8-2013

Vine Copulas as a Way to Describe and Analyze Multi-Variate Dependence in Econometrics: Computational Motivation and Comparison with Bayesian Networks and Fuzzy Approaches

Songsak Sriboonchitta
Chiang Mai University, songsakecon@gmail.com

Jianxu Liu
Chiang Mai University, liujianxu1984@163.com

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Hung T. Nguyen
New Mexico State University - Main Campus, hunguyen@nmsu.edu

Recommended Citation

Vine Copulas as a Way to Describe and Analyze Multi-Variate Dependence in Econometrics: Computational Motivation and Comparison with Bayesian Networks and Fuzzy Approaches

Songsak Sriboonchitta[1], Jianxu Liu[1], Vladik Kreinovich[2], and
Hung T. Nguyen[1,3]

[1] Department of Economics, Chiang Mai University
Chiang Mai, Thailand, songsakecon@gmail.com, liujianxu1984@163.com
[2] Department of Computer Science, University of Texas at El Paso
500 W. University, El Paso, TX 79968, USA, vladik@utep.edu
[3] Department of Mathematical Sciences, New Mexico State University
Las Cruces, New Mexico 88003, USA, hunguyen@nmsu.edu

Abstract. In the last decade, vine copulas emerged as a new efficient techniques for describing and analyzing multi-variate dependence in econometrics; see, e.g., [1–3, 7, 9–11, 13, 14, 21]. Our experience has shown, however, that while these techniques have been successfully applied to many practical problems of econometrics, there is still a lot of confusion and misunderstanding related to vine copulas. In this paper, we provide a motivation for this new technique from the computational viewpoint. We show that other techniques used to described dependence – Bayesian networks and fuzzy techniques – can be viewed as a particular case of vine copulas.

1 Copulas – A Useful Tool in Econometrics: Motivations and Descriptions

Need for studying dependence in econometrics. Many researchers have observed that economics is more complex than physics. In physics, many parameters, many phenomena are independent. As a result, we can observe (and thoroughly study) simple systems which can be described by a small number of parameters. Based on these simple systems, we can separately determine the laws that describe mechanics, electrodynamics, thermodynamics, etc., and then combine these laws to describe more complex phenomena.

In contrast, in economics, most phenomena are interrelated. Thus, to numerically describe economic phenomena, we need to take into account several dependent parameters. So, in econometrics, studying dependence is of utmost importance.

Statistical character of economic phenomena. An additional complexity of economics – as compared to physics – is that while most physical processes are

deterministic, in economics, we can only make statistical predictions. If we repeatedly drop the same object from the Leaning Tower of Pisa (as Galileo did), we will largely observe the exact same behavior every time. In contrast, if several very similar restaurants open in the same area, some of them will survive and some will not, and it is practically impossible to predict which will survive – at best, we can predict the probability of survival. We can deterministically predict the future trajectory of a spaceship, but we can, at best, make statistical predictions about the future values of a stock index.

Conclusion: we need to study dependence between random variables. Because of the statistical character of economic phenomena, each parameter describing the economics is a random variables. Thus, the need to study dependence means that we need to study dependence between random variables.

Simplest case when random variables are independent: reminder. In order to analyze how to describe dependence of random variables, let us recall how *in*dependent random variables can be described.

In general, a random variable X_i can be described by its cumulative distribution function $F_i(x_i) \stackrel{\text{def}}{=} \text{Prob}(X_i \leq x_i)$. If two random variables X_1 and X_2 are independent, this means that their joint distribution function $F(x_1, x_2) \stackrel{\text{def}}{=} \text{Prob}(X_1 \leq x_1 \,\&\, X_2 \leq x_2)$ is equal to the product of the marginal distributions $F_1(x_1)$ and $F_2(x_2)$: $F(x_1, x_2) = F_1(x_1) \cdot F_2(x_2)$.

Towards describing dependence between two random variables: the notion of a copula. In the independent case, general, the joint distribution function $F(x_1, x_2)$ of two random variables X_1 and X_2 is equal to the product $F_1(x_1) \cdot F_2(x_2)$ of the marginal distributions. In general, when the random variables X_1 and X_2 are dependent, the joint distribution function $F(x_1, x_2)$ is different from the product $F_1(x_1) \cdot F_2(x_2)$. It is reasonable to describe this general joint distribution in such a way that we will clearly see how different is the joint distribution from the independent case. In the independent case, $F(x, x_2)$ is the product of the marginal distributions $F_1(x_1)$ and $F_2(x_2)$; to describe deviations from this product, it make sense to consider more general combination functions, i.e., to consider expressions of the type

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \tag{1}$$

Such combination functions $C(a, b)$ are known as *copulas*; see, e.g., [19, 26] (see also [1–3, 7, 9–11, 13, 14, 21]).

The independence case corresponds to the product combination function $C(a, b) = a \cdot b$. The more the combination function $C(a, b)$ is different from the product, the more dependent are the random variables X_1 and X_2.

