}^*(\hat{D}^{*-1} - I)^{-1}\hat{\beta}^{*\prime}K^*)(K^{*\prime}\hat{v}^*\hat{v}^{*\prime}K^*)^{-1}\},$$

where $\hat{D}^* = diag(\hat{\lambda}_1^*,\ldots,\hat{\lambda}_r^*)$, *is* χ^2 *with* $(p-s)r$ *degrees of freedom, here* $K^* = (K',0)'$ *and* $\Pi^* = \alpha\beta^{*\prime}$.

PROOF: The proof is the same as for Theorem 7.4. Note that since $K^{*\prime}\xi = 0$, then $K^{*\prime}\hat{\beta}^{*}$ only involves the estimate of the cointegration vector and not the constant term. A similar remark holds for $K^{*\prime}\hat{v}^{*}$.

We shall apply the result to the special case when r = 1 and we have only one cointegration relation where we want to test some linear constraint on the coefficients. We shall formulate the result as a Corollary.

COROLLARY 7.6: *If only 1 cointegration vector* β *is present* $(r = 1)$, *and if we want to test the hypothesis*

$$K'\beta = 0$$

then , the test statistic

$$T^{1/2}K'\hat{\beta}/\{(\hat{\lambda}_1^{-1} - 1)(K'\hat{v}\hat{v}'K)\}^{1/2} \tag{7.12}$$

is asymptotically normalized Gaussian. Here $\hat{\lambda}_1$ *is the maximal eigenvalue and* $\hat{\beta}$ *the corresponding eigenvector of the equation*

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

The remaining eigenvectors form $\hat{v}$. *A similar result holds for the model with no trend.*

The normalization $\hat{v}'S_{kk}\hat{v} = I$ implies that $\hat{v}'\hat{v}$ is of the order of T^{-1} which shows that $\hat{\beta}$ is really normalized by T.

Thus if there is only one cointegration vector $\hat{\beta}$ one can think of the matrix $(\hat{\lambda}_1^{-1}-1)\hat{v}\hat{v}'/T$ as giving an estimate of the asymptotic "variance" of $\hat{\beta}$.

This result should be interpreted with care since $K'\hat{\beta}$ is not asymptotically Gaussian and may not have an asymptotic variance, but one can still normalize a linear combination of the components of $\hat{\beta}$ in such a way that it becomes asymptotically Gaussian.

REFERENCES

[1] Ahn, S. K., and G. C. Reinsel (1987): "Estimation for Partially Non-stationary Multivariate Autoregressive Models," University of Wisconsin.

[2] Anderson, T. W. (1984): *An introduction to multivariate statistical analysis*. New York: Wiley.

[3] Box, G. E. P., and G. C. Tiao (1981): "A Canonical Analysis of Multiple Time Series with Applications," *Biometrika*, 64, 355-365.

[4] Davidson, J. (1986): "Cointegration in Linear Dynamic Systems," Discussion paper, LSE.

[5] Engle, R. F. and C. W. J. Granger (1987): " Co-integration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55, 251-276.

[6] Fuller W. (1976): "*Introduction to Statistical Time Series*," New York: Wiley.

[7] Granger, C. J. (1981): " Some Properties of Time Series Data and their Use in Econometric Model Specification,"*Journal of Econometrics*, 16, 251-276.

[8] Granger, C. W. J., and A. A. Weiss (1983): "Time Series Analysis of Error Correction Models," in *Studies in economic time series and multivariate statistics*, ed. by S. Karlin, T. Amemiya and L. A. Goodman. New York: Academic Press.

[9] Jeganathan, P. (1988): "Some Aspects of Asympotic Theory with Applications to Time Series Models," Technical Report The University of Michigan.

[10] Johansen, S.(1988a): "The Mathematical Structure of Error Correction Models," *Contemporary Mathematics*, American Mathematical Society, 12, 359-386.

[11] ________(1988b): "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control*, 12, 231-254.

