PROCEEDINGS OF THE ICASSP, HONG-KONG, 2003

KALMAN FILTERING USING PAIRWISE GAUSSIAN MODELS

Wojciech Pieczynski and François Desbouvries

Institut National des Télécommunications, Département CITI
9 rue Charles Fourier, 91011 Evry, France
Wojciech.Pieczynski@int-evry.fr, Francois.Desbouvries@int-evry.fr

ABSTRACT

An important problem in signal processing consists in recursively estimating an unobservable process $\mathbf{x} = \{\mathbf{x}_n\}_{n\in\mathbb{N}}$ from an observed process $\mathbf{y} = \{\mathbf{y}_n\}_{n\in\mathbb{N}}$. This is done classically in the framework of Hidden Markov Models (HMM). In the linear Gaussian case, the classical recursive solution is given by the well-known Kalman filter. In this paper, we consider Pairwise Gaussian Models by assuming that the pair $(\mathbf{x}, \mathbf{y})$ is Markovian and Gaussian. We show that this model is strictly more general than the HMM, and yet still enables Kalman-like filtering.

1. INTRODUCTION

Since its introduction in the 1960s in the control engineering community the Kalman filter has become a major tool in signal processing and automatic control. In Kalman filtering we wish to estimate samples of an unobserved process $\mathbf{x}$, given samples of some observed process $\mathbf{y}$ and a (state-space) dynamic stochastic model for processes $\mathbf{x}$ and $\mathbf{y}$.

Now, it is well known (see e.g. [1] [2]) that the state-space model which underlies the Kalman filter is indeed an HMM (with continuous state process).

In this paper, we propose to extend the Kalman filter by using the general idea that enabled us to successfully generalize Hidden Markov Fields (HMF) to Pairwise Markov Fields (PMF) [3], Hidden Markov Chains (with discrete hidden process) (HMC) to Pairwise Markov Chains (PMC) [4] [5], and Hidden Markov Trees (HMT) to Pairwise Markov Trees (PMT) [6]. More precisely, it is well known that if $(\mathbf{x}, \mathbf{y})$ is a classical HMM, then the pair $(\mathbf{x}, \mathbf{y})$ itself is Markovian. Our aim is to study the converse proposition : what can be said if the pair $(\mathbf{x}, \mathbf{y})$ is a Markov Chain (MC)?

In this paper, we thus directly assume that the pair $(\mathbf{x}, \mathbf{y})$ is a MC, and we show : (i) that a Kalman-like filter can still be computed; and (ii) that such a "Pairwise Markov Model" (PMM) is strictly more general than the classical HMM.

This paper is organized as follows. In section 2 we recall the classical dynamical state-space model, as well as the properties which underly the derivation of the Kalman filter. In section 3 we introduce a generalized stochastic dynamical model in which the pair $(\mathbf{x}, \mathbf{y})$ is Markovian. We derive the general time-update and measurement-update equations for this model, as well as the associated Kalman-like filter which holds in the particular case of a linear and Gaussian PMM. In section 4 we show that Gaussian PMM are strictly more general than Gaussian HMM.

2. CLASSICAL HIDDEN MARKOV MODELS

2.1. General HMM

Let us consider the following classical stochastic dynamical system :

$$\begin{cases} \mathbf{x}_{n+1} = g_n(\mathbf{x}_n, \mathbf{u}_n) \\ \mathbf{y}_n = h_n(\mathbf{x}_n, \mathbf{v}_n) \end{cases} , \quad (1)$$

in which $g_n(.,.)$ is a (possibly nonlinear) function from $\mathbb{R}^N \times \mathbb{R}^p$ to $\mathbb{R}^N$, $h_n(.,.)$ is a (possibly nonlinear) function from $\mathbb{R}^N \times \mathbb{R}^q$ to $\mathbb{R}^q$, and $\mathbf{u} = \{\mathbf{u}_n\}_{n\in\mathbb{N}}$ and $\mathbf{v} = \{\mathbf{v}_n\}_{n\in\mathbb{N}}$ are zero-mean sequences which are independent, jointly independent and independent of $\mathbf{x}_0$.

