Søren Johansen

The Power Function of the Likelihood Ratio Test for Cointegration

Preprint
September
1989
8

Institute of Mathematical Statistics
University of Copenhagen

ISSN 0902-8846

Søren Johansen

THE POWER FUNCTION OF THE LIKELIHOOD RATIO TEST FOR COINTEGRATION

Preprint 1989 No. 8

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

September 1989

University of Copenhagen

Institute of

Mathematical Statistics

THE POWER FUNCTION OF THE LIKELIHOOD RATIO TEST FOR COINTEGRATION

by

Søren Johansen.

September 1989

Abstract.

The distribution of the likelihood ratio test for cointegration under local alternatives is found to be given by a suitable functional of a p-r dimensional Ornstein-Uhlenbeck process. The results are related to those of Phillips (1988), who considered near integrated processes, and derived the power function of a test for unit roots based on a regression estimate. The power function is investigated numerically for a one-dimensional alternative.

1. Introduction

The concept of cointegration was introduced by Granger (1981) in order to define the notion of a stable economic relation among non stable economic variables.

More preceisely he considered *non stationary* economic variables, i.e. a non stationary vector process, and defined a cointegrating relation as a linear combination of the components with the property that it determined a *stationary* stochastic process. This makes precise one of the many meanings of the notion of stability and as such it can be investigated and tested by statistical techniques, see Engle and Granger (1987), Phillips and Park (1986) and Phillips and Ouliaris (1987). Many papers have since then been devoted to finding properties of various regression and eigenvector estimates for cointegration vectors.

This paper deals with vector autoregressive processes with Gaussian errors where maximum likelihood estimators and likelihood ratio tests can be found, see Johansen (1988b). There are now a number of papers that describe this method in detail, and apply it to economic problems, see Johansen and Juselius (1989), Juselius (1989), Kunst (1988) and (1989), Kunst and Neusser (1988), Lütkepohl and Reimers (1989), Hoffman and Rasche (1989), Hall (1988), Garbers (1989). The method will be described briefly below.

Consider therefore the p-dimensional vector autoregressive process X_t, $t = 1,\dots,T$, defined by the equations:

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

where $\epsilon_1,\ldots,\epsilon_T$ are independent Gaussian variables with mean zero and variance matrix Λ and $X_0,\ldots,X_{-k+1}$ are fixed. The parametres are the p×p matrices $(\Pi_1,\ldots,\Pi_k,\Lambda)$. Let $\Pi = -I + \Pi_1 +\ldots+ \Pi_k$ be the total impact matrix, and consider the hypothesis of the existence of (at most) r cointegration vectors formulated as

(1.2) $$H: \quad \Pi = \alpha\beta',$$

where α and β are p×r matrices.

The maximum likelihood estimation and likelihood ratio test of this model were investigated in Johansen (1988), and can be described as follows. First the model (1.1) is rewritten[1] as

(1.3) $$\Delta X_t = \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \alpha\beta' X_{t-k} + \epsilon_t, \quad t = 1,\ldots,T,$$

where $\Gamma_i = -I + \Pi_1 + \ldots + \Pi_i$, $i = 1,\ldots,k-1$. In this model the parameters $(\Gamma_1,\ldots,\Gamma_{k-1},\alpha,\beta,\Lambda)$ are variation independent and one can easily maximize with respect to the parameters $\Gamma_1,\ldots,\Gamma_{k-1}$ by regressing ΔX_t and X_{t-k} on the lagged differences, and obtain residuals R_{0t} and R_{kt}. Then we define

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

and find $\hat{\alpha}(\beta) = S_{0k}\beta(\beta' S_{kk}\beta)^{-1}$. Finally $\hat{\beta}$ is found as the eigenvectors corresponding to the r largest eigenvalues of the equation

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

---

[1]One can also rewrite the model as

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i^* \Delta X_{t-i} + \epsilon_t$$

with $\Gamma_i^* = \Gamma_i - \Pi = -\Pi_{i+1} - \ldots - \Pi_k$. It may be convenient to see the error correction term with one lag, but the analysis remains the same.

giving the maximized likelihood function

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

Hence the likelihood ratio test for the hypothesis H of (at most) r cointegrating relations is given by

$$(1.5) \qquad -2 \ln(Q) = T \sum_{i=r+1}^{p} \ln(1 - \hat{\lambda}_i).$$

The limiting distribution of this statistic is non standard and has been tabulated by simulation in Johansen (1988). The theory for the model which allows a constant term is a bit more involved, and is given in Johansen (1989).

It is the purpose of this paper to derive the power function of the above likelihood ratio test, applying the results of Phillips (1988).

The alternative we are interested in is clearly that there are one or more extra cointegrating relations than assumed under H. If we investigate the power of the test under the assumption that there is in fact another cointegration relation with some non zero loadings, then it is not difficult to see, that the power tends to one. More interesting is it to consider local alternatives of the form

$$(1.6) \qquad H_T: \; \Pi_T = \alpha\beta' + \alpha_1\beta_1'/T$$

where α_1 and β_1 are p×s matrices. Under the alternative H_T we are thus allowing s extra cointegration vectors to enter the model with small weight α_1/T. These extra linear components of X_t are "near integrated" in the terminology of Phillips (1988), and the basic mathematical results are directly taken from that paper and adopted to this slightly different framework of analysing the multivariate Gaussian distribution.

In the next section we describe the asymptotic properties of near integrated series, and apply these to find an asymptotic expression for the power function. In section 3 we investigate by simulation the power function for the local alternative of one extra cointegration vector with a small loading. The appendix contains the proofs of the mathematical results.

