DEPARTMENT OF DEFENCE
DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION
AERONAUTICAL RESEARCH LABORATORY
MELBOURNE, VICTORIA

Technical Note 42

LITERATURE REVIEW OF THE "ERRORS-IN-VARIABLES" APPROACH TO THE PROBLEM OF SYSTEM IDENTIFICATION

by

BETTY EMSLIE

This document has been approved for public release and sale; its distribution is unlimited.

Approved for public release.

© COMMONWEALTH OF AUSTRALIA 1993

JULY 1993
LITERATURE REVIEW OF THE "ERRORS-IN-VARIABLES" APPROACH TO THE PROBLEM OF SYSTEM IDENTIFICATION

by

BETTY EMSLIE

SUMMARY

The papers reviewed here deal with various aspects of the problem of system identification when all the observed variables are contaminated by noise errors. Theoretical issues of identifiability and practical methods for developing algorithms to estimate parameters of a dynamic system are considered. The total least squares approach and the Koopmans-Levin method are among the methods discussed. Both these methods are strongly based on matrix singular value decomposition, which is a computationally robust numerical technique. A book, which is a compilation of papers originating from a workshop on singular value decomposition and signal processing is included in the references.
CONTENTS

1.0 INTRODUCTION 1
2.0 INITIAL REVIEW OF LITERATURE SEARCHES 1
3.0 REVIEW OF SELECTED PAPERS 2
   3.1 Review of papers dealing with the problem of estimating 2
       the parameters in a transfer function model  
       of a dynamic system 
   3.2 Review of papers concerned with the mathematical 7
       and computational aspects of the methods  
       used for system identification 
4.0 CONCLUDING REMARKS 8
REFERENCES 8
DISTRIBUTION 8
DOCUMENT CONTROL DATA
1.0 INTRODUCTION

Estimation of system parameters using noise-corrupted data is of fundamental importance in the problem of system identification. Many identification methods are based on the assumption that there is no input noise in the system. However, this is not true in most practical situations.

A fitting technique, called total least squares (TLS), can be used for parameter estimation when there is noise on all the data, both input and output. TLS has a long history in the statistical literature where the method is also known as orthogonal regression or errors-in-variables regression. In recent years, the TLS approach has been applied in fields outside statistics.

Other methods are also available for system identification when all observed variables are contaminated by noise errors. These methods, together with the TLS approach, are discussed in the reviews in Section 3.0.

The identification problem being considered is that of estimating the parameters in a transfer function model of a dynamic system. This problem, given in its errors-in-variables form is

\[ A(q^{-1})y_0(t) = B(q^{-1})u_0(t) \]
\[ u(t) = u_0(t) + \Delta u(t) \]
\[ y(t) = y_0(t) + \Delta y(t) \]

Here \( u_0(t) \) and \( y_0(t) \) are the unmeasurable noise-free inputs and outputs, \( u(t) \) and \( y(t) \) the noisy measured inputs and outputs while \( \Delta u(t) \) and \( \Delta y(t) \) represent all stochastic disturbances to the inputs and outputs, respectively. \( A(q^{-1}) \) and \( B(q^{-1}) \) are polynomials of order \( n_a \) and \( n_b \), respectively, in the backward shift operator \( q^{-1} \) and have the following form:

\[ A(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a} \]
\[ B(q^{-1}) = b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b} \]

The backward shift operator \( q^{-1} \) is such that

\[ q^{-1} y(t) = y(t-1) \]

2.0 INITIAL REVIEW OF LITERATURE SEARCHES

Several data base searches were carried out by staff of the ARL library. Preliminary study of these literature searches identified four main groups of papers. These are:

1. Papers by B D O Anderson and his co-authors. Anderson is with the Department of Systems Engineering, Australian National University.

2. Papers by S Van Huffel, J Vandewalle and others. Many of these authors are associated with the Katholieke Universiteit Leuven.

3. Papers and a book by G H Golub and C F Van Loan. These authors are from the Departments of Computer Science at Stanford and Cornell Universities respectively.

4. A paper by K V Fernando and H Nicholson. These authors are with the Department of Control Engineering, University of Sheffield.
### 3.0 REVIEW OF SELECTED PAPERS

The papers selected for review fall into two broad categories. Papers in the first category deal with the problem of estimating the parameters in a transfer function model of a dynamic system. (Refs. 1, 2, 3, 4, 10 and 11). Papers in the second category are mainly concerned with the mathematical and computational aspects of the methods used for system identification. (Refs. 5, 6, 7, 8 and 9).

