SLOVENSKÁ TECHNICKÁ UNIVERZITA

STAVEBNÁ FAKULTA

COPULA: CONSTRUCTIONS AND APPLICATION

HABILITAČNÁ PRÁCA

2017 Ing. Tomáš Bacigál, PhD.

SLOVENSKÁ TECHNICKÁ UNIVERZITA

STAVEBNÁ FAKULTA

COPULA: CONSTRUCTIONS AND APPLICATION

HABILITAČNÁ PRÁCA

Študijný odbor: 9.1.9 aplikovaná matematika
Pracovisko: Katedra matematiky a deskriptívnej geometrie
SvF STU v Bratislave

Bratislava 2017 **Ing. Tomáš Bacigál, PhD.**

Many thanks to my three lovely girls for respecting my home office anytime needed, and to my colleagues for making such a human-friendly and inspirational place for research and teaching.

Author

Abstract

The thesis surveys copula as a tool for modeling stochastic dependence in multivariate probability distribution, namely it provides basic analytical and statistical properties, classification to three popular groups, stochastic interpretation and construction methods linked mainly to the author's publication history. The second part is dedicated to practical use including methods for statistical inference, open source software tools and some most significant applications of copula models published by the author. Copies of author's scientific papers that are most frequently cited by the thesis are included as appendix.

Abstrakt

Práca skúma kopulu ako nástroj modelovania stochastickej závislosti vo viacrozmernom rozdelení pravdepodobnosti, konkrétne popisuje jej základné analytické a štatistické vlastnosti, tri najznámejšie triedy, a prevažne tie metódy konštrukcie, ktoré sa dotýkajú autorovej výskumnej činnosti. Druhá časť je venovaná praktickému používaniu a využitiu kopule, najmä popisu metód pre odhad a testovanie modelov závislosti, predstaveniu voľne dostupných softvérových nástrojov a zhrnutiu niektorých najvýznamnejších aplikácií publikovaných autorom. Prílohu tvoria v práci najčastejšie citované autorské zdroje.

Contents

Preface

Copula is a mathematical concept that allows modeling dependence separately from margins in a multivariate probability distribution. During the last ten years it was the main topic of my post-doctoral research at the Department of Mathematics and Constructive Geometry. However, this habilitation thesis does not focus solely on my achievements, it rather tries (a) to give a self-contained (both analytical and statistical) overview of the relevant copula theory with noticeable bias towards topics of my scientific interest, and (b) to provide references to our publications including hard copy of the most referred ones.

The first chapter summarizes copula properties and their most prominent classes (elliptical, Archimedean and extreme-value copulas) with statistical interpretation. The second one provides a survey on construction methods, especially those connected with Archimedean and extreme-value copulas, such as distortion (leading to, e.g., Archimax copulas) and methods used to construct generators and dependence functions. Our main contribution with co-authors include constructions 1–4, 6–11 and 13. The last chapter review practical use of copulas, i.e., statistical inference such as estimation based on observed data and testing goodness of their fit, then (mainly open source) software tools for handling copulas with focus on my own package developed in R, and finally the chapter sums up those our publications where copulas (of various constructions) were used for modeling stochastic dependence between variables from fields like hydrology and economics.

In Bratislava, on 28th November 2017 author

Chapter 1

Introduction to copulas

In recent years copulas turned out to be a promising tool in multivariate modeling, mostly with applications in actuarial sciences and hydrology. In this section we provide fundamental facts about these functions and describe the most popular parametric classes.

1.1 Definition and properties

From statistical point of view, copula is a function[1] $C\colon [0,1]^d \to [0,1]$, $d \geq 2$, which allows modeling dependence structure in stochastic vector $\boldsymbol{X} = (X_1, ..., X_d)$. The main advantage is that the copula approach can split the problem of constructing multivariate distributions into a part containing the marginal distribution functions $F_{X_1}, \dots, F_{X_d}\colon \mathbb{R} \to [0,1]$ and a part containing the dependence structure. These two parts can be studied and estimated separately and then rejoined to form a multivariate distribution function $F_{\boldsymbol{X}}\colon \mathbb{R}^d \to [0,1]$, formally

$$F_{\boldsymbol{X}}(x_1, \dots, x_d) = C\left(F_{X_1}(x_1), \dots, F_{X_d}(x_d)\right), \quad \text{with } x_1, \dots, x_d \in \mathbb{R}$$

as introduced by [Sklar, 1959]. Thus copula can be seen as joint distribution function restricted to $[0,1]^d$ with uniform margins.

From axiomatic point of view, (bivariate) copula is a function $C\colon [0,1]^2 \to [0,1]$ which satisfies

- the boundary conditions $C(u,0) = C(0,v) = 0$ (C is grounded), $C(u,1) = u$, $C(1,v) = v$ (1 is neutral element of C),
- 2-increasing property $C(u,v) + C(u',v') - C(u,v') - C(u',v) \geq 0$ for all $u, v, u', v' \in [0,1]$, $u \leq u'$, $v \leq v'$.

Special cases include copulas

$$\begin{aligned} M(u,v) &= \min(u,v), \\ \Pi(u,v) &= uv, \\ W(u,v) &= \max(u+v-1, 0), \end{aligned}$$

where (the strongest copula) M and (the weakest copula) W satisfy Fréchet-Höffding bounds inequality

$$W(u,v) \leq C(u,v) \leq M(u,v), \quad \text{for any copula C,}$$

[1] As for notation, for instance $[a,b)$ means a left closed and right open interval with endpoints a and b.

and represent perfect positive and negative dependence, respectively. The product copula Π stands for complete independence.

Copula is symmetric if $C(u,v) = C(v,u)$ for all $(u,v) \in [0,1]^2$ and is asymmetric otherwise.

Generally, in d-variate case,

- the boundary conditions means $C(u_1,\ldots,u_d) = 0$ whenever $0 \in \{u_1,\ldots,u_d\}$ and $C(u_1,\ldots,u_d) = u_i$ whenever $u_j = 1$ for each $j \neq i$, and
- C is d-increasing, i.e., for any $\mathbf{u},\mathbf{v} \in [0,1]^d$, $\mathbf{u} \leq \mathbf{v}$ (i.e., $u_1 \leq v_1,\ldots,u_d \leq v_d$), the C-volume $V_C([\mathbf{u},\mathbf{v}])$ of rectangle $[\mathbf{u},\mathbf{v}]$ is nonnegative, where

$$V_C([\mathbf{u},\mathbf{v}]) = \sum_{\varepsilon\in\{-1,1\}^d} \left(C(\mathbf{z}_\varepsilon) \prod_{i=1}^{d} \varepsilon_i \right) \geq 0,$$

  with $\mathbf{z}_\varepsilon = (z_1^{\varepsilon_1},\ldots,z_d^{\varepsilon_d})$, $z_i^1 = v_i$, $z_i^{-1} = u_i$.

Though there is straightforward extension of functions M, W, Π to d dimensions,

$$\begin{aligned}
M(u_1,\ldots,u_d) &= \min(u_1,\ldots,u_d),\\
\Pi(u_1,\ldots,u_d) &= \prod_{j=1}^{d} u_j,\\
W(u_1,\ldots,u_d) &= \max(\sum_{j=1}^{d} u_j - d + 1, 0),
\end{aligned}$$

the lower bound W is not a copula for $d \geq 3$ (M and Π are copulas for any dimension $d \geq 2$).

Besides the above elementary copulas there exist numerous parametric families, possibly grouped in classes. Following [Durante & Sempi, 2015], a *family of copulas* is any subset of $\mathcal{C}_d$ (all d-dimensional copulas) that can be indexed by a suitable set Θ, denoted $(C_\theta)_{\theta\in\Theta}$. Such a Θ may be a subset of $\mathbb{R}^p$ ($p \geq 1$, then θ is usually referred to as a parameter) or a set of functions with suitable properties. Ideally, a family should be *identifiable* (a copula in the family cannot be parametrised in two different ways) and *monotonically ordered* (the order between parameters is reflected by the same, or opposite, $\leq$ order between copulas), moreover for practical purposes also *interpretable* (probabilistic interpretation suggesting natural situation where the family may be considered), *flexible* with wide range of dependence, and should also be *easy to handle* (expressed in a closed form or at least analytically tractable).

The most frequently used classes of parametric families are eliptical, Archimedean and Extreme-value copulas which we will briefly characterize in the following subsections.

1.2 Elliptical copulas

These are copulas of the elliptically contoured distributions such as normal distribution with Gaussian (or normal) copula

$$C_\Phi(x_1,...,x_d) = \Phi\left[\Phi_1^{-1}(x_1),...,\Phi_d^{-1}(x_d)\right]$$

and Student t-distribution with t-copula

$$C_t(x_1, ..., x_d) = t\left[t_1^{-1}(x_1), ..., t_d^{-1}(x_d)\right],$$

where Φ and t are respective joint distribution functions of multivariate normal and Student t- distributions, similarly Φ_i^{-1} and t_i^{-1}, $i = 1, ..., d$ are univariate quantile functions related to X_i. Both parametric families flexibly describe dependence in multidimensional random vector, however only their density has an explicit form.

To see analogy with Archimedean copulas, mainly treated in this thesis, let us recall statistical origin of elliptical copulas [Mai & Scherer, 2012]. The family of elliptical distributions is a generalization of the class of spherical distributions - each elliptical distribution is obtained as a linear transformation of spherical distributions. Formally, the random vector $\boldsymbol{X} = (X_1, \ldots, X_d)'$ has an elliptical distribution if

$$\mathbf{X} \stackrel{d}{=} \mu + A'\mathbf{Y} \stackrel{d}{=} \mu + A'R\mathbf{S}$$

where $\boldsymbol{\mu} \in \mathbb{R}^d$ is a linear shift, $A \in \mathbb{R}^{k\times d}$ is a linear transformation of the k-dimensional *spherically distributed* random variable $\boldsymbol{Y} = R\boldsymbol{S}$, further R is a non-negative random variable (interpreted as *radius*), $\boldsymbol{S}$ is a random vector (interpreted as *direction*) that is independent of R and uniformly distributed on the unit sphere $S_{L^2,k} = \left\{\boldsymbol{x} \in \mathbb{R}^k | \sum_{i=1}^k x_i^2 = 1\right\}$. An elliptical distribution has elliptically contoured density level surfaces, which explains the name. Besides the transformation via A, the random variable R introduces additional dependence to the components and influences in particular the (joint) tail behavior.

Equivalently, spherical distribution can be characterized by means of certain one-dimensional function φ (*characteristic generator*), such that the characteristic function $\phi_{\boldsymbol{Y}}$ of (spherically distributed) $\boldsymbol{Y}$ admits the representation $\phi_{\boldsymbol{Y}}(\boldsymbol{t}) = \mathbb{E}\left[e^{i\boldsymbol{t}'\boldsymbol{Y}}\right] = \varphi(\|\boldsymbol{t}\|_2)$, $\boldsymbol{t} \in \mathbb{R}^k$.

Then, there can be seen the two well-known parametric families as special cases, namely

- $\boldsymbol{X} \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$ (multivariate normal distribution) if $\boldsymbol{Y} \stackrel{d}{=} \boldsymbol{Z} \sim \mathcal{N}_k(\mathbf{0}, I)$ with $\Sigma = A'A \in \mathbb{R}^{d\times d}$ being the (positive-semidefinite) covariance matrix. Note that $R^2 \stackrel{d}{=} Z_1^2 + \ldots + Z_k^2 \sim \chi^2(k)$ and $\varphi(x) = \exp(-x/2)$.

- $\boldsymbol{X} \sim t_d(\boldsymbol{\mu}, \Sigma, \nu)$ (multivariate Student's t-distribution) if $\boldsymbol{Y} \stackrel{d}{=} \sqrt{W}\boldsymbol{Z}$ with $1/W \sim \Gamma(\mu/2, \mu/2)$ (Gamma distribution) and $A = \Sigma^{1/2}$, where $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$ is positive definite, W and $\boldsymbol{Z}$ are independent. Note that $R^2/d \sim \mathrm{F}(d, \nu)$ (F-distribution) and $\varphi(x) = \int_0^\infty e^{-w\,x/2} dF_W(w)$, where F_Z denotes distribution function of W.

Finally, elliptical copulas are obtained by standardizing the univariate marginals of elliptical distributions.

1.3 Archimedean copulas

Archimedean copulas are popular for their easy construction and nice analytical properties, they are characterized by the associativity[2] and the diagonal inequality $C(u, \ldots, u) < u$, $\forall u \in [0, 1]$, represented by the formula

$$C(u_1, \ldots, u_d) = f^{(-1)}\Big(f(u_1) + \ldots + f(u_d)\Big) \tag{1.1}$$

[2] i.e., e.g., $C(C(u_1, u_2), u_3) = C(u_1, C(u_2, u_3))$ for all $u_1, u_2, u_3 \in [0, 1]$

where the so-called additive generator $f\colon [0,1] \to [0,\infty]$ is a continuous strictly decreasing mapping such that $f(1) = 0$, and its pseudo-inverse $f^{(-1)}\colon [0,\infty] \to [0,1]$ given by $f^{(-1)}(u) = f^{-1}\big(\min(f(0),u)\big)$ is d-monotone[3] [McNeil & Nešlehová, 2009]. We denote by $\mathcal{F}_d$ the class of all additive generators that generate d-dimensional copulas (in fact k-dimensional with any $k \le d$) and $f \in \mathcal{F}_\infty$ will be called universal generators, where $\mathcal{F}_\infty = \bigcap_{d=2}^{\infty} \mathcal{F}_d$. Obviously $\mathcal{F}_2 \supset \mathcal{F}_3 \supset \ldots \supset \mathcal{F}_\infty$. Note that for any multiplicative constant $c > 0$, functions f and $c\,f$ generate the same Archimedean copula C.

When convenient, we will denote the inverse function $f^{(-1)}$ by g and its pseudo-inverse[4] will be defined as $g^{(-1)}(u) = \inf\{t \in [0,\infty) | g(t) \ge u\}$. It is not difficult to check that $g = f^{(-1)}$ and $f = g^{(-1)}$, i.e., the information contained in the additive generator f of an Archimedean copula C is the same as the information contained in its pseudo-inverse g, and that $C(u,v) = f^{(-1)}(f(u)+f(v)) = g(g^{(-1)}(u)+g^{(-1)}(v))$. In fact, in the literature both forms are used, the first one (based on f) being preferred in probabilistic areas, while the second one (based on g) is more frequently used in the statistical literature. Analogously, we denote by $\mathcal{G}_d$ the class of all additive generators inverses g that generate d-dimensional copulas and $\mathcal{G}_\infty$ will refer to inverses of universal generators.

Some well known examples include

- $f(u) = 1 - u$, $f \in \mathcal{F}_2 \setminus \mathcal{F}_3$, that generates W,

- $f(u) = -\log u$, $f \in \mathcal{F}_\infty$, the generator of Π,

- $f(u) = (-\log u)^p$ with parameter $p \ge 1$, $f \in \mathcal{F}_\infty$, giving the so-called *Gumbel* family

- $f(u) = \frac{1}{p}(u^{-p} - 1)$ with $p > 0$, $f \in \mathcal{F}_\infty$ (the strict case), and with $\frac{-1}{d-1} \le p < 0$, $f \in \mathcal{F}_d$, representing the so-called *Clayton* family, including the weakest Archimedean copula (for $p = \frac{-1}{d-1}$, moreover W for $d = 2$) and having Π as limiting case ($p \to 0$),

- $f(u) = -\log\left(\frac{e^{-pu}-1}{e^{-p}-1}\right)$ with $p \ne 0$, $f \in \mathcal{F}_\infty$, yielding the so-called *Frank* family that goes to Π as $p \to 0$ and to M as $p \to \infty$ (the same with Gumbel and Clayton family).

The name "Archimedean" comes from the Archimedean property that these copulas share with 1-Lipschitz triangular norms, for which [Ling, 1965] showed they can be represented as in (1.1). See convenient explanation, e.g., in [Mai & Scherer, 2014].

Although Archimedean copulas arose as purely analytical construction, they possess also interesting statistical interpretations, particularly in the context of multiplicative frailty and resource sharing models, briefly depicted in the following subsections. See also [Genest et al., 2011].

[3] A real function g is called d-monotone, $d \ge 2$, if it is differentiable up to the order $d-2$ and the derivatives satisfy $(-1)^k g^{(k)}(u) \ge 0$, $k = 0, 1, \ldots, d-2$, for any u in its domain and further if $(-1)^{d-2} g^{(d-2)}$ is nonincreasing and convex.

[4] Pseudo-inverses are deeply discussed in [Klement et al., 1999]

1.3.1 Frailty models

Let $T_1, ..., T_d$ be lifetimes with survival[5] functions $\bar{F}_{T_i}$ and Z be *positive random variable* called *common frailty* with distribution function F_Z, such that $\mathbb{P}(T_i > t|Z = z) = \bar{G}_i(t)^z$ (i.e., Z lowers T_i), $z > 0$, with some associated survival function $\bar{G}_i$. Hence

$$\bar{F}_{T_i}(t) = \mathbb{P}(T_i > t) = \int_0^\infty \bar{G}_i(t)^z dF_Z(z) = \mathcal{L}_Z\left(-\ln \bar{G}_i(t)\right)$$

$$\mathbb{P}(T_1 > t_1, ...T_d > t_d|Z = z) = \prod_{i=1}^d \bar{G}_i(t_i)^z, \quad \text{(condit. independent)}$$

$$\begin{aligned}\bar{F}_{\boldsymbol{T}}(t_1, ..., t_d) = \mathbb{P}(T_1 > t_1, ...T_d > t_d) &= \mathcal{L}_Z\left(-\sum_{i=1}^d \ln \bar{G}_i(t)\right) \\ &= g\left\{\sum_{i=1}^d f \circ \bar{F}_{T_i}(t_i)\right\}\end{aligned}$$

where $\mathcal{L}_Z$ is the Laplace-Stieltjes transform[6] of Z. Thus survival copula associated with $\boldsymbol{T} = (T_1, ..., T_d)$ is obviously an Archimedean copula with generator $f \in \mathcal{F}_\infty$, $g = f^{(-1)} = \mathcal{L}_Z$. Note that the random vector $\boldsymbol{T}$ admits only positive dependence among its components with this models.

In the special case when $G_i(t) = e^{-t}$ (exponential distribution with parameter $\lambda = 1$), $\forall i \in \{1, \dots, d\}$, is a survival function of a random variable Y_i, then lifetimes can be expressed as $T_i \stackrel{d}{=} Y_i/Z$ with $\bar{F}_{T_i} = g$ and $\bar{F}_{\boldsymbol{T}}(t_1, \dots, t_d) = E[e^{-(t_1+\dots+t_d)Z}] = g(t_1 + \dots + t_d)$.

In dual model, $\mathbb{P}(T_i \le t|Z = z) = G_i(t)^z$ (i.e., Z elevates T_i), $z > 0$, with some associated distributon function G_i, the random variable Z is called *resilience*. For more details and references we recommend [Joe (2015)].

1.3.2 Resource sharing models

Let $R > 0$ be a *common resource* (positive random variable with distribution function F_R) to be distributed among $d \ge 2$ agents, and $S_1, ..., S_d$ be their *shares* uniformly distributed on the standard simplex $\left\{(s_1, ...s_d) \in [0, 1]^d | s_1 + \dots + s_d = 1\right\}$. Then the amounts of the resource,

$$(X_1, ..., X_d) = R \times (S_1, ..., S_d), \tag{1.2}$$

follows a simplex distribution with radial part R and survival functions of $\boldsymbol{X} = (X_1, ..., X_d)$ are

$$\begin{aligned}\bar{F}_{\boldsymbol{X}}(x_1, ..., x_d) &= \int_0^\infty Pr\left[S_1 > \frac{x_1}{r}, ..., S_d > \frac{x_d}{r}\right] dF_R(r) = \\ &= \int_0^\infty \left(1 - \frac{x_1 + ... + x_d}{r}\right)^{d-1} dF_R(r) = \\ &= g(x_1 + ... + x_d) \\ \bar{F}_{X_i}(x_i) &= g(x_i) \qquad i = 1, \dots, d\end{aligned}$$

[5] Survival function is dual to distribution function, $\bar{F}_X(x) = 1 - F_X(x) = \mathbb{P}(X > x)$, similarly joint survival function $\bar{F}(x_1, \dots, x_d) = \hat{C}(\bar{F}_{X_1}(x_1), \dots, \bar{F}_{X_d}(x_d))$ with survival copula $\hat{C}$.

[6] $\mathcal{L}_Z(s) = E[\exp(-sZ)] = \int_0^\infty \exp(-sz)dF_Z(z)$

Thus, as shown by [McNeil & Nešlehová, 2009], a copula C is Archimedean copula iff it is survival copula of a vector $\boldsymbol{X}$ with representation (1.2) whose radial part fulfills $F_R(0) = 0$.

The transformation of F_R to g is called the Williamson d-transform denoted $\mathcal{W}_d$ and defined

$$g(x) = \mathcal{W}_d(F_R)(x) = \int_x^\infty \left(1 - \frac{x}{r}\right)^{d-1} dF_R(r) = E\left[1 - \frac{x}{R}\right]_+^{d-1},$$

$$F_R(r) = \mathcal{W}_d^{-1}(g)(r) = 1 - \sum_{i=0}^{d-2} \frac{(-1)^i}{i!} r^i g^{(i)}(r) - \frac{(-1)^{d-1}}{(d-1)!} r^{d-1} g_+^{(d-1)}(r).$$

If $g = \mathcal{W}_d(F_R) = \mathcal{L}_Z$ then $R \overset{d}{=} E_d/Z$, where $E_d \sim Erlang(d) = \Gamma(d,1)$ and the frailty variable Z are independent. Scale of R has no impact on the generated copula.

The above statistical interpretations of Archimedean copulas show their similarity with spherical distributions. For further interesting details about this analogy, see [Joe, 2015, 152].

1.4 Extreme-value

EV copulas model dependence structure between componentwise maxima $X_n^{\max} = (X_{n,1}^{\max}, \ldots, X_{n,d}^{\max})$, $X_{n,j}^{\max} = \bigvee_{i=1}^n X_{i,j}$, of d-variate stationary stochastic process $\{X_{i1}, \ldots, X_{id})\}_{i=1}^n$ with common distribution function F, margins $F_1, \ldots, F_d$ and copula C_F.

Let C_n be a copula of $X_n^{\max}$, then it holds that $C_n(u_1, \ldots, u_d) = C_F\left(u_1^{1/n}, \ldots, u_d^{1/n}\right)^n$, and an extreme-value copula C is just its limit case

$$C(\boldsymbol{u}) = \lim_{n\to\infty} C_n(\boldsymbol{u})$$

(if it exists). It is said, that C_F lie in the domain of attraction of C, which in turn is called max-stable (lying in its own domain of attraction).

The class of Extreme-value copulas coincide with the set of copulas of extreme-value distributions, therefore they may be represented as

$$C(u_1, \ldots, u_d) = \exp(-\ell(-\log u_1, \ldots, -\log u_d))$$

where the tail dependence function $\ell\colon [0,\infty)^d \to [0,\infty)$ is defined by

$$\ell(x_1, \ldots, x_d) = \int_{\Delta_{d-1}} \bigvee_{j=1}^d (w_j x_j) dH(w_1, \ldots, w_d),$$

the unit simplex in $\mathbb{R}^d$ is given as usually $\Delta_{d-1} = \{(w_1, \ldots, w_d) \in [0,\infty)^d | \sum_j w_j = 1\}$ and the spectral measure H is due to uniformity of copula margins constrained to $\int_{\Delta_{d-1}} w_j dH(w_1, \ldots, w_d) = 1$. The tail dependence function ℓ si convex, positively homogeneous of order one[7] and satisfies $\max(x_1, \ldots, x_d) \le \ell(x_1, \ldots, x_d) \le x_1 + \ldots + x_d$ for all $(x_1, \ldots, x_d) \in [0,\infty]^d$.

By homogeneity, it is characterized by the *Pickands dependence function* $A\colon \Delta_{d-1} \to [1/d, 1]$ (restriction to unit simplex)

$$\ell(x_1, \ldots, x_d) = (x_1 + \ldots + x_d) A(w_1, \ldots, w_d), \quad \text{where } w_j = \frac{x_j}{x_1 + \ldots + x_d}$$

[7] i.e. $\ell(cx_1, \ldots, cx_d) = c\ell(x_1, \ldots, x_d)$ for $c > 0$

(with convention $\frac{0}{0} = \frac{1}{d}$), then extreme-value copula C can be expressed in terms of A via

$$C(u_1, \ldots, u_d) = \exp\left[\left(\sum_{i=1}^{d} \log u_i\right) A\left(\frac{\log u_1}{\sum_{i=1}^{d} \log u_i}, \ldots, \frac{\log u_d}{\sum_{i=1}^{d} \log u_i}\right)\right].$$

In bivariate case, A is most often defined as a function $[0,1] \to [1/2,1]$ of one argument, $A(w) = \ell(w, 1-w)$, characterized by the convexity and boundary conditions $\max(w, 1-w) \leq A(w) \leq 1$.

Some well known cases include

- $\ell(\boldsymbol{x}) = \sum_{i=1}^{d} x_i$ which gives Π (the smallest EV copula),
- $\ell = \max$, the tail dependence function of M,
- $\ell(\boldsymbol{x}) = \left(\sum_{i=1}^{d} x_i^{\theta}\right)^{\frac{1}{\theta}}$ with parameter $\theta \geq 1$ giving the so-called *Gumbel* copulas family[8],
- $\ell(\boldsymbol{x}) = \sum_{\emptyset \neq I \subset \{1,\ldots,d\}} (-1)^{card(I)+1} \left(\sum_{j \in I} x_j^{-\theta}\right)$ that for $\theta > 0$ constitutes *Galambos* family of d-dimensional EV copulas, however the more familiar formula belongs to Pickands dependence function in bivariate case, $A(t) = 1 - (t^{-\theta} + (1-t)^{-\theta})^{-1/\theta}$.
- $A(t) = t\Phi\left(\frac{1}{\theta} + \frac{\theta}{2}\log\frac{t}{1-t}\right) + (1-t)\Phi\left(\frac{1}{\theta} - \frac{\theta}{2}\log\frac{t}{1-t}\right)$, $\theta > 0$, where Φ is CDF of standard normal distribution, and it is known as *Hüsler-Reiss* family.

All the three above parametric families are parametrized to range from Π to M (explicitly or as limiting case) as θ increases.

For more details about Extreme-value copulas we recommend [Gudendorf & Segers, 2010].

[8] Gumbel - also known as Gumbel-Hougaard - copulas are the only family belonging to Archimedean and Extreme-value class at the same time.

Chapter 2

Construction methods

There is a great variety of constructions developed in the history of copulas and it can be quite difficult to summarize it exhaustively, one can consult, e.g., a rather encyclopedic monograph of [Joe, 2015] reviewing the recent state-of-the-art of dependence modeling with copulas classifying parametric families with respect to their construction origins. Here we recall a brief summary given by [Durante & Sempi, 2010], who distinguish three essential kinds of construction, yet, "at an abstract level, all the methods start with some known copulas and/or some auxiliary functions". The first group (a) is made of *copulas with given lower dimensional margins* with the most prominent representatives being pair-copula construction (Vine copulas) and nested construction (hierarchical Archimedean copulas). Then there are (b) *copula-to-copula transformations* like ordinal sums, distortions, pointwise composition and shuffles of copulas, that transform d-dimensional copulas into other d-dimensional copulas having possibly additional features. Finally, (c) *geometric constructions* start with some information about the copula structure, for example support, diagonals, horizontal and vertical sections. Details and references can be find in [Durante & Sempi, 2010].

In the context of our presented work, specifically the distortions are interesting and we dedicate it the next subsection with emphasize on Archimax copulas which at the same time form a superclass combining Archimedean and Extreme-value copulas.

2.1 Distortions of copula

Consider a copula C and an increasing bijection $h\colon [0,1] \to [0,1]$, then distortion of C is defined as

$$C_h(\boldsymbol{u}) = h\left[C(h^{-1}(u_1), \dots, h^{-1}(u_d))\right], \tag{2.1}$$

see [Morillas, 2005] for discussion and examples.

2.1.1 Mixture of max-id

Now, following [Joe2015, p.98], consider a *max-infinitely divisible* (max-id)[1] copula K and a function $g \in \mathcal{G}_\infty$ which is the Laplace transform of some resilience variable

[1] A multivariate cdf F is max-infinitely divisible if F^z is a cdf for all $z > 0$. Analogously, $\bar{F}$ is min-infinitely divisible if $\bar{F}^z$ is a multivariate survival function for all $z > 0$. Max-id condition is satisfied by all EV copulas and both max-id and min-id by the bivariate Archimedean copulas based on $\mathcal{F}_\infty$ generators.

Z (see section 1.3.1). Then

$$C_{h,K}(\boldsymbol{u}) = g\left(-\log K(e^{-g^{-1}(u_1)}, \ldots, e^{-g^{-1}(u_d)})\right) \tag{2.2}$$

is the copula of multivariate distribution $F_{g,K} = \int_0^\infty K^z dF_Z(z) = g(-logK)$, a mixture of max-id distributions. By relation $h = g \circ (-\log)$ it is obviously the distortion kind of construction and it generalizes Archimedean copula families with more flexible positive dependence property than is the exchangeability. For bivariate copula families, there are lots of choices for g and K, the well known are 2-parameter BB1 to BB7 families.

In particular, when

- K is Archimedean copula with generator φ then also (2.2) yields an Archimedean copula with generator $\varphi_g(t) = \varphi\left(e^{g^{-1}(t)}\right)$,
- K is a bivariate EV copula with dependence function ℓ, then

$$C_{g,K}(u_1, u_2) = g\left[\ell(g^{-1}(u_1), g^{-1}(u_2))\right] \tag{2.3}$$

matches the class of so-called (bivariate) *Archimax copulas* introduced in [Capéraà et al., 2000].

As Archimax copulas make an important part of our scientific research, more about them needs to be told which is satisfied in the next subsection.

2.1.2 Archimax

In terms of Pickands dependence function A and generator f the formula (2.3) rewrites to (in the literature a more familiar representation)

$$C_{f,A}(u_1, u_2) = f^{(-1)}\left[(f(u_1) + f(u_2))A\left(\frac{f(u_1)}{f(u_1) + f(u_2)}\right)\right] \tag{2.4}$$

which reduces to Archimedean copula for $A(t) = 1$, to Extreme-value copula for $f(t) = -\log(t)$ and to the comonotonicity copula M with $A(t) = \max(t, 1-t)$. Here the convention $\frac{0}{0} = \frac{\infty}{\infty} = 1$ is considered.

In the general, d-dimensional case, [Charpentier et al., 2014] provide sufficient conditions for the function $C_{g,\ell}$,

$$C_{g,\ell}(\boldsymbol{u}) = g \circ \ell(g^{(-1)}(u_1), \ldots, g^{(-1)}(u_d)), \tag{2.5}$$

to be a copula, namely that g generates d-variate Archimedean copula and ℓ is *d-variate stable tail dependence function* defined by the following properties:

a) ℓ is positively homogeneous of degree one,

b) $\ell(\boldsymbol{e}_1) = \ldots = \ell(\boldsymbol{e}_d) = 1$, where $\boldsymbol{e}_i$ is i-th unit (basis) vector in $\mathbb{R}^d$,

c) ℓ is fully d-max decreasing[2].

[2] i.e., $\forall\ x_1, \ldots, x_d, h_1, \ldots, h_d \in [0, \infty]$ and for any $J \subseteq \{1, \ldots, d\}$ of arbitrary size $|J| = k$

$$\sum_{\iota_1, \ldots, \iota_k \in \{0,1\}} (-1)^{\iota_1 + \ldots + \iota_k} \ell(x_1 + \iota_1 h_1 \mathbf{1}_{1 \in J}, \ldots, x_d + \iota_d h_d \mathbf{1}_{1 \in J}) \leq 0$$

These conditions imply (but are stronger than) then the well-known characteristics of stable tail dependence functions (convexity in each argument and boundary conditions) which are necessary but insufficient.

[Charpentier et al., 2014] also provides stochastic representation for Archimax copula in the storyline given for Archimedean copulas, specifically $C_{g,\ell}$ is survival copula of

- $\boldsymbol{T} \stackrel{d}{=} \boldsymbol{Y}/Z$ where $Y_1, \ldots, Y_d$ are unit exponential random variables (independent of Z) whose survival copula is Extreme-value with stable tail dependence function ℓ,
$$\bar{F}_Y(y_1, \ldots, y_d) = e^{-\ell(y_1,\ldots,y_d)},$$
and g is the Laplace transform of Z, $g = \mathcal{L}_Z$.

- $\boldsymbol{X} \stackrel{d}{=} R \times \boldsymbol{S}$ where $S_1, \ldots, S_d \sim \mathcal{B}(1, d-1)$ (Beta distribution) are independent of R and have joint survival function
$$\bar{F}_S(s_1, \ldots, s_d) = [\max(0, 1 - \ell(s_1, \ldots, s_d))]^{d-1},$$
with support on $\{(s_1, \ldots, s_d) \in [0,1]^d | \ell(s_1, \ldots, s_d) \leq 1\}$, and $g = \mathcal{W}_d(F_R)$.

Whenever $R \sim \mathit{Erlang}(d)$ then $g(x) = e^{-x}$, consequently $C_{g,\ell}$ is EV copula and $\boldsymbol{X} \stackrel{d}{=} \boldsymbol{Y}$. Whenever $g = \mathcal{L}_Z = \mathcal{W}_d(F_{R=E_d/Z})$ where $E_d \sim \mathit{Erlang}(d)$ then $\boldsymbol{X} \stackrel{d}{=} \boldsymbol{Y}/Z$.

2.2 Other constructions based on Archimedean generators

There is one class of bivariate copulas (we cannot assign to any above mentioned groups) that was proposed by [Durante & Jaworski, 2012] to fulfill a certain statistical property and is connected with Archimedean copulas. We summarize it in the next subsection, followed by a generalization brewed by [Mesiar & Pekárová, 2010].

2.2.1 Univariate conditioning stable

Let C be the copula of random vector (X, Y), then by (left) α-conditional (or threshold) copula $C_{[\alpha]}$ we denote the copula associated with conditional distribution function $F(x, y) = \mathbb{P}(X \leq x, Y \leq y | X \leq F_X^{-1}(\alpha))$ which is given by formula

$$C_{[\alpha]}(u, v) = \mathbb{P}(U \leq u, V \leq v | U < \alpha) = \frac{1}{\alpha} C\left(\alpha u, \eta^{(-1)}(v)\right)$$

where $\eta(v) = \frac{1}{\alpha} C(\alpha, v)$ and its right-inverse $\eta^{(-1)}(v) = \sup\{t \in [0,1] |$ $C(\alpha, t) < \alpha v\}$, see [Mesiar et al., 2008]. Further, by $\mathcal{C}_{UCS}$ we denote the set of all *univarate (left) conditioning stable copulas, UCS* (or copulas invariant under left uivariate truncation), i.e, copulas C that are equal to their associated $C_{[\alpha]}$ for all $\alpha \in (0, 1)$, see [Jágr et al., 2010] for some properties and examples[3]. Perhaps it is

[3] For completeness, there exist also bivariate copulas invariant under bivariate conditioning and they coincide with the Clayton family of Archimedean copulas.

worth to note, that also any g-ordinal sum[4] of collection of copulas from $\mathcal{C}_{UCS}$ is stable under univariate conditioning.

To picture the use of such copulas, take for instance the example given by [Durante & Jaworski, 2012]: when X a Y represent returns of two financial markets linked by a copula C that is invariant under univariate conditioning, we can say that the dependence structure does not change when one market is taking on large losses.

A large class of copulas that fulfill such a invariance property are due to [Durante & Jaworski, 2012] constructed by means of a one-dimensional function and it is defined by

$$C_{(f)}(u,v) = xf^{(-1)}\left(\frac{f(y)}{x}\right),$$
$$C_{(\bar{f})}(u,v) = x\left(1 - f^{(-1)}\left(\frac{\bar{f}(y)}{x}\right)\right), \quad x \in (0,1],$$

where - interestingly - $f \in \mathcal{F}_2$, that is the generators of bivariate Archimedean copulas, with flippings $\bar{f}(x) = f(1-x)$ and $C_{(\bar{f})}(x,y) = x - C_{(f)}(x,1-y)$. Note that $C_{(c\,f)} = C_{(f)}$, $c > 0$, moreover $C_{(f)} \leq \Pi$ and $C_{(\bar{f})} \geq \Pi$ (and $\Pi \notin \{C_{(f)}, C_{(\bar{f})}\}_{f\in\mathcal{F}_2}$). For example $f(x) = 1-x$, the generator of W, gives $C_{(f)} = W$ and $C_{(\bar{f})} = M$, while $f_1(x) = ((1-x)^{-p}-1)^{-1/p}$ with $p > 0$ and $f_2(x) = (1-x^{-p})^{-1/p}$ with $p \in [-1,0)$ give $C_{(\bar{f}_1)}$ and $C_{(f_2)}$ which form the Clayton family of Archimedean copulas. The only exchangeable copulas in the class $\mathcal{C}_{UCS}$ are M and Clayton family copulas. Further, [Durante et al., 2011] elaborate more on the similarities with Archimedean copulas: Let (U,V) be a pair of continuous random variables with Archimedean copula (generated by $f \in \mathcal{F}_2$) as distribution function, and let $U' = \frac{f^{(-1)}(V)}{f^{(-1)}(U)+f^{(-1)}(V)}$; then $C_{(f)}$ is the distribution function of the random pair $(U'V)$. This finding is useful e.g., for generating random samples from $C_{(f)}$ as well as for statistical inference.

Subsequently, [Mesiar & Pekárová, 2010] proposed a generalization that is inspired by viewing Archimax copulas as a distortion of Archimedean copulas. We glance through it briefly.

