One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)

NASA TECHNICAL MEMORANDUM

NASA TM X-62,446

(NASA-TM-X-62446) RECURSIVE IDENTIFICATION AND TRACKING OF PARAMETERS FOR LINEAR AND NONLINEAR MULTIVARIABLE SYSTEMS (NASA) 39 p HC $3.75 CSCL 01C

N75-25960

Unclas
G3/08 28032

NASA TM X-62,446

RECURSIVE IDENTIFICATION AND TRACKING OF PARAMETERS FOR LINEAR AND NONLINEAR MULTIVARIABLE SYSTEMS

Menahem Sidar

Ames Research Center
Moffett Field, California 94035

JUL 1975
RECEIVED
NASA STI FACILITY
INPUT BRANCH

May 1975

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TM X-62,446		

4. Title and Subtitle	5. Report Date
RECURSIVE IDENTIFICATION AND TRACKING OF PARAMETERS FOR LINEAR AND NONLINEAR MULTIVARIABLE SYSTEMS	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Menahem Sidar	A-6115
9. Performing Organization Name and Address	10. Work Unit No.
NASA Ames Research Center Moffett Field, Calif. 94035	505-07-11
	11. Contract or Grant No.
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
National Aeronautics and Space Administration Washington, D.C. 20546	Technical Memorandum
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The problem of identifying constant and variable parameters in multi-input, multi-output, linear and nonlinear systems is considered, using the maximum likelihood approach. An iterative algorithm, leading to recursive identification and tracking of the unknown parameters and the noise covariance matrix, is developed. Agile tracking, and accurate and unbiased identified parameters are obtained. Necessary conditions for a globally, asymptotically stable identification process are provided; the conditions proved to be useful and efficient. Among different cases studied, the stability derivatives of an aircraft were identified and some of the results are shown as examples.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Identification Estimation Aircraft stability and control identification	Unclassified - Unlimited STAR Category - 08,59,66

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	39	$3.75

*For sale by the National Technical Information Service, Springfield, Virginia 22151

Recursive Identification and Tracking of Parameters for Linear and Nonlinear Multivariable Systems

Menahem Sidar*

ABSTRACT

The problem of identifying constant and variable parameters in multi-input, multi-output, linear and nonlinear systems is considered, using the maximum likelihood approach. An iterative algorithm, leading to recursive identification and tracking of the unknown parameters and the noise covariance matrix, is developed. Agile tracking, and accurate and unbiased identified parameters are obtained. Necessary conditions for a globally, asymptotically stable identification process are provided; the conditions proved to be useful and efficient. Among different cases studied, the stability derivatives of an aircraft were identified and some of the results are shown as examples.

1. Introduction

It has been recognized during the last several years that on-line recursive identification and tracking of unknown, time-varying parameters in linear and nonlinear dynamical systems is of paramount importance. Valuable algorithms that operate on recorded input-output time histories over finite time intervals have been reported for the batch identification of constant system parameters (Aström

*NRC Senior Research Associate, Ames Research Center, NASA, Moffett Field, California 94035

and Eykhoff, 1971 and Mehra, 1974). The relatively few recursive schemes reported to date (such as Hastings and Sage, 1969 and Gertler and Banyasz, 1974, for example) are applicable only to linear systems and, even there, probably not to high order, multivariable linear systems. Neither the batch nor the available recursive algorithms can be used in many problems of aeronautical interest. For example, identification of aircraft engine parameters, of aircraft stability and control derivatives in high angle-of-attack or accelerated flight, of atmospheric perturbations that vary through the flight regime—all require nonlinear or time-varying parameter tracking capability.

This paper adopts the maximum likelihood approach to provide a robust technique to satisfy the above aeronautical needs. An on-line recursive algorithm is developed for the simultaneous identification and tracking of unknown variable parameters. It is shown that the resulting scheme satisfies several guidelines: it provides and tracks unbiased estimated for multiparameter linear and nonlinear systems; it is computationally simple; and, it is usable on line.

After formulating the identification problem in section 2, the estimation criterion and the identification scheme are selected in section 3. Because our main goal is to develop an on-line identification algorithm, a recursive identification algorithm based on the maximization of the likelihood function is developed

in section 4. The algorithm makes use of certain sensitivity functions for the parameterized system. Section 5 shows how the sensitivity functions are generated and how to compute the sensitivity matrix. This matrix provides information that is also useful with respect to the system's identifiability. In order to provide necessary stability requirements for the recursive algorithm, convergency conditions are obtained in section 6, by the application of Lyapunov's second method and of a stability theorem concerning discrete-time dynamic systems.

Two examples are given and results showing the identification and tracking capabilities of the recursive algorithm are discussed in section 7.

2. Identification problem formulation

The basic problem we deal with here is the recursive identification of a set of (m) unknown parameters $\underline{p} = [p_1, p_2, \ldots, p_m]^T$ of the following (linear or nonlinear) dynamical system:

$$\dot{\underline{x}} = \underline{f}\,[\underline{x}, \underline{p}, \underline{u}(t)] \tag{1}$$

where $\underline{x} = [x_1, x_2, \ldots, x_n]^T$ is the n-dimensional state vector, representing the solution of the set of first order differential equations (1) for given: (i) initial conditions $\underline{x}\,(t=0) = \underline{x}^o$; and (ii) (r) control functions $\underline{u}_h(t)$, $h = 1, 2, \ldots, r$. We are also tacitly assuming in the sequel that a solution vector

$\underline{x}(t) \varepsilon E^n$ exists, is unique, and is stable. The existence of all the derivatives $\partial fi/\partial x_\rho$ and $\partial fi/\partial p_j$, for all $i, \rho = 1, \ldots, n$ and $j = 1, \ldots, m$ is also assumed.

Since in most applications there is no direct way to measure all of the state vector components, the set of measured (or observed) data available for identification purposes, is given by:

$$\underline{y} = C\underline{x} + \underline{v} \tag{2}$$

where $\underline{y}$ is the measured output (qxl) vector, C is the measurement matrix (qxn), and $\underline{v}(t)$ is a (qxl) vector representing the inevitable measurement noise.

With respect to $\underline{v}$, the following two assumptions are made. First, the measurement noise $\underline{v}(t)$ is an additive, Gaussian, white noise process, $\underline{v}(t) \perp \underline{x}(t)$, with zero mean value:

$$E\{\underline{v}(t)\} = 0 \tag{3}$$

At least for the length of the identification experiment T, this condition is considered true on an ensemble and time-average basis. Second, the covariance of the noise $\underline{v}(t)$ is given by:

$$E\left\{\underline{v}(t)\underline{v}^T(t-\tau)\right\} = R_o\delta(t-\tau) \quad , \quad \forall t,\tau \in [0,T] \tag{4}$$

This covariance is defined rather in a generalized, derived-martingale sense, R_o being finite. We suppose in the sequel that R_o is unknown and is to be found during the identification process.

The problem, then, is to use the (q) measured output functions and the (r) known and measured control functions to identify and track uniquely and with an acceptable accuracy, the unknown, possibly time-varying, elements of the parameter vector $\underline{p} = (p_1, \ldots, p_m)$.

The present state of the art makes it feasible to develop the identification algorithm in a discrete time version, since both the data and data processing are available in a discrete form, so that appropriate digital computers can be used for implementation of the recursive parameter identification. Therefore, the discrete versions of the controls, the state vector, the output vector, and the parameter vector are described in this paper as a sequence of discrete numbers, for all $i \varepsilon [0, N]$, $N \triangleq T/\Delta T$, ΔT being the basic time increment:

$$\begin{aligned} \underline{u}(i) &= [u_1(i), u_2(i), \ldots, u_r(i)]^T \\ \underline{x}(i) &= [x_1(i), x_2(i), \ldots, x_n(i)]^T \\ \underline{y}(i) &= [y_1(i), y_2(i), \ldots, y_q(i)]^T \\ \underline{p}(i) &= [p_1(i), p_2(i), \ldots, p_m(i)]^T \end{aligned} \qquad \forall i=0, 1, \ldots, N. \quad (5)$$

Similarly, we describe the measurement noise, as a random, zero mean, independent Gaussian sequence, with:

$$\begin{aligned} &E\{\underline{v}(i)\} = 0 \\ &E\left\{\underline{v}(i)\underline{v}^T(j)\right\} = R_1\delta_{kj} \quad (\delta_{ij} \text{ - Kronecker's delta function}) \end{aligned} \quad (6)$$

for all $i,j \varepsilon [0, N]$, as stated before in eqns. (3) and (4).

It is appropriate to remark at this point, that the identification technique presented here can be easily extended, with slight modifications, for the case when $\underline{v}(i)$ is a random correlated Gaussian sequence, that is, a colored measurement noise with given correlation functions. For clarity of presentation and to avoid cumbersome notations we prefer to treat throughout this paper only the uncorrelated and independent Gaussian measurement noise case.

