IN-37

IDENTIFICATION OF OBSERVER/KALMAN FILTER MARKOV PARAMETERS: THEORY AND EXPERIMENTS

Jer-Nan Juang, Minh Phan, Lucas G. Horta, and Richard W. Longman

June 1991

NASA

IDENTIFICATION OF OBSERVER/KALMAN FILTER MARKOV PARAMETERS: THEORY AND EXPERIMENTS

Jer-Nan Juang,[*] Minh Phan,[**] Lucas G. Horta,[+] and Richard W. Longman[++]
NASA Langley Research Center, Hampton, VA 23665

Abstract

This paper discusses an algorithm to compute the Markov parameters of an observer or Kalman filter from experimental input and output data. The Markov parameters can then be used for identification of a state space representation, with associated Kalman gain or observer gain, for the purpose of controller design. The algorithm is a non-recursive matrix version of two recursive algorithms developed in previous works for different purposes, and the relationship between these other algorithms is developed. The new matrix formulation here gives insight into the existence and uniqueness of solutions of certain equations, and gives bounds on the proper choice of observer order. It is shown that if one uses data containing noise, and seeks the fastest possible deterministic observer, the deadbeat observer, one instead obtains the Kalman filter -- which is the fastest possible observer in the stochastic environment. The results of the paper are demonstrated in numerical studies and in experiments on a ten-bay truss structure.

Introduction

Many future spacecraft such as the space station will be large and flexible and will require control of the vibrational motion induced by internal and external disturbances for fine pointing and shape control. One can classify controllers for flexible structures into two types, namely, model-independent controllers and model-dependent controllers. The model-independent controller[1] is attractive because it provides stability without precise knowledge of the system. However, it generally produces low authority control which may not meet the performance requirements. On the other hand, model-dependent controllers require an accurate model to achieve high

---

[*] Principal Scientist, Spacecraft Dynamics Branch, Fellow AIAA.
[**]National Research Council Research Associate, Spacecraft Dynamics Branch.
[+]Aerospace Engineer, Spacecraft Dynamics Branch.
[++]National Research Council Senior Research Associate, Spacecraft Dynamics Branch. Also Professor of Mechanical Engineering, Columbia University, New York, NY. Fellow AIAA.

performance, but inaccuracies in the model may result in instability of the controlled system.[2] Current results indicate that an accurate model is necessary to design controllers with the needed performance level.

In the past decade, many system identification techniques were developed and/or applied to identify a state space model for modal parameter identification of large flexible space structures. The modal parameters include frequencies, damping, and mode shapes. The identified state space model is also used in controller design. Many satisfactory results were reported in the literature.[3,4] Most techniques are based on sampled pulse or impulse system response histories which are known as Markov parameters. The usual practice uses the Fast Fourier Transforms (FFT) of the inputs and measured outputs to compute the sampled pulse response histories. The discrete nature of the FFT causes one to obtain pulse response rather than impulse response, and a somewhat rich input is required to prevent numerical ill-conditioning in the computation. Another approach is to solve directly in the time domain for the Markov parameters from the input and output data. The drawbacks of this method include the need to invert an input matrix which necessarily becomes particularly large for lightly damped systems.[5]

Recently, an approach was developed[6-8] to address the problem of inverting a large-dimensional input matrix. Reference 8 gives a more detailed presentation of the developments in Refs. 6 and 7 including additional examples. Rather than identifying the system Markov parameters which may exhibit very slow decay, it uses an asymptotically stable observer to form a stable state space discrete model for the system to be identified. The primary purpose of introducing an observer in Ref. 8 is as an artifice to compress the data and improve system identification results in practice. The system identification engineer can assign any poles desired, and hence specify the decay rate of the observer Markov parameters to be determined from the data and simultaneously the number of parameters needed before they have decayed to a negligible level. The desired poles can be real, complex or deadbeat. The deadbeat means that all the poles are zero in the complex plane for a discrete model.

The treatment in Ref. 8 is purely deterministic. When stochastic models are considered, it would be desirable to identify not only the system matrices of a realization, but also the noise or uncertainty characteristics of the model directly from the experimental data. This presumes that the same sensors and actuators used in the identification tests will also be used in the control system which is to be designed from the system identification results. There are basically two ways to characterize system uncertainties including plant and measurement noises. One way is to describe the input and output uncertainties directly in terms of their covariances. Another way is to specify

2

the Kalman filter equation with its steady state Kalman gain which is function of the input and output covariances. Recently, a recursive identification method was presented in Ref. 9 to identify Markov parameters for identification of not only the system matrices, but also the Kalman filter gain. Note that the work in Ref. 9 was motivated by the unsolved problem in Ref. 10 in which single-mode projection filters were developed for modal parameter identification. There exists many unsolved issues such as the relationship between the order of the Kalman filter and the order of the system using the approach of Ref. 9. Furthermore, the computation time and the data length are too long to become attractive in practice. Examination of the mathematics involved in Ref. 8 for accelerated identification using observers, and that in Ref. 9 for direct identification of the Kalman filter Markov parameters, shows striking parallels which are investigated here.

One objective of this paper is to present an algorithm to directly compute the Markov parameters of a steady state Kalman filter from experimental data. From these parameters, one can use various methods to obtain the Kalman filter state space realization. The approach used in this paper is to reformulate in matrix form the equations presented in Ref. 8 for deadbeat observers, and in Ref. 9 for Kalman filters. This matrix form gives added insight into the uniqueness of the transformation from observer or filter Markov parameters to the system Markov parameters, and also allows the development of upper and lower bounds on the choice of observer order. Also, the recursive least squares method of solution of the equations in Refs. 8 and 9 is replaced by a non-recursive least squares solution. This results in an improved rate of convergence for the Kalman filter identification process by comparison to Ref. 9.

Underlying this work is a second objective, to establish the relationship between the observer identification equations in Ref. 8 and the Kalman filter identification equations in Ref. 9. When the observer poles in Ref. 8 are all placed at the origin in the z-plane in order to obtain a deadbeat observer of a sufficiently high order, and then data containing both plant and measurement noise is used to develop the desired Markov parameters, the result is the Kalman filter Markov parameters. Stated in different words, if one uses data containing noise, and seeks the Markov parameters for the fastest possible deterministic discrete time observer, one instead obtains the Markov parameters of the slower Kalman filter -- which is the fastest possible observer in the stochastic environment.

This paper starts by writing the relationship between the input and output histories in terms of system Markov parameters without any observer. An observer is then introduced into the input and output matrix relation, which is solved by a non-recursive least squares approach to compute the observer Markov parameters. Formulations are derived to compute the system Markov parameters and the observer gain from the observer Markov parameters. The relationship between

3

the identified deadbeat observer and a Kalman filter is then established through use of the ergodic property of stationary random processes. The optimal nature of the identified observer is also discussed. Numerical and experimental results are given to illustrate the validity of the algorithm presented in this paper. The experimental results are obtained from a 10-bay truss structure having two accelerometers and two thrusters.

Basic Formulation

Consider a discrete multivariable linear system described by

$$
\begin{aligned}
x(i+1) &= Ax(i) + Bu(i) \\
y(i) &= Cx(i) + Du(i)
\end{aligned}
\tag{1}
$$

where $x(i) \in R^n$, $y(i) \in R^q$, $u(i) \in R^m$. Assuming zero initial conditions, $x(0) = 0$, the set of this equations for a sequence of i can be written as

$$
\underset{q \times m\ell}{\overset{q \times \ell}{y}} = \mathbf{Y} \overset{m\ell \times \ell}{\mathbf{U}}
\tag{2}
$$

where

$$
y = \begin{bmatrix} y(0) & y(1) & y(2) & \cdots & y(\ell-1) \end{bmatrix}
$$
$$
\mathbf{Y} = \begin{bmatrix} D & CB & CAB & \cdots & CA^{\ell-2}B \end{bmatrix}
$$

and

$$
\mathbf{U} = \begin{bmatrix}
u(0) & u(1) & u(2) & \cdots & u(\ell-1) \\
 & u(0) & u(1) & \cdots & u(\ell-2) \\
 & & u(0) & \cdots & u(\ell-3) \\
 & & & \ddots & \vdots \\
 & & & & u(0)
\end{bmatrix}
$$

Equation (2) is a matrix representation of the relationship between input and output histories. The matrix y is a $q \times \ell$ output data matrix where q is the number of outputs and ℓ is the number of data samples. The matrix $\mathbf{Y}$, of dimension $q \times m\ell$ with m the number of inputs, contains all the Markov parameters D, CB, CAB, $\cdots$, $CA^{\ell-2}B$ to be determined. The matrix $\mathbf{U}$ is an $m\ell \times \ell$ upper block triangular input matrix. It is square in the case of a single input system, and otherwise has more rows than columns.

4

Inspection of Eq. (2) indicates that there are $q \times m\ell$ unknowns in the Markov parameter matrix but only $q \times \ell$ equations. For the case where $m > 1$, the solution for **Y** is not unique. However, it is known that, for a finite-dimensional linear system, **Y** must be unique. The matrix **Y** can only be uniquely determined from this set of equation for $m = 1$. Even in this case, if the input has zero initial value, i.e. $u(0) = 0$, or the input signals are not rich enough such as the case with sinusoidal input signals, the matrix **U** becomes ill-conditioned and thus the matrix $Y = yU^{-1}$ cannot be accurately computed.

