JUIN 1985.

N° 8516

TESTING UNKNOWN LINEAR RESTRICTIONS ON PARAMETER FUNCTIONS

C. GOURIEROUX*, A. MONFORT**, E. RENAULT**

* CEPREMAP and University of LILLE I

** Institut National de la Statistique et des Etudes Economiques.

TESTING UNKNOWN LINEAR RESTRICTIONS ON PARAMETER FUNCTIONS

C. GOURIEROUX*, A. MONFORT**, E. RENAULT**

S U M M A R Y

We consider test of hypotheses of the form :

$$H_o = \left\{ \exists \lambda \in R^k : h_o(\theta) = H(\theta)\lambda \right\} ,$$

where θ is the q-dimensional parameter of interest and h_o , H are known functions of size $(p,1)$ and (p,k) . The obtained test procedures are applied to various problems : test of rational expectation, test for common roots, test for the orders of an ARMA process, test for co-integration.

R E S U M E

Nous considérons des tests d'hypothèses s'écrivant sous la forme :

$$H_o = \left\{ \exists \lambda \in R^k : h_o(\theta) = H(\theta)\lambda \right\} ,$$

où θ est un paramètre de dimension q , h_o et H des fonctions connues de tailles respectives $(p,1)$ et (p,k) . Les procédures obtenues sont ensuite appliquées à divers problèmes : test de l'hypothèse d'anticipation rationnelle, test de racines communes, test sur l'ordre d'un processus ARMA, test de co-intégration.

I - INTRODUCTION.

In the standard theory of hypothesis testing the null hypothesis about the parameter θ is formulated in an implicit form, $g(\theta) = o$, or in an explicit form, $\theta = \varphi(\gamma)$, where γ is a parameter whose dimension is smaller than that of θ. However there is a third form, called the mixed form, which naturally appears in many problems ; in this form the null hypothesis is formulated as follows : there exists λ such that $g^*(\theta,\lambda) = o$.

It is easily seen that this general form contains, as particular cases, the implicit form and the explicit form and a test theory based on this general form is developed in Gourieroux-Monfort [8] .

In the present paper, we consider the frequent case where the mixed form is linear in λ. In such a case $g^*(\theta,\lambda) = o$ becomes $h_o(\theta) = H(\theta)\lambda$, where $h_o(\theta)$ is a p-vector and $H(\theta)$ a $p \times K$ matrix. It is clear that, in this case, the null hypothesis is also easily formulated in an implicit form, for instance $g(\theta) = \left\{I - H(\theta)\,[H'(\theta)\,H(\theta)]^{-1}\,H'(\theta)\right\} h_o(\theta) = o$, and the standard theory can, in principle, be used. However we show in this paper that it is, in general, awkward to use this implicit form because it destroys the linear aspect of the mixed form. More precisely, we propose a test, which is asymptotically equivalent to the Wald test, and whose derivation, from the unrestricted estimator of θ, only necessitates the use of O.L.S. and G.L.S. methods. This computational simplicity is appealing since, as mentioned in the examples, the computation of the Wald test based on the implicit form may be very complicated.

In section 2 we give the theoretical results and the other sections are devoted to various applications : test of rational expectation in section 3, test for common roots in section 4, test for the orders of an ARMA process in section 5, and test for co-integration in section 6.

II - THEORETICAL RESULTS.

II.1 - Description of the hypothesis to be tested.

II.1.a - Preliminary results on Wald type tests.

Let us consider an unknown parameter θ belonging to Θ , an open subset of R^q . We assume that $\hat{\theta}_n$ is a sequence of estimators of θ such that :

$$(2.1) \quad \sqrt{n}\,(\hat{\theta}_n - \theta) \xrightarrow[n \to \infty]{d} N[o, Q(\theta)] \quad ,$$

where $Q(\theta)$ is a $q \times q$ non-singular matrix.

Let us also consider the null hypothesis H_o defined by $H_o = \left\{\theta : g(\theta) = o\right\}$, where g is a G-dimensional function which is continuously differentiable on Θ and such that the rank of the $(G \times q)$ jacobian matrix $\frac{\partial g}{\partial \theta'}(\theta)$ is fixed, and denoted by r , for any $\theta \in H_o$.

In this context, a Wald-type test statistic can be defined by :

$$(2.2) \quad \xi_n^W = n\, g'(\hat{\theta}_n) \left[\frac{\partial g}{\partial \theta'}(\hat{\theta}_n)\, Q(\hat{\theta}_n)\, \frac{\partial g'}{\partial \theta}(\hat{\theta}_n) \right]^- g(\hat{\theta}_n) \quad ,$$

where A^- denotes any generalized inverse of A . Under H_o , the asymptotic distribution of ξ_n^W is a chi-square distribution with r degrees of freedom.

The following property provides asymptotically equivalent forms of the previous test statistic.

PROPERTY 2.3 :

> Let $f(g,\theta)$ be a continuously differentiable function such that :
>
> i) $f(o,\theta) = o \quad \forall\, \theta$
>
> ii) for any $\theta \in H_o$, the rank of $\dfrac{\partial f}{\partial g'}(o,\theta)\, \dfrac{\partial g}{\partial \theta'}(\theta)$ is equal to r, the rank of $\dfrac{\partial g}{\partial \theta'}(\theta)$
>
> then, under H_o, the statistic $\overset{\approx}{\xi}{}^W_n$ defined by :
>
> $$\overset{\approx}{\xi}{}^W_n = n\, f[g(\hat\theta_n),\hat\theta_n]' \left[\frac{\partial f}{\partial g'}(o,\hat\theta_n)\, \frac{\partial g}{\partial \theta'}(\hat\theta_n)\, Q(\hat\theta_n)\, \frac{\partial g'}{\partial \theta}(\hat\theta_n) \frac{\partial f'}{\partial g}(o,\hat\theta_n)\right]^{-} f[g(\hat\theta_n),\hat\theta_n]$$
>
> is such that :
>
> $$\overset{\approx}{\xi}{}^W_n = \xi^W_n + o_p(1) \quad .$$
>
> In particular, the asymptotic distribution of $\overset{\approx}{\xi}{}^W_n$, under H_o, is $X^2(r)$.

Proof : Using the Taylor expansion of $f[g(\hat\theta_n),\hat\theta_n]$ around θ we have, for any $\theta \in H_o$:

$$\sqrt{n}\ f[g(\hat\theta_n),\hat\theta_n] = \sqrt{n}\ f[o,\theta] + \sqrt{n}\left[\frac{\partial f}{\partial \theta'}(o,\theta) + \frac{\partial f}{\partial g'}(o,\theta)\, \frac{\partial g'}{\partial \theta}(\theta)\right](\hat\theta_n - \theta) + o_p(1)$$

From assumption i) we immediately have $f(o,\theta) = o$.

Moreover, i) also implies $\frac{\partial f}{\partial \theta'}(o,\theta) = o$ and therefore, we get :

$$\sqrt{n}\; f[g(\hat{\theta}_n),\hat{\theta}_n] = \frac{\partial f}{\partial g'}(o,\theta)\,\frac{\partial g'}{\partial \theta}(\theta)\;\sqrt{n}\, g(\hat{\theta}_n - \theta) + o_p(1)$$

$$= \frac{\partial f}{\partial g'}(o,\theta)\;\sqrt{n}\, g(\hat{\theta}_n) + o_p(1)$$

This implies :

$$\tilde{\xi}^{w}_{n} = g'(\hat{\theta}_n)\,\frac{\partial f'}{\partial g}(o,\theta)\left[\frac{\partial f}{\partial g'}(o,\theta)\,\frac{\partial g}{\partial \theta'}(\theta)\,Q(\theta)\,\frac{\partial g'}{\partial \theta}(\theta)\frac{\partial f'}{\partial g}(o,\theta)\right]^{-}$$

$$\frac{\partial f}{\partial g'}(o,\theta)\; g(\hat{\theta}_n) + o_p(1)$$

The result follows from ii) and from the well-known property (see Rao-Mitra [12] lemma 2.2.5c) : if rank (AB) = rank B, then $A'[AB\,C\,B'A']^{-}A$ is a generalized inverse of $B\,C\,B'$.

□

It is clear that if o is replaced by $g(\hat{\theta}_n)$ in $\frac{\partial f}{\partial g'}$ or, if $\hat{\theta}_n$ is replaced by a consistent estimator of θ in the derivatives or in $Q(\hat{\theta}_n)$, we still obtain an equivalent statistic.

Finally the results shown above can be extended to the case of a sequence of local alternatives. More precisely, let us assume that the parameter value is $\theta_n = \theta + \frac{\delta}{\sqrt{n}}$, where $\theta \in H_o$ and δ is a fixed vector, and that $\sqrt{n}\,(\hat{\theta}_n - \theta_n)$ converges in distribution to $N(o,Q(\theta))$; then the asymptotic distribution of ξ^{w}_{n}, or $\tilde{\xi}^{w}_{n}$ is a non-central chi-square with r degrees of freedom and whose non-centrality parameter is :

$$(2.4)\quad v = \delta'\,\frac{\partial g'}{\partial \theta}(\theta)\left[\frac{\partial g}{\partial \theta'}(\theta)\,Q(\theta)\,\frac{\partial g'}{\partial \theta}(\theta)\right]^{-}\frac{\partial g}{\partial \theta'}(\theta)\,\delta\ .$$

II.1.b - Definition of the null hypothesis.

Let us denote by $h_o(\theta)$, $h_1(\theta),\dots,h_K(\theta)$, K+1 p-dimensional continuously differentiable functions of θ . We assume :

(A1) for any θ there exists at most one vector $\lambda' = (\lambda_1,\dots,\lambda_K)$ such that :

$$h_o(\theta) = \lambda_1 h_1(\theta) + \dots + \lambda_K h_K(\theta) \quad , \tag{2.5}$$

or in matrix notation :

$$h_o(\theta) = H(\theta)\lambda \quad .$$

The null hypothesis H_o considered in this paper is the set of the θ's for which there exists exactly one λ satisfying (2.4).

Moreover, we assume (A2), for any $\theta \in H_o$, the constraints are assumed not to be "redondant" i.e. :

$\dfrac{\partial h_o}{\partial \theta'} - \sum_{i=1}^{K} \lambda_i \dfrac{\partial h_i}{\partial \theta'}$ is assumed to be of rank p , which implies $p \leqslant q$.

This null hypothesis will be called an unknown linear restriction on the K+1 p-dimensional functions $h_o,\dots,h_K$.

Note that, if θ belongs to H_o , the rank of $(p \times K)$ matrix $H(\theta)$ is necessarily K , which implies $K \leqslant p$, and finally, $K \leqslant p \leqslant q$. It is useful to note that H_o can be expressed in several ways by constraints in the implicit form $g(\theta) = o$ and to show that these various forms of H_o provide asymptotically equivalent Wald type statistics.

Let us denote by $M[\Omega(\theta),H(\theta)]$, or simply M_θ , the orthogonal projector on the space spanned by the columns of $H(\theta)$, the orthogonality being the one associated with the symmetric positive matrix $\Omega^{-1}(\theta)$. The null hypothesis H_o can be expressed as :

$$H_o = \left\{\theta : (I - M_\theta)\, h_o(\theta) = o\right\}$$

If $\Omega*(\theta)$ is another symmetric positive matrix and if $M^*_\theta = M[\Omega*(\theta), H(\theta)]$, H_o can also be expressed by :

$$H_o = \left\{\theta : (I - M^*_\theta)\, h_o(\theta) = o\right\}$$

We can use the results of the previous subsection to show that the Wald type statistics based on $(I - M_\theta)\, h_o(\theta) = o$ and on $(I - M^*_\theta)\, h_o(\theta) = o$ are asymptotically equivalent. In effect, denoting $(I - M_\theta)\ h_o(\theta)$ by $g(\theta)$, we have :

$$\begin{aligned}(I - M^*_\theta)h_o(\theta) &= (I - M^*_\theta)(I - M_\theta + M_\theta)h_o(\theta)\\ &= (I - M^*_\theta)g(\theta)\\ &= f[g(\theta),\theta] \quad \text{say} .\end{aligned}$$

It is clear that $f(o,\theta) = o$, $\forall\, \theta$.

Moreover $\frac{\partial f}{\partial g'}(o,\theta) = I - M^*_\theta$, whose kernel is the space spanned by the columns of $H(\theta)$ and it can be shown (see appendix 1) that, for any $\theta \in H_o$,

$\frac{\partial g}{\partial \theta'} = (I - M_\theta)\left[\frac{\partial h_o}{\partial \theta'} - \sum_{i=1}^{K} \lambda_i \frac{\partial h_i}{\partial \theta'}\right]$; this implies that the space spanned by the columns of $\frac{\partial g}{\partial \theta'}$ has a zero intersection with the kernel of $\frac{\partial f}{\partial g'}$; in other words, for any $\theta \in H_o$,

$$\text{rank}\left[\frac{\partial f}{\partial g'}(o,\theta)\ \frac{\partial g}{\partial \theta'}(\theta)\right] = \text{rank}\left[\frac{\partial g}{\partial \theta'}(\theta)\right] = p - K .$$

Applying property 2.3, it is clear that the Wald statistics based on the two forms of H_o : $(I - M_\theta)h_o(\theta) = o$ and $(I - M_\theta^*)h_o(\theta) = o$, are asymptotically equivalent under H_o and under a sequence of local alternatives.

