technische universität
dortmund

Almost opposite regression dependence in bivariate distributions

Karl Friedrich Siburg

Preprint 2013-16 Dezember 2013

Fakultät für Mathematik
Technische Universität Dortmund
Vogelpothsweg 87
44227 Dortmund

tu-dortmund.de/MathPreprints

Almost opposite regression dependence in bivariate distributions

Karl Friedrich Siburg
Fakultät für Mathematik
Technische Universität Dortmund
Vogelpothsweg 87
44227 Dortmund
Germany

Pavel A. Stoimenov
International Energy Company
Vienna
Austria

December 9, 2013

Abstract

Let X, Y be two continuous random variables. Investigating the regression dependence of Y on X, respectively, of X on Y, we show that the two of them can have almost opposite behavior. Indeed, given any $\varepsilon > 0$, we construct a bivariate random vector (X, Y) such that the respective regression dependence measures $r_{2|1}(X,Y), r_{1|2}(X,Y) \in [0,1]$ introduced in Dette et al. (2013) satisfy $r_{2|1}(X,Y) = 1$ as well as $r_{1|2}(X,Y) < \varepsilon$.

1 Introduction and results

Recently, Dette et al. (2013) presented a new approach to the problem of ordering and measuring regression dependence in the bivariate case. Let (X, Y)

be a bivariate random vector. Since regression dependence is a directional relationship, it is first necessary to specify the direction of interest. Without loss of generality, consider the dependence of Y on X. The fundamental idea behind regression is predictability – the more predictable Y is from X, the more regression dependent they are. It is straightforward to single out the two extreme cases: independence and almost sure functional dependence, when there exists a Borel measurable function g such that $Y = g(X)$ with probability one (Lancaster, 1963). In the former case, X provides no information about Y, whereas in the latter case there is perfect predictability of Y from X.

Apart from the two extreme cases, however, there exists a variety of intermediate ones with a certain degree of regression dependence. In order to measure the strength of dependence of Y on X, Dette et al. (2013) defined a nonparametric measure of regression dependence, $r_{2|1}(X,Y) \in [0,1]$. Beside being monotone in a regression dependence order, the measure takes on its extreme values precisely at independence and almost sure functional dependence, respectively, i.e., we have

(i) $r_{2|1}(X,Y) = 1$ if and only if Y is a.s. a Borel function of X.

(ii) $r_{2|1}(X,Y) = 0$ if and only if X and Y are independent.

Analogously, one can define a measure $r_{1|2}(X,Y) = r_{2|1}(Y,X)$ measuring the degree of dependence of X on Y.

We point out that it is important to have equivalences in both of the properties (i) and (ii), because only then the value $r_{2|1}(X,Y)$ can serve as a genuine measure of how much Y is dependent on X. Indeed, if we only had $r_{2|1}(X,Y) = 0$ if (but not only if) X and Y are independent, then an assertion like $r_{2|1}(X,Y) < \varepsilon$ would not imply that Y is 'almost independent' from X.

The following is the main result of the present paper.

Theorem 1. *For any given $\varepsilon > 0$, there is a random vector (X,Y) such that* *the following assertions hold:*

1. $r_{2|1}(X,Y) = 1$, *i.e., Y is a.s. a Borel function of X.*

2. $r_{1|2}(X,Y) < \varepsilon$.

The paper is organized as follows. In Section 2 we give a quick review of the construction in Dette et al. (2013) of the nonparametric measure $r_{2|1}$ of regression dependence. Section 3 then contains the proof of Theorem 1, and relates this result to other problems in the literature.

2 Preliminaries

In this section we recall the basic notion of copula and the definition of the nonparametric measure of regression dependence introduced in Dette et al. (2013). A (two-dimensional) copula is a function $C : I^2 \to I$ with $I := [0,1]$, satisfying the following conditions:

1. $C(x,0) = C(0,y) = 0$ for all $x,y \in I$
2. $C(x,1) = x$ and $C(1,y) = y$ for all $x,y \in I$
3. C is 2-increasing, i.e., $C(x_2,y_2) - C(x_2,y_1) - C(x_1,y_2) + C(x_1,y_1) \geq 0$ for all rectangles $[x_1,x_2] \times [y_1,y_2] \subset I^2$.

