PROD. TYPE: COM
PP: 1-18 (col.fig.: nil)
COMSTA 3510 MODC+
ED: Bibhu
PAGN: Suresh – SCAN: Global

ARTICLE IN PRESS

Available online at www.sciencedirect.com
ScienceDirect

ELSEVIER

Computational Statistics & Data Analysis ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮

COMPUTATIONAL STATISTICS & DATA ANALYSIS

www.elsevier.com/locate/csda

Sampling algorithms for generating joint uniform distributions using the vine-copula method

D. Kurowicka*, R.M. Cooke

Delft Institute for Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628CD Delft, The Netherlands

Received 11 October 2005; received in revised form 17 October 2006; accepted 4 November 2006

Abstract

An n-dimensional joint uniform distribution is defined as a distribution whose one-dimensional marginals are uniform on some interval I. This interval is taken to be [0,1] or, when more convenient $[-\frac{1}{2}, \frac{1}{2}]$. The specification of joint uniform distributions in a way which captures intuitive dependence structures and also enables sampling routines is considered. The question whether every n-dimensional correlation matrix can be realized by a joint uniform distribution remains open. It is known, however, that the *rank* correlation matrices realized by the joint normal family are sparse in the set of correlation matrices. A joint uniform distribution is obtained by specifying conditional rank correlations on a regular vine and a copula is chosen to realize the conditional bivariate distributions corresponding to the edges of the vine. In this way a distribution is sampled which corresponds exactly to the specification. The relation between conditional rank correlations on a vine and correlation matrix of corresponding distribution is complex, and depends on the copula used. Some results for the elliptical copulae are given.

Keywords: Joint uniform; Joint distribution; Dependence modeling; Vines; Sampling

1. Introduction

The problem of constructing and sampling distributions with given continuous invertible marginals and given rank correlation matrix, or equivalently, constructing and sampling joint uniforms with given correlation matrix, remains open. That is, we do not know whether an arbitrary correlation matrix can be realized by a joint uniform distribution. We seek methods for specifying and sampling joint uniform distributions. Existing methods for generating joint uniform distributions appeal to the joint normal transformation (Iman and Conver, 1982) or the dependence tree-copula method (Cooke, 1997) where a *copula* is a bivariate distribution with uniform margins. For background on copulae see Genest and Rivest (1993), Nelsen (1999), Dall'Aglio et al. (1991) and Joe (1997).

Using the joint normal transform method, we start with a correlation matrix R which we would like to realize in a joint uniform distribution. We construct a joint normal distribution with correlation matrix R and then transform the one-dimensional margins to uniform. This transformation is not linear and does not preserve R. Hence, we do not realize a joint uniform with correlation matrix R (Ghosh and Henderson, 2002); indeed, we do not know if such a distribution

* Corresponding author. Tel.: +31 152785756; fax: +31 152687255.
E-mail addresses: d.kurowicka@tudelft.nl (D. Kurowicka), r.m.cooke@tudelft.nl (R.M. Cooke).

0167-9473/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2006.11.043

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

exists. In practice, for high dimensions, the matrix R will be only partially specified, and we face the notorious matrix completion problem: can a partially specified matrix be extended to a positive definite matrix.

The tree-copula method builds high-dimensional distributions from two-dimensional margins whose overlap structure forms a tree. The copulae used were either minimum information copulae or the diagonal band (Meeuwissen and Bedford, 1997; Cooke and Waij, 1986). The tree-copula method yields distributions which exactly correspond to the specification, but for n-dimensional problems we can specify only $n-1$ correlations corresponding to the edges of the tree.

Regular vines (Cooke, 1997) are a graphical tool for specifying conditional bivariate constraints. When these constraints are associated with partial correlations, it has been shown (Bedford and Cooke, 2002) that any assignment of values from $(-1, 1)$ to the partial correlations corresponding to edges on the vine is consistent, and determines a unique correlation matrix, and that every correlation matrix arises in this way. In other words, a partial correlation regular vine provides an algebraically independent parametrization of the set of correlation matrices. However, we do not know how to sample partial correlation vines, and we are unable to construct a joint uniform distribution from a partial correlation vine.

This paper gives algorithms for sampling regular vines when the bivariate constraints are associated with conditional rank correlations. The relation between conditional rank correlation and partial correlation is complex, and depends on the copula. Some results are given in Section 4. There are two advantages to using the vine-copula method with conditional rank correlations. First, we can construct and sample a distribution that exactly corresponds to the vine specification, and second, if some conditional correlations are unspecified, then a minimal information distribution satisfying the incomplete specification can easily be constructed whenever the copula makes uncorrelated margins (conditionally) independent. In this case it is simply a matter of assigning conditional rank correlation zero to the unspecified nodes in the vine (Cooke, 1997; Bedford and Cooke, 2002). The relation of incomplete vine specification to the matrix completion problem is studied in Kurowicka and Cooke (2003).

The second section reviews briefly facts about rank, product moment and partial correlations, copulas and vines. We show that an arbitrary correlation matrix need not be the rank correlation matrix of a joint normal distribution. The third section presents a sampling algorithm to exactly sample a high-dimensional distribution with uniform margins and given conditional rank correlations using the vine-copula method. The fourth section derives results concerning the relationship between conditional rank and partial correlations for copulae. The fifth section presents simulation results for the elliptical, diagonal band and Frank's copulae. The last section gives conclusions.

2. Correlation and vines

2.1. Rank, product moment, and partial correlations

The obvious relationship between product moment and rank correlations follows directly from their definitions as rank correlation is just a product moment correlation of variables transformed to uniforms. Hence for uniform variables rank and product moment correlations are equal but in general they are different.

Pearson (1907), proved that if vector (X_1, X_2) has a joint normal distribution, then the relationship between rank (r) and product moment correlation (ρ) is given by

$$\rho(X_1, X_2) = 2\sin\left(\frac{\pi}{6} r(X_1, X_2)\right). \tag{1}$$

The proof of this fact is based on the property that the derivative of the density function for bivariate normals with respect to correlation is equal to the second order derivative with respect to x_1 and x_2.

The rank correlation has some important advantages over the product moment correlation. It always exists, can take any value in the interval $[-1, 1]$, is independent of the marginal distributions and is invariant under monotone transformations.

In this paper we will study relationships between conditional product moment, conditional rank and partial correlations. We assume that the reader is familiar with first two correlation coefficients and we introduce here only the partial correlation.

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

Let us consider variables X_i with zero mean and standard deviations $\sigma_i, i=1,\ldots,n$. Let the numbers $b_{12;3,\ldots,n},\ldots,b_{1n;2,\ldots,n-1}$ minimize

$$E((X_1 - b_{12;3,\ldots,n}X_2 - \cdots - b_{1n;2,\ldots,n-1}X_n)^2).$$

Definition 1. A partial correlation of X_1 and X_2 with $X_3,\cdots,X_n$ is

$$\rho_{12;3,\ldots,n} = \mathrm{sgn}(b_{12;3,\ldots,n})\sqrt{b_{12;3,\ldots,n}b_{21;3,\ldots,n}}.$$

Partial correlations can be computed from correlations with the following recursive formula (Yule and Kendall, 1965):

$$\rho_{12;3,\ldots,n} = \frac{\rho_{12;3,\ldots,n-1} - \rho_{1n;3,\ldots,n-1}\cdot\rho_{2n;3,\ldots,n-1}}{\sqrt{1-\rho^2_{1n;3,\ldots,n-1}}\sqrt{1-\rho^2_{2n;3,\ldots,n-1}}}. \tag{2}$$

It is well known that partial and conditional correlations coincide for the joint normal distribution but in general they are not equal. Interestingly, conditional independence is not sufficient for zero partial correlation (Kurowicka and Cooke, 2000).

