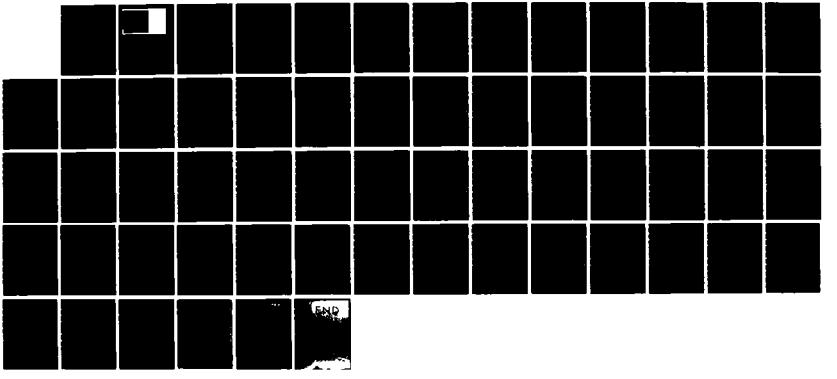
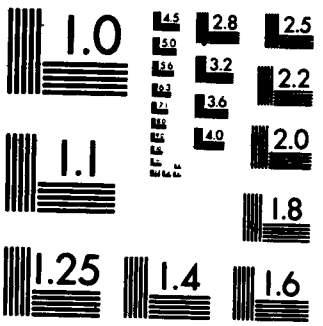
AD-A147 495 A FACTOR ANALYSIS FOR TIME SERIES(U) WISCONSIN 1/1
UNIV-MADISON MATHEMATICS RESEARCH CENTER D PENA ET AL.
JUL 84 MRC-TSR-2722 DAAG29-80-C-0041

UNCLASSIFIED F/G 12/1 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3

AD-A147 495

MRC Technical Summary Report #2722

A FACTOR ANALYSIS FOR TIME SERIES

Daniel Peña and George E. P. Box

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

DTIC FILE COPY

July 1984

(Received May 7, 1984)

Approved for public release
Distribution unlimited

DTIC
ELECTE
NOV 0 9 1984
E

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

84 11 06 052

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

A FACTOR ANALYSIS FOR TIME SERIES

Daniel Peña* and George E. P. Box

Technical Summary Report #2722
July 1984

ABSTRACT

This paper studies how to identify hidden factors in multivariate time series process. It is shown that the number of factors must be equal to the rank of both the covariance matrices and the parameter matrices of the infinite moving average representation of the process. A canonical transformation is derived which can recover such factors. The method is illustrated with several examples. Originator-supplied Keywords include:

AMS (MOS) Subject Classifications: 62M10, 62H25, 60G35

Key Words: Multivariate ARMA process, canonical analysis, eigenvalues, and eigenvectors

Work Unit Number 4 (Statistics and Probability)

*Statistics Department, ETSII, Universidad Politécnica de Madrid, Spain.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. The first author acknowledges support from the United States-Spanish Joint Committee for Educational and Cultural Affairs.

A

SIGNIFICANCE AND EXPLANATION

A central problem in the study of multivariate data is the reduction of dimensionality. This problem is specially acute in modeling multiple time series because if the number of series is large, a huge number of parameters may be needed to obtain an adequate representation of the behavior of the process. This paper explores ways to reduce the dimensionality of the observed process through the extension of the static factor analysis model to the dynamic context. It is shown how the number of factors can be identified, how the loading matrix can be estimated and how a canonical transformation can be built to recover the factors and to obtain a simpler representation of the process. The usefulness of this theory is illustrated with several examples.

Accession For	
NTIS GRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.

A FACTOR ANALYSIS FOR TIME SERIES

Daniel Peña* and George E. P. Box

INTRODUCTION

A central problem in the study of multivariate observations is the reduction of dimensionality. In the time domain study of vectors of time series, Quenoulli (1957) suggested how to use the eigenvectors of the parameter matrices of an autoregressive process to obtain a simpler interpretation and the system, Box and Tiao (1977) showed how to build a canonical transformation of an autoregressive vector process which order the components from least to most predictable, and Reinsel (1983) presented how to estimate index variables to simplify the representation of a vector autoregressive time series using a formulation previously suggested by Sargent and Sims (1977).

Brillinger (1981) extended the standard principal component approach to the frequency domain. Related work is due to Priestley (1981) and Subba Rao (1975). Geweke (1977) and Geweke and Singleton (1981) discussed a frequency domain version of factor analysis. Finally, Hannan (1981) has presented methods to estimate the dimension of a linear system.

In this paper a factor analysis of time series in the time domain is proposed. It is assumed that an observed k-vector of time series z_t could be written as

$$\underset{\sim}{z}_t = \underset{\sim}{P}\underset{\sim}{y}_t + \underset{\sim}{\varepsilon}_t \qquad (1.1)$$

where $\underset{\sim}{P}$ is a $k \times r$ matrix of unknown parameters, $\underset{\sim}{y}_t$ is an unobservable r-dimension vector process and $\underset{\sim}{\varepsilon}_t$ is a k-dimensional white noise sequence with full rank covariance matrix Σ_ε. This model is of course only relevant either if $r < k$ or if $r = k$ but $\Sigma_\varepsilon = 0$. In the first case a reduction of dimensionality could be achieved without loss of information. In the second case, $r = k$, an interesting problem is to find out a linear transformation of the series that allows a simpler representation of the system.

*Statistics Department, ETSII, Universidad Politécnica de Madrid, Spain.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. The first author acknowledges support from the United States-Spanish Joint Committee for Educational and Cultural Affairs.

In Section 2 the general model (1.1) is analyzed and some of its properties are studied. A canonical transformation can then be introduced to recover the factors. In Section 3 the one factor model is studied in detail, and in Section 4 a practical methodology to apply these procedures is developed. Section 5 includes several examples of the application of the suggested models, and in Section 6 this methodology is compared with related approaches in the time domain. Finally, Section 7 contains some concluding remarks.

2. THE GENERAL UNCOUPLE FACTORS MODEL

2.1 Formulation

Let l_t be a vector of k time series and $z_t = l_t - \mu_z$ the vector of deviations from some origin μ_z that will be the mean if the series are stationary. We assume that these series are generated by $r (r \leq k)$ factors, y_t, plus a measurement error ε_t as

$$\underline{z}_t = \underline{P}\underline{y}_t + \underline{\varepsilon}_t \tag{2.1}$$

where $\underline{P}$ is a $k \times r$ matrix of parameters of rank r and $\underline{\varepsilon}_t$ is a white noise sequence with full rank covariance matrix $\underline{\Sigma}_\varepsilon$. The vector $\underline{y}_t$ follows a r-dimensional ARMA (p_y, q_y) process of the form:

$$\underline{\phi}_y(B)\underline{y}_t = \underline{\theta}_y(B)\underline{a}_t \tag{2.2}$$

where

$$\underline{\phi}_y(B) = \underline{I} - \underline{\phi}_y(1)B - \ldots - \underline{\phi}_y(P)B^P$$

$$\underline{\theta}_y(B) = \underline{I} - \underline{\theta}_y(1)B - \ldots - \underline{\theta}_y(q)B^q$$

are matrix polynomials in the backshift operator B, the $\underline{\phi}$'s and the $\underline{\theta}$'s are $r \times r$ matrices and the roots of the determinantal polynomial $|\underline{\phi}_y(B)|$ are on or outside the unit circle, whereas those of $|\underline{\theta}_y(B)|$ are all outside the unit circle. Also, $\{\underline{a}_t\}$ is a sequence of vector Gaussian white noise with zero mean and covariance matrix $\underline{\Sigma}_a$. We shall assume in this section that the r factors are independent, and all the $\underline{\phi}$ and $\underline{\theta}$ matrices will be diagonal. However, we will allow contemporaneous dependency in the noise matrix Σ_a and assume only that is positive definite. In Section 4 this model will be generalized to the case in which the factors have dynamic dependence and the parameter matrices $\underline{\phi}_y$ and $\underline{\theta}_y$ are no longer diagonal.

The matrix $\underline{P}$ will be called the factor loadings or factor weights matrix, and its elements, p_{ij}, represent the weight of the factor j in the observed ith component. If $\underline{C}$ is any $r \times r$ nonsingular matrix, the generating equation could also be written

$$\underline{z}_t = \underline{P}^*\underline{y}_t^* + \underline{\varepsilon}_t$$

where $\underline{P}^* = \underline{P}\underline{C}^{-1}$ is the new rectangular matrix of coefficients and $\underline{y}_t^* = \underline{C}\underline{y}_t$ is a linear transformation of the factors. Multiplying (2.2) by $\underline{C}$:

$$\underline{\underline{C}}\underline{\underline{\phi}}(B)\underline{\underline{C}}^{-1}\underline{\underline{C}}\underline{y}_t = \underline{\underline{C}}\underline{\underline{\theta}}(B)\underline{\underline{C}}^{-1}\underline{\underline{C}}\underline{a}_t$$

$$\underline{\underline{\phi}}^*(B)\underline{y}_t^* = \underline{\underline{\theta}}^*(B)\underline{a}_t^*$$

and so, the model for the new set of factors is again an r-dimensional ARMA (p_y, q_y) model with parameters

$$\underline{\underline{\phi}}^*(B) = \underline{\underline{C}}\underline{\underline{\phi}}(B)\underline{\underline{C}}^{-1}$$

$$\underline{\underline{\theta}}^*(B) = \underline{\underline{C}}\underline{\underline{\theta}}(B)\underline{\underline{C}}^{-1}$$

$$\underline{\underline{\Sigma}}_a^* = \underline{\underline{C}}\underline{\underline{\Sigma}}_a\underline{\underline{C}}'$$

In general, the restriction that the $\underline{\underline{\phi}}$ and $\underline{\underline{\theta}}$ parameter matrices of the components are diagonal imposes a uniqueness condition. However, if $\underline{\underline{\phi}}(B) = \underline{\underline{I}} - \underline{\underline{I}}B$ and $\underline{\underline{\theta}}(B) = \underline{\underline{I}}$, that is, the factors follow and nonstationary AR(1) with all the roots in the unit circle, some indeterminacy appears, because whatever the $\underline{\underline{C}}$ matrix the new factor will be uncorrelated at all lags. A possibility is to choose $\underline{\underline{C}}$ to diagonalize $\underline{\underline{P}}'\underline{\underline{P}}$ and $\underline{\underline{\Sigma}}_a$ at the same time. In this way, not only the factors are uncorrelated to all lags and contemporaneously, but also the columns that transmit their effect to the observed series are orthogonal.

However, if $\underline{\underline{C}}$ is diagonal, that is can be interpreted as a change of scale, the transformed parameter matrices, $\underline{\underline{\phi}}^*(B)$ and $\underline{\underline{\theta}}^*(B)$ are always diagonal and $\underline{\underline{\Sigma}}_a^*$ keeps its basic features. To remove this source of indeterminacy we assume that the columns of the $\underline{\underline{P}}$ matrix of factor loadings are such that

$$\sum_{i=1}^{k} p_{ij}^2 = 1$$

The objective of the analysis will be to estimate these factor loadings and to build a canonical transformation to recuperate the hidden factors.

2.2 The Covariance and Correlation Matrices of the Observed Process

Let us call $\Gamma_z(k) = E[\underline{z}_{t-k}\underline{z}_t']$ the covariance matrices of the process $\underline{z}_t$ and $\underline{\underline{\Gamma}}_y(k)$ the covariance matrices for the generating vector $\underline{y}_t$. Then:

$$\underline{\underline{\Gamma}}_z(0) = \underline{\underline{P}}\underline{\underline{\Gamma}}_y(0)\underline{\underline{P}}' + \underline{\underline{\Sigma}}_\varepsilon \qquad (2.3)$$

$$\Gamma_z(k) = P\Gamma_y(k)P' \qquad k > 1 \tag{2.4}$$

and the rank of $\Gamma_z(k)$ for $k > 1$ will be equal to the number r of common factors. Also, if the factors are independent for all lags, and the matrix Γ_a is diagonal, then all the covariance matrices $\Gamma_y(k)$ will be diagonal and, hence: (1) the matrices $\Gamma_z(k)$ will be symmetric for $k > 1$, (2) the columns of P will be eigenvectors of $\Gamma_z(k)$ with eigenvalues $\gamma_i(k)$, where $\gamma_i(k)$ are the diagonal elements of $\Gamma_y(k)$.

The partial autocorrelation matrices will be given by (see Tiao and Box (1981))

$$P'(\ell) = \begin{cases} \Gamma_z^{-1}(0)\Gamma_z(1) & \ell = 1 \\ [\Gamma_z(0) - b'(\ell)A^{-1}(\ell)b(\ell)]^{-1} & \\ [\Gamma_z(\ell) - b'(\ell)A^{-1}(\ell)\ \ (\ell)] & \ell > 1 \end{cases}$$

where

$$b'(\ell) = [\Gamma_z'(\ell-1)\Gamma_z'(\ell-2) \ldots \Gamma_z'(1)]$$

$$c'(\ell) = [\Gamma_z'(1)\Gamma_z'(2) \ldots \Gamma'(\ell-1)]$$

$$A(\ell) = \begin{bmatrix} \Gamma(0) & \Gamma'(1) \ldots \Gamma'(\ell-2) \\ \Gamma(1) & \Gamma(0) \\ & \ddots \\ \Gamma(\ell-2) & \Gamma(0) \end{bmatrix}$$

and so, the partial autocorrelation matrices can be written

$$P'(\ell) = B(\ell)[\Gamma_z(\ell) - \sum_{i=1}^{\ell-1}\sum_{j=1}^{\ell-1} \Gamma_z'(\ell-i)A^{-1}(i,j,\ell)\Gamma_z(j)]$$

where $B(\ell) = [\Gamma_z(0) - b'(\ell)A^{-1}(\ell)b(\ell)]^{-1}$ and $A^{-1}(i,j,\ell)$ is the (i,j) element of $A^{-1}(\ell)$. Then

$$P'(\ell) = B(\ell)P[\Gamma_y(\ell) - \sum_{i=1}^{\ell-1}\sum_{j=1}^{\ell-1} \Gamma_y'(\ell-1)P'A^{-1}(i,j,\ell)P\Gamma_y(j)]P' \tag{2.5}$$

Now, if we call v the null space of P defined by all the vectors that verify $P'v = 0$,

and if $\underline{V}$ is a $k \times (k - r)$ matrix which columns are a basis for this space ν, then

$$\underline{P}'(\ell)\underline{V} = \underline{H} \cdot \underline{P}'\underline{V} = 0$$

this shows that (1) the rank of $\underline{P}'(\ell)$ is at most r, and (2) the eigenvectors linked to zero eigenvalues of $\underline{P}'(\ell)$ belong to the ν space.

