ated with the eigenvalues of $M(0)$ other than $\lambda^i(0)$. The following lemma gives the derivatives.

Lemma 1: The m first derivatives $\lambda_j^{i'}(0)(j = 1,\cdots, m)$ of a multiplicity m eigenvalue $\lambda^i(0)$ are the m eigenvalues of the matrix

$$(Y^i)^* M'(0) Y^i \tag{7}$$

where $M'(t)$ denotes the derivative of $M(t)$ with respect to t. In addition, if $t = 0$ is a regular point, then the m second derivatives $\lambda_j^{i''}(0)(j = 1,\cdots, m)$ are the eigenvalues of

$$(Y^i)^*\left(M''(0) + 2M'(0)Z^i(\lambda_j^i(0)I - \Lambda_1^i)^{-1} (Z^i)^* M'(0)^*\right) Y^i. \tag{8}$$

Proof: This is a special case of a result proven in [8].

Lemma 2: Let $A(t) = e^{Dt}A(0)e^{-Dt}$, let $M(t) = \dfrac{A(t) + A^*(t)}{2}$, let D be an invertible diagonal matrix $\in \mathscr{R}^{n\times n}$ and let $Y(t)$ be a matrix whose columns span the eigenspace of $M(0)$ corresponding to $\lambda_j^i(0)$. Then

$$(Y^i)^* M''(0) Y^i = (2Y^i)^* D(\lambda_j^i(0)I - M(0))DY^i. \tag{9}$$

Proof: The derivatives $A'(0)$ and $A''(0)$ are give as follows:

$$A'(0) = DA(0) - A(0)D$$

$$A''(0) = D^2A(0) - 2DA(0)D + A(0)D^2.$$

Hence

$$\begin{aligned}(Y^i)^* M''(0) Y^i &= (Y^i)^*(D^2M(0) - 2DM(0)D + M(0)D^2)Y^i\\ &= (2Y^i)^*D(\lambda_j^i(0)I - M(0))DY^i\end{aligned}$$

where the latter equality follows via the identity

$$M(0)Y^i = \lambda_j^i(0)Y^i.$$

Equation (9) follows. Q.E.D.

The main result can now be proved.

Theorem 1: The functional

$$\lambda_{\max}\left(\frac{e^DAe^{-D} + (e^DAe^{-D})^*}{2}\right) \tag{10}$$

is convex in D.

Proof: To establish the convexity of the functional (10) of D, it suffices to prove that for any $D \in \mathscr{R}^{n\times n}$, the function

$$\lambda_{\max}\left(\frac{e^{Dt}Ae^{-Dt} + (e^{Dt}Ae^{-Dt})^*}{2}\right)$$

is convex in t. Using Lemma 2, it can be easily checked that (8) is positive semidefinite for all

$$t \in T_i \triangleq \left\{t \in \mathscr{R} \mid \lambda^i(t) = \max_k \lambda^k(t)\right\}.$$

Hence, if $0 \in T_i$ then $\lambda^{i''}(0) \geq 0$. But $t = 0$ is not a distinguished point, so $\lambda^{i''}(t) \geq 0$ for all $t \in T_i$. It follows that $\lambda^i(t)$ is convex on T_i. Since $\bigcup_i T_i = \mathscr{R}$ and since the maximum of several convex functions is convex it follows that $\lambda_{\max}(t)$ is convex. Q.E.D.

III. Conclusions

The principal conclusion of the note is the proof that the minimization

$$\min_D \lambda_{max}\left(\frac{e^DAe^{-D} + (e^DAe^{-D})^*}{2}\right)$$

is convex in D. In theory, this implies that a steepest descent optimization method for finding the minimizing D should converge globally, through practical considerations such as step size and tolerances for detection of sharp gradient changes could drastically affect convergence rates. Nevertheless, the authors have developed a robust computer algorithm [12] which reliably overcomes these difficulties using a Davidon–Fletcher–Powell-type modification of the generalized-gradient steepest-descent approach. Robustness analysis was performed on lateral directional flight control system designs with large uncertainty [12]. For a system with one-sided perturbations, the one-sided MSM provides a better robustness measure than using the traditional two-sided MSM.

References

[1] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties.* New York: Academic, 1975.

[2] J. C. Doyle, "Robustness of multiloop linear feedback systems," in *Proc. 24th IEEE Conf. Decision Contr.*, Ft. Lauderdale, FL., 1978.

[3] —, "Analysis of feedback systems with structured uncertainties," *IEE Proc.*, vol. 129, p. D, pp. 242–250, 1982.

[4] —, "Structured uncertainty in control system design," in *Proc. IEEE Conf. Decision Contr.*, Ft. Lauderdale, FL, 1985.

