RICE UNIVERSITY

NONLINEAR SYSTEM IDENTIFICATION

by

Arun N. Netravali

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

Thesis Director's Signature:

Rui J. P. de Figueiredo

Houston, Texas

February 1969

ABSTRACT

NONLINEAR SYSTEM IDENTIFICATION

Arun N. Netravali

Stochastic approximation methods for the identification of parameters of nonlinear systems without dynamics have been widely discussed in the literature (see e.g., [1], [2]).

In this work, two classes of discrete-time nonlinear dynamical systems driven by independent noise are considered. The measurements are assumed to be linear scalar and are made in the presence of independent noise. The systems under consideration are identified by the estimation of the parameters appearing in the evolution operator. These parameters are assumed to be constant during the identification time and they are estimated by means of stochastic approximation algorithms.

A computer algorithm based on the above method is used to identify the parameters for some typical examples.

ACKNOWLEDGMENTS

The author wishes to express his gratitude to his advisor, Professor Rui J.P. de Figueiredo for his encouragement and numerous valuable suggestions.

This research was supported by National Science Foundation under Grant GU-1153.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION TO RESEARCH TOPIC

I-1. Statement of the problem

The identification or modeling of a given system plays a central part in the system theory. In linear systems one may identify the system's impulse response, system matrix, state transition matrix or coefficients of the transfer function. Thus, for a linear system all these result in estimation of certain parameters.

When dealing with nonlinear systems, we first face the problem of determining the representation of the evolution operator for the system and then estimating the parameters connected with the particular representation which will be optimum in some sense. Selection of a general class of mathematical models for a useful characterization of all nonlinear systems is difficult and may be impossible.

In this work, no attempt has been made to solve the problem of optimum representation for a nonlinear evolution operator. Thus in Chapters II and III we have considered nonlinear dynamical systems whose evolution operator is represented (or adequately approximated) by a polynomial in the state variables. Specifically, Chapter II treats the case of a scalar-state and Chapter III deals with the vector-state. In Chapter IV, the evolution operator is represented by a sine and cosine series in the state variables. In both cases, parameters associated with the model are adjusted by the use of stochastic approximation.

I-2. Relation to the Results of Other Investigators

Various results for the linear systems have been obtained by other investigators. Sakrison [6] has obtained algorithms for the estimation of the coefficients of numerator and denominator polynomials of the system transfer function. Saridis and Stein [7] have estimated the state transition matrix for a discrete linear system. For nonlinear systems with no dynamics, Fu, et al., [2] have obtained convergent estimates for any arbitrary representation of the memoryless transformation. For the case of nonlinear systems with deterministic dynamics, with noise affecting only the measurements, one may resort to the on-line state estimation algorithms of reference [3] by treating the parameters as augmented states of the system to be identified. In this work, an on-line scheme for the estimation of parameters of two classes of nonlinear dynamical systems when both dynamic noise and measurement noise are present, is developed.

CHAPTER II

IDENTIFICATION OF SYSTEMS WITH A SCALAR STATE

II-1. Introduction and Assumptions

We treat below the single input, single output system which is time invariant and for which power series expansion is exact up to a finite number of terms and which is governed by the following equations. Components of vectors will be denoted by superscripts and their positions in time by subscripts. J will denote the set of nonnegative integers. The superscript T on a vector or matrix will denote its transpose.

$$x_{k+1} = g(x_k) + u_k \quad ; \quad k \in J \tag{2-1}$$

$$z_{k+1} = h\, x_{k+1} + v_{k+1} \ ; \ h \neq 0 \tag{2-2}$$

where x_k and u_k denote the state and the dynamic noise input at time k , v_{k+1} being the additive noise occuring in the measurement z_{k+1}.

$$g(x_k) = a^T \Phi(x_k) \tag{2-3}$$

where

$$\Phi(x_k) = \text{col}\ [x_k,\ (x_k)^2, \ldots, (x_k)^p] \tag{2-4}$$

and

$$a = \text{col}\ [a_1,\ a_2, \ldots a_p] \qquad (2\text{-}5]$$

Let

$$E[(v_k)^m] = \alpha_m \qquad ; \qquad k, m \in J \tag{2-6}$$

$$E[(u_k)^m] = \beta_m \qquad ; \qquad k, m \in J \tag{2-7}$$

We assume the following:

(a) $\{u_k\}$ and $\{v_k\}$ are sample functions from independent stationary random processes, and in addition, u_k, v_k and x_k are stochastically independent for each k.

(b) $\lambda_o = \inf_{k \geq 1} [\lambda_k] > 0$

where λ_k is the minimum eigenvalue of $E[\Phi(x_k)\Phi^T(x_k)]$.

(c) The values of $\beta_1, \alpha_1, \ldots \alpha_{2p}$ are all known and $\beta_2, \alpha_{2p+1} \cdots \alpha_{4p}$, $E[(x_k)^i]_{i=1,\ldots 2p}$ are all bounded from above by a constant for all k.

II-2. Developments Leading to the Algorithm

Let us define

$$z_{k;2p} = E\begin{bmatrix}(z_k)\\(z_k)^2\\ \cdot \\ \cdot \\ \cdot \\ (z_k)^{2p}\end{bmatrix} \quad \text{and} \quad x_{k;2p} = E\begin{bmatrix}(x_k)\\(x_k)^2\\ \cdot \\ \cdot \\ \cdot \\ (x_k)^{2p}\end{bmatrix} \tag{2-8}$$

Then it is easily seen from equation (2-2) that

$$z_{k;2p} = \begin{bmatrix} h & 0 & 0 \cdots 0 \\ \binom{2}{1} h\alpha_1 & (h)^2 & 0 \cdots 0 \\ \binom{3}{2} h\alpha_2 & \binom{3}{1}(h)^2\alpha_1 & (h)^3 \cdots 0 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & 0 \\ \binom{2p}{2p-1} h\alpha_{2p-1} & \cdots\cdots & (h)^{2p} \end{bmatrix} x_{k;2p} + \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_{2p} \end{bmatrix}^{*} \tag{2-9}$$

Let

$$\text{matrix}[a_{ij}] = A = \begin{bmatrix} h & 0 \cdots\cdots & 0 \\ \binom{2}{1} h\alpha_1 & (h)^2 \cdots\cdots & 0 \\ \vdots & \vdots & \vdots \\ \binom{2p}{2p-1} h\alpha_{2p-1} & \cdots\cdots & (h)^{2p} \end{bmatrix}^{-1} \tag{2-10}$$

A is thus a lower triangular matrix whose elements are known by our assumptions. Clearly

$$E[(x_k)^n] = \sum_{j=1}^{n} a_{nj} \{E(z_k)^j - \alpha_j\} \tag{2-11}$$

---

* $\binom{n}{r} = \frac{n!}{r!(n-r)!}$

From (2-4)

$$\Phi(z_k/h)\ \Phi^T(z_k/h) = \Phi(x_k)\ \Phi^T(x_k) + U \tag{2-12a}$$

where U is a matrix whose $i\text{-}j^{th}$ term is

$$U(i,j) = \sum_{r=1}^{i+j} \binom{i+j}{r} (x_k)^{i+j-r} \left(\frac{v_k}{h}\right)^r \qquad i;j = 1,\ldots,p \tag{2-12b}$$

Define the matrix O_k whose $i\text{-}j^{th}$ term is

$$O_k(i;j) = \sum_{r=1}^{i+j} \binom{i+j}{r} \left\{ \sum_{\ell=1}^{i+j-r} a_{(i+j-r);\ell}[(z_k)^\ell - \alpha_\ell] \right\} \frac{\alpha_r}{(h)^r} \tag{2-13}$$

$$i;j = 1,\ldots,p$$

Computation of this symmetric matrix is possible knowing the noise moments, the $[a_{ij}]$ and the measurements z_k. From (2-11), (2-12) and (2-13) we have:

$$E[\Phi(z_k/h)\ \Phi^T(z_k/h)] = E[\Phi(x_k)\ \Phi^T(x_k)] + E[O_k] \tag{2-14}$$

This relation plays an important part in the regression function found in the next section.