As an example, and for further reference, we derive the first partial autocorrelation matrix assuming that the factors follow an AR(1) process with parameter matrix $\underline{\phi}_y$.

$$\underline{P}(1) = \underline{\Gamma}'_z(1)\underline{\Gamma}_z^{-1}(0) = \underline{P}\underline{\Gamma}'_y(1)\underline{P}'(\underline{P}\underline{\Gamma}_y(0)\underline{P}' + \underline{\Sigma}_\epsilon)^{-1}$$

and using the relations

$$(\underline{\Sigma}_\epsilon + \underline{P}\underline{\Gamma}_0(y)\underline{P}')^{-1} = \underline{\Sigma}_\epsilon^{-1} - \underline{\Sigma}_\epsilon^{-1}\underline{P}(\underline{I} + \underline{\Gamma}_y(0)\underline{P}'\underline{\Sigma}_\epsilon^{-1}\underline{P})^{-1}\underline{\Gamma}_y(0)\underline{P}'\underline{\Sigma}_\epsilon^{-1}$$

and

$$\underline{\phi}_y = \underline{\Gamma}'_y(1)\underline{\Gamma}_y^{-1}(0)$$

and calling $\underline{A} = \underline{P}'\underline{\Sigma}_\epsilon^{-1}\underline{P}$, we obtain

$$\underline{P}(1) = \underline{P}\underline{\phi}_y\underline{\Gamma}_y(0)[\underline{I} - \underline{A}(\underline{I} + \underline{\Gamma}_y(0)\underline{A})^{-1}\underline{\Gamma}_y(0)]\underline{P}'\underline{\Sigma}_\epsilon^{-1}$$

now, using the relations (Handerson and Searle (1981))

$$(\underline{I} + \underline{P})^{-1} = \underline{I} - \underline{P}(\underline{I} + \underline{P})^{-1} = \underline{I} - (\underline{I} + \underline{P})^{-1}\underline{P}$$

$$(\underline{I} + \underline{P}\underline{Q})^{-1}\underline{P} = \underline{P}(\underline{I} + \underline{Q}\underline{P})^{-1}$$

we obtain

$$\underline{P}(1) = \underline{P}\underline{\phi}_y\underline{\Gamma}_y(0)(\underline{I} + \underline{A}\underline{\Gamma}_y(0))^{-1}\underline{P}'\underline{\Sigma}_\epsilon^{-1} . \qquad (2.6)$$

The matrix $\underline{A}\underline{\Gamma}_y(0)$ can be considered as a measure of the ratio between the signal introduced by the factors in the observed series and the noise due to $\underline{\epsilon}_t$. For instance, in the one factor case, assuming $\underline{\Sigma}_\epsilon = \sigma_\epsilon^2\underline{I}$, then

$$\underline{A}\underline{\Gamma}_y(0) = \frac{\Sigma p_i^2 \gamma_y(0)}{\sigma_\epsilon^2}$$

where p_i are the components of the vector $\underline{P}$ and $\gamma_y(0)$ the variance of the factor. The above expression shows that $\underline{A}\underline{\Gamma}_y(0)$ compares the variance of every component of $\underline{z}_t$ due to the factor, $p_i^2\gamma_y(0)$, with the variance introduced by the error, σ_ϵ^2.

In the general case, if this matrix is much greater than $\underline{I}$ (that is, the signal-noise ratio is larger than one) then

$$(\underline{I} + \underline{A}\Gamma_y(0))^{-1}\underline{A}\Gamma_y(0) \approx \underline{I}$$

and if we multiply (2.6) by $\underline{P}$

$$\underline{P}(1)\underline{P} = \underline{P}\underline{\phi}_y\Gamma_y(0)(\underline{I} + \underline{A}\Gamma_y(0))^{-1}\underline{A} \approx \underline{P}\underline{\phi}_y$$

and so, approximately, $P(1)$ has eigenvalues equal to the diagonal elements of ϕ_y with corresponding eigenvectors the columns of P.

2.3 The Model for the Observed Process and Its Parameters

The properties we have studied for the covariance and correlation matrices impose strong restriction on the type of model the observed z_t vector of series can follow, and to the characteristics of its parameter values.

Theorem 1. Suppose $z_t = Py_t + \varepsilon_t$ where y_t is a r-dimensional ARMA (p_y,q_y) process, $\underline{P}$ is a $k \times r$ matrix $(k > r)$ of rank r and ε_t is a k-dimensional white noise sequence with covariance matrix Γ_ε. Then, z follow a k-dimensional ARMA (p_z,q_z) with $p_z = p_y$, $q_z = \max(p_y,q_y)$.

Proof.

The covariance matrices of an ARMA (p,q) model satisfy the equation (see Tiao and Box (1981)):

$$\Gamma_y(\ell) = \begin{cases} \sum_{j=\ell-m}^{\ell-1} \Gamma_y(j)\underline{\phi}_y'(\ell - j) - \sum_{j=0}^{m-\ell} \underline{\psi}_y(j)\Gamma_\varepsilon\underline{\theta}_y'(j + \ell) & \ell = 0,\ldots,m \\ \sum_{j=1}^{m} \Gamma_y(\ell - j)\underline{\phi}_y'(\ell - j) & \ell > m \end{cases}$$

where $m = \max(p_y,q_y)$, the ψ's are obtained from the relationship $\underline{\psi}(B) = \underline{\phi}^{-1}(B)\underline{\theta}(B)$, and $\underline{\theta}_0 = -\underline{I}$. It is assumed that (a) if $p < q$, $\underline{\phi}(p + 1) = \ldots = \underline{\phi}(m) = 0$, and (b) if $q < p$, $\underline{\theta}(q + 1) = \underline{\theta}(m) = 0$. Premultiplying the system of equations by $\underline{P}$ and postmultiplying by $\underline{P}'$:

$$\Gamma_z(\ell) = \begin{cases} \sum_{j=-m}^{-1} \underline{P}\Gamma_y(j)\underline{\phi}_y'(\ell - j)\underline{P}' - \sum_{j=0}^{m} \underline{P}\Psi_y(j)\Sigma_a\underline{\theta}_y'(j)\underline{P}' + \Sigma_\varepsilon & \ell = 0 \\ \sum_{\ell-m}^{\ell-1} \underline{P}\Gamma_y(j)\underline{\phi}_y'(\ell - j)\underline{P}' - \sum_{j=0}^{m-\ell} \underline{P}\Psi_y(j)\Sigma_a\underline{\theta}_y'(j + \ell)\underline{P}' & \ell = 1,\ldots m \\ \sum_{j=1}^{m} \underline{P}\Gamma_y(\ell - j)\underline{\phi}_y'(\ell - j)\underline{P}' & \ell > m \end{cases} \quad (2.7)$$

For $\ell > m$, equating the precedent result to the general expression for the covariances of an ARMA model

$$\sum_{j=1}^{m} \underline{P}\Gamma_y(\ell - j)\underline{\phi}_y'(\ell - j)\underline{P}' = \sum_{j=1}^{p_z} \Gamma_z(\ell - j)\underline{\phi}_z'(\ell - j) \qquad \ell = m + 1,\ldots,$$

and so, using (2.4), $p_z = p_y$ and:

$$\underline{\phi}_y'(k)\underline{P}' = \underline{P}'\underline{\phi}_z'(k) \qquad k > 1 \quad (2.8)$$

and substituting (2.8) into (2.7)

$$\Gamma_z(\ell) = \begin{cases} \sum_{j=-m}^{-1} \Gamma_z(j)\underline{\phi}_z'(\ell - j) - \sum_{j=0}^{m} \underline{P}\Psi_y(j)\Sigma_a\underline{\theta}_y'(j)\underline{P}' + \Sigma_\varepsilon & \ell = 0 \quad (2.9) \\ \sum_{j=\ell-m}^{\ell-1} \Gamma_z(j)\underline{\phi}_z'(\ell - j) - \Sigma_\varepsilon\underline{\phi}_z'(\ell) - \sum_{j=0}^{m-\ell} \underline{P}\Psi_y(j)\Sigma_a\underline{\theta}_y'(j + \ell)\underline{P}' & \ell = 1,\ldots,m \quad (2.10) \\ \sum_{j=1}^{m} \Gamma_z(\ell - j)\underline{\phi}_z'(\ell - j) & \ell > m \quad (2.11) \end{cases}$$

If the order of the AR component of $\underline{y}_t$, p_y, is zero, then there will be exactly q_y, $\Gamma_z(\ell)$ covariance matrices different from zero, and the observed process $\underline{z}_t$ will follow an MA (q_y) process. On the other hand, if $q_y = 0$, the system of equations reduce to

$$\underline{\Gamma}_z(0) = \sum_{-m}^{-1} \underline{\Gamma}_z(j)\underline{\phi}_z'(-j) + \underline{\Sigma}_\varepsilon + \underline{P}\underline{\Sigma}_a\underline{P}'$$

$$\underline{\Gamma}_z(\ell) = \sum_{j=\ell-m}^{\ell-1} \underline{\Gamma}_z(j)\underline{\phi}_z'(\ell - j) - \underline{\Sigma}_\varepsilon\underline{\phi}_z'(\ell) \qquad \ell = 1,\ldots,P$$

$$\underline{\Gamma}_z(\ell) = \sum_{j=1}^{m} \underline{\Gamma}_z(\ell - j)\underline{\phi}_z'(\ell - j) \qquad \ell > P$$

and writing

$$\underline{\Sigma}_\varepsilon + \underline{P}\underline{\Sigma}_a\underline{P}' = \sum_{j=0}^{m} \underline{\psi}_z(j)\underline{\Sigma}_u\underline{\theta}_z'(j)$$

$$\underline{\Sigma}_\varepsilon\underline{\phi}_z'(\ell) = \sum_{j=0}^{m-\ell} \underline{\psi}_z(j)\underline{\Sigma}_u\underline{\theta}_z'(j + \ell) \qquad \ell = 1,\ldots,P$$

the process will follow an ARMA (P,P) model. In the general case, if $p_y > q_y$, $m = p_y$, the first p_y covariance matrices will be given by (2.9) and (2.10) and so they will not follow the AR pattern and an MA(p) component will be needed in addition to the AR(p) linked to the pattern of (2.12). So, the required model will be ARMA (p,p). On the other hand, if $q_y > p_y$ the first q_y matrices will follow the nonsystematic structure and the process will be ARMA (p_y,q_y).

The representation obtained for z_t is not unique, as shown in the following:

Theorem 2. If $\underline{z}_t$ is an observed vector time series generated by r common factors as in (2.1) and which follows model (2.2), then $\underline{z}_t$ could be equally represented as $\underline{\phi}^*(B)\underline{z}_t = \underline{\theta}^*(B)\underline{a}_t$ with new parameters $\underline{\phi}_\ell^* = \underline{\phi}_\ell + \underline{A}$, $\underline{\theta}_\ell^* = \underline{\theta}_\ell + \underline{A}$, where $\underline{A}$ is a nonnull matrix of rank = k - r.

Proof.

Let $m = \max(p,q)$, then, assuming $\theta(q + 1) = \ldots = \theta(m) = 0$ if $m > q$ and $\phi(p + 1) = \ldots = \phi(m) = 0$ if $m > p$:

$$\underline{\theta}^*(B) = (\underline{I} - (\underline{\theta}_1 + \underline{A})B - \ldots - (\underline{\theta}_m + \underline{A})B^m) = \underline{\theta}(B) + \underline{A}\underline{I}_m(B)$$

$$\underline{\phi}^*(B) = \underline{\phi}(B) + \underline{A}\underline{I}_m(B)$$

and the condition for the matrices $\theta^*(B)$ and $\phi^*(B)$ to be equally acceptable as

parameter matrices of the process is

$$\underline{\phi}(B)\underline{z}_t + \underline{\Delta}\underline{I}_m(B)\underline{z}_t = \underline{\theta}(B)z_t + \underline{\Delta}\underline{I}_m(B)\underline{a}_t$$

and so, the condition is

$$\underline{\Delta}\underline{z}_t = \underline{\Delta}\underline{a}_t \tag{2.13}$$

and there must be $k - r$ linear combination of the series that are random noise. Suppose that

$$\underline{z}_t = \underline{P}\underline{y}_t + \underline{\varepsilon}_t$$

if we multiply for any matrix $\underline{V}'$ such that $\underline{V}'\underline{P} = 0$, then by Sylvester's law of nullity rank(V) $\leq k - r$, and:

$$\underline{V}'\underline{z}_t = \underline{V}'\underline{\varepsilon}_t \tag{2.14}$$

now we prove that any matrix V' that verifies (2.14) must verify (2.13), as

$$\underline{z}_t = \underline{\phi}(B)^{-1}\underline{\theta}(B)a_t = \underline{\psi}(B)a_t$$

$$\underline{\varepsilon}_t = \underline{\psi}(B)a_t - \underline{P}\underline{y}_t$$

and if $\underline{V}'\underline{P} = 0$

$$\underline{V}'\underline{\varepsilon}_t = \underline{V}'\underline{\psi}(B)\underline{a}_t$$

and, as $\underline{V}'\underline{\varepsilon}_t$ must be a white noise sequence,

$$\underline{V}'\underline{\varepsilon}_t = \underline{V}'\underline{a}_t \tag{2.14}$$

$$\underline{V}'\underline{\psi}_i = \underline{0} \quad \forall\, i > 1 \tag{2.15}$$

and so, if we take $\underline{A} = \underline{V}'$ the proof is complete. Note that there always exist a matrix $\underline{V}'$ such that $\underline{V}'\underline{P} = 0$. We can take, for example, as first $k - r$ rows of $\underline{V}'$ a set of the $k - r$ lineary independent eigenvectors of $\underline{P}\underline{P}'$ that are linked to zero eigenvalues, and put in the leftover r positions any r vectors of the linear subspace generated by the first $k - r$ rows.

An important conclusion from Theorem 1 is that the autoregressive matrices of the multivariate observed process z_t must satisfy (2.8) and so

$$\underline{\phi}_z(k)\underline{P} = \underline{P}\underline{\phi}_y(k) \tag{2.16}$$

which has the general solution

$$\underline{\phi}_z(k) = \underline{P}\underline{\phi}_y(k)\underline{P}^- + \underline{C}(\underline{I} - \underline{P}\underline{P}^-) \tag{2.17}$$

where $\underline{P}^-$ is any generalized inverse of $\underline{P}$ that satisfy $\underline{P}\underline{P}^-\underline{P} = \underline{P}$ and $\underline{C}$ is any

arbitrary matrix with the only restriction that the roots of $|\underline{\phi}_z(B)|$ are on or outside the unit circle. As the matrix $\phi_y(k)$ is diagonal, equation (2.16) shows that the columns of $\underline{P}$ are eigenvectors of $\underline{\phi}_z(k)$, with eigenvalues the diagonal elements of $\underline{\phi}_y(k)$. However, the matrix $\underline{\phi}_z(k)$ can have any rank, due to the presence of the arbitrary matrix $\underline{C}$. On the other hand, according to Theorem 2, for every possible solution $\underline{\phi}_z(k)$ there will be a set of $\underline{\theta}_z(k)$ matrices that must verify restriction (2.15). For example, if $P_y = 1$, $q_y = 0$, and the vector of components follows a multivariate AR(1) process, the system of equations to determine the parameters is obtained as a particular case of (2.10) to (2.12) after some straightforward algebra, to be

$$\underline{\Sigma}_c + \underline{P}\underline{\Sigma}_a\underline{P}' = \underline{\Sigma}_u - \underline{\phi}_z\underline{\Sigma}_u\underline{\theta}_z' + \underline{\theta}_z\underline{\Sigma}_u\underline{\theta}_z' \tag{2.18}$$

$$\underline{\Sigma}_c\underline{\phi}_z' = \underline{\Sigma}_u\underline{\theta}_z' \tag{2.19}$$

$$\underline{\phi}_z\underline{P} = \underline{P}\underline{\phi}_y$$

and any set of matrices $\underline{\Sigma}_u$, $\underline{\phi}_z$, $\underline{\theta}_z$ that satisfy this system is a solution of the system. Although there are infinite solutions, they can be characterized as follows. First, $\underline{\phi}_z$ is given by (2.17):

$$\underline{\phi}_z = \underline{P}\underline{\phi}_y\underline{P}^- + \underline{A}(\underline{I} - \underline{P}\underline{P}^-)$$

and, given ϕ_z, we have in (2.18) $k(k + 1)/2$ equations (because of its symmetry) plus k^2 in (2.19), that is equal to the number of unknown parameters. Furthermore, using result (2.14)

$$\underline{V}'\underline{\psi}_1 = \underline{V}_1'(\underline{\phi}_z - \underline{\theta}_z) = \underline{0} \tag{2.21}$$

which implies that, although $\underline{\phi}_z$ and $\underline{\theta}_z$ can have any rank, their difference must have rank no greater than the number of components in the system. Also, suppose that $\underline{h}$ is an eigenvector of the matrix $\underline{\phi}_z - \underline{\theta}_z$, then $(\phi_z - \theta_z)h = \lambda h$, and assume that λ is nonzero. Then, if $\underline{v}_i'$ is any row of $\underline{V}$ as

$$\underline{v}_i'(\underline{\phi}_z - \underline{\theta}_z)\underline{h} = \lambda\underline{v}_i'\underline{h} = 0$$

and $\underline{h}$ must be orthogonal to the subspace generated by the matrix $\underline{V}$, and, therefore, belongs to the subspace generated by $\underline{P}$.