Let $\mathbf{x}_{0:n} = \{\mathbf{x}_i\}_{i=0}^n$ and $\mathbf{y}_{0:n} = \{\mathbf{y}_i\}_{i=0}^n$. Let also $p(\mathbf{x}_n)$, $p(\mathbf{x}_{0:n})$ and $p(\mathbf{x}_n|\mathbf{y}_{0:n})$, say, denote the probability density function (pdf) of $\mathbf{x}_n$, the pdf of $\mathbf{x}_{0:n}$, and the pdf of $\mathbf{x}_n$, conditional on $\mathbf{y}_{0:n}$, respectively; the other pdf are defined similarly. Then one can check that the following properties hold :

$$p(\mathbf{x}_{n+1}|\mathbf{x}_{0:n}) = p(\mathbf{x}_{n+1}|\mathbf{x}_n) \; ; \quad (2)$$

$$p(\mathbf{y}_{0:n}|\mathbf{x}_{0:n}) = \prod_{i=0}^{n} p(\mathbf{y}_i|\mathbf{x}_{0:n}) \; ; \quad (3)$$

$$p(\mathbf{y}_i|\mathbf{x}_{0:n}) = p(\mathbf{y}_i|\mathbf{x}_i) \text{ for all } i, \; 0 \le i \le n \; . \quad (4)$$

In other words, $\mathbf{x}$ is a MC, and since it is known only through the observed process $\mathbf{y}$, (1) is an HMM.

Now, from (2) to (4) we get

$$p(\mathbf{x}_{n+1}|\mathbf{x}_n, \mathbf{y}_{0:n}) = p(\mathbf{x}_{n+1}|\mathbf{x}_n) \; , \quad (5)$$

$$p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1}, \mathbf{y}_{0:n}) = p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1}). \quad (6)$$

As a consequence, the recursive propagation of the posterior density (i.e., the computing of $p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n+1})$ from $p(\mathbf{x}_n|\mathbf{y}_{0:n})$) under model (1) is given by the following "time-update" and "measurement-update" equations :

$$p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n}) = \int p(\mathbf{x}_{n+1}|\mathbf{x}_n)p(\mathbf{x}_n|\mathbf{y}_{0:n})d\mathbf{x}_n \quad ; \quad (7)$$

$$p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n+1}) = \frac{p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1})p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n})}{p(\mathbf{y}_{n+1}|\mathbf{y}_{0:n})} = \frac{p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1})p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n})}{\int p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1})p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n})d\mathbf{x}_{n+1}} \quad (8)$$

2.2. Linear Gaussian HMM

Let us now consider the important particular case in which (1) reduces to the linear, stochastic dynamical model

$$\left\{ \begin{array}{lcl} \mathbf{x}_{n+1} & = & \mathbf{F}_n\mathbf{x}_n + \mathbf{G}_n\mathbf{u}_n \\ \mathbf{y}_n & = & \mathbf{H}_n\mathbf{x}_n + \mathbf{v}_n \end{array} \right. , \quad (9)$$

in which $\mathbf{F}_n$, $\mathbf{G}_n$ and $\mathbf{H}_n$ are matrices of dimensions $N \times N$, $N \times p$ and $q \times N$, respectively. Let $\mathbf{w}_n = [\mathbf{u}_n^T, \mathbf{v}_n^T]^T$ and $\mathbf{z}_n = [\mathbf{x}_{n+1}^T, \mathbf{y}_n^T]^T$. If furthermore $\mathbf{x}_0$ and $\mathbf{w}_n$ are Gaussian variables, the process $\mathbf{z} = \{\mathbf{z}_n\}_{n\in\mathbb{N}}$ is Gaussian. So the posterior densities are also Gaussian and are thus described by their means and covariance matrices. Propagating $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ amounts to propagating these parameters, and in this case equations (7) and (8) reduce to the well known Kalman filter [7]. On the other hand, if the Gaussian assumption does not hold, or if (9) is replaced by the general, nonlinear model (1), then computing equations (7) and (8) often becomes difficult in practice. Consequently, a number of approximate, Monte Carlo based methods have been derived; see e.g. the recent books [8] [9] or tutorial [10].