2.2.2 Distorted UCS

Let again $f \in \mathcal{F}_2$ and let $d\colon [0,1] \to [0,1]$ be a function such that

$$C_{(f,d)}(u,v) = xf^{(-1)}\left(\frac{f(y)}{d(x)}\right), \text{ and its flipping}$$
$$C_{(\bar{f},d)}(u,v) = x\left(\bar{f}^{(-1)}\left(\frac{\bar{f}(y)}{d(x)}\right)\right), \quad x \in (0,1].$$

Then $C_{(f,d)}$ and $C_{(\bar{f},d)}$ are copulas iff there is a function $\tilde{d}\colon [0,1] \to [0,1]$ dual to d in the sense $\tilde{d}(x)d(x) = x$ and both $\tilde{d}, d$ are nondecreasing on $(0,1]$. The strongest distortion function $d(x) = 1$ leads to $C_{(f,d)} = \Pi$ while the weakest, $d(x) = x$, gives

[4] A g-ordinal sum copula is defined for any disjoint system $\{(a_i,b_i)\}_{i\in\mathcal{I}}$ of open subintervals of $(0,1)$ and any system $\{C_i\}_{i\in\mathcal{I}}$ of copulas by

$$C(u,v) = \begin{cases} a_i v + (b_i - a_i)C_i\left(\frac{u-a_i}{b_i-a_i}, v\right) & \text{if } u \in (a_i,b_i), \\ uv & \text{otherwise} \end{cases}$$

$C_{(f,d)} = C_{(f)}$. Authors denotes these distorted copulas briefly as DUCS copulas and generally they are not univariate conditioning stable.

It can be interesting to see DUCS copula rewritten, $C_{(f,d)}(u,v) = \tilde{d}(u)C_{(f)}(d(u),v) = \Pi(\tilde{d}(u),1)C_{(f)}(d(u),v)$ to notice that this distortion can be applied to a general copula (not just $C_{(f)}$) so that it would become a particular case of construction method proposed by [Liebscher, 2008] to bring asymmetry[5] , i.e., $C(u,v) = A(\alpha_1(u),\alpha_2(v)) \cdot B(\beta_1(u),\beta_2(v))$ (which is a copula whenever A, B are copulas and $\alpha_1,\alpha_2,\beta_1,\beta_2\colon [0,1] \to [0,1]$ are nondecreasing such that $\alpha_1(x)\beta_1(x) = \alpha_2(x)\beta_2(x) = x$, $\forall x \in [0,1]$), with $\alpha_1 = \tilde{d}$, $\beta_1 = d$, $\alpha_2 = 1$ and β_2 is identity.

2.3 Generators of Archimedean copulas

In this subsection we summarize well-known as well as new construction methods of generators of Archimedean copulas, categorized according to whether they arise from a given generator, some given function, by gluing or by agreggating several generators. For more details (proofs, references and examples) see mainly our papers [Bacigál et al., 2015], [Bacigál et al., 2015] and [Bacigál et al., 2010].

2.3.1 From a given generator

In the following constructions we will start with some given generator f of an Archimedean copula.

Construction 1. Let $\varphi\colon [0,1] \to [0,1]$ be an automorphism of $[0,1]$, then the composition

$$f_{c1} = f \circ \varphi$$

is also a generator, in particular a) $f_{c1} \in \mathcal{F}_d$ for all $f \in \mathcal{F}_d$ and some fixed $d \in \{2,3,\dots\}$ iff φ^{-1} has $(d-2)$ derivatives on $(0,1)$, $(\varphi^{-1})^{(k)}(x) \geq 0$ for all $k \in 1,\dots,d-2$ and $x \in (0,1)$, and $(\varphi^{-1})^{(d-2)}$ is a convex function, see Proposition 3 in [Bacigál et al., 2015]. Similarly b) $f_{c1} \in \mathcal{F}_\infty$ for all $f \in \mathcal{F}_\infty$ iff φ^{-1} is absolutely monotonic on $(0,1)$.

Construction 2. Let $\eta\colon [0,\infty] \to [0,\infty]$ be an automorphism such that its inverse $\eta^{-1}\colon [0,\infty] \to [0,\infty]$ has $(d-2)$ derivatives (all derivatives) on $(0,\infty)$, $(\eta^{-1})^{(k)}(x) \geq 0$ for all $x \in (0,\infty)$ and $k \in \{1,\dots,d-2\}$ $(k \in N)$ so that $(\eta^{-1})^{(d-2)}$ is a convex function. Then for any $f \in \mathcal{F}_d$ (any $f \in \mathcal{F}_\infty$) also $f_{c2} \in \mathcal{F}_n$ (F_∞) given

$$f_{c2} = \eta \circ f$$

(see Proposition 5 in [Bacigál et al., 2015]).

Construction 3. Let $f \in \mathcal{F}_d$ and $\lambda \in (0,1]$ then for any $d \in \{2,3,\dots\} \cup \{+\infty\}$ also $f_{c3} \in \mathcal{F}_d$, given definition

$$f_{c3}(x) = f(\lambda x) - f(\lambda)$$

obtained as a result of univariate conditioning in [Mesiar et al., 2008] for $d = 2$ and further generalized in Proposition 6 by [Bacigál et al., 2015].

[5] This is in turn a special case of the construction method *pointwise composition of copulas* already metioned at the begining of the chapter. In fact, so are also distortions described in Section 2.1, yet due to different historical origins they are being distinguished from pointwise composition methods.

2.3.2 From a given function

Construction 4. Let $h\colon [a_1, a_2] \to [-\infty, \infty]$ be a strictly decreasing convex continuous function. Then for any non-trivial bounded interval $[b_1, b_2] \subseteq [a_1, a_2]$ (if $h(a_2) = -\infty$ then $[b_1, b_2] \subset [a_1, a_2)$) the function

$$f_{c4}(x) = h(b_1 + x(b_2 - b_1)) - h(b_2)$$

is an additive generator from $\mathcal{F}_d$ whenever h is d-monotone and an universal generator whenever h is totally monotonic. As an example consider the case $h = \coth$, $b_1 = 0$ and $b_2 > 0$ leading to the generator $f(x) = \coth(b_2 x) - \coth(b_2)$ we characterized in [Najjari et al., 2014]

Obviously, any additive generator satisfies the constraints of Construction 4, and hence the method can be seen as an extension and generalization of Construction 3. Moreover, we can generalize Constructions 1 and 2 to make additive generators by means of Construction 4 in two ways. Either we apply them directly to the introduced additive generators f_{c4}, or we apply them (in modified form) to the generating function h. We illustrate the later approach in Theorem 7 of [Bacigál et al., 2015].

Another construction uses the Williamson and Laplace transforms to get a generator of Archimedean copula in any dimension (see sections 1.3.2 and 1.3.1, respectively).

Construction 5. Let $F_X\colon (-\infty, \infty) -> [0, 1]$ with $F_X(x) = 0$ for all $x \leq 0$ be a distribution function of a positive random variable X, then $f_{c5} = g_{c5}^{(-1)} \in \mathcal{F}_d$ ($\mathcal{F}_\infty$) given that

$$g_{c5} = \mathcal{W}_d(F_X) \qquad (g_{c5} = \mathcal{L}_X)$$

for any $d \in \mathbb{N}$.

The above construction is in details examined by [Bacigál & Ždímalová, 2017], where many examples are given. As shown there, the two transformations mostly yield families with no explicit form of generator f, which complicates its application. As a remedy we proposed to discretize X so that F_X becomes a sum of Dirac functions [6] and the corresponding generator functions g and f are piecewise polynomial (linear segments constitute $g \in \mathcal{G}_2$, quadratic $g \in \mathcal{G}_3$, etc.). We also proved that convergence of such approximation in F_X implies convergence in the resulting copula. The approximation copula is practically a mosaic made of the weakest Archimedean copula scaled to each discretization cell (d-dimensional hypercube).

There is also another interesting aspect of the construction method. Since distribution of a single random variable can generate a stochastic dependence structure, one may ask what is the connection between their statistical properties and how can we use them to design (or identify) a joint distribution of some random variables of our interest? We raise this question in [Bacigál, 2017] as particular open problems. To briefly illustrate the matter, note that, for instance, single-valued X (no more a 'variable') leads to counter monotonic dependence through the Williamson 2-transform and to independence by the Laplace transform. On the other hand, independence is linked by the inverse Williamson d-transform with the sum of d independent and exponentially distributed random variables, which follows the so-called Erlang probability distribution. Further, the exponential distribution with

[6] Dirac function is defined as $\delta_{x_0}(x) = \begin{cases} 0 & x < x_0 \\ 1 & x \geq x_0 \end{cases}$

arbitrary parameter gives the same copula (Clayton with parameter $p = 1$) through the Laplace transform.

Construction 5 can be extended such that we start with an arbitrary generator ending up with generator of Archimedean copula in different dimension, such as proposed in the following construction that formally belongs to Section 2.3.1.

Construction 6. Take, for an arbitrary $m \in \mathbb{N}$, an additive generator $f = g^{(-1)} \in \mathcal{F}_m$, introduce a positive distance function $F = \mathcal{W}_m^{-1}(g)$, possibly modify F into a new positive distance function $\tilde{F}$ (e.g. $\tilde{F}(x) = F(x-a)$ for a fixed constant $a \in (0,\infty)$), finally for a fixed $n \in \mathbb{N}$ apply Construction 5, i.e, $\tilde{g} = \mathcal{W}_n(\tilde{F}_X)$ so that $\tilde{f} \in \mathcal{F}_n$.

2.3.3 Gluing

Another construction method for additive generators from $\mathcal{F}_2$ is based on the gluing of two additive generators from $\mathcal{F}_2$ [Bacigál et al., 2015, Theorem 8]. Note that due to the Williamson transform this approach can be extended to any dimension.

Construction 7. Let $f_1, f_2 \in \mathcal{F}_2$ and $k \in (0,1)$ be given. Then

$$f_{c7}(x) = \begin{cases} \frac{f_1(x)}{f_1(k)} & \text{if } x \in [0,k], \\ \frac{f_2(x)}{f_2(k)} & \text{otherwise} \end{cases}$$

generates 2-dimensional Archimedean copula whenever $\frac{f'_{1-}(k)}{f_1(k)} \leq \frac{f'_{2-}(k)}{f_2(k)}$. Observe that the resulting copula C_{c7} can be viewed as the gluing of copulas C_1, C_2 (generated by f_1,f_2, respectively) via an interpolation method, $C_{c7}(x,y) = C_1(x,y)$ for all $(x,y) \in [0,k]^2$, and $C(x,y) = C_2(x,y)$ for all $(x,y) \in [k,1]^2$ such that $f_2(x) + f_2(y) \leq f_2(k)$. Also the positive multiplicative constants do not influence our gluing method.

2.3.4 Aggregation

In this subsection we briefly characterize aggregation functions[7] preserving the classes $\mathcal{F}_2$ and $\mathcal{G}_2$. The topic is in details covered by our paper [Bacigál et al., 2015].

Note, first of all, that the class $\mathcal{F}$ is convex, i.e., for any $f_1, \dots, f_n \in \mathcal{F}_2$ and $c_1, \dots, c_n \in [0,1]$, $\sum_{i=1}^n c_i = 1$, also $f = \sum_{i=1}^n c_i f_i \in \mathcal{F}_2$. Due to the already mentioned fact that any positive multiple $c\,f$ of an additive generator $f \in \mathcal{F}_2$ is again an additive generator, $c\,f \in \mathcal{F}_2$, we see that one can relax the constraint $\sum_{i=1}^n c_i = 1$ into $\sum_{i=1}^n c_i > 0$, i.e., any non-trivial non-negative linear combination of additive generators from $\mathcal{F}$ is again an element of $\mathcal{F}$. For more details and examples see our older paper [Bacigál et al., 2010].

Construction 8. Let $n \in \{2,3,\dots\}$, $A\colon [0,\infty]^n \to [0,\infty]$ be an aggregation function, and $f_1, \dots, f_n \in \mathcal{F}_2$. Then relation

$$f_{c8} = A(f_1, \dots, f_n)$$

yields a generator from $\mathcal{F}_2$ whenever A is a continuous jointly strictly increasing aggregation function satisfying[8] $2A\left(\mathbf{x} + \frac{\mathbf{y}}{2}\right) \leq A(\mathbf{x}) + A(\mathbf{x}+\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in (0,\infty)^n$,

[7] For $n \in \{2,3,\dots\}$, an aggregation function $A\colon [a,b]^n \to [a,b]$ is characterized by the increasing monotonicity in each coordinate and by boundary conditions $A(a,\dots,a) = a$ and $A(b,\dots,b) = b$.

[8] The following sufficient and necessary condition for A to preserve $\mathcal{F}_2$ can be seen as a weaker form of the so-called Jensen convexity and the ultramodularity, see Remark 1 in [Bacigál et al., 2015].

see Theorem 2 of [Bacigál et al., 2015]. Examples of A count weighted sum, product, maximum and p-norm.

Note that each composition of such aggregation functions will again satisfy all constraints of this construction. Moreover, if A has 0 as its annihilator (i.e., if $x_i = 0$ for some $i \in \{1, \ldots, n\}$ then $A(x_1, \ldots, x_n) = 0$) also new aggregation given by $A(x_1 + c_1, \ldots, x_n + c_n) = 0$, where $c_i \geq 0$ and $\prod_{i=1}^{n} c_i = 0$, will also satisfy the constraints of this construction.

As for aggregating pseudo-inverses, observe first that $g_1, g_2 \in \mathcal{G}_2$ generate the same Archimedean copula if and only if $g_1(x) = g_2(c\,x)$ for some constant $c \in (0, \infty)$. Similarly to class $\mathcal{F}_2$, also the class $\mathcal{G}_2$ is convex, see [Bacigál et al., 2010]. However, one should stress that a non-trivial convex combination of pseudo-inverses $g_1, \ldots, g_n \in \mathcal{G}_2$ related to a given Archimedean copula C yields a pseudo-inverse $g \in \mathcal{G}_2$ linked to some different copula D (contradicting the related convex combination of additive generators).

Construction 9. Let $n \in \{2, 3, \ldots\}$, $A\colon [0, \infty]^n \to [0, \infty]$ be an aggregation function, and $g_1, \ldots, g_n \in \mathcal{G}_2$. Then

$$f_{c9} = A(g_1, \ldots, g_n)$$

is a generator pseudo-inverse from $\mathcal{G}_2$ whenever A is a continuous weak[9] jointly strictly increasing aggregation function satisfying[10] $A\left(\frac{\mathbf{x}+\mathbf{y}}{2}\right) \leq \frac{A(\mathbf{x})+A(\mathbf{y})}{2}$ for all $\mathbf{x}, \mathbf{y} \in (0, 1)^n$, $\mathbf{x} \geq \mathbf{y}$ (such that if $x_i = y_i$ then $x_i = y_i = 0$, $i \in \{1, \ldots, n\}$), see [Bacigál et al., 2015, Theorem 3].

Due to the relaxed joint strict monotonicity in Construction 8, one can construct aggregation functions preserving pseudo-inverses of additive generators of Archimedean copulas by means of aggregation functions preserving additive generators of Archimedean copulas, but not vice-versa.

2.4 Dependence functions

In this subsection we summarize constructions that 1) link dependence functions to powers of generators in Archimax copulas, or 2) combine two or more possibly symmetric dependence functions into an asymmetric one, and finally, 3) are based on partitions.

2.4.1 Linked to power of generator

First, please recall that when $f \in \mathcal{F}_2$ then also its power $f^{\lambda} \in \mathcal{F}_2$, $\lambda \geq 1$ (due to Construction 2).

Construction 10. Let $C_{f,A}$ be an Archimax copula with generator f and dependence function A. Then for any f,A and $\lambda \geq 1$, the Archimax copula $C_{f^{\lambda},A}$ is also an

[9] The monotonicity of pseudo-inverses $g \in \mathcal{G}_2$, i.e., the fact that it is strictly decreasing on $[0, a]$ and that it vanishes on $[a, \infty]$, with $a = Min(x \in (0, \infty)|g(x) = \infty)$ (observe that due to $g(\infty) = 0$ and continuity of g, a is well defined), is preserved by an aggregation function A if and only if $A(\mathbf{x}) > A(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in [0, 1]^n$ such that $x_1 \geq y_1, \ldots, x_n \geq y_n$, $A(\mathbf{y}) > 0$, and if $y_i > 0$ then $x_i \neq y_i$, $i \in \{1, \ldots, n\}$. We will call this property *weak joint strict increasingness* of A. Obviously, each jointly strictly increasing aggregation function A is also weak jointly strictly increasing.

[10] Due to the continuity of A the following property can be seen as an ordered convexity (ultramodularity), now on the $[0, 1]$ scale.

copula based on generator f, specifically $C_{f^\lambda,A} = C_{f,A_{c10}}$ where

$$A_{c10}(t) = A_{(\lambda)}(t)\left[A\left(\left(\frac{t}{A_{(\lambda)}(t)}\right)^\lambda\right)\right]^{1/\lambda},$$

with $A_{(\lambda)}(t) = \left(t^\lambda + (1-t)^\lambda\right)^{1/\lambda}$ (dependence functions of the Gumbel family), see Proposition 2.1 of [Bacigál et al., 2011].

This construction reveals an important fact about the structure of Archimax copulas. For any generator f, classes $\mathcal{A}_{f^\lambda}$ of Archimax copulas based on generators f^λ, $\lambda \in [1,\infty)$, are nested, and $\mathcal{A}_{f^\lambda} \subsetneq \mathcal{A}_{f^\mu}$ whenever $1 \le \mu < \lambda \le \infty$, where $\mathcal{A}_{f^\infty} = \bigcap_{\lambda=1}^{\infty} \mathcal{A}_{f^\lambda} = \{M\}$. Therefore it is important to know the basic form $\dot{f}$ of each generator f, $f = \dot{f}^\lambda$, such that for any $\lambda \in (0,1)$, $\dot{f}^\lambda$ is no more convex. Such generators $\dot{f}$ will be called *basic generators* and they correspond to Archimedean copulas $C_{\dot{f}}$ such that for any $p > 1$, the corresponding L_p-norm $\|C_{\dot{f}}\|_p > 1$ (for more details we recommend [De Baets et al., 2010; Mesiarová, 2007]). Due to Proposition 2.2 of [Bacigál et al., 2011] we may find it through relation $\dot{f} = f^{1-\alpha}$ where $\alpha = \inf\left\{\frac{f(x)f''(x)}{(f'(x))^2} \,\middle|\, x \in]0,1[\text{and } f'(x), f''(x) \text{ exist}\right\}$ and (due to Proposition 2.3 of [Bacigál et al., 2011]) check if some generator f is basic through property $\dot{f}'(1^-) \neq 0$.

2.4.2 Combination

Construction 11. Let $A_1,\dots,A_n$ be dependence functions. Then for any probability vectors $(a_1,\dots,a_n), (b_1,\dots,b_n) \in [0,1]^n$, $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i = 1$, also the function

$$A_{c11}(t) = \sum_{i=1}^{n} (ta_i + (1-t)b_i)\, A_i\left(\frac{ta_i}{ta_i + (1-t)b_i}\right).$$

is a dependence function. For $(a_1,\dots,a_n) = (b_1,\dots,b_n)$ it turns into the standard convex sum $A(t) = \sum_{i=1}^n a_i A_i(t)$.

Given and prooved as Proposition 3.1 of our paper [Bacigál et al., 2011], observe that this method can be deduced by induction from the original formula given in [Khoudraji, 1995], as well as seen as extension of Proposition 3 of [Genest et al., 1998] dealing with A_1, A_2.

Construction 11 was further extended to d-dimensional case by [Mesiar & Jagr, 2013] in terms of tail dependence function ℓ:

Construction 12. Let $\ell_1,\dots,\ell_n$ be tail dependence functions and let $\alpha_j i \ge 0$ and $\sum_{j=1}^n \alpha_{ji} = 1$, $j \in \{1,\dots,n\}$, $i \in \{1,\dots,d\}$. Then the function $\ell \to [0,\infty)^d \to [0,\infty)$ given by

$$l_{c12}(x_1,\dots,x_d) = \sum_{j=1}^{n} \ell_j(\alpha_{j1}x_1,\dots,\alpha_{jd}x_d)$$

is also a tail dependence function. Moreover, if $\alpha_{ji} = \lambda_j$, $j \in \{1,\dots,n\}$, $\forall i,j$, then $\ell_{c12} = \sum_{j=1}^n \lambda_j \ell_j$ is the standard convex sum.

Evidently, it allows to introduce asymmetric Archimax copulas even if starting from symmetric Archimax copulas, similarly to the following construction method.

Construction 13. For a dependence function A denote by B a $[0,1] \to [0.1]$ function given by $B(t) = A(t) - 1 + t$. Each such a B is characterized by its convexity, non-decreasingness and boundary conditions $\max(0, 2t-1) \le B(t) \le t$ with pseudo-inverse $B^{(-1)}(u) = \sup\{t \in [0,1] | B(t) \leq u\}$. Let $A_1, \dots, A_n$ be dependence functions with corresponding functions $B_1, \dots, B_n$ and let $(\lambda_1, \dots, \lambda_n) \in [0,1]^n$ be a probability vector. Then

$$A_{c13}(t) = \left(\sum_{i=1}^{n} \lambda_i B_i^{(-1)} \right)^{(-1)} (t) + 1 - t$$

is a dependence function, see Proposition 3.3 of [Bacigál et al., 2011].

2.4.3 Partitioning

By the following construction we briefly summarize result of [Mesiar & Jagr, 2013], see Theorem 3.1.

Construction 14. For a fixed $d \geq 2$, consider a partition $\mathcal{P} = \{B_1, \dots, B_k\}$ of the set $\{1, \dots, d\}$. Then

$$\ell_{c14}(x_1, \dots, x_d) = \sum_{j=1}^{k} \left(\bigvee_{i \in B_j} x_i \right),$$

where $\bigvee$ denotes maximum, is a tail dependence function. For instance, the finest partition $\mathcal{P}^* = \{\{1\}, \dots, \{d\}\}$ gives $\ell^*(\boldsymbol{x}) = x_1 + \ldots + x_d$ while the coarse one, $\mathcal{P}_* = \{1, \dots, d\}$, results to $\ell_*(\boldsymbol{x}) = \max(x_1, \dots, x_d)$.

The corresponding EV copula $C(u_1, \dots, u_d) = \prod_{j=1}^{k} min(u_i | i \in B_j)$ describes the stochastic dependence structure of random variables $(X_1, \dots, X_d)$ such that for any X_n, X_m, if $n, m \in B_j$ for some j, then X_n and X_m are co-monotone, otherwise they are independent.

Chapter 3

Applications

Previous chapter have provided a theoretical basis - either our results or methods well-known in the literature - necessary for understanding our contributions in practical dependence modeling. Yet in the first subsection we summarize methods of model building and outline the possible outcomes needed in practice. Then software solutions is introduced and finally we discuss several representative study cases.

3.1 Model building and inference

Usually the procedure to build a model of dependence using copulas start with plotting the individual data series $X_{1j}, \dots, X_{nj}$, $j = 1, \dots, d$, (most often time series) to 1) inspect serial dependence - temporal, spatial - and in case of doubts test it formally (test for trend, seasonality, periodicity, autocorrelation, structural breaks etc.). An eventual serial dependence needs to be removed by a suitable model so that residuals are independent and identically distributed, in order to prevent a bias in copula model estimates. The next step is to 2) rescale the individual data series into $[0, 1]$ interval by means of their (marginal) distribution functions, either parametric (if they are known), or empirical

$$F_j(x) = \frac{1}{n+1} \sum_{i=1}^{n} 1_{X_{ij} \leq x}, \quad j = 1, \dots, d,$$

where 1_A is the indicator function which yields 1 whenever A is true and 0 otherwise, so that after the transformation we get pseudo-observations $U_{ij} = F_j(X_{ij})$. After 3) plotting them in a scatter plot (often a matrix of bivariate scatter plots) an inspection of dependence structure helps to choose suitable parametric families of copulas to be 4) estimated. Finally 5) the estimated copula models needs to be verified by a goodness-of-fit (GOF) test.

Basically there are two main methods recently used for estimating one-parameter families. One uses various measures of dependence, such as Kendall's tau through formal relation with parameter θ of a copula C_θ,

$$\tau(\theta) = 4 \iint_{[0,1]^2} C_\theta(u, v) dC_\theta(u, v) - 1,$$

the another is based on maximization of a likelihood function

$$L(\theta) = \sum_{i=1}^{n} \log\left(c_\theta(U_{i1}, \dots, U_{id})\right)$$

employing copula density c_θ (which is n-order mixed derivative with respect to all variables), see [Genest et al., 1995]. For general multi-parameter copulas $C_{\boldsymbol{\theta}}$ (not, e.g., the multivariate normal or pair-copulas) the first method is problematic and to our best knowledge no satisfactory study has been presented so far.

Goodness of fit can be checked by comparing (L_2-norm) squared distances

$$S_n = \sum_{i=1}^{n} \left(C_n(U_{i1}, \ldots, U_{id}) - C_{\boldsymbol{\theta}}(U_{i1}, \ldots, U_{id})\right)^2$$

between estimated parametric copulas $C_{\boldsymbol{\theta}}$ and empirical copula function

$$C_n(u_1, \ldots, u_d) = \frac{1}{n} \sum_{i=1}^{n} 1_{U_{i1} \leq u_1} \ldots 1_{U_{i,n} \leq u_n}, \quad (u_1, \ldots, u_d) \in [0, 1]^d,$$

or alternatively - instead of copula - using the Kendall's distribution function

$$K_C(w) = \mathbb{P}(C(U, V) < w) = \int_{[0,1]^2} 1_{C(u,v) \leq w} dC(u, v), \quad w \in [0, 1],$$

calculated from parametric copula C and estimated from observations. For further details and alternatives see [Genest et al., 2009].

The best fitting copula models can be used for various purposes, e.g., a) by inspecting properties such as tail dependence and various symmetries it can be made inference about the true underlying distribution, calculate b) probabilities that some (or all) variables will (not) exceed certain threshold, or conversely, c) quantiles which will be (not) exceeded with some uncertainty. Consider the following cases, for simplicity with just two variables:

- $\mathbb{P}(U \leq u \wedge V \leq v) = C(u, v)$, probability of simultaneous non-exceedence;
- $\mathbb{P}(U > u \vee V > v) = 1 - C(u, v)$, at least one variable will exceed its threshold;
- $\mathbb{P}(U \leq u \wedge V \leq v) = 1 - u - v + C(u, v)$, simultaneous exceedence;
- $\mathbb{P}(U \leq u | V = v) = \frac{\partial C(u,v)}{\partial v}$, conditional probability, and similarly
- $\mathbb{P}(U \leq u | V \leq v) = \frac{C(u,v)}{v}$,
- $\mathbb{P}(U > u | V > v) = 1 - \frac{u - C(u,v)}{1-v}$.

In practice, a level curves plot is often useful, e.g., for reading the return period of some hydrological phenomenon such as flood or drought.

3.2 Software tools

Since the turn of century when copulas began to attract attention of masses, several software tools arose. The first public yet commercial to mention was EVANESCE library [Carmona & Morrisson, 2000] included in FinMetrics extension to S programming environment (predecessor of R), that provided a rich battery of copula classes, though only bivariate. With growing popularity of R (free software environment for statistical computing and graphics, [R Core Team, 2016]) there emerged open-source packages like *copula* [Hofert et al., 2017] and *VineCopula* [Schepsmeier et al., 2017] that are still under vivid development. For further reading about copula software, both commercial and open-source, see our paper [Bacigál, 2012][1], here we summarize major features of the two above mentioned R packages and introduce our own.

3.2.1 Well established packages for R

The package *VineCopula* provides tools for the statistical analysis of vine copula models, which represent an flexible construction method of higher dimensional copulas based on bivariate copulas, conditioning and a graphical tool for labeling constraints in high-dimensional probability distributions. The package includes tools for parameter estimation, model selection, simulation, goodness-of-fit tests, and visualization as well as tools for estimation, selection and exploratory data analysis of bivariate copula models (several Archimedean, two elliptical and EV copulas, with rotations).

The package *copula* implements commonly used (two elliptical, five Archimedean, four extreme-value and other) copula families, as well as their rotations, mixtures and asymmetrizations, moreover it provides nested Archimedean (hierarchical) copulas with related tools, methods for density, distribution, random number generation, bivariate dependence measures, Rosenblatt transform, Kendall distribution function, perspective and contour plots. Fitting of copula models with potentially partly fixed parameters, including standard errors, serial independence tests, copula specification tests (independence, exchangeability, radial symmetry, extreme-value dependence, goodness-of-fit) and model selection based on cross-validation is provided as well. Finally, empirical copula, smoothed versions, and non-parametric estimators of the Pickands dependence function are mentioned in the description.

3.2.2 R package *acopula*

In the present subsection, which is built on our paper [Bacigál, 2013c] later extended in [Bacigál, 2013b], we introduce our software package *acopula* developed under the environment R and provided for public use since 2013 on the Comprehensive R Archive Network (CRAN) which is an official database of packages for R, see [Bacigál, 2013a]. It extended current offerings (on the imaginary software market) a) by class of *Archimax* copulas and b) by several handy tools to test, modify, manipulate and inference both from them and from *arbitrary* user-defined continuous copulas, thus making copulas ready for *application*. That explains the initial letter of the package name.

The motivation for such a project lies in the lack of inference tools back in the years of its birth, before 2013. The software solutions available around were able to fit copulas and generate random samples, but did not offer calculation of neither conditional probabilities nor quantiles needed in practice mainly for prediction and assessing risk. Also our research was focused on the superclass embracing Archimedean and EV copulas, that had and still has no support in copula software.

[1] The referenced paper may now appear a bit out-dated considering a rapid development common for information technologies, however the truth is that implementation of copula modeling routines in modern commercial software ceased at few popular families such as Gaussian, Gumbel, Frank, Clayton and possibly t-copula, with some bright exceptions including, e.g., experimental support for hierarchical copulas in SAS. Thus the development is concentrated practically in open-source environments of which R became the key player in academic (research and education) as well as commercial space (mainly econometrics and risk management).

We made the package to be relatively simple for researchers and practitioner to analyze their data and even easily insert their own copula definitions, as demonstrated bellow.

Every parametric class/family of copulas is defined within a list, either by its generator (in case of Archimedean copulas), Pickand's dependence function (Extreme-Value copulas) or directly by cumulative distribution function (CDF) with/or only by its density. Example of one such definition list follows[2] for generator of Gumbel-Hougaard family of Archimedean copulas

```
> genGumbel()
$parameters
[1] 4
$pcopula
function (t, pars) exp(-sum((-log(t))^pars[1])^(1/pars[1]))
$gen
function (t, pars) (-log(t))^pars[1]
$gen.der
function (t, pars) -pars[1]*(-log(t))^(pars[1]-1)/t
$gen.der2
function (t, pars) pars[1]*(-log(t))^(pars[1]-2)*(pars[1]-1-log(t))/t^2
$gen.inv
function (t, pars) exp(-t^(1/pars[1]))
$gen.inv.der
function (t, pars) -exp(-t^(1/pars[1]))*t^(1/pars[1]-1)/pars[1]
$gen.inv.der2
function (t, pars)
exp(-t^(1/pars[1]))*t^(1/pars[1]-2)*(pars[1]+t^(1/pars[1])-1)/pars[1]^2
$lower
[1] 1
$upper
[1] Inf
$id
[1] "Gumbel"
```

where, though some items may be fully optional (here `\$pcopula` and `\$id`), they can contribute to better performance or transparency. The user is encouraged to define new parametric families of Archimedean copula generator (likewise dependence function or copula in general) according to his/her needs, bounded only by this convention and allowed to add `pcopula` (stands for probability distribution function or CDF), `dcopula` (density) and `rcopula` (random sample generator) items.

Currently implemented generators can be listed.

```
> ls("package:acopula",pattern="gen")
[1] "genAMH" "genClayton" "generator" "genFrank" "genGumbel" "genJoe" "genLog"
```

Notice the generic function `generator` which links to specified definition lists.

Similarly, Pickand's dependence functions are defined, namely Gumbel-Hougaard, Tawn, Galambos, Hüsler-Reiss (last three form only bivariate EV), extremal dep. functions and generalized convex combination of arbitrary valid dep. functions (see [Mesiar & Jagr, 2013]). So are definition lists available for generic (i.e., not necessarily Archimax) copula, e.g. normal, Farlie-Gumbel-Morgenstern, Plackett and Gumbel-Hougaard parametric family. Their corresponding function names starts with `dep` and `cop`, respectively.

[2]Output printing is simplified whenever contains irrelevant parts.

As the class of Archimax copulas contains Archimedean and EV class as its special cases, the setting `depfu = dep1()` and `generator = genLog()` can distinguish them, respectively.

Any definition list item can be replaced already during the function call as shown bellow. Thus one can set starting value of parameter(s) and their range in estimation routine, for instance.

First thing one would expect from a copula package is to obtain a value of desired copula in some specific point. To show variability in typing commands, consider again Gumbel-Hougaard copula with parameter equal to 3.5 in point (0.2,0.3). Then the following commands give the same result.

```
> pCopula(data=c(0.2,0.3),generator=genGumbel(),gpars=3.5)
> pCopula(data=c(0.2,0.3),generator=genGumbel(parameters=3.5))
> pCopula(data=c(0.2,0.3),generator=generator("Gumbel"),gpars=3.5)
> pCopula(data=c(0.2,0.3),generator=generator("Gumbel",parameters=3.5))
> pCopula(data=c(0.2,0.3),copula=copGumbel(),pars=3.5)
> pCopula(data=c(0.2,0.3),copula=copGumbel(parameters=3.5))
> pCopula(data=c(0.2,0.3),generator=genLog(),depfun=depGumbel(),dpars=3.5)
> pCopula(data=c(0.2,0.3),generator=genLog(),depfun=depGumbel(parameters=3.5))
[1] 0.1723903
```

If we need probabilities that a random vector would not exceed several points, those can be supplied to `data` in rows of matrix or data frame.

Conversely, given an incomplete point and a probability, the corresponding quantile emerge.

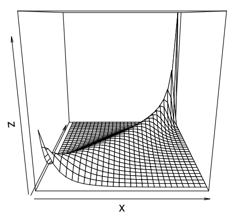
```
> pCopula(c(0.1723903,0.3),gen=genGumbel(),gpar=3.5,quantile=1)
> pCopula(c(NA,0.3),gen=genGumbel(),gpar=3.5,quan=1,prob=0.1723903)
> qCopula(c(0.3),quan=1,prob=0.1723903,gen=genGumbel(),gpar=3.5)
[1] 0.1999985
```

Conditional probability $P(X < x|Y = y)$ of a random vector (X, Y) has similar syntax.

```
> cCopula(c(0.2,0.3),conditional.on=2,gen=genGumbel(),gpar=3.5)
[1] 0.2230437
> qCopula(c(0.3),quan=1,prob=0.2230437,cond=c(2),gen=genGumbel(),gpar=3.5)
[1] 0.200005
```

Density of a copula can be visualized such as in the following example.

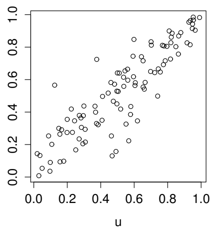
```
x <- seq(0,1,length.out=30)
y <- seq(0,1,length.out=30)
z <- dCopula(expand.grid(x,y),generator=genGumbel(),gpars=3.5)
dim(z) <- c(30,30)
persp(x,y,z)
```



If definition lists do not contain explicit formulas for (constructing) density, the partial derivatives are approximated linearly. This is mostly the case with 3- and more-dimensional copulas.

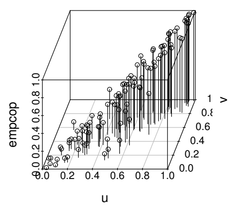
Sampling from the copula is, unsurprisingly, also provided.

```
sample <- rCopula(n=1000,dim=2, generator=genGumbel(), gpars=3.5)
plot(sample)
```



Sometimes no assumption about parametric family of copula is made, instead an empirical distribution is of more interest. Then for a given data, say, the previous random sample, one may ask for value of empirical copula in specific point(s) and more easily in the points of its discontinuity.

```
> pCopulaEmpirical(c(0.2,0.3),base=sample)
[1] 0.14
> empcop <- pCopulaEmpirical(sample)
> scatterplot3d::scatterplot3d(cbind(sample,empcop),type="h",angle=70)
```



Currently, there are two universal methods for parameters estimation implemented in the package (named `technique`): "ML", maximum (pseudo)likelihood method employing copula density, and "LS", least squares method minimizing distance to empirical copula. Each 'technique' supplies function to perform optimization `procedure` over, thus finding those parameters that correspond to an optimum. The 'procedures' are three: "optim", "nlminb" and "grid". First two are system native, based on well-documented smart optimization methods, the third one uses brute force to get approximate global maximum/minimum and can be useful with multi-parameter copulas, at least to provide starting values for the other two 'procedures'. The next few examples sketch various options one has got for copula fitting.

```
> eCopula(sample,gen=genClayton(),dep=depGumbel(),
+ technique="ML",procedure="optim",method="L-BFGS-B")
generator parameters:  0.09357958
depfun parameters:  3.52958
ML function value:  82.63223
convergence code:  0
> eCopula(sample,gen=genClayton(),dep=depGumbel(),tech="ML",proc="nlminb")
```

```
generator parameters:  0.09183014
depfun parameters:  3.533706
ML function value:  82.63228
convergence code:  0
> eCopula(sample,gen=genClayton(),dep=depGumbel(), tech="ML",proc="grid",
+ glimits=list(c(0),c(5)),dlimits=list(c(1),c(10)),pgrid=10)
generator parameters:  0.5555556
depfun parameters:  3
ML function value:  80.63322
convergence code:
```

So far, no precision for copula parameters is provided.