2. Asymptotic properties of the process X_t and the power function under local alternatives.

In order to discuss the asymptotic properties we shall recall some of the properties of the process X_t as given in Theorem 3.1 (Grangers representation theorem) Johansen (1989). We assume that rank(α) = rank(β) = r and let $\alpha_\perp$ and $\beta_\perp$ denote p×(p−r) matrices orthogonal to α and β and of full rank. The condition of "balance", see Johansen (1988a), expressed as a relation between the impact matrix $\Pi = \alpha\beta'$ and the mean lag matrix $\Psi = \sum_{i=1}^{k} i\Pi_i$, is given by

(2.1) $$\text{rank}(\alpha_\perp'\Psi\beta_\perp) = p-r.$$

Under this condition we have the moving average representation $\Delta X_t = C(L)\epsilon_t$ and the expression

(2.2) $$C(1) = \beta_\perp(\alpha_\perp'\Psi\beta_\perp)^{-1}\alpha_\perp' = C,$$

say. From this it is seen that ΔX_t is stationary, $\beta' X_t$ is stationary and X_t is non stationary.

In order to separate the stationary part of the process from the near integrated and integrated components under the model H_T we introduce the r-dimensional process $Y_t^{(T)} = (\beta'\beta)^{-1}\beta' X_t^{(T)}$ and the (p−r)-dimensional process $Z_t^{(T)} = (\beta_\perp'\beta_\perp)^{-1}\beta_\perp' X_t^{(T)}$, such that

(2.3) $$X_t^{(T)} = \beta Y_t^{(T)} + \beta_\perp Z_t^{(T)} .$$

THEOREM 1. *Under the local alternatives*

$$H_T : \Pi_T = \alpha\beta' + \alpha_1\beta_1'/T,$$

and under the condition (2.1) and

(2.4) $$rank\{(\alpha,\alpha_1)_\perp \Psi(\beta,\beta_1)_\perp\} = p-r-s$$

the process $X_t^{(T)}$ *converges weakly to the process* X_t, *and the process* $Z_{[Tt]}^{(T)} T^{-½}$ *converges weakly to the* $(p-r)$*-dimensional process* K_t, which *satisfies the stochastic differential equation*

(2.5) $$- ab'\int_0^t K_u du + K_t = B_t .$$

Here B_t *is a* $(p-r)$*-dimensional Brownian motion, and* $a = \alpha_\perp'\alpha_1$ *and* $b = \beta_\perp'\beta_1$.

The $(p-r)\times(p-r)$ *matrix* ab' *has rank* s *and the solution to* (2.5) *is the Ornstein-Uhlenbeck process in* $p - r$ *dimensions which can be expressed as*

(2.6) $$K_t = \int_0^t exp\{ab'(t-u)\}dB.$$

The proof of this result will be given in the the appendix. We shall apply the result in the proof of the main result about the power function:

THEOREM 2. *The asymptotic distribution of the likelihood ratio test* $-2lnQ$ *for the hypothesis* $H\colon \Pi = \alpha\beta'$ *is, under the local alternative*

$$H_T : \Pi = \alpha\beta' + \alpha_1\beta_1'/T,$$

given by

$$\mathrm{tr}\{\int(dK)K'(\int KK'du)^{-1}\int K(dK)'\}, \tag{2.7}$$

where K *is given by the* $(p - r)$*-dimensional Ornstein-Uhlenbeck process* (2.6).

The proof is given in the appendix. If the (p-r)-dimensional process K_t has coordinates $\{K_i(t),\ i= 1,\dots\ p-r\}$, then the notation $\int(dK)K'$ stands for a (p-r)×(p-r) matrix of stochastic integrals with elements $\int dK_iK_j = \int_0^1 K_j dK_i$. The matrix $\int KK'du$ has elements $\int_0^1 K_i(u)K_j(u)du$.

This result is closely related to the result by Phillips (1988) p.1031 on the power function for a certain test statistic for the hypothesis $\Pi = 0$ based on a normalization of a regression estimate of Π. In this situation, where we want to test $\Pi = 0$, the present test which exploits the Gaussian distribution can not be distinguished by its asymptotic properties for local alternatives from the general test given by Phillips.

It is seen that the asymptotic power function for local alternatives depends on the parameters only through $a = \alpha_\perp'\alpha_1$ and $b = \beta_\perp'\beta_1$. Thus it depends on how the extra loadings (α_1) and cointegrating relations (β_1) are related to the α and β assumed under H. A tabulation of the power function thus involves 2(p-r)×s parameters. It turns out, however, that fewer parameters will do. In order to see this, we shall exploit the invariance of the multivariate Brownian motion under rotation and the invariance of the test statistic.

Let O denote a $(p-r)\times(p-r)$ orthonormal matrix. Then clearly B_t and OB_t have the same distribution and hence multiplying (2.5) by O we see that OK_t solves the equation (2.5) for the Brownian motion OB_t with the $p\times s$ matrices Oa and Ob. The test statistic is independent of this transformation and thus the power function is the same for any alternatives ab' and a_1b_1' such that $Oab'O' = a_1b_1'$. We can simplify still further by choosing new coordinates using the orthonormal vectors

$$e_1 = b(b'b)^{-1/2}$$
$$e_2 = (a - b(b'b)^{-1}b'a)(a'a - a'b(b'b)^{-1}b'a)^{-1/2}$$

and finally e_3 such that (e_1,e_2,e_3) are orthonormal. We define $K_{it} = e_i'K_t$ and $B_{it} = e_i'B_t$ $i = 1,2,3$. then the equation (2.5) is equivalent to

$$- (b'b)^{-1/2}b'a(b'b)^{1/2} \int_0^t K_{1u}du + K_{1t} = B_{1t}, \tag{2.7}$$

$$- (a'a - a'b(b'b)^{-1}b'a)^{1/2}(b'b)^{1/2}\int_0^t K_{1u}du + K_{2t} = B_{2t}, \tag{2.8}$$

$$K_{3t} = B_{3t}. \tag{2.9}$$

This shows that the power function only depends on α_1 and β_1 through the matrices a'a, b'b and a'b. For $s = 1$ this result simplifies still further.