The transformation (1) does not preserve positive definiteness. This makes normal transform method only approximate (Ghosh and Henderson, 2002). Let us consider a following example:

Example 2. Let

$$A = \begin{bmatrix} 1 & 0.7 & 0.7 \\ 0.7 & 1 & 0 \\ 0.7 & 0 & 1 \end{bmatrix}.$$

We can easily check that A is positive definite. However, the matrix B, such that

$$B(i,j) = 2\sin\left(\frac{\pi}{6}A(i,j)\right) \quad \text{for } i,j=1,\ldots,4,$$

that is,

$$B = \begin{bmatrix} 1 & 0.7167 & 0.7167 \\ 0.7167 & 1 & 0 \\ 0.7167 & 0 & 1 \end{bmatrix}$$

is not positive definite.

From this we can conclude that not every symmetric positive definite matrix with ones on the main diagonal is the rank correlation matrix of a joint normal distribution. Simulation studies show that the normal rank correlation matrices become very sparse in the set of correlation matrices as dimension increases (Ghosh and Henderson, 2002; Kurowicka and Cooke, 2001).

2.2. *Vines*

In this section we briefly introduce a graphical model called vines (Cooke, 1997; Bedford and Cooke, 2002) which together with the copulae will be used in specifying dependence in high-dimensional distributions.

A vine $\mathscr{V}(n)$ on n variables is a nested set of trees $\mathscr{V}(n)=(T_1,\ldots,T_{n-1})$ where the edges of tree j are the nodes of tree $j+1$, $j=1,\ldots,n-2$ and each tree has the maximum number of edges. A *regular vine* on n variables is a vine in which two edges in tree j are joined by an edge in tree $j+1$ only if these edges share a common node, $j=1,\ldots,n-2$.

For each edge of a vine we define *constraint*, *conditioned* and *conditioning* sets of this edge as follows: the set of variables reachable from a given edge via the membership relation is called the constraint set of that edge. When two

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

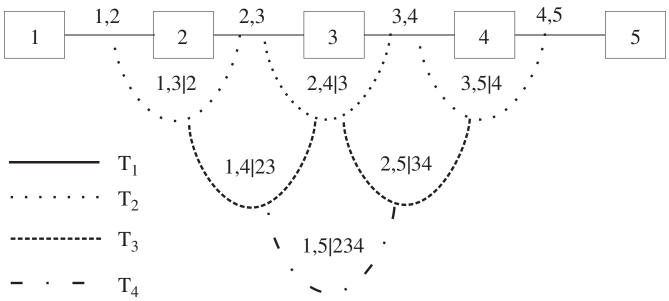

Fig. 1. A D-vine on five elements showing conditioned (before |) and conditioning (after |) sets.

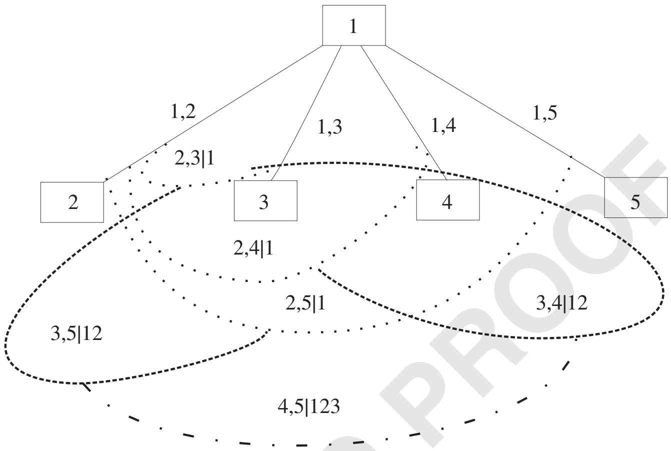

Fig. 2. A canonical vine on five elements showing conditioned and conditioning sets.

edges are joined by an edge of the next tree, the intersection of the respective constraint sets form the conditioning set, and the symmetric difference of the constraint sets is the conditioned set. We denote the conditioning and conditioned sets of an edge e as D_e and $\{C_{e,j}, C_{e,k}\}$, respectively.

Different types of regular vines are available.

Definition 3. A regular vine is called a *canonical vine (C-vine)* if each tree T_i has a unique node of degree $n-i$. The node with maximal degree in T_1 is the *root*. A regular vine is called a *D-vine* if no node has degree greater than two.

The canonical and D-vines are generic in the sense that the order of the variables in the first tree determines the vine completely. D-vines were used in Joe (1997). Figs. 1 and 2 show a canonical and a D-vine on five variables.

Regular vines have some interesting properties that will be used later on in this paper. We collect them in the proposition below that was proven in Bedford and Cooke (2002) and Kurowicka and Cooke (2003):

Proposition 4. *Let* $\mathscr{V}(n) = (T_1, \ldots, T_{n-1})$ *be a regular vine, then*

(1) *the number of edges is* $n(n-1)/2$,
(2) *each conditioned set is a doubleton, each pair of variables occurs exactly once as a conditioned set*,
(3) *if two edges have the same conditioning set, then they are the same edge*,
(4) *if* $e \in E_i$, $i = 1, \ldots, n-1$, *then* $\#D_e = i - 1$.

Defining m-child of a node *f* as a node *e* that is an element of node *f*, the following lemma can be proven (Kurowicka and Cooke, 2006):

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

Lemma 5. *For any edge M in a regular vine, if variable i is a member of the conditioned set of M, then i is a member of the conditioned set of exactly one of the m-children of M, and the conditioning set of an m-child of M is a subset of the conditioning set of M.*

Definition 6. A *partial correlation specification* for a regular vine is an assignment of values in $(-1, 1)$ to each edge of the vine.

The edges in a regular vine may be associated with a set of partial correlations in the following way: for $i=1,\ldots,n-1$, with $e \in E_i$, $e=\{j,k\}$ we associate

$$\rho_{C_{e,j}C_{e,k};D_e}.$$

In Bedford and Cooke (2002) the following is proved:

Theorem 7. *For any regular vine on n elements there is a one-to-one correspondence between the set of $n \times n$ correlation matrices and the set of partial correlation specifications for the vine.*

The above theorem shows that all assignments of numbers between -1 and 1 to partial correlation in regular vine specification are consistent and all full rank correlation matrices can be obtained this way.

This relationship can be used in checking whether a given symmetric matrix with one's on the main diagonal is positive definite (Kurowicka and Cooke, 2003). Furthermore, it allows us to sample a positive definite matrix by simply sampling a vector from $(-1,1)^{\binom{n}{2}}$ and assigning the components to the partial correlations on the regular vine. Values in the correlation matrix are computed with the formula (2).

Definition 8. A *conditional rank correlation specification* for a regular vine is an assignment of values in $[-1, 1]$ to each edge of the vine. For $i=1,\ldots,n-1$, with $e \in E_i$, $e=\{j,k\}$ we associate conditional rank correlation

$$r_{C_{e,j}}, C_{e,k}|D_e.$$

Note that we can use correlations 1 and -1 in a rank correlation specification. In Section 3 we give explicit algorithms for sampling a regular vine.

2.3. Copulae

A copula is a bivariate distribution with uniform margins. We present here only few families of copula that will be used later on in the paper.