Probability density function in terms of the copula. The expression for the probability density function $f(x_1, x_2) = \dfrac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2}$ in terms of the copula can be

obtained by differentiating the above formula with respect to x_1 and x_2. As a result, we get the expression

$$f(x_1, x_2) = c(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2), \qquad (2)$$

where $c(a, b) \overset{\text{def}}{=} \dfrac{\partial^2 C(a, b)}{\partial a \, \partial b}$ and $f_i(x_i) \overset{\text{def}}{=} \dfrac{dF_i(x_i)}{dx_i}$ are probability densities of the marginal distributions.

Can copulas describe all possible dependencies? The expression (1) is a natural generalization of the independence case. At first glance, it may sound that such expressions describe some special class of dependent variables. However, it can be shown that this expression is general enough to capture the general dependence between random variables. Namely, for continuous distributions, e.g., for distributions with well-defined probability density functions, once we know the joint distribution function $F(x_1, x_2)$ and marginal distributions $F_1(x_1)$ and $F_2(x_2)$, we can get the representation (1) if we take $C(a, b) = F(F_1^{-1}(a), F_2^{-1}(b))$, where $F_i^{-1}(a)$ denotes a function which is inverse to the function $F_i(x)$.

Computational advantage of copulas. In many applications of econometrics, it is important not only to have the right models for describing the corresponding phenomena, it is also extremely important to have efficient algorithms which use these models for predicting future values of the corresponding quantities. For example, if several agents have access to the models that can predict the increase in the price of a certain stock, but one of the agents has a faster algorithm for this prediction, then this agent can learn about this future increase before everyone else. This computational advantage will give this agent the opportunity to buy the about-to-increase stock for the current price, and thus, earn a profit when the price of this stock actually increases.

From this viewpoint, it should be noticed that a copula representation indeed speeds up computations. To explain this speed-up, let us start with the case of a single random variable. For a single variable X_i, we can use its observations $x_{i1}, \ldots, x_{iN}$ to estimate the corresponding probability distribution. For example, we can use a histogram distribution, i.e., approximate the probability by the corresponding frequency: $F_i(x_i) = \text{Prob}(X_i \leq x_i) \approx \dfrac{1}{N} \cdot \#\{j : x_{ij} \leq x_i\}$.

Comment. In practice, we rarely use the histogram distribution. Usually, we find a smooth distribution which is sufficient close to the histogram one (e.g., in the sense of the Kolmogorov-Smirnov criterion), so that this smooth distribution is statistically possible, and use the corresponding smooth distribution.

For two random variables X_1 and X_2, we can, in principle, also use the corresponding pairs of observations (x_{1j}, x_{2j}), $1 \leq j \leq N$, and estimate the probability $F(x_1, x_2) = \text{Prob}(X_1 \leq x_1 \,\&\, X_2 \leq x_2)$ as the corresponding frequency $\dfrac{1}{N} \cdot \#\{j : x_{1j} \leq x_1 \,\&\, x_{2j} \leq x_2\}$. From the computational viewpoint, this would

mean, however, that we need to process all N pairs (x_{1j}, x_{2j}) (i.e., all $2N$ numbers x_{1j} and x_{2j}) to find each of the values $F(x_1, x_2)$. Usually, we have a large amount of economic data, so the need to process all the data all the time makes computations longer.

If instead of representing the unknown distribution by its joint distribution function $F(x_1, x_2)$, we use a copula representation, in which a distribution is represented by two marginals $F_1(x_1)$, $F_2(x_2)$, and a copula $C(a, b)$, then, to find each of the marginals $F_i(x_i)$, we only need to process N values x_{ij} ($j = 1, \ldots, N$) (and we only need to process all $2N$ real values to determine the copula $C(a, b)$). This decrease in the number of inputs speed up computations.

Case of three of more variables. As we have mentioned, to adequately describe economic phenomena, we need to use several random variables

$$X_1, \ldots, X_n, \ \ n \gg 2.$$

Each such random tuple can be described by its probability distribution

$$F(x_1, \ldots, x_n) = \mathrm{Prob}(X_1 \le x_1 \& \ldots \& X_n \le x_n). \tag{3}$$

Similarly to the case of two variables, when all the random variables are independent, the joint distribution is equal to the product of all the marginal distributions:

$$F(x_1, \ldots, x_n) = F_1(x_1) \cdot \ldots \cdot F_n(x_n).$$

Similarly to the two-variables case, the general distribution can be obtained by applying an appropriate combination function (copula) $C(a_1, \ldots, a_n)$ to the marginals:

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)). \tag{4}$$

To prove that such a representation is possible for a given joint distribution $F(x_1, \ldots, x_n)$, we can take

$$C(a_1, \ldots, a_n) = F(F_1^{-1}(a_1), \ldots, F_n^{-1}(a_n)). \tag{5}$$

2 From General Copulas to Vine Copulas: Motivations and Descriptions

From the computational viewpoint, additional speed-up is needed. Similarly to the two-variables case, the use of multi-dimensional copulas decreases the computation time. However, this decreased computation time still exponentially increases with the dimension n.

