Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 91 (2016) 141 – 150

Procedia Computer Science

Information Technology and Quantitative Management (ITQM 2016)

Multivariate skew normal copula for non-exchangeable dependence

Zheng Wei[a], Seongyong Kim[b], Daeyoung Kim[a,*]

[a]*Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA, 01003-9305, USA*

[b]*Department of Applied Statistics, Hoseo University, Asan-si, Chungcheongnam-do, 31499, Republic of Korea*

Abstract

The exchangeability assumption on the dependence structure of the multivariate data is restrictive in practical situations where the variables of interest are not likely to be associated to each other in an identical manner. In this paper, we propose a flexible class of multivariate skew normal copulas to model high-dimensional non-exchangeable dependence patterns. The proposed copulas have two sets of parameters capturing non-exchangeable dependence, one for association between the variables and the other for skewness of the variables. In order to efficiently estimate the two sets of parameters, we introduce the block coordinate ascent algorithm. The proposed class of multivariate skew normal copulas is illustrated using a real data set.

Peer-review under responsibility of the Organizing Committee of ITQM 2016

Keywords: Copula; Non-exchangeability; Skew-normal distribution

1. Introduction

The copulas have been increasingly popular for modeling statistical dependence in multivariate data and have been applied to many areas including finance [1], medical research [2], econometrics [3], environmental science [4], actuarial science [5], just to name a few. A key feature of copulas is that they provide flexible representations of the multivariate distribution by allowing for the dependence structure of the variables of interest to be modeled separately from the marginal structure.

Most of the commonly used copulas are exchangeable, which means that the value of the copula is invariant under permutations of its arguments. For some practical situations where one component of the variables influences the other one more than the other way around, exchangeability assumption on copula is not suitable. This is because the dependence based on exchangeable copulas cannot distinguish between components of the variables.

*Corresponding author. Tel.: +1-413-577-0255 ; fax: +1-413-545-1801 .
E-mail address: daeyoung@math.umass.edu

Peer-review under responsibility of the Organizing Committee of ITQM 2016
doi:10.1016/j.procs.2016.07.051

Recently, several researchers have put considerable efforts into developing multivariate non-exchangeable copulas [6, 7, 8, 9]. The hierarchical Archimedean copulas are constructed from the idea of the compositions of simple Archimedean copulas [8]. However, they requires intricate compatibility conditions and there is no easy stochastic interpretation of them. The vine copulas [7, 10, 11], as popular hierarchical graphical models, utilize the bivariate copulas as the building blocks of a tree structure for multivariate distributions and their dependency structure is determined by a cascade of bivariate copulas. Although the vine copulas are highly flexible, they are not closed under the marginalization and the lack of the stochastic representation makes the interpretation of the vine copulas difficult. The Liouville copulas [9] are proposed as a non-exchangeable generalization of the Archimedean copulas. But, they have some limitations in applications due to unavailability of an algebraically tractable form, strictly positive support, and linear relationship between parameters and dimension.

In this paper we propose a flexible class of multivariate skew normal copulas to model high-dimensional non-exchangeable dependence patterns. The proposed skew normal copula derived from the multivariate skew normal distribution [12, 13] has the two sets of parameters capturing non-exchangeable dependence between the variables of interest, one for correlation between the variables and the other for skewness of the variables. Depending on the restrictions on these parameters, the proposed multivariate skew normal copula produces six parsimonious (non-exchangeable/exchangeable) copulas. We also propose using the block coordinate ascent algorithm to efficiently estimate the parameters in the proposed class of multivariate skew normal copulas. Instead of estimating all parameters simultaneously, the introduced algorithm partitions the parameters into two disjoint blocks, one for the correlation matrix and the other for skewness parameters, and update block by block.

This paper is organized as follows. In Section 2, we briefly review the concept of the multivariate copula and its non-exchangeable property. Section 3 proposes a class of the multivariate skew normal copulas that can capture various (non-exchangeable/exchangeable) dependence structures. In Section 4 we introduce the block coordinate algorithm for estimating the proposed copulas and discuss its convergence property. Section 5 illustrates the proposed class of skew normal copulas with a real data example. We end this article with a discussion in Section 6.

2. Review on multivariate non-exchangeable copulas

We here briefly review the multivariate non-exchangeable copula. For the general copula theory, see [14, 15, 16].