### 3.1 Review of papers dealing with the problem of estimating the parameters in a transfer function model of a dynamic system

Reference 11 Söderström T (1981)

The purpose of this paper is to investigate when and how a dynamic system can be identified from noise-corrupted input-output data. After brief reviews of some previous publications, the author provides a formal statement of the identification problem under discussion and a review of general identifiability properties.

The rest of the paper is devoted to three different methods for identifying systems from noisy data. These methods are:

1. Spectral analysis (SA).
2. Correlation techniques (CT).
3. Joint output approach (JO).

Details of these methods and their advantages and disadvantages are discussed in the paper. The CT method can be interpreted as an instrumental variable (IV) method using delayed inputs as instruments.

The problem being considered is:

\[
\begin{align*}
z(t) &= G(q^{-1})w(t) \\
u(t) &= w(t) + \nu(t) \\
y(t) &= z(t) + F(q^{-1})e(t)
\end{align*}
\]

Here \( w(t) \) and \( z(t) \) are the unmeasurable noise-free inputs and outputs, while \( u(t) \) and \( y(t) \) are the noisy measured inputs and outputs.

The input measurement noise, \( \nu(t) \), is assumed to be white and uncorrelated with the output noise. The output noise is assumed to describe both measurement noise and process noise acting on the system. It is assumed that the output noise can be described as \( F(q^{-1})e(t) \) where \( F(q^{-1}) \) is an asymptotically stable filter and \( e(t) \) is white noise.

In some cases, additional assumptions on the input signal are used. In these cases, \( w(t) \) is described as

\[
w(t) = H(q^{-1})f(t)
\]

where \( H(q^{-1}) \) is an asymptotically stable filter and \( f(t) \) is white noise, uncorrelated with \{e(\cdot)\} and \{\nu(\cdot)\}. The variance matrices of \( e(t) \), \( f(t) \) and \( \nu(t) \) are denoted respectively by \( \Lambda_e \), \( \Lambda_f \) and \( \Lambda_\nu \).
To illustrate the properties of the three methods, two numerical examples were considered.

System S1:

\[
G(q^{-1}) = \frac{1.0q^{-1}}{1 - 0.8q^{-1}}
\]

\[
F(q^{-1}) = 1
\]

\[
H(q^{-1}) = \frac{1}{1 - 0.5q^{-1}}
\]

The noise variances were \( \lambda_e^2 = \lambda_f^2 = \lambda_y^2 = 1 \). The undisturbed signals have variances \( Ew(t)^2 = 1.33 \) and \( Ez(t)^2 = 25.9 \). Ten different realizations of 500 data pairs were generated.

System S2:

\[
G(q^{-1}) = \frac{1.0q^{-1} + 0.5q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2}}
\]

\[
F(q^{-1}) = 1
\]

\[
H(q^{-1}) = \frac{1}{1 - 0.9q^{-1}}
\]

The noise variances were \( \lambda_e^2 = \lambda_f^2 = \lambda_y^2 = 1 \). The undisturbed signals have variances \( Ew(t)^2 = 5.26 \) and \( Ez(t)^2 = 320.2 \). Ten different realizations of 200 data pairs were generated.

The numerical results support the following conclusions.

The modified spectral analysis will not always work, since estimated spectral densities are not necessarily decomposable. The computational requirements are far higher than those for correlation techniques but the accuracy (when estimates can be computed) is also considerably better than CT. The main merits of the spectral analysis approach is to conceptually investigate the identifiability properties. The joint output approach with a prediction error criteria requires much computer time. The big advantage is that very accurate parameter estimates are obtained.

Reference 1 Fernando K V and Nicholson H (1985)

The objective of this paper is to develop the Koopmans-Levin (KL) method of identification of linear systems using singular value decomposition (SVD). As Söderström (Ref. 11) has not considered the KL method, the results here complement his review of different methods of identification in the presence of input and output noise.

In this paper, three algorithms have been developed for identification of univariate linear discrete-time systems. The technique is based essentially on spectral decomposition of a covariance matrix formed using input-output data. While it is not necessary to know the variances of the input and output noise explicitly, it is assumed that the ratio of these variances is known.

The three algorithms are:

1. The batch or direct algorithm.
2. Sequential updating of the batch algorithmic solution. This algorithm can be used to update the batch method when one or more new input-output data records are available.
3. Sequential estimation using an information matrix. This algorithm is based on an information
matrix which is defined as the inverse of the covariance matrix.

The authors have tested the batch algorithm extensively on the two numerical examples de-
scribed by Söderström (Ref. 11). They have also tested a square-root implementation of the
information matrix approach using the same examples.

