TRIPLET PARTIALLY MARKOV CHAINS AND TREES

W. Pieczynski

GET/INT Département CITI, CNRS UMR 5157
9, rue Charles Fourier, 91000 Evry, France

ABSTRACT

Hidden Markov models (HMM), like chains or trees considered in this paper, are widely used in different situations. Such models, in which the hidden process X is a Markov one, allow one estimating the latter from an observed process Y, which can be seen as a noisy version of X. This is possible once the distribution of X conditional on Y is a Markov distribution. These models have been recently generalized to Pairwise Markov models (PMM), in which one assumes the markovianity of (X,Y), and Triplet Markov models (TMM), in which the distribution of (X,Y) is the marginal distribution of an Markov model (X,U,Y). In this paper we propose further generalization of TMM by considering that (X,U,Y) is a Markov model with respect to (X,U), but is not necessarily a Markov one with respect to Y. We show that in such models, called "partially Markov", classical restoration algorithms remain valid.

1. INTRODUCTION

Let S be a finite set and $X=(X_s)_{s\in S}$, $Y=(Y_s)_{s\in S}$ two stochastic processes. The problem is to estimate $X=x$ from $Y=y$. The set S can be seen as a set of nodes in a network, and thus $X=x$ is some "hidden" state of the network that we have to estimate from its "observed" state $Y=y$. When the distribution of $Z=(X,Y)$ is simple enough, $X=x$ can be estimated from $Y=y$ by some Bayesian methods. We deal in this paper with two particular cases of increasing generality : Markov chains and Markov trees. The simplest models, which are widely used in various situations, are Hidden Markov Chains (HMC) and Hidden Markov Trees (HMT). Recently they have been first generalized to Pairwise Markov Chains (PMC [2, 9]) and Pairwise Markov Trees (PMT [5, 6]), and then to Triplet Markov Chains (TMC [7]) and Triplet Markov Trees (TMT [10]). Roughly speaking, in a Triplet Markov Model (chain or tree), the distribution of $Z=(X,Y)$ is a marginal distribution of $T=(X,U,Y)$, which is assumed to be a Markov model and where U is a latent process. One can then show that when U is not too complex, classical calculus used in HMC and HMT can be adapted and thus different quantities of interest can be calculated. In particular, $p(x_s|y)$ is computable, which makes possible the application of the Bayesian Maximum; Posterior Mode (PMP) restoration method.

The aim of this paper is to introduce a further generalization of TMC and TMT, called Partially TMC (PTMC) and Partially TMT (PTMT). Roughly speaking, different calculus of interest can be performed once (X,U) is Markovian conditionally on $Y=y$. This is true when $T=(X,U,Y)$ is a Markov process [7]; however, the latter is not necessary and thus, relaxing it, we arrive to more general context.

2. TRIPLET PARTIALLY MARKOV CHAINS

2.1 Pairwise and Triplet Markov chains

Let us briefly recall Pairwise and Triplet Markov chains (PMC and TMC) models. Let $X=(X_1,...,X_n)$, $Y=(Y_1,...,Y_n)$ be two stochastic processes, where $X=x$ is not observable and has to be estimated from $Y=y$. In the classical HMC model the distribution of $Z=(X,Y)$ is $p(z)=p(x_1)p(x_2|x_1)...p(x_n|x_{n-1})p(y_1|x_1)...p(y_n|x_n)$, which means that X is a Markov chain with $p(x)=p(x_1)p(x_2|x_1)...p(x_n|x_{n-1})$, and $p(y|x)=p(y_1|x_1)...p(y_n|x_n)$. In PMC model one assume the markovianity $Z=(X,Y)$, which means that $p(z)=p(z_1)p(z_2|z_1)...p(z_n|z_{n-1})$. PMC is strictly more general that HMC because in PMC X is not necessarily a Markov process [9]. In TMC model one assumes that the ditribution of $Z=(X,Y)$ is the marginal distribution of a Markov chain $T=(X,U,Y)$. In an analogous manner that above, TMC is strictly more general that PMC because in TMC $Z=(X,Y)$ is not necessarily a Markov process [7].

PMC have been studied for discrete X [2] and continuous one [1, 8] as well, and the same has been done for TMC in [5, 7] and [4, 5], respectively.

2.2 Pairwise and Triplet Partially Markov Chains

Let us directly deal with Triplet Partially Markov chains, the Pairwise ones appearing as a particular case of the latter. So, let $T = (X, U, Y) = (X_1, U_1, Y_1, ..., X_n, U_n, Y_n)$ be a triplet process, where $X = (X_1, ..., X_n)$ is the hidden process we look for, $Y = (Y_1, ..., Y_n)$ is the observed process, and $U = (U_1, ..., U_n)$ is an auxiliary, possibly without physical existence, process. To fix things, we will assume that X_i takes its values in a finite set $\Omega = \{\omega_1, ..., \omega_k\}$, each U_i takes its values in a finite set $\Lambda = \{\lambda_1, ..., \lambda_m\}$, and each Y_i takes its values in R. Further, let us put $V = (X, U) = (X_1, U_1, ..., X_n, U_n)$.

