IDENTIFICATION OF MULTIPLE-INPUT TRANSFER FUNCTION MODELS

Lon-Mu Liu*
Department of Biomathematics
University of California
Los Angeles, California 90024

Dominique M. Haussens
Graduate School of Management
University of California
Los Angeles, California 90024

Keywords and phrases: transfer function model; time-series analysis; least-squares estimation; filtering; corner method.

ABSTRACT

This paper proposes a procedure for transfer function identification (specification) based on least-squares estimates of transfer function weights using the original or filtered series. The corner method is then used to identify a parsimonious rational form of the transfer function. The procedure is illustrated in a simulated example; it is shown how this straightforward approach outperforms other identification methods such as Box and Jenkins' prewhitening and Haugh and Box' double prewhitening techniques.

1. INTRODUCTION

The transfer function model is one of the most widely used time-series models in several areas of application such as engineering, economics and management science. Yet in spite of its popularity, the identification or specification stage of transfer function analysis is not sufficiently developed, in particular for multiple-input models.

*Now at the Department of Quantitative Methods, University of Illinois at Chicago Circle, Box 4348, Chicago, Illinois 60680

Copyright © 1982 by Marcel Dekker, Inc.
The best known identification approach is due to Box and Jenkins (1976), who propose a comprehensive procedure for the one-input situation by studying the sample cross-correlation function (SCCF) between the prewhitened input series and the corresponding filtered output series (filtered by the ARMA model of the input series). While their method is promising for the one-input situation, it is difficult to generalize it for multiple-input models. Priestley (1971) recommends prewhitening both input and output series and obtaining the transfer function weights of the prewhitened series using least-squares estimation. Haugh and Box (1977) also suggest studying the SCCF of the prewhitened input and output series from which they identify the transfer function model. We refer to the Priestley, and Haugh and Box approach as the double prewhitening method. Fask and Robinson (1977) generalize the double prewhitening method to multivariate dynamic models. Granger and Newbold (1977a,b) also apply this method for the identification of a two-way causal system. Although double prewhitening has been widely used, it is more difficult to apply than the Box and Jenkins method, even in the one-input situation. The major difficulties are that a model is frequently over-structured due to the prewhitening factors, and that the lag structure is often cumbersome to derive.

Tiao et al (1979) use sample cross-correlation and partial correlation matrices of a set of multiple series for the identification of multiple ARIMA models. Since the transfer function model can be viewed as a special case of the multiple ARIMA model, their method can be applied to transfer function identification. Unfortunately, their method imposes an extra structure between the factors on the residuals and the transfer function, causing the transfer function model to be over-structured.

As an alternative to time-domain analysis, transfer function identification has also been explored using a frequency-domain approach (spectral methods), e.g., Box and Jenkins (1976) and Priestley (1971). However, it is rather difficult to apply spectral methods in practice.
The present paper proposes a procedure which mainly applies linear least-squares estimation on the original or filtered series for transfer function identification, similar to Priestley's (1969) suggestion for the one-input situation. Another least-squares approach was developed by Caines, Sethi and Brotherton (1977), who used a Cholesky least-squares algorithm for transfer function analysis. However, their method is aimed at model estimation rather than identification and their results are difficult to extend to multiple-input models. Although the procedure in the present paper is simple and straightforward, it performs very well. In addition to its simplicity, the method is easily extended to models with intervention components, which are often difficult to handle with other methods.

2. STATISTICAL BACKGROUND

Without loss of generality, we study the following two-input transfer function model:

\[ Y_t = c + \frac{w_1(B)}{\delta_1(B)} X_{1t} + \frac{w_2(B)}{\delta_2(B)} X_{2t} + \varepsilon_t, \]

\[ t = 1, 2, \ldots, N \]  \hspace{1cm} (2.1)

where \( X_{1t} \) and \( X_{2t} \) are input series (or exogenous variables), \( Y_t \) is the output series (or endogenous variable), \( \varepsilon_t \) the noise series, and \( c \) a constant. All series are stationary and follow ARMA (autoregressive-moving average) processes. The ARMA process for the noise series is