The *elliptical* copula was introduced in Kurowicka et al. (2000). The density function of the elliptical copula with correlation $\rho \in (-1,1)$ is

$$f_\rho(x,y)=\begin{cases}\dfrac{1}{\pi}\dfrac{1}{\sqrt{\frac{1}{4}(1-\rho^2)-x^2-y^2+2\rho xy}}, & (x,y)\in B,\\ 0, & (x,y)\notin B,\end{cases}$$

where

$$B=\left\{(x,y)\,\middle|\,x^2+\left(\frac{y-\rho x}{\sqrt{1-\rho^2}}\right)^2<\frac{1}{4}\right\}.$$

Fig. 3 depicts a graph of the density function of the elliptical copula with correlation $\rho=0.8$.

The following properties of the elliptical copula are proven in (Kurowicka et al., 2000) and will be used later on in this paper:

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

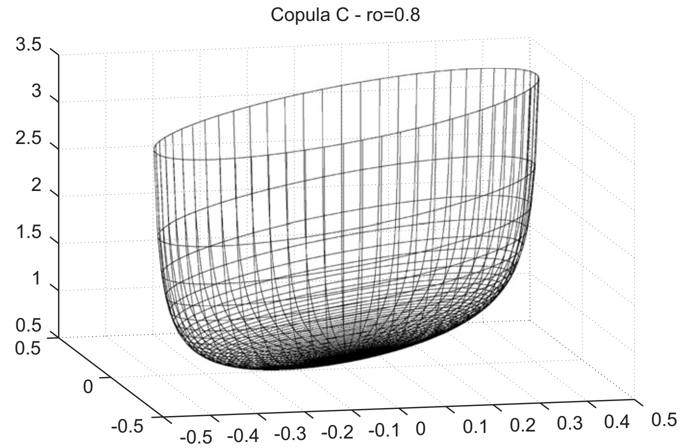

Fig. 3. A density function of an elliptical copula with correlation 0.8.

Theorem 9. *If X, Y uniform on $[-\frac{1}{2}, \frac{1}{2}]$ are joined by the elliptical copula with correlation ρ then*

(a) $E(Y|X) = \rho X$,
(b) $Var(Y|X) = \frac{1}{2}(1-\rho^2)(\frac{1}{4} - X^2)$,
(c) *for* $\rho X - \sqrt{1-\rho^2}\sqrt{\frac{1}{4} - X^2} < y < \rho X + \sqrt{1-\rho^2}\sqrt{\frac{1}{4} - X^2}$

$$F_{Y|X}(y) = \frac{1}{2} + \frac{1}{\pi}\arcsin\left(\frac{y-\rho X}{\sqrt{1-\rho^2}\sqrt{\frac{1}{4}-X^2}}\right),$$

(d) *for* $-\frac{1}{2} \leqslant t \leqslant \frac{1}{2}$

$$F^{-1}_{Y|X}(t) = \sqrt{1-\rho^2}\sqrt{\tfrac{1}{4}-X^2}\,\sin(\pi t) + \rho X.$$

Note that the cumulative and inverse cumulative distribution functions of the conditional distributions have tractable closed form expressions, in spite of the fact that the density is infinite along its boundary.

Theorem 10. *Let X, Y, Z be uniform variables on $[-\frac{1}{2}, \frac{1}{2}]$ and let X, Y and X, Z be joined by elliptical copula with correlations ρ_{XY} and ρ_{XZ}, respectively, and assume that the conditional copula for YZ given X does not depend on X; then the conditional correlation $\rho_{YZ|X}$ is constant in X and*

$$\rho_{YZ|X} = \rho_{YZ;X}.$$

The above result is very specific for the elliptical copulae. For diagonal band and minimum information copulae the conditional correlation $\rho(Y|X, Z|X)$ will depend on X even when $r(Y|X, Z|X)$ does not depend on X but the partial correlation and mean conditional product moment correlation are approximately equal (Kurowicka and Cooke, 2000). This approximation, however, deteriorates as the correlations become more extreme. Theorem 10 can be trivially extended to conditional copulae $(YZ|X)$ which are mixtures of elliptical copulae.

Fig. 4 shows graphs of density functions of the diagonal band (Cooke and Waij, 1986) and Frank's (1979) copulae that will be used later on in this paper.

For both these copulas it is possible to find conditional distribution and its inverse as well as relationship between their parameters and correlation. This will be used in the next section.

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

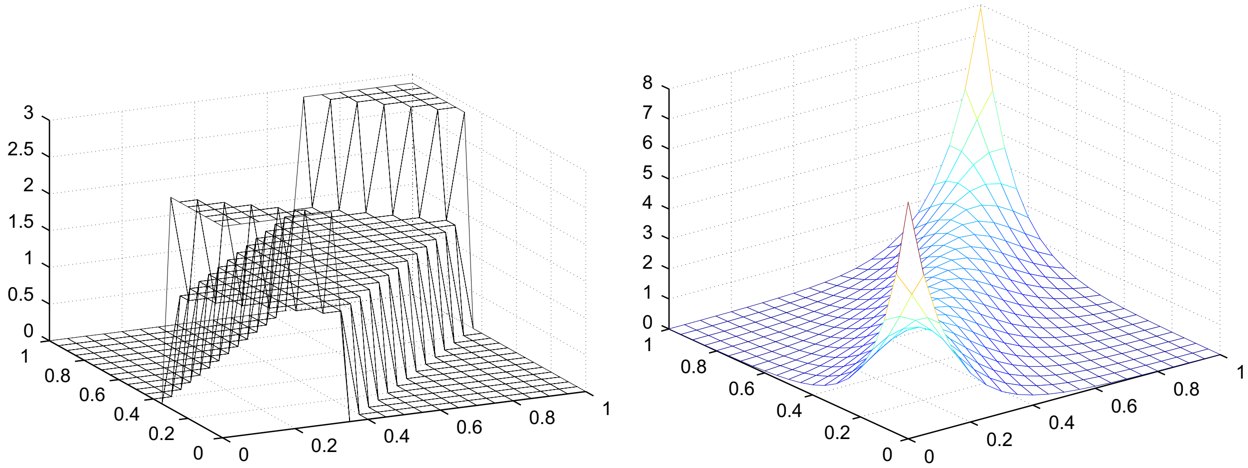

Fig. 4. Densities of the diagonal band (left) and Frank's (right) copulae with correlation 0.8.

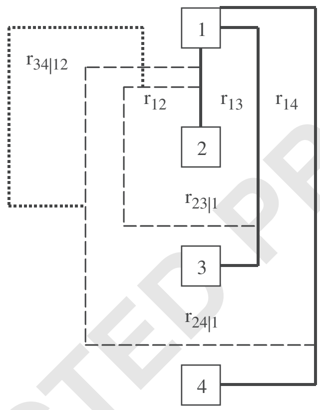

Fig. 5. The canonical vine on four variables with (conditional) rank correlations assigned to the edges.

3. Sampling vines

A rank correlation specification on regular vine plus copula determines the whole joint distribution. The procedure of sampling such a distribution can be written for any regular vine. We first illustrate the procedure for the canonical and D-vines, and then give the general procedure.

3.1. *Canonical vine*

For the canonical vine, the sampling algorithm takes a simple form. We illustrate this algorithm for a canonical vine on four variables, shown in Fig. 5.

