Statistics and Probability Letters 168 (2021) 108929

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Nonparametric recursive estimation of the copula

Felix Camirand Lemyre [a,b,*], Geoffrey Decrouez [c]

[a] Université de Sherbrooke, Canada
[b] Centre Hospitalier Universitaire de l'Université de Sherbrooke, Canada
[c] Zalando SE, Berlin, Germany

ARTICLE INFO

Article history:
Received 12 June 2019
Received in revised form 29 July 2020
Accepted 1 September 2020
Available online 9 September 2020

Keywords:
Asymptotic theory
Copula
Nonparametric estimation
Recursive methods
Stochastic approximation

ABSTRACT

This paper introduces two nonparametric recursive estimators of the copula. These estimators employ a recursive estimation of the quantile achieved using a stochastic approximation algorithm. Their asymptotic properties and numerical performance are investigated in the context of i.i.d. data.

1. Introduction

Copulas have become a major tool in statistics to capture and analyze the dependence structure of multivariate distributions. For a random vector $\mathbf{X} = (X_1, \dots, X_d)^\top \in \mathbb{R}^d$, with joint distribution function F and continuous marginals $F_1, \dots, F_d$, $d \geq 2$, Sklar's theorem ensures the existence of a unique function $C : [0, 1]^d \to [0, 1]$, called the copula, such that $C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$, where F_j^{-1} denotes the generalized inverse of F_j, i.e. $F_j^{-1}(u_j) \equiv q_j^{u_j} = \inf\{x \mid F_j(x) \geq u_j\}$.

In this study, we adopt a nonparametric approach to estimate the copula of $\mathbf{X}$ based on the observation of a sequence of independent and identically distributed random vectors $\mathbf{X}_1, \dots, \mathbf{X}_n$, where $\mathbf{X}_i \equiv (X_{i,1}, \dots, X_{i,d})^\top \in \mathbb{R}^d$ is distributed like $\mathbf{X}$. For $\mathbf{u} = (u_1, \dots, u_d)^\top \in (0, 1)^d$, put $q^{\mathbf{u}} \equiv (q_1^{u_1}, \dots, q_d^{u_d})^\top$. For any function $G : A \subset \mathbb{R}^d \to \mathbb{R}$, we write $G(\mathbf{a}) = G(a_1, \dots, a_d)$ for $\mathbf{a} = (a_1, \dots, a_d)^\top \in A$. In this notation, we have $C(\mathbf{u}) = F(q^{\mathbf{u}})$. The copula C can be estimated nonparametrically by replacing F by its empirical counterpart, and the quantiles $q_j^{u_j}$ by the quantile estimates $Q_{n,j}^{u_j} \equiv \inf_v\{n^{-1}\sum_{i=1}^n \mathbf{1}(X_{i,j} \leq v) \geq u_j\}$, yielding $\widehat{C}_n(\mathbf{u}) \equiv n^{-1}\sum_{i=1}^n \mathbf{1}\{\mathbf{X}_i \leq \mathbf{Q}_i^{\mathbf{u}}\}$, where $\leq$ is taken componentwise and where $\mathbf{Q}_n^{\mathbf{u}} \equiv (Q_{n,1}^{u_1}, \dots, Q_{n,d}^{u_d})^\top$. This estimator was originally studied by Deheuvels (1979) under the assumption of i.i.d. data with independent margins. In what follows, we introduce a recursive version of $\widehat{C}_n$, which can be updated in $O(1)$, so that the total cost after n observations are received is $O(n)$. This is achieved by considering a recursive version of $\mathbf{Q}_n^{\mathbf{u}}$, obtained by means of a stochastic approximation algorithm.

Stochastic approximation algorithms were introduced by Robbins and Monro (1951), as a method for estimating the zeros of a function observed only through noisy observations. The quantile $q_j^{u_j}$ is precisely defined as the root of a function, since it corresponds to the value of q satisfying $\mathbf{E}\{u_j - \mathbf{1}(X_j \leq q)\} = 0$. Estimation of the expected value, together with the approximation of the root is performed recursively via the sequence of estimates $Q_{n,j}^{u_j} = Q_{n-1,j}^{u_j} + \alpha_j n^{-1}\{u_j - \mathbf{1}(X_{n,j}$

* Corresponding author at: Université de Sherbrooke, Canada.
E-mail addresses: felix.camirand.lemyre@usherbrooke.ca (F. Camirand Lemyre), geoffrey.decrouez@zalando.de (G. Decrouez).