Another characterization of H_o is the following.

Let us assume that there exist $p - K$ p-dimensional vectors, denoted by $h_{k+1}(\theta),\dots,h_p(\theta)$, such that the $(p \times p)$ matrix $[h_1(\theta),\dots,h_k(\theta), h_{k+1}(\theta),\dots,h_p(\theta)] = [H(\theta),H_*(\theta)]$, is invertible.

In this case, there is a unique decomposition of $h_o(\theta)$ on the columns of $H(\theta)$ and $H_*(\theta)$ given by :

$$h_o(\theta) = H(\theta)\,\lambda + H_*(\theta)\,\mu \quad .$$

Moreover, θ belongs to H_o if, and only if, $\mu = o$.

Denoting by M_θ the projector whose range is the space spanned by the columns of $H(\theta)$ and whose kernel is the space spanned by the columns of $H_*(\theta)$, we have :

$$\begin{aligned} H_*(\theta)\,\mu &= (I - M_\theta)\,h_o(\theta) \\ &= g(\theta) \quad \text{say} \quad , \end{aligned}$$

and μ can be written, for instance :

$$\begin{aligned} \mu &= [H_*'(\theta)\,\Omega^{-1}(\theta)\,H_*(\theta)]^{-1}\,H_*'(\theta)\,\Omega^{-1}(\theta)\,g(\theta) \\ &= f[g(\theta),\,\theta] \quad \text{say} \quad , \end{aligned}$$

where $\Omega(\theta)$ is the matrix appearing in

$$M_\theta = H(\theta)\,[H(\theta)'\,\Omega^{-1}(\theta)\,H(\theta)]^{-1}\,H'(\theta)\,\Omega^{-1}(\theta)$$

It is clear that $f[o,\theta] = o$, $\forall\, \theta$ and that $\frac{\partial f}{\partial g'}(o,\theta)$ is equal to $[H'_*(\theta)\, \Omega^{-1}(\theta)\, H_*(\theta)]^{-1}\, H'_*(\theta)\, \Omega^{-1}(\theta)$ whose kernel is the space spanned by the columns of $H(\theta)$; the intersection of this space with the space spanned by the columns of $\frac{\partial g}{\partial \theta'}$ is zero for any $\theta \in H_o$, for the same reason as above. Therefore a Wald-type statistics based on μ is asymptotically equivalent to that based on $g(\theta) = (I - M_\theta)\, h_o(\theta)$ where, from the previous equivalence, M_θ may be any projector on the space spanned by the columns of $H(\theta)$.

II.2 - Computation of the test-statistics.

In the previous section we have seen that asymptotically equivalent Wald-type tests can be based on :

$$(I - M_\theta)\, h_o(\theta) = o \quad ,$$

where $M_\theta = H(\theta)\, [H(\theta)'\, \Omega^{-1}(\theta)\, H(\theta)]^{-1}\, H'(\theta)\, \Omega^{-1}(\theta)$

and $\Omega(\theta)$ is any symmetric positive matrix.

Since the choice of $\Omega(\theta)$ is free, we shall choose $\Omega(\theta)$ in order to make the computation of the test statistic as simple as possible. Let us first give useful asymptotic results.

THEOREM 2.6 :

> Denoting respectively by $\hat{M}$, $\hat{h}_o$, $\hat{H}$ the matrices M_θ , $h_o(\theta)$, $H(\theta)$ in which θ is replaced by $\hat{\theta}_n$, $\sqrt{n}\,(I - \hat{M})\, \hat{h}_o$ is asymptotically equivalent to : $\sqrt{n}\,(I - M_\theta)\,(\hat{h}_o - \hat{H}\,\lambda)$, where λ is defined by $h_o(\theta) = H(\theta)\,\lambda$.

Proof : see appendix 2.

It is clear that, under H_o , $\sqrt{n}\,(\hat{h}_o - \hat{A}\,\lambda)$ has a normal asymptotic distribution whose mean is zero and whose covariance matrix is

$$\Sigma(\theta) = \left[\frac{\partial h_o}{\partial \theta'} - \sum_{i=1}^{K} \frac{\partial h_i}{\partial \theta'} \lambda_i \right] Q(\theta) \left[\frac{\partial h'_o}{\partial \theta} - \sum_{i=1}^{K} \frac{\partial h'_i}{\partial \theta} \lambda_i \right]$$

From assumption (A2) , the rank of $\Sigma(\theta)$ is equal to p for any $\theta \in H_o$.

Therefore we deduce immediately the following corollary :

Corollary 2.7 :

> Under H_o , the asymptotic distribution of $\sqrt{n}\,(I - \hat{M})\,\hat{h}_o$ is $N[o,(I - M_\theta)\ \Sigma(\theta)\ (I - M'_\theta)]$, and the rank of $(I - M_\theta)\ \Sigma(\theta)\ (I - M'_\theta)$ is $p - K$.

The Wald-type statistic associated with $(I - M_\theta)\ h_o(\theta) = o$ can be written.

$$\xi_n^W(\Omega) = n\ \hat{h}'_o(I - \hat{M}')\ [(I - \hat{M})\ \hat{\Sigma}(I - \hat{M}')]^-\ (I - \hat{M})\ \hat{h}_o$$

where $\hat{\Sigma}$ is obtained from $\Sigma(\theta)$ by replacing θ by $\hat{\theta}_n$ and λ by $\hat{\lambda}_n$; $\hat{\lambda}_n$ is any consistent estimator of λ , for instance the O.L.S. estimator obtained by a regression of h_o on the columns of $\hat{A}$.

From the previous results we deduce :

Corollary 2.8 :

> Under H_o , the asymptotic distribution of $\xi_n^W(\Omega)$ is a chi-square distribution with $p - K$ degrees of freedom.

$\xi_n^W(\Omega)$ is particularly simple if $\Omega(\theta)$ is equal to $\Sigma(\theta)$ because in this case, it is easily seen that a generalized inverse of $(I - M)\ \Sigma\ (I - M')$ is simply Σ^{-1} and that $\xi_n^W(\Sigma)$, denoted by ξ_n^W, becomes :

$$\xi_n^W = n\ \hat{h}_o'(I - \hat{M}')\ \hat{\Sigma}^{-1}\ (I - \hat{M})\ \hat{h}_o$$

$$= n\ [\hat{h}_o'\ \hat{\Sigma}^{-1}\ \hat{h}_o - \hat{h}_o'\ \hat{M}'\ \hat{\Sigma}^{-1}\ \hat{M}\ \hat{h}_o]\ .$$

The computation of ξ_n^W is then straightforward :

i) first compute $\tilde{\tilde{\lambda}}$, by the O.L.S. method based on $\hat{h}_o$ and $\hat{A}$, and deduce $\hat{\Sigma}$

ii) compute n times the "sum of squared residuals" of the G.L.S. method based on $\hat{h}_o$, $\hat{A}$ and the matrix $\hat{\Sigma}$.

II.3 - Properties of the test procedures.

Up to now, the only assumption which has been made about $\hat{\theta}_n$ is the asymptotic normality assumption (2.1).

If, in addition, we assume that $\hat{\theta}_n$ is a maximum likelihood estimator, the test obtained in the previous subsection is a genuine Wald test. Therefore the test procedure proposed is asymptotically equivalent, under H_o and under a sequence of local alternatives, to the Lagrange multiplier test or to the likelihood ratio test. This also implies that the test procedure proposed has the optimality properties of these three classical tests, in particular it is asymptotically locally more powerful than any Wald type test based on a consistent asymptotically normal (C.A.N.) estimator of θ . In effect, if $I^{-1}(\theta)$ and $Q(\theta)$ are the asymptotic covariance matrices of the maximum likelihood estimator and of the C.A.N. estimator respectively, the associated Wald test statistics are asymptotically distributed, under a sequence of local alternatives $\theta + \frac{\delta}{\sqrt{n}}$, $\theta \in H_o$, as a chi-square with $p - K$ degrees of freedom and with non-centrality parameters (see 2.4) :

$$v^* = \delta' \frac{\partial g'}{\partial \theta}(\theta) \left[\frac{\partial g}{\partial \theta'}(\theta)\ I^{-1}(\theta)\ \frac{\partial g'}{\partial \theta}(\theta) \right]^{-} \frac{\partial g}{\partial \theta'}(\theta)\ \delta$$

$$v = \delta' \frac{\partial g'}{\partial \theta}(\theta) \left[\frac{\partial g}{\partial \theta'}(\theta)\ \ Q(\theta)\ \frac{\partial g'}{\partial \theta}(\theta) \right]^{-} \frac{\partial g}{\partial \theta'}(\theta)\ \delta \quad \text{respectively} ,$$

where $g(\theta) = o$ is any implicit form of H_o .
Since $I^{-1}(\theta) << Q(\theta)$, we have $v^* \geqslant v$ and, therefore, the asymptotic power of first test is greater than that of the second one.

Another property will be used in the applications considered in the following sections.
Let us come back within the general framework of the previous subsections, but let us consider a Wald type statistic only based on $\tilde{\tilde{p}}\ (< p)$ equations of (2.4) ; in other words, we consider a Wald-type test statistic based on the hypothesis $\tilde{\tilde{H}}_o$, included in H_o , defined by

$$\tilde{\tilde{H}}_o = \left\{ \theta \ / \ \exists\ \lambda \ \text{ unique : } \ \tilde{\tilde{h}}_o(\theta) = \sum_{i=1}^{K} \lambda_i\ \tilde{\tilde{h}}_i(\theta) \right\}$$

where $\tilde{\tilde{h}}_i$, $i = o,\ldots,K$, is obtained from h_i by suppressing $p - \tilde{\tilde{p}}$ terms (for instance the last $p - \tilde{\tilde{p}}$ ones). We necessarily have $\tilde{\tilde{p}} \geqslant K$, and

$\dfrac{\partial \tilde{\tilde{h}}_o}{\partial \theta'} - \sum_{i=1}^{K} \lambda_i \dfrac{\partial \tilde{\tilde{h}}_i}{\partial \theta'}$ is of rank $\tilde{\tilde{p}}$. This null hypothesis $\tilde{\tilde{H}}_o$ means that $\tilde{\tilde{h}}_o(\theta)$ belongs to the space spanned by the columns of $\tilde{\tilde{H}}(\theta) = [\tilde{\tilde{h}}_1(\theta),\ldots,\tilde{\tilde{h}}_K(\theta)]$.

<u>PROPERTY 2.9</u> :

> When testing H_o , the Wald type procedures based on $\tilde{\tilde{H}}_o$ are asymptotically locally less powerful than those based on H_o .

<u>Proof</u> : An implicit form of H_o can be obtained by imposing that $\tilde{\tilde{p}} - K$ determinants [of size $(K+1)\times(K+1)$] are equal to zero, that is

$$(2.10) \qquad \tilde{\tilde{g}}(\theta) = o \quad ,$$

where $\tilde{\tilde{g}}$ is $\tilde{\tilde{p}} - K$ dimensional vector, such that $\dfrac{\partial \tilde{\tilde{g}}}{\partial \theta'}$ is of rank $\tilde{\tilde{p}} - K$ for $\theta \in H_o$.

Similarly, an implicit form of H_o can be obtained by adding to (2.10) $p - \tilde{\tilde{p}}$ conditions imposing the nullity of $(K+1)\times(K+1)$ dimensional determinants ; in other words H_o can be written as :

$$(2.11) \qquad g(\theta) = o \quad ,$$

where g is a $p - K$ dimensional vector, whose $\tilde{\tilde{p}} - K$ (first) components are $\tilde{\tilde{g}}$, and such that $\dfrac{\partial g}{\partial \theta'}$ is of rank $p - K$ for $\theta \in H_o$.

The general results of the previous sections show that the test statistics $\tilde{\tilde{\xi}}_n^W$, based on $\tilde{\tilde{H}}_o$ are asymptotically distributed, under H_o , as $\chi^2(\tilde{\tilde{p}} - K)$ and that, under a sequence of local alternatives $\theta + \dfrac{\delta}{\sqrt{n}}$, $\theta \in H_o$, they are asymptotically distributed as a chi-square with $\tilde{\tilde{p}} - K$ degrees of freedom and with non centrality parameter :

$$\tilde{\tilde{v}} = \delta' \frac{\partial \tilde{\tilde{g}}'}{\partial \theta}(\theta) \left[\frac{\partial \tilde{\tilde{g}}}{\partial \theta'}(\theta) \; Q(\theta) \frac{\partial \tilde{\tilde{g}}'}{\partial \theta}(\theta) \right]^{-1} \frac{\partial \tilde{\tilde{g}}}{\partial \theta'}(\theta) \, \delta$$

Since $\tilde{\tilde{g}}$ is a subvector of g , it is easily seen that $\tilde{\tilde{v}}$ is smaller than the non-centrality parameter associated with $\tilde{\tilde{\xi}}_n^W$ i.e. :

$$v = \delta' \frac{\partial g'}{\partial \theta}(\theta) \left[\frac{\partial g}{\partial \theta'}(\theta) \; Q(\theta) \frac{\partial g'}{\partial \theta}(\theta) \right]^{-} \frac{\partial g}{\partial \theta'}(\theta) \, \delta$$

Finally, since the asymptotic distribution of $\tilde{\tilde{\xi}}_n^W$, under a sequence of local alternatives, has less degrees of freedom and a smaller non centrality parameter than that of ξ_n^W , the test procedure based on $\tilde{\tilde{\xi}}_n^W$ is asymptotically locally less powerful than that based on ξ_n^W .

