stichting
mathematisch
centrum

AFDELING MATHEMATISCHE BESLISKUNDE (DEPARTMENT OF OPERATIONS RESEARCH) BW 167/82 AUGUSTUS

J.H. VAN SCHUPPEN

ADAPTIVE STOCHASTIC FILTERING PROBLEMS THE CONTINUOUS TIME CASE

Preprint

kruislaan 413 1098 SJ amsterdam

Printed at the Mathematical Centre, 413 Kruislaan, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

1980 Mathematics subject classification: 93E11, 93E12, 93E03, 60G44

Adaptive stochastic filtering problems - the continuous time case*)

by

J.H. van Schuppen

ABSTRACT

The adaptive stochastic filtering problem for Gaussian processes is considered. The selftuning-synthesis procedure is used to derive two algorithms for this problem. Almost sure convergence for the parameter estimate and the filtering error will be established. The convergence analysis is based on an almost-supermartingale convergence lemma that allows a stochastic Lyapunov like approach.

KEY WORDS & PHRASES: *Adaptive stochastic filtering, selftuning synthesis procedure, almost sure convergence.*

---

*) This report will appear in the Proceedings of the 2nd Bad Honnef Workshop on Stochastic Differential Systems.

1. INTRODUCTION

The goal of this paper is to present some algorithms for a continuous-time adaptive stochastic filtering problem, and to establish almost sure convergence results for these algorithms.

What is the adaptive stochastic filtering problem? The adaptive stochastic filtering problem is to predict or filter a process when the parameter values of the dynamical system representing the process are not known. This problem is highly relevant for practical prediction and filtering problems in engineering and the social sciences.

Why should one consider the continuous-time version of this problem? In discrete-time the adaptive stochastic filtering problem has been investigated by many researchers. Hence the question why? Time is generally perceived to be continuous. In practice a continuous-time signal is sampled and the subsequent data processing is done in a discrete-time mode. One question then is what happens with the predictions when the sampling time gets smaller and smaller? To study these and related questions continuous-time algorithms must be derived, and their relationship with discrete-time algorithms investigated.

The questions that one would like to solve for the adaptive stochastic filtering problem are how to synthesize algorithms and how to evaluate the performance of these algorithms?

The selftuning synthesis procedure will be used in this paper. This procedure suggests first to solve the associated stochastic filtering problem, and secondly to estimate the values of the parameters of the filter system in an on-line fashion. A continuous-time parameter estimation algorithm is thus necessary. Although considerable attention has been given to off-line algorithms [1,2], on-line algorithms are scarce [3,7]. Below two new algorithms are presented.

As to the performance evaluation of the algorithms, the major question is the convergence of the parameter estimates and of the error in the filtering estimate. For these variables one may consider almost sure convergence and the asymptotic distribution. Below an almost sure convergence result for the given algorithms will be presented. This result is based on a convergence lemma that is of independent interest.

A brief outline of the paper follows. The problem formulation is given in section 2. Section 3 is devoted to the statement of the main results. The proofs of the results may be found elsewhere [14].

2. THE PROBLEM FORMULATION

The adaptive stochastic filtering problem is to predict or filter a stochastic process when the parameters of the distribution of this process are not known. The object of this section is to make this problem formulation precise.

Throughout this paper (Ω,F,P) denotes a complete probability space. Let $T = R$. The terminology of Dellacherie, Meyer [4,5] will be used.

Assume given an R-valued Gaussian process with stationary increments. From stochastic realization theory it is known [8] that under certain conditions this process has a minimal realization as the output of what will be called a *Gaussian system*:

$$dx_t = Ax_t dt + Bdv_t, \quad (1)$$

$$dy_t = Cx_t dt + dv_t, \quad (2)$$

where $y : \Omega \times T \to R$, $x : \Omega \times T \to R^n$, $v : \Omega \times T \to R^m$ is a standard Brownian motion process, $A \in R^{n\times n}$, $B \in R^{n\times m}$, $C \in R^{1\times n}$, $D \in R^{1\times m}$. The precise statement on the representation is that the distribution of the output y of this system is the same as that of the given process.