https://doi.org/10.1016/j.spl.2020.108929
0167-7152/

$\leq Q^{u_j}_{n-1,j})\}$. Provided F_j has a continuous derivative f_j at $q_j^{u_j}$, and $0 < \alpha_j^{-1} < 2f_j(q_j^{u_j})$, Sacks (1958) showed that a central limit theorem holds for this sequence, with minimal asymptotic variance obtained by taking $\alpha_j = 1/f_j(q_j^{u_j})$. The constant α_j^{-1} can be replaced with a current estimate of the density f_j evaluated at $q_j^{u_j}$. Here, we follow the approach developed in Amiri and Thiam (2014), by plugging a smooth estimator of $f_j(q_j^{u_j})$, yielding a sequence of estimates $\boldsymbol{\mathcal{Q}}^{\mathbf{u}}_n \equiv (\mathcal{Q}^{u_1}_{n,1}, \dots, \mathcal{Q}^{u_d}_{n,d})^\top$. The smooth recursive quantile estimator $\boldsymbol{\mathcal{Q}}^{\mathbf{u}}_n$ is introduced formally in Section 2. It is then used in the expression of $\widehat{C}_n$ instead of $\mathbf{Q}^{\mathbf{u}}_n$ to define the recursive estimator $C_n(\mathbf{u}) = n^{-1}\sum_{i=1}^n \mathbf{1}(\mathbf{X}_i \leq \boldsymbol{\mathcal{Q}}^{\mathbf{u}}_{i-1})$.

The definition of C_n provides a new framework to estimate copulas nonparametrically. The present contribution does not attempt to cover refinements of it. Instead, it is aimed at deriving fundamental properties of such estimators, opening the way to several research avenues, see e.g. Section 5. It is worth mentioning that, apart from the recursive quantile estimation, nonparametric recursive methods have been studied in other statistical frameworks, including density estimation (Hall and Patil, 1994; Mokkadem et al., 2009) and joint distribution estimation (Slaoui, 2014). Also, copulas have been used in settings where data are collected sequentially, see e.g. Harvey (2010), and Bennafla et al. (2016).

2. Definition of the recursive estimators of the copula

We start with the definition of the recursive quantile. Let K be a probability density called kernel, put $K_a(u) \equiv K(u/a)/a$ for $a > 0$ and $H(z) = \int_{-\infty}^z K(u)du$. Let $b_n \equiv (b_{n,1}, \dots, b_{n,d})$ and $\tilde{h}_n \equiv (\tilde{h}_{n,1}, \dots, \tilde{h}_{n,d})$ be two sequences of bandwidths. Denote by $\mathcal{Q}^{u_j}_{n,j}$ the estimator of $q_j^{u_j}$ at iteration $n \geq 1$. For constants $\mathcal{Q}^{u_j}_{0,j}$, $f^{u_j}_{0,j}$ and $\kappa \in (0, 2)$, we consider

$$\mathcal{Q}^{u_j}_{n,j} = \mathcal{Q}^{u_j}_{n-1,j} + (n\kappa a^{u_j}_{n-1,j})^{-1}\left[u_j - H\left\{b^{-1}_{n,j}(\mathcal{Q}^{u_j}_{n-1,j} - X_{n,j})\right\}\right], \tag{2.1}$$

where $a^{u_j}_{n,j} = \max[\mu_j, \min\{f^{u_j}_{n,j}, \nu(\log n + 1)\}]$, with $f^{u_j}_{n,j} = (1 - n^{-1})f^{u_j}_{n-1,j} + n^{-1}K_{\tilde{h}_{n,j}}(\mathcal{Q}^{u_j}_{n-1,j} - X_{n,j})$, for a given choice of $\nu, \mu_j > 0$.

In the case $\kappa = 1$, it is shown in Amiri and Thiam (2014) that the sequence $n^{1/2}(\mathcal{Q}^{u_j}_{n,j} - q_j^{u_j})$ converges weakly to a normal distribution. In the general case where $\kappa \in (0, 2)$, a simple adaptation of their proof yields that the latter sequence converges in distribution to a centered normal random variable with asymptotic variance $\{\kappa(2-\kappa)f_j(q_j^{u_j})^2\}^{-1}u_j(1-u_j)$.

We make use of the recursion (2.1) on the entries of $\boldsymbol{\mathcal{Q}}^{\mathbf{u}}_n \equiv (\mathcal{Q}^{u_1}_{n,1}, \dots, \mathcal{Q}^{u_d}_{n,d})^\top$ to define a recursive empirical estimator of C evaluated at $\mathbf{u} \in (0, 1)^d$, namely

$$C_n(\mathbf{u}) \equiv \left(1 - n^{-1}\right)C_{n-1}(\mathbf{u}) + n^{-1}\mathbf{1}(\mathbf{X}_n \leq \boldsymbol{\mathcal{Q}}^{\mathbf{u}}_{n-1}). \tag{2.2}$$

Compared with the batch estimator $\widehat{C}_n(\mathbf{u})$, the recursive estimator at (2.2) requires $O(n)$ operations and can be updated as new data are collected.