With regard to the model assumed in eqns. (1) and (2), we should mention here that this model is by no means the most general one (see Mehra, 1970), but we deliberately adopted it for the sake of conciseness and simplicity of notation. Nevertheless, the present algorithm is compatible, with slight modifications, with the whole group of models for aircraft parameter identification, stability and control derivatives, as discussed in Mehra (1970) and in Stepner and Mehra (1973), including the case when process noise may exist (Wingrove, 1974).

3. Parameter estimation criteria and identification scheme

Prior to the development of the recursive identification algorithm we have to restate in this paragraph our main objective and the rationale adopted for this study. There are several possibilities for the choice of the topological scheme of the identification, each of those schemes having certain merits and inherent disadvantages. Discussions of these topics have appeared

in the technical literature, such as in Aström and Eykhoff (1974), IEEE (1974), and Landau (1974). Although the work of Aström and Eykhoff (1974) is probably one of the most objective and comprehensive, the other references are also excellent and up to date. Therefore, we will try to avoid unnecessary repetition and address ourselves to the main points of the rationale, those relevant to the sequel.

The output error method was adopted for the development of our identification algorithm because, taking into account the existence of the measurement noise, our main objective was to obtain unbiased estimates for the identified parameters. Equation error methods are known to produce biased estimates in the same case. Moreover, the algorithm has to be able to cope with the problem of the identification of biases (that possibly exist) in the measurement instrumentation. The output error method can deal successfully with this problem also.

It is unfortunately true that in adopting the output error scheme one has to deal with essentially a nonlinear iterative algorithm procedure. This fact is at the expense of the output error scheme, since in the equation error method the unknown parameters enter the dynamical identification equations in a linear fashion, so that the computational problem is, at least in principle, simple to deal with. Nevertheless, remembering that the original aim was to handle noisy nonlinear system identification problems with, eventually, process noise additive inputs, the

output error does not necessarily represent a prohibitive penalty from the computational point of view. As mentioned before, after judging the appropriate alternatives the output error method was finally adopted.

With respect to the identification criterion, our main objectives were: (1) to choose a mathematically tractable criterion so that the recursive algorithm would be easy to implement; (2) to obtain linear dependence between the parameter variations and the output error vector; (3) to generate the covariance matrix for the measurement noise as a by-product of the identification procedure, in order to check the robustness, the sensitivity, and the convergence properties of the identification process; and (4) to have a test for identifiability.

Checking the different criteria, which were potentially able to correspond more or less to the above rationale, we found the maximum likelihood function as the best choice of criterion (Aström and Eykhoff, 1971, Mehra, 1970 and Kashyap, 1970) because: (1) maximization of the likelihood estimation function leads to unbiased, consistent, and minimum variance estimates as asymptotic, limiting values, under fairly mild conditions; (2) maximization of the likelihood function leads to the evaluation of the covariance matrix for the measurement noise; and (3) the change in the parameter vector is linearly related to the output error vector. Moreover, the recursive-type version of the identification algorithm

is easily obtained from the batch type maximum likelihood estimation (m.l.e) algorithm. The m.l.e. may provide the covariance matrix for the identified parameters, indicating the correlation among them. It may also furnish a useful test and criterion identifiability. The m.l.e. criterion for recursive identification was adopted also by other authors (such as in Kashyap, 1970, Gertler and Banyasz, 1974 and Rogers and Steiglitz, 1967) but, as mentioned before, only for linear systems.

The conceptual diagram for the system identification is shown in figure 1, which uses the following notation:

1. For the (real) nominal, dynamical system, whose parameters are to be identified, we use, instead of eqns. (1) and (2), the equations:

$$\dot{\underline{x}}_{(n)} = \underline{f}\left[\underline{x}_{(n)}, \underline{p}_{(n)}(t), \underline{u}(t)\right] \tag{7a}$$

$$y_{(n)} = C\,\underline{x}_{(n)} + \underline{v} \tag{8a}$$

$\underline{p}_{(n)}(t)$ being the nominal values of the (unknown) parameters.

2. For the adjustable model, for a certain value $\underline{p}(t)$ of the parameter vector, the appropriate set of equations governing the system's model would be denoted by:

$$\dot{\underline{x}} = \underline{f}[\underline{x}, \underline{p}(t), \underline{u}(t)] \tag{7b}$$

$$\underline{y} = C\,\underline{x} \tag{8b}$$

$\underline{x}(t)$ and $\underline{y}(t)$ being the best available model state and model output, respectively. We assume that the exact value of $\underline{x}_{(n)}(0)$ is unknown. The output error is defined by:

$$\underline{\eta}(t) \triangleq \underline{y}_{(n)}(t) - \underline{y}(t)$$

$$= C.x_{(n)}(t) + \underline{v} - \underline{y}(t) \tag{9}$$

As explained before, being interested in the discrete-form identification algorithm, and as a preparatory step toward the recursive version of this algorithm, we will proceed further with the batched-type formulation of the identification problem in discrete version. Hence, for every sampling instant, we obtain the output error:

$$\underline{\eta}(i) = C\,\underline{x}_{(n)}(i) + \underline{v}(i) - \underline{y}(i), \qquad i\varepsilon[0, N] \tag{10}$$

which can be identified as an innovations sequence in the sense defined by Gevers and Kailath (1973). It is well known that as the sampling rate increases, the probability density distribution function of the innovations sequences, $\underline{\eta}(i)$ tends to a Gaussian p.d. distribution (Kailath, 1969).

Making use of the likelihood function approach, one has to find, that is, identify, the best estimate of $\underline{p}$, namely $\underline{p}^*$, based on the measured output sequence $\underline{Y}^k \triangleq [\underline{y}(1), \ldots, \underline{y}(k)]$. The maximum likelihood estimate of $\underline{p}$ is obtained by the maximization of the conditional probability of $\underline{Y}^k$, given $\underline{p}$:

$$\underline{p}^* = \left\{ \underline{p}^* : \max_{\underline{p}} \text{pr}\,(\underline{Y}^k | \underline{p}) \right\} \tag{11}$$

Applying, successively, the Bayes chain rule, one obtains:

$$\max_{\underline{p}} \text{pr}\,(\underline{Y}^k | \underline{p}) = \max_{\underline{p}} \left\{ \prod_{i=1}^{k} \text{pr}\left[\underline{y}(i) | \underline{Y}^{i-1}, \underline{p} \right] \right\} \tag{12}$$

As a matter of convenience, let us now define $[\ln \text{pr}\,(\underline{Y}^k | \underline{p})]$ as the likelihood function and, since the logarithm function is monotonic, we can write instead of eqn. (12):

$$\max_{\underline{p}} \left[\ln \text{pr}(\underline{Y}^k | \underline{p}) \right] = \max_{\underline{p}} \left\{ \sum_{i=1}^{k} \ln \text{pr}\left[\underline{y}(i) | \underline{Y}^{i-1}, \underline{p} \right] \right\} \tag{13}$$

Assuming that $\underline{x}(0)$ is normally distributed around $\underline{x}_{(n)}(0)$, and based on the hypothesis made in eqn. (6), we infer that $\text{pr}[\underline{y}(i) | \underline{Y}^{i-1}, \underline{p}]$ will also be normal and, consequently, described by second order statistics. Therefore, this quantity can be uniquely determined by computing the mean value:

$$E\left[\underline{y}(i) | \underline{Y}^{i-1}, \underline{p} \right] \triangleq \hat{y}(i|i-1) \tag{14}$$

and the covariance matrix:

$$E\{[\underline{y}(i) - \hat{\underline{y}}(i|i-1)][\underline{y}(i) - \hat{\underline{y}}(i|i-1)]^T\} \triangleq R(i|i-1) \tag{15}$$

the quantity $\nu(i) \triangleq [\underline{y}(i) - \hat{\underline{y}}(i|i-1)]$ being the "new information" brought up by $\underline{y}(i)$.

Taking into account eqns. (14) and (15), together with eqn. (13), we finally obtain:

$$\underline{p}^* = \{\underline{p}^* : \max_{\underline{p},R} L(\underline{p},R)\} \tag{16a}$$

where

$$L(\underline{p},R) = -\frac{1}{2}\sum_{i=1}^{k} \underline{\nu}^T(i)R^{-1}(i|i-1)\underline{\nu}(i) + \ln|R(i|i-1)| \tag{16b}$$

and:

$$\max_{\underline{p},R} L(\underline{p},R) = L(\underline{p}^*,\hat{R})$$

In eqn. (16b), $|R(i|i-1)|$ is the determinant of the symmetric covariance matrix $R(i|i-1)$ —to be written for brevity as $R(i)$.

