



HAL
open science

Erratum to “Nonparametric estimation of the stationary density and the transition density of a Markov chain”

Claire Lacour

► **To cite this version:**

Claire Lacour. Erratum to “Nonparametric estimation of the stationary density and the transition density of a Markov chain”. *Stochastic Processes and their Applications*, 2012, 122, pp.2480-2485. 10.1016/j.spa.2012.03.001 . hal-01139399

HAL Id: hal-01139399

<https://hal.science/hal-01139399v1>

Submitted on 4 Apr 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Erratum to

“Nonparametric estimation of the stationary density and the transition density of a Markov chain” Stochastic Process. Appl. 118 (2008)

Claire Lacour

New proof of Proposition 1

The result of Proposition 1 is true but the proof must be modified in the following way. We replace Lemma 10 by

Lemma 10 *Under the assumptions of Proposition 1, and if (X_n) has an atom A ,*

$$\sum_{\lambda \in \Lambda_m} \mathbb{E} |S_j(\varphi_\lambda)|^2 \leq r_0^2 \mathbb{E}_A(\tau^2) D_m.$$

Proof of Lemma ??: Using a convex inequality, we can write

$$\sum_{\lambda \in \Lambda_m} \mathbb{E} |S_j(\varphi_\lambda)|^2 \leq \sum_{\lambda \in \Lambda_m} \mathbb{E}_\mu \left| \sum_{i=\tau+1}^{\tau(2)} \varphi_\lambda(X_i) \right|^2 \leq \sum_{\lambda \in \Lambda_m} \mathbb{E}_\mu \left((\tau(2) - \tau) \sum_{i=\tau+1}^{\tau(2)} \varphi_\lambda^2(X_i) \right)$$

Assumption M2 entails $\| \sum_{\lambda \in \Lambda_m} \varphi_\lambda \|_\infty \leq r_0^2 D_m$. Then

$$\sum_{\lambda \in \Lambda_m} \mathbb{E} |S_j(\varphi_\lambda)|^2 \leq \mathbb{E}_\mu \left((\tau(2) - \tau) \sum_{i=\tau+1}^{\tau(2)} r_0^2 D_m \right) \leq r_0^2 \mathbb{E}_\mu ((\tau(2) - \tau)^2) D_m$$

To conclude, recall that by the Markov property,

$$\begin{aligned} \mathbb{E}_\mu ((\tau(2) - \tau)^2) &= \sum_k \sum_{l>k} (l - k)^2 \mathbb{P}_\mu(\tau = k, \tau(2) = l) \\ &= \sum_k \sum_{l>k} (l - k)^2 \mathbb{P}(X_{k+1} \notin A, \dots, X_{l-1} \notin A, X_l \in A | X_k \in A) \mathbb{P}_\mu(X_1 \notin A, \dots, X_k \in A) \\ &= \sum_k \sum_{l>k} (l - k)^2 \mathbb{P}_A(X_1 \notin A, \dots, X_{l-k-1} \notin A, X_{l-k} \in A) \mathbb{P}_\mu(\tau = k) \\ &= \sum_k \sum_{j>0} j^2 \mathbb{P}_A(\tau = j) \mathbb{P}_\mu(\tau = k) = \mathbb{E}_A(\tau^2). \end{aligned}$$

□

We can then give the bound

$$\sum_{\lambda \in \Lambda_m} \mathbb{E}(\nu_n^{(3)}(\varphi_\lambda)^2) \leq \frac{r_0^2 \mathbb{E}_A(\tau^2) D_m}{n}.$$

Finally $\mathbb{E}\|f_m - \hat{f}_m\|^2 \leq CD_m/n$ with $C = 4[8r_0^2(\mathbb{E}_\mu(\tau^2) + \mu(A)\mathbb{E}_A(\tau^4)) + r_0^2\mathbb{E}_A(\tau^2)]$.