3. PAIRWISE MARKOV MODELS

3.1. General PMM

Let us now turn back to model (1). Let $\mathbf{z}_n = [\mathbf{x}_n^T, \mathbf{y}_{n-1}^T]^T$ and let $\mathbf{z}_0 = \mathbf{x}_0$. $\mathbf{z}_n$ satisfies

$$\mathbf{z}_{n+1} = G_n(\mathbf{z}_n, \mathbf{w}_n) \quad (10)$$

for some function $G_n(.,.)$, where $\mathbf{w}_n = [\mathbf{u}_n^T, \mathbf{v}_n^T]^T$ is a zero-mean process which is independent and independent of $\mathbf{x}_0$. As a consequence (and as is well known), the process $\mathbf{z} = \{\mathbf{z}_n\}_{n\in\mathbb{N}}$ is also a MC.

Throughout this section we will thus consider model (10). As we now see, this model still enables to solve the filtering problem. One can show that equations (5) and (6) are replaced respectively by

$$p(\mathbf{x}_{n+1}|\mathbf{x}_n, \mathbf{y}_{0:n}) = p(\mathbf{x}_{n+1}|\mathbf{x}_n, \mathbf{y}_n, \mathbf{y}_{n-1}), \quad (11)$$

$$p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1}, \mathbf{y}_{0:n}) = p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1}, \mathbf{y}_n). \quad (12)$$

Consequently, (7) and (8) are replaced respectively by the new relations

$$p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n}) = \int p(\mathbf{x}_{n+1}|\mathbf{x}_n, \mathbf{y}_n, \mathbf{y}_{n-1})p(\mathbf{x}_n|\mathbf{y}_{0:n})d\mathbf{x}_n, \quad (13)$$

$$p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n+1}) = \frac{p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1}, \mathbf{y}_n)p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n})}{\int p(\mathbf{y}_{n+1}|\mathbf{x}_{n+1}, \mathbf{y}_n)p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n})d\mathbf{x}_{n+1}}, \quad (14)$$

which enable to compute $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ recursively under model (10).

3.2. Linear Gaussian PMM

We now consider the important particular case where $G_n(.,.)$ is a linear function. Let us consider the following model :

$$\underbrace{\begin{bmatrix} \mathbf{x}_{n+1} \\ \mathbf{y}_n \end{bmatrix}}_{\mathbf{z}_{n+1}} = \underbrace{\begin{bmatrix} \mathbf{F}_n^1 & \mathbf{F}_n^2 \\ \mathbf{H}_n^1 & \mathbf{H}_n^2 \end{bmatrix}}_{\mathcal{F}_n} \begin{bmatrix} \mathbf{x}_n \\ \mathbf{y}_{n-1} \end{bmatrix} + \underbrace{\begin{bmatrix} \mathbf{G}_n^{11} & \mathbf{G}_n^{12} \\ \mathbf{G}_n^{21} & \mathbf{G}_n^{22} \end{bmatrix}}_{\mathcal{G}_n} \underbrace{\begin{bmatrix} \mathbf{u}_n \\ \mathbf{v}_n \end{bmatrix}}_{\mathbf{w}_n} \quad (15)$$

in which $\{\mathbf{w}_n = [\mathbf{u}_n^T, \mathbf{v}_n^T]^T\}_{n\geq 0}$ are random vectors which are zero-mean, independent and independent of $\mathbf{x}_0$. Matrices $\mathbf{F}_n^1$, $\mathbf{F}_n^2$, $\mathbf{H}_n^1$ and $\mathbf{H}_n^2$ are of dimensions $N \times N$, $N \times q$, $q \times N$ and $q \times q$, and $\mathbf{G}_n^{11}$, $\mathbf{G}_n^{12}$, $\mathbf{G}_n^{21}$ and $\mathbf{G}_n^{22}$ of dimensions $N \times p$, $N \times q$, $q \times p$ and $q \times q$, respectively. This model is a particular case of model (10), and a generalization of the classical linear HMM (1), which is obtained by setting $\mathbf{F}_n^2 = \mathbf{0}_{N\times q}$, $\mathbf{H}_n^2 = \mathbf{0}_{q\times q}$, $\mathbf{G}_n^{12} = \mathbf{0}_{N\times q}$, $\mathbf{G}_n^{21} = \mathbf{0}_{q\times p}$ and $\mathbf{G}_n^{22} = \mathbf{I}_{q\times q}$.