$\square$

III - TEST OF THE RATIONAL EXPECTATION HYPOTHESIS.

In macromodels the dynamics is often introduced by means of expectations of some economic variables. Faced with unobserved expectations, economists have to specify the way expectations are formed. In this section we consider a unidimensional dynamic linear model with current expectation of the endogenous variable and, following the same approach as in Wallis [16], Pesaran [11], we derive the constraints on the reduced parameters implied by some general methods of modelling expectations, essentially perfect foresight method and rational method. Since these constraints take the form of some unknown linear restrictions between functions of the reduced parameters, the test procedure developed in section II can be directly applied for determining the kind of expectation scheme.

III.1 - Derivation of the reduced forms.

III.1.a - The structural form.

The model with current expectation of the endogenous variable can be written as :

(3.1) $y_t = a \; {}_{t-1}\hat{y}_t + b' x_t + u_t \quad , \quad t = 1,\dots,T \quad , \quad a \neq o \quad ,$

where ${}_{t-1}\hat{y}_t$ stands for the expectation of y_t held at time $t-1$, $x_t = (x_{1t},\dots,x_{Kt})'$ for the values of K exogenous variables and u_t for the disturbance with variance σ^2 . We assume that the expectation does have an influence on y_t , i.e. $a \neq o$. Moreover the error process is assumed to be an (independent) white noise and to be independent from present and past values of the exogenous.

Since the predictions ${}_{t-1}\hat{y}_t$ are in general unobservable the structural form has to be completed by an expectation scheme.

III.1.b - Reduced form associated with rational expectations.

Under the rational expectation hypothesis ${}_{t-1}\hat{y}_t$ is formed as the expectation of y_t conditional on the available information at time $t-1$:

(3.2) ${}_{t-1}\hat{y}_t = E(y_t / I_{t-1})$

with : $I_{t-1} = \left\{ y_{t-1}, x_{t-1}, y_{t-2}, x_{t-2}, \dots \right\}$

Replacing in the structural form, we obtain :

$$y_t = a\,E(y_t/I_{t-1}) + b'\,x_t + u_t$$

and taking expectations of both sides of this equation conditional on I_{t-1} we derive the expression of y_t in terms of the exogenous and of their predictions :

$$(3.3)\quad y_t = \frac{ab'}{1-a}\,E\,(x_t/I_{t-1}) + b'\,x_t + u_t$$

Therefore we have passed the problem of expectation formation of the endogenous variable on to the expectation formation of the exogenous variables and, to continue the derivation of the reduced form, we have to precise the expression of $E\,(x_t/I_{t-1})$. We assume that this prediction is linear in some observable variables z, which may contain lagged values of x :

$$(3.4)\quad E\,(x_t/I_{t-1}) = C\,z_t\,, \quad \text{(where } C \text{ is a } (K,L) \text{ matrix)}$$

and that the prediction error :

$$(3.5)\quad v_t = x_t - E\,(x_t/I_{t-1})\,,$$

is an independent white noise with a covariance matrix denoted by Σ. Since v is a function of past values of y and of present and past values of x it is easily seen that v_t and u_t are uncorrelated.

Taking into account (3.4) and (3.5), we obtain the reduced form :

$$(3.6)\quad \begin{cases} y_t = \dfrac{b'C}{1-a} z_t + u_t + b' v_t \\ \\ x_t = C z_t + v_t \end{cases}$$

with $\quad V \begin{pmatrix} u_t \\ v_t \end{pmatrix} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \Sigma \end{pmatrix}$

III.1.c - Reduced form associated with perfect foresight.

In the case of perfect foresight :

$${}_{t-1}\hat{y}_t = y_t \quad ,$$

we directly obtain by replacing in the structural form : $y_t = \dfrac{b'}{1-a} x_t + \dfrac{u_t}{1-a}$.
Taking into account the evolution of the exogenous process the reduced form is :

$$(3.7)\quad \begin{cases} y_t = \dfrac{b'C}{1-a} z_t + \dfrac{u_t}{1-a} + \dfrac{b'}{1-a} v_t \\ \\ x_t = C z_t + v_t \end{cases}$$

with $\quad V \begin{pmatrix} u_t \\ v_t \end{pmatrix} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \Sigma \end{pmatrix}$

III.2 - Constraints on the reduced form parameters.

The two previous reduced forms are naturally compared with the unconstrained reduced form :

$$(3.8)\begin{cases} y_t = \Pi_1 z_t + \omega_{1t} = \Pi_{11} z_{1t} + \dots + \Pi_{1L} z_{Lt} + \omega_{1t} \\ x = \Pi_2 z_t + \omega_{2t} = \Pi_{21} z_{1t} + \dots + \Pi_{2L} z_{Lt} + \omega_{2t} \end{cases},$$

where Π_1 and Π_2 satisfy no a priori constraints and where the elements of the covariance matrix $V\begin{pmatrix}\omega_{1t}\\ \omega_{2t}\end{pmatrix} = \begin{pmatrix}\Omega_{11} & \Omega_{12}\\ \Omega_{21} & \Omega_{22}\end{pmatrix}$ are only submitted to the usual conditions of symmetry and non negativity.

When $z'_t = (x'_{t-1},\dots,x'_{t-Q} ; y'_{t-1},\dots,y'_{t-Q})$ this model is simply the usual unconstrained autoregressive time series model.

III.2.a - Case of rational expectations

The reduced parameters Π_1 , Π_2 , Ω_{11} , Ω_{22} , Ω_{12} are related to the structural parameters a , b , C , σ^2 , Σ by :

$$(3.9)\quad \begin{cases} \Pi_2 = C \ , \quad \Pi_1 = \dfrac{b'C}{1 - a} \\ \Omega_{11} = \sigma^2 + b'\,\Sigma\, b \ , \quad \Omega_{12} = b'\,\Sigma \ , \quad \Omega_{22} = \Sigma \end{cases}$$

The parameters a , b , C are first order unidentifiable.

However all the structural parameters are second order identifiable.

The unconstrained reduced form depends on : $L + LK + \frac{K(K+1)}{2} + 1 + K$ independent parameters and the R.E. constrained reduced form depends on $K + LK + 1 + \frac{K(K+1)}{2} + 1$ independent parameters.

Therefore the R.E. hypothesis can be expressed by imposing $L-1$ independent constraints on the reduced parameters. It may also be written as unknown linear restrictions. Two formulations are possible :

$$(3.10)\ H_o = \left\{ \exists\, \mu \in R - \langle 1 \rangle \ : \ \Omega_{12}\, \Omega_{22}^{-1}\, \Pi_2 = \mu\, \Pi_1 \right\}$$

or

$$(3.11)\ H_o = \left\{ \exists\, \lambda_1, \ldots, \lambda_K \in R \ , \ \mu \in R - \langle 1 \rangle \ : \right.$$

$$\left. \Omega_{12} = (\lambda_1, \ldots, \lambda_K)\, \Omega_{22} \ \text{ and } \ (\lambda_1, \ldots, \lambda_K)\, \Pi_2 - \mu\, \Pi_1 = o \right\}$$

(The condition $\mu \in R - \langle 1 \rangle$ is equivalent to $a \neq o$)

III.2.b - Case of perfect foresight.

The relations between reduced and structural parameters are :

$$(3.12) \quad \begin{cases} \Pi_1 = \dfrac{b'C}{1 - a} \ , \quad \Pi_2 = C \\[2ex] \Omega_{11} = \dfrac{\sigma^2}{(1-a)^2} + \dfrac{b'}{1 - a}\, \Sigma\, \dfrac{b}{1 - a} \ , \quad \Omega_{12} = \dfrac{b'}{1 - a}\, \Sigma \ , \quad \Omega_{22} = \Sigma \end{cases}$$

The perfect foresight model is first order observationally equivalent to the R.E. model and in particular a , b , c are first order unidentifiable. Then in order to distinguish between the two expectation schemes it is necessary to carefully examine the second order properties.

In the case of perfect foresight the structural parameters are also second order unidentifiable. The identifiable parameters are simply the functions of C , $\dfrac{b}{1 - a}$, $\dfrac{\sigma^2}{(1-a)^2}$ and Σ .

Therefore it will be necessary for estimation to impose an identification constraint, for instance to fix the structural parameter a to a given value a_o . The appropriate choice of a_o is discussed in the next subsection.

III.2.c - Separation of the two hypothesis

Before considering the problem of test, we have to examine if the two previous hypotheses have an intersection. For this purpose, we have to look for solutions of the system :

$$\begin{cases} C = \bar{C} \\ \dfrac{b'C}{1-a} = \dfrac{\bar{b}'\bar{C}}{1-\bar{a}} \\ \sigma^2 + b'\Sigma b = \dfrac{\bar{\sigma}^2}{(1-\bar{a})^2} + \dfrac{\bar{b}'}{1-\bar{a}}\bar{\Sigma}\dfrac{\bar{b}}{1-\bar{a}} \\ b'\Sigma = \dfrac{\bar{b}'}{1-\bar{a}}\bar{\Sigma} \\ \Sigma = \bar{\Sigma} \end{cases}$$

where a , b , C , Σ , σ^2 [resp $\bar{a}$, $\bar{b}$, $\bar{C}$, Σ , $\bar{\sigma}^2$] are associated with R.E. (resp with perfect foresight).

This system implies : $\dfrac{b'}{1-a} = \dfrac{\bar{b}'}{1-\bar{a}}$ and $b' = \dfrac{\bar{b}'}{1-\bar{a}}$ and therefore $a = o$, which is an impossible value for the structural parameter a .

PROPERTY 3.14 :

The R.E. model and the perfect foresight model are disjoint.

Moreover if parameter a were equal to zero, we would have $\bar{a} = a = o$, $\bar{b} = b$, $\bar{\Sigma} = \Sigma$, $\bar{C} = C$, $\bar{\sigma}^2 = \sigma^2$. This means that the perfect foresight model is observationaly equivalent to a R.E. model computed with $a = o$, which is equivalent to $\mu = 1$.

In summary if the general hypothesis corresponds to the unconstrained reduced form :

$$H \; : \; \left\{ \Pi_1, \Pi_2, \Omega_{11}, \; \Omega_{22}, \; \Omega_{12} \text{ unconstrained} \right\} ,$$

the union of R.E. and perfect foresight models is obtained by imposing :

$$(3.15) \quad H^*_o = \left\{ \exists \, \lambda \in R^K , \; \mu \in R \; : \; \lambda' \Pi_2 - \mu \, \Pi_1 = o \; , \; \Omega_{12} - \lambda' \Omega_{22} = o \right\}$$

and the perfect foresight model corresponds to :

$$(3.16) \quad H^{**}_o = H^*_o - H_o \; = \left\{ \exists \, \lambda \in R^k \; : \; \lambda' \Pi_2 - \Pi_1 = o \; , \; \Omega_{12} - \lambda' \Omega_{22} = o \right\}$$

Since $H^{**}_o \subset H^*_o \subset H$, the two expectation schemes can be tested by a stepwise procedure.

III.3 - Test statistics.

It is now possible to develop some test procedures along the line of section II. We denote $\hat{\Pi}_1$, $\hat{\Pi}_2$, $\hat{\Omega}_{12}$, $\hat{\Omega}_{22}$ the usual least squares estimates of Π_1 , Π_2 , Ω_{12} , Ω_{22} computed on the unconstrained model. It is well known that $\left(\hat{\Pi}_1 , \hat{\Pi}_2\right)$ and $\left(\hat{\Omega}_{12} , \hat{\Omega}_{22}\right)$ are asymptotically uncorrelated.

III.3.a - Test of H^*_o against the general hypothesis.

