Newton's method in discrete-time nonlinear data smoothing

J. S. Meditch
Boeing Scientific Research Laboratories, Seattle, Washington, USA

It is well known that the problem of estimating the state and parameters of discrete-time nonlinear dynamic systems from noisy measurements can be formulated as a nonlinear two-point boundary-value problem. In this paper, an algorithm is developed for solving the above boundary-value problem. The algorithm, which is based on Newton's method, requires very little storage. A detailed discussion of the algorithm is given and certain extensions of it are indicated.

(Received October 1969)

1. Introduction

The purpose of this paper is to present a successive approximation procedure for fixed-interval data smoothing for a particular class of discrete-time nonlinear systems. The corresponding procedure for continuous-time nonlinear systems was developed by Meditch (1969). However, the two procedures differ in certain essential details and it is, therefore, desirable to treat them separately.

The underlying idea here is that of applying Newton's method to the nonlinear two-point boundary-value problem which arises in nonlinear data smoothing when the maximum likelihood estimation viewpoint is adopted. The problem formulation and the development of the resulting two-point boundary-value problem, both of which are due to Cox (1964), are given below in Section 2. The computational procedure is then developed and discussed in Sections 3 and 4, respectively.

2. Problem formulation

Consider the class of systems

\[ x(k) = f(x(k-1), k-1) + u(k-1) \]
\[ z(k) = h(x(k), k) + v(k) \]

for \( k = 1, \ldots, N \) where \( k \) is the discrete-time index; \( x \) is an \( n \)-vector, the state; \( u \) is an \( n \)-vector, the disturbance; \( z \) is an \( m \)-vector, the measurement; and \( v \) is an \( m \)-vector, the measurement error. Further, \( f \) and \( h \) are, respectively, \( n \) and \( m \)-dimensional vector-valued functions of the indicated variables; and \( N \) is a positive integer.

It is assumed that \( \{u(i), i = 0, 1, \ldots, N-1\} \) and \( \{v(j), j = 1, \ldots, N\} \) are independent, zero mean, gaussian sequences with

\[ E[u(i)u^T(k)] = Q(k) \delta_{jk} \]
\[ E[v(j)v^T(k)] = R(k) \delta_{jk} \]
\[ E[u(j)u^T(k)] = 0 \]

for all relevant values of the integers \( j \) and \( k \). In this set of expressions, \( E \) denotes the expectation operator, \( \delta_{jk} \) is the Kronecker delta, prime denotes the transpose, and \( Q(k) \) and \( R(k) \) are, respectively, \( n \times n \) and \( m \times m \) positive definite covariance matrices.

The initial state \( x(0) \) of the system in (1) is taken to be a gaussian random vector which is independent of \( \{u(i), i = 0, 1, \ldots, N-1\} \) and \( \{v(j), j = 1, \ldots, N\} \) and has mean \( \bar{x}(0) \) and positive definite covariance matrix \( P(0) \).

The estimation problem of interest here is the following: Having observed the sequence of measurements \( \{z(1), \ldots, z(N)\} \), obtain an estimate of the sequence of states \( \{x(0), x(1), \ldots, x(N)\} \) which is based on the given measurement data.

The viewpoint adopted here is that of maximum likelihood estimation. Specifically, the estimate is chosen to be that sequence \( \{\hat{x}(0), \ldots, \hat{x}(N)\} \) for which the \textit{a posteriori} conditional probability density function

\[ p(x(0), \ldots, x(N)|z(1), \ldots, z(N)) \]

is maximized, i.e. the estimate is the mode of this density function.

By a direct application of Bayes' rule, it can be shown that

\[ p(x(0), \ldots, x(N)|z(1), \ldots, z(N)) = \mu(Z(N), \zeta(N)) \exp\left(-\frac{1}{2}||x(0) - \bar{x}(0)||^2_{P(0)}\right) \]
\[ - \frac{1}{2} \sum_{i=1}^{N} \left(||z(i) - h(x(i), i)||^2_{R(i)} + ||x(i) - f(x(i-1), i-1)||^2_{Q(i-1)}\right) \]

where \( A^{-1} \) denotes the inverse of the indicated matrix and \( ||y||^2_{B} \) is the quadratic form \( y'By \).

