KALMAN FILTERING FOR TRIPLET MARKOV CHAINS : APPLICATIONS AND EXTENSIONS

B. Ait El Fquih and F. Desbouvries

GET/INT/dépt. CITI and CNRS UMR 5157
9 rue Charles Fourier, 91011 Evry, France
boujemaa.ait_elfquih@int-evry.fr, francois.desbouvries@int-evry.fr

ABSTRACT

An important problem in signal processing consists in estimating an unobservable process $\mathbf{x} = \{\mathbf{x}_n\}_{n\in\mathbb{N}}$ from an observed process $\mathbf{y} = \{\mathbf{y}_n\}_{n\in\mathbb{N}}$. In Linear Gaussian Hidden Markov Chains (LGHMC), the classical recursive solution is given by the Kalman filter. In this paper, we consider Linear Gaussian Triplet Markov Chains (LGTMC) by assuming that the triplet $(\mathbf{x}, \mathbf{r}, \mathbf{y})$ (in which $\mathbf{r} = \{\mathbf{r}_n\}_{n\in\mathbb{N}}$ is some additional process) is Markovian and Gaussian. We first show that this model encompasses and generalizes the classical linear stochastic dynamical models with autoregressive process and / or measurement noise. We next propose (for the regular and for the perfect-measurement cases) restoration Kalman-like algorithms for general LGTMC.

1. INTRODUCTION

Let us consider the classical linear dynamical stochastic system :

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

in which $\mathbf{x}_n \in \mathbb{R}^{n_\mathbf{x}}$ is the state, $\mathbf{y}_n \in \mathbb{R}^{n_\mathbf{y}}$ is the observation, and $\mathbf{F}_n$, $\mathbf{G}_n$, $\mathbf{H}_n$ and $\mathbf{J}_n$ are known deterministic matrices. The input noise $\mathbf{u}_n \in \mathbb{R}^{n_\mathbf{u}}$ and the measurement noise $\mathbf{v}_n \in \mathbb{R}^{n_\mathbf{v}}$ are assumed to be 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) (w.r.t. Lebesgue measure) of $\mathbf{x}_n$, the pdf of $\mathbf{x}_{0:n}$, and the pdf of $\mathbf{x}_n$, conditionally on $\mathbf{y}_{0:n}$, respectively; the other pdf are defined similarly. A fundamental problem associated with (1) (the so-called filtering problem) is the recursive computation of the posterior pdf $p(\mathbf{x}_n|\mathbf{y}_{0:n})$. If furthermore $\mathbf{x}_0$ and $\mathbf{n}_n = [\mathbf{u}_n^T, \mathbf{v}_n^T]^T$ are Gaussian variables, then $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ is also Gaussian and is thus described by its mean and covariance matrix. Propagating $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ amounts to propagating these parameters, and the algorithm we get is the celebrated Kalman filter[1].

Since this pionnering work, the Kalman filter has been generalized in many directions. To name just a few examples, square-root type algorithms have been proposed; the independence assumptions on $\{\mathbf{u}_n\}$ and $\{\mathbf{v}_n\}$ have been dropped; and the extension to non-linear and / or non- Gaussian systems has been considered.

Yet another direction in which it is possible to extend the Kalman filter consists in releasing some conditional independence assumptions among $\mathbf{x}$ and $\mathbf{y}$. Let us come back to model (1). We see that $\mathbf{x}$ is a Markov Chain (MC), and since it is known only through the observed process $\mathbf{y}$, (1) is an HMC. Now, if (1) holds then both $\{(\mathbf{x}_n, \mathbf{y}_n)\}_{n\in\mathbb{N}}$ and $\{(\mathbf{x}_{n+1}, \mathbf{y}_n)\}_{n\in\mathbb{N}}$ are (vector) MC. Conversely, starting only from one of these assumptions (i.e. assuming a so-called "Pairwise" MC (PMC) model) is a more general point of view, which nevertheless enables efficient restoration algorithms; extending the Kalman filter to a model where $\{(\mathbf{x}_n, \mathbf{y}_n)\}$ (resp. $\{(\mathbf{x}_{n+1}, \mathbf{y}_n)\}$) is Markovian has been considered in [1] (resp. [2]).

Let now $\mathbf{r} = \{\mathbf{r}_n\}_{n\in\mathbb{N}}$ be some additional process, and let us set $\mathbf{t}_n = [\mathbf{x}_n^T, \mathbf{r}_n^T, \mathbf{y}_{n-1}^T]^T$. A Triplet MC (TMC) is a model in which we only assume that $\{\mathbf{t}_n\}$ is a (vector) MC. This model generalizes the PMC model, and yet enables (in the LG regular case) the development of an efficient Kalman-like restoration algorithm [3].

Let us turn to the contents of this paper. In section 2, we show that the classical linear models with autoregressive process and/or measurement noise are, among other models, some important particular cases (mostly with perfect, i.e. with unnoisy measurements) of the general linear TMC model; so the triplet model, which initially was designed as a Markovian extension of (1), happens also to encompass (and generalize) the early classical generalizations (as regards processes $\{\mathbf{u}_n\}$ and $\{\mathbf{v}_n\}$) of model (1). In section 3, we propose a restoration algorithm for general LGTMC with perfect measurements. Finally, some applications are considered in section 4.