Consider the case where A is asymptotically stable so that for some sufficiently large p, $A^i \approx 0$ for all time steps $i \geq p$. Equation (2) can then be approximated by

$$\underset{q \times mp}{\overset{q \times \ell}{y} \approx Y \overset{mp \times \ell}{U}} \tag{3}$$

where

$$y = \begin{bmatrix} y(0) & y(1) & y(2) & \cdots & y(p) & \cdots & y(\ell-1) \end{bmatrix}$$

$$Y = \begin{bmatrix} D & CB & CAB & \cdots & CA^{p-1}B \end{bmatrix}$$

$$U = \begin{bmatrix} u(0) & u(1) & u(2) & \cdots & u(p) & \cdots & u(\ell-1) \\ & u(0) & u(1) & \cdots & u(p-1) & \cdots & u(\ell-2) \\ & & u(0) & \cdots & u(p-2) & \cdots & u(\ell-3) \\ & & & \ddots & \vdots & \cdots & \vdots \\ & & & & u(0) & \cdots & u(\ell-p-1) \end{bmatrix}$$

Note that the script U ($mp \times \ell$) and Y ($q \times mp$) refer to truncated versions of the Roman bold **U** and **Y** in Eq. (2). Choose the data length ℓ greater than mp where again m is the number of inputs and p is an integer such that $CA^iB \approx 0$ for $i \geq p$. Equation (3) indicates that there are more equations ($q \times \ell$) than unknowns ($q \times mp$) because $\ell \geq mp$. We conclude that if the data has a realization in the form of Eq. (1), then the first p Markov parameters approximately satisfy $Y = yU^+$ where U^+ is the pseudo-inverse of the matrix U, and the approximation error decreases as p increases.

Unfortunately, for lightly damped space structures, the integer p and thus the ℓ required to make the approximation in Eq. (3) valid becomes impractically large in the sense that the size of the matrix U is too large to solve for its pseudo-inverse U^+ numerically. The question arises, is there any way to artificially increase the damping of the system in order to allow solution of Eq. (3) for

5

the Markov parameters? A control engineer will immediately suggest that a feedback loop can be added to make the system as stable as desired. The same effect can be achieved by considering the following algebraic manipulation as presented in Ref. 8.

Add and subtract the term $My(i)$ to the right hand side of the state equation in Eq. (1) to yield

$$x(i+1) = Ax(i) + Bu(i) + My(i) - My(i)$$
$$= (A + MC)x(i) + (B + MD)u(i) - My(i)$$

or

$$x(i+1) = \overline{A}x(i) + \overline{B}v(i)$$
$$y(i) = Cx(i) + Du(i) \tag{4}$$

where

$$\overline{A} = A + MC$$
$$\overline{B} = [B + MD, -M]$$
$$v(i) = \begin{bmatrix} u(i) \\ y(i) \end{bmatrix} \tag{5}$$

and M is an $n \times q$ arbitrary matrix chosen to make the matrix $\overline{A}$ as stable as desired. Although Eq. (4) is mathematically identical to Eq. (1), it is expressed using different system matrices and has a different input. In fact, Eq. (4) is an *observer* equation if the state *x(i)* is considered as an observer state (see Ref. 7 or Ref. 8). Therefore, the Markov parameters of the system in Eq. (4) will be referred to as the *observer Markov parameters*. The input-output description in matrix form for Eq. (4) becomes

$$\underset{q \times [q(\ell-1)+m]}{\overset{q \times \ell}{y}} = \overline{Y} \overset{[q(\ell-1)+m] \times \ell}{V} \tag{6}$$

where

$$y = \begin{bmatrix} y(0) & y(1) & y(2) & \cdots & y(p) & \cdots & y(\ell-1) \end{bmatrix}$$
$$\overline{Y} = \begin{bmatrix} D & C\overline{B} & C\overline{A}\overline{B} & \cdots & C\overline{A}^{p-1}\overline{B} & \cdots & C\overline{A}^{\ell-2}\overline{B} \end{bmatrix}$$

6

$$
V = \begin{bmatrix}
u(0) & u(1) & u(2) & \cdots & u(p) & \cdots & u(\ell-1) \\
 & v(0) & v(1) & \cdots & v(p-1) & \cdots & v(\ell-2) \\
 & & v(0) & \cdots & v(p-2) & \cdots & v(\ell-3) \\
 & & & \ddots & \vdots & \cdots & \vdots \\
 & & & & v(0) & \cdots & v(\ell-p-1) \\
 & & & & & \ddots & \vdots \\
 & & & & & & v(0)
\end{bmatrix}.
$$

Equation (6) is obtained from Eq. (2) by replacing A by $\overline{A}$, B by $\overline{B}$ and u by v except for the first row partition. Because the $n \times q$ matrix M can be arbitrarily chosen, the eigenvalues of $\overline{A}$ may be arbitrarily assigned for an observable system. Reference 8 considers the identification of observer Markov parameters for any chosen observer pole locations for $\overline{A} = A + MC$. The mathematical development here can be interpreted from the point of view of Ref. 8 as attempting to place all the eigenvalues of $\overline{A}$ at the origin, i.e. a deadbeat observer. This provides that $C\overline{A}^{i}\overline{B} = 0$ for $i \geq p$. When using real data including noise, the eigenvalues of $\overline{A}$ are in fact placed such that $C\overline{A}^{i}\overline{B} \approx 0$ for $i \geq p$ where p is a sufficiently large integer. Alternatively, if $\overline{A}$ represents the state matrix of the Kalman filter including the steady state Kalman filter gain, the same property is satisfied as used in Ref. 7, which will be discussed in a latter section.

When $C\overline{A}^{i}\overline{B} \approx 0$ for $i \geq p$, one can solve for the observer Markov parameters from real data, using the same approach as in Eq. (3):

$$
\overset{q \times \ell}{y} = \overset{[(m+q)p+m] \times \ell}{\underset{q \times [(m+q)p+m]}{\overline{Y}} \quad V} \tag{7}
$$

where

$$
y = \begin{bmatrix} y(0) & y(1) & y(2) & \cdots & y(p) & \cdots & y(\ell-1) \end{bmatrix}
$$

$$
\overline{Y} = \begin{bmatrix} D & C\overline{B} & C\overline{A}\,\overline{B} & \cdots & C\overline{A}^{p-1}\overline{B} \end{bmatrix}
$$

$$
V = \begin{bmatrix}
u(0) & u(1) & u(2) & \cdots & u(p) & \cdots & u(\ell-1) \\
 & v(0) & v(1) & \cdots & v(p-1) & \cdots & v(\ell-2) \\
 & & v(0) & \cdots & v(p-2) & \cdots & v(\ell-3) \\
 & & & \ddots & \vdots & \cdots & \vdots \\
 & & & & v(0) & \cdots & v(\ell-p-1)
\end{bmatrix}
$$

Note that the script V and $\bar{Y}$ refer to truncated versions of the Roman bold **V** and **Ȳ** in Eq. (6). Similar to Eq. (3), if the data has a realization in the form of Eq. (1) or its equivalent, Eq. (4), then the first p Markov parameters approximately satisfy $\bar{Y} = yV^+$ where V^+ is the pseudo-inverse of the matrix V, and the approximation error decreases as p increases. Note that the observer Markov parameters thus identified may not necessarily appear to be asymptotically decaying during the first p-1 steps, although it produces $C\bar{A}^i\bar{B} = 0$ for $i \geq p$ for noise free data. Reference 8 allows one to place the observer poles to produce more typical asymptotic decay of the observer Markov parameters. To solve for $\bar{Y}$ uniquely, all the rows of V must be linearly independent. Furthermore, to minimize any numerical error due to the computation of the pseudo-inverse, the rows of V should be chosen as independent as possible. As a result, the maximum p is the number that maximizes the number, $(m+q)p+m$, of independent rows of V. The maximum p means the upper bound of the order of the deadbeat observer. The lower bound of the order of the observer will be addressed in the next section.

There are many ways of producing the least squares solution to equations such as Eq. (7) for $\bar{Y}$. Reference 5 presents three different approaches to solving equations similar to Eq. (7), including a bootstrapping procedure, the singular value decomposition and a recursive algorithm. However, the recursive least squares algorithm presented in Ref. 8 includes substitution of the desired eigenvalues into Eq. (7) to minimize the unknown parameters in $\bar{Y}$. This is probably a very efficient computational procedure. However, it is not obtaining the least squares solution of Eq. (7) but rather a somewhat modified problem which can be interpreted as a weighted least squares solution. Disadvantages of the recursive formula include a problem-dependent choice of error sensitivity which requires experiences in least squares methods.