We also introduce a smooth version of (2.2). Put $H_a(x) \equiv H(xa^{-1})$ for $a \in \mathbb{R}$, and $W_{\mathbf{a}}(\mathbf{x}) = \prod_{\ell=1}^d H_{a_\ell}(x_\ell)$ for $\mathbf{x}, \mathbf{a} \in \mathbb{R}^d$. Consider the bandwidths $h_n = (h_{n,1}, \dots, h_{n,d})$. The recursive kernel copula estimator is then defined as $\mathcal{C}_n(\mathbf{u}) = (1 - n^{-1})\mathcal{C}_{n-1}(\mathbf{u}) + n^{-1}W_{h_n}(\boldsymbol{\mathcal{Q}}^{\mathbf{u}}_{n-1} - \mathbf{X}_n)$.

3. Asymptotic normality

For $\mathbf{u} = (u_1, \dots, u_d)^\top$, put $\mathbf{u}^{(\ell)}(v) \equiv (u_1, \dots, u_{\ell-1}, u_\ell \wedge v, u_{\ell+1}, \dots, u_d)$ and denote by $\mathbf{u} \wedge \mathbf{v}$ the vector of element-wise minimums. Let $C_{\ell,m}$ denote the copula of (X_ℓ, X_m), $1 \leq \ell, m \leq d$, and for any $\mathbf{u}, \mathbf{v} \in (0, 1)^d$, let $\sigma^{(1)}(\mathbf{u}, \mathbf{v}) = C(\mathbf{u} \wedge \mathbf{v}) - C(\mathbf{u})C(\mathbf{v})$, $\sigma^{(2)}(\mathbf{u}, \mathbf{v}) = \sum_{\ell=1}^d C^{(\ell)}(\mathbf{u})[C\{\mathbf{v}^{(\ell)}(u_\ell)\} - u_\ell C(\mathbf{v})]$ and $\sigma^{(3)}(\mathbf{u}, \mathbf{v}) = \sum_{\ell,m=1}^d C^{(\ell)}(\mathbf{u})C^{(m)}(\mathbf{v})[C_{\ell,m}(u_\ell, v_m) - u_\ell v_m]$, where $C^{(\ell)}$ denotes the partial derivative of C with respect to the ℓth variable. To establish our asymptotic results, we consider the conditions below.

($\mathcal{B}$) For any $j \in \{1, \dots, d\}$, $\tilde{h}_{n,j} \to 0$ as $n \to \infty$, $\sum_{k\geq 1} k^{\delta-1}b^2_{k,j} < \infty$ for any $\delta < 1/2$, and $\sum_k (k\tilde{h}_{k,j})^{-2} < \infty$.

($\mathcal{K}$) K is a Lipschitz continuous density that satisfies $\int uK(u)du = 0$ and $\nu \equiv \int u^2K(u)\,du < \infty$.

($\mathcal{D}$) There exists $[a, b] \subset (0, 1)$ such that *(i)* for $j \in \{1, \dots, d\}$ the density f_j is strictly positive and twice continuously differentiable over $[F_j^{-1}(a), F_j^{-1}(b)]$, and *(ii)* F has continuous second order partial derivatives over $\mathcal{T}_{a,b} \equiv [F_1^{-1}(a), F_1^{-1}(b)] \times \dots \times [F_d^{-1}(a), F_d^{-1}(b)]$.

($\mathcal{B}^\star$) As $n \to \infty$, $n(h_{n,1}, \dots, h_{n,d}) \to \infty$ and $n^{1/2}(h^2_{n,1}, \dots, h^2_{n,d}) \to (z_1, \dots, z_d)$ for some $z_1, \dots, z_d \in [0, \infty)$.

Assumptions ($\mathcal{B}$), ($\mathcal{K}$) and ($\mathcal{D}$)–(i) are multivariate analogs of the ones used in Amiri and Thiam (2014) and guarantee an almost sure consistency rate for the recursive quantile estimators. Assumption ($\mathcal{D}$)–(ii) is satisfied, for example, by the multivariate gaussian distribution and by any distribution obtained by combining a copula that is twice continuously differentiable on any compact set $A \subset (0, 1)^d$, e.g. the Clayton, the normal or the Gumbel copula, with a set of marginals satisfying ($\mathcal{D}$)–(i). Assumption ($\mathcal{B}^\star$) is used to control the bias and the variance of the recursive kernel copula estimator. Regarding Condition ($\mathcal{D}$)–(ii), we note that our proof of Theorem 1 only requires that F has bounded first-order partial derivatives over $\mathcal{T}_{a,b}$.

Theorem 1. *Suppose that $\mu_j < f_j(q_j^{u_j})$, for $1 \le j \le d$. Under Assumptions $(\mathcal{B})$, $(\mathcal{K})$ and $(\mathcal{D})$, for any $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k \in [a,b]^d \subset (0,1)^d$, $k \ge 1$, where $[a,b]$ is as in Condition $(\mathcal{D})$, and any $\kappa \in (0,2)$, $\kappa \neq 1$, the sequence $n^{1/2}\mathbf{C}_n \equiv n^{1/2}(C_n(\boldsymbol{u}_1) - C(\boldsymbol{u}_1), \ldots, C_n(\boldsymbol{u}_k) - C(\boldsymbol{u}_k))^\top$ converges weakly to a centered multivariate normal distribution with covariance matrix $\Sigma^C = \Sigma^C(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ whose entries satisfy $\Sigma^C_{ij} = \sigma^{(1)}_{ij} - \sigma^{(2)}_{ij} - \sigma^{(2)}_{ji} + 2(2-\kappa)^{-1}\sigma^{(3)}_{ij}$, where $\sigma^{(\ell)}_{ij} \equiv \sigma^{(\ell)}(\boldsymbol{u}_i, \boldsymbol{u}_j)$, for $i,j \in \{1,\ldots,k\}$ and $\ell \in \{1,2,3\}$.*

