Biometrika (1977), 64, 2, *pp.* 355–65
With 1 *text-figure*
Printed in Great Britain

A canonical analysis of multiple time series

BY G. E. P. BOX AND G. C. TIAO
Department of Statistics, University of Wisconsin, Madison

SUMMARY

This paper proposes a canonical transformation of a k-dimensional stationary autoregressive process. The components of the transformed process are ordered from least to most predictable. The least predictable components are often nearly white noise which can reflect stable contemporaneous relationships among the original variables. The most predictable can be nearly nonstationary representing the dynamic growth characteristic of the series. The method is illustrated with a series with five variables.

Some key words: Autoregressive process; Canonical variable; Eigenvalue; Eigenvector; Multiple time series; Variance component.

1. INTRODUCTION

Data frequently occur in the form of k related time series simultaneously observed at some constant interval. In particular, economic, industrial and ecological data are often of this kind. Much work has been done on the problem of detecting, estimating and describing relationships of various kinds among such series; see, for example, Quenouille (1957), Hannan (1970), Box & Jenkins (1970), and Brillinger (1975). In this paper we shall consider a particular method for characterizing structure.

Consider a $k \times 1$ vector process $\{\mathscr{Z}_t\}$ and let $z_t = \mathscr{Z}_t - \mu$, where μ is a convenient $k \times 1$ vector of origin which is the mean if the process is stationary. Suppose z_t follows the pth order multiple autoregressive model

$$z_t = \hat{z}_{t-1}(1) + a_t, \tag{1·1}$$

where

$$\hat{z}_{t-1}(1) = E(z_t|z_{t-1}, z_{t-2}, \ldots) = \sum_{l=1}^{p} \pi_l z_{t-l}$$

is the expectation of z_t conditional on past history up to time $t-1$, the π_l are $k \times k$ matrices, $\{a_t\}$ is a sequence of independently and normally distributed $k \times 1$ vector random shocks with mean zero and covariance matrix Σ, and a_t is independent of $\hat{z}_{t-1}(1)$. The model (1·1) can be written

$$\left(I - \sum_{l=1}^{p} \pi_l B^l\right) z_t = a_t, \tag{1·2}$$

where I is the identity matrix and B is the backshift operator such that $Bz_t = z_{t-1}$. The process $\{z_t\}$ is stationary if the determinantal polynomial in B, det $(I - \Sigma \pi_l B^l)$, has all its zeros lying outside the unit circle, and otherwise the process will be called nonstationary.

Now suppose $k = 1$. Then, if the process is stationary,

$$E(z_t^2) = E\{\hat{z}_{t-1}(1)\}^2 + E(a_t^2),$$

that is

$$\sigma_z^2 = \sigma_{\hat{z}}^2 + \sigma_a^2.$$

We can define a quantity λ measuring the predictability of a stationary series from its past as $\lambda = \sigma_{\hat{z}}^2 \sigma_z^{-2} = 1 - \sigma_a^2 \sigma_z^{-2}$.

Suppose that we are considering k different stock market indicators such as the Dow Jones Average, the Standard and Poors index, etc., all of which exhibit dynamic growth. It is natural to conjecture that each might be represented as some aggregate of one or more common inputs which may be nearly nonstationary, together with other stationary or white noise components. This leads one to contemplate linear aggregates of the form $u_t = m'z_t$, where m is a $k \times 1$ vector. The aggregates which depend most heavily on the past, namely having large λ, may serve as useful composite indicators of the overall growth of the stock market. By contrast, the aggregates with λ nearly zero may reflect stable contemporaneous relationships among the original indicators. The analysis given in this paper yields k 'canonical' components from least to most predictable. The most predictable components will often approach nonstationarity and the least predictable will be stationary or independent. Thus we may usefully decompose the k-dimensional space of the observation z_t into independent, stationary and nonstationary subspaces. Variables in the nonstationary space represent dynamic growth while those in the stationary and independent spaces can reflect relationships which remain stable over time.

2. Choice of the canonical variables

2·1. *General considerations*

Suppose that z_t is stationary. Let $\Gamma_j(z) = E(z_{t-j} z_t')$ be the lag j autocovariance matrix of z_t. It follows from (1·1) that

$$\Gamma_0(z) = \sum_{l=1}^{p} \pi_l \Gamma_l(z) + \Sigma = \Gamma_0(\hat{z}) + \Sigma, \tag{2·1}$$

say, where $\Gamma_0(\hat{z})$ is the covariance matrix of $\hat{z}_{t-1}(1)$. Until further notice, we shall assume that Σ, and therefore $\Gamma_0(z)$, are positive-definite.

Now, consider the linear combination $u_t = m'z_t$. We have that $u_t = \hat{u}_{t-1}(1) + v_t$, where $\hat{u}_{t-1}(1) = m'\hat{z}_{t-1}(1)$ and $v_t = m'a_t$. The predictability of u_t from its past is therefore measured by

$$\lambda = \sigma_{\hat{u}}^2 \sigma_u^{-2} = \{m'\Gamma_0(\hat{z})\, m\}/\{m'\Gamma_0(z)\, m\}. \tag{2·2}$$

It follows that for maximum predictability, λ must be the largest eigenvalue of $\Gamma_0^{-1}(z)\,\Gamma_0(\hat{z})$ and m the corresponding eigenvector. Similarly, the eigenvector corresponds to the smallest eigenvalue will yield the least predictable combination of z_t.