COROLLARY. *Under the local alternative $\Pi = \alpha\beta' + \alpha_1\beta_1'/T$, where α_1 and β_1 are $p\times 1$ vectors, the power function for the likelihood ratio test for cointegration depends on α_1 and β_1 only through the quantities f and g given by*

$$f = a'b = \beta_1'C\alpha_1 < 0, \tag{2.10}$$

and

$$(2.11)\qquad g^2 = \alpha'ab'b - (\alpha'b)^2$$
$$= (\alpha_1'\alpha_\perp(\alpha_\perp'\Lambda\alpha_\perp)^{-1}\alpha_\perp'\alpha_1)(\beta_1'C\Lambda C'\beta_1) - (\beta_1'C\alpha_1)^2 ,$$

where $C = \beta_\perp(\alpha_\perp'\Psi\beta_\perp)^{-1}\alpha_\perp'$, *see* (2.2).

Note that the expressions above for C as well as $\alpha_\perp(\alpha_\perp'\Lambda\alpha_\perp)^{-1}\alpha_\perp'$ do not depend on the particular choice for $\alpha_\perp$ and $\beta_\perp$. Note also that for p − r = 1 the power function depends only on f, since (2.7) above describe the process K_t..

The proof of the corollary is given in the appendix.

In order to interprete these results we consider the hypothesis H_T for the simple example we get from (1.3) by letting k = 1, that is we consider the system

$$\Delta X_t^{(T)} = T^{-1}\alpha_1\beta_1'X_{t-1}^{(T)} + \epsilon_t.$$

Thus we have left out the short term dynamics and assumed that $\alpha = \beta = 0$. We then want to investigate if we can find a cointegration vector β_1 in the data. In this case $\Pi = 0$, $\alpha_\perp = \beta_\perp = I$ and $\Psi = I$. Thus C = I, $a = \alpha_1$ and $b = \beta_1$, such that $f = \alpha_1'\beta_1$ and $g^2 = \alpha_1'\alpha_1\beta_1'\beta_1 - (\alpha_1'\beta_1)^2$.

We can express the results as follows: The power for finding a stationary relation β_1 with loadings $T^{-1}\alpha_1$ depends on the position of the vectors α_1 and β_1 through the angle (f/g) between them and the area (g) spanned by them, as well as the dimension (p) of the space in which we are looking for them.

3. Numerical investigation of the power function under the alternative of one extra cointegration vector.

For s = 1 the equations (2.7), (2.8) and (2.9) are of the form

(3.1) $$- f \int_0^t K_{1u} du + K_{1t} = B_{1t},$$

(3.2) $$- g \int_0^t K_{1u} du + K_{2t} = B_{2t},$$

(3.3) $$K_{3t} = B_{3t}.$$

The distribution of the functional (2.7) is too complicated to find analytically but can be found by simulation. In order to simulate the processes we shall apply the discrete version of these equations

(3.4) $$K_{1t} = (1 + f/T)K_{1t-1} + \epsilon_{1t},$$

(3.5) $$K_{2t} = K_{2t-1} + (g/T)K_{1t-1} + \epsilon_{2t},$$

(3.6) $$K_{3t} = K_{3t-1} + \epsilon_{3t},$$

$t = 1,\ldots,T$ starting with $K_0 = 0$.

Note that the actual expression for the limit distribution is not so important. What is being applied is that the limit distribution exists and does not involve the parametres, but for f and g. Thus in order to simulate the system we simply choose the simplest possible system compatible with the given f and g.

We then easily solve the equations (3.4), (3.5) and (3.6) recursively and form the $T\times(p-r)$ matrix M with elements $M_{ti} = K_{it}$. Then we calculate ΔM and M_{-1} i.e. the differences and lagged variables respectively and find the test statistic

$$\text{Test} = \text{tr}\{\Delta M' M_{-1}(M_{-1}' M_{-1})^{-1} M_{-1}' \Delta M\}.$$

The number of observations T has to be chosen so large that the approximation of the random walk to the Brownian motion is sufficiently good. We have chosen T = 400.

We find the results for stationary alternatives and p-r = 1,2, and 3 in Table I and Figures 1, 2 and 3 based on 2000 - 5000 simulations and T = 400. It is seen that, not surprisingly, the power decreases as the dimension increases, i.e. if there are many dimensions to hide in, i.e. it is difficult to find the cointegrating vector, if it has a small loading. It was found that the non stationary alternatives (not shown) are readily picked up by the test with large power. The test appears unbiased as f and g move away from 0.

One can interprete the coefficient $1 + f/T = 1 + \alpha_1'\beta_1/T$ as the autoregressive parameter, see (3.4), in the stationary (or near integrated) relation we are trying to find. Hence we see from Table I, that if for instance we have T = 100 observations and expect an autoregressive coefficient of say .79 then we have f = 100*(.79 - 1.00) = - 21. Now the power of finding such a process depends on the relation between the loadings and the cointegration vector. If α_1 and β_1 are proportional, such that g = 0, then, if p = 1, we have a probability of .998 of rejecting the hypothesis of non statonarity, and hence of finding a stationary relation. If, however, the system is of dimension p = 3, then the probability of rejecting the non stationarity hypothesis is only .346. For a given angle that is for fixed f/g it is seen from the tables that the larger the vectors the easier it is to find them.