We first have to derive the expression of the asymptotic covariance matrix of : $\begin{pmatrix} \hat{\Pi}'_2 \lambda & - & \hat{\Pi}'_1 \mu \\ \hat{\Omega}_{21} & - & \hat{\Omega}_{22} \mu \end{pmatrix}$. This is a bloc diagonal matrix, which will be denoted by

$$\Sigma = \begin{pmatrix} S_1 & o \\ o & S_2 \end{pmatrix}$$

. Then this matrix has to be consistently estimated under H_o^* $\left[\hat{\Sigma} = \begin{pmatrix} \hat{S}_1 & o \\ o & \hat{S}_2 \end{pmatrix} \text{ is such an estimate} \right]$ and the test statistic is simply n times the S.S.R. associated with the linear model :

$$\begin{bmatrix} o \\ \hat{\Omega}_{21} \end{bmatrix} = \begin{bmatrix} \hat{\pi}'_2 & -\hat{\pi}'_2 \\ \hat{\Omega}_{22} & o \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} + \begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \end{bmatrix} .$$

with $E \begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \end{bmatrix} = o$ and $V \begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \end{bmatrix} = \begin{pmatrix} \hat{S}_1 & o \\ o & \hat{S}_2 \end{pmatrix}$

Introducing the following notations :

$$\langle \hat{\pi}_2, \hat{\pi}_1 \rangle = \hat{\pi}_2 \hat{S}_1^{-1} \hat{\pi}'_1 \quad , \quad ||\hat{\pi}_2||^2 = \hat{\pi}_2 \hat{S}_1^{-1} \hat{\pi}'_2 \quad , \quad ||\hat{\pi}_1||^2 = \hat{\pi}_1 \hat{S}_1^{-1} \hat{\pi}'_1$$

$$\langle \hat{\pi}_1, \hat{\pi}_2 \rangle = \hat{\pi}_1 \hat{S}_1^{-1} \hat{\pi}'_2$$

and similarly :

$$||\hat{\Omega}_{22}||^2 = \hat{\Omega}_{22} \hat{S}_2^{-1} \hat{\Omega}_{22} \quad ,$$

$$||\hat{\Omega}_{21}||^2 = \hat{\Omega}_{12} \hat{S}_2^{-1} \hat{\Omega}_{21} \quad ,$$

$$\langle \hat{\Omega}_{22}, \hat{\Omega}_{21} \rangle = \hat{\Omega}_{22} \hat{S}_2^{-1} \hat{\Omega}_{21} = \langle \hat{\Omega}_{21}, \hat{\Omega}_{22} \rangle \quad ,$$

it can be seen that the test statistic is given by n SSR* with :

$$\text{(3.16)} \quad \boxed{SSR* = ||\hat{\Omega}_{21}||^2 - \langle\hat{\Omega}_{12},\hat{\Omega}_{22}\rangle \left\{||\hat{\Pi}_2||^2 - \langle\hat{\Pi}_2,\hat{\Pi}_1\rangle(||\hat{\Pi}_1||^2)^{-1}\langle\hat{\Pi}_1,\hat{\Pi}_2\rangle + ||\hat{\Omega}_{22}||^2\right\} \langle\hat{\Omega}_{22},\hat{\Omega}_{21}\rangle}$$

III.3.b - Test of H^{**}_o against the general hypothesis.

We now have to consider the asymptotic covariance matrix of $\begin{pmatrix} \hat{\Pi}'_2\lambda - \hat{\Pi}'_1 \\ \hat{\Omega}_{21} - \hat{\Omega}_{22}\lambda \end{pmatrix}$, which is directly deduced from Σ by imposing $\mu = 1$; in particular $\hat{\Sigma}$ is also a consistent estimate of this matrix under H^{**}_o.

The S.S.R. is associated with the linear model :

$$\begin{bmatrix} \hat{\Pi}'_1 \\ \hat{\Omega}_{21} \end{bmatrix} = \begin{bmatrix} \hat{\Pi}'_2 \\ \hat{\Omega}_{22} \end{bmatrix} \lambda_o + \begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \end{bmatrix}$$

with $E\begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \end{bmatrix} = o$ and $V\begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \end{bmatrix} = \begin{pmatrix} \hat{S}_1 & o \\ o & \hat{S}_2 \end{pmatrix}$

The test statistic is given by : n SSR** with :

$$\text{(3.17)} \quad \boxed{SSR** = ||\hat{\Pi}_1||^2 + ||\hat{\Omega}_{21}||^2 - [\langle\hat{\Pi}_1,\hat{\Pi}_2\rangle + \langle\hat{\Omega}_{12},\hat{\Omega}_{22}\rangle] \left\{||\hat{\Pi}_2||^2 + ||\hat{\Omega}_{22}||^2\right\}^{-1} [\langle\hat{\Pi}_2,\hat{\Pi}_1\rangle + \langle\hat{\Omega}_{22},\hat{\Omega}_{21}\rangle]}$$

III.3.c - Test of perfect foresight against rational expectation.

This is the test of H_o^{**} against $H_o^{*} - H_o^{**}$. A simple procedure is directly obtained by considering the difference between the two previous statistics :

n [SSR** - SSR*] .

IV - TESTING FOR COMMON ROOTS TO A SET OF POLYNOMIALS.

IV.1 - The "dynamic specification problem".

In two papers [13] , [14], Sargan and Sargan-Mehta have examined the possibility of factoring out an autoregressive error specification from a general lag structure. This specification problem can be easily presented in a single equation context. Let us consider the dynamic linear equation :

(4.1) $\alpha'(L)\, x_t \;=\; u_t \qquad t = 1,\ldots,T$,

where x_t is an $n \times 1$ vector of variables in period t and $\alpha(L)$ is an $n \times 1$ vector of polynomials in the lag operator L . If the error process u has an autoregressive representation :

(4.2) $\varrho(L)\, u_t \;=\; e_t$.

where $\varrho(L)$ is a scalar polynomial of degree r and e is a white noise, we obtain replacing in (4.1) :

(4.3) $\Psi'(L)\, x_t \;=\; \varrho(L)\, \alpha'(L)\, x_t \;=\; e_t$

If we now consider the unconstrained autoregressive representation : $\Psi'(L)\, x_t \;=\; e_t$, it is interesting to develop some test procedures to determine, if it is possible to factorize $\psi'(L)$ into $\varrho(L)\, \alpha'(L)$. This is equivalent to test the existence of a common factor $\varrho(L)$ of degree r to the set of polynomials $\Psi_j(L)$, $j = 1,\ldots,n$, where $\Psi_j(L)$ denotes the j^{th}

component of $\Psi(L)$.

Such a test can be performed along three different lines depending on the form under which the null hypothesis is expressed. To illustrate this point let us consider the simplest case of two polynomials of degree one and of the hypothesis H_o of one common root. These polynomials are :

$$\Psi_1(L) = \Psi_{10} + \Psi_{11}L \quad , \quad \Psi_2(L) = \Psi_{20} + \Psi_{21}L$$

i) The null hypothesis can be written under a parametric form. Under H_o , there exist some parameters γ_o , γ_1 , γ_2 such that : $\Psi_1(L) = \gamma_1 (\gamma_o + L)$, $\Psi_2(L) = \gamma_2 (\gamma_o + L)$. Therefore we have :

$$(4.4) \quad H_o \; : \left\{ \exists \, \gamma_o , \gamma_1 , \gamma_2 \; : \; \Psi_{10} = \gamma_o \gamma_1 \; , \; \Psi_{11} = \gamma_1 \; , \; \Psi_{20} = \gamma_2 \gamma_0 \; , \; \Psi_{21} = \gamma_2 \right\}$$

If $\hat{\Psi} = (\hat{\Psi}_{10}, \hat{\Psi}_{11}, \hat{\Psi}_{20}, \hat{\Psi}_{21})'$ is a consistent asymptotically normal estimator of Ψ , with an asymptotic covariance matrix estimated by $\hat{Q}$, we can determine consistent estimates $\hat{\gamma}_o$, $\hat{\gamma}_1$, $\hat{\gamma}_2$ of γ_o , γ_1 , γ_2 under H_o by solving the optimization problem :

$$\underset{\gamma_o,\gamma_1,\gamma_2}{\text{Min}} \quad S_T(\gamma_o,\gamma_1,\gamma_2) = (\hat{\Psi}_{10} - \gamma_o \gamma_1 , \hat{\Psi}_{11} - \gamma_1 , \hat{\Psi}_{20} - \gamma_2 \gamma_0 , \hat{\Psi}_{21} - \gamma_2)$$

$$\hat{Q}^{-1} \begin{bmatrix} \hat{\Psi}_{10} - \gamma_o \gamma_1 \\ \hat{\Psi}_{11} - \gamma_1 \\ \hat{\Psi}_{20} - \gamma_2 \gamma_0 \\ \hat{\Psi}_{21} - \gamma_2 \end{bmatrix}$$

and then we can base the test on the statistic : $T\, S_T(\hat{\gamma}_o, \hat{\gamma}_1, \hat{\gamma}_2)$.

ii) Another possibility is to express the null hypothesis under an implicit form :

$$(4.5) \quad H_o : \left\{ \det \begin{bmatrix} \psi_{10} & \psi_{20} \\ \psi_{11} & \psi_{21} \end{bmatrix} = \psi_{10}\psi_{21} - \psi_{20}\psi_{11} = o \right\}$$

and to use Wald criterium, i.e to compare

$$\xi_T^W = T \frac{\left(\hat{\psi}_{10}\hat{\psi}_{21} - \hat{\psi}_{20}\hat{\psi}_{11}\right)^2}{\hat{V}[\hat{\psi}_{10}\hat{\psi}_{21} - \hat{\psi}_{20}\hat{\psi}_{11}]}$$

to the critical level $\chi^2_{95\%}(1)$.

iii) Finally the hypothesis of common root can also be written as an unknown linear restriction :

$$H_o = \left\{ \exists\, \lambda : \begin{bmatrix} \psi_{20} \\ \psi_{21} \end{bmatrix} = \lambda \begin{bmatrix} \psi_{10} \\ \psi_{11} \end{bmatrix} \right\} ,$$

and be tested by the procedure introduced in section II.

It is easily seen that these three procedures are asymptotically equivalent under the null.

In the general context of several polynominals ψ_j of any degrees and of the hypothesis of a common factor of degree r , the approach based on the parametric form of H_o leads to the application of complicated nonlinear least squares procedures.

The implicit formulation has been followed by Sargan [13] . The required constraints can be expressed by stating that a certain matrix has a given rank or equivalently that a set of determinants are all zero. The test procedure obtained is rather complex and in particular it is necessary to use numerical approximations of Wald statistic [see Sargan [13] for the description of an algorithmic approach].

As seen below the interpretation of the constraints in term of unknown linear restrictions leads to the application of linear least squares method and therefore has the advantage of simplicity.

IV.2 - A necessary and sufficient condition for a common factor of degree r.

PROPERTY 4.6 :

Let $\Psi_j(L) = \sum_{k=o}^{f_j} \Psi_{jk} L^k$, $j = 1,\dots,n$ be n polynomials of degree f_j , $j = 1,\dots,n$ ($\Psi_{j,f_j} \neq o \;\; \forall j$), normalized by $\Psi_{10} = 1$. Let us assume that these polynomials may have at most r common roots.

A necessary and sufficient condition for these polynomials to possess as common factor a polynomial $\rho(L) = \sum_{k=o}^{r} \rho_k L^k$ of degree r : $\rho_r \neq o$ is the existence of n polynomials

$\beta_j(L) = \sum_{k=o}^{f_j-r} \beta_{jk} L^k$ $j = 1,\dots,n$, of degree $f_j - r$ ($\beta_{j,f_j-r} \neq o$) normalized by $\beta_{10} = 1$ and satisfying :

$$\Psi_j(L)\,\beta_{j+1}(L) = \Psi_{j+1}(L)\,\beta_j(L) \qquad j = 1,\dots,n-1$$

Moreover these polynomials $\beta_j(L)$, $j = 1,\dots,n$ are uniquely defined.

The proof of this property is given in Appendix 3 . The result is a direct generalization of Bezout's identity.

The normalization condition $\Psi_{10} = 1$ is satisfied in practice since equation (4.3) gives the value of the explained variable, for instance x_{1t} , in term of the present values of the other variables and of lagged values of all variables.

If the hypothesis of a common factor of degree r is satisfied, each polynomial $\Psi_j(L)$ can be decomposed in :

$$\Psi_j(L) = \varrho(L)\, \alpha_j(L) ,$$

where $\alpha_j(L) = \sum_{k=o}^{f_j - r} \alpha_{jk} L^k$ is of degree $f_j - r$. Since $\Psi_{10} = 1$, it is possible to retain the decomposition such that $\alpha_{10} = 1$. Then from the uniqueness of $\beta_j(L)$, $j = 1,\dots,n$ in property 4.6, it is directly seen that $\beta_j(L)$ is necessarily equal to $\alpha_j(L)$. In particular the coefficients β_{jk} are real even if the common factor $\varrho(L)$ has complex roots.

IV.3 - The test procedure.

We consider the sets of hypotheses :

$$H_{or} = \left\{ \text{the polynomials } \Psi_j(L),\ j = 1,\dots,n, \text{ have exactly } r \text{ common roots} \right\}$$

$$H_r = \left\{ \text{the polynomials } \Psi_j(L),\ j = 1,\dots,n, \text{ have at most } r \text{ common roots} \right\}$$

for $r = o,\dots,f = \underset{j=1,\dots,n}{\text{Min}} f_j$. These hypotheses satisfy $H_r = \bigcup_{\tilde{r}=o}^{r} H_{o\tilde{r}}$, $H_o = H_{oo}$.