All the above equations assume zero initial conditions, $x(0) = 0$. For nonzero initial conditions, a somewhat different formula should be used. Rewrite Eq. (4) in another matrix form as

$$\bar{y} = C\bar{A}^p x + \bar{Y}V \tag{8}$$

where

$$\bar{y} = \begin{bmatrix} y(p+1) & y(p+2) & \cdots & y(\ell-1) \end{bmatrix} \quad x = \begin{bmatrix} x(0) & x(1) & \cdots & x(\ell-p-2) \end{bmatrix}$$

$$\bar{Y} = \begin{bmatrix} D & C\bar{B} & C\bar{A}\bar{B} & \cdots & C\bar{A}^{p-1}\bar{B} \end{bmatrix} \quad \bar{V} = \begin{bmatrix} u(p+1) & u(p+2) & \cdots & u(\ell-1) \\ v(p) & v(p+1) & \cdots & v(\ell-2) \\ v(p-1) & v(p) & \cdots & v(\ell-3) \\ \vdots & \vdots & \vdots & \ddots \\ v(0) & v(1) & \cdots & v(\ell-p-2) \end{bmatrix}$$

For the case where $\bar{A}^p$ is sufficiently small and all the states in x are bounded, Eq. (8) can be approximated by neglecting the first term on the right hand side,

$$\bar{y} = \bar{Y}\bar{V} \tag{9}$$

which has the following least squares solution

$$\bar{Y} = \bar{y}\bar{V}^T [\bar{V}\bar{V}^T]^{-1} \tag{10}$$

provided that $[\bar{V}\bar{V}^T]^{-1}$ exists. Equation (9) is identical to Eq. (7) except that the y in Eq. (7) is replaced by $\bar{y}$ and V by $\bar{V}$. The matrices $\bar{y}$ and $\bar{V}$ are subsets of y and V respectively produced by deleting the first p columns. For nonzero unknown initial conditions, Eq. (9) must be used in order to eliminate the effect of initial conditions, because the initial conditions become negligible when they are multiplied by $\bar{A}^p$. In other words, the initial conditions have negligible influence on the measured data after p time steps. When there are both system and measurement noise present, the elimination of initial condition dependence makes the system response become stationary, a fact which is used later to obtain the steady state Kalman filter gain.

Computation of Actual System Markov Parameters and Observer Gain

To recover the system Markov parameters in Y from the observer Markov parameters in $\bar{Y}$, partition $\bar{Y}$ such that

$$\bar{Y} = \begin{bmatrix} \bar{Y}_{-1} & \bar{Y}_0 & \bar{Y}_1 & \cdots & \bar{Y}_{p-1} \end{bmatrix} \tag{11}$$

where

9

$$\bar{Y}_k = C\bar{A}^k\bar{B}$$
$$= \left[C(A+MC)^k(B+MD), \; -C(A+MC)^k M \right]$$
$$\equiv \left[\bar{Y}_k^{(1)}, \; \bar{Y}_k^{(2)} \right]; \quad k = 0,1,2,\dots$$

$$\bar{Y}_{-1} = D$$

Note that the Markov parameter $\bar{Y}_{-1}$ has a smaller dimension than the remaining Markov parameters. From the second equation in Eq. (11), the Markov parameter CB of the system is simply

$$Y_0 = CB = C(B+MD) - (CM)D$$
$$= \bar{Y}_0^{(1)} + \bar{Y}_0^{(2)}D \tag{12}$$

To obtain the Markov parameter CAB, first consider the product $\bar{Y}_1^{(1)}$

$$\bar{Y}_1^{(1)} = C(A+MC)(B+MD)$$
$$= CAB + CMCB + C(A+MC)MD$$

Hence,

$$Y_1 = CAB$$
$$= \bar{Y}_1^{(1)} + \bar{Y}_0^{(2)}Y_0 + \bar{Y}_1^{(2)}D \tag{13}$$

Similarly, to obtain the Markov parameter CA^2B, consider the product $\bar{Y}_2^{(1)}$

$$\bar{Y}_2^{(1)} = C(A+MC)^2(B+MD)$$
$$= C(A^2 + MCA + AMC + MCMC)(B+MD)$$
$$= CA^2B + CMCAB + C(A+MC)MCB + C(A+MC)^2MD$$

Therefore,

$$Y_2 = CA^2B$$
$$= \bar{Y}_2^{(1)} - CMCAB - C(A+MC)MCB - C(A+MC)^2MD$$
$$= \bar{Y}_2^{(1)} + \bar{Y}_0^{(2)}Y_1 + \bar{Y}_1^{(2)}Y_0 + \bar{Y}_2^{(2)}D \tag{14}$$

As established in Ref. 8, the general relationship between the actual system Markov parameters and the observer Markov parameters is

10

$$Y_k = \overline{Y}_k^{(1)} + \sum_{i=0}^{k-1} \overline{Y}_i^{(2)} Y_{k-i-1} + \overline{Y}_k^{(2)} D \tag{15}$$

Knowledge of the actual system Markov parameters allows one to obtain a state-space realization of the system of interest. Modal parameters including natural frequencies, damping ratios, and mode shapes can then be found. Note that there are only $p+1$ observer Markov parameters computed as a least squares solution from Eq. (7). By the choice of p, $\overline{Y}_k^{(1)}$ and $\overline{Y}_k^{(2)}$ are considered to be zero for $k > p$.

The relationship between observer Markov parameters and system Markov parameters can be further developed as follows. Let the matrices H, $\overline{Y}^{(2)}$ and $\underline{Y}$ be defined as

$$\overline{Y}^{(2)} = \begin{bmatrix} -\overline{Y}_{p-1}^{(2)} & -\overline{Y}_{p-2}^{(2)} & -\overline{Y}_{p-3}^{(2)} & \cdots & -\overline{Y}_0^{(2)} \end{bmatrix}; \quad H = \begin{bmatrix} Y_1 & Y_2 & Y_3 & \cdots & Y_N \\ Y_2 & Y_3 & Y_4 & \cdots & Y_{N+1} \\ Y_3 & Y_4 & Y_5 & \cdots & Y_{N+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Y_p & Y_{p+11} & Y_{p+2} & \cdots & Y_{N+p-1} \end{bmatrix} \tag{16}$$

and

$$\underline{Y} = \begin{bmatrix} Y_{p+1} & Y_{p+2} & Y_{p+3} & \cdots & Y_{p+N} \end{bmatrix}$$

where N is a sufficiently large arbitrary integer and H is obviously a generalized Hankel matrix consisting of a number of system Markov parameters. From Eq. (16), one obtains

$$\overline{Y}^{(2)} H = \underline{Y} \tag{17}$$

By using the definition of system Markov parameters, the Hankel matrix can be expressed by

$$H = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{p-1} \end{bmatrix} A \begin{bmatrix} B & AB & A^2B & \cdots & A^{N-1}B \end{bmatrix} = VAW \tag{18}$$

11

where A is the system state matrix for a discrete model representation, V the observability matrix, and W the controllability matrix. Then Eq. (17) becomes

$$\bar{Y}^{(2)}H = \bar{Y}^{(2)}[VAW] = \underline{Y} \tag{19}$$

It is known that the rank of a sufficiently large H is the order of the controllable and observable part of the system. From the experimental point of view, the identified state matrix A represents only the controllable and observable part of the system. The size of the matrix H is qp by Nm where N is an arbitrary integer Assuming that $Nm > qp$, the maximum rank of H is thus qp. If p is chosen such that $qp > n$ (the order of the matrix A) and $\bar{Y}^{(2)}$ is obtained uniquely, then a realized state matrix A with order n should exist.

Therefore we conclude that the number of observer Markov parameters computed, p, must be chosen such that qp>n where q is the number of outputs and n the order of the system and obviously p can be smaller than the true order of the system for a multiple output system. For a single output system, the number p must be greater than or equal to the true order of the system. The number p determines the maximum number of independent system Markov parameters as seen from Eq. (15). Therefore, p represents the upper bound on the identified system model. When a Hankel matrix is formed for the purpose of system identification, there is no benefit to include additional system Markov parameters beyond the necessary number to create a full rank Hankel matrix.

Equation (15) can be written in the following matrix form

$$
\begin{bmatrix}
I & & & & \\
-\bar{Y}_0^{(2)} & I & & & \\
-\bar{Y}_1^{(2)} & -\bar{Y}_0^{(2)} & I & & \\
\vdots & \vdots & \vdots & \ddots & \\
-\bar{Y}_{k-1}^{(2)} & -\bar{Y}_{k-2}^{(2)} & -\bar{Y}_{k-3}^{(2)} & \cdots & I
\end{bmatrix}
\begin{bmatrix}
Y_0 \\
Y_1 \\
Y_2 \\
\vdots \\
Y_k
\end{bmatrix}
=
\begin{bmatrix}
\bar{Y}_0^{(1)} + \bar{Y}_0^{(2)}D \\
\bar{Y}_1^{(1)} + \bar{Y}_1^{(2)}D \\
\bar{Y}_2^{(1)} + \bar{Y}_2^{(2)}D \\
\vdots \\
\bar{Y}_k^{(1)} + \bar{Y}_k^{(2)}D
\end{bmatrix}
\tag{20}
$$

Note that I and all $\bar{Y}_i^{(2)}$ $(i=0, 1, ..., k)$ are $q \times q$ square matrices. It is immediately seen that back substitution for $Y_0, Y_1, ..., Y_k$ from Eq. (20) yields Eq. (15). It is known that recursive back substitution without pivoting may result in a significant error accumulation in the solution. For numerical accuracy, it is better to use some type of pivoting procedure to minimize the error accumulation, unless the diagonal terms are dominant. However, the recursive back substitution is

1 2

superior in computational efficiency relative to other methods, particularly when real-time computation is required.

