Bayesian selector of adaptive bandwidth for multivariate gamma kernel estimator with standard bias correction on $[0,\infty)^d$

Somé Sobom M.*
Université Thomas Sankara, LST Ouagadougou, Burkina Faso - sobom.some@gmail.com

Kokonendji Célestin C.
Université Bourgogne Franche-Comté, LMB Besançon, France - celestin.kokonendji@univ-fcomte.fr

Adjabi Smail
Université de Béjaïa, LAMOS Béjaïa, Algérie - adjabi@hotmail.com

Beddek Said
Université de Béjaïa, LAMOS Béjaïa, Algérie - said.beddek@yahoo.fr

Abstract

In this paper, a Bayesian adaptive estimation of bandwidth vector is provided for the multivariate modified gamma kernels. This kernel with standard bias reduction is appropriated to estimate nonnegative orthant densities with support $[0,\infty)^d$. For this purpose, we treat the bandwidth vector as a random vector using the inverse gamma prior. Exact expression of the posterior distribution and the vector of bandwidths are obtained through the usual quadratic loss function. Simulation studies and applications highlight similar performances between the proposed approach and the standard gamma case without bias reduction, and under integrated squared errors.

Keywords: Asymmetric kernel, multivariate boundary kernel, nonnegative data, prior distribution.

1. Introduction

Asymmetric kernels are known to improve smoothing quality for partially or totally bounded supports; e.g. Scaillet (2004) with inverse and reciprocal inverse Gaussian kernels. However, these kernels induces an additional quantity in the bias that needs reduction via modified versions; see for example Chen (1999, 2000), Hirukawa and Sakudo (2014, 2015), Igarashi and Kakizawa (2014, 2015), Malec and Schienle (2015). In the multivariate setting, let $\mathbf{X}_1,\ldots,\mathbf{X}_n$ be independent and identically distributed (iid) d-variate random variables with an unknown probability density function (pdf) f on $\mathbb{T}_d=[0,\infty)^d$, a subset of $\mathbb{R}^d$ with $d\geq 1$. Then, the multiple (standard and modified) gamma kernel estimators $\widehat{f_n}$ and $\widetilde{f_n}$ of f are defined, respectively, for $\mathbf{X}_i=(X_{i1},\ldots,X_{id})^\top$, $i=1,\ldots,n$, by

$$\widehat{f_n}(\boldsymbol{x})=\frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d}G_{x_j,h_j}(X_{ij}) \quad \text{and} \quad \widetilde{f_n}(\boldsymbol{x})=\frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d}G_{\rho(x_j;h_j),h_j}(X_{ij}) \quad \forall \boldsymbol{x}\in\mathbb{T}_d=[0,\infty)^d, \tag{1}$$

where $\boldsymbol{x}=(x_1,\ldots,x_d)^\top$ is the target vector and $\mathbf{h}=(h_1,\ldots,h_d)^\top$ is the vector of smoothing parameters with $h_j>0$, $j=1,\ldots,d$; see, e.g. Bouerzmarni and Rombouts (2010). Both functions $G_{x,h}(\cdot)$ and $G_{\rho(x;h),h}(\cdot)$ are, respectively, the standard and modified gamma kernels given on the support $\mathbb{S}_{x,h}=[0,\infty)=\mathbb{T}_1$ with $x\geq 0$ and $h>0$:

$$G_{x,h}(u)=\frac{u^{x/h}}{\Gamma(1+x/h)\,h^{1+x/h}}\exp\left(-\frac{u}{h}\right)\mathbb{1}_{[0,\infty)}(u) \quad \text{and} \quad G_{\rho(x;h),h}(u)=\frac{u^{\rho(x;h)-1}}{\Gamma(\rho(x;h))\,h^{\rho(x;h)}}\exp\left(-\frac{u}{h}\right)\mathbb{1}_{[0,\infty)}(u), \tag{2}$$

where $\Gamma(v)=\int_0^\infty s^{v-1}\exp(-s)ds$ is the classical gamma function with $v>0$, $\mathbb{1}_E$ denotes the indicator function of any given event E, and the parameter $\rho(x;h)$ is

$$\rho(x;h)=\begin{cases} 1+(x/2h)^2 & \text{if } x\in[0,2h) \\ x/h & \text{if } x\in[2h,\infty). \end{cases} \tag{3}$$

The gamma kernel $G_{x,h}(\cdot)$ appears to be the pdf of the gamma distribution, denoted by $\mathcal{G}(1 + x/h, h)$ with shape parameter $1 + x/h$ and scale parameter h.