The proposed test procedure is a sequential one. We first test H_{of} against $H_f - H_{of}$: if H_{of} is accepted the common factor $\varrho(L)$ is considered to have degree f ; otherwise we can consider that H_{f-1} is satisfied. Then at the second step it is possible to test $H_{o,f-1}$ against $H_{f-1} - H_{o,f-1}$ and so on...until one hypothesis H_{or} is accepted.

It is necessary at each step to change the general hypothesis H_r . In effect without the constraint imposed by H_r , the polynomials $\Psi_j(L)$, $j = 1,\dots,n$ might have more than r common roots and in this case the common factor $\varrho(L)$ of degree r would not be uniquely defined which would imply an identification problem.

At each step, H_{or} will be tested against $H_r - H_{or}$ with a first kind error equal for instance to 5 % . It is clear that the probability of error associated with the whole sequential procedure is difficult to derive and this problem will not be studied in this paper.

When H_r is assumed to be satisfied, the null hypothesis H_{or} can be written under the form of unknown linear restrictions. In effect the necessary and sufficient condition of property 4.6 is equivalent to :

$\exists\,!\ \beta_{jk}$ $k = o,\dots,f_j$, $j = 1,\dots,n$, with $\beta_{10} = 1$, $\forall\, j = 1,\dots,n-1$:

$$\sum_{k=o}^{f_j} \Psi_{jk}\, L^k \sum_{l=o}^{f_{j+1}-r} \beta_{j+1,l}\, L^l = \sum_{k=o}^{f_{j+1}} \Psi_{j+1,k}\, L^K \sum_{l=o}^{f_j-r} \beta_{j,l}\, L^l$$

$\Longleftrightarrow$ $\exists\,!\ \beta_{jk}$ $k = o,\dots,f_j$, $j = 1,\dots,n$, with $\beta_{10} = 1$

$\forall\, j = 1,\dots,n-1$, $\forall\, l = o,\dots,f_j + f_{j+1} - r$:

$$(4.7)\quad \sum_{k=\mathrm{Max}(o,1-f_{j+1}+r)}^{\mathrm{Min}(1,f_j)} \Psi_{jk}\,\beta_{j+1,1-k} = \sum_{k=\mathrm{Max}(o,1-f_{j+1}+r)}^{\mathrm{Min}(1,f_j)} \Psi_{j+1,k}\,\beta_{j,1-k}$$

These constraints are linear in the auxiliary parameters β_{jk} , with constant terms which are introduced through the normalization condition $\beta_{10} = 1$. Therefore the test can be performed by directly applying the general results of section II.

The test statistic has asymptotically a chi-square distribution ; the degree of freedom v is equal to the difference between the number of independent parameters under the general hypothesis and under the null :

$$\left\{ \sum_{j=1}^{n} (f_j+1) - 1 \right\} - \left\{ \sum_{j=1}^{n} (f_j-r+1) + r - 1 \right\}$$

[f_j-r+1 for $\beta_j(L)$, r for $\varrho(L)$, -1 for the condition $\beta_{10} = 1$] .

This gives :

$$(4.8)\qquad \boxed{v = r\,(n-1)}$$

Let us remark that in system (4.7) the number of equations is equal to $\sum_{i=1}^{n-1} [f_i + f_{i+1} - r + 1]$ and the number of independent parameters β_{jk} is $\sum_{i=1}^{n} (f_j-r+1) - 1$. The difference between these two quantities is :

$$\sum_{i=1}^{n-1} (f_i + f_{i+1} - r + 1) - \left[\sum_{i=1}^{n} (f_i-r+1) - 1 \right]$$

$$= \sum_{i=2}^{n-1} f_i + r \qquad \text{(with the convention } \sum_{i=2}^{1} f_i = o \text{)}$$

It is equal to the degree of freedom v iff $n = 2$ or $f_i = r$, $i = 2,\dots,n-1$. In the other cases some equations in (4.7) are functions of the other ones and two approaches are then possible : an application of the test procedure with a generalized inverse of Σ or the research of a subsystem with full rank (see Sargan [13]).

IV.4 - An example.

In the following model containing only one explanatory variable x :

$$y_t + \Psi_{11} y_{t-1} + \dots + \Psi_{1f} y_{t-f} + \Psi_{20} x_t + \dots + \Psi_{2f} x_{t-f} = e_t$$

$\Psi_{1f} \neq 0$, $\Psi_{2f} \neq 0$ the two lag polynomials $\Psi_1(L)$, $\Psi_2(L)$ have the same degree f. The condition for exactly r common roots between $\Psi_1(L)$, $\Psi_2(L)$ is :

$$H_{or} = \left\{ \exists\,!\ \beta_{jk} \quad j = 1,2 \;,\; k = o,\dots,f-r \;,\; \beta_{10} = 1 \;:\; \forall\, l = o,\dots,2f-r \right.$$

$$\left. \sum_{k=\mathrm{Max}(o,l-f+r)}^{\mathrm{Min}(l,f)} \Psi_{1k}\,\beta_{2,l-k} = \sum_{k=\mathrm{Max}(o,l-f+r)}^{\mathrm{Min}(l,f)} \Psi_{2k}\,\beta_{1,l-k} \right\}$$

These equations can also be written under matrix form. Let us denote :

$\Psi_1 = [\Psi_{10}, \Psi_{11}, \dots, \Psi_{1f}]'$, $\Psi_2 = [\Psi_{20}, \Psi_{21}, \dots, \Psi_{2f}]'$,

$\beta_1 = [\beta_{10}, \dots, \beta_{1,f-r}]'$, $\beta_2 = [\beta_{20}, \dots, \beta_{2,f-r}]'$ and 0_s a vector of s zeros, the null hypothesis H_{or} is defined by :

$$\exists\,!\ \beta_1, \beta_2 \text{ with } \beta_{10} = 1$$

$$\begin{bmatrix} 0_o & 0_1 & 0_2 & \cdots\cdots & 0_{f-r} & 0_o & 0_1 & \cdots\cdots & 0_{f-2} \\ -\Psi_2 & -\Psi_2 & -\Psi_2 & \cdots\cdots & -\Psi_2 & \Psi_1 & \Psi_1 & \cdots\cdots & \Psi_1 \\ 0_{f-r} & 0_{f-r-1} & 0_{f-r-2} & \cdots\cdots & 0_o & 0_{f-r} & 0_{f-r-1} & \cdots\cdots & 0_o \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = 0$$

The first matrix has $2f-r+1$ rows and $2f-2r+2$ columns. It depends on Ψ_1, Ψ_2 and it is denoted by $\Phi_r(\Psi_1,\Psi_2)$.

Then the test of H_{or} against $H_r - H_{or}$ can be based on arbitrary consistent asymptotically normal estimates $\hat{\Psi}_1$, $\hat{\Psi}_2$, of Ψ_1, Ψ_2, whose asymptotic covariance matrix is Q.

i) We first derive the asymptotic covariance matrix of $\Phi_r(\hat{\Psi}_1,\hat{\Psi}_2)\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$.

The analytical expression of this matrix is easily obtained ; in effect the derivative of $\Phi_r(\Psi_1,\Psi_2)\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$ with respect to Ψ_{1k} for instance is simply

$$\begin{bmatrix} 0_o & 0_1 & \cdots\cdots & 0_{f-r} \\ e_{(k+1)} & e_{(k+1)} & \cdots\cdots & e_{(k+1)} \\ 0_{f-r} & 0_{f-r-1} & \cdots\cdots & 0_o \end{bmatrix} \beta_2 ,$$

where $e_{(k+1)}$ is a $f+1$ dimensional vector, whose $(k+1)^{th}$ element is equal to one and the others are equal to zero. This covariance matrix depends on Q, β_1, β_2 and is denoted by $\Sigma[Q,\beta_1,\beta_2]$.

ii) Then we look for consistent estimates $\hat{Q},\hat{\beta}_1,\hat{\beta}_2$ of Q,β_1,β_2 and we deduce an estimator $\hat{\Sigma} = \Sigma[\hat{Q},\hat{\beta}_1,\hat{\beta}_2]$.

[$\hat{\beta}_1, \hat{\beta}_2$ can be obtained for instance by OLS applied to $o = \phi_r [\Psi_1, \Psi_2] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + v$

where $E\, v = o$, $V\, v = Id$; constrained by $\beta_{10} = 1$] .

iii) Finally the test statistic is obtained by considering the SSR in the linear model :

$o = \phi_r [\hat{\Psi}_1, \hat{\Psi}_2] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \hat{v}$, where $E\, \hat{v} = o$, $V\, \hat{v} = \hat{\Sigma}$ and where the first element of β_1 : β_{10} is constrained by $\beta_{10} = 1$. This test statistic has asymptotically under the null a $\chi^2(r)$ distribution.

V - IDENTIFICATION OF AN AUTOREGRESSIVE MOVING AVERAGE (ARMA) MODEL.

V.1 - The hypothesis and its specification in terms of unknown linear restrictions.

A real process $(X_t)_{t \in Z}$ has an ARMA representation of orders p and q if there exists a white noise $(\varepsilon_t)_{t \in Z}$ and (p+q) real numbers $\varphi_1, \varphi_2, \ldots, \varphi_p$ $\theta_1, \theta_2, \ldots, \theta_q$ such that :

$$(5.1) \quad X_t = \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t - \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}$$

where lag polynomials $\phi(L) = 1 - \sum_{i=1}^{p} \varphi_i L^i$ and $\theta(L) = 1 - \sum_{j=1}^{q} \theta_j L^j$ have their roots outside the unit circle.

Such an ARMA process has an infinity of ARMA representations, since it is always possible to multiply both sides of : $\phi(L)\, X_t = \theta(L)\, \varepsilon_t$ by a same polynomial. Therefore it is necessary to have a definition of the orders, which gives the uniqueness of p and q . This definition is based on the notion of minimal ARMA (p,q) representation. This ARMA representation is minimal if (5.1) is satisfied and if moreover : $\theta_q \neq 0$, $\varphi_p \neq 0$ and $\Phi(L)$ and $\theta(L)$ have no common roots.

In this section, we are interested in the determination of the minimal orders p and q by a sequential test procedure, whose each step corresponds to the test of unknown linear restrictions.

We introduce the following hypothesis :

$H(p,q)$: "the process X has an ARMA(p,q) representation"

$H^*(p,q)$: "the process X has a minimal ARMA(p,q) representation".

This last hypothesis is related to the previous ones by :

(5.2) $H^*(p,q) = H(p,q) - [H(p-1,q) \cup H(p,q-1)]$

Conversely, we have :

$$H(p,q) = \bigcup_{\substack{p' \leq p \\ q' \leq q}} H^*(p',q')$$

Let us now consider a process X with an ARMA representation. The orders of this representation can be characterized by some conditions on the autocorrelation function :

(5.3) $$\rho(h) = \frac{\mathrm{Cov}(X_t, X_{t+h})}{V X_t} \qquad h = 1,2,\ldots$$

We have the following well-known property :

PROPERTY 5.4 :

> X has an ARMA(p,q) representation, i.e the hypothesis $H(p,q)$ is satisfied, if and only if there exists p real numbers $\varphi_1,\dots,\varphi_p$ such that :
>
> $$\rho(h) = \sum_{i=1}^{p} \varphi_i \, \rho(h-i) \qquad h = q+1, q+2\dots$$

Let us remark that it results from the convergence of $\sum_{h=-\infty}^{+\infty} |\rho(h)|$ that the coefficients $\varphi_1,\dots,\varphi_p$ are such that $\Phi(L)$ has its roots outside the unit circle. Moreover these coefficients give the autoregressive part of an ARMA representation of X .

These coefficients are not necessarily unique and therefore the previous set of conditions cannot be directly interpreted as unknown linear restrictions.

However we have the following result.

PROPERTY 5.5 :

> Let us assume that $H(p-1,q)$ is not satisfied ; then a necessary and sufficient condition for $H(p,q)$ to be satisfied is the condition of property 5.4 . Moreover, if there exist, the real numbers $\varphi_1,\dots,\varphi_p$ are unique.

Proof : This is a direct consequence of the form of the spectral density, which allows to characterize the second order properties of the process by the ratio $\frac{\theta}{\Phi}$.

□

The sequence of tests is performed along the following lines :

i) We first test for an ARMA(o,o) representation, i.e for $H(o,o)$. If this hypothesis is accepted the minimal orders are $p = o$ and $q = o$.

ii) Otherwise we test $H(1,o)$. If $H(1.o)$ is accepted the minimal orders are $p = 1$, $q = o$.

iii) Otherwise we test $H(o,1)$. If $H(o,1)$ is accepted the minimal orders are $p = o$, $q = 1$.

iv) Otherwise we test $H(o,2)$ and so on ....

