RM-77-33

A MULTIVARIATE TIME SERIES APPROACH TO MODELLING MACROECONOMIC SEQUENCES

Johannes Ledolter

May 1977

PREFACE

Modelling of economic systems is an important task in the research program of the System and Decision Sciences area at IIASA (Task 1 of the Research Plan for System and Decision Sciences area for 1977).

The classical econometric approach to modelling and prediction of economic systems uses single and simultaneous equation models to represent the relationships among economic variables which are postulated by economic theory. The present paper explores an alternative approach and investigates whether multivariate time series methodology can provide useful tools for the assessment of economic relationships and the short term prediction of economic variables.

ABSTRACT

In this paper we discuss a multivariate generalization of autoregressive integrated moving average models. A methodology for constructing multivariate time series models is developed and the derivation of forecasts from such models is considered. A bivariate model for Austrian macroeconomic sequences is constructed. Furthermore it is discussed whether multivariate time series methods can be expected to lead to a significant increase in prediction accuracy for macroeconomic series.

A Multivariate Time Series Approach To Modelling Macroeconomic Sequences

1. Introduction

Since 1976 the Austrian Institute of Economic Research (Österreichisches Institut für Wirtschaftsforschung) has been using univariate time series methods, commonly known under the name of Box-Jenkins analysis, to derive short term predictions of macroeconomic series (Ledolter, Schebeck and Thury [19], [28]). The experience with these techniques over the last year has been excellent and confirms the results of many empirical comparisons which show that simple univariate Box-Jenkins forecasts are quite accurate and compare very favourably with predictions from econometric models (Christ [7], Cooper [8], Narasimham and Singpurwalla [20], Naylor, Seaks and Wichern [21], Nelson [22], Prothero and Wallis [25]).

It is not the objective of this paper to add a further study to this long list of empirical comparisons and to establish a priority of one method over the other. Both methods, the relatively simple univariate Box-Jenkins models which use information only from the past history and the elaborate more time consuming econometric models which incorporate latest economic theory, are not seen in competition but as complementing each other.

It is nevertheless surprising how well univariate Box-Jenkins procedures perform considering their relative simplicity and considering that they utilize information only of their own past and do not incorporate the information from other, possibly related series.

Univariate methodology provides the building block for multivariate modelling which considers more series at the same time. A question which arises and which is addressed in this paper is to what extent the predictions from the univariate time series models can be improved by enlarging the information set and considering several series jointly.

The paper consists of several parts. In the second section univariate seasonal and non seasonal ARIMA models are reviewed. In the third section a multivariate generalization of the class of ARIMA model is considered and illustrated with simple examples. The model building methodology for multivariate time series models, in particular their specification, estimation and validation, is discussed. An example using Austrian total private consumption and disposable personal income data is considered in the fourth section of the paper and univariate and bivariate models are given. The last section consists of concluding remarks and a discussion whether multivariate time series methods can be expected to lead to a significant increase in the prediction accuracy for macro-economic series.

2. Univariate time series models

For the analysis of univariate nonseasonal time series z_t Box and Jenkins [3] use the class of autoregressive integrated moving average (ARIMA) models of the form

$$\phi(B)(1-B)^d z_t = \theta_0 + \theta(B)a_t \qquad (2.1)$$

where

$$\phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$$

$$\theta(B) = 1 - \theta_1 B - \ldots - \theta_q B^q$$

B is the backshift operator; $B^m z_t = z_{t-m}$

$\{a_t\}$ is a sequence of independent random variables (white noise sequence)

$$E(a_t) = 0 \quad ; \quad E(a_t a_{t+k}) = \begin{cases} \sigma^2 & \text{for } k = 0 \\ 0 & \text{for } k \neq 0 \end{cases} .$$

It is assumed that the roots of $\phi(B) = 0$ and $\theta(B) = 0$ lie outside the unit circle (stationarity and invertibility condition) and that they have no common roots. The differencing operator $(1 - B)$ is used to represent non stationary processes (i.e.: series which do not have a fixed level, slope etc., but which apart from this exhibit stationary behavior). The polynomial $\phi(B)$ which includes the parameters $\phi_1,\ldots,\phi_p$ is called autoregressive operator; the polynomial $\theta(B)$ with parameters $\theta_1,\ldots,\theta_q$ is called moving average operator. When $d = 0$ (no differencing) the original series is stationary and θ_0 allows for a nonzero mean; for $d \geq 1$ the parameter θ_0 is capable of representing a deterministic trend in the form of a polynomial of degree $d - 1$.

Economic series frequently have non stationary variance and in particular the variation often depends on the level of the series. If the variation is proportional to the level then the logarithmic transformation will stabilize the variance. In other cases, however, the logarithmic transformation might not be suitable and other transformations have to be tried. A particularly useful class of transformations is the class of power transformations introduced by Box and Cox [2]. Use of this parametric class, which includes the logarithmic transformation as a special case, in an economic time series context is, for example, made by Box and Jenkins [4], Tintner and Kadekodi [27].

For observations with a seasonal pattern the model in (2.1) has to be extended. Box and Jenkins [3] introduce multiplicative seasonal models

$$\phi(B)\phi_s(B^s)(1 - B)^d(1 - B^s)^D z_t = \theta_0 + \theta(B)\theta_s(B^s)a_t \quad , \qquad (2.2)$$

where $\phi(B)$ and $\theta(B)$ are as defined above and

$\phi_s(B^s) = 1 - \phi_{1,s}B^s - \ldots - \phi_{P,s}B^{Ps}$ is a polynomial of degree P in B^s

$\theta_s(B^s) = 1 - \theta_{1,s}B^s - \ldots - \theta_{Q,s}B^{Qs}$ is a polynomial of degree Q in B^s.

Since most quarterly economic series show a distinct seasonal pattern ($s = 4$) this class is important for the modelling of macroeconomic series.

Past experience in many fields shows that the class of ARIMA models (2.1) and their seasonal extension (2.2) are capable of representing many series observed in practice, both stationary and non stationary.

