Markov, Semi-Markov Models and Associated Fields (from Theory to Application and back)

1-4 Jul 2025
Paris
France

Table of contents

Piecewise Deterministic Markov Processes and Bacterial Growth

Bertrand Cloez, Benoîte de Saporta , ***Nathalie Krell*** *,Tristan Roget,*
Emeline Schmisser

Abstract

This presentation will focus on the study of piecewise deterministic Markov process (PDMP) models used to describe dynamic phenomena involving occasional random events. In the first part, I will present work done in collaboration with Emeline Schmisser [6] on the nonparametric estimation of the jump rate of a deterministic PDMP based on observations of the embedded Markov chain (Z_k). We construct an adaptive estimator of the jump rate λ and establish an L^2 norm risk bound as well as a quasi-minimax result.

In a second part, I will focus on a bacterial growth model based on a PDMP-type multi-type branching process. Each cell grows exponentially at an individual rate, then divides at a rate depending on its size. This work extends a model introduced by Doumic, Hoffmann, Krell, and Robert [2] by integrating two cell types according to pole position. We show that the process is well defined, satisfies a "many-to-one" formula, and that its average empirical measure converges to a growth-fragmentation equation with multiple state variables. I will conclude with perspectives on estimating the division rate in this framework, in collaboration with Benoîte de Saporta, Bertrand Cloez, and Tristan Roget [1]. To this end, I will rely on a proceedings paper [4] in which I constructed a two-type branching process to model bacterial growth.

References

[1] B. Cloez, B. de Saporta, N. Krell and T. Roger. *Investigation of asymmetry in E. coli growth rate.* Work in progress.

[2] Doumic, M., Homann.,M., Krell, N. and Robert, L. (2015) *Statistical inference across scales for size-strutured models under growth variability.* Bernoulli, 21, 1760-1799.

[3] Doumic, M., Homann.,M., Krell, N., Robert, L., Aymerich S. and Robert J. (2014) *Division Control in Escherichia coli is Based on a Size-sensing rather than Timing Mechanism.* BMC Biology 12:17, 2014.

[4] N. Krell 2025 *Branching processes and bacterial growth.* To appear Proceedings IWBPA24.

[5] N. Krell (2016) *Statistical estimation of the jump rate of a special class of PDMPs. ESAIM: PS 20 196–216.*

[6] N. Krell , and E. Schmisser (2021) *Nonparametric estimation of jump rates for a specific class of piecewise deterministic Markov processes. Bernoulli, 27(4):2362–2388.*

Dmitrii Silvestrov, Stockholm University, Mälardalen University (Sweden)

Title. Limit and Ergodic Theorems for Perturbed Semi-Markov-Type Processes

Abstract.

The lecture aims to present and comment on the results of the recent books on perturbed semi-Markov-type processes and their applications:

[1] Silvestrov, D. (2025). Coupling and Ergodic Theorems for Semi-Markov-Type Processes II: Semi-Markov Processes and Multi-Alternating Regenerative Processes with Semi-Markov Modulation. Springer, Cham, xiv+550 pp.
[2] Silvestrov, D. (2025). Coupling and Ergodic Theorems for Semi-Markov-Type Processes I: Markov Chains, Renewal and Regenerative Processes. Springer, Cham, xiv+590 pp.
[3] Silvestrov, D. (2022). Perturbed Semi-Markov Type Processes II: Ergodic Theorems for Multi-Alternating Regenerative Processes. Springer, Cham, xvii+413 pp.
[4] Silvestrov, D. (2022). Perturbed Semi-Markov Type Processes I: Limit Theorems for Rare-Event Times and Processes. Springer, Cham, xvii+401 pp.
[5] Silvestrov, D., Silvestrov, S. (2017). Nonlinearly Perturbed Semi-Markov Processes. Springer Briefs in Probability and Mathematical Statistics, Springer, Cham, xiv+143 pp.

Multichain Hidden Markov and Semi-Markov Models: Formalization, Inference, and Applications

Peyrard N., Bacave H., Durand J.-B., Franc A., Plancade S., Sabbadin R.

Hidden Markov Models (HMMs) and Hidden Semi-Markov Models (HSMMs) are widely used statistical models for studying dynamic processes that cannot be observed directly or governed by a hidden layer. In many applications, particularly those involving spatial data, the hidden chain and the observed time series are multidimensional, exhibiting structured dependencies between variables at time t and variables at time $t+1$. In this talk, we introduce the framework of multichain HMMs, which provides a unified approach to existing models in the literature and also allows for their generalization. We demonstrate the utility of these models in ecological dynamics modeling. We then discuss inference for multichain HMMs within the context of the Expectation-Maximization (EM) algorithm. We explain why exact inference remains tractable for certain structures of multichain HMMs, while it is computationally intractable for others. Finally, we consider an extension to the semi-Markov case and propose the first rigorous definition of a multichain HSMM.