[5] M. K. Fan and A. L. Tits, "Characterization and efficient computation of the structured singular value," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 734–743, 1986.

[6] T. Kato, *Perturbation Theory of Linear Operators.* New York: Springer-Verlag, 1976.

[7] M. Safonov, A. Laub, and G. Hartmann, "Feedback properties of multivariable systems: The role and use of the return difference matrix," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 47–65, 1981.

[8] M. G. Safonov and J. Doyle, "Minimizing conservativeness of robustness singular values," in *Multivariable Control*, S. G. Tzafestas, Ed., Dordrecht, The Netherlands: Reidel, 1984, pp. 197–207.

[9] M. G. Safonov, "Stability margins of diagonally perturbed multivariable feedback systems," *IEE Proc.*, vol. 129, p. D, pp. 251–256, 1982.

[10] M. G. Safonov and M. Athans, "A multiloop generalization of the circle criterion for stability margin analysis," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 415–422, 1981.

[11] R. S. Sezginer and M. L. Overton, "The largest singular value of $e^XA_0e^{-X}$ is convex on convex sets of commuting matrices," *IEEE Trans. Automat. Contr.*, vol. 35, no. 2, pp. 229–230, 1990.

[12] J. A. Tekawy, M. G. Safonov, and R. Chiang, "Algorithms for computing multivariable stability margin," *Proc. 3rd Annual Conf. Aero. Computational Contr.*, Oxnard, CA, 1989.

[13] N.-K. Tsing, "Convexity of the largest singular value of e^DMe^{-D}: a convexity lemma," *IEEE Trans. Automat. Contr.*, vol. 35, no. 6, p. 748–749, 1990.

[14] G. Zames, "On the input-output stability of time-varying nonlinear feedback systems—Parts I and II," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 228–238 and pp. 465–476, 1966.

Estimation of Dynamical-Varying Parameters by the Internal Model Principle

G. Davidov, A. Shavit, and Y. Koren

***Abstract*—A novel design method of recursive algorithms for identification of linear deterministic SISO stable discrete systems with dynami-**

Manuscript received March 16, 1990; revised January 4, 1991.

The authors are with the Department of Mechanical Engineering, Technion—Israel Institute of Technology, Haifa, Israel.

IEEE Log Number 9105149.

0018-9286/92$03.00

cal-varying parameters is presented. An algorithm for parameter identification of such systems, based on the known internal model principle and on the recursive least squares parameter estimation, is proposed. The system parameters are assumed to satisfy a linear difference equation with constant coefficients. A persistent excitation condition of the measurement vector automatically guarantees exponential stability and therefore there is no need to use any resetting procedures. This condition is similar in form to the observability gramian property of a linear time-varying system. Simulation and practical application of the algorithm on an experimental robot system show good tracking even when the parameters vary drastically and in an abrupt manner.

INTRODUCTION

The applicability of most of the recursive algorithms for the identification of system parameters is limited to the estimation of constant or slowly time-varying parameters.

Several advanced recursive algorithms for identification of linear fast time-varying (LFTV) systems have recently been published [1]–[3]. Basically, the design strategy of these methods is to combine recursive algorithms of time-invariant systems with an approximation of the variation of the parameters as polynomials in time.

The idea of approximating time-varying parameters with time polynomials is not new and was vastly applied to batch identification algorithms. Various kinds of polynomials have been employed such as the Chebyshev polynomial [4], the Legendre polynomial [5], the shifted Laguerre polynomial [6], and the generalized orthogonal polynomial [7]. Such applications suffer from lack of convergence problems, since the polynomials are made to fit a fixed set of measurements. The identification must be carried out, however, off-line on previously obtained data. Batch procedures could not, therefore, be used for the purpose of real-time tracking and predictions of time-varying parameters, neither could they be incorporated in a real-time adaptive control scheme. The need for a recursive algorithm that could be used on-line is obvious.

Xiania and Evans [1] and Zheng [2] proposed recursive algorithms for the identification of time-varying parameters. These algorithms did not guarantee boundedness of estimated parameter vectors even when the persistent excitation (PE) condition was satisfied. In order to get such boundedness, they had to use resetting procedures. In an attempt to overcome the boundedness problem, Hersh [3] used a bounded semiperiodical time functions model. He assumed that the period of those functions was known. This assumption is nontrivial, since such specific and detailed information on the system parameters is not always available.

We propose a novel methodology by which time-varying parameters can be identified in real time without the need for resetting procedures or special time function models. A new algorithm named internal model least squares (IMLS), which is based on the internal model principle [8], was developed. This principle was originally proposed as a solution to the servomechanism problem. It states that a servosystem can track command input signals and suppress input disturbances if its controller includes an internal model which represents the nature of the dynamics of those inputs.