New proof of Theorem 3

The result of Theorem 3 is true but the proof must be modified in the following way. Proposition 12 must be replaced by:

Proposition 12 *Let (X_n) be a Markov chain which satisfies A1–A5 and $(S_m)_{m \in \mathcal{M}_n}$ be a collection of models satisfying M1–M3. We suppose that (X_n) has an atom A . Let $B(m, m') = \{t \in S_m + S_{m'}, \|t\| = 1\}$ and*

$$p(m, m') = K\mu(A)\mathbb{E}_A(\tau^2)r_0^2 \frac{\dim(S_m + S_{m'})}{n}$$

(where K is a numerical constant). Then

$$\sum_{m' \in \mathcal{M}_n} \mathbb{E} \left[\sup_{t \in B(m, m')} \nu_n^2(t) - p(m, m') \right]_+ = O(n^{-1}).$$

Remark 1 *This gives a penalty in Theorem 3 of the form*

$$\text{pen}(m) = K\mu(A)\mathbb{E}_A(\tau^2)r_0^2 \frac{D_m}{n}, \text{ for some } K > K_0$$

with K_0 a numerical constant. Note that this penalty is simpler than in the previous version of this theorem. In particular, it does not depend on $\|f\|_\infty$.

Remark 2 *As it can be seen in the proof, Assumption M1 can be relaxed, it is now sufficient to assume that each S_m is a linear subspace of $(L^\infty \cap L^2)([0, 1])$ with dimension $D_m \leq n$. This entails an improvement on the smoothness assumption for Corollary 5 : $\alpha > 0$ is sufficient. In the same way, M1' can be relaxed and the condition for Corollary 8 is only $\alpha > 0$.*

Proof of Proposition ??: The heart of the proof is to use Theorem 7 in ? which is a concentration inequality for Markov chains. In our case $T_1 = \tau(1) = \tau$ and $T_2 = \tau(2) - \tau(1)$. Let us check that our assumptions allow us to use this theorem.

- We can easily prove that our assumption A4 implies the Minorization Condition with $m = 1$ in ?. Indeed, since $\int h d\mu > 0$, there exists C with measure $\mu(C) > 0$ and $\delta > 0$ such that h is larger than δ on C . Then for all x in C and all event B , $P(x, B) \geq h(x)\nu(B) \geq \delta\nu(B)$. Moreover, fixing $x \in \mathbb{R}$, for n large enough, the ergodicity of the chain gives

$$|P^n(x, C) - \mu(C)| \leq \frac{\mu(C)}{2},$$

which implies $P^n(x, C) \geq \mu(C)/2 > 0$.

- As noticed at the very beginning of his Section 3.5, the assumption of finiteness of Orlicz norm of T_1 and T_2 , which is required to apply the theorem, is equivalent to existence of a number $s > 1$ such that

$$\mathbb{E}_\mu(s^\tau) < \infty, \quad \mathbb{E}_\nu(s^\tau) < \infty. \quad (1)$$

Now, we use condition A5 of geometric ergodicity. Theorem 15.4.2 in ? shows that there exists a full absorbing set S such that S is geometrically regular, i.e. $\sup_{x \in S} \mathbb{E}_x(s^\tau) < \infty$ for some $s > 1$ (depending on A). Since S is full absorbing, and μ is the limit distribution of the chain $\mu(S) = 1$. Moreover $\mu(C \cap S) > 0$, where C is the set introduced in the Minorization Condition. So we can find an $x \in C \cap S$ and $\delta\nu(S^c) \leq P(x, S^c) = 0$. Thus $\nu(S) = 1$ too. This implies condition (??).

Now we write an integrated version of the concentration inequality. We denote $\nu_n(t) = n^{-1} \sum_{i=1}^n [t(X_i) - \langle t, f \rangle]$ where f is the stationary density of the chain and we consider a countable class \mathcal{B} of measurable functions t . Let a and H such that

$$\sup_{t \in \mathcal{B}} \|t - \langle t, f \rangle\|_\infty \leq a, \quad \mathbb{E} \left(\sup_{t \in \mathcal{B}} |\nu_n(t)| \right) \leq H.$$

Let the variance term

$$\sigma^2 = \mathbb{E}_A(\tau)^{-1} \sup_{t \in \mathcal{B}} \mathbb{E}_A \left[\left(\sum_{i=1}^{\tau} t(X_i) - \langle t, f \rangle \right)^2 \right].$$