Predictions from ARIMA models

Given the model and the value of the parameters optimal forecasts (optimal in the minimum mean square error sense, i.e., providing unbiased forecasts which minimize the variance of the forecast error) are readily derived. It can be shown [3] that the minimum mean square error forecast of a future observation $z_{n+\ell}$, given all the information up to time n, is given by the conditional expectation

$$\hat{z}_n(\ell) = E(z_{n+\ell} | z_n, z_{n-1}, \ldots) \quad . \tag{2.3}$$

The predictions are easily interpreted considering the inverted form of model (2.1) (to simplify the discussion we assume that $\theta_0 = 0$)

$$\frac{\phi(B)(1-B)^d}{\theta(B)} z_t = a_t \quad . \tag{2.4}$$

Defining $\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \ldots = \frac{\phi(B)(1-B)^d}{\theta(B)}$ one can write

$$z_{n+\ell} = \sum_{j\geq 1} \pi_j z_{n+\ell-j} + a_{n+\ell} \quad .$$

Since the conditional expectation of any future $a_{n+\ell}$ ($\ell>0$) is zero, the one step ahead prediction error ($\ell=1$) is given by

$$\hat{z}_n(1) = \sum_{j\geq 0} \pi_{j+1} z_{n-j} = \pi_1 z_n + \pi_2 z_{n-1} + \pi_3 z_{n-2} + \ldots \tag{2.5}$$

For general ℓ

$$\hat{z}_n(\ell) = \sum_{j\geq 0} \pi_{j+1}^{(\ell)} z_{n-j} = \pi_1^{(\ell)} z_n + \pi_2^{(\ell)} z_{n-1} + \pi_3^{(\ell)} z_{n-2} + \cdots \tag{2.6}$$

where $\pi_j^{(\ell)}$ are functions of the original π-weights,

$$\pi_j^{(\ell)} = \pi_{j+\ell-1} + \sum_{k=1}^{\ell-1} \pi_k \pi_j^{(\ell-k)} \quad .$$

The π-weights which depend on the structure of the model and on the values of the autoregressive and moving average parameters provide a weight function which discounts past information.

Philosophy of model building

For a given model the forecasts are readily derived. In practice, however, the form of the process is rarely, if ever, known and one has to use past observations to derive adequate models and to estimate their parameters.

Box and Jenkins [3] develop a three stage iterative procedure consisting of model specification, model fitting and model diagnostic checking to find members of the class of ARIMA models which are parsimonious in their parameters, but adequate for the description of the correlation structure of the data.

Since the class of ARIMA models is too extensive to be fitted directly to the data, model specification procedures employ the data (in terms of sample autocorrelations and sample partial autocorrelations) to suggest an appropriate parsimonious subclass of models which may be tentatively entertained. At the estimation stage the parameters of the tentatively entertained model are estimated (the programs which are used for the examples in this paper calculate maximum likelihood estimates conditional on zero starting values). At the third stage, the model validation stage, diagnostic checks are applied with the intent to reveal possible model inadequacies and to achieve improvement. The residuals (observed minus fitted values) contain the information about the adequacy of the fitted model. The sample autocorrelation function of the residuals indicates whether the entertained model is adequately describing the correlation structure of the data or if, and how, the model should be revised. After the model passes the diagnostic checks it can be used for interpretation and prediction.

3. Multivariate time series models

It was pointed out before that univariate time series models frequently face the criticism that they use information only of its own past and do not use the information from other sources.

For example, let us suppose that data on a pair of time series, z_1 and z_2, is available and we have to make a prediction of future values of z_1. One could use the past history of z_1 only and build a univariate model predicting future values of z_1 from its own past. Alternatively, one could use the larger information set $\{z_{1n-j}, z_{2n-j}; j = 0,1,2,\dots\}$ and build a multivariate model. One would hope that in this case superior forecasts can be obtained (superior in terms of Granger's [12] concept of predictability i.e., smaller variance of the forecast error).

If future values of z_1 are better forecast with an information set extended to include both present and past values of z_1 and z_2, but the forecast of z_2 is not improved by the addition of current and past z_1, then the series are said to exhibit no feedback. (Other terminologies such as unidirectional causality from z_2 to z_1, or z_2 being exogeneous relative to z_1 are sometimes used in the literature.) In this case transfer function models (dynamic regression models, distributed lag models) as discussed by Box and Jenkins [3] can be used.

$$z_{1t} = v(B) z_{2t} + n_t \quad ; \qquad (3.1)$$

where $v(B) = (\omega_0 - \omega_1 B - \dots - \omega_s B^s) | (1 - \delta_1 B - \dots - \delta_r B^r)$ and where both z_{2t} and n_t have ARIMA representations and where z_{2t} is independent of n_{t+k} (for all k).

If future values of z_2 as well as z_1 are better predicted by using the extended information set, the pair of series is said to exhibit feedback and multivariate models have to be used.

The m-dimensional time series generalization of univariate stationary autoregressive moving average models was first introduced by Quenouille [26], and further discussed by Hannan [15]

$$\Phi(B)\underset{\sim}{z}_t = \underset{\sim}{\theta}_0 + \Theta(B)\underset{\sim}{a}_t \quad , \qquad (3.2)$$

where

$\underset{\sim}{z}'_t = (z_{1t} z_{2t} \cdots z_{mt})$ is a m-dimensional vector of realizations at time t

$\underset{\sim}{a}'_t = (a_{1t} a_{2t} \cdots a_{mt})$ is a m-dimensional crosscorrelated white noise sequence

$E(\underset{\sim}{a}_t) = \underset{\sim}{0}$; $E(\underset{\sim}{a}_t \underset{\sim}{a}'_{t+k}) = \delta_0^k \Sigma$ where

δ_0^k is the Kronecker delta function

$$\delta_0^k = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases}$$

and Σ is a symmetric, positive definite $[m \times m]$ matrix.

$\underset{\sim}{\theta}'_0 = (\theta_{01} \theta_{02} \cdots \theta_{0m})$

$\Phi(B) = I - \Phi_1 B - \ldots - \Phi_p B^p$ is the autoregressive operator with autoregressive $[m \times m]$ matrix parameters $\Phi_1, \ldots, \Phi_p$.

$\Theta(B) = I - \Theta_1 B - \ldots - \Theta_q B^q$ is the moving average operator with moving average $[m \times m]$ matrix parameters $\Theta_1, \ldots, \Theta_q$.

I is the $[m \times m]$ identity matrix.

It is assumed that the roots of $\det\Phi(B) = 0$ lie outside the unit circle (stationarity condition) and that the roots of $\det\Theta(B) = 0$ lie on or outside the unit circle. Furthermore it is assumed that $\det\Phi(B) = 0$ has no common roots with $\det\Theta(B) = 0$. This condition, together with nonsingularity of Φ_p and Θ_q, will lead to an identified (unique) model. The above condition is further relaxed by Hannan [14] who gives necessary and sufficient conditions for the uniqueness of the parameters.

It is instructive to consider special cases of this general class of models.

