P3

M. Gevers

INT. J. CONTROL, 1975, VOL. 22, NO. 2, 197–213

Optimal sensors' allocation strategies for a class of stochastic distributed systems

S. E. AIDAROUS†, M. R. GEVERS† and M. J. INSTALLÉ†

The minimum-variance state-estimator for a class of linear distributed parameter systems, with both process and measurement disturbances, is derived. A new algorithm is presented for the optimal simultaneous spatial allocation of sensors. The algorithm minimizes recursively the spatial integral of the covariance matrix of the error in the state estimates. For time-invariant systems the algorithm leads to the minimization of the spatial integral of the steady-state error covariance matrix. The influence of the system disturbances and measurement noise on these locations is discussed. An illustrative example is given to demonstrate the numerical performance of the algorithm.

1. Introduction

The problem of estimating the state of distributed parameter systems with noisy disturbances has become of great importance in many engineering applications. Randomness can enter such systems in many ways, the measurements of the plant states may be contaminated with noise, or the plant may be subjected to random disturbances or random parameter variations. Several authors have derived various forms of the minimum-variance state-estimator for such systems under the assumption of a finite number of fixed measurements locations (Falb 1967, Tzafestas and Nightingale 1968 a, b, Thau 1969, Kushner 1970, Sakawa 1972, Aidarous and Ghonaimy 1974).

Less attention has been devoted to observability and optimal measurement location in stochastic distributed systems. The first discussion of observability of such systems was given by Wang (1964); he defined observability in terms of the existence of the inverse of a certain self-adjoint observation operator. Goodson and Klein (1970) gave a definition of observability based on restricted spatial domains and system normal modes. Yu and Seinfeld (1973) examined the effect of measurement locations on observability for a class of linear distributed systems whose solutions can be expressed by eigenfunction expansions. They developed an algorithm for determining a sub-optimal set of measurement locations.

This paper presents a minimum-variance state-estimator algorithm for a class of distributed parameter systems with both process and measurement disturbances. Given a space of admissible measurement locations, a spatial distribution of measurement noise, and a fixed number of sensors, a new algorithm is presented for the optimal determination of the sensor locations that minimize the trace of the spatial integral of the state error-covariance matrix. The algorithm can be applied to both the one-sensor allocation and the simultaneous multi-sensor allocation problems.

Received 2 April 1975.

† Laboratoire d'Automatique, Université Catholique de Louvain, Bâtiment Maxwell, 1348 Louvain-la-Neuve, Belgium.

The estimator algorithm and the sensor allocation algorithm are derived in discrete time, for purposes of digital implementation. The system behaviour at discrete instants is expressed in terms of recursive functional expressions involving the systems Green's function. After expanding all quantities in terms of a finite number of elements taken from a complete set of orthonormal basis, the minimum-variance state-estimator is obtained via the solution of a Riccati-like recursive equation. The solution of this equation is the matrix of coefficients of the expanded state error-covariance matrix, which is a function of the sensor locations.

The optimal sensor allocation algorithm is obtained via the minimization of a non-linear functional of the sensor locations; a modified gradient algorithm is used for this minimization. To simplify the derivations, the one-sensor problem is considered first; the results are subsequently generalized to the N-sensor problem.

2. Problem formulation

Consider a linear distributed system on a fixed and bounded domain Ω, with boundary $\partial\Omega$. The mathematical model of the system is described, for simplicity of notations, by the one-dimensional parabolic equation

$$\frac{\partial u(t, x)}{\partial t} = \mathscr{L}_1 u(t, x) + f(t, x), \quad x \in \Omega, \quad t \in [0, T] \tag{1}$$

with boundary conditions

$$\mathscr{L}_2 u(t, x) = 0, \quad x \in \partial\Omega \tag{2}$$

where $u(t, x)$ is the state of the system, $\mathscr{L}_1$ and $\mathscr{L}_2$ are two spatial operators.

The observations are made only at discrete instants $\tau, 2\tau, \ldots, k\tau, \ldots$. Assume first that a single measuring device is available, at that time $k\tau$ it is placed at position $x_m(k)$. For simplicity of notations we shall not write the dependence of x_m on k unless necessary, although this dependence will be implicitly assumed. The measurement equation at time $k\tau$ will therefore be written as

$$y_k(x_m) = M_k(x_m) u_k(x_m) + v_k(x_m) \tag{3}$$

where $u_k(x_m)$ is used to denote $u(k\tau, x_m)$.

The two noise processes $f(t, x)$ and $v_k(x_m)$ are statistically independent white Gaussian processes with zero mean and covariances

$$E\{f(t, x) f(t', x')\} = C(t, x, x') \delta(t - t') \tag{4}$$

$$E\{v_k(x_m) v_j(x_m)\} = Q_k(x_m) \delta_{kj} \tag{5}$$

We shall first derive the equations for the minimum-variance estimate $\hat{u}_{k|k}(x)$ of the state $u_k(x)$ for all $x \in \Omega$, based on the set of noisy measurements $\{y_j(x_m),\ 0 \leqslant j \leqslant k\}$, where the state-error covariance at time $k\tau$ will be denoted by

$$P_k(x, x') \triangleq E\{[u_k(x) - \hat{u}_{k|k}(x)][u_k(x') - \hat{u}_{k|k}(x')]\} \tag{6}$$

Next, for a given spatial distribution $P_{k-1}(x, x)$ of the state error covariance at time $(k-1)\tau$, we shall find the optimal position $x_m(k)$ of the sensor that minimizes the spatial integral of the error covariance at time $k\tau$,

$$\min_{x_m(k)} \int_{\Omega} P_k(x, x)\, dx \tag{7}$$

This minimization problem can be looked upon as a single-stage non-linear optimal control problem with $x_m(k)$ as the control variable.