II-3 Formulation of the Algorithm

Define

$$P_{k;p} = E\begin{bmatrix} \frac{z_{k+1}}{h}\ (z_k/h)^0 \\ \frac{z_{k+1}}{h}\ (z_k/h)^1 \\ \vdots \\ \frac{z_{k+1}}{h}\ (z_k/h)^p \end{bmatrix} ; \text{ and } \quad Q_{k;p} = E\begin{bmatrix} (x_k)^0 g(x_k) \\ x_k\ g(x_k) \\ \vdots \\ (x_k)^p g(x_k) \end{bmatrix} \tag{2-15}$$

Then

$$P_{k;p} = \begin{bmatrix} 1 & 0 & \cdots\cdots & 0 \\ \frac{\alpha_1}{h} & 1 & & 0 \\ \vdots & & & \vdots \\ \frac{\alpha_p}{(h)^p} & \binom{p}{p-1}\frac{\alpha_{p-1}}{(h)^{p-1}} & \cdots\cdots & 1 \end{bmatrix} Q_{k;p} + \left(\beta_1 + \frac{\alpha_1}{h}\right) \begin{bmatrix} 1 \\ ------ \\ E[\Phi(z_k/h)] \\ \\ \\ \end{bmatrix} \qquad (2\text{-}16)$$

Let

$$\text{matrix } [b_{ij}] = \begin{bmatrix} 1 & 0 & \cdots\cdots\cdots\cdots & 0 \\ \frac{\alpha_1}{h} & 1 & \cdots\cdots\cdots\cdots & 0 \\ \vdots & & & \vdots \\ \frac{\alpha_p}{(h)^p} & \binom{p}{p-1}\frac{\alpha_{p-1}}{(h)^{p-1}} & \cdots\cdots & 1 \end{bmatrix}^{-1} \qquad (2\text{-}17)$$

Clearly matrix $[b_{ij}]$ is a lower triangular matrix with unity diagonal elements and with all the terms known.

Define

$$H_1 = \begin{bmatrix} \sum_{j=0}^{1} b_{1j}(z_{k+1}/h)(z_k/h)^j \\ \vdots \\ \sum_{j=0}^{p} b_{pj}(z_{k+1}/h)(z_k/h)^j \end{bmatrix} \text{ and } H_2 = \begin{bmatrix} \sum_{j=0}^{1} b_{1j}(z_k/h)^j \\ \vdots \\ \sum_{j=0}^{p} b_{pj}(z_k/h)^j \end{bmatrix} \qquad (2\text{-}18)$$

Then from (2-16) and (2-18)

$$E[g(x_k)\Phi(x_k)] = E\{H_1\} - E\{H_2\} \tag{2-19}$$

Equations (2-14) and (2-19) motivate the following algorithm:

$$\hat{a}_{k+1} = \hat{a}_k + \rho(k)[\{H_1\} - \{H_2\} - \{\Phi(z_k/h)\Phi^T(z_k/h) - O_k\}\hat{a}_k] \tag{2-20}$$

$$k = 1,2,\ldots; \|\hat{a}_1\| < \infty$$

where $\rho(j)$ is a sequence of nonnegative real numbers such that

$$\sum_{j=1}^{\infty} \rho^2(j) < \infty \quad \text{and} \quad \sum_{j=1}^{\infty} \rho(j) = \infty \tag{2-21}$$

Since the system under consideration is a special case of the vector problem treated in the following chapter, the convergence of the above algorithm follows from the proof given in Appendix 1.

CHAPTER III

IDENTIFICATION OF NONLINEAR SYSTEMS WITH A VECTOR STATE

III-1. Introduction

In this chapter we consider nonlinear dynamical systems which have n "states". This is essentially the generalization of the previous chapter. Stochastic approximation methods for the identification of discrete linear time invariant systems with vector state, driven by white noise using measurements corrupted by white noise have been discussed by Saridis and Stein .[7] It has been proved that in any on-line estimation of parameters of a linear or nonlinear system using stochastic approximation to minimize a mean-square error criterion, the knowledge of noise moments is essential. Our method is on-line and thus requires the knowledge of certain noise moments.

III-2. Problem Formulation

In what follows, components of vectors will be denoted by superscripts and their positions in time by subscripts. J will denote the set of nonnegative integers.

The class of nonlinear systems that we consider in this chapter is described by the set of equations

$$x_{k+1} = g(x_k) + u_k \quad , \quad k \in J . \tag{3-1a}$$

$$z_{k+1} = h^T x_{k+1} + v_{k+1} \ , \ k \epsilon J \ , \qquad (3\text{-}1b)$$

where $x_k = \mathrm{col}(x_k^1, \ldots, x_k^n)$ and $u_k = \mathrm{col}(0, \ldots, 0, u_k^n)$ denote the state and the dynamic noise input at time k, $h = \mathrm{col}(0, \ldots, 0, 1)$, the scalar variable v_{k+1} represents the additive noise occuring in the measurement z_{k+1}, the superscript T on a vector or matrix denotes its transpose and $g(x_k) \equiv \mathrm{col}\,(x_k^2, \ldots, x_k^n, \Psi(x_k))$ where

$$\Psi(x_k) = \sum_{i_1=1}^{n} a_{i_1} x_k^{i_1} + \sum_{i_1=1}^{n} \sum_{i_2=1}^{i_1} a_{i_1 i_2} x_k^{i_1} x_k^{i_2} + \ldots$$

$$+ \sum_{i_1=1}^{n} \sum_{i_2=1}^{i_1} \cdots \sum_{i_{p-1}=1}^{i_{p-2}} \sum_{i_p=1}^{i_{p-1}} a_{i_1 i_2 \ldots i_p} x_k^{i_1} x_k^{i_2} \ldots x_k^{i_p} . \qquad (3\text{-}2)$$

The preceding equation may be rewritten in the form

$$\Psi(x_k) = a^T \Phi(x_k) \ , \qquad (3\text{-}3a)$$

where

$$a = \mathrm{col}[a_1, \ldots, a_n, \vdots\, a_{11}, a_{21}, a_{22}, \ldots, a_{nn} \vdots \ldots \vdots\, \underbrace{a_{11\ldots1}}_{p\ \text{times}}, a_{21\ldots1}, \ldots, a_{nn\ldots n}] \ ; \qquad (3\text{-}3b)$$

and

$$\Phi(x_k) = \mathrm{col}[x_k^1, \ldots, x_k^n \vdots\, x_k^1 x_k^1, x_k^2 x_k^1, x_k^2 x_k^2, \ldots\ldots, x_k^n x_k^n \vdots \ldots \vdots\, \underbrace{x_k^1 x_k^1 \ldots x_k^1}_{p\ \text{times}}, \ldots, x_k^n x_k^n \ldots x_k^n] \ . \qquad (3\text{-}3c)$$

Let

$$E[(v_k)^m] = \alpha_m \ ; \ k, m \epsilon J \ , \qquad (3\text{-}4)$$

$$E[(u_k^n)^m] = \beta_m \ ; \ n, k, m \epsilon J \ . \qquad (3\text{-}5)$$

We assume the following:

(a) $\{u_k\}$ and $\{v_k\}$ are sample functions from independent stationary random processes, and, in addition, u_k, v_k and x_k are stochastically independent for each k.

(b) $\lambda_0 = \inf_{k \geq n} [\lambda_{k+n+1}] > 0$

where λ_{k+n-1} is the minimum eigen value of $E[\Phi(x_{k+n-1})\Phi^T(x_{k+n-1})]$.

(c) For part I, the values of β_1 $\alpha_1, \alpha_2, \ldots, \alpha_{2p}$ are all known and $\beta_2, \alpha_{2p+1}, \ldots, \alpha_{4p}$, $E[(x_k^i)^\ell (x_k^j)^m]_{\substack{i,j=1,2,\ldots,n \\ \ell+m \leq 4p}}$ are all bounded from above by a constant for all k.

Much of our effort in the remaining part of the paper will be devoted to the development of those quantities required to formulate a convergent stochastic approximation algorithm for the unbiased estimation of the parameter vector a.

III-2. Developments Leading to the Algorithm

Throughout this section we will use the notations I_{mm} for an $m \times m$ identity matrix and $\binom{m}{j}$ for $m!/(j!(m-j)!)$.

