Posterior consistency for partially-observed Markov models

LTCI, CNRS, Telecom ParisTech, Université Paris-Saclay

*Joint work with **Randal Douc** and **Jimmy Olsson**.*

Outline

Introduction

- Models
- Bayesian setting

Outline

Outline

Outline

Introduction
- Models
- Bayesian setting

Main results
- Our setting
- Checking the $\mathbb{P}$-remoteness property: NSC
- Checking the $\mathbb{P}$-remoteness property: AMLE
- Checking the approximate merging property: complete MC.

Some extensions
- Support of the complete chain depending on the parameter
- Toward non-stationarity

Conclusion, summary and remaining questions...

Introduction

- Models
- Bayesian setting

Main results

Some extensions

Conclusion, summary and remaining questions...

Introduction

Models

Bayesian setting

Main results

Some extensions

Conclusion, summary and remaining questions...

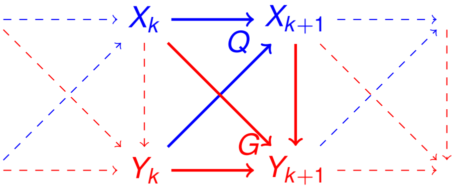
Partially observed Markov models

Definition

A partially observed Markov model is a pairwise homogeneous Markov chain $(Z_n = (X_n, Y_n), \mathcal{F}_n)_{n\geq 0}$ with kernel **K** generally described as

$$
\begin{aligned}
X_{k+1} | \mathcal{F}_k &\sim Q(X_k, Y_k; \cdot) \;, \\
Y_{k+1} | \mathcal{F}_k, X_{k+1} &\sim G(X_k, Y_k, X_{k+1}; \cdot) \;,
\end{aligned}
\tag{1}
$$

- and such that only the $\{Y_k\}'$s are observed.



Two important examples

$$\begin{tikzcd} {} \arrow[r, dashed] & X_k \arrow[r, "Q"'] \arrow[d, dashed] & X_{k+1} \arrow[r, dashed] \arrow[d, "G"] & {} \\ & Y_k & Y_{k+1} & \end{tikzcd}$$

Hidden Markov models (HMM)

$$\begin{tikzcd} {} \arrow[r, dashed] & X_k \arrow[r] \arrow[d, dashed] & X_{k+1} \arrow[r, dashed] \arrow[d, "G"] & {} \\ & Y_k \arrow[ur, "Q"] & Y_{k+1} & \end{tikzcd}$$

Obs.-Driv. models (ODM)

- An ODM moreover requires that

$$\begin{aligned} Q(X_k, Y_k; \cdot) &= \delta_{f_{Y_k}(X_k)}(\cdot) \\ G(X_k, X_{k+1}, Y_k; \cdot) &= G(X_{k+1}; \cdot) \end{aligned} \tag{2}$$

Markov latent variables

As for HMM's, $\{X_k\}$ is Markov with a transition kernel

$$R(x_0; A) = \int G(x_0; \mathrm{d}y_0)\, \mathbb{1}_A\left(f_{y_0}(x_0)\right) \tag{3}$$

- Parametric models are obtained by setting $Q = Q^\theta$ (or for ODMs: $f_y(x) = f_y^\theta(x)$) and, sometimes, $G = G^\theta$, $\theta \in \Theta$
- Partially dominated models: $\dfrac{\mathrm{d}G^\theta(x;\cdot)}{\mathrm{d}\nu}(y) = g^\theta(x; y)$
- Fully-dominated model: $\mathbf{K}_\theta(z_0, \mathrm{d}z_1) = k_\theta(z_0, z_1)\, \mu \otimes \nu(\mathrm{d}z_1)$.