Definition 2.1. *A* $(k+1)$*-dimensional* **copula** *(or* $(k+1)$-**copula***) is a function* $C : [0,1]^{k+1} \mapsto [0,1]$ *satisfying following properties: for* $i = 0, \dots, k$,

(a) $C(u_0, u_1, \dots, u_k) = 0$ *if at least one* $u_i = 0$*;*

(b) $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$ *for every* $u_i \in [0,1]$*;*

(c) C *is* $(k+1)$*-increasing in the sense that, for any* $J = \prod_{i=0}^{k}[u_i, v_i] \subseteq [0,1]^{k+1}$ *with* $u_i, v_i \in [0,1]$, $volC(J) = \sum_{\boldsymbol{a}} sgn(\boldsymbol{a})C(\boldsymbol{a}) \geq 0$*, where the summation is over all vertices* $\boldsymbol{a}$ *of* J, $\boldsymbol{a} = (a_0, a_1, \dots, a_k)^T$ *is the transpose of* $(a_0, a_1, \dots, a_k)$*, and* $a_i = u_i$ *or* v_i,

$$sgn(\boldsymbol{a}) = \begin{cases} 1, & \text{if} \quad a_i = v_i \quad \text{for an even number of } i's, \\ -1, & \text{if} \quad a_i = v_i \quad \text{for an odd number of } i's. \end{cases}$$

From Definition 2.1, we can see that a $(k+1)$-copula is a joint cumulative distribution function (CDF) on $[0,1]^{k+1}$ with standard uniform marginal distributions.

Let $\boldsymbol{X} = (X_0, X_1, \ldots, X_k)^T$ be a $(k+1)$-dimensional random vector with the CDF $H(x_0, x_1, \ldots, x_k)$, and marginal CDF's $F_0(x_0)$, $F_1(x_1)$, …, $F_k(x_k)$. Sklar's theorem [17] states that if the marginals of $\boldsymbol{X}$ are continuous, then there exist a unique copula C such that

$$H(x_0, x_1, \ldots, x_k) = C(F_0(x_0), F_1(x_1), \ldots, F_k(x_k)).$$

The important properties of the copula C of $\boldsymbol{X}$ are that the copula C represents the dependence structure of $\boldsymbol{X}$ on a quantile scale and it is invariant under strictly increasing transformations of the marginals.

For an absolutely continuous copula C, the copula density is defined to be

$$c(u_0, u_1, \ldots, u_k) = \frac{\partial^{k+1} C(u_0, u_1, \ldots, u_k)}{\partial u_0 \partial u_1 \cdots \partial u_k}.$$

Given a random vector $\boldsymbol{X}$, with an absolute continuous H and strictly increasing continuous marginals $F_0(x_0)$, $F_1(x_1)$, …, $F_k(x_k)$, the copula density defined above is given by

$$c(u_0, u_1, \ldots, u_k) = \frac{h(F_0^{-1}(u_0), F_1^{-1}(u_1), \ldots, F_k^{-1}(u_k))}{f_0(F_0^{-1}(u_0)) f_1(F_1^{-1}(u_1)) \ldots f_k(F_k^{-1}(u_k))}, \tag{1}$$

where $h(x_0, x_1, \ldots, x_k)$ is the joint density of $H(x_0, x_1, \ldots, x_k)$ and $f_0, f_1, \ldots, f_k$ are the marginal densities.

Most of the commonly used copulas for applied research (for example, Archimedean copulas and all meta-elliptical copulas) assume that the dependence structure between the variables of interest is exchangeable.

Definition 2.2. *A $(k+1)$-copula C is exchangeable if it is the distribution function of a $(k+1)$-dimensional exchangeable uniform random vector $\boldsymbol{U} = (U_0, U_1, \ldots, U_k)^T$ satisfying $C(u_0, u_1, \ldots, u_k) = C(u_{\sigma(0)}, u_{\sigma(1)}, \ldots, u_{\sigma(k)})$ for any permutation $\sigma \in \Gamma$ where Γ denotes the set of all permutations on the set $\{0, 1, \ldots, k\}$.*

The exchangeability assumption on copulas is too restrictive for some applied situations where one component of the variables influences the other one more than the other way around. In the bivariate case exchangeability means that the conditional probability distribution of the first variable given the second variable equals the conditional probability distribution of the second variable given the first variable.

3. Class of multivariate skew normal copulas

In this section we propose the multivariate skew normal copula and its six nested copulas including the multivariate skew normal factor copula.

3.1. Multivariate skew normal distribution

The class of skew normal distributions was introduced by Azzalini [18], and its extensions and applications have been studied by many authors in past three decades [12, 13, 19, 20, 21]. A random variable X is said to have the skew normal distribution with the parameter $\lambda \in \mathbb{R}$, denoted as $X \sim SN(\lambda)$, if it is continuous with the probability density function (PDF),

$$f(x; \lambda) = 2\phi(x)\Phi(\lambda x), \qquad x \in \mathbb{R}, \tag{2}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the PDF and the CDF of a standard normal variable. [12] introduced the transformation method and the conditioning method for constructing the multivariate skew normal distribution. The transformation method provides the multivariate extension of Eq. (2) utilizing the following lemma.