[1] As is well known, one can equivalently drop the Gaussian assumption and derive the Kalman filter as the recursive solution of a linear minimum mean-square error estimation procedure.

2. LINEAR TMC : DEFINITION & APPLICATIONS

2.1. The linear TMC model

Let $\mathbf{x} = \{\mathbf{x}_n\}_{n\in\mathbb{N}}$ be the hidden state process, $\mathbf{y} = \{\mathbf{y}_n\}_{n\in\mathbb{N}}$ the observed process and $\mathbf{r} = \{\mathbf{r}_n\}_{n\in\mathbb{N}}$ an additional (possibly artificial) process. The process $\mathbf{t} = \{\mathbf{t}_n\}_{n\in\mathbb{N}}$, with $\mathbf{t}_n = [\mathbf{x}_n^T, \mathbf{r}_n^T, \mathbf{y}_{n-1}^T]^T$ is a linear TMC [3] if

$$\underbrace{\begin{bmatrix} \mathbf{x}_{n+1} \\ \mathbf{r}_{n+1} \\ \mathbf{y}_n \end{bmatrix}}_{\mathbf{t}_{n+1}} = \underbrace{\begin{bmatrix} \mathcal{F}_n^{\mathbf{x},\mathbf{x}} & \mathcal{F}_n^{\mathbf{x},\mathbf{r}} & \mathcal{F}_n^{\mathbf{x},\mathbf{y}} \\ \mathcal{F}_n^{\mathbf{r},\mathbf{x}} & \mathcal{F}_n^{\mathbf{r},\mathbf{r}} & \mathcal{F}_n^{\mathbf{r},\mathbf{y}} \\ \mathcal{F}_n^{\mathbf{y},\mathbf{x}} & \mathcal{F}_n^{\mathbf{y},\mathbf{r}} & \mathcal{F}_n^{\mathbf{y},\mathbf{y}} \end{bmatrix}}_{\mathcal{F}_n} \begin{bmatrix} \mathbf{x}_n \\ \mathbf{r}_n \\ \mathbf{y}_{n-1} \end{bmatrix} + \underbrace{\begin{bmatrix} \mathbf{w}_n^{\mathbf{x}} \\ \mathbf{w}_n^{\mathbf{r}} \\ \mathbf{w}_n^{\mathbf{y}} \end{bmatrix}}_{\mathbf{w}_n}, \quad (2)$$

in which $\mathbf{w} = \{\mathbf{w}_n\}_{n\in\mathbb{N}}$ is a zero-mean process which is independent and independent of $\mathbf{t}_0 = [\mathbf{x}_0^T, \mathbf{r}_0^T, \mathbf{0}^T]^T$. We assume that matrix $\mathcal{F}_n$ is known.

2.2. Some particular cases

Let us first see that some classical and widely used models are particular linear TMC; they differ from one another by the matrices $\mathcal{F}_n$ (some submatrices of which are equal to zero); by the physical meaning of the additional process $\mathbf{r} = \{\mathbf{r}_n\}_{n\in\mathbb{N}}$; and/or by independence assumptions among subvectors of $\mathbf{w}_n$.

2.2.1. Linear HMC

The standard state-space model is a very particular linear TMC, since (2) reduces to (1) if we assume that matrices $\mathcal{F}_n^{\mathbf{x},\mathbf{r}}$, $\mathcal{F}_n^{\mathbf{x},\mathbf{y}}$, $\mathcal{F}_n^{\mathbf{r},\mathbf{x}}$, $\mathcal{F}_n^{\mathbf{r},\mathbf{y}}$, $\mathcal{F}_n^{\mathbf{y},\mathbf{r}}$ and $\mathcal{F}_n^{\mathbf{y},\mathbf{y}}$ are all equal to zero, and that $\{\mathbf{w}_n^{\mathbf{x}}\}$, $\{\mathbf{w}_n^{\mathbf{r}}\}$ and $\{\mathbf{w}_n^{\mathbf{y}}\}$ are independent.

2.2.2. Autoregressive process noise

The case where in (1) $\{\mathbf{u_n}\}$ becomes an MC has been introduced in [4] (see also [5]); this model can be written as

$$\begin{bmatrix} \mathbf{x}_{n+1} \\ \mathbf{u}_{n+1} \\ \mathbf{y}_n \end{bmatrix} = \begin{bmatrix} \mathbf{F}_n & \mathbf{G}_n & \mathbf{0}_{n_\mathbf{x}\times n_\mathbf{y}} \\ \mathbf{0}_{n_\mathbf{u}\times n_\mathbf{x}} & \mathbf{A}_n^{\mathbf{u}} & \mathbf{0}_{n_\mathbf{u}\times n_\mathbf{y}} \\ \mathbf{H}_n & \mathbf{0}_{n_\mathbf{y}\times n_\mathbf{u}} & \mathbf{0}_{n_\mathbf{y}\times n_\mathbf{y}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_n \\ \mathbf{u}_n \\ \mathbf{y}_{n-1} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{n_\mathbf{x}\times 1} \\ \xi_n^{\mathbf{u}} \\ \mathbf{J}_n\mathbf{v}_n \end{bmatrix}. \quad (3)$$