Basically, our proposed identification scheme is based on the observer theory, specifically on the fact that an observer for a dynamical system should incorporate the dynamics of the observed system. The problem, however, is how to combine the parameters' dynamics with the system equation. We employed the internal model principle. Accordingly, discrete transfer functions or adequate difference equations, which duplicate the parameters' dynamics, should be merged with the system equations. Thus, our technique differs from the previously mentioned works by the fact that it does not employ time functions to describe the variation of the parameters.

It is shown that an identification algorithm can better track fast time-varying parameters if it includes a suitable internal model for the time variations of these parameters. The model for the dynamical variation of the parameters may accept any function that satisfies a linear difference equation of the form $D(q^{-1})\theta(t) = 0$, where $\theta(t)$ is the modeled parameter, D is a linear difference equation of some order n, and q^{-1} is a unit time delay.

A priori information of the variation of the parameters, such as natural frequency, exponential time constant, etc., may be easily incorporated into the model, and enhance the identification process. A proper formulation of this model, combined with the Kalman filter theory, yields a theory of LFTV parameter identification that is similar in form to the theory for the state estimation of a linear time-varying system. As a consequence, the PE condition which is essential to the quality of the identification process is equivalent to the observability gramian property [9] of a linear time-varying system.

The PE condition automatically ensures global stability of the IMLS algorithm and exponential convergence to zero of the identification error [10], [11]. Furthermore, there is no need to use any stabilizers or parameter resetting techniques. The convergence proof of the IMLS algorithm is based on the assumption that the dynamics of the parameters is known. However, the performance of the IMLS algorithm is demonstrated by simulation and by application to an industrial robot, where successful identification of the parameters was also obtained for an approximate dynamical model of the parameters.

PROBLEM FORMULATION

We assume that the system to be identified is SISO, stable, and can be described by a time-varying deterministic ARMA (DARMA) model of a known order. Following standard notation we write

$$y(t) = \phi^T(t-1)\theta(t-1) \qquad (1)$$

where t is an integer that denotes the time steps, and

$$\phi^T(t-1) = [y(t-1), \cdots, y(t-n), u(t-1), \cdots, u(t-m)] \qquad (2)$$

$$\theta^T(t-1) = [\alpha_1(t-1), \cdots, \alpha_n(t-1), \beta_1(t-1), \cdots, \beta_m(t-1)] \qquad (3)$$

$y(t)$ and $u(t)$ are the system output and input, respectively.

$$\phi^T(t) \in \mathbb{R}^{1,(n+m)} \qquad \text{is the measurement vector}$$

and

$$\theta^T(t) \in \mathbb{R}^{1,(n+m)}$$

is the vector of the time-varying system parameters.

The general variation with time of the parameters $\theta(t)$ is assumed to satisfy a linear difference equation with constant coefficients

$$D(q^{-1})\theta(t) = 0 \qquad (4)$$

where q^{-1} denotes a unit time delay. The matrix $D(q^{-1})$ is of the form

$$D(q^{-1}) = \text{diag}[W_1(q^{-1}), \cdots, W_{n+m}(q^{-1})] \qquad (5)$$

each of the $W_i(q^{-1})$ are polynomials of the order L

$$W_i(q^{-1}) = 1 + \sum_{j=1}^{L} d_{i,j} q^{-j} \qquad i = 1, \cdots, n+m. \qquad (6)$$

The terms $d_{i,j}$ are constants to be selected according to the available *a priori* information, or as an approximation thereof. In addition, in order to prevent singularity problems, the last terms must be nonzero

$$d_{i,L} \neq 0. \qquad (7)$$

The information in (4) may be combined with the DARMA model of (1), using a state-space formulation. Hence, a new measurement vector, $\Phi(t)$, an auxiliary matrix C, and a new parameter vector $\Theta(t)$ are defined as follows:

$$\Phi^T(t) \equiv [\phi^T(t), 0, 0, \cdots, 0] \qquad \Phi \in \mathbb{R}^{1, L(n+m)} \tag{8}$$

$$C \equiv [I, 0, 0, \cdots, 0] \qquad C \in \mathbb{R}^{(n+m), L(n+m)} \tag{9}$$

and

$$\Theta^T(t) \equiv [\theta^T(t), \theta^T(t-1), \cdots, \theta^T(t-L+1)] \qquad \Theta \in \mathbb{R}^{1, L(n+m)}. \tag{10}$$