3. Derivation of the filter equations with arbitrary sensor location x_m

Since the interest is in the state at discrete-time instants, the partial differential eqn. (1) will be replaced by a discrete-time equation whose solution approximates the solution of (1) at times $\tau, 2\tau, \ldots, k\tau, \ldots$. To do this, first $u_k(x)$ will be written as the solution of an integral equation. Assume that the problem is well posed in the sense of Hadamard, i.e. the solution of (1) exists uniquely and depends continuously on the initial and boundary conditions, and there exists a fundamental $G(t, t', x, x')$, which is known as the system Green's function, defined for $t-t' \geqslant 0$, and $x, x' \in \Omega$, such that

$$G(t_0, t_0, x, x') = \delta(x-x') \tag{8}$$

Then the state at $k\tau$ can be expressed as

$$u_k(x) = \int_{\Omega} G(k\tau, \overline{k-1}\tau, x, x') u_{k-1}(x')\, dx' + \int_{\Omega} \int_{t'=\overline{k-1}\tau}^{k\tau} G(k\tau, t', x, x')\, d_{t'}\, \beta(t', x')\, dx' \tag{9}$$

where $d\beta(t, x)$ is the incremental Wiener process with zero mean and incremental covariance

$$E\{d\beta(t, x)\, d\beta(t', x')\} = C(t, x, x')\, dt \tag{10}$$

where $C(t, x, x')$ is the same Kernel given in (4).

Given an optimal estimate $\hat{u}_{k-1|k-1}(x)$ at the time instant $(k-1)\tau$, the estimate at $k\tau$ can be predicted as (Aidarous and Ghonaimy 1974)

$$\hat{u}_{k|k}(x) = \hat{u}_{k|k-1}(x) + K_k(x, x_m)[Y_k(x_m) - M_k(x_m)\hat{u}_{k|k-1}(x_m)] \tag{11}$$

where

$$\hat{u}_{k|k-1}(x) = \int_{\Omega} G_{k-1}(x, x')\hat{u}_{k-1|k-1}(x')\, dx' \tag{12}$$

$$G_{k-1}(x, x') \triangleq G(k\tau, \overline{k-1}\tau, x, x') \tag{13}$$

In order to be complete, the gain function $K_k(x, x_m)$ in (11) needs to be determined. This gain has to be chosen so as to minimize $P_k(x, x)$, the covariance of the error in the estimate of the state at x. To compute this optimal gain, we first derive a recursive equation for $P_k(x_1, x_2)$.

Define

$$\tilde{u}_k(x) = u_k(x) - \hat{u}_{k|k}(x) \tag{14}$$

$$P_k(x_1, x_2) = E\{\tilde{u}_k(x_1)\tilde{u}_k(x_2)\} \tag{15}$$

Using (9), (11) and (12) eqn. (14) can be written as

$$\tilde{u}_k(x) = \int_\Omega [G_{k-1}(x, x') - K_k(x, x_m)M_k(x_m)G_{k-1}(x_m, x')]\tilde{u}_{k-1}(x')\,dx' + \int_\Omega \int_{t=\overline{k-1}\tau}^{k\tau} [G(k\tau, t', x, x') - K_k(x, x_m)M_k(x_m)G(k\tau, t', x_m, x')] \times d_{t'}\,\beta(t', x')\,dx' - K_k(x, x_m)v_k(x_m) \qquad (16)$$

The error covariance function $P_k(x_1, x_2)$ as defined in (15) can then be expressed in terms of $P_{k-1}(x', x'')$ as follows :

$$\begin{aligned} P_k(x_1, x_2) = {} & \int_\Omega \int_\Omega G_{k-1}(x_1, x')[P_{k-1}(x', x'') + \tau C_{k-1}(x', x'')]G_{k-1}(x_2, x'')\,dx'\,dx'' \\ & - \int_\Omega \int_\Omega G_{k-1}(x_1, x')[P_{k-1}(x', x'') \\ & + \tau C_{k-1}(x', x'')]G_{k-1}(x_m, x'')\,dx'\,dx''\,M_k(x_m)K_k(x_2, x_m) \\ & - K_k(x_1, x_m)M_k(x_m) \int_\Omega \int_\Omega G_{k-1}(x_m, x')[P_{k-1}(x', x'') \\ & + \tau C_{k-1}(x', x'')]G_{k-1}(x_2, x'')\,dx'\,dx'' \\ & + K_k(x_1, x_m)\{ \int_\Omega \int_\Omega M_k(x_m)G_{k-1}(x_m, x')[P_{k-1}(x', x'') \\ & + \tau C_{k-1}(x', x'')]G_{k-1}(x_m, x'')M_k(x_m)\,dx'\,dx'' \\ & + Q_k(x_m)\}K_k(x_m, x_2) \end{aligned} \qquad (17)$$

The last step was completed under the assumption that there is no correlation between $f_{k-1}(x)$ and $\hat{u}_{k-1|k-1}(x)$. This is true since the random drive at the instant $(k-1)\tau$ will affect the state (and consequently the error) only at the next instants $k\tau$, $(k+1)\tau$, ...

Notice that in order to derive the gain $K_k(x, x_m)$ for the minimum-variance estimator we only need $P_k(x, x)$, i.e. the value of $P_k(x_1, x_2)$ for $x_1 = x_2 = x$. However, as can be seen from (17), $P_k(x, x)$ depends upon $P_{k-1}(x', x'')$ for all $x', x'' \in \Omega$. This explains why a recursive relation is required for $P_k(x_1, x_2)$ rather than just for $P_k(x, x)$.

4. Solution using orthonormal bases

The minimization of $P_k(x, x)$ as given by (17) w.r.t. $K_k(x, x_m)$ is based upon the expansion of the state error covariance P, the systems Green's function G, the state noise covariance C and the distributed Kalman gains K in terms of orthonormal bases. If G and P have certain regularity properties (e.g. they can be assumed square integrable), then a set of orthonormal functions may be used to express them in series form to any required accuracy. The choice of orthonormal basis simplifies the solution and reduces the amount of required computations.

Assuming a complete set of orthonormal bases $\{z_i(x), i = 1, \infty, \ldots, x \in \Omega\}$, where

$$\int_\Omega z_i(x')z_j(x)\,dx = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases} \qquad (18)$$

and considering that the first r elements of these functions are enough for the required accuracy, then P, G, C and K can be written in the form

$$P_k(x', x'') = Z^{\mathrm{T}}(x') W_k Z(x'') \tag{19}$$

$$G_k(x', x'') = Z^{\mathrm{T}}(x') A_k Z(x'') \tag{20}$$

$$C_k(x', x'') = Z^{\mathrm{T}}(x') H_k Z(x'') \tag{21}$$

$$K_k(x, x_m) = Z^{\mathrm{T}}(x) D_k(x_m) \tag{22}$$

where

$$Z(x) = \begin{bmatrix} z_1(x) \\ \vdots \\ z_r(x) \end{bmatrix} \tag{23}$$

and W, A, H and D are $r \times r$, $r \times r$, r matrices and vector of coefficients.

