C 10

to appear in "Lecture notes in Control and Information Sciences", ed. Prof. Ruberti Springer-Verlag, 1977.

ON THE ASYMPTOTIC BEHAVIOR OF SENSORS' ALLOCATION ALGORITHM IN STOCHASTIC DISTRIBUTED SYSTEMS

S.E.Aidarous*, M.R.Gevers and M.J.Installe
Laboratoire d'Automatique
Universite Catholique de Louvain
B-1348 louvain La Neuve, Belgium

INTRODUCTION

In a recent paper[1], an algorithm is presented for the optimal simultaneous allocation of a finite number of sensors in a stochastic distributed parameter system. The allocation algorithm considered is based on the recursive of a Riccati equation together with the minimization of a nonlinear functional of the sensors' locations. This minimization is performed recursively through a modified gradient algorithm, that operates simultaneously with the Riccati equation. At each iteration certain parameters of the Riccati equation are thereby changed.

When applying the algorithm recursively on a time-invariant system, two important questions will arise for the resulting time-variant Riccati equation. First, the existence of a steady-state solution, i.e. the determination of conditions to be satisfied for such a solution to exist. Secondly, the stability of the algorithm, i.e. does the effect of initial errors become negligible as time evolves.

In this paper, the above two questions will be investigated. First, the existence of a steady-state optimal solution is demonstrated, the necessary conditions for the convergence of the algorithm towards this optimal solution are then discussed.

PROBLEM FORMULATION

In the sensors allocation problem [1], a set of sensors' positions X_s is called optimal if it minimizes the trace of the spatial integral of the steady-state error covariance

* Presently with the computer and control section, Faculty of Engineering, Ain Shams University, Cairo, Egypt.

$$\min_{X_s} \left\{ \mathrm{tr} \int_{\Omega} \underset{\infty}{P}(x,x) \quad d\Omega \right\} \tag{1}$$

which can be reduced to

$$\min_{X_s} \left\{ \mathrm{tr} \ [W_\infty] \right\} \tag{2}$$

where W_∞ is the matrix of expansion coefficients of P_∞. W is obtained via the solution of the following matrix Riccati equation [1]

$$E_{k+1} = A\ E_k\ A^T - A\ E_k\ B_k^T(X_s)\ Q_k^{-1}(X_s)\ B_k(X_s)\ E_k\ A^T + \tau A\ H\ A^T \tag{3}$$

where

$$E_k = A\ [W_k + \tau H]\ A^T \tag{4}$$

$$Q_k(X_s) = B_k(X_s)\ E_k\ B_k^T(X_s) + R_k(X_s) \tag{5}$$

B_k and R_k are functions of the input matrix and the measurement error covariance respectively. Their variation with time depends upon the movement of the sensors as determined by the following recursive equation for the sensors' positions

$$X_{k+1} = X_k + \alpha_k \ \nabla_k\ J(X_k)\ , \qquad X_s \in \Omega_s \tag{6}$$

where

$$J(X_k) = \mathrm{tr}\ [E_k\ B_k^T(X_s)\ Q_k^{-1}(X_s)\ B_k(X_s)\ E_k] \Big|_{X_k} \tag{7}$$

and Ω_s is the space of admissible measurement locations.

It is required to show the existence of an optimal solution to the minimization problem (2). If we call X_s^* this optimal set of sensors' positions, then it must be shown that this set can be obtained by the repeated application of (6) and (3).

<u>EXISTENCE OF SOLUTION FOR THE ALLOCATION PROBLEM</u>

The steady-state set of sensor positions is defined as

$$\lim_{k \to \infty} \quad X_s(k) \tag{8}$$

and $X_s(k)$ is obtained by the repeated application of the algorithm (3) and (6).