These conditions imply further key properties. A copula is Lipschitz continuous and increasing in each argument; therefore, its partial derivatives exist a.e. on I^2. We refer the reader to Nelsen (2006) for more information about copulas.

Given two continuous random variables X and Y with corresponding copula C, the measure of regression dependence $r_{2|1}(X,Y)$ introduced in Dette et al. (2013) is defined by

$$r_{2|1}(X,Y) = 6\|\partial_1 C\|_2^2 - 2 = 6\int_{I^2} |\partial_1 C(x,y)|^2 \mathrm{d}(x,y) - 2 \tag{1}$$

where ∂_1 denotes the partial derivative with respect to the first variable and $\|\cdot\|_2$ is the L^2-norm on I^2. The quantity $r_{2|1}$ measures the degree of dependence of Y on X. It is a measure of regression dependence with respect to two natural regression dependence orders, also introduced in Dette et al. (2013).

Analogously, one can define a measure

$$r_{1|2}(X,Y) = 6\|\partial_2 C\|_2^2 - 2 = r_{2|1}(Y,X)$$

such that this quantity measures the degree of dependence of X on Y.

3 Two proofs of Theorem 1

In this section, we will construct two sequences (X_n, Y_n) of bivariate random vectors such that

$$r_{2|1}(X_n,Y_n) = 1 \text{ for all } n, \tag{2}$$

$$\lim_{n\to\infty} r_{1|2}(X_n,Y_n) = 0. \tag{3}$$

This proves Theorem 1. In fact, we will construct sequences of copulas C_n rather than the random variables themselves. This is sufficient because the measures $r_{2|1}$ and $r_{1|2}$ depend only on the corresponding copula. For the construction of these copulas, we use the so-called gluing method developed in Siburg and Stoimenov (2008a). For the convenience of the reader, we quickly recall its definition.

Given two copulas C_1, C_2 and a parameter $\theta \in (0,1)$, we define the function

$$(C_1 \circledast_{x=\theta} C_2)(x,y) = \begin{cases} \theta C_1\left(\frac{x}{\theta}, y\right) & \text{if } 0 \le x \le \theta \\ (1-\theta) C_2\left(\frac{x-\theta}{1-\theta}, y\right) + \theta y & \text{if } \theta \le x \le 1 \end{cases} \tag{4}$$

Thus, $C_1 \circledast_{x=\theta} C_2$ corresponds to gluing the two copulas C_1 and C_2: it equals C_1, rescaled and fit into the rectangle $[0,\theta]\times I$, and equals $C_2 + \theta y$, rescaled

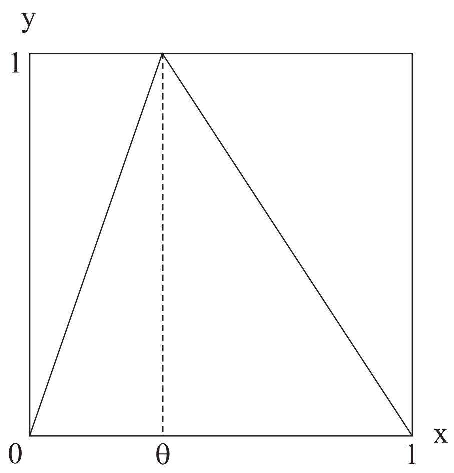

Figure 1: The support of the singular copula C in Example 1

and fit into $[\theta,1]\times I$. It is shown in Siburg and Stoimenov (2008a) that the gluing process yields a copula again, i.e., $C_1 \circledast_{x=\theta} C_2$ is a copula for any parameter θ. For later purposes, we need also the gradient of the resulting copula which is given by

$$\nabla(C_1 \circledast_{x=\theta} C_2)(x,y) = \begin{cases} \left(\frac{\partial C_1}{\partial x}\left(\frac{x}{\theta},y\right), \theta\frac{\partial C_1}{\partial y}\left(\frac{x}{\theta},y\right)\right) & \text{if } 0\le x\le\theta \\ \left(\frac{\partial C_2}{\partial x}\left(\frac{x-\theta}{1-\theta},y\right), (1-\theta)\frac{\partial C_2}{\partial y}\left(\frac{x-\theta}{1-\theta},y\right)+\theta\right) & \text{if } \theta\le x\le 1 \end{cases} \tag{5}$$