Lemma 3.1. *[12] If Z_0 and Z_1 are independent standard normal random variables with mean 0 and variance 1, denoted by $Z_0, Z_1 \sim N(0,1)$, then $X = \delta|Z_0| + \sqrt{1-\delta^2}Z_1 \sim SN(\lambda)$, where $\delta \in (-1,1)$ and $\lambda = \delta/\sqrt{1-\delta^2}$.*

Let $\boldsymbol{Z} = (Z_0, \ldots, Z_k)^T$ be the multivariate normal distributed random vector with standard normal marginals, independent of $Z \sim N(0,1)$. The joint distribution of Z and $\boldsymbol{Z}$ is given by

$$\begin{pmatrix} Z \\ \boldsymbol{Z} \end{pmatrix} \sim N_{k+2}\left(\boldsymbol{0}, \begin{pmatrix} 1 & \boldsymbol{0}^T \\ \boldsymbol{0} & R \end{pmatrix}\right), \tag{3}$$

where R is the $(k+1)\times(k+1)$ correlation matrix of $\boldsymbol{Z}$, and $N_p(\boldsymbol{\mu}, \Omega)$ is the p-dimensional multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Ω. Define the $(k+1)$ random variables $X_0, X_1, \ldots, X_k$ by

$$X_j = \delta_j|Z| + \sqrt{1-\delta_j^2}Z_j, \tag{4}$$

where $j = 0, \ldots, k$ and $\delta_j \in (-1,1)$. Therefore, by Lemma 3.1, X_j follows $SN(\lambda_j)$ where $\lambda_j = \delta_j/\sqrt{1-\delta_j^2}$.

[12] showed that the joint PDF of $\boldsymbol{X} = (X_0, X_1, \ldots, X_k)^T$ is

$$h(\boldsymbol{x}; R, \boldsymbol{\lambda}) = 2\phi_{k+1}(\boldsymbol{x}; \Sigma)\Phi(\boldsymbol{\alpha}^T\boldsymbol{x}), \tag{5}$$

where

$$\boldsymbol{\alpha}^T = \frac{\boldsymbol{\lambda}^T R^{-1}\Delta^{-1}}{(1+\boldsymbol{\lambda}^T R^{-1}\boldsymbol{\lambda})^{1/2}}, \quad \Sigma = \Delta(R+\boldsymbol{\lambda}\boldsymbol{\lambda}^T)\Delta, \quad \Delta = diag((1-\delta_0^2)^{1/2}, \ldots, (1-\delta_k^2)^{1/2}), \quad \boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_k)^T,$$

and $\phi_{k+1}(\boldsymbol{x}; \Sigma)$ denotes the PDF of $N_{k+1}(\boldsymbol{0}, \Sigma)$. The random vector $\boldsymbol{X}$ with the joint PDF in Eq (5) is said to be the multivariate skew normal distribution with the location parameter $\boldsymbol{0}$, the correlation matrix R, and the skewness parameter $\boldsymbol{\lambda}$, denoted by $\boldsymbol{X} \sim SN_{k+1}(R, \boldsymbol{\lambda})$.

3.2. Multivariate skew normal copula

We now define the multivariate non-exchangeable skew normal copula using the multivariate skew normal distribution defined in Section 3.1.

Definition 3.1. *A $(k+1)$-copula C is said to be the* **skew normal copula** *if*

$$C(u_0, u_1, \ldots, u_k; R, \boldsymbol{\lambda}) = H\left(F_0^{-1}(u_0;\lambda_0), F_1^{-1}(u_1;\lambda_1), \ldots, F_k^{-1}(u_k;\lambda_k); R, \boldsymbol{\lambda}\right), \tag{6}$$

where $F_j(x_j, \lambda_j)$ and $F_j^{-1}(u_j;\lambda_j)$ denote the CDF of $X_j \sim SN(\lambda_j)$ and its inverse, respectively, and H is the CDF of $\boldsymbol{X} \sim SN_{k+1}(R, \boldsymbol{\lambda})$.

Note that the skew normal copula in Eq. (6) is exchangeable if and only if $\lambda_0 = \lambda_1 = \ldots = \lambda_k$ and all off-diagonal elements of R equal.

The skew normal copula in Definition 3.1 has two sets of parameters capturing the non-exchangeable dependence between the variables of interest, the correlation matrix R accounting for association between unobservable or latent variables $\boldsymbol{Z}$ in Eq.(3) and $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_k)$ accounting for the differential skewness of the variables involved.

The corresponding skew normal copula density is given by

$$c(u_0, u_1, \ldots, u_k; R, \boldsymbol{\lambda}) = \frac{h(F_0^{-1}(u_0;\lambda_0), F_1^{-1}(u_1;\lambda_1), \ldots, F_k^{-1}(u_k;\lambda_k); R, \boldsymbol{\lambda})}{f_0(F_0^{-1}(u_0;\lambda_0);\lambda_0)f_1(F_1^{-1}(u_1;\lambda_1);\lambda_1)\ldots f_k(F_k^{-1}(u_k;\lambda_k);\lambda_k)}. \tag{7}$$

where $h(x_0, x_1, \ldots, x_k; R, \boldsymbol{\lambda})$ is the multivariate skew normal density in Eq. (5) and $f_j(x_j;\lambda_j)$ is the marginal density of a skew normal variable, $X_j \sim SN(\lambda_j)$.

