IDENTIFIABILITY OF PARAMETRIC MODELS

Edited by

E Walter

Ecole Supérieure d'Electricité, Gif-sur-Yvette, France

Pergamon Press

IDENTIFIABILITY OF PARAMETRIC MODELS

by

E. WALTER
Ecole Supérieure d'Electricité
Gif-sur-Yvette, France

PERGAMON PRESS
OXFORD · NEW YORK · BEIJING · FRANKFURT
SÃO PAULO · SYDNEY · TOKYO · TORONTO

U.K. Pergamon Press, Headington Hill Hall, Oxford OX3 0BW, England

U.S.A. Pergamon Press, Maxwell House, Fairview Park, Elmsford, New York 10523, U.S.A.

PEOPLE'S REPUBLIC OF CHINA Pergamon Press, Qianmen Hotel, Beijing, People's Republic of China

FEDERAL REPUBLIC OF GERMANY Pergamon Press, Hammerweg 6, D-6242 Kronberg, Federal Republic of Germany

BRAZIL Pergamon Editora, Rua Eça de Queiros, 346, CEP 04011, São Paulo, Brazil

AUSTRALIA Pergamon Press Australia, P.O. Box 544, Potts Point, N.S.W. 2011, Australia

JAPAN Pergamon Press, 8th Floor, Matsuoka Central Building, 1-7-1 Nishishinjuku, Shinjuku-ku, Tokyo 160, Japan

CANADA Pergamon Press Canada, Suite 104, 150 Consumers Road, Willowdale, Ontario M2J 1P9, Canada

---

First edition 1987

Library of Congress Cataloging-in-Publication Data
Walter, Eric.
Identifiability of parametric models.
Edited, updated, and expanded papers from the 7th IFAC/IFORS Symposium on Identification and System Parameter Estimation held in July 1985 in York, England.
Includes bibliographies and indexes.
1. System identification—Congresses.
2. Parameter estimation—Congresses. I. IFAC/IFORS Symposium on Identification and System Parameter Estimation (7th: 1985: York, North Yorkshire)
II. Title.
QA402.W365 1987 003 87-2324

British Library Cataloguing in Publication Data
Walter, E.
Identifiability of parametric models
1. System theory
I. Title
003 Q295
ISBN 0-08-034929-3

In order to make this volume available as economically and as rapidly as possible the authors' typescripts have been reproduced in their original forms. This method unfortunately has its typographical limitations but it is hoped that they in no way distract the reader.

Printed in Great Britain by A. Wheaton & Co. Ltd., Exeter

FOREWORD

Why is identifiability receiving so much attention? There are two main reasons in my opinion. The serious one is that it is of great practical importance. Those involved in estimating parameters from measurements would of course like to know whether they stand any chance of succeeding. Whenever a model is not uniquely identifiable, there are several values of the parameters that correspond to exactly the same input-output behavior; and the very meaning of an attempt to estimate them is questionable, although we shall see in the sequel that there may still be some hope. A less serious (but very important!) reason is that it is a fascinating subject. Many researchers who planned to give it only a passing thought have found themselves trapped in a long-term project. Part of the attraction comes from the fact that the basic problem is very simple to explain, part from the fact that no one can claim to have solved it definitively.

The idea of this volume took shape at York during the 7th IFAC/IFORS Symposium on Identification and System Parameter Estimation in July 1985. Many of those working in the field of identifiability were present, and they shared the feeling that the subject was now mature enough for a coordinated presentation. Papers given at York form the backbone of the book. They have been edited, updated and significantly expanded. One paper has been written especially for the occasion. All the authors were aware of the subjects that were to be treated by the other contributors . They were asked to take advantage of the availability of the conference preprints and send each other suggestions for improvement and for maximizing complementarity.

Chapter 1, by Godfrey and DiStefano, is a **tutorial** that recalls the basic methods for structural identifiability testing. In addition it provides tools that can be used to obtain bounds on the possible values of unidentifiable parameters.

Chapters 2 and 3 are devoted to **linear time-invariant models**. Chapter 2, by Delforge, d'Angio and Audoly, presents methods that are especially interesting when dealing with large-scale models and proposes conjectures on global identifiability deserving further consideration. Chapter 3, by Chapman and Godfrey, addresses the problems of initial model selection and of generating the set of all models having exactly the same input-output behavior.

Chapters 4 and 5 deal with **nonlinear models**. Chapter 4 , by Vajda, gives a finite algebraic condition for the structural identifiability of the parameters of a class of **polynomial models**. It also addresses the detection of dependences among the parameters that can result from the effect of measurement noise even when the model considered is structurally identifiable. When such dependences occur, the proposed method suggests simplifying assumptions for removing the ambiguity. Chapter 5, by Lecourtier, Lamnabhi-Laggarigue and Walter, describes in a tutorial way various methods, based on recent results on Volterra and generating power series approaches, that can be used to test nonlinear models for structural identifiability. **Linear time-varying models** and **bilinear models** are treated as special cases.

Chapter 6, by Chavent, is concerned with **infinite-dimensional models**, such as those described by partial differential equations. It investigates the relations between identifiability and the well-posedness of the estimation problem.

Chapter 7, by Lecourtier and Raksanyi, describes the facilities offered by **computer algebra** for performing the algebraic manipulations required for testing a model for structural controllability, observability, identifiability or distinguishability. It also presents interesting conjectures on the relations between the global injectivity of an application and the global properties of its Jacobian.

Chapters 8 to 11 are devoted to the relations between **identifiability and parameter uncertainty**. Chapter 8, by Cobelli and Toffolo, describes methods that can be used to estimate parameters even when the model proves to be structurally unidentifiable. Chapter 9, by Hadaegh and Bekey, investigates the consequences of the fact that the model is only an approximation to the real system under study. Chapter 10, by Happel, Walter and Lahanier, shows by a realistic chemical example how it may be possible to obtain quantitative information on the parameters of interest even when there are several competing model structures that are neither uniquely identifiable nor distinguishable. Finally Chapter 11, by Walter and Pronzato, is devoted to robust experiment design, viewed as the maximization of some measure of identifiability.

I sincerely believe that the book as it stands gives a more comprehensive presentation of identifiability than could have been written by any of the authors alone. I hope you will share this view.