\[ \phi(B) \varepsilon_t = \theta(B) \varepsilon_t \]  \hspace{1cm} (2.2)

with

\[ \phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p, \]

and

\[ \theta(B) = 1 - \theta_1 B - \ldots - \theta_q B^q, \]

where \( B \) is a backward shift operator (\( B \varepsilon_t = \varepsilon_{t-1} \)) and \( \varepsilon_t \)'s are independently identically distributed. The rational polynomial \( \frac{w_1(B)}{\delta_1(B)} \) is the transfer function between \( Y_t \) and \( X_{1t} \) where

\[ w_1(B) = (w_{11} B + \ldots + w_{1s_1} B^{s_1})) B^{t-1} \]

and

\[ \delta_1(B) = (\delta_{11} B + \ldots + \delta_{1s_1} B^{s_1}) B^{t-1} \]

\[ s_1 = \max \{ s_i, s_j \}, \quad i, j = 1, 2 \]
and

\[ \delta_i(B) = 1 - \delta_{i1}B - \ldots - \delta_{ir_i}B^r_i \]  

(2.3)

The roots of the polynomials \( \delta_i(B) \), \( \phi(B) \) and \( \theta(B) \) lie outside the unit circle. The noise series \( \varepsilon_t \) must be independent of each input series, but the input series may be correlated with each other.

The transfer function component \( \omega_i(B)/\delta_i(B) \) can also be expressed as

\[ V_i(B) = v_{i0} + v_{i1}B + v_{i2}B^2 + \ldots, \]

where \( V_i(B) \), a linear form of the transfer function, has a finite number of terms if \( \delta_i(B) = 1 \) and an infinite number of terms if \( \delta_i(B) \neq 1 \). Since all roots of the \( \delta_i(B) \) polynomial lie outside the unit circle, the transfer function \( \omega_i(B)/\delta_i(B) \) can always be approximated by \( V_i(B) \) with a finite number of terms, say \( K_i \), in practice. We refer to \( v_{ij} \) as the \( j \)-th transfer function (or impulse response) weight for the input series \( X_{1t} \).

The task of transfer function identification is to find appropriate estimates of the \( V_i(B) \) polynomials and to express them in rational forms \( \omega_i(B)/\delta_i(B) \). Both elements are discussed below.

2.1 Estimating the Transfer Function Weights

For sufficiently large \( K_i \)'s, (2.1) can be expressed as

\[ Y_t = c + (v_{10} + v_{11}B + \ldots + v_{1K_1}B^{K_1})X_{1t} \]

\[ + (v_{20} + v_{21}B + \ldots + v_{2K_2}B^{K_2})X_{2t} + \varepsilon_t. \]

(2.4)

Using

\[ K = \text{Max}(K_1, K_2), \]
\[ n = N - K, \]
\[ B = [c, v_{10}, v_{11}, \ldots, v_{1K_1}, v_{20}, v_{21}, \ldots, v_{2K_2}], \]

and

\[ Y = [Y_{K+1}, Y_{K+2}, \ldots, Y_{K+n}], \]
\[ X = [1, X_0, X_1, \ldots, X_{K_1}, X_0, X_1, \ldots, X_{K_2}], \]

\[ X_t = [X_0, X_1, \ldots, X_{K_1}, X_0, X_1, \ldots, X_{K_2}], \]
MULTIPLE-INPUT TRANSFER FUNCTION MODELS

where

\[ X_i^j = B^j X_i^0 \quad \text{and} \quad X_i^0 = [X_i(K+1) X_i(K+2) \ldots X_i(K+n)]' \],

the ordinary least-squares (OLS) estimate of \( \beta \) can be expressed as

\[ \hat{\beta} = (X'X)^{-1}X'Y \quad \text{(2.5)} \]

Two problems may arise in the above least-squares estimation. First, if one of the input series contains an autoregressive (AR) factor with roots close to one, i.e., if the series is close to nonstationarity, the \( X'X \) matrix may be near-singular and result in ill-conditioning for matrix inversion. This can be observed as follows: if we standardize the matrix \( X'X \) by its diagonal elements, the new matrix approximately contains the sample autocorrelations and cross-correlations of the \( X_{1t} \) and \( X_{2t} \) as its elements. Thus if one of the input series contains an autoregressive polynomial with roots close to the unit circle, a number of sample autocorrelations will be very close to one. This difficulty is less serious if an input series follows a moving (MA) process since the sample autocorrelations of an MA process will not be close to one even if the root of the MA polynomial is one (Box and Jenkins 1976). Second, the noise series \( \epsilon_t \) may not be white noise, therefore the OLS estimates of \( \beta \) may not be efficient.