The algorithm involves sampling four independent uniform $(0, 1)$ variables $U_1, \ldots, U_4$ (realizations denotes as $u_1, \ldots, u_4$). We assume that the variables $X_1, \ldots, X_4$ are also uniform. Let $r_{i,j|k}$ denote the conditional rank correlation between variables (i, j) given k. Let $F_{r_{i,j|k};U_i}(X_j)$ denote the cumulative distribution function for X_j given U_i under the conditional copula with correlation $r_{i,j|k}$. Then

$$X_j = F^{-1}_{r_{i,j|k};U_i}(U_j)$$

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

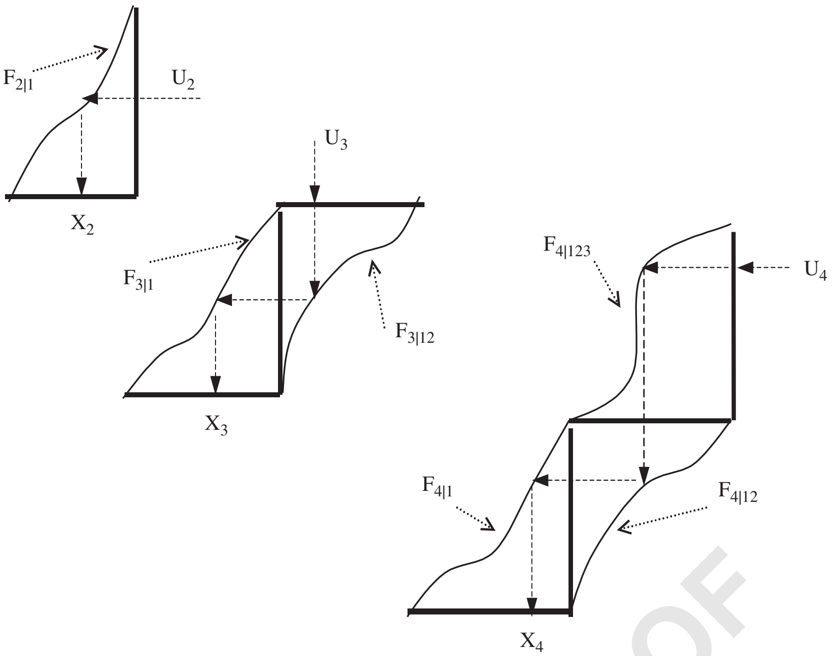

Fig. 6. Staircase graph representation of canonical vine sampling procedure.

expresses X_j as a function of U_j and U_i. The algorithm can now be stated as follows:

$$
\begin{aligned}
x_1 &= u_1,\\
x_2 &= F^{-1}_{r_{12};u_1}(u_2),\\
x_3 &= F^{-1}_{r_{13};u_1}(F^{-1}_{r_{23|1};u_2}(u_3)),\\
x_4 &= F^{-1}_{r_{14};u_1}(F^{-1}_{r_{24|1};u_2}(F^{-1}_{r_{34|12};u_3}(u_4))).
\end{aligned}
\tag{3}
$$

We see that the uniform variables $U_1, \ldots, U_4$ are sampled independently, and the variables $X_1, \ldots, X_4$ are obtained by applying successive inverse cumulative distribution functions.

The "staircase graphs" in Fig. 6 show the sampling procedure graphically. The horizontal and vertical lines represent the (0, 1) interval; the intervals are connected via conditional cumulative distribution functions. Notice that u_1, u_2 and u_3 are values of X_1, $F_{2|1}$ and $F_{3|12}$, respectively, hence conditional distributions $F_{4|1}$, $F_{4|12}$ and $F_{4|123}$ can be easily found by conditionalizing copulae with correlations r_{14}, $r_{24|1}$ and $r_{34|12}$ on values of u_1, u_2 and u_3, respectively. Inverting the value of u_4 through $F_{4|1}$, $F_{4|12}$ and $F_{4|123}$ gives x_4.

In general we can sample an n-dimensional distribution represented graphically by the canonical vine on n variables with (conditional) rank correlations

$$
\begin{array}{lllll}
r_{12}, & r_{13}, & r_{14}, & \ldots & r_{1n},\\
 & r_{23|1}, & r_{24|1}, & \ldots & r_{2n|1},\\
 & & r_{34|12}, & \ldots & r_{3n|12},\\
 & & & \ldots & \\
 & & & & r_{n-1,n|12\ldots n-2},
\end{array}
$$

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

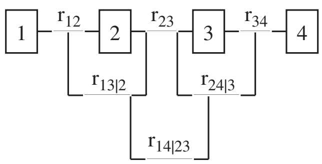

Fig. 7. Rank correlation specification for the D-vine on four variables.

assigned to the edges of the vine, by sampling n independent, uniform (0,1) variables, say $U_1, U_2, \ldots, U_n$, and calculating

$$\begin{aligned}
x_1 &= u_1,\\
x_2 &= F^{-1}_{r_{12};u_1}(u_2),\\
x_3 &= F^{-1}_{r_{13};u_1}(F^{-1}_{r_{23|1};u_2}(u_3)),\\
x_4 &= F^{-1}_{r_{14};u_1}(F^{-1}_{r_{24|1};u_2}(F^{-1}_{r_{34|12};u_3}(u_4))),\\
&\ldots\\
x_n &= F^{-1}_{r_{1n};u_1}(F^{-1}_{r_{2n|1};u_2}(F^{-1}_{r_{3n|12};u_3}(\ldots(F^{-1}_{r_{n-1,n|12\ldots n-2};u_{n-1}}(u_n))\ldots))).
\end{aligned}$$

3.2. D-vine

The sampling algorithm for the D-vine is more complicated than that of the canonical vine. We illustrate the sampling algorithm for a D-vine on four variables, shown in Fig. 7.

For reasons that become apparent in the next section, we choose the sampling order to be X_3, X_2, X_4, X_1. The sampling procedure for the D-vine in Fig. 7 is

$$\begin{aligned}
x_3 &= u_3,\\
x_2 &= F^{-1}_{r_{23};x_3}(u_2),\\
x_4 &= F^{-1}_{r_{43};x_3}(F^{-1}_{r_{42|3};F_{r_{23};x_3}(x_2)}(u_4)),\\
x_1 &= F^{-1}_{r_{12};x_2}(F^{-1}_{r_{13|2};F_{r_{32};x_2}(x_3)}(F^{-1}_{r_{14|23};F_{r_{42|3};F_{r_{32};x_3}(x_2)}(F_{r_{43};x_3}(x_4))}(u_1))).
\end{aligned}$$

Notice that the sampling procedure for D-vine uses conditional cumulative distribution functions as well as inverse conditional cumulative distribution functions, hence will be much slower than the procedure for the canonical vine. To shorten the notation the above algorithm can be stated as

$$\begin{aligned}
x_3 &= u_3,\\
x_2 &= F^{-1}_{2|3:x_3}(u_2),\\
x_4 &= F^{-1}_{4|3:x_3}(F^{-1}_{4|23:F_{2|3}(x_3)}(u_4)),\\
x_1 &= F^{-1}_{1|2:x_2}(F^{-1}_{1|32:F_{3|2}(x_3)}(F^{-1}_{1|432:F_{4|32}(x_4)}(u_1))).
\end{aligned}$$

Fig. 8 gives a staircase graph representation of the D-vine sampling procedure. Notice that for the D-vine values of $F_{3|2}$ and $F_{4|32}$ that are used to conditionalize copulae with correlations $r_{13|2}$ and $r_{14|23}$ to obtain $F_{1|23}$ and $F_{1|432}$, respectively, have to be calculated. This is in contrast to the canonical vine where values of u_i are used.