Eric Walter

CONTENTS

Infinite-Dimensional Models

Computer Algebra

Identifiability and Parameter Uncertainty

Chapter 1
IDENTIFIABILITY OF MODEL PARAMETERS

K. R. Godfrey* and J. J. DiStefano, III**

**Department of Engineering, University of Warwick, Coventry CV4 7AL, UK*
***Biocybernetics Laboratory and Departments of Computer Science and Medicine, University of California, Los Angeles, CA 90024, USA*

Abstract. The problem is whether the parameters of a model can be identified (uniquely or with more than one distinct solution) from a specified input-output experiment. If perfect data are assumed, and the models are linear and time-invariant, there are several approaches for identifiability analysis, alternatively referred to as structural, deterministic or a priori identifiability analysis, and five such approaches are described. Only one, based on the Taylor series expansion of the observations, is directly applicable to nonlinear or time-varying models.

When a model is unidentifiable from a proposed experiment, physically based constraints on the model often provide a means of computing finite bounds for the parameters (interval identifiability). This is illustrated for the class of linear, time-invariant models in which the system matrix is of compartmental form. Under certain conditions, model parameter intervals are sufficiently small for all practical purposes, and they are thus called quasiidentifiable.

The identifiability question in the presence of real, noisy data, often referred to in the literature as numerical or a posteriori identifiability, is classified and treated as a separate problem. In the context of parameter estimation, the numerical identifiability problem is the same as the well known and long studied problem of parameter estimation accuracy or precision, given that the parameters are known to be structurally identifiable. Thus, the identifiability question arises as an issue separate from parameter estimation in the classical or general sense only in the context of specifically structured models; only in this topological sense is it new or different.

Keywords. Identification; Laplace transforms; Linear systems; Modelling; Multivariable systems; Nonlinear systems; Parameter estimation; State-space methods; Time-domain analysis.

1. INTRODUCTION

The notion of identifiability is fundamentally a problem in uniqueness of solutions for specific attributes of certain classes of mathematical models. The identifiability problem usually has meaning in the context of unknown model parameters, although it occasionally has had other meanings. The usual question is whether or not it is possible to find a unique solution for each of the unknown parameters of the model, from data collected in experiments performed on the real system. It is clearly a critical aspect of the modelling process, especially when the parameters are analogs of physical attributes of interest and the model is needed to quantify them. Identifiability analysis is normally used to determine the extent to which a particular model is suitable for reduction of data from a specific parameter estimation experiment. As a consequence, it is also of significant value in experiment design.

If the model is deterministic and the data are noise-free, the problem is generally a nonlinear algebraic one and, unfortunately, solution of this algebraic problem, as we see below, is generally both nontrivial and nonunique for all but the simplest of models. In the presence of real, noisy data, the problem is compounded. Nevertheless, structural identifiability conditions for the noise-free case are minimal, necessary conditions for achieving a successful estimation of model parameters of interest from real input/output data.

Identifiability analysis was put on a formal basis by Bellman and Åström (1970), although specific models were considered a good deal earlier than this, e.g., Skinner et al (1959). Bellman and Åström describe this as *structural identifiability*, and the term *a priori identifiability* also has been used quite widely, on the grounds that the analysis can (and should) be done before a proposed experiment is carried out. The term *deterministic identifiability* also has been proposed, in an attempt to circumvent certain inadequacies of the adjective "structural" in cases where identifiability properties are dependent on the form of the input, normally considered to be an independent variable in a dynamic system model. These issues are treated anew in Section 3.1.

The question of identifiability in the presence of noisy data has come to be known as the *numerical* or *a posteriori identifiability* problem. In the context of model *parameter* estimation, this is none other than the classical problem of parameter estimation accuracy (precision) in disguise. We address the noise-free problem as one that must be resolved, ascertaining identifiability under ideal circumstances, prior to attempting the usually difficult problem of parameter estimation with real data. The latter issue is addressed in Section 5.

2. BASIC CONCEPTS AND IDENTIFIABILITY ANALYSIS FOR NOISE-FREE LINEAR TIME-INVARIANT MODELS

2.1 Basic Concepts and Linear Models

Before turning to formal definitions, we review the basic concept and models, for linear deterministic systems, and present several examples to illustrate some of the subtleties of the identifiability concept.

The following simple first-order model is adapted from Bellman and Åström (1970):

$$\dot{x}(t) = - p_1 x(t) + p_2 u(t) , \tag{1}$$

$$x(0) = 0 \tag{2}$$

$$y(t) = p_3 x(t) . \tag{3}$$

The model has three unknown parameters: p_1 , p_2 , p_3 . For any known u , the explicit solution of equations (1)-(3) is:

$$y(t) = p_2 p_3 \int_0^t e^{-p_1(t-\tau)} u(\tau) d\tau . \tag{4}$$

If the input is an impulse $u(t) = \delta(t)$,

$$y(t) = p_2 p_3 e^{-p_1 t} . \tag{5}$$

It is well known that a semilogarithmic plot of the data, represented as $y(t)$ for this model, yields the coefficient $A \equiv p_2 p_3$ and exponent $\lambda \equiv p_1$ of this model. Thus only p_1 and the product $p_2 p_3$ can be determined and not p_2 or p_3 , i.e. the model is *unidentifiable*. This is also clear from equation (4) for any known $u(t)$. If p_2 or p_3 were known, or if a unique functional relationship between p_2 and p_3 were known, all parameters could be uniquely determined from $y(t)$, and we could say the model (or model parameters) is (are) *uniquely (globally) identifiable*.

This result generalises quite easily for linear constant coefficient systems. A system of n first-order linear ordinary differential equations, depending on a set of P unknown constant parameters $p_1, p_2, \ldots, p_P$, may be written conveniently in vector-matrix form as

$$\dot{\underline{x}}(t,\underline{p}) = A(\underline{p})\underline{x}(t,\underline{p}) + B(\underline{p})\underline{u}(t) \tag{6}$$

with

$$\underline{x}(0^-,\underline{p}) = \underline{x}_0(\underline{p}) \tag{7}$$

$$\underline{y}(t,\underline{p}) = C(\underline{p})\underline{x}(t,\underline{p}) . \tag{8}$$

Suppressing the arguments, the n state variables of the model are denoted by the vector $\underline{x} \equiv [x_1\ x_2\ \ldots\ x_n]^T$, with initial conditions given by (7); $\underline{u} \equiv [u_1\ u_2\ \ldots\ u_r]^T$ is the vector of r known inputs, $\underline{y} \equiv [y_1\ y_2\ \ldots\ y_m]^T$ is the vector of m output (measurement) variables of the model and A , B and C are constant matrices of appropriate dimension, each consisting of some or all of the unknown parameters $\underline{p} \equiv [p_1\ p_2\ \ldots\ p_p]^T$. The superscript T represents the vector transpose operation.