To avoid the difficulty of inverting an ill-conditioned \( X'X \), we suggest to examine the ARMA models of the input series and then take appropriate action. As argued earlier, only the roots of the AR polynomials are important: if the absolute values of all these roots are rather large there is no problem in using the original data. However, if some of the roots are close to one, a common filter on the input and output series is used in order to make \( X'X \) well-conditioned.

Common filters have been used previously, primarily because they do not alter the transfer function weights if the series are stationary (Priestley 1971, Granger and Newbold 1977b). For example, Sims (1972) employed an ad hoc filter \((1-.75B)^2\) in a well-
known study of causality between money and income. Although the choice of a filter can be flexible, we do offer a selection criterion, namely that it eliminates the AR factors with roots close to one. If there are more than one such factors they can be combined in one common filter. As an example, if three input series $X_1$, $X_2$ and $X_3$ follow the ARMA processes

$$(1 - .9B)(1 - .4B^4) X_{1t} = a_{1t},$$

$$(1 - .3B)(1 - .8B^4) X_{2t} = (1 - .7B)a_{2t},$$

and

$$(1 - .5B) X_{3t} = (1 - .9B^4)a_{3t},$$

a choice of $(1 - .9B)(1 - .8B^4)$ is recommended as a common filter.

It should be pointed out that common filtering is done for numerical accuracy rather than statistical efficiency. Its necessity depends on the roots of the AR polynomials and the precision of the computer. In some situations, filtering could be avoided by using double precision computing. Alternatively, the multicollinearity could be handled by ridge regression analysis (e.g., Erickson 1981). Also, since a filter with roots close to unity may create an NA polynomial near non-invertibility when $\varepsilon_t$ is white noise, it should be avoided. For instance, the filter $(1 - .7B)(1 - .7B^4)$ may be safer in the above example.

The second problem can be resolved by estimating $\hat{\beta}$ with generalized least squares (GLS) rather than OLS. If the covariance matrix $\Sigma$ of $\varepsilon_t$ ($t=K, K+1, \ldots, N$) is known the GLS estimator of $\beta$ is:

$$\hat{\beta} = (X'\Sigma^{-1}X')^{-1}X'\Sigma^{-1}Y,$$  \hspace{1cm} (2.6)

which is consistent and efficient. The direct computation of (2.6) is complicated. It may be easier to obtain $\hat{\beta}$ by using OLS estimation on the transformed series $\tilde{X}$ and $\tilde{Y}$. Since $\Sigma^{-1}$ can be expressed as $H'H$, we obtain

$$\tilde{\beta} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{Y},$$  \hspace{1cm} (2.7)
where
\[ \tilde{X} = H \tilde{X}, \]
\[ \tilde{Y} = H \tilde{Y}. \]

Following Ljung and Box (1979) the transformations (2.8) and (2.9) are approximately equivalent to filtering the input and output series by the ARMA model of the noise series. In practice, the noise model is unknown but can be identified from ARMA analysis on the OLS residuals. An iterative process of filtering and OLS estimation could then be started until the estimates converge. Such a procedure is similar to estimating regression coefficients with the well-known Cochrane-Orcutt (1949) method for simple AR(1) situations.

Since the purpose of the least-squares procedures is identification of transfer functions, we find it not necessary to apply the iterative process. In our experience, the OLS estimates based on the filtered series are usually very satisfactory. For more accurate estimates we recommend joint estimation of the transfer function weights and noise parameters by using nonlinear least-squares. This can be done easily with computer programs such as BMDQ2T (Liu 1979).

