Sixth International Conference of Computational Methods in Sciences and Engineering, September, 25-30, Hersonissos, Crete, Grece, 2008.

Exact Filtering In Semi-Markov Jumping System

Noufel Abbassi and Wojciech Pieczynski

Institut TELECOM; TELECOM et Management SudParis; Dépt. CITI; CNRS UMR 5157
9, rue Charles Fourier, 91000 Evry, France
phone 33 1 60 76 44 25, fax 33 1 60 76 44 33, e-mail Wojciech.Pieczynski@it-sudparis.eu

Abstract. The classical hidden linear Gaussian system allows one to use the classical Kalman filter, which calculates some distributions of interest with linear complexity in number of observations. However, such calculations become impossible when adding a Markov jump process. The aim of the paper is to propose two new hidden models with Markov and semi-Markov jump processes in which the exact computation of the Kalman filter is feasible with linear complexity in number of observations.

Keywords: Hidden Markov models, triplet Markov models, Bayesian segmentation, Kalman filtering, Markov jumps, semi-Markov jumps.
PACS: 02.50.Ga, 02.50.Ey

INTRODUCTION

Hidden Markov Chains (HMC) are widely used in different areas and hundred papers are published on the subject each year. Their success is due to the fact that they make possible restoring unobservable random sequences $x = (x_1, ..., x_n)$ from observed random sequences $y = (y_1, ..., y_n)$ for very large n : one million, or more. The classical HMC distribution of the couple $Z = (X, Y)$ is given by

$$p(z) = p(x_1)p(x_2|x_1)\dots p(x_n|x_{n-1})p(y_1|x_1)\dots p(y_n|x_n) \tag{1}$$

Triplet Markov Chains (TMC) model proposed in [10] consists of introducing a third chain $R = (R_1, ..., R_n)$ and of assuming that the triplet $T = (X, R, Y)$ is a Markov chain. Such a model is much more general than HMC and yet allows one to search $x = (x_1, ..., x_n)$ from $x = (x_1, ..., x_n)$ [11, 12]. Let us assume Y continuous: each Y_i takes its values in the set of real numbers R . The hidden chain X can then be either discrete or continuous. When it is discrete, the chain R , taken also discrete, can model different properties of interest like the non stationarity of X [8] or its semi-Markovianity [9]. When it is continuous, one takes continuous R and different processing, like Kalman filtering, can be extended to $T = (X, R, Y)$ [1].

In this paper we propose a new TMC in which both X and Y are continuous, while R is discrete. We are interested on filtering X from Y , and R models the "jumps" of the parameters of the system. We show that the new model is workable and that the Bayesian optimal filter based on it can be computed with linear complexity in number of observations.

TRIPLET MIXED MARKOV CHAIN

Let $X = (X_1, ..., X_n)$, $Y = (Y_1, ..., Y_n)$ be two stochastic chains with each X_i and Y_i taking their values in the set of real numbers R , and let $R = (R_1, ..., R_n)$ be a stochastic chain with each R_i taking its values in a finite set $\Delta = \{1, ..., m\}$. The first TMC we propose is defined by the three following conditions:

(i) R is a Markov chain;

(ii) $p(y|r) = \prod_{i=1}^{n} p(y_i|r_i)$;

(iii) $p(x|r) = \prod_{i=1}^{n} p(x_i|r_i)$.

Its distribution is then written

$$p(t) = p(r_1)p(r_2|r_1)...p(r_n|r_{n-1})p(y_1|r_1)...p(y_n|r_n)p(x_1|r_1)...p(x_n|r_n) \qquad (2)$$

It resembles to the distribution (1); however, it is very different from this. In fact, $R = (R_1,...,R_n)$ is a latent chain and the distribution to be compared to (1) is the distribution of $Z = (X,Y)$, which is a marginal distribution of (2). Now, following an analogous proof that the proof in [12], one could show that the distribution of $Z = (X,Y)$ is not even a Markov distribution. However, in spite of this, we are going to see that the Kalman-like filtering is feasible. In fact, we have

$$p(x_i|y_1^i) = \sum_{r_i \in \Delta} p(x_i|r_i)p(r_i|y_1^i) \qquad (3)$$

Thus $p(x_i|y_1,...,y_i)$ is computable once $p(u_i|y_1,...,y_i)$ is. Now, according to (i) and (ii), the couple (U,Y) is the classical HMC. Then we have the following formulas, which can be seen as the Kalman filter's formulas applied in the discrete case :

$$p(r_{i+1}|y_1^i) = \sum_{r_i \in \Delta} p(r_{i+1}|r_i)p(r_i|y_1^i) \qquad (4)$$

$$p(r_{i+1}|y_1^{i+1}) = \frac{p(y_{i+1}|r_{i+1})p(r_{i+1}|y_1^i)}{\sum_{r_{i+1} \in \Delta} p(y_{i+1}|r_{i+1})p(r_{i+1}|y_1^i)} \qquad (5)$$

Finally, $p(r_{i+1}|y_1^{i+1})$ are computed from $p(r_i|y_1^i)$ using (4) and (5). Having $p(r_{i+1}|y_1^{i+1})$ immediately gives $p(x_{i+1}|y_1^{i+1})$ with (3).