Let us first illustrate the glueing construction with a fundamental example. Recall that a copula C is called singular if its density $\partial^2 C/\partial x\partial y$ vanishes almost everywhere in I^2. Moreover, the support of a copula C is defined as the complement of the union of all (relatively) open subsets of I^2 whose measure, induced by C, is zero. We refer to Nelsen (2006) for more details.

Example 1. Let $\theta\in(0,1)$, and suppose that the probability θ is uniformly distributed along the line segment joining $(0,0)$ and $(\theta,1)$, and the probability $1-\theta$ is uniformly distributed along the segment between $(\theta,1)$ and

$(1,0)$. Consider the resulting singular copula C_θ whose support consists of these two line segments; see Figure 1. It follows (see (Nelsen, 2006, Ex. 3.3)) that

$$C_\theta(x,y) = \begin{cases} x & \text{if } x \le \theta y \\ \theta y & \text{if } \theta y < x < 1-(1-\theta)y \\ x+y-1 & \text{if } 1-(1-\theta)y \le x. \end{cases}$$

Note that C_θ can be written as the gluing

$$C_\theta = C^+ \circledast_{x=\theta} C^-$$

where $C^+(x,y) = \min(x,y)$ and $C^-(x,y) = \max(x+y-1,0)$ are the upper and lower Fréchet-Hoeffding bound, respectively.

Since the support of C_θ is a graph over the x-axis, this copula links random variables X and Y where Y is completely dependent on X. This follows from Dette et al. (2013, Prop. 1) and the fact that a function is Borel measurable if and only if its graph is Borel measurable and has probability one (Buckley, 1974). On the other hand, X is not completely dependent on Y because the support of C_θ is not a graph over the y-axis.

This example will serve as a fundamental building block for our final construction of copulas C_n satisfying (2) and (3).

First proof of Theorem 1. We start with the copula $C^+ \circledast_{x=\theta} C^-$ from Example 1 where, in order to simplify calculations, we set $\theta = 1/2$. Then we define C_n inductively by