3.3. Multivariate skew normal factor copula

Factor copula for modeling the joint distribution of high-dimensional data was studied in [22, 23]. In the factor copula, the dependence between variables is assumed to be captured via one or several common factors. We here derive the 1-factor skew normal copula. As will be shown below, the skew normal factor copula has a much smaller number of parameters than the skew normal copula in Eq. (6) and it is still non-exchangeable.

Let $\mathbf{U} = (U_0, \ldots, U_k)^T$ be a random vector with standard uniform marginals and a $(k+1)$-copula $C(u_0, \ldots, u_k)$ be the joint CDF of $\mathbf{U}$. The random variables $U_0, U_1, \ldots, U_k$ are assumed to be conditionally independent given the latent variable V where $V \sim U(0,1)$. Let $C_{j|V}(u_j|v)$ be the conditional CDF of $U_j|V = v$, $j = 0, 1, \ldots, k$. Then the joint CDF of $\boldsymbol{U}$ is

$$C(u_0, \ldots, u_k) = \int_0^1 \prod_{j=0}^{k} C_{j|V}(u_j|v)dv. \tag{8}$$

The copula given in Eq. (8) is called the **1-factor copula**.

For the skew normal factor copula, we consider the bivariate copula $C(u, v)$ given below,

$$C(u, v) \quad = \quad \int_0^{G^{-1}(v)} \int_{-\infty}^{F^{-1}(u;\lambda)} \frac{2}{\sqrt{1-\delta^2}} \phi(s) \phi\left(\frac{t - \delta s}{\sqrt{1-\delta^2}}\right) dtds.$$

where $G(z)$ and $F(x; \lambda)$ denote the CDF of standard half normal distribution and the CDF of $SN(\lambda)$, respectively, and $\delta = \lambda / \sqrt{1+\lambda^2}$. Set $C_{j,v}(u_j, v) = C(u_j, v)$ for every $j = 0, 1, \ldots, k$ in the 1-factor copula of Eq. (8) and after some calculations, we can define the multivariate skew normal factor copula as follows.

Definition 3.2. *A $(k+1)$-copula C is said to be the* **skew normal factor copula** *if*

$$C(u_0, u_1, \ldots, u_k; \boldsymbol{\lambda}) = \int_0^1 \prod_{j=0}^{k} \Phi\left(\frac{F_j^{-1}(u_j; \lambda_j)) - \delta_j G^{-1}(v)}{\sqrt{1-\delta_j^2}}\right) dv, \tag{9}$$

where $\delta_j = \lambda_j / \sqrt{1+\lambda_j^2}$ and $G(v)$ denote the CDF of the standard half normal distribution.

From Definition 3.2 we can see that the skew normal factor copula in Eq. (9) is the skew normal copula with the identity correlation matrix $R = I_{k+1}$ and the unrestricted $\boldsymbol{\lambda}$. Note that the 1-factor skew normal copula given in Eq. (9) is exchangeable if and only if $\lambda_0 = \lambda_1 = \ldots = \lambda_k$.

Remark 3.1. *Since the skew normal factor copula is parameterized by the skewness parameters λ_0, λ_1, …, λ_k, Σ and $\boldsymbol{\alpha}$ in Eq.(5) can be written as $\Sigma = (\sigma_{ij})$, where $\sigma_{ij} = \delta_i \delta_j$, for $i \neq j$, and $\sigma_{ij} = 1$, for $i = j$, and $\alpha_i = \frac{1}{\sqrt{1+\sum_{i=0}^{k} \lambda_i^2}} \times \frac{\delta_i}{(1-\delta_i^2)}$, for $i = 0, \ldots, k$. Note that the skew normal factor copula has less flexibility than the skew normal copula in Eq. (6). For example, if the skewness parameters λ_i and λ_j, $i \neq j$, are both positive or negative, the corresponding σ_{ij} in Σ under the skew normal factor copula is restricted to be positive.*

3.4. List of multivariate skew normal copulas

In the previous two subsections, we proposed the multivariate skew normal copula and the multivariate skew normal factor copula. Depending on the restrictions of the skewness parameters $\boldsymbol{\lambda} = (\lambda_0, \ldots, \lambda_k)$ and the correlation matrix R, we can further develop several parsimonious skew normal copulas. In the following, we provide a list of multivariate skew normal copulas that can be useful in data analysis.

- Model 0: skew normal copula in Eq. (6) with unrestricted $\boldsymbol{\lambda}$ and unrestricted R;
- Model 1: skew normal copula with the same skewness parameters ($\lambda_0 = \lambda_1 = \ldots = \lambda_k$) and unrestricted R;
- Model 2: skew normal copula with $\boldsymbol{\lambda} = \mathbf{0}$ and unrestricted R, which is the multivariate normal copula;
- Model 3: skew normal copula with unrestricted $\boldsymbol{\lambda}$ and exchangeable $R = [R_{ij}]$ where, for $i, j = 0, \ldots, k$,

$$R_{ij} = \begin{cases} 1 & \text{for} \quad i = j \\ \rho & \text{otherwise.} \end{cases} \tag{10}$$

- Model 4: skew normal copula with the unrestricted $\boldsymbol{\lambda}$ and $R = I_{k+1}$, the skew normal factor copula in Eq. (9);
- Model 5: skew normal copula with the same skewness parameter and the exchangeable R in Eq. (10);
- Model 6: skew normal copula with $\boldsymbol{\lambda} = \mathbf{0}$ and the exchangeable R in Eq. (10).