2.2 Expressing the Transfer Function Weights in Rational Form

In many instances a more parsimonious representation of transfer function weights may be obtained by expressing the \( V_i(B) \) in rational form. This task is equivalent to finding the values \( r_i, s_i \) and \( b_i \) in (2.3). Box and Jenkins (1976) use the pattern of the transfer function weights for this purpose, similar to their ARMA identification method. A cut-off pattern implies that \( r_i = 0 \) and is simple to model. A tail-out pattern implies that a denominator polynomial is present, but it appears difficult to apply Box and Jenkins' guidelines in this case. The authors propose to use a modification of the corner method for ARMA identification (Beguin, Gourieroux and Monfort 1980) for this purpose.

Let \( v_{i, \text{max}} \) be the maximum value of \( |v_{ij}|, j=0,1,2, \ldots, K_i \), where the \( v_{ij} \)'s are the true transfer function weights for the
rational polynomial $\omega_i(B)/\delta_i(B)$ and let $\eta_{ij} = v_{ij}/v_{i,max}$. Then, for each explanatory variable $X_i$ we can construct a $(g \times g)$ matrix $D(f,g)$ and its determinant $\Delta(f,g)$, where

$$D(f,g) = \begin{bmatrix}
\eta_f & \eta_{f-1} & \ldots & \eta_{f-g+1} \\
\eta_{f+1} & \eta_f & \ldots & \eta_{f-g+2} \\
\vdots & \vdots & \ddots & \vdots \\
\eta_{f+g-1} & \eta_{f+g-2} & \ldots & \eta_f
\end{bmatrix}$$

$f>0$, $g>1$, and $\eta_j = 0$ if $j<0$. The subscript $i$ is omitted for simplicity. Now, for any integer $M$ large enough to be greater than $r_i$ and $(s_i+b_i)$, we can build an $(M+1)\times M$ array $C$ with $\Delta(f,g)$ as its $fg$-th element, where $f=0,1,2,\ldots,M$, and $g=1,2,\ldots,M$. Then the transfer function weights $v_{ij}$ have a representation $\omega_i(B)/\delta_i(B)$ with orders $r,s$ and $b$, if and only if the $C$ array has the following structure:

<table>
<thead>
<tr>
<th>$f$</th>
<th>1</th>
<th>2</th>
<th>...</th>
<th>$r$</th>
<th>$r+1$</th>
<th>...</th>
<th>$M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b-1</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td>b</td>
<td>$\Delta(b,1)$</td>
<td>$\Delta(b,2)$</td>
<td>...</td>
<td>$\Delta(b,r)$</td>
<td>$\Delta(b,r+1)$</td>
<td>...</td>
<td>$\Delta(b,M)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>s+b-1</td>
<td>$\Delta(s+b-1,1)$</td>
<td>$\Delta(s+b-1,2)$</td>
<td>...</td>
<td>$x$</td>
<td>$x$</td>
<td>...</td>
<td>$x$</td>
</tr>
<tr>
<td>s+b</td>
<td>$\Delta(s+b,1)$</td>
<td>$\Delta(s+b,2)$</td>
<td>...</td>
<td>$x$</td>
<td>$0$</td>
<td>...</td>
<td>$0$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>$\Delta(M,1)$</td>
<td>$\Delta(M,2)$</td>
<td>...</td>
<td>$x$</td>
<td>$0$</td>
<td>...</td>
<td>$0$</td>
</tr>
</tbody>
</table>
where an "x" means that the term is different from zero. The pattern implies the orders \( r, s \) and \( b \) if and only if the first \( b \) rows and the lower right-hand corner starting at the \((s+b+1)\)th row and the \((r+1)\)th column of the \( \zeta \)-array are all zeroes. Details of the corner method are discussed in Beguin, Gourieroux and Monfort (1980).