Clearly E_k , (and consequently W_k), are functions of E_o (and W_o),

$X_s(0)$ and the sequence

$$X_s^k = \{X_s(j),\ j=1,2,\ldots,k\} \tag{9}$$

Thus we can write

$$E_k = E_k\{E_o, X_s(0), X_s^k\} \tag{10}$$

$$W_k = W_k\{W_o, X_s(0), X_s^k\} \tag{11}$$

In the sequel, if E_1 and E_2 are symmetric matrices, $E_1 > E_2$ [$E_1 \geqslant E_2$] means that $E_1 - E_2$ is positive [semi-positive] definite. This practical order satisfies the conditions:

(i) $E_1 \geqslant E_2$, $E_1' \geqslant E_2'$ implies $E_1 + E_1' \geqslant E_2 + E_2'$,

(ii) $E_1 \geqslant E_2$ implies $E^T E_1 E \geqslant E^T E_2 E$ for any matrix or vector E,

(iii) $E_2^{-1} \geqslant E_1^{-1}$ if $E_1 \geqslant E_2 > 0$.

Definition: We shall say that the system is ρ-observable at X_s if for a fixed set of sensor positions X_s (i.e. $X_s(k)=X_s\ \forall k$), there exists $L>0$, and a fixed $\rho>0$, such that

$$\mathrm{tr}\ P_k(x,x') < \rho < \infty \qquad \forall k \geqslant L \text{ and } x,x' \in \Omega \tag{12}$$

Recall that Ω_s is the subspace of observable positions: $\Omega_s \subseteq \Omega$, i.e. if $X_s \in \Omega_s$, then condition (12) is satisfied for some ρ and some L.

Theorem 1: (Existence theorem)
Let Ω_s be non empty. If A is invertible and if $\forall X_s \in \Omega_s$ $[A, B_k(X_s)]$ is an observable pair, and $R_k(X_s)$ and $A H A^T$ are positive definite, then there exists an optimal set of locations X_s^* that minimizes the trace of the spatial integral of the steady-state error covariance matrix given by the Riccati equation (3).

Proof:
For each fixed $X_s \in \Omega_s$, we can define a time-invariant Riccati equation

$$E_{k+1} = A E_k A^T - A E_k B^T Q_k^{-1} B E_k A^T + \tau A H A^T \tag{13}$$

with initial condition

$$E_o = A [W_o + \tau H] A^T \tag{14}$$

Under the conditions stated above Caines and Mayne [2] have proved that starting with any positive semi-definite initial condition E_o, iteration with the matrix Riccati equation (13) gives a bounded sequence of positive semi-definite matrices E_k, k=1,2,.... This sequence converges to a unique positive definite matrix E_∞ that is independent of the initial condition E_o and could in principle have been found by solving directly the steady-state equation

$$E_\infty = A\ E_\infty\ A^T - A\ E_\infty\ B^T\ Q_\infty^{-1}\ B\ E_\infty A^T + \tau A\ H\ A^T \tag{15}$$

where

$$Q_\infty^{-1} = B\ E_\infty\ B^T + R \tag{16}$$

Therefore, the limit exists

$$\lim_{k \to \infty} E_k = E_\infty \tag{17}$$

or

$$\lim_{k \to \infty} W_k = W_\infty \tag{18}$$

where both E_∞ and W_∞ are functions of the fixed positions X_s, i.e. $E_\infty(X_s)$ and $W_\infty(X_s)$.

Define

$$V_\infty(X_s) = \mathrm{tr}\ [W_\infty(X_s)] = \mathrm{tr} \int_\Omega P_\infty(x,x)\ d\Omega \tag{19}$$

Then for each $X_s \in \Omega_s$ there exists a $V_\infty(X_s)$ with the following properties