In summary when building a multivariate time series model for a vector of series z_t that is driven by a number of common factors as in (2.1) it may appear that a complicated

model with all kinds of dependency is needed. In particular we may finish with a model showing feedback among all the components of the vector z_t. However, the parameter matrices in this case must satisfy many restrictions. There may be summarize as follows: (1) all the autoregressive matrices will have r common eigenvectors that define the subspace $\underline{P}$ of the factors and (2) all the $\underline{\psi}$ matrices must have rank equal to r and their columns must belong to the subspace generated by $\underline{P}$.

2.4 A Canonical Transformation

An important problem is to find a transformation of the observed series z_t that allows us to recover the factors. To explain how to find such a transformation first, note that, if $\underline{P}^-$ is any generalized inverse of $\underline{P}$, if

$$\underline{z}_t = \underline{P}\underline{y}_t + \underline{\varepsilon}_t$$

then

$$\underline{y}_t = \underline{P}^-\underline{z}_t - \underline{P}^-\underline{\varepsilon}_t$$

second, suppose now that we apply a linear transformation $\underline{M}$ to $\underline{z}_t$ where $\underline{M} = \begin{bmatrix} P^- \\ B \end{bmatrix}$, then

$$\underline{x}_t = \underline{M}\underline{z}_t = \begin{bmatrix} \underline{P}^-\underline{z}_t \\ \underline{B}\underline{z}_t \end{bmatrix} = \begin{bmatrix} \underline{y} + \underline{P}^-\underline{\varepsilon}_t \\ \underline{BP}\underline{y}_t + \underline{B}\underline{\varepsilon}_t \end{bmatrix} = \begin{bmatrix} \underline{x}_{1t} \\ \underline{x}_{2t} \end{bmatrix}$$

To recover the r factors $\underline{y}_t$, $\underline{B}$ should be chosen such that $BP = 0$. Then, the first r components of x_t will be equal to the vector $\underline{y}_t$ of common factors plus some added noise, and the second $k - r$ components will be just white noise. A matrix B that satisfies this conditions may be obtained putting as rows of B the $k - r$ eigenvectors linked to zero eigenvalues of $\underline{PP}'$. Calling $\underline{v}_i$ this eigenvectors:

$$\underline{PP}'[v_1,\ldots,v_{k-r}] = \underline{0} = \underline{PP}'\underline{V}$$

that implies,

$$\underline{V}'\underline{P} = \underline{0}$$

and so $\underline{B} = \underline{V}'$ has the desired property. Although the selection of P^- could be arbitrary, it seems sensible to choose a generalized inverse that leads to canonical

variables as simple as possible. A convenient property is that the $\underline{x}_1$ and $\underline{x}_2$ components of the new vector be independent, but this is not in general possible unless $\Sigma_\varepsilon = I$. Assuming this equality, then:

$$E[\underline{x}_{1t}\underline{x}'_{2t}] = \underline{P}^{-}\underline{\Sigma}_\varepsilon\underline{R}' = \underline{P}^{-}\underline{R}'$$

and if we choose $\underline{P}^{-}$ as the Moore-Penroe generalized inverse of $\underline{R}$, given by $(\underline{R}'\underline{R})^{-1}\underline{R}'$, the covariance matrix will be zero and the components will be independent.

To illustrate this transformation let us suppose that $\underline{y}_t$ follows an r-dimensional AR(1) process. Then $\underline{z}_t$ is ARMA(1,1) and so is any linear transformation of $\underline{z}_t$. However, the parameters of the k-dimensional $\underline{x}_t$ process given by $\underline{x}_t = \underline{M}\underline{z}_t$, where:

$$\underline{M} = \left[\begin{array}{c} (\underline{R}'\underline{R})^{-1}\underline{R}' \\ - - - - \\ \underline{V}' \end{array}\right]$$

will be

$$\underline{\Phi}_x = \underline{M}\underline{\Phi}_z\underline{M}^{-1}$$

$$\underline{\Theta}_x = \underline{M}\underline{\Theta}_z\underline{M}^{-1}$$

where:

$$\underline{M}^{-1} = [\underline{R} \mid \underline{V}]$$

as can be verified by direct multiplication $\underline{M}\underline{M}^{-1} = \underline{M}^{-1}\underline{M} = \underline{I}$. So,

$$\underline{\Phi}_x = \left|\begin{array}{cc} (\underline{R}'\underline{R})^{-1}\underline{R}'\underline{\Phi}_z\underline{R} & (\underline{R}'\underline{R})^{-1}\underline{R}'\underline{\Phi}_z\underline{V} \\ \underline{V}'\underline{\Phi}_z\underline{R} & \underline{V}'\underline{\Phi}_z\underline{V} \end{array}\right|$$

and using the general expression (2.20) for $\underline{\Phi}_z$

$$\underline{\Phi}_x = \left|\begin{array}{cc} \underline{\Phi}_y & 0 \\ 0 & \underline{V}'\underline{A}\underline{V} \end{array}\right|$$

in the same way, calling $\underline{\Theta}_0$ the value of $\underline{\Theta}_z$ associated with $\underline{\Phi}_z = \underline{P}\underline{\Phi}_y\underline{P}^{-}$, when $\underline{A} = \underline{0}$, then, for Theorem 2, the general expression of $\underline{\Theta}_z$ is

$$\underline{\Theta}_z = \underline{\Theta}_0 + \underline{A}(\underline{I} - \underline{P}\underline{P}^{-})$$

and so

$$\underline{\theta}_x = \begin{vmatrix} (\underline{P}'\underline{P})^{-1}\underline{P}'\underline{\theta}_0\underline{P} & (\underline{P}'\underline{P})^{-1}\underline{P}'\underline{\theta}_g\underline{V} \\ \underline{0} & \underline{V}'\underline{A}\underline{V} \end{vmatrix}$$

where we have used that, for (2.21) $\underline{V}'\underline{\phi}_g = \underline{V}'\underline{\theta}_g$. So, calling $\underline{x}_t' = [\underline{x}_{1t}'/\underline{x}_{2t}]$ and taking into account that the addition of MA process is a new MA process:

$$(\underline{I} - \underline{\phi}_y B)\underline{x}_{1t} = (\underline{I} - \underline{\theta}_x B)\underline{b}_{1t}$$

$$\underline{x}_{2t} = \underline{b}_{2t}$$

that allows us to recover the components mixed up with noise in x_{1t}.

2.5. Some Simulation Results

We present here two simulations to show how the above theory can be expected to work in practice. The results we show are representative of the many simulations we have made.

CASE 1

A sample of 100 observations of a vector of 4 time series was generated according to the model

$$\underline{z}_t = \begin{pmatrix} .8 & .2 \\ 1.1 & .0 \\ .5 & 1.0 \\ .2 & 1.2 \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} + \underline{\varepsilon}_t$$

with $\Sigma_\varepsilon = I$, $\nabla y_t = \underline{a}_t$ and $\Sigma_a = \begin{pmatrix} 2 & 0 \\ 0 & 1.5 \end{pmatrix}$ Table 1a shows the eigenvalues of the covariance and correlation matrices and the eigenvectors linked to the nonzero eigenvalues. To explain these results, first note that as the two factors are nonstationary their sample variances are much larger than the diagonal elements of Σ_ε and so $\Gamma_z(0)$ has a very similar structure to $\Gamma_z(k)$ for $k > 1$. This means that

$$\underline{\Gamma}_z(\ell) \simeq \underline{P}\underline{\Gamma}_y(\ell)\underline{P}' \qquad \forall\, \ell\ .$$

Second, as the factor are completely independent and with similar variances,

$\underline{\Gamma}_y(\ell) \simeq \gamma_y(\ell)I$, where $\gamma_y(\ell)$ is a covariance scalar function. Then

$$\underline{\Gamma}_y(\ell) \simeq \gamma_y(\ell)\underline{P}\underline{P}' \qquad \forall\, \ell$$

and the eigenvalues of $\underline{\Gamma}_z(\ell)$ are, approximately, proportional to the eigenvalues of $\underline{P}\underline{P}'$ and with the same eigenvectors. The two eigenvectors of $\underline{P}\underline{P}'$ linked to nonzero

eigenvalues are (.42 .72 -.21 -.52) and (.37 .39 .61 .59) that are similar to those found in the covariance matrices.

The rank of the covariance matrices is clearly two and so the number of factors in the system would be estimated correctly.

TABLE 1

a) Largest eigenvalue and eigenvector

lag	0	1	2	3	5
λ	141.6	136.6	132.9	129.7	119.7
vector	.51	.50	.50	.50	.47
	.81	.80	.80	.79	.78
	.00	-.01	-.04	-.06	-.11
	-.28	-.30	-.33	-.35	-.40

covariance

lag	0	1	2	3	5
λ	2.35	2.27	2.21	2.14	1.92
vector	.58	.56	.53	.51	.43
	.61	.59	.57	.55	.49
	.16	-.20	-.26	-.30	-.41
	.51	-.54	-.57	-.60	-.64

correlation

b) Second largest

lag	0	1	2	3	5
λ	36.6	35.6	30.2	28.4	24.2
vector	.19	.19	.18	.17	.16
	.27	.12	.10	.09	.06
	.64	.64	.63	.64	.67
	.73	.73	.74	.75	.77

lag	0	1	2	3	5
λ	1.59	1.41	1.30	1.24	1.10
vector	.35	.29	.27	.23	.16
	.26	.22	.18	.14	.04
	.76	.78	.80	.84	1.02
	.48	.52	.58	.63	.81

c) Third and fourth eigenvalues

lag	0	1	2	3	4
3rd	.89	-.04	-.25	-.04	-.18
4th	.81	.07	-.04	.04	-.04

lag	0	1	2	3	4
3rd	.013	.003	-.009	.0008	-.0025
4th	.04	-.0008	-.0019	-.0001	-.0025

Table 1b shows the eigenvalues and eigenvectors of the first partial autocorrelation matrices. The identification of the rank is here less clear although (1) $P(1)$ is mainly dominated by two large eigenvalues; (2) the eigenvectors linked to these two eigenvalues in $P(1)$ show, roughly, the same structure that the largest two eigenvectors in the

covariance matrices, and (3) $P(2)$ is small and nearly zero. These results suggest that the model for the common factors is probably AR(1) and confirm mildly the hypothesis of two factors leading the system.

TABLE 1b

Partial Autocorrelation Matrices

$P(1)$

.97	.87	.19	.10
.65	.27	-1.39	-.28
.68	.65	.76	.51
.66	-.35	-.89	-.62
.50	-.65	.52	.52

$P(2)$

-.14	.21	.04±.09i
-.57	.30	complex
.62	.60	
.11	-.30	
-.41	-.86	

Note $P(\ell) = 0$ for $\ell > 2$

Table 2 shows the eigenvalues of the estimated parameter matrices and of its differency. Note that, as $\underset{\sim}{\phi}_y = \underset{\sim}{I}$, then

$$\underset{\sim}{\phi}_z \underset{\sim}{P} = \underset{\sim}{PI} = \underset{\sim}{P}$$

and for any nonsingular matrix C

$$\underset{\sim}{\phi}_z \underset{\sim}{PC} = \underset{\sim}{PC}$$

and so, any linear transformation of the columns $\underset{\sim}{P}$ may appear as eigenvalues of $\underset{\sim}{\phi}_z$. The rank of the difference $\underset{\sim}{\phi} - \underset{\sim}{\theta}$ is nearly 2, which confirms the existence of two factors. Note that the two largest eigenvalues of $\underset{\sim}{\phi}_z$ are, in this case, similar to the eigenvalues of the covariance matrices.

Taking this two eigenvalues as factor loadings to build the transformation, the new vector of transformed series will be given by

$$\underset{\sim}{M} = \begin{vmatrix} .74 & -.58 & .24 & -.24 \\ .37 & -.06 & -.74 & .56 \\ .42 & .62 & .01 & -.19 \\ .29 & .36 & .55 & .57 \end{vmatrix}$$

and, using this transformation, the new series $\underset{\sim}{x}_t = \underset{\sim}{M}\underset{\sim}{z}_t$ will have parameters

$$\Phi_x = \begin{vmatrix} . & .18 & . & . \\ .40 & -.70 & . & . \\ .43 & .70 & 1.0 & . \\ . & .80 & . & .98 \end{vmatrix} = \begin{vmatrix} . & . & . & . \\ . & . & . & . \\ . & . & 1.0 & . \\ . & . & . & .98 \end{vmatrix} + \Delta_1$$

$$\theta_x = \begin{vmatrix} . & .13 & . & . \\ .32 & -.85 & . & . \\ .52 & 1.0 & .23 & . \\ . & 1.1 & . & .18 \end{vmatrix} = \begin{vmatrix} . & . & . & . \\ . & . & . & . \\ . & . & .23 & . \\ . & . & . & .18 \end{vmatrix} + \Delta_2$$

TABLE 2

Eigenvalues of the Estimated Parameter Matrices

Φ				θ			$\Phi - \theta$			
1.0	.90	-.77	.20	-.99	-.14	.28±.09i	.96	.73	.25	.06
.64	.16	.02	-.15	.27	.43		.42	.31	-1.15	.71
.97	.15	.41	.96	.58	1.41	complex	.72	.46	0	.67
-.22	.73	1.03	.17	.75	.35		.50	-.42	-.02	-.43
-.67	.87	-.50	-.14	-.12	.16		.22	-.69	.29	.82

with $A_1 = A_2$ has rank equal to two. So we will obtain two random component plus the two stationary factors contaminated with some white noise, as expected.

The main conclusions from this exercise are as follows: (1) Although the determination of the number of factors using the rank of the matrices may be not clear for some matrices, putting together all the different pieces of information that can be obtained, the picture is normally quite clear. (2) The factor loadings can be obtained with reasonable accuracy from the eigenvectors linked to large eigenvalues of the Φ parameter matrices.

CASE 2

The generating equation and the sample size in this second example are the same as in Case 1. However, now only one factor is nonstationary and it has been allowed some correlation among both components. The factor time series model is

$$\left(I - \begin{bmatrix} .6 & 0 \\ 0 & 1.0 \end{bmatrix} B\right) \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} a_{1t} \\ a_{2t} \end{pmatrix}, \quad \Sigma_a = \begin{bmatrix} 2 & 1.5 \\ 1.5 & 3.5 \end{bmatrix}$$

Table 3a displays the eigenstructure of covariance and correlation matrices and Table 3b that of the partial autocorrelation matrices.