The authors demonstrate that the accuracy of the estimates using the KL method is comparable
with that of the joint output method, and superior to the other methods described by Söderström
(Ref. 11). However, numerical requirements of the KL method are very much less than with
the joint output method.


Ref. 10 considers the estimation of the parameters in a transfer function (TF) model, given in
its errors-in-variables form. i.e.

\[ A(q^{-1})y_0(t) = B(q^{-1})u_0(t) \]  
\[ u(t) = u_0(t) + \Delta u(t) \]  
\[ y(t) = y_0(t) + \Delta y(t) \]

Here \( u_0(t) \) and \( y_0(t) \) are the unmeasurable noise-free inputs and outputs, \( u(t) \) and \( y(t) \) the noisy
measured inputs and outputs while \( \Delta u(t) \) and \( \Delta y(t) \) represent all stochastic disturbances to
the inputs and outputs, respectively. \( A(q^{-1}) \) and \( B(q^{-1}) \) are polynomials of order \( na \) and \( nb \),
respectively, in the backward shift operator \( q^{-1} \) and have the following form:

\[ A(q^{-1}) = 1 + a_1 q^{-1} + \ldots + a_{na} q^{-na} \]  
\[ B(q^{-1}) = b_1 q^{-1} + \ldots + b_{nb} q^{-nb} \]

This paper evaluates the total least squares (TLS) approach and the instrumental variable (IV)
techniques with respect to parameter estimation in the TF model (1). The TLS technique is
compared with several IV methods for estimating the parameters in the TF model (1), given
only the normal operating input and output of the system over a limited period of time. It is
assumed that the noise sequences added to the noise-free inputs and outputs are independent,
discrete, stationary and white with zero mean and equal variance.

The authors conclude that TLS is especially useful in the estimation of the parameters of a
TF model for short sample lengths, when the outputs and possibly the inputs are disturbed by
white noise. TLS is particularly recommended when the zeros of the polynomial \( A(z) \) approach
the unit circle or when both the inputs and outputs are disturbed. In other cases, TLS and IV
methods give comparable results.

The authors suggest that the usefulness of the TLS method may be questioned for large sample
lengths when only output noise is present. The computational efficiency is then an important
factor. The TLS solution involves off-line computations. It is thus computationally not so
attractive as the IV method which exists in recursive form, allowing good on-line estimation
schemes.

The TLS method is simpler to use since the IV need not be generated. However, the TLS
method is more restrictive since the IV approach allows arbitrary noise models.
Refs. 5–8 discuss details of the total least squares problem.

**Reference 2** Anderson B D O (1985)
**Reference 3** Green M and Anderson B D O (1986)
**Reference 4** Anderson B D O and Deistler M (1987)

All three papers consider errors-in-variables identification problems, which are problems where all observed variables are contaminated by noise errors. Given noisy measurements of an n-vector process, these papers consider such questions as: "How many independent (possibly linear) relations exist among the non-noisy components of the process?", and "What is the set of such linear relations for which the data are compatible?".

In the main, such questions have been posed under a collection of standing assumptions: stationarity of the underlying processes, linearity of the underlying relations, and availability of covariance data. In Ref. 2, a dynamic problem where \( n = 2 \) is discussed. In Ref. 3, dynamic problems, where the measured vector of dimension \( 2m \) has prescribed partitioning into an \( m \)-dimensional input vector and an \( n \)-dimensional output vector, are discussed. In Ref. 4, the authors focus on the case \( n = 3 \).

In Refs. 2 and 3, the authors hypothesize causality of a transfer function (transfer matrix in Ref. 3) appearing in a dynamic errors-in-variable model and make certain other reasonable assumptions. Then the concern of these two papers is show that it is possible to parameterize the class of transfer functions (transfer matrices in Ref. 3) consistent with the available data in a finite-dimensional way.

In Ref. 2, the author postulates the existence of three real random sequences \( \{x_k^z\}, \{u_k\}, \{v_k\} \), mutually independent and stationary, together with a time-invariant linear system defined by a real bounded linear causal convolution operator \( \{w_k, k \geq 0\} \) mapping \( \{x_k^z\} \) into a sequence \( \{y_k^z\} \) according to

\[
y_k^z = \sum_{-\infty}^{k} w_{k-l} x_l^z
\]

The processes \( \{x_k^z\}, \{y_k^z\} \) are not available for measurement, but rather one can measure for \( k \in (-\infty, \infty) \)

\[
x_k = x_k^z + u_k \quad (3.2a)
\]
\[
y_k = y_k^z + v_k \quad (3.2b)
\]

Attention is confined to Gaussian processes. The concern is not to identify a unique \( \{w_k\} \), but to characterize the class of \( \{w_k\} \) which fit the data.