Let us define

$$Z_{k+n-1} = \begin{bmatrix} z_k \\ z_{k+1} \\ \vdots \\ \vdots \\ z_{k+n-1} \end{bmatrix} \quad ; \quad V_{k+n-1} = \begin{bmatrix} v_k \\ v_{k+1} \\ \vdots \\ \vdots \\ v_{k+n-1} \end{bmatrix} . \tag{3-6,a,b}$$

Then, according to (3.2)

$$Z_{k+n-1} = x_{k+n-1} + V_{k+n-1} \tag{3-7}$$

Let also

$$Q_1(x_{k+n-1}) = \begin{bmatrix} x^1_{k+n-1} \\ x^2_{k+n-1} \\ \vdots \\ \vdots \\ x^n_{k+n-1} \end{bmatrix} \quad ; \quad Q_2(x_{k+n-1}) = \begin{bmatrix} x^1_{k+n-1} \cdot x^1_{k+n-1} \\ x^2_{k+n-1} \cdot x^1_{k+n-1} \\ \vdots \\ \vdots \\ x^n_{k+x-1} \cdot x^n_{k+n-1} \end{bmatrix} \tag{3-8,i; 3-8,ii}$$

$$Q_p(x_{k+n-1}) = \begin{bmatrix} \underbrace{x^1_{k+n-1} \cdots\cdots x^1_{k+n-1}}_{p \text{ times}} \\ x^2_{k+n-1} \cdot x^1_{k+n-1} \cdots\cdots x^1_{k+n-1} \\ \vdots \\ x^n_{k+n-1} \cdots\cdots x^n_{k+n-1} \end{bmatrix} \qquad (3\text{-}8,p)$$

where the dimension of $Q_i(x)$ is $\binom{n+i-1}{i} \times 1$, $1 \leq i \leq p$. Then from (3-3,c)

$$\Phi(x_{k+n-1}) = \begin{bmatrix} Q_1(x_{k+n-1}) \\ \hline Q_2(x_{k+n-1}) \\ \hline \vdots \\ \hline Q_p(x_{k+n-1}) \end{bmatrix} . \qquad (3\text{-}9)$$

Now according to (3-7),

$$Q_1(Z_{k+n-1}) = \widetilde{A}_{11} Q_1(x_{k+n-1}) + V_{k+n-1} . \qquad (3\text{-}10)$$

Also, since

$$\begin{aligned} Z^1_{k+n-1} \cdot Z^1_{k+n-1} &= x^1_{k+n-1} x^1_{k+n-1} + 2V^1_{k+n-1} x^1_{k+n-1} + V^1_{k+n-1} \cdot V^1_{k+n-1} \\ Z^2_{k+n-1} \cdot Z^1_{k+n-1} &= x^2_{k+n-1} x^1_{k+n-1} + V^2_{k+n-1} x^1_{k+n-1} + V^1_{k+n-1} x^2_{k+n-1} + V^2_{k+n-1} V^1_{k+n-1} \\ &\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (3\text{-}11) \\ Z^n_{k+n-1} \cdot Z^n_{k+n-1} &= x^n_{k+n-1} \cdot x^n_{k+n-1} + 2V^n_{k+n-1} x^n_{k+n-1} + V^n_{k+n-1} V^n_{k+n-1} , \end{aligned}$$

it follows from (3-8) that

$$Q_2(Z_{k+n-1}) = \widetilde{A}_{21} Q_1(x_{k+n-1}) + \widetilde{A}_{22} Q_2(x_{k+n-1}) + Q_2(V_{k+n-1}) , \qquad (3\text{-}12)$$

where $\widetilde{A}_{22} = I_{\binom{n+1}{2} \times \binom{n+1}{2}}$ and

$$\widetilde{A}_{21} = \begin{bmatrix} 2V^1_{k+n-1} & 0 \cdots\cdots & 0 \cdots 0 \cdots\cdots 0 \\ V^2_{k+n-1} & V^1_{k+n-1} \cdots & 0 \cdots 0 \cdots\cdots 0 \\ 0 & 2V^2_{k+n-1} & 0 \cdots\cdots\cdots \\ \vdots & \vdots & \vdots \\ 0 \cdots\cdots & \cdots\cdots\cdots & \cdots 2V^n_{k+n-1} \end{bmatrix} \quad . \tag{3-13}$$

By similar procedure, for $2 < i \leq p$,

$$Q_i(Z_{k+n-1}) = \sum_{j=1}^{i} \widetilde{A}_{ij} Q_j(x_{k+n-1}) + Q_i(V_{k+n-1}) \quad , \tag{3-14}$$

where $\widetilde{A}_{ij}$ is a $\binom{n+i-1}{i} \times \binom{n+j-1}{j}$ stochastic matrix having only the noise components as terms, and

$$\widetilde{A}_{jj} = I_{\binom{n+j-1}{j} \times \binom{n+j-1}{j}} \quad .$$

By means of the notation

$$\widetilde{A} = \begin{bmatrix} \widetilde{A}_{11} & 0 \cdots\cdots & 0 \\ \widetilde{A}_{21} & \widetilde{A}_{22} & 0 \\ \vdots & & \ddots \vdots \\ \widetilde{A}_{p1} & \widetilde{A}_{p2} \cdots & \widetilde{A}_{pp} \end{bmatrix} \tag{3-15}$$

finally we arrive at the equation:

$$\Phi(Z_{k+n-1}) = \widetilde{A}\, \Phi(x_{k+n-1}) + \Phi(V_{k+n-1}) \quad . \tag{3-16}$$

Let

$$A = E[\widetilde{A}] \quad . \tag{3-17}$$

According to our assumptions A is known. Also

$$E[\Phi(Z_{k+n-1})] = A\, E[\Phi(x_{k+n-1})] + E[\Phi(V_{k+n-1})] \quad . \tag{3-18}$$

At this point, we need to introduce the tensors $\widetilde{B}$ and $\widetilde{C}$ by means of the equations

$$\widetilde{A}\Phi(x_{k+n-1})\, \Phi^T(x_{k+n-1})\, \widetilde{A}^T = \widetilde{B}\Phi(x_{k+n-1})\Phi^T(x_{k+n-1}) \tag{3-19}$$

and

$$\widetilde{A}\ \Phi(x_{k+n-1})\ \Phi^T(V_{k+n-1}) = \widetilde{C}\ \Phi(x_{k+n-1}) \qquad (3\text{-}20)$$

We note that $\widetilde{B}$ and $\widetilde{C}$ have only noise terms as their components and define

$$E[\widetilde{B}] = B\ ,\ E[\widetilde{C}] = C \qquad (3\text{-}21)$$

All the components of B and C are known by our assumptions.

Clearly

$$\Phi(Z_{k+n-1})\ \Phi^T(Z_{k+n-1}) = \widetilde{B}\ \Phi(x_{k+n-1})\ \Phi^T(x_{k+n-1}) + \widetilde{C}\ \Phi(x_{k+n-1}) + [\widetilde{C}\ \Phi(x_{k+n-1})]^T + \Phi(V_{k+n-1})\ \Phi^T(V_{k+n-1})\ . \qquad (3\text{-}22)$$

Taking expectation of both sides and using the tensor B^{-1} as defined in Appendix II, we obtain

$$\begin{aligned} E[\Phi(x_{k+n-1})\ \Phi^T(x_{k+n-1})] = {} & B^{-1}\ E[\Phi(Z_{k+n-1})\ \Phi^T(Z_{k+n-1})] - B^{-1}CA^{-1}[E(\Phi(Z_{k+n-1}))] \\ & -B^{-1}[CA^{-1}E(\Phi(Z_{k+n-1}))]^T + B^{-1}CA^{-1}E[\Phi(V_{k+n-1})] \\ & +B^{-1}[CA^{-1}E[\Phi(V_{k+n-1})]]^T - B^{-1}[E[\Phi(V_{k+n-1})\Phi^T(V_{k+n-1})]]. \end{aligned} \qquad (3\text{-}23)$$

Now define

$$\begin{aligned} O_{k+n-1} = {} & B^{-1}CA^{-1}[\Phi(Z_{k+n-1})] + B^{-1}[CA^{-1}[\Phi(Z_{k+n-1})]]^T \\ & -B^{-1}CA^{-1}[E\Phi(V_{k+n-1})] - B^{-1}[CA^{-1}\ E(\Phi(V_{k+n-1}))]^T \\ & +B^{-1}E[\Phi(V_{k+n-1})\ \Phi^T(V_{k+n-1})] \end{aligned} \qquad (3\text{-}24)$$

Then clearly,

$$E[\Phi(x_{k+n-1})\Phi^T(x_{k+n-1})] = B^{-1}E[\Phi(Z_{k+n-1})\Phi^T(Z_{k+n-1})] - E[O_{k+n-1}]\ . \qquad (3\text{-}25)$$

This relation plays an important role in the regression function appearing in the following section.