TABLE 3a

Eigenvalues and Eigenvectors of Covariance and Correlation Matrices

a) Largest eigenvalue and its eigenvectors

Covariance

lag	0	1	2	3	4	5
λ	54.10	44.63	40.66	35.87	32.46	30.74
	.22	.48	.17	.15	.14	.13
	.15	.08	.05	.02	.02	.02
	.63	.63	.63	.63	.63	.62
	.73	.75	.75	.76	.76	.77

Correlations

0	1	2	3	4	5
2.98	2.2	1.9	1.6	1.4	1.3
.53	.43	.40	.35	.32	.28
.43	.25	.16	.07	.06	.04
.57	.61	.63	.64	.64	.64
.51	.61	.65	.68	.69	.71

b) Second largest eigenvalue and its eigenvector

lag	0	1	2	3	4	5
λ	5.84	3.07	2.11	1.33	.84	.08
	+.55	.49	.51	.54	.35	-.60
	+.78	.80	.77	.77	.86	.65
	-.08	.03	.07	.01	-.27	.32
	-.27	-.34	-.40	-.36	-.25	-.23

0	1	2	3	4	5
.83	.46	.33	.21	.16	.01
.31	.31	.36	.38	.23	.75
.70	.77	.76	.80	.90	-.44
-.42	-.31	-.31	-.31	-.39	-.47
-.48	-.49	-.53	-.50	-.30	-.11

c) Third and fourth eigenvectors

0	1	2	3	4	5
1.2	+.16	-.14	.04	.07+.06i	.17+.18i
+.80	-.13	+.06	-.16	.07-.06i	.17-.18i

0	1	2	3	4	5
.15	.03	-.02	-.01	.01	-.01
.05	-.005	.002	.005	.005	-.04

TABLE 3b

Eigenvalues and Eigenvectors of the Partial Correlation Matrices

P(1)

.88	.46	.28	-.09
.15	1.5	2	+.07
.03	2.4	1.0	-.17
.67	.59	.16	-.93
.82	-.23	0	.30

P(2)

-.22	-.043	.15±.04i
.68	-.29	Complex
-.52	-.23	
.17	-.79	
-.35	1.13	

$(\ell) \approx 0 \quad \ell > 2$

The main conclusions of both autocovariance and autocorrelation matrices are that (1) there are probably two common components; (2) one is much more important than the other; (3) one seems to be nonstationary taking into account the slow decreasing of the largest eigenvalues, whereas the other seems to be stationary.

It is important to see the differences with respect to Case 1. If we compare Table 1a with Table 3a it can be seen that the second largest eigenvalue in Table 1 decreases slowly and the eigenvector is fairly stable even for lags as far as five. In contrast with this, the second eigenvalue of Table 3a decreases quickly, and is clearly zero at lag five, as shown not only but its small value but as well by the change of structure of its linked eigenvector.

The partial autocorrelation matrices show that, as far as the rank of these matrices is concerned, the observed results can differ from theory. The only clear indication in this case is that the general pattern of the eigenvectors linked to the two largest eigenvalues is not in contradiction with the previous results.

Table 4 presents the eigenstructure of the estimated parameter matrices. As in the previous simulation the final estimated model has many significant coefficients both in the autoregressive and moving average matrices, and it seems to represent a quite complex relationship among the series. However, the matrix $\underset{\sim}{\phi} - \underset{\sim}{\theta}$ has two larger eigenvalues similar to those of $P(1)$.

If we accept the hypothesis of two main factors and choose the two largest eigenvectors of $\underset{\sim}{\phi}$ to define the matrix of factor loadings $\underset{\sim}{P}$, the resulting parameter matrices for the $\underset{\sim}{x}_t = \underset{\sim}{M}\underset{\sim}{z}_t$ series can be decomposed as before as the sum of two matrices, one of which cancels in both terms. The simplified parameter matrices are:

$$\phi_x = \begin{vmatrix} . & . & . & . \\ . & . & . & . \\ . & . & .86 & . \\ . & . & . & 1 \end{vmatrix} \qquad \theta_x = \begin{vmatrix} . & . & . & . \\ . & . & . & . \\ . & . & .60 & . \\ . & . & . & .14 \end{vmatrix}$$

showing one nonstationary factor and a second stationary factor.

TABLE 4

ϕ			
1.0	.85	+.57	-.53
.14	1.32	.37	.65
-.03	2.30	.77	-.34
.80	-2.40	-.16	-.57
.94	-2.90	-.48	-.36

θ		
.99	-.83	.33±.08i
.93	-.67	complex
.57	.03	
-.60	.48	
.67	.53	

$\phi - \theta$			
.77	.52	.11	-.32
.26	1.2	.22	.26
.15	1.0	1.0	-.27
.71	.28	.14	-.84
.77	-.12	.08	.37

3. THE ONE FACTOR MODEL

3.1 Introduction

The one common factor model is, on the one hand, important in its own right because it describes the relationship of several measurements over time of the same dynamic variable. On the other hand, this situation is going to appear approximately in the multifactor situation when one of the factors has much greater variance, and for this reason is more important than the others.

We will assume in this section that in the general representation (1.1) $\underline{P}$ is now a $k \times 1$ vector. We will concentrate here in the case that y_t follows an AR(1) process and will analyze the behavior of the system as the parameter ϕ_y of y_t approaches the unit value.

With this set up, every component z_{it} of $\underline{z}_t$ will follow an ARMA (1,1) model with the same parameter ϕ_y, and with parameters θ_i that, as can be seen equating moments must verify:

$$\theta_i^2\phi_y - \theta_i(\mu_i + 1 + \phi_y^2) + \phi_y = 0 \qquad (3.1)$$

where $\mu_i = \dfrac{p_i^2\sigma_a^2}{\sigma_i^2}$, being p_i the ith component of $\underline{P}$, σ_a^2 the variance of the noise in the model for y_t and σ_i^2 the variance of the ith component of the measurement error vector $\underline{\varepsilon}_t$. It can readily be seen from (3.1) that the parameter θ_i in each series depends on μ_i. If $\mu_i \to \infty$, which means that the measurement error can be ignored, θ_i tends towards zero, and if $\mu_i \to 0$ everything is dominated by the noise and $\theta_i \to \phi$, which means that this component will be white noise.

3.2 The Covariance and Correlation Matrices

Calling $\gamma_y(k)$ to the autocovariance coefficient of order k in series y_t,

$$\underline{\underline{\Gamma}}_z(0) = \underline{\underline{P}}\underline{P}'\gamma_y(0) + \underline{\underline{\Sigma}}_\varepsilon$$

$$\underline{\underline{\Gamma}}_z(k) = \underline{\underline{P}}\underline{P}'\gamma_y(k) \qquad k \geq 1$$

and for $k \geq 1$, (1) all the covariance matrices will be symmetric; (2) all will have the rank equal to one; (3) all will have an eigenvector equal to $\underline{P}$ linked to the nonzero eigenvalue; (4) all nonzero eigenvalues will be proportional to the autocovariance or autocorrelation function of y_t.

So, in the special case in which y_t will follow an AR(1) the nonzero eigenvalues of the covariance matrices must decrease exponentially.

The partial autocorrelation matrices can be obtained from equations (2.5) and (2.6). The first term is

$$\underline{\underline{P}}(1) = \phi_y\gamma_y(0)(1 + a\gamma_y(0))^{-1}\underline{\underline{P}}\underline{P}'\underline{\underline{\Sigma}}_\varepsilon^{-1}$$

where now $a = \underline{P}'\underline{\underline{\Sigma}}_\varepsilon^{-1}\underline{P}$, and so

$$\underline{\underline{P}}(1)\underline{P} = \phi_y\left(\frac{a\gamma_y(0)}{1 + a\gamma_y(0)}\right)\underline{P}$$

that shows clearly that if the signal to noise ratio $a\gamma_y(0)$ is much greater than 1, the first partial autocorrelation matrix will have an eigenvector $\underline{P}$ linked to the eigenvalue ϕ_y. Note that in all cases $\underline{P}$ will be an eigenvector of $P(1)$ to difference from the multifactor situation.

For the others terms, using (2.5)

$$\underline{\underline{P}}(\ell) = \underline{\underline{P}}(\gamma_y(\ell) - \sum_{i=1}^{\ell-1}\sum_{j=1}^{\ell-1}\gamma_y(\ell - i)\gamma_y(j)\underline{P}'\underline{\underline{A}}^{-1}(i,j,\ell)\underline{P})\underline{P}'\cdot\underline{\underline{B}}'(\ell)$$

where

$$\underline{\underline{B}}(\ell) = (\gamma_y(0)\underline{\underline{P}}\underline{P}' + \underline{\underline{\Sigma}}_\varepsilon - \sum_{i=1}^{\ell-1}\sum_{j=1}^{\ell-1}\gamma_y(i)\gamma_y(j)\underline{\underline{P}}\underline{P}'\underline{\underline{A}}^{-1}(i,j,-)\underline{\underline{P}}\underline{P}')^{-1}$$

calling

$$d_\ell = \gamma_y(0) - \sum_{i=1}^{\ell-1}\sum_{j=1}^{\ell-1} \gamma_y(i)\gamma_y(j)\underline{P}'\underline{A}^{-1}(i,j,\ell)\underline{P} \qquad (3.2)$$

$$f_\ell = \gamma_y(\ell) - \sum_{i=1}^{\ell-1}\sum_{j=1}^{\ell-1} \gamma_y(\ell - i)\gamma_y(j)\underline{P}'\underline{A}^{-1}(i,j,\ell)\underline{P} \qquad (3.3)$$

then

$$\underline{P}(\ell) = f_\ell \underline{P}\underline{P}'(d_\ell \underline{P}\underline{P}' + \underline{\Sigma}_\varepsilon)^{-1}$$

and using the expression for the inverse of the sum of two matrices one of them being nonsingular (see Section 2)

$$\underline{P}(\ell) = \left(\frac{f\ell}{1 + d_\ell a}\right)\underline{P}\underline{P}'\underline{\Sigma}_\varepsilon^{-1} \qquad (3.4)$$

and so $\underline{P}(\ell)$ will have an eigenvector equal to $\underline{P}$ with eigenvalue $af_\ell(1 + ad_\ell)^{-1}$. It can be seen from (3.2) that, again, if the signal to noise ratio is large, the eigenvalue linked to the $\underline{P}$ eigenvector of $\underline{P}(\ell)$ will follow the partial autocorrelation function of the component y_t.

In particular, if y_t follows an AR(1), $\underline{P}(\ell)$ will be near zero for $\ell > 1$ if the signal to noise ratio is large. Otherwise, $\underline{P}(\ell)$ will be different from zero and will have a nonzero eigenvalue linked to $\underline{P}$.

3.3 The Parameter Matrices

The parameter matrices of the process can be obtained equating the representation of the covariance matrices implied by the one factor model to the covariance matrices for an ARMA (1,1) vector process. Then

$$\underline{\Gamma}_0 = \underline{P}\underline{P}'\gamma_y(0) + \underline{\Sigma}_\varepsilon = \underline{\Gamma}_1\underline{\Phi}' + \underline{\Sigma}_u - \underline{\Phi}\underline{\Sigma}_u\underline{\theta}' + \underline{\theta}\underline{\Sigma}_u\underline{\theta}' \qquad (3.5)$$

$$\underline{\Gamma}_1 = \underline{P}\underline{P}'\gamma_y(1) = \underline{\Gamma}_0\underline{\Phi}' - \underline{\Sigma}_u\underline{\theta}' \qquad (3.6)$$

$$\underline{\Gamma}_k = \underline{P}\underline{P}'\gamma_y(k) = \underline{\Gamma}_{k-1}\underline{\Phi}' \qquad k > 2 \qquad (3.7)$$

where $\underline{\Phi}$, $\underline{\theta}$ and $\underline{\Sigma}_u$ are the matrices that define the observed vector process z_t, and we use ϕ_y and σ_a^2 to represent the parameters of the AR(1) model for the common factor y_t. Using (3.7) for $k > 2$ it is readily seen that

$$\underline{\Phi}\underline{P} = \phi_y \underline{P} \tag{3.8}$$

and so ϕ_y must be an eigenvalue of $\underline{\Phi}$ linked to the eigenvector $\underline{P}$. The general solution of (3.8) is (see Rao and Mitra (1971)):

$$\underline{\Phi} = (\underline{P}'\underline{P})^{-1}\phi_y \underline{P}\underline{P}' + \phi_y \underline{C}(I - (\underline{P}'\underline{P})^{-1}\underline{P}\underline{P}') \tag{3.9}$$

where $\underline{C}$ is any arbitrary matrix with the only restrictions that the eigenvalues of Φ must be on or inside the unit circle. Eliminating the covariances from (3.5) and (3.6):

$$\underline{\Sigma}_u = \underline{\Sigma}_e + \underline{P}\underline{P}'\sigma_a^2 + \underline{\Phi}\underline{\Sigma}_u\underline{\theta}' - \underline{\theta}\underline{\Sigma}_u\underline{\theta}' \tag{3.10}$$

$$\underline{\Sigma}_e\underline{\Phi}' = \underline{\Sigma}_u\underline{\theta}' \tag{3.11}$$

Equations (3.8), (3.10) and (3.11) define the restrictions which must be obeyed by the parameters of the system. On the whole, the system has not an unique solution because there are more unknowns than equations. However, once $\underline{\Phi}$ is assumed fixed according to (3.9) the parameters matrices $\underline{\theta}$ and $\underline{\Sigma}_u$ are determined by (3.10) and (3.11). It can be verified that the solution of these equations is:

$$\underline{\theta} = \underline{\Phi} - \phi_y \frac{\lambda}{1 + \lambda b}\underline{P}\underline{P}'\underline{\Sigma}_e^{-1} \tag{3.12}$$

$$\underline{\Sigma}_u = \lambda \underline{P}\underline{P}' + \underline{\Sigma}_e \tag{3.13}$$

where $\underline{\Phi}$ is given by (3.9), $b = \underline{P}'\underline{\Sigma}_e^{-1}\underline{P}$ and λ verifies:

$$b\lambda^2 - \lambda(b\sigma_a^2 + \phi_y^2 - 1) - \sigma_a^2 = 0 \tag{3.14}$$

and corresponds to the solution with θ having its eigenvalues on or inside the unit circle. Note that if $\phi_y \to 0$, $\lambda \to \sigma_a^2$, as expected. Equations (3.9), (3.12) and (3.13) summarize the relevant properties of the matrices of the process: (1) $\underline{\Phi}$ may have any rank but must have $\underline{P}$ as an eigenvector with eigenvalue ϕ_y; (2) $\underline{\theta}$ may have any rank but again must have $\underline{P}$ as an eigenvector with eigenvalue:

$$\lambda_\theta = \frac{\phi_y}{1 + \lambda b}$$

and using (3.14), the expression of λ_θ as a function of the parameters of the common factor is:

$$\lambda_\theta \phi_y - \lambda_\theta (b\sigma_a^2 + \phi_y^2 + 1) + \phi_y = 0 \qquad (3.15)$$

and if we compared (3.15) to (3.1) is clear that λ_θ will be similar to the univariate parameter of the series, although smaller; (3) $\underline{\underline{\phi}} - \underline{\underline{\theta}}$ must have rank equal to one, and $\underline{P}$ will be the eigenvector linked to its nonzero eigenvalue; (4) The matrix $\underline{\underline{\Sigma}}_u$ will be, in general, a full rank matrix.