One may construct the asymptotic Kalman-Bucy filter for the system (1,2), say:

$$d\hat{x}_t = A\hat{x}_t dt + K(dy_t - C\hat{x}_t dt),$$

where $\hat{x}_t = E[x_t | F_t^y]$, $F_t^y = \sigma(\{y_s, \forall s \le t\})$.
This filter may be written as a Gaussian system

$$d\hat{x}_t = A\hat{x}_t dt + Kd\bar{v}_t, \quad (3)$$

$$dy_t = C\hat{x}_t dt + d\bar{v}_t, \quad (4)$$

where $\bar{v} : \Omega \times T \to R$ is the innovations process, which is a Brownian motion process, say with variance $\sigma^2.t$. It is a result of stochastic realization theory that the two relizations (1,2) and (3,4) are indistinguishable on the basis of information about y only. For adaptive stochastic filtering one may therefore limit attention to the realization (3,4). That realization has the additional advantage that it is suitable for prediction purposes. The fact that (1,2) is a minimal realization, and hence that (3,4) is a minimal realization, implies that the pair (A,C) is observable and that the spectrum of A is in $C^{-} := \{c \in C \mid Re(c) < 0\}$

2.1. PROBLEM. Assume given an R-valued Gaussian process with stationary increments having a minimal past-output based stochastic realization given by

$$d\hat{x}_t = A\hat{x}_t dt + K d\bar{v}_t, \tag{3'}$$

$$dy_t = C\hat{x}_t dt + d\bar{v}_t, \tag{4'}$$

$$\hat{z}_t := C\hat{x}_t, \tag{5}$$

with the properties mentioned above. Assume further thet the value of the dimension of this system and the value of σ^2 occurring in the variance of $\bar{v}$ are known, but that the values of the parameters A,K,C are not known. The *adaptive stochastic filtering problem* for the above defined Gaussian system is to recursively estimate $\hat{z}$ given y.

For the parameter estimation problem another representation of the Gaussian system (3,4) is needed. Such a representation is derived below. For notational convanience the time set is taken to be $T = R_+$ in the following.

2.2. PROPOSITION. *Given the Gaussian system* (3,4). *The two following representations describe the same relation between* $\bar{v}$ and $\hat{z}$.

a. $$d\hat{x}_t = A\hat{x}_t dt + K d\bar{v}_t,\ \hat{x}_0 = 0,$$

$$\hat{z}_t = C\hat{x}_t,$$

$$dy_t = C\hat{x}_t dt + d\bar{v}_t,\ y_0 = 0.$$

b. $$dh_t = Fh_t dt + G_1 dy_t + G_2(dy_t - h_t^T p dt),\ h_o = 0, \tag{6}$$

$$\hat{z}_t = h_t^T p, \tag{7}$$

$$dy_t = h_t^T p dt + d\bar{v}_t,\ y_0 = 0, \tag{8}$$

where $h : \Omega \times T \to R^{2n}$

$$h_t^T = (y_t^{(1)}, \ldots, y_t^{(n)}, \bar{v}_t^{(1)}, \ldots, \bar{v}_t^{(n)}), \tag{9}$$

$$y_t^{(1)} = y_t$$

$$y_t^{(i)} = \int_0^t y_s^{(i-1)} ds, \text{ for } i = 2,3,\ldots,n,$$

$p \in R^{2n}$ is related to A,K,C as indicated in the proof,

$$F_1 = \begin{pmatrix} 0 & \cdots & 0 \\ & & \vdots \\ I_{n-1} & & 0 \end{pmatrix} \in R^{n\times n}, \quad F = \begin{pmatrix} F_1 & 0 \\ 0 & F_1 \end{pmatrix} \in R^{2n\times 2n},$$

$$G_1 = e_1 \in R^{2n},\ G_2 = e_{n+1} \in R^{2n},$$

where e_i *is the i-th unit vector.*

The proof of this result is given in [14].

3. THE MAIN RESULTS

In this section two algorithms are presented for the continuous-time adaptive stochastic filtering problem, and convergence results are provided. The proofs of these results may be found in [14].

In the following attention is restricted from the Gaussian system (3,4) or (6,8), to the auto regressive case described by

$$y_t = \sum_{i=1}^{n} a_i y_t^{(i+1)} + \bar{v}_t, \tag{10}$$

or

$$dy_t = h_t^T p\, dt + d\bar{v}_t,\ y_0 = 0, \tag{11}$$

where now $h : \Omega \times T \to R^n$, $p \in R^n$,

$$h_t^T = (y_t^{(1)},\dots,y_t^{(n)}) \tag{12}$$

$$p^T = (a_1,\dots,a_n).$$

Then

$$dh_t = \begin{pmatrix} a_1 \cdots & a_n \\ & 0 \\ I_{n-1} & \vdots \\ & 0 \end{pmatrix} h_t dt + \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} d\bar{v}_t,\ h_0 = 0, \tag{13}$$

and one concludes that asymptotically h is a stationary Gauss-Markov process. Since the interest here is in the stationary situtation, it will henceforth be assumed that h is a stationary Gauss-Markov process. By the stability of the Gaussian system the covariance function of h is integrable, thus h is an ergodic process [15, p. 69].