The estimators (1) were originally introduced in the univariate case by Chen (2000) and then used in multivariate case by Bouerzmarni and Rombouts (2010). Notice that the formulation (3) allows to distinguish the boundary region from the interior one with respect to the bias corrections; see Zhang (2010) and Libengué Dobélé-Kpoka and Kokonendji (2017) for another choices of the boundary and interior regions and also, Funke and Kawka (2015) for several nonnegative kernels. A generalized form of (1) is recently introduced by Kokonendji and Somé (2018).

The performance of the estimators in (1) depends crucially on the diagonal bandwidth matrix $\mathbf{h} := diag(h_1, \ldots, h_d)$. Remark that this matrix with d real parameters is a particular case of the full symmetric one with $d(d + 1)$ independent parameters. The global bandwidth matrices selections such as cross-validation, plug-in and recently global Bayesian are known to have lower performances than their variable counterparts namely adaptive and local. The reader can see for example Ziane et al. (2015), Somé (2021), Somé and Kokonendji (2021) and Zougab et al. (2014) for Bayesian adaptive cases using univariate Birnbaum-Saunders kernel, (univariate and multiple) standard gamma kernel and multivariate Gaussian kernel, respectively.

The purpose of this paper is to propose an explicit selector of adaptive Bayesian bandwidths in multivariate pdf estimation on $[0, \infty)^d$ using the product of d univariate modified gamma kernels (2) with parametrization (3). The inverse gamma is used as prior to obtain the exact formula of the posterior distribution and the vector of bandwidth. The performances of this Bayesian approach for multiple modified gamma kernels are finally compared to the one with multiple standard gamma kernels (Somé and Kokonendji, 2021) using simulated and real data. See also Kokonendji and Somé (2021) for a Bayesian selector of adaptive bandwidth in a semiparametric topic with these multiple standard gamma kernels.

2. Bayesian adaptive bandwidth selector for multiple modified gamma kernels

Following the multiple standard gamma case of Somé and Kokonendji (2021), the modified gamma kernel estimator is constructed from (1) with (2) and (3) by considering a variable bandwidth vector $\mathbf{h}_i = (h_{i1}, \ldots, h_{id})^\top$ for each observation $\mathbf{X}_i = (X_{i1}, \ldots, X_{id})^\top$ in place of the fixed bandwidth vector $\mathbf{h} = (h_1, \ldots, h_d)^\top$. Thus, we treat $\mathbf{h}_i$ as a random vector with a prior distribution $\pi(\cdot)$.

The estimator (1) with parametrization (3) for multiple modified gamma kernel and variable vector of bandwidth $\mathbf{h}_i$ is written in $\boldsymbol{x} = (x_1, \ldots, x_d)^\top \in [0, \infty)^d$ as

$$\widehat{f_n}(\boldsymbol{x}) = \frac{1}{n} \sum_{i=1}^{n} \prod_{\ell=1}^{d} G_{\rho(x_\ell, h_{i\ell}), h_{i\ell}}(X_{i\ell}). \tag{4}$$

The leave-one-out kernel estimator of $f(\mathbf{X}_i)$ is deduced from (4) as

$$\widehat{f_{n,\mathbf{h}_i,-i}}(\mathbf{X}_i) := \frac{1}{n-1} \sum_{j=1, j\neq i}^{n} \prod_{\ell=1}^{d} G_{\rho(x_\ell, h_{i\ell}), h_{i\ell}}(X_{j\ell}). \tag{5}$$

Let $\pi(\mathbf{h}_i)$ be the prior distribution of $\mathbf{h}_i$, then the posterior distribution for each variable bandwidth vector $\mathbf{h}_i$ given $\mathbf{X}_i$ provided from the Bayesian rule is expressed as follow

$$\pi(\mathbf{h}_i \mid \mathbf{X}_i) = \frac{\widehat{f_{n,\mathbf{h}_i,-i}}(\mathbf{X}_i)\pi(\mathbf{h}_i)}{\int_{\chi} \widehat{f_{n,\mathbf{h}_i,-i}}(\mathbf{X}_i)\pi(\mathbf{h}_i) d\mathbf{h}_i}, \tag{6}$$

where χ is the space of positive vectors. The Bayesian estimator $\widetilde{\mathbf{h}_i}$ of $\mathbf{h}_i$ is obtained through the usual quadratic loss function as

$$\widetilde{\mathbf{h}_i} = \mathbb{E}\left(\mathbf{h}_i \mid \mathbf{X}_i\right) = \left(\mathbb{E}(h_{i1} \mid \mathbf{X}_i), \ldots, \mathbb{E}(h_{id} \mid \mathbf{X}_i)\right)^\top. \tag{7}$$