Examples of ODM

- GARCH(1, 1) by [Bollerslev(1986)]:

$$G(x;\cdot) = \mathcal{N}(0, x) \quad \text{and} \quad f_y(x) = \omega + ax + by^2$$

- In-GARCH, see [Davis et al.(2003)]:
  $G(x;\cdot) = \mathrm{Poi}(x) \quad \text{and} \quad f_y(x) = \omega + ax + by$

Examples of ODM

- GARCH(1, 1) by [Bollerslev(1986)]:

$$G(x;\cdot) = \mathcal{N}(0, x) \quad \text{and} \quad f_y(x) = \omega + ax + by^2$$

- In-GARCH, see [Davis et al.(2003)]:
  $G(x;\cdot) = \mathrm{Poi}(x) \quad \text{and} \quad f_y(x) = \omega + ax + by$
- Log-In-GARCH, see [Fokianos and Tjøstheim (2011)]:

$$G(x;\cdot) = \mathrm{Poi}(\exp x) \quad \text{and} \quad f_y(x) = \omega + ax + b\log(1+y)$$

- NBINGARCH(1, 1) by [Zhu(2011)]:

$$G(x;\cdot) = \mathcal{NB}\left(r, \frac{1}{1+x}\right) \quad \text{and} \quad f_y(x) = \omega + ax + by$$

- NM(d)-GARCH(1,1), etc.

Basic assumption and definitions

Ergodicity assumption

We assume that $\mathbf{K}_\theta$ is ergodic for all $\theta \in \Theta$, and denote by π^θ the *unique* stationary distribution.

For all initial distribution η,

- let $\mathbb{P}^\theta_\eta$ be the probability on $((\mathsf{X} \times \mathsf{Y})^{\mathbb{N}}, (\mathcal{X} \otimes \mathcal{Y})^{\otimes \mathbb{N}})$ induced by $\mathbf{K}_\theta$ starting from $(X_0, Y_0) \sim \eta$
- $\mathbb{P}^\theta$ is $\mathbb{P}^\theta_{\pi^\theta}$ extended to the negative time indices
- $\tilde{\mathbb{P}}^\theta$ is $\mathbb{P}^\theta$ restricted to $\mathsf{Y}^{\mathbb{Z}}$ components

Maximum likelihood inference

- Maximum Likelihood Estimator (MLE) $\theta_{\eta,n}$ is defined as

$$\boxed{\theta_{\eta,n} \in \operatorname{argmax}_{\theta\in\Theta} \left\{ \mathsf{p}_\eta^\theta(Y_{1:n}) \right\}} \tag{4}$$

  for some arbitrary initial dist. ξ.

- In well-specified models, a standard consistency result consists in showing that

$$\boxed{\lim_{n\to\infty} \theta_{\eta,n} = \theta_\star\,, \quad \tilde{\mathbb{P}}^{\theta_\star}\text{-a.s.}} \tag{5}$$

  where $\tilde{\mathbb{P}}^{\theta_\star}$ is the stationary distribution of $Y_{1:\infty}$.

Introduction

Models

Bayesian setting

Main results

Some extensions

Conclusion, summary and remaining questions...

Posterior distribution

Consider the posterior distribution:

$$\boxed{\lambda_n(A) = \frac{\int_A \lambda(\mathrm{d}\theta) p_{\theta,n}}{\int_\Theta \lambda(\mathrm{d}\theta) p_{\theta,n}}} .$$

where

1. $p_{\theta,n} = p_\theta(Y_1, \ldots, Y_n)$ where $Y_1, \ldots, Y_n$ are the observations
2. λ is a (possibly infinite) measure, called the prior.

- $p_n^\star$ denote the real density of the observations $Y_{1:n}$.
- the parameter set (Θ, d) is a metric space and denote by $\mathcal{T}$ its Borel σ-field.

The true value of the parameter

Definition

$p_n^\star$ and $p_{\theta_\star,n}$ merge with probability 1 if and only if

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{p_{\theta_\star,n}}{p_n^\star} = 0 \quad \mathbb{P}\text{-a.s.}$$

$\theta_\star$ is called the "true value" of the parameter.

The true value of the parameter

Definition

$p_n^\star$ and $p_{\theta_\star,n}$ merge with probability 1 if and only if

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{p_{\theta_\star,n}}{p_n^\star} = 0 \quad \mathbb{P}\text{-a.s.}$$

$\theta_\star$ is called the "true value" of the parameter.

