Statistics and Probability Letters 117 (2016) 72–79

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

A computable bound of the essential spectral radius of finite range Metropolis–Hastings kernels

Loïc Hervé, James Ledoux*

INSA de Rennes, IRMAR, F-35042, France
CNRS, UMR 6625, Rennes, F-35708, France
Université Bretagne Loire, France

ARTICLE INFO

Article history:
Received 4 April 2016
Received in revised form 12 May 2016
Accepted 12 May 2016
Available online 24 May 2016

MSC:
60J10
47B07

Keywords:
Markov chain operator
Metropolis–Hastings algorithms
Spectral gap

ABSTRACT

Let π be a positive continuous target density on $\mathbb{R}$. Let P be the Metropolis–Hastings operator on the Lebesgue space $\mathbb{L}^2(\pi)$ corresponding to a proposal Markov kernel Q on $\mathbb{R}$. When using the quasi-compactness method to estimate the spectral gap of P, a mandatory first step is to obtain an accurate bound of the essential spectral radius $r_{ess}(P)$ of P. In this paper a computable bound of $r_{ess}(P)$ is obtained under the following assumption on the proposal kernel: Q has a bounded continuous density $q(x, y)$ on $\mathbb{R}^2$ satisfying the following finite range assumption : $|u| > s \Rightarrow q(x, x+u) = 0$ (for some $s > 0$). This result is illustrated with Random Walk Metropolis–Hastings kernels.

1. Introduction

Let π be a positive distribution density on $\mathbb{R}$. Let $Q(x, dy) = q(x, y)dy$ be a Markov kernel on $\mathbb{R}$. Throughout the paper we assume that $q(x, y)$ satisfies the following finite range assumption: there exists $s > 0$ such that

$$|u| > s \Longrightarrow q(x, x+u) = 0. \tag{1}$$

Let $T(x, dy) = t(x, y)dy$ be the nonnegative kernel on $\mathbb{R}$ given by

$$t(x, y) := \min\left(q(x, y), \frac{\pi(y)\, q(y, x)}{\pi(x)}\right) \tag{2}$$

and define the associated Metropolis–Hastings kernel:

$$P(x, dy) := r(x)\, \delta_x(dy) + T(x, dy) \qquad \text{with } r(x) := 1 - \int_{\mathbb{R}} t(x, y)\, dy, \tag{3}$$

where $\delta_x(dy)$ denotes the Dirac distribution at x. The associated Markov operator is still denoted by P, that is we set for every bounded measurable function $f : \mathbb{R} \to \mathbb{C}$:

$$\forall x \in \mathbb{R}, \quad (Pf)(x) = r(x)f(x) + \int_{\mathbb{R}} f(y)\, t(x, y)\, dy. \tag{4}$$

* Corresponding author at: INSA de Rennes, IRMAR, F-35042, France.
E-mail addresses: Loic.Herve@insa-rennes.fr (L. Hervé), James.Ledoux@insa-rennes.fr (J. Ledoux).

http://dx.doi.org/10.1016/j.spl.2016.05.007

In the context of Monte Carlo Markov Chain methods, the kernel Q is called the proposal Markov kernel. We denote by $(\mathbb{L}^2(\pi), \|\cdot\|_2)$ the usual Lebesgue space associated with the probability measure $\pi(y)dy$. For convenience, $\|\cdot\|_2$ also denotes the operator norm on $\mathbb{L}^2(\pi)$, namely: if U is a bounded linear operator on $\mathbb{L}^2(\pi)$, then $\|U\|_2 := \sup_{\|f\|_2=1}\|Uf\|_2$. Since

$$t(x,y)\pi(x) = t(y,x)\pi(y), \tag{5}$$

we know that P is reversible with respect to π and that π is P-invariant (e.g. see Roberts and Rosenthal, 2004). Consequently P is a self-adjoint operator on $\mathbb{L}^2(\pi)$ and $\|P\|_2 = 1$. Now define the rank-one projector Π on $\mathbb{L}^2(\pi)$ by

$$\Pi f := \pi(f)1_{\mathbb{R}} \quad \text{with } \pi(f) := \int_{\mathbb{R}} f(x)\,\pi(x)\,dx.$$

Then the spectral radius of $P - \Pi$ is equal to $\|P - \Pi\|_2$ since $P - \Pi$ is self-adjoint, and P is said to have the spectral gap property on $\mathbb{L}^2(\pi)$ if

$$\varrho_2 \equiv \varrho_2(P) := \|P - \Pi\|_2 < 1.$$

In this case the following property holds:

$$\forall n \geq 1,\ \forall f \in \mathbb{L}^2(\pi), \quad \|P^n f - \Pi f\|_2 \leq \varrho_2^n\,\|f\|_2. \tag{SG$_2$}$$

The spectral gap property on $\mathbb{L}^2(\pi)$ of a Metropolis–Hastings kernel is of great interest, not only due to the explicit geometrical rate given by (SG$_2$), but also since it ensures that a central limit theorem (CLT) holds true for additive functional of the associated Metropolis–Hastings Markov chain under the expected second-order moment conditions, see Roberts and Rosenthal (1997). Furthermore, the rate of convergence in the CLT is $O(1/\sqrt{n})$ under third-order moment conditions (as for the independent and identically distributed models), see details in Hervé and Pène (2010) and Ferré et al. (2012).