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

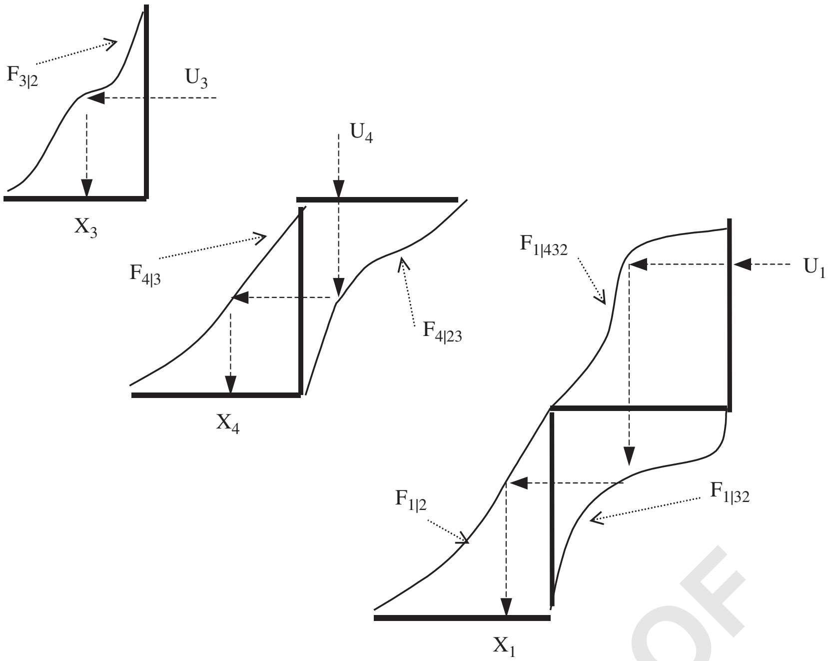

Fig. 8. Staircase graph representation of D-vine sampling procedure.

Rather than giving the general algorithm for the D-vine, we turn directly to the sampling procedure for an arbitrary regular vine.

3.3. *Regular vines*

A regular vine on n nodes will have a single node in tree $n-1$. It suffices to show how to sample one of the conditioned variables in this node, say n. Assuming we have sampled all the other variables we proceed as follows:

(1) By Lemma 5, the variable n occurs in trees $1, \ldots, n-1$ exactly once as a conditioned variable. The variable with which it is conditioned in tree j is called its "j-partner". We define an ordering for n as follows: index the j-partner of variable n as variable j. We denote the conditional bivariate constraints corresponding to the partners of n as

$$(n, 1|\emptyset), (n, 2|D_2^n), (n, 3|D_3^n) \cdots (n, n-1|D_{n-1}^n).$$

Again by Lemma 5, variables $1, \ldots, n-1$ appear first as conditioned variables (to the left of "|") before appearing as conditioning variables (to the right of "|"). Also,

$$0 = \#D_1^n < \#D_2^n < \cdots < \#D_{n-1}^n = n-2.$$

(2) Assuming we have sampled all variables except n, sample one variable uniformly distributed on the interval (0,1), denoted u_n. We use the general notation $F_{a|b,C}$ to denote $F_{a,b|C:F_{b|C}}$; that is, the conditional copula for $\{a, b|C\}$ conditional on a value of the cumulative conditional distribution $F_{b|C}$. Here, $\{a, b|C\}$ is the conditional bivariate constraint corresponding to a node in the vine.

(3) Sample x_n as follows:

$$x_n = F^{-1}_{n|1,D_1^n}(F^{-1}_{n|2,D_2^n}(\ldots(F^{-1}_{n|n-1,D_{n-1}^n}(u_n))\ldots)). \quad (4)$$

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

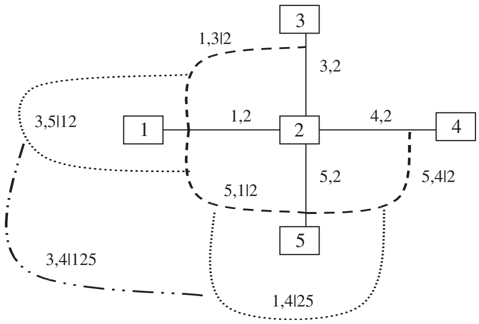

Fig. 9. Regular vine with five variables.

The innermost term of (4) is

$$F^{-1}_{n|n-1,D^n_{n-1}} = F^{-1}_{n,n-1|D^n_{n-1}:F_{n-1|D^n_{n-1}}} = F^{-1}_{n,n-1|D^n_{n-1}:F_{n-1,n-2|D^{n-1}_{n-2}:F_{n-2|D^{n-1}_{n-2}}}}.$$

Example 11. We illustrate the sampling procedure for the regular vine on five variables in Fig. 9 which is neither canonical nor a D-vine. The top node is $\{34|512\}$. Assuming we have sampled variables 2,5,1,3 already the sampling procedure for x_4 is

$$x_4 = F^{-1}_{4|2}(F^{-1}_{4|5,2}(F^{-1}_{4|1,52}(F^{-1}_{4|3,152}(u_4)))).$$

4. Relationship between conditional rank and partial correlations assigned to the edges of a vine

For a vine-copula method we would like a copula for which rank correlations are equal to corresponding partial correlations on the vine or at least for which the relationship between them is known. We now examine some copulae from this perspective.

4.1. Elliptical copula

A copula for which partial and constant conditional product moment correlations are equal is the elliptical copula. More precisely, it is known that when (X, Y) and (X, Z) are joined by elliptical copulae (Section 2.3), and when the conditional copula of (Y, Z) given X does not depend on X, then the conditional correlation of (Y, Z) given X is equal to the partial correlation (Theorem 10). We can find a relationship between partial and conditional rank correlation by incorporating the sampling algorithm for a canonical vine from the previous section and then compute:

$$\rho_{23;1} = \frac{12\int\int\int_{I^3} x_2 x_3 \, \mathrm{d}u_1 \, \mathrm{d}u_2 \, \mathrm{d}u_3 - r_{12}r_{13}}{\sqrt{(1-r_{12}^2)(1-r_{13}^2)}}, \tag{5}$$

where $I = [-\frac{1}{2}, \frac{1}{2}]$. Calculating the above integral with x_2 and x_3 given by the sampling procedure (3) with inverse conditional distributions of the elliptical copula (Theorem 9) and simplifying the expression we get

$$\rho_{23;1} = 2\int\int_{I^2} \sin(\pi u_2) \sin\left(\pi\left[\sqrt{1-r_{23|1}^2}\sqrt{\frac{1}{4}-u_2^2}\,\sin(\pi u_3) + r_{23|1}u_2\right]\right) \mathrm{d}u_2 \, \mathrm{d}u_3. \tag{6}$$

Notice that $\rho_{23;1}$ does not depend on r_{12}, r_{13}. It depends only on $r_{23|1}$. This is very specific for the elliptical copula. We denote the relationship (6) as

$$\rho_{23;1} = \psi(r_{23|1}). \tag{7}$$

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

Now we can easily show that when $r_{23|1}=1$ then

$$\rho_{23;1}=2\int\int_{I^2}\sin(\pi u_2)\ \sin(\pi u_2)\,\mathrm{d}u_2\,\mathrm{d}u_3=1$$

and when $r_{23|1}=-1$

$$\rho_{23;1}=2\int\int_{I^2}\sin(\pi u_2)\ \sin(-\pi u_2)\,\mathrm{d}u_2\,\mathrm{d}u_3=-1.$$

From the above result, Hoeffding's (1940) theorem and Theorem 7 we get that there exists trivariate uniform distribution realizing any correlation structure.