Anomaly Detection based on Markov Data:
A Statistical Depth Approach

Carlos A. Fernández[1] and Stephan Clémençon[1]

[1] *LTCI, Télécom Paris, Institut Polytechnique de Paris, France*

Abstract

The purpose of this article is to extend the notion of statistical depth to the case of sample paths of a Markov chain. Initially introduced to define a center-outward ordering of points in the support of a multivariate distribution, depth functions permit to generalize the notions of quantiles and (signed) ranks for observations in $\mathbb{R}^d$ with $d > 1$, as well as statistical procedures based on such quantities. Here we develop a general theoretical framework for evaluating the depth of a Markov sample path and recovering it statistically from an estimate of its transition probability with (non-) asymptotic guarantees. We also detail some of its applications, focusing particularly on unsupervised anomaly detection. Beyond the theoretical analysis carried out, numerical experiments are displayed, providing empirical evidence of the relevance of the novel concept we introduce here to quantify the degree of abnormality of Markov paths of variable length.

Keywords: statistical depth functions, anomaly detection, non-parametric statistic, Markov chains

References

Carlos A. Fernández and Stephan Clémençon. Anomaly detection based on markov data: A statistical depth approach. Submitted to LOD 2025, 2025. URL https://arxiv.org/abs/2406.16759.

Non-linear monotone models for general Harris recurrent Markov chains

Patrice Bertail[1], Cécile Durot[1], and Carlos A. Fernández[2]

[2]*MODAL'X, UMR CNRS 9023, Université Paris Nanterre, France*
[2]*LTCI, Télécom Paris, Institut Polytechnique de Paris, France*

Abstract

In this paper, we study a nonlinear cointegration-type model of the form $Z_t = f_0(X_t) + W_t$ where f_0 is a continuous monotone function and X_t is a Harris recurrent Markov chain, which may be either positive or β-null recurrent. We use a nonparametric Least Square Estimator based on the concave majorant to locally estimate f_0. Under mild conditions, we show its strong consistency and obtain its rate of convergence. In the positive recurrent case, the rate is similar to the optimal one obtained in the i.i.d case that is of order $n^{1/3}$. In the β-recurrent case, the rate of convergence is essentially linked to the local properties of the β -null recurrent chain X_t and typically of the order $n^{\beta/3}$, up to some slowly varying function. The main advantage of our estimator is that we do not need any smoothing parameters, neither in a stationary nor in a non-stationary context and the rate does not depend on any additional smoothness assumptions. New results (of the Glivenko-Cantelli type) for localized null recurrent Markov chains are also proved.

Keywords: monotone regression, isotonic regression, nonlinear cointegration, nonparametric estimation, null recurrent Markov chains

References

Patrice Bertail, Cécile Durot, and Carlos A. Fernández. Harris Recurrent Markov Chains and Nonlinear Monotone Cointegrated Models. 2024. URL https://arxiv.org/abs/2407.05294.

Bayes-Adaptive Impulse Control of Piecewise-Deterministic Markov Processes

Alice Cleynen[1,2], Benoîte de Saporta[1], Orlane Rossini[1], Régis Sabbadin[3], and Meritxell Vinyals[3]

[1]IMAG, Univ Montpellier, CNRS, Montpellier, France
[2]John Curtin School of Medical Research, Australian National University, Canberra, ACT, Australia
[3]Univ Toulouse, INRAE-MIAT, Toulouse, France

March 20, 2025

1 Abstract

In this paper, we propose a Bayes-adaptive approach for controlling Piecewise-Deterministic Markov Processes (PDMPs) under partial observability and incomplete model knowledge. PDMPs are a form of continuous time (semi-)Markov processes which enable the description of hybrid (discrete-continuous) process dynamics. They allow for the modeling of very general dynamics with a minimal set of interpretable parameters. When some parameters of the controlled PDMP are poorly known, we demonstrate that a Bayes-adaptive approach can provide a *learning while managing* control method. To achieve this, we show how to formulate the problem as a Bayes-Adaptive Partially Observable Markov Decision Process (BAPOMDP), which itself can be modeled as a higher-dimensional Partially Observable Markov Decision Process (POMDP). Subsequently, deep reinforcement learning algorithms can be employed to solve the resulting problem off-line. We illustrate the various steps of our approach with a medical patient follow-up application and utilize the Proximal Policy Optimization (PPO) algorithm to solve the final POMDP model.