Having set of observations, it is often of great interest to test whether the estimated copula suffices to describe dependence structure in the data. For this purpose many goodness-of-fit tests were proposed, yet the principle remains to use different criterion than was employed with estimation of the copula parameters. Here we implement one of the 'blanket' tests described in [Genest et al., 2009] that is based on Kendall's transform. In the example below normal copula is tested on the Gumbel copula sample data.

```
> gCopula(sample,cop=copNormal(),
+ etechnique="ML",eprocedure="optim",ncores=1,N=100)
Loading required package: mvtnorm
|================================================================| 100%

Blanket GOF test based on Kendall's transform

statistic       q95   p.value
0.1195500 0.1658125 0.1800000
-----------------------------
data:  sample
copula:  normal
estimates:
pars     fvalue
0.9155766 80.3420886
```

Although the p-value does not lead to rejection of the copula adequacy, its low value and small data length arouse suspicion. As for the other arguments, `N` sets number of bootstrap cycles whereas their parallel execution can be enabled by setting number of processor cores in `ncores`. Package *mvtnorm* has been loaded to assist with simulation from normal copula, and when missing, internal but slower routine would be performed instead.

The traditional parametric bootstrap-based procedure to approximate p-value, when theoretical probability distribution of the test statistic is unknown, is reliable yet computationally very exhaustive, therefore recently a method based on multiplier central limit theorem and proposed by [Kojadinovic et al., 2011] becomes popular with large-sample testing. Its implementation to testing goodness of parametric copula fit is scheduled for next package update. Nevertheless, the multiplier method takes part here in another test comparing two empirical copulas, i.e. dependence structure of two data sets, see [Rémillard & Scaillet, 2009]. In the following example, random sample of the above Gumbel-Hougaard copula is tested for sharing common dependence structure with sample simulated from Clayton copula, parameter of which corresponds to the same Kendall's rank correlation ($\tau = 0.714$).

```
> sampleCl <- rCopula(n=100,dim=2,generator=genClayton(),gpars=5)
```

```
> gCopula(list(sample,sampleCl),ncores=1,N=100)
|==========================================================| 100%

Test of equality between 2 empirical copulas

statistic        q95    p.value
0.09791672 0.52893392 0.66000000
-----------------------------
data:  sample sampleCl
copula:
estimates:
NULL
```

Obviously, the test fails to distinguish copulas with differing tail dependence, at least having small and moderate number of observations, however it is sensitive enough to a difference in rank correlation.

The last procedure to mention checks the properties of a d-dimensional copula ($d \geq 2$), that is, being d-increasing as well as having 1 as neutral element and 0 as annihilator. The purpose is to assist approval of new copula constructs when theoretical proof is too complicated. The procedure examines every combination of discrete sets of copula parameters, in the very same fashion as within "grid" procedure of `eCopula`, by computing a) first differences recursively over all dimensions of an even grid of data points,i.e., C-volumes of subcopulas, b) values on the margin where one argument equals zero and c) where all arguments but one equals unity. Then whenever the result is a) negative, b) non-zero or c) other than the one particular argument, respectively, a record is made and first 5 are printed as shown below. In the example we examine validity of an assumed Archimedean copula generated by Gumbel-Hougaard generator family, only with a parameter being out of bounds.

```
> isCopula(generator=genGumbel(lower=0),dim=3,glimits=list(0.5,2),
+ dagrid=10,pgrid=4,tolerance=1e-15)

Does the object appears to be a copula(?):  FALSE

Showing 2 of 2 issues:

dim property      value gpar
1   2    monot -0.1534827  0.5
2   3    monot -0.1402209  0.5
```

Three parameter values $(0.5, 1, 1.5, 2)$ were used, each supposed copula were evaluated in 10^3 grid nodes, and every violation of copula properties (the most extremal value per dimension and exceeding `tolerance`) were reported. Thus it is seen, that parameter value 0.5 does not result in copula because 3-monotonicity is not fulfilled (negative difference already in the second-dimension run). Note that without redefinition of lower bound the parameter value 0.5 would be excluded from the set of Gumbel-Hougaard copula parameters.

For the *acopula* package to work many utility functions were created during development that were neither available in the basic R libraries nor they were found in contributed package under CRAN. Most of them are hidden within the procedures described above, however the two following are accessible on demand. The first to mention is a linear approximation of partial derivative of any-dimensional function and of any order with specification of increment (theoretically fading to zero) and area (to allow semi-differentiability)

```
> fun <- function(x,y,z) x^2*y*exp(z)
> nderive(fun,point=c(0.2,1.3,0),order=c(2,0,1),difference=1e-04,area=0)
[1] 2.600004
```

whereas the second utility function numerically approximates integration (by trapezoidal rule) such as demonstrated on example of joint standard normal density with zero correlation parameter

```
> nintegrate(function(x,y) mvtnorm::dmvnorm(c(x,y)),
+            lower=c(-5.,-5.),upper=c(0.5,1),subdivisions=30)
[1] 0.5807843
> pnorm(0.5)*pnorm(1)
[1] 0.5817583
```

fine-tuned by number of subdivisions.

To conclude, all the introduced and exemplified procedures are (a) extendible to arbitrary dimension, which is one of the significant contributions of the package. If explicit formulas are unavailable (through definition lists) then numerical approximation does the job. Another significant benefit is brought by (b) conditional probability and quantile function of the copula, as well as estimation methods based on least squares and grid complementing the usual maximum-likelihood method. Together with implementing (c) generalization of Archimedean and Extreme-Value by Archimax class with a (d) construction method of Pickand's dependence function, (e) test of equality between two empirical copulas, (f) numerical check of copula properties useful in new parametric families development, and (g) parallelized goodness-of-fit test based on Kendall's transform, these all (and under one roof) make the package competitive among both proprietary and open-source software tools for copula based analysis, to the date.

That the package has found its place among people we can show by monthly download statistics aggregated from logs (records) that are provided by the CRAN mirror of RStudio - an integrated development environment for R which has practically become a standard (with no relevant competitor) [RStudio Team, 2016]. Figure 3.1 shows roughly 250 downloads of *acopula* per months comparing to approx. 1000 downloads of *VineCopula* and 2500 downloads of *copula* package. Considering the other two are still actively developed and focusing on multidimensional copula constructions (either Vine or hierarchical) for practical application while *acopula* aimed at providing framework for new (mostly bivariate) Archimax copula development, these are nice results. Needless to say, the numbers are not all the new users, some significant amount can be attributed to repetitive downloads due to annual releases of R main subversions, that by default (but not necessarily) requires re-installation of extra packages, and to updates of the existing package installations.

3.3 Case studies

Though copulas are known since the late fifties, it took approx 40 years to start gaining considerable popularity in applied sciences. The major part of the pie is held by fields where earning or saving substantial amount of money is involved such as quantitative finance, risk management and actuarial sciences, followed by fields where human health and environmental hazards are in stake such as biostatistics, hydrology, climatology and the like. In this subsection, we will catalog not even the

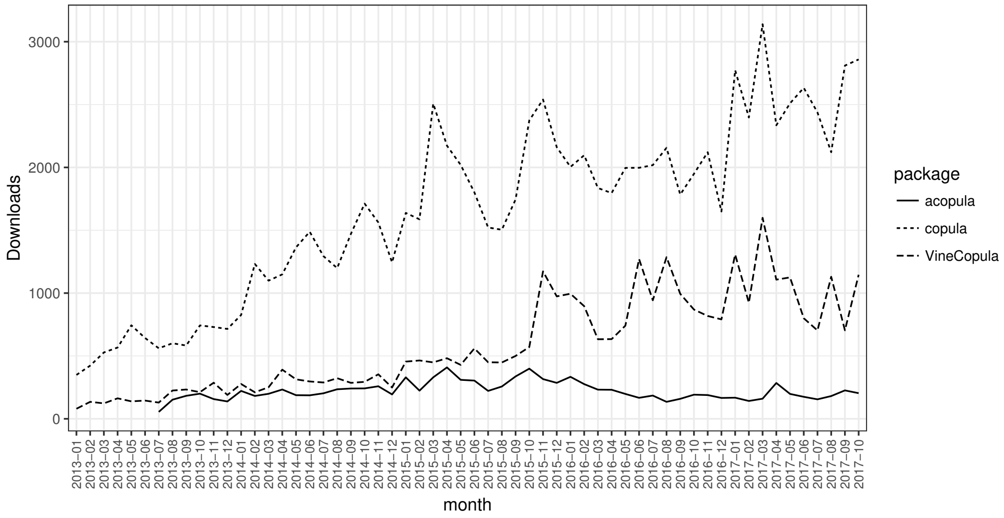


Figure 3.1: Monthly downloads of three copula packages.

most outstanding works [3], instead we directly proceed to studies we have carried out during the last ten years, either a) to show that our particular construction methods are able to give copulas that explain real data dependence structure better than the usual ones, or b) to analyze patterns in relation between variables of interest.

In our early works, we showed suitability of Archimax copulas for modeling non-exchangeable random variables with Archimax copulas, either focusing on a new way of parameters estimation [Capri2008, SvF2008] or employing newly constructed dependence functions based on Constructions 11 to 13 [Bacigál et al., 2011] and Construction 14 [Bacigál & Mesiar, 2012]. As the particular data displaying non-exchangeability property we used the monthly discharges of the river Danube and its tributary Inn. The natural causality of river flow was also used in [Pekárová & Bacigál, 2011] to demonstrate appropriateness of another asymmetric copula model, the (distorted) univariate conditioning stable copulas described in Section 2.2, alongside with rather symmetric actuarial data Loss and ALAE[4].

Superior fit of a particular Archimedean copula family based on Construction 4 to hydrologic (peak river flow and the corresponding flood volume) and meteorologic data (temperature maxima and minima) was demonstrated in [Najjari et al., 2014].

On the other hand in the series of hydrological papers we jointly analyzed two random variables associated with extremal events - maximal river flow (flood peak) and volume of water flown within the same event (flood volume) - to both characterize changes in the relationship and to predict amount of volume occurred during a flood with observed peak.

Specifically, in [Bacigál et al., 2012] we built up joint distribution on summer discharge maxima series recorded during 1877-2002 in the Vltava-Kamýk dam profile and the corresponding volumes, also examined effect of marginal distributions

[3] If interested, please see for instance [Genest et al., 2009] who provide bibliometric evidence to illustrate the development of copula theory in statistics, actuarial science and finance and identifies challenges for he future, further the recent paper [Hao & Singh, 2016] review dependence modeling methods and applications in hydrology, climatology and water resources, we also recommend to read Preface in [Durante & Sempi, 2015] for a quick and critical survey on the use of copulas in practice.

[4] Loss represents amounts that an insurance company pays for claims made under the insurance contract, ALAE stands for alocated loss adjustment expenses - the claim settlement process not attributable to specific claims, e.g., the cost of hiring an attorney to defend a specific claim.

and volume construction on the fitting results, consequently in [Szolgay et al., 2012] the volume was estimated as a quantile (alternatively an expected or mean value) of distribution function conditional on the value of maximum discharge with 10 000 years return period. Employing several alternative settings, such as two constructions of flood wave volume, five copulas, three location measures, marginal distributions and return period values, the results varied hugely. Thus for similar analysis we recommended paying the greatest attention to choice of the copula shape and parametric marginals, in both cases counting in upper-tail behavior, supported by interpretability. Prediction of flood wave volume is used to design hydromechanic structures.

Then in [Szolgay et al., 2016] we analyse the bivariate relationship between flood peaks and volumes in regional context with a focus on flood generation processes in general, the regional differentiation of these and the effect of the sample size on reliable discrimination among models. A total of 72 catchments in North-West of Austria are analyzed for the period 1976–2007. From the hourly runoff data set, 25 697 flood events were isolated and assigned to one of three flood process types: synoptic floods (including long- and short-rain floods), flash floods or snowmelt floods (both rain-on-snow and snowmelt floods). The first step of the analysis examines whether the empirical peak-volume copulas of different flood process types are regionally statistically distinguishable, separately for each catchment and the role of the sample size on the strength of the statements. The results indicate that the empirical copulas of flash floods tend to be different from those of the synoptic and snowmelt floods. The second step examines how similar are the empirical flood peak-volume copulas between catchments for a given flood type across the region. Empirical copulas of synoptic floods are the least similar between the catchments, however with the decrease of the sample size the difference between the performances of the process types becomes small. The third step examines the goodness-of-fit of different commonly used copula types to the data samples that represent the annual maxima of flood peaks and the respective volumes both regardless of flood generating processes (the traditional engineering approach) and also considering the three process-based classes. Extreme-value copulas show the best performance both for synoptic and flash floods, while the Frank copula (which is also radially symmetric) shows the best performance for snowmelt floods. It is concluded that there is merit in treating flood types separately when analysing and estimating flood peak-volume dependence copulas.
The paper summarize and is extended in partial works [Szolgay et al., 2015, 2016; Kohnová et al., 2016; Gaál et al., 2016].

In [Papaioannou et al., 2016] the suitability of various copula families for a bivariate analysis of peak discharges and flood volumes has been examined on streamflow data from selected gauging stations along the whole Danube River. The methodology is applied to two different data samples: 1) annual maximum flood (AMF) peaks combined with annual maximum flow volumes of fixed durations at 5 to 60 days, respectively (which can be regarded as a regime analysis of the dependence between the extremes of both variables in a given year), and 2) annual maximum flood (AMF) peaks with corresponding flood volumes (which is a typical choice for engineering studies).

Finally, [Bacigál et al., 2016] focus on 3-dimensional copula models of returns of US financial markets indices (various bond indices have been investigated in the literature much less than stock indices). Although, for our particular data (comprising two triples of bond indices: US Investment Bond indices and US Corporate Bond indices), the global dominance of more traditional classes of elliptic (especially Student type) 3-dimensional copulas was demonstrated (and some conclusions concerning optimizations of investment portfolios can be based on fairly simple arguments), the optimal local Vine copulas helps to obtain more insight in the detailed development of the investigated triples of investments.

Conclusion

We have shown copulas as a recent and complex mathematical tool for describing stochastic dependence in random vector, summarized their properties, ways of constructions, software tools and some few applications. Since the research in this area is still in its growth phase and there is a loud call for meaningful multivariate models from applied sciences, we may expect that the focus in (theoretical) scientific community will be put on construction of multidimensional yet flexible copulas with well statistically interpretable structure and computationally feasible estimation procedures available.

Bibliography

Bacigál, T. (2012). Recent tools for modelling dependence with copulas and r. In *Forum Statisticum Slovacum*, volume 8, (pp. 62–67).

Bacigál, T. (2013a). *acopula: Modelling dependence with multivariate Archimax (or any user-defined continuous) copula*. R package version 0.9.2.

Bacigál, T. (2013b). Modelling dependence with copulas and r package acopula ver. 0.9.2. In Mesiar, R. & Bacigál, T. (Eds.), *Uncertainty Modelling 2013: Proceedings*, (pp. 7–16). Nakladateľstvo STU. `http://www.math.sk/wiki/SeminarModelovanieNeurcitosti`.

Bacigál, T. (2013c). R package to handle archimax or any user-defined continuous copula construction: acopula. In *Aggregation Functions in Theory and in Practise* (pp. 75–84). Springer.

Bacigál, T. (2017). On some applications of williamson's transform in copula theory. In Torra V., Mesiar R., B. B. (Ed.), *Aggregation Functions in Theory and in Practice. AGOP 2017*, volume 581 of *Advances in Intelligent Systems and Computing*, (pp. 21–30). Springer.

Bacigál, T., Jágr, V., & Mesiar, R. (2011). Non-exchangeable random variables, archimax copulas and their fitting to real data. *Kybernetika*, *47*(4), 519–531.

Bacigál, T., Juráňová, M., & Mesiar, R. (2010). On some new constructions of archimedean copulas and applications to fitting problems. *Neural network world*, *20*(1), 81.

Bacigál, T., Kohnová, S., Szolgay, J., & Hlavčová, K. (2012). Návrhová povodeň: výber typu a odhad parametrov združeného rozdelenia pravdepodobnosti prietoku a objemu pomocou kopule. *Acta Hydrologica Slovaca*, *13*(2), 297–305.

Bacigál, T., Komorník, J., Komorníková, M., & Nguyen, C. (2016). Dynamics of parallel development of the bond market indices in the us market and its multidimensional copula models. *Global and Stochastic Analysis*, *3*(2), 47–59.

Bacigál, T. & Mesiar, R. (2012). 3-dimensional archimax copulas and their fitting to real data. In *Proc. 20th International conference on computational statistics COMPSTAT*, (pp. 81–88).

Bacigál, T., Mesiar, R., & Najjari, V. (2015). Generators of copulas and aggregation. *Information Sciences*, *306*, 81–87.

Bacigál, T., Najjari, V., Mesiar, R., & Bal, H. (2015). Additive generators of copulas. *Fuzzy Sets and Systems*, *264*, 42–50.

Bacigál, T. & Ždímalová, M. (2017). Convergence of linear approximation of archimedean generator from williamson's transform in examples. *Tatra Mountains Mathematical Publications*, *69*, 1–18. to appear.

Capéraà, P., Fougères, A.-L., & Genest, C. (2000). Bivariate distributions with given extreme value attractor. *Journal of Multivariate Analysis*, *72*(1), 30–49.

Carmona, R. & Morrisson, J. (2000). Evanesce, an s-plus library for heavy tail distributions and copulas. *Dept. of Operations Research & Financial Engineering, Princeton University.*

Charpentier, A., Fougères, A.-L., Genest, C., & Nešlehová, J. (2014). Multivariate archimax copulas. *Journal of Multivariate Analysis*, *126*, 118–136.

De Baets, B., De Meyer, H., & Mesiar, R. (2010). Lipschitz continuity of copulas wrt lp-norms. *Nonlinear Analysis: Theory, Methods & Applications*, *72*(9), 3722–3731.

Durante, F. & Jaworski, P. (2012). Invariant dependence structure under univariate truncation. *Statistics*, *46*(2), 263–277.

Durante, F., Jaworski, P., & Mesiar, R. (2011). Invariant dependence structures and archimedean copulas. *Statistics & Probability Letters*, *81*(12), 1995–2003.

Durante, F. & Sempi, C. (2010). Copula theory: an introduction. In *Copula theory and its applications* (pp. 3–31). Springer.

Durante, F. & Sempi, C. (2015). *Principles of copula theory.* CRC press.

Gaál, L., Szolgay, J., Bacigál, T., Kohnová, S., Hlavčová, K., Výleta, R., Parajka, J., & Blöschl, G. (2016). Similarity of empirical copulas of flood peak-volume relationships: a regional case study of north-west austria. *Contributions to Geophysics and Geodesy*, *46*(3), 155–178.

Genest, C., Gendron, M., & Bourdeau-Brien, M. (2009). The advent of copulas in finance. *The European Journal of Finance*, *15*(7-8), 609–618.

Genest, C., Ghoudi, K., & Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, *82*(3), 543–552.

Genest, C., Ghoudi, K., & Rivest, L.-P. (1998). "understanding relationships using copulas," by edward frees and emiliano valdez, january 1998. *North American Actuarial Journal*, *2*(3), 143–149.

Genest, C., Nešlehová, J., & Ziegel, J. (2011). Inference in multivariate archimedean copula models. *Test*, *20*(2), 223–256.

Genest, C., Rémillard, B., & Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and economics*, *44*(2), 199–213.

Gudendorf, G. & Segers, J. (2010). Extreme-value copulas. *Copula theory and its applications*, 127–145.

Hao, Z. & Singh, V. P. (2016). Review of dependence modeling in hydrology and water resources. *Progress in Physical Geography*, *40*(4), 549–578.

Hofert, M., Kojadinovic, I., Maechler, M., & Yan, J. (2017). copula: Multivariate dependence with copulas. *R package version 0.999-17, URL http://CRAN. R-project. org/package= copula.*

Jágr, V., Komorníková, M., & Mesiar, R. (2010). Conditioning stable copulas. *Neural Network World*, *20*(1), 69.

Joe, H. (2015). Dependence modeling with copulas. no. 134 in monographs on statistics and applied probability.

Khoudraji, A. (1995). *Contributions à l'étude des copules et à la modélisation des valeurs extrémes bivariées.* PhD thesis, Université Laval, Québec.

Klement, E. P., Mesiar, R., & Pap, E. (1999). Quasi-and pseudo-inverses of monotone functions, and the construction of t-norms. *Fuzzy Sets and Systems*, *104*(1), 3–13.

Kohnová, S., Gaál, L., Bacigál, T., Szolgay, J., Hlavčová, K., Valent, P., Parajka, J., & Blöschl, G. (2016). Process-based selection of copula types for flood peak-volume relationships in northwest austria: a case study. *Contributions to Geophysics and Geodesy*, *46*(4), 245–268.

Kojadinovic, I., Yan, J., & Holmes, M. (2011). Fast large-sample goodness-of-fit tests for copulas. *Statistica Sinica*, 841–871.

Liebscher, E. (2008). Construction of asymmetric multivariate copulas. *Journal of Multivariate analysis*, *99*(10), 2234–2250.

Ling, C. (1965). Representation of associative functions. *Publicationes Mathematicae Debrecen*, *12*, 189–212.

Mai, J. & Scherer, M. (2012). Simulating copulas, series in quantitative finance, vol. 4.

Mai, J. & Scherer, M. (2014). *Financial engineering with copulas explained*. Springer.

McNeil, A. J. & Nešlehová, J. (2009). Multivariate archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions. *The Annals of Statistics*, 3059–3097.

Mesiar, R. & Jagr, V. (2013). d-dimensional dependence functions and archimax copulas. *Fuzzy Sets and Systems*, *228*, 78–87.

Mesiar, R., Jágr, V., Juráňová, M., & Komorníková, M. (2008). Univariate conditioning of copulas. *Kybernetika*, *44*(6), 807–816.

Mesiar, R. & Pekárová, M. (2010). Ducs copulas. *Kybernetika*, *46*(6), 1069–1077.

Mesiarová, A. (2007). k-lp-lipschitz t-norms. *International Journal of Approximate Reasoning*, *46*(3), 596–604.

Morillas, P. M. (2005). A method to obtain new copulas from a given one. *Metrika*, *61*(2), 169–184.

Najjari, V., Bacigál, T., & Bal, H. (2014). An archimedean copula family with hyperbolic cotangent generator. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *22*(05), 761–768.

Papaioannou, G., Kohnová, S., Bacigál, T., Szolgay, J., Hlavčová, K., & Loukas, A. (2016). Joint modelling of flood peaks and volumes: A copula application for the danube river. *Journal of Hydrology and Hydromechanics*, *64*(4), 382–392.

Pekárová, M. & Bacigál, T. (2011). Distorted univariate conditioning stable copulas. In B. De Baets, R. Mesiar, L. T. (Ed.), *Proceedings of the 6th International Summer School on Aggregation Operators*, (pp. 35–43)., Benevento, Italy. University of Sannio.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rémillard, B. & Scaillet, O. (2009). Testing for equality between two copulas. *Journal of Multivariate Analysis*, *100*(3), 377–386.

RStudio Team (2016). *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, Inc.

Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B., Nagler, T., & Erhardt, T. (2017). *VineCopula: Statistical Inference of Vine Copulas*. R package version 2.1.2.

Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, *8*, 229–231.

Szolgay, J., Gaál, L., Bacigál, T., Kohnová, S., Hlavčová, K., Výleta, R., & Blöschl, G. (2016). A regional look at the selection of a process-oriented model for flood peak/volume relationships. *Proceedings of the International Association of Hydrological Sciences*, *373*, 61–67.

Szolgay, J., Gaál, L., Bacigál, T., Kohnová, S., Hlavčová, K., Výleta, R., Parajka, J., & Blöschl, G. (2016). A regional comparative analysis of empirical and theoretical flood peak-volume relationships. *Journal of Hydrology and Hydromechanics*, *64*(4), 367–381.

Szolgay, J., Gaál, L., Kohnová, S., Hlavcová, K., Výleta, R., Bacigál, T., & Blöschl, G. (2015). A process-based analysis of the suitability of copula types for peak-volume flood relationships. *Proceedings of the International Association of Hydrological Sciences*, *370*, 183.

Szolgay, J., Kohnová, S., Bacigál, T., & Hlavčová, K. (2012). Návrhová povodeň: združená pravdepodobnostná analýza maximálnych prietokov a im prislúchajúcich objemov. *Acta Hydrologica Slovaca*, *13*(2), 289–296.

Index

Appendix

Reprints of the papers:

A.1 Bacigál, T., Juráňová, M., Mesiar, R. (2010). On some new constructions of Archimedean copulas and applications to fitting problems. *Neural network world*, 20 (1), 81–90.

A.2 Bacigál, T., Jágr, V., Mesiar, R. (2011). Non-exchangeable random variables, Archimax copulas and their fitting to real data. *Kybernetika*, 47(4), 519–531.

A.3 Bacigál, T., Najjari, V., Mesiar, R., Bal, H. (2015). Additive generators of copulas. *Fuzzy Sets and Systems*, 264, 42-50.

A.4 Bacigál, T., Mesiar, R., Najjari, V. (2015). Generators of copulas and aggregation. *Information Sciences*, 306, 81-87.

A.5 Bacigál, T., Ždímalová, M. (2017). Convergence of linear approximation of Archimedean generator from Williamson's transform in examples. *Tatra Mountains Mathematical Publications*, 69, 1–18 (in print).

On some new constructions of Archimedean copulas and applications to fitting problems

Tomáš Bacigál, Monika Juráňová and Radko Mesiar *

Abstract: Several constructions of additive generators of binary Archimedean copulas are introduced and discussed. Extension to general Archimedean copulas is also included. Applications to fitting of copulas to real data are given and examplified.

Key words: *Additive generator, Archimedean copula, fitting of copulas*

Received: July 15, 2009
Revised and accepted: January 29, 2010

1. Introduction

Copulas are an improtant tool in modelling dependence structure of multidimensional random vectors. They have numerous applications in finance [6], hydrology [2, 8] etc., see also a recent monograph [18].

Definition 1. For $n \in N$, $n > 1$, an n-dimensional copula is a mapping $C\colon [0,1]^n \to [0,1]$ which is

i) grounded, i.e., $C(x_1, ..., x_n) = 0$ whenever $0 \in \{x_1, ..., x_n\}$ (0 is annihilator of C);

ii) having neutral element 1, i.e., $C(x_1, ..., x_n) = x_i$ whenever all $x_j = 1$ for $j = i$;

iii) n-increasing, i.e., for all $\boldsymbol{x}, \boldsymbol{y} \in [0,1]^n$, $\boldsymbol{x} \leq \boldsymbol{y}$, it holds

$$\sum_{\boldsymbol{\varepsilon} \in \{-1,1\}^n} \left(\prod_{i=1}^{n} \varepsilon_i \right) C(u_1^{(\varepsilon_1)}, ..., u_n^{(\varepsilon_n)}) \geq 0$$

where $u_i^{(1)} = y_i$ and $u_i^{(-1)} = x_i$.

The relationship of copulas and random vectors is clarified by Sklar theorem [20].

*Faculty of Civil Engineering, Slovak University of Technology in Bratislava, bacigal@math.sk, juranova@math.sk, mesiar@math.sk

Neural Network World 1/10, 81–90

Theorem 1 (Sklar). *Let $V = (X_1, ...X_n)$ be a random vector with marginal distributions $F_{X_i}: R \to [0,1]$, $i = 1, ..., n$, $n \geq 2$. Then $F_V: R^n \to [0,1]$ is a joint distribution function of V if and only if there is a copula $C\colon [0,1]^n \to [0,1]$ such that for all $(x_1, ..., x_n) \in R^n$ it holds*

$$F_V(x_1, ..., x_n) = C(F_{X_1}(x_1), ..., F_{X_n}(x_n)). \tag{1}$$

Copula C is unique if and only if V is a continuous random vector.

More details about copulas can be found in Nelsen [15]. Sklar theorem allows to examine rank-dependent characteristics of V directly on the corresponding copula C (for example, rank correlation, i.e., Spearmann's rho). Morewer, due to Sklar theorem, fitting of multivariate distributions to real data can be split into two steps:

i) fitting of marginal (1-dimensional) distributions

ii) fitting of (n-dimensional) copula.

A good fitting of copulas requires a rich backlog of possible candidates for copulas modelling real data dependence structure.

In this paper, we will focus on Archimedean copulas and new construction methods for them. The paper is organised as follows. In the next section, Archimedean copulas are introduced and discussed. Section 3 brings new construction methods for binary Archimedean copulas. In Section 4, new construction methods for general Archimedean copulas are given. Section 5 is devoted to application of introduced methods to the fitting of Archimedean copula to real data. Finally, some concluding remarks are given.

2. Archimedean copulas

Among several classes of copulas, most popular copulas for fitting purposes are Archimedean copulas, i.e., associative copulas with no non-trivial idempotent elements. Note that the classical definition of associativity was related to binary functions, and then n-ary associative copulas were linked to binary copulas. New definition of associativity for n-ary functions [3] allows to introduce n-ary Archimedean copulas straightforwardly.

Definition 2. An n-ary copula $C\colon [0,1]^n \to [0,1]$ is called Archimedean whenever it is associative and it has no non-trivial idempotent elements, i.e., it satisfies:

i) $C(C(x_1, ..., x_n), x_{n+1}, ..., x_{2n-1}) = C(x_1, C(x_2, ..., x_{n+1}), x_{n+2}, ..., x_{2n-1}) = \cdots = C(x_1, ..., x_{n-1}, C(x_n, ..., x_{2n-1}))$ for all $x_1, \ldots, x_{2n-1} \in [0,1]$;

ii) $C(x, ..., x) < x$ for all $x \in]0,1[$.

In this paper we will deal with binary Archimedean copulas and general Archimedean copulas.

Definition 3. Let $C\colon [0,1]^2 \to [0,1]$ be a binary Archimedean copula. A mapping $C\colon \bigcup_{n\geq 2}[0,1]^n \to [0,1]$ (we use the same notation as for the binary C, as there is no confussion with this convention) given by $C(x_1,...,x_n) = C(x_1, C(x_2,..., C(x_{n-1},x_n)\,)$ is called a general Archimedean copula whenever the restriction $C|\,[0,1]^n$ is an n-copula for all $n \geq 2$.

Note that $W\colon [0,1]^2 \to [0,1]$, given by $W(x_1,x_2) = \max(0, x_1+x_2-1)$, is the weakest 2-copula. It is also Archimedean copula. However, its ternary extension, given by $W(x_1,x_2,x_3) = \max(0, x_1+x_2+x_3-2)$, is not a 3-copula and thus W is not a general Archimedean copula. On the other hand, the independence copula $\Pi\colon \bigcup_{n\geq 2}[0,1]^n \to [0,1]$, given by

$$\Pi(x_1,...x_n) = \prod_{i=1}^{n} x_i$$

is a general Archimedean copula. Due to results of Moynihan [14] based on the fact that binary copulas are 1-Lipschitz triangular norms [12], the next representation of Archimedean copulas is valid.

Theorem 2. *A function $C\colon [0,1]^2 \to [0,1]$ is a binary Archimedean copula if and only if there is a convex continuous strictly decreasing function $f\colon [0,1] \to [0,\infty]$ satisfying $f(1) = 0$ so that for all $x_1,x_2 \in [0,1]$ it holds*

$$C(x_1,x_2) = f^{-1}\big[\min(f(0), f(x_1)+f(x_2))\big]. \tag{2}$$

Function f satisfying all requirements of Theorem 2 is called an additive generator (of copula C), and the set of all additive generators of binary copulas we denote as $\mathcal{F}$. Observe that, for a given binary Archimedean copula C, the corresponding additive generator is unique up to a positive multiplicative constant. General Archimedean copulas were characterised by Kimberling [13].

Theorem 3. *A maping $C\colon \bigcup_{n\geq 2}[0,1]^n \to [0,1]$ is a general Archimedean copula if and only if there is an absolutely monotone decreasing bijection $g\colon [0,\infty] \to [0,1]$ (i.e., g has derivatives of any order $m \in N$ in each point $x \in\ ]0,\infty[$ and $\operatorname{sign} g^{(m)}(x) = (-1)^m$ for each $m \in \mathcal{N}, x \in\ ]0,\infty[$) so that for any $n \geq 2$, $x_1,...,x_n \in [0,1]$, it holds*

$$C(x_1,...,x_n) = g\left(\sum_{i=1}^{n} g^{-1}(x_i)\right). \tag{3}$$

For a general Archimedean copula C, the function g^{-1} is called a (general) additive generator, and also it is unique up to a positive multiplicative constant. The set of all absolutely monotone functions linked to general Archimedean copulas we denote as $\mathcal{G}$. Note that the weakest general Archimedean copula is the independence copula Π.

Example 1.

Neural Network World 1/10, 81–90

i) Clayton family $(C_\lambda^{Cl})_{\lambda\geq -1}$ of binary Archimedean copulas is determined by the next additive generators:

$$f_\lambda^{Cl}(x) = 1 - x^{-\lambda} \qquad \text{if } \lambda \in [-1, 0[,$$
$$f_0^{Cl}(x) = -\log x,$$
$$f_\lambda^{Cl}(x) = x^{-\lambda} - 1 \qquad \text{if } \lambda \in]0, \infty[.$$

The weakest Clayton copula $C_{-1}^{Cl} = W$ is the weakest binary copula, while the strongest (binary) copula M, $M(x_1, x_2) = \min(x_1, x_2)$, is the limit member of Clayton family, $\lim_{\lambda\to\infty} C_\lambda^{Cl}(x_1, x_2) = M(x_1, x_2)$. Moreover, $C_0^{Cl} = \Pi$ is the independence copula, and C_1^{Cl}, given by $C_1^{Cl}(x_1, x_2) = \frac{x_1 x_2}{x_1 + x_2 - x_1 x_2}$, whenever $(x_1, c_2) \notin (0, 0)$, is also called Ali-Mikhail-Haq copula. In the language of triangular norms, W is called Lukasiewicz t-norm, while C_1^{Cl} is called Hamacher product.

ii) Gumbel family $(C_\lambda^G)_{\lambda\geq 1}$ of binary Archimedean copulas is determined by the additive generators $f_\lambda^G(x) = (-\log x)^\lambda$, $\lambda \in [1, \infty[$. Note that $C_1^G = \Pi$ and $\lim_{\lambda\to\infty} C_\lambda^G = M$.

iii) Frank family $(C_p^F)_{p\in[-\infty,\infty[}$ of binary Archimedean copulas is determined by the additive generators $f_{-\infty}^F(x) = 1-x$, $f_0^F(x) = -\log x$, $f_p^F(x) = \log \frac{e^{-p}-1}{e^{-px}-1}$, $p \in]-\infty, 0[\cup]0, \infty[$. Note that $C_{-\infty}^F = W$, $C_0^F = \Pi$ and $\lim_{p\to\infty} C_p^F = M$.

Remark 1. Some of binary Archimedean copulas given in Example 1 are binary forms of general Archimedean copulas. This is the case of Clayton's copulas C_λ^{Cl} for $\lambda \geq 0$, of all Gumbel's copulas, and of Frank copulas C_p^F for $p \geq 0$. Observe that recently McNeil and Nešlehová [10] have shown that, for $n \geq 2$, the n-ary extension of $C_{\frac{-1}{n-1}}^{Cl}$ is the weakest Archimedean n-copula.

3. Construction methods for binary Archimedean copulas

Due to Theorem 3, construction methods for binary Archimedean copulas are linked to construction method for elements of the set $\mathcal{F}$ of all additive generators of binary copulas. Two methods of construction of a new additive generator from a given one are known from recent investigations.

Construction 1. *[12] Let $\varphi\colon [0,1] \to [0,1]$ be a concave increasing bijection. Then, for any additive generator $f \in \mathcal{F}$, also $f \circ \varphi\colon [0,1] \to [0,\infty]$, $(f \circ \varphi)(x) = f(\varphi(x))$, is an additive generator from $\mathcal{F}$.*

Note that some further generalisations of Construction 1 (φ need not be a bijection) can be found in Durante at al. [5, 4]. As a typical example, take $\varphi_p\colon [0,1] \to [0,1]$ given by $\varphi_p(x) = x^p$. Then φ_p is concave increasing bijection if and only if $p \in]0,1]$. Take $f_W \in \mathcal{F}$, $f_W(x) = 1 - x$. Then $(f_W \circ \varphi_p)(x) = 1 - x^p = f_{-p}^{Cl}$, i.e., Construction 1 applied to f_W has resulted to a part of Clayton's family, see Example 1i).

Construction 2. *[12, 19] Let $f \in \mathcal{F}$. Then for any $\lambda \in [1, \infty[$, also $f^\lambda \in \mathcal{F}$.*

Typical example of an application of Construction 2 is the Gumbel family. Indeed, starting from $f = f_\Pi \in \mathcal{F}$, $f(x) = -\log x$, we have $f_\lambda^G = f^\lambda$, $\lambda \in [1, \infty[$.

Surprisingly, several simple methods for constructing new additive generators are missing in the literature. We propose three such new methods.

Construction 3. *Let $\eta\colon [0, \infty] \to [0, \infty]$ be a convex increasing bijection. Then, for any $f \in \mathcal{F}$, also $\eta \circ f \in \mathcal{F}$.*

Proof. Continuity of $\eta \circ f$ follows from the continuity of both η and f. Similarly, increasingness of η and decreasingness of f ensures the decreasingness of $\eta \circ f$. Moreover, $\eta(0) = 0$ forces $(\eta \circ f)(1) = 0$. Finally, both η and f are convex, and then $(\eta \circ f)(\lambda x + (1-\lambda)y) \le \eta(\lambda f(x) + (1-\lambda)f(y)) \le \lambda\,(\eta \circ f)(x) + (1-\lambda)\,(\eta \circ f)(y)$ for all $x, y, \lambda \in [0, 1]$, where the first inequality follows from the increasingness of η and convexity of f, while the second inequality is guaranteed by the convexity of η. □

Observe that the set $\mathcal{N}$ of all convex increasing bijections $\eta\colon [0, \infty] \to [0, \infty]$ equipped with the composition of functions (as a binary operation of $\mathcal{N}$) is a cancellative semigroup with neutral element $\eta = \mathrm{id}_{[0,\infty]}$.