It is obvious from the above definitions that the system model (1) can be rewritten as

$$y(t) = \Phi^T(t-1)\Theta(t-1) \tag{11}$$

and that the parameter vector $\Theta(t)$ is related to the previous $\Theta(t-1)$ by a constant matrix A such that

$$\Theta(t) = A\Theta(t-1) \tag{12}$$

where A is the internal model matrix, defined as

$$A = \begin{Bmatrix} -D_1 & -D_2 & \cdots & -D_L \\ I & 0 & & 0 \\ \vdots & & & \vdots \\ 0 & ,\cdots, & I & 0 \end{Bmatrix} \tag{13}$$

where $A \in \mathbb{R}^{(n+m)L, (n+m)L}$ and $I \in \mathbb{R}^{n+m, n+m}$. The elements D_i of the matrix A are

$$D_i = \operatorname{diag}[d_{i,1}, \cdots, d_{i,n+m}]. \tag{14}$$

The identification algorithm must be properly synthesized. In cases where the internal model describes the dynamical variations of the parameters perfectly, it is sufficient to estimate only the initial parameter vector $\Theta(t_o)$, where t_o is the initial time. In cases where the internal model is just an approximation of the real dynamical variation of $\theta(t)$, the problem becomes to estimate recursively $\theta(t)$ with minimal identification error.

The IMLS Algorithm

The problem of parameter identification, as described by (11), (12) has a similar form to the problem of the state vector estimation of a linear time-varying system [12]. The vector $\Theta(t)$ in (12) stands for the state vector A in the system matrix. Similarly, (11) may be viewed as an output equation, where $y(t)$ is the output variable and $\Phi(t)$ is the output system matrix. The observer theory of D. Luenberger [12] can now be used to create the following equation for observing the "state vector" $\Theta(t)$:

$$\hat{\Theta}(t) = A\hat{\Theta}(t-1) + K(t-1)e(t) \tag{15}$$

where the error $e(t)$ is

$$e(t) = y(t) - \Phi^T(t-1)\hat{\Theta}(t-1) \tag{16}$$

and $K(t)$ is the observer gain, selected such that the process, $z(t)$, given by

$$z(t) = (A - K(t-1)\Phi^T(t-1))z(t-1); \qquad z(t) \in \mathbb{R}^{1, L(n+m)}; \quad z(0) = z_0 \tag{17}$$

is stable. Such an observer may estimate the state vector $\Theta(t)$ if the system has the observability gramian property. Following the development of Meditch [13], this property can be easily shown to be

$$\sum_{j=0}^{s-1} A^{-T(j+1)}\Phi(t-j)\Phi^T(t-j)A^{-(j+1)} \geq a_1 I > 0 \tag{18}$$

where s is a positive scalar and a_1 is a positive constant.

A stabilizing gain vector $K(t)$ is proposed in the present IMLS algorithm as

$$K(t) = AP(t-2)\Phi(t-1)\gamma(t-1) \tag{19}$$

where

$$P(t-1) = A[P(t-2) - \gamma(t-1)P(t-2)\Phi(t-1) \cdot \Phi^T(t-1)P(t-2)]A^T + cI, \tag{20}$$

c is an arbitrary positive constant and γ is defined below

$$\gamma(t-1) = (1 + \Phi^T(t-1)P(t-2)\Phi(t-1))^{-1} \tag{21}$$

$$P(0) = k_o I \tag{22}$$

where k_o is a positive scalar, and

$$\hat{\Theta}(0) = \hat{\Theta}_0 \tag{23}$$

is the best *a priori* knowledge of the vector $\Theta(0)$.

The gain vector $K(t)$ is similar in form to the Kalman filter gain vector [14]. The term cI in (20) takes the place of the covariance system noise matrix Q of the Kalman filter theory. The inclusion of the term cI in the equation guarantees that the covariance matrix, P, does not shrink to zero.

It was shown in [11], that c is related to the lower bound of the covariance matrix P. The constant c in (20) is arbitrary but still care must be exercised in its selection. If c is selected too low then the gain vector $K(t)$ is also low and the tracking ability is limited. If, on the other hand, c is selected too high then the algorithm becomes sensitive to noise and to numerical errors due to the high gain. Thus, a trial and error procedure is indicated for the best selection.

The main properties of the algorithm are summarized in the following theorem.

Theorem: If the following PE condition holds:

$$\sum_{j=0}^{s-1} A^{-T(j+1)}\Phi(t-j)\Phi^T(t-j)A^{-(j+1)} \geq a_1 I > 0 \tag{24}$$

where s is a positive scalar and a_1 is a positive constant (note that similar conditions were given by Zheng [2] and by Hersh [3]), then

1) the covariance matrix $P(t)$ is bounded and does not shrink to zero;

2) $$\lim_{t\to\infty} \|\tilde{\Theta}(t)\| = 0 \tag{25}$$

where $\tilde{\Theta}(t)$ is the parameter identification error, defined by

$$\tilde{\Theta}(t) \triangleq \hat{\Theta}(t) - \Theta(t) \tag{26}$$

3) $$\lim_{t\to\infty} \gamma(t-1)e(t) = 0 \tag{27}$$

which implies that the normalized error $e(t)$ shrinks to zero.