[12] Johansen, S. and K. Juselius (1988): "Hypothesis Testing for Cointegration vectors - with an application to the Demand for Money in Denmark and Finland," Preprint University of Copenhagen

[13] ________(1989): "Full Information Maximum Likelihood Procedure for Inference on Cointegration - with Applications," Preprint University of Copenhagen

[14] Pena, D. and G. E. P. Box (1987): "Identifying a Simplifying Structure in Time Series," *Journal of the American Statistical Association*, 82, 836-843.

[15] Phillips, P. C. B. (1987): "Multiple Regression with Integrated Time Series," Cowles Foundation discussion paper No. 852.

[16] ————(1988): "Optimal inference in cointegrated systems," Cowles Foundation discussion paper No.866.

[17] Phillips, P. C. B., and S. N. Durlauf (1986): "Multiple Time Series Regression with Integrated Processes," *Review of Economic Studies*, 53, 473-495.

[18] Phillips, P. C. B., and S. Ouliaris (1986): " Testing for Cointegration," Cowles Foundation discussion paper No 809.

[19] ————(1987): "Asymptotic Properties of Residual based Tests for Cointegration," Cowles Foundation discussion paper No.847.

[20] Phillips, P. C. B., and J. Y. Park (1986a): " Asymptotic Equivalence of OLS and GLS in Regression with Integrated Regressors," Cowles Foundation discussion paper No.802.

[21] ————(1986b): "Statistical Inference in Regressions with Integrated Processes : Part 1," Cowles Foundation discussion paper No. 811.

[22] ————(1987): "Statistical Inference in Regressions with Integrated Processes : Part 2," Cowles Foundation discussion paper No. 819.

[23] Sims, A., J. H. Stock, and M. W. Watson (1986): " Inference in Linear Time Series Models with some Unit Roots," Preprint.

[24] Stock, J.H. (1987): "Asymptotic Properties of Least Squares Estimates of Cointegration Vectors," *Econometrica*, 55, 1035-1056.

[25] Stock, J. H., and M. W. Watson (1987): " Testing for Common Trends," Working paper in Econometrics, the Hoover Institution, Stanford.

[26] Tso, M.K-S (1981): "Reduced-rank Regression and Canonical Analysis," Journal of the Royal Statistical Society B, 43,183-189.

[27] Velu, R. P., G. C. Reinsel, and D. W. Wichern (1986): "Reduced Rank Models for multiple time series," *Biometrika*, 73, 105-118.

[28] Velu, R. P., D. W. Wichern, and G. C. Reinsel (1987): "A Note on Non-stationary and Canonical Analysis of Multiple Time Series Models," *Journal of Time Series Analysis*, 8, 479-487.

PREPRINTS 1988

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 Jacobsen, Martin: Discrete Exponential Families: Deciding when the Maximum Likelihood Estimator Exists and Is Unique.

No. 2 Johansen, Søren and Juselius, Katarina: Hypothesis Testing for Cointegration Vectors - with an Application to the Demand for Money in Denmark and Finland.

No. 3 Jensen, Søren Tolver, Johansen, Søren and Lauritzen, Steffen L.: An Algorithm for Maximizing a Likelihood Function.

No. 4 Bertelsen, Aksel: On Non-Null Distributions Connected with Testing that a Real Normal Distribution Is Complex.

No. 5 Tjur, Tue: Statistical Tables for Personal Computer Users.

No. 6 Tjur, Tue: A New Upper Bound for the Efficiency of a Block Design.

No. 7 Bunzel, Henning, Høst, Viggo and Johansen, Søren: Some Simple Non-Parametric Tests for Misspecification of Regression Models Using Sign Changes of Residuals.

No. 8 Brøns, Hans and Jensen, Søren Tolver: Maximum Likelihood Estimation in the Negative Binomial Distribution.

No. 9 Andersson, S.A. and Perlman, M.D.: Lattice Models for Conditional Independence in a Multivariate Normal Distribution.

PREPRINTS 1989

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 Bertelsen, Aksel: Asymptotic Expansion of a Complex Hypergeometric Function.

No. 2 Davidsen, Michael and Jacobsen, Martin: Weak Convergence of Twosided Stochastic Integrals, with an Application to Models for Left Truncated Survival Data.

No. 3 Johansen, Søren: Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models.