Proceedings of 23rd Conference
on Decision and Control
Las Vegas, NV, December 1984

AN EFFICIENT FILTER FOR ABRUPTLY CHANGING SYSTEMS

H.A.P. Blom

National Aerospace Laboratory NLR
P.O. Box 90502, 1006 BM Amsterdam
The Netherlands

Abstract

For a linear discrete time system with Markovian coefficients a new filtering algorithm is given, which is called the Interacting Multiple Model (IMM) algorithm. The mathematical support for this algorithm is outlined and a qualitative comparison with other known filtering algorithms is made. The main conclusion is that the ratio between performance and computational complexity is far better for the IMM algorithm.

1 Introduction

An important practical problem is filtering for a linear system with Markovian coefficients. Therefore this problem has had considerable attention in the past, mostly in a discrete time setting. By now there are some families of discrete time filtering algorithms to come arbitrary close to MMSE performance (Ref. 1). Despite of all those algorithms the situation is not satisfactory. The computational complexity of the algorithms that come close to MMSE performance is often too high for practical application. At the same time from a mathematical point of view some important families of those algorithms become less relevant when the time-lag between measurement updates becomes smaller and smaller. These two considerations formed the reason for a further study of the problem. During the first phase of the study the attention has been directed to the continuous time problem. The exploitation of the powerful tools for stochastic differential equations (Refs. 2-7) yielded additional insight into filtering for systems with Markovian coefficients (Ref. 8). During the second phase of the study this insight formed the guide to obtain a new discrete time filtering algorithm. This new algorithm is called the IMM (Interacting Multiple Model) algorithm and is defined in sections 2 and 3. Subsequently in section 4 a qualitative comparison is made between the IMM algorithm and the main other algorithms.

2 The problem

The problem considered is filtering an increasing information field Y_t of real valued observations on the discrete time line t,

$$y_t = h(c_t)x_t+g(c_t)w_t \tag{1}$$

where c is a discrete time finite Markov chain assuming values in $M=\{m_1,..,m_N\}$ according to a deterministic transition probability matrix A, w is discrete time standard white Gaussian noise and x is a real valued process which satisfies the stochastic difference equation,

$$x_t = a(c_t)x_{t-1}+b(c_t)v_t \tag{2}$$

where v is discrete time standard white Gaussian noise. Further w and v are mutually independent, h, g^2, a and b^2 are finite real functions on M and for simplicity time-invariant. Also restriction to scalar x and y is for the sake of notational simplicity only.

Before going to the IMM algorithm take a closer look at this filtering problem. First of all verify that a sufficient statistic for the solution of the filtering problem is a set of closed form equations for the evolution of the conditional probability measure $P^{Y_t}\{x_t \in X, c_t=m\}$ on $R \times M$. Such closed form equations are respectively a forward Chapman-Kolmogorov equation for the time extrapolation and Bayes rule for the measurement update equation. As a probability measure $P^B\{x_t \in X, c_t=m\}$ is on a continuous-discrete space $R \times M$ both the time extrapolation equations and the measurement update equations can be decomposed in equations for a measure $P^B\{c_t=m\}$ on the discrete space M and for all $m \in M$ in equations for a measure $P^B\{x_t \in X|c_t=m\}$ on the continuous subspace $R \times M$.

After this decomposition the remaining problem is to approximate each of the measures on the N continuous subspaces sufficiently accurate by a measure that can be described by a finite number of parameters, and to derive the time extrapolation equations and the measurement update equations for these parameters. The measures, known for such an approximation, range from a Gaussian measure up to an arbitrary accurate finite element constructed measure. In this paper we evaluate the filter algorithm which assumes a Gaussian measure on each of the N continuous subspaces $R \times M$.

3 The IMM algorithm

Definition *The IMM (Interacting Multiple Model) algorithm for the filtering problem given by (1) and (2) consists of time extrapolation equations between t-1 and t and measurement update equations on moment t. These equations are for the weights* $p(m)$, *the means* $\hat{x}(m)$ *and the variances* $R(m)$ *for all* $m \in M$.
The time extrapolation equations between the measurement moments $t-1$ *and* t *are for all* $m \in M$:

i. *The jump extrapolation equations*

$$p_+(m) = \sum_{n\in M} A_{nm}p(n) \quad , \tag{3}$$

$$\hat{x}_+(m) = \sum_{n\in M} A_{nm}p(n)\hat{x}(n)/p_+(m) \quad \text{and} \tag{4}$$

$$R_+(m) = \sum_{n\in M} A_{nm}p(n)(R(n)+(\hat{x}(n)-\hat{x}_+(m))^2)/p_+(m) \tag{5}$$

under the condition that $p_+(m)>0$.