4) $$\lim_{t\to\infty} \|\hat{\Theta}(t) - A\hat{\Theta}(t-1)\| = 0 \tag{28}$$

which implies that the identified parameters behave as the real parameters do. The proof of the theorem is given in [11]. It is pretty standard and is omitted here.

Simulation and Application to a Robot

The capabilities of the IMLS algorithm are demonstrated via a simulation and via application to an industrial robot.

Simulation

The following one-step-ahead predictor model was used:

$$y(t) = [\alpha(t-1), \beta(t-1)][y(t-1), u(t-1)]^T. \tag{29}$$

The input $u(t)$ is a random Gaussian sequence with 6 Hz bandwidth and the sampling time $T = 0.05$ s. The following parameters have been selected to reflect rapid variation:

$$\alpha(t) = 0.5 \cdot \sin(0.4t) \qquad \text{for } 0 \le t \le 120 \tag{30}$$

where $t = \text{int}(\tau/T)$ and τ is the elapsed time

$$\beta(t) = \begin{cases} 5 + 1.5t & \text{for} \quad 0 \le t \le 40 \\ 65 & \text{for} \quad 40 < t < 80 \\ 65 - 0.0375(t-80)^2 & \text{for} \quad 80 \le t \le 120 \end{cases}. \tag{31}$$

For purpose of comparison, the simulations of two cases were conducted: ordinary recursive least squares with constant covariance modification (RLS/CCM) algorithm, [10], [11] and the proposed IMLS algorithm.

The initial values for the two cases were selected as follows:

Case 1: RLS/CCM algorithm $\hat{\theta}(0) = 0$, $c = 1$, and $P(-2) = I$;

Case 2: IMLS algorithm $\hat{\Theta}(t) = 0$, $c = 1$, and $P(-2) = I$.

Using the information in (30), (31) leads to the selection of the following 6×6 matrix A for the internal model

$$A = \begin{bmatrix} 2.34 & 0 & -2.34 & 0 & 1 & 0 \\ 0 & 3 & 0 & -3 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Matrix A is obtained by using difference equations which describe the dynamics of the parameters $\alpha(t)$ and $\beta(t)$. The technique is presented in the section on problem formulation. It can be shown that the difference polynomial operator for parameter $\alpha(t)$ is

$$(1 - q^{-1})(1 - 1.34q^{-1} + q^{-2})$$

where the term $(1 - q^{-1})$ was added in order to prevent singularity problems (see (7)). The difference polynomial operator for parameter $\beta(t)$ is

$$(1 - q^{-1})^3.$$

For the purpose of demonstration of the capability of the algorithm to a more general case, this *a priori* information is disregarded and the parameters α and β are approximated by ramp functions leading to the following 4×4 approximate internal model matrix A:

$$A = \begin{Bmatrix} 2I & -I \\ I & 0 \end{Bmatrix} \qquad I, 0 \in R^{2,2}. \tag{32}$$

Simulation results of Case 1 are given in Figs. 1–3. The parameter α in (30) and the estimated $\hat{\alpha}$ are shown in Fig. 1, the parameter β in (31) and the estimated $\hat{\beta}$ are shown in Fig. 2, and the trace of matrix $P(t)$ is given in Fig. 3. The corresponding IMLS simulation results (Case 2) are given in Figs. 4–6 and demonstrate excellent tracking ability. It can be seen from these results that when β is constant the difference between the two cases is insignificant. However, if β varies more rapidly, such as ramp or acceleration functions, the IMLS performs much better.

Selecting a higher order model to describe the parameters α and β could further reduce the tracking errors as compared to those presented in Figs. 4 and 5. It also increases, however, the computational effort and possibly the sensitivity to noise. Figs. 3 and 6 show that, for rich input, the trace of the covariance matrix is bounded which is consistent with the theorem.

Implementation on a Robot

The algorithm was applied to identify the parameters of a two link Scara-type robot of the Hirata company, which is shown in Fig. 7. Each link was driven by a dc motor whose velocity was controlled by a servo loop with PI (proportional and integrator) controller. A mass of about 5 kg was mounted on the end effector. The bandwidth

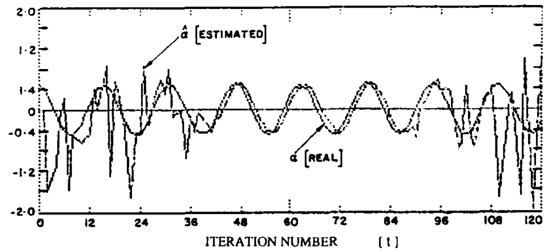

Fig. 1. Parameter α: Estimated versus real (RLS/CCM).