Example 2.

i) The next functions are elements of the set $\mathcal{N}$:

$$\eta_\lambda(x) = x^\lambda, \quad \lambda \in [1, \infty[\quad \text{(see Construction 2)};$$
$$\eta_c^{(\lambda)}(x) = (1 + cx)^\lambda - 1, \quad c \in \,]0, \infty[\,,\ \lambda \in [1, \infty[\,;$$
$$\text{composite } \left(\eta_c^{(\lambda)} \circ \eta_n\right)(x) = (1 + cx^n)^\lambda - 1.$$

ii) Take the additive generator $\varphi_\rho^{Cl} \in \mathcal{F}$, $\rho \ge -1$. Then also $\eta_c^{(\lambda)} \circ \varphi_\rho^{Cl} \in \mathcal{F}$ for any $c \in \,]0, \infty[$ and $\lambda \in [1, \infty[$. For $c = 1$ and $\rho > 0$ it holds $\left(\eta_1^{(\lambda)} \circ \varphi_\rho^{Cl}\right)(x) = x^{-\rho\lambda} - 1 = \varphi_{\rho\lambda}^{Cl}$. However, for $c = 1$, Archimedean copulas generated by $\eta_c^{(\lambda)} \circ \varphi_\rho^{Cl}$ are not known in the literature.

The set $\mathcal{F}$ is a convex set, yielding one more construction method.

Construction 4. *Let $f_1, ..., f_k \in \mathcal{F}$ be additive generators, and let $c_1, ..., c_k \in [0, 1]$, $\sum_{i=1}^k c_i = 1$. Then $f = \sum_{i=1}^k c_i f_i \in \mathcal{F}$.*

Proof. It is enough to observe that convex combinations of functions preserves continuity, decreasingness, convexity, as well as fixed value in a given point, i.e., $f(1) = \sum_{i=1}^k c_i f_i(1) = 0$. □

Example 3. Let $f_1, f_2 \in \mathcal{F}$, $f_1 = f_2$. Then the system $(f_{(c)})_{c\in[0,1]}$ of functions given by $f_{(c)} = cf_1 + (1-c)f_2$ generates a parametric family $(C_{(c)})_{c\in[0,1]}$ of binary Archimedean copulas, with extremal elements $C_{(0)}$ generated by f_1, and $C_{(1)}$ generated by f_2. Put $f_1 = f_1^{Cl}$(i.e., $f_1(x) = \frac{1}{x} - 1$), and $f_2 = f_{-1}^{Cl}$(i.e., $f_2(x) = 1 - x$).

Neural Network World 1/10, 81–90

Then $f_{(c)}^{(}x) = \frac{c+(1-2c)x-(1-c)x^2}{x}$, $c \in [0,1]$, and the corresponding parametric system $(C_{(c)})_{c\in[0,1]}$ connects the Ali-Mikhail-Haq copula C_1^{Cl} and the weakest copula W.

Remark 2.

i) Example 3 shows how we can improve known methods for fitting Archimedean copulas [8]. More details will be given in Section 5..

ii) It is well-known that (binary) copulas form a convex class. The problem when a convex combination of Archimedean copulas is an Archimedean copula was posed in [1] and it was particully solved in [16, 17]. Till now, no non-trivial positive example is known. Obviously, convex sums of copulas and convex sums of additive generators do not commute. Though this result (Construction 4) is trivial and most probably observed by many copula researchers, it seems so that its explicit formulation as given in Construction 4 appears for the first time in this paper. Similar is the case of Construction 5 which will follow.

Formula (2) involves the pseudo-inverse $f^{(-1)}\colon [0,\infty] \to [0,1]$ of an additive generator $f \in \mathcal{F}$, $f^{(-1)}(x) = f^{-1}(\min(f(0),x))$. For more details about pseudo-inverses we recommend [11]. It is not difficult to check that a function $h\colon [0,\infty] \to [0,1]$ is a pseudo-inverse of some additive generator from f if and only if h is non-increasing, strictly decreasing on $[0,a]$ where $a = \inf\{x \in [0,\infty]\,|h(x) = 0\}$, continuous, convex and $h(0) = 1$. Observe that then f is the inverse function to $h|_{[0,a]}$. Similarly as it was possible to show the convexity of the class $\mathcal{F}$, also the class of all pseudo-inverses of additive generators from $\mathcal{F}$ is a convex class. This fact proves the next result.

Construction 5. *Let $f_1, ...f_k \in \mathcal{F}$ and $c_1, ..., c_k \in [0,1]$,$\sum_{i=1}^k c_i = 1$. Then $f \in \mathcal{F}$, where $f^{(-1)} = \sum_{i=1}^k c_i f_i^{(-1)}$ (i.e., pseudo-inverse of f is a convex combination of pseudo-inverses of $f_1, ..., f_k$).*

Example 4. Similarly as in Example 3 , for any $f_1, f_2 \in \mathcal{F}$, $f_1 = f_2$, we can introduce a parametric family of binary Archimedean copulas $(C_{[c]})_{c\in[0,1]}$ included by additive generators $(f_{[c]})_{c\in[0,1]}$ such that $(f_{[c]})^{(-1)} = cf_1^{(-1)} + (1-c)f_2^{(-1)}$.

Continuing Example 3, using the same f_1 and f_2, we have $f_{[c]}^{(-1)} = \frac{c}{1+x} + (1-c)\max(0,1-x)$, i.e., for $c > 0$,

$$f_{[c]}(x) = \begin{cases} \frac{c}{x} - 1 & \text{if} \quad x \in [0,c]\,, \\ \frac{\sqrt{(2-x)^2-4c(1-x)}-x}{2(1-c)} & \text{if} \quad x \in\ ]c,1]\,. \end{cases}$$

4. Construction methods for general Archimedean copulas

In the case of general Archimedean copulas, observe first that $\mathcal{G}^{-1} = \{g^{-1}|g \in \mathcal{G}\} \subsetneqq \mathcal{F}$.

When checking construction methods given in Section 3., only Construction method 4 and 5 can be applied straightforwardly also for general Archimedean copulas. Indeed, this fact follows from the convexity of classes $\mathcal{G}$ and $\mathcal{G}^{-1}$.

Construction $1_\mathcal{G}$. *For any functions $g_1, ..., g_k \in \mathcal{G}$ and constants $c_1, ..., c_k \in [0,1]$, $\sum_{i=1}^k c_i = 1$, also $g = \sum_{i=1}^k c_i g_i \in \mathcal{G}$.*

Construction $2_\mathcal{G}$. *For any functions $g_1, ..., g_k \in \mathcal{G}$ and constants $c_1, ..., c_k \in [0,1]$, $\sum_{i=1}^k c_i = 1$, also $g = \ (\sum_{i=1}^k c_i g_i^{-1})^{-1} \in \mathcal{G}$.*

To examplify these two methods, it is enough to repeat Example 3 and 4 (and replace $W = C_{-1}^{Cl}$ by $\Pi = C_o^{Cl}$, for example).

Construction 2 was a special case of Construction 3. In the case of general Archimedean copulas, counterpart of Construction 2 is still valid, i.e., for each $g \in \mathcal{G}$ also $g_\lambda \in \mathcal{G}$, where $g_\lambda(x) = g(x^{\frac{1}{\lambda}})$ and $\lambda \geq 1$. We give now the counterpart of Construction 3, which should by modified accordingly.

Construction $3_\mathcal{G}$. *Let $\tau : [0,\infty] \to [0,\infty]$ be an increasing bijection which is antiderivative to some absolutely monotone integrable function, i.e. τ has derivatives of all orders on $]0,\infty[$ and sign $\tau^{(k)}(x) = (-1)^{k+1}$ for all $x \in]0,\infty[$ and $k \geq 1$. Then for any $g \in \mathcal{G}$, also $g \circ \tau \in \mathcal{G}$.*

Proof. It is enough to check the signs of derivatives of $g \circ \tau$. □

Example 5.

i) It is evident that $\tau(x) = x^p$ fulfils the requirements of Construction method $3_\mathcal{G}$ if and only if $p \in]0,1]$, what yields the counterpart of Construction method 2 for general Archimedean copulas.

ii) For $c > 0$, $\lambda \geq 1$, let $\tau(x) = \frac{(1+x)^{\frac{1}{\lambda}}-1}{c}$. Then τ satisfies requirements of Construction method $3_\mathcal{G}$, and it is a counterpart of Example 2i).

iii) Define $\tau(x) = e^x - 1$. Then τ satisfies reguirements of Construction method $3_\mathcal{G}$. For independence copula Π with additive generator $\varphi_\Pi(x) = -\log x$, $(\tau \circ \varphi_\Pi)(x) = \frac{1}{x} - 1$, i.e., we have obtained additive generator of Ali-Mikhail-Haq copula C_1^{Cl}.

iv) Define $\tau(x) = \min(x, \frac{x+1}{2})$. Then its inverse is given by $\tau^{-1}(x) = \max(x, 2x-1)$ and it satisfies requirements of Construction method 3. Thus, for any $f \in \mathcal{F}$, $\tau^{-1} \circ f \in \mathcal{F}$. However, τ does not fit Construction $3_\mathcal{G}$, i.e., it cannot be applied to construction of additive generators of general Archimedean copulas. Indeed, take $f_1^{Cl} \in \mathcal{F}$, $f_1^{Cl}(x) = \frac{1}{x} - 1$.
Then $g = (f_1^{Cl})^{-1} \in \mathcal{G}$. Define $h = g \circ \tau$,

$$h(x) = \begin{cases} \frac{1}{1} + x & \text{if} \quad x \in [0,1]\,, \\ \frac{2}{3} + x & \text{if} \quad x \in [1,\infty]\,, \end{cases}$$

Then $h \notin \mathcal{G}$. To see this fact observe that $f = h^{-1} \in \mathcal{F}$,

$$f(x) = \begin{cases} \frac{2}{x} - 3 & \text{if} \quad x \in \left[0, \frac{1}{2}\right[\,, \\ \frac{1}{x} - 1 & \text{if} \quad x \in \left[\frac{1}{2}, 1\right]\,, \end{cases}$$

Neural Network World 1/10, 81–90

is an additive generator of a binary copula but not of a ternary copula. Take $\mathbf{x} = (\frac{5}{7}, \frac{5}{7}, \frac{5}{7})$ and $\mathbf{y} = (\frac{6}{7}, \frac{6}{7}, \frac{6}{7})$.
The left-hand side of the inequality (1) equals $-\frac{1}{56}$, it is negative, i.e., the corresponding ternary function C is not 3-increasing and thus not a copula.

The next method is linked to Construction 1.

Construction 4$_{\mathcal{G}}$. *Let $\psi\colon [0,1] \to [0,1]$ be an absolutely monotone bijection (i.e., all derivatives of ψ on $]0,1[$ exist and they are non-negative). Then for any $g \in \mathcal{G}$, also $\psi \circ g \in \mathcal{G}$.*

Proof. Again it is enough to check the signs of derivatives of $\varphi \circ g$ □

Remark 3.

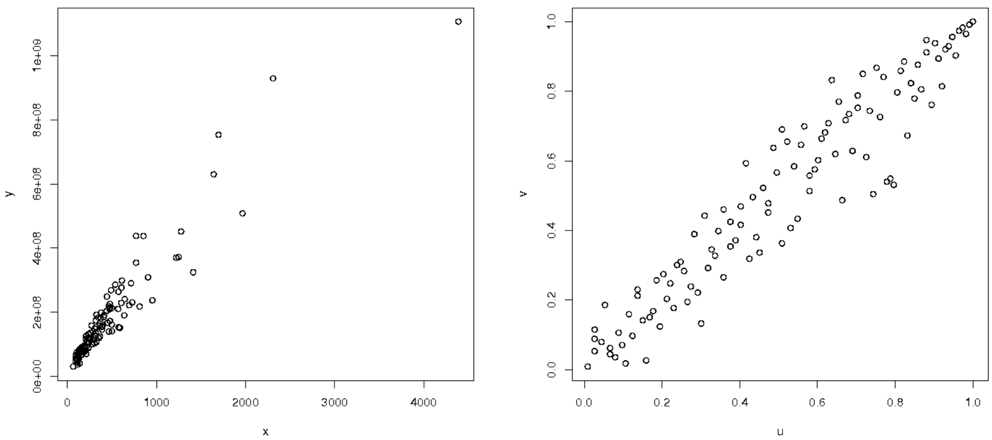
i) Put $\Psi(x) = x^{\frac{3}{2}}$. Then ψ does not satisfy the requirements of Construction $4_{\mathcal{G}}$. However, for all g linked to general Archimedean copulas introduced in Example 2 also $g^{\frac{3}{2}} \in \mathcal{G}$. An open problem arises. Are requirements on ψ in Construction $4_{\mathcal{G}}$ also necessary? In particular, is there any $g \in \mathcal{G}$ so that $g^{\frac{3}{2}} \notin \mathcal{G}$?

ii) Any ψ satisfying requirements of Construcion $4_{\mathcal{G}}$ is a (possibly infinite) convex sum of power functions, $\psi(x) = \sum_{n=1}^{\infty} a_n x^n$, with $a_n \geq 0$, $\sum_{n=1}^{\infty} a_n = 1$. Then $\psi \circ g = \sum_{n=1}^{\infty} a_n g^n$, i.e., Construction $4_{\mathcal{G}}$ can be seen as a corollary of Construction $1_{\mathcal{G}}$ because of compactness of the class of general copulas, and of the fact that for each $g \in \mathcal{G}$, $n \in N$, also $g^n \in \mathcal{G}$ (as a canequence of Construction 4_G.

5. Application to the fitting of copulas

For demonstrating the experiment we chose data of anual river flow peaks and their corresponding volumes of flood waves [21], see Fig. 1.

Fig. 1 shows $n = 114$ pairs of data and its ranks divided by $(n + 1)$, i.e., no assumptions about marginal distributions was made. Three Archimedean families was chosen for estimation and construction of pairwise convex combinations, particularly the so-called Clayton, Gumbel-Hougaard and Frank families. First we estimated parameters for each of the three copulas, then - using the estimates - we estimated weighting parameter (of convex combination) for every possible couple of the above three Archimedean generators and their inverses, separately. Estimation was performed by means of maximum pseudo-likelihood method [7] and optimisation routines in R. Subsequently we checked goodnes of fit (GoF) by one of the "blanket" tests [9] that measure distance between tested parametric copula and so-called empirical copula. P-value of the GoF test had to be simulated by a bootstrap method (with 5000 replications), which is computationaly the most time-consuming procedure to do, otherwise all routines performs well even for larger sets of data. The results are sumarized in Tab. 1, where we put the parameters estimates and simulations of corresponding p-values of GoF test. The improvement in performance is visible from the higher p-values since null hypothesis assumes inadequacy of a model.



a b

Fig. 1 *Scatter plot of flow peaks (x,u) to flood volumes (y,v) before (a) and after (b) probability integral transformation.*

Tab. I *Best copula parameters with simulated p-values of the GoF test*

Copula	Parameters	GoF test
Gumbel	4.728	0.4432
Clayton	4.143	0.0125
Frank	18.385	0.3196
Generator		
Clayton – Gumbel	0	0.4428
Clayton – Frank	1	0.0130
Frank – Gumbel	0.113	0.4915
Inverse of generator		
Clayton – Gumbel	0	0.4399
Clayton – Frank	1	0.0118
Frank – Gumbel	0.128	0.4920

Neural Network World 1/10, 81–90

All computations were performed in R, the open-source environment for statistical computing and visualisation, the algorithms will be soon and freely available on the author webpage `www.math.sk/bacigal` under the section "Research".

Acknowledgement

The research was supported by APVV LPP - 0004-07, VEGA 1/0496/08, APVV-0443-07.

References

[1] Alsina L., Frank M.J., Schweizer B.: Problems on associative functions. Aequationes Math. 66, 2003, pp. 128–140.

[2] Bárdossy A. (2006): Copula-based geostatistical models for groundwater quality parameters, Water Resour. Res., 42, W11416, doi:10.1029/2005WR004754.

[3] Couceiro M., Marichal J.-L.: Associative n-ary lattice polynomial. Submitted.

[4] Durante F., Foschi R., Sarkoci P.: Distorted copulas: constructions and tail dependence, Comm. Statist. Theory and Method, in press (2010).

[5] Durante F., Sempi C.: Copula and semicopula transforms. Int. J. Math. Sci., 2005, pp. 645–655.

[6] Embrechts P., McNeil A.J., and Straumann D.(2001): Correlation and Dependency in Risk Management: Properties and Pitfalls, Cambridge Univ. Press, New York.

[7] Genest C., Ghoudi K., and Rivest L.-P., (1995): A semiparametric estimation procedure of dependence parameters in multivariate families of distribution. Biometrika, 40, pp. 543-552.

[8] Genest C., Favre A.-C.: Everything you always wanted to know about copula modeling but were afraid to ask. Journal of Hydrologic Engineering. 12(4), 2007, pp. 347–368.

[9] Genest C., Rémillard B. , and Beaudoin D.(2008): Omnibus goodnessoffit tests for copulas: A review and a power study. Insurance: Mathematics and Economics 42.

[10] McNeil A.J., Nešlehová J.: Multivariate Archimedean Copulas, D-monotone functions and L_1-norm symmetric distributions. Ann. Statist. Volume 37, Number 5B (2009) 3059–3097.

[11] Klement E.P., Mesiar R., and Pap E.: Quasi- and pseudo- inverses of monotone functions, and the construction of t-norms. Fuzzy Sets and Systems 104, 1999, pp. 3–13.

[12] Klement E.P., Mesiar R., and Pap E.: Triangular Norms. Kluwer Academic Publishers,Dordrecht 2000.

[13] Kimberling C.H.: On a class of associative functions. Publ. Math. Debrecen 20, 1973, pp. 21–39.

[14] Moynihan R.: On τ_T-semigroups of probability distributions. Aequationes Math. 17, 1978, pp. 19–40.

[15] Nelsen R.B.: An Introduction to Copulas. Springer, New York 2006.

[16] Ouyang Y., Fang J.X.: Some observations about the convex combination of continuous triangular norms. Nonlinear Anal. 68, 2008, pp. 3382–3387.

[17] Petrík M., Sarkoci P.: Convex combinations of nilpotent triangular norms. J. Math. Anal. Appl. 350, 2009, pp. 271–275.

[18] Salvadori G., De Michele C., Kottegoda N.T., and Rosso R.: Extremes in Nature. An Approach Using Copulas. (WTS Library Series, Vol. 56.) SpringerVerlag, Berlin 2007.

[19] Schweizer B., Sklar A.: Probabilistic Metric Spaces, Nord-Holland, 1983.

[20] Sklar A.: Fonctions de repartition 'a n dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris 8, 1959, pp. 229-231.

[21] Szolgay J., Šurek P.: Calibration of a Multilinear Transformation Model with a Hypothetical Flood Wave by a Genetic Algorithm. Acta Hydrologica Slovaca, Vol. 9, 2008, No. 2, pp. 236–242.

KYBERNETIKA — VOLUME 47 (2011), NUMBER 4, PAGES 519–531

NON-EXCHANGEABLE RANDOM VARIABLES, ARCHIMAX COPULAS AND THEIR FITTING TO REAL DATA

Tomáš Bacigál, Vladimír Jágr and Radko Mesiar

The aim of this paper is to open a new way of modelling non-exchangeable random variables with a class of Archimax copulas. We investigate a connection between powers of generators and dependence functions, and propose some construction methods for dependence functions. Application to different hydrological data is given.

Keywords: Archimax copula, dependence function, generator

Classification: 93E12, 62A10

1. INTRODUCTION

In recent years copulas turned out to be a promising tool in multivariate modelling, mostly with applications in actuarial sciences and hydrology.

In short, copula is a function which allows modelling dependence structure between stochastic variables. The main advantage is that the copula approach can split the problem of constructing multivariate probability distributions into a part containing the marginal one-dimensional distribution functions and a part containing the dependence structure. These two parts can be studied and estimated separately and then rejoined to form a joint distribution function.

Restricting ourselves to bivariate case, copula is a function $C\colon [0,1]^2 \to [0,1]$ which satisfies the boundary conditions, $C(t,0) = C(0,t) = 0$ and $C(t,1) = C(1,t) = t$ for $t \in [0,1]$ (uniform margins), and the 2-increasing property, $C(u_2,v_2)-C(u_2,v_1)-C(u_1,v_2)+C(u_1,v_1) \geq 0$ for all $u_1 \leq u_2, v_1 \leq v_2 \in [0,1]$. Copula is symmetric if $C(u,v) = C(v,u)$ for all $(u,v) \in [0,1]^2$ and is asymmetric otherwise. By $[a,b]$ we mean a closed interval with endpoints a and b, while $]a,b[$ will denote an open interval.

There are several approaches how to model exchangeable random variables. Most of them refer to Archimedean copulas [15], i. e., copulas $C_\varphi\colon [0,1]^2 \to [0,1]$ expressible in the form

$$C_\varphi(u,v) = \varphi^{(-1)}\left(\varphi(u) + \varphi(v)\right), \tag{1}$$

where $\varphi\colon [0,1] \to [0,\infty]$ is a continuous strictly decreasing convex function satisfying

$\varphi(1) = 0$, and its pseudo-inverse $\varphi^{(-1)}\colon [0,\infty] \to [0,1]$ is given by

$$\varphi^{(-1)}(x) = \varphi^{-1}\left(\min(\varphi(0), x)\right). \tag{2}$$

Among several approaches allowing to fit copulas to real data we recall [5] and references therein.

The aim of this paper is to open a new way of modelling non-exchangeable random variables which are related to asymmetric copulas. In the next section we recall Archimax copulas, a special class of copulas which are non-symmetric, in general. After some new theoretical results about the structure of Archimax copulas, in Section 3 we propose new construction methods for one part of this class of copulas. Section 4 gives a short overview of estimation methods used in the application to modelling hydrological data in Section 5.

2. ARCHIMAX COPULAS

Since there are much more cases in the nature when we feel the causality among stochastic processes flows in certain direction rather than the cases when we observe random variables equally affected by common underlying process, we find symmetry of the most used copulas quite restrictive. Among few classes of asymmetric copulas, convenient enough to model non-exchangeable random variables, we focus on the class of Archimax copulas [2] built up from a convex continuous decreasing function $\varphi\colon [0,1] \to [0,\infty]$, $\varphi(1) = 0$, called generator and a convex function $A\colon [0,1] \to [0,1]$, $\max(t, 1-t) \le A(t) \le 1$ for all $t \in [0,1]$, called dependence function. Then the corresponding Archimax copula is given by

$$C_{\varphi,A}(u,v) = \varphi^{(-1)}\left[(\varphi(u) + \varphi(v))\, A\left(\frac{\varphi(u)}{\varphi(u) + \varphi(v)}\right)\right] \quad \text{for all } u, v \in [0,1] \tag{3}$$

(with conventions $0/0 = \infty/\infty = 0$, where $\varphi^{(-1)}$ is given by (2)). Observe that Archimax copulas contains as special subclasses all Archimedean copulas (then $A \equiv A^* = 1$) and all extreme value copulas [16], in short EV copulas (then $\varphi(t) = -\log(t)$). For the weakest dependence function $A = A_*$,

$$A_*(t) = \max(t, 1-t),$$

we have $C_{\varphi,A_*} = M$, the strongest copula of co-monotone dependence, independently of the generator φ.

Moreover, it is easy to check that an Archimax copula $C_{\varphi,A}$ is symmetric if and only if $A(t) = A(1-t)$ for all $t \in [0,1]$ (i. e., A is symmetric wrt. axis $x = 1/2$). Recent results on measuring asymmetry can be found in [4].

Suppose that φ is a generator of a copula C_φ. Then also φ^λ, $\lambda > 1$, is a generator of a copula C_{φ^λ}. As an example recall the Gumbel family of copulas $\left(C^G_{(\lambda)}\right)_{\lambda \in [1,\infty[}$, generated by generators $\varphi^G_{(\lambda)}\colon [0,1] \to [0,\infty]$, $\varphi^G_{(\lambda)}(x) = (-\log x)^\lambda$, which bears from the product copula Π generated by $\varphi^G_{(1)}$.

Proposition 2.1. Let $\varphi\colon [0,1] \to [0,\infty]$ be a generator of a copula C_φ. For any dependence function A, and any $\lambda \ge 1$, the Archimax copula $C_{\varphi^\lambda,A}$ is also an Archimax copula based on generator φ, i. e., $C_{\varphi^\lambda,A} = C_{\varphi,B_{(A,\lambda)}}$, where $B_{(A,\lambda)} : [0,1] \to [0,1]$ is a dependence function given by

$$B_{(A,\lambda)} = A_{(\lambda)}(t)\left[A\left(\left(\frac{t}{A_{(\lambda)}(t)}\right)^{\lambda}\right)\right]^{1/\lambda}, \tag{4}$$

with $A_{(\lambda)} :\ [0,1] \to [0,1]$, $A_{(\lambda)}(t) = \left(t^\lambda + (1-t)^\lambda\right)^{1/\lambda}$.

Proof. Formula (4) is a matter of processing of the equality $C_{\varphi^\lambda,A} = C_{\varphi,B_{(A,\lambda)}}$. Indeed,

$$\begin{aligned} C_{\varphi^\lambda,A}(u,v) &= \varphi^{(-1)}\left(\left[(\varphi^\lambda(u)+\varphi^\lambda(v))A\left(\frac{\phi^\lambda(u)}{\phi^\lambda(u)+\phi^\lambda(v)}\right)\right]^{1/\lambda}\right)\\ &= \varphi^{(-1)}\left((\varphi(u)+\varphi(v))\left[\frac{\phi^\lambda(u)+\phi^\lambda(v)}{(\phi(u)+\phi(v))^\lambda}A\left(\frac{\phi^\lambda(u)}{\phi^\lambda(u)+\phi^\lambda(v)}\right)\right]^{1/\lambda}\right), \end{aligned}$$

while

$$\begin{aligned} &C_{\varphi^\lambda,B(A,\lambda)}(u,v)\\ &\quad= \varphi^{(-1)}\left((\varphi(u)+\varphi(v))\left[\left(\frac{\phi(u)}{\phi(u)+\phi(v)}\right)^\lambda + \left(\frac{\phi(v)}{\phi(u)+\phi(v)}\right)^\lambda\right]^{1/\lambda}\right.\\ &\qquad\qquad \left. A\left(\frac{\left(\frac{\phi(u)}{\phi(u)+\phi(v)}\right)^\lambda}{\left(\frac{\phi(u)}{\phi(u)+\phi(v)}\right)^\lambda + \left(\frac{\phi(v)}{\phi(u)+\phi(v)}\right)^\lambda}\right)\right)\\ &\quad= \varphi^{(-1)}\left((\varphi(u)+\varphi(v))\left[\frac{\phi^\lambda(u)+\phi^\lambda(v)}{(\phi(u)+\phi(v))^\lambda}A\left(\frac{\phi^\lambda(u)}{\phi^\lambda(u)+\phi^\lambda(v)}\right)\right]^{1/\lambda}\right). \end{aligned}$$

To see that $B_{(A,\lambda)}$ is indeed a dependence function, note that based on Gumbel family, we have also $C_{\varphi^G_{(\lambda)},A} = C_{\varphi^G_{(1)},B_{(A,\lambda)}}$. Due to [2], $C_{\varphi^G_{(\lambda)},A}$ is a copula. Moreover, for any power $p \in]0,\infty[$,

$$\begin{aligned} &C_{\varphi^G_{(\lambda)},A}(u^p,v^p)\\ &= \exp\left(-\left[\left((-\log u^p)^\lambda + (-\log v^p)^\lambda\right)A\left(\frac{(-\log u^p)^\lambda}{(-\log u^p)^\lambda+(-\log v^p)^\lambda}\right)\right]^{1/\lambda}\right)\\ &= \exp\left(-p\left[\left((-\log u)^\lambda + (-\log v)^\lambda\right)A\left(\frac{(-\log u)^\lambda}{(-\log u)^\lambda+(-\log v)^\lambda}\right)\right]^{1/\lambda}\right)\\ &= \left(C_{\varphi^G_{(\lambda)},A}(u,v)\right)^p, \end{aligned}$$

i. e., $C_{\varphi^G_{(\lambda)},A}$ is an EV-copula [15, 16]. Due to representation of EV-copulas, there is a dependence function $B:\ [0,1] \to [0,1]$ such that

$$C_{\varphi^G_{(\lambda)},A} = C_{\varphi^G_{(1)},B}$$

and evidently $B = B_{A,\lambda}$. □

Dependence functions $A_{(\lambda)}$, $\lambda \in [0,1]$, are called Gumbel dependence functions due to the fact that $C^G_{(\lambda)} = C_{\varphi^G_{(1)},A_{(\lambda)}}$. Observe that the Archimedean copula C_{φ^λ} is just an Archimax copula based on φ and $A_{(\lambda)}$, $C_{\varphi^\lambda} = C_{\varphi,A_{(\lambda)}}$, independently of the generator φ. Proposition 2.1 has an important impact for the structure of Archimax copulas. For any generator $\varphi:\ [0,1] \to [0,\infty]$, classes $\mathcal{A}_{\varphi^\lambda}$ of Archimax copulas based on generators φ^λ, $\lambda \in [1,\infty[$, are nested, and $\mathcal{A}_{\varphi^\lambda} \subsetneqq \mathcal{A}_{\varphi^\mu}$ whenever $1 \le \mu < \lambda \le \infty$, where $\mathcal{A}_{\varphi^\infty} = \bigcap_{\lambda=1}^{\infty} \mathcal{A}_{\varphi^\lambda} = \{M\}$. Therefore it is important to know the basic form η of each generator φ, $\varphi = \eta^\lambda$ with $\lambda \ge 1$, where $\eta:\ [0,1] \to [0,\infty]$ is a generator such that for any $\lambda \in]0,1[$, η^λ is no more convex. Such generators η will be called basic generators and they correspond to Archimedean copulas C_η such that for any $p > 1$, the corresponding L_p-norm $\|C_\eta\|_p > 1$ (for more details we recommend [3, 14]).

Proposition 2.2. Let $\varphi:\ [0,1] \to [0,\infty]$ be a generator. Let

$$\alpha = \inf\Big\{\frac{\varphi(x)\varphi''(x)}{(\varphi'(x))^2} \mid x \in]0,1[\text{and } \varphi'(x), \varphi''(x) \text{ exist}\Big\}.$$

Then

$$\eta = \varphi^{1/p}, \quad \text{where } p = \frac{1}{1-\alpha},$$

is a basic generator.

P r o o f . The convexity of a generator $\varphi(\eta)$ is equivalent to the non-negativity of the derivatives $\varphi''(\eta'')$ in all points where they exist. Let $\varphi = \eta^p$, $p \ge 1$, where η is a basic generator. Then $\eta = \varphi^{1/p}$, $\eta' = \frac{1}{p}\varphi^{1/p-1}\varphi'$ and

$$\begin{aligned}\eta'' &= \frac{1}{p}\left(\frac{1}{p}-1\right)\varphi^{\frac{1}{p}-2}(\varphi')^2 + \frac{1}{p}\varphi^{\frac{1}{p}-1}\varphi'' \\ &= \frac{1}{p}\varphi^{\frac{1}{p}-2}\left(\left(\frac{1}{p}-1\right)(\varphi')^2 + \varphi'\varphi''\right) \ge 0\end{aligned}$$

if and only if $\alpha \le \varphi\varphi''/(\varphi')^2$, where $\alpha = 1 - 1/p$, where the last inequality should be satisfied in each point from $]0,1[$ where φ' and φ'' exist. The result follows. □

Based on Propositions 2.1 and 2.2, we propose to fit Archimax copulas based on basic generators η only. Thus before choosing the appropriate candidates for fitting of a generator, one should check their basic forms. The next lemma gives a sufficient condition for a generator η to be basic.

Lemma 2.3. Let $\eta: [0,1] \to [0,\infty]$ be a generator and let $\eta'(1^-) \neq 0$. Then η is a basic generator.

Proof. Due to continuity of η and $\eta(1) = 0$, if $\eta'(1^-) \neq 0$ then $\alpha = \inf \left\{ \frac{\eta(x)\eta''(x)}{(\eta'(x))^2} \mid x \in]0,1[\text{ and } \eta'(x), \eta''(x) \text{ exist} \right\} = 0$ and thus $p = 1$. □

Example 2.4.

(i) For each Gumbel generator $\varphi^G_{(\lambda)}$, the corresponding basic generator is $\eta = \varphi^G_{(1)}$ (the generator of the product copula).

(ii) The weakest copula $C^{(p)}$ which has minimal L_p-norm, $\|C^{(p)}\|_p = 1$, $p \in [1,\infty[$, is an Archimedean copula generated by a generator $\varphi^Y_{(p)} : [0,1] \to [0,\infty]$, $\varphi^Y_{(p)}(x) = (1-x)^p$ (Y stands for Yager family, see [18], more details on L_p-norms and copulas can be found in [3]). Again, for any $p \in [1,\infty[$, the corresponding basic generator $\eta = \varphi^Y_{(1)}$ is unique (generator of the lower Frechet-Hoeffding bound W).

(iii) Based on Lemma 2.3 one can quickly check that the families of Clayton, Frank, Ali–Mikhail–Haq (see [10, 15]), are generated by basic generators only.

(iv) Taking generators from some two-parameter families given in [10], one may easily verify that

— BB1 generator $\varphi(t) = (t^{-a} - 1)^b$ with $a > 0, b \geq 1$ gains its basic form only for $b = 1$, while BB3 with $\varphi(t) = e^{b(-\log t)^a} - 1$ and $a \geq 1, b > 0$ only for $a = 1$. Then, both would result in strict Clayton copula;

— BB2 generator $\varphi(t) = e^{b(t^{-a}-1)} - 1$ with $a > 0, b > 0$ is basic for any a, b;

— BB6 generator $\varphi(t) = \left[-\log(1-(1-t)^a)\right]^b$ with $a \geq 1, b > 0$ reduces to basic form if $b = 1/a$;

— each BB7 generator $\varphi(t) = \left[\left(1-(1-t)^a\right)^{-b} - 1\right]^{1/a}$ with $a \geq 1, b > 0$ is a basic generator.

3. SOME CONSTRUCTION METHODS FOR DEPENDENCE FUNCTIONS

Based on some known dependence functions, it is desirable to be able to construct new dependence functions to increase the fitting potential of our Archimax copulas buffer. Recall that for dependence functions $A_1, \ldots, A_n$ also their convex sum $A = \sum_{i=1}^n \lambda_i A_i$ (with $\sum_{i=1}^n \lambda_i = 1$) is a dependence function. Inspired by the bivariate construction [11] and based on the recent results [13], consider dependence functions $A_1, \ldots, A_n$. Then the corresponding EV copulas $C_{A_1}, \ldots, C_{A_n} : [0,1]^2 \to [0,1]$ are given by

$$C_{A_i}(u,v) = \exp\left((\log u + \log v) A_i \left(\frac{\log u}{\log u + \log v} \right) \right). \tag{5}$$

Take arbitrary two probability vectors $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in [0,1]^n$, $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i = 1$. Then due to [13] the function $C\colon\ [0,1]^2 \to [0,1]$ given by

$$C(u,v) = \prod_{i=1}^{n} C_i\left(u^{a_i}, v^{b_i}\right) \tag{6}$$

is also a copula. Note that EV copulas are characterized by the power stability $C\left(u^\lambda, v^\lambda\right) = (C(u,v))^\lambda$ for any $\lambda \in]0,\infty[$, $u,v \in [0,1]$. It is then easy to see that C given by (6) is also an EV copula, and thus there is a dependence function A so that $C = C_A$. For processing purposes, denote $t = \frac{\log u}{\log u + \log v}$. Then $\log v = \frac{1-t}{t}\log u$ and $\log u + \log v = \frac{\log u}{t}$. Moreover,

$$C(u,v) = \exp\left((\log u + \log v) A\left(\frac{\log u}{\log u + \log v}\right)\right) = \exp\left(\frac{\log u}{t} A(t)\right). \tag{7}$$

On the other hand, due to (6),

$$\begin{aligned} C(u,v) &= \prod_{i=1}^{n} \exp\left(\left(a_i \log u + b_i \frac{1-t}{t}\log u\right) A_i\left(\frac{a_i \log u}{a_i \log u + b_i \frac{1-t}{t}\log u}\right)\right) \\ &= \exp\left(\frac{\log u}{t} \sum_{i=1}^{n} \left(ta_i + (1-t)b_i\right) A_i\left(\frac{ta_i}{ta_i + (1-t)b_i}\right)\right). \end{aligned} \tag{8}$$

Comparing (7) and (8), we see that

$$A(t) = \sum_{i=1}^{n} \left(ta_i + (1-t)b_i\right) A_i\left(\frac{ta_i}{ta_i + (1-t)b_i}\right). \tag{9}$$

What was just shown is the following construction method.

Proposition 3.1. Let $A_1, \ldots, A_n$ be dependence functions. Then for any probability vectors $(a_1, \ldots, a_n), (b_1, \ldots, b_n) \in [0,1]^n$, also the function $A\colon\ [0,1] \to [0,1]$ given by (9) is a dependence function.

