ADAPTIVE FILTERING FOR LARGE SPACE STRUCTURES—A CLOSED-FORM SOLUTION

H. E. Rauch and **D. B. Schaechter**
Lockheed Palo Alto Research Laboratory
Palo Alto, CA 94304

ABSTRACT

In a previous paper Schaechter proposes using an extended Kalman filter to estimate adaptively the (slowly varying) frequencies and damping ratios of a large space structure. The present paper shows that the time-varying gains for estimating the frequencies and damping ratios can be determined in closed-form so it is not necessary to integrate the matrix Riccati equations. After certain approximations, the time-varying adaptive gain can be written as the product of a constant matrix times a matrix derived from the components of the estimated state vector. This is an important savings of computer resources and allows the adaptive filter to be implemented with approximately the same effort as the non-adaptive filter. The success of this new approach for adaptive filtering has been demonstrated using synthetic data from a two mode system.

I. INTRODUCTION

Adaptive estimation and control techniques are being studied for their future application to the real-time control of large space structures, where uncertain or changing parameters may destabilize standard control system designs. In a recent paper Schaechter [1] proposes using an extended Kalman filter to estimate adaptively the (slowly varying) frequencies and damping ratios of a large space structure. For a system with N states and M (slowly varying) parameters the extended Kalman filter requires integration of an N+M by N+M nonlinear matrix Riccati equation to determine the covariance and gain for the filter. Schaechter introduces approximations which allow the integration of the nonlinear matrix Riccati equation to be replaced by integration of a smaller set of linear matrix equations. The N states of the system are estimated using constant gains determined off-line. The time-varying gains for estimating the (slowly varying) set of M parameters are determined on-line by integrating an M by N set of linear matrix equations.

The contribution of the work presented here is to show that the time-varying gains for estimating the (slowly varying) frequencies and damping ratios can be determined in closed-form so it is not necessary to integrate the M by N set of linear matrix equations. This is an important savings of computer resources and allows the adaptive filter to be implemented with approximately the same effort as the non-adaptive filter. In particular, after certain approximations the time-varying adaptive gain can be written as the product of a constant matrix times a matrix derived from the components of the estimated state vector. The constant matrix is determined off-line just as the constant gains for estimating the state are determined off-line.

The success of this new approach for adaptive filtering has been demonstrated on a computer simulation using synthetic data from a two mode system. Work in progress is applying the new approach to a much larger system using experimental data. The theoretical development and preliminary experimental results are presented in the paper.

II. FORMULATION WITHOUT ADAPTIVE FILTERING

The standard state variable formulation of the dynamic equations of motion are shown below where the dot indicates derivative, x is the state vector, u is the control vector, z is the measurement vector, and w and v are dynamic noise and measurement noise. [2]

$$\dot{x} = Fx + Gu + \Gamma w \tag{1}$$

$$z = Hx + v$$

When the dynamic system is precisely known, a state estimator of the following form may be constructed where $\hat{x}$ indicates the estimate of the state x and K is the gain matrix.

$$\dot{\hat{x}} = F\hat{x} + Gu + K(z - H\hat{x}) \tag{2}$$

The differential equation for the estimation error $\tilde{x} = x - \hat{x}$ is obtained by subtracting Eq. (2) from Eq. (1).

$$\dot{\tilde{x}} = (F - KH)\tilde{x} + \Gamma w - Kv \tag{3}$$

The differential matrix equation for the covariance of the estimation error P follows where R and Q are from the covariance of the measurement noise v an the dynamic noise w.

$$P = E(\tilde{x}\tilde{x}^T)$$

$$\dot{P} = (F - KH)P + P(F - KH)^T + \Gamma Q \Gamma^T + KRK^T \tag{4}$$

The optimal gain matrix K is chosen to minimize the trace of the estimate error covariance to give the usual result

$$K = PH^T R^{-1} \tag{5}$$

Notice that for a precisely known dynamic system, the estimation gains may be precomputed, even in the event of a time varying system. The analysis used with the adaptive filter closely parallels the development without adaptive filtering.