Note that the Model 0 - Model 4 are non-exchangeable and the Model 5-Model 6 are exchangeable.

3.5. Kendall's tau

Given a bivariate random vector (X, Y), Kendall's tau, which is one of the commonly-used concordance measures, is defined as the probability of concordance minus the probability of discordance [14], i.e.,

$$\tau = P[(X - X')(Y - Y') \geq 0] - P[(X - X')(Y - Y') \leq 0],$$

where (X', Y') is an identical and independent distributed as (X, Y). The Kendall's tau is invariant to monotone transformations of random variables, while the Pearson's correlation is unaffected only by the linear transformation. Thus, the Kendall's tau is useful in comparing association/dependence structure in the data computed under different copulas. Note that in this paper we numerically compute Kendall's tau for the skew normal copulas using the following formula and the numerical integration function 'adaptIntegrate' [24] in R [25],

$$\tau = 4 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(x, y; \rho, \lambda_1, \lambda_2) h(x, y; \rho, \lambda_1, \lambda_2) dx dy - 1.$$

4. Parameter estimation using block coordinate ascent algorithm

This section discusses the estimation method and the numerical algorithm for the proposed class of skew normal copulas.

4.1. Maximum pseudo-likelihood method

In the copula literature there are three commonly used estimation methods, the maximum likelihood (ML), the inference functions for margins (IFM) [15], and the maximum pseudo-likelihood (MPL) [26, 27]. The ML and IFM methods require the specification of parametric models for the marginals. On the other hand, the MPL method uses the rank-based estimators for the marginals and is robust against misspecification of the marginal models [28]. In this paper we employ the MPL method for estimating the proposed class of skew normal copulas.

Given a sample of n observations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ from a continuous random vector $\boldsymbol{X} = (X_0, \ldots, X_k)^T$, we first compute the normalized ranks or the rescaled empirical distributions for the j-th variable X_j, $u_{ij} = \frac{r_{ij}}{n+1}$ where $i = 1, \ldots, n$, $j = 0, 1, \ldots, k$, and r_{ij} is the rank of x_{ij} among n data points from X_j, $x_{1j}, \ldots, x_{nj}$. The pseudo log-likelihood function for the parameters in the skew normal copula is

$$\ell^p(R, \boldsymbol{\lambda}) = \log \prod_{i=1}^{n} c(u_{i0}, \ldots, u_{ik}; R, \boldsymbol{\lambda}) = \sum_{i=1}^{n} \log c(u_{i0}, \ldots, u_{ik}; R, \boldsymbol{\lambda}) \tag{11}$$

where $c(u_0, \ldots, u_k; R, \boldsymbol{\lambda})$ is the skew normal copula copula density in Eq. (7). We can obtain the maximum pseudo-likelihood estimators (MPLE) for the parameters by maximizing Eq. (11) with respect to $(\boldsymbol{\lambda}, R)$. Note that the number of the parameters to be estimated is $k + 1 + k(k + 1)/2$, $(k + 1)$ in $\boldsymbol{\lambda}$ and $k(k + 1)/2$ in R.

4.2. 2-block coordinate ascent algorithm

The estimation of R through Eq. (11) is a difficult numerical problem, as the estimate of R must be positive definite. Furthermore, for high-dimensional data, the simultaneous estimation of λ and R is difficult to implement.

To estimate the skew normal copulas proposed in Section 3, we introduce the 2-block coordinate ascent algorithm [29]. A main idea is to decompose the problem of simultaneously estimating all parameters into two simpler estimation subproblems. First, partition the parameters into two disjoint blocks, one for $\boldsymbol{\lambda}$ and the other for R. Second, maximize $\ell^p(R, \boldsymbol{\lambda})$ over one of the blocks, while keeping the other block fixed at their current values, and replace the values of the active block by the maximizer of $\ell^p(R, \boldsymbol{\lambda})$ with one block fixed, and then proceed by choosing the other block to become active. The details of the proposed algorithm are given as follows:

Algorithm 1 2-block coordinate ascent algorithm for the skew normal copulas

---

Step 1. Initialize two blocks with $\hat{\boldsymbol{\lambda}}_0$ and $\hat{R}_0$.