Observe that the formula (9) can be deduced by induction from the original formula given in [11], as well as seen as extension of Proposition 3 of [9] dealing with A_1, A_2 and $\alpha, \beta \in [0,1]$. Then the function $A\colon\ [0,1]^2 \to [0,1]$ given by

$$\begin{aligned} A(t) &= \left(\alpha t + \beta(1-t)\right) A_1\left(\frac{\alpha t}{\alpha t + \beta(1-t)}\right) \\ &+ \left((1-\alpha)t + (1-\beta)(1-t)\right) A_2\left(\frac{(1-\alpha)t}{(1-\alpha)t + (1-\beta)(1-t)}\right) \end{aligned}$$

is a dependence function. Moreover, if $(a_1, \ldots, a_n) = (b_1, \ldots, b_n)$, then the formula (9) turns into the standard convex sum $A(t) = \sum_{i=1}^n a_i A_i(t)$. Evidently, this method allows to introduce asymmetric Archimax copulas even if starting from symmetric Archimax copulas.

Example 3.2. Consider dependence functions A_1, A_2. Let $A_2 = A_{(2)}$, see (4), $a_1 = a_2 = 1/2$, $b_1 = 0, b_2 = 1$. Then the dependence function A given by (9) does not depend on A_1, and it holds

$$A(t) = \frac{t}{2} + \frac{(2-t)}{2} A_2\left(\frac{t}{2-t}\right) = \frac{t}{2} + \sqrt{\left(\frac{t}{2}\right)^2 + (1-t)^2}.$$

Observe that $A(1/3) = (1+\sqrt{17})/6 = 0.85385$ and $A(2/3) = (1+\sqrt{2})/3 = 0.80474$, proving the asymmetry of any relevant Archimax copula $C_{\phi,A}$ (recall that $C_{\phi,A}$ is symmetric if and only if $A(t) = A(1-t)$ for all $t \in [0,1]$).

Inspired by [1] where construction methods for generators of Archimedean copulas were discussed, we propose one more new construction method for dependence function. For a dependence function A, denote by B a $[0,1] \to [0,1]$ function given by $B(t) = A(t) - 1 + t$. Each such B is characterized by its convexity, non-decreasingness and boundary conditions

$$\max(0, 2t-1) \le B(t) \le t.$$

The pseudo-inverse $B^{(-1)}\colon\ [0,1] \to [0,1]$ of B is given by

$$B^{(-1)}(u) = \sup\{t \in [0,1] \,|\, B(t) \le u\},$$

and it is characterized by concavity, non-decreasingness and boundary conditions

$$u \le B^{(-1)}(u) \le \frac{u+1}{2}. \tag{10}$$

Consider dependence functions $A_1, \dots, A_n$ and related functions $B_1^{(-1)}, \dots, B_n^{(-1)}$. Then the convex combination $\sum_{i=1}^n \lambda_i B_i^{(-1)}$ is concave, non-decreasing and satisfy the boundary conditions (10), and thus there is a dependence function A such that its related function $B^{(-1)}$ is just equal to $\sum_{i=1}^n \lambda_i B_i^{(-1)}$. This fact proves our next construction method.

Proposition 3.3. Let $A_1, \dots, A_n$ be dependence functions and let $(\lambda_1, \dots, \lambda_n) \in [0,1]^n$ be a probability vector. Then the function $A\colon\ [0,1] \to [0,1]$ given by

$$A(t) = \left(\sum_{i=1}^n \lambda_i B_i^{(-1)}\right)^{(-1)} (t) + 1 - t \tag{11}$$

is a dependence function.

Example 3.4. Consider the extremal dependence functions $A_1 = A_*$ and $A_2 = A^* = 1$. Then $B_1(t) = \max(0, 2t-1) = B_*(t)$ and $B_2(t) = t = B^*(t)$. Moreover, $B_1^{(-1)}(u) = \frac{u+1}{2}$ and $B_2^{(-1)}(u) = u$. For a fixed $\lambda \in [0,1]$, $\big(\lambda B_1^{(-1)} + (1-\lambda) B_2^{(-1)}\big)(u) = \big(1 - \frac{\lambda}{2}\big)(u) + \frac{\lambda}{2} = B_\lambda^{(-1)}(u)$, and thus $B_\lambda(t) = \big(B_\lambda^{(-1)}\big)^{(-1)}(t) = \max\big(0, \frac{2t-\lambda}{2-\lambda}\big)$ and $A_\lambda(t) = B_\lambda(t) + 1 - t = \max\big(1-t, \frac{2-2\lambda+\lambda t}{2-\lambda}\big)$. Note that both A_1 and A_2 are symmetric (wrt. axis $x = 1/2$), but not A_λ for any $\lambda \in \,]0,1[$.

4. ESTIMATION METHOD

Basically there are two main methods recently used for estimating one-parameter families. One uses various measures of dependence, such as Kendall's tau through formal relation with copula parameter θ, the another is based on maximization of a likelihood function [8]. For general multi-parameter copulas (not e. g. multivariate normal or pair-copulas) the first method is problematic and to our best knowledge no satisfactory study has been presented so far. Given a sample of n-dimensional random vectors $\boldsymbol{X}_1, \dots, \boldsymbol{X}_m$, here we use pseudo-loglikelihood function

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{m} \log \big(c_{\boldsymbol{\theta}}\big(F_1(X_{1,i}), \dots, F_n(X_{n,i})\big)\big)$$

employing copula density $c_{\boldsymbol{\theta}}$ (which is n-order mixed derivative with respect to all variables) with vector parameter $\boldsymbol{\theta}$ and marginal empirical distribution functions

$$F_j(x) = \frac{1}{m+1} \sum_{i=1}^{m} 1(X_{i,j} \leq x), \quad j = 1, \dots, n,$$

where $1(\cdot)$ is the indicator function which yields 1 whenever $\cdot$ is true and 0 otherwise. The marginal empirical distribution functions transform $\boldsymbol{X}_i$ into pseudo-observations $\boldsymbol{U}_i$, $i = 1, \dots, m$. Goodness of fit can be checked by comparing (L_2-norm) squared distances

$$S_n = \sum_{i=1}^{m} \left(C_n(U_{i,1}, \dots, U_{i,n}) - C_{\boldsymbol{\theta}}(U_{i,1}, \dots, U_{i,n})\right)^2$$

between estimated parametric copulas $C_{\boldsymbol{\theta}}$ and empirical copula function

$$C_n(u_1, \dots, u_n) = \frac{1}{m} \sum_{i=1}^{m} 1(U_{i,1} \leq u_1, \dots, U_{i,n} \leq u_n), \quad (u_1, \dots, u_n) \in [0,1]^n .$$

However because of computational intensity of simulation related to bootstrap method, we do not perform GOF test [7] here (unless one would be interested in classifying copulas according to their competence to describe particular data). Instead, for comparison purposes we employ model selection criterion, in particular the Bayesian information criterion defined

$$BIC = -2L(\boldsymbol{\theta}) + k \log(m)$$

where k denotes number of parameters. The model preference grows with decreasing BIC.

5. APPLICATION

To examine performance of new models we consider two kinds of bivariate ($n = 2$) hydrological data. One is constituted by monthly average flow rate of two rivers

– Danube at Nagymaros (Hungary) $\{X_{i,1}\}$, $i = 1, \dots m$, and Inn measured at Schärding (Austria) $\{X_{i,2}\}$ (Inn is tributary to Danube, Nagymaros lies about 570 km downstream) comprising $m = 660$ realisations recorded for 55 years until 1991, see [17]. Another sequence of $m = 113$ entries comes from annual summer term maxima of the Vltava river (Bohemia) flow rate $\{X_{i,1}\}$ (measured above the dam Kamyk until 2007) with corresponding flood volume $\{X_{i,2}\}$, which is total amount of water run within 8 days starting three days before the corresponding flow rate peak.

Due to temporal manner of the monthly river discharge, the data were found not being i.i.d. After filtering the lowest frequencies, seasonal component and applying AR(1) model, the residuals were tested by Ljung-Box test and test of serial independence (based on copulas and proposed by [6]) with positive result.

Both bivariate data transformed to unit square are shown in Figure 1.

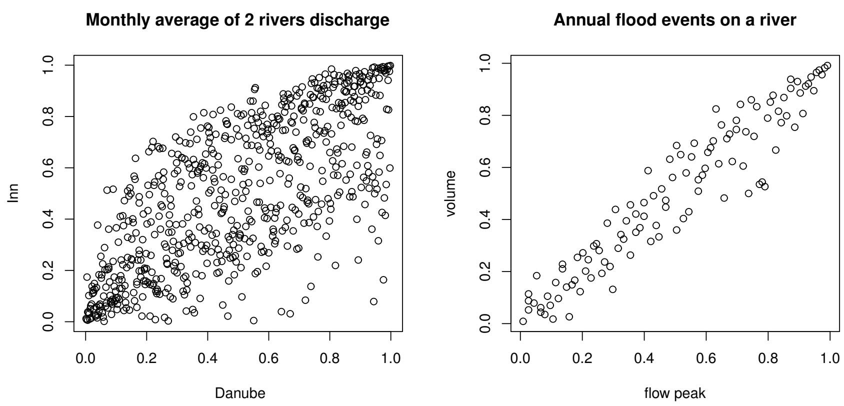


Fig. 1. Scatter plot of data after transformation by its marginal empirical distribution function.

Tables 2 and 3 summarize competition of new construction methods alongside well-established models (for overview see Table 1, [10, 15]) and related construction methods [1]. Besides parameters and maximized value of log-likelihood function we provide the corresponding estimation time and BIC criterion. Parameters were found by box-constrained optimisation (method L-BFGS-B implemented in R) which, if failed to find global maximum, was helped by pre-search over a grid. This happens mostly with more-parameter piece-wise construction of dependence function such as LPL. Parameters other than bounded by unit interval were rounded to one decimal place. Values in parentheses are fixed during estimation, square brackets indicate construction method of dependence function, in particular [bi] denotes biconvex combination given by Proposition 3.1 for $n = 2$, [li] represents special case of [bi] when $a_i = b_i$ $(i = 1, 2)$, and [inv] refers to Proposition 3.3. So far we implemented construction procedures for two dependence functions only and their individual parameters are estimated separately (in advance) from weighting parameters of their combination.

family	generator $\varphi_{\boldsymbol{\theta}}(t)$	parameter range	limiting case (Archimed.)
Gumbel	$(-\log(t))^{\theta_1}$	$[1, \infty]$	$\{1\}\ \Pi$, $\{\infty\}\ M$
Clayton	$t^{-\theta_1} - 1$	$]0, \infty]$	$\{0\}\ \Pi$, $\{\infty\}\ M$
Frank	$-\log\left(\frac{e^{-\theta_1 t} - 1}{e^{-\theta_1} - 1}\right)$	$\Re$	$\{-\infty\}\ W$, $\{0\}\ \Pi$, $\{\infty\}\ M$
Joe	$-\log\left(1 - (1-t)^{\theta_1}\right)$	$[1, \infty]$	$\{1\}\ \Pi$, $\{\infty\}\ M$
BB1	$(t^{-\theta_1} - 1)^{\theta_2}$	$]0, \infty] \times [1, \infty]$	$\{0, 1\}\ \Pi$, $\{\infty, \infty\}\ M$
	dependence function $A_{\boldsymbol{\theta}}(t)$		limiting case (EV)
Mixed	$\theta_1 t^2 - \theta_1 t + 1$	$[0, 1]$	$\{0\}\ \Pi$
Gumbel (logistic)	$\left(t^{\theta_1} + (1-t)^{\theta_1}\right)^{1/\theta_1}$	$[1, \infty]$	$\{1\}\ \Pi$, $\{\infty\}\ M$
Hüsler Reiss	$t * \Phi\left(\frac{1}{\theta_1} + \frac{\theta_1}{2\log(t/(1-t))}\right) +$ $+(1-t)\Phi\left(\frac{1}{\theta_1} - \frac{\theta_1}{2\log(t/(1-t))}\right)$ Φ is CDF of standard normal	$[0, \infty]$	$\{1\}\ \Pi$, $\{\infty\}\ M$
Tawn (asymmetric logistic)	$1 - \theta_2 + (\theta_2 - \theta_1)t +$ $+\left((\theta_1 t)^{\theta_3} + (\theta_2(1-t))^{\theta_3}\right)^{\frac{1}{\theta_3}}$	$[0,1]\times$ $[0,1]\times$ $[0,\infty]$	$\{0, 0, 1\}\ \Pi$, $\{1, 1, \infty\}\ M$
LPL (linear-parabolic-linear)	$\begin{cases} 1 - \frac{1-b}{a}t & t \le a - c \\ \frac{b-a}{1-a} + \frac{1-b}{1-a}t & t \ge a + c \\ At^2 + Bt + C & \text{otherwise} \end{cases}$ $A = \frac{(1-b)}{4(1-a)ac}$ $B = \frac{2(1-b)(2ac-a-c)}{4(1-a)ac}$ $C = \frac{2(1+b)ac + (1-b)c^2 - (b+4c-1)a^2}{4(1-a)ac}$ $a = \theta_1,\ c = \theta_3 \min(a, 1-a)$ $b = \max(a, 1-a)(1-\theta_2) + \theta_2$	$[0,1]\times$ $[0,1]\times$ $[0,1]$	$\{0, ., .\}\ \{1, ., .\}\ \{., 1, .\}\ \Pi$ $\{0.5, 0, 0\}\ M$

Table 1. Overview of parametric families used to construct Archimax copula.

All procedures are implemented in R and freely available[1].

6. CONCLUSION

As seen from our results, given the two different data sets, the newly proposed construction methods do not give significantly better fit according to the selection criterion (which penalizes inclusion of additional parameters), however in case of dependence functions with roughly equal fitting performance they elevate the maximized likelihood. Note that the best results for fixed number of parameters are given by Archimax construction with both generator and dependence function, from which we may judge that the majority of well-established models in Archimedean and EV class capture mutually different dependence structure, in other words, they complement one another. The few exceptions that follow from Proposition 2.1 are equivalences of Archimedean copula with Gumbel generator and EV copula with Gumbel dependence function, or equivalence of BB1 and Archimax copula with Clayton generator and Gumbel dependence function.

In our software actually the estimation of Archimedean part is generally faster which may evoke a demand for some alternative to Proposition 2.1 in reverse order.

[1] www.math.sk/wiki/bacigal

generator		dependence function		log-lik	time	criterion
family	par.	family	par.	$L(\boldsymbol{\theta})$	[sec]	BIC
Gumbel	2.1			278.1	3	-549.8
Clayton	1.2			162.3	2	-318.1
Frank	6.6			255.2	3	-504.0
Joe	2.6			249.2	3	-492.0
BB1	0.1 2.1			278.5	11	-544.0
		Mixed	1.00	254.2	5	-502.0
		Gumbel	2.1	278.1	16	-549.8
		HüslerReiss	1.9	272.0	78	-537.7
		LL	0.56 0.70 (0.05)	66.9	1905	-120.9
		LPL sym.	(0.50) 0.05 0.80	274.0	1540	-528.5
		LPL	0.50 0.05 0.80	274.0	3122	-528.5
		Tawn	0.92 1.00 2.3	281.9	328	-544.3
Gumbel	2.1	Mixed	0.00	278.1	42	-543.3
Gumbel	1.5	Gumbel	1.5	278.1	36	-543.3
Gumbel	2.1	HüslerReiss	0.2	278.1	410	-543.3
Clayton	0.3	Mixed	1.00	269.1	84	-525.3
Clayton	0.1	Gumbel	2.1	278.5	55	-544.0
Clayton	0.9	HüslerReiss	1.8	272.9	190	-532.9
Frank	2.3	Mixed	0.97	280.8	86	-548.5
Frank	1.4	Gumbel	1.9	280.2	85	-547.4
Frank	1.8	HüslerReiss	1.5	276.1	255	-539.2
Joe	1.4	Mixed	0.98	273.9	81	-534.8
Joe	1.0	Gumbel	2.1	272.0	70	-531.2
Joe	1.0	HüslerReiss	1.9	272.1	348	-531.2
BB1	0.1 2.1	Mixed	0.00	278.5	71	-537.5
BB1	0.1 1.4	Gumbel	1.4	278.5	78	-537.5
BB1	0.1 2.1	HüslerReiss	0.1	278.5	40	-537.5
Gum–Cla	0.99			279.0	95	-538.5
Gum–Fra	0.46			278.7	53	-537.9
Gum–Joe	1.00			278.1	14	-536.7
Cla–Fra	0.00			256.2	335	-492.9
Cla–Joe	0.01			263.8	188	-508.1
Fra–Joe	0.74			271.8	64	-524.1
BB1–Gum	1.00			278.5	43	-537.5
BB1–Cla	1.00			278.5	14	-537.5
BB1–Fra	1.00			278.8	440	-537.5
BB1–Joe	1.00			278.5	16	-538.1
		[li] Mix–Gum	0.00	278.1	29	-536.7
		[inv]	0.00	278.1	322	-536.7
		[bi]	0.05 0.00	280.8	364	535.6
		[li] Mix–Hüs	0.10	276.3	314	-533.1
		[inv]	0.00	272.6	2049	-525.7
		[bi]	0.05 0.00	281.1	687	536.2
		[li] Gum–Hüs	0.71	278.5	125	-537.5
		[inv]	1.00	278.1	624	-536.7
		[bi]	0.92 0.99	279.2	1059	-532.4

Table 2. Estimation summary for **2 rivers flow rate**. Families denoted by [bi] and [li] (special case with $a_i = b_i$) refers to new construction method from Proposition 3.1 while [inv] to Proposition 3.3.

ACKNOWLEDGEMENT

This work was partially supported by the Slovak Research and Development Agency under the contract No. APVV-LPP-0004-07, SK-AT-0020-08 and APVV-0496-10, and by Program na podporu mladých výskumníkov, STU v Bratislave, projekt IMOPAR.

(Received August 31, 2010)

generator		dependence function		log-lik	GOF
family	par.	family	par.	$L(\boldsymbol{\theta})$	S_n
Gumbel	4.8			128.4	-252.1
Clayton	4.5			94.0	-183.3
Frank	18.4			123.9	-243.1
Joe	6.2			112.1	-219.5
BB1	0.3 4.3			129.5	-249.5
		Mixed	1.00	66.5	-128.3
		Gumbel	4.8	128.4	-252.1
		HüslerReiss	5.0	128.7	-252.7
		LL	0.50 0.50 (0.05)	57.3	-105.2
		LPL sym.	(0.50) 0.00 0.50	112.4	-215.3
		LPL	0.50 0.00 0.50	112.4	-210.6
		Tawn	1.00 1.00 4.8	128.4	-242.6
Gumbel	2.8	Mixed	1.00	128.6	-247.8
Gumbel	2.2	Gumbel	2.2	128.4	-247.2
Gumbel	3.1	HüslerReiss	1.3	128.9	-248.3
Clayton	2.0	Mixed	1.00	108.5	-207.5
Clayton	0.3	Gumbel	4.3	129.5	-249.6
Clayton	0.3	HüslerReiss	4.5	129.9	-250.5
Frank	10.0	Mixed	1.00	128.5	-247.6
Frank	4.0	Gumbel	3.2	131.3	-253.1
Frank	4.0	HüslerReiss	3.2	131.9	-254.4
Joe	3.4	Mixed	1.00	117.7	-225.9
Joe	1.0	Gumbel	4.8	128.4	-247.4
Joe	1.0	HüslerReiss	5.0	128.7	-247.9
BB1	0.3 2.5	Mixed	1.00	129.8	-245.4
BB1	0.3 1.7	Gumbel	2.5	129.5	-244.8
BB1	0.2 2.2	HüslerReiss	1.7	130.0	-245.8
Gum–Cla	1.00			128.4	-242.6
Gum–Fra	1.00			128.4	-242.6
Gum–Joe	1.00			128.4	-242.6
Cla–Fra	0.00			106.0	-197.8
Cla–Joe	0.00			112.1	-210.0
Fra–Joe	0.00			112.1	-210.0
BB1–Gum	1.00			129.4	-244.6
		[li] Mix–Gum	0.00	128.4	-242.6
		[inv]	0.00	128.4	-242.6
		[bi]	0.00 0.00	128.4	-237.9
		[li] Mix–Hüs	0.00	128.7	-243.2
		[inv]	0.00	128.7	-243.2
		[bi]	0.00 0.00	128.7	-238.5
		[li] Gum–Hüs	0.22	128.7	-243.2
		[inv]	0.26	128.6	-243.4
		[bi]	0.86 0.92	129.7	-240.5

Table 3. Estimation summary for **summer flood data**.

REFERENCES

[1] T. Bacigál, M. Juráňová, and R. Mesiar: On some new constructions of Archimedean copulas and applications to fitting problems. Neural Network World *1* (2010), 10, 81–90.

[2] P. Capéraá, A.-L. Fougéres, and G. Genest: Bivariate distributions with given extreme value attractor. J. Multivariate Anal. *72* (2000), 1, 30–49.

[3] B. De Baets, H. De Meyer, and R. Mesiar: Lipschitz continuity of copulas w.r.t. L_p-norms Nonlinear Anal., Theory, Methods Appl. *72(9-10)* (2010), 3722–3731.

[4] F. Durante and R. Mesiar: L^∞-measure of non-exchangeability for bivariate extreme value and Archimax copulas. J. Math. Anal. Appl. *369* (2010), 2, 610–615.

[5] G. Genest and A.-C. Favre: Everything you always wanted to know about copula modeling but were afraid to ask. J. Hydrolog. Engrg. *12* (2007), 4, 347–368.

[6] C. Genest and B. Rémillard: Tests of independence or randomness based on the empirical copula process. Test *13* (2004), 335–369.

[7] C. Genest, B. Rémillard, and D. Beaudoin: Goodness-of-fit tests for copulas: A review and a power study. Insurance Math. Econom. *44* (2009), 199–213.

[8] C. Genest, K. Ghoudi, and L.-P. Rivest: A semi-parametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika *82* (1995), 543–552.

[9] C. Genest, K. Ghoudi, and L.-P. Rivest: Discussion of the paper by Frees and Valdez (1998). North Amer. Act. J. *2* (1998), 143–149.

[10] H. Joe: Multivariate Models and Dependence Concepts. Chapman and Hall, London 1997.

[11] A. Khoudraji: Contributions à l'étude des copules et à la modélisation des valeurs extrémes bivariées. PhD. Thesis, Université Laval, Québec 1995.

[12] E. P. Klement, R. Mesiar, and E. Pap: Triangular Norms. Kluwer, Dordrecht 2000.

[13] E. Liebscher: Construction of Asymmetric multivariate copulas. J. Multivariate Anal. *99* (2008), 10, 2234–2250.

[14] A. Mesiarová: k-l_p-Lipschitz t-norms. Internat. J. Approx. Reasoning *46* (2007), 3, 596–604.

[15] R. B. Nelsen: An Introduction to Copulas. Second edition. Springer–Verlag, New York 2006.

[16] J. A. Tawn: Bivariate extreme value theory: Models and estimation. Biometrika *75* (1988), 397–415.

[17] C. J. Vorosmarty, B. M. Fekete, B. A. and Tucker: Global River Discharge, 1807-1991, V. 1.1 (RivDIS). Data set. 1998. Available on-line [http://www.daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, TN, U.S.A. doi:10.3334/ORNLDAAC/199.

[18] R. R. Yager: On a general class of fuzzy connectives. Fuzzy Sets and Systems *4* (1980), 3, 235–242.

Tomáš Bacigál, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, Bratislava. Slovak Republic.
e-mail: bacigal@math.sk

Vladimír Jágr, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, Bratislava. Slovak Republic.
e-mail: jagr@math.sk

Radko Mesiar, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, Bratislava, Slovak Republic, and Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.
e-mail: mesiar@math.sk

Available online at www.sciencedirect.com

ScienceDirect

Fuzzy Sets and Systems 264 (2015) 42–50

FUZZY sets and systems

www.elsevier.com/locate/fss

Additive generators of copulas

Tomáš Bacigál [a], Vadoud Najjari [d,*], Radko Mesiar [a,b], Hasan Bal [c]

[a] Faculty of Civil Engineering, Slovak University of Technology in Bratislava, Slovakia
[b] Institute for Research and Application of Fuzzy Modeling, University of Ostrava, Czech Republic
[c] Gazi University, Faculty of Science, Statistics Department, Ankara, Turkey
[d] Department of Mathematics, Islamic Azad University, Maragheh branch, Maragheh, Iran

Received 24 January 2014; received in revised form 27 August 2014; accepted 1 October 2014

Available online 7 October 2014

Abstract

In this study, we discuss additive generators of copulas with a fixed dimension $n \geq 2$ and additive generators that yield copulas for any dimension $n \geq 2$. We review the reported methods used to construct additive generators of copulas, and we introduce and exemplify some new construction methods.

Keywords: Additive generator; Archimedean copula; Copula; Joint distribution function

1. Introduction

Since their introduction by Sklar in [24], copulas have become an important tool for modeling the stochastic dependence of random vectors and thus the modeling of real data, which can be viewed as outcomes of some n-dimensional random experiment, $n \geq 2$. Thus, copulas can be considered to be basic tools in statistics, but also in related sciences, including economics, information sciences, and sociology. We recall that copulas aggregate one-dimensional marginal distribution functions into n-dimensional joint distribution functions. As a typical example, we recall the case of independent random variables where the stochastic dependence is captured by the product copula Π and the joint distribution function is simply the product of the corresponding continuous marginal one-dimensional distribution functions.

From an axiomatic viewpoint, a function $C: [0, 1]^n \to [0, 1]$ is called a (n-dimensional) copula whenever it satisfies the boundary conditions (C1) and it is an n-increasing function (C2), as follows.

(C1) $C(x_1, \ldots, x_n) = 0$ whenever $0 \in \{x_1, \ldots, x_n\}$, i.e., 0 is an annihilator of C, and $C(x_1, \ldots, x_n) = x_i$ whenever $x_j = 1$ for each $j \neq i$ (i.e., 1 is a neutral element of C),

* Corresponding author.
E-mail addresses: tomas.bacigal@stuba.sk (T. Bacigál), vnajjari@gazi.edu.tr (V. Najjari), radko.mesiar@stuba.sk (R. Mesiar), hasanbal@gazi.edu.tr (H. Bal).

http://dx.doi.org/10.1016/j.fss.2014.10.001

(C2) For any $\mathbf{x}, \mathbf{y} \in [0,1]^n$, $\mathbf{x} \leq \mathbf{y}$ (i.e., $x_1 \leq y_1, \dots, x_n \leq y_n$), it holds that

$$V_C\big([\mathbf{x},\mathbf{y}]\big) = \sum_{\varepsilon \in \{-1,1\}^n} \left(C(\mathbf{z}_\varepsilon) \prod_{i=1}^{n} \varepsilon_i \right) \geq 0,$$

where $\mathbf{z}_\varepsilon = (z_1^{\varepsilon_1}, \dots, z_n^{\varepsilon_n})$, $z_i^1 = y_i$, $z_i^{-1} = x_i$.

Note that $V_C([\mathbf{x},\mathbf{y}])$ is called the C-volume of the rectangle $[\mathbf{x},\mathbf{y}]$.

As mentioned earlier, the main interest in copulas is due to Sklar's theorem [24]: for a random vector $Z = (X_1, \dots, X_n)$, $F_Z \colon R^n \to [0,1]$ is a joint distribution of Z if and only if there is a copula $C \colon [0,1]^n \to [0,1]$ such that

$$F_Z(x_1, \dots, x_n) = C\big(F_{X_1}(x_1), \dots, F_{X_n}(x_n)\big), \tag{1}$$

where $F_{X_i} \colon R \to [0,1]$ is a distribution function related to the random variable X_i, $i = 1, \dots, n$. The copula C in (1) is unique whenever the random variables are continuous. For more details of copulas, we recommend [11] and [21].

A highly prominent class of binary copulas is the class of Archimedean copulas characterized by the associativity of C and the diagonal inequality $C(x,x) < x$ for all $x \in \,]0,1[$. Note that although Archimedean copulas are necessarily symmetric, i.e., they can model the stochastic dependence of exchangeable random variables (X,Y) only, they comprise most of the copula families employed in financial, hydrological, and other application areas. For fitting purposes, these copulas are used in most of the software systems that deal with copulas, such as [1,9,28,29]. The popularity of Archimedean copulas is explained by their representation using one-dimensional functions, which are generally called additive generators of (binary) copulas. This crucial result is attributed to Moynihan [20].

Theorem 1. *A function $C \colon [0,1]^2 \to [0,1]$ is an Archimedean copula if and only if there is a convex strictly decreasing function $f \colon [0,1] \to [0,\infty]$, $f(1) = 0$, such that*

$$C(x,y) = f^{(-1)}\big(f(x) + f(y)\big), \tag{2}$$

where the pseudo-inverse $f^{(-1)} \colon [0,\infty] \to [0,1]$ is given by

$$f^{(-1)}(u) = f^{-1}\big(\min\big(u, f(0)\big)\big).$$

The function f is called an additive generator of the copula C and it is unique up to a positive multiplicative constant.

We denote $\mathcal{F}_2$ as the class of all additive generators of the binary copulas characterized in the theorem above. Many families of these generators can be found in numerous previous studies, such as [11,13,21]. Several studies have been devoted to methods for constructing additive generators, which we review in Section 3. We also recall an important link between additive generators of copulas and positive distance functions based on the Williamson transform, as observed and discussed by McNeil and Nešlehová in [17].

However, copulas of higher dimensions can also be generated using additive generators. Thus, previous studies inspired us to review the known details for additive generators of copulas and to introduce some new methods for generating them. This paper is organized as follows. In the next section, we summarize known results for additive generators of n-ary copulas (copulas of any dimension). In Section 3, we review some previously reported methods for constructing additive generators of copulas and we also propose a new construction method based on the Williamson transform (see [17]). Section 4 proposes some new construction methods and we present examples. Finally, we provide some concluding remarks.

2. Additive generators of copulas

For any binary Archimedean copula $C \colon [0,1]^2 \to [0,1]$ generated by an additive generator $f \colon [0,1] \to [0,\infty]$, C is also a triangular norm [13,20,23] and thus it can be extended univocally to an n-ary function (we retain the original notation for this extension) $C \colon [0,1]^n \to [0,1]$ given by

$$C(x_1, \dots, x_n) = f^{(-1)}\left(\sum_{i=1}^{n} f(x_i) \right). \tag{3}$$

Obviously, for any $n \geq 2$, C satisfies the boundary conditions (C1). However, for $n > 2$, (C2) may fail to be satisfied. For example, the smallest binary copula $W: [0,1]^2 \to [0,1]$ is generated by an additive generator $f_W: [0,1] \to [0,\infty]$, $f(x) = 1 - x$, and $W(x, y) = \max(0, x + y - 1)$. Its n-ary extension is given by

$$W(x_1, \ldots, x_n) = 1 - \min\left(1, \sum_{i=1}^{n}(1 - x_i)\right) = \max\left(0, \sum_{i=1}^{n} x_i - (n-1)\right).$$

Consider $\mathbf{x}, \mathbf{y} \in [0,1]^n$, $\mathbf{x} = (\frac{1}{2}, \ldots, \frac{1}{2})$, $\mathbf{y} = (1, \ldots, 1)$. Then, $V_W([\mathbf{x}, \mathbf{y}]) = 1 - \frac{n}{2}$, i.e., this volume is negative whenever $n > 2$, which shows that W is a copula only for $n = 2$. A complete description of additive generators of binary copulas such that the corresponding n-ary function generated given by (3) is also an n-ary copula, $n > 2$, was given by McNeil and Nešlehová in [17].

Theorem 2. *Let $f: [0,1] \to [0,\infty]$ be a continuous strictly decreasing function such that $f(1) = 0$ (i.e., f is an additive generator of a continuous Archimedean t-norm; see [13]). Then, the n-ary function $C: [0,1]^n \to [0,1]$ given by (3) is an n-ary copula if and only if the function $g: [-\infty, 0] \to [0,1]$ given by $g(u) = f^{(-1)}(-u)$ is $(n-2)$ times differentiable with non-negative derivatives $g', \ldots, g^{(n-2)}$ on $]-\infty, 0[$, and $g^{(n-2)}$ is convex.*

We denote $\mathcal{F}_n$ as the class of all additive generators that generate n-ary copulas, as characterized in Theorem 2.

Note that the smallest n-ary generated copula is the non-strict Clayton copula [11,21] $C^{Cl}_{-\frac{1}{n-1}}: [0,1]^n \to [0,1]$ generated by an additive generator $f^{Cl}_{-\frac{1}{n-1}}: [0,1] \to [0,\infty]$, $f^{Cl}_{-\frac{1}{n-1}}(x) = 1 - x^{\frac{1}{n-1}}$.

Universal additive generators yield an n-ary copula for any $n \geq 2$. We obtain the next result due to Theorem 2 (see also [17]).

Corollary 1. *Let $f: [0,1] \to [0,\infty]$ be an additive generator of a binary copula $C: [0,1]^2 \to [0,1]$. Then, the n-ary extension $C: [0,1]^n \to [0,1]$ given by (3) is an n-ary copula for each $n \geq 2$ if and only if the function $g: [-\infty, 0] \to [0,1]$ given by $g(u) = f^{(-1)}(-u)$ is absolutely monotonic, i.e., $g^{(k)}$ exists and it is non-negative for each $k \in N = \{1, 2, \ldots\}$.*

The class of all universal additive generators characterized in the corollary above is denoted by $\mathcal{F}_\infty$. It is not difficult to check that $\mathcal{F}_2 \supset \mathcal{F}_3 \supset \ldots \supset \mathcal{F}_\infty$. As a typical example to illustrate Corollary 1, we consider the product copula Π and its corresponding additive generator $f_\Pi: [0,1] \to [0,\infty]$ given by $f_\Pi(x) = -\log x$. Obviously, the corresponding n-ary copula is the n-ary product and it captures the stochastic dependence structure of n-dimensional random vectors with independent marginals, provided that they are continuous. As another example of a universal additive generator, we consider the function $f: [0,1] \to [0,\infty]$ given by $f(x) = \frac{1}{x} - 1$ with convention $f(0) = \infty$ (the additive generator of the Ali–Mikhail–Haq copula, which is also called the Hamacher product in the t-norms area). Then, $f^{(-1)}(u) = \frac{1}{u+1}$, $u \in [0,\infty]$, i.e., $g: [-\infty, 0] \to [0,1]$ is given by $g(u) = (1-u)^{-1}$. Then, for any $k \in N$, $g^{(k)}(u) = k!(1-u)^{-(k+1)}$, i.e., g is absolutely monotonic. The corresponding n-ary copula $C: [0,1]^n \to [0,1]$ is then given by

$$C(x_1, \ldots, x_n) = \frac{1}{\sum_{i=1}^{n} \frac{1}{x_i} - (n-1)}$$

(obviously, if some values of x_i equal 0, because of (C1), $C(x_i, \ldots, x_n) = 0$).

Recently, the reverse problem of characterizing n-ary copulas generated by an additive generator was solved by Stupňanová and Kolesárová [25].

Theorem 3. *Let $C: [0,1]^n \to [0,1]$ be an n-ary copula, $n > 2$. Then, C is generated by an additive generator $f: [0,1] \to [0,\infty]$ if and only if C satisfies the diagonal inequality $C(x, \ldots, x) < x$ for all $x \in]0,1[$, and C is associative in the sense of Post, i.e., for any $x_1, \ldots, x_{2n-1} \in [0,1]$, it holds that*

$$\begin{aligned} &C\big(C(x_1, \ldots, x_n), x_{n+1}, \ldots, x_{2n-1}\big) \\ &= C\big(x_1, C(x_2, \ldots, x_{n+1}), x_{n+2}, \ldots, x_{2n-1}\big) = \ldots \\ &= C\big(x_1, \ldots, x_{n-1}, C(x_n, \ldots, x_{2n-1})\big). \end{aligned}$$

3. Known methods for constructing additive generators of copulas

Many families of (binary) generated copulas have been obtained as solutions to various problems, including Frank copulas [6], Plackett copulas [22], Clayton copulas [4], and Gumbel copulas [8]. Several other ad hoc families have also been proposed, such as Yager copulas [27] (a subfamily of Yager t-norms), and a small number of families of generated copulas were summarized in [11,21]. Recently, [3] discussed an aggregation function that preserves the classes of additive generators (of binary copulas) and their pseudo-inverses. These methods facilitate the construction of new additive generators (or their pseudo-inverses) from k a priori given additive generators, $k \geq 2$. In the present study, we discuss methods for constructing additive generators of copulas (binary, n-ary, universal) from some a priori given function. Previously reported methods deal mostly with an a priori given additive generator f and they aim to construct new additive generators using f. One of the first results of this type is attributed to [14].

Proposition 1. *Let $\varphi\colon [0,1] \to [0,1]$ be a concave automorphism (i.e., increasing bijection) of $[0,1]$. Then, for any $f \in \mathcal{F}_2$, and $f \circ \varphi \in \mathcal{F}_2$.*

This result was extended in [5] by relaxing the bijectivity of f, i.e., by considering concave continuous strictly increasing transforms $\varphi\colon [0,1] \to [0,1]$ such that $\varphi(1) = 1$, and then again $f \circ \varphi \in \mathcal{F}_2$ for any $f \in \mathcal{F}_2$.

Example 1. Consider $f_\Pi \in \mathcal{F}_2$, $f_\Pi(x) = -\log x$ (recall the convention $f(0) = \infty$), and $\varphi\colon [0,1] \to [0,1]$ given by $\varphi(x) = a + (1-a)x$, $a \in\,]0,1[$. Then, $f_\Pi \circ \varphi(x) = -\log(a + (1-a)x)$, $x \in [0,1]$, and the corresponding copula $C\colon [0,1]^2 \to [0,1]$ is given by

$$C(x,y) = \max\left(0, \frac{(a + (1-a)x)(a + (1-a)y) - a}{1-a}\right).$$

In [2], Proposition 1 was modified for the class $\mathcal{F}_\infty$.

Proposition 2. *Let $\varphi\colon [0,1] \to [0,1]$ be an automorphism of $[0,1]$ such that its inverse $\varphi^{-1}\colon [0,1] \to [0,1]$ is absolutely monotonic on $]0,1[$ (i.e., $(\varphi^{-1})^{(k)}(x) \geq 0$ for any $k \in N$ and $x \in [0,1]$). Then, for any $f \in \mathcal{F}_\infty$ also $f \circ \varphi \in \mathcal{F}_\infty$.*

Using similar ideas to those described in [2,14], we can also prove the next result.