In practice the \( \tilde{V}_{ij} \) are estimated by the \( \tilde{V}_{ij} \), which are subject to random errors. Consequently, one will find some small values in the \( \zeta \)-array for the zeroes indicated above. However, the upper and lower-right hand corner will show a sudden drop in array values. As an illustration, the method is applied to the gas furnace data in Box and Jenkins (1976). The estimated transfer function weights are:

\[
\begin{array}{cccccccc}
1 & 0 & 1 & 2 & 3 & 5 & 6 & 8 & 9 & 10 & 11 & 12 & 13 \\
\tilde{q}_1 & 02 & 10 & 36 & 53 & 83 & 88 & 12 & 12 & 06 & 36 & 10 & 36 & 74 & 12 \\
\end{array}
\]

The \( \zeta \) array is:

<table>
<thead>
<tr>
<th>( \zeta )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-.02</td>
<td>.00</td>
<td>-.00</td>
<td>.00</td>
<td>-.00</td>
<td>.00</td>
<td>-.00</td>
</tr>
<tr>
<td>1</td>
<td>.11</td>
<td>.01</td>
<td>.00</td>
<td>.00</td>
<td>-.00</td>
<td>-.00</td>
<td>-.00</td>
</tr>
<tr>
<td>2</td>
<td>-.07</td>
<td>.07</td>
<td>-.03</td>
<td>.01</td>
<td>-.00</td>
<td>.00</td>
<td>-.00</td>
</tr>
<tr>
<td>3</td>
<td>-.60</td>
<td>.31</td>
<td>-.17</td>
<td>.12</td>
<td>-.08</td>
<td>.05</td>
<td>-.03</td>
</tr>
<tr>
<td>4</td>
<td>-.72</td>
<td>-.09</td>
<td>.29</td>
<td>-.01</td>
<td>-.07</td>
<td>.02</td>
<td>.00</td>
</tr>
<tr>
<td>5</td>
<td>-.100</td>
<td>.58</td>
<td>-.33</td>
<td>.15</td>
<td>-.07</td>
<td>.01</td>
<td>-.00</td>
</tr>
<tr>
<td>6</td>
<td>-.59</td>
<td>-.01</td>
<td>.09</td>
<td>.01</td>
<td>-.04</td>
<td>.01</td>
<td>.00</td>
</tr>
<tr>
<td>7</td>
<td>-.36</td>
<td>.09</td>
<td>-.02</td>
<td>.02</td>
<td>-.02</td>
<td>.01</td>
<td>-.01</td>
</tr>
</tbody>
</table>

and indicates clearly that \( b=3 \), \( s=3 \) and \( r=1 \), which is the final model obtained in Box and Jenkins (1976).

In recent work, de Gooijer and Heuts (1981) examine the application of the corner method to the identification of ARMA(p,q) models and report that it fails to work well in practice. However, we find that the \( \zeta \)-arrays in their paper indicate the simulated models quite clearly. The corner method indicates a few
candidates for the values of b, s and r, thereby providing an efficient guideline to transfer function identification. The principle of parsimony should be applied in cases where more than one combination of parameters is feasible. Since the 5(B) polynomials are usually simple in practice, the use of the corner method to transfer function identification seems to be less difficult than for ARMA models.

3. A PROCEDURE FOR THE IDENTIFICATION OF A TRANSFER FUNCTION MODEL

Based on the results in Section 2, we now present a five-step procedure which introduces filtering and least-squares estimation in the transfer function identification.

Step 1:

Build ARMA models for all input series after the series are appropriately differenced to achieve stationarity. If no AR factors are found or the roots of the AR factors are large (not close to 1), proceed to Step 2. If there are processes with AR roots close to 1, choose a common filter from the AR factors. Apply this filter to all input series and the output series.

Step 2:

Perform least-squares estimation of the transfer function weights for the series obtained from Step 1. The values $K_i$ should be chosen from subject-matter considerations and should be sufficiently large to avoid truncation bias. It is also important to check the SACT of the residuals since they provide information about the reliability of the usual least-squares hypothesis testing. It is recommended to omit the unnecessary terms in (2.4) if it is clear that they can be deleted.