III-3. Formulation of the Algorithm

From the fact that

$$E[Z^n_{k+n}] = E[\Psi(x_{k+n-1})] + \beta_1 + \alpha_1 \qquad (3\text{-}26)$$

and

$$E[Z^n_{k+n}\Phi(Z_{k+n-1})] = AE[\Psi(x_{k+n-1})\Phi(x_{k+n-1})] + (\beta_1 + \alpha_1)E[\Phi(Z_{k+n-1})]$$
$$+E[Z^n_{k+n}]\ E[\Phi(V_{k+n-1})] - (\beta_1 + \alpha_1)E[\Phi(V_{k+n-1})] \tag{3-27}$$

we are led to the definitions:

$$\{H_1\} = A^{-1}Z^n_{k+n}[\Phi(Z_{k+n-1})-E(\Phi(V_{k+n-1}))] \tag{3-28a}$$
$$\{H_2\} = (\beta_1 + \alpha_1)A^{-1}[\Phi(Z_{k+n-1})-E(\Phi(V_{k+n-1}))]. \tag{3-28b}$$

Then, according to (3-27) and (3-28),

$$E[\Psi(x_{k+n-1})\Phi(x_{k+n-1}) - (\{H_1\} - \{H_2\})] = 0 \tag{3-29}$$

Now if we combine (3-25) with the expression inside square brackets in (3-29),we obtain the desired regression function. The resultant algorithm then takes the form

$$\hat{a}_{k+n} = \hat{a}_{k-1} + \rho(\tfrac{k-1}{n+1})\ [\{H_1\} - \{H_2\} - \{B^{-1}\Phi(Z_{k+n-1})\Phi^T(Z_{k+n-1}) - O_{k+n-1}\}\hat{a}_{k-1}], \tag{3-30}$$

$k=1\ ,\ n+2\ ,\ 2n+3\ ,\ \ldots\ldots\ ;\ \|\hat{a}_o\| < \infty$,

where $\rho(j)$ is a sequence of nonnegative numbers such that

$$\sum_{j=1}^{\infty} \rho(j) = \infty \quad \text{and} \quad \sum_{j=1}^{\infty} \rho^2(j) < \infty\ . \tag{3-31}$$

As shown in Appendix I, the sequence of $\left(\sum_{i=1}^{p} \binom{n+i-1}{i}\right)$ - dimensional vectors $\{\hat{a}_{k(n+1)};\ k\epsilon J\}$ tends to the true parameter vector a in mean square and with probability one.

CHAPTER IV

TRIGONOMETRIC EVOLUTION OPERATOR-SYSTEMS

IV-1. Introduction and Algorithm Development

Nonlinear dynamical systems whose evolution operators are represented (or adequately approximated) by a finite number of terms in their expansion by a sine and cosine- series in state variables are treated in this part. The class of nonlinear systems is represented by:

$$x_{k+1} = g(x_k) + u_k \quad ; k \in J \tag{4-1}$$

$$z_{k+1} = x_{k+1} + v_{k+1} \tag{4-2}$$

where x_k (scalar), u_k, v_k are the state, input noise, and measurement noise respectively.

$$g(x_k) = a^T \Phi(x_k) \tag{4-3}$$

where

$$a = \text{col.}\ [a_1, a_2, \ldots, a_{2n}] \text{ and} \tag{4-4a}$$

$$\Phi(x_k) = \text{col.}\ [\sin x_k, \cos x_k, \ldots \sin nx_k, \cos nx_k] \tag{4-4b}$$

Let

$$E[\cos nv_k] = c_n \quad ; \quad E[\sin nv_k] = s_n$$

Then we assume that α_1, β_1, $c_1, \ldots, c_{2n}$, $s_1, \ldots, s_{2n}$ are all known, and β_2, α_2, $c_{2n+1}, \ldots, c_{4n}$, $s_{2n+1}, \ldots, s_{4n}$, $E[\sin ix_k]_{i=1,\ldots,4n}$, $E[\cos ix_k]_{i=1,\ldots,4n}$ are finite.

Consider the following set of equations:

$$\begin{bmatrix} \sin z_k = \sin x_k \cos v_k + \cos x_k \sin v_k \\ \cos z_k = \cos x_k \cos v_k - \sin x_k \sin v_k \end{bmatrix} \tag{4-5,1}$$

$$\vdots$$

$$\begin{bmatrix} \sin 2nz_k = \sin 2nx_k \cos 2nv_k + \cos 2nx_k \sin 2nv_k \\ \cos 2nz_k = \cos 2nx_k \cos 2nv_k - \sin 2nx_k \sin 2nv_k \end{bmatrix} \tag{4-5,2n}$$

Define

$$\tilde{\bar{A}}^{1}_{4n\times 4n} = \begin{bmatrix} \cos v_k & \sin v_k & 0 & 0 & & & \\ -\sin v_k & \cos v_k & 0 & 0 & & 0 & \\ 0 & 0 & \cos 2v_k & \sin 2v_k & & & \\ 0 & 0 & -\sin 2v_k & \cos 2v_k & & & \\ & & & & \ddots & & \\ & 0 & & & & \cos 2nv_k & \sin 2nv_k \\ & & & & & -\sin 2nv_k & \cos 2nv_k \end{bmatrix} \quad \text{and} \qquad (4\text{-}6)$$

$$\psi^{4n}(z_k) = \text{col. } [\sin z_k, \cos z_k \ldots \sin 2nz_k, \cos 2nz_k] \qquad (4\text{-}7)$$

Then the set of equations (4-5) can be written as:

$$\psi^{4n}(z_k) = \tilde{\bar{A}}^{1}_{4n\times 4n} \cdot \psi^{4n}(x_k) \qquad (4\text{-}8)$$

Let

$$E[\tilde{\bar{A}}^{1}_{4n\times 4n}] = \bar{A}^{1}{}_{4n\times 4n} \qquad (4\text{-}9)$$

Then (4-8) reduces to

$$E[\psi^{4n}(x_k)] = A_{4n\times 4n} \cdot E[\psi^{4n}(z_k)] \qquad (4\text{-}10)$$

where

$$A_{4n\times 4n} = \begin{bmatrix} c_1 & -s_1 & 0 & 0 & & & \\ s_1 & c_1 & 0 & 0 & & 0 & \\ 0 & 0 & c_2 & -s_2 & & & \\ 0 & 0 & s_2 & c_2 & & & \\ & & & & \ddots & & \\ & 0 & & & & c_{2n} & -s_{2n} \\ & & & & & s_{2n} & c_{2n} \end{bmatrix} \qquad (4\text{-}11)$$

This implies that

$$E[\sin ix_k] = c_i\ E[\sin iz_k] - s_i\ E[\cos iz_k] \tag{4-12a}$$

$$E[\cos ix_k] = s_i\ E[\sin iz_k] + c_i\ E[\cos iz_k] \tag{4-12b}$$

From (4-4b)

$$E[\Phi(x_k)\Phi^T(x_k)]=E\begin{bmatrix} \sin^2 x_k & \sin x_k \cos x_k & \cdots & \sin x_k \cos nx_k \\ \sin x_k \cos x_k & \cos^2 x_k & \cdots & \cos x_k \cos nx_k \\ \vdots & & \ddots & \vdots \\ \sin x_k \cos nx_k & \cos x_k \cos nx_k & \cdots & \cos^2 nx_k \end{bmatrix} \tag{4-13}$$

This matrix can be written as the sum of two matrices by the use of trigonometric identities.

$$E[\Phi(x_k)\Phi^T(x_k)]=\tfrac{1}{2}E\left\{\begin{bmatrix} 1 & \sin 2x_k & \cdots & -\sin(n-1)x_k \\ \sin 2x_k & 1 & \cdots & \cos(n+1)x_k \\ \vdots & & & \vdots \\ \sin(n+1)x_k & \cdots & & 1 \end{bmatrix} + \begin{bmatrix} -\cos 2x_k & 0 & \sin(n+1)x_k \\ 0 & \cos 2x_k & \cos(n-1)x_k \\ \vdots & & \vdots \\ -\sin(n-1)x_k & \cdots & \cos 2nx_k \end{bmatrix}\right\} \tag{4-14}$$

Define

$$\widetilde{P}_1 = \begin{bmatrix} 1 & c_2 \sin 2z_k & \cdots & -c_{n-1}\sin(n-1)z_k \\ c_2 \sin 2z_k & 1 & \cdots & s_{n+1}\sin(n+1)z_k \\ \vdots & & \ddots & \vdots \\ c_{n+1}\sin(n+1)z_k & \cdots & & 1 \end{bmatrix} + \tag{4-15}$$

cont'd next pg.

$$+\begin{bmatrix} 0 & -s_2\cos 2z_k \cdots & s_{n-1}\cos(n-1)z_k \\ -s_2\cos 2z_k & 0 \cdots & c_{n+1}\cos(n+1)z_k \\ \vdots & \ddots & \vdots \\ -s_{n+1}\cos(n+1)z_k \cdots & & 0 \end{bmatrix} \tag{4-15}$$

$$\tilde{P}_2 = \begin{bmatrix} -s_2\sin 2z_k & 0\cdots & c_{n+1}\sin(n+1)z_k \\ 0 & s_2\sin 2z_k \cdots & s_{n-1}\sin(n-1)z_k \\ \vdots & & \vdots \\ -c_{n-1}\sin(n-1)z_k \cdots & & s_{2n}\sin 2nz_k \end{bmatrix}$$

$$+\begin{bmatrix} -c_2\cos 2z_k & 0\cdots & -s_{n+1}\cos(n+1)z_k \\ 0 & c_2\cos 2z_k & c_{n-1}\cos(n-1)z_k \\ \vdots & & \vdots \\ s_{n-1}\cos(n-1)z_k \cdots & & c_{2n}\cos 2nz_k \end{bmatrix}. \tag{4-16}$$

Then clearly from (4-12), (4-14), (4-15), (4-16) we have

$$E[\tfrac{1}{2}(\tilde{P}_1 + \tilde{P}_2)] = E[\Phi(x_k)\Phi^T(x_k)] \tag{4-17}$$

This important relation is used in the regression function developed in the next section.