We are interested in finding conditions under which the posterior consistency property holds, i.e.

$$\boxed{\lambda_n \underset{n\to\infty}{\Longrightarrow} \delta_{\theta_\star} \quad \mathbb{P}\text{-a.s.}}$$

Posterior consistency condition

By the Portmanteau Lemma, it is equivalent to show that $\mathbb{P}$-a.s.,

$$\limsup_n \lambda_n(C) \le \delta_{\theta_\star}(C) , \quad C \text{ closed set in } \Theta$$

i.e. for all $A_p = \{\theta \in \Theta \;:\; d(\theta, \theta_\star) \ge 1/p\}$

$$\limsup_n \lambda_n(A_p) = 0 , \quad \mathbb{P}\text{-a.s.}$$

Posterior consistency condition

By the Portmanteau Lemma, it is equivalent to show that $\mathbb{P}$-a.s.,

$$\limsup_n \lambda_n(C) \le \delta_{\theta_\star}(C)\,, \quad C \text{ closed set in } \Theta$$

i.e. for all $A_p = \{\theta \in \Theta \;:\; d(\theta, \theta_\star) \ge 1/p\}$

$$\limsup_n \lambda_n(A_p) = 0\,, \quad \mathbb{P}\text{-a.s.}$$

Note that

$$\lambda_n(A) = \frac{\int_A \lambda(\mathrm{d}\theta) p_{\theta,n}}{\int_\Theta \lambda(\mathrm{d}\theta) p_{\theta,n}}\,.$$

Posterior consistency condition

By the Portmanteau Lemma, it is equivalent to show that $\mathbb{P}$-a.s.,

$$\limsup_n \lambda_n(C) \leq \delta_{\theta_\star}(C)\,, \quad C \text{ closed set in } \Theta$$

i.e. for all $A_p = \{\theta \in \Theta \,:\, d(\theta, \theta_\star) \geq 1/p\}$

$$\limsup_n \lambda_n(A_p) = 0\,, \quad \mathbb{P}\text{-a.s.}$$

Note that

$$\lambda_n(A) = \frac{\int_A \lambda(\mathrm{d}\theta) p_{\theta,n}/p_n^\star}{\int_\Theta \lambda(\mathrm{d}\theta) p_{\theta,n}/p_n^\star}\,.$$

Remoteness

$$\lambda_n(A) = \frac{\boxed{\int_A \lambda(\mathrm{d}\theta) p_{\theta,n}/p_n^\star}}{\int_\Theta \lambda(\mathrm{d}\theta) p_{\theta,n}/p_n^\star} \, .$$

Definition

We say that a set $A \in \mathcal{T}$ is $\mathbb{P}$-remote if and only if

$$\limsup_{n\to\infty} n^{-1} \log \left(\int_A \frac{p_{\theta,n}}{p_n^\star} \lambda(\mathrm{d}\theta) \right) < 0 \quad \mathbb{P}\text{-a.s.}$$

Approximate remoteness

$$\lambda_n(A) = \frac{\boxed{\int_A \lambda(\mathrm{d}\theta) p_{\theta,n}/p_n^\star}}{\int_\Theta \lambda(\mathrm{d}\theta) p_{\theta,n}/p_n^\star} = \boxed{\frac{\int_A \lambda(\mathrm{d}\theta) p_{\theta,n}}{\int_\Theta \lambda(\mathrm{d}\theta) p_{\theta,n}}} .$$

Definition

Moreover, we say that a set A is approximately $\mathbb{P}$-remote if and only if for all $\varepsilon > 0$ there exists a set $K_\varepsilon \in \mathcal{T}$ such that

(i) $A \cap K_\varepsilon$ is $\mathbb{P}$-remote;

(ii) $\limsup_{n\to\infty} \lambda_n(K_\varepsilon^c) \leq \varepsilon \quad \mathbb{P}\text{-a.s.}$

Typically, K_ϵ is a compact set.