III. ADAPTIVE FORMULATION AND SOLUTION

Adaptive control may be required when the model in Eq. (1) is unknown, uncertain, or dependent upon a changing system configuration. The modifications that need to be made in Eq. (1) in order to include the effects of an

uncertain parameter are given below where the vector parameter a has a dynamics matrix C with dynamic noise w_a,

$$\begin{aligned} \dot{x} &= F(a)x + Gu + \Gamma w \\ \dot{a} &= \quad Ca + w_a \\ z &= Hx + v \end{aligned} \tag{6}$$

As can be seen from Eq. (6), the system dynamics are now a function of the vector parameter a. In this formulation, the vector parameter a represents small changes from a nominal value so the average value of a is zero. These parameters are assumed to be slowly varying so that they may be adjoined to the state vector. An adaptive state estimator may be written so both the state vector and the vector of parameters are updated using the measurements.

$$\begin{aligned} \dot{\hat{x}} &= F(a)\hat{x} + Gu + K_x(z - H\hat{x}) \\ \dot{\hat{a}} &= C\hat{a} + K_a(z - H\hat{x}) \end{aligned} \tag{7}$$

Let the symmetric matrices P_x and P_a represent the covariance of the error in the estimates for x and a, respectively, and let the rectangular matrix P_{ax} represent the cross-covariance of the errors in the estimates of x and a. It is necessary to calculate these covariance matrices in order to determine the optimal gains K_x and K_a. The optimal gains are selected to minimize the trace of the covariance of the estimation error and have the following values,

$$\begin{aligned} K_x &= P_x H^T R^{-1} \\ K_a &= P_{ax} H^T R^{-1} \end{aligned} \tag{8}$$

Proceeding as before, and assuming the estimation error $\tilde{a} = a - \hat{a}$ is small, gives the vector differential equation for the error,

$$\begin{aligned} \dot{\tilde{x}} &= (F - K_x H)\tilde{x} + \left(\frac{\partial F}{\partial \hat{a}}\hat{x}\right)\tilde{a} + \Gamma w - K_x v \\ \dot{\tilde{a}} &= - K_a H \tilde{x} + C\tilde{a} + w_a - K_a v \end{aligned} \tag{9}$$

The matrix differential equations for the covariance are:

$$
\begin{aligned}
dP_x/dt &= (F-K_xH)\,P_x + P_x\,(F-K_xH)^T \\
&\quad + Q + K_xRK_x^T \\
&\quad + (F_a\hat{x})\,P_{ax} + P_{ax}^T(F_a\hat{x})^T \\
dP_{ax}/dt &= CP_{ax} + P_{ax}(F-K_xH)^T \\
&\quad + P_a(F_a\hat{x})^T - K_aHP_x \\
&\quad + K_aRK_x^T \\
dP_a/dt &= CP_a + P_aC^T + Q_a \\
&\quad - K_aHP_{ax}^T - P_{ax}(K_aH)^T + K_a\,RK_a^T
\end{aligned}
\tag{10}
$$

where $F_a = \partial F/\partial a$

and R, Q, and Q_a are covariances of v, Γw, and w_a (without delta function).

The remainder of this analysis will show approximations which can be used to reduce the computational effort needed to calculate the covariance matrices and the optimal gains when the covariance matrix P_a is very small (of order ε) and the covariance matrix P_{ax} is also very small (of order ε). The gain K_a will be very small (of order ε) because it is calculated from P_{ax}.

The differential equation for the covariance matrix P_x will involve some small terms, but most of the terms are larger and constant. If the last two terms in the differential equation for P_x are neglected (because they are small terms of order ε), it is possible to calculate the steady-state constant value of the covariance P_x. From the constant value of the covariance P_x the constant gain K_x can be determined. As one might suspect, the constant gain K_x has the same value as it would have if there were no errors in estimating the parameters a.

Because the covariance matrices P_{ax} and P_a are of order ε, many of the terms in the differential equation for P_a are of order ε squared. If the last terms in the differential equation for P_a are neglected (because they are very small terms of order ε squared), it is possible to calculate the steady-state value of the covariance P_a (to order ε). As one might suspect, the constant steady-state value obtained for P_a is the same value which would have been obtained if K_a were zero.