In (3), \( \mu(Z(N), \zeta(N)) \) is a scalar-valued function of the sequence of measurements, \( Z(N) \triangleq \{z(1), \ldots, z(N)\} \), and the set of known statistical parameters \( \zeta(N) \triangleq \{P(0); Q(0), \ldots, Q(N-1); R(1), \ldots, R(N)\} \).
Utilizing (1), it is clear that maximizing \( p(x(0), \ldots, x(N)|z(1), \ldots, z(N)) \) with respect to \( \{x(0), \ldots, x(N)\} \) is equivalent to minimizing

\[
V_N = \frac{1}{2} ||x(0) - \bar{x}(0)||^2_{p(x(0))} + \frac{1}{2} \sum_{i=1}^{N} ||z(i) - h(x(i), i)||^2_{_{\Sigma(i)}} + ||u(i - 1)||^2_{Q(i-1)} + \psi(i - 1)[(x(i) - f(x(i - 1), i - 1) - u(i - 1)) (4)
\]

with respect to both \( \{x(0), \ldots, x(N)\} \) and \( \{u(0), \ldots, u(N - 1)\} \), subject to the set of constraints

\[
x(i) - f(x(i - 1), i - 1) - u(i - 1) = 0 (5)
\]

where \( i = 1, \ldots, N \).

With the introduction of the set of Lagrange multipliers \( \{\psi(0), \ldots, \psi(N)\} \) where each \( \psi(i) \) is an \( n \)-vector, the problem becomes one of minimizing the function

\[
V_N^* = \frac{1}{2} ||x(0) - \bar{x}(0)||^2_{p(x(0))} + \frac{1}{2} \sum_{i=1}^{N} ||z(i) - h(x(i), i)||^2_{_{\Sigma(i)}} + ||u(i - 1)||^2_{Q(i-1)} + \psi(i - 1)[(x(i) - f(x(i - 1), i - 1) - u(i - 1)) (4)
\]

with respect to \( \{x(0), \ldots, x(N)\} \) and \( \{u(0), \ldots, u(N - 1)\} \).

Setting the gradient of \( V_N^* \) with respect to \( u(k - 1) \) equal to zero for each \( k = 1, \ldots, N \) leads to the condition

\[
u(k - 1) = Q(k - 1)\psi(k - 1). (6)
\]

Hence, the set of relations in (5) can now be written

\[
x(k) = f(x(k - 1), k - 1) + Q(k - 1)\psi(k - 1) (6)
\]

where \( k = 1, \ldots, N \), and the notation \( \bar{x}(k) \) is used to denote the minimizing value of \( x(k) \).

Similarly, setting the gradient of \( V_N^* \) with respect to \( x(0) \) and also with respect to \( x(k), k = 1, \ldots, N \) yields the two relations

\[
P^{-1}(0)[\bar{x}(0) - x(0)] - f_x(x(0), 0)\psi(0) = 0 \tag{7}
\]

and

\[
\psi(k - 1) - h_x(x(k), k)\psi(k)
\]

\[
- h_x(x(k), k)R^{-1}(k)[z(k) - h(x(k), k)] = 0,
\]

respectively.

In these relations, \( f_x(x(k), k) \) and \( h_x(x(k), k) \) are, respectively, the \( n \times n \) and \( n \times m \) Jacobian matrices

\[
f_x(x(k), k) = \left[ \frac{\partial f(x(k), x)}{\partial x_j} \right]_{x = \bar{x}(k)} i = 1, \ldots, n
\]

and

\[
h_x(x(k), k) = \left[ \frac{\partial h(x(k), x)}{\partial x_j} \right]_{x = \bar{x}(k)} j = 1, \ldots, m.
\]

If \( f_x(x(j), j) \) is nonsingular for all \( j = 0, 1, \ldots, N \), the above two equations can be put in the form

\[
\psi(0) = [P(0)f_x(x(0), 0)]^{-1}[\bar{x}(0) - x(0)] \tag{7}
\]

and

\[
\psi(k) - h_x(x(k), k)R^{-1}(k)[z(k) - h(x(k), k)] = 0
\]

for each iteration \( k = 1, \ldots, N \).