ii. *The diffusion extrapolation equations*

$$\hat{x}_+(m) := a(m)\hat{x}_+(m) \quad \text{and} \tag{6}$$

$$R_+(m) := a^2(m)R_+(m)+b^2(m) \ . \tag{7}$$

The equations to process a measurement y *at moment* t *are for all* $m \in M$:

iii. *The measurement update equations*

$$\hat{x}(m) = \hat{x}_+(m) + K(m)\nu(m) \quad , \tag{8}$$

$$R(m) = R_+(m) - K(m)h(m)R_+(m) \quad , \tag{9}$$

with: $$Q(m) = h^2(m)R_+(m)+g^2(m) \tag{10}$$

$$\nu(m) = y-h(m)\hat{x}_+(m) \tag{11}$$

$$K(m) = h(m)R_+(m)/Q(m) \quad , \text{ and} \tag{12}$$

$$p(m) = Cp_+(m)\exp(-\tfrac{1}{2}\nu^2(m)/Q(m))/Q^{\frac{1}{2}}(m) \tag{13}$$

with C *independent of* m *such that* $\sum_{m\in M} p(m)=1$.

After step iii at moment t *the estimates* $p(m), \hat{x}(m)$ *and* $R(m)$ *for all* $m \in M$ *represent an approximation* $\bar{P}^{Y_t}\{x_t \in X, c_t=m\}$ *of the conditional probability measure*

$P^{Y_t}\{x_t\in X, c_t=m\}$ *in the following way:*

$$\frac{\delta}{\delta x}\bar{P}^{Y_t}\{x_t<x, c_t=m\} = p(m).N\{\hat{x}(m),R(m)\}, \qquad (14)$$

with N a Gaussian.

Of the equations given in the definition above (3) and (13) are well known for a Markov chain, (6) through (12) are exactly the Kalman equations for (1) and (2) with deterministic c, while (4) and (5) are unusual. A schematic representation of these equations is given in the figure below.

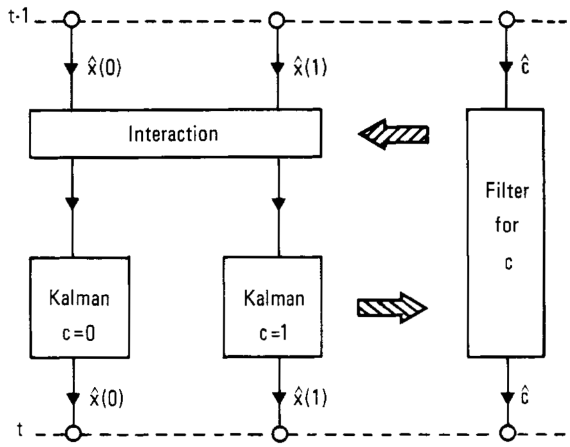

Fig. 1 The IMM algorithm for M={0,1}

Obviously when A=I equations (3), (4) and (5) can be deleted, and the IMM algorithm reduces to the well known MM (Multiple Model) algorithm.

With some ad hoc modifications of the filter for c, the MM algorithm has also been proposed for the case A≠I (Ref. 9). Contrary to the IMM algorithm this modified MM algorithm does not reflect the jumping property of c by interaction between the Kalman filters in the bank. But (4) and (5), do they adequately reflect the influence of the jumping property of c on the evolution of x? A positive answer to that question is given by the following theorem.

<u>Theorem</u> *The unconditional evolution of the first and second central moments of x_{t-1} given $c_{t-1}=m$ to those of x_t given $c_t=m$ satisfies (4) through (7).*

The proof of the theorem is given in the appendix.

4 Comparison with other algorithms

For the filtering of (1) and (2) several methods to approximate the MMSE filter have been developed during the last fifteen years. Inherent to the problem for these methods there is a trade-off between performance and complexity. Some recent publications (Refs. 1, 10 and 11) give overviews and comparisons of these methods. The common views in these publications are:

- The general structure of all methods is a bank of Kalman-like filters for x, cooperating with a filter for the Markov chain c.
- The methods can be classified in decision directed and merging directed methods. The latter are often referred to as Bayesian.
- On the average a merging directed method performs slightly better than a decision directed method of the same complexity.
- Of the merging directed methods the most important ones are the so-called Generalized Pseudo Bayes (GPB) algorithms.

On this basis the GPB algorithms are the main concurrents of the IMM algorithm. Therefore we will make a qualitative comparison between the IMM algorithm and the GPB algorithms of comparable complexity or performance.