Then we prove the existence of a numerical constant $c > 0$ such that

$$\mathbb{E}[\sup_{t \in \mathcal{B}} |\nu_n(t)|^2 - cH^2]_+ \leq K_1 \left(\frac{1}{n^2} + \frac{\sigma^2}{n} e^{-K_2 \frac{nH^2}{\sigma^2}} + \frac{a^2(\log n)^2}{n^2} e^{-K_3 \frac{nH}{a \log n}} \right) \quad (2)$$

where K_1, K_2, K_3 are depending on the chain. Indeed we compute, for $c = 8K^2$,

$$\begin{aligned} \mathbb{E} \left[\sup_{t \in \mathcal{B}} |\nu_n(t)|^2 - cH^2 \right]_+ &= \int_0^\infty P \left(\sup_{t \in \mathcal{B}} |\nu_n(t)|^2 \geq cH^2 + x \right) dx \\ &\leq \int_0^\infty P \left(\sup_{t \in \mathcal{B}} |\nu_n(t)| \geq \sqrt{c/2}H + \sqrt{x/2} \right) dx \leq \int_0^\infty P \left(Z \geq \sqrt{c/2}\mathbb{E}Z + n\sqrt{\frac{x}{2}} \right) dx \\ &\leq \int_0^\infty P \left(Z \geq K\mathbb{E}Z + K\mathbb{E}Z + n\sqrt{\frac{x}{2}} \right) dx \end{aligned}$$

where $Z = n \sup_{t \in \mathcal{B}} |\nu_n(t)|$. If $x \geq 2n^{-2}$, $t = K\mathbb{E}Z + n\sqrt{x/2} \geq 1$ so that we can apply Theorem 7. Moreover

$$\int_0^{2n^{-2}} P \left(Z \geq K\mathbb{E}Z + K\mathbb{E}Z + n\sqrt{\frac{x}{2}} \right) dx \leq 2n^{-2}.$$

Thus

$$\begin{aligned} & \mathbb{E} \left[\sup_{t \in \mathcal{B}} |\nu_n(t)|^2 - cH^2 \right]_+ \leq \frac{2}{n^2} \\ & + \int_0^\infty K \exp \left(-\frac{1}{K'} \min \left(\frac{[K\mathbb{E}Z + n\sqrt{x/2}]^2}{n\sigma^2}, \frac{K\mathbb{E}Z + n\sqrt{x/2}}{a \log n} \right) \right) dx \\ & \leq \frac{2}{n^2} + \frac{1}{K_2} e^{-\frac{K_2(\mathbb{E}Z)^2}{n\sigma^2}} \int_0^\infty e^{-\frac{K_2 n x}{\sigma^2}} dx + \frac{1}{K_3} e^{-\frac{K_3 \mathbb{E}Z}{a \log n}} \int_0^\infty e^{-\frac{K_3 n \sqrt{x}}{a \log n}} dx \\ & \leq \frac{2}{n^2} + K_4 \frac{\sigma^2}{n} e^{-\frac{K_2 n H^2}{\sigma^2}} + K_5 \frac{(a \log n)^2}{n^2} e^{-\frac{K_3 n H}{a \log n}} \end{aligned}$$

This gives inequality (??). This result can be extended to a non-countable class \mathcal{B} with classical density arguments. So we apply it with $\mathcal{B} = B(m, m')$. Moreover, the result of ? is also true when replacing $\mathbb{E}Z = n\mathbb{E}(\sup_{t \in \mathcal{B}} |\nu_n(t)|)$ by $n\mathbb{E}(\sup_{t \in \mathcal{B}} |\nu'_n(t)|)$ with

$$\nu'_n(t) = \frac{1}{n} \sum_{j=1}^{\lfloor 3n/\mathbb{E}_A(\tau) \rfloor} S_j(t)$$

(see in the proof of Theorem 7, p 1020). Thus (??) is also valid with $H \geq \mathbb{E}(\sup_{t \in \mathcal{B}} |\nu'_n(t)|)$. It remains to compute a, H and σ^2 . We denote $D(m, m') = \max(D_m, D_{m'})$ the dimension of the space $S_m + S_{m'}$ (recall that the models are nested) and $(\varphi_\lambda)_{\lambda \in \Lambda(m, m')}$ an orthonormal basis of $S_m + S_{m'}$.