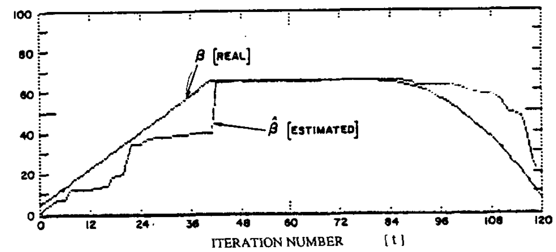

Fig. 2. Parameter β: Estimated versus real (RLS/CCM).

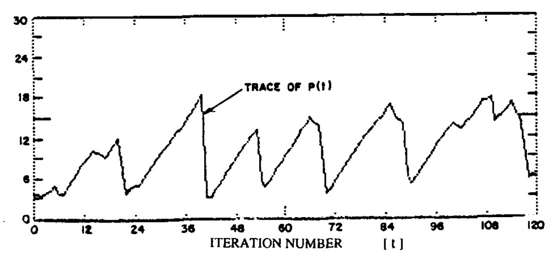

Fig. 3. Trace of covariance matrix $P(t)$ (RLS/CCM).

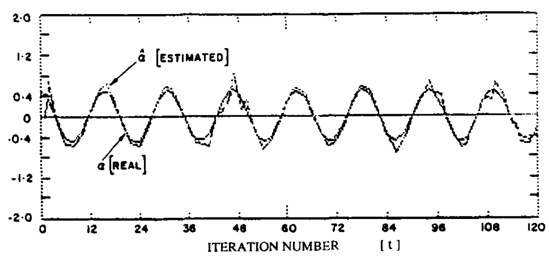

Fig. 4. Parameter α: Estimated versus real (IMLS).

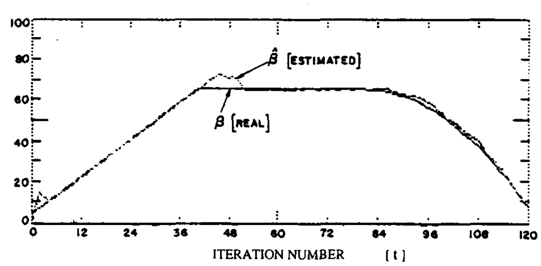

Fig. 5. Parameter β: Estimated versus real (IMLS).

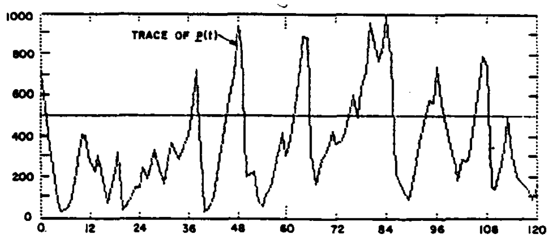

Fig. 6. Trace of covariance matrix $P(t)$ (IMLS).

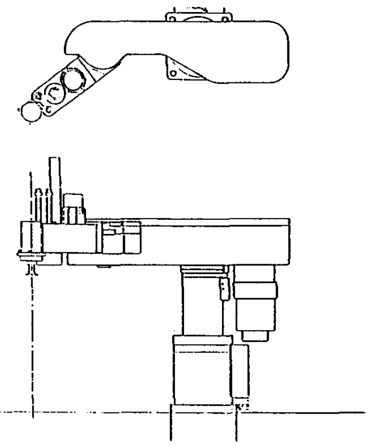

Fig. 7. The Hirata robot.

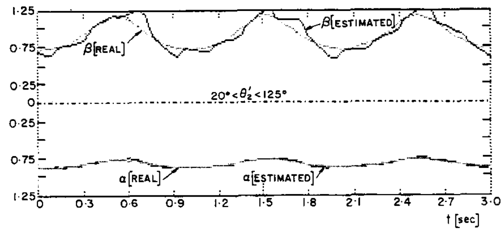

Fig. 8. Simulation results of Hirata robot parameter identification by IMLS algorithm.

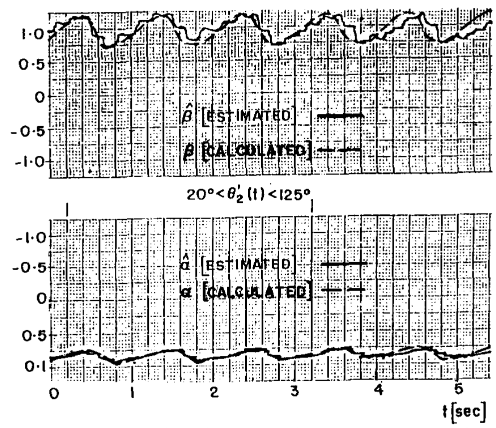

Fig. 9. Laboratory results of Hirata robot parameter identification by IMLS algorithm.