Indeed, a full knowledge about a function $f(x)$ of one variable defined on an interval $[0, 1]$ would mean that we know infinitely many values of this function, corresponding to infinitely many real numbers $x \in [0, 1]$. In practice, we can only

store finitely many values. So, to describe a function in a computer, we select a small step h and only consider $\dfrac{1}{h}$ values

$$f(0), f(h), f(2h), \ldots, f(k \cdot h), \ldots, f(1), \quad k = 1, 2, \ldots, \frac{1}{h}. \tag{6}$$

Similarly, to describe a copula $C(a_1, \ldots, a_n)$, we need to store values

$$C(k_1 \cdot h, \ldots, k_n \cdot h)$$

corresponding to all possible combinations of integers $k_1, \ldots, k_n$ corresponding to $k_i = 1, \ldots, \dfrac{1}{h}$. For each of n variables k_i, we have $\dfrac{1}{h}$ possible values. Thus, the total number of tuples $(k_1, \ldots, k_n)$ is equal to $\dfrac{1}{h^n}$.

Each of these values needs to be estimated and processed. Thus, the resulting computation time is proportional to $\dfrac{1}{h^n}$ and hence, exponentially grows with the number of variables n. For large n, this computation time becomes unrealistically large (see, e.g., [22]) – especially in view of the above-mentioned fact that in econometrics, we need computations to be as fast as possible. Thus, an additional speed-up is needed.

We already know that for two variables, a copula-based description – which only uses functions of two variables – is realistic and practically useful. From this viewpoint, it is desirable to only use functions of two variables in our description of multi-variate distributions. Such a description is possible if we use *vine copulas*. Let us explain how the corresponding vine copula techniques naturally emerge from the analysis of our problem.

Main idea: using conditional probabilities. Our objective is to represent dependence. To arrive at the copula techniques, we started with the description of independence, and we used this description to come up with a general copula-based description of dependence. From the mathematical viewpoint, this copula-based description is sufficient to describe an arbitrary dependence. However, from the computational viewpoint, we need to go beyond the general copula-based formula. To move forward, let us go back to the independence case, and see if there are some other independence-related techniques that we can generalize to the general dependence case.

Our previous analysis was based on the fact that independence between random variables can be described in terms of the product of the corresponding probabilities: $F(x_1, x_2) = F_1(x_1) \cdot F_2(x_2)$. There is, however, an equivalent (and probably more intuitive) description of independence, a representation in term of conditional probabilities: $F_{1|2}(x_1 \mid x_2) = F_1(x_1)$, where

$$F_{1|2}(x_1 \mid x_2) \stackrel{\text{def}}{=} \text{Prob}(X_1 \leq x_1 \mid X_2 = x_2). \tag{7}$$

To relate this representation to the previous one, let us describe the conditional probability in terms of the copula. By definition of the conditional prob-

ability, we have

$$F_{1|2}(x_1 \,|\, x_2) = \text{Prob}(X_1 \le x_2 \,|\, X_2 = x_2) =$$

$$\lim_{\varepsilon \to 0} \text{Prob}(X_1 \le x_2 \,|\, x_2 - \varepsilon \le X_2 \le x_2 + \varepsilon) =$$

$$\lim_{\varepsilon \to 0} \frac{\text{Prob}(X_1 \le x_2 \,\&\, x_2 - \varepsilon \le X_2 \le x_2 + \varepsilon)}{\text{Prob}(x_2 - \varepsilon \le X_2 \le x_2 + \varepsilon)}. \tag{8}$$

The probability in the numerator $\mathcal{N}$ of the corresponding fraction can be described as

$$\mathcal{N} = \text{Prob}(X_1 \le x_1 \,\&\, X_2 \le x_2 + \varepsilon) - \text{Prob}(X_1 \le x_1 \,\&\, X_2 \le x_2 - \varepsilon) =$$

$$F(x_1, x_2 + \varepsilon) - F(x_1, x_2 - \varepsilon). \tag{9}$$

In terms of the corresponding copula $C_{12}(a, b)$ and the marginals $F_1(x_1)$ and $F_2(x_2)$, we get

$$\mathcal{N} = C_{12}(F_1(x_1), F_2(x_2 + \varepsilon)) - C_{12}(F_1(x_1), F_2(x_2 - \varepsilon)). \tag{10}$$

Since ε is small, we get

$$\mathcal{N} \approx 2\varepsilon \cdot \frac{\partial C_{12}(F_1(x_1), F_2(x_2))}{\partial x_2} = 2\varepsilon \cdot C_{1|2}(F_1(x_1), F_2(x_2)) \cdot f_2(x_2), \tag{11}$$

where we denoted $C_{1|2}(a, b) \stackrel{\text{def}}{=} \dfrac{\partial C_{12}(a, b)}{\partial b}$, and $f_2(x_2) = \dfrac{dF_2(x_2)}{dx_2}$ is the probability density of the second marginal distribution.