2.2 The Laplace Transform or Transfer Function Approach

It is convenient to employ the Laplace transform of equations (6)-(8) for further analysis of the identifiability properties of linear constant coefficient models. We assume $\underline{x}_0(\underline{p}) = \underline{0}$ here, for convenience, but hasten to emphasise that the approach applies for any initial or other boundary conditions. With $\underline{x}_0 = \underline{0}$,

$$\underline{Y}(s,\underline{p}) = C(\underline{p})[sI - A(\underline{p})]^{-1} B(\underline{p})\underline{U}(s) . \tag{9}$$

The identifiability properties are established by examining the expressions in the powers of s in the numerators and denominators of the measured outputs, together with any other information available about $\underline{p}$.

A parameter p_i is *uniquely (globally) identifiable* from a set of measurements $y_k(t,\underline{p})$, $k = 1, 2 \ldots q$ if it can be evaluated uniqely from the equations for $y_k(t,\underline{p})$, plus any other information about p_i ; it is *locally identifiable* from $y_k(t,\underline{p})$ if it has a countable number (≥ 1) of solutions or *nonuniquely identifiable* if this number > 1 , while it is *unidentifiable* from $y_k(t,\underline{p})$ if it has an infinite number of solutions. If all p_i of a model are identifiable, the model is said to be identifiable. The definitions are easily extended for subsets of the parameter vector $\underline{p} \equiv [p_1\ p_2\ \ldots\ p_p]^T$. Also, it is often of interest to know which *combinations* of individually unidentifiable parameters are identifiable (uniquely or otherwise). We formalise and extend these definitions in Section 3.1, following presentation of several examples that illustrate specific problems.

Identifiability analyses of linear, single-input, single-output (SISO) models may be performed directly from the impulse responses. These generally have the form

$$y(t,\underline{p}) = \sum_{k=1}^{n} A_k(\underline{p}) \exp[\lambda_k(\underline{p})]t . \tag{10}$$

The n coefficients A_k and n exponents λ_k are typically determined by fitting this model output function to the data in a least squares sense and the $2n$ relationships among the A_k , λ_k and p_i determine the extent to which the p_i are identifiable.

Taking the Laplace transform of equation (10) for the case with $n = 2$ and distinct eigenvalues,

$$\begin{aligned} Y(s,\underline{p}) &= \frac{A_1}{s-\lambda_1} + \frac{A_2}{s-\lambda_2} \\ &= \frac{(A_1+A_2)s - (A_1\lambda_2 + A_2\lambda_1)}{s^2 - (\lambda_1+\lambda_2)s + \lambda_1\lambda_2} \\ &\equiv \frac{\beta_1 s + \beta_2}{s^2 + \alpha_1 s + \alpha_2} \end{aligned} \tag{11}$$

If there are no common factors in these numerator and denominator polynomials, it is clear that all of the $2n$ α and β coefficients (often called *moment invariants*) can be determined from the experiment (data), just as all of the $2n$ coefficients A_k and λ_k of the sum of expo-

nentials solution can be determined from the $y(t,\underline{p})$ data.

Identifiability has to be qualified as being for almost all parameter values. For example, particular (isolated) combinations of parameter values or particular input functions which give rise to pole-zero cancellations in $Y(s,\underline{p})$ do not invalidate the general analysis.

The following model is used to illustrate many of the concepts and problems introduced in this and the next several sections:

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}b_1 & 0\\ 0 & b_2\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix} \tag{12}$$

$$\begin{bmatrix}y_1\\ y_2\end{bmatrix} = \begin{bmatrix}c_1 & 0\\ 0 & c_2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} \tag{13}$$

It is easy to show that the transfer function matrix for this model is:

$$H(s) \equiv C(sI-A)^{-1}B$$

$$\equiv \begin{bmatrix}Y_1(s)/U_1(s) & Y_1(s)/U_2(s)\\ Y_2(s)/U_1(s) & Y_2(s)/U_2(s)\end{bmatrix}$$

$$= \frac{1}{\Delta(s)}\begin{bmatrix}c_1b_1(s-a_{22}) & c_1b_2a_{12}\\ c_2b_1a_{21} & c_2b_2(s-a_{11})\end{bmatrix} \tag{14}$$

where $\Delta(s) = s^2-(a_{11}+a_{22})s+a_{11}a_{22}-a_{12}a_{21}$. (15)

We now consider several individual experiments in turn.

Experiment 1: SISO Case. $u_1 \neq 0$, $u_2 = 0$ and only y_1 is measured. From equations (14) and (15),

$$Y_1(s) = K_1\frac{s+\beta_1}{s^2+\alpha_1 s+\alpha_2}U_1(s) \tag{16}$$

where K_1 , β_1 , α_1 and α_2 are assumed known and, from equation (14)

$$c_1b_1 = K_1 \tag{17A}$$

$$a_{22} = -\beta_1 \tag{17B}$$

while from equation (15),

$$a_{11} = -\alpha_1 + \beta_1 \tag{18A}$$

$$a_{12}a_{21} = \beta_1(\alpha_1-\beta_1)-\alpha_2 \ . \tag{18B}$$

Thus from this experiment a_{11}, a_{22} and the products c_1b_1 and $a_{12}a_{21}$ are uniquely identifiable, but the parameters a_{12} and a_{21} are unidentifiable individually. It is of interest to see whether a_{12} and a_{21} could be identified by either observing an additional output or by stimulating an additional input.

Experiment 2: Single-Input Dual-Output. $u_1 \neq 0$, $u_2 = 0$, y_1 and y_2 measured. From the transfer function matrix (14), the additional observation is of the form

$$Y_2(s) = \frac{K_2}{s^2+\alpha_1 s+\alpha_2} \tag{19}$$

so that the extra equation, in addition to equations (17) and (18), is

$$c_2b_1a_{21} = K_2 \tag{20}$$

which only gives a_{21} if c_2b_1 is known.