To identify the observer gain M, first recover the sequence of parameters

$$Y_k^o = CA^k M \; ; \; k = 0, 1, 2, \ldots \tag{21}$$

in terms of the observer Markov parameters. In fact, the first parameter in the sequence is simply

$$Y_0^o = CM = -\overline{Y}_0^{(2)} \tag{22}$$

The next parameter in the sequence is obtained by considering $\overline{Y}_1^{(2)}$

$$\overline{Y}_1^{(2)} = -C\overline{A}M = -(CAM + CMCM) = -Y_1^o + \overline{Y}_0^{(2)}Y_0^o$$

which yields

$$Y_1^o = -\overline{Y}_1^{(2)} + \overline{Y}_0^{(2)}Y_0^o \tag{23}$$

Similarly,

$$\overline{Y}_2^{(2)} = -C\overline{A}^2 M = -\left(CA^2 M + CMCAM + C\overline{A}MCM\right) = -Y_2^o + \overline{Y}_0^{(2)}Y_1^o + \overline{Y}_1^{(2)}Y_0^o$$

which yields

$$Y_2^o = -\overline{Y}_2^{(2)} + \overline{Y}_0^{(2)}Y_1^o + \overline{Y}_1^{(2)}Y_0^o \tag{24}$$

By induction, the general relationship is

$$Y_k^o = -\overline{Y}_k^{(2)} + \sum_{i=0}^{k-1} \overline{Y}_i^{(2)}Y_{k-i-1}^o \tag{25}$$

Having obtained the sequence $Y_k^o = CA^k M \; ; \; k = 0, 1, 2, \ldots$ where C and A can be realized by an identification method[11-12] from the Markov parameter sequence $Y_k = CA^k B \; ; \; k = 0, 1, 2, \ldots$ obtained from Eq. (20), the observer gain M can be computed from

$$M = \left(O^T O\right)^{-1} O^T Y^o \tag{26}$$

where

13

$$O = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^k \end{bmatrix} , \quad Y^o = \begin{bmatrix} Y_0^o \\ Y_1^o \\ Y_2^o \\ \vdots \\ Y_k^o \end{bmatrix} = \begin{bmatrix} CM \\ CAM \\ CA^2M \\ \vdots \\ CA^kM \end{bmatrix} \tag{27}$$

Equation (25) can be written in matrix form as

$$\begin{bmatrix} I & & & & \\ -\overline{Y}_0^{(2)} & I & & & \\ -\overline{Y}_1^{(2)} & -\overline{Y}_0^{(2)} & I & & \\ \vdots & \vdots & \vdots & \ddots & \\ -\overline{Y}_{k-1}^{(2)} & -\overline{Y}_{k-2}^{(2)} & -\overline{Y}_{k-3}^{(2)} & \cdots & I \end{bmatrix} \begin{bmatrix} Y_0^o \\ Y_1^o \\ Y_2^o \\ \vdots \\ Y_k^o \end{bmatrix} = \begin{bmatrix} -\overline{Y}_0^{(2)} \\ -\overline{Y}_1^{(2)} \\ -\overline{Y}_2^{(2)} \\ \vdots \\ -\overline{Y}_k^{(2)} \end{bmatrix} \tag{28}$$

The above italicized statement about Eq. (15) regarding the number of independent system Markov parameters also applies to the observer gain Markov parameters, Y_k^o, in Eq. (25) or (28). Note that I and all $\overline{Y}_i^{(2)}$ ($i = 0, 1, \ldots$) are $q \times q$ square matrices. Therefore, the left most matrix in Eq. (28) is square and full rank and identical to that in Eq. (20). Hence, Y^o is determined uniquely from an identified set of observer Markov parameters. Equation (26) implies that the observer gain M computed from Eq. (26) is automatically in the same coordinates as those for a set of A, C, (and B) resulted from any realization. Recall that the the set of system Markov parameters used for realization of the system is also uniquely determined from the same identified set of observer Markov parameters, Eq. (15) or Eq. (20). Computationally, Eqs. (15) and (25) or Eqs. (20) and (28) can be combined as a single matrix equation to solve for Y_k and Y_k^o simultaneously, i.e.

$$\begin{aligned} P_k = \begin{bmatrix} Y_k & Y_k^o \end{bmatrix} &= \begin{bmatrix} CA^kB & CA^kM \end{bmatrix} = CA^k \begin{bmatrix} B & M \end{bmatrix} \\ &= \begin{bmatrix} \overline{Y}_k^{(1)} + \overline{Y}_k^{(2)}D & -\overline{Y}_k^{(2)} \end{bmatrix} + \sum_{i=0}^{k-1} \overline{Y}_i^{(2)} \begin{bmatrix} Y_{k-i-1} & Y_{k-i-1}^o \end{bmatrix} \end{aligned} \tag{29}$$

Conventional system identification methods would use only the impulse response history, Y_k to determine A, B, C and D. Here, the combined system and observer gain Markov parameters, P_k, are used in a Hankel matrix to identify $A, [B \ M], C$ and D by some time domain method such as ERA[11] or ERA/DC[12]. There are several advantages for this approach. First, the observer gain,

14

M, is obtained directly, which will be shown to be related to the Kalman filter gain in the next section. Second, the number of independent Markov parameters has been compressed by using the observer. This allows one to use a smaller Hankel matrix and thus reduce the computational effort in the identification algorithm. Third, one can identify the number of independent system Markov parameters from a single set of data for lightly damped systems with multiple inputs and multiple outputs. This is a result of increased stability produced by adding an observer gain which allows one to use a smaller p in Eq. (7) than in Eq. (3).

Relationship Between the Identified Observer and a Kalman Filter

Let Eq. (1) be extended to include process and measurement noise described as

$$x(i+1) = Ax(i) + Bu(i) + w(i)$$
$$y(i) = Cx(i) + Du(i) + v(i)$$

(30)

where $w(k)$ is the process noise assumed to be Gaussian, zero mean and white with the covariance matrix Q, and $v(i)$ is the measurement noise with the same assumption as $w(k)$ but a different covariance matrix R. The sequence $w(i)$ and $v(i)$ are assumed statistically independent of each other.

A typical Kalman filter for the above equation can then be written as

$$\hat{x}^+(i) = \hat{x}^-(i) + K[y(i) - \hat{y}(i)] \overset{\Delta}{=} \hat{x}^-(i) + K\varepsilon_r(i)$$
$$\hat{x}^-(i) = A\hat{x}^+(i-1) + Bu(i-1)$$
$$\hat{y}(i) = C\hat{x}^-(i) + Du(i)$$

(31)

where $\hat{x}^+(i)$ is the estimated state. The term $\varepsilon_r(i)$ is called the residual and is defined as the difference between the real measurement $y(i)$ and the predicted measurement $\hat{y}(i)$. Combination of the first two equations in Eq. (31) yields

$$\hat{x}^-(i+1) = A[I - KC]\hat{x}^-(i) + [B - AKD]u(i) + AKy(i)$$

or

$$\hat{x}^-(i+1) = \tilde{A}\hat{x}^-(i) + \tilde{B}v(i)$$
$$y(i) = C\hat{x}^-(i) + Du(i) + \varepsilon_r(i)$$

(32)

15

where

$$\tilde{A} = A[I - KC]$$

$$\tilde{B} = [B - AKD, \ AK]$$

$$v(i) = \begin{bmatrix} u(i) \\ y(i) \end{bmatrix}$$

Comparison of Eqs. (4) and (32) reveals that they are identical if $M = -AK$ and $\varepsilon_r(i) = 0$, and so are their Markov parameters. An question immediately arises as to whether $K = -A^{-1}M$ if M is computed using the computational procedure developed above. It is known that the Kalman filter gain K depends on the process covariance Q and the measurement covariance R. There must exist some conditions such that the equation $M = -AK$ is valid due to the fact that the same equations are used to solve for the Kalman filter gain K and for the observer gain M. The key is the error term. The conditions will be derived in the following.

Equation (32) can be written in the following matrix form

$$\bar{y} = \tilde{Y}\bar{V} + \varepsilon + C\tilde{A}^p\hat{x}^- \tag{33}$$

where $\bar{y}$ and $\bar{V}$ are defined in Eq. (8) and

$$\tilde{Y} = \begin{bmatrix} D & C\tilde{B} & C\tilde{A}\tilde{B} & \cdots & C\tilde{A}^{p-1}\tilde{B} \end{bmatrix}$$

$$\hat{x}^- = \begin{bmatrix} \hat{x}^-(0) & \hat{x}^-(1) & \hat{x}^-(2) & \cdots & \hat{x}^-(\ell-p-2) \end{bmatrix}$$

$$\varepsilon = \begin{bmatrix} \varepsilon_r(p+1) & \varepsilon_r(p+2) & \varepsilon_r(p+3) & \cdots & \varepsilon_r(\ell-1) \end{bmatrix}$$

and ε is the residual error as defined in Eq. (31), and ℓ is the data length. This equation applies to any equation with the same observer structure as Eq. (31). If the observer happens to be a Kalman filter, then the residual is white, zero-mean, and Gaussian.