Similarly, the denominator $\mathcal{D}$ has the form

$$\mathcal{D} = \text{Prob}(X_2 \le x_2 + \varepsilon) - \text{Prob}(X_2 \le x_2 - \varepsilon) = F_2(x_2 + \varepsilon) - F(x_2 - \varepsilon). \tag{12}$$

Since ε is small, we get

$$\mathcal{N} \approx 2\varepsilon \cdot f_2(x_2).$$

Thus, the ratio $F_{1|2}(x_1 \,|\, x_2)$ is equal to:

$$F_{1|2}(x_1 \,|\, x_2) = C_{1|2}(F_1(x_1), F_2(x_2)). \tag{13}$$

The corresponding conditional probability density $f_{1|2}(x_1 \,|\, x_2)$ can be obtained by differentiating both sides of this equation with respect to x_1:

$$f_{1|2}(x_1 \,|\, x_2) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1), \tag{14}$$

where

$$c_{12}(a, b) = \frac{\partial C_{1|2}(a, b)}{\partial a} = \frac{\partial}{\partial a}\left(\frac{\partial C_{12}(a, b)}{\partial b}\right) = \frac{\partial^2 C_{12}(a, b)}{\partial a \, \partial b}.$$

There are two ways to use conditional probabilities to speed up our computations. Let us illustrate both of them on the example of trivariate distributions.

First idea: D-vine copulas. We know how to describe bivariate distributions in terms of copulas: namely, each pair of random variables X_1 and X_2 with a joint distribution $F(x_1, x_2)$ can be represented as $F(x_1, x_2) = C_{12}(F_1(x_1), F_2(x_2))$. We would like to use this idea to describe *three* random variables X_1, X_2, and X_3. A natural idea is to fix the value x_3, and to consider corresponding *conditional* distributions. For each x_3, we can have a similar representation of the corresponding conditional distribution

$$F_{12|3}(x_2, x_2 \,|\, x_3) \stackrel{\text{def}}{=} \text{Prob}(X_1 \leq x_1 \,\&\, X_2 \leq x_2 \,|\, X_2 = x_3) =$$

$$C_{12|3}(F_1(x_1 \,|\, x_3), F_2(x_2 \,|\, x_3), x_3). \tag{15}$$

In general, for different values x_3, we can have different copulas $C(a, b) = C_{12|3}(a, b, x_3)$. These copula describe the dependence between X_1 and X_2. In many practical situations, it makes sense to assume that the dependence between X_1 and X_2 does not depend on the value of X_3. In such situations, the copula $C_{12|3}(a, b)$ which describes this dependence does not depend on x_3: $C_{12|3}(a, b, x_3) = C_{12|3}(a, b)$. Then, the formula (14) takes the simplified form

$$F_{12|3}(x_1, x_2 \,|\, x_3) = C_{12|3}(F_{1|3}(x_1 \,|\, x_3), F_{2|3}(x_2 \,|\, x_3)). \tag{16}$$

We already know how to describe conditional distributions $F_{1|3}(x_1 \,|\, x_3)$ and $F_{2|3}(x_2 \,|\, x_3)$ in terms of bivariate copulas and marginals: specifically, we can use the formula (13). Thus, we can describe the conditional probabilities $F_{12|3}(x_1, x_2 \,|\, x_3)$ in terms of bivariate copulas and marginals.

Our goal is to compute the distribution function $F(x_1, x_2, x_3)$. To describe the corresponding probabilities $F(x_1, x_2, x_3)$ in terms of conditional probabilities $F_{12|3}(x_1, x_2 \,|\, x_3)$, we can use the formula of total probability:

$$F(x_1, x_2, x_3) = \int_{-\infty}^{x_3} F_{12|3}(x_1, x_2 \,|\, z) \cdot f_3(z) \, dz. \tag{17}$$

Combining formulas (13), (16), and (17), we get the following expression of the multivariuate distribution in terms of bivariate copulas and marginal distributions:

$$F(x_1, x_2, x_3) = \int_{-\infty}^{x_3} C_{12|3}(F_1(x_1 \,|\, z), F_{2|3}(x_2 \,|\, z)) \, dz, \tag{18}$$

where

$$F_{1|3}(x_1 \,|\, z) = C_{1|3}(F_1(x_1), F_3(z)), \quad F_{2|3}(x_2 \,|\, z) = C_{2|3}(F_2(x_2), F_3(z)), \tag{19}$$

$C_{1|3}(a, b) \stackrel{\text{def}}{=} \dfrac{\partial C_{13}(a, b)}{\partial b}$, and $C_{2|3}(a, b) \stackrel{\text{def}}{=} \dfrac{\partial C_{23}(a, b)}{\partial b}$. This description is a particular case of a *D-vine copula*.

Second idea: C-vine copulas. The idea behind C-vine copulas comes from considering not directly probabilities and conditional probabilities (as for D-vine copulas), but rather probability *densities* and conditional probability *densities*. A multivariate probability density can be described in terms of conditional probability densities, as

$$f(x_1, x_2, x_3) = f_{1|23}(x_1 \,|\, x_2, x_3) \cdot f_{23}(x_2, x_3). \tag{20}$$

The probability density $f_{23}(x_2, x_3)$ can also be similarly represented as $f_{2|3}(x_2 \,|\, x_3) \cdot f_3(x_3)$, so we conclude that

$$f(x_1, x_2, x_3) = f_{1|23}(x_1 \,|\, x_2, x_3) \cdot f_{2|3}(x_2 \,|\, x_3) \cdot f_3(x_3). \tag{21}$$