Main assumption

$$\lambda_n(A) = \frac{\int_A \lambda(\mathrm{d}\theta) p_{\theta,n}/p_n^\star}{\boxed{\int_\Theta \lambda(\mathrm{d}\theta) p_{\theta,n}/p_n^\star}} .$$

Assumption (A1)

For all $\delta > 0$, there exists a set $\Theta_\delta \in \mathcal{T}$ such that $\lambda(\Theta_\delta) > 0$ and for all $\theta \in \Theta_\delta$,

$$\liminf_{n\to\infty} n^{-1} \log \frac{p_{\theta,n}}{p_n^\star} \geq -\delta \quad \mathbb{P}\text{-a.s.} \tag{6}$$

▶This can be seen as an asymptotic merging property since (6) implies

$$0 \geq \limsup_{n\to\infty} n^{-1} \log \frac{p_{\theta,n}}{p_n^\star} \geq \liminf_{n\to\infty} n^{-1} \log \frac{p_{\theta,n}}{p_n^\star} \geq -\delta \quad \mathbb{P}\text{-a.s.}$$

Immediate consequence

Theorem ([Barron et al.(1999)], adapted)

Assume (A1) . Then all approximately $\mathbb{P}$-remote sets A satisfy

$$\lim_{n\to\infty} \lambda_n(A) = 0 \;, \quad \mathbb{P}\text{-a.s.} \tag{7}$$

Immediate consequence

Theorem ([Barron et al.(1999)], adapted)

Assume (A1) . Then all approximately $\mathbb{P}$*-remote sets A satisfy*

$$\lim_{n \to \infty} \lambda_n(A) = 0 \;, \quad \mathbb{P}\text{-a.s.} \tag{7}$$

Objectives:

1. Give sufficient and handy conditions for getting the posterior consistency for partially observed Markov models.

Immediate consequence

Theorem ([Barron et al.(1999)], adapted)

Assume (A1) . Then all approximately $\mathbb{P}$*-remote sets A satisfy*

$$\lim_{n\to\infty} \lambda_n(A) = 0 \;, \quad \mathbb{P}\text{-a.s.} \tag{7}$$

Objectives:

1. Give sufficient and handy conditions for getting the posterior consistency for partially observed Markov models.
2. Treat the case of non compact parameter spaces

Immediate consequence

Theorem ([Barron et al.(1999)], adapted)

Assume (A1) . Then all approximately $\mathbb{P}$-remote sets A satisfy

$$\lim_{n\to\infty} \lambda_n(A) = 0 , \quad \mathbb{P}\text{-a.s.} \tag{7}$$

Objectives:

1. Give sufficient and handy conditions for getting the posterior consistency for partially observed Markov models.
2. Treat the case of non compact parameter spaces
3. Treat the case of non stationary observations

Immediate consequence

Theorem ([Barron et al.(1999)], adapted)

Assume (A1) . Then all approximately $\mathbb{P}$-remote sets A satisfy

$$\lim_{n \to \infty} \lambda_n(A) = 0 , \quad \mathbb{P}\text{-a.s.} \tag{7}$$

Objectives:

1. Give sufficient and handy conditions for getting the posterior consistency for partially observed Markov models.
2. Treat the case of non compact parameter spaces
3. Treat the case of non stationary observations
4. Give explicit Θ_δ such that Assumption (A1) can be easily checked.

Introduction

Main results

Our setting

Checking the $\mathbb{P}$-remoteness property: NSC

Checking the $\mathbb{P}$-remoteness property: AMLE

Checking the approximate merging property: complete MC.

Some extensions

Conclusion, summary and remaining questions...

Fully dominated partially observed Markov chains

Assumption

The Markov kernel $\mathbf{K}_{\theta_\star}$ is Fully-dominated model with $k_\theta(z, z') > 0$ and has a unique stationary distribution $\pi_\star$.

In this case, we set

- The true distribution is $\mathbb{P} = \mathbb{P}^{\theta_\star}_{\pi_\star}$;
- The true density $p^\star_n$ is the corresponding density applied to $Y_{1:n}$, $p^\star_n = p_{\theta_\star,\pi_\star}(Y_{1:n})$;
- The target density $p_{\theta,n}$ with parameter θ is given by $p_{\theta,n} = p_{\theta,\eta}(Y_{1:n})$ for some arbitrary initial distribution η.