Definition 1

The triplet process $T = (X, U, Y)$ will be called "Triplet Partially Markov Chain" (TPMC) if for each $1 \le i \le n-1$,

$$p(t_{i+1} | t_1, ..., t_i) = p(t_{i+1} | v_i, y_1, ..., y_i) \qquad (2.1)$$

We see how a TPMC is more general that a TMC: in the latter, we have $p(t_{i+1} | t_1, ..., t_i) = p(t_{i+1} | t_i)$.
To show that TPMC has good properties we will need the following lemma.

Lemma 1

Let $W = (W_1, ..., W_n)$ be a process with each W_i taking its values in a finite set. W is a Markov chain if and only if there exist positive functions q_2, …, q_n such that

$$p(w_1, ..., w_n) \propto q_2(w_1, w_2) ... q_n(w_{n-1}, w_n) \qquad (2.2)$$

Further, q_2, …, q_n define the transitions $p(w_{i+1} | w_i)$ and the marginals $p(w_i)$ via the following "forward" and "bacward" quantities, calculated recursively with :

$$\alpha_1(w_1) = 1,\ \alpha_{i+1}(w_{i+1}) = \sum_{w_{i+1}} q_{i+1}(w_i, w_{i+1}) \alpha_i(w_i)\ ; \qquad (2.3)$$

$$\beta_n(w_n) = 1,\ \beta_i(w_i) = \sum_{w_{i+1}} q_{i+1}(w_i, w_{i+1}) \beta_{i+1}(w_{i+1})\,. \qquad (2.4)$$

We have

$$p(w_{i+1} | w_i) = q_{i+1}(w_i, w_{i+1}) \beta_{i+1}(w_{i+1}) / \beta_i(w_i) \qquad (2.5)$$

and

$$p(w_i) = \alpha_i(w_i) \beta_i(w_i) / \sum_{w_i'} \alpha_i(w_i') \beta_i(w_i') \qquad (2.5)$$

This lemma, whose proof is quite classical, allow us to state the following proposition

Proposition 1

Let $T = (X, U, Y)$ be a TPMC verifying (2.1). Then $V = (X, U)$ is a Markov chain conditionally on Y. Further, the transitions $p(v_{i+1} | v_i, y)$ and the marginal distributions $p(v_i | y)$ are computable.

Proof. According to (2.1) we can write

$$p(t_1, ..., t_n) = p(t_1) p(t_2 | t_1) p(t_3 | t_1, .t_2,) ... p(t_n | t_1, ..., t_{n-1}) =$$

$$p(v_1, y_1, v_2, y_2) p(v_3, y_3 | v_2, y_1, y_2) ... p(v_n, y_n | v_{n-1}, y_1, ..., y_{n-1})$$

On the one hand, we have $p(v | y) \propto p(t_1, ..., t_n)$ and, on the other hand, putting

$q_2(v_1, v_2) = p(v_1, y_1, v_2, y_2)$,

$q_3(v_2, v_3) = p(v_3, y_3 | v_2, y_1, y_2)$, …,

$q_n(v_{n-1}, v_n) = p(v_n, y_n | v_{n-1}, y_1, ..., y_{n-1})$

we have

$p(t_1, ..., t_n) = q_2(v_1, v_2) q_3(v_2, v_3) ... q_n(v_{n-1}, v_n)$.

So $p(v | y) \propto q_2(v_1, v_2) q_3(v_2, v_3) ... q_n(v_{n-1}, v_n)$, and thus it is a Markov chain, with calculable transitions and marginals, from Lemma 1.

Remarks

1. According to (2.5) and the definition of $q_{i+1}(v_i, v_{i+1})$ above we see that in a TPMC the transition $p(v_{i+1} | v_i, y)$ depends on all $y_1, ..., y_n$. In classical HMC, PMC and TMC it depends on $y_i, ..., y_n$, but does not depend on $y_1, ..., y_{i-1}$.

2. Pairwise Partially Markov chain (PPMC) is a particular case of TPMC in which $X = U$ (this means that $V = X$ and thus there is no latent process).

3. If we consider a PPMC (in which (2.1) is replaced by $p(z_{i+1} | z_1, ..., z_i) = p(z_{i+1} | x_i, y_{i-r}, ..., y_i)$; recall that $V = X$ and $Z = (X, Y)$), we find a particular case of the high-

order hidden Markov chains, successfully applied in image segmentation in [3].