of the base link velocity loop is influenced by the position of the second link. When the angle between the links was approximately 10° the bandwidth was measured to be 10 Hz and for an angle of 170° the bandwidth was 4 Hz. For sufficiently fast sampling rates the dynamics can be approximately described by the following first-order time-varying discrete equation:

$$w(k+1) = \alpha(k)w(k) + \beta(k)u(k) \tag{33}$$

where $w(k)$ is the base link velocity measured by the tachometer, and $\alpha(k)$ and $\beta(k)$ are the time-varying system parameters. Equation (33) was derived by first analyzing by FFT the system response and obtaining the transfer function of the robot at different values of the relative angle between the links $\psi(k)$. Then discretization of the analog transfer function using the Z-transform.

The parameters $\alpha(k)$ and $\beta(k)$ were found to be best fitted by the following relations

$$\begin{aligned} \alpha(k) &= \exp\left[-T/\tau(k)\right] \\ \beta(k) &= 4.6\left[1-\alpha(k)\right] \\ \tau(k) &= 0.0536 + 0.0443\cos\psi(k) \end{aligned} \tag{34}$$

The IMLS algorithm was capable to identify the parameters $\alpha(k)$ and $\beta(k)$ even for a case where the angle $\psi(k)$ was varied between 20° to 125° with a frequency of 1 Hz with a sampling rate of 10 ms and the constant [of (20)] set to $c = 10^{-3}$. The IMLS algorithm was implemented in assembler language on a Digital PDP-11/23 minicomputer. The internal model matrix A was selected to have the same ramp functions as given in (32). The input to the system $u(k)$ was selected to be harmonic of amplitude of 18°/s and 3 Hz. The frequency and amplitude of the excitation signal must be carefully selected. The frequency must be higher than that of the variation of the angle $\psi(k)$ but not too high as to cause rattling in the gear teeth. The amplitude need also be high enough to minimize the backlash effects but not too high as to cause saturation of the control system.

Simulation results are given in Fig. 8 and results obtained from the robot are given in Fig. 9. It was observed that a single frequency harmonic input function was sufficient to excite the system and the algorithm to yield satisfactory identification quality. Identification results on the robot are quite compatible with those of the simulations.

Discussion

It was shown in [11], that c is related to the lower bound of the covariance matrix P. The constant c in (20) is arbitrary but still care must be exercised in its selection. If c is selected too low then the gain vector $K(t)$ is also low and the tracking ability is limited. If, on the other hand, c is selected too high then the algorithm becomes sensitive to noise and to numerical errors due to the high gain. Thus, a trial and error procedure is indicated for the best selection.

The theorem was strictly proven for the case when the internal model represents accurately the time variation of the parameters and the condition for PE was satisfied.

If the PE condition is not satisfied, the algorithm can not be used to accurately identify the real parameters. In the abscence of PE, the algorithm would be unstable if the matrix A is not stable. Stability of the algorithm may still be achieved by selecting a stable matrix A or by using resetting procedures.

In cases where the model does not describe exactly the actual variations in the parameters, so that certain deviations do exist, then it is still possible to apply an approximate internal model and obtain reasonable results. Several options for applying the approximate model follow.

1) The internal model (12) can be extended to the stochastic case as:

$$\hat{\Theta}(t) = A\hat{\Theta}(t-1) + W(t-1). \tag{35}$$

In this equation the deviation from the exact internal model is described by a white noise $W(t-1)$ with a covariance matrix $c\boldsymbol{I}$.

2) If no *a priori* information (such as time constant, natural frequency, etc.) about the parameters variation is given, A may be selected such that the parameters have a polynomial form of some order. The algorithm will then find the best coefficients of the polynomial that approximate the parameters. For example, if a ramp function (polygonal lines) is selected as an approximation, the algorithm will then find the best slope in a finite time interval using the last two successive samples of each parameter. If the variations in the slopes are large then identification errors may occur. These errors can be minimized by increasing the sampling rate.

3) The identification quality may be improved by selecting polynomials of higher order for the approximation and by increasing the sampling rate. For example, the eigen polynomial of an harmonic function, namely, $[1 - 2 \cdot \exp(-\omega T)q^{-1} + q^{-2}]$, where ω is the frequency and T is the sampling time, can be approximated by the eigen polynomial of a ramp function, if T is sufficiently small, such that $[\text{ext}(-\omega T) \to 1]$. If the high-sampling rate still does not improve the identification quality, then the order of the approximating polynomials has to be increased.

4) Minimization of the identification errors may be achieved by a selection of higher order of the approximating polynomials. Hence, similar to the results of Zheng [2], a selection of very high-order will increase the computational burden, large numerical errors may occur, and the identification may be resonative. It is obvious that the appropriate sampling rate and the order of the approximating polynomials must be selected by a trial and error procedure.