The quasi-compactness approach can be used to compute the rate $\varrho_2(P)$. This method is based on the notion of essential spectral radius. Indeed, first recall that the essential spectral radius of P on $\mathbb{L}^2(\pi)$, denoted by $r_{ess}(P)$, is defined by (e.g. see Wu, 2004 for details):

$$r_{ess}(P) := \lim_n (\inf \|P^n - K\|_2)^{1/n} \tag{6}$$

where the above infimum is taken over the ideal of compact operators K on $\mathbb{L}^2(\pi)$. Note that the spectral radius of P is one. Then P is said to be quasi-compact on $\mathbb{L}^2(\pi)$ if $r_{ess}(P) < 1$. Second, if $r_{ess}(P) \leq \alpha$ for some $\alpha \in (0,1)$, then P is quasi-compact on $\mathbb{L}^2(\pi)$, and the following properties hold: for every real number κ such that $\alpha < \kappa < 1$, the set $\mathcal{U}_\kappa$ of the spectral values λ of P satisfying $\kappa \leq |\lambda| \leq 1$ is composed of finitely many eigenvalues of P, each of them having a finite multiplicity (e.g. see Hennion, 1993 for details). Third, if P is quasi-compact on $\mathbb{L}^2(\pi)$ and satisfies usual aperiodicity and irreducibility conditions (e.g. see Meyn and Tweedie, 1993), then $\lambda = 1$ is the only spectral value of P with modulus one and $\lambda = 1$ is a simple eigenvalue of P, so that P has the spectral gap property on $\mathbb{L}^2(\pi)$. Finally the following property holds: either $\varrho_2(P) = \max\{|\lambda|, \lambda \in \mathcal{U}_\kappa, \lambda \neq 1\}$ if $\mathcal{U}_\kappa \neq \emptyset$, or $\varrho_2(P) \leq \kappa$ if $\mathcal{U}_\kappa = \emptyset$.

This paper only focuses on the preliminary central step of the previous spectral method, that is to find an accurate bound of $r_{ess}(P)$. More specifically, we prove that, if the target density π is positive and continuous on $\mathbb{R}$, and if the proposal kernel $q(\cdot,\cdot)$ is bounded continuous on $\mathbb{R}^2$ and satisfies (1) for some $s > 0$, then

$$r_{ess}(P) \leq \alpha_a \quad \text{with } \alpha_a := \max(r_a, r'_a + \beta_a) \tag{7}$$

where, for every $a > 0$, the constants r_a, r'_a and β_a are defined by:

$$r_a := \sup_{|x|\leq a} r(x), \qquad r'_a := \sup_{|x|>a} r(x), \qquad \beta_a := \int_{-s}^{s} \sup_{|x|>a} \sqrt{t(x,x+u)\,t(x+u,x)}\,du. \tag{8}$$

This result is illustrated in Section 2 with Random Walk Metropolis–Hastings (RWMH) kernels for which the proposal Markov kernel is of the form $Q(x,dy) := \Delta(|x-y|)\,dy$, where $\Delta : \mathbb{R} \to [0,+\infty)$ is an even continuous and compactly supported function.

In Atchadé and Perron (2007) the quasi-compactness of P on $\mathbb{L}^2(\pi)$ is proved to hold provided that (A) the essential supremum of the rejection probability $r(\cdot)$ with respect to π is bounded away from unity; (B) the operator T associated with the kernel $t(x,y)dy$ is compact on $\mathbb{L}^2(\pi)$. Assumption (A) on the rejection probability $r(\cdot)$ is a necessary condition for P to have the spectral gap property (SG$_2$) (see Roberts and Tweedie, 1996). But this condition, which is quite generic from the definition of $r(\cdot)$ (see Remark 3), is far to be sufficient for P to satisfy (SG$_2$). The compactness Assumption (B) of Atchadé and Perron (2007) is quite restrictive, for instance it is not adapted for random walk Metropolis–Hastings kernels. Here this compactness assumption is replaced by the condition $r'_a + \beta_a < 1$. As shown in the examples of Section 2, this condition is adapted to RWMH.

In the discrete state space case, a bound for $r_{ess}(P)$ similar to (7) has been obtained in Hervé and Ledoux (in press). Next a bound of the spectral gap $\varrho_2(P)$ has been derived in Hervé and Ledoux (in press) from a truncation method for which the control of the essential spectral radius of P is a central step. It is expected that, in the continuous state space case, the bound (7) will provide a similar way to compute the spectral gap $\varrho_2(P)$ of P. This issue, which is much more difficult than in the discrete case, is not addressed in this work.