Step 2. At the t-th iteration, update the estimate $\hat{\boldsymbol{\lambda}}_t$ by maximizing $\ell^p(R, \boldsymbol{\lambda})$ over $\boldsymbol{\lambda}$ with a fixed R $(R = \hat{R}_{t-1})$:

$$\hat{\boldsymbol{\lambda}}_t = \underset{\boldsymbol{\lambda}}{\operatorname{argmax}} \{\ell^p(\hat{R}_{t-1}, \boldsymbol{\lambda})\}.$$

Step 3. At the t-th iteration, update the estimate $\hat{R}_t$ by maximizing $\ell^p(R, \boldsymbol{\lambda})$ over R with a fixed $\boldsymbol{\lambda}$ $(\boldsymbol{\lambda} = \hat{\boldsymbol{\lambda}}_t)$:

$$\hat{R}_t = \underset{R}{\operatorname{argmax}} \{\ell^p(R, \hat{\boldsymbol{\lambda}}_t)\}.$$

Step 4. Repeat Steps 2 and 3 until the algorithm converges.

Most convergence results of the block coordinate algorithm require either the convexity, quasi-convexity property or the unique maximization point assumptions of the target function. However, [29] showed that 2-block coordinate algorithm is globally convergent towards stationary points, even in the absence of convexity or uniqueness assumptions. The following theorem establish the convergence of the Algorithm 1.

Theorem 4.1. *[29] Suppose that the global maximization with respect to each block is well defined and let $\hat{\boldsymbol{\theta}}_t$ denote the estimates of the vectorized R and $\boldsymbol{\lambda}$ in the skew normal copula generated from the t-th iteration of the Algorithm 1. Then, every limit point of the infinite sequence $\{\hat{\boldsymbol{\theta}}_t\}$ is a stationary point.*

The estimation of R in Step 3 of Algorithm 1 involves the optimization of Eq. (11) over R and $\hat{R}_t$ (the estimate of R) should be positive definite at each iteration. To ensure the positive definiteness of $\hat{R}_t$, we utilize the spherical parametrization. Because R is a correlation matrix, it can be uniquely factored as $R = L^T L$ where L is an upper triangular matrix, which gives the Cholesky parameterization of R. Let L_j denote the j-th column of L, l_{ij} denote the i-th elements of L_j, and θ_{ij} denote the spherical coordinates of the first j elements of L_j where $j = 2, \ldots, k + 1$ and $i = 1, \ldots, j$. We then reparameterize the Cholesky decomposed triangular matrix with trigonometric functions:

$$\begin{aligned} &l_{1j} = cos(\theta_{1j}), \quad l_{2j} = sin(\theta_{1j})cos(\theta_{2j}), \quad \cdots, \\ &l_{(j-1)j} = sin(\theta_{1j}) \ldots sin(\theta_{(j-2)j})cos(\theta_{(j-1)j}), \quad l_{jj} = sin(\theta_{1j}) \ldots sin(\theta_{(j-2)j})sin(\theta_{(j-1)j}). \end{aligned} \tag{12}$$

The inverse of the transformation in Eq. (12) will also be needed in the Algorithm 1, which is given below,

$$\theta_{1j} = arccot(l_{1j}/\sqrt{l_{jj}^2 + \ldots + l_{2j}^2}), \quad \theta_{2j} = arccot(l_{2j}/\sqrt{l_{jj}^2 + \ldots + l_{3j}^2}), \ \cdots,$$

$$\theta_{(j-2)j} = arccot(l_{(j-2)j}/\sqrt{l_{jj}^2 + l_{(j-1)j}^2}), \quad \theta_{(j-1)j} = 2arccot\left(\frac{l_{(j-1)j} + \sqrt{l_{jj}^2 + l_{(j-1)j}^2}}{l_{jj}}\right),$$

where *arccot* is the inverse of the cotangent function. Furthermore, the unconstrained estimation is obtained by defining $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_{\frac{k(k+1)}{2}})^T$ where $\phi_{(j-1)(j-2)/2+i-1} = \log\left(\frac{\theta_{ij}}{\pi-\theta_{ij}}\right)$.

For the positive definiteness of the estimate of the exchangeable correlation matrix R in the Models 3, 5, and 6, we employ the Fisher transformation for ρ and ensure the estimate of ρ to be between $-1/k$ and 1. To avoid time-consuming computation for evaluating the marginal skew normal quantile functions $F_j^{-1}(u; \lambda_j)$, we utilize the monotone cubic Hermit interpolation [15, 30] on points (x_s, u_s) where $F_j(x_s; \lambda_j) = u_s$ and $s = 1, \ldots, 150$. Note that we employed the multiple starting value strategy in the Algorithm 1.

5. Data analysis

This section presents an example concerning a study of nutrient intake conducted by the U.S. Department of Agriculture in 1985. We consider three variables measured from 737 women, daily calcium intake, daily iron intake and daily protein intake. Figure 1 shows the scatterplot of the normalized ranks of three variables. As pointed out in [9], asymmetric patterns shown in the bivariate marginals of Figure 1 (e.g., calcium-protein margins with respect to the diagonal line) and the different values of Kendall's rank correlations between three pairs (calcium-iron 0.32, calcium-protein 0.36, iron-protein 0.52) indicate the need of the non-exchangeable copula.