Let us further assume that the process $\mathbf{w} = \{\mathbf{w}_n\}_{n\in\mathbb{N}}$ is Gaussian and that $p(\mathbf{x}_0) \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0)$. Then $\mathbf{z}$ is a Gaussian process and consequently the pdf $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ and $p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n})$ are also Gaussian. Let us set

$$p(\mathbf{x}_n|\mathbf{y}_{0:n}) \sim \mathcal{N}(\hat{\mathbf{x}}_{n|n}, \mathbf{P}_{n|n}) \quad , \quad (16)$$

$$p(\mathbf{x}_{n+1}|\mathbf{y}_{0:n}) \sim \mathcal{N}(\hat{\mathbf{x}}_{n+1|n}, \mathbf{P}_{n+1|n}) \quad , \quad (17)$$

and let

$$E(\mathbf{w}_n\mathbf{w}_m^T) = \begin{bmatrix} \mathbf{Q}_n & \mathbf{S}_n \\ \mathbf{S}_n^T & \mathbf{R}_n \end{bmatrix} \delta_{n,m} = \mathcal{Q}_n \, \delta_{n,m} \, , \quad (18)$$

$$\begin{bmatrix} \widetilde{\mathbf{G}}_n^{11} & \widetilde{\mathbf{G}}_n^{12} \\ \widetilde{\mathbf{G}}_n^{21} & \widetilde{\mathbf{G}}_n^{22} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_n^{11} & \mathbf{G}_n^{12} \\ \mathbf{G}_n^{21} & \mathbf{G}_n^{22} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_n & \mathbf{S}_n \\ \mathbf{S}_n^T & \mathbf{R}_n \end{bmatrix} \begin{bmatrix} \mathbf{G}_n^{11} & \mathbf{G}_n^{12} \\ \mathbf{G}_n^{21} & \mathbf{G}_n^{22} \end{bmatrix}^T = \widetilde{\mathcal{G}}_n \, . \quad (19)$$

The following result is an extension to model (15) of the classical Kalman filter :

Proposition 1 (Pairwise Kalman Filter) *Let us assume that model (15) holds. Suppose that* $p(\mathbf{x}_0) \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0)$ *and that* $p(\mathbf{w}_n) \sim \mathcal{N}(\mathbf{0}, \mathcal{Q}_n)$*. Then* $\hat{\mathbf{x}}_{n+1|n+1}$ *and* $\mathbf{P}_{n+1|n+1}$

can be computed from $\hat{\mathbf{x}}_{n|n}$ *et* $\mathbf{P}_{n|n}$ *via :*

$$\hat{\mathbf{x}}_{n+1|n} = [\mathbf{F}_n^1 - \widetilde{\mathbf{G}}_n^{12}(\widetilde{\mathbf{G}}_n^{22})^{-1}\mathbf{H}_n^1]\hat{\mathbf{x}}_{n|n} + \widetilde{\mathbf{G}}_n^{12}(\widetilde{\mathbf{G}}_n^{22})^{-1}\mathbf{y}_n + [\mathbf{F}_n^2 - \widetilde{\mathbf{G}}_n^{12}(\widetilde{\mathbf{G}}_n^{22})^{-1}\mathbf{H}_n^2]\,\mathbf{y}_{n-1} \quad (20)$$

$$\mathbf{P}_{n+1|n} = [\widetilde{\mathbf{G}}_n^{11} - \widetilde{\mathbf{G}}_n^{12}(\widetilde{\mathbf{G}}_n^{22})^{-1}\widetilde{\mathbf{G}}_n^{21}] + [\mathbf{F}_n^1 - \widetilde{\mathbf{G}}_n^{12}(\widetilde{\mathbf{G}}_n^{22})^{-1}\mathbf{H}_n^1]\;\mathbf{P}_{n|n}\,[\mathbf{F}_n^1 - \widetilde{\mathbf{G}}_n^{12}(\widetilde{\mathbf{G}}_n^{22})^{-1}\mathbf{H}_n^1]^T \quad (21)$$