2·2. *The canonical transformation*

Let $\lambda_1, \ldots, \lambda_k$ be the k real eigenvalues of the matrix $\Gamma_0^{-1}(z)\,\Gamma_0(\hat{z})$. Suppose that the λ_j are ordered with λ_1 the smallest, and that the k corresponding linearly independent eigenvectors, $m_1', \ldots, m_k'$, form the k rows of a matrix M. Then, we can construct a transformed process $\{y_t\}$, where

$$y_t = \hat{y}_{t-1}(1) + b_t, \tag{2·3}$$

with

$$y_t = Mz_t, \quad b_t = Ma_t, \quad \hat{y}_{t-1}(1) = \sum_{l=1}^{p} \tilde{\pi}_l y_{t-l}, \quad \tilde{\pi}_l = M\pi_l M^{-1}.$$

Corresponding to (2·1), we now have

$$\Gamma_0(y) = \Gamma_0(\hat{y}) + \tilde{\Sigma}, \tag{2·4}$$

where $\Gamma_0(y) = M\Gamma_0(z)\, M'$, $\Gamma_0(\hat{y}) = M\Gamma_0(\hat{z})\, M'$ and $\tilde{\Sigma} = M\Sigma M'$.

Note that: (i)

$$M'^{-1}\Gamma_0^{-1}(z)\,\Gamma_0(\hat{z})\, M' = \Lambda, \quad M'^{-1}\Gamma_0^{-1}(z)\,\Sigma M' = I - \Lambda, \tag{2·5}$$

where Λ is the $k \times k$ diagonal matrix with elements $(\lambda_1, \ldots, \lambda_k)$; (ii) $0 \leqslant \lambda_j < 1$ $(j = 1, \ldots, k)$; and (iii) for $i \neq j$, $m_i' \Sigma m_j = m_i' \Gamma_0(\hat{z}) m_j = 0$. In other words, $M\Sigma M'$, $M\Gamma_0(\hat{z})M'$ and, therefore, $M\Gamma_0(z)M'$ are all diagonal. Thus, the transformation (2·3) produces k components series $\{y_{1t}, \ldots, y_{kt}\}$ which (i) are ordered from least predictable to most predictable, (ii) are contemporaneously independent, (iii) have predictable components $\{\hat{y}_{1(t-1)}(1), \ldots, \hat{y}_{k(t-1)}(1)\}$ which are contemporaneously independent, and (iv) have unpredictable components $\{b_{1t}, \ldots, b_{kt}\}$ which are contemporaneously and temporally independent.

2·3. *Zero roots*

Special interest may attach to situations where certain of the λ_j approach zero. When the k_1 roots, $\lambda_1, \ldots, \lambda_{k_1}$, are zero, the matrix $\Gamma_0(\hat{y})$ in (2·4) then can be written

$$\Gamma_0(\hat{y}) = \begin{bmatrix} 0 & 0 \\ 0 & D \end{bmatrix}, \tag{2·6}$$

where D is an $k_2 \times k_2$ diagonal matrix. With

$$y_t' = [y_{1t}' \vdots y_{2t}'], \quad b_t' = [b_{1t}' \vdots b_{2t}'], \tag{2·7}$$

where y_{1t} and b_{1t} are $k_1 \times 1$ vectors, it readily follows that, with probability one, $y_{1t} = b_{1t}$. For $l = 1, \ldots, p$, partitioning the $k \times k$ matrix $\tilde{\pi}_l$ into

$$\tilde{\pi}_l = \begin{bmatrix} \tilde{\pi}_{11}^{(l)} & \tilde{\pi}_{12}^{(l)} \\ \tilde{\pi}_{21}^{(l)} & \tilde{\pi}_{22}^{(l)} \end{bmatrix}, \tag{2·8}$$

where $\tilde{\pi}_{11}^{(l)}$ is a $k_1 \times k_1$ matrix, we have the transformed series $\{y_t\}$ expressed as

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \sum_{l=1}^{p} \begin{bmatrix} 0 & 0 \\ \tilde{\pi}_{21}^{(l)} & \tilde{\pi}_{22}^{(l)} \end{bmatrix} B^l \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \begin{bmatrix} b_{1t} \\ b_{2t} \end{bmatrix}. \tag{2·9}$$

Thus, the canonical transformation decomposes the original $k \times 1$ vector process $\{z_t\}$ into two parts: (i) a part $\{y_{1t}\}$ which follows a k_1-dimensional white noise process, and (ii) a part $\{y_{2t}\}$ which is stationary but whose predictable part depends on both $y_{1(t-l)}$ and $y_{2(t-l)}$ for $l = 1, \ldots, p$.

The practical importance of (2·9) is that it implies that there are k_1 relationships between the original variables of the 'static' form

$$m_{j1}\mathscr{X}_{1t} + \ldots + m_{jk}\mathscr{X}_{kt} = \eta_j + b_{jt} \quad (j = 1, \ldots, k_1),$$

where the b_{jt} are contemporaneously and temporally independent. We shall later illustrate this situation with an example.