Using the previous notations and making use of (18) eqn. (17) can be written in the form

$$\begin{aligned} P_k(x_1, x_2) = {} & Z^{\mathrm{T}}(x_1) E_{k-1} Z(x_2) - Z^{\mathrm{T}}(x_1) E_{k-1} Z(x_m) M_k(x_m) D_k{}^{\mathrm{T}}(x_m) Z(x_2) \\ & - Z^{\mathrm{T}}(x_1) D_k(x_m) M_k(x_m) Z^{\mathrm{T}}(x_m) E_{k-1} Z(x_2) \\ & + Z^{\mathrm{T}}(x_1) D_k(x_m) [M_k(x_m) Z^{\mathrm{T}}(x_m) E_{k-1} Z(x_m) M_k(x_m) \\ & + Q_k(x_m)] D_k{}^{\mathrm{T}}(x_m) Z(x_2) \end{aligned} \tag{24}$$

where

$$E_{k-1} = A_{k-1} [W_{k-1} + \tau H_{k-1}] A_{k-1}{}^{\mathrm{T}} \tag{25}$$

Define

$$B_k(x_m) = M_k(x_m) Z^{\mathrm{T}}(x_m) \tag{26}$$

$$\theta_{k-1}(x_1, x_2) = Z^{\mathrm{T}}(x_1) E_{k-1} Z(x_2) \tag{27}$$

$$\Psi_k(x, x_m) = Z^{\mathrm{T}}(x) E_{k-1} B_k{}^{\mathrm{T}}(x_m) \tag{28}$$

$$\Phi_k(x_m) = B_k(x_m) E_{k-1} B_k{}^{\mathrm{T}}(x_m) + Q_k(x_m) \tag{29}$$

Then eqn. (24) can be written in the compact form

$$\begin{aligned} P_k(x_1, x_2) = {} & \theta_{k-1}(x_1, x_2) - \Psi_k(x_1, x_m) K_k{}^{\mathrm{T}}(x_2, x_m) - K_k(x_1, x_m) \\ & \times \Psi_k{}^{\mathrm{T}}(x_2, x_m) + K_k(x_1, x_m) \Phi_k(x_m) K_k{}^{\mathrm{T}}(x_2, x_m) \end{aligned} \tag{30}$$

The optimal gain for the minimum-variance filter can now be derived from the last expression. For $x_1 = x_2 = x$ and for a given $P_{k-1}(x, x)$ (i.e. for a given θ_{k-1}), the error covariance $P_k(x, x)$ will be minimized for

$$K_k(x, x_m) = \Psi_k(x, x_m) \Phi_k{}^{-1}(x_m) \tag{31}$$

where $\Psi_k(x, x_m)$ and $\Phi_k(x_m)$ are given by (28) and (29), and E_k obeys the following recursive equation :

$$\begin{aligned} E_k = {} & A_k E_{k-1} A_k{}^{\mathrm{T}} - A_k E_{k-1} B_k{}^{\mathrm{T}}(x_m) [B_k(x_m) E_{k-1} B_k{}^{\mathrm{T}}(x_m) + Q_k(x_m)]^{-1} \\ & \times B_k(x_m) E_{k-1} A_k{}^{\mathrm{T}} + \tau A_k H_k A_k{}^{\mathrm{T}} \end{aligned} \tag{32}$$

With this optimal gain the corresponding state error covariance is given by

$$P_k(x, x) = \theta_{k-1}(x, x) - \Psi_k(x, x_m)\Phi_k^{-1}(x_m)\Psi_k^{\mathrm{T}}(x, x_m) \tag{33}$$

which, of course, depends upon the sensor position x_m, through $\Psi_k(x, x_m)$ and $\Phi_k(x_m)$.

In the following section the spatial integral of the state error covariance $P_k(x, x)$, which is already minimized w.r.t. the filter gains, will be minimized w.r.t. the sensor location x_m, for a given $P_{k-1}(x, x)$.

5. Gradient algorithm for the calculation of x_m

On examination of the definitions of $\Psi_k(x, x_m)$ and $\Phi_k(x_m)$ (see eqns. (28), (29) and (26)), it can be seen that $P_k(x, x)$ depends upon x_m via the output matrix $M_k(x_m)$ and the measurement noise covariance function $Q_k(x_m)$, which was to be expected.

In order to find the optimal sensor location the spatial integral of $P_k(x, x)$ will be minimized w.r.t. x_m, assuming that $P_{k-1}(x_1, x_2)$ is given,

$$\min_{x_m} \int_{\Omega} P_k(x, x)\, dx \tag{34}$$

This integral is a measure of the total error covariance.

By expression (33) this is equivalent to

$$\min_{x_m} \int_{\Omega} [\theta_{k-1}(x, x) - \Psi_k(x, x_m)\Phi_k^{-1}(x_m)\Psi_k^{\mathrm{T}}(x, x_m)]\, dx \tag{35}$$

For a given $\theta_{k-1}(x, x)$ this is equivalent to

$$\max_{x_m} \int_{\Omega} \Psi_k(x, x_m)\Phi_k^{-1}(x_m)\Psi_k^{\mathrm{T}}(x, x_m)\, dx \tag{36}$$

Substituting for $\Psi_k(x, x_m)$ from (28) this can be written in the form

$$\max_{x_m} \int_{\Omega} Z^{\mathrm{T}}(x)E_{k-1}B_k^{\mathrm{T}}(x_m)\Phi_k^{-1}(x_m)B_k(x_m)E_{k-1}Z(x)\, dx \tag{37}$$

Using the orthonormality properties of $Z(x)$, the above criterion reduces to

$$\max_{x_m} \mathrm{tr}\,[E_{k-1}B_k^{\mathrm{T}}(x_m)\Phi_k^{-1}(x_m)B_k(x_m)E_{k-1}] \tag{38}$$

The last quantity will be defined by $J(x_m)$,

$$J(x_m) = \mathrm{tr}\,[E_{k-1}B_k^{\mathrm{T}}(x_m)\Phi_k^{-1}(x_m)B_k(x_m)E_{k-1}] \tag{39}$$

Clearly $J(x_m)$ is a non-linear function of x_m. To find the maximum of this function a modified conjugate gradient algorithm will be used, which is effective in such a case (Polak 1971).