The empirical copula process $C^{\mathbb{E}}_n \equiv n^{1/2}(\widehat{C}_n - C)$, a function on $[0,1]^d$, converges in distribution to a centered gaussian process, see Segers (2012). A consequence of this result is that the random vector $(C^{\mathbb{E}}_n(\mathbf{u}_1), \ldots, C^{\mathbb{E}}_n(\mathbf{u}_k))^\top$ is asymptotically normally distributed with covariance matrix $\Sigma^{\mathbb{E}}$ satisfying $\Sigma^{\mathbb{E}}_{ij} = \sigma^{(1)}_{ij} - \sigma^{(2)}_{ij} - \sigma^{(2)}_{ji} + \sigma^{(3)}_{ij}$. Since the convergence of $n^{1/2}\mathbf{C}_n$ requires $\kappa > 0$, each entry of their asymptotic covariance matrix differs by $\kappa(2-\kappa)^{-1}\sigma^{(3)}_{ij}$. Therefore, under our assumptions, the asymptotic variance of $C_n(\mathbf{u}_i)$ is always larger than the one of $\widehat{C}_n(\mathbf{u}_i)$ since $\sigma^{(3)}_{ii} > 0$. This also suggests that computing C_n with a small value for κ might lead to a lower-variance estimator, at least when n is large.

Theorem 2. *Suppose that K is compactly supported and vanishes outside of $[-1,1]$. Under the Assumptions of Theorem 1 and of $(\mathcal{B}^\star)$, the random vector $n^{1/2}\boldsymbol{\mathcal{C}}_n \equiv n^{1/2}(\mathcal{C}_n(\boldsymbol{u}_1) - C(\boldsymbol{u}_1), \ldots, \mathcal{C}_n(\boldsymbol{u}_k) - C(\boldsymbol{u}_k))^\top$ with $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k \in [a,b]^d \subset (0,1)^d$, $k \ge 1$, converges in distribution to a multivariate normal distribution with covariance matrix Σ^C and mean $(B(q^{\boldsymbol{u}_1}), \ldots, B(q^{\boldsymbol{u}_k}))^\top$, where $B(\mathbf{x}) \equiv (\nu/2)\sum_{j=1}^d z_j\, \partial^2 F(\mathbf{x})/(\partial x_j^2)$.*

The proof of Theorem 2 is similar to that of Theorem 1 and is presented in the supplementary material accompanying this paper, see Appendix B.

4. Numerical study

We compare the finite sample behavior of the empirical and kernel recursive copula estimators for different values of $\kappa \in \{0.5, 0.2, 0.1\}$ in a context where the recursive estimators are initiated with $n_0 = 15$ observations and then continuously updated until the total sample size reaches $n \in \{100, 1000\}$, using the Epanechnikov kernel $K(x) = 3/4(1-x^2)\mathbf{1}(|x| \le 1)$. To assess their performance, we use 100 Monte Carlo simulations. Specifically, for each estimator, we calculate its mean square errors (MSE) for $\mathbf{u} \in \{0.1, 0.2, \ldots, 0.9\}^d$ and consider the average of these values. These results are presented in light of the one of the empirical batch copula $\widehat{C}_n$. The data is generated from a selection of bivariate ($d = 2$) parametric copulas, namely the Gaussian and Clayton copulas, and the marginals are taken as standard normal distributions. Regarding the choice of bandwidths, note that Theorems 1 and 2 suggest that b_n and $\tilde{h}_n$ only have a second order effect on the asymptotic behavior of the recursive copula estimators; this was also supported by many numerical experiments not reported here. Hence, we use Silverman's rule of thumb and set $\tilde{h}_{n,j} = 1.06 \times \sigma_{n,j} \times n^{-1/5}$, where $\sigma_{n,j}$ is the standard deviation of the sample $\{X_{1,j}, \ldots, X_{n,j}\}$ and we put $b_{n,j} = \tilde{h}_{n,j} \times n^{-0.06}$ so that resulting bandwidths satisfy condition $(\mathcal{B})$. To ensure that the recursive kernel copula estimator is asymptotically unbiased, we set $h_n = b_n$. Other bandwidth selection methods have been investigated for h_n, such as plug-in strategies and copula-based rules of thumb, but none of them led to conclusive improvements. The constants $\mathcal{Q}^{u_j}_{n_0,j}$ and $f^{u_j}_{n_0,j}$ are respectively taken as the empirical quantile and the corresponding kernel density estimator computed through the n_0 first observations, and ν is set to 1. Finally, instead of taking a fixed value for μ_j, at iteration n, we set $\mu_j = 0.1 \times n^{-1/4}$. It can be shown that this modification has no impact on our theoretical results. The above procedure ensures that $O(1)$ operations are required at each iteration.