Proposition 3. *Let $n \in \{2,3,\ldots\}$. Let $\varphi\colon [0,1] \to [0,1]$ be an automorphism of $[0,1]$ such that its inverse $\varphi^{-1}\colon [0,1] \to [0,1]$ has $(n-2)$ derivatives on $]0,1[$, $(\varphi^{-1})^{(k)}(x) \geq 0$ for all $k \in 1,\ldots,n-2$ and $x \in\,]0,1[$, and $(\varphi^{-1})^{(n-2)}$ is a convex function. Then, for all $f \in \mathcal{F}_n$ also $f \circ \varphi \in \mathcal{F}_n$.*

An alternative approach to the transformation of additive generators was proposed in [2].

Proposition 4. *Let $\eta\colon [0,\infty] \to [0,\infty]$ be a convex automorphism of $[0,\infty]$. Then, for any $f \in \mathcal{F}_2$ also $\eta \circ f \in \mathcal{F}_2$.*

An important consequence of Proposition 4 is linked to the power functions $\eta(u) = u^\lambda$, $\lambda \geq 1$, which ensure that any $f \in \mathcal{F}_2$, and the family $(f^\lambda)_{\lambda \geq 1} \subset \mathcal{F}_2$. For example, we recall the Gumbel family of copulas linked to the family $((-\log x)^\lambda)_{\lambda \geq 1}$ of additive generators from $\mathcal{F}_2$.

Due to Theorem 2 and Corollary 1, we can derive the next generalization of Proposition 4 for additive generators from $\mathcal{F}_n$, $n \geq 2$, and from $\mathcal{F}_\infty$.

Proposition 5. *Let $n \in \{2,3,\ldots\}$. Let $\eta\colon [0,\infty] \to [0,\infty]$ be an automorphism such that its inverse $\eta^{-1}\colon [0,\infty] \to [0,\infty]$ has $(n-2)$ derivatives (all derivatives) on $]0,\infty[$, $(\eta^{-1})^{(k)}(x) \geq 0$ for all $x \in\,]0,\infty[$ and $k \in \{1,\ldots,n-2\}$ ($k \in N$) such that $(\eta^{-1})^{(n-2)}$ is a convex function. Then, for any $f \in \mathcal{F}_n$ (any $f \in \mathcal{F}_\infty$) also $\eta \circ f \in \mathcal{F}_n$ ($\eta \circ f \in \mathcal{F}_\infty$).*

Another method based on an a priori given additive generator $f \in \mathcal{F}_2$ that yields a parametric family of additive generators was obtained as a result of univariate conditioning in [10,16], which can be generalized for any $n \in \{3,4,\ldots\} \cup \{+\infty\}$.

Proposition 6. *Let $f \in \mathcal{F}_n$, $n \in \{2, 3, \dots\} \cup \{+\infty\}$. Then, $(f_\lambda)_{\lambda \in]0,1]} \subset \mathcal{F}_n$, where $f_\lambda\colon [0, 1] \to [0, \infty]$ is given by*

$$f_\lambda(x) = f(\lambda x) - f(\lambda).$$

The parametric family $(f_\lambda)_{\lambda \in]0,1]}$ is non-trivial (i.e., its members generate different copulas with different parameters) if and only if f do not belong to the Clayton family of additive generators. In the case of Clayton's additive generators, $f_\lambda = f$ for each $\lambda \in]0, 1]$.

Note that Proposition 1, Proposition 4, and Proposition 6 for $n = 2$ were also considered by [7] (as the right composition, left composition, and scaling rule, respectively) and described further in [19].

Several construction methods are available and we recall some of them. Junker and May [12] showed that for any $f \in \mathcal{F}_2$ and $a \in]1, \infty[$, also $g = a^f - 1 \in \mathcal{F}_2$, for any $f \in \mathcal{F}_2$ and $a \in]0, 1[$, also $h = a^{-f} - 1 \in \mathcal{F}_2$. However, if we consider $\eta\colon [0, \infty] \to [0, \infty]$ given by $\eta(u) = a^u - 1$, $a \in]1, \infty[$, it is evident that η is a convex automorphism of $[0, \infty]$, thus this result is a special instance of Proposition 4. Similarly, $\eta(u) = a^{-u} - 1$ defines a convex automorphism of $[0, \infty]$ whenever $a \in]0, 1]$. Another result is attributed to Michiels and Schepper [18,19]. By considering $f_1, f_2 \in \mathcal{F}_2$ such that $(f_2')^2 \le f_2'$, they also showed that $f_1(e^{-f_2}) \in \mathcal{F}_2$. However, if $(f_2')^2 \le f_2'$, then e^{-f_2} is a concave continuous strictly increasing function that satisfies the modified form of Proposition 1 (due to Durante and Sempi [5]).

An interesting link between additive generators of copulas and positive distance functions [15], i.e., distribution functions with support in $]0, \infty[$, was described in detail in [17]. Based on the results of Williamson [26], we recall the next important result.

Theorem 4. *(See [17, Corollary 3.1].) The following claims are equivalent for an arbitrary $n \in \{2, 3, \dots\}$:*

(i) $f \in \mathcal{F}_n$
(ii) *Under the notation of Theorem 2, the function $F\colon]-\infty, \infty[\to [0, 1]$ given by $F(x) = 0$ if $x \le 0$, and for $x > 0$,*

$$F(x) = 1 - \sum_{k=0}^{n-2} \frac{x^k g^{(k)}(-x)}{k!} - \frac{x^{n-1} g_-^{(n-1)}(-x)}{(n-1)!} \tag{4}$$

is a distribution function of a positive random variable X (i.e., $P(X \le 0) = 0$), where $g_-^{(n-1)}$ is the left-derivative of order $n - 1$.

We observe that due to [26], if F is a positive distance function, i.e., a distribution function of a positive random variable X, then for a fixed $n \in \{2, 3, \dots\}$, the Williamson n-transform provides an inverse transformation to (4),

$$g(x) = \int_{-x}^{\infty} \left(1 + \frac{x}{t}\right)^{n-1} dF(t), \tag{5}$$

where $x \in]-\infty, 0]$, $g(-\infty) = 0$.

Due to the transformations (4) and (5) described above, we can construct new additive generators of (n-dimensional) copulas as follows:

- For an arbitrary $m \in \{2, 3, \dots\}$, take an additive generator $f \in \mathcal{F}_m$;
- Introduce a positive distance function F using the (4) transform;
- Possibly modify F into a new positive distance function $\tilde{F}$ (e.g., $\tilde{F}(x) = F(x - a)$ for a fixed constant $a \in]0, \infty[$);
- Apply the Williamson transform (5) to $\tilde{F}$ by considering a fixed $n \in \{2, 3, \dots\}$, to obtain a function $\tilde{g}\colon [-\infty, 0] \to [0, 1]$;
- $\tilde{f}$ linked to $\tilde{g}$ is an additive generator from $\mathcal{F}_n$.

Note that a similar relationship can be demonstrated between additive generators from $\mathcal{F}_\infty$ and positive distance functions based on the Laplace transform. For further details, we recommend [17].

Example 2. Consider $f_W \in \mathcal{F}_2$. Then, $g: [-\infty, 0] \to [0, 1]$ is given by $g(x) = \max(0, x + 1)$. Based on transform (4), we define a positive distance function $F:]-\infty, \infty[\to [0, 1]$ given by $F(x) = 0$ if $x \le 0$, but otherwise by

$$F(x) = 1 - g(-x) - xg'_{-}(-x) = 1 - \max(0, 1 - x) - x\mathbf{1}_{]-1,0]}(-x) = \begin{cases} 0 & x \le 1 \\ 1 & x > 1 \end{cases},$$

i.e., F is the Dirac distribution function focused on point $x_0 = 1$. For an arbitrary $n \in \{2, 3, \ldots\}$, the Williamson transform (5) defines a function $\tilde{g}: [-\infty, 0] \to [0, 1]$ given by $\tilde{g}(-\infty) = 0$, and for $x \in]-\infty, 0]$, by

$$\tilde{g}(x) = \int_{-x}^{\infty} \left(1 + \frac{x}{t}\right)^{n-1} dF(t) = (1 + x)^{n-1}.$$

Then, the related pseudo-inverse $\tilde{f}_{n-1}^{(-1)}: [0, \infty] \to [0, 1]$ is given by $\tilde{f}^{(-1)}(x) = (1 - x)^{n-1}$, and the related additive generator $\tilde{f}: [0, 1] \to [0, \infty]$ given by $\tilde{f}(x) = 1 - x^{\frac{1}{n-1}}$ belongs to $\mathcal{F}_n$. We observe that $\tilde{f}$ is an additive generator of a non-strict Clayton copula with parameter $\lambda = \frac{1}{n-1}$ (the weakest n-dimensional Archimedean copula).

In the next section, we introduce some new methods for constructing additive generators of copulas.

4. New methods for constructing additive generators of copulas

The next result is a generalization of Proposition 6.

Theorem 5. *Let $h: [a, b] \to [-\infty, \infty]$ be a strictly decreasing convex continuous function. Then, for any non-trivial bounded $[c, d] \subseteq [a, b]$ (if $h(b) = -\infty$, then $[c, d] \subset [a, b[$), the function $f_{c,d}: [0, 1] \to [0, \infty]$ given by*

$$f_{c,d}(x) = h\big(c + x(d - c)\big) - h(d)$$

is an additive generator from $\mathcal{F}_2$.

Proof. First, we observe that $c + x(d - c)$ defines a strictly increasing linear transformation of the arguments from $[0, 1]$. Thus, $h(c + x(d - c))$ defines a continuous, strictly decreasing convex function on $[0, 1]$, and the same holds for its shift $f_{c,d}$. Moreover, $f_{c,d}(1) = 0$, i.e., $f_{c,d} \in \mathcal{F}_2$. □

Example 3.

(i) Consider $h: [-\infty, \infty] \to [-\infty, \infty]$ given by $h(x) = e^{-x}$. Then, for any $c, d \in]-\infty, \infty[$, $c < d$, $f_{c,d}: [0, 1] \to [0, \infty]$ is given by

$$f_{c,d}(x) = e^{-(c+x(d-c))} - e^{-d} = e^{-c}\big(e^{-x(d-c)} - e^{-(d-c)}\big).$$

We observe that e^{-c} is a positive multiplicative constant and thus $f_{c,d}$ depends on $\lambda = d - c > 0$ only. Then, the additive generator $f_\lambda \in \mathcal{F}_2$ given by $f_\lambda(x) = e^{-\lambda x} - e^{-\lambda}$ generates the same binary copula as the additive generator $f_{c,d}$.

(ii) Consider $h: [0, \infty] \to [-\infty, \infty]$, $h(x) = \frac{1}{\arctan x}$. Then, h satisfies the constraints of Theorem 5 and thus $f_{c,d}: [0, 1] \to [0, \infty]$ given for any $[c, d] \subset [[0, \infty[$ by $f_{c,d}(x) = \frac{1}{\arctan(c+x(d-c))} - \frac{1}{\arctan d}$ is an additive generator from $\mathcal{F}_2$. In particular, an interesting parametric family $(f_{0,d})_{d\in]0,\infty[}$ of additive generators from $\mathcal{F}_2$ is given by

$$f_{0,d}(x) = \frac{1}{\arctan dx} - \frac{1}{\arctan d}.$$

(iii) Consider $h: [-\infty, \infty] \to [-\infty, \infty]$ given by $h(x) = -\log x$. Due to Theorem 5, for any $0 \le c < d < \infty$, $f_{c,d}: [0, \infty] \in [0, 1]$ given by

$$f_{c,d}(x) = -\log\big(c + x(d - c)\big) - \log d = -\log\big(a + (1 - a)x\big),$$

where $a = \frac{c}{d} \in [0, 1[$ is an additive generator from $\mathcal{F}_2$ (also compare Example 1)

Theorem 5 can be generalized to obtain additive generators from $\mathcal{F}_n$, $n \in \{3, 4, \ldots\}$, and from $\mathcal{F}_\infty$.

Theorem 6. *Let $h: [a, b] \to [-\infty, \infty]$ satisfy the constraints of Theorem 5. Then,*

i) *If for $n \in \{3, 4, \ldots\}$, the inverse function h^{-1} has $(n-2)$ derivatives on $]h(b), h(a)[$ such that $(h^{-1})^{(k)}(x) \cdot (-1)^k \geq 0$ for all $k \in \{1, \ldots, n-2\}$ and $x \in]h(b), h(a)[$, and $(h^{-1})^{(n-2)}(-1)^k$ is convex, then for any bounded interval $[c, d] \subset [a, b]$ (if $h(b) = -\infty$, then $d < b$), the function $f_{c,d}$ given in Theorem 5 is an additive generator from $\mathcal{F}_n$.*

ii) *If the inverse function h^{-1} is totally monotonic on $]h(b), h(a)[$, then $f_{c,d}$ given in Theorem 5 belongs to $\mathcal{F}_\infty$.*

As examples, we observe that $f_{c,d}$ given in Example 3 i) and iii) belong to $\mathcal{F}_\infty$, i.e., for any dimension $n \geq 2$, $f_{c,d}$ generates an n-ary copula.

Example 4. Define $h: [0, \infty] \to [-\infty, \infty]$ by $h(x) = -x^{0.4}$. Then, $h^{-1}: [-\infty, 0] \to [0, \infty]$ is given $h^{-1}(u) = (-u)^{2.5}$. It is evident that $(h^{-1})'(u) = -\frac{5}{2}(-u)^{1.5}$, $(h^{-1})''(u) = -\frac{15}{4}(-u)^{0.5}$, $(h^{-1})'''(u) = -\frac{15}{8}(-u)^{0.5}$, $(h^{-1})^{(4)}(u) = -\frac{15}{16}(-u)^{-1.5}$, i.e., h satisfies the constraints of Theorem 6 i) for $n = 3$ but not for $n = 4$. Thus, $f_{c,d}: [0, 1] \to [0, \infty]$ given that for $0 \leq c < d < \infty$, by $f_{c,d}(x) = d^{0.4} - (c + (d - c)x)^{0.4}$ generates a three-dimensional copula but not a four-dimensional copula. We observe that if $c = 0$, then $f_{0,d}(x) = d^{0.4}(1 - x^{0.4})$ is the Clayton copula with parameter -0.4 (see [11,21]).

Obviously, any additive generator $f \in \mathcal{F}_2$ ($\mathcal{F}_n$, $\mathcal{F}_\infty$) satisfies the constraints of Theorem 5 (Theorem 6), thus our results can be viewed as an extension and generalization of Proposition 6. Moreover, we can generalize Propositions 1–5 to construct additive generators using Theorems 5 and 6 in two ways: we can either apply them directly to the introduced additive generators $f_{c,d}$, or we can apply them (in modified form) to the generating function h. We illustrate the latter approach with the modified Proposition 1.

Theorem 7. *Let $h: [a, b] \to [-\infty, \infty]$ satisfy the constraints of Theorem 5. Let $\varphi: [\alpha, \beta] \to [a, b]$ be a concave increasing bijection. Then, the function $h \circ \varphi: [\alpha, \beta] \to [-\infty, \infty]$ also satisfies the constraints of Theorem 5, i.e., for any bounded interval $[\gamma, \delta] \subseteq [\alpha, \beta]$ (if $h(0) = -\infty$, then $\delta < \beta$), the function $f_{\gamma,\delta}: [0, 1] \to [0, \infty]$ given by*

$$f_{\gamma,\delta}(x) = h\big(\varphi\big(\gamma + (\delta - \gamma)x\big)\big) - h\big(\varphi(\delta)\big)$$

is an additive generator from $\mathcal{F}_2$.

Example 5. By continuing Example 4, let $\varphi: [0, \infty] \to [0, \infty]$ be given by $\varphi(x) = \sqrt{x}$. Then, φ is a concave increasing bijection. In addition, $\varphi_{|[0,1]}$ is a concave increasing $[0, 1] \to [0, 1]$ bijection. For the additive generator $f_{c,d}$, by applying Proposition 1, $f_{c,d} \circ \varphi_{|[0,1]}(x) = d^{0.4} - (c + (d - c)\sqrt{x})^{0.4}$, while after applying Theorem 7 and by considering $h \circ \varphi: [0, \infty] \to [0, \infty]$ given by $h \circ \varphi(x) = -x^{0.2}$, for $0 \leq \gamma < \delta < \infty$, we have $f_{\gamma,\delta}(x) = \delta^{0.2} - (\gamma + (\delta - \gamma)x)^{0.2}$. Note that both $f_{c,d} \circ \varphi_{|[0,1]}$ and $f_{\gamma,\delta}$ are additive generators from $\mathcal{F}_2$.

Another new method for constructing additive generators from $\mathcal{F}_2$ is based on the gluing of two additive generators from $\mathcal{F}_2$. Note that this approach can be extended for any dimension n due to the Williamson transform (5) and transform (4).

Theorem 8. *Let $f_1, f_2 \in \mathcal{F}_2$ and $k \in]0, 1[$ be given. Define a function $f: [0, 1] \to [0, \infty]$, which is also denoted by $f = f_1 *_k f_2$, as*

$$f(x) = \begin{cases} \frac{f_1(x)}{f_1(k)} & \text{if } x \in [0, k], \\ \frac{f_2(x)}{f_2(k)} & \text{otherwise} \end{cases}$$

whenever $\frac{f'_{1-}(k)}{f_1(k)} \leq \frac{f'_{2-}(k)}{f_2(k)}$. If $\frac{f'_{1-}(k)}{f_1(k)} > \frac{f'_{2-}(k)}{f_2(k)}$, then

$$f(x) = \begin{cases} \frac{f_2(x)}{f_2(k)} & \text{if } x \in [0,k], \\ \frac{f_1(x)}{f_1(k)} & \text{otherwise}. \end{cases}$$

Then, $f \in \mathcal{F}_2$.

Proof. Evidently, $f(k) = 1 = \frac{f_1(k)}{f_1(k)} = \frac{f_2(k)}{f_2(k)}$, thus f is continuous and strictly decreasing, and $f(1) = 0$. We only need to show the convexity of f. Note that both f_1, f_2 are convex, thus their left (right) derivatives exist at each point $x \in]0,1[$ and they are decreasing (not necessarily strictly decreasing). Moreover, $f'_{1-}(k) \le f'_{1+}(k) \le f'_{1-}(x)$ for all $x \in]k,1[$, and similarly for f_2. Consider $\frac{f'_{1-}(k)}{f_1(k)} \le \frac{f'_{2-}(k)}{f_2(k)}$. Then, $f'_-(x) = \frac{f'_{1-}(x)}{f_1(k)}$ for all $x \in]0,k]$ and $f'_-(x) = \frac{f'_{2-}(x)}{f_2(k)}$ for all $x \in]k,1[$. Consequently, f_- is decreasing. Similarly, f_+ is decreasing because $f'_+(x) = \frac{f'_{1+}(x)}{f_1(k)}$ for all $x \in]0,k[$ and $f'_+(x) = \frac{f'_{2+}(x)}{f_2(k)}$ for all $x \in [k,1[$. Hence, f is convex. The remaining case yields the convexity of f in a similar manner. □

In general, we observe that the operation $*_k$ acting on $\mathcal{F}_2$ is neither commutative nor associative. The next result is not difficult to confirm so its proof is omitted.

Proposition 7. *Let* $f_1, f_2 \in \mathcal{F}_2$ *generate the Archimedean copulas* $C_1, C_2 \in \mathcal{C}_2$. *Fix* $k \in]0,1[$, *and let* $f = f_1 *_k f_2$ *generate an Archimedean copula* C. *Suppose that* $\frac{f'_{1-}(k)}{f_1(k)} \le \frac{f'_{2-}(k)}{f_2(k)}$, *then:*

i) $C(x,y) = C_1(x,y)$ *for all* $(x,y) \in [0,k]^2$,
ii) $C(x,y) = C_2(x,y)$ *for all* $(x,y) \in [k,1]^2$ *such that* $f_2(x) + f_2(y) \le f_2(k)$.

Due to Proposition 7, copula C obtained by the gluing method $*_k$ can be viewed as the gluing of copulas C_1 and C_2 via an interpolation method. We observe that the positive multiplicative constants do not influence our gluing method, i.e., for each $c, d \in]0,\infty[$ and $f_1, f_2 \in \mathcal{F}_2$, $(cf_1) *_k (df_2) = f_1 *_k f_2$.

Example 6. Consider two basic Archimedean copulas $W, \Pi \in \mathcal{C}_2$ and their additive generators $f_W, f_\Pi \in \mathcal{F}_2$, $f_W(x) = 1 - x$, $f_\Pi(x) = -\log x$. For any fixed $k \in]0,1[$,

$$\frac{f'_W(k)}{f_W(k)} = \frac{-1}{1-k} \ge \frac{1}{k \log k} = \frac{f'_\Pi(k)}{f_\Pi(k)}.$$

Therefore, $f_k = f_W *_k f_\Pi$ is given by

$$f_k(x) = \begin{cases} \log_k(x) & \text{if } x \in [0,k], \\ \frac{1-x}{1-k} & \text{otherwise}. \end{cases}$$

The corresponding Archimedean copula $C_k \in \mathcal{C}_2$ is given by

$$C(x,y) = \begin{cases} xy & \text{if } (x,y) \in [0,k]^2, \\ x + y - 1 & \text{if } x + y > k + 1, \\ x \cdot k^{\frac{1-y}{1-k}} & \text{if } x \le k < y, \\ y \cdot k^{\frac{1-x}{1-k}} & \text{if } y \le k < x, \\ k^{\frac{2-x-y}{1-k}} & \text{otherwise}. \end{cases}$$

The family of Archimedean copulas $(C_k)_{k \in]0,1[}$ is continuous and strictly increasing in parameter k, with the limit members $C_0 = W$ and $C_1 = \Pi$.

5. Concluding remarks

In this study, we reviewed some previously reported methods for constructing additive generators of copulas (two-dimensional, n-dimensional, and for any dimension), including a method based on the Williamson transform. These methods are based on a priori knowledge of some additive generators, whereas we introduced a rather general construction method based on a given special real function h, which yields two-parameter families of additive generators. Moreover, we introduced a parametric family of methods for gluing two additive generators from $\mathcal{F}_2$ into a new additive generator from $\mathcal{F}_2$. We also illustrated these construction methods by providing several examples.

The connecting methods considered in this study are based on the aggregation of additive generators, as discussed previously [3], and they have high potential for fitting purposes when modeling the stochastic dependence structure of real data from many different domains, including finance, hydrology, sociology, economics, engineering, and information sciences.

Acknowledgement

The research was supported by the Slovak Research and Development Agency under projects APVV-0073-10 and APVV-0496-10, which is kindly acknowledged.

References

[1] T. Bacigál, Acopula: modelling dependence with multivariate Archimax, (or any user-defined continuous) copula. R package version 0.9.2, http://CRAN.R-project.org/package=acopula, 2013.
[2] T. Bacigál, M. Juráňová, R. Mesiar, On some new constructions of Archimedean copulas and applications to fitting problems, Neural Netw. World 20 (1) (2010) 81.
[3] T. Bacigál, R. Mesiar, V. Najjari, Generators of copulas and aggregation, Inf. Sci., submitted for publication.
[4] D.G. Clayton, A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, Biometrika 65 (1978) 141–151.
[5] F. Durante, C. Sempi, Copula and semicopula transforms, Int. J. Math. Sci. (2005) 645–655.
[6] M.J. Frank, On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$, Aequ. Math. 19 (1979) 194–226.
[7] C. Genest, K. Ghoudi, L.P. Rivest, "Understanding relationships using copulas," by Edward Frees and Emiliano Valdez, N. Am. Actuar. J. 2 (3) (1998) 143–149.
[8] E.J. Gumbel, Distributions des valeurs extrêmes en plusiers dimensions, Publ. Inst. Stat. Univ. Paris 9 (1960) 171–173.
[9] M. Hofert, I. Kojadinovic, M. Maechler, J. Yan, Copula: multivariate dependence with copulas, R package version 0.999-7, http://CRAN.R-project.org/package=copula, 2013.
[10] V. Jágr, M. Komorníková, R. Mesiar, Conditioning stable copulas, Neural Netw. World 20 (1) (2010) 69–79.
[11] H. Joe, Multivariate Models and Dependence Concepts, Chapman and Hall, London, 1997.
[12] M. Junker, A. May, Measurement of aggregate risk with copulas, Econom. J. 8 (3) (2005) 428–454.
[13] E.P. Klement, R. Mesiar, E. Pap, Triangular Norms, Kluwer Academic Publishers, Dordrecht, 2000.
[14] E.P. Klement, R. Mesiar, E. Pap, Transformations of copulas, Kybernetika 41 (4) (2005) 425–434.
[15] K. Menger, Statistical metrics, Proc. Natl. Acad. Sci. USA 28 (12) (1942) 535.
[16] R. Mesiar, V. Jágr, M. Juráňová, M. Komorníková, Univariate conditioning of copulas, Kybernetika 44 (6) (2008) 807–816.
[17] A.J. McNeil, J. Nešlehová, Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions, Ann. Stat. 37 (5B) (2009) 3059–3097.
[18] F. Michiels, A. De Schepper, Understanding copula transforms: a review of dependence properties, Working paper No. 2009012 2009.
[19] F. Michiels, A. De Schepper, How to improve the fit of Archimedean copulas by means of transforms, Stat. Pap. 53 (2) (2012) 345–355.
[20] R. Moynihan, On τ_T semigroups of probability distributions II, Aequ. Math. 17 (1978) 19–40.
[21] R.B. Nelsen, An Introduction to Copulas, Springer, New York, 2006.
[22] R.L. Plackett, A class of bivariate distributions, J. Am. Stat. Assoc. 60 (1965) 516–522.
[23] B. Schweizer, A. Sklar, Probabilistic Metric Spaces, North-Holland, 1983.
[24] A. Sklar, Fonctions de répartition à n dimensions et leurs marges, Publ. Inst. Stat. Univ. Paris 8 (1959) 229–231.
[25] A. Stupňanová, A. Kolesárová, Associative n-dimensional copulas, Kybernetika 47 (1) (2011) 93–99.
[26] R.E. Williamson, Multiply monotone functions and their Laplace transforms, Duke Math. J. 23 (1956) 189–207.
[27] R. Yager, An approach to inference in approximate reasoning, Int. J. Man-Mach. Stud. 13 (1980) 323–338.
[28] MathWorks Matlab documentation center: probability distributions used for multivariate modeling, http://www.mathworks.com/help/stats/probability-distributions-used-for-multivariate-modeling.html#bqttfgl-8.
[29] Wolfram mathematica documentation center: copula distribution, http://reference.wolfram.com/mathematica/ref/CopulaDistribution.html.

Information Sciences 306 (2015) 81–87

Contents lists available at ScienceDirect

Information Sciences

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

Generators of copulas and aggregation

Tomáš Bacigál [a], Radko Mesiar [a,b], Vadoud Najjari [c,*]

[a] Faculty of Civil Engineering, Slovak University of Technology in Bratislava, Slovakia
[b] Institute for Research and Application of Fuzzy Modeling, University of Ostrava, Czech Republic
[c] Department of Mathematics, Islamic Azad University, Maragheh Branch, Maragheh, Iran

ARTICLE INFO

Article history:
Received 27 June 2013
Received in revised form 18 December 2014
Accepted 31 January 2015
Available online 7 February 2015

2010 MSC:
62E

Keywords:
Additive generator
Aggregation function
Archimedean copula
Copula
Pseudo-inverse of an additive generator
Convexity

ABSTRACT

The most applied class of copulas for fitting purposes is undoubtedly the class of Archimedean copulas due to their representation by means of single functions of one variable, i.e., by means of additive generators or the corresponding pseudo-inverses. In this paper, we characterize aggregation functions preserving additive generators (pseudo-inverses of additive generators) of Archimedean copulas. As a by-product, we obtain an efficient method to construct new additive generators from some given ones.

1. Introduction

Since their proposal by Sklar [22], copulas became an important tool for the study and modelling problems dealing with random vectors. In this paper we will consider only bivariate random vectors, and hence only bivariate copulas. From the axiomatic point of view, copula is a function $C: [0,1]^2 \to [0,1]$ satisfying

(a) the boundary conditions $C(u,0) = C(0,v) = 0$ (C is grounded), $C(u,1) = u$, $C(1,v) = v$ (1 is neutral element of C)
(b) 2-increasing property $C(u,v) + C(u',v') - C(u,v') - C(u',v) \geqslant 0$ for all $u,v,u',v' \in [0,1]$, $u \leqslant u'$, $v \leqslant v'$.

From the statistical point of view, a copula $C: [0,1]^2 \to [0,1]$ is a function such that for any marginal distribution functions $F_X, F_Y: R \to [0,1]$ of random variables X and Y, the function $F_Z: R^2 \to [0,1]$ given by

$$F_Z(x,y) = C(F_X(x), F_Y(y)),$$

is a joint distribution function of some random vector $Z = (X,Y)$. Copula C describes here the dependence structure of the random vector Z. For more details we recommend [11,20].

* Corresponding author.
E-mail addresses: tomas.bacigal@stuba.sk (T. Bacigál), radko.mesiar@stuba.sk (R. Mesiar), vnajjari@gazi.edu.tr (V. Najjari).

http://dx.doi.org/10.1016/j.ins.2015.01.032
0020-0255/

A special class of copulas is characterized by associativity, $C(C(u,v),w) = C(u,C(v,w))$ and diagonal inequality $C(u,u) < u$ (for all $u \in]0,1[$). This class appearing first in the study of probabilistic metric spaces [21], is called the class of Archimedean copulas, and due to Moynihan [19] we have the next important characterization result.

Theorem 1 (Moynihan [19]). *A function $C : [0,1]^2 \to [0,1]$ is an Archimedean copula if and only if there is a continuous strictly decreasing convex function $f : [0,1] \to [0,\infty]$ satisfying $f(1) = 0$, called an additive generator, so that*

$$C(u,v) = g(f(u) + f(v)),$$

where $g : [0,\infty] \to [0,1]$ is given by $g(x) = f^{-1}(Min(f(0),x))$, and it is called a pseudo-inverse of f.

Note that pseudo-inverses were deeply discussed in [14], and for a decreasing non-constant function $h : [c,d] \to [a,b]$ the corresponding pseudo-inverse is given by

$$h^{(-1)}(x) = \sup\{t \in [a,b] | h(t) > x\},$$

with the convention $\sup \emptyset = a$. It is not difficult to check that, when considering the functions f, g from Theorem 1, then $g = f^{(-1)}$ and $f = g^{(-1)}$, i.e., the information contained in the additive generator f of an Archimedean copula C is the same as the information contained in its pseudo-inverse g, and that $C(u,v) = f^{(-1)}(f(u) + f(v)) = g(g^{(-1)}(u) + g^{(-1)}(v))$. In fact, in the literature both forms are used, the first one (based on an additive generator f) being preferred in probabilistic areas, while the second one (based on g) is more frequently used in the statistical literature.

We denote by $\mathcal{F}$ the class of all additive generators f, i.e., of all functions $f : [0,1] \to [0,\infty]$ which are continuous, strictly decreasing, convex and satisfying $f(1) = 0$. Observe that if $f \in \mathcal{F}$ then also $cf \in \mathcal{F}$ for any positive constant $c \in]0,\infty[$, and that if an Archimedean copula C is generated by two additive generators f_1 and f_2, then necessarily $f_1 = cf_2$ for some $c \in]0,\infty[$.

We denote by $\mathcal{G}$ the class of all pseudo-inverses g of additive generators $f \in \mathcal{F}$.

Lemma 1. *A function $g : [0,\infty] \to [0,1]$ belongs to $\mathcal{G}$ if and only if it is continuous, convex, $g(0) = 1$, and there is a constant $a \in]0,\infty]$ such that g is strictly decreasing on $[0,a]$, and $g(x) = 0$ for all $x \in [a,\infty]$.*

Proof. Suppose $g \in \mathcal{G}$, i.e., there is $f \in \mathcal{F}$ such that $g = f^{(-1)}$. Then clearly $g(0) = 1$, and due to results from [14], g is continuous and decreasing. Moreover, denoting $a = f(0)$, g is vanishing on $[a,\infty]$. Then the convexity of g on $[0,\infty]$ is guaranteed by the convexity of g on $[0,a]$. Finally, the function $h : [0,a] \to [0,1]$, $h(x) = g(x)$, is the inverse of f, $h = f^{-1}$ and thus strictly decreasing and convex, concluding the necessity part of this lemma. To see the sufficiency, it is enough to consider $f = h^{-1}$, where $h : [0,a] \to [0,1]$ is given as above, $h(x) = g(x)$. Then evidently $f \in \mathcal{F}$ and $g = f^{(-1)}$, i.e., $g \in \mathcal{G}$. □

Observe that Lemma 1 allows to consider the class $\mathcal{G}$ independently of the class $\mathcal{F}$.

The aim of this paper is to study construction methods for additive generators (or their pseudo-inverses) of Archimedean copulas from some a priori given additive generators (pseudo-inverses) by means of aggregation functions. In the next section, we characterize aggregation functions preserving the class $\mathcal{F}$ of all additive generators of Archimedean copulas. Section 3 brings a characterization of aggregation functions preserving the class $\mathcal{G}$ of all pseudo-inverses of additive generators of Archimedean copulas. Finally, some concluding remarks are given.

2. Aggregation of additive generators of Archimedean copulas

Note, first of all, that the class $\mathcal{F}$ is convex, i.e., for any $f_1,\ldots,f_n \in \mathcal{F}$ and $c_1,\ldots,c_n \in [0,1]$, $\sum_{i=1}^n c_i = 1$, also $f = \sum_{i=1}^n c_i f_i \in \mathcal{F}$. Due to the already mentioned fact that any positive multiple cf of an additive generator $f \in \mathcal{F}$ is again an additive generator, $cf \in \mathcal{F}$, we see that one can relax the constraint $\sum_{i=1}^n c_i = 1$ into $\sum_{i=1}^n c_i > 0$, i.e., any non-trivial non-negative linear combination of additive generators from $\mathcal{F}$ is again an element of $\mathcal{F}$. For more details and examples see [1].

Recall that, for $n \in \{2,3,\ldots\}$, an aggregation function $A : [a,b]^n \to [a,b]$ is characterized by the increasing monotonicity in each coordinate and by boundary conditions $A(a,\ldots,a) = a$ and $A(b,\ldots,b) = b$. For more details we recommend [5,8,3], see also [9,10].

To aggregate additive generators $f_1,\ldots,f_n \in \mathcal{F}$ into an additive generator $f \in \mathcal{F}$, $f = A(f_1,\ldots,f_n)$, i.e., for all $x \in [0,1]$, $f(x) = A(f_1(x),\ldots,f_n(x))$, obviously one should consider the interval $[a,b] = [0,\infty]$ for inputs/output domain of considered values, i.e., we look for appropriate aggregation functions $A : [0,\infty]^n \to [0,\infty]$. It is immediate that due to the boundary condition for additive generators and agg. functions, $f(1) = A(f_1(1),\ldots,f_n(1)) = A(0,\ldots,0) = 0$, independently of A and $f_1,\ldots,f_n \in \mathcal{F}$. To ensure the continuity of f, A should be continuous. Similarly, to ensure the strict monotonicity of f, A should be jointly strictly increasing, i.e., $A(\mathbf{x}) < A(\mathbf{y})$ whenever $\mathbf{x},\mathbf{y} \in [0,\infty]^n$ and $x_i < y_i, i = 1,\ldots,n$. To ensure $f \in \mathcal{F}$ we have to ensure the convexity of f.

The next theorem gives a complete characterization of aggregation functions preserving the class $\mathcal{F}$ of all additive generators of copulas.

Theorem 2. *Let $n \in \{2, 3, \ldots\}$ and $A: [0,\infty]^n \to [0,\infty]$ be an aggregation function. Then the following are equivalent:*

(i) *for any $f_1, \ldots, f_n \in \mathcal{F}$ also $f \in \mathcal{F}$, where $f = A(f_1, \ldots, f_n)$*
(ii) *A is a continuous jointly strictly increasing aggregation function satisfying, for all $\mathbf{x}, \mathbf{y} \in]0,\infty[^n$,*

$$2A\left(\mathbf{x} + \frac{\mathbf{y}}{2}\right) \leqslant A(\mathbf{x}) + A(\mathbf{x} + \mathbf{y}). \tag{1}$$

Proof. We have to show only the fact that the convexity of $f = A(f_1, \ldots, f_n)$, for any $f_1, \ldots, f_n \in \mathcal{F}$, is equivalent to the fact that inequality (1) is satisfied for all $\mathbf{x}, \mathbf{y} \in]0,\infty[^n$, and thus, due to the continuity of A, also to all $\mathbf{x}, \mathbf{y} \in [0,\infty]^n$. Observe that our condition (1) is weaker than the standard (Jensen) convexity.

Suppose the inequality (1) holds true and consider any points $x, y \in]0,1[, x > y$. The convexity of f is equivalent to its Jensen's convexity, which is valid because of

$$\begin{aligned} f\left(\frac{x+y}{2}\right) &= A\left(f_1\left(\frac{x+y}{2}\right), \ldots, f_n\left(\frac{x+y}{2}\right)\right) \leqslant A\left(\frac{f_1(x)+f_1(y)}{2}, \ldots, \frac{f_n(x)+f_n(y)}{2}\right) \\ &= A\left(f_1(x) + \frac{f_1(y)-f_1(x)}{2}, \ldots, f_n(x) + \frac{f_n(y)-f_n(x)}{2}\right) \leqslant \frac{A(f_1(x), \ldots, f_n(x)) + A(f_1(y), \ldots, f_n(y))}{2} = \frac{f(x)+f(y)}{2}, \end{aligned}$$

considering $\mathbf{x} = (f_1(x), \ldots, f_n(x))$, $\mathbf{y} = \left(\frac{f_1(y)-f_1(x)}{2}, \ldots, \frac{f_n(y)-f_n(x)}{2}\right) \in]0,\infty[^n$, where the first inequality follows from the monotonicity of A and Jensen's convexity of $f_1, \ldots, f_n$.