Discrete-time Markov Switching Hawkes Process

Anna Bonnet, Stéphane Robin

Sorbonne Université, Laboratoire de Probabilités, Statistique et Modélisation

`<anna.bonnet@sorbonne-universite.fr>`, `<stephane.robin@sorbonne-universite.fr>`

Over the last few decades, the Hawkes process has become an important framework for modeling successions of self-exciting (or self-inhibiting) events that occur over time. It has been used in a wide variety of applications, such as earth sciences (seismology, vulcanology), economics, genomics, ecology, to name but a few. The conditional intensity of a Hawkes process comprises a base function (or immigration rate) and the sum of the influences of all past events, encoded in the so-called kernel function.

We consider here a discrete-time version of the Hawkes process, with an exponential kernel, where the immigration term varies according to a latent Markov chain. We prove that this model is identifiable and that, thanks to the specificities of the exponential kernel, it can be reformulated in terms of a hidden Markov model, with Poisson emissions. Based on these properties, we show that maximum likelihood inference of the model's parameters can be performed using an EM algorithm, which involves a recursive M-step.

We evaluate the accuracy of parameter estimates through a simulation study when signal level and discretization step vary. We also address the selection of the number of hidden states. Finally, we illustrate the use of the proposed modeling to distinguish different bat behaviors, based on the recording of their cries.

Title : Review of estimation algorithms for HMM/HSMM with mixed effects.

Authors : Mathieu Valdeyron (PhD Student, MIAT INRAE), Jean-Baptiste Durand (AMAP CIRAD), Nathalie Peyrard (MIAT INRAE), Sandra PLancade (MIAT INRAE).

Abstract : Several articles in the literature consider HMM or HSMM with mixed effects in the law of the observed process and of the hidden process. The covariates are introduced via the composition of a potentially non-linear link function and a linear expression with fixed and random effects, in the image of Generalised Linear Mixed Models (GLMM). The presentation will consist of a review of these models, in particular their modelling and estimation methods.

Unlike an HMM or HSMM without covariates, the Expectation-Maximisation (EM) algorithm cannot be realised exactly. The maximisation step M generally no longer has an analytical solution, and the presence of random effects prevents the exact calculation of the E step. Numerical approximations must be considered to overcome this additional difficulty. It is also possible to bypass the EM algorithm by directly maximising the log-likelihood after approximating it numerically. Bayesian approaches have also been considered, which rely on computational methods. We will review how these different approaches have been mobilized for H(S)MM with mixed effects.

Gaussian Markov random fields for MRF sampling

Jean-Baptiste Courbot
IRIMAS UR 7499, Université de Haute-Alsace
Mulhouse, France
`jean-baptiste.courbot@uha.fr`

Hugo Gangloff
Université Paris-Saclay, AgroParisTech, INRAE
UMR MIA Paris-Saclay, Palaiseau, France
`hugo.gangloff@inrae.fr`

Markov random fields are classical models for Bayesian image modeling [1], but their sampling relies traditionally on Gibbs techniques, implying a non-negligible computational cost. On the other hand, Gaussian Markov random fields (GMRF), which are used to model continuous and spatially-homogeneous variables, can benefit from highly efficient sampling techniques [2]. In this contribution, we propose a new discrete Markov field, based on a unit-simplex geometry, and coined Gaussian unit-simplex Markov field (GUM). The core of our proposition follows results from the three following definitions.

Definition 1 (Unit simplex)**.** *A unit P-simplex is a regular simplex belonging in $\mathbb{R}^P$, whose $P+1$ vertices lie on a unit sphere. Let us denote $\mathbf{U}_P = \{\mathbf{v}_1, \ldots, \mathbf{v}_{P+1}\}$ the set of vertices of the unit P-simplex.*

Definition 2 (Gaussian Unit-simplex Markov random field (GUM))**.** *Let $\mathbf{U}_{K-1}$ be the K vertices of a unit $(K-1)$-simplex (Definition 1), and $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ a GMRF taking values in $\mathbb{R}^{n(K-1)}$, such that $\mathbf{Z} = \{\mathbf{Z}_s\}_{s\in\mathcal{S}}$ and $\mathbf{Z}_s$ takes values in $\mathbb{R}^{K-1}$. We define the mapping $\phi_{K,c}\colon \mathbb{R}^{n(K-1)} \mapsto \mathbb{R}^n$ such that:*

$$\phi_{K,c}(\mathbf{Z}) = \sum_{i=1}^{K} \omega_i \pi_i^c(\mathbf{Z}) \text{ and } \forall s \in \mathcal{S},\ \pi_i^c(\mathbf{Z}_s) = \frac{\exp(-c^{-2}\|\mathbf{Z}_s - \mathbf{v}_i\|^2)}{\sum_{k=1}^{K}\exp(-c^{-2}\|\mathbf{Z}_s - \mathbf{v}_k\|^2)} \tag{1}$$

with $c > 0$, $\mathbf{v}_k, \mathbf{v}_i \in \mathbf{U}_{K-1}$ unit simplex vertices, and $\omega_i \in \Omega \subset \mathbb{N}$. $\phi_{K,c}(\mathbf{Z})$ is named a GUM random field*. π_i^c is designed to indicate, site-wise, the distance between $\mathbf{Z}$ and the $i-$th vertices $\mathbf{v}_i$ of $\mathbf{U}_{K-1}$. We can show that $\phi_{K,c}(\mathbf{Z})$ is also Markovian.*