The results are presented in Table 1. As expected, the larger the sample size n, the smaller the MSE. In cases where the strength of the dependence is low or mild, the MSE of the recursive estimators with $\kappa = 0.1$ is smaller than the one of the empirical copula. In these cases, in agreement with Theorems 1 and 2, the difference gets smaller as the sample size increases. Also, the accuracy is of the same order across simulations, irrespectively of the underlying copula used to generate the data. This small numerical study shows the suitability of the recursive estimators, in particular when used with a small value of κ.

5. Discussion

This paper investigated two nonparametric recursive estimators for a copula function. The methodology could be improved, for e.g. by designing a data-driven selection method for the parameters involved. In nonparametric frameworks, such schemes often involve minimizing an approximation of the corresponding MSE with respect to the targeted parameters. The latter task is especially challenging in our context, since the whole procedure has to be executed recursively, and it is not obvious how typical approaches such as cross-validation or iterative plug-in methods could be adapted here. Our methodology could also be extended, for example, to the problem of estimating recursively a copula density, a copula functional or a conditional copula. Also, to allow for the recursive estimation of a time-varying copula, linear filtering strategies could be employed to render the proposed estimators suitable to dynamic settings. All of the above considerations are quite challenging and are the subject of current research by the authors.

Table 1
Average MSE$\times 10^4$ of the empirical (C_n) and kernel ($\mathcal{C}_n$) recursive copula estimators compared to the batch empirical copula, for $n \in \{100, 1000\}$ and $\kappa \in \{0.5, 0.2, 0.1\}$.

			$\kappa = 0.5$		$\kappa = 0.2$		$\kappa = 0.1$		Batch
			C_n.	$\mathcal{C}_n$	C_n	$\mathcal{C}_n$	C_n	$\mathcal{C}_n$	$\widehat{C}_n$.
$n = 100$	Gaussian	$\rho = 0.05$	5.22	4.16	3.88	4.10	2.90	2.60	3.22
		$\rho = 0.2$	5.49	4.16	3.78	4.23	2.87	2.66	3.21
		$\rho = 0.4$	5.33	3.78	3.70	4.24	2.74	2.71	2.95
		$\rho = 0.6$	5.11	3.45	3.43	4.05	2.50	2.68	2.31
	Clayton	$\theta = -0.33$	4.83	4.07	4.08	3.68	2.87	2.99	3.13
		$\theta = 0.1$	5.55	4.59	4.49	4.29	3.29	3.17	3.61
		$\theta = 1.3$	4.97	4.32	4.77	3.94	3.19	3.79	2.62
		$\theta = 3.0$	4.02	3.14	3.98	3.26	2.49	3.59	1.49
$n = 1000$	Gaussian	$\rho = 0.05$	0.48	0.40	0.38	0.37	0.29	0.26	0.31
		$\rho = 0.2$	0.49	0.39	0.37	0.36	0.28	0.26	0.30
		$\rho = 0.4$	0.47	0.38	0.36	0.35	0.26	0.26	0.28
		$\rho = 0.6$	0.41	0.33	0.33	0.31	0.24	0.26	0.21
	Clayton	$\theta = -0.33$	0.45	0.39	0.39	0.35	0.29	0.29	0.31
		$\theta = 0.1$	0.48	0.43	0.42	0.36	0.31	0.30	0.34
		$\theta = 1.3$	0.40	0.36	0.39	0.30	0.26	0.31	0.24
		$\theta = 3.0$	0.32	0.27	0.34	0.25	0.22	0.31	0.13

CRediT authorship contribution statement

Felix Camirand Lemyre: Methodology, Investigation , Writing - original draft, Writing - review & editing. **Geoffrey Decrouez:** Methodology, Investigation, Writing - original draft, Writing - review & editing.

Acknowledgments

This research was undertaken thanks in part to the Natural Sciences and Engineering Research Council of Canada, the Canada First Research Excellence Fund and by Australian Research Council Discovery Project DP140100125.