On the other hand, consider that (i) holds true. For a fixed couple $\mathbf{x}, \mathbf{y} \in]0,\infty[^n$, denote $v = \frac{Min\left(\frac{y_1}{x_1}, \ldots, \frac{y_n}{x_n}\right)}{1 + Min\left(\frac{y_1}{x_1}, \ldots, \frac{y_n}{x_n}\right)}$. Clearly, $v \in]0,1[$. For $i = 1, \ldots, n$, define $f_i: [0,1] \to [0,\infty]$ by

$$f_i(x) = \begin{cases} \frac{x_i}{1-v}(1-x) & \text{if } x \in [v,1], \\ x_i + 2y_i - \frac{2y_i}{v}x & \text{elsewhere.} \end{cases}$$

Observe that each f_i is a piecewise linear function with slopes $\frac{-x_i}{1-v}$ on $[v,1]$ and $\frac{-2y_i}{v}$ on $[0,v]$. Moreover

$$\frac{-x_i}{1-v} - \left(\frac{-2y_i}{v}\right) = \left(1 + Min\left(\frac{y_1}{x_1}, \ldots, \frac{y_n}{x_n}\right)\right)\left(\frac{2y_i}{Min\left(\frac{y_1}{x_1}, \ldots, \frac{y_n}{x_n}\right)} - x_i\right) \geqslant x_i > 0,$$

and thus each f_i is an additive generator from $\mathcal{F}$. Then also $f = A(f_1, \ldots, f_n) \in \mathcal{F}$, and hence necessarily

$$2f\left(\frac{3}{4}v\right) \leqslant f\left(\frac{v}{2}\right) + f(v).$$

Note that, for $i = 1, \ldots, n$,

$$f_i(v) = x_i,$$
$$f_i\left(\frac{v}{2}\right) = x_i + y_i,$$
$$f_i\left(\frac{3}{4}v\right) = x_i + \frac{y_i}{2}.$$

Consequently,

$$2f\left(\frac{3}{4}v\right) = 2A\left(f_1\left(\frac{3}{4}v\right), \ldots, f_n\left(\frac{3}{4}v\right)\right) = 2A\left(\mathbf{x} + \frac{\mathbf{y}}{2}\right) \leqslant A(\mathbf{x}) + A(\mathbf{x}+\mathbf{y}) = f(v) + f\left(\frac{v}{2}\right),$$

providing the validity of (ii). □

Remark 1. Theorem 2 deals with the presentation of convexity of composite functions. Evidently, for any convex function $F: [0,\infty]^n \to [0,\infty]$ and convex functions $f_1, \ldots, f_n: [0,1] \to [0,\infty]$, also the function $f: [0,1] \to [0,\infty]$ given by $f(x) = F(f_1(x), \ldots, f_n(x))$ is convex. Note that the convexity of F means that $F(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leqslant \lambda F(\mathbf{x}) + (1-\lambda)F(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in [0,\infty]^n$ and $\lambda \in [0,1]$ (or equivalently, $2F(\frac{\mathbf{x}+\mathbf{y}}{2}) \leqslant F(\mathbf{x}) + F(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in [0,\infty]^n$, i.e., convexity coincides with the Jensen convexity). However, convexity of F is sufficient but not necessary, see [4]. As a typical example it is enough to recall the *Max* function on $[0,1]^n$, which is not convex but it preserves the convexity of inner functions $f_1, \ldots, f_n$, i.e., $f = Max(f_1, \ldots, f_n)$, is convex whenever $f_1, \ldots, f_n$ are convex.

Another sufficient but not necessary condition on F to preserve the convexity is the directional convexity [17,6], known also as ultramodularity [16,13]. Recall that $F: [0,\infty]^n \to [0,\infty]$ is ultramodular (directionally convex) if and only if for each $\mathbf{x},\mathbf{y},\mathbf{z} \in [0,\infty]^n$ it satisfies the inequality

$$F(\mathbf{x}+\mathbf{y}+\mathbf{z}) + F(\mathbf{x}) \geqslant F(\mathbf{x}+\mathbf{y}) + F(\mathbf{x}+\mathbf{z}).$$

Again, *Max* is not ultramodular.

Our sufficient and necessary condition (1) for a continuous jointly strictly monotone aggregation function A to preserve the class $\mathcal{F}$ of all additive generators of 2-copulas can be seen as a weaker form of both above types of convexity and it can be called an *ordered convexity*. Indeed, the inequality (1) can be seen as the Jensen convexity $2F(\frac{\mathbf{u}+\mathbf{v}}{2}) \leqslant F(\mathbf{u}) + F(\mathbf{v})$ valid for all ordered (i.e., comparable) pairs $\mathbf{u},\mathbf{v} \in [0,\infty]^n$ (and then (1) is obtained putting $\mathbf{x} = Min(\mathbf{u},\mathbf{v})$ and $\mathbf{y} = |\mathbf{u}-\mathbf{v}|$).

Recall also the link between additive generators of copulas and distribution functions of positive random variable through the Williamson transform (Laplace transform) as given in [18,12]. Then one can apply any aggregation preserving such distribution functions (e.g., mixtures). However, such an approach is out of the framework of this paper, where the point-wise aggregation of additive generators (or their pseudo-inverses) are considered.

Example 1

(i) Consider $A(x_1,\ldots,x_n) = Min(x_1,\ldots,x_n)$ which is continuous, jointly strictly increasing aggregation function on $[0,\infty]$. However, fix $x_2 = \ldots = x_n = 1$. Then $A(x,1,\ldots,1) = Min(x,1)$ is not a convex function (indeed, it is concave piecewise linear function). Define $f_1, f_2 \in \mathcal{F}$ by

$$f_1(x) = \begin{cases} 1-2x & \text{if } x \in [0,\frac{1}{3}], \\ \frac{1-x}{2} & \text{elsewhere}, \end{cases}$$

$$f_2(x) = \begin{cases} 2-9x & \text{if } x \in [0,\frac{1}{6}], \\ \frac{3(1-x)}{5} & \text{elsewhere}. \end{cases}$$

Then $f = Min(f_1, f_2)$ is given by

$$f(x) = \begin{bmatrix} 1-2x & \text{if } x \in [0,\frac{1}{7}], \\ 2-9x & \text{if } x \in [\frac{1}{7},\frac{1}{6}], \\ \frac{3(1-x)}{5} & \text{if } x \in [\frac{1}{6},\frac{2}{7}], \\ 1-2x & \text{if } x \in [\frac{2}{7},\frac{1}{3}], \\ \frac{1-x}{2} & \text{if } x \in [\frac{1}{3},1], \end{bmatrix}$$

which clearly is not convex and thus $f \notin \mathcal{F}$.

(ii) As a positive example of aggregation functions satisfying the constraints of Theorem 1 we can consider:

- The weighted sum $A(x_1,\ldots,x_n) = \sum_{i=1}^n c_i x_i, c_i \geqslant 0, \sum_{i=1}^n c_i > 0$ (see also the discussion at the beginning of this section);
- the product $A(x_1,\ldots,x_n) = \prod_{i=1}^n x_i$;
- the maximum $A(x_1,\ldots,x_n) = Max(x_1,\ldots,x_n)$;
- the p-sum $A(x_1,\ldots,x_n) = \left(\sum_{i=1}^n x_i^p\right)^{1/p}, \ p \in [1,\infty[$.

Note that each composition of aggregation functions (including the discussed one-dimensional transformation φ) satisfying Theorem 1 will again satisfy all constraints of Theorem 1. For example, consider $A(x_1,x_2,x_3) = (x_1^4 + x_2^2 + x_3^3)^{1/2}$, or $A(x_1,\ldots,x_5) = Max(x_1^2 x_3^3, 2x_2 + 3x_4 x_5^2)$. Moreover, if A satisfying the constraints of Theorem 1 has 0 as its annihilator, i.e., if $x_i = 0$ for some $i \in \{1,\ldots,n\}$ then $A(x_1,\ldots,x_n) = 0$, also new aggregation given by $A(x_1+c_1,\ldots,x_n+c_n) = 0$, where $c_i \geqslant 0$ and $\prod_{i=1}^n c_i = 0$, will also satisfy the constraints of Theorem 1. As an example one can consider $A(x_1,x_2) = x_1(x_2+1)$.

3. Aggregation of pseudo-inverses from $\mathcal{G}$

Observe first that two pseudo-inverses $g_1, g_2 \in \mathcal{G}$ generate the same Archimedean copula if and only if $g_1(x) = g_2(cx)$ for some constant $c \in]0,\infty[$. Similarly to class $\mathcal{F}$, also the class $\mathcal{G}$ is convex, see [1]. However, one should stress that a non-trivial convex combination of pseudo-inverses $g_1,\ldots,g_n \in \mathcal{G}$ related to a given Archimedean copula C yields a pseudo-inverse $g \in \mathcal{G}$ linked to some different copula D (contradicting the related convex combination of additive generators).

Example 2. Consider $c,d \in]0,\infty[, c > d$. Then the pseudo-inverses $g_1, g_2 \in \mathcal{G}$ given by $g_1(x) = Max(1-cx, 0)$ and $g_2(x) = Max(1-dx, 0)$ generates the smallest copula $W, W(u,v) = Max(0, u+v-1)$. The arithmetic mean $g = \frac{g_1+g_2}{2} \in \mathcal{G}$ is given by

$$g(x) = \begin{cases} 1 - \frac{c+d}{2}x & \text{if } x \in [0, \frac{1}{c}], \\ \frac{1-dx}{2} & \text{if } x \in [\frac{1}{c}, \frac{1}{d}], \\ 0 & \text{elsewhere}, \end{cases}$$

and it generates an Archimedean copula $C \neq W$. Obviously, for each $\alpha \in]0, \infty[$, the pair $c' = \alpha c,\ d' = \alpha d$ generates the same copula. For a fixed parameter $\lambda = \frac{c}{d} \in]1, \infty[$ we can then introduce a parametric family $(g_\lambda)_{\lambda \in]1,\infty[} \subset \mathcal{G}$ given by

$$g_\lambda(x) = Max\left(1 - \frac{1+\lambda}{2}x, \frac{1-x}{2}, 0\right).$$

Note that the corresponding family $(f_\lambda)_{\lambda \in]1,\infty[} \subset \mathcal{F}$ of additive generators is given by

$$f_y(x) = Max\left(1 - 2x, \frac{2}{1+\lambda}(1-x)\right).$$

To aggregate functions $g_1, \ldots, g_n \in \mathcal{G}$, we should consider aggregation functions $A: [0,1]^n \to [0,1]$. Evidently, the function $g = A(g_1, \ldots, g_n)$ satisfies $g(0) = A(g_1(0), \ldots, g_n(0)) = A(1, \ldots, 1) = 1$.

The monotonicity of pseudo-inverses $g \in \mathcal{G}$, i.e., the fact that it is strictly decreasing on $[0, a]$ and that it vanishes on $[a, \infty]$, with $a = Min(x \in]0, \infty] | g(x) = \infty)$ (observe that due to $g(\infty) = 0$ and continuity of g, a is well defined), is preserved by an aggregation function A if and only if $A(\mathbf{x}) > A(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in [0,1]^n$ such that $x_1 \geqslant y_1, \ldots, x_n \geqslant y_n,\ A(\mathbf{y}) > 0$, and if $y_i > 0$ then $x_i \neq y_i, i \in \{1, \ldots, n\}$. We will call this property *weak joint strict increasingness* of A. Obviously, each jointly strictly increasing aggregation function A is also weak jointly strictly increasing. Again, only the preservation of convexity of elements of $\mathcal{G}$ remains to be guaranteed.

Theorem 3. *Let $n \in \{2, 3, \ldots\}$ be fixed and let $A: [0,1]^n \to [0,1]$ be an aggregation function. Then the following are equivalent.*

(i) *for any $g_1, \ldots, g_n \in \mathcal{G}$, also $g \in \mathcal{G}$, where $g = A(g_1, \ldots, g_n)$;*
(ii) *A is continuous, weak jointly strictly increasing, and for all $\mathbf{x}, \mathbf{y} \in [0,1[^n$, $\mathbf{x} \geqslant \mathbf{y}$ such that if $x_i = y_i$ then $x_i = y_i = 0, i \in \{1, \ldots, n\}$, then*

$$A\left(\frac{\mathbf{x}+\mathbf{y}}{2}\right) \leqslant \frac{A(\mathbf{x}) + A(\mathbf{y})}{2}. \tag{2}$$

Proof. Due to the above discussion, only the convexity issues should be shown. Evidently, (2) ensures the convexity of g, independently of $g_1, \ldots, g_n \in \mathcal{G}$. On the other side, suppose (i) holds true, and consider $\mathbf{x}, \mathbf{y} \in [0,1[^n$, $\mathbf{x} \geqslant \mathbf{y}$ such that $x_i = y_i$ only if $x_i = y_i = 0, i \in \{1, \ldots, n\}$. Obviously, if $\mathbf{x} = \mathbf{y} = \mathbf{0}$, (2) holds. Hence suppose $\mathbf{x} \neq \mathbf{0}$ and put $v = 2 \max \frac{1-y_i}{1-x_i} > 1$.

If $x_i = y_i = 0$, we define $g_i \in \mathcal{G}$ by $g_i(x) = Max(0, 1-x)$. In the remaining cases, we introduce $g_i: [0, \infty] \to 0$ by

$$g_i(x) = Max\left(0, 1 - (1-x_i)x, \frac{x_i v - y_i - (x_i - y_i)x}{v-1}\right).$$

The slope of g on $[0,1]$ is $-(1-x_i)$, on $\left[1, \frac{x_i v - y_i}{x_i - y_i}\right]$ it is $-\frac{x_i - y_i}{v-1}$, and on $\left[\frac{x_i v - y_i}{x_i - y_i}, \infty\right]$, it is 0. Then the convexity of g_i is equivalent to the validity of the inequality

$$-(1-x_i) \leqslant -\frac{x_i - y_i}{v-1},$$

which is valid for each $v \geqslant \frac{1-y_i}{1-x_i}$. Hence $g_i \in \mathcal{G}$. Moreover $g_i(1) = x_i,\ g_i(v) = y_i$ and $g_i(\frac{v+1}{2}) = \frac{x_i+y_i}{2}$. Due to the validity of (i), and the convexity of $g \in \mathcal{G}$, it holds

$$A\left(\frac{\mathbf{x}+\mathbf{y}}{2}\right) = A\left(\frac{x_1+y_1}{2}, \ldots, \frac{x_n+y_n}{2}\right) = A\left(g_1\left(\frac{v+1}{2}\right), \ldots, g_n\left(\frac{v+1}{2}\right)\right) = g\left(\frac{v+1}{2}\right) \leqslant \frac{g(v)+g(1)}{2} = \frac{A(\mathbf{x})+A(\mathbf{y})}{2},$$

i.e., (2) is satisfied. □

Observe that due to the continuity of A, property (2) can be seen as an ordered convexity discussed after Theorem 2 (now on the $[0,1]$ scale). Due to the relaxed joint strict monotonicity in (ii) of Theorem 2, one can construct aggregation functions preserving pseudo-inverses of additive generators of Archimedean copulas by means of aggregation functions preserving additive generators of Archimedean copulas, but not vice versa.

Proposition 1. *Let $A: [0, \infty]^n \to [0, \infty]$ be an aggregation function satisfying the constraints of Theorem 2. Then $f = A(1, \ldots, 1) \in]0, \infty[$, and for any $c \in [0, b[$, the function $C: [0,1]^n \to [0,1]$ given by*

$$C(\mathbf{x}) = \frac{Max(0, A(\mathbf{x}) - c)}{b - c} \tag{3}$$

is an aggregation function satisfying all constraints of Theorem 3.

Proof. The only non-trivial problem is to show the ordered convexity of C. However, this is guaranteed due to the fact that both constant function 0 and $A - c$ are ordered convex, and max operator preserves the ordered convexity. □

Example 3

(i) Consider $A: [0,\infty]^2 \to [0,\infty]$, $A(x,y) = x + y$. Obviously, A satisfies the constraints of Theorem 2. Then, for any $c \in [0, 2[$, also the function $C: [0,1]^2 \to [0,1]$ given by

$$C(x,y) = Max\left(0, \frac{x+y-c}{2-c}\right),$$

satisfies $C(g_1, g_2) \in \mathcal{G}$ for any $g_1, g_2 \in \mathcal{G}$. In particular, for $c = 1$, $C(x,y) = Max(0, x+y-1)$ is well known Fréchet–Hoeffding lower bound of copulas, i.e., $C = W$.

(ii) For $A(x,y) = xy$, the corresponding aggregation functions preserving the class $\mathcal{G}$ are given by

$$C(x,y) = Max\left(0, \frac{xy-c}{1-c}\right), \quad c \in [0, 1[,$$

(obviously, $c = 0$ gives the standard product).

Similarly, for 1-dimensional functions, a function $\varphi: [0,1] \to [0,1]$ preserves the class $\mathcal{G}$, $\varphi(g) \in \mathcal{G}$ for any $g \in \mathcal{G}$, if and only if φ is increasing convex surjection which is constant only on the pre-image of 0, i.e., $\varphi(x) = \varphi(y)$ implies $x = y$ or $\varphi(x) = \varphi(y) = 0$. Then the construction method given in Proposition 1 can be seen as a composition of $\varphi: [0,1] \to [0,1]$ given by $\varphi(x) = Max(0, \frac{x-\alpha}{1-\alpha})$ and $B: [0,1]^n \to [0,1]$ given by $B(\mathbf{x}) = \frac{A(\mathbf{x})}{b}$, where $\alpha = \frac{c}{b} \in [0, 1[$, $C = \varphi \circ B$.

4. Concluding remarks

We have completely characterized aggregation functions preserving the classes $\mathcal{F}$ and $\mathcal{G}$ of additive generators of Archimedean copulas and their pseudo-inverses, vice versa. Observe that though in both cases the crucial problem is the preserving of convexity, there are several important differences: due to different ranges, types of monotonicity and relationship of additive generators (pseudo-inverses) generating the same Archimedean copula C. Our results extends the buffer of potential copulas for fitting purposes. Observe, for example, approach of [2] based on fitting piecewise linear additive generators. This class has as its counterpart the class of piecewise linear pseudo-inverses from $\mathcal{G}$. Any such function g is determined by points $(v_i, x_i), i = 1, \ldots, n$, such that $x_n = 0$, and $\left(\frac{x_i - x_{i-1}}{v_i - v_{i-1}}\right)_{i=1}^{n}$ is an increasing sequence (to ensure the convexity of g), where $(v_0, x_0) = (0, 1)$. It is not difficult to check that then $g = \sum_{i=1}^{n} c_i g_i$ is a convex combination of pseudo-inverses $g_1, \ldots, g_n$, all of them generating the smallest copula W, given by $g_i(x) = Max(0, 1 - \frac{x}{v_i})$.

Moreover, the corresponding additive generator $f \in \mathcal{F}$ is a piecewise linear function determined by points $(1, 0), (x_1, v_1), \ldots, (x_{n-1}, v_{n-1}), (0, v_n)$.

In majority of fitting problems, special families of Archimedean copulas, such as Gumbel, Clayton, and Frank, are considered (see, e.g., [7] among first applications, and [15] providing also software tools). Our approach can improve the fitting power of existing tools, simply considering appropriate aggregation of the best additive generators (pseudo-inverses) from considered families. To avoid the possible discrepancies caused by non-uniqueness in the relationship of generators and copulas, we can consider additive generators constrained by $f(\frac{1}{2}) = 1$, and in the case of pseudo-inverses by $g(1) = \frac{1}{2}$. Then there is a one-to-one correspondence between Archimedean copulas and additive generators (pseudo-inverses of additive generators). Obviously, to preserve such additive generators by means of an aggregation function $A: [0,\infty]^n \to [0,\infty]$, one should consider $A(1, \ldots, 1) = 1$, i.e., 1 should be an idempotent element of A. Then, for A satisfying the constraints of Theorem 2, it is enough to consider $A^* = \frac{A}{A(1,\ldots,1)}$. Typical examples: product $A(x_1, \ldots, x_n) = \prod_{i=1}^{n} x_i$, maximum $A(x_1, \ldots, x_n) = Max(x_1, \ldots, x_n)$, weighted p-sum $A(x_1, \ldots, x_n) = \left(\sum_{i=1}^{n} c_i x_i^p\right)^{\frac{1}{p}}$ with weights $c_i, \sum_{i=1}^{n} c_i = 1$.

If, for example, Clayton's best fitting copula to modelled data is related to (constrained) additive generator f_1, and from Gumbel family we have obtained additive generator f_2, we can introduce a 1-parameter family $h_\lambda = \sqrt{\lambda f_1^2 + (1-\lambda) f_2^2}, \lambda \in [0,1]$ and check its best fitting member. Our result cannot be worse than the original one. A similar consideration for pseudo-inverses requires to consider aggregation functions $A: [0,1]^n \to [0,1]$ satisfying the constraints of Theorem 3, and such appropriate aggregation function should be necessarily idempotent, $A(x, \ldots, x) = x$ for each $x \in [0,1]$. However, one can consider an arbitrary A fitting Theorem 3, and then "normalize" the output of pseudo-inverse.

Acknowledgements

The support of APVV-0073-10 and APVV-0496-10 is kindly acknowledged. We also thank two anonymous reviewers for their valuable suggestions.

References

[1] T. Bacigál, M. Juráňová, R. Mesiar, On some new constructions of Archimedean copulas and applications to fitting problems, Neural Netw. World 20 (1) (2010) 81.
[2] G. Beliakov, T. Calvo, Construction of k-Lipschitz triangular norms and conorms from empirical data, IEEE Trans. Fuzzy Syst. 17 (5) (2009) 1217–1220.
[3] G. Beliakov, A. Pradera, T. Calvo, Aggregation Functions: A Guide for Practitioners, Springer, Berlin, 2007.
[4] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, Cambridge, 2004.
[5] T. Calvo, A. Kolesárová, M. Komorníková, R. Mesiar, Aggregation operators: properties, classes and construction methods, in: Aggregation Operators, Physica-Verlag, Heidelberg, 2002, pp. 3–104.
[6] M.M. Denuit, M. Mesfioui, Generalized increasing convex and directionally convex orders, J. Appl. Probab. 47 (1) (2010) 264–276.
[7] E.W. Frees, E.A. Valdez, Understanding relationships using copulas, North Am. Actuarial J. 2 (1) (1998) 1–25.
[8] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, Aggregation Functions, Encyclopedia of Mathematics and its Applications, vol. 127, Cambridge University Press, 2009.
[9] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, Aggregation functions: means, Inf. Sci. 181 (2011) 1–22.
[10] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, Aggregation functions: construction methods, conjunctive, disjunctive and mixed classes, Inf. Sci. 181 (2011) 23–43.
[11] H. Joe, Multivariate Models and Dependence Concepts, Chapman and Hall, London, 1997.
[12] C.H. Kimberling, A probabilistic interpretation of complete monotonicity, Aequationes Math. 10 (2) (1974) 152–164.
[13] E.P. Klement, M. Manzi, R. Mesiar, Ultramodular aggregation functions, Inf. Sci. 181 (2011) 4101–4111.
[14] E.P. Klement, R. Mesiar, E. Pap, Quasi-and pseudo-inverses of monotone functions, and the construction of t-norms, Fuzzy Sets Syst. 104 (1) (1999) 3–13.
[15] I. Kojadinovic, J. Yan, Modeling multivariate distributions with continuous margins using the copula R package, J. Stat. Softw. 34 (9) (2010) 1–20.
[16] M. Marinacci, L. Montrucchio, Ultramodular functions, Math. Oper. Res. 30 (2) (2005) 311–332.
[17] A.W. Marshall, I. Olkin, B.C. Arnold, Inequalities: Theory of Majorization and its Applications, Springer Series in Statistics, second ed., Springer, New York, 2011.
[18] A.J. McNeil, J. Nešlehová, Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions, Ann. Stat. 37 (5B) (2009) 3059–3097.
[19] R. Moynihan, Infinite τ_T products of distribution functions, J. Austral. Math. Soc. Ser. A 26 (1978) 227–240.
[20] R.B. Nelsen, An Introduction to Copulas, Springer, New York, 2006.
[21] B. Schweizer, A. Sklar, Probabilistic Metric Spaces, Courier Dover Publications, 1983.
[22] A. Sklar, Fonctions de répartition à n dimensions et leurs marges, 8, Publ. Inst. Statist. Univ. Paris, 1959. pp. 229–231.

DE GRUYTER OPEN

TATRA MOUNTAINS
Mathematical Publications
DOI: 10.1515/tmmp-2017-0000
Tatra Mt. Math. Publ. **69** (2017), 1–18

CONVERGENCE OF LINEAR APPROXIMATION OF ARCHIMEDEAN GENERATOR FROM WILLIAMSON'S TRANSFORM IN EXAMPLES

Tomáš Bacigál — Mária Ždímalová

ABSTRACT. We discuss a new construction method for obtaining additive generators of Archimedean copulas proposed by McNeil, A. J.—Nešlehová, J.: *Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions,* Ann. Statist. **37** (2009), 3059–3097, the so-called Williamson n-transform, and illustrate it by several examples. We show that due to the equivalence of convergences of positive distance functions, additive generators and copulas, we may approximate any n-dimensional Archimedean copula by an Archimedean copula generated by a transformation of weighted sum of Dirac functions concentrated in certain suitable points. Specifically, in two dimensional case this means that any Archimedean copula can be approximated by a piece-wise linear Archimedean copula, moreover the approximation of generator by linear splines circumvents the problem with the non-existence of explicit inverse.

1. Introduction

Copulas form an important class of multivariate dependence models. They have a lot of practical applications, including multivariate survival modelling. Recall that copulas aggregate 1-dimensional marginal distribution functions into n-dimensional ($n \geq 2$) joint distribution functions. For more details we recommend [13].

We first define a copula. A function $C\colon [0,1]^n \to [0,1]$ is called a (n-dimensional) copula whenever it satisfies the boundary conditions (C1) and it is an n-increasing function, see (C2), where:

2010 Mathematics Subject Classification: 60E05.
Keywords: Archimedean copula, Williamson, approximation.
The work on this paper was supported by grant APVV-14-0013.

(C1) $C(x_1,\dots,x_n)=0$ whenever $0\in\{x_1,\dots,x_n\}$, i.e., 0 is an annihilator of C, and $C(x_1,\dots,x_n)=x_i$ whenever $x_j=1$ for each $j\neq i$ (i.e., 1 is a neutral element of C).

(C2) For any $\mathbf{x},\mathbf{y}\in[0,1]^n$, $\mathbf{x}\le\mathbf{y}$, it holds

$$V_C([\mathbf{x},\mathbf{y}]) = \sum_{\varepsilon\in\{-1,1\}^n}\left(C(\mathbf{z}_\varepsilon)\prod_{i=1}^n \varepsilon_i\right)\ge 0,$$

where $\mathbf{z}_\varepsilon=\left(z_1^{\varepsilon_1},\dots,z_n^{\varepsilon_n}\right)$, $z_i^1=y_i$, $z_i^{-1}=x_i$.

Note that $V_C([\mathbf{x},\mathbf{y}])$ is called the C-volume of the n-dimensional interval (n-box) $[\mathbf{x},\mathbf{y}]$.

Due to Sklar's theorem [16] for a random vector $Z=(X_1,\dots,X_n)$, a function $F_Z\colon R^n\to[0,1]$ is a joint distribution function of Z if and only if there is a copula $C\colon[0,1]^n\to[0,1]$ so that

$$F_Z(x_1,\dots,x_n)=C\big(F_{X_1}(x_1),\dots,F_{X_n}(x_n)\big), \tag{1}$$

where $F_{X_i}\colon R\to[0,1]$ is a distribution function related to the random variable X_i, $i=1,\dots,n$. The copula C in (1) is unique whenever random variables $X_1,\dots,X_n$ are continuous. For some other details on copulas see [5] and [13].

Hereafter we will consider a class of copulas named Archimedean copulas. In the simplest case, Archimedean 2-copulas are characterized by the associativity of C and the diagonal inequality $C(x,x)<x$ for all $x\in]0,1[$. They are necessarily symmetric, i.e., they can model the stochastic dependence of exchangeable random variables (X,Y) only, yet their popularity in practice (hydrology, financial, and other applied areas) is indisputable, mainly due to the representation using one-dimensional functions called generators as shown in the next result, attributed to Moynihan [12].

Theorem 1. *A function $C\colon[0,1]^2\to[0,1]$ is an Archimedean copula if and only if there is a convex (i.e., a 2-monotone) continuous strictly decreasing function $f\colon[0,1]\to[0,\infty]$, $f(1)=0$, so that*

$$C(x,y)=f^{(-1)}\big(f(x)+f(y)\big), \tag{2}$$

where the pseudo-inverse $f^{(-1)}\colon[0,\infty]\to[0,1]$ is given by

$$f^{(-1)}(u)=f^{-1}\Big(\min\big(u,f(0)\big)\Big).$$

The function f is called an additive generator of the copula C, and it is unique up to a positive multiplicative constant.

Let $\mathcal{F}_2$ be the class of all additive generators of binary copulas characterized in the above theorem. More details about the generators can be found in [5], [6], [13] and about construction methods for additive generators in [1], [2], [4], [7], [11].

Before we review several known facts for additive generators of copulas, let us briefly recall a link between copula C and Spearman's correlation coefficient ρ,

$$\rho = 12E[UV] - 3 = 12\iint\limits_{[0,1]^2} uv dC(u,v) - 3 = 12\iint\limits_{[0,1]^2} C(u,v) du dv - 3 \tag{3}$$

as well as Kendall's correlation coefficient τ,

$$\tau = 4E\big[C(U,V)\big] - 1 = 4\iint\limits_{[0,1]^2} C(u,v)\, dC(u,v) - 1, \tag{4}$$

where $U = F_X(X)$ and $V = F_Y(Y)$ are uniformly distributed random variables that are connected by the same copula as are X and Y. Alternatively, Kendall's tau can be computed directly from copula generator,

$$\tau = 1 + 4\int_0^1 \frac{f(t)}{f'(t)}\, dt = 1 - 4\int_0^\infty t\big(f^{(-1)\prime}(t)\big)^2 dt$$

which is far more convenient.

Any binary Archimedean copula $C\colon [0,1]^2 \to [0,1]$ generated by an additive generator $f\colon [0,1] \to [0,\infty]$, is also a triangular norm [6], [15] and thus, it can be univocally extended to an n-ary function (we keep the original notation also for this extension) $C\colon [0,1]^n \to [0,1]$ given by

$$C(x_1,\dots,x_n) = f^{(-1)}\left(\sum_{i=1}^n f(x_i)\right). \tag{5}$$

Obviously, for any $n \geq 2$, C satisfies the boundary conditions (C1). However, for $n > 2$, (C2) may fail. For example, the smallest binary copula $W\colon [0,1]^2 \to [0,1]$ given by $W(x,y) = \max(0, x+y-1)$ is generated by the additive generator $f_W\colon [0,1] \to [0,\infty]$, $f_W(x) = 1 - x$. Its n-ary extension is given by

$$W(x_1,\dots,x_n) = 1 - \min\left(1, \sum_{i=1}^n (1-x_i)\right) = \max\left(0, \sum_{i=1}^n x_i - (n-1)\right).$$

Consider $\mathbf{x},\mathbf{y} \in [0,1]^n$, $\mathbf{x} = \left(\frac{1}{2},\dots,\frac{1}{2}\right)$, $\mathbf{y} = (1,\dots,1)$. Then $V_W([\mathbf{x},\mathbf{y}]) = 1 - \frac{n}{2}$, i.e., this volume is negative whenever $n > 2$, which shows that W is a copula only for $n = 2$. A complete description of additive generators of binary copulas such that the corresponding generated n-ary function is also an n-ary copula, $n > 2$, was given by McNeil and Nešlehová in [8] and is recalled in the next theorem.

Theorem 2. *Let $f\colon [0,1] \to [0,\infty]$ be a continuous strictly decreasing function such that $f(1) = 0$ (i.e., f is an additive generator of a continuous Archimedean t-norm, see [6]). Then the n-ary function $C\colon [0,1]^n \to [0,1]$ given by (5)* *is an n-ary copula if and only if the function $g\colon [-\infty, 0] \to [0,1]$ given by $g(u) = f^{(-1)}(-u)$ is $(n-2)$-times differentiable with non-negative derivatives $g', \dots, g^{(n-2)}$ on $]-\infty, 0[$ (or equivalently, $(-1)^n (f^{(-1)})^{(n)}(u) \geq 0$), and $g^{(n-2)}$ is a convex function (see Figure 1).*

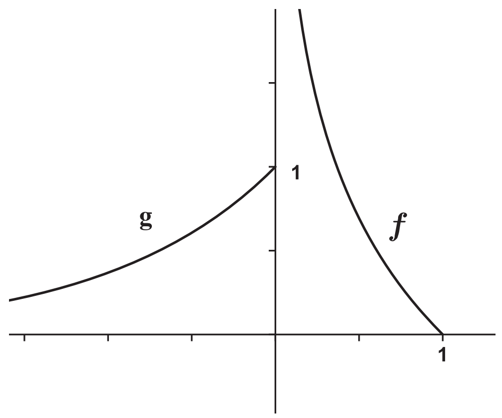


FIGURE 1. Illustration of a generator f and its corresponding function g.

We denote by $\mathcal{F}_n$ the class of all additive generators that generate n-ary copulas as characterized in Theorem 2.

Additive generators, which generate an n-ary copula for any $n \geq 2$, are called universal generators. Due to Theorem 2, we have the next result, see also [8].

COROLLARY 1. *Let $f\colon [0,1] \to [0,\infty]$ be an additive generator of a binary copula $C\colon [0,1]^2 \to [0,1]$. Then the n-ary extension $C\colon [0,1]^n \to [0,1]$ given by (5) is an n-ary copula for each $n \geq 2$ if and only if the function $g\colon [-\infty, 0] \to [0,1]$ given by $g(u) = f^{(-1)}(-u)$ is absolutely monotone, i.e., $g^{(k)}$ exists and is non-negative for each $k \in N = \{1, 2, \dots\}$.*

The class of all universal additive generators will be denoted by $\mathcal{F}_\infty$. It is not difficult to check that $\mathcal{F}_2 \supset \mathcal{F}_3 \supset \cdots \supset \mathcal{F}_\infty$.

The reverse problem of characterization of n-ary copulas which are generated by an additive generator was solved by Stupňanová and Kolesárová [17].

THEOREM 3. *Let $C\colon [0,1]^n \to [0,1]$ be an n-ary copula, $n > 2$. Then C is generated by an additive generator $f\colon [0,1] \to [0,\infty]$ if and only if C satisfies the diagonal inequality $C(x, \dots, x) < x$ for all $x \in]0,1[$, and C is associative in the Post sense, i.e., for any $x_1, \dots, x_{2n-1} \in [0,1]$ it holds*

$$C\big(C(x_1, \dots, x_n), x_{n+1}, \dots, x_{2n-1}\big) = \\ = C\big(x_1, C(x_2, \dots, x_{n+1}), x_{n+2}, \dots, x_{2n-1}\big) = \dots \\ \dots = C\big(x_1, \dots, x_{n-1}, C(x_n, \dots, x_{2n-1})\big).$$

The n-monotone Archimedean copula generators may be characterized using a little known integral transform introduced by Williamson in 1956, see [18]. In McNeil and Nešlehová [8] there is a description of this transform, which, for a fixed $n \geq 2$, will be called the Williamson n-transform. In what follows, we discuss the Williamson n-transform and illustrate it by examples.

2. The Williamson n-transform

An interesting link between additive generators of copulas and positive distance functions [9], i.e., distribution functions with support in $]0, \infty[$, was described in details in [8]. Based on the results of Williamson [18], we recall the next important result.

Theorem 4 (McNeil and Nešlehová [8], Corollary 3.1). *The following claims are equivalent for an arbitrary $n \in \{2, 3, \ldots\}$:*

(i) $f \in \mathcal{F}_n$.

(ii) Under the notation of Theorem 2, the function $F\colon]-\infty, \infty[\to [0, 1]$ given by $F(x) = 0$ if $x \leq 0$, and for $x > 0$,

$$F(x) = 1 - \sum_{k=0}^{n-2} \frac{(-1)^k x^k (f^{(-1)})^{(k)}(x)}{k!} - \frac{(-1)^{n-1} x^{n-1} (f^{(-1)})_+^{(n-1)}(x)}{(n-1)!} \tag{6}$$

is a distribution function of a positive random variable X (i.e., $P(X \leq 0) = 0$), where $\cdot_+^{(n-1)}$ denotes the right-derivative of order $n - 1$.

Note that due to [18], if F is a positive distance function, i.e., a distribution function of a positive random variable X, then for a fixed $n \in \{2, 3, \ldots\}$ the Williamson n-transform provides an inverse transformation to (6),

$$f^{(-1)}(x) = \int_x^\infty \left(1 - \frac{x}{t}\right)^{n-1} dF(t) = \begin{cases} \max\left(0, E\left[1 - \frac{x}{X}\right]^{n-1}\right), & x > 0, \\ 1 - F(0), & x = 0, \end{cases}$$

$$\text{where} \quad x \in [0, \infty[\quad \text{and} \quad f^{(-1)}(\infty) = 0. \tag{7}$$

Note that a similar relationship can be shown between additive generators from $\mathcal{F}_\infty$ and positive distance functions, based on the Laplace transform, i.e,

$$f^{(-1)}(x) = \int_0^\infty e^{-xt}\, dF(t). \tag{8}$$

For more and interesting details we recommend [8].