Three remarks concerning the above formulation are in order before continuing to the computational procedure. First, the notation \( \bar{x}(k) \) is used here to denote the fixed-interval smoothed estimate of \( x(k) \) rather than \( \hat{x}(k|N) \). The latter is commonly employed in the literature to emphasize that each estimate is based on all of the measurement data which is available over the fixed interval \([0, N]\) and to distinguish it from other types of estimates, e.g. \( \hat{x}(k|k) \), the filtered estimate, and \( \hat{x}(k + 1|k) \), the predicted estimate. The former notation is chosen for present purposes because of its simplicity, the understanding being that only fixed-interval smoothed estimates are of concern here.

Second, the requirement that \( f_x(x(k), k) \) be nonsingular is always satisfied for an important class of problems which are of practical interest. In particular, if (1) is the sampled-data model of a continuous-time nonlinear system, \( f_x(x(k), k) \) is the state transition matrix for the system's variational equations with the linearization carried out about the current estimate \( \bar{x}(k) \). It is well known that this transition matrix is nonsingular.

Third, if the system in equations (1) and (2) is linear, the above two-point boundary-value problem, which is also linear, can be solved in closed-form and recursive relations obtain for both filtering and fixed-interval smoothing (see Cox, 1964).

3. Computational procedure

The nonlinear two-point boundary-value problem of Section 2 can be expressed in the form

\[
\bar{x}(k) = \alpha(x(k - 1), \psi(k - 1), k - 1) \tag{10}
\]

\[
\psi(k) = \beta(x(k), \psi(k - 1), k) \tag{11}
\]

\[
\psi(0) = \gamma(x(0), \Pi) \tag{12}
\]

and

\[
\psi(N) = 0 \tag{13}
\]

where \( k = 1, \ldots, N \), and \( \alpha, \beta, \gamma \) are the \( n \)-dimensional vector-valued functions which correspond to the right-hand sides of equations (6), (8) and (7), respectively. Further, \( \Pi \) in (12) denotes the given parameters \( P(0) \) and \( x(0) \) in (7), and the explicit time dependence in equations (10) and (11) includes the parameters \( Q(k - 1) \) and \( R^{-1}(k) \) in equations (6) and (8), respectively, as well as the known measurements \( z(k) \) in the latter relation. The problem which is defined by equations (10)–(13) can be viewed as one in which \( \bar{x}(0) \) is to be determined such that when it and its corresponding value of \( \psi(0) \), as given by (12), are used as initial conditions in equations (10) and (11), the resulting solution satisfies (13), \( \psi(N) = 0 \). This can be done in the following way.

For each iteration \( i, i = 1, 2, \ldots \), let \( \bar{x}(k) \) and \( \psi(k) \) be the solutions of

\[
\frac{d}{dk}\bar{x}(k) = \alpha(\bar{x}(k - 1), \psi(k - 1), k - 1) \tag{14}
\]

\[
\psi(k) = \beta(\bar{x}(k), \psi(k - 1), k) \tag{15}
\]
where \( k = 1, \ldots, N \) for a given \( \xi(0) \) with \( \psi(0) = \gamma(\xi(0), \Pi) \). 

Now let 
\[
\xi^{i+1}(k) = \xi(k) + \delta \xi(k)
\]
and 
\[
\psi^{i+1}(k) = \psi(k) + \delta \psi(k).
\]

Then, the system of variational equations for equations (14)–(16) are 
\[
\delta \dot{\xi}(k) = \alpha_{s}(\xi(k-1), \psi(k-1), k - 1) \delta \xi(k - 1) + \alpha_{p}(\xi(k-1), \psi(k-1), k - 1) \delta \psi(k - 1) \\
\delta \dot{\psi}(k) = \beta_{s}(\xi(k), \psi(k-1), k) \delta \xi(k) + \beta_{p}(\xi(k), \psi(k-1), k) \delta \psi(k - 1)
\]
where \( k = 1, \ldots, N \); and \( \alpha_{s}, \alpha_{p}, \beta_{s}, \beta_{p}, \gamma_{s} \) are \( n \times n \) Jacobian matrices which are evaluated at the indicated arguments, e.g., 
\[
\alpha_{s}(\xi(k-1), \psi(k-1), k - 1) = \left. \frac{\partial \alpha_{s}}{\partial \xi_{j}} \right|_{\xi(k-1), \psi(k-1), k - 1}
\]
where \( j, l = 1, \ldots, n \). Note that the values of the variables which are needed to evaluate these Jacobian matrices are obtained from the solution of equations (14)–(16) for a given \( \xi(0) \).