It is obvious from the nature of eqns. (16), that $L(\underline{p},R)$ is not explicit in $\underline{p}$ or R and hence it is not possible to determine directly the value of the vector $\underline{p}^*$, unless a proper iterative maximization procedure is used. In order to do that, let us assume that we are generating a trajectory $\underline{x}_{(\gamma)}$ corresponding to a certain value of the parameter vector $\underline{p}_{(\gamma)}$, considered as the best value known by us at this stage, and the nominal control $\underline{u}(t)$. In this case we infer that, $\underline{\eta}(i)$ and $\underline{\nu}(i)$ are equivalent:

$$\underline{\eta}(i) = \underline{y}_{(n)}(i) - \underline{y}_{(\gamma)}(i) \triangleq \underline{\nu}_{(\gamma)}(i) \tag{17}$$

$\underline{y}_{(\gamma)}$ being the current best estimate of $\underline{y}_{(n)}$. The likelihood function then is:

$$L\left(\underline{p}_{(\gamma)}, R_{(\gamma)}\right) = -\frac{1}{2}\sum_{i=1}^{k}\left\{\underline{\eta}^T(i)R^{-1}_{(\gamma)}(i)\underline{\eta}(i) + \ln|R_{(\gamma)}(i)|\right\} \tag{18}$$

Suppose now that being at the stage (γ), we ask for an improved value of the parameter vector $\underline{p}_{(\gamma+1)} = \underline{p}_{(\gamma)} + \Delta\underline{p}_{(\gamma)}$, ($\Delta\underline{p}_{(\gamma)}$ being a small perturbation) in order to maximize the likelihood function. Accordingly, a new trajectory $\underline{x}_{(\gamma+1)}$ is generated, corresponding to the new parameter vector $\underline{p}_{(\gamma+1)}(t)$, $\underline{x}_{(\gamma+1)}$ being close to $\underline{x}_{(\gamma)}(t)$ in a first order closeness sense. Let us now define:

$$\underline{\Delta x}(t) \triangleq \underline{x}_{(\gamma+1)}(t) - \underline{x}_{(\gamma)}(t) = S_{(\gamma)}(t) \cdot \underline{\Delta p}_{(\gamma)} \tag{19}$$

assuming that the small trajectory deviation $\underline{\Delta x}(t)$ can be represented as a linear transformation with respect to $\underline{\Delta p}_{(\gamma)}$. As in eqn. (17) we can write the following:

$$\begin{aligned}\underline{\nu}_{(\gamma+1)}(i) &= \underline{y}_{(n)}(i) - \underline{y}_{(\gamma+1)}(i)\\ &= \underline{y}_{(n)}(i) - \underline{y}_{(\gamma)}(i) - CS_{(\gamma)}(i)\underline{\Delta p}_{(\gamma)}\\ &= \underline{\eta}(i) - CS_{(\gamma)}(i)\underline{\Delta p}_{(\gamma)}\end{aligned} \tag{20}$$

We will call the (nxm) matrix $S_{(\gamma)}$, the sensitivity matrix, in the sense defined by Sidar (1968) and Larson (1968) and we will show in the sequel how this matrix is computed.

The value of the likelihood function is actually a function of $\Delta\underline{p}$ and R:

$$L(\underline{p} + \underline{\Delta p}, R) = -\frac{1}{2} \sum_{i=1}^{k} \left\{ \left[\underline{\eta}(i) - CS_{(\gamma)}(i)\underline{\Delta p}_{(\gamma)} \right]^T R^{-1}(i) \left[\underline{\eta}(i) - CS_{(\gamma)}(i)\Delta \underline{p}_{(\gamma)} \right] + \ln|R(i)| \right\} \tag{21}$$

In order to maximize this function with respect to $\Delta \underline{p}$ and R, we shall calculate the partial derivatives of $L(\underline{p},R)$ with respect to the vector $\Delta \underline{p}$ and with respect to the matrix $R(i)$. Setting those derivatives equal to zero, one obtains, after a few algebraic manipulations, the following necessary conditions for maximization of the likelihood function:

$$\sum_{i=1}^{k} \left\{ S^T(i)C^T\hat{R}^{-1}(i)CS(i) \right\} \underline{\Delta p}^*_{(\gamma)} = \sum_{i=1}^{k} \left[S^T(i)C^T\hat{R}^{-1}(i)\eta(i) \right] \tag{22}$$

and

$$\sum_{i=1}^{k} \hat{R}(i) = \sum_{i=1}^{k} \left\{ \left[\underline{\eta}(i) - CS(i)\Delta \underline{p}^*_{(\gamma)} \right] \left[\underline{\eta}(i) - CS(i)\Delta \underline{p}^*_{(\gamma)} \right]^T \right\} \tag{23}$$

Those two coupled equations allow us, in principle, to establish a batch-iterative procedure, in order to compute $\Delta \underline{p}^*_{(\gamma)}$ and $\hat{R}(i)$ (Grove, Bowles, and Mayhew, 1972 and Aubrun, 1971). Computationally, this can be done only by assuming some simplifications as did Gertler and Banyasz (1974) and Grove, Bowles, and Mayhew (1972); otherwise one has to deal with very laborious manipulations, including some intermediate iterative algorithms. Since we are not interested here in obtaining an explicit form

for the identification algorithm in the batch-formulated case, we will proceed further, in order to obtain the recursive version of the maximum likelihood identification algorithm.

4. The recursive identification algorithm

In this section, the recursive version of the identification algorithm is derived from the previous necessary conditions (22, 23). In particular, the optimal parameter vector increment would be obtained from eqn. (22) and we shall assume in the sequel that: $k \gg 1$ and that the sampling rate is high enough, that is, ΔT is small in comparison with the process dynamics. From eqn. (22), defining $\Delta \underline{p}_k \triangleq \Delta \underline{p}_{(\gamma)}$, we note that for the interval $[0,k]$ one has:

$$\sum_{i=1}^{k-1} \left\{ S^T(i) C^T \hat{R}^{-1}(i) C S(i) \right\} \underline{\Delta p}_k^* + \left\{ S^T(k) C^T \hat{R}^{-1}(k) C S(k) \right\} \underline{\Delta p}_k^*$$

$$= \sum_{i=1}^{k-1} \left[S^T(i) C^T \hat{R}^{-1}(i) \underline{\eta}(i) \right] + \left[S^T(k) C^T \hat{R}^{-1}(k) \underline{\eta}(k) \right] \tag{24a}$$

Since eqn. (22) is the m.l.e. necessary condition, valid for all k, it certainly holds as well for the interval $[0,k-1]$, so that:

$$\sum_{i=1}^{k-1} \left\{ S^T(i) C^T \hat{R}^{-1}(i) C S(i) \right\} \underline{\Delta p}_{k-1}^* = \sum_{i=1}^{k-1} \left[S^T(i) C^T \hat{R}^{-1}(i) \underline{\eta}(i) \right] \tag{24b}$$

To establish a recursive algorithm we ask for the following condition:

$$\underline{\Delta p}_k^* = \begin{cases} \underline{\Delta p}_{k-1}^* & \text{for } [0,k-1] \\ \underline{\delta p}_k^* & \text{for } [k-1,k] \end{cases} \tag{25}$$

Conditions (25) and eqn. (24b) imply that (24a) has the form:

$$\sum_{i=1}^{k-1} \left[S^T(i)C^T\hat{R}^{-1}(i)CS(i)\right] \Delta\underline{p}_{k-1}^* + \left[S^T(k)C^T\hat{R}^{-1}(k)CS(k)\right] \delta\underline{p}_k^*$$

$$= \sum_{i=1}^{k-1} \left[S^T(i)C^T\hat{R}^{-1}(i)CS(i)\right] \Delta\underline{p}_{k-1}^* + \left[S^T(k)C^T\hat{R}^{-1}(k)\underline{\eta}(k)\right] \tag{24c}$$

From (24c) we obtain the change $\delta\underline{p}^*(k)$ in $\underline{p}(k)$ for the interval $[k-1,k]$ needed to identify and perhaps track the parameter vector:

$$\delta\underline{p}_k^* = \left[S^T(k)C^T\hat{R}^{-1}(k)CS(k)\right]^{-1} \cdot \left[S^t(k)C^T\hat{R}^{-1}(k)\ \underline{\eta}(k)\right] \tag{26}$$

or, in a recursive form, the following identification/tracking algorithm is obtained:

$$\underline{p}(k) = \underline{p}(k-1) + [A]^{-1}[B]\underline{\eta}(k) \tag{27}$$

where:

$$[B] = [B(k)] \triangleq S^T(k)C^T\hat{R}^{-1}(k)$$

is an (mxq) matrix and:

$$[A] = [A(k)] \triangleq S^T(k)C^T\hat{R}^{-1}(k)CS(k)$$

- an (mxn) matrix, is the incremental Fisher information matrix for the unknown parameter vector.