The algorithm will be presented here for the single-sensor case, while a generalization for the multi-sensor case will be given later,

$$x_m(n+1) = x_m(n) + \alpha(n)\gamma(n) \tag{40}$$

The direction vector $\gamma(n)$ (scalar in this case) is given by

$$\gamma(n+1) = -g(n+1) + \eta\gamma(n)S(n) \tag{41}$$

$$S(n) = \frac{g(n+1)g(n+1)}{g(n)g(n)} \tag{42}$$

where η has the values 0 or 1, whether we are using the gradient or the conjugate gradient (to accelerate convergence).

The gradient $g(n)$ is given by

$$g(n) = \nabla_{x_m} J(x_m)\big|_{x_m = x_m(n)} \tag{43}$$

where

$$\nabla J(x_m) = \mathrm{tr}\,[E_{k-1}\{\nabla B_k^{\mathrm{T}}(x_m)\Phi_k^{-1}(x_m)B_k(x_m) + B_k^{\mathrm{T}}(x_m)\nabla\Phi_k^{-1}(x_m)B_k(x_m) + B_k^{\mathrm{T}}(x_m)\Phi^1(x_m)\nabla B_k(x_m)\}E_{k-1}] \tag{44}$$

The initial position $x(0)$ should be selected as the best *a priori* estimate of the position. The initial direction $\gamma(0)$ is specified arbitrarily as is commonly chosen as

$$\gamma(0) = -g(0) \tag{45}$$

The gain coefficient $\alpha(n)$ is adapted automatically to accelerate the convergence following a procedure that is described below. The practical implementation of the above algorithm is as follows (see the flow chart of Fig. 1) :

1. Choose the orthonormal basis so as to satisfy the imposed boundary conditions and incorporate all available *a priori* information about the system.
2. Fix the initial estimates of x_m and W, as well as the stopping variables S_1, S_2, S_3 and S_4 (see Appendix).
3. Solve the filter Riccati equation until the transients due to the filtering algorithm are removed.
4. Apply the allocation algorithm to get a new position of the sensor, and check that the new position is in the spatial domain Ω. If not, take a new position halfway between the old and new value of x_m, and so on until the new position is in Ω.
5. If the distance between the old and new positions of x_m is smaller than some prespecified value δ_2, then stop and choose the last position for the sensor location. If not, go to step 6.
6. Iterate the filter Riccati equation with the updated position x_m.
7. If the cost corresponding to the next position is greater than that of the previous one, then take a new x_m halfway between the two most recent values and go to step 6.
8. If the gradient is near zero (i.e. smaller than some prespecified δ_3) then stop and take the optimum value of x_m.

   If the gradient changes its sign (i.e. we have bypassed the optimum value), then decrease the gain (e.g. divide by 2), and go to step 4. If the gradient has the same sign (i.e. we are still going towards the optimum value), then increase the gain (e.g. multiply by 2), and go to step 4.

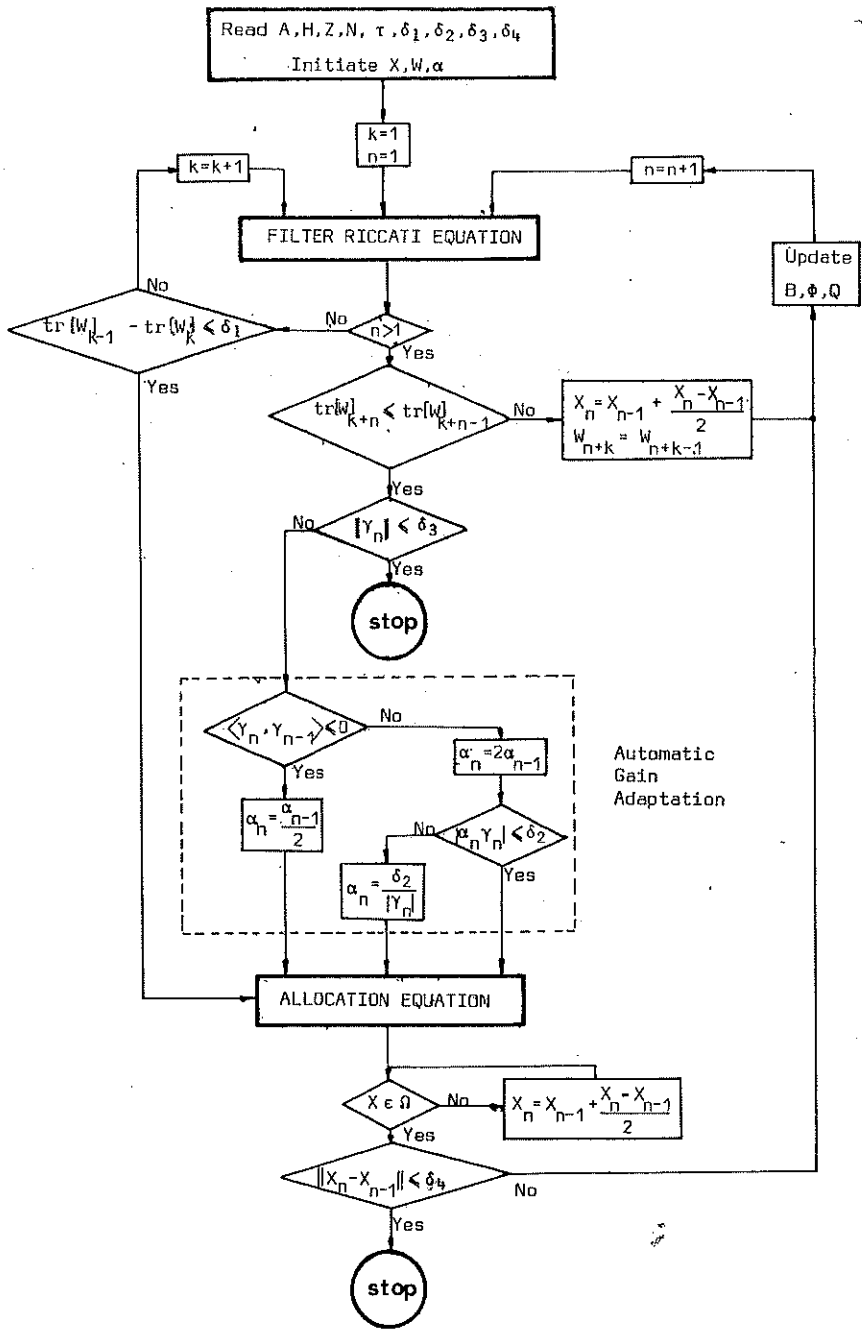

Figure 1. Flow chart of the allocation algorithm.