4. Appendix.

We have here collected the proofs of Theorem 1, Theorem 2 and the Corollary. We first define the variance and covariance matrices conditional on the lagged differences by

$$\Sigma_{00} = \mathrm{Var}(\Delta X_t | \Delta X_{t-1}, \ldots, \Delta X_{t-k+1}),$$
$$\beta'\Sigma_{kk}\beta = \mathrm{Var}(\beta' X_{t-k} | \Delta X_{t-1}, \ldots, \Delta X_{t-k+1}),$$
$$\Sigma_{0k}\beta = \mathrm{Cov}(\Delta X_t, \beta' X_{t-k} | \Delta X_{t-1}, \ldots, \Delta X_{t-k+1}).$$

It follows from the representation (1.3) that, since ϵ_t is independent of the past values of the process X_t, one can find the relations

$$\Sigma_{00} = \alpha\beta'\Sigma_{kk}\beta\alpha' + \Lambda, \tag{4.1}$$

and

$$\Sigma_{0k}\beta = \alpha\beta'\Sigma_{kk}\beta, \tag{4.2}$$

which can be solved for α

$$\alpha = \Sigma_{0k}\beta(\beta'\Sigma_{kk}\beta)^{-1}. \tag{4.3}$$

One can apply these identities to prove

LEMMA 1. *The following relations hold*

$$\Lambda^{-1} - \Lambda^{-1}\alpha(\alpha'\Lambda^{-1}\alpha)^{-1}\alpha'\Lambda^{-1} = \alpha_\perp(\alpha_\perp'\Lambda\alpha_\perp)^{-1}\alpha_\perp' = \tag{4.4}$$
$$\alpha_\perp(\alpha_\perp'\Sigma_{00}\alpha_\perp)^{-1}\alpha_\perp' = \Sigma_{00}^{-1} - \Sigma_{00}^{-1}\alpha(\alpha'\Sigma_{00}^{-1}\alpha)^{-1}\alpha'\Sigma_{00}^{-1}$$

PROOF. The second equality sign holds by the relation (4.1) and the two other are just the well known relation between a block matrix M and its inverse

$$M^{22} - M^{21}(M^{11})^{-1}M^{12} = M_{22}^{-1},$$

expressed in the coordinates given by α and $\alpha_\perp$.

We now choose $\alpha_\perp$, such that

(4.5) $$\alpha_\perp' \Lambda \alpha_\perp = I.$$

This can always be done for if $\tilde{\alpha}_\perp$ is any choice of p-r vectors orthogonal to α, then we just take $\alpha_\perp = \tilde{\alpha}_\perp(\tilde{\alpha}_\perp' \Lambda \tilde{\alpha}_\perp)^{-½}$.

We shall further choose $\beta_\perp$ such that

(4.6) $$\alpha_\perp' \Psi \beta_\perp = I.$$

This is always possible since for any choice of $\tilde{\beta}_\perp$ of p-r vectors orthogonal to β we just take $\beta_\perp = \tilde{\beta}_\perp(\alpha_\perp' \Psi \tilde{\beta}_\perp)^{-1}$ which satisfies (4.6).

These choices will simplify the calculations later, but the results about the power function are of course formulated such that they do not depend on this particular choice.

We now give the proof of Theorem 1. The result follows from the results in Phillips (1988), and we shall here only sketch the calculations involved in the proof.

We write the model (1.1) under H_T in the form

(4.7) $$-(\alpha\beta' + \alpha_1\beta_1'/T)X_t^{(T)} + \Psi\Delta X_t^{(T)} + \Pi_2(L)\Delta^2 X_t^{(T)} = \epsilon_t.$$

The basic idea is to go to the limit in this stochastic difference equation and derive either another stochastic difference equation or a limiting stochastic differential equation. It turns out that only Π and Ψ will be relevant in this last approximation, such that near integrated processes can be approximated by AR(1) processes, see (3.4).

If we let $T \to \infty$, in (4.7) we get the equation

$$-\alpha\beta' X_t + \Psi\Delta X_t + \Pi_2(L)\Delta^2 X_t = \epsilon_t,$$

which shows that $X_t^{(T)}$ converges weakly to X_t. If we multiply (4.7) by $\alpha_\perp'$ we get instead

$$- T^{-1}\alpha_{\perp}'\alpha_1\beta_1'X_t^{(T)} + \alpha_{\perp}'\Psi\Delta X_t^{(T)} + \alpha_{\perp}'\Pi_2(L)\Delta^2X_t^{(T)} = \alpha_{\perp}'\epsilon_t.$$

Summing over $t = 1,\ldots,[Tt]$, and dividing by $T^{1/2}$ we obtain, apart from some initial values,

$$- aT^{-1}\sum_{t=1}^{[Tt]}\beta_1'X_t^{(T)}/T^{1/2} + \alpha_{\perp}'\Psi X_{[Tt]}^{(T)}/T^{1/2} + \alpha_{\perp}'\Pi_2(L)\Delta X_{[Tt]}^{(T)}/T^{1/2} = \sum_{t=1}^{[Tt]}\alpha_{\perp}'\epsilon_t/T^{1/2}. \tag{4.8}$$

Now apply the decomposition (2.3) and use the result that $\beta Y_{[Tt]}^{(T)}/T^{1/2} = \beta(\beta'\beta)^{-1}\beta' X_{[Tt]}^{(T)}/T^{1/2} \overset{P}{\to} 0$ since $\beta' X_t$ is stationary. Hence we can replace X_t by $\beta_{\perp}Z_t$.

The right hand side converges to a Brownian motion with covariance $\alpha_{\perp}'\Lambda\alpha_{\perp} = I$, by the choice of $\alpha_{\perp}$, and the first term is $- ab'T^{-1}\sum_{t=1}^{[Tt]}Z_t^{(T)}/T^{1/2}$, i.e. a Riemann sum of the process $Z_t^{(T)}/T^{1/2}$, which converges weakly to the integral $- ab'\int_0^t K_u du$, whereas the second term converges to $\alpha_{\perp}'\Psi\beta_{\perp}K_t = K_t$ by the choice of $\beta_{\perp}$.