$$\widetilde{\mathbf{y}}_{n+1} = \mathbf{y}_{n+1} - \mathbf{H}_{n+1}^1\hat{\mathbf{x}}_{n+1|n} - \mathbf{H}_{n+1}^2\mathbf{y}_n \quad (22)$$

$$\mathbf{L}_{n+1} = \widetilde{\mathbf{G}}_{n+1}^{22} + \mathbf{H}_{n+1}^1\mathbf{P}_{n+1|n}(\mathbf{H}_{n+1}^1)^T \quad (23)$$

$$\mathbf{K}_{n+1|n+1} = \mathbf{P}_{n+1|n}(\mathbf{H}_{n+1}^1)^T\mathbf{L}_{n+1}^{-1} \quad (24)$$

$$\hat{\mathbf{x}}_{n+1|n+1} = \hat{\mathbf{x}}_{n+1|n} + \mathbf{K}_{n+1|n+1}\,\widetilde{\mathbf{y}}_{n+1} \quad (25)$$

$$\mathbf{P}_{n+1|n+1} = \mathbf{P}_{n+1|n} - \mathbf{K}_{n+1|n+1}\mathbf{L}_{n+1}\mathbf{K}_{n+1|n+1}^T \quad (26)$$

Proof. From (15) and (19), we get

$$p(\mathbf{x}_{n+1}, \mathbf{y}_n|\mathbf{x}_n, \mathbf{y}_{n-1}) \sim \mathcal{N}(\mathcal{F}_N \begin{bmatrix} \mathbf{x}_n \\ \mathbf{y}_{n-1} \end{bmatrix}, \widetilde{\mathcal{G}}_n) \,. \quad (27)$$

By combining (13), (14), (16), (17) and (27), we compute $p(\mathbf{x}_{n+1}, \mathbf{x}_n|\mathbf{y}_{0:n})$ and $p(\mathbf{x}_{n+1}, \mathbf{y}_{n+1}|\mathbf{y}_{0:n})$, from which equations (20) to (26) are deduced. ■

We check that if $\mathbf{F}_n^2 = \mathbf{0}_{N\times q}$, $\mathbf{H}_n^2 = \mathbf{0}_{q\times q}$, $\mathbf{G}_n^{12} = \mathbf{0}_{N\times q}$, $\mathbf{G}_n^{21} = \mathbf{0}_{q\times p}$ et $\mathbf{G}_n^{22} = \mathbf{I}_{q\times q}$, then (15) reduces to the classical model (9), and equations (20) to (26) reduce to classical Kalman filter equations (for example, (20), (21), (25) and (26) coincide respectively with [11, eq. (5.5) p. 115], [11, eq. (5.12) p. 117], [11, eq. (5.6) p. 116] and [11, eq. (5.11) p. 117]).

Remarks.

The introduction of Pairwise Models in the context of Kalman filtering is not entirely new. A closely related model was introduded independently in the Gaussian case [12] (see also [13, Corollary 1 p. 72]). In this model, the pair $\mathbf{z}_n = [\mathbf{x}_n^T, \mathbf{y}_n^T]^T$ satisfies a linear equation similar to (15) and thus is Markovian. Optimal filtering equations for this model have also been derived [13, eqs. (13.56) et (13.57)] (in fact, due to the difference in times indices, these equations extend the classical Kalman one-step ahead prediction algorithm $(\hat{\mathbf{x}}_{n|n-1}, \mathbf{P}_{n|n-1}) \rightarrow (\hat{\mathbf{x}}_{n+1|n}, \mathbf{P}_{n+1|n})$). However, to our best knowledge, equations (20) to (26) of Proposition 1 (which form the optimal filter for model (15) in the Gaussian case); equations (13) and (14), which generalize (7) and (8), on the one hand, and (20) to (26), on the other hand; and section 4, which specifies relationships between PMM and classical HMM, are original.