Checking the $\mathbb{P}$-remoteness property: a necessary and sufficient condition

Proposition (A necessary and sufficient condition)

The set $A \in \mathcal{T}$ is $\mathbb{P}$-remote if and only if there exists a sequence $(B_n)_{n\in\mathbb{N}}$ of sets in $\mathcal{F}$ such that $B_n \in \mathcal{F}_n$ for all $n \in \mathbb{N}$ and

$$\limsup_{n\to\infty} n^{-1} \log \int_A \lambda(\mathrm{d}\theta)\mathbb{P}_{\theta,n}(B_n) < 0 \;,$$

$$\mathbb{P}\left(\liminf_{n\to\infty} B_n\right) = 1 \;.$$

The Necessary and Sufficient condition is obtained with

$$B_n = \left\{ \int_A \lambda(\mathrm{d}\theta) \frac{p_{\theta,n}}{p_n^\star} \le \varrho^n \right\} ,$$

The Necessary and Sufficient condition is obtained with

$$B_n = \left\{ \int_A \lambda(\mathrm{d}\theta) \frac{p_{\theta,n}}{p_n^\star} \leq \varrho^n \right\} ,$$

while, in concrete examples, the condition is used with

$$B_n = \left\{ \left| \frac{1}{n-m} \sum_{k=1}^{n-m} 1_C(Y_{k:k+m}) - \mathbb{P}_{\theta_\star,n}(Y_{0:m} \in C) \right| \leq \epsilon \right\}$$

Introduction

Main results

Our setting

Checking the $\mathbb{P}$-remoteness property: NSC

Checking the $\mathbb{P}$-remoteness property: AMLE

Checking the approximate merging property: complete MC.

Some extensions

Conclusion, summary and remaining questions...

Checking the $\mathbb{P}$-remoteness property: Approximate MLE

Definition

Let K be a compact subset of Θ. We say that a random sequence $(\hat{\theta}_n)_{n\in\mathbb{N}}$ adapted to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ are Approximate Maximum Likelihood Estimators (AMLE) on K if it is valued in K and, for all $n \in \mathbb{N}$,

$$n^{-1} \log p_{\hat{\theta}_n,n} \geq n^{-1} \log p_n^\star + \epsilon_n \quad \text{with} \quad \lim_{n\to\infty} \epsilon_n = 0 \quad \mathbb{P}\text{-a.s.}$$

Checking the $\mathbb{P}$-remoteness property: Approximate MLE

Definition

Let K be a compact subset of Θ. We say that a random sequence $(\hat{\theta}_n)_{n\in\mathbb{N}}$ adapted to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ are Approximate Maximum Likelihood Estimators (AMLE) on K if it is valued in K and, for all $n \in \mathbb{N}$,

$$n^{-1} \log p_{\hat{\theta}_n,n} \geq n^{-1} \log p_n^\star + \epsilon_n \quad \text{with} \quad \lim_{n\to\infty} \epsilon_n = 0 \quad \mathbb{P}\text{-a.s.}$$

Proposition

Let K be a compact subset of Θ such that $\lambda(K) < \infty$. If

- *All sequences $(\hat{\theta}_n)_{n\in\mathbb{N}}$ of AMLE on K are strongly consistent,*

then, for all closed set A not containing $\theta_\star$, the set $A \cap K$ is $\mathbb{P}$-remote.

Introduction

Main results

Our setting

Checking the $\mathbb{P}$-remoteness property: NSC

Checking the $\mathbb{P}$-remoteness property: AMLE

Checking the approximate merging property: complete MC.

Some extensions

Conclusion, summary and remaining questions...

Checking the approximate merging property

- How to prove (A1) ?

Checking the approximate merging property

▶How to prove ▸ (A1) ?

For all $\theta \in \Theta_\delta$, we write

$$\liminf_{n\to\infty} n^{-1} \log \frac{p_{\theta,\eta}(Y_{1:n})}{p_{\theta_\star,\pi_\star}(Y_{1:n})} \geq \liminf_{n\to\infty} n^{-1} \log \frac{p_{\theta,\eta}(Z_{1:n})}{p_{\theta_\star,\pi_\star}(Z_{1:n})} \quad \mathbb{P}\text{-a.s.} \ ,$$

- By the Birkhoff ergodic theorem, the right-hand side converges to

$$\blacksquare(\theta_\star, \theta) := \mathbb{E}^{\theta_\star}_{\pi_\star}\left[\mathrm{KL}\left(\mathbf{K}_{\theta_\star}(Z_0,\cdot) \| \mathbf{K}_\theta(Z_0,\cdot)\right)\right].$$

Checking the approximate merging property

▶How to prove (A1) ?