$$C_1 = C^+ \circledast_{x=1/2} C^-$$
$$C_{n+1} = C_n \circledast_{x=1/2} C_n$$

for $n \ge 1$. We claim that

$$\int_{I^2} |\partial_1 C_n(x,y)|^2 \mathrm{d}(x,y) = \frac{1}{2} \tag{6}$$

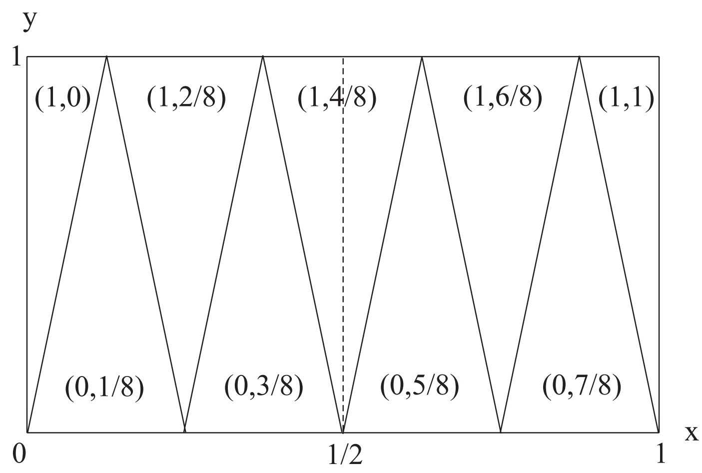

Figure 2: The gradient of the copula C_3 in the first proof of Theorem 1

for all $n \geq 1$, as well as

$$\int_{I^2} |\partial_2 C_n(x,y)|^2 \mathrm{d}(x,y) \to \frac{1}{3} \tag{7}$$

as $n \to \infty$. These relations imply that

$$r_{2|1}(X,Y) = 6 \int_{I^2} |\partial_1 C_n(x,y)|^2 \mathrm{d}(x,y) - 2 = 1$$

for all n, as well as

$$r_{1|2}(X,Y) = 6 \int_{I^2} |\partial_2 C_n(x,y)|^2 \mathrm{d}(x,y) - 2 \to 0$$

as $n \to \infty$, which are precisely the assertions (2) and (3) that we wanted to prove.

For the proof of (6) and (7), we have to calculate the gradient ∇C_n. Using (5) and the fact that $1 - \theta = \theta = 1/2$, we see that $\partial C_n / \partial x = 1$ in the upper and $\partial C_n / \partial x = 0$ in the lower triangles formed by the line segments of the support of C_n, and the second component $\partial C_n / \partial y$ takes the values $0, 1/2^n, 2/2^n, \ldots, (2^n - 1)/2^n, 1$ respectively; see Figure 2 for the case $n = 3$.

Since the gradient of C_n is constant on each triangle, the integration reduces to multiplying the square of the respective constant with the area of the corresponding triangle. Thus, considering the first component of the gradient, we obtain

$$\int_{I^2} |\partial_1 C_n(x,y)|^2 \mathrm{d}(x,y) = \frac{1}{2}$$

for each $n \geq 1$, proving (6).

The integral for the second component amounts to

$$\int_{I^2} |\partial_2 C_n(x,y)|^2 \mathrm{d}(x,y) = \left[\sum_{i=1}^{2^n-1} \left(\frac{i}{2^n} \right)^2 \cdot \frac{1}{2^n} \right] + 1^2 \cdot \frac{1}{2^{n+1}}$$

where the last term stems from the triangle containing the vertex $(1,1)$ which is just half as big as the other ones. Using the formula

$$\sum_{i=1}^{k-1} i^2 = k^3/3 + \mathscr{O}(k^2)$$

we conclude that

$$\int_{I^2} |\partial_2 C_n(x,y)|^2 \mathrm{d}(x,y) = \frac{1}{2^{n+1}} + \left(\frac{1}{2^n} \right)^3 \cdot \sum_{i=1}^{2^n-1} i^2 = \frac{1}{3} + \mathscr{O}\left(\frac{1}{2^n} \right)$$

as $n \to \infty$, proving also our claim (7). □

We conclude this section with a second proof of Theorem 1 where we use an even simpler building block than in the previous one.

Second proof of Theorem 1. Choosing C^+ as a building block instead of $C^+ \circledast_{x=1/2} C^-$, we consider the copulas

$$C_1 = C^+$$
$$C_{n+1} = C_n \circledast_{x=1/2} C_n$$

for $n \geq 1$. We claim that both (6) and (7) hold also for this choice of copula C_n.

Setting $\theta = 1/2$ in (5), one sees that $\partial_1 C_n$ takes the values 0 and 1, each in 2^{n-1} triangles of area $1/2^n$; compare Figure 3 indicating the gradient of

$$C_3 = (C^+ \circledast_{x=1/2} C^+) \circledast_{x=1/2} (C^+ \circledast_{x=1/2} C^+).$$