Appendix A. Proof of Theorem 1

For any $\mathbf{u} \in \mathbb{R}^d$, let $U_n^{\mathbf{u}} \equiv C_n(\mathbf{u}) - C(\mathbf{u})$, $U_{n,j}^{u_j} \equiv \mathcal{Q}_{n,j}^{u_j} - q_j^{u_j}$ and $\bar{U}_n^{\mathbf{u}} \equiv \kappa(1-\kappa)^{-1}\sum_{j=1}^d F^{(j)}(q^{\mathbf{u}})U_{n,j}^{u_j}$. We first prove that for $\mathbf{U}_n^{\mathbf{u}_j} \equiv [U_n^{\mathbf{u}_j} + \bar{U}_n^{\mathbf{u}_j}, \bar{U}_n^{\mathbf{u}_j}]^\top$, the random vector $n^{1/2}\mathbf{U}_n \equiv n^{1/2}[(\mathbf{U}_n^{\mathbf{u}_1})^\top \dots (\mathbf{U}_n^{\mathbf{u}_k})^\top]^\top$ converges in distribution to a multivariate centered normal distribution with $2k \times 2k$ covariance matrix $\Sigma^U = [\Sigma_{ij}^U]$ defined by blocks, where, for $i, j \in \{1, \dots, k\}$, Σ_{ij}^U is a 2×2 symmetric matrix with entries $\Sigma_{ij}^U(1,1) = \sigma_{ij}^{(1)} - (1-\kappa)^{-1}\sigma_{ij}^{(2)} - (1-\kappa)^{-1}\sigma_{ji}^{(2)} + (1-\kappa)^{-2}\sigma_{ij}^{(3)}$, $\Sigma_{ij}^U(1,2) = \Sigma_{ji}^U(2,1) = \kappa\{-(1-\kappa)^{-1}\sigma_{ij}^{(2)} + (1-\kappa)^{-2}\sigma_{ij}^{(3)}\}$ and $\Sigma_{ij}^U(2,2) = \kappa(2-\kappa)^{-1}(1-\kappa)^{-2}\sigma_{ij}^{(3)}$. To do this, we use a result due to Fabian (1968) that allows to establish the asymptotic normality of random vectors satisfying a type of recurrence relation that is typically encountered in stochastic approximation algorithms. For completeness, a version of Fabian's theorem adapted to our context can be found in the supplementary material of this paper, see Appendix B. To apply that result to $n^{1/2}\mathbf{U}_n$, we establish in *Step I* a recurrence relation for $\mathbf{U}_n$ in (see (A.1)). Then, we verify in *Step II* that the terms involved in the recurrence relation satisfy the conditions required to apply Fabian's Theorem. *Step III* derives the asymptotic normality of $\mathbf{C}_n$ from that of $\mathbf{U}_n$.

Step I. Let $\mathcal{F}_n$ be the σ–algebra generated by $\mathbf{X}_1, \dots, \mathbf{X}_n$. For $j \in \{1, \dots, d\}$, set $V_{n,j}^{u_j} \equiv \mathbf{1}(X_{n+1,j} \le \mathcal{Q}_{n,j}^{u_j}) - F_j(\mathcal{Q}_{n,j}^{u_j})$ and $G_{n,j}^{u_j} \equiv F_j(\mathcal{Q}_{n,j}^{u_j}) - u_j$, and put $V_n^{\mathbf{u}} \equiv \mathbf{1}(\mathbf{X}_{n+1} \le \mathcal{Q}_n^{\mathbf{u}}) - F(\mathcal{Q}_n^{\mathbf{u}})$, $\bar{V}_n^{\mathbf{u}} \equiv (1-\kappa)^{-1}\sum_{j=1}^d C^{(j)}(\mathbf{u})V_{n,j}^{u_j}$ and $G_n^{\mathbf{u}} \equiv F(\mathcal{Q}_n^{\mathbf{u}}) - C(\mathbf{u})$. Note that $\mathbf{E}(V_n^{\mathbf{u}} \mid \mathcal{F}_n) = \mathbf{E}(V_{n,j}^{u_j} \mid \mathcal{F}_n) = 0$ and that $\mathbf{E}\left(V_n^{\mathbf{u}}V_n^{\mathbf{v}} \mid \mathcal{F}_n\right) = F(\mathcal{Q}_n^{\mathbf{u}} \wedge \mathcal{Q}_n^{\mathbf{v}}) - F(\mathcal{Q}_n^{\mathbf{u}})F(\mathcal{Q}_n^{\mathbf{v}})$. We treat the terms $\bar{U}_n^{\mathbf{u}}$ and $U_n^{\mathbf{u}}$ appearing in the expression of $\mathbf{U}_n^{\mathbf{u}}$ separately, and we first deal with $\bar{U}_n^{\mathbf{u}}$. Under our assumptions, we deduce from Theorem 1 in Amiri and Thiam (2014) that $a_{n,j}^{u_j} = f_j(q_j^{u_j}) + o_{a.s}(1)$. A Taylor expansion of order 1 of F_j around $q_j^{u_j}$ entails $(a_{n,j}^{u_j})^{-1}G_{n,j}^{u_j} = (\mathcal{Q}_{n,j}^{u_j} - q_j^{u_j})\{1 + o_{a.s}(1)\} = U_{n,j}^{u_j}(1 + o_{a.s}(1))$. Making use of (2.1), straightforward computations shows that $U_{n,j}^{u_j} = \{1 - n^{-1}(\kappa^{-1} + o_{a.s}(1))\}U_{n-1,j}^{u_j} - \kappa^{-1}(na_{n-1,j}^{u_j})^{-1}V_{n-1,j}^{u_j}$, yielding $\bar{U}_n^{\mathbf{u}} = \{1 - n^{-1}(\kappa^{-1} + o_{a.s}(1))\}\bar{U}_{n-1}^{\mathbf{u}} - n^{-1}(1-\kappa)^{-1}\sum_{j=1}^d F^{(j)}(q^{\mathbf{u}})(a_{n-1,j}^{u_j})^{-1}V_{n-1,j}^{u_j}$. Using the fact that $a_{n,j}^{u_j} = f_j(q_j^{u_j}) + o_{a.s}(1)$, we deduce that $F^{(j)}(q^{\mathbf{u}})(a_{n-1,j}^{u_j})^{-1} = F^{(j)}(q^{\mathbf{u}})(f_j(q^{\mathbf{u}}))^{-1} + o_{a.s}(1) = C^{(j)}(\mathbf{u}) + o_{a.s}(1)$. Since $\bar{V}_n^{\mathbf{u}} = (1-\kappa)^{-1}\sum_{j=1}^d C^{(j)}(\mathbf{u})V_{j,n}^{u_j}$, we obtain from these computations that $\bar{U}_n^{\mathbf{u}} = \{1 - n^{-1}(\kappa^{-1} + o_{a.s}(1))\}\bar{U}_{n-1}^{\mathbf{u}} - n^{-1}\{1 + o_{a.s}(1)\}\bar{V}_{n-1}^{\mathbf{u}}$. Turning our attention to $U_n^{\mathbf{u}}$, we have from (2.2) the decomposition $U_n^{\mathbf{u}} = (1 - n^{-1})U_{n-1}^{\mathbf{u}} + n^{-1}G_{n-1}^{\mathbf{u}} + n^{-1}V_{n-1}^{\mathbf{u}}$. Under Assumption $(\mathcal{D})$, a Taylor expansion and Theorem 1 in Amiri and Thiam (2014) give $G_{n-1}^{\mathbf{u}} = \kappa^{-1}(1-\kappa)\{1 + o_{a.s}(1)\}\bar{U}_{n-1}^{\mathbf{u}}$. It therefore follows that $U_n^{\mathbf{u}} = \left(1 - n^{-1}\right)U_{n-1}^{\mathbf{u}} + n^{-1}\kappa^{-1}(1-\kappa)\{1 + o_{a.s}(1)\}\bar{U}_{n-1}^{\mathbf{u}} + n^{-1}V_{n-1}^{\mathbf{u}}$. The latter recursive expressions for $\bar{U}_n^{\mathbf{u}}$ and $U_n^{\mathbf{u}}$