4. PAIRWISE MARKOV MODELS VS. HIDDEN MARKOV MODELS

In section 3, we introduced models (10) and (15), which are generalizations of (1) and (9), respectively. The aim of this section is to make relations between HMM and PMM clearer. To that end, we are looking for conditions under which the marginal process $\mathbf{x}$ of a Markovian process $\mathbf{z} = (\mathbf{x}, \mathbf{y})$ is itself Markovian.

Proposition 2 *Let* $\mathbf{z}_n = [\mathbf{x}_n^T, \mathbf{y}_{n-1}^T]^T$ *(with* $\mathbf{z}_0 = \mathbf{x}_0$*) and* $\mathbf{z} = \{\mathbf{z}_n\}_{n\in\mathbb{N}}$*. Assume that* $\mathbf{z}$ *is a MC. Further assume that either*

$$\text{for all } n,\;\; p(\mathbf{y}_n|\mathbf{x}_{n+1}, \mathbf{x}_{n+2}) = p(\mathbf{y}_n|\mathbf{x}_{n+1}) \;\;, \quad (28)$$

or

$$\text{for all } n,\;\; p(\mathbf{y}_n|\mathbf{x}_{n+1}, \mathbf{x}_n) = p(\mathbf{y}_n|\mathbf{x}_{n+1}) \;\;. \quad (29)$$

Then $\{\mathbf{x}_n\}_{n\geq 0}$ *is a MC.*

Proof. Since $\mathbf{z}$ is a MC,

$$p(\mathbf{z}_{0:n}) = \frac{p(\mathbf{z}_0, \mathbf{z}_1)\cdots p(\mathbf{z}_{n-1}, \mathbf{z}_n)}{p(\mathbf{z}_1)\cdots p(\mathbf{z}_{n-1})} = \underbrace{\frac{p(\mathbf{y}_0|\mathbf{x}_0, \mathbf{x}_1)\cdots p(\mathbf{y}_{n-2}, \mathbf{y}_{n-1}|\mathbf{x}_{n-1}, \mathbf{x}_n)}{p(\mathbf{y}_0|\mathbf{x}_1)\cdots p(\mathbf{y}_{n-2}|\mathbf{x}_{n-1})}}_{A} \times \underbrace{\frac{p(\mathbf{x}_0, \mathbf{x}_1)\cdots p(\mathbf{x}_{n-1}, \mathbf{x}_n)}{p(\mathbf{x}_1)\cdots p(\mathbf{x}_{n-1})}}_{B} \,.$$

$\mathbf{x}$ is a MC if and only if $\int A d\mathbf{y}_{0:n-1} = 1$, which is ensured under (28) or under (29). ■

Conversely, we are now looking for local conditions implied if $\mathbf{x}$ is Markovian. In the Gaussian case, the following result holds :

Proposition 3 *Let* $\mathbf{z}_n = [\mathbf{x}_n^T, \mathbf{y}_{n-1}^T]^T$ *(with* $\mathbf{z}_0 = \mathbf{x}_0$*) and* $\mathbf{z} = \{\mathbf{z}_n\}_{n\in\mathbb{N}}$*. Assume that* $\mathbf{z}$ *is a MC. Further assume that* $\mathbf{z}$ *is zero-mean and Gaussian, and that* $y_n \in \mathbb{R}$ *(i.e. that* $q = 1$*). If* $\{\mathbf{x}_n\}_{n\geq 0}$ *is a MC, then for all* n*, either* $p(y_n|\mathbf{x}_{n+1}, \mathbf{x}_{n+2}) = p(y_n|\mathbf{x}_{n+1})$*, or* $p(y_n|\mathbf{x}_{n+1}, \mathbf{x}_n) = p(y_n|\mathbf{x}_{n+1})$*.*