6. Generalization of results

The generalization of the foregoing results to the simultaneous optimal allocation of N sensors is straightforward. Suppose that we have N sensors placed at $x_i \in \Omega$, $i = 1, \ldots, N$, and the measurements are given by

$$Y_k(x_i) = M_k(x_i)u_k(x_i) + v_k(x_j), \quad i = 1, \ldots, N \tag{46}$$

with the assumption

$$E\{v_k(x_j)v_1(x_j)\} = 0, \quad k \neq 1, i \neq j \tag{47}$$

The optimal state estimate will have the form

$$\hat{u}_{k|k}(x) = \hat{u}_{k|k-1}(x) + \sum_{i=1}^{N} K_k(x_2, x_i)[Y_k(x_i) - M_k(x_j)\hat{u}_{k|k-1}(x_i)] \tag{48}$$

The gains $K_k(x, x_i)$, $i = 1, \ldots, N$ will be computed from

$$K_k(x) = \Psi_k(x)\Phi_k^{-1} \tag{49}$$

where

$$K_k(x) = [K_k(x, x_1) \ \ldots \ K_k(x, x_N)] \tag{50}$$

$$\Psi_k(x) = Z^{\mathrm{T}}(x)E_{k-1}B_k{}^{\mathrm{T}} \tag{51}$$

$$\Phi_k = B_k E_{k-1} B_k{}^{\mathrm{T}} + Q_k \tag{52}$$

$$B_k = [M_k(x_1)Z^{\mathrm{T}}(x_1) \ \ldots \ M_k(x_N)Z^{\mathrm{T}}(x_N)] \tag{53}$$

$$Q_k = \mathrm{diag}\,[Q_k(x_1) \ \ldots \ Q_k(x_N)] \tag{54}$$

The matrix E_{k-1} that is necessary to compute Φ and Ψ obeys the same recursive relation (32), taking into account the new dimensions of B, Q, and D. The same holds for the error covariance matrix $P_k(x_1, x_2)$ which will have the same expression as in (30).

The criterion to be optimized in this case will be (cf. (39))

$$J(X) = \mathrm{tr}\,[E_{k-1}B_k{}^{\mathrm{T}}\Phi_k{}^{-1}B_k E_{k-1}] \tag{55}$$

where X is an N vector of the sensor's positions given by

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \tag{56}$$

and the gradient algorithm will have the form

$$X(n+1) = X(n) - \Gamma(n)\gamma(n) \tag{57}$$

where the direction vector is governed by

$$\gamma(n+1) = -g(n+1) + \eta\gamma(n)S(n) \tag{58}$$

and

$$S(n) = \frac{g^{\mathrm{T}}(n+1)g(n+1)}{g^{\mathrm{T}}(n)g(n)} \tag{59}$$

The gradient vector g is given by

$$g(n)=\left.\begin{matrix}\dfrac{\partial J(X)}{\partial x_1}\\ \vdots\\ \dfrac{\partial J(X)}{\partial x_N}\end{matrix}\right|_{X=X(n)} \tag{60}$$

and the gain matrix $\Gamma(n)$ will have the form

$$\Gamma(n)=\text{diag}\,[\alpha(n,1)\ \ldots\ \alpha(n,N)] \tag{61}$$

In this case Φ is a matrix and the standard formulae for the derivation of the inverse of a matrix (Derusso *et al.* 1967) is needed. Interchanging the derivative and trace operators, the following is obtained :

$$\nabla_X J(X)=\text{tr}\,[E_{k-2}\{\nabla B_k{}^{\mathrm{T}}\Phi_k{}^{-1}B_k+B_k{}^{\mathrm{T}}\nabla\Phi_k{}^{-1}B_k+B_k{}^{\mathrm{T}}\Phi_k{}^{-1}\nabla B_k\}E_{k-1}] \tag{62}$$

and

$$\nabla_X\Phi_k{}^{-1}=-\Phi_k{}^{-1}\nabla\Phi_k\Phi_k{}^{-1} \tag{63}$$

The search procedure is depicted in Fig. 1. The choice of the stopping rules is discussed in the Appendix.

7. Illustrative example

Consider a slab-type nuclear reactor in which the time behaviour of the neutron flux spatial distribution is represented by the one-dimensional diffusion equation,

$$\frac{\partial u}{\partial t}=\gamma_1\frac{\partial^2 u}{\partial x^2}+\gamma_2 u+\xi(t,x),\quad x\in[0,h]$$

with boundary conditions

$$u(t,0)=u(t,h)=0,\quad \forall t$$

and $\xi(t,x)$ is a white Gaussian noise with zero mean and covariance

$$E\{\xi(t,x)\xi(t,x')\}=C(x)\delta_D(x,x')$$

Neutron flux measurement transducers are movable in their ionization chambers during the reactor operation. These movements are between prespecified finite positions which are usually determined *a priori* from safety considerations.

We shall consider both the one-sensor and two-sensor case.

For the latter, the measurement equation is given by

$$y_k(x_i)=u_k(x_i)+v_k(x_i),\quad i=1,2$$

where the measurement noise $v_k(x)$ is a white Gaussian noise with zero mean and covariance

$$E\{v_k(x_i)v_1(x_j)\}=q(x_i)\quad \text{if } \|x_i-x_j\|\leqslant\delta$$
$$=0\qquad \text{otherwise}$$

for some small δ.