3. The first-order autoregressive process

3·1. *The canonical model*

In this section we discuss some properties of the canonical transformation when $\{z_t\}$ follows an k-dimensional autoregressive process of order one. Thus with $p = 1$ and $\pi_1 = \phi$, (1·1) yields

$$z_t = \hat{z}_{t-1}(1) + a_t = \phi z_{t-1} + a_t. \tag{3·1}$$

Since $\Gamma_1'(z) = \phi\Gamma_0(z)$ it follows from (2·1) that $\Gamma_0(z) = \phi\Gamma_0(z)\phi' + \Sigma$ and the required roots λ_j and vectors m_j are the k eigenvalues and eigenvectors of the matrix

$$Q = \Gamma_0^{-1}(z)\,\phi\Gamma_0(z)\,\phi'. \tag{3·2}$$

If $\tilde{\phi} = M\phi M^{-1}$ the transformed process can now be written

$$y_t = \tilde{\phi} y_{t-1} + b_t. \tag{3·3}$$

3·2. *Nonstationary series and unit roots*

In the above we have assumed that z_t is stationary. In practice, many time series exhibit nonstationary behaviour. A useful class of models to represent nonstationary series may be obtained by allowing the zeros of the det $(I-\Sigma \pi_l B^l)$ of (1·2) to lie on the unit circle. For the model (3·1), let $\alpha_1, \ldots, \alpha_k$ be the eigenvalues of the matrix ϕ. Then

$$\det(I-\phi B) = \prod_{j=1}^{k}(1-\alpha_j B),$$

so that the zeros of det $(I-\phi B)$ are simply $\alpha_1^{-1}, \ldots, \alpha_k^{-1}$. If one of more of the α_j are on the unit circle, then $\Gamma_0(z)$ does not exist and the canonical transformation method will break down. However, it is of interest to study the limiting situation when k_2 of the α_j approach values on the unit circle. Letting

$$y'_t = [y'_{1t} \vdots y'_{2t}], \quad b'_t = [b'_{1t} \vdots b'_{2t}], \quad \mathring{\phi} = \begin{bmatrix} \mathring{\phi}_{11} & \mathring{\phi}_{12} \\ \mathring{\phi}_{21} & \mathring{\phi}_{22} \end{bmatrix},$$

where y_1 and b_{1t} are $k_1 \times 1$ vectors and $\mathring{\phi}_{11}$ is a $k_1 \times k_1$ matrix with $k_1 = k - k_2$, we show in the Appendix that:

(i) if, and only if, k_2 of the eigenvalues α_j of ϕ approach values on the unit circle, then k_2 of the eigenvalues λ_j of Q in (3·2) approach unity;

(ii) the transformed model for y_t is, in the limit,

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \mathring{\phi}_{11} & 0 \\ \mathring{\phi}_{21} & \mathring{\phi}_{22} \end{bmatrix} B \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \begin{bmatrix} b_{1t} \\ b_{2t} \end{bmatrix}. \tag{3·4}$$

The canonical transformation therefore decomposes z_t into two parts:

(i) a part y_{1t} which follows a stationary first-order autoregressive process, and

(ii) a part y_{2t} which is approaching nonstationarity and also depends on $y_{1(t-1)}$.

The practical significance of this result is that the components y_{2t} can serve as useful composite indicators of the overall dynamic growth of the original series.

3·3. *Zero and unit roots*

For the model (3·1), suppose that k_1 of the λ_j are zero, k_3 of them approach unity and the remaining $k_2 = k - k_1 - k_3$ are intermediate in size. Then, from the results in (2·9) and (3·4), and upon partitioning y_t, b_t and $\mathring{\phi}$ into

$$y'_t = [y'_{1t} \vdots y'_{2t} \vdots y'_{3t}], \quad b'_t = [b'_{1t} \vdots b'_{2t} \vdots b'_{3t}],$$

$$\mathring{\phi} = \begin{bmatrix} \mathring{\phi}_{11} & \mathring{\phi}_{12} & \mathring{\phi}_{13} \\ \mathring{\phi}_{21} & \mathring{\phi}_{22} & \mathring{\phi}_{23} \\ \mathring{\phi}_{31} & \mathring{\phi}_{32} & \mathring{\phi}_{33} \end{bmatrix}, \tag{3·5}$$

where y_{1t} and b_{1t} are $k_1 \times 1$ vectors, y_{2t} and b_{2t} are $k_2 \times 1$ vectors, and $\mathring{\phi}_{11}$ and $\mathring{\phi}_{22}$ are, respectively, $k_1 \times k_1$ and $k_2 \times k_2$ matrices, the transformed process $\{y_t\}$ takes the form

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ \mathring{\phi}_{21} & \mathring{\phi}_{22} & 0 \\ \mathring{\phi}_{31} & \mathring{\phi}_{32} & \mathring{\phi}_{33} \end{bmatrix} B \begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} + \begin{bmatrix} b_{1t} \\ b_{2t} \\ b_{3t} \end{bmatrix}. \tag{3·6}$$

Thus there are: (i) a k_1-dimensional white noise process $\{y_{1t}\}$, (ii) a k_2-dimensional stationary process $\{y_{2t}\}$ such that the predictable part of y_{2t} depend only on $y_{1(t-1)}$ and $y_{2(t-1)}$, and (iii) a k_3-dimensional near nonstationary process $\{y_{3t}\}$ such that the predictable part of y_{3t} depends on $y_{1(t-1)}$, $y_{2(t-1)}$ and $y_{3(t-1)}$.

3·4. *Variance components for the first-order process*

Whatever the scaling of the transformed process $\{y_t\}$ in (3·3), since the jth element y_{jt} is $y_{jt} = \Sigma_i \tilde{\phi}_{ji} y_{i(t-1)} + b_{jt}$, where $(\tilde{\phi}_{j1}, \ldots, \tilde{\phi}_{jk})$ is the jth row of $\tilde{\phi}$, and since $y_{1(t-1)}, \ldots, y_{k(t-1)}$ and b_{jt} are independent, it follows that $\sigma^2_{y_j} = \Sigma_i \tilde{\phi}^2_{ji} \sigma^2_{y_i} + \sigma^2_{b_j}$. The contributions of $y_{1(t-1)}, \ldots, y_{k(t-1)}$ and b_{jt} to the variance of y_{jt} are, therefore, $\tilde{\phi}^2_{j1} \sigma^2_{y_1}, \ldots, \tilde{\phi}^2_{jk} \sigma^2_{y_k}$ and $\sigma^2_{b_j}$, respectively. It is convenient to consider these variance components in terms of their proportional contribution to $\sigma^2_{y_j}$, that is to consider $(\phi_{ji} \sigma^2_{y_i})/\sigma^2_{y_j}$ and $\sigma^2_{b_j}/\sigma^2_{y_j} = 1 - \lambda_j$. This can be done conveniently by arranging the canonical variables with scaling such that the variances of y_{jt} are all unity.