2.2.3. Autoregressive measurement noise

The case where in (1) $\{\mathbf{v_n}\}$ becomes an MC has been first addressed in [6], then generalized in [7] (see also [4] and [5]). This model is widely used in a lot of applications, and in particular in speech enhancement and coding, see e.g. [8]; it can be rewritten as

$$\begin{bmatrix} \mathbf{x}_{n+1} \\ \mathbf{v}_{n+1} \\ \mathbf{y}_n \end{bmatrix} = \begin{bmatrix} \mathbf{F}_n & \mathbf{0}_{n_\mathbf{x}\times n_\mathbf{v}} & \mathbf{0}_{n_\mathbf{x}\times n_\mathbf{y}} \\ \mathbf{0}_{n_\mathbf{v}\times n_\mathbf{x}} & \mathbf{A}_n^{\mathbf{v}} & \mathbf{0}_{n_\mathbf{v}\times n_\mathbf{y}} \\ \mathbf{H}_n & \mathbf{J}_n & \mathbf{0}_{n_\mathbf{y}\times n_\mathbf{y}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_n \\ \mathbf{v}_n \\ \mathbf{y}_{n-1} \end{bmatrix} + \begin{bmatrix} \mathbf{G}_n\mathbf{u}_n \\ \xi_n^{\mathbf{v}} \\ \mathbf{0}_{n_\mathbf{y}\times 1} \end{bmatrix}. \quad (4)$$

2.2.4. Autoregressive model noise

Sorenson [9] introduced a model which extends the two previous ones by assuming that $\{\mathbf{u}_n\}$ and $\{\mathbf{v}_n\}$ are simultaneously Markovian (but still independent); see also [10] for a full algorithmic treatment and applications to radar tracking. This model can be further generalized by assuming that

$$\underbrace{\begin{bmatrix} \mathbf{u}_{n+1} \\ \mathbf{v}_{n+1} \end{bmatrix}}_{\mathbf{n}_{n+1}} = \underbrace{\begin{bmatrix} \mathbf{A}_n^{\mathbf{u}} & \mathbf{A}_n^{\mathbf{u},\mathbf{v}} \\ \mathbf{A}_n^{\mathbf{v},\mathbf{u}} & \mathbf{A}_n^{\mathbf{v}} \end{bmatrix}}_{\mathbf{A}_n} \begin{bmatrix} \mathbf{u}_n \\ \mathbf{v}_n \end{bmatrix} + \underbrace{\begin{bmatrix} \xi_n^{\mathbf{u}} \\ \xi_n^{\mathbf{v}} \end{bmatrix}}_{\xi_n}, \quad (5)$$

where $\xi = \{\xi_n\}_{n\in\mathbb{N}}$ is zero mean, independent and independent of $\mathbf{n}_0$. The associated triplet model is

$$\begin{bmatrix} \mathbf{x}_{n+1} \\ \mathbf{n}_{n+1} \\ \mathbf{y}_n \end{bmatrix} = \begin{bmatrix} \mathbf{F}_n & \overline{\mathbf{G}}_n & \mathbf{0}_{n_\mathbf{x}\times n_\mathbf{y}} \\ \mathbf{0}_{n_\mathbf{n}\times n_\mathbf{x}} & \mathbf{A}_n & \mathbf{0}_{n_\mathbf{n}\times n_\mathbf{y}} \\ \mathbf{H}_n & \overline{\mathbf{J}}_n & \mathbf{0}_{n_\mathbf{y}\times n_\mathbf{y}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_n \\ \mathbf{n}_n \\ \mathbf{y}_{n-1} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{n_\mathbf{x}\times 1} \\ \xi_n \\ \mathbf{0}_{n_\mathbf{y}\times 1} \end{bmatrix}, \quad (6)$$

with $\overline{\mathbf{G}}_n = [\mathbf{G}_n, \mathbf{0}_{n_\mathbf{x}\times n_\mathbf{v}}]$ and $\overline{\mathbf{J}}_n = [\mathbf{0}_{n_\mathbf{y}\times n_\mathbf{u}}, \mathbf{J}_n]$; it reduces to the model introduced by Sorenson if $\mathbf{A}_n^{(\mathbf{u},\mathbf{v})} = \mathbf{0}_{n_\mathbf{u}\times n_\mathbf{v}}$, $\mathbf{A}_n^{(\mathbf{v},\mathbf{u})} = \mathbf{0}_{n_\mathbf{v}\times n_\mathbf{u}}$ and $\{\xi_n^{\mathbf{u}}\}$ and $\{\xi_n^{\mathbf{v}}\}$ are independent.