The inverse covariance matrix $\hat{R}^{-1}(k)$ can not be obtained directly from eqn. (23). Therefore, we will first derive a recursive equation for $\hat{R}(k)$. From (23), we note:

$$\sum_{i=1}^{k} \hat{R}(i) = \sum_{i=1}^{k-1} \hat{R}(i) + \underline{\eta}(k)\underline{\eta}(k)^T \tag{28a}$$

Defining the matrix $\hat{R}_o(k) \triangleq \frac{1}{k} \sum_{i=1}^{k} \hat{R}(i)$, we write, from (28a):

$$k\hat{R}_o(k) = (k-1)\hat{R}_o(k-1) + \underline{\eta}(k)\underline{\eta}(k)^T \tag{28b}$$

and finally:

$$\hat{R}_o(k) = \frac{k-1}{k} \hat{R}_o(k-1) + \frac{1}{k} \underline{\eta}(k)\underline{\eta}^T(k) \tag{28c}$$

This recursive equation is of utmost interest because in the steady state $(k >> 1)$ of the recursive identification, assuming satisfactory parameter tracking, it allows us to identify the covariances of the instrumentation noises. That is so because for large enough values of k, we have already obtained a good estimate for the parameter $\underline{p}$ which is very close to $\underline{p}_{(n)}$, so that:

$$\lim_{k>>1} \hat{R}_o(k) = \hat{R}(k) = R_1$$

(see eqn. (18))

In all the cases we studied, this result was verified, a verification that represents by itself a good test of and criterion for the accuracy of the identification process.

Since eqn. (27) needs $\hat{R}^{-1}(k)$, one is interested in the recursive equation of the inverse covariance matrix. This matrix can be obtained either directly, by inversion of $\hat{R}_o(k)$ obtained from the recursive algorithm (28c), or by making use of the matrix inversion lemma. Assuming steady state conditions, we obtain the following recursive algorithm:

$$\hat{R}^{-1}(k) = \frac{k}{k-1}\hat{R}^{-1}(k-1) - \frac{k}{(k-1)^2}\hat{R}^{-1}(k-1)\underline{\eta}\underline{\eta}^T\hat{R}^{-1}(k-1) \qquad (29)$$

It is clear from eqn. (27) that the existence (for all k) of the inverse of the incremental information matrix, $[A(k)]^{-1}$, is necessary for the implementation of the recursive identification algorithm. The computation of $[A(k)]^{-1}$ is by itself an excellent test for identifiability. Its existence constitutes a good criterion for detection ill-posed identification problems (see Aubrun, 1971 and Glover, 1973). We therefore assume from now on that the inverse matrix $[A(k)]^{-1}$ exists, such that the identifiability conditions are met.

Note that there is a formal similarity between (27) and other linear least-square gradient-type iterative algorithms, although our algorithms (27) and (29) are different and able to handle non-linear systems in recursive identification problems.

For the start-up procedure of the recursive algorithm, we are assuming that <u>a priori</u> estimates $\underline{p}(0)$ of the parameters exist,

obtained from early measurements, or from such experiments as wind tunnel measurements in the case of stability and control derivatives for aircraft. Beginning with those initial values, the following sequence of computations is done for each interval [k-1,k]:

1. $y_{(n)}(k)$ is obtained either directly from real experiments (on-line or off-line) or, in the case of simulation studies, from the integration of the differential equations of the nominal systems (7a and 8a),

2. in order to carry out the computation of $\underline{p}(k)$, the computation of $S(k)$ is necessary, following the procedure shown in the next section,

3. $\underline{y}(k)$ is calculated from eqns. (7b and 8b), the value of the parameter vector being $\underline{p}(k-1)$. The vector $\underline{\eta}(k)$ is calculated from (10) and is used in eqns. (27,29) in order to compute the new value of the parameter vector $\underline{p}(k)$.

4. $\underline{p}(k)$ is calculated from eqns. (27) and (29), which are the basic recursive equations for the parameter identification and tracking, with $\underline{p}(0)$ as the initial value (the best available guess on an a priori knowledge basis). As will be shown later, convergence of the identification algorithm can be guaranteed under the hypothesis that $\underline{p}(0)$ is close enough to $\underline{p}^*$ (although the convergence region in the parameter space is quite large).

The schematic flow diagram of the computations necessary for implementation of the recursive identification algorithm is shown in figure 2.

From eqn. (26), it is easy to prove that for $k>>1$ and given $E[\underline{n}(k)] = 0$, one obtains: $E[\delta p_k^*] = 0$. This result shows that the parameter vector $\underline{p}(k)$, identified by the algorithm (26), is unbiased. Furthermore, as we mentioned before, one observed from eqn. (28c) that the iterated covariance matrix $\hat{R}(k)$ tends to the noise covariance matrix R_i as k increases. Those are two valuable results obtained as a direct consequence of using the maximum likelihood estimation approach.

5. The sensitivity matrix

As we noted before, use of the recursive algorithm (27) requires the computation of the sensitivity matrix $S(t)$. The $(n.m)$ elements of $S(t)$, the sensitivity functions $\partial x_i/\partial p_j$, $\forall\ i=1, \ldots, n$ and $j=1, \ldots, m$, are by themselves solutions of (nxm) differential equations. Those differential equations are obtained from the system's dynamic equations (7b), assuming the current value of the parameter vector by taking the derivatives of the equations with respect to each parameter p_j:

$$\frac{\partial\left(\frac{dx_i}{dt}\right)}{\partial p_j} = \sum_{\beta=1}^{n} \frac{\partial f_i}{\partial x_\beta} \cdot \frac{\partial x_\beta}{\partial p_j} + \frac{\partial f_i}{\partial p_j} \tag{30}$$

Since eqn. (30) deals with a combined limiting process with respect to dt and ∂p_j (both tend to zero), and since we already assumed the continuity of the solutions with respect to $\underline{p}$ and t, we may invert the order of differentation (see Sidar, 1968, Larson, 1968), obtaining:

$$\frac{d}{dt}\left(\frac{\partial x_i}{\partial p_j}\right) = \sum_{\beta=1}^{n} \frac{\partial f_i}{\partial x_\beta} \cdot \left(\frac{\partial x_\beta}{\partial p_j}\right) + \frac{\partial f_i}{\partial p_j} \tag{31}$$

Denoting $\partial x_i/\partial p_j \triangleq s_{ij}$ as the sensitivity function, we have:

$$\dot{s}_{ij} = \sum_{\beta=1}^{n} \frac{\partial f_i}{\partial x_\beta} \cdot s_{\beta j} + \frac{\partial f_i}{\partial p_j} \tag{32}$$

or in a compact matrix form:

$$\dot{S}(t) = F_x S + F_p \tag{33}$$

where:

$$S \triangleq \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1m} \\ \vdots & & & \\ s_{n1} & s_{n2} & \cdots & s_{nm} \end{bmatrix} \tag{34}$$

F_x, the system's Jacobian (for the current value of $\underline{p}$) and F_p, are defined below:

$$F_x \triangleq \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}_{(nxn)} \qquad F_p \triangleq \begin{bmatrix} \frac{\partial f_1}{\partial p_1} & \cdots & \frac{\partial f_1}{\partial p_m} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial p_1} & & \frac{\partial f_n}{\partial p_m} \end{bmatrix}_{(nxm)} \tag{35}$$

Therefore, to obtain the matrix S(i) it is necessary to integrate (n.m) additional differential equations, an operation that demands a certain computing effort, but does not represent a difficult task for a high speed digital computer. Gupta and Mehra (1974) and Denery (1971a) have shown that, for linear systems, some savings in the computational effort of solving eqns. (33), are possible.

The initial conditions for the integration of the differential system (33) are $s_{ij} = 0$.

6. Conditions for convergence

The best available policy of changing $\underline{p}(k-1)$ in order to track, in an optimal way, the trajectory $\underline{x}_{(n)}(k)$ and the measured output vector $\underline{y}(k)$, is given by applying the recursive algorithm (27). In the parameter space p^m, this is equivalent to the choice of the best direction in order to converge sequentially to $\underline{p}_{(n)}(t)$. The size of the step $\underline{\delta p}_k$, in the optimal direction, has to be chosen in a way such that convergence can be guaranteed. That means that the trajectory tracking error (discussed below) has to

be, at least, a non-increasing function. The fundamental problem of choosing the optimal step in nonlinear programming iterative schemes has been widely treated in the literature (see, for instance, Luenberger, 1972, Bard, 1970, and Mehra and Gupta, 1974) and we will not treat it again here.