- Computation of a . If $t \in S_m + S_{m'}$, $\|t\|_\infty \leq r_0 \sqrt{D(m, m')} \|t\|$. Then $a = 2r_0 \sqrt{D(m, m')}$.
- Computation of H^2 . Since any $t \in B(m, m')$ can be written $t = \sum_{\lambda \in \Lambda(m, m')} a_\lambda \varphi_\lambda$,

$$\mathbb{E} \left(\sup_{t \in B(m, m')} \nu'_n(t)^2 \right) \leq \sum_{\lambda \in \Lambda(m, m')} \mathbb{E}(\nu'_n(\varphi_\lambda)^2) \leq \sum_{\lambda \in \Lambda(m, m')} \mathbb{E} \left(\left(\frac{1}{n} \sum_{j=1}^{\lfloor 3n/\mathbb{E}_A(\tau) \rfloor} S_j(\varphi_\lambda) \right)^2 \right).$$

Recall that the $S_j(t)$ are independent identically distributed and centered. Then, using (new) Lemma 10,

$$\mathbb{E} \left(\sup_{t \in B(m, m')} \nu'_n(t)^2 \right) \leq \frac{\lfloor 3n/\mathbb{E}_A(\tau) \rfloor}{n^2} r_0^2 \mathbb{E}_A(\tau^2) D(m, m').$$

Finally, since $\mu(A) = \mathbb{E}_A(\tau)^{-1}$, we set $H^2 = CD(m, m')/n$ with $C = 3\mu(A)\mathbb{E}_A(\tau^2)r_0^2$.

- Computation of σ^2 . We use the following inequality, given in ?, subsection 17.4.3:

$$\mu(A)\mathbb{E}_A \left[\left(\sum_{i=1}^{\tau} t(X_i) - \langle t, f \rangle \right)^2 \right] = 2 \int (t - \langle t, f \rangle) \hat{t} d\mu - \int (t - \langle t, f \rangle)^2 d\mu$$

where

$$\hat{t}(x) := \mathbb{E}_x \left(\sum_{i=0}^{\sigma_A} t(X_i) - \langle t, f \rangle \right)$$

and $\sigma_A = \inf\{n \geq 0, X_n \in A\}$. Then, since $\mu(A) = \mathbb{E}_A(\tau)^{-1}$,

$$\sigma^2 \leq \sup_{t \in B(m, m')} 2 \int (t - \langle t, f \rangle) \hat{t} d\mu \leq \sup_{t \in B(m, m')} 2 \left(\int (t - \langle t, f \rangle)^2 d\mu \int \hat{t}^2 d\mu \right)^{1/2}.$$

But $\int (t - \langle t, f \rangle)^2 d\mu \leq \int t^2 f \leq \|f\|_{\infty} \|t\|^2$ and

$$\hat{t}^2(x) \leq \mathbb{E}_x \left(\left(\sum_{i=0}^{\sigma_A} t(X_i) - \langle t, f \rangle \right)^2 \right) \leq 4\|t\|_{\infty}^2 \mathbb{E}_x((\sigma_A + 1)^2).$$

with $\mathbb{E}_x((\sigma_A + 1)^2) \leq \mathbb{E}_x((\tau + 1)^2)$. Then

$$\sigma^2 \leq 4\sqrt{\mathbb{E}_{\mu}((\tau + 1)^2)} \sqrt{\|f\|_{\infty}} \sup_{t \in B(m, m')} \|t\|_{\infty} \|t\|$$

so that

$$\sigma^2 \leq 4\sqrt{\mathbb{E}_{\mu}((\tau + 1)^2)} \sqrt{\|f\|_{\infty}} r_0 \sqrt{D(m, m')}.$$

Now, we can use inequality (??): it implies existence of positive constants K'_1, K'_2, K'_3 such that

$$\mathbb{E}[\sup_{t \in \mathcal{B}} |\nu_n(t)|^2 - cCD(m, m')/n]_+ \leq K'_1 \left(\frac{1}{n^2} + \frac{\sqrt{D(m, m')}}{n} e^{-K'_2 \sqrt{D(m, m')}} + \frac{D(m, m')(\log n)^2}{n^2} e^{-K'_3 \frac{\sqrt{n}}{\log n}} \right).$$

Using that $D(m, m') = \max(D_m, D'_m) \leq n$, we obtain that $\sum_{m' \in \mathcal{M}_n} \sqrt{D(m, m')} e^{-K'_2 \sqrt{D(m, m')}}$ and $\sum_{m' \in \mathcal{M}_n} D(m, m')(\log n)^2 n^{-1} e^{-K'_3 \frac{\sqrt{n}}{\log n}}$ are bounded. Moreover $|\mathcal{M}_n| n^{-2} = O(n^{-1})$. Thus

$$\sum_{m' \in \mathcal{M}_n} \mathbb{E}[\sup_{t \in \mathcal{B}} |\nu_n(t)|^2 - cCD(m, m')/n]_+ = O(n^{-1})$$