$$V_\infty(X_s) \geqslant 0 \tag{20}$$

$$V_\infty(X_s) \leqslant \rho S_\Omega \tag{21}$$

where

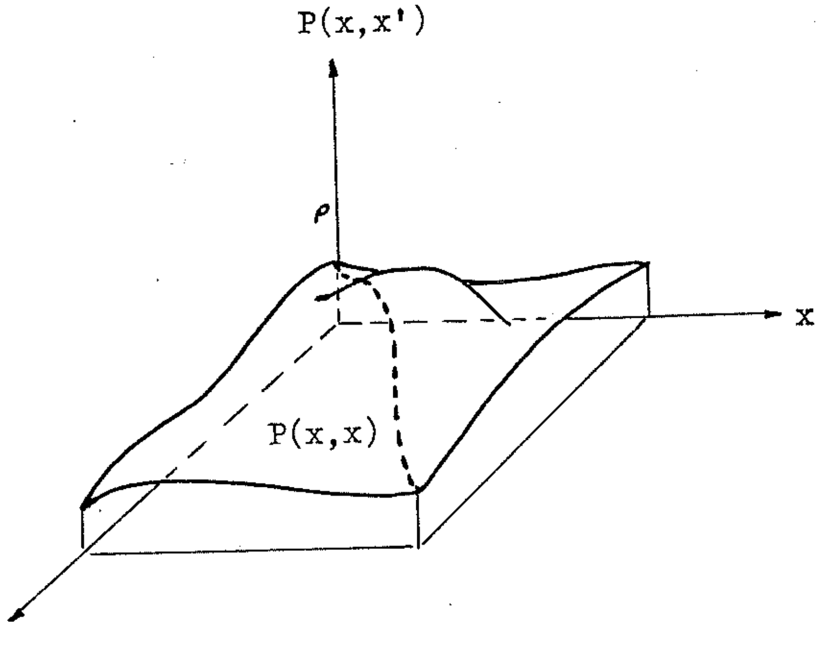
$$S_\Omega = \int_\Omega d\Omega \tag{22}$$

See Fig.1 for the one-dimensional single sensor case; notice that $P(x,x)$ is only the diagonal section of this surface.

To each position X_s in Ω_s there corresponds a steady-state solution of the Riccati equation. The set of all these steady-state solutions defines a curve $\mathscr{P}$ which is the locus of $V_\infty(X_s)\ \forall X_s \in \Omega_s$ (cf. Fig.2 for the one-dimensional single sensor case).

Therefore, there exists X_s^* such that

$$V_\infty(X_s^*) < V_\infty(X_s) \qquad \forall X_s \neq X_s^*, \quad X_s \in \Omega_s \tag{23}$$



Fig,1 Typical error covariance profile.

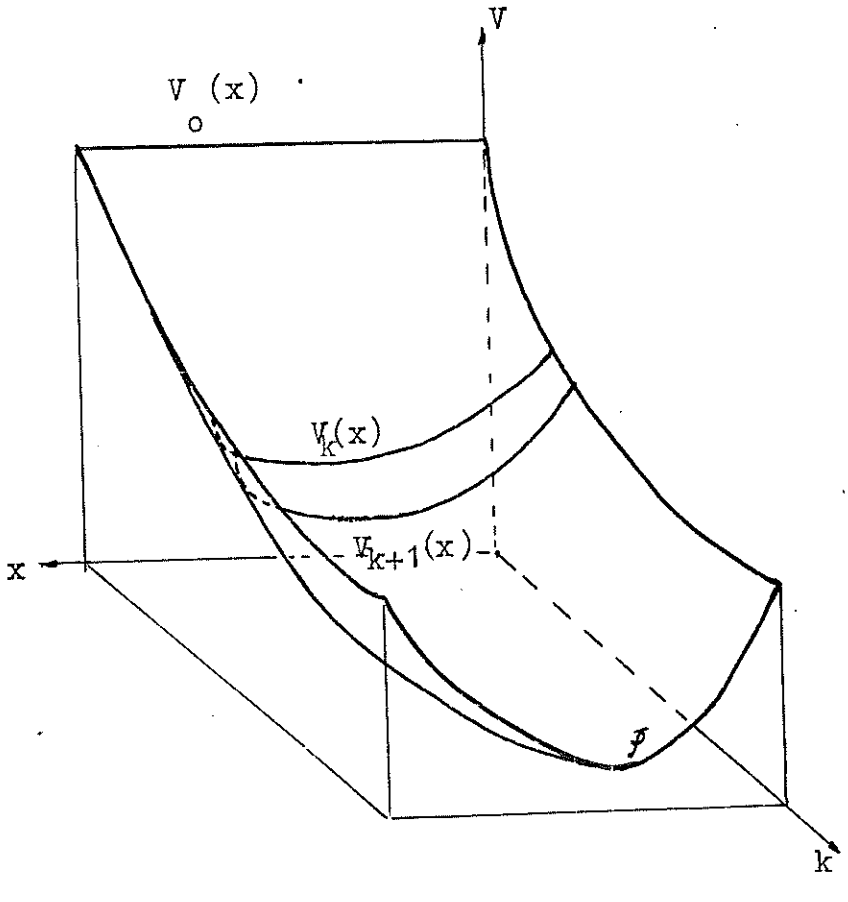


Fig.2 Steady-state solution for different values of X_s and fixed W_o.