(i) Bivariate first order autoregressive model

$$(I - \Phi B)\underset{\sim}{z}_t = \underset{\sim}{a}_t \tag{3.3}$$

where

$$\underset{\sim}{z}'_t = (z_{1t} z_{2t}) \qquad \underset{\sim}{a}'_t = (a_{1t} a_{2t})$$

$$\Phi = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \quad \text{and}$$

$$E(\underset{\sim}{a}_t \underset{\sim}{a}'_{t+k}) = \delta_0^k \Sigma \quad \text{with} \quad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} .$$

The model can be written as

$$\begin{cases} (1 - \phi_{11}B)z_{1t} - \phi_{12}z_{2t-1} = a_{1t} \\ (1 - \phi_{22}B)z_{2t} - \phi_{21}z_{1t-1} = a_{2t} \end{cases} . \tag{3.4}$$

The univariate first order autoregressive process is characterized by an exponentially decreasing autocovariance function. A similar pattern also holds for the multivariate AR(1) process, except that the dimensionality (now matrices instead of scalars) makes it more difficult to recognize the pattern.

$$\Gamma(k) = \Gamma(k-1)\Phi' \qquad k \geq 1$$
$$\Gamma(0) = \textstyle\sum + \Gamma(1)'[\Gamma(0)]^{-1}\Gamma(1) \tag{3.5}$$

where $\Gamma(k) = E(\underset{\sim}{z}_t \underset{\sim}{z}'_{t+k})$ is the lag k autocovariance matrix (note that here and in the following we assume that $\underset{\sim}{z}_t$ are already deviations from their means).

If all elements in Φ are different from zero the two series exhibit feedback and multivariate techniques have to be applied. In the case when one off diagonal element, let us take ϕ_{21}, is zero, the second series z_2 influences z_1, but in turn is not influenced by z_1 (no feedback from z_1 to z_2)

$$\begin{cases} (1-\phi_{11}B)z_{1t} - \phi_{12}Bz_{2t} = a_{1t} \\ (1-\phi_{22}B)z_{2t} = a_{2t} \end{cases} \tag{3.6}$$

a_{1t} and a_{2t} are in general correlated ($\sigma_{12} \neq 0$). However, it is possible to express

$$a_{1t} = a^*_{1t} + c\,a_{2t}$$

where $c = \frac{\sigma_{12}}{\sigma_{22}}$ and where a^*_{1t} and a_{2t} are uncorrelated. Substituting for a_{1t} in (3.6) gives

$$\begin{cases} (1-\phi_{11}B)z_{1t} - \phi_{12}B\,z_{2t} - c(1-\phi_{22}B)z_{2t} = a^*_{1t} \\ (1-\phi_{22}B)z_{2t} = a_{2t} \end{cases}$$

or

$$z_{1t} = \frac{1}{1-\phi_{11}B}\{c + (\phi_{12} - c\phi_{22})B\}z_{2t} + \frac{a^*_{1t}}{1-\phi_{11}B} \tag{3.7}$$

where z_{2t} is independent of a^*_{1t+k} for all k.

Model (3.7) is a transfer function (dynamic regression, distributed lag) model of the form (3.1) where z_{2t} can be considered input for z_{1t} and where the input is independent of the noise.

(ii) Bivariate first order moving average model

$$\underset{\sim}{z}_t = (I - \Theta B)\underset{\sim}{a}_t \tag{3.8}$$

where $\Theta = \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix}$ and $\underset{\sim}{z}_t$ and $\underset{\sim}{a}_t$ as above. The model can be written as

$$\begin{cases} z_{1t} = (1 - \theta_{11}B)a_{1t} - \theta_{12}a_{2t-1} \\ z_{2t} = (1 - \theta_{22}B)a_{2t} - \theta_{21}a_{1t-1} \end{cases} . \tag{3.9}$$

It can be shown that

$$\Gamma(0) = \Sigma + \Theta\Sigma\Theta' \tag{3.10}$$

$$\Gamma(1) = - \Sigma\Theta'$$

$$\Gamma(k) = 0 \quad \text{for} \quad k \geq 2 \quad .$$

The autocovariances of z_1 and z_2 as well as the cross covariances between z_1 and z_2 are zero from lag 2 onwards, a fact which is helpful in the specification stage of modelling. Similar to (i), it can be shown that there is no feedback from z_1 to z_2 iff $\theta_{21} = 0$.

(iii) Bivariate first order autoregressive moving average model

$$(I - \Phi B)\underset{\sim}{z}_t = (I - \Theta B)a_t \tag{3.11}$$

where $\Phi = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}$; $\Theta = \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix}$.

It can be shown that

$$\Gamma(0) = \Phi\Gamma(1) + \sum - \Theta\sum\Phi' + \Theta\sum\Theta' \tag{3.12}$$

$$\Gamma(1) = \Gamma(0)\Phi' - \sum\Theta'$$

$$\Gamma(k) = \Gamma(k-1)\Phi' \quad \text{for } k \geq 2 \quad .$$

In (3.11) we assume that the roots of det $\Phi(B) = 0$ lie outside the unit circle (stationarity condition) and that the roots of det $\Theta(B) = 0$ lie on or outside the unit circle.

These conditions, however, are not enough to identify the model parameters (i.e., there may be other values of Φ and Θ which lead to the same covariance structure and forecast weights). Simple cases for nonidentified models are, for example, given when $\Phi = \Theta$ or more general when there exists a matrix A for which $A\Phi = A\Theta = 0$.

If the model is not identified, not all parameters are estimable. For example, when $\Phi = \Theta$, only the difference of the elements of Φ and Θ is estimable, but not Φ and Θ individually. In practice, non identified models (near non identified models) will lead to an ill defined estimation problem resulting in high correlations among the parameter estimates.

If the model is used for prediction purposes the question of identifiability is not a critical one, since any $\Phi(=\Theta)$ will lead to the same prediction weights.

(iv) <u>Extension to seasonal and non stationary models</u>

Multiplicative seasonal models in m-dimensions can be written as

$$\Phi(B)\Phi_s(B^s)\underset{\sim}{z}_t = \underset{\sim}{\theta}_0 + \Theta(B)\Theta_s(B^s)\underset{\sim}{a}_t \tag{3.13}$$

where $\Phi(B)$ and $\Theta(B)$ are as in (3.2) and where

$$\Phi_s(B^s) = I - \Phi_{1s}B^s - \dots - \Phi_{P,s}B^{Ps}$$

$$\Theta_s(B^s) = I - \Theta_{1s}B^s - \dots - \Theta_{Q,s}B^{Qs}$$

In Section 2 it was shown how simplifying operators such as ordinary differences $(1-B)^d$, seasonal differences $(1-B^s)^D$, or in general operators with roots on the unit circle can be used to transform non stationary into stationary sequences.