For all $\theta \in \Theta_\delta$, we write

$$\liminf_{n\to\infty} n^{-1} \log \frac{p_{\theta,\eta}(Y_{1:n})}{p_{\theta_\star,\pi_\star}(Y_{1:n})} \geq \liminf_{n\to\infty} n^{-1} \log \frac{p_{\theta,\eta}(Z_{1:n})}{p_{\theta_\star,\pi_\star}(Z_{1:n})} \geq -\delta \quad \mathbb{P}\text{-a.s.} \ ,$$

- By the Birkhoff ergodic theorem, the right-hand side converges to

$$\blacksquare(\theta_\star,\theta) := \mathbb{E}^{\theta_\star}_{\pi_\star}\left[\mathrm{KL}\left(\mathbf{K}_{\theta_\star}(Z_0,\cdot)\|\mathbf{K}_\theta(Z_0,\cdot)\right)\right].$$

Then it suffices to check that, for all $\delta > 0$,

$$\lambda(\Theta_\delta) > 0 \text{ where } \boxed{\Theta_\delta := \{\theta \in \Theta \ : \ \blacksquare(\theta_\star,\theta) \leq \delta\}}.$$

Introduction

Main results

Some extensions

Support of the complete chain depending on the parameter

Toward non-stationarity

Conclusion, summary and remaining questions...

Where the approximate merging bound fails

If $X_k = \psi_\theta(X_{k-1}) + U_k$ and $Y_k = \phi_\theta(X_k) + V_k$, where

1. $U_k \overset{\text{i.i.d}}{\sim} f_\theta$ where f_θ is a density with respect to ν with a bounded support.
2. $V_k \overset{\text{i.i.d}}{\sim} g_\theta$ where g_θ is a positive density with respect to μ.

François Roueff

Where the ▸ approximate merging bound fails

If $X_k = \psi_\theta(X_{k-1}) + U_k$ and $Y_k = \phi_\theta(X_k) + V_k$, where

1. $U_k \overset{\text{i.i.d}}{\sim} f_\theta$ where f_θ is a density with respect to ν with a bounded support.
2. $V_k \overset{\text{i.i.d}}{\sim} g_\theta$ where g_θ is a positive density with respect to μ.

More generally, assume that

$$\boxed{\mathbf{K}_\theta(z_0, \mathrm{d}z_1) = Q_\theta(x_0, \mathrm{d}x_1) g_\theta(x_1, y_1) \mu(\mathrm{d}y_1) \ ,}$$

with $\qquad g_\theta > 0$.

TELECOM ParisTech

Extension of the approximate merging property

$$\liminf_{n\to\infty} n^{-1} \log \frac{p_{\theta,\eta}(Y_{1:n})}{p_{\theta_\star,\pi_\star}(Y_{1:n})}$$

$$\geq \liminf_{n\to\infty} n^{-1} \log \frac{\prod_{i=1}^n g_\theta(X_i', Y_i)}{\prod_{i=1}^n g_{\theta_\star}(X_i, Y_i)} \geq -\delta \quad \mathbb{P}\text{-a.s.}\ ,$$

where

($\star$) $(X_i, X_i')_{i\in\mathbb{N}}$ is a Markov kernel coupling of $(Q_{\theta_\star}, Q_\theta)$ starting with an initial distribution γ of marginals $\pi_\star$ and η, and $Y_i \sim g_{\theta_\star}(X_i, \cdot)$.

In that case, we have used that $\theta \in \Theta_\delta$ satisfies

$$\blacksquare(\theta_\star, \theta) := \mathbb{E}_\gamma^{\theta_\star,\theta}\left[\mathrm{KL}\left(g_{\theta_\star}(X_0,\cdot)\|g_\theta(X_0',\cdot)\right)\right] \geq \delta\ .$$

Toward non-stationary observation processes

Many arguments rely on the fact that the observation process is stationary and ergodic under $\mathbb{P}$.