We shall rather present a specific criterion for convergence based on Lyapunov's second method. Somewhat similar results have also been obtained by us, as shown later, by developing an alternative criterion based on the location of the eigenvalues of the output vector tracking difference equation. A successful application of Lyapunov's direct method in order to obtain stable adaptive schemes for system identification is due to Narendra and Kudva (1974). Mendel (1968) studied the stability properties of a parameter vector difference equation by applying Lyapunov stability theory. Here we deal with the stability and convergence properties of the state vector tracking error difference equation. The stability study of this tracking error sequence by the means of Lyapunov's method reveals it to be a very valuable tool for choosing the scalar $\rho(k)$ in the modified identification algorithm:

$$\underline{p}(k) = \underline{p}(k-1) + \rho(k)[A(k)]^{-1}[B(k)]\underline{\eta}(k) \tag{36}$$

Let us define the following positive definite function:

$$V_k = \underline{\delta x}^T(k) \cdot \underline{\delta x}(k) \tag{37}$$

where:

$$\underline{\delta x}(k) \triangleq \underline{x}_{(n)}(k) - \underline{x}(k)$$

as a possible Lyapunov function. For V_k to be a Lyapunov function for the recursive iterative algorithm, the necessary condition is:

$$\Delta V_k \triangleq \underline{\delta x}^T(k+1) \cdot \underline{\delta x}(k+1) - \underline{\delta x}^T(k)\underline{\delta x}(k) \leq 0 \tag{38}$$

The quantity V_k can be easily calculated from (7,8,9, and 36). By neglecting higher-order infinitesimal terms containing ΔT^2 (as ΔT tends to zero), and assuming a noise-free case, we finally obtain:

$$\Delta V_k = - \Delta T \; \underline{\delta x}^T(k) \; M(\rho) \; \underline{\delta x}(k) \tag{39}$$

with $M(\rho) = P + \rho Q$, P and Q being two symmetric (nxn) matrices:

$$P \triangleq F_x + F_x^T \tag{40a}$$

$$Q \triangleq F_p \; A^{-1} \; BC + \left[F_p \; A^{-1} \; BC\right]^T \tag{40b}$$

The proper choice of the iterative gain ρ, has to guarantee that $M(\rho)$ is a positive definite matrix, such that V_k is indeed a Lyapunov function for the identification algorithm and the iterative process is globally asympotically stable. Making use of Sylvester's Theorem, on positive definitness test, the choce of ρ and be made either on a per step basis, or globally prior to the

identification test. The second approach is much easier to apply because it is based on an a priori data base, avoiding unnecessary additional computations.

The stability analysis is based on two assumptions: (1) we are assuming fixed or slowly changing parameters and (2) as mentioned earlier, zero output noise. The impact of the noise covariance R_1 on ρ is presently under study by us, but is not discussed here.

As mentioned before, an alternative way to insure convergence of the identification algorithm is by studying the stability of the following difference equation, obtained from (7, 8, 9, and 36):

$$\underline{\delta x}(k+1) = [I - \Delta T \cdot N(\rho)]\ \underline{\delta x}(k) = \mathcal{N}(\rho)\ \underline{\delta x}(k) \tag{41}$$

where:

$$N(\rho) \triangleq [F + \rho F_p A^{-1} BC] \tag{42}$$

is a real, square (nxn) matrix. The scalar ρ is to be chosen such that $\mathcal{N}(\rho)$ (41) has all its eigenvalues inside the unit circle. A theorem in Jury (1974) provides the necessary and sufficient conditions for stability for the system (41,42). The proper value of the scalar ρ can be calculated from these conditions. Again, as mentioned before, ρ can be determined for each interval or, prior to the identification test, in a very simple way, avoiding cumbersome computations.

Applying either the Lyapunov stability approach or the stability criterion for the difference eqn. (41), we have been able to obtain an algorithm which: (1) is stable, (2) provides high accurate identification, and (3) rapidly tracks changing parameters.

7. Computational results

In order to illustrate the utilization of the recursive algorithm and the results which have been obtained, although taking into account the lack of space, two examples among other cases studied, will be presented.

Example 1:

The identification and tracking of two unknown parameters in a first order, nonlinear system:

$$\dot{x} + ax + bx^3 = u(t)$$

$$y = x + v \tag{43}$$

with:

$$x(o) = 0.5 \quad , \quad a_{nom} = 0.4 \quad , \quad b_{nom} = 0.2 \tag{43}$$

Although this is a relatively simple case, the purpose is to allow the reader to better evaluate the utilization of our algorithm. Twelve different cases were analyzed, including the use of various inputs and various combinations of variable parameters. We identified a and b assuming that:

1. both a_{nom} and b_{nom} are unknown, but constant parameters.
2. a_{nom} varies, but b_{nom} remains constant.
3. a_{nom} is constant but b_{nom} changes.
4. both a_{nom} and b_{nom} are changing, simultaneously.

In each case the identification was performed with three different input functions: (1) $u(t) = 2.0$ (step); (2) $u(t) = 2 + \sin t$; and (3) $u(t) = 2 + n(t)$, $n(t)$ being a white, Gaussian noise input with zero mean and covariance 0.4. The covariance of the output noise was 0.04.

The identification period was 20 sec (real time). The results obtained were excellent from all points of view: (1) convergence, (2) accuracy of unbiased estimates, (3) tracking and, (4) insensitivity to initial guesses for a and b.

In figure 3, as an illustration, the recursive identification and tracking of b is shown for the case where a_{nom} is constant but b_{nom} changes and $u(t) = 2 + \sin t$. Note the simultaneous identification of the noise covariance $R_1(K)$.

Example 2

The recursive (on-line) identification (and tracking) of the stability derivatives of the longitudinal dynamics typical of aircraft. In this case, the nominal system is the third-order linear system, represented by the transfer function:

$$\frac{\theta(S)}{\delta(S)} = \frac{2.857(S + 0.4)}{S(S^2 + 0.685S + 0.53)} \tag{44}$$

where $\theta(t)$ is the aircraft's pitch angle (deg); $\delta(t)$ is the elevator angle (deg); the longitudinal natural frequency is $\omega_o = 0.73$ rad/sec, and the damping ratio is $\xi = 0.47$. In state variables notation, similar to the notation of (7) and (8), we have:

$$\dot{\theta} = \underline{q} \tag{45a}$$

$$\dot{q} = M_q q + M_w U\alpha + M_{\delta_e} \delta_e \tag{45b}$$

$$\dot{\alpha} = q + Z_w \alpha + \frac{Z_{\delta_e}}{U} \cdot \delta_e \tag{45c}$$

We note that $X^T(t) \triangleq [\theta, q, \alpha]$ is the state vector for this example. The numerical values of the coefficients are: $U = 215$ ft/sec, $M_{q_{nom}} = 0.2778$, $Z_{w_{nom}} = 0.4075$, $M_{\delta_e} = -2.857$, $M_w = 0.0019$, and $Z_{\delta_e} = -13.24$.

For identification purposes, one measures the pitch rate q with a rate gyro and the angle of attack α with an angle of attack sensor (vane). Both the rate gyro and the vane have measurement noise assumed to be a Gaussian white noise with zero mean and covariances of 0.1 and 0.2, respectively.

The identification interval was 22 sec with 100 discrete intervals per second. For the sake of simplicity and clarity we show only the case where M_q and Z_w are to be identified. Twelve different combinations were analyzed:

1. both M_q and Z_w are unknown but constant
2. M_q is variable, Z_w is constnat

3. M_q is constant, Z_w is variable

4. both M_q and Z_w are varying simultaneously

Parameter identification and tracking was done for three different inputs:

1. $\delta(t) = -1.0$ deg. (step at $t = 0$)

2. $\delta(t) = -1.0 - 0.25 \sin 0.3t - 0.1 \sin 0.9t - 0.02 \sin 1.5t$

This input was considered as the optimal input for batch-identification and calculated according to Mehra (1973).

3. $\delta(t) = -1.0 + n(t)$

$n(t)$ being a white, zero mean, Gaussian noise with covariance equal to 0.5.

Because it would be impractical to show all our results here, only those obtained for recursive identification and tracking in three cases are shown in Figs. 4, 5, and 6. Inputs (1), (2), and (3) above, each in conjunction with the combination that has M_q and Z_w varying simultaneously, make up the three cases. No sensible differences are observed with respect to the various input functions. It is worth remarking that the identified values of M_q and Z_w are unbiased. Furthermore, the figures show how accurately the algorithm identifies and tracks even the varying nominal parameter, after a short transient.

8. Conclusions

An iterative recursive algorithm for parameter identification and tracking, based on the maximum likelihood approach, was developed. This algorithm allows sequential parameter and instrument noise identification in both linear and nonlinear systems. The estimates are shown to be unbiased and accurate and the results already obtained in several cases show a good ability to track variations of systems parameters. The conditions for a stable-iterative process are analyzed, leading to the choice of the scalar iterative gain ρ.

Establishing bounds on the value of ρ or determining relationships between it (or perhaps a gain matrix G) and the noise covariance R, is a topic for further research. Also, it may be possible, on-line, to generate input, or adequate probing sequences, functions that would maximize identification accuracy, perhaps exploiting the theory developed by Lopez-Toledo and Athans (1974) for linear systems. Maximizing the information matrix may provide necessary conditions for generating such inputs for the identification, even of nonlinear systems.