(v) An interesting observation is originally due to Quenouille [26]. He shows that in general individual series from a multivariate autoregressive model do not follow univariate autoregressive, but ARMA, models. For example, individual series from a bivariate first order autoregressive model follow a second order autoregressive model with correlated residuals.

Individual series from a multivariate moving average model, however, can be shown to follow again univariate moving average models of the same (or lower) order.

Multivariate model building

(i) Model specification:

One important problem in the analysis of time series is the specification of a particular model within the class of multivariate ARIMA models for further analysis. It was pointed out in Section 2 that in the case of univariate time series the sample autocorrelation function can suggest tentative models for estimation. A similar approach can sometimes be applied for multivariate time series data; however, due to increased dimensionality (now matrices instead of scalars) it will be more difficult to recognize the covariance structure (see for example equations (3.5), (3.10), (3.12)).

Specifying a model for multivariate time series data is an extremely difficult task and no simple solution exists. Various approaches have been put forward in the literature. Parzen [23] points out that for the tentative specification of the multivariate model it is essential to first model each component separately. A similar strategy is adopted by Haugh and Box [6] who suggest a two stage specification procedure. The basic idea involved is to identify the relationship between the series by first characterizing their univariate models and secondly modelling the relationship of the two residual series driving each univariate model. The task at the second stage is made more tractable by the fact that one is crosscorrelating two individually not autocorrelated (white noise) sequences and hence the sample cross correlation function is easier to interpret. A similar approach is adopted by Jenkins [18], Granger and Newbold [13].

The approach which is used in this paper uses the information from the univariate analysis. The multivariate model is specified to be of the same form as the univariate models, but now with matrices replacing scalar parameters. For example, if the univariate series follow moving average models with maximum order q, the multivariate model is specified to be moving average of the same order. If the individual series follow ARMA models with maximum orders p and q, the initial model considered for estimation is a multivariate ARMA (p,q) model.

For multivariate models the number of parameters increases very rapidly and the suggested procedure will in general lead to overspecification (i.e., including parameters which are not necessary). Nevertheless the overspecified model provides valuable information since the parameter estimates together with their standard errors and their correlation matrix indicate which parameters can be deleted in the revised model.

(ii) Estimation

After specifying the structure of the model one has to estimate the parameters from past data. A procedure to derive maximum likelihood estimates in the case of normally distributed shocks $\tilde{a}_t$ is discussed by Wilson [29] who uses an iterative method to estimate the parameters in multivariate ARMA models. This method is a generalization of the procedure suggested by Box and Jenkins [3] for the univariate case and is outlined in the Appendix of this paper.

Computer programs for the implementation of this estimation procedure were written for the UNIVAC 1106 at the Austrian Institute of Economic Research. Implementation of the program requires a nonlinear regression routine and matrix routines for eigenvalues and eigenvectors of symmetric positive definite matrices. In the context of iterative nonlinear regression routines restrictions on the parameters such as setting certain elements equal to a constant (for example zero or one) are easily incorporated.

(iii) Diagnostic checks

After fitting the model diagnostic checks look at the residuals to detect lack of fit. If both the model is correctly specified and its parameters are known, the shocks $\tilde{a}_t$ are independently distributed with mean zero and covariance matrix Σ. Then it can be shown (Box and Jenkins [3], Box and Pierce [6], Jenkins [18]) that the estimated autocorrelations $r_{ii}(1),\ldots,$ $r_{ii}(K)$ of a_i and the estimated crosscorrelations $r_{ij}(1),\ldots,$ $r_{ij}(K)$ of a_i and a_j are asymptotically independent and normally distributed with mean zero and variance n^{-1} (where n is the number of observations).

This above result can be used to assess the statistical significance of departures of the estimated autocorrelations and crosscorrelations from zero and thus detect lack of fit. This can be achieved by plotting and comparing the correlations

$r_{ij}(1),\ldots,r_{ij}(K)$ with confidence bands $\pm\ 2n^{-\frac{1}{2}}$.

A useful yardstick for overall lack of fit (portmanteau lack of fit test [3],[18]) computes

$$n \sum_{k=1}^{K} r_{ij}^2(k) \qquad \text{for} \quad 1 \leq i,j \leq m \quad .$$

Under the null hypothesis of no lack of fit this statistic is approximately χ^2 with K degrees of freedom.

<u>Predictions from multivariate time series models</u>

As in the univariate case predictions are best interpreted from the inverted form of the model. The Π-weights, which are matrices now, are defined by

$$\Pi(B) = I - \Pi_1 B - \Pi_2 B^2 - \ldots = [\Theta(B)]^{-1}\Phi(B) \quad .$$

Then

$$\underset{\sim}{z}_{n+\ell} = \Pi_1 \underset{\sim}{z}_{n+\ell-1} + \Pi_2 \underset{\sim}{z}_{n+\ell-2} + \ldots\ldots + \underset{\sim}{a}_{n+\ell}$$

and in general the vector of ℓ-step ahead predictions is given by

$$\hat{\underset{\sim}{z}}_n(\ell) = \sum_{j\geq 0} \Pi_{j+1}^{(\ell)} \underset{\sim}{z}_{n-j} = \Pi_1^{(\ell)} \underset{\sim}{z}_n + \Pi_2^{(\ell)} \underset{\sim}{z}_{n-1} + \ldots \qquad (3.14)$$

where

$$\Pi_j^{(\ell)} = \Pi_{j+\ell-1} + \sum_{k=1}^{\ell-1} \Pi_k \Pi_j^{(\ell-k)} \quad .$$

Forecasts (3.14) are optimal in the sense that they have the smallest possible covariance matrix of their ℓ-step ahead forecast errors $\underset{\sim}{e}_n(\ell) = \underset{\sim}{z}_{n+\ell} - \hat{\underset{\sim}{z}}_n(\ell)$. (If V_1 and V_2 are two real $[m\times m]$ positive definite matrices, then V_1 is said to be smaller than V_2, provided that $\underset{\sim}{d}'V_1\underset{\sim}{d} < \underset{\sim}{d}'V_2\underset{\sim}{d}$ for every non zero $(m \times 1)$ vector $\underset{\sim}{d}$.)

(3.14) shows that in general (full Π-matrices) the prediction of a particular series utilizes the past history of all other components.