All that remains is to calculate the time-varying covariance P_{ax} so that the needed gain K_a can be determined. Because the gain K_x has been chosen to equal $P_xH^TR^{-1}$, the last two terms in the differential equation for P_{ax} cancel out. For the remaining analysis it will be assumed there are N state variables so the first N/2 variables (designated the N/2 length vector x*) cor-

respond to mode position, and the last N/2 variables (designated by the N/2 length vector x**) correspond to velocity of mode position. The differential equations for the dynamics of the mode variables without any forcing or disturbing terms are presented below where A* corresponds to the damping terms $(-2\xi\omega)$ and A** corresponds to the frequency terms $(-\omega^2)$. Notice both A* and A** are diagonal N/2 by N/2 matrices.

$$\begin{aligned} dx/dt &= Fx \\ dx^*/dt &= x^{**} \\ dx^{**}/dt &= A^*x^*+A^{**}x^{**} \end{aligned} \tag{11}$$

Let there be N parameters in the vector a and arrange the order of the parameters a so that the first N/2 parameters are the same as the elements of the diagonal matrix A* and the last N/2 parameters are the same as the elements of the diagonal matrix A**. Furthermore, assume the N-by-N symmetric covariance matrix P_a associated with these parameters is diagonal and composed of diagonal sub-matrices P_a^* and P_a^{**}. With these assumptions, the partial derivative can be written in a particularly simple way where x^* and x^{**} represent diagonal matrices with the diagonal elements equal to the vectors x^* and x^{**}

$$\begin{aligned} \hat{Fx} &= \begin{bmatrix} 0 & I \\ A^* & A^{**} \end{bmatrix} \begin{bmatrix} x^* \\ x^{**} \end{bmatrix} \\ F_a x &= \partial[\hat{Fx}]/\partial_a \\ &= \begin{bmatrix} 0 & 0 \\ [x^*] & [x^{**}] \end{bmatrix} \\ P_a &= \begin{bmatrix} P_a^* & 0 \\ 0 & P_a^{**} \end{bmatrix} \\ F_a x\, P_a &= \begin{bmatrix} 0 & 0 \\ P_a^* & P_a^{**} \end{bmatrix} \begin{bmatrix} [x^*] & 0 \\ 0 & [x^{**}] \end{bmatrix} = P_a^{***}\cdot[\hat{x}] \end{aligned} \tag{12}$$

One further assumption is that the dynamics matric C (for the parameters a) is diagonal and equal to the scalar c_0 times the identity matrix I. With those assumptions, the differential equation for the cross covariance P_{ax} can be written as follows where $\bar{x}$ is a diagonal matrix made up of the elements of x.

$$dP_{ax}/dt = P_{ax}(F-K_xH+C)^T + (P_a***[\bar{x}])^T$$

where $$C = c_0 I$$

and $$P_a*** = \begin{bmatrix} 0 & 0 \\ P_a* & P_a** \end{bmatrix} \quad (13)$$

The remainder of the analysis will deal with the cross-covariance matrix P_{xa} which is the transpose of the covariance matrix P_{ax}. The differential equation for the cross-covariance P_{xa} can be written as follows:

$$dP_{xa}/dt = F* P_{xa} + P_a*** \left[\hat{x}\right]$$

where $$F* = F - K_x H + C \quad (14)$$

The linear matrix differential equation for P_{xa} has particularly desirable characteristics. All the terms in the differential equation are known constants (because the gain K_x and the covariance P_a are known and constant) except for driving terms due to estimates of the state $\hat{x}$. If the approximation is made that the derivative of the forcing terms $\hat{x}$ is equal to the dynamics matrix F times $\hat{x}$, then, except for transient terms, the solution to the linear matrix differential equation for P_{xa} can be written in closed form as a linear combination of the forcing terms $\hat{x}$. This is similar to the result in elementary linear differential equations where the general solution is composed of the sum of the homogeneous solution due to the unforced differential equation and the particular solution due to the forcing function.