4. The continuous state case, where $V_i = (U_i, X_i)$ take their values in R^N and Y_i take their values in R^q, could also be considered. For example, the recursive calculus of $p(x_n | y_1, ..., y_n)$ (either Kalman or particle filtering) is possible in PMC [1, 8] and TMC [4] as well. So, these different methods would be generalized to TPMC in an analogous manner that the generalization described above in the case of discrete hidden state.

3. TRIPLET PARTIALLY MARKOV TREES

Let S be a finite set of points and $X = (X_s)_{s \in S}$, $Y = (Y_s)_{s \in S}$ two stochastic processes indexed on S. Each X_s takes its values in a finite set $\Omega = \{\omega_1, ..., \omega_k\}$, and each Y_s takes its values in the set of real numbers R. Let S^1, …, S^n be a partition of S representing different « generations ». Each $s \in S^i$ admits $s^+ \subset S^{i+1}$ (called his « children ») in such a way that every element of $t \in S^{i+1}$ has a unique « parent » $t^- \in S^i$. We assume that S^1 is a singleton (its element s_r is called « root »). Then we have four models with increasing generality [4, 5, 10]:

(i) the classical Hidden Markov Tree (see [11] for a rich bibliography) with independent noise (HMT-IN), in which $p(x) = p(x_{s_r}) \prod_{s \in S - S^1} p(x_s | x_{s^-})$ (which means that X is a Markov tree), and $p(y|x) = \prod_{s \in S} p(y_s | x_s)$. Thus

$$p(z) = p(x_{s_r}) p(y_{s_r} | x_{s_r}) \prod_{s \in S - S^1} p(x_s | x_{s^-}) p(y_s | x_s) ; \tag{3.1}$$

(ii) the Hidden Markov Tree (HMT), in which X is a Markov tree as above and the pairwise process $Z = (Z_s)_{s \in S}$, where $Z_s = (X_s, Y_s)$, is a Pairwise Markov Tree (PMT) [5], which means that its distribution verifies

$$p(z) = p(z_{s_r}) \prod_{s \in S - S^1} p(z_s | z_{s^-}) ; \tag{3.2}$$

(iii) the PMT $Z = (Z_s)_{s \in S}$ verifying (3.2);

(iv) the Triplet Markov Trees (TMT [7]), in which one introduces a latent variable $U = (U_s)_{s \in S}$ and assumes that the triplet $T = (X, U, Y)$ is a Markov tree (i.e., verifies (3.2) with $t = (x, u, y)$ instead of $z = (x, y)$).

Remark

Let us remark that the greater generality of PMC with respect to HMT-IN appears locally at the transition probability level. In fact, as $p(z_s | z_{s^-})$ in (3.2) can be written $p(z_s | z_{s^-}) = p(x_s, y_s | x_{s^-}, y_{s^-}) =$ $= p(x_s | x_{s^-}, y_{s^-}) p(y_s | x_s, x_{s^-}, y_{s^-})$, we see that HMT-IN is a PMT such that $p(x_s | x_{s^-}, y_{s^-}) = p(x_s | x_{s^-})$ and $p(y_s | x_s, x_{s^-}, y_{s^-}) = p(y_s | x_s)$.

As above, we will directly consider TMT and generalize them to Triplet Partially Markov Trees (TPMT). As above, we will assume that each U_i takes its values in a finite set $\Lambda = \{\lambda_1, ..., \lambda_m\}$.

Definition 2

The triplet process $T = (X, U, Y) = (X_s, U_s, Y_s)_{s \in S}$ will be called "Triplet Partially Markov Tree" (TPMT) if for each $1 \le i \le n-1$ we have:

$$\begin{gathered} p(t^{i+1} | t^1, ..., t^i) = p(t^{i+1} | v^i, y^1, ..., y^i) \\ p(t^{i+1} | v^i, y^1, ..., y^i)) = \prod_{s \in S^{i+1}} p(v_s, y^{i+1} | v_{s^-}, y^1, ..., y^i) \end{gathered} \tag{3.3}$$

We see how a TPMT generalizes TMT: in the latter, we have $p(t^{i+1} | t^1, ..., t^i) = p(t^{i+1} | v^i, y^i)$), and $p(t^{i+1} | v^i, y^i)) = \prod_{s \in S^{i+1}} p(v_s, y_s | v_{s^-}, y_{s^-})$.

As for TPMC above, we have the following lemma.