It is quite obvious that, in general, the solution of equations (14) and (15) for an arbitrary \( \xi(0) \) and the corresponding \( \psi(0) \) from (16) will not satisfy the boundary condition \( \psi(N) = 0 \). However, the variational equations can be utilized to compute corrections \( \delta \dot{\xi}(0) \) and \( \delta \dot{\psi}(0) \) to \( \dot{\xi}(0) \) and \( \dot{\psi}(0) \), respectively, so that, to within first-order terms, the solution of equations (14) and (15) subject to the initial conditions 
\[
\dot{\xi}(0) = \xi(0) + \delta \dot{\xi}(0) \\
\dot{\psi}(0) = \psi(0) + \delta \dot{\psi}(0)
\]
will satisfy \( \dot{\psi}(N) = 0 \).

For notational convenience, let equations (18) and (19) be written 
\[
\delta \dot{\xi}(k) = \Phi_{11}(k) \delta \xi(k - 1) + \Phi_{12}(k) \delta \psi(k - 1) \\
\delta \dot{\psi}(k) = \Phi_{21}(k) \delta \xi(k) + \Phi_{22}(k) \delta \psi(k - 1)
\]
where the definitions of \( \Phi_{ij}(k) \), \( i, j = 1, 2 \), follow from a direct comparison of equations (21) and (22) with equations (18) and (19), respectively. Substituting into (22) from (21), it follows that 
\[
\delta \psi(k) = \Phi_{21}(k) \delta \xi(k) + \Phi_{22}(k) \delta \psi(k - 1) = \Phi_{21}(k) \delta \xi(k) + \Phi_{22}(k) \delta \psi(k - 1) \\
\delta \psi^i(k) = \Phi_{21}(k) \delta \xi(k) + \Phi_{22}(k) \delta \psi(k - 1) = \Phi_{21}(k) \delta \xi(k) + \Phi_{22}(k) \delta \psi(k - 1)
\]
where, by definition, 
\[
\Phi_{21}(k) = \Phi_{21}(k) \Phi_{11}(k)
\]
and 
\[
\Phi_{22}(k) = \Phi_{21}(k) \Phi_{12}(k) + \Phi_{22}(k).
\]

Equations (21) and (23) may now be combined and written 
\[
\begin{bmatrix} \delta \dot{\xi}(k) \\ \delta \dot{\psi}(k) \end{bmatrix} = \begin{bmatrix} \Phi_{11}(k) & \Phi_{12}(k) \\ \Phi_{21}(k) & \Phi_{22}(k) \end{bmatrix} \begin{bmatrix} \delta \xi(k - 1) \\ \delta \psi(k - 1) \end{bmatrix}
\]
where \( k = 1, \ldots, N \).

Recursive application of (24) leads to the result 
\[
\begin{bmatrix} \delta \dot{\xi}(N) \\ \delta \dot{\psi}(N) \end{bmatrix} = \begin{bmatrix} \Theta_{11}(N) & \Theta_{12}(N) \\ \Theta_{21}(N) & \Theta_{22}(N) \end{bmatrix} \begin{bmatrix} \delta \dot{\xi}(0) \\ \delta \dot{\psi}(0) \end{bmatrix}
\]
where 
\[
\begin{bmatrix} \Theta_{11}(N) & \Theta_{12}(N) \\ \Theta_{21}(N) & \Theta_{22}(N) \end{bmatrix} = \begin{bmatrix} \Phi_{11}(N) & \Phi_{12}(N) \\ \Phi_{21}(N) & \Phi_{22}(N) \end{bmatrix} \cdots \begin{bmatrix} \Phi_{11}(1) & \Phi_{12}(1) \\ \Phi_{21}(1) & \Phi_{22}(1) \end{bmatrix}
\]