Because the forcing function $[\hat{x}]$ is a diagonal matrix, the first element $\hat{x}_1$ is the forcing term for the first column of the solution for the matrix P_{ax}, the second element $\hat{x}_2$ is the forcing term for the second column of the matrix P_{ax}, and so on. Let P_i be a vector which represents the i-th column of the matrix P_{ax}. The linear matrix-vector differential equation for the i-th column can be written as follows where P_{ai} is a scalar which is the i-th element of the diagonal matrix P_a and $\hat{x}_i$ is a scalar which is the i-th element of $\hat{x}$ and P_i^* is is the i-th column of the matrix P_a*** which is all zeroes except for entries equal to the diagonal elements of P_a

$$dP_i/dt = F^* P_i + P_i^* \hat{x}_i \tag{15}$$

The solution for the vector P_i is assumed to be composed of the sum of two vectors. The first vector is the constant vector E_i times the scalar $\hat{x}_j$ (corresponding to the estimate of the position of the mode) and the second vector is the constant vector G_i times the scalar $\hat{x}_k$ (corresponding to the estimate of the velocity of the appropriate mode).

$$P_i = E_i \hat{x}_j + G_i \hat{x}_k \tag{16}$$

where for $i \leq N/2$ then $j = i$ and $k = i + N/2$

for $i > N/2$ then $j = i - N/2$ and $K = i$

The derivative of the vector P_i can be calculated directly if it assumed the derivative of the vector $\hat{x}$ is equal to $F\hat{x}$ with A_i^* and A_j^{**} being scalars which represent the j-th element of the representative diagonal matrices which make up F.

$$\begin{aligned} dP_i/dt &= E_i d\hat{x}_j/dt + G_i d\hat{x}_k/dt \\ &= E_i \hat{x}_k + G_i A_j^* \hat{x}_j + G_i A_j^{**} \hat{x}_k \end{aligned} \tag{17}$$

Substituting the expression for the assumed form of the vector P_i and the expression for the derivative of the vector P_i into the differential equation, gives the following equations where δ_{ij} is a discrete delta function which is unity if i equals j and zero otherwise.

$$\begin{aligned} & G_i A_j^* \hat{x}_j + (E_i + G_i A_j^{**}) \hat{x}_k \\ &= F^* E_i \hat{x}_j + F^* G_i \hat{x}_k + \delta_{ij} P_j^* \hat{x}_j + \delta_{ik} P_k^* \hat{x}_k \end{aligned} \tag{18}$$

Collecting all terms which multiply the scalar $\hat{x}_j$ gives one vector equation and collecting all terms which multiply the scalar $\hat{x}_k$ gives a second vector equation. There are two vector equations and two unknown vectors E_i and G_i.

$$\begin{aligned} G_i A_j^* &= F^* E_i + \delta_{ij} P_j^* \\ E_i + G_i A_j^{**} &= F^* G_i + \delta_{ik} P_k^* \end{aligned} \tag{19}$$

The expression for E_i obtained from the second equation is substituted into the first equations to give a single equation with the unknown vector G_i

$$G_i A_j^* = F^*(F^* G_i - G_i A_j^{**} + \delta_{ik} P_k^*) + \delta_{ij} P_j^* \tag{20}$$

Since A_j^* and A_j^{**} are both scalars, it is possible to solve directly for the unknown vector G_i where I is the identity.

$$G_i = (IA_j^* + F^* A_j^{**} - F^* F^*)^{-1} (\delta_{ij} P_j^* + \delta_{ik} F^* P_k^*) \tag{21}$$

In the same way, the expression for G_i obtained from the first equation is substituted into the second equation to give a single equation with the unknown vector E_i.

$$A_j^* E_i = (F^* - IA_j^{**})\ (F^* E_i + \delta_{ij}\ P_j^*) + A_j^* \delta_{ik} P_k^* \tag{22}$$

It is again possible to solve directly for E_i.

$$E_i = (IA_j^* + F^* A_j^{**} - F^* F^*)^{-1} \left[\delta_{ij} (F^* - IA_j^{**}) P_j^* + \delta_{ik} A_j^* P_k^* \right] \tag{23}$$

Thus the two unknown vector quantities E_i and G_i can be determined from known quantities so the covariance vector P_i and the approximation for the covariance matrix P_{ax} can be determined.