Step 3:

Build an ARMA model for the residuals computed from the linear model selected in Step 2. If the residuals are white noise, proceed to Step 5. If not, go to Step 4.
Step 4:

Using the Step 3 ARMA model as a filter, perform OLS estimation of the transfer function weights based on the filtered series. Alternatively, we can also estimate the full transfer function-noise model jointly by nonlinear least squares. The significance tests of the weights can be carried out in the usual regression manner.

Step 5:

If no prefiltering was used in Step 1, the noise model is the one obtained in Step 4. Otherwise compute the noise of the original output series by using the transfer function weights from Step 2 or 4 and identify an ARMA model for the noise. Then, obtain a rational form \( w_i(B)/\delta_i(B) \) for input series \( X_i \) by using the corner method on \( V_i(B) \), if necessary. Note that the corner method should be used only if some of the impulse response weights are significant.

4. Example

To illustrate the identification method a simulated example is presented in this section.

The simulation model is

\[
Y_t = (2B^3 + 4B^4)X_{1t} + \frac{1.5B^2 + 3B^3}{1-B + .24B^2}X_{2t} + \varepsilon_t, \quad t=1, \ldots, 100 \tag{4.1}
\]

The models for \( \varepsilon_t, X_{1t}, \) and \( X_{2t} \) are

\[
(1 - 1.3B + .4B^2)\varepsilon_t = e_t, \quad e_t \sim N(0,2),
\]

\[
(1 - 1.4B + .48B^2)X_{1t} = b_t, \quad b_t \sim N(0,1),
\]

and

\[
(1 - .7B)X_{2t} = c_t, \quad c_t \sim N(0,2),
\]

where \( e_t \) is independent of \( b_t \) and \( c_t \), and \( b_t \) and \( c_t \) are contemporaneously correlated with correlation 0.7.

Following Box and Jenkins' guidelines, the ARMA models for \( Y_t, X_{1t} \) and \( X_{2t} \) are obtained as

(1)
\( (1 - 2.17 B + 1.95 B^2 - 1.02 B^3 + 0.28 B^4) Y_t = \hat{a}_t, \quad \hat{\sigma}_a^2 = 42.377, \)
\( (1 - 1.34 B + 0.37 B^2) X^*_t = \hat{a}_t, \quad \hat{\sigma}_b^2 = 1.012, \)
and
\( (1 - 0.69 B) X^*_t = \hat{c}_t, \quad \hat{\sigma}_c^2 = 2.118. \)

The AR polynomial for \( X^*_t \) can be factored into \((1 - 0.95B)(1 - 0.9B)\). Comparing the AR factors in the input series, we choose \((1 - 0.95B)\) as a common filter. Table I lists the OLS and

<table>
<thead>
<tr>
<th>TABLE I: Estimates of Transfer Function Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>----</td>
</tr>
<tr>
<td>( \phi_1 )</td>
</tr>
<tr>
<td>( \phi_2 )</td>
</tr>
<tr>
<td>( \phi_3 )</td>
</tr>
<tr>
<td>( \phi_4 )</td>
</tr>
<tr>
<td>( \phi_5 )</td>
</tr>
<tr>
<td>( \phi_6 )</td>
</tr>
<tr>
<td>( \phi_7 )</td>
</tr>
<tr>
<td>( \phi_8 )</td>
</tr>
<tr>
<td>( \phi_9 )</td>
</tr>
<tr>
<td>( \phi_{10} )</td>
</tr>
<tr>
<td>( \phi_{11} )</td>
</tr>
<tr>
<td>( \phi_{12} )</td>
</tr>
<tr>
<td>( \phi_{13} )</td>
</tr>
<tr>
<td>( \phi_{14} )</td>
</tr>
<tr>
<td>( \phi_{15} )</td>
</tr>
<tr>
<td>( \phi_{16} )</td>
</tr>
<tr>
<td>( \phi_{17} )</td>
</tr>
<tr>
<td>( \phi_{18} )</td>
</tr>
<tr>
<td>( \phi_{19} )</td>
</tr>
<tr>
<td>( \phi_{20} )</td>
</tr>
<tr>
<td>( \phi_{21} )</td>
</tr>
<tr>
<td>( \phi_{22} )</td>
</tr>
<tr>
<td>( \phi_{23} )</td>
</tr>
<tr>
<td>( \phi_{24} )</td>
</tr>
<tr>
<td>( \phi_{25} )</td>
</tr>
<tr>
<td>( \phi_{26} )</td>
</tr>
</tbody>
</table>

*The t-values in columns 1, 2 and 4 are provided for information only. They do not necessarily reflect the correct significance levels of the estimates.
(1) The backcasting method is used in the estimation of ARIMA models.
GLS estimates of the transfer function weights and their t-values for the original and the prefiltered series.