2.2.5. Linear PMC model and its extensions

The PMC model introduced in [2] reads

$$\underbrace{\begin{bmatrix} \mathbf{x}_{n+1} \\ \mathbf{y}_n \end{bmatrix}}_{\mathbf{Z}_{n+1}} = \underbrace{\begin{bmatrix} \mathbf{F}_n^1 & \mathbf{F}_n^2 \\ \mathbf{H}_n^1 & \mathbf{H}_n^2 \end{bmatrix}}_{\mathbf{F}_n} \begin{bmatrix} \mathbf{x}_n \\ \mathbf{y}_{n-1} \end{bmatrix} + \underbrace{\begin{bmatrix} \mathbf{G}_n^{11} & \mathbf{G}_n^{12} \\ \mathbf{G}_n^{21} & \mathbf{G}_n^{22} \end{bmatrix}}_{\mathbf{G}_n} \underbrace{\begin{bmatrix} \mathbf{u}_n \\ \mathbf{v}_n \end{bmatrix}}_{\mathbf{n}_n}, \quad (7)$$

where $\mathbf{n} = \{\mathbf{n}_n\}_{n\in\mathbb{N}}$ is a zero-mean process which is independent and independent of $\mathbf{x}_0$. This model can be seen as a linear TMC. If we now assume that $\mathbf{n}$ is Markovian, then the model becomes

$$\begin{bmatrix} \mathbf{x}_{n+1} \\ \mathbf{n}_{n+1} \\ \mathbf{y}_n \end{bmatrix} = \begin{bmatrix} \mathbf{F}_n^1 & \overline{\mathbf{G}}_n^1 & \mathbf{F}_n^2 \\ \mathbf{0}_{n_\mathbf{n}\times n_\mathbf{x}} & \mathbf{A}_n & \mathbf{0}_{n_\mathbf{n}\times n_\mathbf{y}} \\ \mathbf{H}_n^1 & \overline{\mathbf{G}}_n^2 & \mathbf{H}_n^2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_n \\ \mathbf{n}_n \\ \mathbf{y}_{n-1} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{n_\mathbf{x}\times 1} \\ \xi_n \\ \mathbf{0}_{n_\mathbf{y}\times 1} \end{bmatrix}, \quad (8)$$

with $\overline{\mathbf{G}}_n^1 = [\mathbf{G}_n^{11}, \mathbf{G}_n^{12}]$ and $\overline{\mathbf{G}}_n^2 = [\mathbf{G}_n^{21}, \mathbf{G}_n^{22}]$, which, again, is a particular linear TMC.

3. LGTMC : RESTORATION ALGORITHMS

The aim of this section is to derive an algorithm for computing recursively $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ in the case of an LGTMC.

Let us first gather the unobserved variables $\mathbf{x}_n$ and $\mathbf{r}_n$ into a commun vector $\mathbf{x}_n^* = [\mathbf{x}_n^T, \mathbf{r}_n^T]^T$. Then (2) can be rewritten more compactly as

$$\underbrace{\begin{bmatrix} \mathbf{x}_{n+1}^* \\ \mathbf{y}_n \end{bmatrix}}_{\mathbf{t}_{n+1}} = \underbrace{\begin{bmatrix} \mathcal{F}_n^{\mathbf{x}^*,\mathbf{x}^*} & \mathcal{F}_n^{\mathbf{x}^*,\mathbf{y}} \\ \mathcal{F}_n^{\mathbf{y},\mathbf{x}^*} & \mathcal{F}_n^{\mathbf{y},\mathbf{y}} \end{bmatrix}}_{\mathcal{F}_n} \begin{bmatrix} \mathbf{x}_n^* \\ \mathbf{y}_{n-1} \end{bmatrix} + \underbrace{\begin{bmatrix} \mathbf{w}_n^{\mathbf{x}^*} \\ \mathbf{w}_n^{\mathbf{y}} \end{bmatrix}}_{\mathbf{w}_n}. \quad (9)$$

Let

$$E(\mathbf{w}_n \mathbf{w}_m^T) = \begin{bmatrix} \mathcal{Q}_n^{\mathbf{x}^*,\mathbf{x}^*} & \mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}} \\ \mathcal{Q}_n^{\mathbf{y},\mathbf{x}^*} & \mathcal{Q}_n^{\mathbf{y},\mathbf{y}} \end{bmatrix} \delta_{n,m} = \mathcal{Q}_n \delta_{n,m} \ . \quad (10)$$

Model (2) is indeed a partially observed vector MC, in which we observe some components $\{\mathbf{y}_n\}$, and we want to restore (part of) the remaining ones $\{\mathbf{x}_n^*\}$; so our algorithm computes $p(\mathbf{x}_n^*|\mathbf{y}_{0:n})$, and next $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ is obtained by marginalization. Let us remark that $p(\mathbf{x}_n^*|\mathbf{y}_{0:n})$ can be computed efficiently for, even though TMC are not necessarily HMC (since $\mathbf{x}$ is not necessarily an MC), the conditional law of $\mathbf{x}^*$ given $\mathbf{y}$ is Markovian; this key computational property, in turn, enables the derivation of fast algorithms.

3.1. Regular LGTMC

Let us first address the case where $\mathcal{Q}_n^{\mathbf{y},\mathbf{y}}$ is positive definite. In this case a Kalman-like filtering algorithm has been proposed in [3]; it is recalled here for convenience of the reader.

