Chapter 3
introduction to copulas

Seminar *Spatio-temporal dependence*,
07.02.2011 - 11.02.2011

Benedikt Gräler
Institute for Geoinformatics
University of Muenster

ifgi
Institute for Geoinformatics
University of Münster

Outline

1 Copulas

- Bounds
- Derivatives
- Sklar's Theorem
- Symmetry
- Tail Dependence
- Elliptical Copulas
- Archimedean Copulas
- Practical

Outline

1 Copulas

- Bounds
- Derivatives
- Sklar's Theorem
- Symmetry
- Tail Dependence
- Elliptical Copulas
- Archimedean Copulas
- Practical

2 Spatial and Spatio-temporal Copulas

- one family approach
- mixed family approach

Outline

1. **Copulas**
   - Bounds
   - Derivatives
   - Sklar's Theorem
   - Symmetry
   - Tail Dependence
   - Elliptical Copulas
   - Archimedean Copulas
   - Practical
2. **Spatial and Spatio-temporal Copulas**
   - one family approach
   - mixed family approach
3. **Estimation of Copulas**
   - Maximum Likelihood
   - Moment based
   - GOF
   - Practical

ifgi
Institute for Geoinformatics
University of Münster

Outline

The d-dimensional Copula

Definition

Denote $[0,1]$ by $\mathbf{I}$. A *copula* of dimension d is a function $C : \mathbf{I}^d \to \mathbf{I}$ with the following properties:

1. $C(u_1, \ldots, u_d) = 0,$ if $\exists\ k \in \{1, \ldots, d\}$ with $u_k = 0$
2. $C(1, \ldots, 1, u_k, 1, \ldots, 1) = u_k\ \forall\ u_k \in \mathbf{I}$ and $\forall\ k \in \{1, \ldots, d\}$
3. $\forall\ \mathbf{u} = (u_1, \ldots, u_d), \mathbf{v} = (v_1, \ldots, v_d) \in \mathbf{I}^d$ with $u_k \leq v_k\ \forall\ k \in \{1, \ldots, d\}$ (called *d-increasing*):

$$\sum_{\mathbf{c} \in \mathcal{V}} \operatorname{sgn}(\mathbf{c}) C(\mathbf{c}) \geq 0$$

while

$$\mathcal{V} := \left\{ c \in \mathbf{I}^d \middle| c_k \in \{u_k, v_k\}\ \forall\ k \in \{1, \ldots, d\} \right\}$$

$$\operatorname{sgn}(\mathbf{c}) := \begin{cases} 1, & if\ c_k = u_k \text{ for an even number of ks} \\ -1, & if\ c_k = u_k \text{ for an odd number of ks} \end{cases}$$

ifgi
Institute for Geoinformatics
University of Münster

The 2-dimensional Copula I

Definition

A *copula* is a function $C : \mathbf{I}^2 \to \mathbf{I}$ with the following properties:

1. $C(u,0) = 0 = C(0,v) \ \forall \ (u,v) \in \mathbf{I}^2$
2. $C(u,1) = u$ *and* $C(1,v) = v \ \forall \ (u,v) \in \mathbf{I}^2$
3. $\forall \ (u_1, v_1), (u_2, v_2) \in \mathbf{I}^2$ *with* $u_1 \leq u_2$ *and* $v_1 \leq v_2$:

$$\begin{aligned} & V_C\big([u_1, u_2] \times [v_1, v_2]\big) \\ =& C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \\ \geq & 0 \end{aligned}$$

   This property is referred to as *2-increasing*. We call the value of the alternating sum $V_C\big([u_1, u_2] \times [v_1, v_2]\big)$ the *C-volume* of the square $[u_1, u_2] \times [v_1, v_2] \subset \mathbf{I}^2$.

See [Nelsen 2006] for a more detailed introduction.

The 2-dimensional Copula II

The graph of a copula can be imagined as a surface in the 3-dimensional unit cube $\mathbf{I}^3$.
The intersection of any copula with the unit cube is given by the edges of the skewed polygon

$$\Big\{(u,0,0),(1,v,v),(u,1,u),(0,v,0)\Big|u,v\in\mathbf{I}\Big\}.$$

The Fréchet-Hoeffding bounds I

ifgi
Institute for Geoinformatics
University of Münster

Example

- *The Fréchet-Hoeffding bounds*
  For $M(u,v) := \min(u,v)$ and $W(u,v) := \max(u+v-1,0)$ the following inequality holds for every copula C:

$$W(u,v) \leq C(u,v) \leq M(u,v)$$

- *The product copula*
  The 2-place function Π defined by $\Pi(u,v) := uv$ is a copula.

The Fréchet-Hoeffding bounds II

ifgi
Institute for Geoinformatics
University of Münster

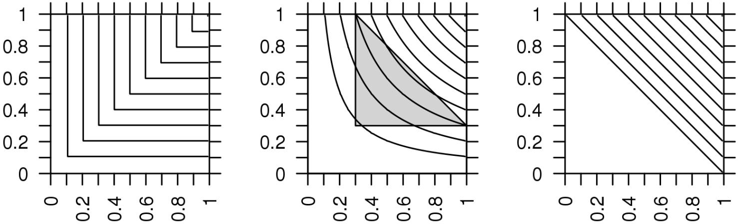


Figure: Contour plots of M, Π and W for the level set $\{0, 0.1, \ldots, 0.9\}$. The light grey triangle represents the Fréchet-Hoeffding bounds for $a_0 = 0.3$.