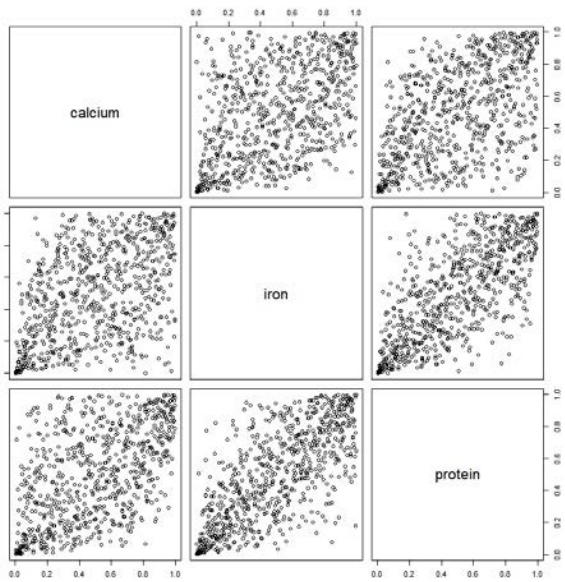

Fig. 1. Normalized ranks of the data from the study of nutrient intake in 737 women conducted by the U.S. Department of Agriculture

[9] applied the non-exchangeable Clayton-Liouville copula to this data and showed a better fit compared to the fully exchangeable Clayton copula. We here illustrate the performance of a list of skew normal copulas with

the same data and compare it to the Clayton-Liouville copula.

We first fit the normalized ranks of the data with the proposed class of trivariate skew normal copulas, Model 0 - Model 6, and the Clayton-Liouville copula. In order to obtain copula models with better fit, we employ the Akaike's information criterion (AIC) [31] and Bayesian information criterion (BIC) [32] based on the normalized ranks [33, 34]: AIC = $2\ell^p(\hat{\boldsymbol{\theta}}) - 2dim(\hat{\boldsymbol{\theta}})$ and BIC = $2\ell^p(\hat{\boldsymbol{\theta}}) - dim(\hat{\boldsymbol{\theta}})\log(n)$ where $\boldsymbol{\theta}$ is the set of the parameters in the copulas, $\hat{\boldsymbol{\theta}}$ is the MPLE for $\boldsymbol{\theta}$, $dim(\boldsymbol{\theta})$ is the dimension of $\boldsymbol{\theta}$ and n is the sample size.

Table 5 compares the class of skew normal copulas with the Clayton-Liouville copula in terms of the maximized pseudo log-likelihood value, AIC, BIC, and the estimated Kendall's tau for each pair of three variables, calcium-iron τ_{ci}, calcium-protein τ_{cp} and iron-protein τ_{ip}. We can see that two skew normal copulas, Model 0 and Model 1, show much better improvement in fit than the Clayton-Liouville copula. The Model 1 and 3 have the same number of parameters as the Clayton-Liouville copula, but the Model 3 has slightly lower AIC and BIC than Clayton-Liouville copula. These results indicate that the copula with the unrestricted R better characterizes the non-exchangeable dependence in this date set than the copula with the unrestricted $\boldsymbol{\lambda}$.

Model	MPLoglik	AIC	BIC	τ_{ci}	τ_{cp}	τ_{ip}
Clayton-Liouville	417.58	827.16	808.75	0.32	0.33	0.48
Model 0	431.60	851.2	823.59	0.31	0.35	0.51
Model 1	429.49	850.98	832.57	0.34	0.38	0.52
Model 2	409.35	812.7	798.89	0.33	0.38	0.51
Model 3	415.78	823.57	805.16	0.30	0.36	0.50
Model 4	314.15	622.3	608.49	0.20	0.32	0.43
Model 5	384.41	764.82	755.61	0.41	0.41	0.41
Model 6	365.95	729.9	725.30	0.40	0.40	0.40

Table 1. Maximum pseudo log-likelihood value (MPLoglik), AIC, BIC, and the estimated Kendall's tau for each pair calcium-iron τ_{ci}, calcium-protein τ_{cp}, and iron-protein τ_{ip} for the Clayton-Liouville copula and the class of skew normal copulas.

6. Conclusion

The non-exchangeable copulas are the fundamental tools to analyze the non-exchangeable/asymmetric dependence structure between multiple variables. We proposed a flexible class of multivariate skew normal copulas to model the high-dimensional data with non-exchangeable dependence patterns and developed the efficient block coordinate ascent algorithm for the parameters estimation. As the future work, we will apply the class of skew normal copulas to analysis of directional dependence using the regression models. The non-exchangeable copula based regression enables us to study the directional dependence stemming from not only marginal behavior of variables, but also the joint behavior of them.

References

[1] U. Cherubini, E. Luciano, W. Vecchiato, Copula methods in finance, Wiley Finance Series, John Wiley & Sons, Ltd., Chichester, 2004.