Let $p(\mathbf{x}_0^*) \sim \mathcal{N}(\widehat{\mathbf{x}}_0^*, \mathbf{P}_0^*)$ and $p(\mathbf{w}_n) \sim \mathcal{N}(\mathbf{0}, \mathcal{Q}_n)$. Then $p(\mathbf{x}_n^*|\mathbf{y}_{0:n})$ and $p(\mathbf{x}_{n+1}^*|\mathbf{y}_{0:n})$ are Gaussian. Let

$$p(\mathbf{x}_n^*|\mathbf{y}_{0:n}) \sim \mathcal{N}(\widehat{\mathbf{x}}_{n|n}^*, \mathbf{P}_{n|n}^*), \quad (11)$$

$$p(\mathbf{x}_{n+1}^*|\mathbf{y}_{0:n}) \sim \mathcal{N}(\widehat{\mathbf{x}}_{n+1|n}^*, \mathbf{P}_{n+1|n}^*). \quad (12)$$

Then $\widehat{\mathbf{x}}_{n+1|n+1}^*$ and $\mathbf{P}_{n+1|n+1}^*$ can be computed from $\widehat{\mathbf{x}}_{n+1|n}^*$ and $\mathbf{P}_{n+1|n}^*$ via the following equations :

$$\begin{aligned}
\widehat{\mathbf{x}}_{n+1|n}^* &= [\mathcal{F}_n^{\mathbf{x}^*,\mathbf{x}^*} - \mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}}(\mathcal{Q}_n^{\mathbf{y},\mathbf{y}})^{-1}\mathcal{F}_n^{\mathbf{y},\mathbf{x}^*}]\, \widehat{\mathbf{x}}_{n|n}^* \\
&+ \mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}}(\mathcal{Q}_n^{\mathbf{y},\mathbf{y}})^{-1}\mathbf{y}_n \\
&+ [\mathcal{F}_n^{\mathbf{x}^*,\mathbf{y}} - \mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}}(\mathcal{Q}_n^{\mathbf{y},\mathbf{y}})^{-1}\mathcal{F}_n^{\mathbf{y},\mathbf{y}}]\, \mathbf{y}_{n-1}, \quad (13) \\
\mathbf{P}_{n+1|n}^* &= [\mathcal{Q}_n^{\mathbf{x}^*,\mathbf{x}^*} - \mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}}(\mathcal{Q}_n^{\mathbf{y},\mathbf{y}})^{-1}\mathcal{Q}_n^{\mathbf{y},\mathbf{x}^*}] \\
&+ [\mathcal{F}_n^{\mathbf{x}^*,\mathbf{x}^*} - \mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}}(\mathcal{Q}_n^{\mathbf{y},\mathbf{y}})^{-1}\mathcal{F}_n^{\mathbf{y},\mathbf{x}^*}] \times \mathbf{P}_{n|n}^* \times \\
&\quad [\mathcal{F}_n^{\mathbf{x}^*,\mathbf{x}^*} - \mathcal{Q}_n^{\mathbf{x}^*,\mathbf{y}}(\mathcal{Q}_n^{\mathbf{y},\mathbf{y}})^{-1}\mathcal{F}_n^{\mathbf{y},\mathbf{x}^*}]^T, \quad (14) \\
\mathbf{K}_{n+1|n+1}^* &= \mathbf{P}_{n+1|n}^*(\mathcal{F}_{n+1}^{\mathbf{y},\mathbf{x}^*})^T \\
&\times [\mathcal{Q}_{n+1}^{\mathbf{y},\mathbf{y}} + \mathcal{F}_{n+1}^{\mathbf{y},\mathbf{x}^*}\mathbf{P}_{n+1|n}^*(\mathcal{F}_{n+1}^{\mathbf{y},\mathbf{x}^*})^T]^{-1}, \quad (15) \\
\widehat{\mathbf{x}}_{n+1|n+1}^* &= \widehat{\mathbf{x}}_{n+1|n}^* + \mathbf{K}_{n+1|n+1}^* \\
&\times [\mathbf{y}_{n+1} - \mathcal{F}_{n+1}^{\mathbf{y},\mathbf{x}^*}\widehat{\mathbf{x}}_{n+1|n}^* - \mathcal{F}_{n+1}^{\mathbf{y},\mathbf{y}}\mathbf{y}_n], \quad (16) \\
\mathbf{P}_{n+1|n+1}^* &= \mathbf{P}_{n+1|n}^* - \mathbf{K}_{n+1|n+1}^*[\mathcal{Q}_{n+1}^{\mathbf{y},\mathbf{y}} + \\
&\quad \mathcal{F}_{n+1}^{\mathbf{y},\mathbf{x}^*}\mathbf{P}_{n+1|n}^*(\mathcal{F}_{n+1}^{\mathbf{y},\mathbf{x}^*})^T]\,(\mathbf{K}_{n+1|n+1}^*)^T. \quad (17)
\end{aligned}$$

3.2. Perfect measurement LGTMC

As we have seen in section 2.2, some useful models are particular linear TMC with perfect (i.e., unnoisy) measurements. In this section we thus address the restoration problem in case $\mathbf{w}_n^{\mathbf{y}} = \mathbf{0}_{n_{\mathbf{y}}\times 1}$. Adapting a classical method used in LGHMC, we shall first perform a state-space transformation in order to reduce the dimension of $\mathbf{x}_n^*$; we will then obtain a new stochastic linear dynamical system, and will propose an estimation algorithm for that system.