For the general process (1·1), to arrange for this scaling the matrix M must be chosen such that $M\Gamma_0(z) M' = I$. Corresponding to (2·3), let the transformed series in this scaling be written as

$$x_t = \hat{x}_{t-1}(1) + d_t. \tag{3·7}$$

Then, $\Gamma_0(x) = \Gamma_0(\hat{x}) + \bar{\Sigma}$, where $\Gamma_0(x) = I$, $\Gamma_0(\hat{x}) = \Lambda$, $\bar{\Sigma} = I - \Lambda$ and Λ is the diagonal matrix in (2·5). For the process (3·1), $\hat{x}_{t-1}(1) = \bar{\phi} x_{t-1}$, and hence

$$\bar{\phi}^2_{ji} = (\tilde{\phi}^2_{ji} \sigma^2_{y_i})/\sigma^2_{y_j}, \quad \bar{\phi}\bar{\phi}' = \Lambda. \tag{3·8}$$

In this scaling, then, the rows of $\bar{\phi}$ are orthogonal and the sum of squares of the jth row is λ_j.

The preceding canonical analysis will now be illustrated by an example.

4. An example

4·1. *U.S. hog, corn and wage series*

Quenouille (1957, pp. 88–101) studied a time series with 5 variates and containing 82 yearly observations from 1867–1948. He made adjustments where necessary, logarithmically transformed each variate and then linearly coded the logs, so as to produce numbers of comparable magnitude in the different series. His resulting five series denoted by $\mathscr{X}_{1t}, \ldots, \mathscr{X}_{5t}$ are plotted in Fig. 1 (*a*) and are identified in Table 4·1.

4·2. *The first-order autoregressive model*

Quenouille fitted the data to a first-order autoregressive process but was doubtful as to the adequacy of the model. We found, however, that the fit can be improved by appropriately shifting series 2 and 5 backward by one period as indicated in Table 4·1.

With the model $z_t = \phi z_{t-1} + a_t$ in (3·1), where $z_t = \mathscr{Z}_t - \mu$, the sample means $\hat{\mu}$ of $\mathscr{Z}_t$ and sample cross-covariance matrices C_j needed in our analysis are as follows:

$$10^{-3}\hat{\mu}' = (0·6989,\ 0·8949,\ 0·7714,\ 1·3281,\ 0·9956),$$

$$10^{-4}C_0 = \begin{bmatrix} 0·6831 & 1·2523 & 0·6535 & 0·9533 & 1·5224 \\ & 6·1939 & 3·7845 & 2·0209 & 5·5708 \\ & & 3·6877 & 0·2633 & 3·4746 \\ & & & 2·1407 & 2·1925 \\ & & & & 5·7206 \end{bmatrix},$$

$$10^{-4}C_1 = \begin{bmatrix} 0·5864 & 1·3670 & 0·7513 & 0·8632 & 1·5151 \\ 1·2038 & 5·2334 & 3·1639 & 1·8849 & 5·0392 \\ 0·4616 & 3·5820 & 2·7173 & 0·5605 & 3·0633 \\ 1·0108 & 1·8972 & 0·8338 & 1·6260 & 2·2508 \\ 1·3993 & 5·1586 & 3·2153 & 1·9817 & 5·3246 \end{bmatrix}.$$