Acknowledgment

I wish to thank Mr. M. Nordstrom for his valuable help in programming the digital simulation and Mr. B. F. Doolin for his valuable comments and review.

The author is also grateful to the National Research Council for a postdoctoral fellowship.

REFERENCES

Aström, K. J. and Eykhoff, P., 1971 Automatica, Vol. 7.

Aubrun, J. N., 1971, NASA TN D-6467.

Y. Bard, 1970, Parameter Estimation Problems, SIAM J. Numer. Anal., Vol. 7.

Denery, D., 1971a, IEEE Trans. on AC.

Denery, D., 1971b, NASA TN D-6468.

Durand, T. and Teper, G., 1968, Tech. Rep. Systems Technology, Inc., TR-137-1.

Gertler, J. and Banyasz, C., 1974 IEEE Trans. on AC, Vol. 19, No. 6.

Gevers, M. and Kailath, T., 1973 IEEE Trans. on AC, Vol. AC-18.

Glover, K., 1973, Rep. ESL-R-516, MIT.

Grove, R. D., Bowles, R. L., and Mayhew, S. C., 1972, NASA TN D-6735.

Gupta, N. K., and Mehra, R. K., 1974, IEEE Trans. on AC, Vol. AC-19, No. 6.

Hastings, R., and Sage, James and M. W., 1969, Proc. IEEE, 116.

IEEE Trans. on A. C., Dec. 1974, Special Issue on System Identification: Mehra pp. 753, Aström and Soderstrom, pp. 769.

Jury, E. J., 1974, Inners and Stability of Dynamical Systems, Wiley-Interscience, pp. 300-301.

Kailath, T., 1969, IEEE Trans. Info. Theory, Vol. IT-15.

Kashyap, R. L., 1970, IEEE Trans. on AC, Vol. AC-15.

Landau, I. D., 1974, Automatica, Vol. 10, No. 4.

Larson, R., 1968, CAL Rep. No. 164.

Luenberger, D. G., 1972, "Introduction to Linear and Nonlinear Programming," Addison-Wesley.

Mehra, R. K., 1970, JACC 1970, Preprints, Paper 18-C.

Mehra, R. K., 1973, Tech. Rep. TR-645, Harvard University.

Mendel, J. M., 1968, Inform. Sciences, Vol. 1.

Narendra, K. S. and Kudva, P., 1974, IEEE Trans. Vol. SMC-4, No. 6.

Rogers, A. E. and Steiglitz, K., 1967, Proc. of JACC.

Sidar, M., 1968, Int. J. Non-Linear Mechanics, Vol. 3, pp. 1-16.

Stepner, D. and Mehra, R. K., 1973, NASA CR-2200.

Lopez-Toledo, A. A. and Athans, M., 1974, Paper No. THA-4.2, Preprints, IEEE Conf. on Dec. and Control.

Wingrove, R., 1974, AGARD meeting on Identification, NASA Langley.

PRECEDING PAGE BLANK NOT FILMED

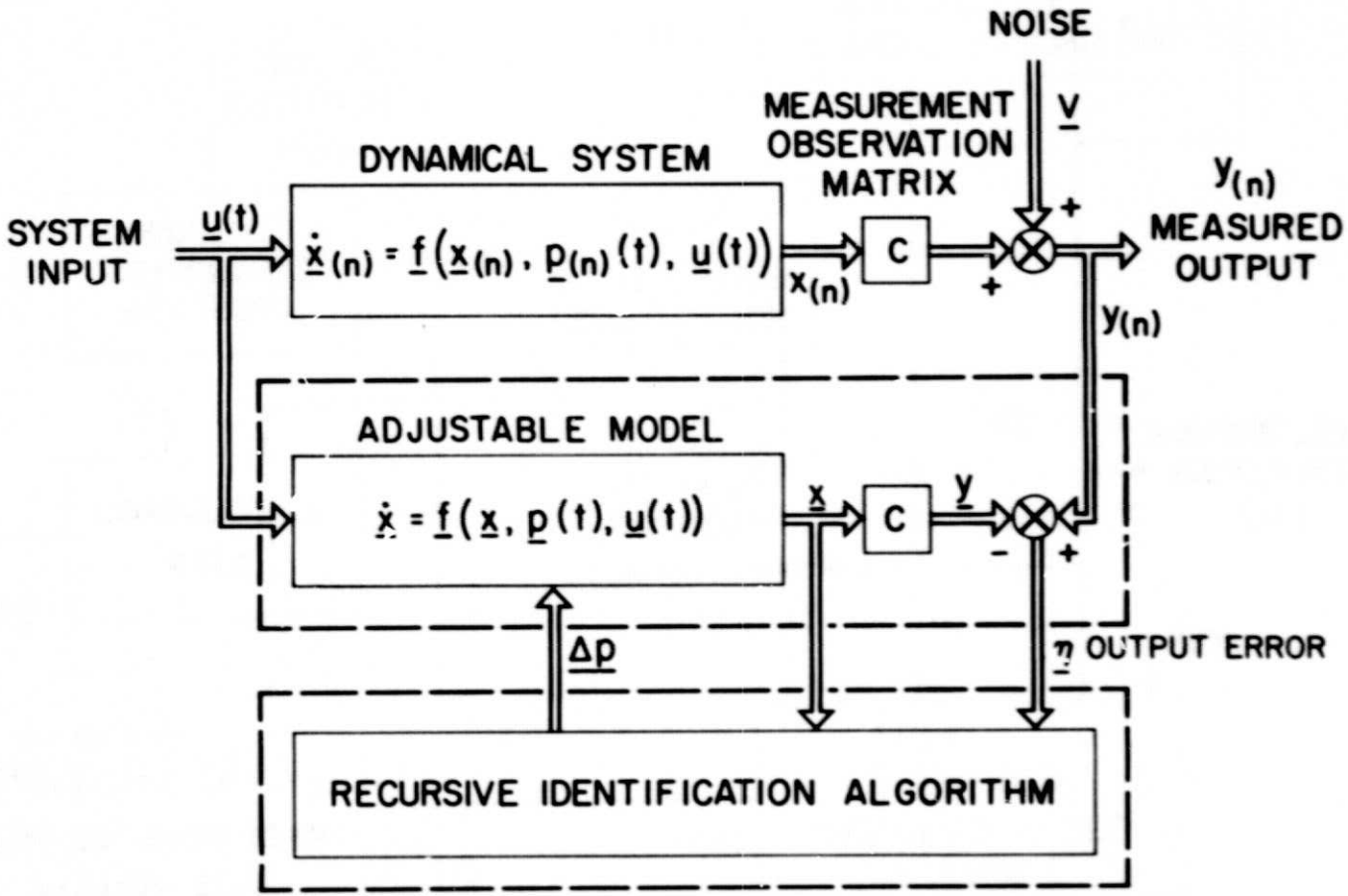

Figure 1. System identification—conceptual design.