The Fréchet-Hoeffding bounds III

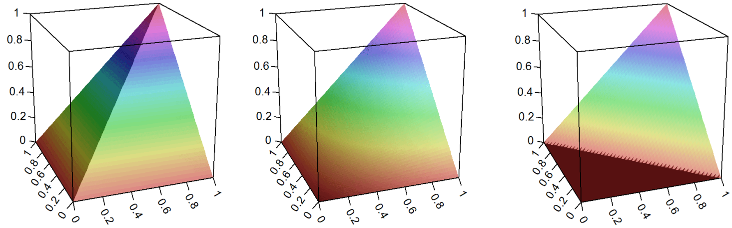


Figure: 3D-plots of M, Π and W

Derivatives of Copulas

ifgi
Institute for Geoinformatics
University of Münster

Theorem

For a given copula C and any $v \in \mathbf{I}$ the partial derivative $\partial C(u,v)/\partial u$ exists for almost all u with respect to the Lebesgue-measure. For such u and v we have

$$0 \leq \frac{\partial}{\partial u} C(u,v) \leq 1 \tag{1}$$

and also

$$v \mapsto \frac{\partial}{\partial u} C(u,v) \tag{2}$$

is defined and non-decreasing almost everywhere on $\mathbf{I}$. The similar result with interchanged roles of u and v holds as well.

The partial derivatives correspond to a conditional distribution function.

Sklar's Theorem

Theorem

Let X and Y be random variables with distribution functions F and G respectively and joint distribution function H. Then there exists a copula C such that for all $(x, y) \in \overline{\mathbb{R}} \times \overline{\mathbb{R}}$:

$$H(x, y) = C\big(F(x), G(y)\big)$$

C is unique if F and G are continuous; otherwise, C is uniquely determined on $\operatorname{ran}(F) \times \operatorname{ran}(G)$. Conversely, if C is a copula and F and G are distribution functions then the function H defined as above is a joint distribution function with margins F and G.

The meaning of Sklar's Theorem

For any joint distribution function $H(x, y)$ exists a copula C that describes the dependence of the two random variables X and Y completely. The distributions of the margins are not of any relevance and are removed by applying their distribution functions F and G respectively.

$$H(x, y) = C\big(F(x), G(y)\big)$$

ifgi
Institute for Geoinformatics
University of Münster

The Gaussian Copula I

ifgi
Institute for Geoinformatics
University of Münster

Example (of Sklar's Theorem)

Consider a bivariate standard Gaussian (X, Y) random variable with mean $\mu := (0, 0)$, a correlation ρ and covariance matrix $\Sigma := \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$.

We denote its distribution function by $H_\rho(x, y)$. The margins X and Y posses univariate standard Gaussian distributions $N(0, 1)$ with distribution function Φ. Following Sklar's Theorem we define:

$$C_\rho^N(u, v) := H_\rho\big(\Phi^{-1}(u), \Phi^{-1}(v)\big)$$

This is the definition of the *Gaussian Copula* C_ρ^N with parameter ρ.

The Gaussian Copula II

ifgi
Institute for Geoinformatics
University of Münster

Note:
C^N_ρ is as well the copula of *any* bivariate non-standard Gaussian (i.e. $\mu_x, \mu_y \neq 0$, $\sigma_x, \sigma_y \neq 1$) random variable and *many* non Gaussian random variables as well.

Attention:
There are "more" bivariate random variables having Gaussian margins but *do not* posses a Gaussian dependence structure (a Gaussian copula).

Copulas & dependence

A copula $C : \mathbf{I}^2 \to \mathbf{I}$ can be understood as bivariate joint distribution function of some distribution over the unit square $\mathbf{I}^2$. As such, they posses a bivariate density function:

$$c : \mathbf{I}^2 \to [0, \infty)$$

Copulas & dependence

A copula $C : \mathbf{I}^2 \to \mathbf{I}$ can be understood as bivariate joint distribution function of some distribution over the unit square $\mathbf{I}^2$. As such, they posses a bivariate density function:

$$c : \mathbf{I}^2 \to [0, \infty)$$

This density is what we are really interested in!

The copula's density reflects the strength of dependence of the two margins.

ifgi
Institute for Geoinformatics
University of Münster

A copula's density

It holds $C(u,v) = \int_{[0,u]\times[0,v]} c(x,y)d(x,y)$

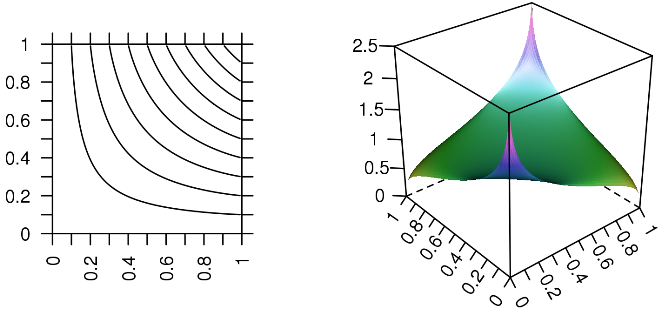


Figure: Contour plot and 3D density plot of a Gaussian Copula for $\rho = 0.2$.

About symmetry

Definition

We will introduce two kinds of symmetry:

- (plain) symmetry: $C(u,v) = C(v,u) \ \forall \ u,v \in \mathbf{I}$
- radial symmetry:
  $C(u,v) = u + v - 1 + C(1-u, 1-v) \ \forall \ u,v \in \mathbf{I}$

About symmetry

Definition

We will introduce two kinds of symmetry:

- (plain) symmetry: $C(u,v) = C(v,u) \; \forall \; u,v \in \mathbf{I}$
- radial symmetry:
  $C(u,v) = u + v - 1 + C(1-u, 1-v) \; \forall \; u,v \in \mathbf{I}$

Example

- The product copula $\Pi(u,v) = uv$ is (obviously) symmetric.
- The Gaussian Copula C_ρ^N is radial symmetric, as any copula deduced from an *elliptical distribution*.