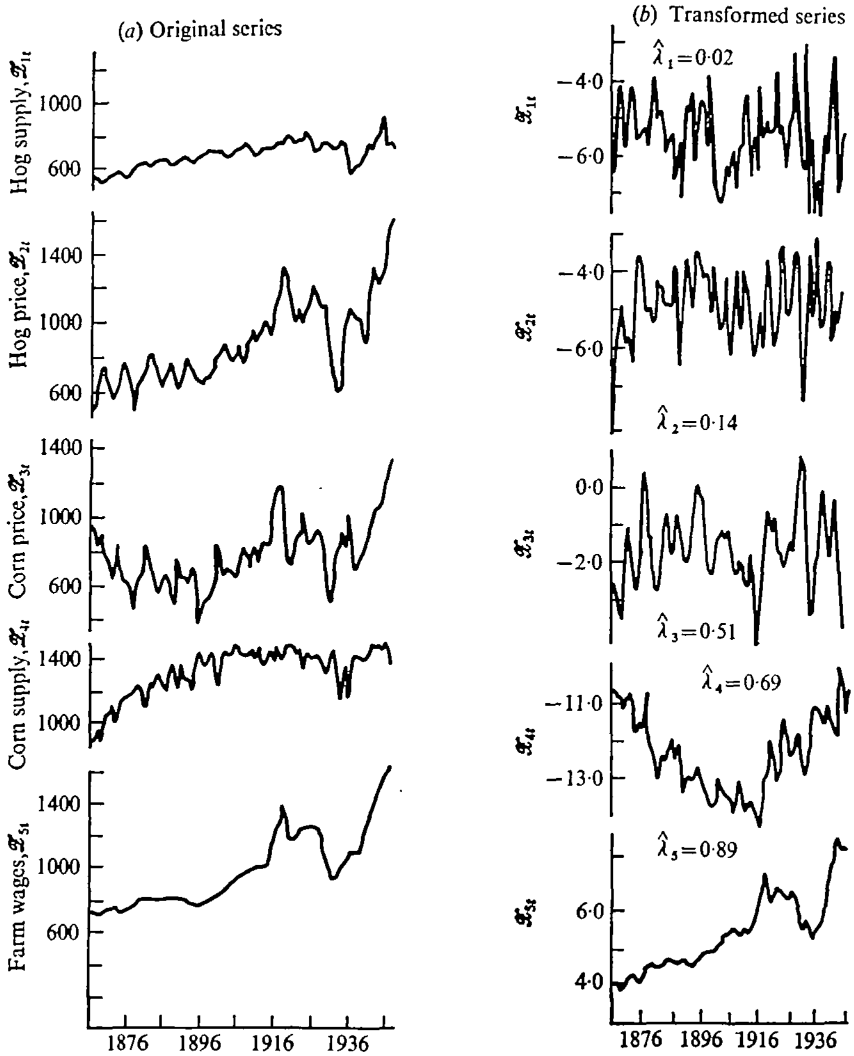

Fig. 1. U.S. hog data. (a) Original series, (b) transformed series.

Table 4·1. *Notation for Quenouille's U.S. hog series*

Variate	Symbol	As logged and linearly coded by Quenouille	Used in our analysis
Hog supply	H_s	$\mathscr{X}_{1t}$	$\mathscr{Z}_{1t} = \mathscr{X}_{1t}$
Hog price	H_p	$\mathscr{X}_{2t}$	$\mathscr{Z}_{2t} = \mathscr{X}_{2(t+1)}$
Corn price	R_p	$\mathscr{X}_{3t}$	$\mathscr{Z}_{3t} = \mathscr{X}_{3t}$
Corn supply	R_s	$\mathscr{X}_{4t}$	$\mathscr{Z}_{4t} = \mathscr{X}_{4t}$
Farm wages	W	$\mathscr{X}_{5t}$	$\mathscr{Z}_{5t} = \mathscr{X}_{5(t+1)}$

Table 4·2. *Estimated eigenvalues and eigenvectors for the hog data*

j	λ_j	H_s	H_p	R_p	R_s	W		
1	0·0232	(1·0000	0·3876	−0·2524	−0·5896	−0·2665)	×	0·0284
2	0·1421	(0·2080	1·0000	−0·8614	−0·3382	−0·3655)	×	0·0111
3	0·5061	(0·8925	−0·6433	−0·8277	−0·4784	1·0000)	×	0·0074
4	0·6901	(−0·9358	−0·2410	−0·4391	−0·5614	1·0000)	×	0·0129
5	0·8868	(0·6687	−0·1206	−0·0134	0·0396	1·0000)	×	0·0039

Table 4·2 gives the estimated eigenvalues and eigenvectors of Q in (3·2). The latter are scaled according to (3·7) so that all the components of the transformed process $\{x_t\}$ have unit estimated variances. The transformed process is $x_t = \bar{\phi} x_{t-1} + d_t$ with the estimated $\bar{\phi}$ given by

$$\begin{bmatrix} 0\cdot1213 & -0\cdot0778 & 0\cdot0465 & -0\cdot0110 & 0\cdot0113 \\ 0\cdot2215 & 0\cdot2766 & -0\cdot1241 & -0\cdot0309 & 0\cdot0119 \\ -0\cdot0321 & 0\cdot3167 & 0\cdot6334 & 0\cdot0444 & -0\cdot0404 \\ 0\cdot0885 & -0\cdot0025 & -0\cdot0492 & 0\cdot8235 & 0\cdot0416 \\ -0\cdot0801 & 0\cdot0378 & 0\cdot0396 & -0\cdot0363 & 0\cdot9360 \end{bmatrix},$$

and the resulting series $\mathscr{X}_t = \hat{M}\mathscr{Z}_t$ are shown in Fig. 1 (*b*).

The estimated proportional contributions of $x_{1(t-1)}, \ldots, x_{5(t-1)}$ and d_{jt} to the variance of x_{jt} are given in Table 4·3.

Table 4·3. *Component variances of the transformed series*

	$x_{1(t-1)}$	$x_{2(t-1)}$	$x_{3(t-1)}$	$x_{4(t-1)}$	$x_{5(t-1)}$	d_{jt}
x_{1t}	0·015	0·006	0·002	0·000	0·000	0·977
x_{2t}	0·049	0·077	0·015	0·001	0·000	0·858
x_{3t}	0·001	0·100	0·401	0·002	0·002	0·494
x_{4t}	0·008	0·000	0·002	0·678	0·002	0·310
x_{5t}	0·006	0·001	0·002	0·001	0·876	0·113

We see from the above calculations that there is very little contribution to x_{1t} and x_{2t} from history. These two transformed series are essentially white noise. The remarkable feature of x_{3t}, x_{4t} and x_{5t} is their heavy dependence on their own past, and this is especially so for the latter two components. It is almost true that x_{4t} and x_{5t} can be expressed as two independent univariate first-order autoregressive processes

$$x_{4t} = 0\cdot82 x_{4(t-1)} + d_{4t}, \quad x_{5t} = 0\cdot94 x_{5(t-1)} + d_{5t}. \tag{4·1}$$