The complete design is summarized below, where the numbers between parentheses are the orders :

$$
\begin{tikzcd}
H(o,o) \arrow[d] & H(o,1) \arrow[r] & H(o,2) \arrow[dl] & H(o,3) \arrow[r] & {} \\
H(1,o) \arrow[ur] & H(1,1) \arrow[dl] & H(1,2) \arrow[ur] & & \\
H(2,o) \arrow[d] & H(2,1) \arrow[ur] & & & \\
H(3,o) \arrow[ur] & & & &
\end{tikzcd}
$$

The advantages of this recursive approach are the following :

i) it results from property 5.2 that the first time an hypothesis $H(p^*,q^*)$ is accepted, then the minimal orders are p^* and q^* ;

ii) it results from properties 5.4 and 5.5 that the constraints defining $H(p,q)$ are linear in parameters $\varphi_1,\ldots,\varphi_p$ and that these parameters are uniquely defined (since $H(p,q)$ is only considered if $H(p-1,q)$ has been rejected) .

Let us finally remark that from a practical viewpoint, we can only consider a finite set of restrictions : $\varrho(h) = \sum_{i=1}^{p} \varphi_i \varrho(h-i)$, $h = q+1, q+2, \dots, q+r$, where r is an a priori given number.

V.2 - Determination of the test statistics.

The "unknown linear restrictions" test procedure can be applied at each step for testing :

$$H_o = \left\{ H(p,q) - H(p-1,q) \right\}$$

in the context of the general hypothesis $H = {}^{c}H(p-1,q)$. In fact this is only possible if $p \geqslant 1$. For testing M A hypothesis an approach based on the sample autocorrelation can be directly used.

The null hypothesis can be approximated by :

$$H_{o,r} = \left\{ \exists \, ! \, \varphi_1, \dots, \varphi_p \; : \; \varrho(h) = \sum_{i=1}^{p} \varphi_i \, \varrho(h-i) \quad h = q+1, \dots, q+r \right\}$$

The autocorrelations $\varrho(h)$ are estimated by their sample counterpart $\hat{\varrho}(h)$. The asymptotic properties of $\hat{\varrho}(h)$ are for instance given in Anderson [1] . In particular, the asymptotic covariance between $\sqrt{T}\,[\hat{\varrho}(h) - \varrho(h)]$ and $\sqrt{T}\,[\hat{\varrho}(k) - \varrho(k)]$, where T is the number of observations is :

$$(5.6) \quad \mathrm{Cov}_{as} \left\{ \sqrt{T}\,[\hat{\varrho}(h) - \varrho(h)] \, , \, \sqrt{T}\,[\hat{\varrho}(k) - \varrho(k)] \right\}$$

$$= \sum_{l=-\infty}^{+\infty} \Big\{ \varrho(l)\, \varrho(l-h+k) + \varrho(l+k)\, \varrho(l-h) - 2\varrho(k)\, \varrho(l)\, \varrho(l-h)$$

$$- 2\varrho(h)\, \varrho(l)\, \varrho(l-k) + 2\varrho(h)\, \varrho(k)\, \varrho^2(l) \Big\}$$

We can use this formula to derive the asymptotic covariance of

$\hat{\varrho}(h) - \sum_{i=1}^{p} \varphi_i \hat{\varrho}(h-i)$, $h = q+1,\dots,q+r$ under the null hypothesis H_o .

PROPERTY 5.7 :

> If process $(X_t)_{t\in Z}$ has an ARMA (p,q) representation :
>
> $$\Phi(L) X_t = \theta(L) \varepsilon_t ,$$
>
> then the asymptotic covariance matrix of :
>
> $$\sqrt{T} \left[\hat{\varrho}(h) - \sum_{i=1}^{p} \varphi_i \hat{\varrho}(h-i)\right]_{q+1\leqslant h\leqslant q+r}$$
>
> is equal to the covariance matrix of $(Y_t)_{q+1\leqslant t\leqslant q+r}$, where Y has the following ARMA(p,q^2) representation :
>
> $\Phi(L) Y_t = [\theta(L)]^2 \eta_t$, with a white noise η_t , such that
>
> $$V \eta_t = \left(\frac{V\varepsilon_t}{VX_t}\right)^2 = \frac{\sigma^4}{\gamma(o)^2} \quad \text{(Say)}$$

Proof : See Appendix.

The previous property shows that for $h \geqslant q+1$ the process $\hat{\varrho}(h) - \sum_{i=1}^{p} \varphi_i \varrho(h-i)$ can be asymptotically considered as an ARMA(p,q^2) process. In particular it appears to be asymptotically stationary. There exist several ways to estimate the asymptotic covariance matrix Σ of section II. For instance, we can in a first step consider the solutions $\tilde{\tilde{\varphi}}_i$, $i = 1,\dots,p$ of the Yule-Walker equations and then replace the (k,l) element of Σ (with $k \geqslant l$) by :

$$\hat{\Sigma}_{kl} = \frac{1}{H} \sum_{h=q+1}^{q+H} \left[\hat{\varrho}(h) - \sum_{i=1}^{p} \tilde{\tilde{\varphi}}_i \hat{\varrho}(h-i)\right] \left[\hat{\varrho}(h+k-l) - \sum_{i=1}^{p} \tilde{\tilde{\varphi}}_i \hat{\varrho}(h+k-l-i)\right]$$

using the asymptotic stationarity of property 5.7.

V.3 - Comparison with an approach proposed by Glasbey.

In a recent paper, Glasbey [7] proposes an identification procedure based on the notion of generalized partial autocorrelations. As shown in Tuan [15] and Berlinet [3], this approach is related to the corner method of Beguin - Gourieroux - Monfort [2] and to the S. Array method of Gray - Kelley - MacIntire [10]. The generalized partial autocorrelations are defined in the following way :

i) We first consider the solution $\varphi_1^*, \ldots, \varphi_p^*$ of the theoretical Yule-Walker equations :

$$\begin{cases} \varrho(q+1) = \varphi_1^* \varrho(q) + \ldots + \varphi_p^* \varrho(q-p+1) \\ \varrho(q+2) = \varphi_1^* \varrho(q+1) + \ldots + \varphi_p^* \varrho(q-p+2) \\ \vdots \\ \varrho(q+p) = \varphi_1^* \varrho(q+p-1) + \ldots + \varphi_p^* \varrho(q) \end{cases}$$

ii) Then the generalized partial autocorrelation is defined as the discrepancy :

$$\varrho(q+p+1) - \varphi_1^* \varrho(q+p) \ldots \varphi_p^* \varrho(q+1)$$

Under the null hypothesis of an ARMA (p,q) representation this discrepancy is equal to zero.

Therefore Glasbey [7] proposes to test for the significativity of this discrepancy. All the theoretical derivations are replaced by sample ones :

i) $\overset{\approx}{\varphi}_1, \ldots, \overset{\approx}{\varphi}_p$ are defined as the solutions of :

$$\left\{\begin{array}{lcl}\hat{\varrho}(q+1) & = & \tilde{\tilde{\varphi}}_1\ \hat{\varrho}(q) \quad +\ldots+\ \tilde{\tilde{\varphi}}_p\ \hat{\varrho}(q-p+1)\\ \hat{\varrho}(q+2) & = & \tilde{\tilde{\varphi}}_1\ \hat{\varrho}(q+1) \quad +\ldots+\ \tilde{\tilde{\varphi}}_p\ \hat{\varrho}(q-p+2)\\ \vdots & & \\ \hat{\varrho}(q+p) & = & \tilde{\tilde{\varphi}}_1\ \hat{\varrho}(q+p-1) +\ldots+\ \tilde{\tilde{\varphi}}_p\ \hat{\varrho}(q)\end{array}\right.$$

ii) The discrepancy is estimated by :

$$y_{pq} = \hat{\varrho}(q+p+1) - \tilde{\tilde{\varphi}}_1\ \hat{\varrho}(q+p) \ \ldots - \tilde{\tilde{\varphi}}_p\ \hat{\varrho}(q+1)$$

iii) Finally if $\hat{\sigma}_{y_{pq}}$ denotes the estimated asymptotic standard error of y_{pq} the test consists in comparing the student ratio $\left|\dfrac{y_{pq}}{\hat{\sigma}_{y_{pq}}}\right|$ to the usual critical level 1.96 .

This approach can be easily related to the "unknown linear restrictions" test procedure of subsection V.2 . In effect, let us introduce the restrictions $H_{o,r}$ associated with $r = p+1$. The conditions

$$H_{o,p+1} = \left\{ \exists\ !\ \varphi_1,\ldots,\varphi_p\ :\ \varrho(h) = \varphi_1\ \varrho(h-1) +\ldots+ \varphi_p\,\varrho(h-p), h = q+1,\ldots,q+p+1 \right\}$$

are equivalent to :

$H_{o,p+1}$ = ⟨ In the unique decomposition

$$\left\{\begin{array}{lcl}\varrho(q+1) & = & \varphi_1\ \varrho(q) \quad +\ldots+\ \varphi_p\ \varrho(q-p+1) \ + \ o\\ \vdots & & \\ \varrho(q+p) & = & \varphi_1\ \varrho(q+p-1) +\ldots+\ \varphi_p\ \varrho(q) \quad + \ o\\ \varrho(q+p+1) & = & \varphi_1\ \varrho(q+p) \quad +\ldots+\ \varphi_p\ \varrho(q+1) \quad + \ \mu\end{array}\right.$$

the parameter $\mu = o$ ⟩ .

We have seen in II.1.b that the asymptotic test of $H_{o,p+1}$ under the first form is equivalent to the test of $\mu = o$. Moreover it is easily seen that the G.L.S. estimator of $\varphi_1,\ldots,\varphi_p$, μ coincides with the OLS one and that in particular the OLS estimator of μ is equal to the estimated discrepancy y_{pq} . Therefore we have the following property.

PROPERTY 5.8 :

> The test procedure proposed in Glasbey is asymptotically equivalent under the null to the unknown linear restrictions test procedure based on $r = p+1$ restrictions.

Moreover, from property 2.9 we immediately deduce the following corollary :

Corollary 5.9 :

> An unknown linear restrictions test procedure based on more than $p+1$ restrictions : $r > p+1$ is asymptotically locally most powerful than Glasbey's test procedure.

V.4 - Comparaison with Box-Pierce test.

An important particular case concerns the pure autoregressive representations AR(p) .

In this case the estimators $\hat{\varphi}_1,\ldots,\hat{\varphi}_p$ of $\varphi_1,\ldots,\varphi_p$ deduced from the Yule-Walker equations are under the null asymptotically equivalent to the maximum likelihood estimators of $\varphi_1,\ldots,\varphi_p$ (under the assumption of gaussian white noise).

Moreover, the process : $\sqrt{T}\,[\hat{\varrho}(h) - \varphi_1\,\hat{\varrho}(h-1)\,\dots - \varphi_p\,\hat{\varrho}(h-p)]_{1\leqslant h}$ has asymptotically an AR(p) representation :

$$\phi(L)\,\sqrt{T}\,[\hat{\varrho}(h) - \varphi_1\,\hat{\varrho}(h-1)\,\dots - \varphi_p\,\hat{\varrho}(h-p)] = \eta_h \qquad h \geqslant 1 \quad ,$$

$$\text{with} \quad V\,\eta_h = \frac{\sigma^4}{[\gamma(o)]^2} \quad .$$

Therefore this process has also a forward representation :

$$\phi(F)\,\sqrt{T}\,[\hat{\varrho}(h) - \varphi_1\,\hat{\varrho}(h-1)\,\dots - \varphi_p\,\hat{\varrho}(h-p)] = \xi_h \qquad h \geqslant 1$$

$$\text{with} \quad V\,\xi_h = \frac{\sigma^4}{[\gamma(o)]^2} \quad ,$$

where F is the forward operator.

Then for r sufficiently large the unknown linear restriction test statistic is equivalent to :

$$\xi_r = T\,\frac{\hat{\gamma}(o)^2}{\hat{\sigma}^4}\,\sum_{h=1}^{r}\left(\hat{\phi}(F)\,[\hat{\varrho}(h) - \hat{\varphi}_1\,\varrho(h-1)\,\dots - \hat{\varphi}_p\,\varrho(h-p)]\right)^2$$

and is under the null asymptotically distributed as a $\chi^2(r-p)$.

It is possible to give an equivalent form for ξ_r . We have :

$$\xi_r = T\,\frac{\hat{\gamma}(o)^2}{\hat{\sigma}^4}\,\sum_{h=1}^{r}\frac{[\hat{\phi}(F)\,\hat{\phi}(L)\,\hat{\gamma}(h)]^2}{\hat{\gamma}(o)^2} = T\,\sum_{h=1}^{r}\frac{[\hat{\phi}(F)\,\hat{\phi}(L)\,\hat{\gamma}(h)]^2}{\hat{\sigma}^4} \quad ,$$

$$\text{where} \quad \hat{\gamma}(h) = \frac{1}{T}\sum_{t=h+1}^{T} X_t\,X_{t-h} \quad .$$

Since : $\hat{\phi}(F)\,\hat{\phi}(L)\,\hat{\gamma}(h) = \hat{\gamma}_{\hat{\varepsilon}}(h)$, where $\hat{\varepsilon}_t = \phi(L)\,X_t$

is the residual at time t, it is seen that ξ_r coïncides with the usual Box-Pierce statistic [5].

$$\xi_r = T \sum_{h=1}^{r} [\hat{\rho}_{\hat{\varepsilon}}(h)]^2$$

VI - TESTING FOR CO-INTEGRATION.