IV-3. Formulation of Algorithm

From (4-2) and (4-4) we have

$$z_{k+1}\psi^{2n}(z_k) = [g(x_k) + u_k + v_{k+1}][\tilde{A}^{-1}_{2nx2n}\psi^{2n}(x_k)] \tag{4-18}$$

Taking expectation of both sides and rearranging

$$E[z_{k+1}\psi^{2n}(z_k)] = \bar{A}^{-1}_{2nx2n}\, E[g(x_k)\psi^{2n}(x_k)] + (\beta_1+\alpha_1)\dot{E}[\psi^{2n}(z_k)] \quad (4\text{-}19)$$

and since

$$\psi^{2n}(x_k) = \Phi(x_k) \quad (4\text{-}20)$$

$$E[g(x_k)\Phi(x_k)] = A_{2nx2n}\cdot E[z_{k+1}\psi^{2n}(z_k)] - (\beta_1+\alpha_1)A_{2nx2n}E[\psi^{2n}(z_k)] \quad (4\text{-}21)$$

As in the first part, we combine the relations (4-17) and (4-21) to get the desired regression function which results in the following algorithm

$$\hat{a}_{k+1} = \hat{a}_k + \rho(k)[A_{2nx2n}(z_{k+1}\psi^{2n}(z_k)) - (\beta_1+\alpha_1)A_{2nx2n}\psi^{2n}(z_k) - \tfrac{1}{2}[\tilde{P}_1+\tilde{P}_2]\hat{a}_k] \quad (4\text{-}22)$$

$$k = 1,2,\ldots;\ \|\hat{a}_o\| < \infty$$

where $\rho(k)$ is the sequence non-negative real numbers satisfying (2-21). By the use of equations (4-17) and (4-21), the proof of convergence of $\hat{a}_k$ to a in the mean square and probability one is same as in Appendix 1.

CHAPTER V

COMPUTER SIMULATIONS

The performance of the identification algorithms given in (2-20) and (3-32) is demonstrated on the following two systems. These algorithms were programmed on the I.B.M. 7040 digital computer by means of Fortran IV language.

Example 1.

This is an example with scalar state and therefore uses the algorithm given in (2-20). The system treated is:

$$x_{k+1} = a_1 x_k + a_2 (x_k)^2 + a_3 (x_k)^3 + u_k \qquad (5\text{-}1)$$

$$a^T = [.1, .01, .001] \qquad (5\text{-}2)$$

$$z_{k+1} = x_{k+1} + v_{k+1} \qquad (5\text{-}3)$$

where u_k and v_k were Gaussianly distributed noises with zero mean and unity covariance. A plot of the normalized value of error shows that the normalized error goes to zero as the number of iterations increase. However, the convergence is not very fast as shown in the graph.

Example 2.

This is an example with vector state and therefore uses the algorithm given in (3-32). The system treated is:

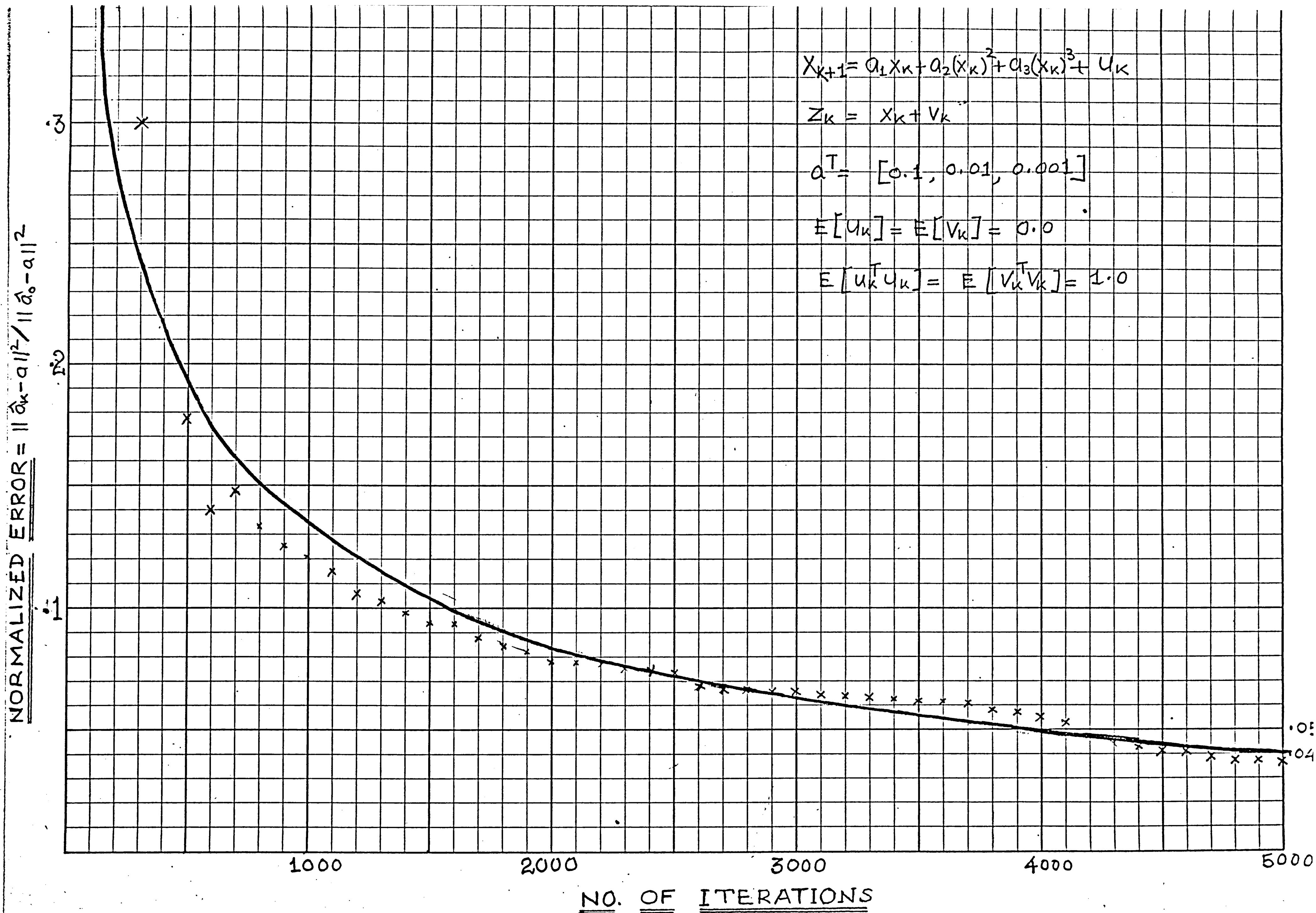
$$x^1_{k+1} = x^2_k$$

$$x^2_{k+1} = \sum_{i_1=1}^{2} a_{i_1} x_k^{i_1} + \sum_{i_1=1}^{2} \sum_{i_2=1}^{i_1} a_{i_1 i_2} x_k^{i_1} x_k^{i_2} + u_k \qquad (5\text{-}4)$$