[2] A. R. de Leon, B. Wu, Copula based regression models for a bivariate mixed discrete and continuous outcome, Statistics in Medicine 30 (2) (2011) 175–185.

[3] A. J. Patton, Modelling asymmetric exchange rate dependence, International Economic Review 47 (2) (2006) 527–556.
[4] L. Zhang, V. P. Singh, Bivariate rainfall frequency distributions using archimedean copulas, Journal of Hydrology 332 (1) (2007) 93–109.
[5] E. W. Frees, E. A. Valdez, Understanding relationships using copulas, North American Actuarial Journal 2 (1) (1998) 1–25.
[6] E. Liebscher, Construction of asymmetric multivariate copulas, Journal of Multivariate Analysis 99 (10) (2008) 2234–2250.
[7] T. Bedford, R. M. Cooke, Vines: a new graphical model for dependent random variables, Annals of Statistics (2002) 1031–1068.
[8] H. Joe, Multivariate models and multivariate dependence concepts, Chapman and Hall/CRC, London, 1997.
[9] A. J. McNeil, J. Nešlehová, From archimedean to liouville copulas, Journal of Multivariate Analysis 101 (8) (2010) 1772–1790.
[10] H. Joe, Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters, Lecture Notes-Monograph Series (1996) 120–141.
[11] J. Stoeber, H. Joe, C. Czado, Simplified pair copula constructionslimitations and extensions, Journal of Multivariate Analysis 119 (2) (2013) 101–118.
[12] A. Azzalini, A. Dalla Valle, The multivariate skew-normal distribution, Biometrika 83 (4) (1996) 715–726.
[13] A. Azzalini, A. Capitanio, Statistical applications of the multivariate skew normal distribution, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 61 (3) (1999) 579–602.
[14] R. Nelsen, An Introduction to Copulas(second edition), Springer, New york., 2006.
[15] H. Joe, Dependence modeling with copulas, CRC Press, 2014.
[16] Z. Wei, T. Wang, P. Nguyen, Multivariate dependence concepts through copulas, International Journal of Approximate Reasoning 65 (2015) 24–33.
[17] A. Sklar, Fonctions de r*é*partition *à n* dimensions et leurs marges, Publications de l'Institut de Statistique de L'Universit*é* de Paris 8 (1959) 229–231.
[18] A. Azzalini, A class of distributions which includes the normal ones, Scandinavian Journal of Statistics 12 (2) (1985) 171–778.
[19] T. Wang, B. Li, A. K. Gupta, Distribution of quadratic forms under skew normal settings, Journal of Multivariate Analysis 100 (3) (2009) 533–545.
[20] R. Ye, T. Wang, A. K. Gupta, Distribution of matrix quadratic forms under skew-normal settings, Journal of Multivariate Analysis 131 (2014) 229–239.
[21] A. Gupta, G. Gonzalez-Farias, J. Dominguez-Molina, A multivariate skew normal distribution, Journal of Multivariate Analysis 89 (1) (2004) 181–190.
[22] P. Krupskii, H. Joe, Factor copula models for multivariate data, Journal of Multivariate Analysis 120 (2013) 85–101.
[23] P. Krupskii, H. Joe, Structured factor copula models: Theory, inference and computation, Journal of Multivariate Analysis 138 (2015) 53–73.
[24] C. code by Steven G. Johnson, R. by Balasubramanian Narasimhan, Cubature: adaptive multivariate integration over hypercubes, r package version 1.1-2 (2013).
[25] R. C. Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2015). URL `https://www.R-project.org/`
[26] D. Oakes, Multivariate survival distributions, Journal of Nonparametric Statistics 3 (3-4) (1994) 343–354.
[27] K. Genest, C.and Ghoudi, L. P. Rivest, A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, Biometrika 82 (3) (1995) 543–552.
[28] G. Kim, M. Silvapulle, P. Silvapulle, Comparison of semiparametric and parametric methods for estimating copulas, Computational Statistics & Data Analysis 51 (2007) 2836–2850.
[29] L. Grippof, M. Sciandrone, Globally convergent block-coordinate techniques for unconstrained optimization, Optimization methods and software 10 (4) (1999) 587–637.
[30] D. Kahaner, C. B. Moler, , S. Nash, Numerical Methods and Software, Prentice Hall, Englewood Cliffs, NJ, 1989.
[31] H. Akaike, A new look at the statistical model identification, IEEE Transactions on Automatic Control 19 (6) (1974) 716–723.
[32] G. Schwarz, Estimating the dimension of a model, The Annals of Statistics 6 (2) (1978) 461–464.
[33] L. A. Jordanger, D. Tjøstheim, Model selection of copulas: Aic versus a cross validation copula information criterion, Statistics & Probability Letters 92 (2014) 249–255.
[34] U. Schepsmeier, Efficient information based goodness-of-fit tests for vine copula models with fixed margins: A comprehensive review, Journal of Multivariate Analysis 138 (2015) 34–52.