IV. SIMULATION RESULTS WITH TWO MODES

The new, simplified adaptive formulation was first tested with a single mode system. After encouraging results were obtained with one mode, a two-mode system was examined. The two-mode system used in the simulation studies is shown in Figure 1.

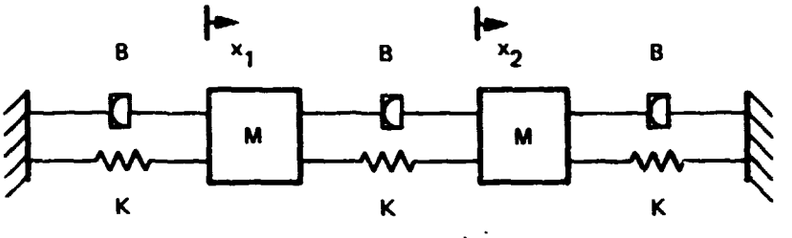

FIG. 1 TWO-MODE SYSTEM

The system consists of two masses, M, three springs, K, and three viscous dampers, B. For this study, M=1, K=1, and B=0.10. Control forces may be applied to both masses, random external forces disturb both masses, and noisy measurements of the position of both masses are available. The measurements are used for estimating the state vector, and for estimating the parameter vector. The differential equations representing this system are:

$$\begin{aligned}
M\ddot{x}_1 + 2B\dot{x}_1 - B\dot{x}_2 + 2Kx_1 - Kx_2 &= f_1 + w_1 \\
M\ddot{x}_2 + 2B\dot{x}_2 - B\dot{x}_1 + 2Kx_2 - Kx_1 &= f_2 + w_2 \\
z_1 &= x_1 + v_1 \\
z_2 &= x_2 + v_2
\end{aligned} \tag{24}$$

The natural frequencies and damping ratios of this system are:

$$\omega_1 = 1 \qquad \zeta_1 = 0.05$$

$$\omega_2 = 1.732 \qquad \zeta_2 = 0.0869$$

where the low frequency mode is the common mode motion of the two masses. The spectral densities of both the process and measurement disturbances (Q and R) are 0.0163. Two hundred position measurements of both the masses were made during a sixty second computer simulation. This sixty second duration was

selected to assure about ten oscillations of the lowest frequency mode. The sample rate was selected to give about ten samples per cycle of the highest frequency mode. The value of the correlation time constant for the parameters that are to be estimated was 250 seconds (so C_o is 1/250). This value is much larger than the time constants of the system. The selection of a "large" value is important in order to allow the adaptive filter to average values over several cycles of the system. The following table gives a summary of the test cases. In each case, both the standard, non-linear extended Kalman filter, and the simplified extended Kalman filter described in this paper were run in order to make comparisons. In all of the cases studied thus far, these two cases were indistinguishable, except for a small, initial transient. This transient effect is attributed to beginning the standard extended Kalman filter covariance integration with values slightly different from the steady state values.

CASE 1	ω_1 unknown ω_2 unknown	ζ_1 known ζ_2 known	no noise
CASE 2	ω_1 unknown ω_2 unknown	ζ_1 unknown ζ_2 unknown	no noise
CASE 3	ω_1 unknown ω_2 unknown	ζ_1 unknown ζ_2 unknown	noise

The results are shown in the following figures and are discussed below. In Case I the starting estimates for the natural frequencies were chosen to be 10% in error with ω_1 estimated to be 0.9 (rather than 1.0, and ω_2 estimated to be 1.559 rather than 1.732). The damping parameters were exact, and no noise was present in the system. The results for the estimate of ω_1 (Fig. 2) show that the modal frequency is very readily identified from the measurements, inspite of the 10% initial error in the estimate. As the system response diminishes, less information is available for updating the parameters. Consequently, with no new information coming into the system, the parameter estimate begins to return to its nominal value (0.9) with the selected time constant of 250 sec. The estimate of ω_2 behaves similarly.