The following numerical values are taken :

$$\gamma_1 = 1600$$

$$\gamma_2 = 0{\cdot}252$$

$$h = 250 \text{ cm}$$

The steady-state pattern of such a system is known to be a half sine wave with an amplitude that is a function of the shape of the disturbance (Hassan *et al.* 1971). Therefore, it is reasonable to choose the orthonormal sets of basis functions as

$$\left\{ \sqrt{\left(\frac{2}{h}\right)} \sin \frac{\pi}{h} ix \; ; \quad i = 1, 2, \ldots \right\}$$

For the example considered,

$$G(t, t', x, x') = \frac{2}{h} \sum_{i=1}^{\infty} \exp\left(-\frac{\pi^2 \gamma_1}{h^2} i^2 + \gamma_2 \right) (t - t') \, . \sin \frac{\pi}{h} ix \sin \frac{\pi}{h} ix'$$

and the matrix A is a constant diagonal matrix with the ith diagonal element given by

$$a_{ii} = \exp\left(-\frac{\pi^2 \gamma_1}{h^2} i^2 + \gamma_2 \right) \tau$$

where τ is the sampling period between observations which was taken as 0·1 sec.

Before applying the allocation algorithm to determine the optimal sensor locations the space was normalized such that $x \in [0, 1]$.

(*a*) *One-sensor case*

The algorithm was applied to the above example assuming that only one sensor was to be allocated. When choosing constant and equally distributed system and measurement noise covariances (i.e. $c(x)$ and $q(x)$ constants), the optimal sensor position converges to the centre of the space ($x^* = 0{\cdot}5$) as shown in Fig. 2.

Next a more selective noise covariance profile is chosen, e.g.

$$c(x) = 0{\cdot}1(1 - 0{\cdot}9 \sin \pi x)$$

$$q(x) = 0{\cdot}1(1 - 0{\cdot}9 \sin 3\pi x)$$

The sensor location now converges to either one of two equivalent optimal positions as shown in Fig. 3.

In the table different combinations of measurement and distributed noise covariances are considered and the results when applying the algorithm are depicted. From this table we can see that the profile of the distributed system noise has less influence on the optimal sensor positions than the profile of the measurement noise ; this is due to the low diffusion coefficient considered in this example.

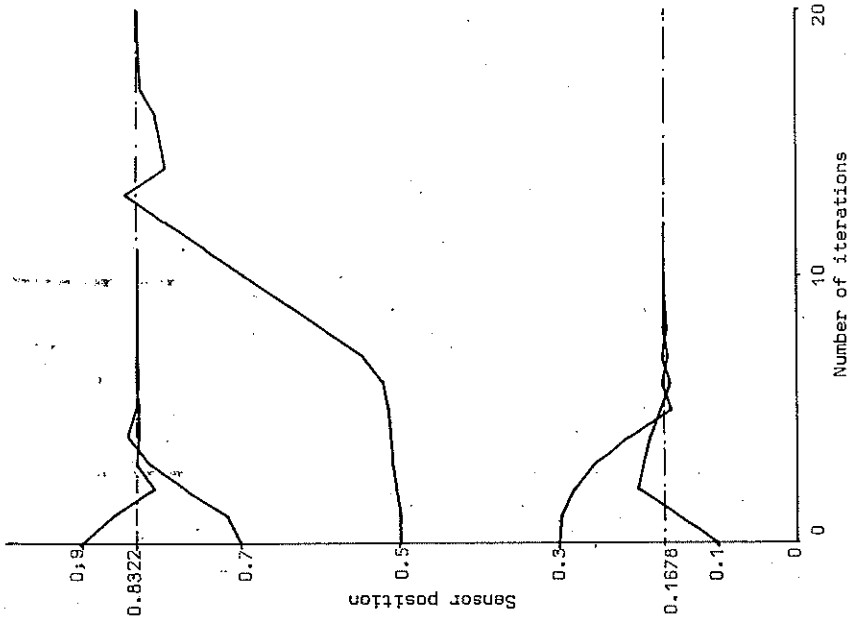

Figure 3. Convergence of sensor position for a constant process noise and a selective measurement noise covariances.

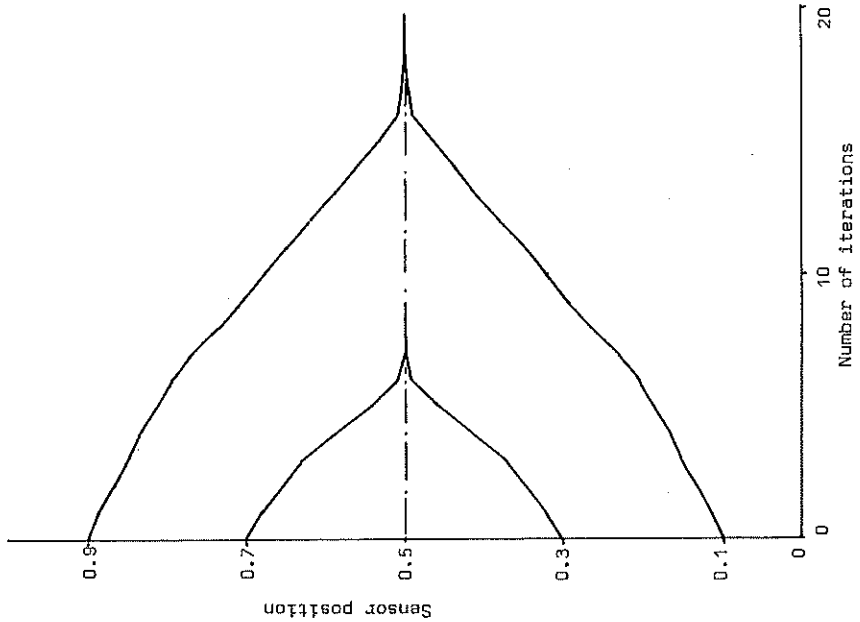

Figure 2. Convergence of sensor position for constant process and measurement noise covariances.

$c(x)$ \ $q(x)$		1	$1-0{\cdot}9\sin\pi x$	$1-0{\cdot}9\sin 2\pi x$	$1-0{\cdot}9\sin 3\pi x$
0·1	x^*	0·5	0·5	0·42	0·5
	P^*	0·1680	0·1245	0·1584	0·1934
$0{\cdot}1(1-0{\cdot}9\sin\pi x)$	x^*	0·5	0·5	0·31	0·23 ; 0·77
	P^*	0·0772	0·0538	0·0641	0·0775
$0{\cdot}1(1-0{\cdot}9\sin 2\pi x)$	x^*	0·62	0·6	0·575	0·7
	P^*	0·1385	0·0979	0·1550	0·1433
$0{\cdot}1(1-0{\cdot}9\sin 3\pi x)$	x^*	0·5	0·5	0·43	0·5
	P^*	0·1175	0·0724	0·1110	0·1439

$$P^* = \min_{x_m} \int_\Omega P_\infty(x, x)\, dx.$$

(*b*) *Two-sensor case*

Consider now the problem of optimally placing two sensors at x_1 and x_2. A constant covariance for the system noise is considered, namely $c(x) = 0{\cdot}1$ and for each sensor, a measurement noise covariance function of the form

$$q(x_i) = 1 - \sin 3\,\pi x_i, \quad i = 1, 2$$

Figure 4 shows the contours of the cost function and the performance of the allocation algorithm for different initial values of x_1 and x_2.