Lemma 2

Let S be a finite set structured as above and let $W = (W^1, ..., W^n)$ be a process with each $W^i = (W_s^i)_{s \in S^i}$, and each W_s^i taking its values in a finite set. W is a Markov tree if and only if there exist positive functions q_2, …, q_n such that

$$p(w^1, ..., w^n) \propto q_2(w^1, w^2) ... q_n(w^{n-1}, w^n) \text{, with} \tag{3.4}$$

$$q_{i+1}(w^i, w^{i+1}) = \prod_{s \in S^i, t \in s^=} q_{i+1}(w_s^i, w_t^{i+1}) \tag{3.5}$$

Further, the transitions $p(w_s | w_{s^-})$ and the marginals $p(w_s)$ can be calculated from q_2, .., q_n in the following way.

Let $\beta(w_s) = 1$ for $s^+ = \varnothing$ ($s \in S^n$), and

$$\beta(w_s) = \prod_{t \in s^+} (\sum_{w_t} q(w_s, w_t) \beta(w_t) \text{ for } s^+ \neq \varnothing \qquad (3.6)$$

The transitions $p(w_s | w_{s^-})$ are then given by

$$p(w_s | w_{s^-}) = \frac{q(w_{s^-}, w_s) \beta(w_s)}{\sum_{w_s} q(w_{s^-}, w_s) \beta(w_s)} \qquad (3.7)$$

Otherwise, the marginals $p(w_s)$ are given by

$$p(w_s) = \frac{\alpha(w_s) \beta(w_s)}{\sum_{w_s} \alpha(w_s) \beta(w_s)} \qquad (3.8)$$

where $\alpha(w_s)$ is calculated by (2.4), using the unique sequence of nodes s_1, .., s_j such that $s_j = s$, .., $s_{i-1} = s_i^-$, .., $s_1 = s_r$ (the sequence is the sequence of successive parents, until the root).

Proposition 2

Let $T = (X, U, Y)$ be a TPMT verifying (3.3). Then $V = (X, U)$ is a Markov tree conditionally on Y. Further, the transitions $p(v_{i+1} | v_i, y)$ and the marginal distributions $p(v_i | y)$ are computable.

As for Proposition 1, the proof uses Lemma 2.

4. CONCLUSIONS

We proposed in this paper two new models, called Triplet Partially Markov Chains (TPMC) and Triplet Partially Markov Trees (TPMT). More general that the Triplet Markov Chains (TMC) and Triplet Markov Trees (TMT), they still allow one to estimate the hidden state from the observed one.

As further research we may mention the possibilities of extending the proposed models to more complex Markov Graphical models, with the associated methods of hidden process restoration.

5. REFERENCES

[1] N. Caylus, A. Guyader, and F. LeGland, Particle filters for partially observed Markov chains, *IEEE Workshop on Statistical Signal Processing (SSP 2003)*, Saint Louis, Missouri, September, 28-October 1, 2003.

[2] S. Derrode and W. Pieczynski, SAR image segmentation using generalized Pairwise Markov Chains, *SPIE's International Symposium on Remote Sensing*, September 22-27, Crete, Greece, 2002.

[3] S. Derrode, C. Carincotte, and S. Bourennane, Unsupervised image segmentation based on high-order hidden Markov chains, *International Conference on Acoustics, Speech and Signal Processing* (*ICASSP 04)*, Montréal, Canada, 2004.

[4] F. Desbouvries and W. Pieczynski, Particle Filtering in Pairwise and Triplet Markov Chains, *Proceedings of the IEEE – EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP 2003)*, Grado-Gorizia, Italy, June 8-11, 2003.

[5] E. Monfrini, J. Lecomte, F. Desbouvries, and W. Pieczynski, Image and Signal Restoration using Pairwise Markov Trees, *IEEE Workshop on Statistical Signal Processing (SSP 2003)*, Saint Louis, Missouri, September, 28-October 1, 2003.

[6] W. Pieczynski, Arbres de Markov Couple – Pairwise Markov Trees, *CRAS – Mathématique*, Paris, Ser. I 335, pp. 79-82, 2002.

[7] W. Pieczynski, Chaînes de Markov Triplet *CRAS – Mathématique*, Série I, Vol. 335, Issue 3, pp. 275-278, 2002.

[8] W. Pieczynski and F. Desbouvries, Kalman Filtering using Pairwise Gaussian Models, *International Conference on Acoustics, Speech and Signal Processing (ICASSP 03)*, Hong-Kong, April 2003.

[9] W. Pieczynski, Pairwise Markov chains, *IEEE Trans. on PAMI*, Vol. 25, No. 5, pp. 634-639, 2003.

[10] W. Pieczynski, Arbres de Markov Triplet et fusion de Dempster-Shafer, *CRAS – Mathématique*, Série I, Vol. 336, Issue 10, pp. 869-872, 2003.

[11] A. S. Willsky, Multiresolution Markov models for signal and image processing, *Proceedings of IEEE*, Vol. 90, No. 8, pp. 1396-1458, 2002.