3.2.1. State-space transformation

Let us first consider the following alternate partition of $\mathbf{x}_n^*$:

$$\mathbf{x}_n^* = \begin{bmatrix} (\mathbf{x}_n)_{n_{\mathbf{x}}\times 1} \\ (\mathbf{r}_n)_{n_{\mathbf{r}}\times 1} \end{bmatrix} = \begin{bmatrix} (\overline{\mathbf{x}}_n)_{(n_{\mathbf{x}}+n_{\mathbf{r}}-n_{\mathbf{y}})\times 1} \\ (\overline{\mathbf{r}}_n)_{n_{\mathbf{y}}\times 1} \end{bmatrix}; \quad (18)$$

Let us partition $\mathcal{F}_n^{\mathbf{y},\mathbf{x}^*}$ as

$$\mathcal{F}_n^{\mathbf{y},\mathbf{x}^*} = \left[(\mathcal{F}_n^{\mathbf{y},\overline{\mathbf{x}}})_{n_{\mathbf{y}}\times n_{\bar{\mathbf{x}}}} \ , \ (\mathcal{F}_n^{\mathbf{y},\overline{\mathbf{r}}})_{n_{\mathbf{y}}\times n_{\mathbf{y}}} \right], \quad (19)$$

and let us assume that $\mathcal{F}_n^{\mathbf{y},\overline{\mathbf{r}}}$ is invertible. In this case the following transformation

$$\underbrace{\begin{bmatrix} \mathbf{I}_{n_{\mathbf{x}}^*-n_{\mathbf{y}}} & \mathbf{0}_{(n_{\mathbf{x}}^*-n_{\mathbf{y}})\times n_{\mathbf{y}}} \\ \mathcal{F}_n^{\mathbf{y},\overline{\mathbf{x}}} & \mathcal{F}_n^{\mathbf{y},\overline{\mathbf{r}}} \end{bmatrix}}_{\mathbf{T}_n} \underbrace{\begin{bmatrix} \overline{\mathbf{x}}_n \\ \overline{\mathbf{r}}_n \end{bmatrix}}_{\mathbf{x}_n^*} = \begin{bmatrix} \overline{\mathbf{x}}_n \\ \mathbf{y}_n - \mathcal{F}_n^{\mathbf{y},\mathbf{y}}\mathbf{y}_{n-1} \end{bmatrix} \quad (20)$$

is invertible, and thus defines the state-space transformation:

$$\underbrace{\begin{bmatrix} \mathbf{T}_{n+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_{\mathbf{y}}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{n+1}^* \\ \mathbf{y}_n \end{bmatrix}}_{\tilde{\mathbf{t}}_{n+1}} = \begin{bmatrix} \mathbf{T}_{n+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_{\mathbf{y}}} \end{bmatrix} \mathcal{F}_n \begin{bmatrix} \mathbf{T}_n^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_{\mathbf{y}}} \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{T}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_{\mathbf{y}}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_n^* \\ \mathbf{y}_{n-1} \end{bmatrix}}_{\tilde{\mathbf{t}}_n}$$

$$+ \begin{bmatrix} \mathbf{T}_{n+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_{\mathbf{y}}} \end{bmatrix} \begin{bmatrix} \mathbf{w}_n^{\mathbf{x}^*} \\ \mathbf{0}_{n_{\mathbf{y}}\times 1} \end{bmatrix}. \quad (21)$$

The first $n_{\mathbf{x}^*}$ equations read :

$$\begin{bmatrix} \overline{\mathbf{x}}_{n+1} \\ \mathbf{y}_{n+1} \end{bmatrix} = \begin{bmatrix} \overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\overline{\mathbf{x}}} & \overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\mathbf{y}} \\ \overline{\mathcal{F}}_n^{\mathbf{y},\overline{\mathbf{x}}} & \overline{\mathcal{F}}_n^{\mathbf{y},\mathbf{y}} \end{bmatrix} \begin{bmatrix} \overline{\mathbf{x}}_n \\ \mathbf{y}_n \end{bmatrix} + \begin{bmatrix} \overline{\mathcal{G}}_n^{\overline{\mathbf{x}}} \\ \overline{\mathcal{G}}_n^{\mathbf{y}} \end{bmatrix} \mathbf{y}_{n-1} + \overline{\mathbf{w}}_n, \quad (22)$$

in which

$$\begin{bmatrix} \overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\overline{\mathbf{x}}} & \overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\mathbf{y}} \\ \overline{\mathcal{F}}_n^{\mathbf{y},\overline{\mathbf{x}}} & \overline{\mathcal{F}}_n^{\mathbf{y},\mathbf{y}} \end{bmatrix} = \mathbf{T}_{n+1}\mathcal{F}_n^{\mathbf{x}^*,\mathbf{x}^*}\mathbf{T}_n^{-1} + \begin{bmatrix} \mathbf{0}_{n_{\overline{\mathbf{x}}}\times n_{\overline{\mathbf{x}}}} & \mathbf{0} \\ \mathbf{0} & \mathcal{F}_{n+1}^{\mathbf{y},\mathbf{y}} \end{bmatrix}, \quad (23)$$

$$\begin{bmatrix} \overline{\mathcal{G}}_n^{\overline{\mathbf{x}}} \\ \overline{\mathcal{G}}_n^{\mathbf{y}} \end{bmatrix} = \mathbf{T}_{n+1}\mathcal{F}_n^{\mathbf{x}^*,\mathbf{y}} - \begin{bmatrix} \overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\mathbf{y}} \\ \overline{\mathcal{F}}_n^{\mathbf{y},\mathbf{y}} - \mathcal{F}_{n+1}^{\mathbf{y},\mathbf{y}} \end{bmatrix} \mathcal{F}_n^{\mathbf{y},\mathbf{y}}, \quad (24)$$

$$\overline{\mathbf{w}}_n = \mathbf{T}_{n+1}\mathbf{w}_n^{\mathbf{x}^*}. \quad (25)$$