The power spectrum matrix of \( [x \ y]^T \)

\[
\begin{bmatrix}
\sigma_{xx}(\omega) & \sigma_{xy}(\omega) \\
\sigma_{yx}(\omega) & \sigma_{yy}(\omega)
\end{bmatrix}
\]

is termed the standard data.

The standing assumptions are set out below.

The random sequences \( \{x_k^z\}, \{u_k\} \) and \( \{v_k\} \) are independent, zero mean processes of which measurements are available according to (3.2). Further, they are regular processes, so that each possesses a bounded spectral density.
In addition, it is assumed that the \( x^*_k \) process is generated by

\[
x^*_k = \sum_{-\infty}^{k} w^*_i \epsilon_i
\]

where \( \{\epsilon_k\} \) is a zero mean, stationary, white noise sequence and \( \{w^*_k, k \geq 0\} \) is a causal impulse response satisfying

\[
\sum_{k=0}^{\infty} \rho^k |w^*_k| < \infty \text{ for some } \rho > 1.
\]

It is also assumed that the impulse response \( \{w_k, k \geq 0\} \) satisfies the strengthened stability requirement

\[
\sum_{k=0}^{\infty} \rho^k |w_k| < \infty \text{ for some } \rho > 1.
\]

It is supposed that \( W(z) \) is the transfer function associated with the sequence \( \{w_k\} \), defined by

\[
W(z) = \sum_{0}^{\infty} w_k z^k
\]

Here, \( z \) is the backward shift operator, and the standing assumptions ensure that \( W(z) \) is analytic inside \( |z| < \rho \).

The problem is to define, to the fullest extent possible, what \( W(z) \) is, under the standing assumptions and given the standard data.

Under the standing assumptions together with the standard data, the following relationships are derived.

\[
\arg\left(\frac{\sigma_{xy}(\omega)}{\sigma_{zz}(\omega)}\right) = \arg[W(e^{j\omega})] = \arg\left[\frac{\sigma_{yy}(\omega)}{\sigma_{yz}(\omega)}\right] \quad (3.3a)
\]

and

\[
\left|\frac{\sigma_{xy}(\omega)}{\sigma_{zz}(\omega)}\right| \leq |W(e^{j\omega})| \leq \left|\frac{\sigma_{yy}(\omega)}{\sigma_{yz}(\omega)}\right| \quad (3.3b)
\]

Then the remainder of the paper is devoted to constructing \( W(e^{j\omega}) \) satisfying (3.3) and parameterizing the solution set of (3.3). The key idea is to construct the magnitude of \( W(e^{j\omega}) \) from the phase, which is known from (3.3a), by using a formula from analytic function theory relating the real and imaginary parts of analytic functions.

The information in (3.3) coupled with an assumption that \( W(e^{j\omega}) \) is causal determines the number of zeros of \( W(e^{j\omega}) \) inside the unit circle. If there are \( N \) such zeros, the solution set is shown to be an \( N + 1 \) parameter family. The simplest case is that of minimum phase \( W(z) \) — those where there are no zeros inside the unit circle. In the case of minimum phase transfer functions, the parameter is a scalar.

The work of Anderson discussed above has been extended, by Green and Anderson, to multivariable errors-in-variables models in Ref. 3. In the multivariable case, it is still possible to obtain formulas analogous to (3.3). However, there are difficulties involved in applying the scalar
solution technique outlined above. The scalar technique proceeds from phase information, but what is the phase of a matrix?

The solution technique presented in Ref. 3 is based on the factorization of matrix valued functions, a special case of which is the better known spectral factorization. The factorization theory is well developed and is used extensively in the theory of integral equations. A summary of matrix factorization theory is presented in Section III of the paper. Then in Section IV, the authors completely characterize the solution set of the multivariable errors-in-variable problem in terms of the factors of the cross-spectrum matrix $\sum_{yx}(\omega)$.

In Ref. 4 the authors deal with two different, though related, problems. The first is a complex version of a static errors-in-variables problem. This problem has been analysed in detail and a solution has been given for the case of three scalar variables. In the second problem, the authors have searched for causal solutions of a dynamic three variable problem. A checkable necessary condition for a solution to exist has been exhibited, and an indication given of how a solution might (numerically) be obtained.

3.2 Review of papers concerned with the mathematical and computational aspects of the methods used for system identification

Reference 8 Golub G H and Van Loan C F (1980)

In this paper a singular value decomposition analysis of the total least squares (TLS) problem is presented and an algorithm for solving the TLS problem is proposed.