Proof. Since $\mathbf{z}$ is a MC, for all n $[\mathbf{x}_n^T, y_{n-1}]^T$ and $[\mathbf{x}_{n+2}^T, y_{n+1}]^T$ are independent conditionally on $[\mathbf{x}_{n+1}^T, y_n]^T$. Consequently, $\mathbf{x}_n$ et $\mathbf{x}_{n+2}$ are also independent conditionally on $[\mathbf{x}_{n+1}^T, y_n]^T$. Let

$$E(\begin{bmatrix} \mathbf{x}_{n+1} \\ y_n \\ \mathbf{x}_n \\ \mathbf{x}_{n+2} \end{bmatrix}(.)^T) = \begin{bmatrix} \mathbf{A}_n & \mathbf{B}_n^T & \mathbf{C}_n^T & \mathbf{D}_n^T \\ \mathbf{B}_n & e_n & \mathbf{F}_n^T & \mathbf{G}_n^T \\ \mathbf{C}_n & \mathbf{F}_n & \mathbf{H}_n & \mathbf{J}_n^T \\ \mathbf{D}_n & \mathbf{G}_n & \mathbf{J}_n & \mathbf{K}_n \end{bmatrix} . \quad (30)$$

Conditionally on $[\mathbf{x}_{n+1}^T, y_n]^T$, the pdf of $[\mathbf{x}_n^T, \mathbf{x}_{n+2}^T]^T$ is Gaussian with covariance matrix

$$\begin{bmatrix} \mathbf{H}_n & \mathbf{J}_n^T \\ \mathbf{J}_n & \mathbf{K}_n \end{bmatrix} - \begin{bmatrix} \mathbf{C}_n & \mathbf{F}_n \\ \mathbf{D}_n & \mathbf{G}_n \end{bmatrix} \begin{bmatrix} \mathbf{A}_n & \mathbf{B}_n^T \\ \mathbf{B}_n & e_n \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{C}_n^T & \mathbf{D}_n^T \\ \mathbf{F}_n^T & \mathbf{G}_n^T \end{bmatrix} ;$$

the variables $\mathbf{x}_n$ et $\mathbf{x}_{n+2}$ are independent conditionnally on $[\mathbf{x}_{n+1}^T, y_n]^T$ if and only if this matrix is block-diagonal, i.e. if and only if

$$\begin{aligned}\text{(i) } \mathbf{0}_{N\times N} &= (\mathbf{J}_n - \mathbf{D}_n\mathbf{A}_n^{-1}\mathbf{C}_n^T) - (\mathbf{G}_n - \mathbf{D}_n\mathbf{A}_n^{-1}\mathbf{B}_n^T) \\ &\quad \times (e_n - \mathbf{B}_n\mathbf{A}_n^{-1}\mathbf{B}_n^T)^{-1}(\mathbf{F}_n^T - \mathbf{B}_n\mathbf{A}_n^{-1}\mathbf{C}_n^T).\end{aligned}$$

Now, further assume that $\mathbf{x}_n$ is a MC. Then for all n, $\mathbf{x}_n$ and $\mathbf{x}_{n+2}$ are independent conditionally on $\mathbf{x}_{n+1}$, which is equivalent to

$$\text{(ii)} \quad \mathbf{J}_n - \mathbf{D}_n\mathbf{A}_n^{-1}\mathbf{C}_n^T = \mathbf{0}_{N\times N} \ .$$

Consequently, under condition (i), (ii) holds if and only if

$$\underbrace{(\mathbf{G}_n - \mathbf{D}_n\mathbf{A}_n^{-1}\mathbf{B}_n^T)}_{N\times 1}\underbrace{(e_n - \mathbf{B}_n\mathbf{A}_n^{-1}\mathbf{B}_n^T)}_{1\times 1}{}^{-1}\underbrace{(\mathbf{F}_n^T - \mathbf{B}_n\mathbf{A}_n^{-1}\mathbf{C}_n^T)}_{1\times N}$$

is equal to $\mathbf{0}_{N\times N}$, i.e., since $q = 1$, if and only if

$$\mathbf{G}_n - \mathbf{D}_n\mathbf{A}_n^{-1}\mathbf{B}_n^T = \mathbf{0}_{N\times 1} \text{ or } \mathbf{F}_n^T - \mathbf{B}_n\mathbf{A}_n^{-1}\mathbf{C}_n^T = \mathbf{0}_{1\times N} \ .$$

