Journal of Multivariate Analysis 137 (2015) 61–81

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Maximum entropy copula with given diagonal section[☆]

Cristina Butucea[a], Jean-François Delmas[b], Anne Dutfoy[c], Richard Fischer[a,b,c,*]

[a] Université Paris-Est, Laboratoire d'Analyse et de Mathématiques Appliquées (UMR 8050), UPEM, UPEC, CNRS, F-77454 Marne-la-Vallée, France

[b] Université Paris-Est, CERMICS (ENPC), 6-8 Avenue Blaise Pascal, Cité Descartes, F-77455 Marne-la-Vallée, France

[c] EDF Research & Development, Industrial Risk Management Department, 1 Av. Général de Gaulle, F-92141 Clamart, France

ARTICLE INFO

Article history:
Received 9 January 2014
Available online 17 January 2015

AMS 2010 subject classifications:
62H05
60E05

Keywords:
Copula
Entropy
Diagonal section

ABSTRACT

We consider copulas with a given diagonal section and compute the explicit density of the unique optimal copula which maximizes the entropy. In this sense, this copula is the least informative among the copulas with a given diagonal section. We give an explicit criterion on the diagonal section for the existence of the optimal copula and give a closed formula for its entropy. We also provide examples for some diagonal sections of usual bivariate copulas and illustrate the differences between these copulas and the associated maximum entropy copula with the same diagonal section.

1. Introduction

Dependence of random variables can be described by copula distributions. A copula is the cumulative distribution function of a random vector $U = (U_1, \ldots, U_d)$ with U_i uniformly distributed on $I = [0, 1]$. For an exhaustive overview on copulas, we refer to Nelsen [16]. The diagonal section δ of a d-dimensional copula C, defined on I as $\delta(t) = C(t, \ldots, t)$ is the cumulative distribution function of $\max_{1\leq i\leq d} U_i$. The function δ is non-decreasing, d-Lipschitz, and verifies $\delta(t) \leq t$ for all $t \in I$ with $\delta(0) = 0$ and $\delta(1) = 1$. It was shown that if a function δ satisfies these properties, then there exists a copula with δ as diagonal section (see Bertino [2] or Fredricks and Nelsen [12] for $d = 2$ and Cuculescu and Theodorescu [6] for $d \geq 2$).

Copulas with a given diagonal section have been studied in different papers, as the diagonal sections are considered in various fields of application. Beyond the fact that δ is the cumulative distribution function of the maximum of the marginals, it also characterizes the tail dependence of the copula (see Joe [14, p. 33] and references in Nelsen et al. [18], Durante and Jaworski [8], Jaworski [13]) as well as the generator for Archimedean copulas (Sungur and Yang [26]). For $d = 2$, Bertino in [2] introduces the so-called Bertino copula B_δ given by $B_\delta(u, v) = u \wedge v - \min_{u\wedge v\leq t\leq u\vee v}(t - \delta(t))$ for $u, v \in I$. Fredricks and Nelsen in [12] give the example called diagonal copula defined by $K_\delta(u, v) = \min(u, v, (\delta(u) + \delta(v))/2)$ for $u, v \in I$. In Nelsen et al. [17,18] lower and upper bounds related to the pointwise partial ordering are given for copulas with a given diagonal section. They showed that if C is a symmetric copula with diagonal section δ, then for every $u, v \in I$, we have:

$$B_\delta(u, v) \leq C(u, v) \leq K_\delta(u, v).$$

☆ This work was financially supported by the French "Association Nationale de la Recherche et de la Technologie", contract no 1531/2012.

* Corresponding author at: Université Paris-Est, CERMICS (ENPC), 6-8 Avenue Blaise Pascal, Cité Descartes, F-77455 Marne-la-Vallée, France.
E-mail addresses: cristina.butucea@univ-mlv.fr (C. Butucea), delmas@cermics.enpc.fr (J.-F. Delmas), anne.dutfoy@edf.fr (A. Dutfoy), fischerr@cermics.enpc.fr (R. Fischer).

http://dx.doi.org/10.1016/j.jmva.2015.01.003
0047-259X/

Durante et al. [10] provide another construction of copulas for a certain class of diagonal sections, called MT-copulas named after Mayor and Torrens and defined as $D_\delta(u, v) = \max(0, \delta(x \vee y) - |x - y|)$. Bivariate copulas with given sub-diagonal sections $\delta_{x_0} : [0, 1 - x_0] \to [0, 1 - x_0]$, $\delta_{x_0}(t) = C(x_0 + t, t)$ are constructed from copulas with given diagonal sections in Quesada-Molina et al. [22]. Durante et al. [9,18] introduce the technique of diagonal splicing to create new copulas with a given diagonal section based on other such copulas. According to [8] for $d = 2$ and Jaworski [13] for $d \geq 2$, there exists an absolutely continuous copula with diagonal section δ if and only if the set $\Sigma_\delta = \{t \in I; \delta(t) = t\}$ has zero Lebesgue measure. de Amo et al. [7] is an extension of [8] for given sub-diagonal sections. Further construction of possibly asymmetric absolutely continuous bidimensional copulas with a given diagonal section is provided in Erdely and González [11].