Thus we can state

Proposition 1

Let (X,R,Y) be a TMC verifying

(i) R is a Markov chain;

(ii) $p(y|r) = \prod_{i=1}^{n} p(y_i|r_i)$; and

(iii) $p(x|r) = \prod_{i=1}^{n} p(x_i|r_i)$.

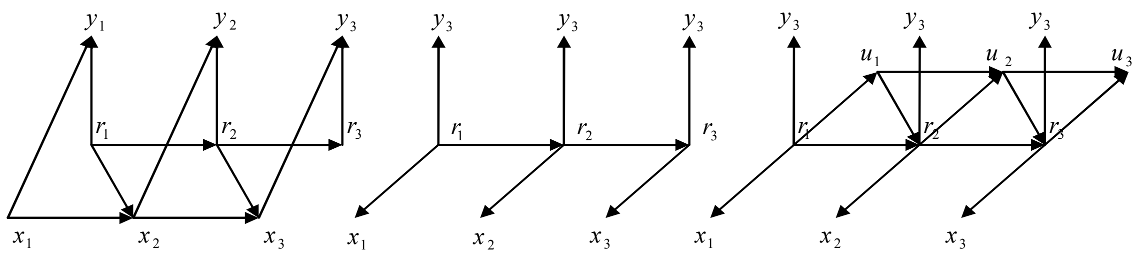
Then $E[X_{n+1}|Y_1^{n+1} = y_1^{n+1})$ and $p(r_{n+1}|y_1^{n+1})$ are computable with linear complexity in number of observations.

Remark 1

In the classical linear model with Markov jumps we have the following hypotheses [6, 7]:

(a) R is a Markov chain;
(b) $X_{n+1} = F_n(R_n)X_n + W_n$;
(c) $Y_n = H_n(R_n)X_n + Z_n$,

with (W_n) and (Z_n) appropriate "noises". In such a model the exact calculation of the Kalman filter with linear complexity in number of observations is not feasible and different approximations, like particle filters, are to be used [2, 5]. We can see, according to the dependence graphs in Figure 1, that the first new model presents a kind of simplification. Thus we lose in generality, and we win in the possibility of exact calculation. The crux point, which allows the exact calculations, is the independence of X and Y conditionally on R (of course, it does not mean that X and Y are independent). As we are going to see in the next section, this conditional independence will still be verified when generalizing the Markovianity of R to its semi-Markovianity, which is at the origin of the second new model proposed.



Classical model First proposed model Semi-Markov jumping model

Figure 1. Dependence graphs of the classical Markov switching linear system, the proposed Markov switching system, and the proposed semi- Markov switching system.

TRIPLET SEMI-MARKOV JUMPING PROCESS

Let us replace the Markov jumps by semi-Markov jumps. One can see that R is a semi-Markov chain if its distribution is the marginal distribution of an appropriate Markov chain (R,U). Let us consider the following family of semi-Markov chains R, which have been successfully used in the segmentation context in [9]. Each U_i takes its values in $\Upsilon = \{0,1,...,m\}$, so that (R,U) is a finite Markov chain. For $(R_n,U_n) = (r_n,u_n)$, the number u_n denotes the minimal sojourn time of the next variables R_{n+1}, …. in r_n. Therefore, if $u_n = j > 0$, we have $(r_{n+1},u_{n+1}) = (r_n,u_n - 1)$, …, $(r_{n+j},u_{n+j}) = (r_n,0)$. If $u_n = 0$, the distribution of R_{n+1} is a given transition $p(r_{n+1}|r_n,u_n = 0)$. Finally, the transition $p(r_{n+1},u_{n+1}|r_n,u_n) = p(r_{n+1}|r_n,u_n)p(u_{n+1}|r_{n+1},r_n,u_n)$ of the Markov chain (R,U) is defined by

$$p(r_{n+1}|r_n,u_n) = \delta_{r_n}(r_{n+1}) \text{ if } u_n > 0 \text{, and } p(r_{n+1}|r_n) \text{ if } u_n = 0 \text{ ;} \qquad (6)$$