The problem of symmetry

- symmetry is nice as long as your process is symmetric

The problem of symmetry

- symmetry is nice as long as your process is symmetric
- there are (many) natural processes that posses an asymmetric dependence structure
  - elevation: valleys are usually smoother than mountains
  - amount of toxics: the increase of a toxic is usually much steeper than its decay

ifgi
Institute for Geoinformatics
University of Münster

The problem of symmetry

- symmetry is nice as long as your process is symmetric
- there are (many) natural processes that posses an asymmetric dependence structure
  - elevation: valleys are usually smoother than mountains
  - amount of toxics: the increase of a toxic is usually much steeper than its decay
- unfortunately, most copula families in the literature are symmetric

ifgi
Institute for Geoinformatics
University of Münster

An asymmetric copula

Example

$$C_{ab}^A(u,v) = uv + uv(1-u)(1-v)\big((a-b)v(1-u)+b\big)$$

for all $|b| \leq 1$ and $\left(b-3-\sqrt{9+6b-3b^2}\right)/2 \leq a \leq 1$ with $a \neq b$ (see Example 3.16 in [Nelsen 2006]).

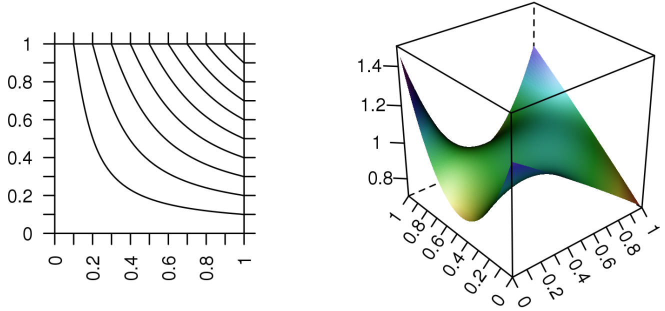


Figure: Contour plot and density plot of the asymmetric copula with $a = -0.5$ and $b = 0.3$.

An asymmetric pseudo-sample

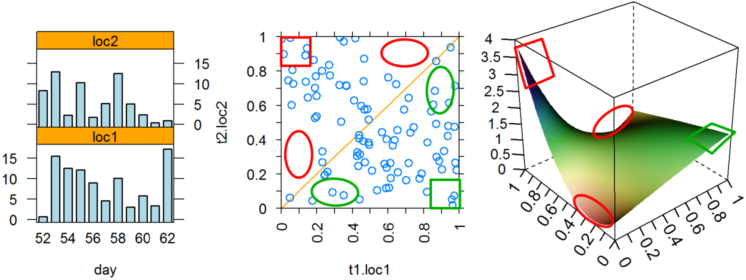


Figure: An eleven day subset of some asymmetric pseudo-sample.

Dependencies of extremes I

In cases of extreme events one is interested in the probability to see joint extremes.
This is $\mathbb{P}\left(Y > G^{-}(t) \middle| X > F^{-}(t)\right)$ for some t close to 1 or 0. We define the upper and lower tail dependence:

$$\lambda_U = \lim_{t \nearrow 1} \mathbb{P}\left(Y > G^{-}(t) \middle| X > F^{-}(t)\right)$$
$$= 2 - \lim_{t \nearrow 1} \frac{1 - C(t,t)}{1 - t}$$
$$\lambda_L = \lim_{t \searrow 0} \mathbb{P}\left(Y \leq G^{-}(t) \middle| X \leq F^{-}(t)\right)$$
$$= \lim_{t \searrow 0} \frac{C(t,t)}{t}$$

Some examples of tail dependence

Example

- any radial symmetric copula has equivalent upper and lower tail dependence
- the family of Gaussian copulas does not exhibit any tail dependence
  (even Gaussian Copulas with correlation coefficients ρ very close to 1 generate (almost) independent extremes)
- the copulas W (perfect negative dependence) and Π (independence) do not exhibit any tail dependence
- for the copula M (perfect positive dependence) we get $\lambda_U = \lambda_L = 1$

Gaussian Copulas I

We have seen the Gaussian Copula before:

$$C_\rho^N(u,v) := \Phi_\rho\big(\Phi^{-1}(u), \Phi^{-1}(v)\big)$$

Its density evaluates to:

$$c_\rho^N(u,v) = \frac{\varphi_\rho\big(\Phi^{-1}(u), \Phi^{-1}(u)\big)}{\varphi\big(\Phi^{-1}(u)\big)\varphi\big(\Phi^{-1}(v)\big)}$$

With $-1 \leq \rho \leq 1$ (Pearson's correlation coefficient)

Gaussian Copulas II

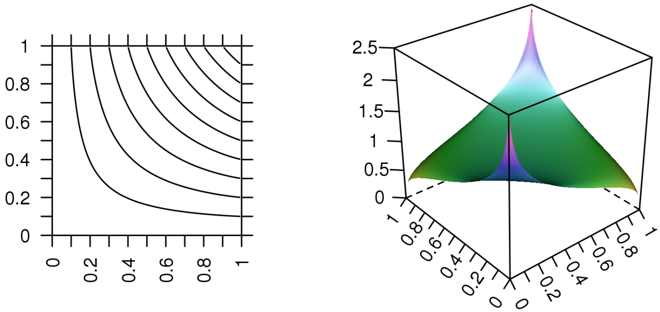


Figure: Contour plot and 3D density plot of a Gaussian Copula for $\rho = 0.2$.