Definition 3 (Discrete GUM)**.** *Let $\mathbf{Z}$ be a GMRF. From Definition 2, we have:*

$$\phi_{K,c}(\mathbf{Z}) \xrightarrow[c\to 0]{} \sum_{i=1}^{K} \omega_i \delta_{\left[\|\mathbf{Z}-\mathbf{v}_i\|_2 \le \|\mathbf{Z}-\mathbf{v}_k\|_2,\ \forall \mathbf{v}_k \in \mathbf{U}_{K-1}\right]} \tag{2}$$

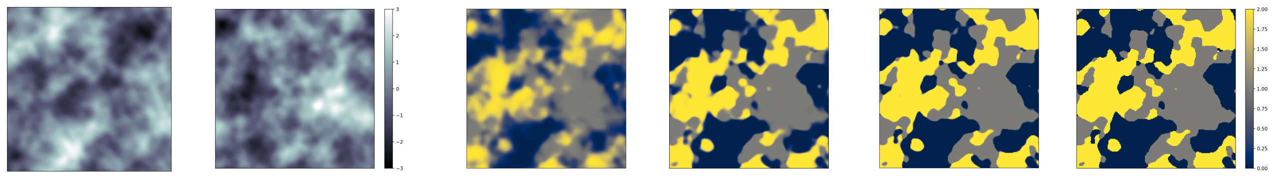
Denoting $\lim_{c\to 0} \phi_{K,c}(\mathbf{Z}) = \mathbf{X} = \{X_s\}_{s\in\mathcal{S}}$, we have $\forall s \in \mathcal{S}$:

$$X_s = \omega_{k^*} \text{ with } k^* \text{ chosen such that } \mathbf{v}_{k^*} = \operatorname*{arg\,min}_{\mathbf{v}\in\mathbf{U}_{K-1}} \|\mathbf{Z}_s - \mathbf{v}\|_2 \tag{3}$$

This discrete limit process will be referred to as a Discretized GUM *or* DGUM*.*

A depiction of GUM sampling is provided in Fig. 1. The computational complexity of the proposed method is that of GMRF sampling, so the proposed model can be sampled notably faster than Ising/Potts based models relying on Gibbs sampling. Our computation time is indeed improved by a factor 14 to 68 for the GPU implementation, and by a factor 200 to 290 in the CPU implementation, depending on the sample size.

Our main perspective is the development of an unsupervized inference similar based on the same models, for which the computational improvement should leverage the use of several latent field, within large images and/or 3D volumes.



(a) Realization $\mathbf{z}$: first component

(b) Realization $\mathbf{z}$: second component

(c) $\phi_{K,c}(\mathbf{z})$ with $c = 1$.

(d) $\phi_{K,c}(\mathbf{z})$ with $c = 0.5$.

(e) $\phi_{K,c}(\mathbf{z})$ with $c = 0.25$.

(f) DGUM of $\mathbf{z}$.

Figure 1: Illustration of the DGUM sampling for $K = 3$ classes, starting from the GMRF realizations (a-b), to the GUM (c-e) and its limit DGUM (f).

References.

[1] Z. Kato, J. Zerubia *et al.*, “Markov random fields in image segmentation,” *Foundations and Trends® in Signal Processing*, vol. 5, no. 1–2, pp. 1–155, 2012.

[2] H. Rue and L. Held, *Gaussian Markov random fields: theory and applications.* Chapman and Hall/CRC, 2005.

Asymptotic properties of the K-in-a-row design

Nancy Flournoy[1], José Moler[2], Maider Sada[2]

[1] University of Missouri. Columbia, USA. flournoyn@missouri.edu
[2] Universidad Pública de Navarra, Pamplona, Spain. jmoler@unavarra.es,
maider.sada@unavarra.es

Abstract

Searching for the level of a stimulus that has a prescribed percentage of success is a common goal for experiments in many fields. We consider experiments in which only a few values of the stimulus can be applied to the experimental units. We assume that the success response probability increases with the stimulus. For instance, in the development of a new drug, the dose which has a toxicity Γ is of interest. When a design with this objective adapts to concentrate experimental units around the unknown target percentile it is called a dose-finding design.