□

New proof of Theorem 9

The result of Theorem 9 is true but the proof must be modified in the following way. Recall that we denote $E_n = \{\|f - \tilde{f}\|_\infty \leq \chi/2\}$ and E_n^c its complementary. We have

$$\mathbb{E}\|\pi - \tilde{\pi}\|^2 \leq \frac{8}{\chi^2} \left(\mathbb{E}\|g - \tilde{g}\|^2 + \|\pi\|_\infty^2 \mathbb{E}\|f - \tilde{f}\|^2 \right) + (a_n + \|\pi\|_\infty)^2 P(E_n^c)$$

so that it is sufficient to bound $(a_n + \|\pi\|_\infty)^2 P(E_n^c)$. We have proven that, for n large enough,

$$P(E_n^c) \leq P\left(\|f_{\hat{m}} - \hat{f}_{\hat{m}}\|_\infty > \frac{\chi}{4}\right) \leq P\left(\|f_{\hat{m}} - \hat{f}_{\hat{m}}\| > \frac{\chi}{4r_0\sqrt{D_{\hat{m}}}}\right).$$

But

$$\|f_{\hat{m}} - \hat{f}_{\hat{m}}\| = \sup_{t \in S_{\hat{m}}, \|t\| \leq 1} \int t(\hat{f}_{\hat{m}} - f_{\hat{m}}) = \sup_{t \in S_{\hat{m}}, \|t\| \leq 1} \nu_n(t).$$

Let S_{m_0} the largest model with dimension $D_{m_0} \leq n^{1/4}$.

$$P(E_n^c) \leq P\left(\sup_{t \in S_{\hat{m}}, \|t\| \leq 1} \nu_n(t)^2 > \frac{\chi^2}{16r_0^2 D_{\hat{m}}}\right) \leq P\left(\sup_{t \in S_{m_0}, \|t\| \leq 1} \nu_n(t)^2 > \frac{\chi^2}{16r_0^2 D_{m_0}}\right).$$

As shown in the (new) proof of Proposition 12, our assumptions allow us to use Theorem 7 in ?. Then, reasoning as in the proof of Proposition 12, we can show the existence of a numerical constant $c > 0$ and constants depending on the chain $K_1, K_2, K_3 > 0$ such that

$$P\left(\sup_{t \in S_{m_0}, \|t\| \leq 1} \nu_n(t)^2 \geq \frac{c}{2} H^2\right) \leq K_1 \left(e^{-K_2 \sqrt{D_{m_0}}} + e^{-K_3 \sqrt{n}/\log(n)} \right)$$

where $H^2 = 3\mu(A)\mathbb{E}_A(\tau^2)r_0^2 D_{m_0}/n$. Now, for n large enough, since $D_{m_0}^2 = o(n)$,

$$\frac{\chi^2}{16r_0^2 D_{m_0}} \geq \frac{3c\mu(A)\mathbb{E}_A(\tau^2)r_0^2 D_{m_0}}{2n}.$$

Then

$$P(E_n^c) \leq P\left(\sup_{t \in S_{m_0}, \|t\| \leq 1} \nu_n(t)^2 \geq \frac{c}{2} H^2\right) \leq K_1 \left(e^{-K_2 \sqrt{D_{m_0}}} + e^{-K_3 \sqrt{n}/\log(n)} \right)$$

so that $(a_n + \|\pi\|_\infty)^2 P(E_n^c) = o(n^{-1})$. Note that it is sufficient to have $D_{m_0} = \lfloor n^{1/2-\epsilon} \rfloor$ to obtain the result.

Acknowledgements

I am grateful to Mathieu Sart for bringing to my attention the mistake in the original proof of Lemma 10.

References

Adamczak, Radosław. *A tail inequality for suprema of unbounded empirical processes with applications to Markov chains*. Electron. J. Probab. 13, pages 1000–1034 (2008).

Meyn, S. P., Tweedie, R. L., 1993. *Markov chains and stochastic stability*. Springer-Verlag, London.