Student Copulas I

ifgi
Institute for Geoinformatics
University of Münster

The Student Copula (or t-Copula) is derived from the t-distribution:

$$C^t_{\nu,\rho}(u,v) = t_{\nu,\rho}\left(t_\nu^{-1}(u), t_\nu^{-1}(v)\right)$$

Where $t_{\nu,\rho}$ is the cumulative distribution function of a bivariate $t_{\nu,\rho}$ distribution and ρ is the correlation coefficient. Its density evaluates to:

$$c^t_{\nu,\rho}(u,v) = \frac{f_{\nu,\rho}\left(f_\nu\left(t_\nu^{-1}(u)\right), f_\nu\left(t_\nu^{-1}(v)\right)\right)}{f_\nu\left(t_\nu^{-1}(u)\right) f_\nu\left(t_\nu^{-1}(v)\right)}$$

Where $f_{\nu,\rho}$ is the joint density of a bivariate t_ν-distribution.

Student Copulas II

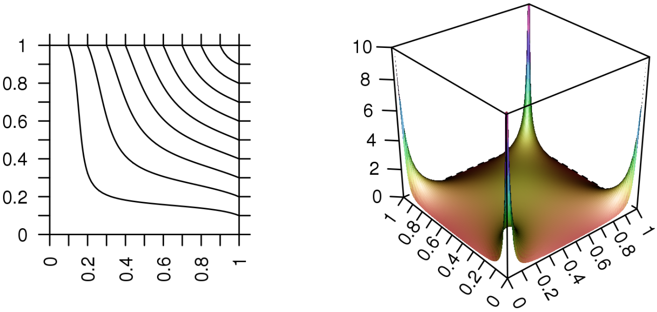


Figure: Contour plot and density plot of a t-Copula with $\rho = 0.2$ and $\nu = 1$. The density graph is limited to a level of 10 (the values for $(u, v) = (0, 0)$ and $(u, v) = (1, 1)$ reach up to 24.2).

Student Copulas III

A t-Copula's tail dependence can be evaluated by

$$\lambda^t_{\nu,\rho} = 2t_{\nu+1}\left(\frac{-\sqrt{(1+\nu)(1-\rho)}}{\sqrt{1+\rho}}\right).$$

Surprisingly, a t-Copula exhibits tail dependence even for negative correlation coefficients.

Student Copulas IV

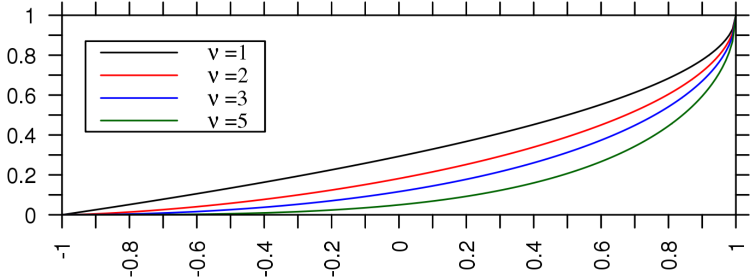


Figure: Comparison of the relation of the linear correlation parameter ρ and the tail dependence λ for different values of ν.

Archimedean Copulas

ifgi
Institute for Geoinformatics
University of Münster

A vast and flexible class of copulas are the *Archimedean Copulas*. They are defined by:

Definition

$C(u,v) = \varphi^{[-]}(\varphi(u) + \varphi(v))$ is an *Archimedean Copula* for any strictly decreasing convex function φ with $\varphi(1) = 0$ - its *generator*. $\varphi^{[-]}$ is defined as the *pseudo-inverse* of φ:

$$\varphi^{[-]}(t) := \begin{cases} \varphi^{-1}(t) & , if\ 0 \leq t \leq \varphi(0) \\ 0 & , if\ \varphi(0) \leq t \leq \infty \end{cases}$$

Some explicit Archimedean Copulas I

Example

The *Frank family* $\mathcal{C}_F := \{C_\theta^F | \theta \in \Theta_F\}$:
For any parameter $\theta \in \Theta_F := (-\infty, \infty) \setminus \{0\}$ and the corresponding generator $\varphi_\theta^F(t) = -\ln(\frac{\exp(-\theta t)-1}{\exp(-\theta)-1})$ one achieves

$$C_\theta^F(u,v) = -\tfrac{1}{\theta} \ln\left(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right).$$

They posses the lower and the upper Fréchet-Hoeffding bounds as limiting cases as θ approaches $-\infty$ and ∞ respectively. For $\theta \to \pm 0$ it takes the product copula as its limit $C_{F,\pm 0} = \Pi$. For all Frank copulas $C_\theta^F(u,v) = \hat{C}_\theta^F(u,v) = u + v - 1 + C_\theta^F(1-u, 1-v)$ holds and it is $\lambda_U^F = \lambda_L^F = 0$. This family is the only Archimedean radially symmetric one.

Some explicit Archimedean Copulas II

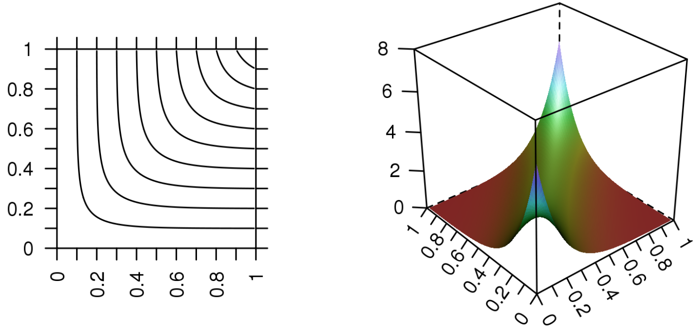


Figure: Contour plot and density plot of a Frank Copula C_7^F.