We assume that each component $h_{i\ell} = h_{i\ell}(n)$, $\ell = 1, \ldots, d$, of $\mathbf{h}_i$ has the univariate inverse gamma prior $\mathcal{IG}(\alpha, \beta_\ell)$ distribution with same shape parameters $\alpha > 0$ and, eventually, different scale parameters $\beta_\ell > 0$ such that $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_d)^\top$. We here recall that the pdf of $\mathcal{IG}(\alpha, \beta_\ell)$ with $\alpha, \beta_\ell > 0$ is defined by

$$IG_{\alpha,\beta_\ell}(u) = \frac{\beta_\ell^\alpha}{\Gamma(\alpha)} u^{-\alpha-1} \exp(-\beta_\ell/u) \mathbb{1}_{(0,\infty)}(u), \quad \ell = 1, \ldots, d, \tag{8}$$

where $\Gamma(\cdot)$ is the usual gamma function. From those considerations, the closed form of the posterior density and the Bayesian estimator of the vector $\mathbf{h}_i$ are given in the following proposition.

Proposition 0.1 *For fixed $i \in \{1, 2, \ldots, n\}$, consider each observation $\mathbf{X}_i = (X_{i1}, \ldots, X_{id})^\top$ with its corresponding vector $\mathbf{h}_i = (h_{i1}, \ldots, h_{id})^\top$ of univariate bandwidths and defining the subset $\mathbb{I}_i = \{k \in \{1, \ldots, d\}; X_{ik} = [0, 2h_{ik})\}$ and its complementary set $\mathbb{I}_i^c = \{\ell \in \{1, \ldots, d\}\,; X_{i\ell} \in [2h_{ik}, \infty)\}$. Using the inverse gamma prior IG_{α,β_ℓ} of (8) for each component $h_{i\ell}$ of $\mathbf{h}_i$ in the multiple gamma estimator (4) with $\alpha > 1/2$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_d)^\top \in (0, \infty)^d$, then:*
(i) there exists $\lambda_{ik} > 0$ for $k \in \mathbb{I}_i$ such that the posterior density (6) is the following weighted sum of inverse gamma

$$\pi(\mathbf{h}_i \mid \mathbf{X}_i) = \frac{1}{D_i(\alpha, \boldsymbol{\beta})} \sum_{j=1, j\neq i}^{n} \left(\prod_{k\in\mathbb{I}_i} A_{ijk}(\alpha, \beta_k)\, IG_{\lambda_{ik}+\alpha+1, X_{jk}+\beta_k}(h_{ik}) \right) \left(\prod_{\ell\in\mathbb{I}_i^c} B_{ij\ell}(\alpha, \beta_\ell)\, IG_{\alpha+1/2, C_{ij\ell}(\beta_\ell)}(h_{i\ell}) \right),$$

with $A_{ijk}(\alpha, \beta_k) = [\Gamma(\lambda_{ik}+\alpha+1) X_{jk}^{\lambda_{ik}}]/[\beta_k^{-\alpha} \Gamma(\lambda_{ik}+1)(X_{jk}+\beta_k)^{\lambda_{ik}+\alpha+1}]$, $B_{ij\ell}(\alpha, \beta_\ell) = [X_{j\ell}^{-1} \Gamma(\alpha+1/2)]/(\beta_\ell^{-\alpha} X_{i\ell}^{-1/2} \sqrt{2\pi}[C_{ij\ell}(\beta_\ell)]^{\alpha+1/2})$, $C_{ij\ell}(\beta_\ell) = X_{i\ell} \log(X_{i\ell}/X_{j\ell}) + X_{j\ell} - X_{i\ell} + \beta_\ell$, and $D_i(\alpha, \boldsymbol{\beta}) = \sum_{j=1, j\neq i}^{n} \left(\prod_{k\in\mathbb{I}_i} A_{ijk}(\alpha, \beta_k) \right) \left(\prod_{\ell\in\mathbb{I}_i^c} B_{ij\ell}(\alpha, \beta_\ell) \right)$;
(ii) under the quadratic loss function, the Bayesian estimator $\widetilde{\mathbf{h}}_i = \left(\widetilde{h}_{i1}, \ldots, \widetilde{h}_{id} \right)^\top$ of $\mathbf{h}_i$, introduced in (7), is

$$\widetilde{h}_{im} = \frac{1}{D_i(\alpha, \boldsymbol{\beta})} \sum_{j=1, j\neq i}^{n} \left(\prod_{k\in\mathbb{I}_i} A_{ijk}(\alpha, \beta_k) \right) \left(\prod_{\ell\in\mathbb{I}_i^c} B_{ij\ell}(\alpha, \beta_\ell) \right) \left(\frac{X_{jm} + \beta_m}{\lambda_{ik} + \alpha} \mathbb{1}_{[0, 2h_{im})}(X_{im}) + \frac{C_{ijm}(\beta_m)}{\alpha - 1/2} \mathbb{1}_{[2h_{im}, \infty)}(X_{im}) \right),$$

for $m = 1, 2, \ldots, d$, with the previous notations of $B_{ij\ell}(\alpha, \beta_\ell)$, $A_{ijk}(\alpha, \beta_k)$, $C_{ijm}(\beta_m)$ et $D_i(\alpha, \boldsymbol{\beta})$.