4·3. *Interpretation*

In terms of the original observations $\mathscr{Z}_t$, the model for the most predictable component $\mathscr{X}_{5t}$ is approximately,

$$\mathscr{X}_{5t} - 0\cdot94 \mathscr{X}_{5t-1} = 0\cdot35 + d_{5t}. \tag{4·2}$$

The autoregressive parameter is close to unity, indicating that the series is nearly nonstationary Also, it is readily seen that the estimated standard deviation of the mean of $\mathscr{X}_{5t} - 0\cdot94\mathscr{X}_{5t-1}$ is 0·04 so that the term 0·35 on the right-hand side of (4·2) is real. Thus, what we have is nearly a random walk with a fixed increment of 0·35 per year. Now $x_{5t} = m_5' z_t$ and from the estimated eigenvector $\hat{m}_5$, x_{5t} is essentially a linear combination of the farm wages, W, and hog supply, H_s,

$$x_{5t} \simeq 0\cdot0039\,(z_{5t} + 0\cdot67 z_{1t}). \tag{4·3}$$

This is then the linear combination which serves as an indicator of the overall dynamic growth pattern in the original series.

The nearly random components x_1 and x_2, omitting the subscript t, associated with small values of λ are also of considerable interest. Their existence implies that any linear combination of the component series in the hyperplane

$$Z = \alpha x_1 + \beta x_2 = c_1 z_1 + c_2 z_2 + c_3 z_3 + c_4 z_4 + c_5 z_5 \tag{4·4}$$

varies nearly independently about fixed means. In choosing the component it is natural to seek combinations which are scientifically meaningful.

Now the dollar value of the hogs sold is proportional to $H_p H_s$ and the dollar value of the corn needed to feed them is $R_p R_s$. If then a Z exists involving these dollar values it will be such that approximately $c_1 = c_2$ and $c_3 = c_4$. By least squares or otherwise it is easy to find the linear combination for which this is nearly true. Specifically, by setting $\alpha = 30{\cdot}01$ and $\beta = 59{\cdot}51$ we obtain the relationship

$$Z = z_1 + z_2 - 0{\cdot}78z_3 - 0{\cdot}73z_4 - 0{\cdot}48z_5. \qquad (4{\cdot}5)$$

That is to say Z in (4·5) is approximately independently distributed about a fixed mean.

Now the average estimated variance of $(z_{1t}, \ldots, z_{5t})$ is $3{\cdot}69 \times 10^4$. For comparison, we normalize the linear combination (4·5) by letting $u = l'z$, where $l = (1{\cdot}84)^{-1}[1, 1, -0{\cdot}78, -0{\cdot}73, -0{\cdot}48]$, so that $l'l = 1$. Since x_1 and x_2 have unit variance and are independent, the variance of u is $0{\cdot}1326 \times 10^4$. Compared with the average variability of the original series, we have thus obtained a remarkably stable contemporaneous relationship among the 5 original variables.

Taking antilogs of (4·5) this implies that

$$I_1 = \frac{H_p H_s}{(R_p R_s)^{0{\cdot}75} W^{0{\cdot}50}} \qquad (4{\cdot}6)$$

is approximately constant. The numerator is obviously a measure of return to the farmer and the denominator a measure of his expenditure. The analysis points to the near constancy of this relation reminding us of the 'economic law' that a viable business must operate so as to balance expenditure and income.

Suppose we choose $\alpha = 46{\cdot}51$ and $\beta = -137{\cdot}80$, we then obtain

$$Z = 1{\cdot}00z_1 - 1{\cdot}02z_2 + 0{\cdot}98z_3 - 0{\cdot}26z_4 + 0{\cdot}21z_5. \qquad (4{\cdot}7)$$

Again, if we normalized the combination by expressing $u = l'z$ with $u = (1{\cdot}76)^{-1}Z$ such that $l'l = 1$, the variance of u would be $0{\cdot}68 \times 10^4$.

Upon taking antilogs, this implies that, very approximately,

$$I_2 = H_s R_p / H_p \qquad (4{\cdot}8)$$

is constant, indicating that a stable relationship existed between hog supply and the price ratio (Wallace & Bressman, 1937, p. 342–50).

In addition, we note that the percentage coefficients of variation of I_1 and I_2, given approximately by $100 \log\{10\sigma(Z)\}$ are 16 % and 18 %, respectively. Thus, both indices are remarkably stable when it is remembered that over the time period studied, the individual elements in the indices underwent massive changes. For example, hog prices increased tenfold.

4·4. *Differencing of the data*

For the hog data, since each of the original series exhibit a growth pattern, questions might be raised as to whether one should difference the data first and then perform a canonical analysis of the differenced series. Indeed, if one were to analyze these series individually, one would be led to consider differencing z_1, z_2, z_4 and z_5. However, in analyzing multiple time series of this kind, it is useful to entertain the possibility that the dynamic pattern in the data may be due to a small subset of nearly nonstationary components and that there may exist stable contemporaneous linear relationships among the variables. If this is so, then differencing all the original series could lead to complications in the analysis. To illustrate, suppose we have the bivariate model

$$z_{1t} = z_{1(t-1)} + a_{1t}, \quad z_{2t} = \beta z_{1t} + a_{2t}, \qquad (4{\cdot}9)$$

so that each series individually will be nonstationary. If we considered the two differenced series, $w_{1t} = (1-B)z_{1t}$ and $w_{2t} = (1-B)z_{2t}$, we would have

$$w_{1t} = a_{1t}, \quad w_{2t} = \beta a_{1t} + a_{2t} - a_{2(t-1)}. \tag{4·10}$$

It is readily shown that while (4·9) can be expressed in the form of a bivariate first-order autoregressive process, the differenced series cannot be put into the autoregressive form (1·1) making the analysis more complicated.