Experiment 3: Dual-Input Single-Output. $u_1 \neq 0$, $u_2 \neq 0$; y_2 not measured. If the two test-inputs are applied at different times, so that the two responses are completely distinguishable, then the extra equation is

$$c_1b_2a_{12} = K_3 \tag{21}$$

where K_3 is a known constant. In this case, a_{12} could be found only if c_1b_2 were known a priori. If, on the other hand, the perturbations are applied simultaneously, then from equations (14) and (15),

$$Y_1(s) = c_1\frac{sb_1U_1(s)-a_{22}b_1U_1(s)+a_{12}b_2U_2(s)}{s^2-(a_{11}+a_{22})s+a_{11}a_{22}-a_{12}a_{21}} \tag{22}$$

We now distinguish between two different input types and two different corresponding identifiability results. If the two input waveforms are the same, the numerator gives c_1b_1 and $(-c_1b_1a_{22}+c_1b_2a_{12})$, providing neither a_{22} nor a_{12} individually. But, suppose the input waveforms were different, e.g., with $u_1(t)$ a unit impulse and $u_2(t)$ a unit step, so that

$$Y_1(s) = c_1\frac{b_1s^2-b_1a_{22}s+b_2a_{12}}{s[s^2-(a_{11}+a_{22})s+a_{11}a_{22}-a_{12}a_{21}]} \ . \tag{23}$$

Now the numerator gives c_1b_1 and a_{22} and, if c_1b_2 were known, a_{12} .

Remark 1: The example illustrated in Experiment 3 is important because it illustrates that, for linear systems with more than one input, identifiability results may depend on whether the inputs are applied simultaneously or separatesly. And, if they are applied simultaneously, the result also can depend on the shape of the input waveforms. Thus, for multiple inputs, it is essential to examine the Laplace transform of the observations (available to the experimenter) rather than just the individual entries in the transfer function matrix.

Remark 2: All of the above examples have illustrated either unique identifiability or unidentifiability results. One way in which our two-state example can give a locally identifiable (nonunique) result is when the denominator of the output transform has intrinsic parameters as its roots, in which case they are indistinguishable. Consider the following example.

Experiment 4: SISO, $u_1 \neq 0$, $u_2 = 0$, y_1 not measured; prior knowledge that $a_{12} = 0$. From equations (14) and (15),

$$Y_2(s) = \frac{c_2 b_1 a_{21}}{(s-a_{11})(s-a_{22})} \cdot U_1(s) \qquad (24)$$

and we see that intrinsically unknown parameters a_{11} and a_{22} cannot be distinguished from the denominator, i.e. a_{11} and a_{22} are locally identifiable, with two solutions, from this experiment. For a system with three or more states, there are many other possibilities for non-uniqueness; see Section 2.7 for an example.

Remark 3: The Laplace transform method is conceptually simple and derivation of the equations relating observations to system parameters is straightforward. Unfortunately, these equations are not linear so that it is difficult to see whether there are multiple solutions or whether redundancy exists. It is not clear how to modify the model structure, input and observed variables to achieve identifiability for a configuration resulting in unidentifiability and it is necessary to re-work for each trial modification. No consistent simple structure carries over from one case to another, so that it is difficult to generalise conclusions drawn from specific cases.

Other examples of the Laplace transform approach are given by Skinner et al (1959), DiStefano et al (1975), Milanese and Molino (1975), Cobelli et al (1979b), Norton (1982) and Godfrey (1983), Chapter 6.

2.3 Taylor Series Expansion of the Observations

In this approach, the output waveforms are expanded in a Taylor series about $t = 0^+$, the successive terms of the expansion being expressed as functions of the model unknowns (Pohjanpalo, (1978)). Specifically, for an observation $y_i(t)$,

$$y_i(t) = y_i(0^+) + t\dot{y}_i(0^+) + \frac{t^2}{2!}\ddot{y}_i(0^+) + \ldots \qquad (25)$$

Successive derivatives are, in principle, measurable and contain information about the parameters to be identified. If the Laplace transform of equation (25) is taken,

$$Y_i(s) = \frac{1}{s}y_i(0^+) + \frac{1}{s^2}\dot{y}_i(0^+) + \frac{1}{s^3}\ddot{y}_i(0^+) + \ldots \qquad (26)$$

from which it may be seen that the test is equivalent to expanding the Laplace transform of the observation vector in a power series in s^{-1}.

To illustrate the approach, consider again Experiments 1 and 2 above, with only input 1 applied ($u_2 = 0$) and let the input be impulsive, $u_1(t) = D.\delta(t)$ with D known. This input can be incorporated as an initial condition, so that the equations may be written

$$\dot{x}_1(t) = a_{11}x_1(t) + a_{12}x_2(t), \quad t > 0 \qquad (27A)$$

$$\dot{x}_2(t) = a_{21}x_1(t) + a_{22}x_2(t), \quad t > 0 \qquad (27B)$$

with initial conditions $x_1(0^+) = b_1 D \qquad (28A)$

$$x_2(0^+) = 0 . \qquad (28B)$$

From equation (27A),

$$\dot{x}_1(0^+) = a_{11}x_1(0^+) + a_{12}x_2(0^+)$$

$$= a_{11}b_1 D \qquad (29A)$$

and similarly from equation (27B),

$$\dot{x}_2(0^+) = a_{21}b_1 D . \qquad (29B)$$

Differentiating equation (27A),

$$\ddot{x}_1(0^+) = a_{11}\dot{x}_1(0^+) + a_{12}\dot{x}_2(0^+)$$

$$= (a_{11}^2 + a_{12}a_{21})b_1 D . \qquad (30A)$$

Differentiating equation (27B),

$$\ddot{x}_2(0^+) = a_{21}\dot{x}_1(0^+) + a_{22}\dot{x}_2(0^+)$$

$$= a_{21}(a_{11} + a_{22})b_1 D . \qquad (30B)$$

Further differentiation of equation (27A) gives

$$\dddot{x}_1(0^+) = a_{11}\ddot{x}_1(0^+) + a_{12}\ddot{x}_2(0^+)$$

$$= [a_{11}(a_{11}^2 + a_{12}a_{21}) + a_{12}a_{21}a_{11} +$$

$$+ a_{12}a_{21}a_{22}]b_1 D . \qquad (31)$$

For Experiment 1, with only y_1 observed, information is obtained only from equations (28A), (29A), (30A) and (31), from which we see that, at the successive stages of differentiation, c_1b_1, a_{11}, $a_{12}a_{21}$ and a_{22} may be identified. With observation restricted to y_1, further differentiation does not yield a_{12} and a_{21} individually, as the reader is invited to confirm.