We know, from the formula (14), that

$$f_{2|3}(x_2 \,|\, x_3) = c_{23}(F_2(x_2), F_3(x_3)) \cdot f_2(x_2). \tag{22}$$

For dependence $f_{1|23}(x_1 \,|\, x_2, x_3)$, we have a similar formula for each x_3:

$$f_{1|23}(x_1 \,|\, x_2, x_3) = c_{12|3}(F_{1|3}(x_1 \,|\, x_3)), F_{2|3}(x_2 \,|\, x_3), x_3) \cdot f_{1|3}(x_1 \,|\, x_3). \tag{23}$$

In general, the corresponding copula $c_{12|3}$ depends on x_3. However, in many practical situations, it makes sense to assume that this copula – describing the dependence – does not depend on x_3, i.e., that we have

$$f_{1|23}(x_1 \,|\, x_2, x_3) = c_{12|3}(F_{1|3}(x_1 \,|\, x_3)), F_{2|3}(x_2 \,|\, x_3)) \cdot f_{1|3}(x_1 \,|\, x_3). \tag{24}$$

We already know how to describe conditional distributions $F_{1|3}(x_1 \,|\, x_3)$ and $F_{2|3}(x_2 \,|\, x_3)$ and conditional probability density $f_{1|3}(x_1 \,|\, x_3)$ in terms of bivariate copulas and marginals: specifically, we can use the formulas (13) and (14). Thus, we can describe the conditional probability density $f_{1|23}(x_1 \,|\, x_2, x_3)$ in terms of bivariate copulas and marginals. By combining the formulas (21), (22), and (24), we get

$$f(x_1, x_2, x_3) = c_{12|3}(F_{1|3}(x_1 \,|\, x_3)), F_{2|3}(x_2 \,|\, x_3)) \cdot f_{1|3}(x_1 \,|\, x_3) \cdot$$

$$c_{23}(F_2(x_2), F_3(x_3)) \cdot f_2(x_2) \cdot f_3(x_3), \tag{25}$$

where

$$F_{1|3}(x_1 \,|\, x_3) = C_{1|3}(F_1(x_1), F_3(x_3)); \quad F_{2|3}(x_2 \,|\, x_3) = C_{2|3}(F_2(x_2), F_3(x_3));$$

$$f_{1|3}(x_1 \,|\, x_3) = c_{13}(F_1(x_1), F_3(x_3)) \cdot f_1(x_1). \tag{26}$$

This description is a particular case of a *C-vine copula*.

Comment. Similar expressions can be obtained for any number of variables. To get such an expression, we need to make some assumptions about copula independence. Depending on which assumptions we make, we get different expressions. For example, the above expression (25)–(26) corresponds to the case when we assume that the copula combining:

- the conditional dependence $F_{1|3}(x_1 \,|\, x_3)$ of x_1 on x_3 and
- the conditional dependence $F_{2|3}(x_2 \,|\, x_3)$ of x_2 on x_3

into a conditional joint dependence $F_{12|3}(x_1, x_2 \,|\, x_3)$ of x_1 and x_2 on x_3 does not depend on x_3. Alternatively, we could assume that the copula combining:

- the conditional dependence $F_{2|1}(x_2 \,|\, x_1)$ of x_2 on x_1 and
- conditional dependence $F_{3|1}(x_3 \,|\, x_1)$ of x_3 on x_1

into a conditional joint dependence $F_{23|1}(x_2, x_3 \,|\, x_1)$ of x_2 and x_3 on x_1 does not depend on x_1; this would lead to a different expression of the type (25)–(26).

How do we select a model? In some cases, from the econometric context, we know which dependencies are independent in each variables. In many practical situations, however, such an information is not available. In such situations, out of models corresponding to different dependencies, we need to select the model which is the best fit for the observations.

3 Comparing Vine Copulas with Other Techniques for Describing Multi-Variate Dependence

Vine copulas vs. general copulas. Vine copulas are a practically important class of copulas: they only use bivariate functions to describe a multi-variate dependence and are, thus, computationally easier (and more feasible) to implement.

It is important to remember, however, that vine copulas *do not* describe a general dependence. As we have mentioned earlier, vine copulas are based on certain *independence* assumptions: e.g., that the copula that transforms the conditional distributions $F_{1|3}(x_1 \,|\, x_3)$ and $F_{2|3}(x_2 \,|\, x_3)$ into a joint conditional distribution $F_{12|3}(x_1, x_2 \,|\, x_3)$ does not depend on the value x_3.

It is worth mentioning that vine copulas' inability to represent a general function of three or more variables is not a drawback of any particular scheme, but rather a general property of smooth (differentiable) functions. Namely, as part of the work on D. Hilbert's 13th problem – one of the famous 23 problems presented in 1900 as a challenge to 20 century mathematics – a Russian mathematician A. G. Vitushkin proved that for any given integer N, it is not possible to represent (or even approximate) a general smooth function of three (or more) variables as a composition of functions of two or fewer variables; see, e.g., [5, 16, 27–29].