allow us to derive a recursion for $U_n + \bar{U}_n$. Indeed, as $\kappa^{-1}(1-\kappa) - \kappa^{-1} = -1$, adding these terms entails $U_n^{\mathbf{u}} + \bar{U}_n^{\mathbf{u}} = \left(1 - n^{-1}\{1 + o_{\mathrm{a.s}}(1)\}\right)\left(U_{n-1}^{\mathbf{u}} + \bar{U}_{n-1}^{\mathbf{u}}\right) + n^{-1}\{1 + o_{\mathrm{a.s}}(1)\}\left(V_{n-1}^{\mathbf{u}} - \bar{V}_{n-1}^{\mathbf{u}}\right)$. Collecting the recursive formulas for $\bar{U}_n^{\mathbf{u}}$ and $U_n^{\mathbf{u}} + \bar{U}_n^{\mathbf{u}}$ derived above entails $\mathbf{U}_n^{\mathbf{u}} = (\mathcal{I}_2 - n^{-1}\{\overline{\Gamma} + o_{\mathrm{a.s}}(1)\})\mathbf{U}_{n-1}^{\mathbf{u}} + n^{-1}\{\mathcal{I}_2 + o_{\mathrm{a.s}}(1)\}\mathbf{V}_{n-1}^{\mathbf{u}}$, where $\mathbf{V}_n^{\mathbf{u}} \equiv [(V_n^{\mathbf{u}} - \bar{V}_n^{\mathbf{u}})\ (-\bar{V}_n^{\mathbf{u}})]$ and $\overline{\Gamma} = \mathrm{diag}[1\ \kappa^{-1}]$. To obtain a recursion equation for $\mathbf{U}_n$, we introduce $\widetilde{\mathbf{V}}_n \equiv [\mathbf{V}_n^{\mathbf{u}_1}\ \ldots\ \mathbf{V}_n^{\mathbf{u}_k}]^\top$ and $\Gamma \equiv \mathrm{diag}[1\ \kappa^{-1}\ \ldots\ 1\ \kappa^{-1}]$, where in the definition of Γ, the vector $[1\ \kappa^{-1}]$ is repeated k times. Since the recursive relation for $\mathbf{U}_n^{\mathbf{u}}$ holds for any $\mathbf{u} \in \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$, we finally obtain

$$\mathbf{U}_n = \left(\mathcal{I}_{2k} - n^{-1}\{\Gamma + o_{\mathrm{a.s}}(1)\}\right)\mathbf{U}_{n-1} + n^{-1}\{\mathcal{I}_{2k} + o_{\mathrm{a.s}}(1)\}\widetilde{\mathbf{V}}_{n-1}. \tag{A.1}$$