Example 12. Construct a trivariate distribution with the following rank correlation structure:

$$A=\begin{bmatrix}1 & 0.7 & 0.7\\ 0.7 & 1 & 0\\ 0.7 & 0 & 1\end{bmatrix}. \tag{8}$$

The correlation specification on a regular vine with elliptical copulae provides a very convenient way of sampling a distribution with rank correlation matrix (8). We have $\rho_{12}=\rho_{13}=0.7$ and $\rho_{23}=0$. The partial correlation $\rho_{23;1}$ can be calculated from (2) as -0.96. From (6) we find conditional rank correlation $r_{23|1}=-0.9635$ that corresponds to partial correlation of -0.96.

Using the sampling algorithm for the canonical vine we can sample a distribution with rank correlation matrix (8) very efficiently. Notice that this rank correlation matrix could not be realized by the joint normal transformation method (see Example 2).

Unfortunately the above result cannot be generalized to higher dimensions. We show now how the relationship between partial and conditional correlations on a vine can be established for a four variate distribution obtained by the vine-elliptical copula method. For a given 4×4 correlation matrix, using the recursive formula 2, a partial correlation specification on the canonical vine on four variables can be obtained

$$\begin{array}{lll}\rho_{12}, & \rho_{13}, & \rho_{14},\\ & \rho_{23;1}, & \rho_{24;1},\\ & & \rho_{34;12}.\end{array} \tag{9}$$

We must find a rank correlation specification on the canonical vine that corresponds to the partial correlation specification (9)

$$\begin{array}{lll}r_{12}, & r_{13}, & r_{14},\\ & r_{23|1}, & r_{24|1},\\ & & r_{34|12}.\end{array} \tag{10}$$

Clearly, $r_{1,i}=\rho_{1i}$, $i=2,3,4$. The $r_{23|1}, r_{24|1}$ that correspond to $\rho_{23;1}, \rho_{24;1}$ respectively can be found from (7), hence $r_{23|1}=\psi^{-1}(\rho_{23;1})$ and $r_{24|1}=\psi^{-1}(\rho_{24;1})$. Now we must only find $r_{34|12}$ that corresponds to $\rho_{34;12}$. Using the sampling procedure for a canonical vine the correlation ρ_{34} can be calculated as

$$\rho_{34}=12\int\int\int\int_{I^4}x_3x_4\,\mathrm{d}u_1\,\mathrm{d}u_2\,\mathrm{d}u_3\,\mathrm{d}u_4.$$

Simplifying and using partial correlation formula (2) we get

$$\rho_{34;1}=2\int\int\int_{I^3}g(r_{23|1},u_2,u_3)g(r_{24|1},u_2,g(r_{34|12},u_3,u_4))\,\mathrm{d}u_2\,\mathrm{d}u_3\,\mathrm{d}u_4, \tag{11}$$

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

where

$$g(r,u,v)=\sin\left[\pi\left(\sqrt{1-r^2}\sqrt{\tfrac{1}{4}-u^2}\,\sin(\pi v)+ru\right)\right].$$

Hence $\rho_{34;1}$ depends on $r_{23|1}, r_{24|1}$, which are already chosen, and $r_{34|12}$ which we want to find. Denoting this relationship by $\rho_{34;1}=\Phi(r_{34|12},r_{23|1},r_{24|1})$ and using partial correlation formula (2) the relationship between $\rho_{34;12}$ and $r_{34;12}$ is

$$\rho_{34;12}=\frac{\Phi(r_{34|12},r_{23|1},r_{24|1})-\rho_{23;1}\rho_{24;1}}{\sqrt{(1-\rho_{23;1}^2)(1-\rho_{24;1}^2)}}. \tag{12}$$

This cannot be solved analytically but using numerical integration we can search for $r_{34|12}$ corresponding to given $\rho_{34;12}$.

Notice that for

$$\rho_{34;12}=\frac{\rho_{34;1}-\rho_{23;1}\rho_{24;1}}{\sqrt{(1-\rho_{23;1}^2)(1-\rho_{24;1}^2)}},$$

where $\rho_{34;1}$ given by (11) and $\rho_{23;1}$, $\rho_{24;1}$ by (6), if $r_{34|12}=1$ then

$$\rho_{34;12}=\frac{2h(r_{23|1},r_{24|1},u_2,u_3)-4t(r_{23|1},u_2,u_3)t(r_{24|1},u_2,u_3)}{\sqrt{(1-4t(r_{23|1},u_2,u_3)^2)(1-4t(r_{24|1},u_2,u_3)^2)}},$$

where

$$h(r_{23|1},r_{24|1},u_2,u_3)=\int_{I^2} g(r_{23|1},u_2,u_3)g(r_{24|1},u_2,u_3)\,\mathrm{d}u_2\,\mathrm{d}u_3,$$

and

$$t(r_{ij|k},u_2,u_3)=\int_{I^2}\sin(\pi u_2)g(r_{ij|k},u_2,u_3)\,\mathrm{d}u_2\,\mathrm{d}u_3.$$

This in general is not equal to 1 but if one additionally assumes that $r_{23|1}=r_{24|1}$ then $\rho_{34;12}=1$.

Suffices to show that

$$2\int_{I^2} g(r_{23|1},u_2,u_3)^2\,\mathrm{d}u_2\,\mathrm{d}u_3=1.$$

We get

$$\begin{aligned}2\int_{I^2} g(r_{23|1},u_2,u_3)^2\,\mathrm{d}u_2\,\mathrm{d}u_3&=2\int_{I^2}\sin^2\left(\pi\left[\sqrt{1-r_{23|1}^2}\sqrt{\frac{1}{4}-u_2^2}\,\sin(\pi u_3)+r_{23|1}u_2\right]\right)\mathrm{d}u_2\,\mathrm{d}u_3\\&=\int_{I^2}1-\cos\left(2\pi\left[\sqrt{1-r_{23|1}^2}\sqrt{\frac{1}{4}-u_2^2}\,\sin(\pi u_3)+r_{23|1}u_2\right]\right)\mathrm{d}u_2\,\mathrm{d}u_3.\end{aligned}$$

Using the formula for cosine of a sum of two angles and noticing that

$$\int_{I^2}\sin(2\pi r_{23|1}u_2)\,\mathrm{d}u_2\,\mathrm{d}u_3=0$$

we obtain

$$=1-\int_{I^2}\cos\left(2\pi\sqrt{1-r_{23|1}^2}\sqrt{\frac{1}{4}-u_2^2}\,\sin(\pi u_3)\right)\cos(2\pi r_{23|1}u_2)\,\mathrm{d}u_2\,\mathrm{d}u_3.$$

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

Integrating the above integral by parts and simplifying we get

$$=1-\frac{1}{2\pi r_{23|1}}\int_{I^2}\sin\left(2\pi\sqrt{1-r_{23|1}^2}\sqrt{\frac{1}{4}-u_2^2}\,\sin(\pi u_3)\right)\sin(2\pi r_{23|1}u_2)\frac{u_2}{\sqrt{\frac{1}{4}-u_2^2}}\,\mathrm{d}u_2\,\mathrm{d}u_3.$$

Since the integrand in above integral is such that $f(u_2,u_3)=-f(-u_2,-u_3)$ and $f(-u_2,u_3)=-f(u_2,-u_3)$ then $\rho_{34;12}=1$.