Postmultiplying Eq. (33) by $\bar{V}^T$ yields

$$\bar{y}\bar{V}^T = \tilde{Y}\bar{V}\bar{V}^T + \varepsilon\bar{V}^T + C\tilde{A}^p\hat{x}^-\bar{V}^T \tag{34}$$

Let $\bar{V}$ be partitioned in rows as

$$\bar{V}^T = \begin{bmatrix} v_{p+1}^T & v_p^T & \cdots & v_0^T \end{bmatrix}$$

Equation (34) can then be rewritten as

1 6

$$\begin{bmatrix} \bar{y}v_{p+1}^T & \bar{y}v_p^T & \cdots & \bar{y}v_0^T \end{bmatrix} - \tilde{Y}\begin{bmatrix} v_{p+1}v_{p+1}^T & v_{p+1}v_p^T & \cdots & v_{p+1}v_0^T \\ v_pv_{p+1}^T & v_pv_p^T & \cdots & v_pv_0^T \\ \vdots & \vdots & \ddots & \vdots \\ v_0v_p^T & v_0v_{p-1}^T & \cdots & v_0v_0^T \end{bmatrix}$$

$$= \begin{bmatrix} \varepsilon v_{p+1}^T & \varepsilon v_p^T & \cdots & \varepsilon v_0^T \end{bmatrix} + C\tilde{A}^p\begin{bmatrix} \hat{x}^- v_{p+1}^T & \hat{x}^- v_p^T & \cdots & \hat{x}^- v_0^T \end{bmatrix};$$

Let us examine a term from $\varepsilon \bar{V}^T$

$$\varepsilon v_i^T = \sum_{j=1}^{\ell-p-1} \varepsilon_r(p+j)v^T(i+j-1); \quad i = 0,...,p+1 \tag{35}$$

If this term is divided by ℓ-p-1, it represents the time average of the product $\varepsilon(k)v^T(k-i)$ from $k=p+1$ to ℓ-1. By the ergodic property[13], if the product is a sample function of a stationary random process, it can be replaced by its ensemble average provided ℓ goes to infinity, $\ell \to \infty$,

$$E[\varepsilon_r(k)v^T(k-i)] = \lim_{\ell\to\infty}\frac{1}{\ell-p-1}\sum_{j=p+1}^{\ell-1}\varepsilon_r(j)v^T(j-i); \quad k > p. \tag{36}$$

Physically, ergodicity implies that a sufficiently long record of a stationary random process contains all the statistical information about the random phenomenon. In practical applications, the ergodic property makes it possible to obtain the noise-related moment functions of a stationary random process from a single long record. The conversion from a time average to an expected value is performed similarly for the other terms including $\bar{y}v_j^T$, $\hat{x}^- v_j^T$, $v_iv_j^T (i,j = 0,...,p+1)$.

The concept of stationarity in a random process is analogous to the steady-state behavior of a deterministic process. In practice, no random process can be truly stationary. However, a long segment of a random process exhibiting uniform characteristics can be treated as stationary. One must allow sufficient time for the system transients to decay before the data sequence starts. In addition, the choice of p in Eq. (33) has to be sufficiently large that the transients of the Kalman filter are negligible.

Equation (34) can now be written as

$$\lim_{\ell \to \infty} \frac{1}{\ell - p - 1} [\bar{y}\bar{V}^T - \tilde{Y}\bar{V}\bar{V}^T] = E\big[\varepsilon_r(k)v^T(k) \quad \varepsilon_r(k)v^T(k-1) \quad \cdots \quad \varepsilon_r(k)v^T(k-p-1)\big]$$

$$+ C\tilde{A}^p E\big[\hat{x}^-(k)v^T(k+p+1) \quad \hat{x}^-(k)v^T(k+p) \quad \cdots \quad \hat{x}^-(k)v^T(k)\big] \tag{37}$$

for all $k > p$. If we choose the observer such that

$$\tilde{Y} = \bar{y}\bar{V}^T[\bar{V}\bar{V}^T]^{-1} \tag{38}$$

in the limit $\ell \to \infty$, then

$$E\big[\varepsilon_r(k)v^T(k) \quad \varepsilon_r(k)v^T(k-1) \quad \cdots \quad \varepsilon_r(k)v^T(k-p-1)\big]$$

$$= -C\tilde{A}^p E\big[\hat{x}^-(k)v^T(k+p+1) \quad \hat{x}^-(k)v^T(k+p) \quad \cdots \quad \hat{x}^-(k)v^T(k)\big] \tag{39}$$

Because $\tilde{A}$ for an observer is asymptotically stable, let p be chosen sufficiently large that the right hand side of the above equation is negligible, i.e.,

$$E[\varepsilon_r(k)v^T(k-i)] = 0 \tag{40}$$

for $i=0, ..., p+1$ and $k>p$. Substitution of the definition for $v(k-i)$ from Eq. (32) in Eq. (40) yields

$$E[\varepsilon_r(k)u^T(k-i)] = 0; \quad i = 0,...,p+1$$

$$E[\varepsilon_r(k)y^T(k-j)] = 0; \quad j = 1,...,p+1 \tag{41}$$

for $k>p$, which implies that the residual error $\varepsilon_r(k)$ at any time k is orthogonal to the input function $u(k-i)$ with the time delay i from 1 up to $p+1$, and the output function $y(k-j)$ with the time delay j from 0 up to $p+1$. In other words, if we choose the observer with the observer Markov parameters which satisfy the least square Eq. (38), the residual describing the difference between the estimated output measurement and real measurement is orthogonal to the given input and the measured output with time delay. This has application to model reduction based on the orthogonality of the output measurements and the residuals, representing the output errors between the full model and the reduced model.[14]

Now given a set of data from a finite-dimensional system of Eq. (30), there exists a Kalman filter with the property that the residual is white, zero-mean, and Gaussian, i.e.

$$E[\varepsilon_r(k)] = 0; \quad E[\varepsilon_r(j)\varepsilon_r^T(k)] = 0; \quad j \neq k \tag{42}$$

18

and satisfies the principle of orthogonality

$$E[\varepsilon_r(k)y^T(k-i)] = 0; \quad i = 1,...,k \tag{43}$$

If the experimental process is stationary and random, the Kalman filter gain is a constant which produces the Kalman filter Markov parameters in the limit $l \to \infty$ satisfying the least squares Eq. (36) provided the inverse $[\overline{V}\overline{V}^T]^{-1}$ exists. For a sufficiently rich input, the inverse always exists.

We conclude that any observer satisfying Eq. (10), or its equivalent Eq. (38), produces the same input-output map as a Kalman filter does if the data length is sufficiently long and the order of the observer is sufficiently large so that the truncation error is negligible. Therefore, when reduced to the system order, the identified observer has to be a Kalman filter and thus the M computed from the combined Markov parameters of Eq. (29) gives the steady state Kalman filter gain

$$K = -A^{-1}M \tag{44}$$

Computational Algorithm

Given a set of experimental input and output data, the identification algorithm proceeds as follows.

Step 1: Choose a value of p (see Eq. (7)) which determines the number of observer Markov parameters to be identified from the given set of input and output data. In general, p is required to be sufficiently larger (at least four or five times) than the effective order of the system for identification of the Kalman filter gain with accuracy.

Step 2: Form the two data matrices y and V as shown in Eq. (7) for zero initial conditions, or $\overline{y}$ and $\overline{V}$ as shown in Eq. (9) for nonzero initial conditions, and compute the least squares solution of the observer Markov parameter matrix $\overline{Y}$.

Step 3: Recover the combined system and Kalman filter Markov parameters, P_k, from the identified observer Markov parameters using Eq. (29). To solve for more Markov parameters than the number of identified observer Markov parameters, simply set the extra observer Markov parameters to zero.

19

Step 4: Realize a state space model of the system and the corresponding Kalman filter gain from the recovered sequence P_k using ERA or ERA/DC.

Numerical Example

As an example, a spring/mass three-degree-of-freedom system is used to simulate data with known noise properties. The simulated system, used in Ref. 9, has one input and two outputs. The continuous system is discretized at a sampling frequency of 10 Hz. The discrete time model for this system is

$$A = diag\left\{\begin{bmatrix} 0.9856 & 0.1628 \\ -0.1628 & 0.9856 \end{bmatrix}, \begin{bmatrix} 0.8976 & 0.4305 \\ -0.4305 & 0.8976 \end{bmatrix}, \begin{bmatrix} 0.8127 & 0.5690 \\ -0.5690 & 0.8127 \end{bmatrix}\right\}$$

$$B = \begin{bmatrix} 0.0011 & 0.0134 & -0.0016 & -0.0072 & 0.0011 & 0.0034 \end{bmatrix}^T$$

$$C = \begin{bmatrix} 1.5119 & 0.0000 & 2.0000 & 0.0000 & 1.5119 & 0.0000 \\ 1.3093 & 0.0000 & 0.0000 & 0.0000 & -1.3093 & 0.0000 \end{bmatrix} \tag{45}$$

$$D = \begin{bmatrix} 0.0000 & 0.0000 \end{bmatrix}^T$$

where the matrix A is in block diagonal form for later comparison with the identification results. The process noise and measurement noise covariances are specified respectively to be

$$Q = diag[0.0242 \quad 3.5920 \quad 0.0534 \quad 1.034 \quad 0.0226 \quad 0.2279] \times 10^{-4}$$
$$R = diag[2.785 \quad 2.785] \times 10^{-2} \tag{46}$$

These covariances were chosen by the following procedure. First a simulation was performed using random $u(k)$ with a standard deviation of 20 to determine the noise free sequences $Bu(k)$ and $y(k)$. The standard deviation of the process noise was computed to be 5% that of the sequence $Bu(k)$. Similarly, the standard deviation of the measurement noise was chosen as 5% that of the sequence $y(k)$. To examine the stochastic properties of the system, one must assume that the sample histories are infinitely long but in practice they are not. Therefore, the effect of short time records must be examined. Also in the theoretical development the observer order p is specified a priori. In the simulations, two different values for the observer order parameter p are used and the results are compared.