Then

$$V^* = V_\infty(X_s^*) = \min_{X_s \in \Omega_s} V_\infty(X_s) \qquad (24)$$

and this completes the proof.

Remark: It has been shown that, under certain observability conditions, there exists an optimal set of locations that minimizes the trace of the spatial integral of the steady-state error covariance matrix. Notice that this optimal set is not neccessarily unique, namely $V_\infty(X_s)$ could have more than one equi-valued minimum.

CONVERGENCE OF THE ALLOCATION ALGORITHM

Nothing can be said about the convergence properties of conjugate gradient methods when applied to the minimization of a non convex function [3]. However empirical results justify its practicability in these cases. According to P. Wolfe [4] just what makes the method work well or poorly on a given problem is not well understood, even in the quadratic case, in the absence of round off error.

In this section, we will study some of the convergence properties of the aforementioned allocation algorithm when applied to time-invarient systems. In the allocation algorithm, at each stage, we minimize tr W_{k+1} w.r.t. $X_s(k)$,

$$\min_{X_s(k)} \left\{ \mathrm{tr}[W_{k+1}] \,/\, W_k \right\} \qquad (25)$$

Equeivalently,

$$\min_{X_s(k)} \left\{ \mathrm{tr}\,[E_k - E_k\, B_k^T(X_s)\, Q_k^{-1}(X_s)\, B_k(X_s)\, E_k] \,/E_k \right\} \qquad (26)$$

or

$$\max_{X_s(k)} \left\{ \mathrm{tr}\,[E_k\, B_k^T(X_s)\, Q_k^{-1}(X_s)\, B_k(X_s)\, E_k] \,/E_k \right\} \qquad (27)$$

Recall that the relation between E_k and W_k is given by the linear transformation (4).

Let the nonlinear operation on E_k in (3) be represented by

$$E_{k+1} = \mathcal{L}_k(E_k) \qquad (28)$$

It will be assumed that the allocation algorithm is applied right from the beginning.

By applying at each step the Riccati operator (28) and the minimization step (25) and by starting with E_o and $X_s(0)$, we can thereby define a sequence

$$E_k\{E_o, X_s(0), X_s^k\}, \quad k=1,2,\ldots \tag{29}$$

and two corresponding sequences

$$X_s^k, \quad k=1,2,\ldots \tag{30}$$

$$I_k\{E_o, X_s(0), X_s^k\}, \quad k=1,2,\ldots \tag{31}$$

where

$$I_k\{E_o, X_s(0), X_s^k\} = \operatorname{tr}\left[W_k\{W_o, X_s(0), X_s^k\}\right] \tag{32}$$

Assuming that the given optimal allocation problem does have a solution, the following two fundamental questions arise pertaining to convergence of the Modified Gradient Algorithm (MGA) used [1].

Let I_k be the cost after k iterations :

1. Does I_k converge ?
2. Assuming $I_k \rightarrow I_\infty$, is I_∞ the optimal cost and do the corresponding sequences of positions X_s^k, $k \rightarrow \infty$ converge to an optimal set of positions ? In particular, are I_∞ and $X_s(\infty)$ independent of the initial values E_o and $X_s(0)$?

If the answer to these questions is affirmative, then the sequence (31) defined by the allocation algorithm (25) and (28) converges to the optimum value, i.e.

$$\lim_{k \rightarrow \infty} I_k\{E_o, X_s(0), X_s^k\} = V^* \tag{33}$$

where V^* is defined by (24).

Before proceeding to discuss these equations, we shall need the following theorem,

<u>Theorem 2</u>:
Given the gradient algorithm (6), a set of positions $X_s(k)$ at time instant k, and the gradient,

$$g_k = \nabla_{X_s} J(X_s) \Big|_{X_s(k)} \tag{34}$$

If $g_k \neq 0$, then

$$J(X_s)\Big|_{X_s(k+1)} > J(X_s)\Big|_{X_s(k)} \qquad (35)$$

where $X_s(k+1)$ is the set of positions that results from the application of the gradient algorithm (6).