VI.1 - Co-integrated processes.

Let us consider a unidimensional process y and a k-dimensional one x, with a joint ARIMA representation :

$$(6.1) \qquad (1 - L)\begin{bmatrix} \phi_{11}(L) & \phi_{12}(L) \\ \phi_{21}(L) & \phi_{22}(L) \end{bmatrix} \begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \theta_1(L) \\ \theta_2(L) \end{bmatrix} \varepsilon_t ,$$

where (ε_t) is a k+1 dimensional white noise and the lag polynomials

$\det \begin{bmatrix} \phi_{11}(L) & \phi_{12}(L) \\ \phi_{21}(L) & \phi_{22}(L) \end{bmatrix}$, $\det \begin{bmatrix} \theta_1(L) \\ \theta_2(L) \end{bmatrix}$ have their roots outside the unit circle. Each component of $\begin{bmatrix} y_t \\ x_t \end{bmatrix}$ has an ARIMA (p,d,q) representation with an integration order $d = 1$.

Granger [9] has introduced the concept of co-integrated time series, which transposes in mathematical terms the idea of non stationary processes in stationary relationship.

Definition 6.2 :

y is co-integrated with x, if there exist some constants $\lambda_1,\ldots,\lambda_k$ such that the process :

$z_t = y_t - (\lambda_1 x_{1t} \ldots - \lambda_k x_{kt}) = y_t - \lambda' x_t$

has an ARMA representation.

Under the co-integration hypothesis, structural model (6.1) takes the form of an error correction model (E.C.M.) and the relation $y_t = \lambda' x_t$ can be interpreted as a kind of long term equilibrium [Davidson - Hendry - Srba - Yeo [6]].

VI.2 - Expression of the co-integration hypothesis.

Multiplying both sides of (6.1) by $[1 : - \lambda']$, we obtain :

$$(1 - L) \; z_t = [1 : - \lambda'] \begin{bmatrix} \Phi_{11}(L) & \Phi_{12}(L) \\ \Phi_{21}(L) & \Phi_{22}(L) \end{bmatrix}^{-1} \begin{bmatrix} C_1(L) \\ C_2(L) \end{bmatrix} \varepsilon_t \; .$$

Therefore (z_t) has a stationary ARMA representation iff the lag-polynomial :

$$\Psi(L) = [1 : - \lambda'] \begin{bmatrix} \Phi_{11}(L) & \Phi_{12}(L) \\ \Phi_{21}(L) & \Phi_{22}(L) \end{bmatrix}^{-1} \begin{bmatrix} C_1(L) \\ C_2(L) \end{bmatrix}$$

can be divided by $1 - L$, or equivalently iff $\Psi(1) = 0$, i.e iff the long term effects of (ε_t) on (z_t) are all equal to zero.

We have :

$$\Psi(1) = o \iff [1 : -\lambda'] \begin{bmatrix} \phi_{11}(1) & \phi_{12}(1) \\ \phi_{21}(1) & \phi_{22}(1) \end{bmatrix}^{-1} \begin{bmatrix} C_1(1) \\ C_2(1) \end{bmatrix} = o$$

Introducing the bloc elements $\phi^{ij}(L)$ of the inverse of the autoregressive matrix, we obtain :

$$[1 : -\lambda'] \begin{bmatrix} \phi^{11}(1) & \phi^{12}(1) \\ \phi^{21}(1) & \phi^{22}(1) \end{bmatrix}^{-1} \begin{bmatrix} C_1(1) \\ C_2(1) \end{bmatrix} = o$$

$$\iff \quad [1 : -\lambda'] \begin{bmatrix} \phi^{11}(1)\, C_1(1) + \phi^{12}(1)\, C_2(1) \\ \phi^{21}(1)\, C_1(1) + \phi^{22}(1)\, C_2(1) \end{bmatrix} = o$$

$$\iff (6.3) \quad \phi^{11}(1)\, C_1(1) + \phi^{12}(1)\, C_2(1) - \lambda'[\phi^{21}(1)\, C_1(1) + \phi^{22}(1)\, C_2(1)] = o$$

These conditions are expressed as "unknown linear restritions". The auxiliary parameter λ gives the coefficients of the long term equilibrium relationship.

Since λ is of size k and there are $k+1$ relations in the system, the degree of freedom for the test statistic is equal to 1 .

To apply test procedure of section II, it is necessary to study the uniqueness of λ . Let us assume that there exist two different vectors λ and μ such that :

$$z_t = y_t - \lambda' x_t \qquad \text{and} \qquad \tilde{\tilde{z}}_t = y_t - \mu' x_t$$

are both stationary. We deduce that :

$z_t - \tilde{\tilde{z}}_t = (\mu' - \lambda')\, x_t$ is also stationary, which means that the processes component of x are co-integrated. Therefore a necessary and sufficient condition for λ to be unique is that the components of x are not co-integrated.

To satisfied this additional condition we are now led to a sequential procedure, which is here given for $K = 3$

i) Firstly test for co-integration of (x_1, x_2) , (x_1, x_3) , (x_2, x_3) .

ii) If these co-integration hypotheses are rejected, then test for co-integration of (x_1, x_2, x_3) .

iii) Finally if this second test rejects the hypothesis, test for co-integration of (y, x_1, x_2, x_3) .

In fact this approach simply consists before examining the possible relations between y and (x_1, x_2, x_3) to eliminate the possible "quasi-collinearities" between the exogenous variables.

APPENDIX 1

COMPUTATION OF $\frac{\partial g}{\partial \theta'}(\theta)$, $\theta \in H_o$

We have :

$$g(\theta) = h_o(\theta) - H(\theta)[H'(\theta)\Omega^{-1}(\theta)H(\theta)]^{-1} H'(\theta)\Omega^{-1}(\theta)\, h_o(\theta) \quad ,$$

or

$$g = h_o - H[H'\Omega^{-1}H]^{-1} H'\Omega^{-1}h_o \quad .$$

This implies

$$dg = dh_o - dH[H'\Omega^{-1}H]^{-1} H'\Omega^{-1}h_o - Hd[(H'\Omega^{-1}H)^{-1} H'\Omega^{-1}h_o]$$

or, denoting $[H'\Omega^{-1}H]^{-1} H'\Omega^{-1}h_o$ by λ :

$$dg = dh_o - dH\lambda - Hd[(H'\Omega^{-1}H)^{-1} H'\Omega^{-1}h_o]$$

$$= dh_o - dH\lambda - Hd[(H'\Omega^{-1}H)^{-1}] H'\Omega^{-1}h_o - H(H'\Omega^{-1}H)^{-1}d(H'\Omega^{-1}h_o)$$

or, using $d(A^{-1}) = - A^{-1}\, dA\, A^{-1}$:

$$dg = dh_o - dH\lambda + H(H'\Omega^{-1}H)^{-1} d(H'\Omega^{-1}H)\, (H'\Omega^{-1}H)^{-1} H'\Omega^{-1}h_o$$

$$- H(H'\Omega^{-1}H)^{-1} d(H'\Omega^{-1}h_o)$$

$$= dh_o - dH\lambda + H(H'\Omega^{-1}H)^{-1} [dH'\Omega^{-1}H + H'd(\Omega^{-1})H + H'\Omega^{-1}dH] \lambda$$

$$- H(H'\Omega^{-1}H)^{-1} [dH'\Omega^{-1}h_o + H'd(\Omega^{-1})h_o + H'\Omega^{-1}dh_o]$$

Since, on H_o, we have $h_o = H\lambda$, this reduces to :

$$dg = dh_o - dH\lambda + M_\theta dH\lambda - M_\theta dh_o ,$$

where $M_\theta = H(H'\Omega^{-1}H)^{-1}H'\Omega^{-1}$ is the orthogonal projector considered in II.1.b.

Finally, on H_o we have :

$$dg = (I - M_\theta)(dh_o - dH\lambda)$$

This implies :

$$\frac{\partial g}{\partial \theta'} = (I - M_\theta) \left(\frac{\partial h_o}{\partial \theta'} - \frac{\partial H}{\partial \theta'} \lambda \right)$$

APPENDIX 2

ASYMPTOTIC EXPANSION OF $\sqrt{n}\, g(\hat{\theta}_n)$

Let us consider :

$$\sqrt{n}\,(I - \hat{M})\,\hat{h}_o \quad = \quad \sqrt{n}\, g(\hat{\theta}_n)$$

Since, under H_o , $g(\theta) = o$ the previous appendix gives :

$$\sqrt{n}\, g(\hat{\theta}_n) \quad = \quad \frac{\partial g}{\partial \theta'}\, \sqrt{n}\,(\hat{\theta}_n - \theta) \;+\; o_p(1)$$

$$= \quad (I - M_\theta)\left(\frac{\partial h_o}{\partial \theta'} - \frac{\partial H}{\partial \theta'}\,\lambda\right)\sqrt{n}\,(\hat{\theta}_n - \theta) \;+\; o_p(1)$$

or

$$\sqrt{n}\, g(\hat{\theta}_n) \quad = \quad \sqrt{n}\,(I - M_\theta)\,(\hat{h}_o - \hat{H}\lambda) \;+\; o_p(1)$$

APPENDIX 3

PROOF OF PROPERTY 4.6

i) Necessary part

Let us first consider n polynomials $\Psi_j(L)$, $j = 1,\dots,n$ with r common roots and $\Psi_{10} = 1$. There exist polynomials $\varrho(L) = \sum_{k=o}^{r} \varrho_k L^k$ with $\varrho_o = 1$, $\varrho_r \neq o$ and $\alpha_j(L)$ $j = 1,\dots,n$ with degrees $f_j - r$ and $\alpha_{10} = 1$ such that :

$$\Psi_j(L) = \varrho(L)\, \alpha_j(L) \qquad j = 1,\dots,n \quad .$$

We deduce that :

$$\Psi_j(L)\, \alpha_{j+1}(L) = \varrho(L)\, \alpha_j(L)\, \alpha_{j+1}(L) = \Psi_{j+1}(L)\, \alpha_j(L)$$

and we have exhibited a set of polynomials $\beta_j(L) = \alpha_j(L)$, $j = 1,\dots,n$ satisfying the conditions of property 4.6

ii) Sufficient part

Let us now consider a set of polynomials $\Psi_j(L)$, $j = 1,\dots,n$ such that there exist $\beta_j(L)$ $j = 1,\dots,n$ of degrees $f_j - r$ satisfying :

$$\Psi_j(L)\, \beta_{j+1}(L) = \Psi_{j+1}(L)\, \beta_j(L) \quad j = 1,\dots,n-1 \quad .$$

a) It is always possible to divide all these equations by the greatest common factor of the β_j's , denoted by $cf[\beta_1,\dots,\beta_n]$. We then obtained a set of relations of the same kind : $\Psi_j(L)\, \bar{\beta}_{j+1}(L) = \Psi_{j+1}(L)\, \bar{\beta}_j(L)$, $j = 1,\dots,n-1$

with $\bar{\beta}_j(L) = \dfrac{\beta_j(L)}{cf[\beta_1,\dots,\beta_n]}$ of degrees $f_j - \bar{r} = f_j - r - \overset{\circ}{d}\, cf[\beta_1,\dots,\beta_n]$.

b) It is then possible to prove by a recursive approach that $\bar{\beta}_1(L)$ is a divider of $\Psi_1(L)$. To simplify the exposition, we assume without loss of generality that $\bar{\beta}_1(L)$ has no multiple roots.

. If $\gamma_1(L) = \dfrac{\bar{\beta}_1(L)}{cf[\bar{\beta}_1(L),\bar{\beta}_2(L)]}$, where $cf[\bar{\beta}_1(L),\bar{\beta}_2(L)]$ is the greatest common factor of $\bar{\beta}_1$ and $\bar{\beta}_2$, we deduce from the first equation $\Psi_1(L)\,\bar{\beta}_2(L) = \Psi_2(L)\,\bar{\beta}_1(L)$ that $\gamma_1(L)$ is a divider of $\Psi_1(L)$.

. If we now introduce $\gamma_2(L) = \dfrac{cf[\bar{\beta}_1(L),\bar{\beta}_2(L)]}{cf[\bar{\beta}_1(L),\bar{\beta}_2(L),\bar{\beta}_3(L)]}$, we can examine the first and second equations :

$$\begin{cases} \Psi_1(L)\,\bar{\beta}_2(L) = \Psi_2(L)\,\bar{\beta}_1(L) \\ \\ \Psi_2(L)\,\bar{\beta}_3(L) = \Psi_3(L)\,\bar{\beta}_2(L) \end{cases}$$

From the second one, we deduce that $\gamma_2(L)$ is a divider of $\Psi_2(L)$ and, using the first one, that $\gamma_2(L)^2$ is a divider of $\Psi_2(L)\,\bar{\beta}_1(L) = \Psi_1(L)\,\bar{\beta}_2(L)$. This implies that $\gamma_2(L)$ divides $\Psi_1(L)$.