20

The computational algorithm is applied to identify the system and the corresponding Kalman filter gain in the presence of the prescribed noise levels. Examination of Table 1 shows that the frequencies are accurately identified in all cases to within 0.2%. Damping estimates, however, vary up to 28% with improved results obtained when the number of data samples is increased. Computation of the frequencies and damping values is based on a realization of the system matrix A from the Markov parameter sequence. The realization algorithm presented in Ref. 12 is used, and a minimum order realization is obtained from the Hankel correlation matrix HH^T, where H is as in Eq. (16). For deterministic systems, the rank of the correlation matrix is equal to the system order. For stochastic systems the problem of rank determination is not as clear and the method of singular value decomposition is used to determine the system order. Retaining only those singular values with significant contribution to the correlation matrix renders a model of the same order as the number of retained singular values. The value of p is chosen to be either 40 or 50 and only the first 6 singular values are retained.

Table 1. Comparison of identified modal parameters

Case No.	Mode 1		Mode 2		Mode 3	
	Freq. (Hz)	Damp. (%)	Freq. (Hz)	Damp. (%)	Freq. (Hz)	Damp. (%)
0	0.261	0.63	0.712	1.01	0.972	1.30
1	0.261	0.55	0.712	0.96	0.970	1.65
2	0.261	0.56	0.712	0.95	0.970	1.67
3	0.261	0.59	0.712	0.99	0.971	1.52
4	0.261	0.59	0.712	1.00	0.971	1.51
5	0.261	0.65	0.712	0.99	0.971	1.52
6	0.261	0.65	0.712	0.98	0.971	1.53

Case 0: True Values
Case 1: 1000 data points, p=40
Case 2: 1000 data points, p=50
Case 3: 2000 data points, p=40

Case 4: 2000 data points, p=50
Case 5: 4000 data points, p=40
Case 6: 4000 data points, p=50

Kalman filter gains are shown in Table 2. Although the numerical comparison in terms of frequencies and damping values is good, the estimated Kalman filter gains for the different cases could be quite different from the true value because of the finiteness of the data lengths. Table 2 shows that as the number of data points used in the identification is increased, the identified Kalman filter gains approach the true value.

Table 2. Comparison of Kalman filter gains

Case No.	Kalman Filter Gain Matrix
0	$K = \begin{bmatrix} 0.0293 & -0.0012 & 0.0025 & 0.0000 & 0.0241 & 0.0047 \\ 0.0295 & 0.0012 & 0.0000 & 0.0005 & -0.0251 & 0.0042 \end{bmatrix}^T$
2	$K = \begin{bmatrix} 0.0261 & -0.0166 & 0.0198 & 0.0046 & 0.0236 & 0.0086 \\ 0.0253 & -0.0020 & -0.0010 & 0.0116 & -0.0328 & 0.0241 \end{bmatrix}^T$
6	$K = \begin{bmatrix} 0.0255 & -0.0066 & 0.0196 & -0.0047 & 0.0246 & 0.0071 \\ 0.0265 & -0.0052 & 0.0008 & 0.0037 & -0.0298 & 0.0062 \end{bmatrix}^T$

The corresponding realized system matrices for Case No. 6 are,

$$A = diag\left\{ \begin{bmatrix} 0.9856 & 0.1629 \\ -0.1629 & 0.9856 \end{bmatrix}, \begin{bmatrix} 0.8977 & 0.4306 \\ -0.4306 & 0.8977 \end{bmatrix}, \begin{bmatrix} 0.8120 & 0.5676 \\ -0.5676 & 0.8120 \end{bmatrix} \right\}$$

$$B = \begin{bmatrix} 0.0011 & 0.0132 & -0.0016 & -0.0068 & 0.0011 & 0.0034 \end{bmatrix}^T$$

$$C = \begin{bmatrix} 1.5107 & 0.0000 & 2.0000 & 0.0000 & 1.5133 & 0.0000 \\ 1.3107 & 0.0011 & -0.0087 & -0.0104 & -1.3073 & 0.0288 \end{bmatrix} \tag{47}$$

$$D = \begin{bmatrix} 0.0007 & 0.0012 \end{bmatrix}^T$$

Comparing the B and C matrices with the true ones in Eq. (45) shows excellent agreement but a nonzero direct transmission term is picked up by the identification. The reconstruction (not shown) of the system response using the identified observer parameters in Eq. (47) and the identified Kalman filter gain in Table 2 when compared to the actual response shows excellent agreement.

Experimental Results

To demonstrate the identification procedure using real experimental data, the structure shown in the photograph in Fig. (1) is used. The truss is 100 inches long with a square cross section of 10 in x 10 in. All the tubing (longerons, battens, and diagonals) and ball joints are made of aluminum. The structure is in a vertical configuration attached from the top using an L-shaped fixture to a backstop. Two cold air thrusters acting in the same direction are placed at the beam tip. The thrusters which are used for excitation and control have a maximum thrust of 2.2 lb each. A mass

of approximately 20 lb is attached at the beam tip to lower the fundamental frequency of the truss. Two servo accelerometers located at a corner of the square cross section provides the in-plane tip acceleration.

The structure was excited using random inputs to both thrusters for 30 seconds. The input signals were filtered to concentrate the energy in the low frequency range. A total of 7498 data points at a sampling rate of 250 Hz is used for identification. The two output acceleration signals were filtered using a three pole Bessel filter with a break frequency of 20 Hz. The value of p is set to 10. Recall that the maximum identified system order can not exceed the number of outputs times the value of p. Therefore, the system order is chosen is to be 14 for realization of the system matrices, using the identified Markov parameters. Since the noise properties of the test structure are unknown, comparison of the identified Kalman filter gain is not possible. Instead the identified Kalman filter model is used to reconstruct the output. Figure (2) shows overlaping 8 seconds of the reconstruction and the test data for the two accelerometers used. The dashed line is reconstruction and solid is test data. Only one plot is seen because the prediction error for both outputs is less that 0.1 %. If one uses the predictor part of the Kalman filter which includes only the matrices A,B,C and D, the reconstruction is shown in Fig. (3). There are some visible differences between test and reconstruction but overall the model is very accurate. It is important that the predictor part be accurate because it is this part that is used as a model for control design. There have been cases where the identified model including Kalman filter gain showed good agreement but the predictor part did not. In those cases it is important to examine the realization parameters, truncation error and residual error to ensure that the proper order has been selected. The identified discrete system matrices are given in the appendix. This identified system was obtained using a single pair of time histories. To check the validity of the model with different tests, a second experiment is performed and the time histories compared with the prediction from this model. The results are as good as those in Fig. (2) and (3). The test structure has three dominant modes with frequencies at 5.8, 7.3, and 48.5 Hz and damping values of 0.5, 0.5, and 0.9 %, respectively.

Concluding Remarks

An algorithm for the direct computation of observer/Kalman filter Markov parameters, and from them the observer/Kalman filter matrices, has been presented. The matrix formulation developed here allows one to establish the uniqueness and invertability of the transformation from observer/Kalman filter Markov parameters to the system Markov parameters. The matrix formulation also establishes bounds on the choice of the observer/Kalman filter order for the data. The algorithm is a non-recursive matrix version of two previous algorithms, one is a recursive

23

algorithm for Kalman filter identification, and the other is an algorithm for direct identification of observers with chosen pole locations, specialized to have all poles at the origin (the deadbeat observer). The non-recursive form of the least squares solution used here results in a substantial improvement in the convergence rate to the true Kalman gain that was reported before. The relationship between the deadbeat observer and the Kalman filter Markov parameter identification problems is established here. It is shown that using the equations for deterministic deadbeat observer parameters on noisy data, results in obtaining the Kalman filter parameters, in the limit as the amount of data used tends to infinity. When a finite set of data is used, the resulting filter satisfies an optimality condition indicating that it is the best filter that can be obtained with the data length available.

References

[1] Juang, J.-N. and Phan, M., "Robust Controller Designs for Second-Order Dynamic Systems: A Virtual Passive Approach," *NASA Technical Memorandum TM-102666*, May 1990.

[2] Juang, J.-N. and Lew, J.-S., "Integration of System Identification and Robust Controller Designs for Flexible Structures in Space," Paper No. AIAA-90-3467-CP, *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Portland, Oregon, August 1990, pp. 1361-1375.

[3] Juang, J.-N., and Pappa, R. S., "A Comparative Overview of Modal Testing and System Identification for Control of Structures," *Shock and Vibration Digest*, Vol. 20, No. 5, May 1988.