Step II. In view of (A.1) and of the theorem presented in Section 2.2 of Fabian (1968), the asymptotic normality of $n^{1/2}\mathbf{U}_n$ follows if we show that $\mathbf{E}\{\mathbf{V}_n^{\mathbf{u}_i}(\mathbf{V}_n^{\mathbf{u}_j})^\top \mid \mathcal{F}_n\} = \Sigma_{ij} + o_{\mathrm{a.s}}(1)$, where Σ_{ij} is a matrix with entries $\Sigma_{ij}(1,1) = \Sigma_{ij}^U(1,1)$, $\Sigma_{ij}(2,1) = \kappa^{-1}\Sigma_{ij}^U(2,1)$, $\Sigma_{ij}(1,2) = \kappa^{-1}\Sigma_{ij}^U(1,2)$ and $\Sigma_{ij}(2,2) = (2-\kappa)\kappa^{-1}\Sigma_{ij}^U(2,2)$. To do this, note that $\mathbf{E}\{V_n^{\mathbf{u}_i}V_n^{\mathbf{u}_j} \mid \mathcal{F}_n\} = F(\mathcal{Q}_n^{\mathbf{u}_i} \wedge \mathcal{Q}_n^{\mathbf{u}_j}) - F(\mathcal{Q}_n^{\mathbf{u}_i})F(\mathcal{Q}_n^{\mathbf{u}_j}) = C(\mathbf{u}_i \wedge \mathbf{u}_j) - C(\mathbf{u}_i)C(\mathbf{u}_j) + o_{\mathrm{a.s}}(1) = \sigma_{ij}^{(1)} + o_{\mathrm{a.s}}(1)$, since $\mathcal{Q}_n^{\mathbf{u}} = q^{\mathbf{u}} + o_{\mathrm{a.s}}(1)$. Proceeding similarly for $\mathbf{E}(\bar{V}_n^{\mathbf{u}_i}\bar{V}_n^{\mathbf{u}_j} \mid \mathcal{F}_n)$ and $\mathbf{E}(\bar{V}_n^{\mathbf{u}_i}\bar{V}_n^{\mathbf{u}_j} \mid \mathcal{F}_n)$ combined with straightforward algebra manipulations yields the desired result. Thus, the asymptotic normality of $n^{1/2}\mathbf{U}_n$ is ensured.

Step III. Put $\mathbf{e} = [1\ -1]$. From the definition of $\mathbf{U}_n^{\mathbf{u}}$, we have $U_n^{\mathbf{u}} = \mathbf{e}\,\mathbf{U}_n^{\mathbf{u}}$. Let $\mathbf{R}$ be the $k \times 2k$ matrix such that $[\mathbf{R}_{i,2i-1}\ \mathbf{R}_{i,2i}] = \mathbf{e}$ and zero elsewhere, so that $\mathbf{C}_n = \mathbf{R}\,\mathbf{U}_n$. The result follows from the continuous mapping theorem and the asymptotic normality of $n^{1/2}\mathbf{U}_n$. □

Appendix B. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spl.2020.108929.

References

Amiri, A., Thiam, B., 2014. A smoothing stochastic algorithm for quantile estimation. Statist. Probab. Lett. 93, 116–125.

Bennafla, D., Bouchentouf, A., Rabhi, A., Sabri, K., 2016. On the recursive estimation using copula function in theregression model. Appl. Intell. 4 (1), 25–37.

Deheuvels, P., 1979. La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. Acad. Roy. Belg. Bull. Cl. Sci. (5) 65, 274–292.

Fabian, V., 1968. On asymptotic normality in stochastic approximation. Ann. Math. Stat. 39 (4), 1327–1332.

Hall, P., Patil, P., 1994. On the efficiency of on-line density estimators. IEEE Trans. Inf. Theory 40 (5), 1504–1512.

Harvey, A., 2010. Tracking a changing copula. J. Empir. Finance 17 (3), 485–500.

Mokkadem, A., Pelletier, M., Slaoui, Y., 2009. The stochastic approximation method for the estimation of a multivariate probability density. J. Statist. Plann. Inference 139, 2459–2478.

Robbins, H., Monro, S., 1951. A stochastic approximation method. Ann. Math. Stat. 22, 400–407.

Sacks, J., 1958. Asymptotic distribution of stochastic approximation procedures. Ann. Math. Stat. 29, 373–405.

Segers, J., 2012. Weak convergence of empirical copula processes under nonrestrictive smoothness assumptions. Bernoulli 18 (3), 764–782.

Slaoui, Y., 2014. The stochastic approximation method for estimation of a distribution function. Math. Methods Statist. (4), 306.