It must be noticed that in this example the optimal position could have been obtained by first optimizing w.r.t. the first sensor, and then optimizing w.r.t. the second, while keeping the first fixed.

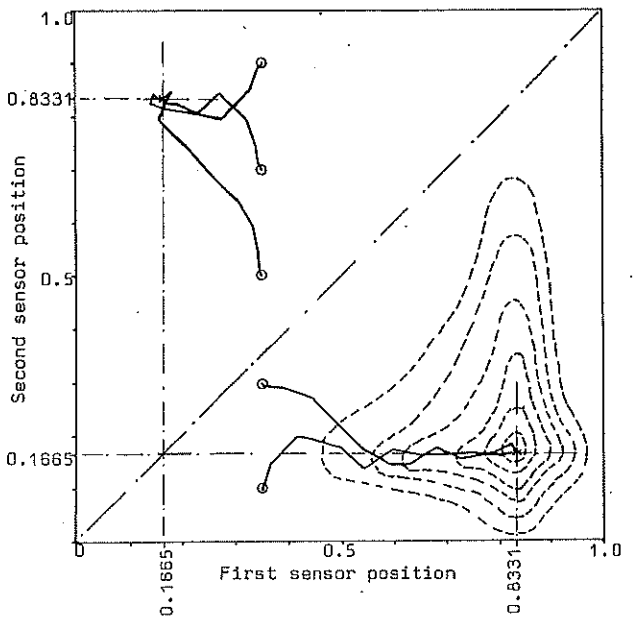

Figure 4. Two-sensor case with symmetrical covariance distributions [⊙ initial point, * optimal point).

When we take an asymmetrical distribution for the system noise, e.g.

$$c(x) = 0{\cdot}1(1 - 0{\cdot}5 \sin 2\pi x)$$

while for each sensor, a measurement noise covariance of the form

$$g(x_i) = 0{\cdot}2(1 - 0{\cdot}8 \sin 2\,\pi x_i), \quad i = 1, 2$$

the cost will have the form given in Fig. 5. In this case the location of the two sensors must be optimized simultaneously. The simulations of the predetermined example are carried out using an IBM 370/158. The average running times for the single and two-sensor cases are 0·9 and 2·0 sec respectively.

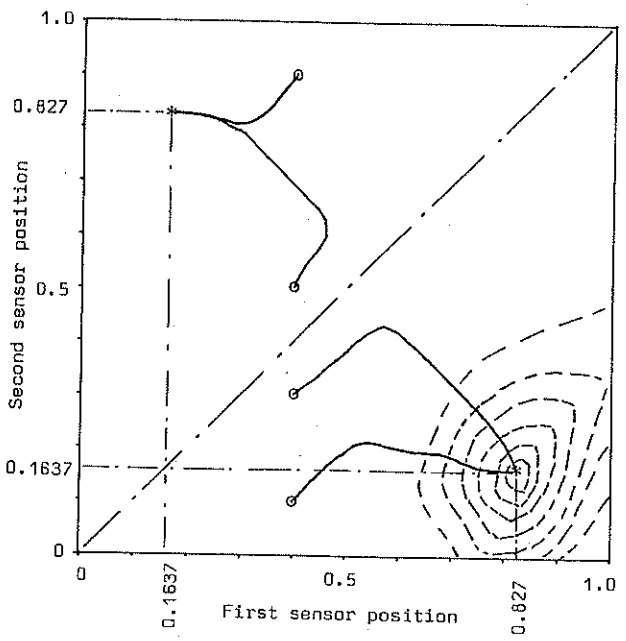

Figure 5. Two-sensor case with asymmetrical covariance distributions [○ initial point, * optimal point).

8. Conclusion

The minimum-variance estimator is designed for a class of linear distributed-parameter systems with state and measurement disturbances.

An iterative sensor allocation algorithm has been presented for the simultaneous optimal allocation of the sensors. The simplicity of the algorithm makes it possible for practical implementation since the variational problem has been replaced by the solution of recursive algebraic equations and the minimization of a non-linear function that is easily computable. The numerical performance of the proposed algorithm has been studied as well as the influence of different disturbance processes. An important feature of the algorithm is the rapid convergence towards the optimum value, as well as the prevention of possible oscillations. This is due to the automatic adaptation of the gains in the gradient algorithm. The algorithm has been applied to a diffusion process, and the convergence was achieved after a low number of iterations.

ACKNOWLEDGMENTS

The authors wish to thank Professor M. A. R. Ghonaimy for reading the manuscript and suggesting improvements.

Appendix

In the following we will discuss briefly the four stopping rules used in the aforementioned algorithm.

(i) δ_1 should be a simple linear function of the previous cost, e.g.

$$\delta_1|_k = c_1 \text{ tr } [W_{k-1}]$$

where c_1 is a small positive constant. Typical values used in the example are between 0·05 and 0·1, i.e. the transients in the solution of the Riccati equation are considered removed when their values are below 5–10% of the current value of the cost.

(ii) δ_2 is upper bounded by

$$\delta_2 \leqslant \min (c_2, c_3)$$

where c_2 is the maximum mechanically allowable step of the sensor, taking into account speed limitation of the sensor movement, and c_3 is the maximum step allowed for the sensor so that its response in its new position is not affected by the dynamics of the movement.

(iii) δ_3, which stands as a hypothetical zero, is taken as a very small positive number. It must be noticed that there is high correlation between δ_3 and δ_4, this is due to the fact that both of them appoint the optimum value of the function to be optimized. However, the influence of δ_4 on the cost is relatively high when the region of search is very narrow or banana-shaped.