4. Example

To illustrate the methodology we consider total private consumption at current prices (CINSGN ... gesamter privater Konsum, nominell) and disposable personal income at current prices (EMNNQ9 ... Masseneinkommen, netto + verfügbares persönliches Einkommen aus Besitz und Unternehmung, netto) for Austria in the period from 1954/1 to 1976/2. First univariate analyses are reported.

(i) Total private consumption

Plot of the data indicates nonstationary variance of the series which can be stabilized by considering the natural logarithm of the series. The original series shows a strong seasonal pattern and is nonstationary. The first regular and first seasonal (s = 4) difference, however, lead to a stationary sequence.

Various models for $w_{1t} = (1-B)(1-B^4)\log z_{1t}$ are specified and estimated.

$$(1-B)(1-B^4)\log z_{1t} = (1-\theta_1 B)(1-\theta_4 B^4)a_{1t} \quad . \qquad \text{(M11)}$$

The parameter estimates are given by

$$\theta_1 = .53 \qquad \text{and} \quad \theta_2 = .52$$

The variance of the one step ahead forecast error* is

$$\sigma^2_{a_1} = .000487 \quad .$$

The autocorrelation function of the residuals is calculated and the χ^2 value (which is to be compared with 18 degrees of freedom) is 19.5.

---

*The estimated variance of the one step ahead forecast error is the sum of squares of residuals divided by the number of residuals.

There is one rather large contribution at lag 3 ($r_3 = .22$) and this leads to consider the revised model

$$(1 - B)(1 - B^4)\log z_{1t} = (1 - \theta_1 B)(1 - \theta_3 B^3)(1 - \theta_4 B^4)a_{1t} . \quad \text{(M12)}$$

Estimates of the parameters are significant and are given by

$$\theta_1 = .56 \quad ; \qquad \theta_3 = -.23 \; ; \qquad \theta_4 = .52 \quad .$$

The variance of the one step ahead forecast error is

$$\sigma^2_{a_1} = .000463 \quad .$$

The χ^2 value of 13.8 (compared to 17 degrees of freedom) and the plot of the autocorrelation of the residuals gives no indication of departure from randomness in the residuals and leads to acceptance of the fitted model.

(ii) <u>Disposable personal income</u>

Also in this case the logarithmic transformation and a regular and seasonal difference is necessary to achieve a stationary sequence. Several models were entertained for $w_{2t} = (1 - B)(1 - B^4)\log z_{2t}$

$$(1 - B)(1 - B^4)\log z_{2t} = (1 - \theta_1 B)(1 - \theta_4 B^4)a_{2t} \qquad \text{(M21)}$$

with estimates $\theta_1 = .33$ and $\theta_2 = .26$.

The variance of the one step ahead forecast error is

$$\sigma^2_{a_2} = .000320 \quad .$$

The autocorrelation function of the residuals has one particularly large contribution at lag 10 ($r_{10} = -.28$ compared to standard deviation of .12) and the χ^2 value of 20.6 is rather high (compared to a χ^2 distribution with 18 degrees of freedom).

The large autocorrelation at lag 10 leads to a revised model with one additional parameter

$$(1-B)(1-B^4)\log z_{2t} = (1-\theta_1 B)(1-\theta_4 B^4)(1-\theta_{10}B^{10})a_{2t} \quad \text{(M22)}$$

with estimates

$$\theta_1 = .26 \quad ; \quad \theta_4 = .30 \quad , \quad \theta_{10} = .43$$

$$\sigma^2_{a_2} = .000281 \quad .$$

The χ^2-value is considerably lower, $\chi^2 = 13.4$ and gives no reason to doubt the adequacy of the model.

(iii) <u>Bivariate model</u>

The structure of (M11) and (M21) is used to specify the multivariate model

$$\underset{\sim}{w}_t = (I - \Theta_1 B)(I - \Theta_4 B^4)\underset{\sim}{a}_t \quad *) \qquad \text{(M31)}$$

where

$$\underset{\sim}{w}'_t = (w_{1t}\, w_{2t}) \quad ; \quad w_{1t} = (1-B)(1-B^4)\log z_{1t} \quad ,$$

$$w_{2t} = (1-B)(1-B^4)\log z_{2t} \quad .$$

Using the multivariate estimation procedure, which is described in the Appendix, estimates of the elements in Θ_1 and Θ_4, together with their standard errors in brackets, are calculated

$$\Theta_1 = \begin{bmatrix} .68 & -.34 \\ (.09) & (.11) \\ -.16 & .50 \\ (.08) & (.10) \end{bmatrix} \quad ; \quad \Theta_4 = \begin{bmatrix} .47 & .003 \\ (.12) & (.14) \\ -.14 & .34 \\ (.10) & (.12) \end{bmatrix}$$

---

*Instead of taking first regular and seasonal differences a model with autoregressive operators was considered; $\underset{\sim}{w}_t = (I - \Phi_1 B)(I - \Phi_4 B^4)\log \underset{\sim}{z}_t$. The estimates of Φ_1 and Φ_4, however, were close to I.

$$\Sigma = \begin{bmatrix} .000465 & .000148 \\ & .000296 \end{bmatrix} .$$

Since the off-diagonal elements in Θ_4 are not significantly different from zero, the model (M31) is respecified by setting these two elements equal to zero.

The estimates of the remaining elements are given by

$$\Theta_1 = \begin{bmatrix} .65 & -.33 \\ (.10) & (.11) \\ -.22 & .52 \\ (.09) & (.11) \end{bmatrix} \quad \Theta_4 = \begin{bmatrix} .53 & 0 \\ (.10) & \\ & .31 \\ 0 & (.10) \end{bmatrix} \qquad \text{(M32)}$$

$$\Sigma = \begin{bmatrix} .000469 & .000150 \\ & .000300 \end{bmatrix} .$$

The decrease in the one step ahead prediction errors of (M32) compared with (M11) and (M21) is 4% for the private consumption series (from .000487 to .000469) and 7% for personal income (from .000320 to .000300).

We already noted that the univariate models (M11) and (M21) showed some shortcomings and that they could be improved by an additional parameter. The same shortcomings become evident when the residuals from the multivariate model (M32) are analyzed.