We show now how the rank correlation specification on the canonical vine on four variables can be found to realize a given correlation structure.

Example 13. Let us consider a matrix

$$A=\begin{bmatrix}1.0000 & -0.3609 & 0.3764 & -0.3254\\ -0.3609 & 1.0000 & 0.6519 & -0.3604\\ 0.3764 & 0.6519 & 1.0000 & -0.2919\\ -0.3254 & -0.3604 & -0.2919 & 1.0000\end{bmatrix}. \tag{13}$$

The partial correlation specification on the canonical vine is

$$\begin{matrix}\rho_{12}, & \rho_{13}, & \rho_{14},\\ & \rho_{23;1}, & \rho_{24;1},\\ & & \rho_{34;12},\end{matrix} \quad = \quad \begin{matrix}-0.3609, & 0.3764, & -0.3254,\\ & 0.9117, & -0.5419,\\ & & 0.8707.\end{matrix} \tag{14}$$

The corresponding rank correlation specification is the following:

$$\begin{matrix}r_{12}, & r_{13}, & r_{14},\\ & r_{23|1}, & r_{24|1},\\ & & r_{34|12},\end{matrix} \quad = \quad \begin{matrix}-0.3609, & 0.3764, & -0.3254,\\ & 0.9170, & -0.5557,\\ & & 0.9392.\end{matrix} \tag{15}$$

where $r_{23|1}$, $r_{24|1}$ are found from (6) and $r_{34|12}$ from (12).

Using sampling procedure (3) with the above rank correlations we obtain a distribution with rank correlation matrix A up to sampling and numerical errors.

The matrix A in above example was chosen such that after transformation (1) becomes non-positive definite. Hence A is not a rank correlation matrix of joint normal distribution but can be realized with the vine-elliptical copula method.

Example 14. The following-four dimensional correlation matrix cannot be realized with the vine-elliptical copula method:

$$A=\begin{bmatrix}1.0000 & 0.8000 & 0.6000 & -0.3000\\ 0.8000 & 1.0000 & 0.2400 & -0.6979\\ 0.6000 & 0.2400 & 1.0000 & 0.5178\\ -0.3000 & -0.6979 & 0.5178 & 1.0000\end{bmatrix}.$$

The partial correlation specification on the canonical vine is

$$\begin{matrix}\rho_{12}, & \rho_{13}, & \rho_{14},\\ & \rho_{23;1}, & \rho_{24;1},\\ & & \rho_{34;12},\end{matrix} \quad = \quad \begin{matrix}0.8, & 0.6, & -0.3,\\ & -0.5, & -0.8,\\ & & 0.99.\end{matrix}$$

We can find that $r_{23|1}=-0.5137$, $r_{24|1}=-0.8101$ which correspond to $\rho_{23;1}=-0.5$, $\rho_{24;1}=-0.8$, respectively. However, assigning $r_{34|12}=1$ yields $\rho_{34;12}$ equal only to 0.9892. The above matrix is also not a rank correlation matrix for joint normal.

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

4.2. Other copulae

Using techniques presented in the previous subsection we can find a relationship between partial and conditional rank correlations for other copula. If in (5) when calculating the integral we use x_2, x_3 with the inverse conditional distributions of, e.g., the diagonal band or Frank's copula then the relationship between partial correlation and a parameter of the copula that corresponds to $r_{23|1}$ can be established. However, in contrast to the elliptical copula, $\rho_{23;1}$ will also depend on r_{12}, r_{13}.

Interestingly, the correlation matrix in Example 12 cannot be realized with the vine method with the diagonal band copula. We may use formula (5) to search for $r_{23|1}$ that corresponds to partial correlation -0.96 and we find that $r_{23|1} = -1$ yields partial correlation equal only to -0.9403.

At the moment there are no analytical results concerning the relationship between partial and constant conditional rank correlations for the diagonal band or the Frank's copula. It would be interesting to know whether there exists a copula for which partial and constant conditional rank correlations are equal. We could also consider incorporating non-constant conditional rank correlations assigned to the edges of a regular vine and see how they relate to partial correlations.

5. Predicting correlation matrices—simulation results

Having specified a joint uniform distribution with the vine-copula method, we would like to predict the resulting correlation matrix. It was shown in the previous section how this can be done for the canonical vine with elliptical copula. In general we cannot calculate the correlation matrix, but if we *pretend* that the conditional rank correlations satisfy the recursive relations for partial correlations, we can predict the correlation matrix with some error. To indicate how big this error can be, we performed a simulation experiment for the matrix in Example 13. We run the sampling procedure (3) for 10 000 samples using the elliptical copula and conditional rank correlations on the canonical vine which are either:

- [*rank-vine*] given in (15),
- [*partial-vine*] equal to partial correlations hence given in (14).

For fair comparison we have used the same independent uniform samples for both simulations. Sample correlation matrices for both distributions were calculated and compared to the original matrix (13). The sum of absolute differences (d) between matrix (13) and sampled matrices was taken as a measure of difference between matrices. This procedure was repeated 500 times and results in a form of average ($E(d)$) and standard deviation ($\sigma(d)$) of these differences as well as average and standard deviation of the differences per cell (denoted as $E(d_{\text{cell}})$, $\sigma(d_{\text{cell}})$, respectively) are presented in Table 1.

We see that the differences between both errors are not too big but the vine with rank correlation indeed predicts the target correlation matrix better than the one where we have assumed equality of partial and conditional rank correlations.

We now estimate this error for the elliptical, diagonal band and Frank's copulae with simulation. The procedure is as follows:

(1) Sample a correlation matrix of size n with onion method (Ghosh and Henderson, 2002).
(2) Find the partial correlation specification on the canonical vine on n variables.
(3) Pretend that conditional rank and partial correlation specifications are equal.

Table 1
Average and standard deviation of differences and differences per cell between matrix (13) and sampling correlation matrices for distributions [rank-vine] and [partial-vine]

Dist.	$E(d)$	$\sigma(d)$	$E(d_{\text{cell}})$	$\sigma(d_{\text{cell}})$
[rank-vine]	0.0665	0.0024	0.055	0.0024
[partial-vine]	0.0828	0.0288	0.069	0.0024

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

Table 2
Performance of different copulae (E, DB, F) in the vine-copula method

Copula/dimension		3	4	5	6	7	8	9	10
E	Min. err. (%)	60.6	53.6	28.6	8.4	1	0	0	0
	Max. err. (%)	0.115	0.213	0.451	0.702	1.093	1.516	2.021	2.807
	Aver. err. (%)	0.035	0.099	0.216	0.414	0.696	1.053	1.504	1.992
DB	Min. err. (%)	25.4	33.6	53.4	62.6	61	46.4	27.4	9.4
	Max. err. (%)	0.129	0.226	0.372	0.556	0.765	0.997	1.237	1.5739
	Aver. err. (%)	0.041	0.103	0.190	0.317	0.467	0.645	0.869	1.045
F	Min. err. (%)	14	12.8	18	29	38	53.6	72.6	90.6
	Max. err. (%)	0.268	0.292	0.418	0.672	0.825	1.049	1.280	1.484
	Aver err. (%)	0.054	0.126	0.219	0.345	0.483	0.642	0.830	0.945
N	Max. err. (%)	0.200	0.364	0.571	0.642	0.873	1.002	1.195	1.550
	Aver. err. (%)	0.013	0.158	0.265	0.381	0.525	0.690	0.863	1.068

Simulation results for joint normal method (N).