At the j^{th} step the same approach is followed introducing

$\gamma_j(L) = \dfrac{cf[\bar{\beta}_1(L),\dots,\bar{\beta}_j(L)]}{cf[\bar{\beta}_1(L),\dots,\bar{\beta}_{j+1}(L)]}$ and considering the first j equations and this for $j = 1,\dots,n-1$. This proves that $\Psi_1(L)$ can be divided by $\gamma_1(L),\gamma_2(L),\dots,\gamma_{n-1}(L)$ or, using the assumption of no multiples roots, by

$\gamma_1(L)\gamma_2(L),\dots,\gamma_{n-1}(L) = \bar{\beta}_1(L)$. Therefore, there exists a polynomial $\varrho(L)$ of degree $\bar{r}$ such that $\Psi_1(L) = \varrho(L)\,\bar{\beta}_1(L)$.

c) From the equations we immediately deduce that :

$$\frac{\Psi_j(L)}{\bar{\beta}_j(L)} = \frac{\Psi_1(L)}{\bar{\beta}_1(L)} = \varrho(L) \qquad \forall\, j \quad ,$$

and we have exhibited a common factor $\varrho(L)$ of degree $\bar{r} \geqslant r$.

Finally remark that the assumption of property 4.6 that there exists no more than r common roots implies that $\bar{r} = r$ and that the $\bar{\beta}_j$ are such that : $\overset{\circ}{d}\,cf[\bar{\beta}_1,\dots,\bar{\beta}_n] = o$.

iii) <u>Uniqueness condition</u>

If there were two different sets of polynomials $\beta_j(L)$ and $\beta^*_j(L)$ satisfying conditions of property 4.6, we would have :

$$\Psi_j(L)\,\beta_{j+1}(L) = \Psi_{j+1}(L)\,\beta_j(L) \quad , \quad \overset{\circ}{d}\,\beta_j(L) = f_j - r \quad ,$$

and
$$\Psi_j(L)\,\beta^*_{j+1}(L) = \Psi_{j+1}(L)\,\beta^*_j(L) \quad , \quad \overset{\circ}{d}\,\beta^*_j(L) = f_j - r \quad , \qquad \forall\, j$$

Therefore we would have :

$$\Psi_j(L)\left[\beta_{j+1}(L) - \frac{\beta_{1,f_1-r}}{\beta^*_{1,f_1-r}}\,\beta^*_{j+1}(L)\right] = \Psi_{j+1}(L)\left[\beta_j(L) - \frac{\beta_{1,f_1-r}}{\beta^*_{1,f_1-r}}\,\beta^*_j(L)\right]$$

Since $\beta_1(L) - \dfrac{\beta_{1,f_{1-r}}}{\beta^*_{1,f_1-r}}\,\beta^*_1(L)$ is of degree at most $f_1 - r - 1$, the polynomials $\gamma_j(L) = \beta_j(L) - \frac{\beta_{1,f_1-r}}{\beta^*_{1,f_1-r}} \beta^*_j(L)$ is of degree at most f_j-r-1 and it results from part ii) that the polynomials $\Psi_j(L)$ $j = 1,\ldots,n$ have as common factor a polynomial of degree at less $r + 1$. This is in contradiction with the assumption of no more than r common roots made in property 4.6 .

APPENDIX 4

PROOF OF PROPERTY 5.7

We denote by

$$\mathrm{Cov}_{as}\left[\sqrt{T}\sum_{i=o}^{p}\varphi_i\,\hat{\varrho}(h-i)\quad,\quad\sqrt{T}\sum_{j=o}^{p}\varphi_j\,\hat{\varrho}(k-j)\right]$$

where $\varphi_o = -1$, the (h,k) element of the asymptotic covariance matrix of

$$\sqrt{T}\left[\hat{\varrho}(h)\quad-\quad\sum_{i=o}^{p}\varphi_i\,\hat{\varrho}(h-i)\right]_{q+1\leq h\leq q+r}$$

From (5.6) , we deduce that :

$$\mathrm{Cov}_{as}\left[\sqrt{T}\sum_{i=o}^{p}\varphi_i\,\hat{\varrho}(h-i)\quad,\quad\sqrt{T}\sum_{j=o}^{p}\varphi_j\,\hat{\varrho}(h-j)\right]$$

$$=\sum_{i=o}^{p}\sum_{j=o}^{p}\varphi_i\,\varphi_j\,\mathrm{Cov}_{as}\left[\sqrt{T}\,[\hat{\varrho}(h-i)-\varrho(h-i)]\;,\;\sqrt{T}\,[\hat{\varrho}(k-j)-\varrho(k-j)]\right]$$

$$=\sum_{l=-\infty}^{+\infty}\sum_{i=o}^{p}\sum_{j=o}^{p}\varphi_i\,\varphi_j\,\varrho(l)\varrho(l-h+i+k-j)$$

$$+\sum_{l=-\infty}^{+\infty}\sum_{i=o}^{p}\sum_{j=o}^{p}\varphi_i\,\varphi_j\,\varrho(l+k-j)\varrho(l-h+i)$$

$$-\;2\sum_{l=-\infty}^{+\infty}\sum_{i=o}^{p}\sum_{j=o}^{p}\varphi_i\,\varphi_j\,\varrho(k-j)\varrho(l)\varrho(l-h+i)$$

$$-\;2\sum_{l=-\infty}^{+\infty}\sum_{i=o}^{p}\sum_{j=o}^{p}\varphi_i\,\varphi_j\,\varrho(h-i)\varrho(l)\varrho(l-k+j)$$

$$+ \; 2 \sum_{l=-\infty}^{+\infty} \sum_{i=o}^{p} \sum_{j=o}^{p} \varphi_i \, \varphi_j \, \varrho(h-i)\varrho(k-j)\varrho^2(l) \quad .$$

But :

$$\sum_{l=-\infty}^{+\infty} \sum_{i=o}^{p} \sum_{j=o}^{p} \varphi_i \, \varphi_j \, \varrho(k-j)\varrho(l)\varrho(l-h+i)$$

$$= \quad \sum_{l=-\infty}^{+\infty} \sum_{i=o}^{p} \varphi_i \, \varrho(l)\varrho(l-h+i) \sum_{j=o}^{p} \varphi_j \, \varrho(k-j) \; = \; o \quad ,$$

since, from property 5.4 :

$$\sum_{j=o}^{p} \varphi_j \, \varrho(k-j) \; = \; o \qquad \forall \; k \geqslant q+1 \quad .$$

In a similar way, we have :

$$\sum_{l=-\infty}^{+\infty} \sum_{i=o}^{p} \sum_{j=o}^{p} \varphi_i \, \varphi_j \, \varrho(h-i)\varrho(l)\varrho(l-k+j) \; = \; o$$

and :

$$\sum_{l=-\infty}^{+\infty} \sum_{i=o}^{p} \sum_{j=o}^{p} \varphi_i \, \varphi_j \, \varrho(h-i)\varrho((k-j) \; \varrho^2(l) \; = \; o$$

Finally :

$$\sum_{l=-\infty}^{+\infty} \sum_{i=o}^{p} \sum_{j=o}^{p} \varphi_i \, \varphi_j \, \varrho(l+k-j)\varrho(l-h+i)$$

is a sum $\sum_{l=-\infty}^{+\infty} \alpha_l$ with :

$$\alpha_l \; = \; \sum_{i=o}^{p} \varphi_i \, \varrho(h-l-i) \; \sum_{j=o}^{p} \varphi_j \, \varrho(l+k-j)$$

From property 5.4 , α_l is null as soon as :

$$h - l \geq q + 1 \quad \text{or} \quad l + k \geq q + 1 \quad .$$

But, since h and k are both greater than q :

$$h - l < q + 1 \implies l > o \implies l + k > q + 1$$

Hence $\alpha_l = o$ for every l .

Thus :

$$\mathrm{Cov}_{as} \left[\sqrt{T} \sum_{i=o}^{p} \varphi_i \hat{\varrho}(h-i) \ , \ \sqrt{T} \sum_{j=o}^{p} \varphi_j \hat{\varrho}(k-j) \right]$$

$$= \sum_{l=-\infty}^{+\infty} \sum_{i=o}^{p} \sum_{j=o}^{p} \varphi_i \varphi_j \varrho(l)\varrho(l-h+i+k-j)$$

Moreover, the asymptotic covariance matrix of

$\sqrt{T} \left[\hat{\varrho}(h) - \sum_{i=o}^{p} \varphi_i \hat{\varrho}(h-i) \right]_{q+1 \leq h \leq q+r}$ is equal to the covariance matrix

of $(Y_t)_{q+1 \leq t \leq q+r}$, where $(Y_t)_{t \in Z}$ is a stationary process whose spectral density is defined by :

$$f_Y(\omega) = \frac{1}{2\pi} \sum_{m=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} \sum_{i=o}^{p} \sum_{j=o}^{p} \varphi_i \varphi_j \varrho(l)\varrho(l+m+i-j) e^{-\alpha\omega m}$$

with $\alpha^2 = -1$.

If we note that :

$$\frac{f_X(\omega)}{\gamma(o)} = \frac{1}{2\pi} \sum_{l=-\infty}^{+\infty} \varrho(l) e^{-\alpha l \omega} \quad ,$$

we obtain, by using the following factorisation :

$$e^{-\alpha\omega m} = e^{-\alpha i\omega} e^{\alpha j\omega} e^{\alpha\omega(l+i-m-j)} e^{-\alpha\omega l} \quad ,$$

a factorisation of $f_Y(\omega)$:

$$f_Y(\omega) = \phi(e^{\alpha\omega})\,\bar{\phi}(e^{\alpha\omega})\,\frac{f_X(\omega)}{\gamma(o)}\,\frac{\bar{f}_X(\omega)}{\gamma(o)}\,.\,2\pi$$

But, it is known that :

$$f_X(\omega) = \frac{\sigma^2}{2\pi}\,\frac{|\,\theta\,(e^{\alpha\omega})\,|^2}{|\,\phi\,(e^{\alpha\omega})\,|^2}\,,$$

where σ^2 is the variance of the white noise innovation process ε_t .

Hence :

$$f_Y(\omega) = \frac{\sigma^4}{2\pi}\,\frac{1}{[\gamma(o)]^2}\,\frac{|\,\theta\,(e^{\alpha\omega})\,|^4}{|\,\phi\,(e^{\alpha\omega})\,|^2}\,,$$

which proves that $(Y_t)_{t\in Z}$ has the following ARMA(p,q^2) representation :

$$\phi(L)\,Y_t = [\,\theta\,(L)\,]^2\,\eta_t\,,$$

with $V\,\eta_t = \dfrac{\sigma^4}{[\gamma(o)]^2}$.

REFERENCES

[1] ANDERSON, T.W. : The statistical analysis of time series, New York, Wiley, 1971.

[2] BEGUIN, J.M., C. GOURIEROUX and A. MONFORT : "Identification of a mixed autoregressive moving average process : The corner method, in Time Series, edited by O.D. ANDERSON, North-Holland 1980, 423-436.

[3] BERLINET, A. : Estimating the degrees of an ARMA model, Compstat Lectures, 1985.

[4] BOX, G.E. and G.M. JENKINS : "Time series analysis, Forecasting and control", San Francisco, Holden Day, 1970.

[5] BOX, G.E. and D. PIERCE : "Distribution of residual autocorrelation in autoregressive integrated moving average time series models", JASA 65, 1970, 1509-1529.

[6] DAVIDSON J., D. HENDRY, F. SRBA and S. YEO : "Econometric modelling of the aggregate time series relationship between consumers'expenditure and income in the U.K.", Economic Journal 88, 19 , 661-692.

[7] GLASBEY, C.A. : "A generalisation of partial autocorrelations useful in identification of ARMA models", Technometrics, 24, 1982, 223-228.

[8] GOURIEROUX, C. and A. MONFORT : "A general theory for asymptotic tests", Document de travail INSEE, 1985.

[9] GRANGER, C.W. : "Some properties of time series data and their use in econometric model specification", Journal of Econometrics 16, 1981, 121-136.

[10] GRAY, H.L., G.D. KELLEY and D.D. McINTIRE : "A new approach to ARMA modelling", Commun. Statist. Simul. Comp. B7, 1978, 1-77.

[11] PESARAN, M.H. : "Identification of rational expectations models", Journal of Econometrics 16, 1981, 375-398.

[12] RAO, C.R. and S.K. MITRA : Generalized inverses of matrices and its applications, Wiley, New-York, 1982.

[13] SARGAN, J.D. : "Some tests of dynamic specification for a single equation", Econometrica 48, 1980, 879-898.

[14] SARGAN, J.D. and F. MEHTA : "A generalization of the Durbin significance test and its application to dynamic specification", Econometrica 51, 1983, 1551-1568.

[15] TUAN, P.D. : "A note on some statistics useful in identifying the order of autoregressive moving average model", Journal of Time Series Analysis 5, 1984, 273-279.

[16] WALLIS, K. : "Econometric implications of the rational expectations hypothesis", Econometrica 48, 1980, 49-73.