The first algorithm to be presented is based on the least-squares method.

3.1.DEFINITION. The adaptive stochastic filtering algorithm for the autoregressive representation based in the least-squares method is defined to be

$$d\hat{p}_t = Q_t h_t\, \sigma^{-2}[dy_t - h_t^T \hat{p}_t dt]\ p_0 = 0,$$

$$dQ_t = -Q_t h_t h_t^T Q_t \sigma^{-2} dt,\ Q_0 \in R^{n\times n},\ Q_0 = Q_0^T > 0,$$

$$\hat{\hat{z}}_t = h_t^T \hat{p}_t,$$

where $\hat{p} : \Omega \times T \to R^n$, $Q : \Omega \times T \to R^{n \times n}$, $\hat{\hat{z}} : \Omega \times T \to R$.
Here $\hat{\hat{z}}$ is the desired estimate of $\hat{z}$, and $\hat{p}$ is an estimate of the parameter p.

The algorithm of 3.1 may be derived as follows. One has the representation

$$dp_t = 0, \; p_0 = p$$
$$dy_t = h_t^T p_t dt + d\bar{v}t, \; y_0 = 0,$$

where it is now assumed that $\bar{v}$ is a Brownian motion process, $p : \Omega \times T \to R^n$, p is Gaussian $G(0,Q_0)$ and that p and $\bar{v}$ are independent objects. With (12) one concludes that $(h_t, F_t^y, t \in T)$ is adapted. The conditional Kalman-Bucy filter [9, 12.1] applied to the above representation then yields the algorithm given in 3.1.

3.2. THEOREM. *Given the stochastic dynamic system of* 2.1 *restricted to the autoregressive case as indicated above. Consider the adaptive stochastic filtering algorithm of* 3.1. *Under these conditions:*

a. $$\text{as} - \lim_{t\to\infty} \hat{p}_t = p;$$

b. $$\text{as} - \lim_{t\to\infty} t^{-1} \int_0^t (\hat{z}_s - \hat{\hat{z}}_s)^2 ds = 0.$$

The above result shows that under the stated conditions the difference between the filter estimate $\hat{z}$ with known parameters and the adaptive filter estimate goes to zero in the above stated sense. In addition the parameter estimate converges to the actual value.

One might conjecture that a result like 3.2 also holds if the restriction to the autoregressive case is relaxed to that of (6,8) and a recursive extended least-squares algorithm is used. An investigation indicates that this is unlikely. The reason why this is the case is not yet fully understood.

The second algorithm is related to that of Goodwin, Ramadge, and Caines [6], and that of Chen [3]. The latter also provides a continuous-time algorithm not only for the autoregressive case but also for the general case of 2.1.

3.3. DEFINITION. The adaptive stochastic filtering algorithm for the autoregressive representation (11, 12) based on the parameter estimation algorithm AML2 [6] is defined to be

$$d\hat{p}_t = h_t r_t^{-1} \sigma^{-2} [dy_t - h_t^T \hat{p}_t dt], \; \hat{p}_0 = 0,$$
$$dr_t = \sigma^{-2} h_t^T h_t dt, \; r_0 = 1,$$
$$\hat{\hat{z}}_t = h_t^T \hat{p}_t,$$

where $\hat{p} : \Omega \times T \to R^n$, $r : \Omega \times T \to R$, $\hat{\hat{z}} : \Omega \times T \to R$ and h is given in (12).

3.4. THEOREM. *Given the stochastic dynamic systems of 2.1. restricted to the autoregressive case as indicated above. Consider the adaptive stochastic filtering algorithm 3.3. Under these conditions*

$$\text{as} - \lim_{t\to\infty} t^{-1} \int_0^t (\hat{z}_s - \hat{\hat{z}}_s)^2 ds = 0.$$

In [3] a convergence result is given for the representation 2.1 with an algorithm that has the same structure as that of 3.3. There the convergence is obtained under an unnatural assumption. One possible reason for this assumption is that the second innovation process

$$d\bar{\bar{v}}_t = dy_t - \hat{h}_t^T \hat{p}_t dt$$

is directly used in $\hat{h}$ and not prefiltered.