For multivariate models we use the diagnostic checks discussed in Section 2. The χ^2 values for the residuals of model (M32) which have to be compared with a χ^2 distribution with 20 degrees of freedom are for the

first series: $\chi^2 = 18.7$ with a large contribution at lag 3 ($r_3 = .23$ compared to standard deviation of .11) second series: $\chi^2 = 23.6$ with a large contribution at lag 10 ($r_{10} = -.33$ compared to standard deviation of .12)

crosscorrelations
with lag on series 2: $\chi^2 = 17.9$
with lag on series 1: $\chi^2 = 10.5$

The rather high values at lag 3 (for series 1) and lag 10 (for series 2) lead to the revised model

$$\underset{\sim}{w}_t = (I - \Theta_1 B)(I - \Theta_4 B^4)(I - \Theta_3 B^3 - \Theta_{10} B^{10})\underset{\sim}{a}_t \qquad \text{(M33)}$$

where Θ_3 and Θ_{10} have only one nonzero element. The estimates of the parameters are given by

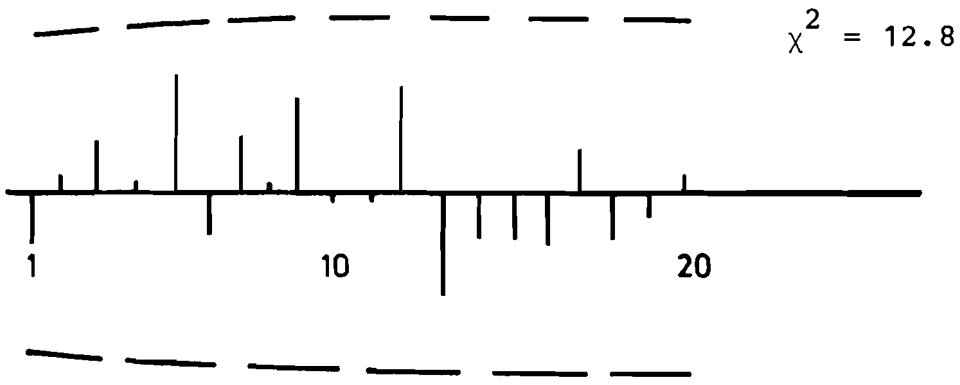
$$\Theta_1 = \begin{bmatrix} .66 & -.34 \\ (.07) & (.11) \\ -.27 & .51 \\ (.07) & (.11) \end{bmatrix} \qquad \Theta_4 = \begin{bmatrix} .50 & 0 \\ (.10) & \\ & .38 \\ 0 & (.10) \end{bmatrix}$$

$$\Theta_3 = \begin{bmatrix} -.19 & 0 \\ (.07) & \\ 0 & 0 \end{bmatrix} \qquad \Theta_{10} = \begin{bmatrix} 0 & \\ & .53 \\ 0 & (.10) \end{bmatrix}$$

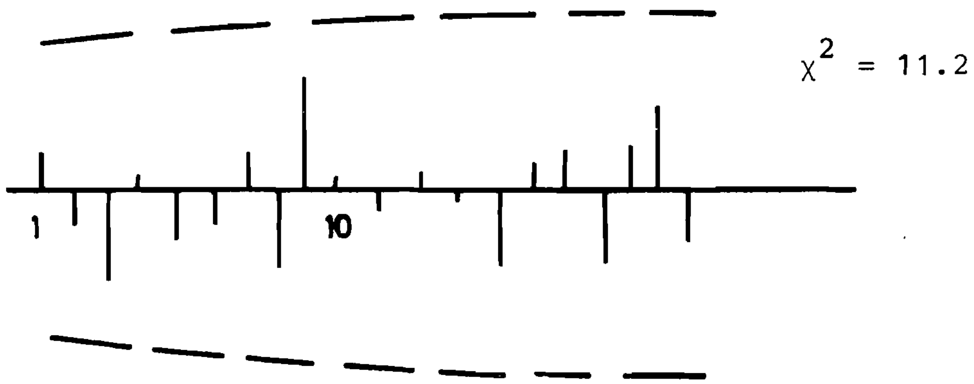
$$\Sigma = \begin{bmatrix} .000443 & .000126 \\ & .000240 \end{bmatrix} .$$

All the parameters are significant and the diagnostic checks give no reason to doubt the validity of the model. The χ^2 values are considerably lower. The residual auto- and cross-correlations together with the χ^2-values, are plotted in Figure 1.

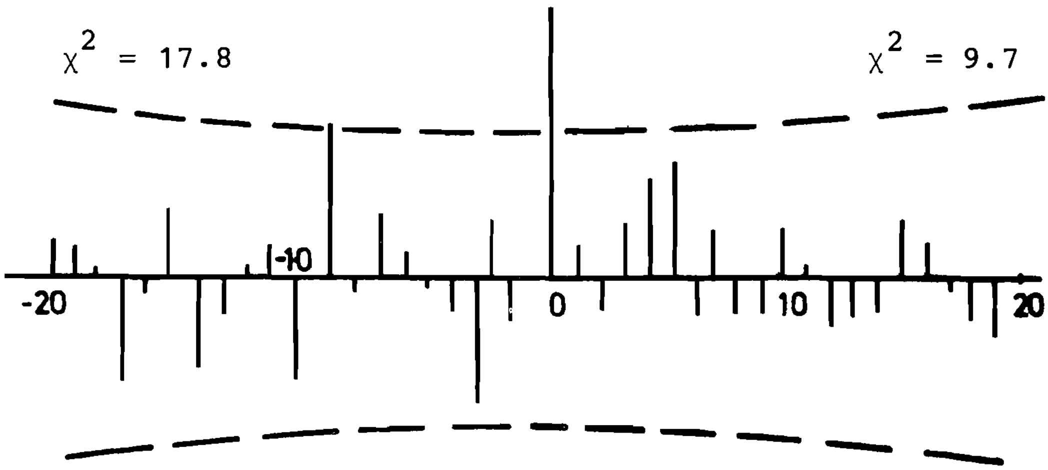
The improvement in the one step ahead prediction errors is 5% for the private consumption series and 17% for the personal income series.