For Experiment 2, in which y_2 also is observed, equation (29B) gives $c_2b_1a_{21}$ uniquely, which if the product c_2b_1 is known, gives a_{21}, and then a_{12}, uniquely. Thus, as expected, the result is the same as for the Laplace transform analysis.

Remark: This example has illustrated that the method suffers from the same drawbacks as the Laplace transform approach as described in Remark 3 of Section 2.2. The method has not been used much for linear, time-invariant systems. It has the decided advantage, however, that it is applicable to nonlinear and time-varying systems, for which the Laplace transform approach is not. Applications to some nonlinear models are given in Section 3.

2.4 Markov parameter matrix approach

Grewal and Glover (1976) described a technique for testing whether two sets of parameter values can give the same observed responses for all admissible forms of excitation. First, the Markov parameter matrix

$$G = [(CB)^T(CAB)^T(CA^2B)^T \ldots (CA^{2n-1}B)^T] \qquad (32)$$

(where n is the model order) is formed. It is possible to use G itself for a global identifiability test by determining whether

$$G(\underline{p}) = G(\underline{p}') \Rightarrow \underline{p} = \underline{p}'$$

but this again suffers from the same drawbacks as the Laplace transform approach. What Grewal and Glover (1976) did was to find the rank of the

Jacobian of G with respect to the unknown model parameters. If the rank is equal to the number of parameters, the system is identifiable, but not necessarily uniquely. The rank test is also applicable to Jacobians from other approaches, for example, the equations resulting from the Laplace transform or Taylor series approaches.

Taking our two-state example,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$A^2 = \begin{bmatrix} a_{11}^2+a_{12}a_{21} & a_{12}(a_{11}+a_{22}) \\ a_{21}(a_{11}+a_{22}) & a_{22}^2+a_{12}a_{21} \end{bmatrix}$$

$$A^3 = \begin{bmatrix} a_{11}^3+2a_{11}a_{12}a_{21}+a_{22}a_{12}a_{21} & \\ (a_{11}^2+a_{22}^2)a_{21}+a_{11}a_{22}a_{21}+a_{12}a_{21}^2 & \\ (a_{11}^2+a_{22}^2)a_{12}+a_{11}a_{22}a_{12}+a_{12}^2a_{21} & \\ a_{22}^3+2a_{22}a_{12}a_{21}+a_{11}a_{12}a_{21} & \end{bmatrix}$$

If only u_1 is applied and only y_1 is observed,

$$G(\underline{p}) = \begin{bmatrix} c_1b_1 \\ c_1b_1a_{11} \\ c_1b_1(a_{11}^2+a_{12}a_{21}) \\ c_1b_1(a_{11}^3+2a_{11}a_{12}a_{21}+a_{22}a_{12}a_{21}) \end{bmatrix}$$

where the parameter vector $\underline{p} = [c_1b_1\ a_{11}\ a_{12}\ a_{21}\ a_{22}]^T$. It is clear that $\frac{\partial G}{\partial \underline{p}}$ cannot possibly be of rank 5, so let us consider the case where there is prior knowledge that $a_{12} = \alpha (\neq 0)$. Then $\underline{p} = [c_1b_1\ a_{11}\ a_{21}\ a_{22}]^T$,

$$\frac{\partial G}{\partial \underline{p}} = \begin{bmatrix} 1 & 0 \\ a_{11} & c_1b_1 \\ (a_{11}^2+a_{12}a_{21}) & c_1b_1 2a_{11} \\ (a_{11}^3+2a_{11}\alpha a_{21}+a_{22}\alpha a_{21}) & c_1b_1(3a_{11}^2+2\alpha a_{21}) \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ c_1b_1\alpha & 0 \\ c_1b_1(2a_{11}\alpha+a_{22}\alpha) & c_1b_1\alpha a_{21} \end{bmatrix}$$

and it is readily seen that rank $\frac{\partial G}{\partial \underline{p}} = 4$ provided c_1b_1, α and a_{21} are not zero.

Further examples are given by Grewal and Glover (1976) and by Carson, Cobelli and Finkelstein (1983, Section 7.5).

Remark 1: This method has the advantage that it is computationally convenient and amenable to computer implementation.

Remark 2: It is instructive at this point to consider the general time domain solution of the system described by equations (6), (7) and (8):

$$\underline{x}(t) = e^{At}\underline{x}(0^-) + \int_{0^-}^{t} e^{A(t-\tau)}B\underline{u}(\tau)d\tau \qquad (33)$$

and in particular the form for a zero initial state experiment. Then parameter sets $\underline{p}$ and $\tilde{\underline{p}}$ are indistinguishable if and only if

$$C(\underline{p})A^k(\underline{p})B(\underline{p}) = C(\tilde{\underline{p}})A^k(\tilde{\underline{p}})B(\tilde{\underline{p}}) \qquad (34)$$

for $k = 0,1, \ldots$. It is well known that we can restrict consideration to the first $2n$ equations of the form (34). The upper bound $k = (2n-1)$ is readily arrived at through the use of Tether's continuation lemma (Tether, 1970), which states that, given matrices M_1, M_2 and M_3 such that

$$\text{rank } M_1 = \text{rank}(M_1\ M_2) = \text{rank}\begin{pmatrix} M_1 \\ M_3 \end{pmatrix} \qquad (35A)$$

there exists, at most, one matrix M_4 for which

$$\text{rank}\begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix} = \text{rank } M_1 . \qquad (35B)$$