[4] Pappa, R. S., and Juang, J.-N., "Some Experiences with the Eigensystem Realization Algorithm," *Sound and Vibration*, Vol. 22, No. 1, Jan. 1988, pp. 30-35.

[5] Juang, J.-N., Horta, L.G., Phan, M., and Longman, R.W., "Learning System Identification for Flexible Structures,"*Proceedings of the 2nd USAF/NASA Workshop on System Identification and Health Monitoring of Precision Space Structures*, California Institute of Technology, Pasadena, California, March 1990.

[6] Phan, M., Juang, J.-N., Longman, R. W., "Identification of Linear Systems from a Single Set of Data by Identification of Observers with Assigned Real Eigenvalues," Paper No. 91-0949, *Proceedings of the AIAA 32nd Structures, Structural Dynamics, and Materials Conference*, Baltimore, Maryland, April 1991.

[7] Phan, M., Horta, L. G., Juang, J.-N., Longman, R. W., "Linear System Identification Via an Asymptotically Stable Observer," *Submitted for possible presentation at the AIAA Guidance, Navigation and Control Conference*, New Orleans, Louisiana, August 1991.

[8] Phan, M., Horta, L. G., Juang, J.-N., Longman, R. W., "Identification of Linear Systems Via an Asymptotically Stable Observer," to appear as a *NASA Technical Paper*, April 1991.

[9] Chen, C.-W., Huang, J.-K. Phan, M., and Juang, J.-N., "Integrated System Identification and Modal State Estimation for Control of Large Flexible Space Structures," Paper No. AIAA-90-3470-CP, *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Portland, Oregon, August 1990, pp. 1396-1404. Also accepted for publication in the *Journal of Guidance, Control and Dynamics*.

[10] Huang, J.-K., Juang, J.-N. and Chen, C.-W., "Single-Mode Projection Filters for Modal Parameter Identification for Flexible Structures," *Journal of Guidance, Control and Dynamics*, Vol. 12, No. 4, July-Aug. 1989, pp. 568-576.

[11] Juang, J.-N., and Pappa, R. S., "An Eigensystem Realization Algorithm for Modal Parameter Identification and Model Reduction," *Journal of Guidance, Control and Dynamics*, Vol. 8, No. 5, Sept.-Oct. 1985, pp. 620-627.

[12] Juang, J.-N, Cooper, J.E., and Wright J.R, "An Eigensystem Realization Algorithm Using Data Correlations (ERA/DC) for Modal Parameter Identification," *Control-Theory and Advanced Technology*, Vol. 4, No. 1, pp. 5-14, 1988.

[13] Nigam, N. C., *Introduction to Random Vibration*, The MIT Press, Cambridge, Massachusetts, 1983.

[14] Gawronski, W. and Juang, J.-N., "Model Reduction for Flexible Structures," *Advances in Large Scale Systems Dynamics*, edited by C. T. Leondes, Academic Press, Inc., New York, 1990, pp. 143-222.

Appendix

Identified discrete time system matrices for truss structure. Sampling frequency is 250 Hz.

$$A = diag\left\{\begin{bmatrix} -0.8500 & 0.2936 \\ -0.2936 & -0.8500 \end{bmatrix}, \begin{bmatrix} -0.5531 & 0.6185 \\ -0.6185 & -0.5531 \end{bmatrix}, \begin{bmatrix} -0.2663 & 0.9207 \\ -0.9207 & -0.2663 \end{bmatrix}, \begin{bmatrix} 0.9885 & 0.1466 \\ -0.1466 & 0.9885 \end{bmatrix}, \right.$$
$$\left. \begin{bmatrix} 0.9824 & 0.1819 \\ -0.1819 & 0.9824 \end{bmatrix}, \begin{bmatrix} 0.2519 & 0.8203 \\ -0.8203 & 0.2519 \end{bmatrix}, \begin{bmatrix} 0.3410 & 0.9286 \\ -0.9286 & 0.3410 \end{bmatrix} \right\}$$

$$B = \begin{bmatrix} 0.4262 & -0.5070 \\ 0.2259 & -0.5824 \\ 0.1531 & 0.2496 \\ -0.3463 & -0.1154 \\ -0.1863 & 0.2382 \\ 0.0792 & 0.0805 \\ 0.0999 & -0.1017 \\ -0.0438 & 0.0484 \\ 0.0630 & -0.0676 \\ -0.0138 & 0.0196 \\ 0.1496 & -0.0304 \\ -0.4487 & -0.0154 \\ 0.0270 & 0.0903 \\ 0.2419 & 0.1165 \end{bmatrix} \quad C^T = \begin{bmatrix} 0.5828 & 1.8975 \\ 0.0000 & 0.2447 \\ 1.5925 & -1.1030 \\ 0.0000 & -0.4972 \\ 0.6539 & 1.8091 \\ 0.0000 & 0.5474 \\ 1.8031 & 0.8654 \\ 0.0000 & 0.0034 \\ 0.8830 & -1.7941 \\ 0.0000 & -0.0393 \\ 1.2509 & -1.3068 \\ 0.0000 & 0.8530 \\ 1.3514 & -1.4671 \\ 0.0000 & 0.1463 \end{bmatrix} \quad M = \begin{bmatrix} -0.2955 & -0.1710 \\ -0.1296 & -0.0384 \\ -0.0176 & -0.0638 \\ 0.2171 & -0.0307 \\ -0.0281 & 0.0050 \\ -0.1451 & 0.1529 \\ 0.2175 & 0.2670 \\ 0.2234 & 0.1295 \\ 0.0684 & 0.0661 \\ 0.0439 & 0.0436 \\ 0.1461 & 0.1301 \\ -0.0104 & 0.0195 \\ -0.0198 & 0.0210 \\ 0.0092 & -0.1252 \end{bmatrix}$$

$$D = \begin{bmatrix} -0.1260 & 0.0044 \\ 0.0362 & 0.1865 \end{bmatrix}$$
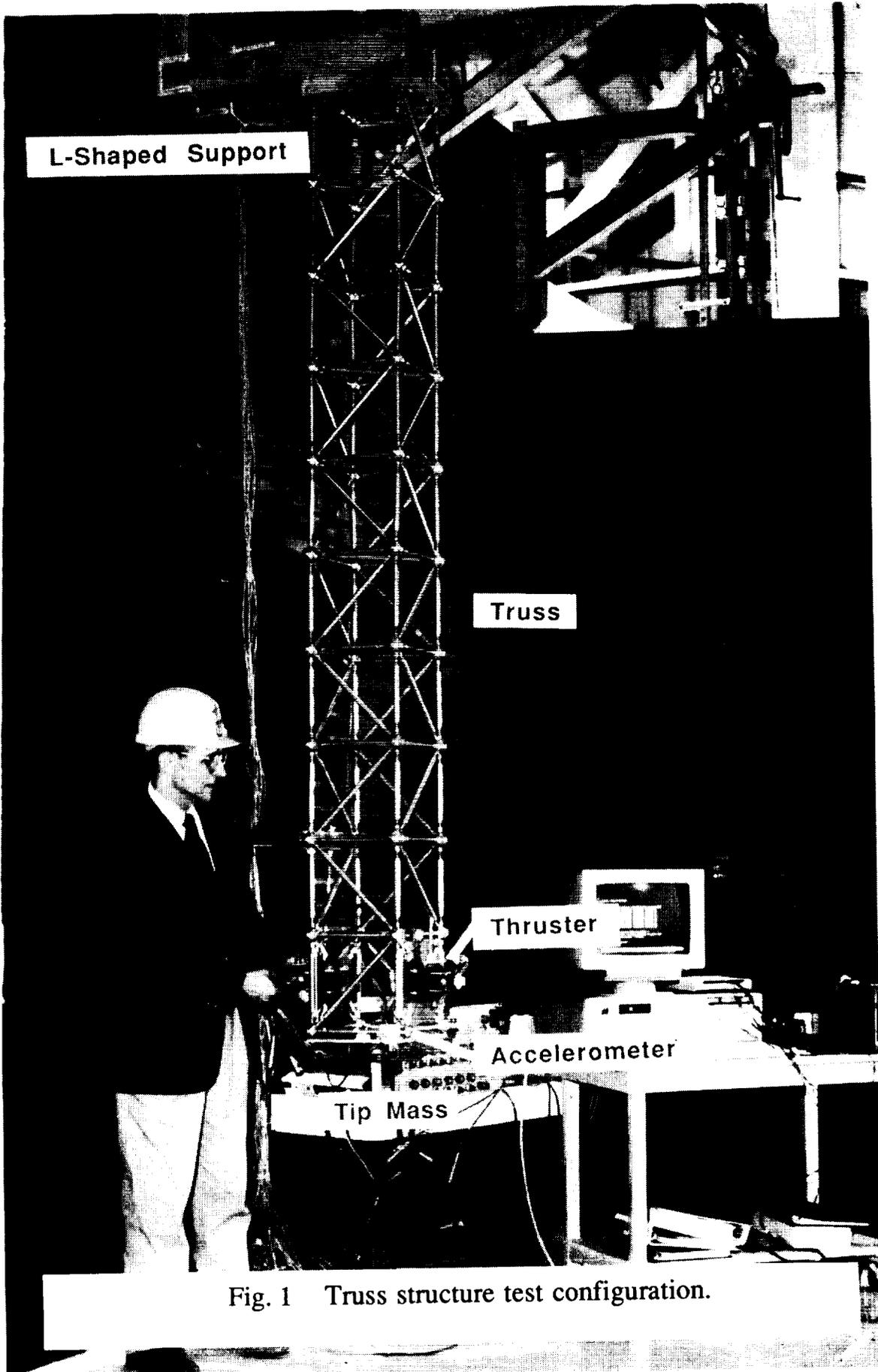
L-Shaped Support

Truss

Thruster

Accelerometer

Tip Mass

Fig. 1   Truss structure test configuration.