In Case II, the objective was the same as in Case I with the additional problem of simultaneously estimating the damping parameters. The initial estimates of the damping parameters were zero. The results of the poor initial guess of the damping parameter are evident in Fig. 3. The estimate of the modal frequency tends to be lightly damped, but in all other aspects, the estimate of ω_1 appears to have the same features that were present in Case I. As has been found in past studies [1], the estimate of the damping

parameter itself is quite poor. This is due to the fact that the position measurement contains very little damping information.

Case III is identical to Case II with the addition of both process and measurement disturbances. Surprisingly, this case yielded the best results, as is evident in Figures 4 and 5. The effects of the noise are clearly visible in the figures. However, in contrast with the previous two cases, the process noise continues to excite the system after the transient effect of the initial conditions have subsided. The result is that the measurements continue to provide information on the parameters for the duration of the simulation. Since the higher frequency mode is more heavily damped, and is less perturbed by the external disturbance, the improvement in the natural frequency estimate of mode two is not as dramatic.

CONCLUSIONS

This paper has developed approximations which allow dramatic reductions in the on-line computational requirements of the extended Kalman filter. Numerical simulations of this technique have validated the approach for two simple spring-mass systems. It was found that the full non-linear extended Kalman filter and the closed-form adaptive filter developed in this paper gave virtually identical results. Work is currently in progress to apply this approach to a much larger system using experimental data, rather than simulated data.

REFERENCES

(1) SCHAECHTER, D. B., "Adaptive Control for Large Space Structures", AIAA Guidance and Control Conference, Gatlinburg, Tennessee, paper no. 83-2246, August 15-17, 1983.

(2) BRYSON, A. E. and HO, Y. C., Applied Optimal Control, Blaisdell, Waltham Massachusetts, 1969.

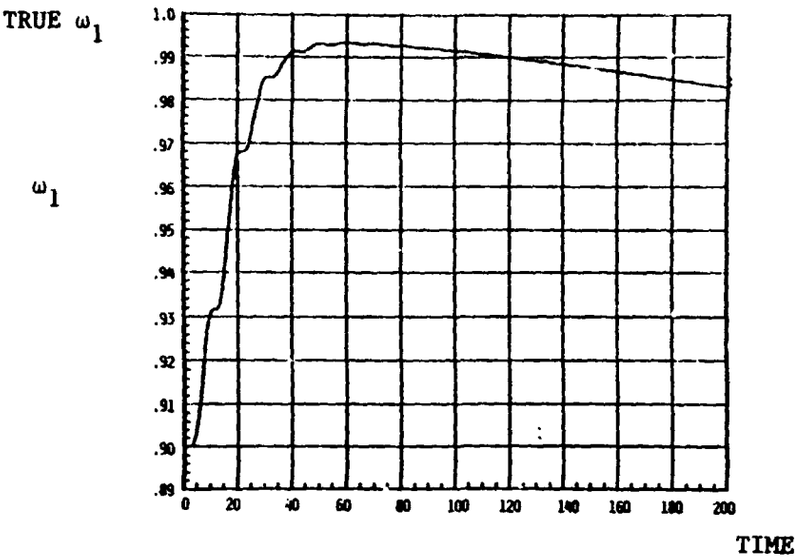

FIGURE 2 Estimate of ω_1 (CASE I)