One can remark that $\lambda_{ik} \to 0$ with $k \in \mathbb{I}_i$ provides the Bayesian adaptive bandwidth vector for the multiple gamma kernel of Somé and Kokonendji (2021). Similarly to Somé and Kokonendji (2021) and Kokonendji and Somé (2021) for nonparametric and semiparametric approaches respectively, we have to consider $\alpha = \alpha_n = n^{2/5} > 2$ and $\beta_\ell = 1 > 0$, $\ell = 1, \ldots, d$ in numerical illustrations. These previous choices are not necessarily the optimal for obtaining the best smoothing quality. Simulations, applications to real datasets with the R (2021) software, and comparisons to the Bayesian adaptive bandwidth with standard gamma kernels will point out the efficiency of this modified version.

References

Bouezmarni, T. & Rombouts, J.V.K. (2010), Nonparametric density estimation for multivariate bounded data, *Journal of Statistical Planning and Inference* **140**, 139–152.

Chen, S.X. (1999) A beta kernel estimation for density functions, *Computational Statistics and Data Analysis* **31**, 131–145.

Chen, S.X. (2000). Probability density function estimation using gamma kernels, *Annals of the Institute of Statistical Mathematics* **52**, 471–480.

Funke, B. and Kawka, R. (2015) Nonparametric density estimation for multivariate bounded data using two non-negative multiplicative bias correction methods, *Computational Statistics and Data Analysis* **92**, 148–162.

Hirukawa, M. & Sakudo, M. (2014) Nonnegative bias reduction methods for density estimation using asymmetric kernels, *Computational Statistics and Data Analysis* **75**, 112–123.

Hirukawa, M. & Sakudo, M. (2015) Family of the generalised gamma kernels: a generator of asymmetric kernels for nonnegative data, *Journal of Nonparametrics Statistics* **27**, 41–63.

Igarashi, G. & Kakizawa, Y. (2014) Re-formulation of inverse Gaussian, reciprocal inverse Gaussian, and Birnbaum-Saunders kernel estimators, *Statistics and Probability Letters* **84**, 235–246.

Igarashi, G. & Kakizawa, Y. (2015). Bias correction for some asymmetric kernel estimators, *Journal of Statistical Planning and Inference* **159**, 37–63.

Kokonendji, C.C. & Somé, S.M. (2018). On multivariate associated kernels for smoothing some density function, *Journal of the Korean Statistical Society* **47**, 112–126.

Kokonendji, C.C. & Somé, S.M. (2021). Bayesian bandwidths in semiparametric modelling for nonnegative orthant data with diagnostics, *Stats* **4**, 162–183.

Libengué Dobélé-Kpoka, F.G.B. & Kokonendji, C.C. (2017) The mode-dispersion approach for constructing continuous associated kernels, *Afrika Statistica*. **12**, 1417–1446.

Malec, P. & Schienle, M. (2014). Nonparametric kernel density estimation near the boundary, *Computational Statistics and Data Analysis* **72**, 57–76.

R Development Core Team (2021). R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, URL `http://cran.r-project.org/`.

Scaillet, O. (2004). Density estimation using inverse and reciprocal inverse Gaussian kernels, *Journal of Nonparametric Statistics* **16**, 217–226.

Somé, S.M. (2021). Bayesian selector of adaptive bandwidth for gamma kernel estimator on $[0, \infty)$: simulations and applications, *Communications in Statistics-Simulation and Computation* forthcoming, doi:10.1080/03610918.2020.1828921.

Somé, S.M. & Kokonendji, C.C. (2021). Bayesian selector of adaptive bandwidth for multivariate gamma kernel estimator on $[0, \infty)^d$, *Journal of Applied Statistics* forthcoming, doi:10.1080/02664763.2021.1881456.

Zhang, S. (2010). A note on the performance of the gamma kernel estimators at the boundary, *Statistics and Probability Letters* **80**, 548–557.

Ziane, Y., Zougab, N. & Adjabi, S. (2015) Adaptive Bayesian bandwidth selection in asymmetric kernel density estimation for nonnegative heavy-tailed data, *Journal of Applied Statistics* **42**, 1645–1658.

Zougab, N., Adjabi, S. & Kokonendji, C.C. (2014). Bayesian estimation of adaptive bandwidth matrices in multivariate kernel density estimation, *Computational Statistics and Data Analysis* **75**, 28–38.