For any integer \( l = 2, \ldots, N \), it is obvious that 
\[
\begin{bmatrix} \Theta_{11}(l) & \Theta_{12}(l) \\ \Theta_{21}(l) & \Theta_{22}(l) \end{bmatrix} = \begin{bmatrix} \Phi_{11}(l) & \Phi_{12}(l) \\ \Phi_{21}(l) & \Phi_{22}(l) \end{bmatrix} \cdots \begin{bmatrix} \Phi_{11}(1) & \Phi_{12}(1) \\ \Phi_{21}(1) & \Phi_{22}(1) \end{bmatrix}
\]
Equation (26) provides an efficient means for determining the required \( 2n \times 2n \) matrix in (25) as computations proceed along any iteration. Clearly, values of the submatrices \( \Phi_{11}(k), \Phi_{12}(k), \Phi_{21}(k) \) and \( \Phi_{22}(k) \) for a given \( k \) need not be stored once they have been incorporated into the computation in (26).

For \( k = N \), the second relation in (17) is 
\[
\psi^{i+1}(N) = \psi(N) + \delta \psi(N).
\]
Since it is desired that \( \psi^{i+1}(N) = 0 \), it follows that 
\[
\delta \psi(N) = - \psi(N)
\]
where \( \psi(N) \) is a known vector, viz. the solution of (14) and (15) for an arbitrary \( \dot{\xi}(0) \) and its corresponding \( \dot{\psi}(0) \) from (16). From the above relation and (25), it is clear that 
\[
\delta \dot{\xi}(0) = \Phi_{21}(N) \delta \xi(0) + \Phi_{22}(N) \delta \psi(0).
\]

With the aid of (20), it then follows that the required correction to \( \dot{\xi}(0) \) to obtain \( \dot{\xi}^{i+1}(0) \) is 
\[
\delta \dot{\xi}(0) = - \left[ \begin{bmatrix} \Theta_{21}(N) \\ \Theta_{22}(N) \end{bmatrix} \gamma_{s}(\dot{\xi}(0), \Pi) \right]^{-1} \psi(N)
\]
under the assumption that the indicated inverse exists.

The computational procedure can now be summarized as follows:

1. Substitute \( \dot{\xi}(0) \) as obtained from the preceding iteration, or as specified by the ‘initial guess’ for \( i = 1 \), into (16) to obtain \( \psi(0) \).
2. Solve (14) and (15) for \( \dot{\xi}(k) \) and \( \psi(k) \), \( k = 1, \ldots, N \), subject to the initial conditions given above.
3. Concurrently with Step 2, evaluate the matrices \( \Phi_{11}(l), \Phi_{12}(l), \Phi_{21}(l) \) and \( \Phi_{22}(l) \) and use (26) to determine 
\[
\begin{bmatrix} \Theta_{11}(l) & \Theta_{12}(l) \\ \Theta_{21}(l) & \Theta_{22}(l) \end{bmatrix}
\]
for \( l = 1, \ldots, N \).

4. At the completion of Steps 2 and 3, retain only \( \hat{x}(0), \Theta_2(N), \Theta_3(N) \) and \( \psi(N) \), and substitute these quantities into (27) to obtain \( \delta \hat{x}(0) \).

5. Set

\[
\hat{x}^{i+1}(0) = \hat{x}(0) + \delta \hat{x}(0)
\]

and repeat the process.

Alternatively, it should be noted that \( \psi^{i+1}(0) \) for each succeeding iteration can be determined from the relation

\[
\psi^{i+1}(0) = \psi(0) + \delta \psi(0)
\]

where \( \delta \psi(0) \) is obtained by substituting \( \delta \hat{x}(0) \) from Step 4 into (20).

4. Discussion

The procedure developed above is applicable to post-experimental data analysis problems where it is desired to determine the time history of a system’s state and the parameters of the system model from normal operating records. Fields of application include orbit determination, guidance system post-flight error analysis, chemical plant parameter estimation, and seismology.