Toward non-stationary observation processes

Many arguments rely on the fact that the observation process is stationary and ergodic under $\mathbb{P}$.

What can be said if $\mathbb{P} = \mathbb{P}^{\theta_\star}_{\eta^*}$ instead of $\mathbb{P}^{\theta_\star}_{\pi_\star}$?

Toward non-stationary observation processes

Many arguments rely on the fact that the observation process is stationary and ergodic under $\mathbb{P}$.
What can be said if $\mathbb{P} = \mathbb{P}^{\theta_\star}_{\eta^*}$ instead of $\mathbb{P}^{\theta_\star}_{\pi_\star}$?

Proposition

If $k_{\theta_\star} > 0$, then for all $A \in \sigma(Z_{1:\infty})$ and all initial distribution η^*,

$$\mathbb{P}^{\theta_\star}_{\eta^*}(A) = \mathbb{E}^{\theta_\star}_{\pi_\star}\left(\frac{\int \eta^*(\mathrm{d}z_0) k_{\theta_\star}(z_0, Z_1)}{\pi_\star(Z_1)} \mathbb{1}_A \right) ,$$

As a consequence, $(\mathbb{P}^{\theta_\star}_{\eta^*}(\lambda_n \Longrightarrow_{n\to\infty} \delta_{\theta_\star}) = 1)$ if and only if $(\mathbb{P}^{\theta_\star}_{\pi_\star}(\lambda_n \Longrightarrow_{n\to\infty} \delta_{\theta_\star}) = 1)$.

Introduction

Main results

Some extensions

Conclusion, summary and remaining questions...

1. The posterior consistency relies on two main properties: $\mathbb{P}$-remoteness of closed sets avoiding $\theta_\star$ and approximate merging.

1. The posterior consistency relies on two main properties: $\mathbb{P}$-remoteness of closed sets avoiding $\theta_\star$ and approximate merging.
2. Fairly general sufficient conditions can be derived using explicit contrast functions.

1. The posterior consistency relies on two main properties: $\mathbb{P}$-remoteness of closed sets avoiding $\theta_\star$ and approximate merging.
2. Fairly general sufficient conditions can be derived using explicit contrast functions.
3. Geometric rates of convergence or known results on MLE can be directly applied to prove remoteness of (closed/compact) sets.

1. The posterior consistency relies on two main properties: $\mathbb{P}$-remoteness of closed sets avoiding $\theta_\star$ and approximate merging.
2. Fairly general sufficient conditions can be derived using explicit contrast functions.
3. Geometric rates of convergence or known results on MLE can be directly applied to prove remoteness of (closed/compact) sets.
4. Approximate remoteness for a more general parameter space can be derived (not mentioned here).

1. The posterior consistency relies on two main properties: $\mathbb{P}$-remoteness of closed sets avoiding $\theta_\star$ and approximate merging.
2. Fairly general sufficient conditions can be derived using explicit contrast functions.
3. Geometric rates of convergence or known results on MLE can be directly applied to prove remoteness of (closed/compact) sets.
4. Approximate remoteness for a more general parameter space can be derived (not mentioned here).
5. The class of observation driven models need a specific treatment (not fully dominated).

References

- Barron, A., Chervish, M. and Wasserman, L. (1999) The consistency of posterior distributions in non parametric problems. *Ann. Statist.*, 27 (2), 536–561.
- Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* 31, 307–327.
- Davis, R. and Dunsmuir, W. and Streett, S. (2003). Observation-driven models for Poisson counts. *Biometrika* 90 (4), 777–790.
- Fokianos, K. and Tjøstheim, D. (2011). Log-linear Poisson autoregression. *Journal of Multivariate Analysis* 102 (3), 563–578.
- Leroux, B. G., 1992. Maximum-likelihood estimation for hidden Markov models. *Stoch. Proc. Appl.* 40, 127–143.
- Zhu, F., 2011. A negative binomial integer-valued GARCH model. *J. Time Series Anal.* 32 (1), 54–67.