Vine copulas vs. Bayesian networks. Another approach actively used in applications to represent multivariate dependence is the approach of Bayesian networks, initiated by Judea Pearl; see, e.g., [18, 23–25]. Bayesian newtorks are based on the assumption that for some variables, the corresponding conditional distributions are independent. For example, for the case of three variables, a typical assumption is that the conditional distributions $F_{1|3}(x_1 \,|\, x_3)$ and $F_{2|3}(x_2 \,|\, x_3)$ are independent, i.e., that

$$F_{12|3}(x_1, x_2 \,|\, x_3) = F_{1|3}(x_1 \,|\, x_3) \cdot F_{2|3}(x_2 \,|\, x_3). \tag{27}$$

One can easily see that the resulting formula is a particular case of the vine copula formula (16), corresponding to $C_{1|2}(a, b) = a \cdot b$. Thus, the Bayesian network approach can be viewed as a particular case of the general vine copula approach.

Vine copulas vs. fuzzy techniques. Another practically successful approach for describing and analyzing multivariate dependence is an approach of fuzzy techniques; see, e.g., [12, 20, 30].

One of the main ideas behind fuzzy techniques is that

- while we can extract, from the experts, their degrees of confidence (= subjective probability) in different possible statements $S_1, S_2, \ldots, S_n$ about their domain of expertise,
- it is not realistically possible to extract, from the users, their degrees of confidence in different logical combinations of such statements, such as $S_i \,\&\, S_j$ or $S_i \,\&\, S_j \,\&\, S_k$ – since there are, in general, exponentially many (2^n) such combinations.

Since we cannot elicit all the values, we need to estimate the degree of confidence in a statement $S \,\&\, S'$ based on the known degrees of confidence $d(S)$ and $d(S')$ in component statements S and S'. The algorithm $f_\&(a, b)$ which transforms the known degrees $a = d(S)$ and $b = d(S')$ into an estimate $f_\&(d(S), d(S'))$ for the desired degree $d(S \,\&\, S')$ is known as an *"and"-operation* or a *t-norm*.

From the mathematical viewpoint, there are many possible t-norms. In practice, a t-norm is selected empirically, based on the cases when we do elicit the expert's degree of confidence $d(S \,\&\, S')$ in the composite statement $S \,\&\, S'$. Once these values are known, we select a function $f_\&(a, b)$ for which $f_\&(d(S), d(S')) \approx d(S \,\&\, S')$ for all such pairs of statements.

The resulting "and"-operation depends on the domain. Such an empirical determination was first implemented for the world's first practically successful expert system, a medical expert system MYCIN intended for diagnosing rare blood diseases; see, e.g., [6]. It is worth mentioning that the authors of the corresponding empirical study initially thought that the resulting "and"-operation is a general description of human reasoning. Alas, when they applied their idea to geophysics, it turned out that the medically best "and"-operation is not appropriate for geophysics at all. After the fact, it makes sense: e.g., in search for oil, it makes sense to start drilling a well once there is a reasonable expectation that this well will be productive – and it is OK that a large portion of

these wells do not produce, as long as on average, we are successful. In contrast, in medicine, we do not want to perform a serious surgery on a patient unless we are absolutely sure about the diagnosis. In short, in medicine, experts use very conservative estimates, while in geophysics, they use more optimistic ones. As a result, different application domains use different "and"-operations – but the same "and"-operation is useful for all statements within a given application domain.

The main problem that we solve by using copulas can be described in similar terms. Namely, we have two statements $S = $ "$X_1 \leq x_1$" and $S' = $ "$X_2 \leq x_2$", whose probabilities are values of the marginal distributions $d(S) = F_1(x_1)$ and $d(S') = F_2(x_2)$. The logical combination $S \& S'$ is the statement

$$X_1 \leq x_1 \& X_2 \leq x_2$$

whose probability is equal to $F(x_1, x_2)$. Our objective is to transform the known degrees $d(S) = F_1(x_1)$ and $d(S') = F_2(x_2)$ into an estimate $f_\&(d(S), d(S')) = f_\&(F_1(x_1), F_2(x_2))$ for $F(x_1, x_2)$:

$$F(x_1, x_2) \approx f_\&(F_1(x_1), F_2(x_2)). \tag{28}$$

From this viewpoint, the copula is an "and"-operation.

The fuzzy approach can be viewed as a particular case of the vine copula approach, The main difference between fuzzy approach and the general vine copula approach is that:

- in the fuzzy case, the same "and"-operation is used to combine the probabilities corresponding to different variables, while
- in the general vine copula approach, we can use different copulas to combine the probabilities of different pairs of variables.

Summarizing our analysis. Vine copulas are a particular case of general copulas, and Bayesian network and fuzzy approaches can be viewed as particular cases of the vine copula approach:

General copulas
$\downarrow$
Vine copulas
$\swarrow \qquad \searrow$
Bayesian $\qquad$ Fuzzy
networks $\qquad$ techniques
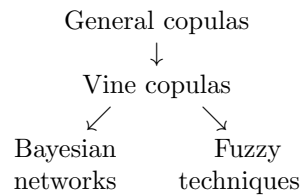
Vine copula approach combines advantages of Bayesian and fuzzy approaches. Both Bayesian networks and fuzzy techniques have numerous successful applications. The very fact that both techniques have been successful means that for each of these techniques, there is an application areas where this particular

technique works well. The fact that both techniques co-exist seems to indicate that for each of these techniques, there are application areas where the other technique works better.