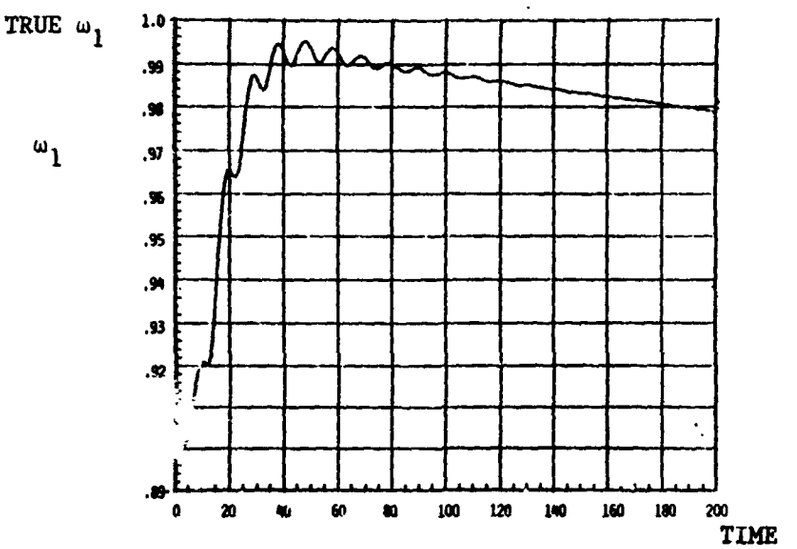

FIGURE 3 Estimate of ω_1 (CASE II)

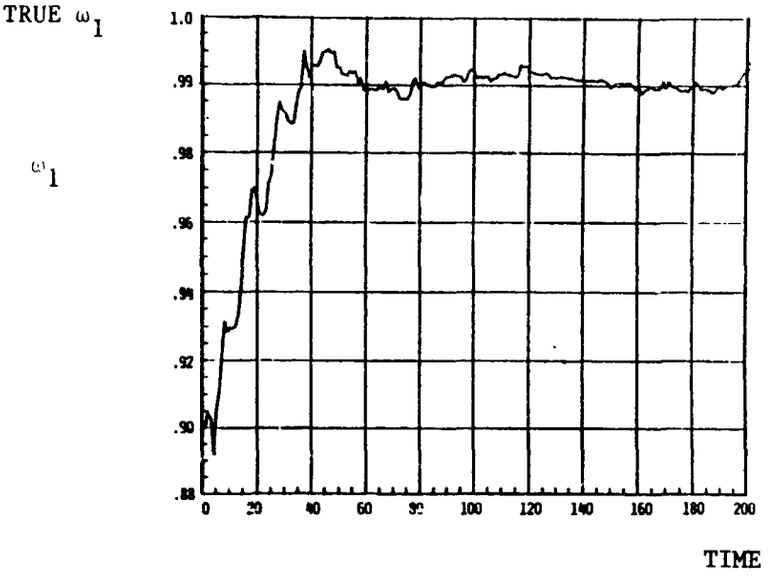

FIGURE 4 Estimate of ω_1 (CASE III)

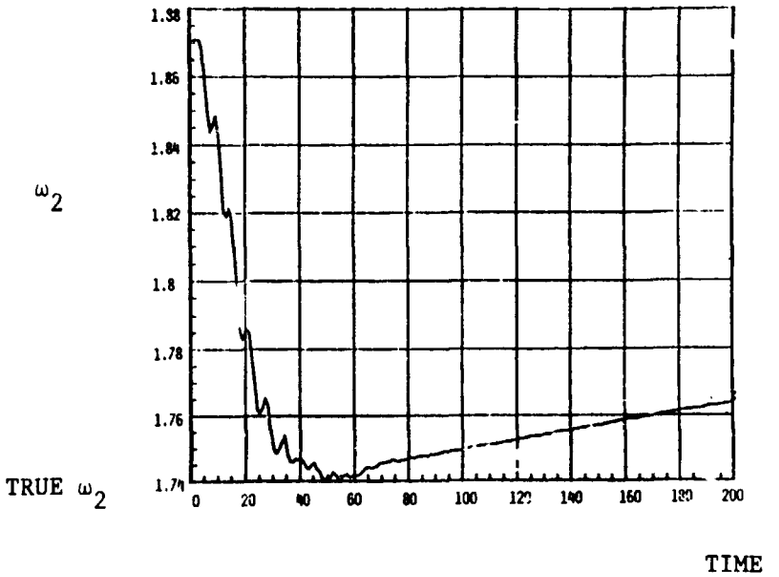

FIGURE 5 Estimate of ω_2 (CASE III)