Some explicit Archimedean Copulas III

ifgi
Institute for Geoinformatics
University of Münster

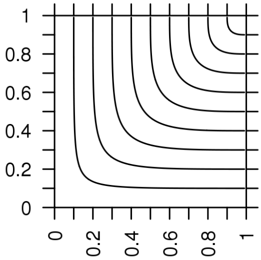
Example

The *Gumbel family* $\mathcal{C}_G := \{C_\theta^G | \theta \in \Theta_G\}$:
For a parameter $\theta \in \Theta_G := [1, \infty)$ and the generator $\varphi_\theta^G(t) = (-\ln(t))^\theta$ one achieves

$$C_\theta^G(u,v) = \exp\left(-\left((-\ln(u))^\theta + (-\ln(v))^\theta\right)^{1/\theta}\right)$$

These copulas range from the product copula Π for $\theta = 1$ to the upper Fréchet-Hoeffding bound as limiting case while θ approaches ∞.
The tail dependence parameters evaluate to $\lambda_U^G = 2 - 2^{1/\theta}$ and $\lambda_L^G = 0$.

Some explicit Archimedean Copulas IV



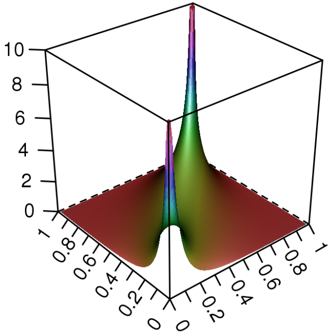


Figure: Contour plot and density plot of a Gumbel Copula C_3^G.

Some explicit Archimedean Copulas V

Example

The *Clayton family* $\mathcal{C}_C := \{C^C_\theta | \theta \in \Theta_C\}$:
For the parameter space $\Theta_C := [-1, \infty)\backslash\{0\}$ and generators of the form $\varphi^C_\theta(t) = t^{-\theta} - 1$ with $\theta \in \Theta_C$ one achieves

$$C^C_\theta(u, v) = \left(\max\left(u^{-\theta} + v^{-\theta} - 1, 0\right)\right)^{-1/\theta}.$$

These copulas range from the lower to almost the upper Fréchet-Hoeffding bounds as θ equals -1 or approaches ∞ respectively. For θ tending towards ± 0 the family converges to the product copula Π.
The tail dependence parameters evaluate to $\lambda^C_U = 0$ and $\lambda^G_L = 2^{-1/\theta}$.

Some explicit Archimedean Copulas VI

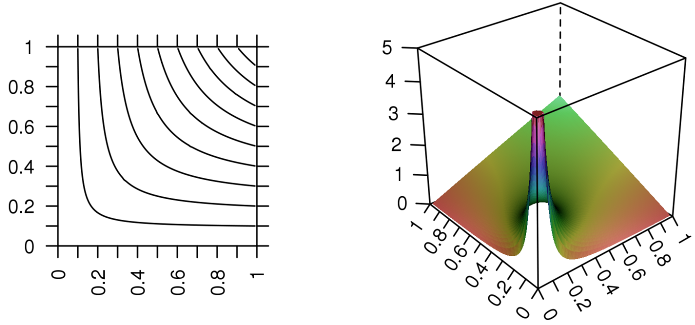


Figure: Contour plot and density plot of a Clayton Copula C_1^C.

TASK - estimate some copulas

1. plot a couple of copulas for a different set of parameters as contour, 3D their densities ...
2. How will the density of the Fréchet Hoeffding bounds look like?
3. How does the density of the product copula Π look like?
4. Which of the introduced families intersect? For which parameters? (Look at the parameter space beforehand.)
5. Plot the difference of two copula densities (or copula) to study their different strength of dependence.
6. Compare in this way the product copula with the *normal* copula for a small parameter (≈ 0.2).

Explain the plots, the differences and the meaning of positive/negative values as well in terms of *strength of dependence* (where appropriate).

General thoughts

A spatial (or spatio-temporal) copula shall describe the spatial (or spatio-temporal) dependence of two locations s_1, s_2 (or (s_1, t_1), (s_2, t_2)) of a random process $\mathcal{Z}$ defined over some region S (or $S \times T$). Thus, instead of the bivariate process of *wind speed and temperature* at *one* location, we look at *wind speed or temperature* at *two* different locations.

- we expect the dependence structure to change for different aligned points

General thoughts

A spatial (or spatio-temporal) copula shall describe the spatial (or spatio-temporal) dependence of two locations s_1, s_2 (or (s_1, t_1), (s_2, t_2)) of a random process $\mathcal{Z}$ defined over some region S (or $S \times T$). Thus, instead of the bivariate process of *wind speed and temperature* at *one* location, we look at *wind speed or temperature* at *two* different locations.

- we expect the dependence structure to change for different aligned points
- we have to make the copula aware of location/distance and (direction)

ifgi
Institute for Geoinformatics
University of Münster

General thoughts

A spatial (or spatio-temporal) copula shall describe the spatial (or spatio-temporal) dependence of two locations s_1, s_2 (or (s_1, t_1), (s_2, t_2)) of a random process $\mathcal{Z}$ defined over some region S (or $S \times T$). Thus, instead of the bivariate process of *wind speed and temperature* at *one* location, we look at *wind speed or temperature* at *two* different locations.

- we expect the dependence structure to change for different aligned points
- we have to make the copula aware of location/distance and (direction)
- we need some function $h : S \to \Theta$ from S into the copula's parameter space Θ

General thoughts

A spatial (or spatio-temporal) copula shall describe the spatial (or spatio-temporal) dependence of two locations s_1, s_2 (or (s_1, t_1), (s_2, t_2)) of a random process $\mathcal{Z}$ defined over some region S (or $S \times T$). Thus, instead of the bivariate process of *wind speed and temperature* at *one* location, we look at *wind speed or temperature* at *two* different locations.