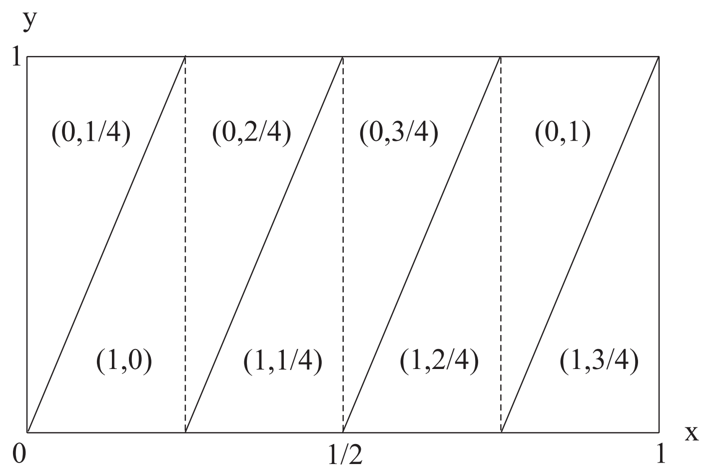

Figure 3: The gradient of the copula C_3 in the second proof of Theorem 1

Therefore,

$$\int_{I^2} |\partial_1 C_n(x,y)|^2 \mathrm{d}(x,y) = \frac{1}{2}$$

for each $n \geq 1$, proving (6).

The second component $\partial_2 C_n$ takes the values $0, 1/2^{n-1}, 2/2^{n-1}, \ldots, (2^{n-1} - 1)/2^{n-1}, 1$ respectively, so that we obtain

$$\begin{aligned}\int_{I^2} |\partial_2 C_n(x,y)|^2 \mathrm{d}(x,y) &= \left[\sum_{i=1}^{2^n-1} \left(\frac{i}{2^n}\right)^2 \cdot \frac{1}{2^n}\right] + 1^2 \cdot \frac{1}{2^{n+1}} \\ &= \frac{1}{2^{n+1}} + \left(\frac{1}{2^n}\right)^3 \cdot \sum_{i=1}^{2^n-1} i^2 \\ &= \frac{1}{3} + \mathscr{O}\left(\frac{1}{2^n}\right)\end{aligned}$$

for $n \to \infty$, proving also (7). □

Finally, we would like to point out that these examples also provide a positive answer to a question stated in Siburg and Stoimenov (2008b). Namely, for both our examples above we have

$$\lim_{n\to\infty} \int_{I^2} |\nabla C_n(x,y)|^2 \mathrm{d}(x,y) = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}, \tag{8}$$

which shows that the bound $5/6$ given in Siburg and Stoimenov (2008b, Thm. 18(ii)) and Siburg and Stoimenov (2010, Thm. 4.3(ii)) is sharp.

References

Buckley, J. J. (1974). Graphs of measurable functions. *Proc. Amer. Math. Soc. 44*(1), 78–80.

Dette, H., K. F. Siburg, and P. A. Stoimenov (2013). A copula-based non-parametric measure of regression dependence. *Scand. J. Stat. 40*(1), 21–41.

Lancaster, H. O. (1963). Correlation and complete dependence of random variables. *Ann. Math. Statist. 34*, 1315–1321.

Nelsen, R. B. (2006). *An introduction to copulas* (second ed.). New York: Springer.

Siburg, K. F. and P. A. Stoimenov (2008a). Gluing copulas. *Comm. Statist. Theory Methods 37*(19), 3124–3134.

Siburg, K. F. and P. A. Stoimenov (2008b). A scalar product for copulas. *J. Math. Anal. Appl. 344*(1), 429–439.

Siburg, K. F. and P. A. Stoimenov (2010). A measure of mutual complete dependence. *Metrika 71*(2), 239–251.

Preprints ab 2012/01

2013-16 **Karl Friedrich Siburg**
Almost opposite regression dependence in bivariate distributions

2013-15 **Christian Palmes and Jeannette H. C. Woerner**
The Gumbel test and jumps in the volatility process

2013-14 **Karl Friedrich Siburg, Katharina Stehling, Pavel A. Stoimenov, Jeannette H. C. Woerner**
An order for asymmetry in copulas, and implications for risk management

2013-13 **Michael Voit**
Product formulas for a two-parameter family of Heckman-Opdam hypergeometric functions of type BC

2013-12 **B. Schweizer and M. Veneroni**
Homogenization of plasticity equations with two-scale convergence methods