5. Further considerations

5·1. *Singularity of the matrix* Σ

So far it has been assumed that the covariance matrix of a_t, Σ, in (2·1) is positive-definite. Situations occur when $\Gamma_0(z)$ is positive-definite but Σ is singular. Specifically, suppose that the rank of Σ is $k_1 < k$. This means that the $k \times 1$ vector process $\{z_t\}$ is in fact driven by a k_1-dimensional nonsingular random shock process. Then, it is readily seen that the k_2 roots $\lambda_{k_1+1}, \ldots, \lambda_k$ of $\Gamma_0^{-1}(z)\,\Gamma_0(\hat{z})$ are exactly equal to one, and the transformed covariance $\tilde{\Sigma}$ matrix of b_t in (2·3) takes the form

$$\tilde{\Sigma} = \begin{bmatrix} D_1 & 0 \\ 0 & 0 \end{bmatrix}, \tag{5·1}$$

where D_1 is an $k_1 \times k_1$ diagonal matrix with positive elements. Partitioning y_t, b_t and $\tilde{\pi}_l$ as given in (2·7) and (2·8), we see that $b_{2t} = 0$ with probability one. Thus the transformed model is

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \sum_{l=1}^{p} \begin{bmatrix} \tilde{\pi}_{11}^{(l)} & \tilde{\pi}_{12}^{(l)} \\ \tilde{\pi}_{21}^{(l)} & \tilde{\pi}_{22}^{(l)} \end{bmatrix} B^l \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} + \begin{bmatrix} b_{1t} \\ 0 \end{bmatrix}. \tag{5·2}$$

In other words, the k_2-dimensional vector y_{2t} is completely predictable from the past values $y_{1(t-l)}$ and $y_{2(t-l)}$ $(l = 1, \ldots, p)$.

In practice, situations may occur where Σ is nearly singular. From the results here and those discussed earlier in § 3·2, we see that for the first-order autoregressive process, certain of the roots λ_j will be nearly equal to one either when some of the eigenvalues of ϕ approach values on the unit circle or when Σ is nearly singular. The problem of how to distinguish between these two cases is currently being investigated.

5·2. *Singularity of the matrix* $\Gamma_0(z)$

Examples can also occur when $\Gamma_0(z)$ is singular. It is not unusual to find exact and quite complex linear relationships imposed by the method in which the data is put together so that $\Gamma_0(z)$ will necessarily be singular (Box *et al.*, 1973). Two situations can occur depending on whether or not the nature of any exact linear relationships existing in the data is already known. If known, then the problem may be avoided by eliminating, in advance of the analysis, any dependencies and applying the analysis to a linearly independent subset of r of the k series. When the nature of exact relationships in the data which might exist are not known, a principle component analysis of the estimate $C_0(z)$ of $\Gamma_0(z)$ should be first conducted. The existence of $k-r$ roots which are nearly zero indicates the existence of $k-r$ linearly independent exact relationships which define a hyperplane in the k space given by the $k-r$ corresponding eigenvectors. The canonical analysis of this paper may now be carried out on any subset of r linearly independent series which lie in the nonsingular space.

The work was supported by U.S. Army Research Office and the Wisconsin Alumni Research Foundation.

APPENDIX

Eigenvalues of first-order autoregressive process

We here sketch the proof of the results given in § 3·1 concerning the situation where k_2 of the eigenvalues of ϕ approach values on the unit circle.

THEOREM. *Suppose that z_t follows the stationary model in* (3·1), *where the covariance matrix Σ of a_t is positive-definite. A sufficient and necessary condition for k_2 of the eigenvalues of*

$$\Gamma_0(z)^{-1}\phi\Gamma_0(z)\phi'$$

to tend to unity is that k_2 of the eigenvalues of ϕ approach values on the unit circle.

Proof. Let $k = k_1+k_2$ and the eigenvalues of ϕ be divided into two sets $\alpha_1^* = (\alpha_1, \ldots, \alpha_{k_1})$ and $\alpha_2^* = (\alpha_{k_1+1}, \ldots, \alpha_k)$ with no common element and such that α_j and its complex conjugate belong to the same set. The characteristic polynomial of ϕ can be written as the product

$$f(\alpha) = f_{k_1}(\alpha)\, f_{k_2}(\alpha), \tag{A 1}$$

where

$$f_{k_1}(\alpha) = \alpha^{k_1} - \gamma_1\alpha^{k_1-1} - \ldots - \gamma_{k_1}, \quad f_{k_2}(\alpha) = \alpha^{k_2} - s_1\alpha^{k_2-1} - \ldots - s_{k_2}$$

are real polynomials of degrees k_1 and k_2 with roots α_1^* and α_2^*, respectively. Now there exists a $k\times k$ real nonsingular matrix C such that $C\phi C^{-1}$ is of the block diagonal form

$$C\phi C^{-1} = \begin{bmatrix} R & 0 \\ 0 & S \end{bmatrix}, \tag{A 2}$$

where R and S are, respectively, $k_1\times k_1$ and $k_2 \times k_2$ matrices such that

$$R = \left[\begin{array}{c:c} 0 & I \\ \hdashline \gamma_{k_1} & \gamma_{k_1-1} \cdots \gamma_1 \end{array}\right], \quad S = \left[\begin{array}{c:c} 0 & I \\ \hdashline s_{k_2} & s_{k_2-1} \cdots s_1 \end{array}\right].$$