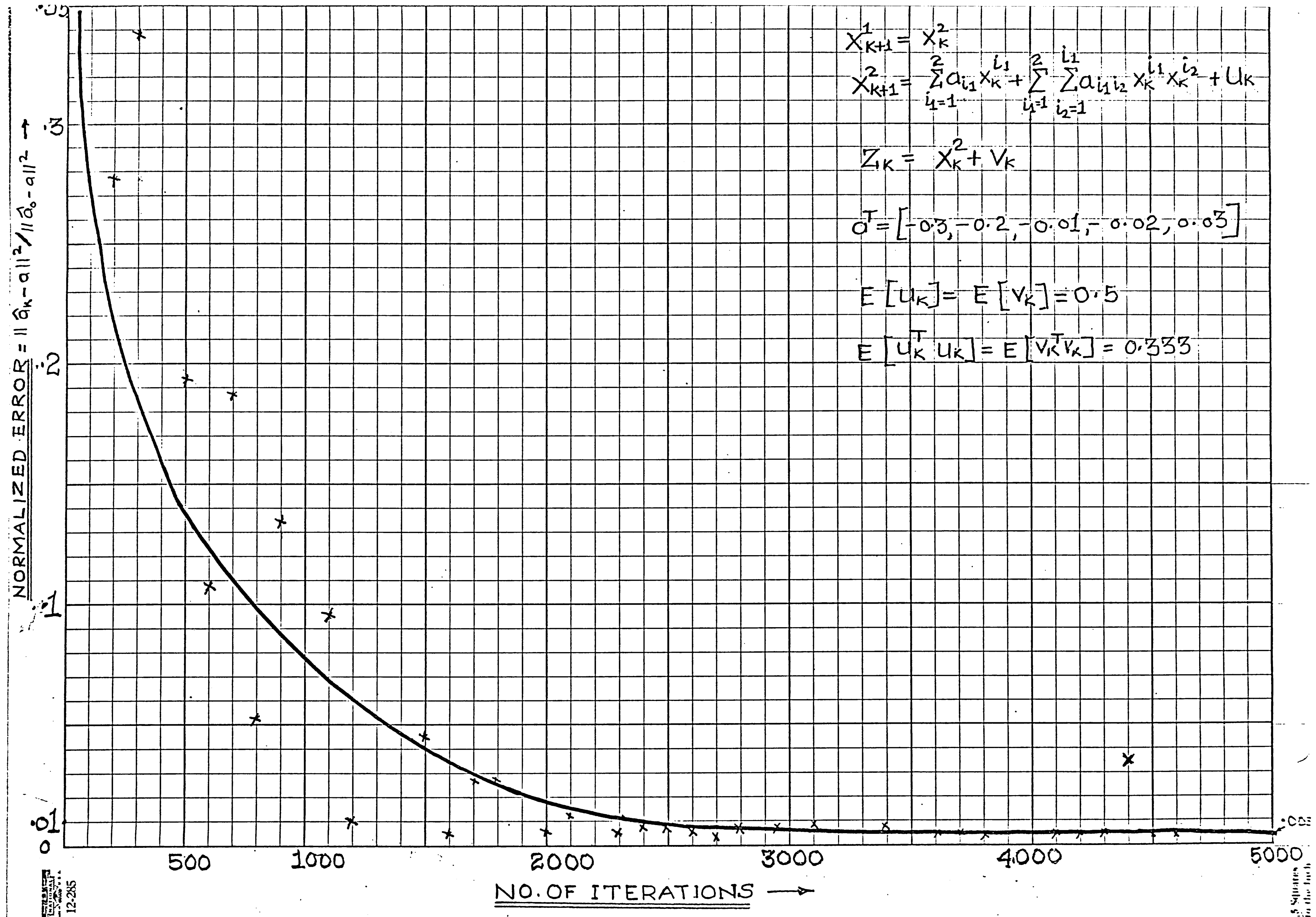
$$a^T = [-0.3, -0.2, -0.01, -0.02, +0.03] \quad (5\text{-}5)$$

$$z_{k+1} = x^2_{k+1} + v_{k+1} \quad (5\text{-}6)$$

where u_k and v_k were uniformly distributed noises from 0 to one. A plot of the normalized value of error shows that it goes to zero as the number of iterations increase thus achieving successful identification for this system.



$X_{k+1} = a_1 X_k + a_2 (X_k)^2 + a_3 (X_k)^3 + U_k$
$Z_k = X_k + V_k$
$a^T = [0.1, 0.01, 0.001]$
$E[U_k] = E[V_k] = 0.0$
$E[U_k^T U_k] = E[V_k^T V_k] = 1.0$
·3
·2
·1
·05
·04
1000
2000
3000
4000
5000
NO. OF ITERATIONS
NORMALIZED ERROR = $\|\hat{a}_k - a\|^2 / \|\hat{a}_0 - a\|^2$



NORMALIZED ERROR = $\|\hat{a}_k - a\|^2 / \|\hat{a}_0 - a\|^2$ →
.3
.2
.1
.01
0
500
1000
2000
3000
4000
5000
NO. OF ITERATIONS →
$X^1_{k+1} = X^2_k$
$X^2_{k+1} = \sum_{i_1=1}^{2} a_{i_1} X_k^{i_1} + \sum_{i_1=1}^{2} \sum_{i_2=1}^{i_1} a_{i_1 i_2} X_k^{i_1} X_k^{i_2} + U_k$
$Z_k = X^2_k + V_k$
$a^T = [-0.3, -0.2, -0.01, -0.02, 0.03]$
$E[U_k] = E[V_k] = 0.5$
$E[U_k^T U_k] = E[V_k^T V_k] = 0.333$

APPENDIX I

Convergence Proof

Define

$$G_{k+n-1} = B^{-1}\Phi(Z_{k+n-1})\Phi^T(Z_{k+n-1}) - O_{k+n-1} - \Phi(x_{k+n-1})\Phi^T(x_{k+n-1}) \quad \text{(A-1)}$$

$$L_{k+n-1} = \{H_1\} - \{H_2\} - \Psi(x_{k+n-1})\Phi(x_{k+n-1}) . \quad \text{(A-2)}$$

Clearly,

$$E[G_{k+n-1}] = 0 = E[L_{k+n-1}] . \quad \text{(A-3)}$$

Subtracting a from both sides of (3-30) and letting $\hat{e}_k = \hat{a}_k - a$, (A-4)

we are led to

$$\hat{e}_{k+n} = \hat{e}_{k-1} - \rho(\tfrac{k-1}{n+1})[\{\Phi(x_{k+n-1})\Phi^T(x_{k+n-1})\}\hat{a}_{k-1} - \Psi(x_{k+n-1})\Phi(x_{k+n-1})]$$
$$+ \rho(\tfrac{k-1}{n+1})[L_{k+n-1} - G_{k+n-1}\hat{a}_{k-1}] . \quad \text{(A-5)}$$

Then

$$\hat{e}^T_{k+n}\hat{e}_{k+n} = \hat{e}^T_{k-1}\hat{e}_{k-1} + \rho^2(\tfrac{k-1}{n+1})\gamma - 2\rho(\tfrac{k-1}{n+1})\ \hat{e}^T_{k-1}[\Phi(x_{k+n-1})\Phi^T(x_{k+n-1})\hat{a}_{k-1} - \Psi(x_{k+n-1})\cdot\Phi(x_{k+n-1})$$
$$+2\rho(\tfrac{k-1}{n+1})[L_{k+n-1} - G_{k+n-1}\hat{a}_{k-1}] , \quad \text{(A-6)}$$

where, by our assumptions, γ is such that

$$E[|\gamma| \,|\hat{e}_{k-1} = f] \leq \gamma_1\|f\|^2 + \gamma_2 , \qquad 0 \leq \gamma_1 ; \gamma_2 < \infty . \quad \text{(A-7)}$$

Hence,

$$E[\hat{e}^T_{k+n}\hat{e}_{k+n}|\hat{e}_{k-1} = f] \leq \|f\|^2 + \rho^2(\tfrac{k-1}{n+1})[\gamma_1\|f\|^2 + \gamma_2]$$
$$-2\rho(\tfrac{k-1}{n+1})f^T\ E[\Phi(x_{k+n-1})\Phi^T(x_{k+n-1})\hat{a}_{k-1}|\hat{e}_{k-1}=f]$$
$$+2\rho(\tfrac{k-1}{n+1})f^T E[\Psi(x_{k+n-1})\Phi(x_{k+n-1})|\hat{e}_{k-1}=f]$$
$$+2\rho(\tfrac{k-1}{n+1})f^T E[L_{k+n-1} - G_{k+n-1}\hat{a}_{k-1}|\hat{e}_{k-1}=f], \quad \text{(A-8)}$$

where, according to (A-3),

$$E[L_{k+n-1} - G_{k+n-1}\hat{a}_{k-1}|\hat{e}_{k-1}=f] = E[L_{k+n-1} - G_{k+n-1}\hat{a}_{k-1}|\hat{a}_{k-1}=f+a] = 0 . \quad \text{(A-9)}$$

By our assumptions equation (A-8) reduces to

$$E[\hat{e}^T_{k+n}\hat{e}_{k+n}] \leq [1 - 2\rho(\frac{k-1}{n+1})\lambda_o + \rho^2(\frac{k-1}{n+1})\gamma_1]E[\hat{e}^T_{k-1}\hat{e}_{k-1}] + \rho^2(\frac{k-1}{n+1})\gamma_2 \qquad \text{(A-10)}$$

since k = 1, n+2 ; 2n+3, ...