An important feature of the present method is that it deals directly with the data smoothing problem without first requiring solution of the filtering problem as is the present practice (see Sage, 1968; Bryson and Ho, 1969; Gura and Henrikson, 1969). In the latter case, the two-point boundary-value problem in (6)–(9) is converted to an initial-value problem via invariant imbedding (Sage, 1968) or the sweep method (Clelland and Fomin, 1963). Both approaches lead to a sequential computational algorithm for determining the filtered estimate \( \hat{x}(k) \) for \( k = 0, 1, \ldots, N \). This algorithm is first-order in system dynamics and second-order in measurement nonlinearity. In addition, an extra \( n(n+1)/2 \) elements must be computed at each time point to obtain \( \hat{x}(k) \).

At the conclusion of this procedure, \( \hat{x}(N) \) and \( \psi(N) = 0 \) are used as final conditions for (6) and (8), respectively, to obtain \( \hat{x}(k|N) \) and \( \psi(k) \) sequentially in reverse-time, viz., \( k = N, N-1, \ldots, 0 \). It should be noted that this requires that \( \Gamma \) in (6) be an invertible function, a situation which is seldom true in practice. This point was apparently overlooked by Gura and Henrikson (1969).

Both the invariant imbedding and sweep procedures lead to series expansions in which only first and second-order coefficients are equated to obtain computational algorithms. The approximations are thus introduced at the outset and solution of the two-point boundary-value problem is approached indirectly whereas the procedure developed in this paper deals directly with the problem as already noted. The advantages of the latter formulation over the former are obvious.

The method in this paper is relatively simple to apply because of its low storage requirements. In particular, only \( \hat{x}^{i+1}(0) \) and \( \psi^{i+1}(0) \) which are computed at the end of one iteration need be stored to initiate the next iteration. By contrast, quasilinearization techniques (Bellman and Kalaba, 1965) generally require storage of entire time histories to proceed from one iteration to the next.

Since Newton’s method underlies the procedure here, convergence is theoretically quadratic. As with any Newton-type method, one has here also the question or problem of a ‘good’ initial estimate. If such an estimate is not available from a consideration of the physics of the particular system being investigated, or if divergence occurs, then it may be necessary to utilize a gradient method to obtain a ‘good’ initial estimate. There is at present no computational experience to report in this connection.

Any one of a number of criteria can be used to terminate the iterations. Since the procedure is designed to force satisfaction of the boundary condition \( \psi(N) = 0 \), one reasonable criterion for this purpose is \( ||\psi(N)|| < \epsilon \) where \( \epsilon \) is a small positive number.

For cases where the \( a \) priori information is ‘poor’ or unavailable, one can consider \( P(0) \) arbitrarily large such that \( P^{-1}(0) \rightarrow 0 \). Under this condition, the first term in \( V^* \) vanishes, and from (7),

\[
\psi(0) = 0
\]

for all \( i = 1, \ldots \).

The choice of \( \hat{x}(0) \) to initialize the procedure is quite important and in general dictates whether or not the procedure converges. However, in the absence of a priori information other than \( \hat{x}(0) \), it seems reasonable to choose \( \hat{x}(0) = \hat{x}(0) \). In this connection, it is remarked that \( \hat{x}(0) \) and \( P(0) \) reflect all of the a priori information including \( z(0) \) if there is a measurement at \( k = 0 \).

It is noted from equations (6) and (8) that \( Q(k-1) \) and \( R(k) \) are needed in the computations of \( \hat{x}(k) \) and \( \psi(k) \). This would seem to indicate that, for this purpose, \( Q(k-1) \) need not be positive definite as assumed initially, but that \( R(k) \) must be. However, \( Q(k-1) \) must be positive definite for all \( k = 1, \ldots, N \) in order for \( V^* \) to be defined. One possible way to circumvent this difficulty is to replace \( Q^{-1}(i-1) \) in \( V^* \) by \( S(i-1) \) which may be positive semidefinite. Setting the gradient of \( V^* \) with respect to \( u(k-1) \) equal to zero gives

\[
S(k-1)u(k-1) = \psi(k-1)
\]

for \( k = 1, \ldots, N \). Then, one can use

\[
u(k-1) = S^T(k-1)\psi(k-1)
\]

where \( S^T \) is the pseudoinverse of \( S \).

Similarly, one can use \( R^T(k) \) in (8) if \( R(k) \) is positive semidefinite. Cases where \( R(k) \) may be singular occur when one of the measurement variables is perfect, i.e. there is no measurement error in that component of the measurement vector, or when one of the measurement errors is a linear combination of some or all of the others.