The K-in-a-Row design (KR design) is a rule that sequentially allocates experimental units to one of the permissible stimuli. As we will prove, the KR design tends to concentrate the allocations around the targeted stimulous. In the context of drug development, assume that M doses are permissible and the target toxicity is $\Gamma < 0.5$. Given the last subject experiences the toxicity at dose d_i, the next subject receives the next lower dose. Otherwise, dose d_i is repeated until K consecutive non-toxic responses are observed; in which case the next subject receives d_{i+1}. Even though this kind of experiment is applied to a small number of subjects, the asymptotic properties of the allocation rule are of interest to characterize the allocations sequences, e.g., to explain its bias and suggest improvements.

In this work we prove that the KR design is a Semi-Markov process in discrete time. A new proof, with fewer conditions, of the strong unimodality of the asymptotic allocation proportions is provided. Moreover, although it was previously known that the allocation mode is adjacent to the target dose, we explicitly identify it.

Our proofs do not require embedding the discrete process into a continuous one. The only assumption we make is that the M probabilities of toxicity increase with dose. This more precisely describes the statistical framework as non-parametric. Our statistical model accommodates contexts in which the response function is unknown, and even when the experimenter is suspicious about irregularities in the response function (non-continuity, non-derivability or, even, sudden changes in the slope of the response function).

Spatio-temporal generation of precipitations using a spatially correlated Bernoulli and hidden Markov model

Caroline Cognot [1&2] & Liliane Bel [2] & David Métivier [3] & Sylvie Parey [1]

[1] *EDF R&D, 91120, Palaiseau, France - caroline.cognot@edf.fr*
[2] *Université Paris-Saclay, AgroParisTech, INRAE, UMR MIA Paris-Saclay, 91120, Palaiseau, France*
[3] *UMR MISTEA, 34000, Montpellier, France*

Abstract. This study proposes a spatio-temporal model for precipitation across multiple stations, extending the multivariate hidden Markov model used in weather generation to a spatial framework for both occurrence and intensity. In previous works, rain occurrence at each site is modeled independently, conditioned on a shared hidden state representing a weather type. In our approach, spatial rain occurrence is modeled using a Gaussian field with a Matérn covariance structure, thresholded at each location into a Bernoulli variable based on the local rain probability. We incorporate this spatial model in the hidden Markov framework, with the aim of identifying interpretable states, such as those characterized by lower rain probability or higher spatial correlation. Additionally, we incorporate periodic parameterizations to capture seasonal patterns efficiently, rather than fitting separate models for each month. The rain quantity is modeled conditionally to the weather state using appropriate distributions. Model parameters are estimated using composite pairwise likelihood and the EM algorithm. Simulations from the fitted model are compared to observed data using several metrics such as the rain occurrence ratio (ROR), which represents the daily proportion of stations experiencing rainfall. The model is applied to 37 weather stations in France, using 51 years of daily precipitation data. Initial results from the spatial model indicate a higher estimated spatial range in winter compared to summer. This aligns with meteorological knowledge that attributes winter rainfall to large-scale phenomena and summer rainfall to smaller-scale processes. Our model allows the generation of high-resolution rainfall with realistic weather patterns reproducing key events such as droughts.

Keywords. Weather Generator, Hidden Markov Model, Spatio-temporal, EM algorithm

Yury A. Kutoyants

Le Mans University, Le Mans

Hidden Markov Processes and Adaptive Filttering

We present several models of partially observed diffusion processes and describe the construction of adaptive Kalman-type filters in the situations where the systems depend on some unknown finite-dimensional parameters. The presented algorithms for the filters and parameter estimators have recurrent structure and the questions of their asymptotic optimality are discussed. The properties of the filters and estimators are studied in the asymptotics of small noise and large samples. For some nonlinear partially observed systems the construction and properties of the corresponding extended adaptive Kalman filters are discussed too.

1. Kutoyants, Y.A. (2025) *Hidden Markov Processes and Adaptive Filtering,* Springer Series in Statistics, Cham.

Walking the mammoth

A reanalysis of isotope ratios in a 17 000 years old tusk to reconstruct the life trajectory of a woolly mammoth

Matthieu Pluntz[1], Juliette Meyniel[2], Hervé Perdry[2]

[1] NeuroDiderot, Inserm & Université Paris-Cité, Paris
[2] CESP, Inserm & Université Paris-Saclay, Villejuif

Elephantid tusks grow throughout the whole animal's life. Their isotopic composition reflect their alimentation, creating a record of their surrounding environment. In 2021, Matthew J. Wooller *et al* obtained the strontium and oxygen isotope ratios along the tusk of Kik, a mammoth who lived in Alaska 17 000 years ago. Using isotopic maps of Alaska, they reconstructed a path consisting of 1 133 successive positions, corresponding to the putative path followed by Kik along the 28 years of his life, with a time resolution of nine days.