5) If the measurements are noisy (the stochastic case), such that the output equation (11) can be rewritten as follows:

$$y(t) = \Phi^T(t-1)\Theta(t-1) + w(t) \tag{36}$$

where $w(t)$ is a white noise with variance r then the IMLS algorithm can be suited to this case by replacing in (9) the variance r instead of 1. The proof of the convergence is based on the Kalman filter theory [14] and it is omitted for the sake of brevity.

Conclusion

A new algorithm for identification of dynamically varying linear SISO systems was investigated, using an internal model to represent the system parameter variation. A new parameterization of the system equations yields, with the aid of Kalman filter theory, the desired algorithm. Boundedness of the covariance matrix, $\boldsymbol{P}(t)$, and the exponential convergence of the algorithm were proven under conditions of persistent excitation of the measurement vector in the time-varying case (21). The capability of the algorithm were demonstrated through results from a robot and by simulation results both achieved good tracking of fast and irregular time-varying parameters.

References

[1] X. Xiania and R. J. Evans, "Discrete-time adaptive control for deterministic time-varying systems," *Automatica*, vol. 20, no. 3, 1984.

[2] L. Zheng, "Discrete-time adaptive control of deterministic fast time-varying systems," *IEEE Trans. Automat. Contr.*, vol. AC-32, no. 5, May 1987.

[3] M. A. Hersh, "Adaptive identification and control of time-varying stochastic systems using 'Semi-Periodic' functions," *Int. J. Contr.*, vol. 47, no. 5, pp. 1243-1257, 1988.

[4] J. Chou and I. Hong, "Application of Chebyshev polynomials to the optimal control of time-varying linear systems," *Int. J. Contr.*, vol. 41, 1985.

[5] C. Huang and Y. P. Shih, "Parameter identification of time delay systems via shifted Legendre polynomials," *Int. J. Contr.*, vol. 41, 1985.

[6] C. Huang and M. Y. Chen, "Analysis and parameter identification via Laguerre polynomials," *Int. J. Contr.*, vol. 41, 1985.

[7] R. Y. Chang, S. Y. Yang, and M. L. Wang, "A new approach for parameters identification of time-varying systems via generalized orthogonal polynomials," *Int. J. Contr.*, vol. 44, no. 6, 1986.

[8] B. A. Francis and W. M. Wonham, "The internal model principle of control theory," *Automatica*, vol. 12, pp. 457-465, 1976.

[9] T. Kailath, *Linear Systems.* Englewood Cliffs, NJ: Prentice-Hall, 1980.

[10] G. Davidov, A. Shavit, and Y. Koren, "An internal model approach to the identification of time-varying systems," in *Proc. ICCON'89*, Jerusalem, Israel, Apr. 1989.

[11] G. Davidov, A. Shavit, N. Shimkin, and Y. Koren, "Identification of time-varying systems by internal model principle," presented at the 16th Conf. IEEE, Israel, Mar. 1989.

[12] D. Luenberger, "Observers for multivariable systems," *IEEE Trans. Automat. Contr.*, vol. 11, no. 2, pp. 190-197, 1966.

[13] J. S. Meditch, *Stochastic Optimal Linear Estimation and Control.* New York: McGraw-Hill, 1969.

[14] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *J. Basic Eng. Trans. ASME*, Ser. D., vol. 83, pp. 95-108, 1961.

Full and Reduced-Order Observer Design for Discrete Stochastic Bilinear Systems

Engin Yaz

***Abstract*—We consider a general discrete-time stochastic bilinear system model and derive the mean square optimal linear unbiased observer equations to be used in reconstructing a prespecified linear combination of state variables based on noisy output measurements.**

Introduction and Problem Statement

This note presents a reduced-order observer design procedure for system and measurement models which contain multiplicative as well as additive noise components. Such models are alternatively called stochastic bilinear, state-dependent noise models or multiplicative noise models. Due to the several application areas such as satellite attitude control [1], chemical reactor control [2], macroeconomics [3], population dynamics [4], time-sharing and random round-off errors in computer operation [5], and recently, robustness studies [6]-[8], there has been interest in estimator design [9]-[14] for various such models. In this study, we consider the general stochastic bilinear model used in the works [12]-[14] and extend the mean square optimal unbiased reduced-order observer results of [15] derived for deterministic parameter systems to this case.

Let us consider the discrete stochastic bilinear system

$$x_{k+1} = A_k(\omega)x_k + v_k(\omega) \tag{1}$$

Manuscript received March 23, 1990; revised August 10, 1990.

The author is with the Department of Electrical Engineering, University of Arkansas, Fayetteville, AR 72701.

IEEE Log Number 9105151.

0018-9286/92$03.00