27
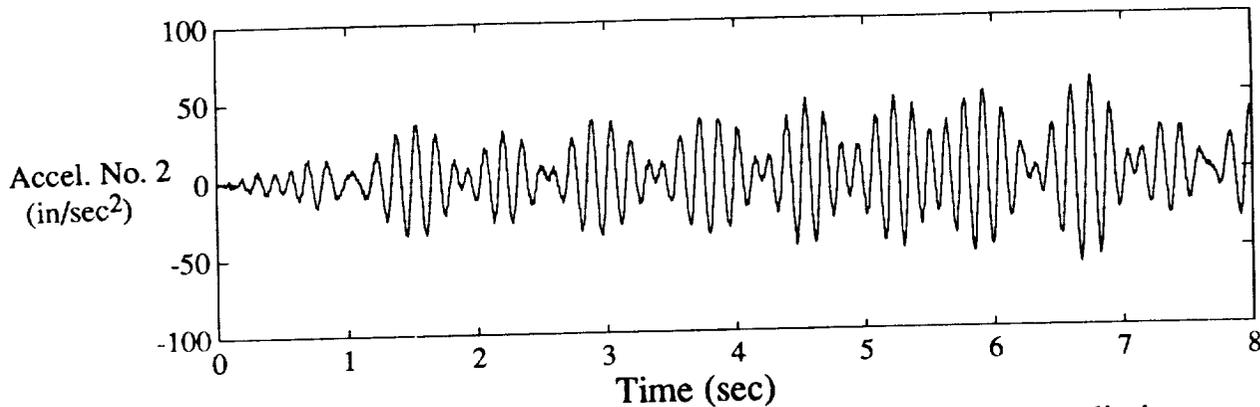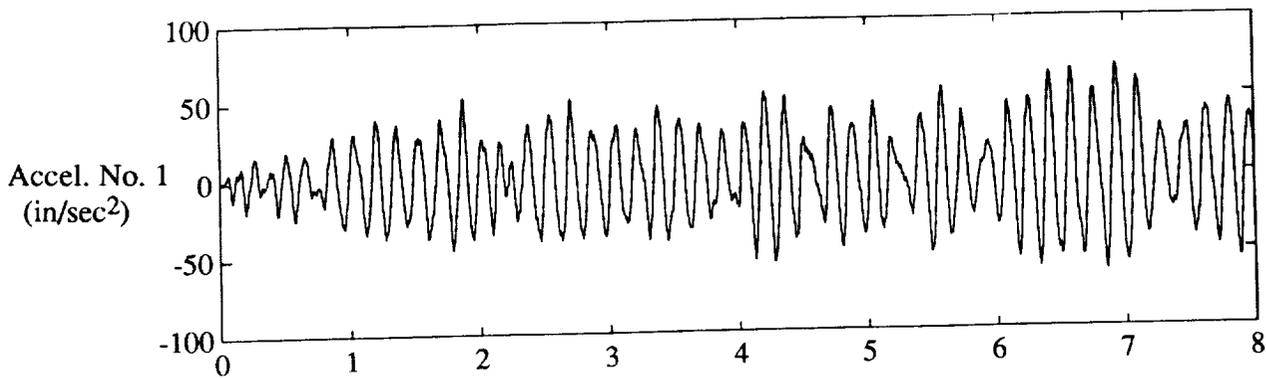
Fig. 2   Comparison of test data (solid) with Kalman filter prediction (dashed).
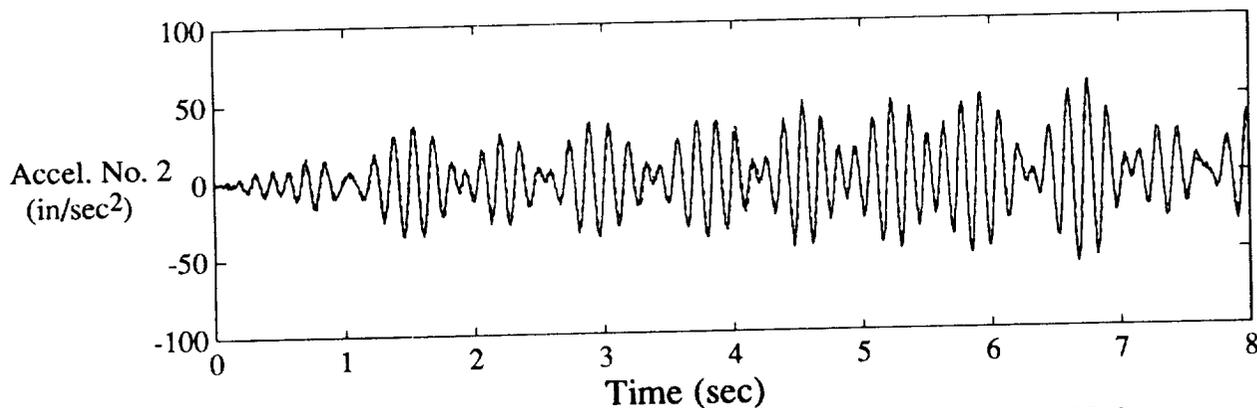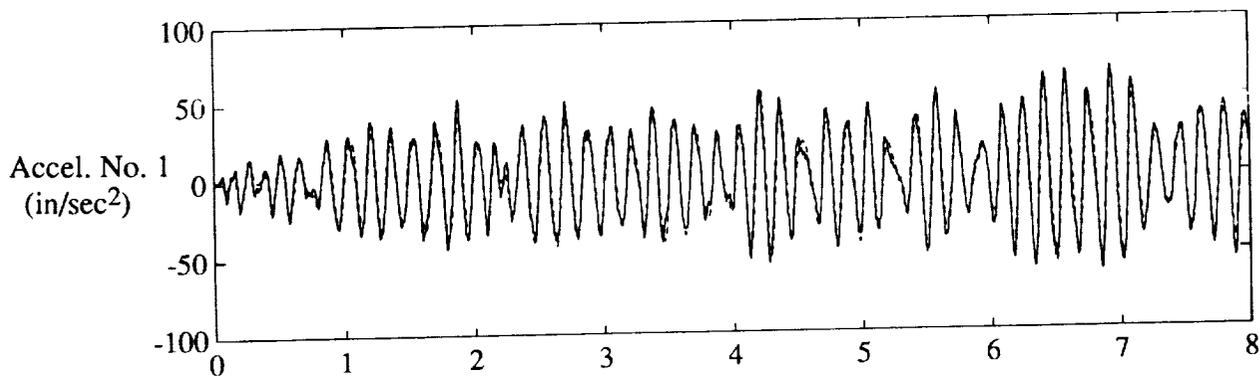
Fig. 3   Comparison of test data (solid) with predictor part of Kalman filter (dashed).

1. Report No. NASA TM-104069	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Identification of Observer/Kalman Filter Markov Parameters: Theory and Experiments		5. Report Date June 1991
		6. Performing Organization Code
7. Author(s) Jer-Nan Juang, Minh Phan, Lucas G. Horta, and Richard W. Longman		8. Performing Organization Report No.
		10. Work Unit No. 590-14-21-01
9. Performing Organization Name and Address NASA Langley Research Center Hampton, VA 23665-5225		11. Contract or Grant No.
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546-0001		13. Type of Report and Period Covered Technical Memorandum
		14. Sponsoring Agency Code

15. Supplementary Notes

Presented at the AIAA Guidance, Navigation, and Controls Conference
New Orleans, LA, August 12-14, 1991.
Jer-Nan Juang and Lucas G. Horta: Langley Research Center, Hampton, Virginia.
Minh Phan and Richard W. Longman: National Research Council Research Associates,
   Langley Research Center, Hampton, Virginia.

16. Abstract

This paper discusses an algorithm to compute the Markov parameters of an observer or Kalman filter from eperimental input and output data. The Markov parameters can then be used for identification of a state space representation, with associated Kalman gain or observer gain, for the the purpose of controller design. The algorithm is a non-recursive matrix version of two recursive algorithms developed in previous works for different purposes, and the relationship between these other algorithms is developed. The new matrix formulation here gives insight into the existence and uniqueness of solutions of certain equations, and gives bounds on the proper choice of observer order. It is shown that if one uses data containing noise, and seeks the fastest possible deterministic observer, the deadbeat observer, one instead obtains the Kalman filter--which is the fastest possible observer in the stochastic environment. The results of the paper are demonstrated in numerical studies and in experiments on a ten-bay truss structure.

17. Key Words (Suggested by Author(s)) Observer Kalman Filter System Identification State Esitmator		18. Distribution Statement Unclassified--Unlimited Subject Category 39	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of pages 29	22. Price A03