It is easily checked that the solution is given by (2.6). In order to see that the matrix ab' has rank s we have to apply the conditions (2.1) and (2.4) which guarantee that the process X_t is at most I(1) both under H and H_T. Let $b_1 = \beta_{\perp}(\beta_{\perp}'\beta_{\perp})^{-1}\beta_{\perp}'\beta_1$ and let b_2 be orthogonal to (β,b_1) such that (β,b_1,b_2) has full rank. Then $\beta_{\perp} = (b_1,b_2)$ and $(\beta,\beta_1)_{\perp} = b_2$. Similarly we define a_1 and a_2 from α_1 and $\alpha_{\perp}$. The condition (2.1) can now be expressed as

$$\text{rank}\begin{bmatrix} a_1'\Psi b_1 & a_1'\Psi b_2 \\ a_2'\Psi b_1 & a_2'\Psi b_2 \end{bmatrix} = p-r,$$

whereas the condition (2.4) can be expressed as

$$\mathrm{rank}(a_2'\Psi b_2) = p-r-s.$$

These conditions clearly imply that

$$A = a_1'\Psi b_1 - a_1'\Psi b_2(a_2'\Psi b_2)^{-1}a_2'\Psi b_1$$

is of full rank s. This therefore also holds for $a_1'\Psi b_1$ and hence for a_1 and b_1. Now

$$ba' = \beta_\perp'\beta_1\alpha_1'\alpha_\perp = \beta_\perp' b_1 a_1'\alpha_\perp = \begin{bmatrix} b_1' \\ b_2' \end{bmatrix} b_1 a_1'(a_1 a_2) = \begin{bmatrix} b_1 b_1 a_1 a_1 & 0 \\ 0 & 0 \end{bmatrix}$$

which has rank s.

Next we shall turn to the proof of Theorem 2. This follows closely the proof given in Johansen (1988) for the asymptotic distribution under H, and it will therefore not be given in full technical detail. The likelihood ratio test is calculated from (1.5) solving the eigenvalue problem (1.4) The basic idea in the proof is to study what happens as T $\to \infty$ in the equation (1.4) under different normalizations, and then apply the continuity of the ordered eigenvalues as a function of the coefficients.

First multiply the matrix equation (1.4) by $(\beta, T^{-1/2}\beta_\perp)$ and its transposed and apply the stationarity of $\beta' X_t$ to see that $\beta' S_{kk}\beta$, $\beta' S_{kk}\beta_\perp$ are $O_P(1)$. By Theorem 1, $\beta_\perp' S_{kk}\beta_\perp/T$ converges weakly to $\beta_\perp'\beta_\perp \int KK' d\beta_\perp'\beta_\perp$ = Φ, say, and this shows that in the limit λ has to satisfy the equation

$$\left|\lambda \begin{bmatrix} \beta'\Sigma_{kk}\beta & 0 \\ 0 & \Phi \end{bmatrix} - \begin{bmatrix} \beta'\Sigma_{k0} \\ 0 \end{bmatrix} \Sigma_{00}^{-1} \begin{bmatrix} \beta'\Sigma_{k0} \\ 0 \end{bmatrix}'\right| = 0$$

or

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

which has r positive solutions and p-r null solutions. Thus the r largest eigenvalues of (1.4) converge to those of

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

and the p-r smallest tend to zero. The r largest eigenvalues correspond to the stationary components $\beta' X_t$, whereas the p-r smallest correspond to the near integrated as well as the non stationary components.

Now normalize $\hat{\lambda}_i$ by T and define $\hat{\rho}_i = T\hat{\lambda}_i$. We multiply (1.4) from boths sides by $(\beta,\beta_\perp)$ and let $T \to \infty$, then we get that the limiting value of $\hat{\rho}$ will satisfy the equation

$$\left| \begin{bmatrix} 0 & 0 \\ 0 & \rho\Phi \end{bmatrix} - \begin{bmatrix} \beta'\Sigma_{k0} \\ F' \end{bmatrix} \Sigma_{00}^{-1} \begin{bmatrix} \beta'\Sigma_{k0} \\ F \end{bmatrix}' \right| = 0,$$

where F' is the weak limit of $\beta_\perp' S_{k0}$. This can be written

$$|\beta'\Sigma_{k0}\Sigma_{00}^{-1}\Sigma_{0k}\beta|\,|\rho\Phi - F'(\Sigma_{00}^{-1} - \Sigma_{00}^{-1}\Sigma_{0k}\beta(\beta'\Sigma_{k0}\Sigma_{00}^{-1}\Sigma_{0k}\beta)^{-1}\beta'\Sigma_{k0}\Sigma_{00}^{-1})F|=0.$$

Since $\alpha = \Sigma_{0k}\beta(\beta'\Sigma_{kk}\beta)^{-1}$ one can express the matrix between F' and F as

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

which by Lemma 1 equals

$$\alpha_\perp(\alpha_\perp'\Lambda\alpha_\perp)^{-1}\alpha_\perp' = \alpha_\perp\alpha_\perp'$$

by the definition of $\alpha_\perp$, see (4.5).

Thus ρ is the solution to the equation

$$|\rho\Phi - F'\alpha_\perp\alpha_\perp' F|=0 \tag{4.9}$$

The equation (1.3) gives the following relation between the residuals and the residual sum of squares after the preliminary regressions,

$$S_{k0} - S_{kk}(\beta\alpha' + \beta_1\alpha_1' T^{-1}) = T^{-1}\sum_{t=1}^{T} R_{kt}\epsilon_t'.$$

It follows, that $\beta_\perp' S_{k0}\alpha_\perp = F'\alpha_\perp$ has the same limit as

$$\beta_\perp' S_{kk}\beta_1 a' T^{-1} + T^{-1}\sum_{t=1}^{T}\beta_\perp' R_{kt}\epsilon_t'\alpha_\perp,$$

which by Theorem 1 converges to

$$\beta_\perp'\beta_\perp(\int KK'du\, ba' + \int K(dB)'),$$

since $V(\alpha_\perp'\epsilon_t) = \alpha_\perp'\Lambda\alpha_\perp = I$, by the choice of $\alpha_\perp$.