- we expect the dependence structure to change for different aligned points
- we have to make the copula aware of location/distance and (direction)
- we need some function $h : S \to \Theta$ from S into the copula's parameter space Θ
- we need to ensure that the spatial (or spatio-temporal) copula respects *Tobler's first law of geography*

ifgi
Institute for Geoinformatics
University of Münster

within one family

One way of defining a spatial (or spatio-temporal) copula is to look at a single copula family. In this case, we only need to find a function $h : S \to \Theta$ which reproduces the changing dependence over space. The one copula family we choose needs to have two properties:

- takes the product copula Π and the upper Fréchet-Hoeffding bound M for some parameter $\theta \in \Theta$ (or at least as limiting cases). The product copula Π can then be chosen for independent far distant locations and M describing perfect positive dependence for very close locations.
- is flexible enough to represent all different dependence structures

setting up a one family spatial copula

Example

Assume you have an isotropic data set of temperature measurements over a given region. Group the data into a set of lag-classes, transform the margins to uniform distributed variables and take a look at the corresponding scatter plots (using for instance `hscat()`).
Choose a suitable copula family $\mathcal{C}$, estimate the parameter(s) for each lag and fit some function $h : [0, \infty) \to \Theta_{\mathcal{C}}$ of distance through them:

$$C_{h(d)}(u, v)$$

multiple families

A second approach considers multiple copula families and grounds on the fact, that any linear convex combination of copulas is a copula.

- Now, we might even change the copula family according to location/distance and (direction).
- The spatial (or spatio-temporal) copula is then a convex combination of a set of copulas (luckily, any convex combination of copulas is a copula).
- In case of very close points we can simply add the copula M and in case of far distant points we can add the product copula Π to the convex combination.

ifgi
Institute for Geoinformatics
University of Münster

setting up a multiple family spatial copula

Example

Assume you have an isotropic data set of temperature measurements over a given region. Group the data into a set of lag-classes, transform the margins to uniform distributed variables and take a look at the corresponding scatter plots (using for instance `hscat()`).
Choose a suitable copula family $\mathcal{C}$ *for each lag-class* and estimate their parameter(s). For any distance d pick the two fitted copulas from the neighboring lag-classes d_l, d_u and define $\lambda := (d_u - d)/(d_u - d_l)$:

$$C_d(u,v) := \lambda \cdot C_{d_l}(u,v) + (1-\lambda) \cdot C_{d_u}(u,v)$$

While $C_0 = M$ and $C_r = \Pi$ for some maximal distance r.

Transform of margins

There are several possibilities to estimate a copula within a family (the choice of family has to be achieved upon inherited properties, by smart guessing or afterwards by GOF-tests). But before, we need to transform the margins by

- knowing the marginal distributions
- estimating the marginal distributions
- approximating the marginal distributions by a rank-order transformation

ifgi
Institute for Geoinformatics
University of Münster

rank-order Transformation

ifgi
Institute for Geoinformatics
University of Münster

In the transformed dataset $\tilde{\mathbf{Z}}$ any observation $z_i \in \mathbf{Z}$ is replaced by its rank divided by the number of observations$+1$:

$$\tilde{\mathbf{Z}} := \left\{ \frac{\mathrm{rank}(z_i)}{n+1} \middle| \; 1 \leq i \leq n \right\}$$

$\tilde{\mathbf{Z}}$ is uniformly distributed. This approach does not alter the copula as a copula is invariant under strictly increasing transformations of the margins. (= As long as you do not alter the ranks in the sample, you do not alter the copula.)

Empirical Copula

For a sample $(\mathbf{X}, \mathbf{Y})$ with transformed margins the *empirical copula* is defined as:

$$C_n(u, v) = \frac{\#\{k \in \{1, \dots, n\} \big| x_k \leq u \wedge y_k \leq v\}}{n}$$

A two-dimensional step function.
We will denote the *empirical copula frequency (empirical density)* by c_n. It is given by:

$$c_n\left(\frac{i}{n}, \frac{j}{n}\right) = \begin{cases} 1/n & , if \ (x_{(i)}, y_{(j)}) \in (\mathbf{X}, \mathbf{Y}) \\ 0 & , otherwise \end{cases}$$

ifgi
Institute for Geoinformatics
University of Münster

different estimation procedures

Copulas can be estimated by

- a Maximum Likelihood approach
- a moment based approach incorporating measures of association like *Kendall's tau* or *Spearman's rho* (does not apply in a general way)
- mixtures of both
- a Bayesian approach
- others

Maximum Likelihood estimation

Assume a bivariate dataset with uniform distributed margins $\mathbf{U} = (u_1, \ldots, u_n)$ and $\mathbf{V} = (v_1, \ldots, v_n)$. For a given copula family $\mathcal{C}$ with parameter space $\Theta_{\mathcal{C}}$ we define its log-likelihood function by:

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} \log\left(c_\theta\left(u, v\right)\right)$$

$$\hat{\theta} = \arg\max_{\theta \in \Theta} \mathcal{L}(\theta)$$

This approach can easily be extend to copulas of higher dimensions.

Maximum Liklihood in R

The library copula[1] offers a build-in method `fitCopula()` to estimate copulas. The data needs to be provided as matrix. In order to choose a copula family one member needs to be provided to the function.

```
fitCopula(copula, data, method="ml")
```

```
uranium.biv <- as.matrix(uranium[c("U","Li")])
fitCopula(frankCopula(.4),uranium.biv,method="ml")
The estimation is based on the maximum likelihood
and a sample of size 655.
      Estimate   Std. Error  z value Pr(>|z|)
param 1.206623 0.0007958563 1516.131        0
The maximized loglikelihood is  599.156
```

[1] see `http://cran.r-project.org/web/packages/copula/`

Copulas and Kendall's tau and Spearman's rho

The two measures of association *Kendall's tau* and *Spearman's rho* can be derived from any copula. Some exhibit a nice functional relation between their parameter and one or both measure(s) of association above.