The above remarks concerning \( Q(i-1) \) also pertain to \( P(0) \). In this case, (7) would become

\[
\psi(0) = [\Gamma(z(0), 0)]^{-1} P(0)[z(0) - \hat{x}(0)]
\]

The computational procedure of Section 3 can be modified to handle cases where the two sequences \( \{u(i), i = 0, 1, \ldots, N-1 \} \) and \( \{u(j), j = 1, \ldots, N \} \) are correlated with each other, i.e.

\[
E[u(j - 1)u(k)] = M(k)\delta_{jk}
\]

where \( M(k) \) is the \( m \times m \) cross-covariance matrix, and \( k = 1, \ldots, N \). This introduces a quadratic form involving \( [x(i) - h(x(i), i)] \) and \( u(i-1), i = 1, \ldots, N \), into \( V^* \). The details are somewhat involved and will
not be pursued here.

If the system disturbance \(u(k - 1), k = 1, \ldots, N\), is not present in (1), the term involving \(\|u(i - 1)\|^2_{2^N-1}\) in \(V^*_N\) does not appear and the set of constraints in (5) is replaced by the set

\[x(i) - f(x(i - 1), i - 1) = 0\]

where \(i = 1, \ldots, N\). The procedure in Section 3 still applies in this case with certain obvious simplifications.

If the statistical description which was given in the problem formulation is not available, the computational method of Section 3 can still be applied by adopting a modified least-squares approach as was done by Detchmendy and Sridhar (1966) for continuous-time nonlinear filtering. In this case, the estimate is constrained to be of the form

\[\hat{x}(k) = f(\hat{x}(k - 1), k - 1) + u(k - 1)\]

for \(k = 1, \ldots, N\), where \(\hat{x}(0), \hat{x}(N)\) and \(u(k - 1)\) are determined so that

\[V_N = \frac{1}{2}\|x(0) - \hat{x}(0)\|^2_{4(0)} + \frac{1}{2} \sum_{i=1}^{N} \| x(i) - h(\hat{x}(i), i) \|^2_{R(i)} + \|u(i - 1)\|^2_{C(i - 1)}\]

is minimized. Here \(\hat{x}(0)\) is a preliminary estimate of \(x(0)\), and \(A(0), B(i)\) and \(C(i - 1)\) are arbitrary weighting matrices. All of these parameters are chosen on the basis of engineering considerations associated with the particular physical system of concern.

References


**Book review**


In recent years considerable efforts have been made to develop algorithms of wide application in the field of linear algebra. It is important that such algorithms should be subjected to very stringent tests and for this purpose a well designed set of test matrices is required.

In this book the authors have assembled a wide selection of matrices which were distributed throughout the literature, together with collections provided by J. Elliott, S. Voigt, R. Greenwood and J. Westlake. The test matrices are divided into two main classes. The first are appropriate for algorithms for solving systems of linear equations and computing determinants and inverses, and the second for computing the eigensystems of matrices, both symmetric and asymmetric.

For anybody involved in developing algorithms in this field it is an invaluable collection of material. I have only two minor criticisms. First, the set of matrices for testing algorithms for the unsymmetric eigenvalue problem includes results for the FRANK matrix of order twelve. It would have been far more useful if the authors had shown how to use algorithms for finding eigenvalues of symmetric tridiagonal matrices to find accurate eigenvalues of the unsymmetric FRANK matrices of all orders, thus providing the reader with a whole series of extremely valuable tests.

Secondly the book concludes with a specific DOLPH-LEWIS complex matrix of order twenty. This matrix has played an important role in the history of the development of eigenvalue algorithms. (Both the authors and the reviewer have been concerned with it.) However, the matrices are 'full' and the elements are given as 19 digit decimal numbers; this makes it a prohibitively tiresome matrix to use. As far as possible test matrices should be easy to generate and should usually contain parameters, so that a large number of tests is provided in a simple manner. This is true of most of the examples given. I hope this book will be regarded as a first attempt in this important field and that its success will stimulate the authors and others to further efforts.

J. H. WILKINSON (Teddington)