Proof:
The proof is similar to that given by Phillipson [5].

Assuming that $E_o > m\ I$, for some large scalar m, then for a fixed $X_s \in \Omega_s$ we can show that I_k is a monotone decreasing sequence that converges to a steady-state value $I_\infty(X_s)$. The existence of this steady-state solution is proved in the previous section, and the monotone behavior arises from the optimality of the Kalman-filter, i.e. using the Kalman-filter gains minimizes the error covariance (see theorem 2.1 in [2]).

In addition

$$I_k\{E_o, X_s(0), X_s^k\} \geqslant 0 \qquad \forall k \qquad (36)$$

since I_k is the spatial integral of covariance matrix.

Then if we fix $X_s(k) = X_s\ \forall k$, change E_o (or W_o) and integrate the Riccati equation, the resultant sequences will converge to the same steady-state value $I_\infty(X_s)$ as shown in Fig.3 (see the proof of theorem 1), of course this value is a function of X_s.

Changing X_s, the curves in Fig.3 will produce a family of surfaces similar to that depicted in Fig.2, and all of these surfaces converge to the same steady-state curve $\not{f}$. Such family of surfaces will be represented, for simplicity, by its projection over different cutting planes at each time instant as shown in Fig.4.

To understand geometrically what the algorithm does, let us assume that the point "a" which represents a value of $X_s(k)$ together with $I_k\{W_o', X_s'(0), X_s^k\}$. When applying the algorithm at this time instant we will get point "b" which represents a value of $X_s(k+1)$ together with $I_{k+1}\{W_o', X_s'(0), X_s^{k+1}\}$, i.e. we moved from point "a" on the first surface to point "b" on another surface as shown in Fig.4. What the algorithm gurantees is that the point "b" :

1. has a lower cost w.r.t. point "a", i.e.

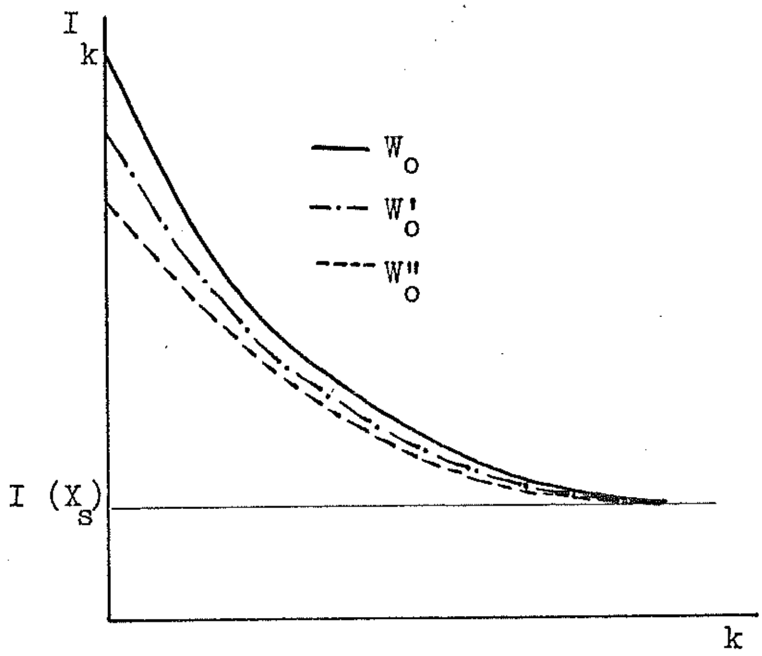


Fig.3 Convergence for different values of W_o and a fixed X_s.