Table 3
Average error per cell for vine-copula method with E, DB, F copulae and for joint normal method (N)

Dist./dim.	3	4	5	6	7	8	9	10
E	0.0058	0.0083	0.0108	0.0138	0.0166	0.0188	0.0209	0.0221
DB	0.0069	0.0086	0.0095	0.0106	0.0111	0.0115	0.0121	0.0116
F	0.0091	0.0105	0.0110	0.0115	0.0115	0.0115	0.0115	0.0105
N	0.0795	0.0132	0.0132	0.127	0.0125	0.0123	0.0120	0.0119

(4) Draw 10 000 samples of the n-dimensional distribution described by the rank correlation specification with the elliptical (E), the diagonal band (DB) and Frank's (F) copula (in the sampling procedure for all copulae the same independent samples are used).
(5) Calculate correlation matrices from these samples and compare with the target correlation matrix (as a measure of difference between matrices we took sum of absolute differences between elements of the target and sampled matrix).

The above procedure was repeated 500 times for each dimension. Error is defined per matrix as the sum over all cells of the absolute difference between the target and the sampled matrix. In Table 2 we record the maximum absolute error over the 500 matrices, the average absolute error over the 500 matrices, and percentage of the 500 simulated matrices for which the given copula led to the minimum error. To compare these results with the joint normal method, we sampled 500 random correlation matrices with the "onion" method, and used these matrices as product moment correlation matrices of the joint normal. We drew 10 000 samples from the joint normal and calculated rank correlation matrix from these samples. Because we sample from normal distributions instead of uniform distributions, the results per matrix cannot be compared with the vine-copula matrices, but the maximum and average error can be compared.

Notice that for three and four-dimensional matrices the elliptical copula gives the smallest error. Five, six and seven-dimensional matrices are best approximated by the vine-diagonal band method and then Frank's copula starts to produce the smallest error.

Of course, the average error over 500 matrices is affected by the fact that number of cells over which we sum increases with dimension. Table 2 shows average error, divided by the number off-diagonal cells. The average error per cell for the vine-copula method and the joint normal method is presented in Table 3. Notice that the average error per cell using the normal method is always bigger than the average error per cell for the vine method with the diagonal band and Frank's copula. However, the difference decreases with matrix dimension.

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

6. Conclusions

The problem of representing and exactly sampling a distribution with given margins and a given rank correlation matrix arises in uncertainty and sensitivity analysis. Up to now, only approximate results (joint normal and Iman and Conver, 1982 method) were available. The vine-copula method presented in this paper is a generalization of the existing tree-copula approach. The tree structure imposes conditional independencies which severely constrains the dependence structure that can be realized. The copula-vine method uses conditional dependence to construct a multidimensional distribution from two-dimensional and conditional two-dimensional distributions of uniform variables. This approach combined with the copula provides on-the-fly sampling algorithms which are fast and accurate. Sampling procedures for a D-vine and a canonical vine (C-vine) have been implemented in the uncertainty analysis program UNICORN, that is free to download at http://ssor.twi.tudelft.nl/∼ risk. We can construct and sample a distribution that exactly corresponds to the vine specification, moreover, rank correlations in vine specifications are algebraically independent. Hence, every rank correlation specification on a regular vine is consistent and can be realized.

It is shown in Section 4 that not every symmetric positive definite matrix with one's on the main diagonal can be realized with the vine elliptical copula method. Using the elliptical copula in vines assures existence of a trivariate joint uniform with any correlation structure. However, there are four-dimensional correlation matrices which cannot be realized in this way. Nonetheless, this method can realize more than joint normal method (Example 13).

At the moment it is not clear if every n-dimensional correlation matrix can be realized by a joint uniform distribution. If it is not a case then we would like to know which correlation matrices are correlation matrices of joint uniform distribution and how this set relates to the set of matrices that can be realized by vine-copula method.

Acknowledgment

The authors would like to thank Prof. M. de Bruin, Prof. P. Groeneboom and Prof. S. Kotz for ideas, comments and discussions about topics covert in this article. We would like to express our gratitude to D. Lewandowski for a significant help in obtaining simulation results in Section 5.

References

Bedford, T.J., Cooke, R.M., 2002. Viness—a new graphical model for dependent random variables. Ann. Statist. 30 (4), 1031–1068.
Cooke, R.M., 1997. Markov and entropy properties of tree and vines-dependent variables. In: Proceedings of the ASA Section of Bayesian Statistical Science.
Cooke, R.M., Waij, R., 1986. Monte Carlo sampling for generalized knowledge dependence with application to human reliability. Risk Anal. 6, 335 –343.
Dall'Aglio, G., Kotz, S., Salinetti, G., 1991. Probability Distributions with Given Marginals; Beyond the Copulas. Kulwer Academic Publishers, Dordrecht.
Frank, M.J., 1979. On the simultaneous associativity of $f(x, y)$ and $x + y - f(x, y)$. Aequationes Math. 19, 194–226.
Genest, C., Rivest, L.P., 1993. Statistical inference procedure for bivariate Archimedean copula's. J. Amer. Statist. Assoc. 88 (423), 1034–1043.
Ghosh, S., Henderson, S.G., 2002. Properties of the notra method in higher dimensions. In: Proceedings of the 2002 Winter Simulation Conference.pp. 263–269.
Hoeffding, W., 1940. Masstabinvariante korrelationstheorie. Schrijtfen Math. Inst. Inst. Angew. Math. Univ. Berlin 5, 179–233.
Iman, R., Conver, W., 1982. A distribution-free approach to inducing rank correlation among input variables. Comm. Statist. Simulation Comput. 11 (3), 311–334.
Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman & Hall, London.
Kurowicka, D., Cooke, R.M., 2000. Conditional and partial correlation for graphical uncertainty models. In: Recent Advances in Reldiability Theory.Birkhauser, Boston, pp. 259–276.
Kurowicka, D., Cooke, R.M., 2001. Conditional, partial and rank correlation for elliptical copula; dependence modeling in uncertainty analysis. Proceedings of ESREL 2001.
Kurowicka, D., Cooke, R.M., 2003. A parametrization of positive definite matrices in terms of partial correlation vines. Linear Algebra Appl. 372, 225–251.
Kurowicka, D., Cooke, R.M., 2006. Completion problem with partial correlation vines. Linear Algebra Appl., to appear.
Kurowicka, D., Misiewicz, J., Cooke, R.M., 2000. Elliptical copulae. In: Proceedings of the International Conference on Monte Carlo Simulation—Monte Carlo.pp. 209–214.

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043

Meeuwissen, A., Bedford, T.J., 1997. Minimally informative distributions with given rank correlation for use in uncertainty analysis. J. Statist. Comput. Simulation 57 (1–4), 143–175.
Nelsen, R.B., 1999. An Introduction to Copulas. Springer, New York.
Pearson, K., 1907. Mathematical contributions to the theory of evolution. Biometric, VI. Series.
Yule, G.U., Kendall, M.G., 1965. An Introduction to the Theory of Statistics. 14th ed. Charles Griffin & Co, Belmont, CA.

Please cite this article as: Kurowicka, D., Cooke, R.M., Sampling algorithms for generating joint uniform distributions using the vine-copula method. Computational Statistics & Data Analysis (2007), doi: 10.1016/j.csda.2006.11.043