The above satisfies the convergence conditions given by Dvoretzky [8, section 8], and

$$\lim_{i\to\infty} E[\hat{e}^T_{i(n+1)}\hat{e}_{i(n+1)}] = \lim_{i\to\infty} E[\|\hat{e}_{i(n+1)}\|^2] = 0 \,. \qquad \text{(A-11)}$$

Hence

$$\lim_{i\to\infty} E[\|\hat{a}_{i(n+1)} - a\|^2] = 0 \,. \qquad \text{(A-12)}$$

and

$$P\{\lim_{i\to\infty} \hat{a}_{i(n+1)} = a\} = 1 \,. \qquad \text{(A-13)}$$

APPENDIX II

On the Tensors $\tilde{B}$ and $\tilde{C}$

1. The tensors $\tilde{B}$ and $\tilde{B}^{-1}$.

Let

$$s = \sum_{i=1}^{p} \binom{n+i-1}{i} \tag{A-14}$$

and consider

$$y = \tilde{A}\,\Phi(x_{k+n-1}) , \tag{A-15}$$

where y and $\Phi(x_{k+n-1})$ are s-vectors and $\tilde{A}$ is the $s\times s$ matrix defined in equation (3-15). Then

$$y^i y^k = \left[\sum_{j=1}^{s} \tilde{a}_{ij}\Phi^j(x_{k+n-1})\right]\left[\sum_{\ell=1}^{s} \tilde{a}_{k\ell}\,\Phi^{\ell}(x_{k+n-1})\right] \tag{A-16}$$

where $\tilde{a}_{ij}$ denote the elements of $\tilde{A}$. We define the tensor $\tilde{B}$ by assigning to its components the values

$$\tilde{B}_{ijk\ell} = \tilde{a}_{ij}\tilde{a}_{k\ell} . \tag{A-17}$$

Since $\tilde{A}$ is a lower triangular matrix with unity diagonal elements, we have

$$\tilde{B}_{ijk\ell} = \begin{cases} 0 & j > i \text{ or } \ell > k \\ 1 & i = j \text{ and } \ell = k \\ \tilde{B}_{ijk\ell} & \text{otherwise .} \end{cases} \tag{A-18}$$

Then

$$y^i y^k = \sum_{\ell=1}^{s}\sum_{j=1}^{s} \tilde{B}_{ijk\ell}\Phi^j(x_{k+n-1})\Phi^{\ell}(x_{k+n-1}) . \tag{A-19}$$

Letting $Y = \text{col}\,[y^1y^1, y^2y^1, \ldots\ldots y^sy^s]$ and

$$x = \text{col}[\Phi^1(x_{k+n-1})\Phi^1(x_{k+n-1}), \ldots\ldots \Phi^s(x_{k+n-1})\Phi^s(x_{k+n-1})] ,$$

(A-19) may be written in the form

$$Y = \underline{B}x \tag{A-20}$$

where $\underline{B}$ is a $\binom{s+1}{2} \times \binom{s+1}{2}$ lower triangular matrix with unity diagonal elements; hence it can be inverted and thus

$$x = \underline{B}^{-1} Y \; . \tag{A-21}$$

Now the components of $\widetilde{B}^{-1}$ may be defined via (A-21) by

$$\Phi^{j}(x_{k+n-1})\Phi^{\ell}(x_{k+n-1}) = \sum_{i=1}^{s} \sum_{k=1}^{s} \widetilde{B}^{-1}_{ijk\ell} \, y^{i} y^{k} \tag{A-22}$$

and $B^{-1} = E[\widetilde{B}^{-1}]$.

2. The Tensor $\widetilde{C}$

The i-j^{th} term of $\widetilde{A}\,\Phi(x_{k+n-1})\,\Phi^{T}(V_{k+n-1})$ is given by

$$\sum_{k=1}^{s} \widetilde{a}_{ik} \, \Phi^{k}(x_{k+n-1}) \, \Phi^{j}(V_{k+n-1}) \; . \tag{A-23}$$

we introduce tensor $\widetilde{C}$ by defining its components as

$$\widetilde{C}^{j}_{ik} = \widetilde{a}_{ik} \Phi^{j}(V_{k+n-1}) \; . \tag{A-24}$$

Then clearly

$$\widetilde{A}\,\Phi(x_{k+n-1})\Phi^{T}(V_{k+n-1}) = \widetilde{C}\,\Phi(x_{k+n-1}) \; . \tag{A-25}$$

APPENDIX III

Acceleration of Convergence

Most stochastic approximation algorithms can be accelerated if the noise variances (or the bounds on these variances) entering the error equation are known.

Let

$$k_j = 1 + (n+1)j \quad ; \quad j = 1,2,\ldots \tag{A-26}$$

and

$$\rho_{k_j} = \rho\left(\frac{k_j}{n+1}\right) \equiv \rho(j) \quad ; \quad j = 1,2,\ldots \tag{A-27}$$

Then the equation (A-10) can be written as

$$v_{k_j} \leq [1-2\rho(j-1)\lambda_o + \rho^2(j-1)\gamma_1]V_{k_{j-1}} + \rho^2(j-1)\gamma_2 \tag{A-28}$$

where

$$v_{k_j} = E[\hat{e}_{k_j}^T e_{k_j}] \tag{A-29}$$

For the acceleration of convergence, we choose $\rho(j-1)$ which minimizes the right hand side of (A-28). By simple differentiation we get the optimum $\rho^*(j-1)$

$$\rho^*(j-1) = \frac{\lambda_o V_{k_{j-1}}}{\gamma_1 V_{k_{j-1}} + \gamma_2} \tag{A-30}$$

which by iteration leads to

$$\rho^*(0) = \frac{\lambda_o V_{k_o}}{\gamma_1 V_{k_o} + \gamma_2} \tag{A-31}$$

$$\rho^*(j) = \lambda_o \frac{1}{\gamma_1 + \gamma_2/[\rho(j-1)\{1 - \lambda_o\rho(j-1)\}]} \tag{A-32}$$

In general the constants λ_o, γ_1, γ_2, V_{k_o} may be unknown but, if they are known, the above gain sequence may be used to obtain accelerated convergence.

APPENDIX IV

Comparison of the Algorithm of (2-20) with the on-line Stochastic Approximation Algorithm of Reference 7 [Saridis and Stein]

In this appendix it is shown that the algorithm of (2-20) reduces to the algorithm given by Saridis and Stein for the case of a linear system.

Consider the case of a linear system with scalar state:

$$x_{k+1} = a\, x_k + u_k \tag{A-26}$$

$$z_k = x_k + v_k \tag{A-27}$$

where

a is the parameter which has to be estimated to identify the system. In the notation of Ref. [7], the quantities D, d* are all identically zero and hence the algorithm reduces to

$$\hat{a}_{k+1} = \hat{a}_k + \rho(k)\ \{z_k\ [z_{k+1} - \hat{a}_k z_k] + \alpha_2 \hat{a}_k\} \tag{A-28}$$

In the algorithm of (2-20) since $h = 1$, we have:

$$\Phi(z_k/h)\Phi(z_k/h) = (z_k)^2 \tag{A-29}$$

$$O_k = 2(z_k - \alpha_1)\alpha_1 + \alpha_2 \tag{A-30}$$

$$H_1 = z_{k+1}[z_k - \alpha_1] \tag{A-31}$$

$$H_2 = (\beta_1 + \alpha_1)[z_k - \alpha_1] \tag{A-32}$$

Thus the algorithm becomes

$$\hat{a}_{k+1} = \hat{a}_k + \rho(k)[\{H_1\} - \{H_2\} - \{\Phi(z_k/h)\Phi^T(z_k/h) - O_k\}\hat{a}_k] \quad \text{(A-33)}$$