Autocorrelation function of residuals-series 1



Autocorrelation function of residuals-series 2



Crosscorrelation function of residuals

Figure 1: <u>Autocorrelations and crosscorrelations of residuals; together with 2σ limits; model (M33)</u>

5. Interpretation of results and concluding remarks

The analysis in Section 4 shows that model (M33) and also model (M32)

(i) represent a truly multivariate (feedback) relationship between private consumption and disposable personal income,

(ii) lead to a decrease in the variance of the one step ahead prediction error which compared to the univariate models is relatively small.

ad i. The two considered series are an example of a truly multivariate (feedback) relationship; i.e., a relationship where the past of both series is needed for the prediction of future values. A feedback relationship is concluded since both off-diagonal elements in Θ_1, θ_{12} and θ_{21}, are significantly different from zero (compared to their standard error). The signs of the coefficients θ_{12} and θ_{21} are both negative and are consistent with economic theory. For example, model (M32) can be written

$$(1-B)(1-B^4)\log z_{1t} = (1-.65B)(1-.53B^4)a_{1t} + .33(1-.31B^4)a_{2t-1} \quad (5.1)$$

$$(1-B)(1-B^4)\log z_{2t} = (1-.52B)(1-.31B^4)a_{2t} + .22(1-.53B^4)a_{1t-1} \quad (5.2)$$

An additional increase in todays income (variable z_2) which is measured by the increase of todays observed income compared to its last prediction (a_{2t}) will lead to an increase in tomorrow's private consumption (variable z_1) and vice versa. Equations (5.1) and (5.2) show that the additional change in today's income (consumption) is always seen in relation to the income (consumption) 4 quarters ago.

ad ii. The phenomenon that predictions of many economic time series, once effective use of their own past has been made, can be little improved by using, in addition, past values of other available series, has been discussed, for example, by Pierce [24], Cramer and Miller [9], Feige and Pearce [11]. They conclude that in general the variance reduction is rather small. Similar conclusions are reached with Austrian macroeconomic series and they will be reported in the near future.

The above results which might be surprising to some economists must be reconciled with the information from regression like analyses which traditionally have always shown strong relationships among macroeconomic variables. Various explanations which help in reconciling the time series results and the fact that certain economic causes and effects are known to exist are given below.

(a) The empirical regression results which show strong relationships among macroeconomic series may be ill-founded due to carelessness about the effect of the time series properties of the data. For example, not accounting for the serial correlation among the observations tends to find relationships which actually don't exist (Box and Newbold [5], Granger and Newbold [13]).

(b) The conclusion from time series studies should not necessarily lead to doubt the existence of economic relationships. It may be concluded that they are perhaps inherently not verifiable. Reasons for not being able to verify economic relations using empirical data over the last 20 years are:

(1) economic data are happenstance data, as far as experimental design is concerned, and usually subject to large measurement error;

(2) any deterministic series can be perfectly predicted from its own past and there is no room for improvement by using another variable. If, for example, one series grows by a constant percentage it will show up as unrelated to

any other variable, regardless what its actual relationship might be.

(3) The series may appear independent only because of a common but opposite association with a third variable.

REFERENCES

[1] Anderson, T.W., *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, 1958.

[2] Box, G.E.P. and D.R. Cox, An Analysis of Transformations, *Journal of Royal Stat. Soc., B*, 26 (1964), 211-243.

[3] Box, G.E.P. and G.M. Jenkins, *Time Series Analysis, Forecasting and Control*, Holden Day, San Francisco, 1970.

[4] Box, G.E.P. and G.M. Jenkins, Some comments on a paper by Chatfield and Prothero and on a review by Kendall, *Journal of Royal Stat. Soc., A*, 135 (1973), 337-345.

[5] Box, G.E.P. and P. Newbold, Some comments on a paper of Coen, Gomme and Kendall, *Journal of Royal Stat. Soc., A*, 134 (1971), 229-240.

[6] Box, G.E.P. and D.A. Pierce, Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models, *JASA*, 65 (1970), 1509-1526.

[7] Christ, C.F., Judging the Performance of Econometric Models of the U.S. Economy, *Int. Econ. Rev.*, 16 (1975), 54-74.

[8] Cooper, R.L., The Predictive Performance of Quarterly Econometric Models of the United States, in B.G. Hickman, ed., *Econometric Models of Cyclical Behavior*, Columbia Univ. Press, New York, 1972.

[9] Cramer, R.H. and R.B. Miller, Dynamic Modelling of Multivariate Time Series for Use in Bank Analysis, *Journal of Money, Credit and Banking*, 8 (1976), 85-96.

[10] Draper, N.R. and H. Smith, *Applied Regression Analysis*, Wiley, New York, 1966.

[11] Feige, E.L. and D.K. Pearce, *The Causality Relationship between Money and Income: A Time Series Approach*, paper presented at the Annual Meeting of the Midwest Economic Association, Chicago, April, 1974.

[12] Granger, C.W.J., Investigating Causal Relations by Econometric Models and Cross Spectral Methods, *Econometrica*, 37 (1969), 424-438.

[13] Granger, C.W.J. and P. Newbold, *Forecasting Economic Time Series*, Academic Press, New York, 1977.

[14] Hannan, E.J., The Identification of Vector Mixed Autoregressive Moving Average Systems, *Biometrika*, 56 (1969), 223-225.

[15] Hannan, E.J., *Multiple Time Series*, Wiley, New York, 1970.

[16] Haugh, L.D. and G.E.P. Box, Identification of Dynamic Regression (Distributed Lag) Models Connecting Two Time Series, *JASA*, 72 (1977), 121-130.

[17] Hillmer, S.C., *Time Series: Estimation, Smoothing and Seasonal Adjusting*, Ph.D. Thesis, Department of Statistics, University of Wisconsin, Madison, 1976.

[18] Jenkins, G.M., The Interaction Between the Muscrat and Mink Cycles in North Canada, in Editura Akademiei Republicii Socialiste Romania, *Proceedings of the 8th Biometric Conference*, 57-71, Bucharest, 1975.

[19] Ledolter, J., Schebeck, F. and G. Thury, Box-Jenkins Methoden - Alternative Verfahren zur Prognose ökonomischer Zeitreihen; *Empirica*, 1977, forthcoming.

[20] Narasimham, G.V.L. and N.D. Singpurwalla, Comparison of Box-Jenkins and BEA Quarterly Econometric Model Predictive Performance, *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, Washington, D.C., 501-504, 1974.

[21] Naylor, T.H., T.G. Seaks, and D.W. Wichern, Box-Jenkins Methods: An Alternative to Econometric Models, *Int. Stat. Rev.*, 40 (1972), 123-139.

[22] Nelson, C.R., The Predictive Performance of the FRB-MIT-PENN Model of the U.S. Economy, *Amer. Econ. Rev.*, 62 (1972), 902-917.

[23] Parzen, E., Multiple Time Series Modelling, in P. Krishnaiah, ed., *Multivariate Analysis II*, Academic Press, New York, 1969.

[24] Pierce, D.A., Relationships-and the Lack thereof-Between Economic Time Series, with Special Reference to Money and Interest Rates, (with discussion), *JASA*, 72 (1977), 11-26.

[25] Prothero, D.L. and K.F. Wallis, Modelling Macroeconomic Time Series (with discussion), *Journal of Royal Stat. Soc.*, *A*, 139 (1976), 468-486.