The population version of Kendal's tau is given by:

$$\tau_C = 4 \int_{\mathbf{I}}^{2} C(u,v) \, \mathrm{d}C(u,v) - 1 =_{Arch.Cop.} 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} \, \mathrm{d}t$$

The population version of Spearman's rho is given by:

$$\rho_C = 12 \int_{\mathbf{I}}^{2} C(u,v) - uv \, \mathrm{d}(u,v)$$

estimating Kendall's tau

Definition

Let $(\mathbf{X}, \mathbf{Y})$ denote the n observations drawn from a continuous random vector (X, Y). We denote the number of concordant pairs of observations by c and the number of discordant pairs of observations by d. The empirical version of Kendall's tau is given by :

$$\hat{\tau}(\mathbf{X}, \mathbf{Y}) := \frac{c-d}{c+d} = \frac{c-d}{\binom{n}{2}}$$

In case the sample contains any ties we use the following corrected version

$$\hat{\tau}(\mathbf{X}, \mathbf{Y}) := \frac{c-d}{\sqrt{c+d+t_x}\sqrt{c+d+t_y}}.$$

Where t_x and t_y are the number of ties in $\mathbf{X}$ or $\mathbf{Y}$ only while ties that happen to occur in both margins simultaneously are not counted at all.

estimating Spearman's rho

Definition

For a given sample $(\mathbf{X}, \mathbf{Y})$ of size n, drawn from a continuous random vector, we define the empirical version of Spearman's rho by

$$\hat{\rho}(\mathbf{X}, \mathbf{Y}) := 1 - \frac{6 \sum_{i=1}^{n} \Delta_i^2}{n(n^2 - 1)}$$

where $\Delta_i := \mathrm{rank}(x_i) - \mathrm{rank}(y_i)$ for $(x_i, y_i) \in (\mathbf{X}, \mathbf{Y})$, $1 \leq i \leq n$. In case of ties within the sample we consider the averaged ranks and adjust $\hat{\rho}$ by

$$\hat{\rho}(\mathbf{X}, \mathbf{Y}) := \frac{n \sum (r_i s_i) - \sum r_i \sum s_i}{\sqrt{n \sum r_i^2 - (\sum r_i)^2} \sqrt{n \sum s_i^2 - (\sum s_i)^2}}.$$

while all sums are taken over $i = 1, \ldots, n$. The variables r_i and s_i are given as $r_i := \mathrm{rank}(x_i)$ and $s_i := \mathrm{rank}(y_i)$.

Spearman's rho

Spearmn's rho can be thought of as the standard correlation coefficient (Pearson) applied to the ranks of a sample.

Spearman's rho assigns 1 to a perfect monotonic dependence structure which need not be linear in any sense.

In general, it is less sensitive to outliers than Pearson's correlation coefficient.

ifgi
Institute for Geoinformatics
University of Münster

Kendall’s tau and Spearman’s rho in R

The function `cor()` provides an argument `method` that takes `"pearson"`, `"kendall"` or `"spearman"`. Where `"pearson"` is the default vlaue.

Estimating copulas with $\hat{\tau}$ or $\hat{\rho}$

Definition

We define the inverse tau esitmator as

$$\hat{\theta}_K = \arg\min_{\theta \in \Theta} \left(\hat{\tau}\big(\mathbf{X}, \mathbf{Y}\big) - \tau_\theta\right)^2.$$

Definition

We define the inverse rho esitmator as

$$\hat{\theta}_S = \arg\min_{\theta \in \Theta} \left(\hat{\rho}\big(\mathbf{X}, \mathbf{Y}\big) - \rho_\theta\right)^2.$$

Note:

In cases where $\hat{\tau}$ or $\hat{\rho}$ take values which cannot be represented by a given copula family the estimates might be rather missleading.

ifgi
Institute for Geoinformatics
University of Münster

Some nice copula families

Example

$$\hat{\theta}_K(\mathbf{X},\mathbf{Y}) := f\big(\hat{\tau}(\mathbf{X},\mathbf{Y})\big).$$

While $f(x)$ takes one of the following forms and we will give $\hat{\theta}_K$ a superscript accordingly:

$f_G(x) := 1/(1-x),\ 0 \le x < 1$	Gumbel, $\Theta_G = [1,\infty)$
$f_C(x) := 2x/(1-x),\ x < 1$	Clayton, $\Theta_C = [-1,\infty)$
$f_N(x) := \sin\left(\frac{1}{2}\pi x\right)$	Gaussian, $\Theta_N = [-1,1]$
$f_t(x) := f_N(x)$	Student, $\Theta_t = [-1,1]$

ifgi
Institute for Geoinformatics
University of Münster

Moment based estimator in R

ifgi
Institute for Geoinformatics
University of Münster

The function `fitCopula()` provides both estimation methods as well. The argument `method` needs to be changed to `"itau"` or `"irho"` respectively.

```
fitCopula(frankCopula(.4),uranium.biv,method="itau")
The estimation is based on the inversion of Kendall's tau
and a sample of size 655.
      Estimate Std. Error  z value     Pr(>|z|)
param 1.210628  0.3102544 3.902051 9.538113e-05

fitCopula(frankCopula(.4),uranium.biv,method="irho")
The estimation is based on the inversion of Spearman's rho
and a sample of size 655.
      Estimate Std. Error  z value   Pr(>|z|)
param 1.198800  0.6369692 1.882037 0.05983096
```

GOF for Copulas I

ifgi
Institute for Geoinformatics
University of Münster

For an empirical copula C_n and $C_{\hat{\theta}}$ the *Cramér-von Mises test-statistic* for $H_0 : C = C_{\hat{\theta}}$ is given by:

$$S_n := \int_{\mathbf{I}^2} n\big(C_n(u,v) - C_{\hat{\theta}}(u,v)\big)^2 \mathrm{d}C_n(u,v)$$

For numerical evaluation purposes the Riemann sum approximate can be used:

$$\tilde{S}_n := \sum_{i=0}^{n} \big(C_n(u_i,v_i) - C_{\hat{\theta}}(u_i,v_i)\big)^2$$

Where $\big((u_1,v_1),\ldots,(u_n,v_n)\big)$ is the transformed sample with margins on $(0,1)$.