Letting $V = C\Gamma_0(z)C'$ and $W = C\Sigma C'$ and partitioning V and W correspondingly, so that, for example, V_{11} and W_{11} are $k_1\times k_1$ matrices, we obtain

$$V_{11} = RV_{11}R' + W_{11}, \quad V_{12} = RV_{12}S' + W_{12}, \quad V_{22} = SV_{22}S' + W_{22}, \tag{A 3}$$

where we use the relation $\Gamma_0(z) = \phi\Gamma_0(z)\phi' + \Sigma$. By writing $V_{22} = AA'$, it is readily seen from (3·2) that the λ_j are the roots of the determinantal polynomial

$$\det\left\{(1-\lambda)\begin{bmatrix} V_{11} & V_{12}A'^{-1} \\ A^{-1}V_{12}' & I \end{bmatrix} - \begin{bmatrix} W_{11} & W_{12}A'^{-1} \\ A^{-1}W_{12}' & A^{-1}W_{22}A'^{-1} \end{bmatrix}\right\} = 0. \tag{A 4}$$

To prove sufficiency, we need to show that if the k_2 eigenvalues

$$\alpha_j \to e^{i\omega_j} \quad (j = k_1+1, \ldots, k), \tag{A 5}$$

then (A 4) will tend to

$$(1-\lambda)^{k_2}\det\{(1-\lambda)V_{11} - W_{11}\} = 0. \tag{A 6}$$

It suffices to prove that (A 5) implies that

$$A^{-1} \to 0, \quad A^{-1}V_{12}' \to 0. \tag{A 7}$$

From (A 3)

$$V_{22}^{-1}SV_{22}S' = \{I + S'^{-1}V_{22}^{-1}S^{-1}W_{22}\}^{-1} \tag{A 8}$$

so that $\det S^2 = \det(I+S'^{-1}V_{22}^{-1}S^{-1}W_{22})^{-1}$. When (A 5) holds, $\det(S^2) = s_{k_2}^2 \to 1$. Since S is nonsingular and W_{22} is positive-definite, it follows that $V_{22}^{-1}\to 0$ and hence $A^{-1}\to 0$.

To show $A^{-1}V_{12}'\to 0$, we have, from (A 3),

$$I = PP' + A^{-1}W_{22}A'^{-1}, \tag{A 9}$$

$$A^{-1}V_{12}' = PA^{-1}V_{12}'R' + A^{-1}W_{12}', \tag{A 10}$$

where $P = A^{-1}SA$. Thus, when (A 5) holds, in (A 9), $P \to P_0$ where P_0 is an orthogonal matrix, and hence (A 10) becomes

$$P_0' A^{-1} V_{12}' = A^{-1} V_{12}' R. \tag{A 11}$$

Since by supposition, S and R have no common eigenvalues, it follows (Gantmacher, 1959, p. 220) that $A^{-1}V_{12}' \to 0$. This completes the proof of the sufficiency.

Next to show necessity, recall that $|\alpha_j| < 1$ and $0 \leqslant \lambda_j < 1$, for $j = 1, \ldots, k$. Thus, if k_2 of the λ_j tends to one, then exactly k_2 of the α_j must approach values on the unit circle. For, if otherwise, and suppose $k' \neq k_2$ of the α_j approach values on the unit circle, then from the sufficiency part of the theorem which we have just proved, k' of the λ_j must approach one, which contradicts the supposition. The theorem thus follows.

To study the eigenvectors and the transformed matrix $\tilde{\phi}$, it is easy to see that the systems of equations $(Q - \lambda I)\, m = 0$ is equivalent to

$$\{(1-\lambda) I - V^{-1} W\} h = 0, \tag{A 12}$$

where $C'h = m$. When (A 5) holds, by using (A 7) it is straightforward to verify that the matrix of eigenvectors M' must be of the form

$$C'^{-1} M' = \begin{bmatrix} H_{11}' & -W_{11}^{-1} W_{12} \\ 0 & I \end{bmatrix}, \tag{A 13}$$

where the columns of H_{11}' are the eigenvectors of $V_{11}^{-1} W_{11}$.

It follows from (A 13) that the transformed matrix $\tilde{\phi}$ takes the form

$$\tilde{\phi} = M \phi M^{-1} = \begin{bmatrix} \tilde{\phi}_{11} & 0 \\ \tilde{\phi}_{21} & \tilde{\phi}_{22} \end{bmatrix}, \tag{A 14}$$

where $\tilde{\phi}_{11} = H_{11} R H_{11}^{-1}$, $\tilde{\phi}_{22} = S_0$, $\tilde{\phi}_{21} = (S_0 W_{12}' W_{11}^{-1} - W_{11}^{-1} R) H_{11}^{-1}$ and S_0 is the limiting matrix of S when all its roots approach values on the unit circle.

References

Box, G. E. P., Erjavec, J., Hunter, W. G. & MacGregor, J. F. (1973). Some problems associated with the analysis of multiresponse data. *Technometrics* **15**, 33–51.

Box, G. E. P. & Jenkins, G. M. (1970). *Time Series Analysis Forecasting and Control*. San Francisco: Holden-Day.

Brillinger, D. R. (1975). *Time Series Data Analysis and Theory*. New York: Holt, Rinehart and Winston.

Gantmacher, F. R. (1959). *The Theory of Matrices*, Vol. I. New York: Chelsea.

Hannan, E. J. (1970). *Multiple Time Series*. New York: Wiley.

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

Wallace, H. A. & Bressman, E. N. (1937). *Corn and Corn Growing*, 4th edition. New York: Wiley.

[*Received November* 1975. *Revised November* 1976]