If we consider the matrices

$$M_1 = \begin{bmatrix} CB & CAB & \ldots. & CA^{n-1}B \\ CAB & CA^2B & \ldots. & CA^nB \\ \ldots.. & \ldots. & & \ldots. \\ CA^{n-1}B & CA^nB & \ldots. & CA^{2n-2}B \end{bmatrix}$$

$$M_2 = \begin{bmatrix} CA^nB \\ CA^{n+1}B \\ \ldots. \\ CA^{2n-1}B \end{bmatrix}$$

$$M_3 = [CA^nB \quad CA^{n+1}B \ \ldots. \ CA^{2n-1}B]$$

$$M_4 = [CA^{2n}B]$$

we know that the rank conditions (35A) and (35B) follow from the Cayley-Hamilton theorem and so from Tether's lemma, $CA^{2n}B$ is uniquely defined. Hence if equations (34) hold for $k = 0,1,\ldots(2n-1)$, then they also hold for any k and no new information is forthcoming from examining higher powers of A. For our two-state example, with only state 1 perturbed and observed, the (1 1) element of A^4 is $a_{11}^4 + 3a_{11}^2a_{12}a_{21} + 2a_{11}a_{22}a_{12}a_{21} + a_{22}^2a_{12}a_{21} + (a_{12}a_{21})^2$ so that a_{12} and a_{21} still do not appear individually.

Also, if we consider a unit impulse input $b_ju_j(t) = b_j\delta(t)$, with $\underline{x}(0^-) = 0$, then

$$\underline{x}(t) = e^{At}\, b_j\underline{\delta}_j$$

where $\underline{\delta}_j$ is a vector zero except for a 1 in row j. Then

$$y_i(t) = c_i\, e^{At}b_j$$
$$= c_ib_j + c_iAb_jt + c_iA^2b_j\frac{t^2}{2!} + \ldots$$
$$+ c_iA^kb_j\frac{t^k}{k!} + \ldots. \qquad (36)$$

This makes it clear that, for a unit impulse

input, exactly the same parameter combinations occur as successive derivatives of $y_i(t)$ at $t = 0^+$ in the Taylor series approach and as successive elements of $c_i A^k b_j$, $k = 0, 1, \ldots$.

Remark 3: Some identifiability approaches have examined properties of the transition matrix e^{At} , for example that described in Chapter 4 of Walter (1982). Generally these techniques prove rather complicated algebraically and are not easy to apply for $n > 2$. A more systematic time domain approach, using a modal expansion is described in Section 2.5.

2.5 Modal Matrix Approach

The methods presented in Sections 2.2 and 2.3 result in equations nonlinear in the unknown parameters, so it is difficult to see whether redundancy exists. There are general methods for solving these equations, some of which are presented by Lecourtier and Raksanyi in this book, but the problem is that sometimes the computations get so complex that no result is obtained. Another method, based on the modal matrix and its inverse, results in bilinear equations. Many (often all) of these equations are then reduced to linear equations by the incorporation of information about outputs and prior knowledge of the elements of A (Norton, 1980a; Norton, Brown and Godfrey, 1980).

The eigenvalues of A , λ_i and, up to a scaling factor, the eigenvectors $\underline{m}_i$ are defined by

$$A\underline{m}_i = \lambda_i \underline{m}_i , \qquad i = 1,2,\ldots n . \tag{37}$$

The eigenvectors are collected together as columns of the modal matrix M :-

$$AM = M\Lambda$$

where Λ is a diagonal matrix having λ_i as principal diagonal element i . The eigenvalues are assumed distinct; systems such as distillation columns which have simple structure but repeated eigenvalues are probably best treated by a specialised analysis exploiting their particular structures. The *modal matrix equations*

$$\underline{r}_i^T \underline{n}_j = 1 , \qquad i = j \tag{38A}$$

$$= 0 , \qquad i \neq j \tag{38B}$$

where $\underline{r}_i^T$ is row i of M and $\underline{n}_j$ is column j of $N \equiv M^{-1}$ are bilinear in the unknown elements of $\underline{r}_i^T$ and $\underline{n}_j$.

The system response is given by equation (33). If we consider an impulse input $b_j u_j(t) = b_j \delta(t)$ with $\underline{x}(0^-) = \underline{0}$,

$$\underline{x}(t) = e^{At} b_j \underline{\delta}_j = M e^{\Lambda t} N b_j \underline{\delta}_j \tag{39}$$

where $\underline{\delta}_j$ is a vector zero except for a 1 in row j . Hence

$$x_i(t) = \underline{r}_i^T e^{\Lambda t} \underline{n}_j b_j , \qquad t > 0$$

and

$$y_i(t) = c_i \underline{r}_i^T e^{\Lambda t} \underline{n}_j b_j , \qquad t > 0 \tag{40}$$

for square, diagonal B and C matrices, as in equations (12) and (13). These equations are known as the *input-output equations* and similar expressions can be obtained for other forms of input. For example, if $u_j(t)$ had been a unit step (again with $\underline{x}(0^-) = \underline{0}$) , then

$$y_i(t) = c_i \underline{r}_i^T \Lambda^{-1} (e^{\Lambda t} - I) \underline{n}_j b_j , \qquad t > 0 . \tag{41}$$

Prior knowledge of any elements of A can be incorporated through the *prior knowledge equations*

$$a_{ij} = \underline{r}_i^T \Lambda \underline{n}_j , \quad i = 1,2\ldots n \; ; \; j = 1,2,\ldots n . \tag{42}$$

As noted above, the objective of the method is to reduce the bilinear matrix equations to linear equations by using the input-output and prior knowledge equations and so to solve for $\underline{r}_i^T$ and $\underline{n}_j$. Let us now consider the four experiments detailed in Section 2.2. For all the two-state examples, the modal matrix equations are:

$$\underline{r}_1^T \underline{n}_1 = 1 \tag{43A}$$

$$\underline{r}_1^T \underline{n}_2 = 0 \tag{43B}$$

$$\underline{r}_2^T \underline{n}_1 = 0 \tag{43C}$$

$$\underline{r}_2^T \underline{n}_2 = 1 \tag{43D}$$

Experiment 1: SISO case, $u_1 = \delta(t)$, $u_2 = 0$, $\underline{y}_2 = \underline{0}$. Since both modes appear in x_1 and the scaling of eigenvectors is arbitrary, $\underline{r}_1^T$ can be taken as $[1 \quad 1]$. The input-output equation is

$$y_1(t) = c_1 \underline{r}_1^T e^{\Lambda t} \underline{n}_1 b_1 , \qquad t > 0 . \tag{44}$$

This gives the proportions of $\underline{n}_1$ from the ratio of the normal modes and the scaling is then given by equation (43A). In the absence of prior knowledge, the only other equations in the remaining unknowns $\underline{r}_2$ and $\underline{n}_2$ are (43B) and (43C) which are linear and (43D) which is bilinear. The product $c_1 b_1$ is obtained uniquely from equations (43A) and (44). Note that equation (43A) is redundant if there is prior knowledge of $c_1 b_1$.