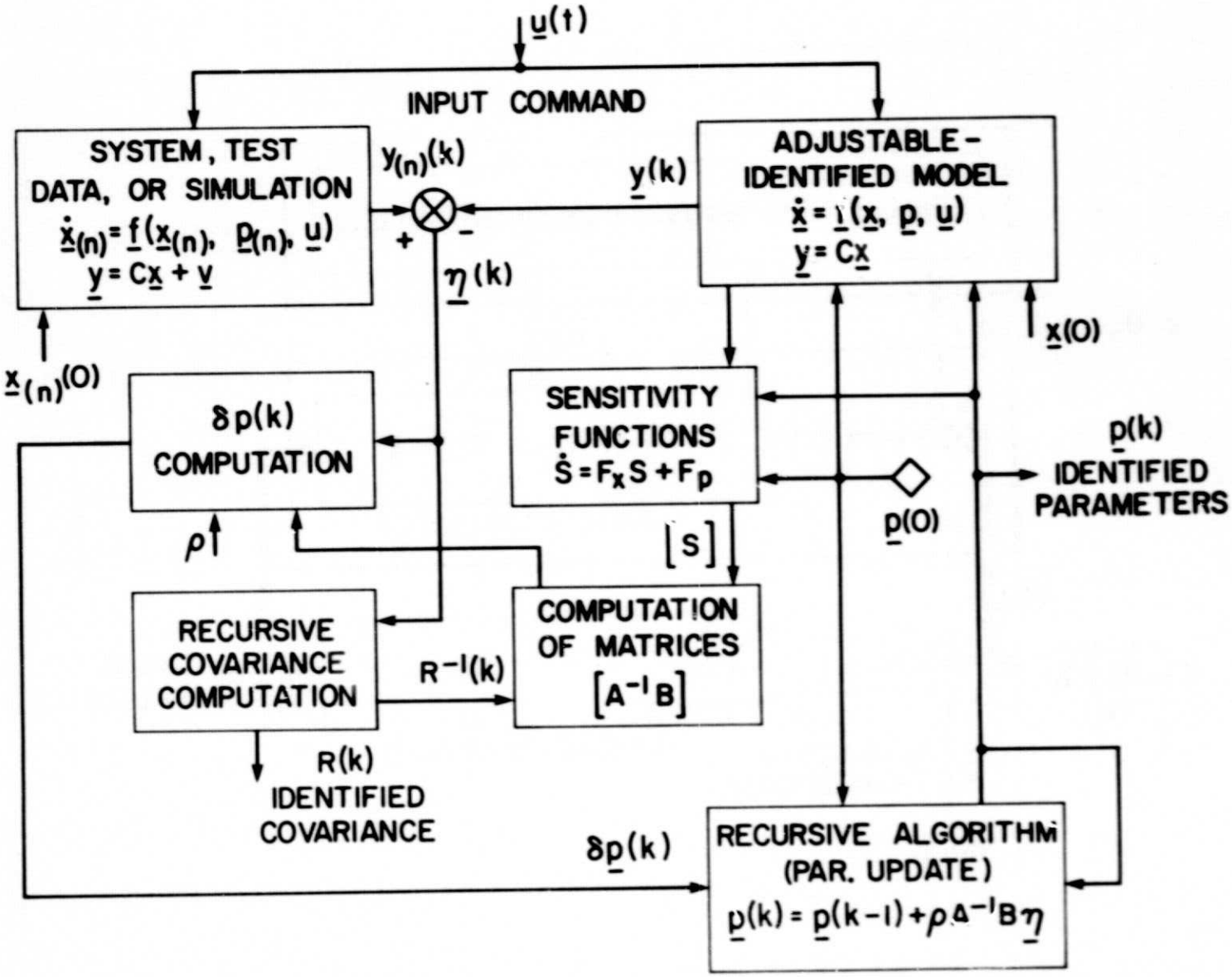

Figure 2. Simplified computational flow diagram.

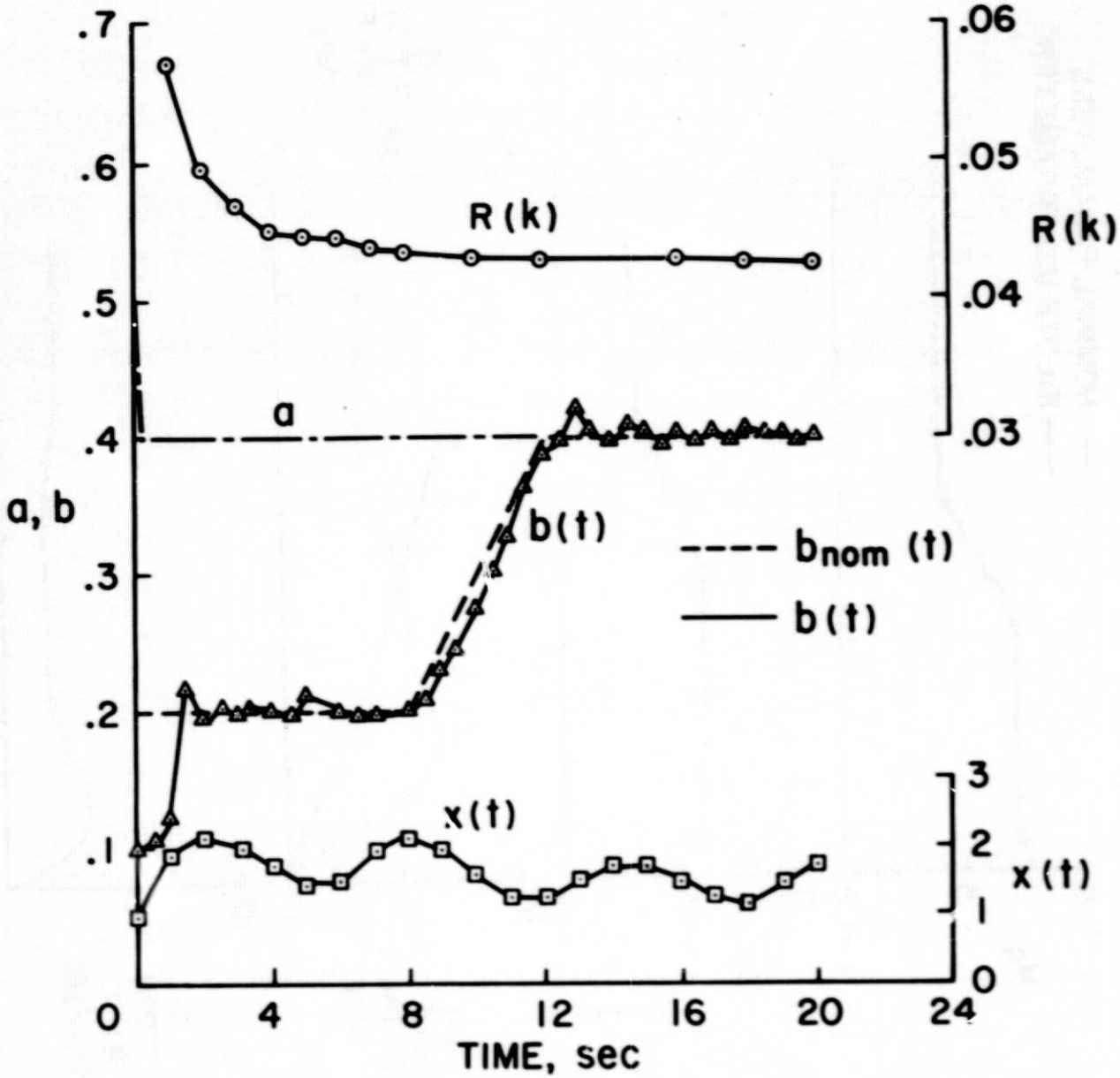

Figure 3. Example 1: identification of a and b(t).

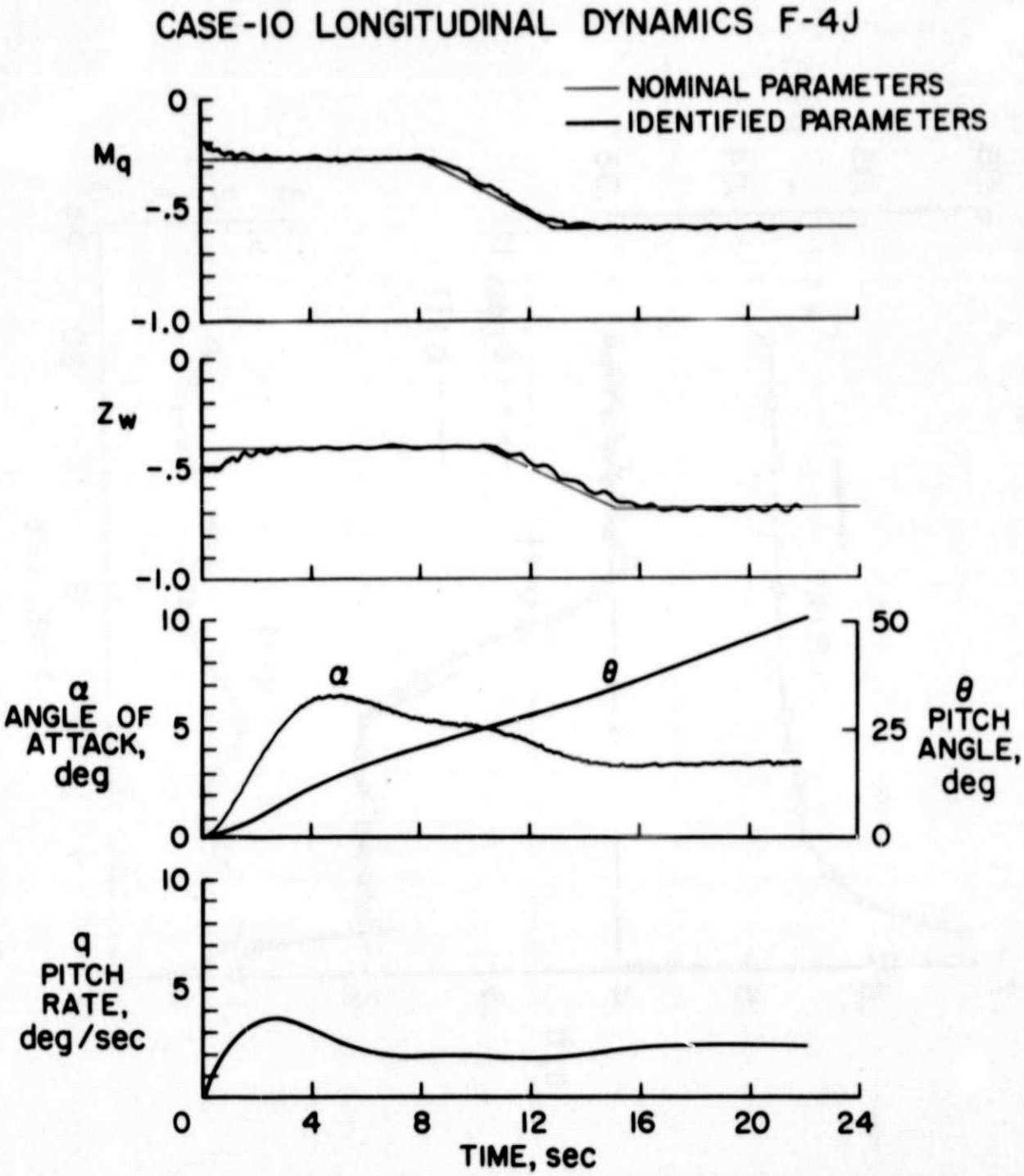

Figure 4. Example 2: Identification of the stability derivatives with input $\delta(t) = -1.0$.