The computations in TABLE I (except for the first column) are performed in single precision on IBM 3032 using BMDQ2T (Liu 1979). When single precision is used in computing the OLS estimates of transfer function weights for the original series, the computation is rejected because the $X'X$ matrix is too close to singularity. The estimates can be obtained (as shown in column 1) by using double precision. The results in column 1 seem to exhibit the simulated pattern well even though $X_{1t}$ is highly autocorrelated. However, in other simulation tests we found that erroneous results are obtained even if double precision is used. When prefiltering is performed prior to OLS estimation (column 2), the transfer function weights can be estimated in single precision. These weights also exhibit the simulated pattern well. Since the residuals of the model in column 2 follow an AR(1) process, the transfer function weights are also estimated by using nonlinear least squares, incorporating this noise process (column 3). Longer lags are used since they are shown to be necessary in the previous two analyses. It is easy to identify the orders $r=0$, $s=2$ and $b=3$ from these estimates. As far as the model for $X_{2t}$ is concerned, we may use the corner method described in Section 2.2. The $\zeta$-array for the transfer function weights of $X_{2t}$ is shown in TABLE II and indicates that $r=2$, $s=2$ and $b=2$. An examination of the residuals of the model in column 3 revealed an AR(2) process, consis-

<table>
<thead>
<tr>
<th>$\zeta$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-0.03</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
<td>-0.00</td>
</tr>
<tr>
<td>1</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>0.29</td>
<td>0.09</td>
<td>0.01</td>
<td>-0.00</td>
<td>-0.00</td>
<td>-0.00</td>
<td>-0.00</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>0.72</td>
<td>0.50</td>
<td>0.35</td>
<td>0.24</td>
<td>0.17</td>
<td>0.12</td>
</tr>
<tr>
<td>4</td>
<td>0.94</td>
<td>0.23</td>
<td>-0.00</td>
<td>-0.05</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>5</td>
<td>0.66</td>
<td>0.08</td>
<td>0.02</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>6</td>
<td>0.38</td>
<td>-0.04</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>7</td>
<td>0.27</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
tent with the originally simulated model. The final estimates and
standard errors (in parentheses) of the parameters are:

\[
\begin{align*}
\hat{\omega}_{11} &= 1.698 (.213) \\
\hat{\omega}_{12} &= 3.987 (.191) \\
\hat{\omega}_{21} &= 1.404 (.194) \\
\hat{\omega}_{22} &= 3.176 (.194) \\
\hat{\sigma}_e^2 &= 2.466
\end{align*}
\]

\[
\begin{align*}
\hat{\delta}_{21} &= .965 (.043) \\
\hat{\delta}_{22} &= -.228 (.038) \\
\phi_1 &= 1.271 (.094) \\
\phi_2 &= -.429 (.093)
\end{align*}
\]

For the purpose of comparison, we also performed two other
identification methods on the simulated data. First, following
Priestley (1971), and Haugh and Box (1977), we may regress \( \hat{a}_t \) on
\( \hat{b}_t \) and \( \hat{c}_t \) using the model in (2.4). The least-squares estimates
of the transfer function weights and their t-values are listed in
the last column of TABLE I. From these weights we obtain the
model for the prewhitened series as

\[
\hat{a}_t = (w_{11} + w_{12}B + w_{13}B^2 + w_{14}B^3 + w_{15}B^4) B^3 \hat{b}_t
\]

\[
+ \left( \frac{w_{21} + w_{22}B + w_{23}B^2 + w_{24}B^3}{1 - \delta_{21} B} \right) B^2 \hat{c}_t + \hat{e}_t
\]

Substituting \( \hat{a}_t, \hat{b}_t, \) and \( \hat{c}_t \) with the corresponding prewhitening
models, we obtain a tentative model for the system. Note that
this tentative model is much more complicated than the simulated
model. It requires a great deal of experience and effort to
simplify it to the actual model.