A GPB algorithm of order n (memory length n-1) has N^n Kalman filters in its bank (for n=2 see figure 2). Complete descriptions of the 1st and 2nd order GPB's are respectively given by Athans e.a. (Ref. 12) and Chang and Athans (Ref. 13), with the comment of Tugnait (Ref. 14). The computational complexity of the IMM algorithm is comparable to that of the GPB algorithm of

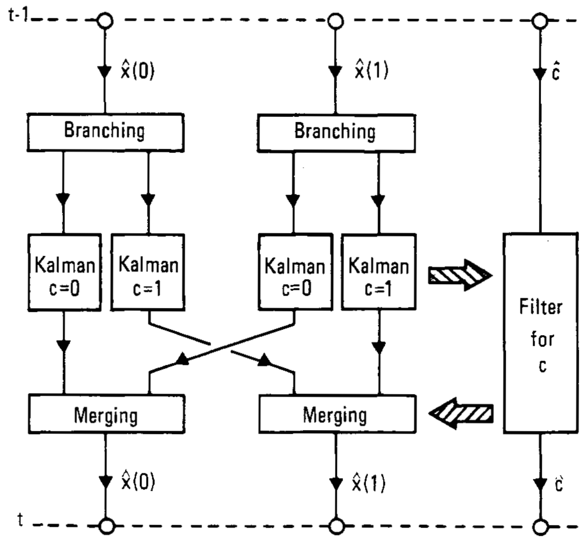

Fig. 2 The GPB algorithm of second order for M={0,1}

first order and about N times lower than that of the GPB algorithm of second order. Nevertheless the performance of the IMM algorithm is not far from that of the GPB algorithm of second order. This can easily be verified by making the time-lag between measurements smaller and smaller. When the time-lag approaches zero both the IMM algorithm and the GPB algorithm of second order converge to the continuous time filter that is derived in ref. 8. The performance of the GPB algorithm of first order is significantly below that of the GPB algorithm of second order (Ref. 1,10). Also when the time-lag approaches zero the GPB-algorithm of first order does not converge to the continuous time filter of ref. 8.

From these qualitative arguments it can be concluded that the IMM algorithm is up to N times more efficient than two of the most efficient other algorithms; the GPB algorithms of first and second order.

5 Final remarks

For a discrete time system with coefficients that are governed by an N-state Markov chain a new filtering algorithm has been defined. It consists of a bank of N interacting Kalman-like filters which communicate with a filter for the Markov chain. This so-called IMM (Interacting Multiple Model) algorithm has a sound basis of which the strongest part is stated by the theorem in section 3. When the transition probabilities of the Markov chain are zero the IMM algorithm reduces to the well-known MM algorithm.

A qualitative comparison with the best other algorithms showed an improvement of the ratio between performance and computational complexity up to a factor N. More specifically the performance of the IMM algorithm reaches for small time lags that of an N times more complex GPB (Generalized Pseudo Bayes) algorithm.

The study of the above filtering problem was motivated by the well-known problem of tracking aircraft by a surveillance radar for Air Traffic Control. There the sudden starting, resp. stopping, of aircraft accelerations forms one of the main tracking subproblems. Obviously for each of these subproblems we need a filtering method, which can be combined with the methods for the other subproblems. Being a Bayesian method the IMM algorithm satisfies this need. Together with Bayesian methods for the other tracking subproblems the IMM algorithm has been exploited to realize a real-time tracking algorithm which outperforms state-of-the-art tracking algorithms (Ref. 15). For completeness we have to mention that to reach this we needed an additional result: the extension of the system (1) and (2) to include sudden additive jumps. The IMM algorithm for such an extended system will be presented in a future publication.

APPENDIX Proof of the theorem

Assume at moment t-1 the probability measure on RxM of the pair $\{x,c\}$ is $P^B\{x_{t-1} \in X, c_{t-1}=m\}$ with B an arbitrary information field. Verify that $\{x,c\}$ is a Markov process which satisfies the Chapman-Kolmogorov equation,

$$P^B\{x_t \in X, c_t=m\} = \sum_{n \in M} \int_R p((u,n),(X,m)) P^B\{x_{t-1} \in du, c_{t-1}=n\} \quad (15)$$

with Markov kernal,

$$p((u,n),(X,m)) = P\{x_t \in X, c_t=m \mid x_{t-1}=u, c_t=n\} \quad (16)$$

With this the first step is evaluating the kernal, by using that x satisfies (2) and c is a Markov chain with transition matrix A,

$$\begin{aligned} p((u,n),(X,m)) &= P\{x_t \in X \mid c_t=m, x_{t-1}=u, c_{t-1}=n\} . \\ &\quad . P\{c_t=m \mid x_{t-1}=u, c_{t-1}=n\} = \\ &= F(X,u,m)\, A_{nm}, \end{aligned} \quad (17)$$

where $F(.,u,m)$ is for all $(u,m) \in RxM$ a Gaussian probability measure on R,

$$F(X,u,m) = \int_X (\sqrt{2\pi}\, b(m))^{-1} \exp\{-\tfrac{1}{2}(a(m)u-x)^2/b^2(m)\} .dx \quad (18)$$