The authors indicate that in some TLS problems, called nongeneric, the algorithm proposed by Golub and Van Loan (Ref. 8) fails to compute a finite TLS solution. This paper generalizes their TLS computations in order to solve these nongeneric TLS problems. An algorithm is presented which includes the proposed generalizations.


Total least squares (TLS) is one method of solving overdetermined sets of linear equations $AX \approx B$ that is appropriate when there are errors in both the observation matrix $B$ and the data matrix $A$. In some linear parameter estimation problems, some of the variables may be observed without error. This implies that some of the columns of $A$ are assumed to be known exactly. To deal with this case, a computationally efficient and numerically reliable Generalized TLS algorithm GTLS, based on the Generalized SVD (GSVD), is developed.


This paper compares least squares (LS) and total least squares (TLS) methods for solving over- determined sets of linear equations, $AX \approx B$, arising in linear parameter estimation problems. The authors examine, via simulations, how perturbations on both the data matrix $A$ and the observation matrix $B$ affect the accuracy of the TLS and LS problem.

Reference 9 Deprettere F, editor (1988)

This book owes its origins to the Workshop on SVD AND SIGNAL PROCESSING, which was held at the Les Houches summer school for Physics, September 21–23, 1987.

A selection of papers is presented, dealing with matrix singular value decomposition, applications of SVD in signal processing and system identification and related algorithms and implementation architectures.
4.0 CONCLUDING REMARKS

The papers reviewed in Section 3.0 discuss both theoretical issues of identifiability and practical methods for system identification when all observed variables are contaminated by noise errors. Details are provided for a number of different methods of estimating the parameters in a transfer function model of a dynamic system. Advantages and disadvantages of each method are discussed, and in some cases comparisons between methods have been carried out. Lists of references in the papers reviewed provide further avenues for information as required.

REFERENCES


DISTRIBUTION

AUSTRALIA

Defence Organisation

Defence Central
  Chief Defence Scientist
  AS Science Corporate Management
  FAS Science Policy
  Counsellor, Defence Science, London (Doc Data sheet only)
  Counsellor, Defence Science, Washington (Doc Data sheet only)
  Scientific Adviser, Defence Central
  OIC TRS, Defence Central Library
  Document Exchange Centre, DSTIC (8 copies)
  Defence Intelligence Organisation
  Librarian, Defence Signals Directorate, (Doc Data sheet only)

Aeronautical Research Laboratory
  Director
  Library
  Chief Airframes and Engines Division
  Author: Betty Emslie
  T. Ryall
  P. Farrell
  J. D'Cruz

Materials Research Laboratory
  Director/Library

Defence Science & Technology Organisation - Salisbury
  Library

Navy
  Navy Scientific Adviser (3 copies Doc Data sheet only)

Army
  Scientific Adviser - Army (Doc Data sheet only)

Air Force
  Air Force Scientific Adviser (Doc Data sheet only)
  OIC ATF, ATS, RAAFSTT, WAGGA (2 copies)

Other Organisations
  NASA (Canberra)
  AGPS

Universities and Colleges
  Adelaide
    Professor Mechanical Engineering
Flinders
Library

LaTrobe
Library

Melbourne
Engineering Library

Monash
Hargrave Library

Newcastle
Library
Professor R. Telfer, Institute of Aviation

New England
Library

Sydney
Engineering Library

NSW
Physical Sciences Library
Head, Mechanical Engineering
Library, Australian Defence Force Academy

Queensland
Library

Tasmania
Engineering Library

Western Australia
Library
Head, Mechanical Engineering

RMIT
Library
Mr M.L. Scott, Aerospace Engineering

University College of the Northern Territory
Library

SPARES (6 COPIES)

TOTAL (50 COPIES)
The papers reviewed here deal with various aspects of the problem of system identification when all the observed variables are contaminated by noise errors. Theoretical issues of identifiability and practical methods for developing algorithms to estimate parameters of a dynamic system are considered. The total least squares approach and the Koopmans-Levin method are among the methods discussed. Both these methods are strongly based on matrix singular value decomposition, which is a computationally robust numerical technique. A book, which is a compilation of papers originating from a workshop on singular value decomposition and signal processing is included in the references.
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>16. Abstract (200 words)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Imprint</td>
<td>Aeronautical Research Laboratory, Melbourne</td>
<td></td>
</tr>
<tr>
<td>18. Document Series and Number</td>
<td>19. WA Number</td>
<td>20. Type of Report and Period Covered</td>
</tr>
<tr>
<td>Technical Note 42</td>
<td>21 224E</td>
<td></td>
</tr>
<tr>
<td>21. Computer Programs Used</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Establishment File Ref(s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. Additional Information (as required)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>