$$p(u_{n+1}|r_{n+1},r_n,u_n) = \delta_{u_n - 1}(u_{n+1}) \text{ if } u_n > 0 \text{, and } p(u_{n+1}|r_{n+1}) \text{ if } u_n = 0 \text{ ;} \qquad (7)$$

with $\delta_a(b) = 1$ for $a = b$, and $\delta_a(b) = 0$ otherwise. Therefore we have four chains X, R, U, and Y and the problem is the same as in the first section : calculate $E[X_{n+1} | Y_1^{n+1} = y_1^{n+1})$ from $E[X_n | Y_1^n = y_1^n)$.
Let us define the distribution $p(x,r,u,y)$ by

$$p(x,r,u,y) = p(r,u)p(x,y|r,u) = p(r,u)p(x,y|r) = p(r,u)p(x|r)p(y|r), \quad (8)$$

where $p(r,u)$ is a Markov chain, $p(x|r)$ is defined by (iii), and $p(y|r)$ is defined by (ii). Replacing the Markov chain R by the Markov chain (R,U), we can use the Proposition 1 to directly state:

Proposition 2

Let (X,R,Y) be a triplet chain such that $p(x,r,y)$ is the Marginal distribution of the Markov distribution $p(x,r,u,y)$ given by (8), with $p(r,u)$ Markov chain distribution, $p(x|r)$ given by (iii), and $p(y|r)$ given by (ii). Then $E[X_{n+1} | Y_1^{n+1} = y_1^{n+1})$ can be computed from $E[X_n | Y_1^n = y_1^n]$ by (4)-(5), in which $p(r_{n+1}|r_n)$ are replaced by $p(r_{n+1},u_{n+1}|r_n,u_n)$. In particular, when $p(r_{n+1},u_{n+1}|r_n,u_n)$ are defined by (6) and (7), we have a system with semi-Markov jumps.
As a consequence, $E[X_{n+1} | Y_1^{n+1} = y_1^{n+1}]$ is computable with linear complexity in number of observations.
The dependence graph of the semi-Markov jumping model is presented in Figure 1.

.

REFERENCES

1. B. Ait-el-Fquih and F. Desbouvries, Kalman filtering in triplet Markov chains, *IEEE Trans. on Signal Processing*, Vol. 54, No. 8, pp. 2957-2963, 2006.
2. C. Andrieu, M. Davy, and A. Doucet, Efficient particle filtering for jump Markov systems. Application to time-varying autoregressions, IEEE Trans on Signal Processing, Vol. 51, No. 7, pp. 1762-1770, 2003.
3. D. Benboudjema and W. Pieczynski, Unsupervised statistical segmentation of non stationary images using triplet Markov fields, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 8, pp. 1367-1378, 2007.
4. D. Benboudjema and W. Pieczynski, Unsupervised image segmentation using triplet Markov fields, *Computer Vision and Image Understanding*, Vol. 99, No. 3, pp. 476-498, 2005.
5. O. L. V. Costa, M. D. Fragoso, and R. P. Marques, Discrete time Markov jump linear systems, New York, Springer-Verlag, 2005
6. Y. Ephraim and N. Merhaw, Hidden Markov processes, *IEEE Trans. on Information Theory,* Vol. 48, No. 6, pp. 1518-1569, 2002.
7. A. H. Jazwinski, *Stochastic processes and filtering theory*, Vol. 64, Mathematics in Science and Engeneering, Academic Press, San Diego, 1970.
8. P. Lanchantin and W. Pieczynski, Unsupervised non stationary image segmentation using triplet Markov chains, Advanced Concepts for Intelligent Vision Systems (ACVIS 04), Aug. 31-Sept. 3, Brussels, Belgium, 2004.
9. J. Lapuyade-Lahorgue and W. Pieczynski, Unsupervised segmentation of hidden semi-Markov non stationary chains, Twenty six International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, MaxEnt 2006, Paris, France, July 8-13, 2006.
10. W. Pieczynski, C. Hulard, and T. Veit, Triplet Markov Chains in hidden signal restoration, SPIE's International Symposium on Remote Sensing, September 22-27, Crete, Greece, 2002.
11. W. Pieczynski and F. Desbouvries, On triplet Markov chains, International Symposium on Applied Stochastic Models and Data Analysis, (ASMDA 2005), Brest, France, May 2005.
12. W. Pieczynski, Multisensor triplet Markov chains and theory of evidence, *International Journal of Approximate Reasoning*, Vol. 45, No. 1, pp. 1-16, 2007.