3.4 The Canonical Transformation

In the one factor case, the canonical transformation is the matrix of eigenvectors of $\underline{\underline{P}}\underline{\underline{P}}'$. Calling $\underline{\underline{M}}'$ this matrix, the covariances of the transformed process $\underline{x}_t = \underline{\underline{M}}\underline{z}_t$ will be

$$\underline{\underline{\Gamma}}_x(k) = \underline{\underline{M}}\underline{\underline{\Gamma}}_z(k)\underline{\underline{M}}'$$

and so

$$\underline{\underline{\Gamma}}_x(k) = \begin{vmatrix} \gamma_k \Sigma p_i^2 & \underline{0}' \\ \underline{0} & \underline{0} \end{vmatrix} \qquad k > 1$$

$$\underline{\underline{\Gamma}}_x(0) = \begin{vmatrix} \gamma_0 \Sigma p_i^2 & \underline{0}' \\ \underline{0} & \underline{0} \end{vmatrix} + \underline{\underline{M}}\underline{\underline{\Sigma}}_\epsilon\underline{\underline{M}}'$$

and the residual covariance matrix will be

$$\underline{\underline{\Sigma}}_x = \underline{\underline{M}}\underline{\underline{\Sigma}}_u\underline{\underline{M}}' = \begin{vmatrix} \lambda \Sigma p_i^2 & \underline{0}' \\ \underline{0} & \underline{0} \end{vmatrix} + \underline{\underline{M}}\underline{\underline{\Sigma}}_\epsilon\underline{\underline{M}}'$$

The covariance of the factors and the canonical components $\underline{x}_t$ is

$$\underline{\underline{\Gamma}}_{xz}(0) = E[\underline{x}_t\underline{z}_t'] = \underline{\underline{M}}\underline{\underline{P}}\underline{\underline{P}}'\gamma_0 + \underline{\underline{M}}\underline{\underline{\Sigma}}_\epsilon$$

and, in the particular case in which the variance of all the measurement errors are the same, $\underline{\underline{\Sigma}}_\epsilon = \sigma_\epsilon^2\underline{\underline{I}}$

$$\underset{\sim}{\Gamma}_{xz}(0) = \begin{vmatrix} \gamma_0 \Sigma p_i^2 + \sigma_\varepsilon^2 & \underset{\sim}{0}' \\ \underset{\sim}{0} & \sigma_\varepsilon^2 \underset{\sim}{I}_{k-1} \end{vmatrix} \quad \underset{\sim}{M} = \begin{vmatrix} (\gamma_0 \Sigma p_i^2 + \sigma_\varepsilon^2)\underset{\sim}{m}_1' \\ \sigma_\varepsilon^2 \underset{\sim}{m}_2' \\ \vdots \\ \sigma_\varepsilon^2 \underset{\sim}{m}_k' \end{vmatrix}$$

where $\underset{\sim}{I}_{k-1}$ is the unit matrix of dimension $k - 1$ and $\underset{\sim}{m}_i'$ are the rows of M, that is, the eigenvectors of $\underset{\sim}{P}\underset{\sim}{P}'$. This equation shows that the covariances between the factors and the series are proportional to the eigenvalues of $\underset{\sim}{P}\underset{\sim}{P}'$.

4. THE FACTOR MODEL WITH DEPENDENT FACTORS

4.1 Formulation

If we remove the restriction that the parameter matrices of the ARMA model for the factors are diagonal, some uniqueness conditions are needed to determine the model. One suitable normalization is to impose $\underline{P}'\underline{P} = \underline{I}$, that means that of all the C matrices:

$$\underline{z}_t = \underline{P}\underline{C}^{-1}\underline{C}\underline{y}_t + \underline{\varepsilon}_t$$

we choose $\underline{C}$ such as $(\underline{C}^{-1})'\underline{P}'\underline{P}\underline{C}^{-1} = \underline{I}$. Other possible selections are to choose $\underline{C}$ to diagonalize some of the parameter matrices of the factor model. For instance, if $\underline{y}_t$ follow a multivariate AR(1), $\underline{C}$ could be chosen to diagonalize $\underline{\Phi}$ which leads back to the uncoupled case. However, in the general case in which y_t follows a general ARMA (p_y,q_y) model, the condition that $\underline{C}$ has to diagonalize any particular parameter matrix is arbitrary. Thus, it seems more natural to link $\underline{C}$ to the properties of the factor loading matrix.

4.2 Properties of the model

Table 4.1 summarizes the differences between the basic properties of the uncouple factors and the couple or dependent factors model. The main difference arises related to the eigenvectors of the covariance and autoregressive parameter matrices that are no longer columns of the factor loading matrix. Equations (2.4) and (2.16) do hold in this case but the conclusions we have to draw from them are different. For instance, starting with (2.4)

$$\underline{\Gamma}_z(k) = \underline{P}\underline{\Gamma}_y(k)\underline{P}' \qquad k > 1$$

and assuming that $\underline{\Gamma}_y(k) = \underline{U}_k\underline{D}_k\underline{U}_k^{-1}$ where $\underline{D}_k$ is diagonal, is clear that $\underline{\Gamma}_z(k)$ has the same eigenvalues than $\underline{\Gamma}_y(k)$ but with eigenvectors $\underline{P}\underline{U}_k$, $\underline{V}$, where the columns of $\underline{V}$ belongs to the null space of $\underline{P}\underline{P}'$. To see this write

$$\underline{\Gamma}_z(k) = [\underline{P}\underline{U}_k : \underline{V}] \begin{vmatrix} \underline{D}_k & \underline{0} \\ \underline{0} & \underline{0} \end{vmatrix} \begin{vmatrix} \underline{U}_k^{-1}\underline{P}' \\ \underline{V}' \end{vmatrix}$$

that shows clearly the eigenstructure of $\underline{\Gamma}_z(k)$. Of course if $\underline{\Gamma}_y(k)$ is diagonal $\underline{U}_k = \underline{I}$ and we obtain the uncoupled case.

As for the autoregressive parameter matrices, the condition is

$$\underline{\Phi}_z(k)\underline{P} = \underline{P}\underline{\Phi}_y(k)$$

and assuming that $\underline{\Phi}_y(k)$ has lineary independent eigenvectors and writing $\underline{\Phi}_y(k) = \underline{W}_k\underline{\Gamma}_k\underline{W}_k^{-1}$, where $\underline{\Gamma}_k$ is diagonal and contains the eigenvalues of $\underline{\Phi}_y(k)$.

TABLE 4.1

Property	Uncouple Factors	Couple Factors
rank $(\underline{\Gamma}_z(k)) = r$; $k > 1$	Yes	Yes
$\underline{\Gamma}_z(k)$ symmetric	Yes	No
eigenvectors of $\underline{\Gamma}_z(k)$ columns of $\underline{P}$	Yes	No
eigenvalues of $\underline{\Gamma}_z(k)$ eigenvalues of $\underline{\Gamma}_y(k)$	Yes	Yes
rank $(\underline{P}(1)) = r$	Yes	Yes
$z_t \sim$ ARMA $(p_z = p_y, q_z = \max(p_y, q_y))$	Yes	Yes
rank $(\underline{\psi}_i) = r$ $i > 1$	Yes	Yes
eigenvalues of $\underline{\Phi}_z(k)$ are eigenvalues of $\underline{\Phi}_y(k)$	Yes	Yes
eigenvectors of $\underline{\Phi}_2(k)$ are eigenvectors of $\underline{\Phi}_y(k)$	Yes	No

$$\underline{\Phi}_z(k)\underline{PW}_k = \underline{PW}_k\underline{\Gamma}_k$$

that shows that the eigenvalues of $\underline{\Phi}_y(k)$ are eigenvalues of $\underline{\Phi}_z(k)$ as well, with eigenvectors $\underline{PW}_k$.

To build the canonical transformation we need to determine the null space of $\underline{PP}'$. Suppose that $\underline{H}$ is the matrix of eigenvectors of the first autoregressive parameter matrices of the observed series z_t. Then, r columns of H are $\underline{PW}_1$, whereas the other $k - r$ columns are arbitrary. We can partition $\underline{H}$ and $\underline{H}^{-1}$ as

$$\underline{H} = [\underline{PW}_1 \quad \underline{R}_1] \qquad \underline{H}^{-1} = \begin{vmatrix} \underline{W}_1^{-1}\underline{P}' \\ \underline{R}_2 \end{vmatrix}$$

and obtain $\underline{PW}_1$ and $\underline{W}_1^{-1}\underline{P}'$. The determination of these columns can be made because (a) r is the rank of $\underline{\Gamma}_z(k)$, $P(1)$ and $\underline{\psi}_i'$ and can be determined; (b) these columns are

linked to the eigenvalues of $\underline{\underline{\phi}}_y(1)$ that are approximately those of $P(1)$. Then as

$$(\underline{\underline{P}}\underline{\underline{W}}_1)(\underline{\underline{W}}_1^{-1}\underline{\underline{P}}') = \underline{\underline{P}}\underline{\underline{P}}'$$

the null space of $(\underline{\underline{P}}\underline{\underline{W}}_1)(\underline{\underline{W}}_1^{-1}\underline{\underline{P}}')$ is the same as that of $\underline{\underline{P}}\underline{\underline{P}}'$ and $\underline{\underline{V}}'$ can be obtained taking the eigenvectors linked to zero eigenvalues of this matrix. The transformation

$$\underline{\underline{M}} = \begin{vmatrix} \underline{\underline{W}}_1^{-1}\underline{\underline{P}}' \\ \underline{\underline{V}}' \end{vmatrix}$$

will produce, as the uncoupled case, a decomposition of the z_t vector into two components: The first will contain the factors mixed up with noise, and the second component will be a vector of white noise.

$$\underline{x}_t = \underline{\underline{M}}\underline{z}_t = \begin{vmatrix} \underline{\underline{W}}_1^{-1}\underline{y}_t + \underline{\underline{W}}_1^{-1}\underline{\varepsilon}_t \\ \underline{\underline{V}}'\underline{\varepsilon}_t \end{vmatrix} = \begin{vmatrix} \underline{x}_{1t} \\ \underline{x}_{2t} \end{vmatrix}$$

An interesting property of the transformation is that the components obtain in x_{1t} will be, in practice, weekly related. For example if y_t follows an AR(1), $\underline{\underline{W}}_1^{-1}\underline{y}_t$ will follow an AR(1) process too, but with a diagonal autoregressive parameter matrix.

5. SOME PRACTICAL COMMENTS ON THE APPLICATION OF THE THEORY

There are three main decisions that have to be taken to apply the following theory to a real data set of time series. First, the number of factor has to be identified; second, the factor loading matrix $\underline{P}$ has to be estimated, and third, the decoupling transformation M has to be built. We will briefly comment on these subjects on the basis of our experience.

To identify the number of factors we have four different tools at our disposal: The covariance matrices, the correlation matrices, the partial autocorrelation matrices and the $\underline{\Psi}$ parameter matrices. We have shown that the rank of all these matrices should be equal to the number of common factors driving the system.

However, when the variance of one of the factors is very large compared to the variance of all the other factors, the identification of the dimension of the factor space using the rank of the covariance and correlation matrices can be misleading. Suppose for instance that the factors follow an AR(1) model. Then if

$$(\underline{I} - \underline{\Phi}_y B)\underline{y}_t = \underline{a}_t$$

the covariance of $\underline{y}_t$ will be

$$\underline{\Gamma}_y(0) = \underline{\Gamma}_y(1)'\underline{\Phi}_y' + \underline{\Sigma}_a$$

$$\underline{\Gamma}_y(\ell) = \underline{\Gamma}_y(\ell - 1)\underline{\Phi}_y' \qquad \ell > 1$$

Suppose that $\underline{\Sigma}_a$ is diagonal. Then, using the spectral decomposition of $\underline{\Gamma}_y(0)$

$$\underline{\Gamma}_y(0) = \sum_{i=1}^{r} \mu_i \underline{v}_i \underline{v}_i'$$

where μ_i are the eigenvalues of $\underline{\Gamma}_y(0)$ and $\underline{v}_i$ its eigenvectors. As $\underline{\Gamma}_y(0)$ is diagonal, μ_i are also the variance of the components, and if the variance of the first factor is much larger than the variance of the others

$$\underline{\Gamma}_y(0) = \mu_{1,0}\underline{v}_{1,0}\underline{v}_{1,0}'$$

and in general

$$\underline{\Gamma}_y(k) = \mu_{1,k}\underline{v}_{1,k}\underline{v}_{1,k}' \tag{5.1}$$

where $\mu_{1,k}$ is the largest eigenvalue and $\underline{v}_{1,k}$ its associated eigenvector. Then

$$\underline{\Gamma}_z(0) = \mu_{1,0}\underline{h}_{1,0}\underline{h}_{1,0}' + \underline{\Sigma}_c$$

$$\underline{\Gamma}_z(k) = \mu_{1,k}\underline{h}_{1,k}\underline{h}'_{1,k} \qquad (k > 1)$$

where $\underline{h}_{1,j} = \underline{P}\underline{v}_{1,j}$. So, all the covariance matrices will be dominated by a large eigenvalue and will have rank equal to one. In particular, if all the components are nonstationary and $\underline{\phi}_y = \underline{I}$, then $h_{1,0} = h_{1,k}$ for all k and the covariance matrices will display the structure we have obtained for the one factor model.

However, as the columns of $\underline{P}$ are eigenvectors of $\underline{\phi}_z$ whatever the variance of the factors, if $\underline{\phi}_y = \underline{I}$, $\underline{\phi}_z$ will still have at least r-roots on the unit circle and at least r eigenvalues equal to one, linked to the $\underline{P}$ vectors. For this reason the identification of the number of factors looking only at the rank of the covariance matrices could be misleading. Note that this problem is not resolved by looking at the correlation matrices because, calling $\underline{R}_z(k)$ the autocorrelation matrix of order k.

$$\underline{R}_z(k) = \text{Diag}(\underline{\Gamma}_z(0))^{-1/2}\underline{\Gamma}_z(k)\text{Diag}(\underline{\Gamma}_z(0))^{-1/2}$$

and using (5.1)

$$\underline{R}_z = \mu_{1,k}\underline{h}_{1,k}\underline{h}'_{1,k}$$

where

$$\underline{h}_1 = \text{Diag}(\underline{\Gamma}_z(0))^{-1/2}\underline{P}\underline{v}_{1,k} .$$

and the same situation will appear.

The rank of the nonzero partial autocorrelation matrices can also be misleading when $\underline{\Gamma}_z(0)$ is almost singular. Then the precision of the computation of these matrices is very low and the determination of how many roots of the partial autocorrelation matrix can be considered equal to zero is not easy. As a general rule, the simulations we have made have shown that estimated eigenvalues as large as .5 can appear for theoretically zero values in the partial autocorrelation matrices when the variance of the noise is small compared to the variance of the signal.

The rank of the $\underline{\Psi}$ matrices seems to be the safest way to determine the number of factors. When there is disagreement between the three features we have analyzed, a conservative strategy is to start assuming a small number of factors and proceed iteratively, building the components and analyzing the noise for further structure.

The factor loading matrix $\underline{P}$ should be determined from the eigenvectors of the maximum likelihood estimator of the autoregressive parameter matrix $\underline{\phi}$. Only in the one factor case the column $\underline{P}$ could be obtained directly from the covariance or partial autocorrelation matrices. A well known rule in matrix computation (see Stewart (1973)) is that the eigenvectors of a matrix are well-conditioned only when the matrix has distinct and well-separated eigenvalues, whereas if the matrix has repeated or clustered eigenvalues the estimation of its eigenvectors may be ill-conditioned. This means that, on the one hand, we can expect a good agreement in the one factor case among the different estimations that we have of the vector of factor loadings; on the other hand, when we have several nonstationary factors, we have to expect this same number of unit roots in the autoregressive matrix and the estimation of the eigenvectors is not expected to be very accurate. For this reason, even if the factors are independent, we can finish with factors that have small correlation or with final noise processes that are not exactly white noise processes.