This path was obtained with a Monte Carlo method, consisting in generating a large number of random walks based on strontium isotope ratios, and keeping among them the path which best matched the oxygen isotope ratio. The resuting path fits reasonnably well the strontium isotopic profile along the tusk, but fits poorly the oxygen isotopic profile.

We reanalyze these data using a Hidden Markov Model with the geographic position of Kik as hidden state. The transition probabilities are those of a Gaussian random walk truncated at a distance of 100 km. This truncature makes the transition matrix sparse, which mitigates the large number of hidden states (more than 20 000), and makes the problem computationally tractable. The observed states are the isotopic ratios, with emission probabilities computed from the ratios given by the isotopic maps.

We obtain paths that fit well both the strontium and oxygene isotopic profile. They differ notably from the path obtained by Wooller *et al,* generating intriguing hypotheses on Kik's death.

References

Wooller MJ *et al* (2021) Lifetime mobility of an Artic woolly mammoth *Science* 373:806-808

Generalized Viterbi algorithm for identifying the multiple most likely paths of hidden states in a HMM and application to detection of homozygosity by descent

Authors:
Matthieu Pluntz[2,3,*], Sidonie Foulon[1,2,3], Anne-Louise Leutenegger[2] et Hervé Perdry[3]

*corresponding author: matthieu.pluntz@inserm.fr

Affiliations:
1. Université Paris-Saclay, UVSQ, Inserm, Gustave Roussy, CESP, 94805, Villejuif, France
2. NeuroDiderot, Inserm, Université Paris Cité, UMR1141, 48 bd Sérurier, 75019, Paris, France
3. Université Paris-Saclay, UVSQ, Inserm, CESP, 94807, Villejuif, France

Keywords:
Hidden Markov Models, Viterbi algorithm, best paths, homozygosity-by-descent

Abstract:

In Hidden Markov models (HMM), inference of the unobserved Markov chain across hidden states is known as decoding. The Viterbi algorithm is an efficient decoding algorithm which finds the single most probable path of hidden states conditional on the observations. It contains a forward step which computes recursively for each state the maximal probability, conditional on the observations up to a step, of a path reaching that state at that step, as well as the previous state in that path (traceback). Then comes the backward reconstruction of the optimal path by concatenating tracebacks.

We propose a generalization of the Viterbi algorithm (α-Viterbi) which, for a given probability ratio α $\in]0,1]$, computes the set of paths having a posterior probability at least α times that of the optimal path computed by the classical Viterbi algorithm. This is done by computing in the forward part of the algorithm, instead of a single traceback, a probability ratio for each pair of consecutive states. Then we use this ratio to construct the set of paths in the backward stage, where we can limit the set to its *k* most probable paths.

In genomics, we use those decodings to detect simulated homozygous-by-descent (HBD) DNA segments in inbred individuals, where HBD status at a locus is unobserved and might randomly differ from HBD status at a neighbouring locus on the DNA chain due to the recombination that occurred since the common ancestor.

Parameter estimation of hidden Markov models: comparison of EM and quasi-Newton methods with a new hybrid algorithm.

Authors :
Sidonie Foulon[1,2,3], Thérèse Truong[1], Anne-Louise Leutenegger[2] et Hervé Perdry[3]

Affiliations :
1. Université Paris-Saclay, UVSQ, Inserm, Gustave Roussy, CESP, 94805, Villejuif, France
2. NeuroDiderot, Inserm, Université Paris Cité, UMR1141, 48 bd Sérurier, 75019, Paris, France
3. Université Paris-Saclay, UVSQ, Inserm, CESP, 94807, Villejuif, France

Keywords :
Hidden Markov models, Baum-Welch, quasi-Newton, L-BFGS-B, SQUAREM, optimisation.

Abstract :

Estimating the parameters of a Hidden Markov Model (HMM) prior to hidden states reconstruction is often a computationally demanding step. A natural method to estimate the parameters of HMM is the Baum-Welch algorithm, a particular case of the Expectation-Maximisation (EM) algorithm. However, other methods exist, particularly the direct maximisation of the likelihood, often employing quasi-Newton methods. The Baum-Welch algorithm always converges to a local maximum, but its convergence is slow. Convergence of “classical” quasi-Newton methods (ex. L-BFGS-B) is generally fast. Still, they can be complicated to implement, and their trajectories can be complex, particularly when the likelihood reaches multiple local maxima. We propose a hybrid algorithm, QNEM, which combines the Baum-Welch algorithm and a quasi-Newton method. The likelihood convexity guides the switch from one method to another.

We compare QNEM with the Baum-Welch algorithm, an accelerated EM algorithm named SQUAREM (Varadhan, 2008, Scand J Statist), and L-BFGS-B. We apply the four algorithms to four examples constructed on different models. Two examples are based on the eruption data of the Old Faithful geyser. The last example is built on a genetic model, where we aim to analyse genome-wide data to identify portions of the genome inherited twice from a common ancestor of an individual's parents. We estimate the parameters of each model with the four algorithms and evaluate their performances.