The subsequent step is to introduce the decomposition,

$$P^B\{x_t \in X, c_t=m\} = P^B\{x_t \in X \mid c_t=m\} \, . \, P^B\{c_t=m\}. \quad (19)$$

Substitution of (17) and (19) in (15) yields for $P^B\{c_t=m\}>0$:

$$\begin{aligned} P^B\{x_t \in X \mid c_t=m\} &= \int_R F(X,u,m) \sum_{n \in M} A_{nm} P^B\{c_{t-1}=n\} . \\ &\quad . P^B\{x_{t-1} \in du \mid c_{t-1}=n\} / P^B\{c_t=m\} \end{aligned} \quad (20)$$

Finally by defining,

$p(m)=P^{Y_{t-1}}\{c_{t-1}=m\}$, $p_+(m)=P^{Y_{t-1}}\{c_t=m\}$,

$\hat{x}(m)=E^{Y_{t-1}}\{x_{t-1} \mid c_{t-1}=m\}$, $\hat{x}_+(m)=E^{Y_{t-1}}\{x_t \mid c_t=m\}$,

$R(m)=E^{Y_{t-1}}\{(x_{t-1}-\hat{x}(m))^2 \mid c_{t-1}=m\}$ and

$R_+(m)=E^{Y_{t-1}}\{(x_t-\hat{x}_+(m))^2 \mid c_t=m\}$,

the relations between $(\hat{x}(m),R(m))$ and $(\hat{x}_+(m),R_+(m))$, such as given by (4) through (7), follow easily from (18) and (20).

Q.E.D.

References

1. Tugnait J.K., "Detection and estimation for abruptly changing systems", Automatica, 18 (1982) pp. 607-615.
2. Gihman I.I., Skorohod A.V., "Stochastic differential equations", Springer, Berlin, 1972.
3. Brockett R.W., Blankenship G.L., "A representation theorem for linear differential equations with Markovian coefficients", Proc. 1977, Allerton Conf. on Circuits and Systems Theory, Urbana, pp. 671-679.
4. Lipster R.S., Shiryayev A.N., "Statistics of random processes", Parts I and II, Springer, 1977, 1978.
5. Kallianpur G., "Stochastic filtering theory", Springer, 1980.
6. Hazewinkel M., Willems J.C. (eds.),"Stochastic systems: The mathematics of filtering and identification and applications", Reidel, 1981.
7. Baras J.S., Blankenship G.L., Hopkins (Jr.)W.E., "Existence, uniqueness and asymptotic behaviour of solutions to a class of Zakai equations with unbounded coefficients", IEEE Tr. on Automatic Control, 28 (1983), pp. 203-214.
8. Blom H.A.P., "Markov jump-diffusion models and decision-making-free filtering", Proceedings of the 6th Int. Conf. Analysis and Optimization of Systems, June 1984, Eds: Bensoussan A., Lions J.L., Springer Berlin.
9. Willsky A.S., Chow E.Y., Gershwin S.B., Greene C.S., Houpt P.K., Kurkjian A.L.,"Dynamic model-based techniques for the detection of incidents on freeways", IEEE Tr. on Automatic Control, 25 (1980), pp. 347-360.
10. Smith A.F.M., Makov U.E., "Bayesian detection and estimation of jumps in linear systems", Eds: Jacobs O.L.R., e.a. "Analysis and optimization of stochastic systems", Academic Press, London, 1980, pp. 333-345.
11. Weiss J.L., Upadhyay T.N., Tenney R., "Finite computable filters for linear systems subject to time-varying model uncertainty", Proc. of the NAECON, 1983, pp. 349-355.
12. Athans M., Whiting R.H., Gruber M.,"A suboptimal estimation algorithm with probabilistic editing for false measurements with applications to target tracking with wake phenomena", IEEE Tr. on Automatic Control, 22 (1977), pp. 372-384.
13. Chang C.B., Athans M.,"State estimation for discrete systems with switching parameters", IEEE Tr. on Aerospace and Electronic Systems, 14 (1978), pp. 418-425.
14. Tugnait J.K., "Comments on ... (Ref. 13)", IEEE Tr. on Aerospace and Electronic Systems, 14 (1978), pp. 418-425.
15. Blom H.A.P., "A sophisticated tracking algorithm for Air Traffic Control surveillance radar data", Proc. of the Int. Conf. on Radar, Paris, May 1984, pp. 393-398.