6. EXAMPLES

6.1 The Series of Wheat Prices in Castille

Figure 1 shows the series of the price of wheat in five provinces of Castillia, Spain from July 1880 to December 1890. Therse provinces are Avila, Palencia, Segovia, Valladolid and Zamora. The source of the data is Sanchez-Albornoz (1975) and Peña and Sánchez Albornoz (1983) have studied some of these series during the second half of the XIXth century, pointing out their lack of seasonality.

The analysis that follows is made with the series in logs, although the conclusion are not depending on the particular transformation used.

Table 5 shows the univariate model for the five series. These models are broadly consistent with the hypothesis of some common factors following an AR(1) model. Table 6 displays the eigenvalues and eigenvectors of the covariance matrices up to lag 6. There is a strong eigenvalue that decreases slowly linked to a fairly stable eigenvector. The second eigenvalue is rather small and the eigenvector is not completely stable. Finally, the other three eigenvectors are practically zero. Then, it seems there are at least one important common component and perhaps a second one although much less important than the first.

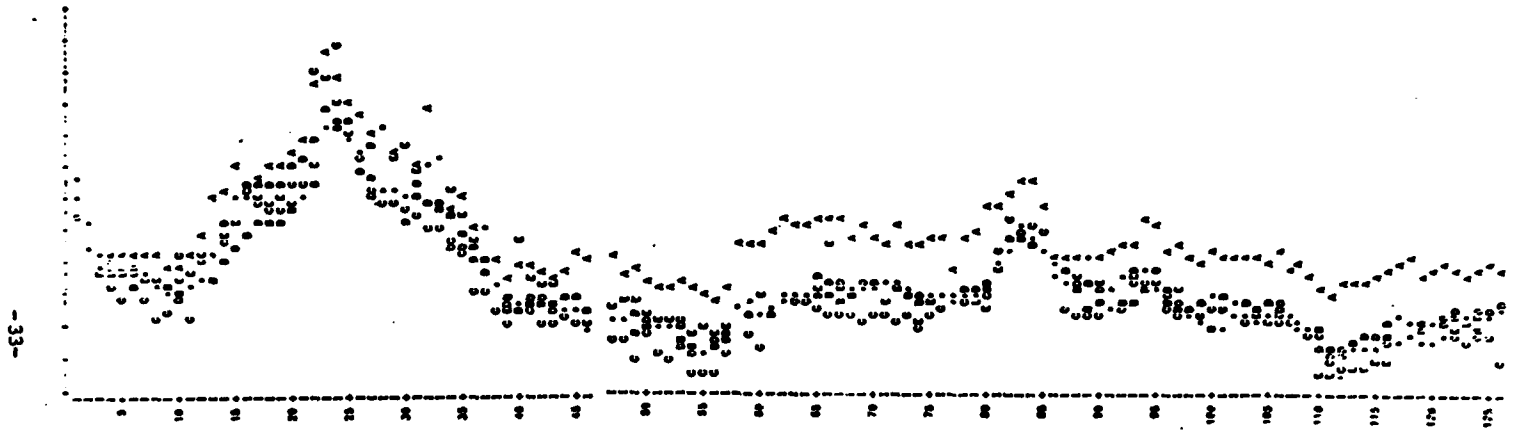


Figure 1

Monthly Wheat Price series (1880-1890) of Avila (A), Palencia (B), Segovia (C), Valladolid (D) and Zaragoza (E).

TABLE 5

Univariate model for the series

∇ ℓn Avila = (1 + .11B) a_t (1.30)	$\hat{\sigma}_a$ = 0.039	Q(23) = 32.8
∇ ℓn Palencia = (1 + .18B) a_t (2.02)	$\hat{\sigma}_a$ = 0.030	Q(23) = 24
∇ ℓn Segovia = (1 - .16B) a_t (1.77)	$\hat{\sigma}_a$ = 0.057	Q(23) = 16
∇ ℓn Valladolid = (1 + .15B) a_t (1.70)	$\hat{\sigma}_a$ = 0.031	Q(23) = 36.9
(1 - .15B - .20B^2) ∇ ℓn Valladolid = a_t (1.70) (2.30)	$\hat{\sigma}_a$ = 0.029	Q(23) = 34.6
∇ ℓn Zamora = (1 - .16B) (1.84)	$\hat{\sigma}_a$ = .049	Q(23) = 23.9

The values under the estimated parameters are t-values. Q(ℓ) is the statistics of Ljung-Box with ℓ degrees of freedom. $\chi^2_{23}(0,05)$ = 35.2.

TABLE 6

Eigenvalues λ and Eigenvectors v of Covariance Matrices

r_0

810	20	10	10	10
.36	.27	.80	.33	.24
.43	-.14	.06	-.80	.40
.50	.57	-.58	.22	.19
.44	.10	.12	-.21	-.86
.49	-.76	-.15	.41	.05

a) Largest eigenvector and eigenvalue

Lag	1	2	3	4	5	6
λ	777	735	693	647	604	560
v	.35	.35	.34	.33	.32	.32
	.44	.44	.45	.45	.45	.45
	.49	.49	.48	.47	.47	.46
	.44	.44	.44	.44	.43	.43
	.49	.50	.51	.52	.53	.55

b) Second largest eigenvector and eigenvalue

lag	1	2	3	4	5	6
λ	12	13	11	8	5	7
v	.65	.40	.08	.21	.38	.18
	-.18	-.17	.01	-.09	.26	.23
	.28	.46	.18	.31	.05	.34
	.06	.16	.04	.16	.10	.43
	-.68	-.77	-.31	-.57	-.68	-1.15

c) Leftover eigenvalues

lag	1	2	3	4	5	6
3rd	0	1	4	3	5	1
4th	5	4	4	3	4	2
5th	5	4	4	1	0	2

Note: All the eigenvalues have been multiplied by 10^4.

Table 7 shows the eigenvalues and eigenvectors of the first two partial correlation matrices. The two main eigenvalues of the covariance matrices appear again linked to the largest and third eigenvalues of $P(1)$ and, less clearly, in the two negative elements of $P(2)$. Table 8 shows the pattern of significant coefficients in the exact maximum likelihood estimation of the parameters of an ARMA (1,1). It can be concluded that there is a very complex relationship among the series with all kinds of feedback present. The eigenvalues of the matrices $\underline{\Phi}$, $\underline{\theta}$ and $\underline{\Phi} - \underline{\theta}$ are shown in Table 9. The five eigenvalues of $\underline{\Phi}$ are close to the unit circle, what can wrongly lead to the conclusion that there are five nonstationary factors. However, the rank of $\underline{\Phi} - \underline{\theta}$ is nearly two, leading to the conclusion that there are 2 factors in the system. The eigenvectors linked to these main eigenvalues are again very close to the previous eigenvectors of the covariance and partial correlation matrices, and these two eigenvectors appear approximately in the $\underline{\Phi}$ and $\underline{\theta}$ matrices, as shown.

In summary, there is strong evidence of at least one, and probably two, common factors driving the system. The first is linked to an eigenvalue that is broadly speaking a mean of the series. The second is roughly the ratio among Avila and Zamora. This result makes sense because this second factor affects the two more western provinces that are closer to the Portugal border.

TABLE 7

Eigenstructure of the Partial Correlation Matrices

.95	.63	.69	.41	.04
.37	.69	-.94	.51	.02
.49	-.16	2.79	-.16	.05
.53	-.05	1.26	-.20	-.93
.48	-.31	1.96	-.16	.27
.57	-.74	4.94	.90	-.24

P(1)

-.47	-.09	.41	.17	.20
.58	.77	.04	.23	.20
.32	-.23	.51	-.36	-.77
.48	-.59	-.53	.95	.43
.38	-.05	-.34	-.37	-.14
.43	.16	.67	.38	1.40

P(2)

TABLE 8

Significant Values in the Parameter Matrices

$$\underset{\sim}{\Phi}: \begin{vmatrix} + & . & - & + & . \\ . & + & . & + & . \\ . & - & + & + & . \\ . & . & . & + & . \\ . & . & . & + & + \end{vmatrix} \qquad \underset{\sim}{\Theta}: \begin{vmatrix} + & - & - & . & + \\ . & . & . & . & . \\ . & - & + & . & . \\ . & . & - & . & . \\ . & - & - & . & + \end{vmatrix}$$

TABLE 9

Eigenvalues of the Parameter Matrices and Main Eigenvector

$\underset{\sim}{\Phi}$	$.99^{(1)}$	.98+.08i	.98-.08i	.84+.14i	.84-.14i
$\underset{\sim}{\Theta}$	$-.51^{(1a)}$	$.51^{(2a)}$	.98+.18i	.98-.18i	.85
$\underset{\sim}{\Phi} - \underset{\sim}{\Theta}$	$1.46^{(1e)}$	$.34^{(2e)}$	.14	-.05+.03i	-.05-.03i

We have chosen the eigenvectors (.37 .49 .53 .48 .57) and (.69 0 0 0 -.74) to build the transformation. Any other election does not seem to change the results. With this two generators, the matrix $\underset{\sim}{M}$ turns out to be

$$\underset{\sim}{M} = \begin{vmatrix} .52 & .07 & -.51 & -.48 & .49 \\ .26 & -.88 & .09 & .31 & .24 \\ .04 & -.15 & .70 & -.70 & .03 \\ .41 & .41 & .45 & .41 & .38 \\ .74 & .07 & .07 & .07 & -.66 \end{vmatrix}$$

The resulting five series are plotted in Figure 2. The new $\underset{\sim}{\Phi}_x$ and $\underset{\sim}{\Theta}_x$ for the transformed $\underset{\sim}{x}$ series obtained transforming the parameter matrices of $\underset{\sim}{z}$ turn out to be

$$\underset{=x}{\phi} = \underset{==}{M\phi M}^{-1} = \begin{vmatrix} . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & -.71 & .94 & .18 \\ . & . & . & . & .92 \end{vmatrix} + \begin{vmatrix} .96 & . & . & . & . \\ . & .98 & -.19 & . & . \\ . & .16 & .84 & . & . \\ -.30 & .82 & -.24 & . & . \\ . & .25 & -.27 & . & . \end{vmatrix}$$

$$\underset{=x}{\theta} = \underset{==}{M\theta M}^{-1} = \begin{vmatrix} . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & -.55 & . \\ . & . & . & .20 & .50 \end{vmatrix} + \begin{vmatrix} .90 & .19 & .15 & . & . \\ . & .93 & -.23 & . & . \\ . & .16 & 1.0 & . & . \\ .12 & .83 & -.12 & . & . \\ . & .33 & -.19 & . & . \end{vmatrix}$$

where we have used a dot to indicate a parameter value below .1. This decomposition shows that there is a matrix A of rank equal to three that can be subtracted from both sides. All the rows of this matrix are orthogonal to the vectors (00010) and (00001) and so, according to Theorem 2, could be subtracted to obtain three independent sequences of white noise plus 2 factors. The first factor is, according to the fourth raw of M, a mean of all the series and describes the general trend of the market. The second factor is stationary and affects to the most westerly provinces, Avila and Zamora, and although stationary seems to have an ARMA (1,1) structure. The factors are not completely uncorrelated and a weak relationahip seems to exits between them.

To check all these results we have built a model for the factors from scratch. Table 10 shows the parameters of the model and Table 11 the correlation and partial correlations of the residual. The first two component of the decomposition are stationary AR(1) process uncorrelated at all lags and the third is white noise. The fourth is the nonstationary component that describes the general evolution of the market and the fourth is the differential effect of the occidental provinces. This factor turns out to be uncorrelated with the main trend.

The results are sensible and confirms that the five series are driven by a main factor that explains their nonstationary behavior. The strongest correlation of the main factor and the five series is with Valladolid $r = .99$ that was the greatest producer of wheat in Castillia and is consistent with the expected behavior "a priori".

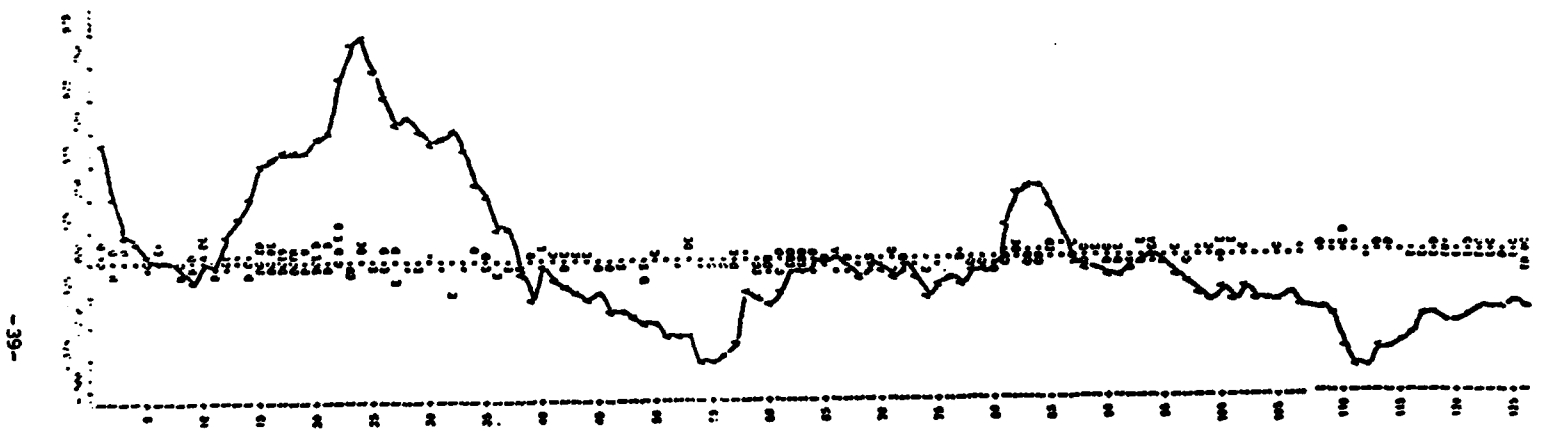


Figure 2

Main Factor (A) and other components of the original series.

TABLE 10

$$\phi_x = \begin{vmatrix} .45 & . & . & . & . \\ (.07) & & & & \\ . & .56 & . & . & . \\ & (.06) & & & \\ . & . & . & . & . \\ -.43 & . & -.74 & .95 & . \\ (.14) & & (.12) & (.02) & \\ . & . & . & . & .76 \\ & & & & (.07) \end{vmatrix} \qquad \theta_x = \begin{vmatrix} . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & -.50 & . \\ & & & (.06) & \\ . & . & . & . & . \end{vmatrix}$$

TABLE 11

$$R_x = \begin{vmatrix} 1 & & & & \\ .09 & 1 & & & \\ -.45 & .12 & 1 & & \\ .00 & .49 & .28 & 1 & \\ -.26 & -.04 & .08 & .24 & 1 \end{vmatrix} \qquad 10^4\Sigma_x = \begin{vmatrix} 10.34 & & & & \\ .59 & 4.54 & & & \\ -4.40 & .76 & 9.21 & & \\ .08 & 5.21 & 4.34 & 25.25 & \\ 2.53 & -.25 & .71 & 3.79 & 9.56 \end{vmatrix}$$

Parameters of the transformed series and correlation among residuals.

As a further check of the methodology we have analyzed the wheat price series of three provinces of different economic behavior in which a unique common factor is not expected. The provinces we have chosen are La Coruña, that is not wheat producer and is the North of Spain, Zaragoza, that although producer of wheat supplied to Aragon and Cataluña, and Valladolid again as representative of Castillia. Table 12 displays the eigenvalues of the covariance matrices and although there is a large eigenvalue linked to the general mean, the other two eigenvalues are fairly stable. The partial autocorrelation coefficients show 3 strong roots and the estimation of the model in Table 13 shows that $\underline{\phi} - \underline{\theta}$ has rank three and it is not possible to reduce the dimension of the system. There are three factors, and the three are nonstationary as expected.