Our results show that the highest performing algorithm depends on the model, the Baum-Welch and SQUAREM algorithms being sometimes faster than L-BFGS-B and QNEM. The latter has globally good performances, at least as fast as the L-BFGS-B, making it a promising alternative to existing algorithms.

The calculation of the probability of observations in a HMM knowing the parameters of the model can be done classically with the forward or the forward-backward algorithm. Knowing that, we formalize the notion of an un-normalized heterogeneous Markov distributions (UHMD), as a tensor associated to a joint distribution, where these algorithms correspond to an elimination and marginalisation algorithm respectively, based on a series of matrix x vector products on a tensor network. It leads to a complexity in $O(TK^2)$ for normalising constant and computing all unary marginals. We show how the sparsity of the transition matrix permits to decrease the complexity of the calculation. We use this dictionary to extend within a common framework the evaluation of the complexity of forward and forward-backward algorithms for a diversity of Markov based Models with hidden variables. We show how to map the joint law of a HMM into a UHMD for computing the distribution of the observations after marginalisation over the hidden variables. This is extended to multichain HMM, and leads to bound the complexity according to the type of couplings using the decomposition or not of the transition matrix as a Kronecker product of univariate matrices, including the Factorial Model. We use the generative model for an ED-HMM to rewrite it as a UHMD, show that the transition matrix is sparse, and show how the sparsity can lower the upper bound of the complexity of the forward-backward algorithm. We present ways to possibly extend this approach towards some cases of multichains HSMM. Finally, we recall that the field property of **R**, i.e. the inversion of the multiplication, and even of the addition, is not required in matrix x vector product in the definition of an UHMD. This means that all these calculations can be implemented in a semi-ring. Such an implementation is classical in max-plus semi-ring and leads to Viterbi algorithm, corresponding to an elimination algorithm, which can be deployed this way on all mentioned models of the Markov family to recover most likely hidden states.

Modeling the impact of co-variates upon the transition from one hidden state to another may improve both estimation and prediction in hidden hidden Markov and semi-Markov models. However, the problem of the interpretability of the parametrization and feasability of the computation becomes crippling as the number of interest covariates increases.

I propose, in a Bayesian perspective, an architecture I hope to be as complex as needed to accommodate complex configurations of the data, but as simple as possible for the sake of usability. A few simulated examples are given to illustrate the flexibility of the model. A Bayesian strategy to fit the model is discussed.

Asymptotic comparison of nonparametric jump rate estimators for one-dimensional piecewise-deterministic Markov processes

Romain Azaïs (joint work with Solune Denis)

Inria Lyon

Piecewise-deterministic Markov processes (PDMPs) offer a powerful stochastic modeling framework that combines deterministic trajectories with random perturbations at random times. Estimating their local characteristics (particularly the jump rate) is an important yet challenging task. Several nonparametric methods for jump rate estimation have been introduced in recent years [2, 3, 4, 5], each grounded in distinct theoretical settings, which has made direct and systematic comparison difficult.

In this context, a unified framework is introduced to standardize and consolidate state-of-the-art approaches. Within this setting, new results are established concerning consistency and asymptotic normality of the estimators, allowing for a rigorous comparison in terms of convergence rates and asymptotic variances. Notably, it is shown that no single method uniformly outperforms the others, even within the same model. These theoretical insights are validated through numerical simulations using a representative PDMP application: the TCP model. Finally, the estimators are applied to experimental data describing cell growth and division in *Escherichia coli*, highlighting their practical relevance.

This abstract is based on the preprint [1], submitted in early 2025.

References

[1] Azaïs, R., and Denis, S. Asymptotic analysis and practical evaluation of jump rate estimators in piecewise-deterministic markov processes. arXiv:2502.14621, 2025.

[2] Azaïs, R., and Muller-Gueudin, A. Optimal choice among a class of nonparametric estimators of the jump rate for piecewise-deterministic Markov processes. *Electron. J. Statist. 10*, 2 (2016), 3648–3692.

[3] Fujii, T. Nonparametric estimation for a class of piecewise-deterministic Markov processes. *Journal of Applied Probability 50*, 4 (2013), 931–942.

[4] Krell, N. Statistical estimation of jump rates for a piecewise deterministic Markov processes with deterministic increasing motion and jump mechanism. *ESAIM: Probability and Statistics 20* (2016), 196–216.

[5] Krell, N., and Schmisser, É. Nonparametric estimation of jump rates for a specific class of piecewise deterministic Markov processes. *Bernoulli 27*, 4 (2021), 2362 – 2388.