(iv) The choice of δ_4 is not an easy task, since it is influenced by many factors which differ from one problem to another depending on the geometry of the search region. It is bounded by $C_4 < \delta_4 < C_5$, where C_4 is governed by the physical considerations, namely the mechanical accuracy in setting the transducers, while C_5 is given as the maximum allowable error in calculating the position. Usually both of these two values are known from the problem data or from previous experience. The choice criterion of δ_4 between those two limits is a compromise between the computational effort and the accuracy in the position resulting in a cost reduction (see Fig. 6), i.e.

$$\min_{\delta_4} (g+f)$$

where g is the cost of computational effort. It can be expressed as a linear function in the number of iterations n,

$$g = a_1 + b_1 n$$

f is the cost resulting from a loss in the achieved accuracy, hence in the performance index. f is proportional to the relative change of the cost w.r.t. the unknown optimum value W^*

$$f = b_2 \frac{\text{tr } [W_{n-1}] - \text{tr } [W_n]}{\text{tr } [W^*]}$$

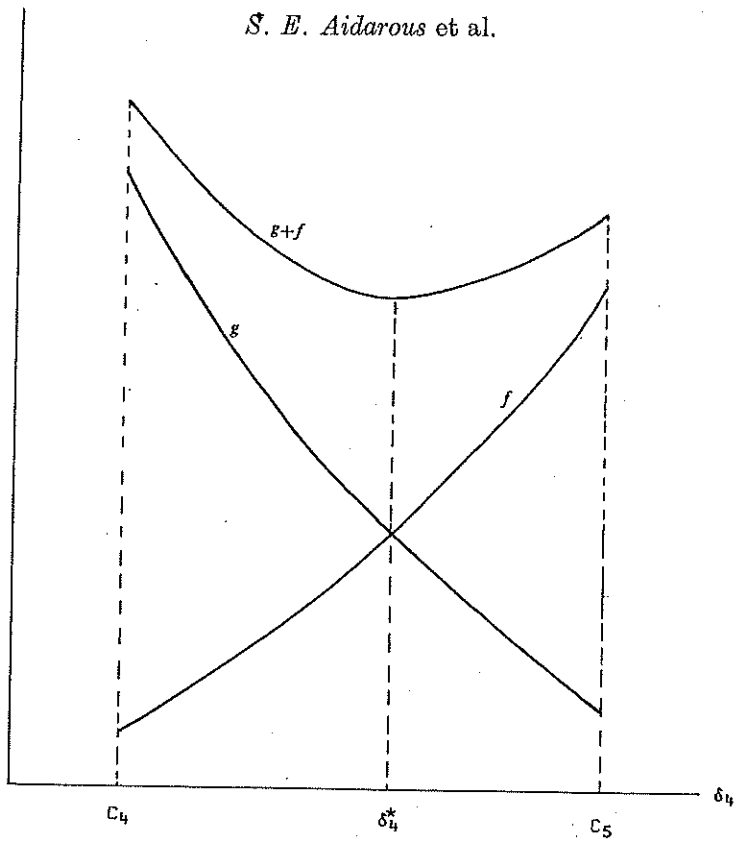

Figure 6. Choice criterion for δ_4.

Assume that the maximum allowable error c_5 is achieved after $n = n_{\max}$ iterations. If c_5 is reduced to a new value δ_s, it is required that $(g+f)$ decreases as n is increased from $n_{\max}$ to n_s :

$$\Delta(g+f) < 0$$

$$\Delta g < -\Delta f$$

where

$$\Delta g = b_1(n_s - n_{\max})$$

$$\Delta f \simeq b_2 \sum_{n=n_{\max}}^{n_s} \left\{ \frac{\operatorname{tr}[W_n] - \operatorname{tr}[W_{n-1}]}{\operatorname{tr}[W_n]} - \frac{\operatorname{tr}[W_{n-1}] - \operatorname{tr}[W_{n-1}]}{\operatorname{tr}[W_{n-1}]} \right\}$$

$$= b_2 \sum_{n=n_{\max}}^{n_s} \left\{ \frac{\operatorname{tr}[W_{n-2}]}{\operatorname{tr}[W_{n-1}]} - \frac{\operatorname{tr}[W_{n-1}]}{\operatorname{tr}[W_n]} \right\} \quad = b_2 \left\{ \frac{\operatorname{tr}[W_{n_{\max}-2}]}{\operatorname{tr}[W_{n_{\max}-1}]} - \frac{\operatorname{tr}[W_{n_s-1}]}{\operatorname{tr}[W_{n_s}]} \right\}$$

Since we are approaching the optimum δ_4 from the right (see Fig. 6), Δf will be negative. The numerical test to be carried out is :

1. Calculate Δg and Δf corresponding to an increase on the number of iterations from $n_{\max}$ to n_s.
2. If $\dfrac{\Delta f}{\Delta g} > 1$ continue to decrease δ_4.
3. If $\dfrac{\Delta f}{\Delta g} \leqslant 1$ The optimum δ_4 is reached.

References

Aidarous, S. E., and Ghonaimy, M. A. R., 1974, *12th Annual Allerton Conference on Circuit and System Theory*, Monticello, Illinois, October.

Derusso, P. M., Roy, R. J., and Close, C. M., 1967, *State Variables for Engineers* (London : Wiley).

Falb, P. L., 1967, *Inf. Control*, **11**, 102.

Goddson, R. E., and Klein, R. E., 1970, *I.E.E.E. Trans. autom. Control*, **15**, 165.

Hassan, M. A., Ghonaimy, M. A. R., and Abd El-Shaheed, M. A., 1971, *Proc. IFAC Symp. on Control of Distributed Parameter Systems*, Banff, Canada.

Kushner, H. K., 1970, *J. SIAM Control*, **8**, 346.

Polak, E., 1971, *Computational Methods in Optimization : A Unified Approach* (New York : Academic Press).

Sakawa, Y., 1972, *Int. J. Control*, **16**, 115.

Thau, F. E., 1969, *J. bas. Engng*, D, **91**, 173.

Tzafestas, S. G., and Nightengale, J. M., 1968 a, *Proc. Instn elect. Engrs*, **115**, 1207 ; 1968 b, *Ibid.*, **115**, 1737.

Wang, P. K. C., 1964, *Advances in Control Systems*, Vol. 1, edited by C. T. Leondes (New York : Academic Press).

Yu, T. K., and Seinfeld, J. H., 1973, *Int. J. Control*, **18**, 787.