Now apply (2.5) to obtain

$$\beta_{\perp}'\beta_{\perp}(\int Kd(B' + \int K'du\dot{b}a')) =$$

$$\beta_{\perp}'\beta_{\perp}\int K(dK)'.$$

Thus inserting this result in (4.9) we find that the limiting value of $\hat{\rho}$ is a solution to the equation

$$|\rho\int KK'du - \int K(dK)'\int(dK)K'| = 0,$$

and hence the result (4.9) holds, since

$$-2\ln Q = -T\sum_{i=r+1}^{p}\ln(1-\hat{\lambda}_i) \simeq \sum_{i=r+1}^{p}T\hat{\lambda}_i = \sum_{i=r+1}^{p}\hat{\rho}_i \simeq tr\{\int(dK)K'(\int KK'du)^{-1}\int K(dK)'\}.$$

Finally we shall give the proof of the Corollary. This result follows from the equations (2.7), (2.8) and (2.9), and the expressions for the coefficients f and g are easily deduced by applying the expression for C $= \beta_{\perp}(\alpha_{\perp}\Psi\beta_{\perp})^{-1}\alpha_{\perp}'$ together with the identities $\alpha_{\perp}'\Lambda\alpha_{\perp} = I$ and $\alpha_{\perp}'\Psi\beta_{\perp} = I$. We find

$$f = b'a = \beta_1'\beta_{\perp}\alpha_{\perp}'\alpha_1 = \beta_1'\beta_{\perp}(\alpha_{\perp}'\Psi\beta_{\perp})^{-1}\alpha_{\perp}'\alpha_1 = \beta_1'C\alpha_1,$$

which is non zero by the result of Theorem 1. Since we are interested in stationary alternatives only it follows from (2.7) that $a'b < 0$. Furthermore

$$a'a = \alpha_1'\alpha_{\perp}\alpha_{\perp}'\alpha_1 = (\alpha_1'\alpha_{\perp}(\alpha_{\perp}'\Lambda\alpha_{\perp})^{-1}\alpha_{\perp}'\alpha_1)$$

and

$$b'b = \beta_1'\beta_{\perp}\beta_{\perp}'\beta_1 = \beta_1'\beta_{\perp}(\alpha_{\perp}'\Psi\beta_{\perp})^{-1}(\alpha_{\perp}'\Lambda\alpha_{\perp})(\beta_{\perp}'\Psi'\alpha_{\perp})^{-1}\beta_{\perp}'\beta_1 = \beta_1'C\Lambda C'\beta_1$$

5. Acknowledgement.

The author thanks Neil Ericsson for some interesting discussions on the topic at an early stage of the work.

6. References

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

Garbers, H. (1989): The relationship between spot- and forward exchange rates and the relative money supply: the Swiss franc case. Tech. Report. University of Zürich.

Granger, C.W.J. (1981) : Some propertes of times series and their use in econometric model specification. *Journal of Econometrics*, 16, 121-130.

Hall, S.G. (1988): Maximum likelihood estimator of cointegrating vectors: an example of the Johansen procedure. *Oxford Bulletin of Economics and Statistics* (1988)

Hoffman, D.L. and Rasche, R. (1989): Long run income and interest elasticities of money demand in the United States. Tech. Report. Arizona State University.

Johansen, S. (1988a): The mathematical structure of error correction models. *Contemporary Mathematics*, 80, 359-386.

Johansen, S. (1988b): Statistical Analysis of Cointegration. *Journal of Economic Dynamics and Control*, 12, 231-254.

Johansen, S. (1989): Estimation and Hypothesis testing of Cointegration Vectors in Gaussian Autoregressive Models. Preprint University of Copenhagen.

Johansen, S. and Juselius, K. (1989): Maximum likelihood estimation and inference on cointegration - with applications to the demand for money. Oxford Bulletin of Economics and Statistics. to appear.

Juselius, K. (1989): Long-run relations in a well defined statistical model for the DGP. Cointegration analysis of the interest parity relations. Preprint. University of Copenhagen.

Kunst, R.M. (1989): Cointegration in Macroeconomic systems: seasonality and expolosive roots. Tech report. Institute for Advances Studies Vienna

Kunst, R.M. (1988): Cointegration in a Macro-economic system: sensitivity and prediction Tech report. Institute for Advances Studies. Vienna.

Kunst, R.M. and Neusser, K. (1988): Cointegration in a Macro-Economic System. Discussion paper. University of Vienna

Lütkepohl, H. and Reimers, H.-E. (1989): Impulse response analysis of co-integrated systems. Tecnical report. Christian-Albrechts Universität.

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

Phillips, P.C.B. (1988): Regression Theory for Near-Integrated Time Series. Econometrica, 56, 1021-1044.

Phillips, P.C.B. and Ouliaris, S. (1987): Asymptotic properties of residual based tests for cointegration .Cowles Foundation Discussion paper. No 847.

Table I

The asymptotic power of the likelihood ratio test at 5% for r cointegration vectors among p variables under the local alternative of one extra cointegration vector. The quantities f and g are defined in (2.10) and (2.11). The number of simulations are 2000-5000 and T = 400.

$p - r = 1$

f	0	-3	-6	-9	-12	-15	-18	-21
power	.052	.141	.350	.620	.820	.945	.987	.998

g				$p - r = 2$				
0	.060	.069	.105	.175	.272	.416	.565	.714
6	.615	.335	.269	298	.390	.513	.642	.772
12	.944	.850	.760	.716	.717	.760	.830	.874

g				$p - r = 3$				
0	.054	.059	.073	.105	.136	.187	.266	.346
6	.510	.211	.143	.142	.188	.233	.305	.383
12	.834	.694	.536	.424	.407	.429	.476	.543

Legend to Figures.

Fig. 1.

The power function of the likelihood ratio test for cointegration, as a function of f, see (2.10), for g = 0, see (2.11), and different values of the dimension p - r.

Fig. 2.

The power function for the likelihood ratio test for cointegration as a function of f and g, see (2.10) and (2.11) and p-r = 2.

Fig. 3

The power function for the likelihood ratio test for cointegration as a function of f and g, see (2.10) and (2.11) and p-r = 3.