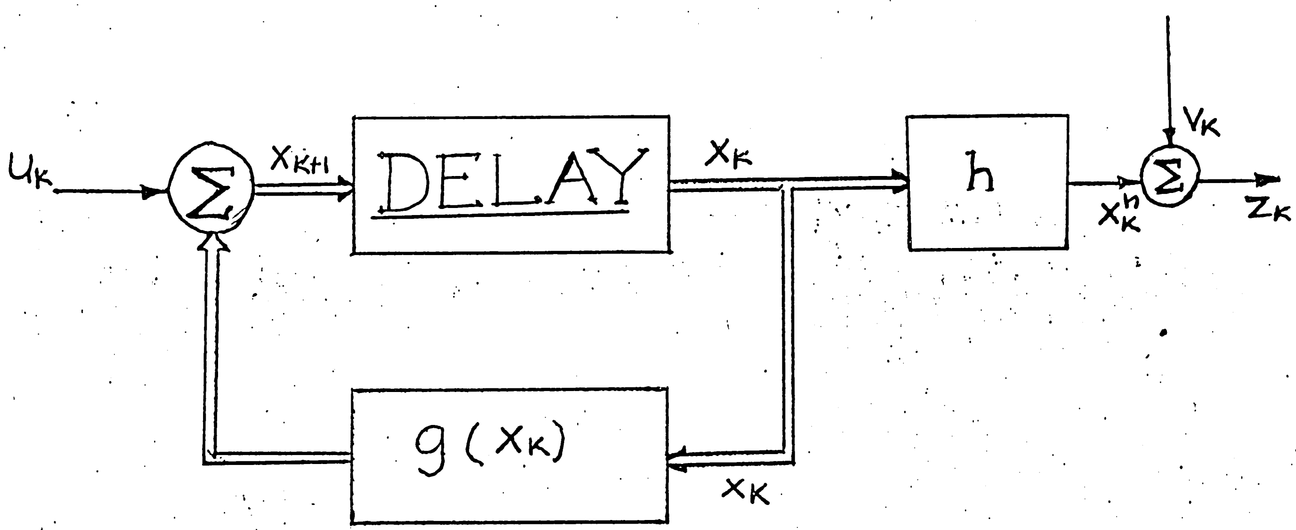
Using (A-29), (A-30), (A-31), (A-32)

$$= \hat{a}_k + \rho(k)[z_{k+1}(z_k-\alpha_1)-(\beta_1+\alpha_1)[z_k-\alpha_1]-\{(z_k)^2-2(z_k-\alpha_1)\alpha_1-\alpha_2\}\hat{a}_k] \quad \text{(A-34)}$$

Since in reference [7] the first moments of u_k and v_k are assumed to be zero, we have

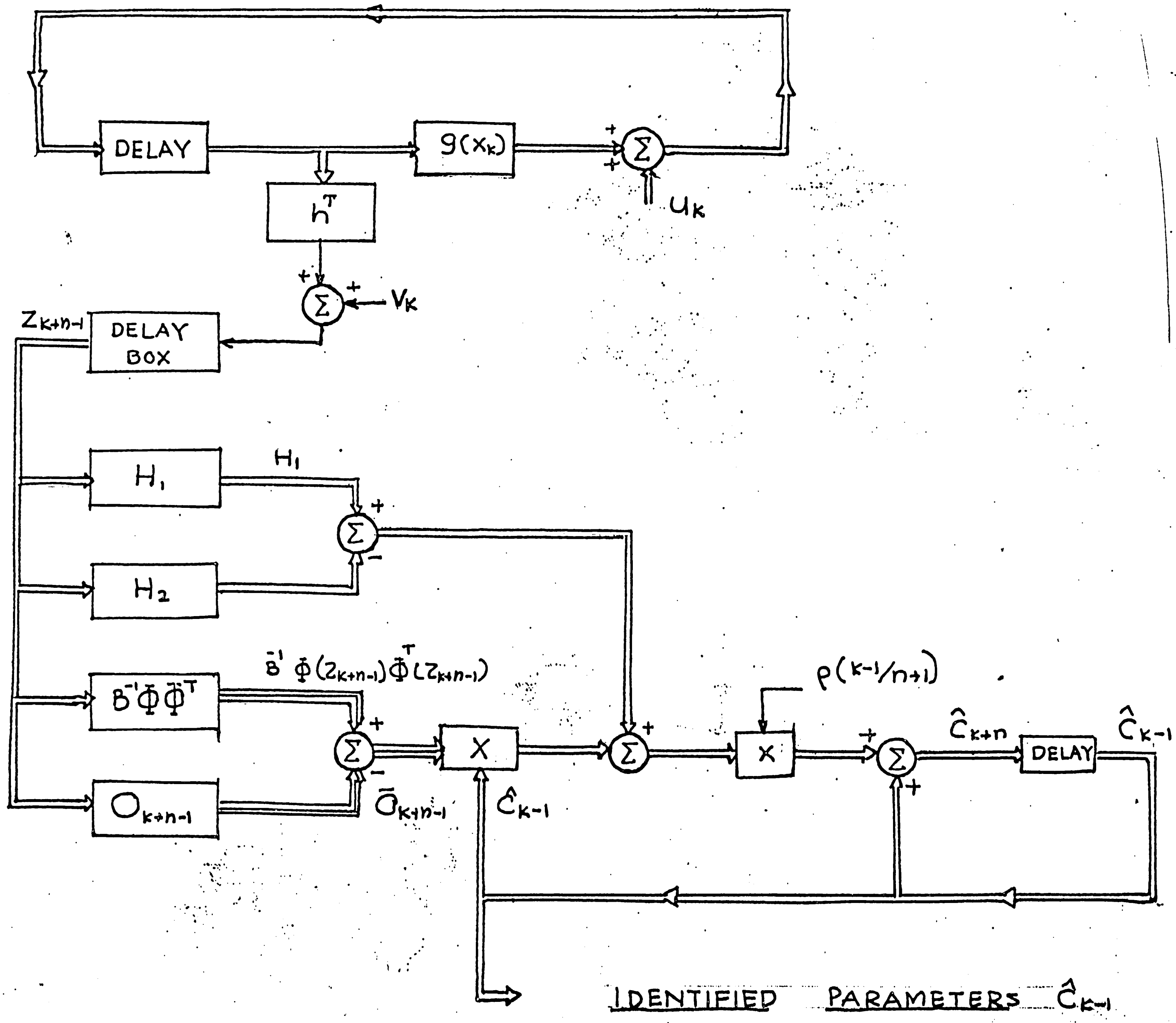
$$\hat{a}_{k+1} = \hat{a}_k + \rho(k)[z_k(z_{k+1} - \hat{a}_k z_k) + \alpha_2\hat{a}_k] \quad \text{(A-35)}$$

(A-35) and (A-28) are identical, which shows that the on-line algorithm of Saridis and Stein is a special case of the algorithm presented here.



THE SYSTEM TO BE IDENTIFIED

FIG. 1



$h^T = [0, 0, \cdots, 1]$

$U_k^T = [0, 0, \ldots\ldots u_k]$

→ Scalar

⇒ Vector

⇛ Matrix

ON LINE IDENTIFICATION SCHEME

FIG. 2

REFERENCES

1. Albert, A.E. and L.A. Gardner, Jr., "Stochastic Approximation and Nonlinear Regression," M.I.T. Press Research Monograph, No. 42, 1967.

2. Saridis, G.N., Z.J. Nikolic and K.S. Fu,"Stochastic Approximation Algorithms for System Identification, Estimation and Decomposition of Mixtures," Proc. of the 5th Annual Allerton Conf., pp. 374-384, Urbana, Ill., 1967.

3. de Figueiredo, R.J.P. and L.W. Dyer, "Extensions of Discrete Stochastic Approximation with Dynamics and Applications to Nonlinear Filtering," Rice University Technical Report, EE-67, No. 2, August 1967.

4. Ho, Y.C. and R.C.K. Lee, "Identification of Linear Dynamic Systems," Journal of Information and Control, vol. 8, pp. 93-110, February 1965.

5. Lee, R.C.K., "Optimal Estimation, Identification and Control," M.I.T. Press Research Monograph, No. 28, 1964.

6. Sakrison, D.J., "The Use of Stochastic Approximation to Solve the System Identification Problem," IEEE Transactions on Automatic Control, vol. AC-12, No. 5, pp. 563-567, October 1967.

7. Saridis, G.N. and G. Stein, "Stochastic Approximation Algorithms for Linear Discrete System Identification," IEEE Transactions on Automatic Control, vol AC-13, No. 5, October 1968 (to appear).

8. Dvoretzky, A. "On Stochastic Approximation," Proc. of the Third Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, vol. 1, pp. 39-55, 1956.