3.2.2. Restoration algorithm

Let us finally address the restoration of $\overline{\mathbf{x}}_n$ from $\{\mathbf{y}_{0:n}\}$ in system (22) (the subsequent restoration of $\mathbf{x}_n$ is immediate if $n_{\mathbf{r}} \geq n_{\mathbf{y}}$, since in this case $\mathbf{x}_n$ is a subvector of $\overline{\mathbf{x}}_n$; it can also be considered if $n_{\mathbf{r}} < n_{\mathbf{y}}$, but this point is omitted here due to lack of space). From (25) and (10), we get

$$\begin{aligned}
E(\overline{\mathbf{w}}_n \overline{\mathbf{w}}_m^T) &= \mathbf{T}_{n+1}\mathcal{Q}_n^{\mathbf{x}^*,\mathbf{x}^*}\mathbf{T}_{n+1}^T\delta_{n,m} \quad (26) \\
&= \begin{bmatrix} \overline{\mathcal{Q}}_n^{\overline{\mathbf{x}},\overline{\mathbf{x}}} & \overline{\mathcal{Q}}_n^{\overline{\mathbf{x}},\mathbf{y}} \\ \overline{\mathcal{Q}}_n^{\mathbf{y},\overline{\mathbf{x}}} & \overline{\mathcal{Q}}_n^{\mathbf{y},\mathbf{y}} \end{bmatrix} \delta_{n,m} = \overline{\mathcal{Q}}_n \, \delta_{n,m} \ .
\end{aligned}$$

Let $p(\overline{\mathbf{x}}_0) \sim \mathcal{N}(\widehat{\overline{\mathbf{x}}}_0, \overline{\mathbf{P}}_0)$ and $p(\overline{\mathbf{w}}_n) \sim \mathcal{N}(\mathbf{0}, \overline{\mathcal{Q}}_n)$. Then $p(\overline{\mathbf{x}}_n|\mathbf{y}_{0:n})$ and $p(\overline{\mathbf{x}}_{n+1}|\mathbf{y}_{0:n})$ are also Gaussian. Let

$$p(\overline{\mathbf{x}}_n|\mathbf{y}_{0:n}) \sim \mathcal{N}(\widehat{\overline{\mathbf{x}}}_{n|n}, \overline{\mathbf{P}}_{n|n}), \quad (27)$$

$$p(\overline{\mathbf{x}}_{n+1}|\mathbf{y}_{0:n}) \sim \mathcal{N}(\widehat{\overline{\mathbf{x}}}_{n+1|n}, \overline{\mathbf{P}}_{n+1|n}). \quad (28)$$

Then $\widehat{\overline{\mathbf{x}}}_{n+1|n+1}$ and $\overline{\mathbf{P}}_{n+1|n+1}$ can be calculated from $\widehat{\overline{\mathbf{x}}}_{n+1|n}$ and $\overline{\mathbf{P}}_{n+1|n}$ via the following equations[2] (the proof is omitted for want of space) :

$$\widehat{\overline{\mathbf{x}}}_{n+1|n} = \overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\overline{\mathbf{x}}} \widehat{\overline{\mathbf{x}}}_{n|n} + \overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\mathbf{y}} \mathbf{y}_n + \overline{\mathcal{G}}_n^{\overline{\mathbf{x}}} \mathbf{y}_{n-1}, \quad (29)$$

$$\overline{\mathbf{P}}_{n+1|n} = \overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\overline{\mathbf{x}}} \overline{\mathbf{P}}_{n|n} (\overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\overline{\mathbf{x}}})^T + \overline{\mathcal{Q}}_n^{\overline{\mathbf{x}},\overline{\mathbf{x}}}, \quad (30)$$

$$\widehat{\overline{\mathbf{y}}}_{n+1|n} = \overline{\mathcal{F}}_n^{\mathbf{y},\overline{\mathbf{x}}} \widehat{\overline{\mathbf{x}}}_{n|n} + \overline{\mathcal{F}}_n^{\mathbf{y},\mathbf{y}} \mathbf{y}_n + \overline{\mathcal{G}}_n^{\mathbf{y}} \mathbf{y}_{n-1}, \quad (31)$$

$$\overline{\mathbf{K}}_{n+1|n+1} = \overline{\mathcal{F}}_n^{\overline{\mathbf{x}},\overline{\mathbf{x}}} \overline{\mathbf{P}}_{n|n} (\overline{\mathcal{F}}_n^{\mathbf{y},\overline{\mathbf{x}}})^T + \overline{\mathcal{Q}}_n^{\overline{\mathbf{x}},\mathbf{y}}, \quad (32)$$