Fig. 1

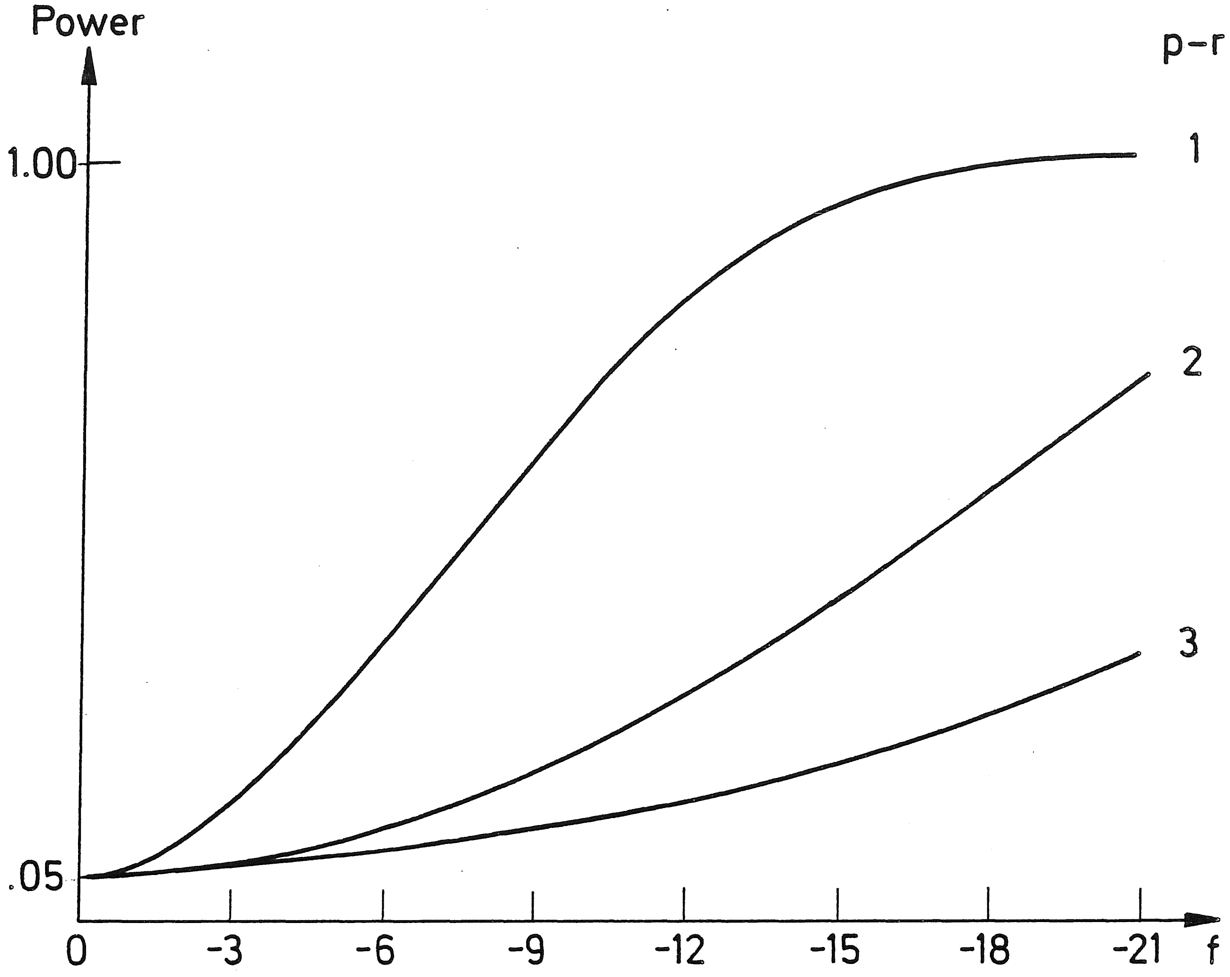


Fig. 2

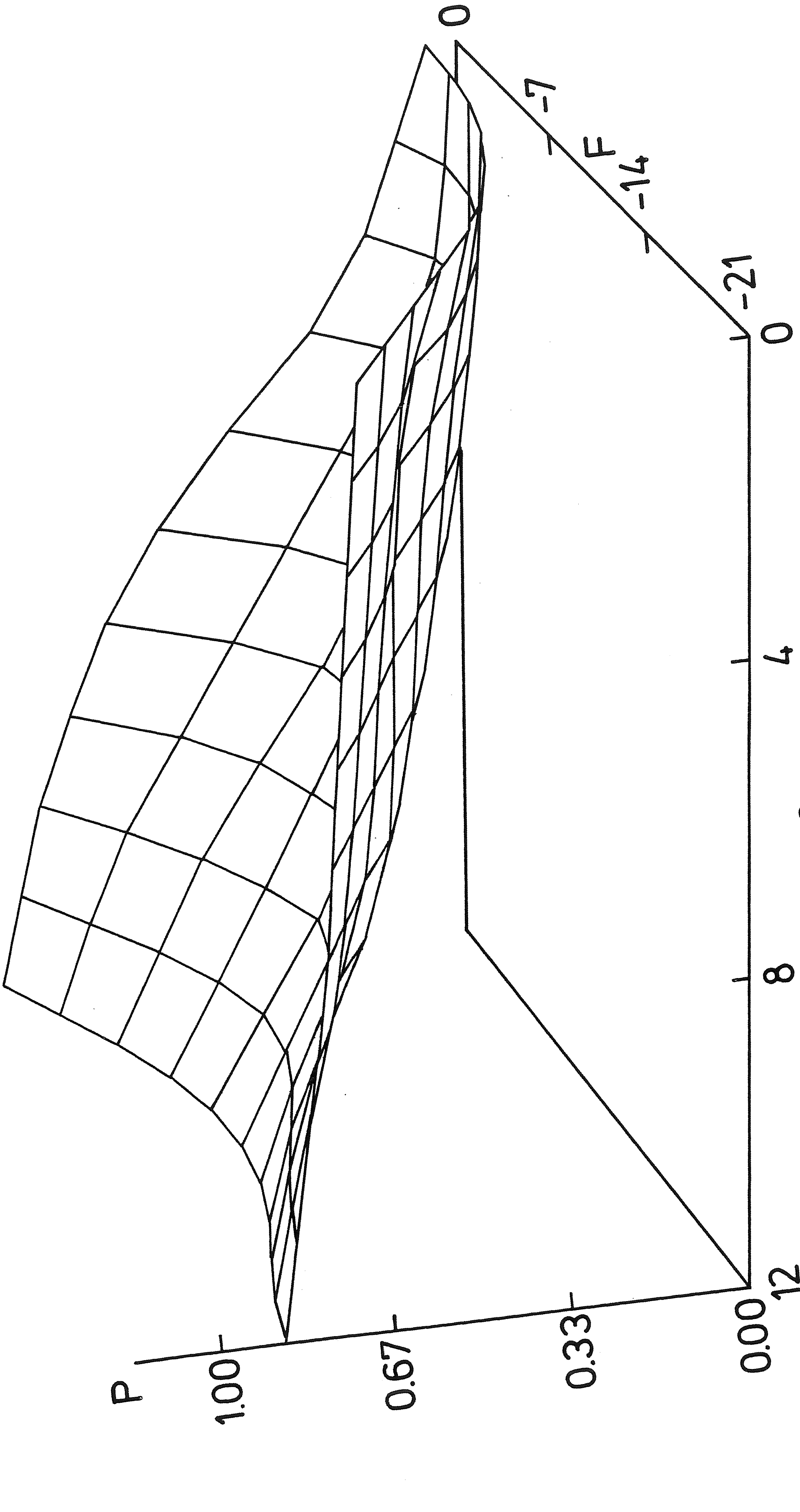
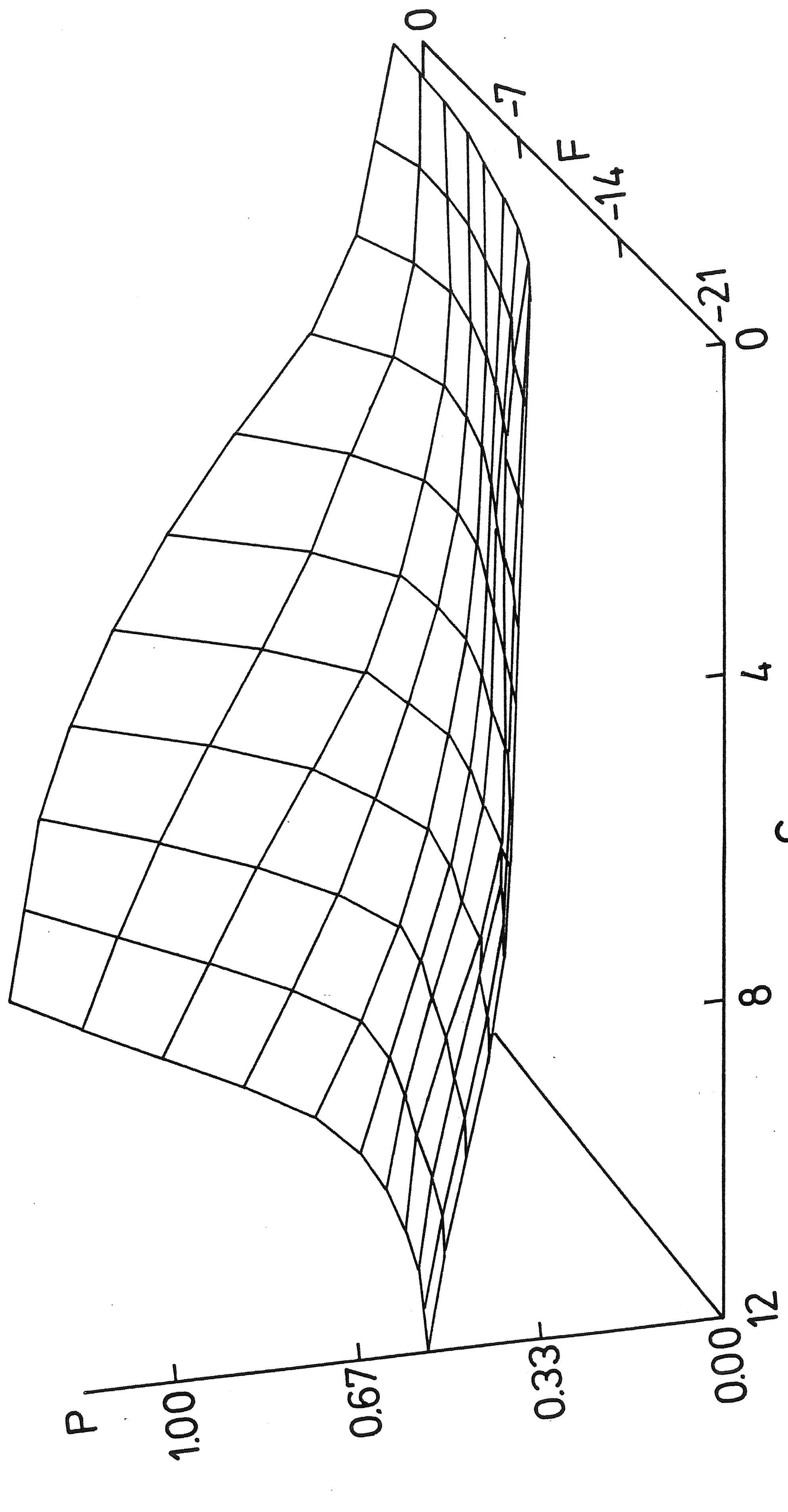


Fig. 3



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.

No. 4 Johansen, Søren and Juselius, Katarina: The Full Information Maximum Likelihood Procedure for Inference on Cointegration - with Applications.

No. 5 Thompson, Steven K.: Adaptive Cluster Sampling.

No. 6 Thompson, Steven K.: Adaptive Cluster Sampling: Designs with Primary and Secondary Units.

No. 7 Thompson, Steven K.: Stratified Adaptive Cluster Sampling.

No. 8 Johansen, Søren: The Power Function of the Likelihood Ratio Test for Cointegration.