The above convergence results for adaptive stochastic filtering are based on a convergence theorem to be provided below. As some of the other concepts and results of system identification the convergence theorem is also inspired by the statistics literature, in particular by the area of stochastic approximation. H. Robbins and D. Siegmund [11] have established a discrete-time convergence result for use in stochastic approximation theory. A simplified version of that result is given as an exercise in [10, II-4]. V. Solo [12, 13] has been the first to use this result in the system identification literature and since then it has become rather popular [6]. This popularity is due not only to the ease with which convergence results are proven but also to the formulation in terms of martingales which show up naturally in stochastic filtering problems. Below the continuous-time analog of [11,th.1.] is given.

A few words about notation follow. $(F_t, t \in T)$ denotes a σ-algebra family satisfying the usual conditions, A^+ is the set of increasing processes, M_{1uloc} the set of locally uniformly integrable martingales, and $\Delta x_t = x_t - x_{t-}$ is the jump of the process x at time $t \in T$.

3.5. THEOREM. *Let* $x : \Omega \times T \to R_+$, $a : \Omega \times T \to R_+$, $b : \Omega \times T \to R_+$, $e : \Omega \times T \to R_+$ *and* $m : \Omega \times T \to R$ *be stochastic processes. Assume that*

1. $x_0 : \Omega \to R_+$ *is* F_0 *measurable;*
2. $(a_t, F_t, t \in T) \in A^+$, $a_0 = 0$, $a_\infty < \infty$ a.s., *and for all* $t \in T$ $\Delta a_t \le c_1 \in R_+$, $(b_t, F_t, t \in T) \in A^+$, $b_0 = 0$;
3. $(e_t, F_t, t \in T)$ *is adapted and* $\int_0^\infty e_s ds < \infty$ a.s.;
4. $(m_t, F_t, t \in T) \in M_{1uloc}$, $m_0 = 0$;
5. x *satisfies the stochastic differential equation*

$$dx_t = e_t x_t dt + da_t - db_t + dm_t,\ x_0.$$

Then

a. $x_\infty := \text{as-}\lim_{t\to\infty} x_t$ *exists in* R_+, *hence* $x_\infty < \infty$ a.s.;

b. $b_\infty := \text{as-}\lim_{t\to\infty} b_t$ *exists and* $b_\infty < \infty$ a.s.

REFERENCES

1. A. Bagchi, Consistent estimates of parameters in continuous time systems, in: Analysis and Optimization of Stochastic Systems, O.L.R. Jacobs, M.H.A. Davis, M.A.H. Dempster, C.J. Harris, P.C. Parks., Academic Press, London, 1980, pp. 437-450.
2. A.V. Balakrishnan, Stochastic differential systems I, Filtering and Control, A function space approach, Lecture Notes in Economics and Mathematical Systems, volume 84, Springer-Verlag, Berlin, 1973.
3. Chen Han Fu, Quasi-least-squares identification and its strong consistency, Int. J. Control, 34 (1981), pp. 921-936.
4. C. Dellacherie, P.A. Meyer, Probabilités et Potentiel, Chapitres I à IV, Hermann, Paris, 1975.
5. C. Dellacherie, P.A. Meyer, Probabilités et Potentiel, Chapitres V à VIII, Theorie des Martingales, Hermann, Paris, 1980.
6. G.C. Goodwin, P.J. Ramadge, P.E. Caines, A globally convergent adaptive predictor, Automatica, 17 (1981), pp. 135-140.
7. Y.D. Landau, Adaptive control, Marcel Dekker, New York, 1979.
8. A. Lindquist, G. Picci, On the stochastic realization problem, SIAM J. Control and Optim., 17 (1969), pp. 365-389.
9. R.S. Liptser, A.N. Shiryayev, Statistics of random processes, I General theory, II Applications, Springer-Verlag, Berlin, 1977,1978.
10. J. Neveu, Martingales à temps discrets, Masson et Cie, Paris, 1972; english translation: Discrete parameter martingales, North-Holland, Amsterdam, 1975.
11. H. Robbins, D. Siegmund, A convergence theorem for non negative almost supermartingales and some applications, in 'Optimizing methods in statistics', J.S. Rustagi ed., Academic Press , New York, 1971, pp. 233-256.
12. V. Solo, The convergence of AML, IEEE Transactions on Automatic Control, 24 (1979), pp. 958-962.
13. V. Solo, The convergence of an instrumental-variable-like recursion, Automatica, 17 (1981), pp. 545-547.
14. J.H. van Schuppen, Convergence results for continuous-time adaptive stochastic filtering algorithms, preprint, 1982.
15. E. Wong, Stochastic processes in information and dynamical systems, McGraw-Hill, New York, 1971.