TOMÁŠ BACIGÁL — MÁRIA ŽDÍMALOVÁ

Let F be a distance function related to a positive random variable X. For any $c > 0$, the random variable $c.X$ possesses the distance function F_c given by $F_c(x) = F\left(\frac{x}{c}\right)$. Then, for any $n \in \{2, 3, \ldots\}$,

$$\begin{aligned} f_c^{(-1)}(x) &= \int_x^\infty \left(1 - \frac{x}{t}\right)^{n-1} dF_c(t) \\ &= \int_x^\infty \left(1 - \frac{x}{t}\right)^{n-1} dF\left(\frac{t}{c}\right) \\ &= \int_{\frac{x}{c}}^\infty \left(1 - \frac{x}{cu}\right)^{n-1} dF(u) \\ &= f^{(-1)}\left(\frac{x}{c}\right). \end{aligned}$$

Obviously, for the related additive generators it holds that $f_c = c.f$, i.e., they generate the same copula. Vice versa, clearly from (6) it follows that if two generators generate the same (n-ary) Archimedean copula, the corresponding positive random variables differ only in a positive multiplicative constant. The next result follows.

Theorem 5. *For each $n \in \{2, 3, \ldots\}$, there is an one-to-one correspondence between the class $\mathcal{F}_n$ and the class $\mathcal{H}$ of all factor classes of positive distance functions related to the equivalence $F \sim G$ if and only if $G(x) = F\left(\frac{x}{c}\right)$ for some $c > 0$.*

In the following, we illustrate the construction method by few examples.

Example 1. Let F be equal to a Dirac function[1] focused at point $x_0 = 1$,

$$F(x) = \delta_1(x) = \begin{cases} 0, & x < 1, \\ 1, & 1 \le x, \end{cases}$$

then, as it is also shown in [8], by the Williamson n-transform we get generator $f_n(x) = 1 - x^{\frac{1}{n-1}}$ of the weakest n-dimensional Archimedean copula, i.e., the non-strict Clayton copula with parameter $\lambda = \frac{-1}{n-1}$, see Figure 2. By rescaling generator to $\tilde{f}_n(x) = \frac{f(x)}{f(1/2)}$, $x \in [0, 1]$, the copula would not change, yet such a generator is fixed to the value $\tilde{f}_n\left(\frac{1}{2}\right) = 1$, which we will use later to show convergence.

[1]Dirac function is defined as $\delta_{x_0}(x) = \begin{cases} 0, & x < x_0, \\ 1, & x \ge x_0. \end{cases}$

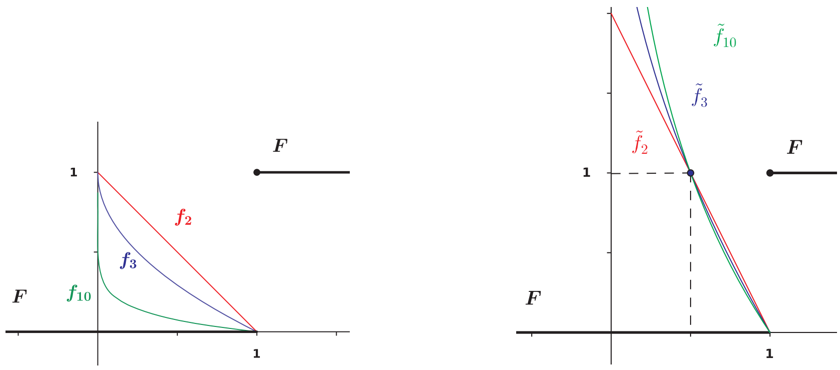


FIGURE 2. Dirac function F, the corresponding generators f_n for different n and rescaled generators $\tilde{f}_n$.

EXAMPLE 2. Let F be a uniform probability distribution function

$$F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x < b \\ 1, & b \leq x. \end{cases} \quad \text{with } \ 0 \leq a < b,$$

Then for dimension $n = 2$ we get

$$f_2^{(-1)}(x) = \int_x^\infty \left(1 - \frac{x}{t}\right)^{2-1} F'(t) dt$$

$$= \begin{cases} \int_a^b \left(1 - \frac{x}{t}\right) \frac{1}{b-a}\, dt, & x < a, \\ \int_x^b \left(1 - \frac{x}{t}\right) \frac{1}{b-a}\, dt, & a \leq x < b, \\ \int_x^\infty \left(1 - \frac{x}{t}\right) 0\, dt, & b \leq x, \end{cases}$$

$$= \begin{cases} \frac{1}{b-a} \left[t - x \log t\right]_a^b = 1 - \frac{x \log\left(\frac{b}{a}\right)}{b-a}, & x < a, \\ \frac{1}{b-a} \left[t - x \log t\right]_x^b = \frac{b}{b-a} - \frac{x + x \log\left(\frac{b}{x}\right)}{b-a}, & a \leq x < b, \\ 0, & b \leq x \end{cases}$$

(where F' denotes a first derivative of F) from which the corresponding generator can be obtained only numerically, and so is the case also with the higher dimensions, e.g.,

$$f_3^{(-1)}(x) = \begin{cases} 1 - \frac{2x \log\left(\frac{b}{a}\right)}{b-a} + \frac{x^2}{ab}, & x < a, \\ \frac{b}{b-a} - 2x \log\left(\frac{b}{x}\right) - \frac{x^2}{(b-a)b}, & a \leq x < b, \\ 0, & b \leq x, \end{cases}$$

displayed in Figure 3. Setting $a = 0$ we get $\tau = 0$ regardless of parameter b, which is in clear accordance with Theorem 5.

TOMÁŠ BACIGÁL — MÁRIA ŽDÍMALOVÁ

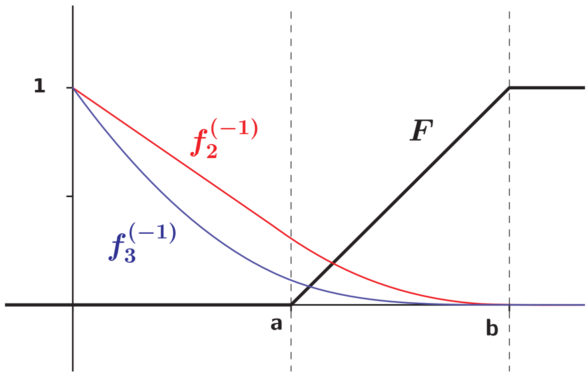


FIGURE 3. Uniform U(a,b) probability distribution function F and pseudo-inverses of the corresponding generators f_n.

We continue with the examples of constructing generators of non-strict Archimedean copulas while restricting the support of univariate distribution in the unit interval. By applying a suitable increasing transformation (such as power function) to a positive distance function on $[0,1]$ we obtain a new distribution.

EXAMPLE 3. Consider a positive distance function $F(x) = \min(1, x^2)$ and the corresponding density $F'(x) = 2x$ on $[0,1]$. Then

$$
\begin{aligned}
f_2^{(-1)}(x) &= \int_x^{\infty} \left(1 - \frac{x}{t}\right)^{2-1} dF(t) \\
&= \begin{cases} \int_x^1 (t-x)\frac{2t}{t}\, dt = (1-x)^2, & 0 \le x \le 1, \\ 0, & 1 < x \end{cases} \\
&= \max(1-x, 0)^2.
\end{aligned}
$$

Then the generator $f_2(x) = 1 - \sqrt{x}$, $x \in [0,1]$, is the generator of Clayton copula for parameter $\lambda = -\frac{1}{2}$. Nevertheless, in higher dimensions, $n \ge 3$, the generator has no closed form, e.g., $f_3^{(-1)}(x) = 1 - 4x + x^2(3 - 2\log x)$ for $x \in [0,1]$ and 0, otherwise (see Figure 4).

EXAMPLE 4. Let us generalize Example 3 and start with a parametric family $F(x) = \min(1, x^p)$, where $p > 0$. Observe that $\lim_{x\to 0} F(x) = \delta_0(x)$ while $\lim_{x\to\infty} F(x) = \delta_1(x)$. Then

$$
f_2^{(-1)}(x) = \begin{cases} \frac{x^p - px + p - 1}{p-1}, & 0 \le x \le 1 \wedge p \ne 1, \\ x(\log x - 1) + 1, & 0 \le x \le 1 \wedge p = 1, \\ 0, & 1 < x. \end{cases}
$$

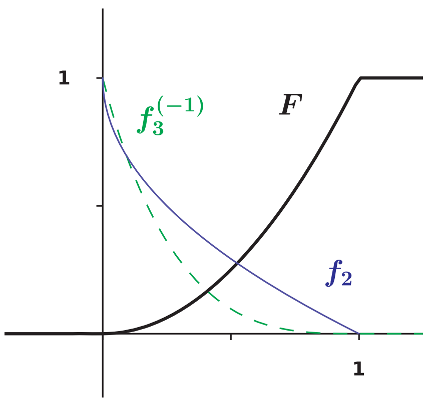
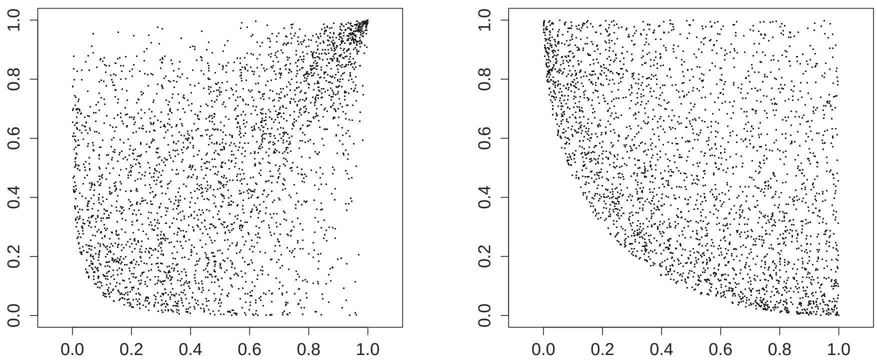


FIGURE 4. Illustration of Example 3 with non-invertible case $n = 3$.

Though it lacks an explicit inverse, the copulas that it generates cover almost whole dependence range with $\tau = 1 - \frac{2p}{1+p}$, $\tau \in (0, 1)$, and we will use it later to demonstrate approximation approach. Figure 5 shows simulations from this parametric copula family for $p = 0.5$ and $p = 2$. The only tail dependence is present at the upper tail for $p \in (0, 1)$.

It is interesting to illustrate also the inverse Williamson n-transform.



$p = 0.5$ $p = 2$

FIGURE 5. Sampling from copula family constructed in Example 4.

EXAMPLE 5. Let us take a generator of:

- the Ali-Mikhail-Haq copula $f(x) = \frac{1}{x} - 1$ corresponding to the parameter $\lambda = 1$ and denote by F_n, $n = 2, 3, \ldots$, a positive distance function related to f through (6). Then $F_n(x) = 1 - \frac{1}{1+x} - \frac{x}{(1+x)^2} - \ldots - \frac{x^{n-1}}{(1+x)^n} = \left(\frac{x}{1+x}\right)^n$ which can be viewed as a parametric subfamily of all positive valued distribution functions F_p with any positive parameter p.

- the product copula $f(x) = -\frac{1}{p}\log x$ with constant $p > 0$ and inverse $f^{-1}(x) = \exp(-px)$. From (6) for $n = 2$ we get $F(x) = 1 - \exp(-px)(1 - px)$. By comparing the density $\frac{\partial F(x)}{\partial x} = p^2 x \exp(-px)$ and the convolution of two exponential distribution $\mathcal{D}_\lambda$ densities with parameter $\lambda > 0$,
$$\int_0^x \lambda \exp(-\lambda t)\lambda \exp\big(-\lambda(x-t)\big)dt = \lambda^2 x \exp(-\lambda x)$$
it becomes clear that the resulting distribution is a distribution of the random variable $Y = X_1 + X_2$, where $X_1, X_2 \sim \mathcal{D}_\lambda$ are independent (and identically distributed) random variables. The relation holds for any $n \geq 2$, thus (6) yields a cumulative distribution function of the sum of i.i.d. random variables $X_1, \ldots, X_n \sim \mathcal{D}_p$, $F_{X_1+\cdots+X_n}(x) = 1 - \exp(-px)\sum_{i=1}^{n} \frac{(px)^{i-1}}{(i-1)!}$ with $p > 0$ which defines the Erlang distribution with rate parameter p and shape parameter n.

To complete the examples, let us illustrate also the Laplace transform.

Example 6. Starting with positive distance function of:

- discrete random variable with probability mass concentrated in $\lambda > 0$, i.e., Dirac function $F(x) = 0$ for $x < \lambda$ and 1 otherwise, then the Laplace transform leads through $g(x) = \exp(\lambda x)$ to the product copula Π.
- exponential distribution $F(x) = 1 - \exp(-\lambda x)$, $\lambda > 0$, we get $f^{-1}(x) = (\frac{\lambda}{x+\lambda})$ and $f(x) = \lambda(\frac{1}{x} - 1)$ which generates the same copula (Clayton copula with parameter equal to 1) regardless of the choice of λ.

Now we focus on the Dirac function since it can be viewed as a building block for distribution functions of a random variable with probability mass concentrated in l discrete points. Immediately a question arises: if such a distribution functions can approximate distribution of a continuous r.v. (for any l, going possibly to infinity), does this convergence imply also a convergence of the corresponding generators and even a convergence of the generated copulas?

3. Convergence theorems

Definition 1. Let $(F_m)_m$ be a sequence of distribution functions and let F be a distribution function. We say that the sequence of distribution functions F_m, $m = 1, 2, \ldots$, weakly converge to distribution function F if
$$\lim_{m\to\infty} F_m(x) = F(x)$$
holds for any point $x \in R$ in which F is continuous. The weak convergence will be denoted by $F_m \xrightarrow{w} F$.

Recall that the Lévy-Cramér continuity theorem [14] ensures the convergence $\int_{-\infty}^{\infty} h(t)dF_m(t) \underset{m\to\infty}{\longrightarrow} \int_{-\infty}^{\infty} h(t)dF(t)$, where $h\colon]-\infty,\infty[\to]-\infty,\infty[$ is a continuous bounded real function and $F_m w \to F$. Obviously, for any $n \geq 2$ and $x < 0$, the function $h\colon]-\infty,\infty[\to]-\infty,\infty[$ given by $h(t) = \left(1+\frac{x}{t}\right)^{n-1} \delta_{-x}(t)$ is continuous and bounded. This fact proves the next important result.

Theorem 6. *Let a sequence $(F_m)_m$ of distance functions converge weakly to a distance function F, $F_m \overset{w}{\to} F$. Then, for any $n \geq 2$, the corresponding additive generators f, f_m, $m = 1, 2, \ldots,$ of n-dimensional Archimedean copulas are related by the pointwise convergence $f_m \underset{m\to\infty}{\longrightarrow} f$.*

P r o o f. Based on the Williamson n-transform (7) and the Lévy-Cramér continuity theorem, $g_m \to g$ pointwisely, and all functions g, g_m, $m = 1, 2, \ldots$ are convex, continuous and strictly increasing. Then also the related additive generators f, f_m, $m = 1, 2, \ldots,$ satisfy $f_m \to f$ pointwisely. □

The reverse of Theorem 6 is based on the next lemma.

Lemma 1. *Let $(f_m)_m$, f be convex real functions defined on a real interval $]\alpha,\beta[$ such that $f_m \to f$ pointwisely. Then for any point $a \in]\alpha,\beta[$, where $f'(a)$ exists it holds* $\lim_{m\to\infty} f'(a^-) = f'(a) = \lim_{m\to\infty} f'_m(a^+)$.

P r o o f. Note that due to the convexity, the left derivatives $f'_m(a^-)$, $f''_m(a^-)$ and the right derivatives $f'_m(a^+)$, $f''_m(a^+)$ exist at each point $a \in]\alpha,\beta[$. Moreover, the convexity ensures also that $f_m(x) \geq f_m(a) + (x-a)f'_m(a^-)$ and $f_m(x) \geq f_m(a) + (x-a)f'_m(a^+)$ for any $m = 1, 2, \ldots$ and $x \in]\alpha,\beta[$. Fix $a \in]\alpha,\beta[$. Then, for any $x > a$, $f'_m(a^-) \leq \frac{f_m(x)-f_m(a)}{x-a}$ and thus $\limsup f'_m(a^-) \leq \frac{f(x)-f(a)}{x-a}$. Therefore $\limsup f'_m(a^-) \leq \lim_{x\to a^+} \frac{f(x)-f(a)}{x-a} = f'(a^+)$. Similarly, $\liminf f'_m(a^-) \geq f'(a^-)$, which implies the existence of the limit of $\left(f'_m(a^-)\right)_m$, whenever $f'(a^-) = f'(a^+) = f'(a)$, $\lim_{m\to\infty} f'_m(a^-) = f'(a)$ if $f'(a)$ exists. Using similar arguments, if $f'(a)$ exists, then also $\lim_{m\to\infty} f'_m(a^+) = f'(a)$. □

Based on Lemma 1, the next result follows directly.

Theorem 7. *Let $f_m, f \in \mathcal{F}_n$, $m = 1, 2, \ldots,$ be additive generators of n-dimensional Archimedean copulas, such that $f_n \to f$ pointwisely on $]0,1]$. Let F_m, F, $m = 1, 2, \ldots,$ be the related distance function obtained by means of the transform (6). Then $F_m \overset{w}{\to} F$.*

P r o o f. Based on Theorem 2, $g_m^{(k)}$, $m = 1, 2, \ldots,$ and $g^{(k)}$ are convex functions for $k = 0, 1, \ldots, n-2$. The pointwise convergence $f_m \to f$ on $]0,1]$ implies the pointwise convergence $g_m \to g$ on $]-\infty,0[$, and due to Lemma 1, repeated $(n-2)$--times, it holds $\lim_{m\to\infty} g'_m(x) = g'(x)$, $\ldots$, $\lim_{m\to\infty} g_m^{(n-2)}(x) = g^{(n-2)}(x)$ and

$\lim_{m\to\infty} g_m^{(n-1)}(x^-) = g^{(n-1)}(x)$ at each point $x \in]-\infty, 0[$, where $g^{(n-1)}(x)$ exists. Therefore, $\lim_{m\to\infty} F_m(x) = F(x)$ at each point $x \in]0, \infty[$, where the function $g^{(n-1)}(x^-)$ is continuous, i.e., where $F(x)$ is continuous. Thus $F_m w{\to}F$. □

The next result from [10] was shown for (n-ary) continuous Archimedean triangular norms (however, any (n-ary) Archimedean copula is also a continuous Archimedean t-norm) and later for Archimedean copulas in [3, Proposition 2].

Theorem 8. *Let $C, C_m\colon [0,1]^n \to [0,1]$, $m = 1, 2, \ldots$, be continuous Archimedean copulas, generated by additive generators $f, f_m\colon [0,1] \to [0,\infty]$, $m = 1, 2, \ldots$, respectively. Then the following are equivalent.*

i) $C_m \underset{m\to\infty}{\longrightarrow} C$ *pointwisely.*

ii) *There are positive constants c_m, $m = 1, 2, \ldots$, so that $c_m f_m \to f$ pointwisely.*

Combining Theorems 6, 7, 8 we have the next result which can be exploited when approximating Archimedean copulas.

Corollary 2. *The following convergences of related objects are equivalent (for any $n \geq 2$):*

i) for distance functions, $F_m \overset{w}{\to} F$;

ii) for additive generators from $\mathcal{F}_n$, $f_m \to f$ pointwisely on $]0, 1]$;

iii) for n-dimensional Archimedean copulas, $C_m \to C$ pointwisely.

Recall that each distance function F can be obtained as a weak limit of (bounded) discrete distance functions F_m, and that each bounded discrete distance function is, in fact, a convex combination of Dirac distance functions.

4. Approximation

In this section we are interested mainly in ($n = 2$)-dimensional case, since it is of most benefit in practice. Therefore hereafter the subscript with generator f gains a different meaning: the number of pieces f is approximated by.

Example 7. Let $F(x) = \min(1, x^2)$ be the positive distance function from the Example 3 and function

$$F_2(x) = F\left(\frac{1}{2}\right)\delta_{\frac{1}{2}}(x) + \left(F(1) - F\left(\frac{1}{2}\right)\right)\delta_1(x) = \begin{cases} 0, & x < \frac{1}{2}, \\ \frac{1}{4}, & \frac{1}{2} \leq x < 1, \\ 1, & 1 \leq x \end{cases}$$

approximates F by means of a sum of $m = 2$ Dirac functions concentrated in respective points $\left(\frac{1}{2}, \frac{1}{4}\right)$, $\left(1, \frac{3}{4}\right)$. Then the Williamson transform with $n = 2$ yields

$$f_2^{(-1)}(x) = \frac{1}{4}\max\left(0, 1 - \frac{x}{\frac{1}{2}}\right) + \frac{3}{4}\max\left(0, 1 - \frac{x}{1}\right) = \begin{cases} 1 - \frac{5}{4}x, & x < \frac{1}{2}, \\ \frac{3}{4} - \frac{3}{4}x, & \frac{1}{2} \le x < 1, \\ 0, & 1 \le x. \end{cases}$$

From Example 7 illustrated in Figure 6 we see that for $n = 2$ the additive generator inverse $f_2^{(-1)}$ is piecewise linear and does not coincide with $f^{(-1)}$ in the interval $]0, 1[$.

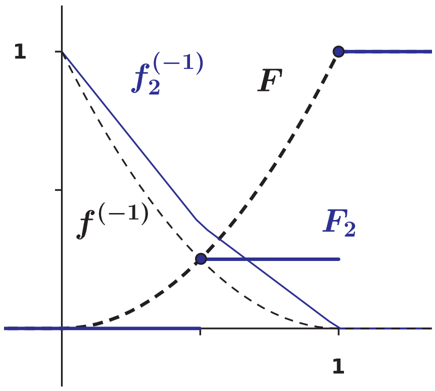


FIGURE 6. Approximation by the sum of $m = 2$ Dirac functions.

Dividing an interval $[a_0, a_m]$ by points $\{a_i\}_{i=1,\ldots m}$, $a_0 < a_1 < \cdots a_m$, with concentration of probability given by some probability mass function $p(x)$, the approximate positive distance function

$$F_m(x) = \sum_{i=1}^{m} p(a_i)\delta_{a_i}(x)$$

is then transformed by (7) to the generator inverse (related to some n-dimensional Archimedean copula)

$$f_m^{(-1)}(x) = \sum_{x<a_i} p(a_i)\left(1 - \frac{x}{a_i}\right)^{n-1} = \sum_{i=1}^{m} p(a_i)\max\left(0, 1 - \frac{x}{a_i}\right)^{n-1}. \tag{9}$$

Observe that the function $f_m^{(-1)}$ in Equation (9) is an $(n–1)$-dimensional spline. For $n = 2$, both $f_m^{(-1)}$ and the corresponding additive generator f_m are linear splines, and the related Archimedean copula C_m is piece-wise linear, as shown in Example 10. In the opposite direction, denote $b_i = f_m^{(-1)}(a_i)$ and $p_i = p(a_i)$ for $i = 1, 2 \ldots m$ with $b_0 = 1$ corresponding to $a_0 = 0$ and, clearly, $b_m = 0$.

Having points $\{(a_i, b_i)\}_{i=1,\dots m}$, their corresponding probabilities can be found by solving exquations (9) with $x = a_1, \dots, a_{m-1}$ written in the form (for $n = 2$)

$$\begin{pmatrix} 1-\frac{a_1}{a_2} & 1-\frac{a_1}{a_3} & \cdots & 1-\frac{a_1}{a_m} \\ 0 & 1-\frac{a_2}{a_3} & \cdots & 1-\frac{a_2}{a_m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1-\frac{a_{m-1}}{a_m} \end{pmatrix} \begin{pmatrix} p_2 \\ p_3 \\ \vdots \\ p_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{m-1} \end{pmatrix}.$$

The solution is $p_1 = 1 - (p_2 + \cdots + p_m)$ and

$$p_i = \frac{a_i \left[b_{i-1}(a_{i+1} - a_i) - b_i(a_{i+1} - a_{i-1} + b_{i+1}(a_i - a_{i-1}))\right]}{(a_{i+1} - a_i)(a_i - a_{i-1})} \quad \text{for} \quad i = 2, \dots m,$$

with auxiliary point (a_{m+1}, b_{m+1}), where $a_{m+1} \geq a_m$ and thus $b_{m+1} = 0$.

Example 8. Let $F(x) = \min(1, x^p)$ be the parametric family from the Example 4, then Figure 7 shows contour plots and samples from copulas generated by (9) with $a_i = F(u_i)$, where u_i is sampled from uniform distribution U(0,1), and $p(a_i) = F(a_i) - F(a_{i-1})$, where $i = 1, \dots m$, and $a_0 = 0$.

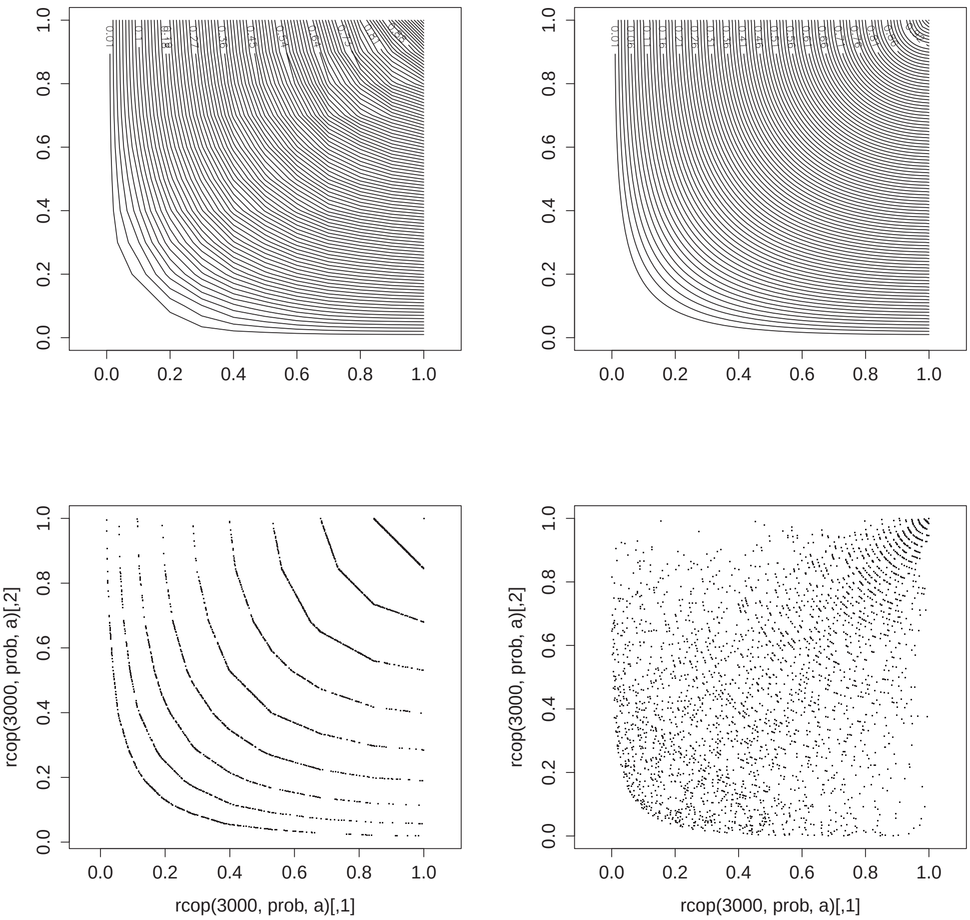
In the following examples we exercise pointwise convergence and show a piecewise linear copula corresponding to the simplest non-trivial case $n = m = 2$.

Example 9. For the simplest case, $n = 2$, $a_i = \frac{i}{m}$ and $p(a_i) = \frac{1}{m}$, $i = 1, \dots m$ (evenly spaced and uniformly distributed), we get

$$f_m^{(-1)}(x) = \sum_{i=1}^{m} \frac{1}{m} \max\left(0, 1 - \frac{mx}{i}\right).$$

If $f_m^{(-1)}(x)$ is to converge to $f^{(-1)}(x) = 1 - x + x \log x$ for $x < 1$ and 0 elsewhere, it needs to converge in any point $x \in]0, 1[$. Let us examine the convergence, say, in $x = \frac{1}{2}$, where

$$\begin{aligned} f_m^{(-1)}\left(\frac{1}{2}\right) &= \frac{1}{m} \sum_{i=1}^{m} \max\left(0, 1 - \frac{m\frac{1}{2}}{i}\right) \\ &= \frac{1}{m} \sum_{i=\lfloor \frac{m}{2} \rfloor + 1}^{m} \left(1 - \frac{m}{2i}\right) \\ &= \frac{1}{m} \sum_{i=1}^{\frac{m}{2}} \frac{i}{i + \frac{m}{2}} \\ &= \frac{1}{m} \sum_{i=1}^{\frac{m}{2}} \left(1 - \frac{\frac{m}{2}}{i + \frac{m}{2}}\right) \\ &= \frac{1}{2} - \frac{1}{2} \sum_{i=\lfloor \frac{m}{2} \rfloor + 1}^{m} \frac{1}{i}. \end{aligned}$$



$m = 10$ $m = 100$

FIGURE 7. Sampling from approximation of copula family constructed in Example 4 with $p = 0.5$.

Then indeed

$$\lim_{m\to\infty} f_m^{(-1)}\left(\frac{1}{2}\right) = \frac{1}{2} - \frac{1}{2}\int_{\frac{m}{2}}^{m} \frac{1}{x}\,dx = \frac{1}{2} - \frac{1}{2}\left[\ln x\right]_{\frac{m}{2}}^{m} = \frac{1}{2} - \frac{1}{2}\ln 2 = f^{(-1)}\left(\frac{1}{2}\right).$$

EXAMPLE 10. Following Example 9, it might help to picture the approximation copula on a simple setting. Due to Example 2 we already know that the trivial case $m = 1$ leads to the weakest copula W. With $m = 2$ we get

$$F_2(x) = \begin{cases} 0, & x < \frac{1}{2}, \\ \frac{1}{2}, & \frac{1}{2} \le x < 1, \\ 1, & 1 \le x, \end{cases}$$

thus

$$f_2^{(-1)}(x) = \begin{cases} 1 - \frac{3}{2}x, & x < \frac{1}{2}, \\ \frac{1}{2} - \frac{1}{2}x, & \frac{1}{2} \le x < 1, \\ 0, & 1 \le x \end{cases} \quad \text{and} \quad f_2(x) = \begin{cases} 1 - 2x, & 0 \le x \le \frac{1}{4}, \\ \frac{2}{3}(1-x), & \frac{1}{4} < x \le 1 \end{cases}$$

shown in Figure 8 a), which leads to copula C_2 expressed in Figure 8 b).

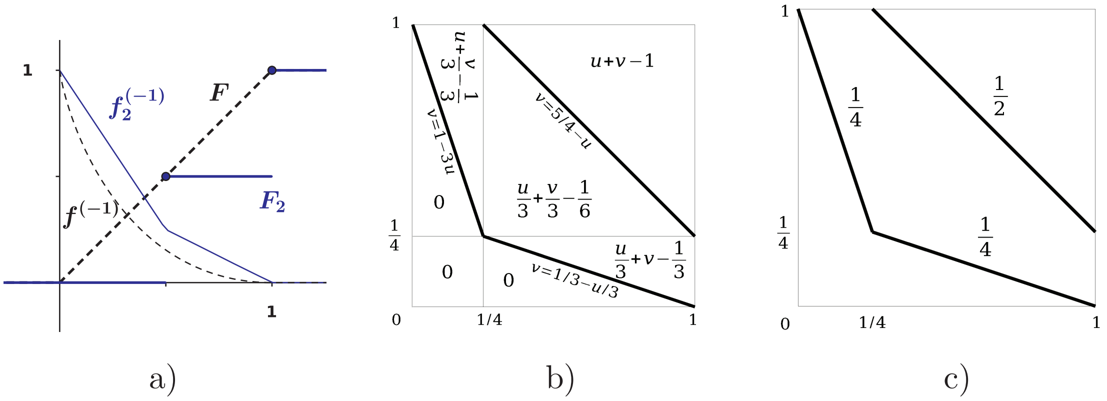


FIGURE 8. a) Distance function, generator (inverse) and b) copula, that correspond to uniform distribution approximated in $m = 2$ equally spaced points. c) Probability mass concentrated on copula support.

To compute measures of dependence (concordance) such as Spearman's rho and Kendall's tau corresponding to singular copula it is generally a challenge, yet for this simple settings it might be an interesting exercise. Since the copula C_2 is piecewise linear, the whole probability mass is concentrated on its support, thus to evaluate the expected values (especially in (4)) one need to find out distribution of the probability. In our case, it is depicted in Figure 8 c). By expressing variable v in terms of u the double integral reduces to one-dimensional integral, then

$$E[UV] = 2\int_0^{1/4} u(1-3u)\frac{\frac{1}{4}}{\frac{1}{4}}\,du + \int_{1/4}^{1} u\left(\frac{5}{4} - u\right)\frac{\frac{1}{2}}{\frac{3}{4}}\,du = \frac{2}{64} + \frac{11}{64} = -\frac{13}{64}$$

and

$$E[C(U,V)] = 2\int_0^{1/4} \max\left(0, u + \frac{1-3u-1}{3}\right)\frac{\frac{1}{4}}{\frac{1}{4}}\,du$$
$$+\int_{1/4}^{1} \max\left(\frac{1}{3}\left(u + \frac{5}{4} - u - \frac{1}{2}\right), u + \frac{5}{4} - u - 1\right)\frac{\frac{1}{2}}{\frac{3}{4}}\,du = 0 + \frac{1}{8}$$

thus $\rho_2 = 12\frac{13}{64} - 3 = -\frac{9}{16}$ and $\tau_2 = 4\frac{1}{8} - 1 = -\frac{1}{2}$, where the subscript 2 conforms the notation of generator. Although we cannot find explicit form of the original generator f (that corresponds to uniform distribution U[0,1]) and analytically calculate ρ, we still can get $\tau = 1 - \int_0^1 t\big((1-t+x\ln t)'\big)^2 dt = 1 - 4\int_0^1 t\ln^2 t dt = 0$ to measure accuracy of our $m = 2$ approximation.

5. Conclusion

We have discussed a new construction method for obtaining additive generators proposed by McNeil and Nešlehová [8], the so-called Williamson n-transform, and illustrated it by some examples. Some of the generators were shown to not have an explicit form due to non-invertability. Thus a natural approach to utilize any such parametric family is to approximate it by piecewise (in 2D case) linear functions with sufficiently dense breakpoints. We showed that due to the equivalence of convergences of positive distance functions, fixed additive generators and copulas, we may approximate any n-dimensional Archimedean copula by a transformation of convex sum of Dirac functions (though feasible mainly in 2D). We showed some simple examples, including calculation of correlation coefficients related to a singular copula.

REFERENCES

[1] BACIGÁL, T.—JURÁŇOVÁ, M.—MESIAR, R.: *On some new constructions of Archimedean copulas and applications to fitting problems,* Neural Netw. World **20** (2010), 81–90.

[2] BACIGÁL, T.—MESIAR, R.—NAJJARI, V.: *Generators of copulas and aggregation,* Inform. Sci. (submitted).

[3] CHARPENTIER, A.—SEGERS, J.: *Convergence of Archimedean copulas,* Statist. Probab. Lett. **78** (2008), 412–419.

[4] JÁGR, V.—KOMORNÍKOVÁ, M.—MESIAR, R.: *Conditioning stable copulas,* Neural Netw. World **20** (2010), 69–79.

[5] JOE, H.: *Multivariate Models and Dependence Concepts.* Chapman and Hall, London, 1997.

[6] KLEMENT, E. P.—MESIAR, R.—PAP, E.: *Triangular norms.* Kluwer Acad. Publ., Dodrecht, 2000.

[7] KLEMENT, E. P.—MESIAR, R.—PAP, E.: *Transformations of copulas,* Kybernetika **41** (2005), 425–434.

[8] MCNEIL, A. J.—NEŠLEHOVÁ, J.: *Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions,* Ann. Statist. **37** (2009), 3059–3097.

[9] MENGER, K.: *Statistical metrics,* Proc. Natl. Acad. Sci. USA **28**(12) (1942), 535–537.

[10] MESIAR, R.: *On the pointwise convergence of continuous Archimedean t-norms and the convergence of their generators,* BUSEFAL **75** (1998), 39–45.

[11] MICHIELS, F.—DE SCHEPPER, A.: *How to improve the fit of Archimedean copulas by means of transforms,* Statist. Papers **53** (2012), 345–355.
[12] MOYNIHAN, R.: *On τ_T semigroups of probability distributions II,* Aequationes Math. **17** (1978), 19–40.
[13] NELSEN, R. B.: *An Introduction to Copulas.* Springer, New York, 2006.
[14] RÉNYI, A.: *Wahrscheinlichketsrechnung mit einem Anhang über Informationstheorie.* Deutsche Verlag der Wissenschaften, Berlin, 1962.
[15] SCHWEIZER, B.—SKLAR, A.: *Probabilistic metric spaces.* Courier Dover Publ., 1983.
[16] SKLAR, A.: *Fonctions de répartition à n dimensions et leurs marges,* Publ. Inst. Statist. Univ. Paris **8** (1959), 229–231.
[17] STUPŇANOVÁ, A.—KOLESÁROVÁ, A.: *Associative n-dimensional copulas,* Kybernetika **47** (2011), 93–99.
[18] WILLIAMSON, R. E.: *Multiply monotone functions and their Laplace transforms,* Duke Math. J. **23** (1956), 189–207.

Received April 13, 2017

Faculty of Civil Engineering
Slovak University of Technology
in Bratislava
Radlinského 11
SK–810-05 Bratislava
SLOVAKIA
E-mail: bacigal@math.sk
zdimalova@math.sk