$$\overline{\mathbf{L}}_{n+1|n+1} = \overline{\mathcal{F}}_n^{\mathbf{y},\overline{\mathbf{x}}} \overline{\mathbf{P}}_{n|n} (\overline{\mathcal{F}}_n^{\mathbf{y},\overline{\mathbf{x}}})^T + \overline{\mathcal{Q}}_n^{\mathbf{y},\mathbf{y}}, \quad (33)$$

$$\widehat{\overline{\mathbf{x}}}_{n+1|n+1} = \widehat{\overline{\mathbf{x}}}_{n+1|n} + \overline{\mathbf{K}}_{n+1|n+1} \overline{\mathbf{L}}_{n+1|n+1}^{-1} (\mathbf{y}_{n+1} - \widehat{\overline{\mathbf{y}}}_{n+1|n}) \quad (34)$$

$$\overline{\mathbf{P}}_{n+1|n+1} = \overline{\mathbf{P}}_{n+1|n} - \overline{\mathbf{K}}_{n+1|n+1} \overline{\mathbf{L}}_{n+1|n+1}^{-1} \overline{\mathbf{K}}_{n+1|n+1}^T \quad (35)$$

4. APPLICATIONS

4.1. Speech Enhancement and Coding

As we have seen in section 2.2, the linear TMC model encompasses some classical models. It happens that the algorithm of section 3.2 also includes some classical algorithms as particular cases.

Let us for instance consider the case of a classical model with autoregressive measurement noise (we consider this example because of its wide applicability, in particular in speech enhancement and coding, see e.g. [8]). If the linear TMC reduces to (4), then $\overline{\mathcal{G}}_n^{\overline{\mathbf{x}}}$ and $\overline{\mathcal{G}}_n^{\mathbf{y}}$ vanish, so in (29) to (35) the dependency on $\mathbf{y}_{n-1}$ vanishes, and these equations reduce to equations of [8]. More precisely, equations (30), (32) and (34) reduce respectively to [8, eq. 51 p. 1736], [8, eq. 57 p. 1737] and [8, eq. 56 p. 1736]; while (29) (resp. (35)) reduces to an equation which can be obtained as part of [8, eq. 54 p. 1736] (resp. [8, eq. 52 p. 1736]).

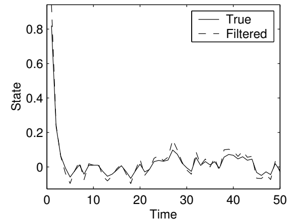
4.2. A numerical example

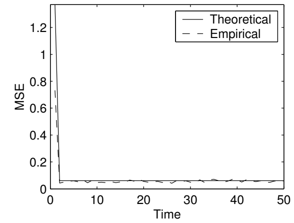
Let us finally provide a numerical example of a general LGTMC with perfect measurement. Let us set

$$\mathcal{F}_n = \begin{bmatrix} .12 & .10 & .11 \\ .11 & .10 & .12 \\ .10 & .11 & .12 \end{bmatrix}, \; \mathcal{Q}_n = \begin{bmatrix} .125 & .015 & 0 \\ .015 & .125 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and let $p(\mathbf{x}_0^*) \sim \mathcal{N}([0.5, 0.5]^T, 2.5 \; \mathbf{I}_2)$. The first figure shows the true and filtered states, and the second one the theoretical and empirical mean square errors; both figures are averaged over 100 realizations.

[2]inverses in (34) and (35) should be replaced by a generalized inverse if $\overline{\mathbf{L}}_{n+1|n+1}$ is not invertible.





5. CONCLUSION

The linear TMC model encompasses and generalizes some important extensions (colored process and/or measurement noise) of the standard state-space model. A restoration algorithm for general LGTMC with unnoisy measurements has been proposed; this algorithm is itself a generalization of some classical Kalman-like algorithms.

6. REFERENCES

[1] 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.

[2] W. Pieczynski and F. Desbouvries, "Kalman filtering using pairwise gaussian models," in *Proceedings of the ICASSP*, Hong-Kong, 2003.

[3] F. Desbouvries and W. Pieczynski, "Modèles de markov triplet et filtrage de kalman," *CRAS-Mathématiques*, vol. Vol. 336-8, pp. 667–670, 2003.

[4] A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communications and Control*, New York: McGraw-Hill, 1971.

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

[6] Jr. A. E. Bryson and L. J. Henrikson, "Estimation using sampled data containing sequentially correlated noise," *J. Spacecr. Rockets*, vol. 5, pp. 662–65, 1968.

[7] P. S. Maybeck, *Stochastic Models, Estimation and Control*, vol. 1, New York: Academic Press, 1979.

[8] J. D. Gibson, B. Koo, and S. D. Gray, "Filtering of colored noise for speech enhancement and coding," *IEEE Tr. Sig. Proc.*, vol. 39-8, pp. 1732–42, August 1991.

[9] H. W. Sorenson, "Kalman filtering techniques," in *Advances in Control Systems Theory and Appl.*, C. T. Leondes, Ed., vol. 3, pp. 219–92. Acad. Press, 1966.

[10] C.K. Chui and G. Chen, *Kalman Filtering with Real-Time Applications*, Berlin, DE: Springer, 1999.