GOF for Copulas II

Kendall's Cramér-von Mises test-statistic is defined as:

$$S_n^K := \int_{\mathbf{I}} n(K_n(v) - K_{\hat{\theta}}(v))^2 \mathrm{d}K_{\hat{\theta}}(v)$$

For ease of numerical evaluation its Riemann sum approximate can be used

$$\begin{aligned}\tilde{S_n^K} := \frac{n}{3} \quad &+ \quad n \sum_{i=1}^{n-1} K_n(u_i)^2 \big(K_{\hat{\theta}}(u_{i+1}) - K_{\hat{\theta}}(u_i)\big) \\ &- \quad n \sum_{i=1}^{n-1} K_n(u_i) \big(K_{\hat{\theta}}(u_{i+1})^2 - K_{\hat{\theta}}(u_i)^2\big)\end{aligned}$$

while $u_1 \leq \ \ldots \ \leq u_n$ are the ordered values of $\{V_1, \ldots, V_n\}$, $V_i := C_n(F_n(x_i), G_n(y_i)),\ i = 1, \ldots, n$, $K_n(v) := \frac{1}{n} \# \left\{ k \in \{1, \ldots, n\} \middle| V_k \leq v \right\},\ v \in \mathbf{I}$ and $K_\theta(t) := \int_{\mathbf{I}^2} 1_{C_\theta(u,v) \leq t} \, \mathrm{d}C_\theta(u, v)$.

ifgi
Institute for Geoinformatics
University of Münster

GOF for Copulas III

An approximate p-value can be achieved by:

i) Estimate θ from the sample through one of the given estimators and calculate its empirical copula C_n (Kendall distribution K_n).

ii) Compute the test-statistic $s_0 := \tilde{S}_n$ ($s_0 := \tilde{S}_n^K$).

iii) Simulate a sample $\bar{\mathbf{Z}}$ from the copula C_θ of the same size as the original one and calculate their corresponding rank statistics.

iv) Estimate $\bar{\theta}$ from $\bar{\mathbf{Z}}$ through the same estimator as above and calculate its empirical copula $\bar{C}_n$ (Kendall distribution $\bar{K}_n$).

v) Repeat the steps iii) and iv) for a large integer N and compute its test-statistic $s_i := \tilde{S}_n$ ($s_i := \tilde{S}_n^K$) for any $1 \leq i \leq N$.

vi) The approximate p-value is $\#\{1 \leq i \leq N \mid s_i > s_o\}/N$.

Further details are discussed in [Genest 2007].

ifgi
Institute for Geoinformatics
University of Münster

Graphical GOF I

A graphical tool to decide which copula fits best can as well be deduced from the Kendall distribution function (proposed e.g. in [Genest 2006]). We define an empirical and theoretical version of a function $\lambda : \mathbf{I} \to [-1, 1]$ respectively by

$$\lambda_n(v) := v - K_n(v)$$

and

$$\lambda_\theta(v) := v - K_\theta(v).$$

A comparison of λ_n with (maybe multiple) λ_θ in a single plot may give some guidance.

Graphical GOF II

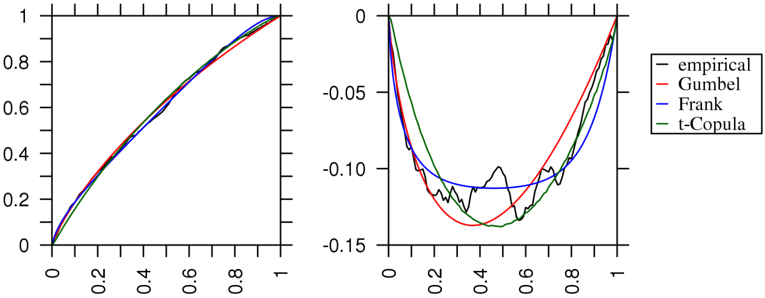


Figure: Comparison of empirical and theoretical Kendall distribution $K(v)$ (left) and $\lambda(v)$ (right).

TASK - estimate some copulas

1. compare Kendall's tau, Spearman's rho and Pearson's correlation coefficient with each other for some bivariate random numbers generated by a copula, and some data set (zinc with lead, ...).
2. plot all three correlation measures for a set of generating parameters (≈ 10).
3. estimate a bivariate copula using the `copula` package for the `uranium` dataset, compare different families and choose the best fitting one

ifgi
Institute for Geoinformatics
University of Münster

- Genest C., Quessy J.-F., Rémillard B. (2006): Goodness-of-fit Procedures for Copula Models Based on the Probability Integral Transformation. Scandinavian Journal of Statistics Vol. 33, pp. 337-366
- Genest C. et al. (2007): Goodness-of-fit tests for copulas: A review and a power study. Insurance: Mathematics and Economics
- Nelsen R. B. (2006): An Introduction to Copulas, 2nd Edition, Springer Science+Buisness, New York