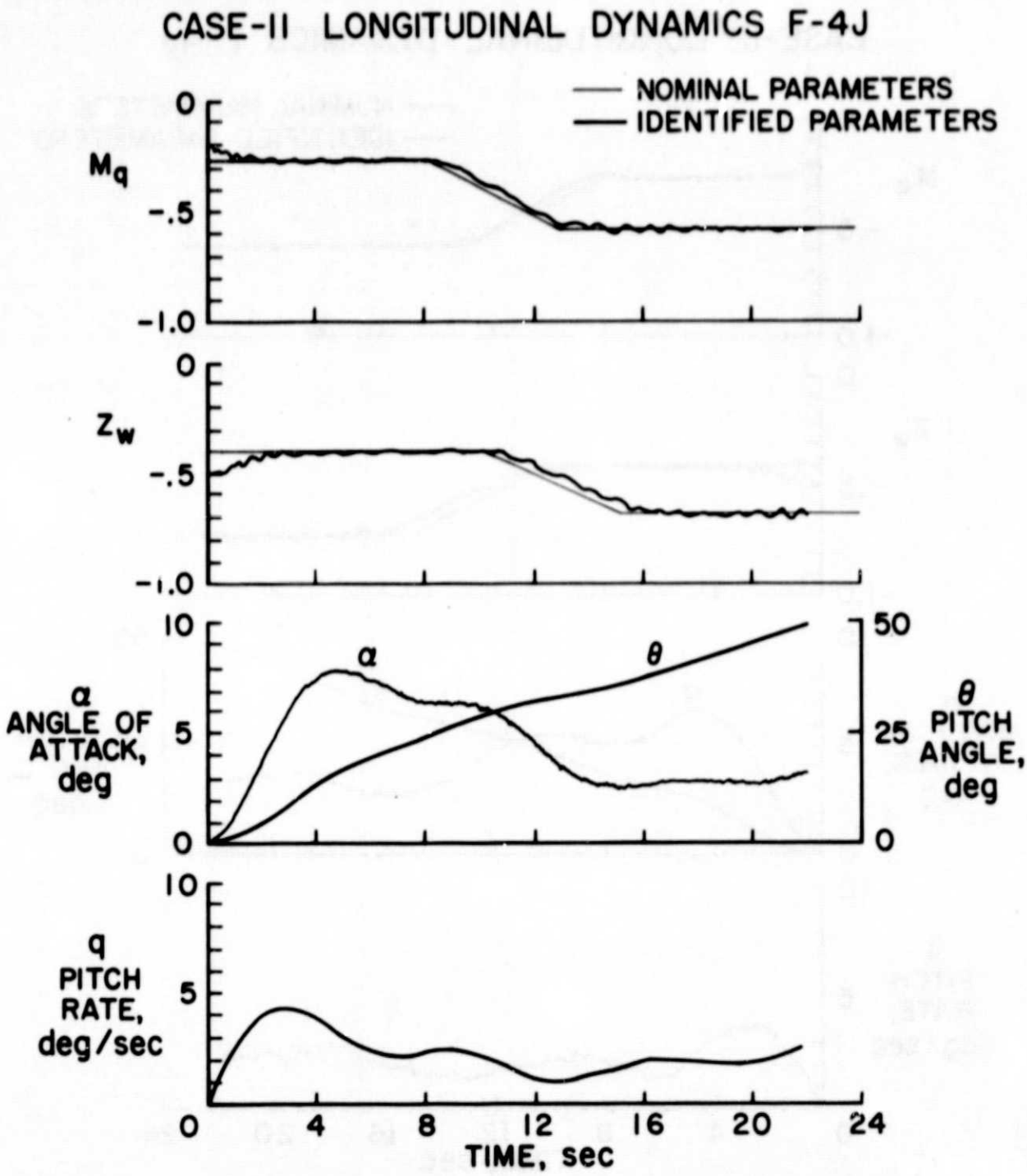

Figure 5. Example 2: Identification of the stability derivatives with input $\delta(t) = -1.0 - 0.25 \sin 0.3t - 0.1 \sin 0.9t - 0.02 \sin 1.5t$.

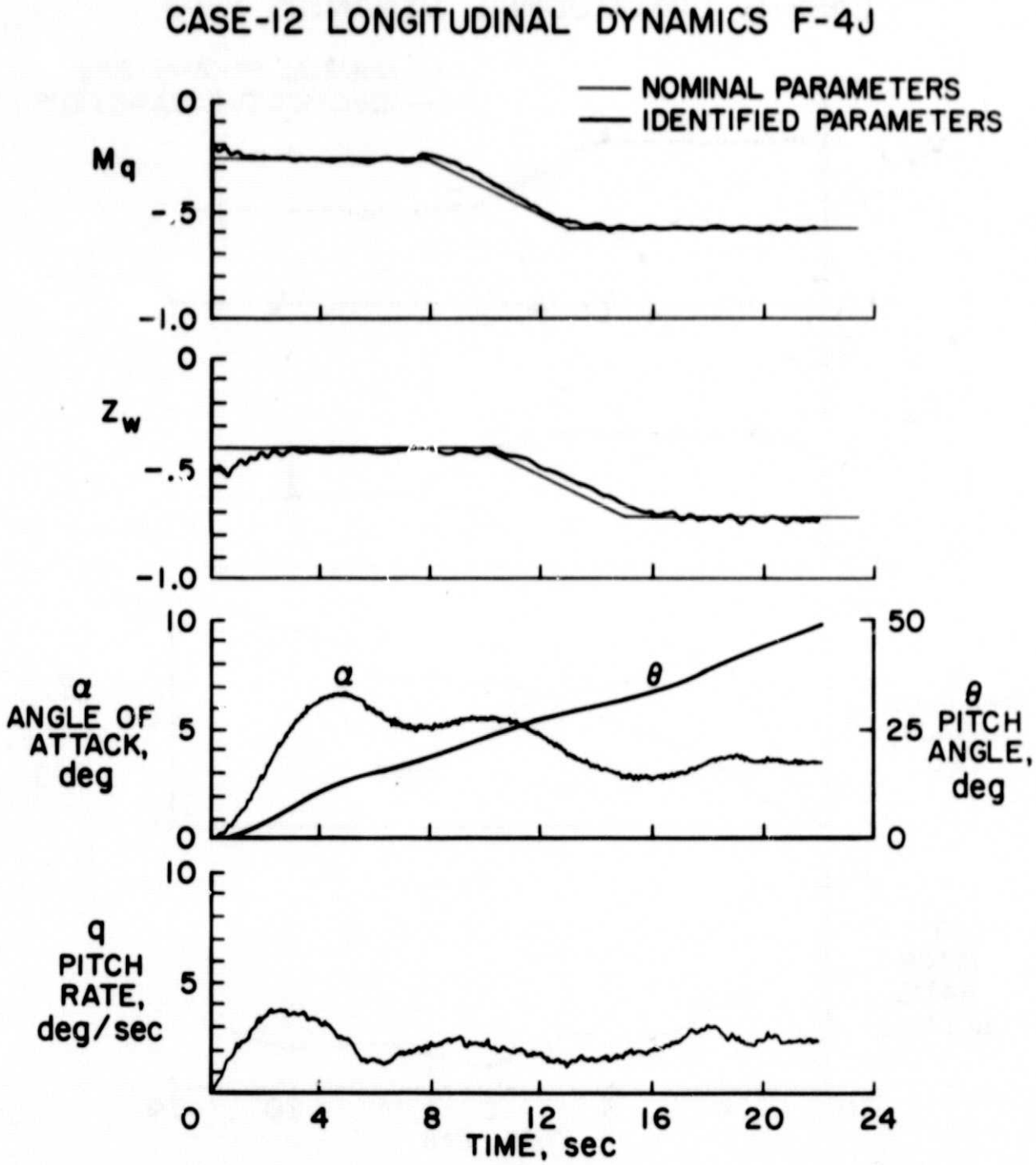

Figure 6. Example 2: Identification of the stability derivatives with input $\delta(t) = -1.0 + n(t)$.