TABLE 12

Eigenvalues and Eigenvectors of the Covariance Matrices of La Coruna, Valladollid and Zaragoza

lag	1	2	3	4	5	6
λ_1	486	479	459	441	422	402
	.42	.43	.44	.46	.47	.48
	.49	.47	.45	.43	.40	.38
	.76	.77	.78	.78	.79	.79

	1	2	3	4	5	6
λ_2	50	47	43	43	41	39
	.64	.64	.66	.66	.68	.66
	-.75	-.75	-.74	-.74	-.73	-.75
	.18	.19	.19	.20	.20	.22

	1	2	3	4	5	6
λ_3	9	7	4	5	4	3
	.67	.68	.69	.68	.68	.68
	.39	.37	.37	.38	.37	.37
	-.63	-.62	-.63	-.63	-.63	-.63

TABLE 13

Eigenvalues of the Parameter Matrices of the Model for La Coruna, Valladolid and Zaragoza

	.93	1	.93
	1	.78	-4.5
$\underline{\Phi}$ =	0	.70	0
	0	.71	1.41

	.21	.14	-.14
	1	1.71	.38
$\underline{\Theta}$ =	0	.78	.79
	0	-.62	.64

	1	.92	.73
	-.6	-1.5	1
$\underline{\Phi} - \underline{\Theta}$ =	1	2	0
	0	1.0	0

6.2 The Series of Stock Prices Indexes

The three series studied are monthly averages of the Dow Jones Index of 30 Industrial Stocks (DJ Series), the Standard and Poor Industrial Stock Index (ID series), that includes 400 stocks, and the Standard and Poor Stock Price Index (SP series) that includes 500 companies chosen from all sectors of the U.S. economy. The sample period is from January 1966 to December 1981. Figure 3 shows the evolution of the three series that have been standarized removing their means and dividing by their standard deviation to show the three in a common scale. Table 4 shows the univariate models.

We will report here the analysis on the logarithm of the series although the conclusion are robust to the scale of measurement. Table 15 displays the eigenstructure of the covariance matrices and, as the spread of Dow-Jones is quite different from the other

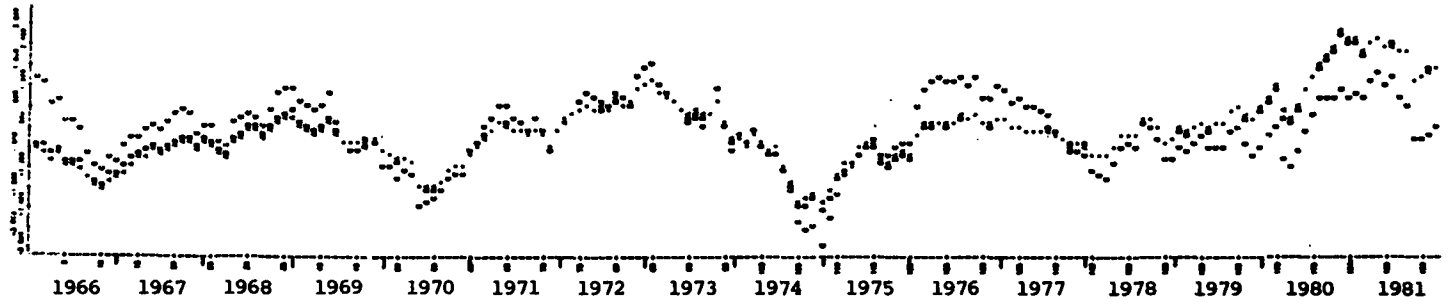


Figure 3

Series of Industrial Standard and Poor (A), Standard and Poor Index (B) and Dow-Jones Index (C).

two, the eigenvalues and eigenvectors of the correlation matrices are also presented. Both set of matrices show three very stable eigenvectors. The eigenvector linked to the greatest eigenvalue indicates that the first factor is a weighted mean of the three series. Thus, it can be interpreted as the general trend of the stock market. The second factor seems to take into account the behavior of large industrial corporations (as measured by the Dow Jones Index) with respect to the rest of the market. The third eigenvalue is very small and, consequently, the system seems to be driven by two factors. The third eigenvalue represents a constant relationship among the three series in the whole sample period. This relationship is

$$.66 \ln \frac{LI}{LS} + .09 \ln \frac{LD}{LS}$$

and is a weighted combination of the performance of the industrial sector (defined by LI and LD) with respect to the whole of the market (LS). The fact that the eigenvalue linked to this eigenvector is equal to zero means that this relationship has been stable in the stock market.

Identification suggests a multivariate ARMA (1,1) model for the vector of the three indexes. When fitted by maximum likelihood, and after deleting the nonsignificant parameters, model A of Table 16 is obtained. The $\underline{\theta}$ parameter matrix has the remarkeable property that rows one and two are nearly identical and the rank of θ is two. Actually, the $\underline{\theta}$ matrix could be decomposed as the addition of

TABLE 14

Univariate Models for Stocks Series

Model	$\hat{\sigma}_a$	Q(35)
$\nabla LI = \underset{(.07)}{(1 + .26B)}a_t$	.251	33.5
$\nabla LS = \underset{(.07)}{(1 + .26B)}a_t$	.262	34.5
$\nabla LD = \underset{(.07)}{(1 + .22B)}a_t$	.377	31.9

LI = ln of Industrial Standard and Poor Index; LS = ln of Standard and Poor Index; LD = ln of Dow Jones Index. Q(k) is the Lunj-Box statistic with k degrees of freedom.

TABLE 15

Eigenvalues of Covariances and Correlation Matrices
of the Series $Z_t' = (LI, LS, LD)$

a) Covariance (eigenvalues × 10^4)

Largest Eigenvalue

lag	0	1	2	3	4	5
λ	440	420	389	363	330	299
	.68	.68	.68	.69	.69	.70
	.64	.64	.64	.69	.66	.65
	.37	.37	.36	.34	.35	.30

Second Eigenvalue

lag	0	1	2	3	4	5
λ	30	32	29	26	23	20
	.33	.33	.32	.30	.30	.30
	.19	.18	.17	.17	.15	.12
	-.92	-.92	-.93	-.94	-.94	-.95

b) Correlations

lag	0	1	2	3	4	5
λ	2.66	2.53	2.33	2.15	1.93	1.72
	.60	.60	.60	.61	.63	.63
	.60	.60	.60	.61	.63	.64
	.54	.53	.52	.50	.47	.44

lag	0	1	2	3	4	5
λ	.34	.30	.27	.25	.22	.19
	-.37	-.41	-.40	-.39	-.38	-.37
	-.37	-.33	-.31	-.30	-.28	-.26
	.85	.85	.86	.87	.88	.89

c) Third eigenvalue

Covariances (eigenvalues × 10^3

lag	0	1	2	3	4	5
λ	0	2	2	3	1	0
	.66	.66	.66	.66	.66	.67
	-.75	-.75	-.75	-.75	-.75	-.74
	.09	.09	.09	.08	.08	.06

Correlations

lag	0	1	2	3	4	5
λ	.009	.009	.009	.004	.004	.004
	.68	.68	.68	.68	.68	.69
	-.73	-.73	-.73	-.73	-.73	-.72
	.06	.06	.06	.05	.06	.04

TABLE 16

a) Model A

$$\underset{\phi}{\begin{vmatrix} 1 & . & . \\ (.001) & & \\ . & 1 & . \\ & (.001) & \\ . & . & 1 \\ & & (.001) \end{vmatrix}} \quad \underset{\theta}{\begin{vmatrix} 2.52 & -3.80 & .92 \\ (.09) & (.06) & (.13) \\ 2.68 & -3.96 & .93 \\ (.10) & (.04) & (.03) \\ 2.56 & -3.44 & .55 \\ (.17) & (-.17) & (.14) \end{vmatrix}} \quad \underset{\Sigma_u \times 10^4}{\begin{vmatrix} 13.12 & & \\ 12.84 & 12.65 & \\ 12.39 & 12.20 & 12.94 \end{vmatrix}}$$

b) Model B

same $\underset{\sim}{\phi} = \underset{\sim}{I}$

$$\underset{\theta}{\begin{vmatrix} -.16 & . & . \\ (.05) & & \\ . & -.14 & . \\ & (.05) & \\ . & .23 & -.36 \\ & (.09) & (.07) \end{vmatrix}} \quad \begin{vmatrix} 12.99 & & \\ 12.80 & 12.60 & \\ 12.42 & 12.22 & 13.05 \end{vmatrix}$$

c) Model C

$\phi = I$

$$\underset{\theta}{\begin{vmatrix} -.23 & . & . \\ .04 & & \\ . & -.22 & . \\ & (.04) & \\ . & & -.21 \\ & & (.04) \end{vmatrix}} \quad \approx -.22I \quad \begin{vmatrix} 12.99 & & \\ 12.69 & 12.47 & \\ 11.26 & 12.05 & 12.86 \end{vmatrix}$$

a matrix similar to the $\underset{\sim}{\theta}$ matrix of model B (see Table 16B) plus a matrix with the three rows equal to (2.7 -3.8 .92). Calling this rank one matrix A, this implies

$$\underset{\sim}{A}\underset{\sim}{a}_t \approx \underset{\sim}{0}$$

That means that a noise sequence can be expressed as a linear combination of the other two.

An interesting fact is that the vector that defines this combination is, approximately, proportional to the eigenvalue linked to the smallest eigenvalue in the covariance and autocorrelation matrices. Thus, the same restriction affects to both the noises and the vector of series. Calling m' this common vector we have

$$\underset{\sim}{m}'\underset{\sim}{z}_t \approx 0$$

$$\underset{\sim}{m}'\underset{\sim}{a}_t \approx 0 \ .$$

For these results to be in agreement with the model for the series, as

$$m'\nabla z_t = m'(I - \theta B)a_t$$

then

$$m'\theta a_t = 0$$

which means that θ must be equal to kI, where k is a scalar. The matrix θ of model B suggest that this can be a reasonable model. The maximum likelihood estimation with the restriction that the matrix θ be diagonal leads to model C. Note that the fit of the model as measured by the diagonal elements of Σ_u has not worsened, and the diagnostic checks does not indicate any problems in the estimated residuals. The three diagonal elements of θ in model C are, approximately, equal and so, the final model for the series is:

$$\nabla z_t = (I + .22BI)a_t .$$

As the ϕ matrix is the identity it does not provide information about eigenvectors in this case. However, we noticed that (a) the model for the common factors must be $\nabla y_t = a_t$ because the eigenvalues of ϕ, that are all equal to one, must match the eigenvalues of the factors. (b) In this particular case the factors are undetermined because any orthogonal transformation of y_t will produce exactly the same model for the factors, (c) we can choose the orthogonal transformation in order to result in orthogonal columns in the vector P, (d) If we do that, the columns of P will be eigenvectors of the covariance matrices.

Choosing this representation, we first notice that the two main eigenvectors of the covariance matrices (Table 15) are orthogonal. Then taking these eigenvectors as generators of the transformation the matrix M

$$M = \begin{vmatrix} .68 & .64 & .37 \\ .32 & .18 & -.92 \\ +.66 & -.75 & +.08 \end{vmatrix}$$

This produces three new aggregates $x_t = My_t$. The first could represent the general tendency of the stock market. The second could be interpreted as the behavior of large corporation versus the rest of the market. The third is the ratio of the industrial sector versus the general market. This third component is approximately the linear combination of

the series that has a variance very close to zero. These three components are shown in Figure 4. The variance of the first is the largest eigenvalue of Γ_0 (.0440) the variance of the second less than 7% of the first (.0030) and the third has a variance that is 0,0023% of the variance of the mean factor and could be considered nearly constant.

However, when we apply the transformation M to the matrices $\underline{\phi}$ and $\underline{\theta}$, since $\underline{\phi} = \underline{I}$ and $\underline{\theta} = -.22\underline{I}$ these parameter matrices do not change. Then, the vector $\underline{x}_t$ follows the same model as $\underline{z}_t$, which means that the third component of $\underline{x}_t$ although almost zero, is a nonstationary factor. This factor is displayed in Figure 5 in a very augmented scale and in fact it is clear that it is nonstationary. It increases slowly but steadily until 1974, suggesting that the stock prices of Industrial corporations were gaining against the general market until 1974. Then, this tendency broke down and industry began loosing steadily position against the general market. This component may indicate a relative deterioration of U.S. industry after the 74 oil crisis.

The conclusion from this example is that we need to carry out the complete analysis to obtain a clear picture of the system. As we stressed in Section 5, if the variance of one factor is very small in relation to all the others, this factor may be overlooked in the study of the rank of the covariance and partial autocorrelation matrices. However, it will show up in the final multivariate ARIMA model and will be easily detected when applying the transformation to recover the factors.

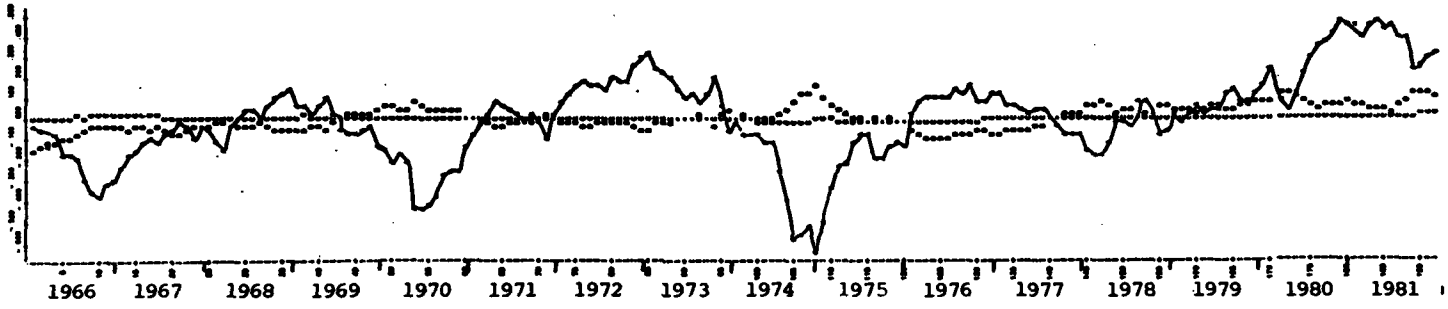


Figure 4

The series in gross is the first component, the second is represented by B and the third by C.

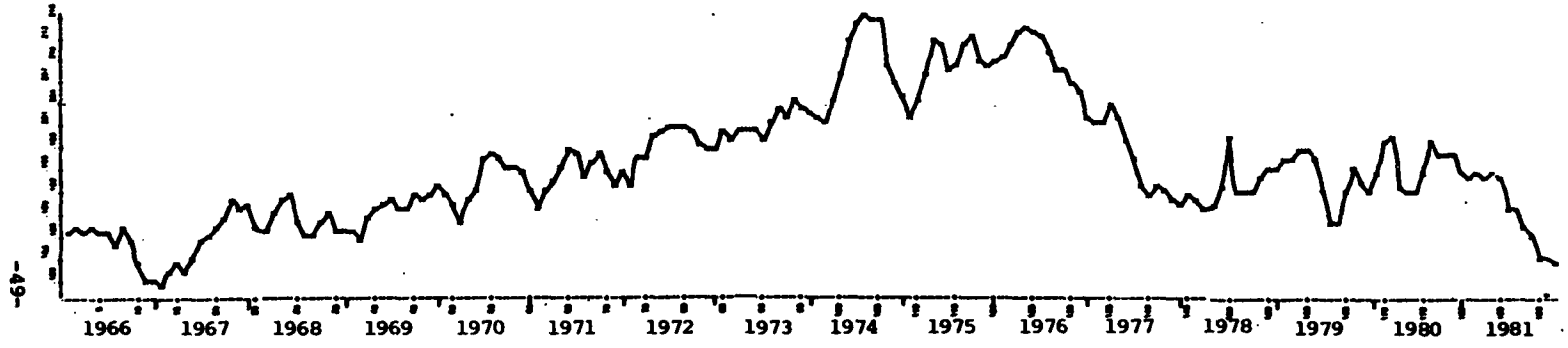


Figure 5

Third component: Industrial versus rest of the market.