As we see from (30), this condition means that $\mathbf{x}_{n+2}$ and y_n are independent conditionally on $\mathbf{x}_{n+1}$, or that $\mathbf{x}_n$ and y_n are independent conditionally on $\mathbf{x}_{n+1}$, which can be written as

$$\begin{aligned} p(y_n|\mathbf{x}_{n+1}, \mathbf{x}_{n+2}) &= p(y_n|\mathbf{x}_{n+1}) \quad \text{or} \\ p(y_n|\mathbf{x}_{n+1}, \mathbf{x}_n) &= p(y_n|\mathbf{x}_{n+1}) \ . \end{aligned}$$

■

Remarks

The sufficient condition of Proposition 2 is local and can thus easily be checked in the framework of a dynamic stochastic model. For instance, let us come back to the Pairwise linear model (15). We check that if $\mathbf{F}_n^2 = \mathbf{0}$, then $p(\mathbf{x}_{n+1}|\mathbf{x}_n, \mathbf{y}_{n-1}) = p(\mathbf{x}_{n+1}|\mathbf{x}_n)$ and so the process $\mathbf{x}$ is a MC. Note that the case $\mathbf{F}_n^2 = \mathbf{0}$, $\mathbf{H}_n^2 \neq \mathbf{0}$ provides a model which is more general than (9), and in which $\mathbf{x}$ remains Markovian.

On the other hand, we can easily verify that there exist models for which the necessary condition of Proposition 3 is not satisfied (consider for instance the model

$$\mathbf{z}_{n+1} = \begin{bmatrix} .5 & .1 \\ 1 & 0 \end{bmatrix}\mathbf{z}_n + \mathbf{w}_n, \mathcal{Q}_n = \begin{bmatrix} 1 & .3 \\ .3 & 1 \end{bmatrix}, p(\mathbf{x}_0) \sim \mathcal{N}(0,1).)$$

This shows that we can find PMM for which $\mathbf{x}$ is not a MC, and thus that model (10) is strictly more general than model (1).

5. REFERENCES

[1] P. E. Caines, *Linear Stochastic Systems*, Wiley series in Probability and Mathematical Statistics. Wiley, New York, 1988.

[2] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov Models, Estimation and Control*, vol. 29 of *Applications of mathematics - Stochastic Modelling and Applied Probability*, Springer-verlag, New York, 1995.

[3] W. Pieczynski and A. N. Tebbache, "Pairwise markov random fields and segmentation of textured images," *Machine Graphics and Vision*, vol. 9, no. 3, pp. 705–718, 2000.

[4] W. Pieczynski, "Pairwise markov chains," *Accepted for publication, IEEE Tr. Pattern Analysis and Machine Intelligence*, 2002.

[5] S. Derrode and W. Pieczynski, "Sar image segmentation using generalized pairwise markov chains," in *Proceedings of SPIE International Symposium on Remote Sensing*, Crete, Greece, September 22-27, 2002.

[6] W. Pieczynski, "Arbres de markov couple - pairwise markov trees," *Comptes Rendus de l'Académie des Sciences - Mathématiques*, vol. 335, pp. 79–82, 2002, Ser. I (*in French*).

[7] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng., Trans. ASME, Series D*, vol. 82, no. 1, pp. 35–45, 1960.

[8] H. Tanizaki, *Nonlinear Filters, Estimation and Applications*, Springer, Berlin, 2nd edition, 1996.

[9] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, Statistics for Engineering and Information Science. Springer Verlag, New York, 2001.

[10] M. Sanjeev Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear / non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, February 2002.

[11] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice Hall, Englewood Cliffs, New Jersey, 1979.

[12] R. S. Lipster and A. N. Shiryaev, "Statistics of conditionally gaussian random sequences," in *Sixth Berkeley Symposium on Mathematics, Statistics and Probability*. 1972, vol. 2, pp. 389–422, University of California Press.

[13] R. S. Lipster and A. N. Shiryaev, *Statistics of Random Processes, Vol. 2 : Applications*, chapter 13 : "Conditionally Gaussian Sequences : Filtering and Related Problems", Springer Verlag, Berlin, 2001.