Prior knowledge of a_{12} $(\neq 0)$ would give

$$\underline{r}_1^T \Lambda \underline{n}_2 = a_{12} . \tag{45}$$

Assuming distinct eigenvalues, $\Lambda \underline{n}_2$ is independent of $\underline{n}_2$ so with $\underline{r}_1^T$ known, equations (43B) and (45) give $\underline{n}_2$; $\underline{r}_2$ is then obtained from equations (43C) and (43D).

In some cases, prior knowledge can result in two equations bilinear in the unknowns, but these can be rearranged to give a singularity equation in one of the unknown rows or columns (Norton, 1980a; Norton, Brown and Godfrey, 1980); an example will be given in Experiment 4.

Experiment 2: One Input, Two Outputs, $u_1(t) = \delta(t)$, $u_2(t) = 0$. The available equations are (43A) to (43D) and (44), from which $\underline{r}_1$ and $\underline{n}_1$ are known, and

$$y_2(t) = c_2\underline{r}_2^T e^{\Lambda t}\underline{n}_1 b_1 , \quad t > 0 \qquad (46)$$

which, with equation (43C), gives $\underline{r}_2$ only if c_2b_1 is known; $\underline{n}_2$ is then obtained from equations (43B) and (43D).

Experiment 3: Two Inputs, One Output. $y_2 = 0$.
If the inputs are applied separately, then for impulsive inputs, the available equations are (43A) to (43D), (44) and there is a second input-output equation:

$$y_1(t) = c_1\underline{r}_1^T e^{\Lambda t}\underline{n}_2 b_2 , \quad t > 0 . \qquad (47)$$

By analogy with Experiment 2, the elements of A are uniquely identifiable if c_1b_2 is known.

Now consider the two inputs applied simultaneously. The observation, for impulsive forcing of both states, is

$$y_1(t) = c_1\underline{r}_1^T e^{\Lambda t}(\underline{n}_1 b_1 + \underline{n}_2 b_2) , \quad t > 0 \qquad (48)$$

which does not give $\underline{n}_1$ or $\underline{n}_2$ individually even if c_1b_2 were known. The system is unidentifiable from this experiment and from any two inputs of the same shape. If $u_1(t)$ is a unit impulse and $u_2(t)$ is a unit step, the input-output equation becomes

$$y_1(t) = c_1[\underline{r}_1^T e^{\Lambda t}(\underline{n}_1 b_1 + \Lambda^{-1}\underline{n}_2 b_2) - \underline{r}_1^T\Lambda^{-1}\underline{n}_2 b_2] , \quad t > 0 . \qquad (49)$$

Once $\underline{r}_1^T$ is chosen, the normal mode amplitudes give

$$c_1\underline{n}_1 b_1 + c_1\Lambda^{-1}\underline{n}_2 b_2 = \underline{f} \quad \text{(say)}.$$

Then, since

$$c_1\underline{r}_1^T\Lambda^{-1}\underline{n}_2 b_2 = \underline{r}_1^T\underline{f} - c_1\underline{r}_1^T\underline{n}_1 b_1 , \qquad (50)$$

$\underline{n}_2$ can be found from equations (43B) and (50) provided c_1b_2 is known. Then $\underline{n}_1$ can be found from $\underline{f}$ and $\underline{r}_2$ from the modal matrix equations. The elements of A are uniquely identifiable from this experiment.

Experiment 4: SISO experiment with $u_1 \neq 0$, $u_2 = 0$, $y_1 = 0$, prior knowledge that $a_{12} = 0$.
For a unit impulse perturbation, the input-output equation is

$$y_2(t) = c_2\underline{r}_2^T e^{\Lambda t}\underline{n}_1 b_1 , \quad t > 0 . \qquad (51)$$

Since both modes appear in x_2, $\underline{r}_2^T$ can be taken as $[1 \quad 1]$ and equations (43C) and (51) then give $\underline{n}_1$ provided c_2b_1 is known. The prior information equation

$$a_{12} = \underline{r}_1^T\Lambda\underline{n}_2 = 0 \qquad (52)$$

and the modal matrix equation (43B) are both bilinear in the unknowns $\underline{r}_1$ and $\underline{n}_2$ but may be combined to give a singularity equation

$$\det\begin{bmatrix}\underline{r}_1^T \\ \underline{r}_1^T\Lambda\end{bmatrix} = 0 \qquad (53)$$

which gives an expression for the non-zero element of $\underline{r}_1$, r_{11} being zero since one of the modes does not appear in y_1. This expression has two solutions, since the ordering of the eigenvalues (a_{11} and a_{22}) in Λ is arbitrary. The model is thus locally identifiable, with two solutions, from this experiment.

Other examples are given in Norton (1980a), Norton, Brown and Godfrey (1980) and Godfrey et al (1982).

Remark 1: The modal matrix approach has the merit that when all the equations can be reduced from bilinear to linear equations, global identifiability is easily checked. In contrast to the Laplace transform or Taylor series approaches, the effects of adding prior information of elements of A, or changing or adding to the inputs or observed states, are seen with a minimum of reworking, which makes the approach particularly attractive if several different cases are to be examined. One drawback of the method is that many more equations have to be examined, with prior knowledge adding to the number of equations, rather than simplifying existing equations. In some cases, local identifiability is not particularly easy to spot. Also, since the method employs the modal matrix rather than the A matrix, identifiable single parameters or combinations of parameters in a model in which not all parameters are identifiable are obtained less readily than with other approaches.