In other words, each of these techniques has its own advantages and limitations. Numerous researchers have expressed the desire to come up with a new technique that would combine the advantages of both techniques – and have none of their limitations. From this viewpoint, the vine copula approach, an approach of which both Bayesian network and fuzzy techniques are particular cases, seems like the desired combination:

- in contrast to Bayesian techniques, vine copula can handle dependence between variables, not just independence;
- in contrast to fuzzy techniques, where the same "and"-operation (t-norm) is applied for combining all pieces of information, the vine copulas allow the use of different "and"-operations (copulas) to combine information about different variables.

4   How Vine Copulas Are Used in Econometrics

Main challenge: econometric processes are dynamic. Vine copulas describe dependence between a few random *variables* $X_1, \ldots, X_n$. In econometrics, however, processes are highly dynamic, so what we have is random *processes* $X_1(t), \ldots, X_n(t)$, not random variables. How can we use vine copulas to describe the dependence between random processes?

Main idea: use known models to describe the dynamics of each variable. For each of the econometric dynamic variables $r_t \stackrel{\text{def}}{=} X_i(t)$, there are known ways to describe its dynamics. One of the most (and probably *the* most) adequate models for such a dynamics are described by an appropriate combination of the Auto-Regressive Moving-Average Model (ARMA) and the Glosten-Jagannathan-Runkle (GJR) form of a Generalized Auto-Regressive Conditional Heteroskedasticity (GARCH) model [4]; see, e.g., [8, 15]. The corresponding ARMA(p, q)-GJR(k, ℓ) model has the form

$$r_t = c + \sum_{i=1}^{p} \varphi_i \cdot r_{t-i} + \varepsilon_i \sum_{j=1}^{q} \psi_j \cdot \varepsilon_{t-j}, \tag{29}$$

$$\varepsilon_t = h_t \cdot \eta_t, \tag{30}$$

$$h_t^2 = \omega + \sum_{i=1}^{k} \alpha_i \cdot \varepsilon_{t-i}^2 + \sum_{i:\ \varepsilon_{t-i}<0} \gamma_i \cdot \varepsilon_{t-i}^2 + \sum_{j=1}^{\ell} \beta_j \cdot h_{t-j}^2, \tag{31}$$

where ε_t and h_t are auxiliary variables, c, φ_i, ψ_j, ω, α_i, and β_j are real-valued constants (which need to be determined based on the observations), and residuals η_t corresponding to different moments of time t are independent identically distributed random variables.

The distribution of the residuals is usually assumed to be distributed according to skewed student-t or skewed Generalized Error Distribution (GED). A skewed t-distribution means that we combine, with fixed weights, t-distributions $f_1(x)$ and $f_2(x)$ with different scalar parameters limited to, correspondingly, positive and negative values x_i: $f(x) = w_1 \cdot f_1(x)$ when $x \geq 0$ and $f(x) = w_2 \cdot f_2(x)$ when $x < 0$.

A GED distribution is a distribution with a probability density proportional to $\exp\left(-\dfrac{|x|^\nu}{\sigma^\nu}\right)$; it generalizes Gaussian distribution – which corresponds to $\nu = 2$. A skewed GED distribution is a combination of two GED distributions $f_1(x)$ and $f_2(x)$ corresponding to different values σ (but the same value ν): $f(x) = w_1 \cdot f_1(x)$ when $x \geq 0$ and $f(x) = w_2 \cdot f_2(x)$ when $x < 0$, where w_i are appropriate weights.

Resulting solution: copula describes the joint distribution of residuals. Copulas in general (and vine copulas in particular) are a good technique for describing the dependence between several random variables $X_1, \ldots, X_n$. In the dynamical case, instead of n variables $X_1, \ldots, X_n$, we have, in effect, a much larger number of dependent random variables $X_i(t)$ corresponding to different values i and different moments of time t. Not only are variables $X_i(t)$ and $X_j(t)$ corresponding to the same moment of time depending on each other, the values $X_i(t)$ and $X_i(t')$ corresponding to different moments of time also depend on each other – and thus, we also have dependence between $X_i(t)$ and $X_j(t')$.

We have already observed, in our motivation for the use of vine copulas, that the larger the number of dependent variables to consider, the more computationally complex the resulting problem, the more computation time it takes to process this data. We have econometric data corresponding to dozens of years, hundreds of months, thousands of days, so we have thousands of dependent quantities corresponding to different values of i and t. Thus, to be able to describe and process the dependence between different econometric quantities within a reasonable amount of computation time, we need to be able to reduce this dependence between thousands of variables to a dependence between a much smaller number of variables.

Good news is that such a reduction is possible: for such a reduction, we can use the above dynamical equations. Indeed:

– while the values $X_i(t)$ and $X_i(t')$ of the original quantity at different moments of time t and t' are, in general,
– the *residuals* η_t and $\eta_{t'}$ corresponding to different moments of time are *independent* (so all the dependence between $X_i(t)$ and $X_i(t')$ is described by the dynamical equations themselves).