Secondly, if we apply Box and Jenkins' (1976) single pre-
whitening method to each pair of input and output series, we obtain
the SCCF between the prewhitened input series and the corre-
ponding filtered output series (i.e., \( r_{\hat{a},(2)} \) and \( r_{\hat{c},(2)} \)) as
shown in TABLE III. The patterns of the SCCF in this situation
are somewhat misleading due to the high correlation between the
two input series.
MULTIPLE-INPUT TRANSFER FUNCTION MODELS

TABLE III
SCCF's for the Simulated Example

<table>
<thead>
<tr>
<th>Lag</th>
<th>-7</th>
<th>-6</th>
<th>-5</th>
<th>-4</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>r_{xy}(2)</td>
<td>.05</td>
<td>.01</td>
<td>.09</td>
<td>.07</td>
<td>.10</td>
<td>.06</td>
<td>.13</td>
<td>.09</td>
<td>.06</td>
<td>.06</td>
<td>.12</td>
<td>.42*</td>
<td>.36*</td>
<td>.08</td>
<td>.05</td>
</tr>
<tr>
<td>r_{xy}(4)</td>
<td>.13</td>
<td>.20</td>
<td>.16</td>
<td>.13</td>
<td>.08</td>
<td>.07</td>
<td>.07</td>
<td>.06</td>
<td>.06</td>
<td>.12</td>
<td>.41*</td>
<td>.51*</td>
<td>.33*</td>
<td>.20*</td>
<td>.16</td>
</tr>
</tbody>
</table>

*The standard errors of the sample cross-correlations are about 10. ** is used to denote significance at \( p < .05 \). The definition of \( r_{xy}(k) \) is the same as in Box and Jenkins (1976).

In conclusion the simulated example illustrates the performance and simplicity of the proposed identification procedure while the other two methods provide not as good results.

6. DISCUSSION

This paper proposes filtering and least-squares estimation to obtain transfer function weights and the corner method to identify the rational form of the model. In practice, it is not necessary to prefilter the data unless one or more input series have autoregressive roots close to one in absolute value. In that case, the choice of a common filter is flexible because it does not alter the transfer function of a stationary system.

Unlike the proposed procedure, the double prewhitening method produces results which are heavily dependent on the choice of a prewhitening model. A simple prewhitening model may not whiten the series completely, thereby casting doubt on the use of the nice theoretical results in this approach. An elaborate model, on the other hand, will typically cause the transfer function to be over-structured.

The procedure is simpler when applied to original as opposed to prefiltered series. Therefore, it may be advisable in some cases to analyze the original data by using double precision computations or even ridge regression techniques. However, since model identification is an exploratory process, it is desirable to work with stationary as opposed to near-nonstationary input series.
Efficient estimates of the transfer function weights may be obtained by applying OLS estimation on the series, filtered by the ARMA model of the output residuals, or by performing joint estimation of transfer function weights and ARMA coefficients by using nonlinear least-squares. In most cases the OLS and GLS estimates of the weights are rather close even if the OLS residuals are not white noise, but their t-values may be quite different.

The corner method provides a mechanical way of finding the orders of $w_i(B)$ and $\delta_i(B)$ when $\delta_i(B)$ is not 1. Other methods, such as the S-array technique by Gray, Kelley and McIntire (1978), may also be applied to identify a rational form. However, the corner method appears to be very informative and easy to use.

ACKNOWLEDGEMENTS

The authors wish to thank the editor and the referees' helpful comments and suggestions on an earlier version of this paper. This research was supported in part by National Science Foundation Grant MCS78-16827 and by a grant from the UCLA Academic Senate.

BIBLIOGRAPHY


Received January, 1980; Revised November, 1980 and February, 1981; Retyped June, 1981.

Recommended Anonymously and Refereed Anonymously.