Parameter Estimation and State Inference in Hidden Drifting Markov Models

Vlad Barbu*[1], Aguemon Tehungue†[1], and Nicolas Vergne‡[1]

[1]Laboratoire de Mathématiques Raphaël Salem, UMR CNRS 6085, Université de Rouen Normandie

Abstract

This work explores the application of Hidden Drifting Markov Models (HDMMs) in the context of time series data with hidden structures. We first introduce the theoretical foundations of HMMs, including the components and assumptions that define them. We then present two key algorithms: the expectation minimization algorithm (Em) for parameter estimation and the modified Viterbi algorithm to deduce the least risky hidden state sequence. Finally, we illustrate these methods through numerical experiments on simulated data. This work highlights the efficiency of HDMMs in capturing hidden dynamics in sequential data and the usefulness of Em and modified Viterbi in model training and interpretation.

Keywords: Drifting Markov Models, Hidden Markov Models, Inference, Expectation-Maximization algorithm, Viterbi algorithm

---

*vladstefan.barbu@univ-rouen.fr
†aguemon.tehungue@univ-rouen.fr
‡nicolas.vergne@univ-rouen.fr

Review of HSMM R and Python software

Caroline Bérard[1], Marie-Josée Cros[2], Jean-Baptiste Durand[3], Corentin Lothodé[4],
Sandra Plancade[2], Ronan Trépos[2], and Nicolas Vergne[4]

[1]LITIS, Université de Rouen Normandie, Rouen, France
[2]Université de Toulouse, INRAE-MIAT, Toulouse, France
[3]AMAP, CIRAD, Montpellier, France
[4]Université d'Angers, CNRS, LAREMA, SFR MATHSTIC, F-49000 Angers, France

Abstract

This work aims to provide a comprehensive review of the various software and packages available for Hidden Semi-Markov Models (HSMM), or with a particular focus on those related to the following themes: multichain models, coupled chains, and the incorporation of covariates. A thorough survey has been conducted to compile both R and Python packages that are specifically designed to handle HSMMs, along with those related to these key themes. The comparison of these packages is based on their functionalities, particularly in relation to the modeling of sojourn time distributions and emission distributions for HSMMs. To further illustrate the practical application of these tools, we demonstrate the use of two selected packages. A Python package is applied to discrete emission distributions in a case study involving squirrels, while an R package is employed for conitnuous emission distributions in a case study on deers. These examples highlight the diversity and the range of options available when working with HSMMs across different software environments.

Keywords: Hidden Semi Markov Models, R, Python

This presentation consists of three different parts. First, we focus on semi-Markov chains and we present different explicit formulas to describe reliability indicators such as failure rates. The Markov chains are defined in a discrete state space and the indicators are estimated by means of empirical plug-in type estimators. Our objective is to compare the different expressions, study the asymptotic properties of the estimators and finally compare them in terms of asymptotic variance. The strong consistency and the asymptotic normality of the estimators are proved. The asymptotic study is made in the large-sample context. Second, we move to continuous-time semi-Markov processes and present bootstrapped, kernel-type estimators and bootstrapped kernel-type estimators of reliability indicators. These last estimators are very advantageous in the case of real data studies, since the evaluation of their asymptotic confidence intervals is feasible due to the fact that they do not depend on theoretical quantities. The third part concerns reliability indicators for hidden Markov and hidden semi-Markov models. Explicit formulas are derived and estimated by means of plug-in type estimators.

List of participants

- Azaïs Romain
- Bacave Hanna
- Berard Caroline
- Bouzebda Salim
- Cognot Caroline
- Coron Camille
- Coube Sébastien
- Courbot Jean-Baptiste
- Davari Soheil
- De Saporta Benoîte
- Denis Solune
- Dousson Lys Gaspard
- Durand Jean-Baptiste
- Erard Adélie
- Farinetto Christian
- Fernández Carlos
- Foulon Sidonie
- Franc Alain
- Gangloff Hugo
- Gravrand Victor
- Guindo Abdoulaye
- Hamza Abderaouf
- Karagrigoriou Alex
- Koutoyants Youri
- Krell Nathalie

- Leite Alessandro
- Lepère Dominique
- Maillard Odalric-Ambrym
- Matias Catherine
- Mews Sina
- Micollet Leo
- Moler José Antonio
- Naulet Zacharie
- Pandolfi Silvia
- Perdry Hervé
- Peyrard Nathalie
- Plancade Sandra
- Pluntz Matthieu
- Porciani Angelique
- Robin Stéphane
- Rossini Orlane
- Sabbadin Régis
- Sadaka Georges
- Sidous Mellina
- Silvestrov Dmitrii
- Skandari Reza
- Sugier Pierre-Emmanuel
- Tehungue Aguêmon Léon
- Valdeyron Mathieu
- Vergne Nicolas
- Vernay Amélie

Author Index