2013-11 **Sven Glaser**
A law of large numbers for the power variation of fractional Lévy processes

2013-10 **Christian Palmes and Jeannette H.C. Woerner**
The Gumbel test for jumps in stochastic volatility models

2013-09 **Agnes Lamacz, Stefan Neukamm, and Felix Otto**
Moment bounds for the corrector in stochastic homogenization of a percolation model

2013-08 **Frank Klinker**
Connections on Cahen-Wallach spaces

2013-07 **Andreas Rätz and Matthias Röger**
Symmetry breaking in a bulk-surface reaction-diffusion model for signaling networks

2013-06 **Gilles Francfort and Ben Schweizer**
A doubly non-linear system in small-strain visco-plasticity

2013-05 **Tomáš Dohnal**
Traveling Solitary Waves in the Periodic Nonlinear Schrödinger Equation with Finite Band Potentials

2013-04 **Karl Friedrich Siburg, Pavel Stoimenov, and Gregor N.F. Weiß**
Forecasting Portfolio-Value-at-Risk with Nonparametric Lower Tail Dependence Estimates

2013-03 **Martin Heida**
On thermodynamics of fluid interfaces

2013-02 **Martin Heida**
Existence of soulutions for two types of generalized versions of the Cahn-Hilliard equation

2013-01 **T. Dohnal, A. Lamacz, B. Schweizer**
Dispersive effective equations for waves in heterogeneous media on large time scales

2012-19 **Martin Heida**
On gradient flows of nonconvex functionals in Hilbert spaces with Riemannian metric and application to Cahn-Hilliard equations

2012-18 **R.V. Kohn, J. Lu, B. Schweizer, and M.I. Weinstein**
A variational perspective on cloaking by anomalous localized resonance

2012-17 **Margit Rösler and Michael Voit**
Olshanski spherical functions for infinite dimensional motion groups of fixed rank

2012-16 **Selim Esedoḡlu, Andreas Rätz, Matthias Röger**
Colliding Interfaces in Old and New Diffuse-interface Approximations of Willmore-flow

2012-15 **Patrick Henning, Mario Ohlberger and Ben Schweizer**
An adaptive multiscale finite element method

2012-14 **Andreas Knauf, Frank Schulz, Karl Friedrich Siburg**
Positive topological entropy for multi-bump magnetic fields

2012-13 **Margit Rösler, Tom Koornwinder, and Michael Voit**
Limit transition between hypergeometric functions of type BC and Type A

2012-12 **Alexander Schnurr**
Generalization of the Blumenthal-Getoor Index to the Class of Homogeneous Diffusions with Jumps and some Applications

2012-11 **Wilfried Hazod**
Remarks on pseudo stable laws on contractible groups

2012-10 **Waldemar Grundmann**
Limit theorems for radial random walks on Euclidean spaces of high dimensions

2012-09 **Martin Heida**
A two-scale model of two-phase flow in porous media ranging from porespace to the macro scale

2012-08 **Martin Heida**
On the derivation of thermodynamically consistent boundary conditions for the Cahn-Hilliard-Navier-Stokes system

2012-07 **Michael Voit**
Uniform oscillatory behavior of spherical functions of GL_n/U_n at the identity and a central limit theorem

2012-06 **Agnes Lamacz and Ben Schweizer**
Effective Maxwell equations in a geometry with flat rings of arbitrary shape

2012-05 **Frank Klinker and Günter Skoruppa**
Ein optimiertes Glättungsverfahren motiviert durch eine technische Fragestellung

2012-04 **Patrick Henning, Mario Ohlberger, and Ben Schweizer**
Homogenization of the degenerate two-phase flow equations

2012-03 **Andreas Rätz**
A new diffuse-interface model for step flow in epitaxial growth

2012-02 **Andreas Rätz and Ben Schweizer**
Hysteresis models and gravity fingering in porous media

2012-01 **Wilfried Hazod**
Intrinsic topologies on H-contraction groups with applications to semistability