[26] Quenouille, M.H., *The Analysis of Multiple Time Series*, Griffin, London, 1957.

[27] Tintner, G. and G. Kadekodi, A Note on the Use of Transformations and Differences in the Estimation of Econometric Relations, *Sankhya*, *B*, 35 (1971), 263-277.

[28] WIFO Monatsberichte 9/1976, Österreichisches Institut für Wirtschaftsforschung.

[29] Wilson, G., The Estimation of Parameters in Multivariate Time Series Models, *Journal of Royal Stat. Soc., B*, 35 (1973), 76-85.

Appendix

Parameter estimation in multivariate ARMA models

The multivariate (m-dimensional) ARMA model is given by

$$(I - \Phi_1 B - \ldots - \Phi_p B^p) \underset{\sim}{z}_t = (I - \Theta_1 B - \ldots - \Theta_q B^q) \underset{\sim}{a}_t \quad . \qquad \text{(A.1)}$$

The unknown parameters $\Phi_1,\ldots,\Phi_p,\Theta_1,\ldots,\Theta_q$, which for convenience are arranged in a column vector $\underset{\sim}{\beta}$, and the elements of the covariance matrix of the white noise sequence $\underset{\sim}{a}_t$, Σ, have to be estimated from the observations $\underset{\sim}{z}_1,\ldots,\underset{\sim}{z}_n$.

Assuming joint normality for $\underset{\sim}{a}_t$ and neglecting the effect of starting values for $\underset{\sim}{a}_t$ (i.e., setting the starting values for $\underset{\sim}{a}_t$ equal to zero -- for relaxing this condition see Hillmer [17]), the likelihood of the parameters $\underset{\sim}{\beta}$ and Σ is given by

$$\ell(\underset{\sim}{\beta},\Sigma|\underset{\sim}{z}_1,\ldots,\underset{\sim}{z}_n) \propto |\Sigma|^{-\frac{n-p}{2}} \exp\left\{-\tfrac{1}{2}\sum_{t=p+1}^{n} \underset{\sim}{a}_t' \Sigma^{-1} \underset{\sim}{a}_t\right\} \qquad \text{(A.2)}$$

where $\underset{\sim}{a}_t$ is a function of $\underset{\sim}{\beta}$

$$\underset{\sim}{a}_t = \underset{\sim}{z}_t - \Phi_1 \underset{\sim}{z}_{t-1} - \ldots - \Phi_p \underset{\sim}{z}_{t-p} + \Theta_1 \underset{\sim}{a}_{t-1} + \ldots + \Theta_q \underset{\sim}{a}_{t-q} \qquad \text{(A.3)}$$

for $t \geq p+1$, and $\underset{\sim}{a}_t = \underset{\sim}{0}$ for $t \leq p$. The log likelihood function is given by

$$L(\underset{\sim}{\beta},\Sigma|\underset{\sim}{z}_1,\ldots,\underset{\sim}{z}_n) \propto -\frac{n-p}{2}\left[\log|\Sigma| + \frac{1}{n-p}\sum_{t=p+1}^{n} \underset{\sim}{a}_t' \Sigma^{-1} \underset{\sim}{a}_t\right] . \qquad \text{(A.4)}$$

To derive maximum likelihood (ML) estimates we have to minimize the function

$$F(\underset{\sim}{\beta},\Sigma) = \log|\Sigma| + \frac{1}{n-p}\sum_{t=p+1}^{n} \underset{\sim}{a}_t' \Sigma^{-1} \underset{\sim}{a}_t \qquad \text{(A.5)}$$

with respect to $\underset{\sim}{\beta}$ and Σ.

Conditional estimation of Σ

It can be shown (Wilson [29]) that the derivative of $F(\underset{\sim}{\beta},\Sigma)$ with respect to elements of $\Sigma^{-1} = \{\sigma^{ij}\}$ is proportional to

$$-\sigma_{ij} + \frac{1}{n-p} \sum_{t=p+1}^{n} a_{it}\, a_{jt}$$

Thus, for given values of $\underset{\sim}{\beta}$, the ML estimate of the elements of $\Sigma = \{\sigma_{ij}\}$ is given by

$$\sigma_{ij} = \frac{1}{n-p} \sum_{t=p+1}^{n} a_{it}\, a_{jt} \quad . \tag{A.6}$$

Conditional estimation of $\underset{\sim}{\beta}$

In order to derive the conditional estimate of $\underset{\sim}{\beta}$ given the value of Σ one has to minimize the second part in (A.5)

$$\sum_{t=p+1}^{n} \underset{\sim}{a}_t' \Sigma^{-1} \underset{\sim}{a}_t = \sum_{t=p+1}^{n} \underset{\sim}{h}_t' \underset{\sim}{h}_t = \sum_{t=p+1}^{n} \sum_{j=1}^{m} h_{jt}^2 \tag{A.7}$$

where

$$\underset{\sim}{h}_t' = \underset{\sim}{a}_t' P \qquad \underset{\sim}{h}_t' = (h_{1t}, h_{2t}, \ldots, h_{mt})$$

and

$$PP' = \Sigma^{-1} \quad \text{or} \quad P'\Sigma^{-1}P = I \ .$$

It can be shown (for example Anderson [1], Appendix 1) that

$$P = H D^{-\frac{1}{2}}$$

where H is the matrix of normalized characteristic vectors of Σ^{-1} and D is a diagonal matrix with corresponding characteristic roots in the diagonal.

A nonlinear regression routine is used to derive the estimates in $\underset{\sim}{\beta}$ such that expression (A.7) becomes as small as possible. A good introduction to nonlinear regression methods is given in Draper and Smith [10].

Simultaneous estimation of $\underset{\sim}{\beta}$ and Σ

The strategy to estimate the parameters $\underset{\sim}{\beta}$ and Σ is to apply the conditional estimation schemes alternately

$$\begin{cases} \Sigma_n = \Sigma(\underset{\sim}{\beta}_n) \\ \underset{\sim}{\beta}_{n+1} = \underset{\sim}{\beta}(\Sigma_n) \end{cases} \qquad n = 0,1,\ldots$$

Since each of the steps is a conditional minimization the above procedure will converge to the overall minimum. Furthermore, as shown by Wilson [29], the estimates $\underset{\sim}{\beta}$ and Σ are consistent and asymptotically uncorrelated. The asymptotic distribution of $\underset{\sim}{\beta}$ is normal.

Computer programs which implement this iterative estimation procedure have been written and are available from the author. As starting value for Σ one usually chooses a diagonal matrix with variance estimates from the univariate models in its diagonal.