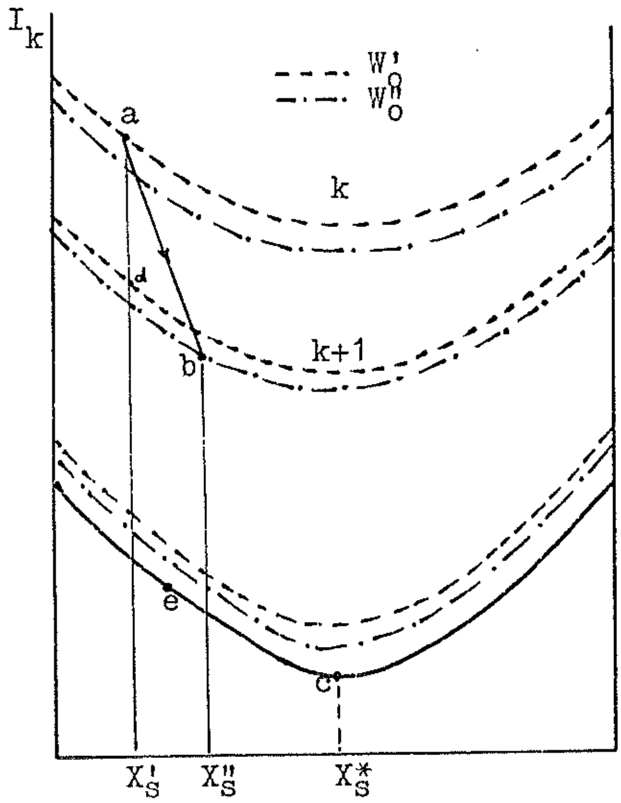


Fig.4 Surfaces resulting from changing W_o and X_s.

$$I_{k+1}\{W_o',X_s'(0),X_s^{k+1}\} \leq I_k\{W_o',X_s'(0),X_s^k\} \quad (37)$$

2. has a lower cost w.r.t. point "d" which results from iterating the Riccati equation once while keeping $X_s(k+1) = X_s(k)$, i.e.

$$I_{k+1}\{W_o',X_s'(0),X_s^{k+1}\} < I_{k+1}\{W_o',X_s'(0),X_s^k\} \quad (38)$$

The algorithm will work until we reach a point on the steady-state curve $\mathcal{S}$. What cannot be guaranteed is that the minimum "c" (see Fig.4) of the curve $\mathcal{S}$ will be reached. The difficulty results basically from the fact that the monotone behavior of the time-varying Riccati equation cannot be proved.

Indeed suppose that for a given initial E_o and $X_s(0)$, the algorithm has reached the steady-state point "e" on the curve $\mathcal{S}$ for $k = k^*$, to which there corresponds :

$$X_s(k^*),\ E_{k^*}\{E_o,X_s(0),X_s^{k^*}\} \quad \text{and} \quad I_{k^*}\{E_o,X_s(0),X_s^{k^*}\} \quad (39)$$

Even though "e" is not the optimal point, it cannot be proved that there exists $X_s(k^*+1)$ such that

$$I_{k^*+1} < I_{k^*} \quad (40)$$

because it cannot be proved that

$$E_{k^*+1} = \mathcal{L}_{k^*}(E_{k^*}) < E_{k^*} \quad (41)$$

In other words, along the section X_s' (see Fig.4) we have arrived at time instant k at a point with I_k corresponding to a value E_k. It is not sure that at the same time instant, there exists in section X_s'', point with I_k lying on a monotone decreasing surface (see Fig.2) and corresponding to a same matrix E_k of which I_k is the trace. There may not exist such a matrix along the section X_s''.

CONCLUSION

The existence of an optimal set of sensor locations and the convergence of the allocation algorithm to this optimal solution have been discussed. It has been shown that under certain observability conditions an optimal solution exists. The problems related to the convergence of the allocation algorithm have been examined.

References

1. S.E. Aidarous, M.R. Gevers and M.J. Installé, "Optimal sensors' allocation strategies for a class of stochastic distributed systems", Int. J. Control, Vol. 22, pp 197-213, 1975.

2. P.E. Caines and D.Q. Mayne, " On the discrete time matrix Riccati equation for optimal control", Int. J. Control, Vol. 12, pp 785-794, 1970.

3. E. Polak, "Computational methods in optimization: a unified approach", Academic Press, New York, 1971.

4. P. Wolfe, "Conjugate gradient methods in optimization", IEEE Conf. on Decision and Control, San Diego, Calif., 1973.

5. G.A. Phillipson, "Identification of distributed systems", American Elsevir, New York, 1971.