Our aim is to find the most uninformative copula with a given diagonal section δ. We choose here to maximize the relative entropy to the uniform distribution on I^d, among the copulas with given diagonal section. This is equivalent to minimizing the Kullback–Leibler divergence with respect to the independent copula. The Kullback–Leibler divergence is finite only for absolutely continuous copulas. The previously introduced bivariate copulas B_δ, K_δ and D_δ are not absolutely continuous, therefore their Kullback–Leibler divergence is infinite. Possible other entropy criteria, such as Rényi, Tsallis, etc. are considered for example in Pougaza and Mohammad-Djafari [21]. We recall that the entropy of a d-dimensional absolutely continuous random vector $X = (X_1, \dots, X_d)$ can be decomposed as the sum of the entropy of the marginals and the entropy of the corresponding copula (see Zhao and Lin [27]):

$$H(X) = \sum_{i=1}^{d} H(X_i) + H(U),$$

where $H(Z) = -\int f_Z(z) \log f_Z(z) dz$ is the entropy of the random variable Z with density f_Z, and $U = (U_1, \dots, U_d)$ is a random vector with U_i uniformly distributed on I, such that U has the same copula as X; namely U is distributed as $\left(F_1^{-1}(X_1), \dots, F_d^{-1}(X_d)\right)$ with F_i the cumulative distribution function of X_i. Maximizing the entropy of X with given marginals therefore corresponds to maximizing the entropy of its copula. The maximum relative entropy approach for copulas has an extensive literature. Existence results for an optimal solution on convex closed subsets of copulas for the total variation distance can be derived from Csiszár [5]. A general discussion on abstract entropy maximization is given by Borwein et al. [3]. This theory was applied for copulas and a finite number of expectation constraints in Bedford and Wilson [1]. Some applications for various moment-based constraints include rank correlation (Meeuwissen and Bedford [15], Chu [4], Piantadosi et al. [20]) and marginal moments (Pasha and Mansoury [19]).

We shall apply the theory developed in [3] to compute the density of the maximum entropy copula with a given diagonal section. We show that there exists a copula with diagonal section δ and finite entropy if and only if δ satisfies: $\int_I |\log(t - \delta(t))| dt < +\infty$. Notice that this condition is stronger than the condition of Σ_δ having zero Lebesgue measure which is required for the existence of an absolutely continuous copula with diagonal section δ. Under this condition, and in the case of $\Sigma_\delta = \{0, 1\}$, the optimal copula's density c_δ turns out to be of the form, for $x = (x_1, \dots, x_d) \in I^d$:

$$c_\delta(x) = b(\max(x)) \prod_{x_i \neq \max(x)} a(x_i),$$

with the notation $\max(x) = \max_{1 \leq i \leq d} x_i$, see Proposition 2.4. The optimal copula's density in the general case is given in Theorem 2.5. Notice that c_δ is symmetric: it is invariant under the permutation of the variables. This provides a new family of absolutely continuous symmetric copulas with given diagonal section enriching previous work on this subject that we discussed, see [2,8–12,18]. We also calculate the maximum entropy copula for diagonal sections that arise from well-known families of bivariate copulas.

The rest of the paper is organized as follows. Section 2 introduces the definitions and notations used later on, and gives the main theorems of the paper. In Section 3 we study the properties of the feasible solution c_δ of the problem for a special class of diagonal sections with $\Sigma_\delta = \{0, 1\}$. In Section 4, we formulate our problem as an optimization problem with linear constraints in order to apply the theory established in [3]. Then in Section 5 we give the proof for our main theorem showing that c_δ is indeed the optimal solution when $\Sigma_\delta = \{0, 1\}$. In Section 6 we extend our results for the general case when Σ_δ has zero Lebesgue measure. We give in Section 7 several examples with diagonals of popular bivariate copula families such as the Gaussian, Gumbel or Farlie–Gumbel–Morgenstern copulas among others. In the Gaussian case, we illustrate how different the Gaussian copula and the corresponding maximum entropy copula can be, by calculating conditional extreme event probabilities.

2. Main results

Let $d \geq 2$ be fixed. We recall a function C defined on I^d, with $I = [0, 1]$, is a d-dimensional copula if there exists a random vector $U = (U_1, \dots, U_d)$ such that U_i are uniform on I and $C(u) = \mathbb{P}(U \leq u)$ for $u \in I^d$, with the convention that $x \leq y$ for $x = (x_1, \dots .x_d)$ and $y = (y_1, \dots, y_d)$ elements of $\mathbb{R}^d$ if and only if $x_i \leq y_i$ for all $1 \leq i \leq d$. We shall say that C is the copula of U. We refer to [16] for a monograph on copulas. The copula C is said absolutely continuous if the random variable U has a density, which we shall denote by c_C. In this case, we have for all $u \in I^d$:

$$C(u) = \int_{I^d} c_C(v) \mathbf{1}_{\{v \leq u\}} \, dv.$$