Since residuals corresponding to different moments of time are independent of each other, it is sufficient to consider, for each moment of time t, the dependence between n residuals corresponding to this moment of time; see, e.g., [17]. Thus, for each t, we use a multi-variate copula to describe the dependence between the n residuals corresponding to the original n quantities $X_1, \ldots, X_n$.

Acknowledgments

References

1. Aas, K., Czado, C., Frigessi, A., Bakken, H.: Pair-copula constructions of multiple dependence, Insurance: Mathematics and Economics, 44, 182–198 (2009)
2. Bedford, T. Cooke, R.M.: Monte Carlo simulation of vine dependent random variables for applications in uncertainty analysis, Proceedings of European Safety and Reliability Conference ESREL'2001, Turin, Italy (2001)
3. Bedford, T. Cooke, R.M.: Vines-a new graphical model for dependent random variables, Annals of Statistics, 30(4), 1031–1068 (2002)
4. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics, 31, 307–327 (1986)
5. Browder, F.E. (editor): Mathematical Developments Arising from Hilbert Problems, American Mathematical Society, Providence, Rhode Island (1976)
6. Buchanan, B.G., Shortliffe, E.H.: Rule Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project, Addison-Wesley, Reading, Massachusetts (1984)
7. Czado, C.: Pair-copula constructions of multivariate copulas, In: Jaworski, P. (Ed.), Copula Theory and Its Applications, Lecture Notes in Statistics, Vol. 198, Springer-Verlag, Berlin, Heidelberg, 93–109 (2010)
8. Glosten, L.R., Jagannathan, R., Runkle, D.E.: On the relation between the expected value and the volatility of the nominal excess return on stocks, The Journal of Finance, 48(5), 1779–1801 (1993)
9. Joe, H., Hu, T.: Multivariate distributions from mixtures of max-infinitely divisible distributions, Journal of Multivariate Analysis, 57(2), 240–265 (1996)
10. Joe, H., Dependence comparisons of vine copulae with four or more variables. In: Kurowicka, D., Joe, H. (Eds.). Dependence Modeling: Vine Copula Handbook, World Scientific, Singapore, 2010.
11. Joe, H., Li, H., Nikoloulopoulos, A.K.: Tail dependence functions and vine copulas. Journal of Multivariate Analysis, 101, 252–270 (2010)
12. Klir, G.J., Yuan, B.: Fuzzy Sets and Fuzzy Logic, Prentice Hall, Upper Saddle River, New Jersey (1995)
13. Kurowicka, D., Cooke, R.M.: Uncertainty Analysis with High Dimensional Dependence Modelling, Wiley, New York (2006)
14. Kurowicka, D., Joe, H. (Eds.). Dependence Modeling: Vine Copula Handbook, World Scientific, Singapore, 2010.
15. Ling, S.: Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/IGARCH models, Journal of Econometrics, 140, 849–873 (2007)
16. Lorenz, G.G.: Approximation of Functions, American Mathematical Society, Providence, Rhode Island (1966)

17. Manner, H., Reznikova, O: A survey on time-varying copulas: Specification, simulations and application, Econometric Reviews, 31(6), 654–687 (2012)
18. Neapolitan, R.E.: Learning Bayesian networks, Prentice Hall, Upper Saddle River, New Jersey (2004)
19. Nelsen, R.B.: An Introduction to Copulas, Springer, New York
20. Nguyen, H.T., Walker, E.A.: First Course In Fuzzy Logic, CRC Press, Boca Raton, Florida (2006)
21. Nikoloulopoulos, A.K.: Vine copulas with asymmetric tail dependence and applications to financial return data, Computational Statistics and Data Analysis, 56, 3659–3673 (2012)
22. Papadimitriou, C. H.: Computational Complexity, Addison Wesley, San Diego (1994)
23. Pearl, J.: Probabilistic Reasoning in Intelligent Systems, Morgan Kaufmann, San Francisco, California (1988)
24. Pearl, J.: Causality: Models, Reasoning, and Inference, Cambridge University Press, Cambridge, Massachusetts (2000)
25. Pourret, O., Naim, P., Marcot, B.: Bayesian Networks: A Practical Guide to Applications, Wiley, Chichester, UK (2008)
26. Sklar, A.: Fonctions de répartition á n dimensions et leurs marges, Publ. Inst. Statist. Univ. Paris, 8, 229–231 (1959)
27. Vitushkin, A.G.: On Hilbert's thirteenth problem, Soviet Math. Doklady (Dokl. Akad. Nauk SSSR), 96, 701–704 (1954)
28. Vitushkin, A.G.: Estimating Complexity of the Tabulation Problem, Fizmatgiz, Moscow, 1959 (in Russian)
29. Vitushkin, A.G.: On Hilbert's thirteenth problem and related questions, Russian Mathematical Surveys, 51(1), 11–25 (2004)
30. Zadeh, L.A.: Fuzzy sets, Information and Control, 8, 338–353 (1965)