7. COMPARISON WITH OTHER APPROACHES

Quenoulli (1957) suggested to use the eigenvectors and eigenvalues of $\hat{\Gamma}_u$, $\Gamma_0(z)$ to build transformations that allow simpler interpretation of the fitted model and/or to identify the possible number of factors. We now revise briefly the properties of these transformations.

Starting with the simplest case, the one factor model it has been shown that both matrices have the representation

$$k_1 \underline{P}\underline{P}' + \Sigma_\varepsilon$$

and so (1) procedures based on either the rank or the eigenvalues of $\hat{\Gamma}_u$ will be similar to those based on these same properties of $\Gamma_0(z)$; (2) both matrices are full rank matrices if Σ_ε is not degenerate; (3) the eigenvectors of either $\hat{\Gamma}_u$ or $\Gamma_z(0)$ will only coincide with those of $\underline{P}\underline{P}'$ when $\Sigma_\varepsilon = \sigma_\varepsilon^2 I$. It turns out then that in the one factor case only when $\Sigma_\varepsilon = \sigma_\varepsilon^2 I$ and σ_ε^2 is very small compared to $\gamma_0(y)$ (see Section 5) principal components based on either $\Gamma_z(0)$ or $\hat{\Gamma}_u$ will obtain the same results as the method suggested in this paper.

In the multifactor situation, the eigenvectors of $\Gamma_0(z)$ and $\hat{\Gamma}_u$ needn't be the same but the above considerations still apply. As $\Gamma_z(0) = \underline{P}\Gamma_y(0)\underline{P}' + \Sigma_\varepsilon$ principal components in the original data will depend on the measurement error matrix Σ_ε, and the same is true for the matrix $\hat{\Gamma}_u$.

Box and Tiao (1977) suggested the use of the eigenvectors of the matrix

$$Q = \Gamma_z^{-1}(0)\hat{\Gamma}_z(0) \tag{7.1}$$

where it is assumed that z_t follows an autoregressive process $z_t = \sum_{\ell=1}^{p} \Phi_\ell z_{t-\ell} + a_t$, and

$$\hat{\Gamma}_z(0) = \sum_{\ell=1}^{p} \Phi_\ell \Gamma_z(\ell)$$

and they proved that a transformation based on the eigenvectors of Q produces a set of new canonical variables that are ordered from least to most predictable. To compare this canonical transformation wih the one recommended in this paper let us assume first, that the model equation for the factors is AR(1) and second, that the model identified and

fitted for the observed series is AR(1). Then the estimation of the autoregressive parameter $\underline{\Phi}_z$, will be

$$\underline{\Phi}_z = \underline{\Gamma}_z'(1)\underline{\Gamma}_z^{-1}(0) = \underline{P}\underline{\Gamma}_y'(1)\underline{P}'(\underline{P}\underline{\Gamma}_y(0)\underline{P}' + \underline{\Sigma}_e)^{-1}$$

and using the expression for the first partial autoregression matrix obtained in (2.6), and inserting it into (7.1),

$$\underline{Q} = \underline{\Gamma}_z^{-1}(0)\underline{\Phi}_z\underline{\Gamma}_z(0)\underline{\Phi}_z'$$

if we now apply that

$$\underline{P}\underline{\Sigma}_e^{-1}(\underline{\Sigma}_e + \underline{P}\underline{\Gamma}_0(y)\underline{P}')\underline{\Sigma}_e^{-1}\underline{P} = \underline{A}(\underline{I} + \underline{\Gamma}_0(y)\underline{A})$$

where

$$\underline{A} = \underline{P}'\underline{\Sigma}_e\underline{P}$$

the general expression for $\underline{Q}$ is:

$$\underline{Q} = (\underline{\Sigma}_e + \underline{P}\underline{\Gamma}_0(y)\underline{P}')^{-1}\underline{P}\underline{\Phi}_y\underline{\Gamma}_y(0)(\underline{I} + \underline{A}\underline{\Gamma}_y(0))^{-1}\underline{A}\underline{\Gamma}_y(0)\underline{\Phi}_y\underline{P}'$$

and using the well known expression for the inversion of the sum of two matrices one of them being nonsingular,

$$\underline{Q} = (\underline{\Sigma}_e^{-1} - \underline{\Sigma}_e^{-1}\underline{P}(\underline{I} + \underline{\Gamma}_y(0)\underline{A})^{-1}\underline{\Gamma}_y(0)\underline{P}'\underline{\Sigma}_e^{-1})\underline{P}\underline{\Phi}_y\underline{\Gamma}_y(0)(\underline{I} + \underline{A}\underline{\Gamma}_y(0))^{-1}\underline{A}\underline{\Gamma}_y(0)\underline{\Phi}_y\underline{P}'$$

and, after some manipulations and taking into account that

$$I - (I + H)^{-1}H = (I + H)^{-1}$$

we finally obtain

$$\underline{Q} = \underline{\Sigma}_e^{-1}\underline{P}(\underline{I} + \underline{\Gamma}_y(0)\underline{A})^{-1}\underline{\Phi}_y\underline{\Gamma}_y(0)(\underline{I} + \underline{A}\underline{\Gamma}_y(0))^{-1}\underline{A}\underline{\Gamma}_y(0)\underline{\Phi}\underline{P}' \qquad (7.2)$$

Then, if $\underline{V}'$ is such that $\underline{V}'\underline{P} = \underline{0}$, then

$$\underline{Q}\underline{V} = \underline{G}\cdot\underline{P}'\underline{V} = \underline{0}$$

and so the columns of $\underline{V}$ are eigenvectors of $\underline{Q}$ linked to the zero eigenvalues of $\underline{Q}$. However, expression (7.2) shows that neither $\underline{P}$ nor $\underline{C}\underline{P}$ when $\underline{C}$ is an orthogonal matrix, will be in general eigenvectors of $\underline{Q}$ and so the canonical transformation will fail to recuperate the factors. The reason is that the linear combinations of greater predictability depends not only on the factors, but on the variance of the signal as well.

In the one factor model, equation (7.2) reduces to

$$\underline{Q} = \underline{\Sigma}_{\varepsilon}^{-1}\underline{P}\underline{P}' \cdot \lambda_Q$$

with

$$\lambda_Q = \frac{a\gamma_0^2(y)\phi_y^2}{(1 + a\gamma_0(y))^2}$$

and $a = \underline{P}'\underline{\Sigma}_{\varepsilon}^{-1}\underline{P}$ as before. Then $\underline{\Sigma}_{\varepsilon}^{-1}\underline{P}$ is the vector linked to the nonzero eigenvalue of $\underline{Q}$ that is $a\lambda_Q$. As ϕ_y approaches the unity, if $a\gamma_0(y)$ is much greater than one, λ_Q will approach unity too, in agreement with Box and Tiao (1977) theorem. In addition, the eigenvalues linked to zero eigenvectors of $\underline{Q}$ must be linear combinations of those of $\underline{P}\underline{P}'$ and, calling $\underline{V}'$ the matrix whose rows are the eigenvectors of $\underline{Q}$, $\underline{V}'\underline{P} = \underline{0}$. Therefore both procedures will come out with a different component plus $k - 1$ white noise. In the canonical transformation this single component depends on $\underline{P}$ and on $\underline{\Sigma}_{\varepsilon}$ whereas the factor analysis decomposition depends only on $\underline{P}$.

8. CONCLUSION

It has been shown that the multivariate time series model for an observed vector of time series may appear very complicated, when in fact a simple representation using a small set of common factor is adequate. Besides, failing to notice this structure can be dangerous, because the interpretation of the internal dynamic that we have to draw from the model will be completely different in both cases.

Unlike in the factor analysis of static variables, the restrictions that the presence of common factors put on the covariance, correlation, partial autocorrelation and parameter matrices make the identification of the number of factors and the estimation of the factor loading matrix relatively easy.

Box and Tiao (1977) and Tiao and Box (1981) recommend the computation of eigenvalues and eigenvectors of different matrices as a useful source of valuable information. It has been shown the relationship between this eigenvalues and eigenvectors in the context of a common number of factors and how a transformation can be built to obtain a parsimonious and intelligible representation of the observed vector of time series.

REFERENCES

[1] Brillinger, D. R. (1981). Time Series, Data Analysis and Theory, expanded edition. Holden-Day.

[2] Box, G. E. P. and Tiao, G. C. (1977). "A canonical analysis of multiple time series". Biometrika, 64, 355-365.

[3] Geweke, J. (1977). "The Dynamic Factor Analysis of Economic Time Series Models", in Latent Variables in Socio-Economic Models, eds. D. J. Aigner and A. S. Goldberger. North-Holland.

[4] Geweke, J. and Singleton, K. (1981). "Maximum likelihood confirmatory factor analysis of economic time series". International Economic Review, 22, 37-54.

[5] Handerson, H. V. and Searle, S. R. (1981). "On deriving the inverse of a sum of matrices". SIAM REVIEW, 23, 1, 53-60.

[6] Hannan, E. J. (1981). "Estimating the dimension of a linear system". Journal of Multivariate Analysis, 11, 459-473.

[7] Pena, D. and Sanchez-Albornoz, N. (1983). "Wheat Prices in Spain, 1857-1890: An application of the Box-Jenkins Methodology". Journal of European Economics History, (to appear).

[8] Priestley, M. B. (1981). Spectral Analysis and Time Series (Vol. 2) Academic Press.

[9] Quenouille, M. H. (1957). The Analysis of Multiple Time Series, London: Griffin.

[10] Reinsel, G. (1983). "Some results on multivariate autoregressive index models". Biometrika, 70, 145-156.

[11] Sanchez-Albornoz, N. (1975). Los precios agricolas durante la segunda mitad del siglo XIX. Banco de España.

[12] Sargent, T. J. and Sims, G. A. (1977). "Business cycle modeling without pretending to have too much a priori economic theory", in New Methods in Business Cycle Research: Proceeding from a Conference, ed. C. A. Sims, Minneapolis, MN; Federal Reserve Bank of Minneapolis.

[13] Stewart, G. W. (1973). Introduction to Matrix Computations, Academic Press.

[14] Subba Rao, T. (1975). "An innovation approach to the reduction of the dimensions in a multivariate stochastic system". Int. J. Control, 21, 673-680.

[15] Tiao, G. C. and Box, G. E. P. (1981). "Modeling Multiple Time Series with Applications". JASA, 76, 376, 802-816.

DP/GEPB/ed

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2722	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER AD-A149 495
4. TITLE (and Subtitle) A FACTOR ANALYSIS FOR TIME SERIES		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Daniel Peña and George E. P. Box		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Statistics and Probability
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE July 1984 13. NUMBER OF PAGES 55
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Multivariate ARMA process, canonical analysis, eigenvalues, eigenvectors

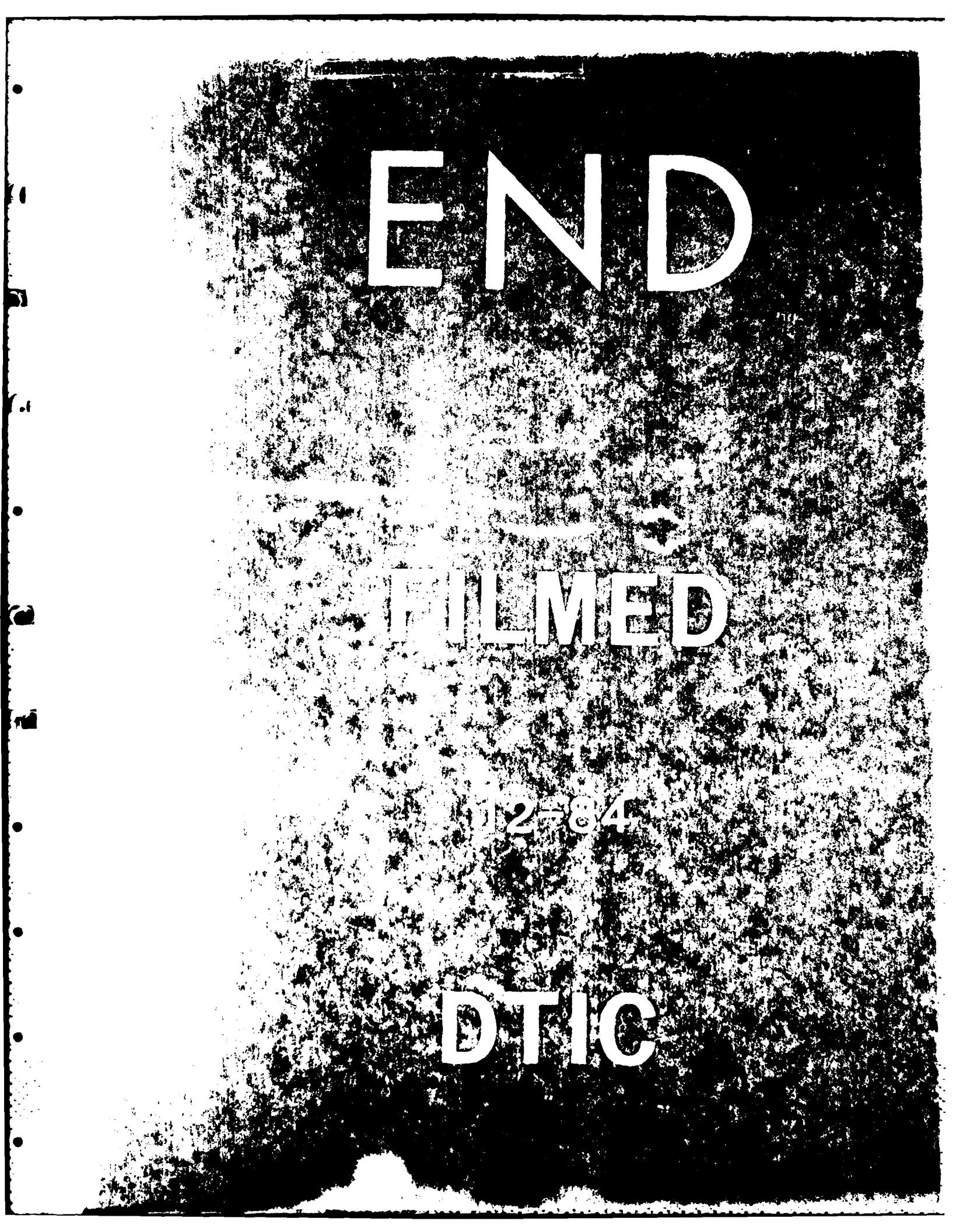
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper studies how to identify hidden factors in multivariate time series process. It is shown that the number of factors must be equal to the rank of both the covariance matrices and the parameter matrices of the infinite moving average representation of the process. A canonical transformation is derived which can recover such factors. The method is illustrated with several examples.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)


END
FILMED
12-84
DTIC