Remark 2: The modal matrix approach has also been used by Delforge (1980, 1981) to count the number of independent equations once redundancies have been determined. An upper limit on the number of solutions from their degree in the unknown elements of M or N is then calculated. As pointed out by Norton (1982), the limit is sometimes not tight enough to be useful and it is possible to overlook nonuniqueness due to inability to choose the ordering of the observed eigenvalues to be consistent with the pattern of zero elements in M and N (Norton, 1980b).

2.6 Exhaustive Modelling Approach

The objective of this method is to generate the set of all models which are output indistinguishable and compatible with the assumptions on the model structure. Starting from a model with system matrix A, input matrix B and observation matrix C, all equivalent systems must have corresponding matrices A', B', C' related to A, B, C by a similarity transformation:

$$A' = TAT^{-1} \qquad (54)$$

$$B' = TB \qquad (55)$$

$$C' = CT^{-1} . \qquad (56)$$

(Strictly speaking, the model must be structurally controllable and structurally observable for this method to be applicable.) The approach is to apply the known constraints on A, B and C to determine the unknown elements of T (Walter and Lecourtier, 1981; Walter, 1982, Chapter 5). If

T is unique, the system is uniquely identifiable; if there is a finite set T, the system is locally identifiable; otherwise, it is unidentifiable. The method was anticipated, but not fully developed, by Berman and Schoenfeld (1956) and Rubinow and Winzer (1971).

If the constraints on A, B and C result in any zero elements in T, then the corresponding elements in T^{-1} are also zero. This can be seen by simply rearranging equations (54), (55) and (56) to give

$$A = T^{-1}A'T$$

$$B = T^{-1}B'$$

$$C = C'T \ .$$

The approach is now illustrated using the Experiments on the two-state model described earlier. For each experiment, the elements of T and T^{-1} are denoted by:

$$T = \begin{bmatrix} t_1 & t_2 \\ t_3 & t_4 \end{bmatrix} \quad \text{and} \quad T^{-1} = \begin{bmatrix} \hat{t}_1 & \hat{t}_2 \\ \hat{t}_3 & \hat{t}_4 \end{bmatrix} \ .$$

Experiment 1: SISO Case, $u_2 = 0$, $y_2 = 0$.

Since $B = [b_1 \quad 0]^T$,

$$B' = \begin{bmatrix} t_1 & t_2 \\ t_3 & t_4 \end{bmatrix} \begin{bmatrix} b_1 \\ 0 \end{bmatrix} = \begin{bmatrix} t_1 b_1 \\ t_3 b_1 \end{bmatrix}$$

and because any zero elements in B must also be zero in B',

$t_3 = 0$ and, from the above, $\hat{t}_3 = 0$.

Similarly since $C = [c_1 \quad 0]$,

$$C' = [c_1 \quad 0] \begin{bmatrix} \hat{t}_1 & \hat{t}_2 \\ \hat{t}_3 & \hat{t}_4 \end{bmatrix} = [c_1\hat{t}_1 \quad c_1\hat{t}_2]$$

so $\hat{t}_2 = 0$ and, from the above, $t_2 = 0$. From $TT^{-1} = I$,

$$\hat{t}_1 = 1/t_1$$

and

$$\hat{t}_4 = 1/t_4 \ .$$

This gives $B' = [t_1 b_1 \quad 0]^T$ (57)

$$C' = [c_1/t_1 \quad 0] \qquad (58)$$

and $$A' = \begin{bmatrix} t_1 & 0 \\ 0 & t_4 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1/t_1 & 0 \\ 0 & 1/t_4 \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11} & \frac{t_1}{t_4} a_{12} \\ \frac{t_4}{t_1} a_{21} & a_{22} \end{bmatrix} \ . \qquad (59)$$

From equation (59), it is seen that a_{11} and a_{22} are uniquely identifiable as is the product $a_{12}a_{21}$. We need to examine both equations (57) and (58) to see that the product $c_1 b_1$ is also uniquely identifiable.

If there is prior knowledge of a_{12}, then a'_{12} must equal a_{12} so that $t_1/t_4 = 1$. Then $A' = A$ and the parameters of A are uniquely identifiable.

Experiment 2: Single-Input, Dual-Output, $u_2 = 0$.

C is now $\begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}$

and since $$C' = \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix} \begin{bmatrix} \hat{t}_1 & \hat{t}_2 \\ \hat{t}_3 & \hat{t}_4 \end{bmatrix} = \begin{bmatrix} c_1\hat{t}_1 & c_1\hat{t}_2 \\ c_2\hat{t}_3 & c_2\hat{t}_4 \end{bmatrix}$$

from which $\hat{t}_2 = \hat{t}_3 = 0$ and hence $t_2 = t_3 = 0$. No new information is obtained because we already knew that $t_3 = 0 = \hat{t}_3$ from $B' = TB$. The approach does not readily indicate identifiable combinations of parameters and these have to be sought from relationships such as $C'A'B' = CT^{-1}TAT^{-1}TB = CAB$. In this case,

$$C'A'B' = \begin{bmatrix} c_1/t_1 & 0 \\ 0 & c_2/t_4 \end{bmatrix} \begin{bmatrix} a_{11} & \frac{t_1}{t_4} a_{12} \\ \frac{t_4}{t_1} a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} t_1 b_1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} c_1 b_1 a_{11} \\ c_2 b_1 a_{21} \end{bmatrix}$$

so that $c_2 b_1 a_{21}$ is identified uniquely. Thus if there is prior knowledge of $c_2 b_1$, a_{21} is known and since a'_{21} must then equal a_{21}, $t_4/t_1 = 1$ and the parameters of A are identified uniquely.

Experiment 3: Dual-Input, Single-Output, $y_2 = 0$.

For inputs well separated in time, then by analogy with Experiment 2, the elements of A can be identified uniquely if there is prior knowledge of $c_1 b_2$.

For simultaneous inputs, $B = [b_1 \quad b_2]^T$ if the inputs are of the same shape, whereas $B = \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix}$ if the inputs are different. In either case, $C = [c_1 \quad 0]$ so $\hat{t}_2 = 0$ and hence, $t_2 = 0$. For $B = [b_1 \quad b_2]^T$,

$$B' = \begin{bmatrix} t_1 & 0 \\ t_3 & t_4 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} b_1 t_1 \\ b_1 t_3 + b_2 t_4 \end{bmatrix}$$

which provides no new information.

By contrast, when the two inputs are different,