Pair-copula Constructions of Multiple Dependence

Claudia Czado
Technische Universität München
cczado@ma.tum.de

September, 2007
Reference: Aas, Czado, Frigessi, and Bakken (2007)

Overview

Introduction

- Bedford and Cooke (2001) and Bedford and Cooke (2002) gave a probabilistic construction of multivariate distributions based on simple building blocks called pair-copulas. See also Kurowicka and Cooke (2006).
- Extends work by Joe (1996).

Introduction

- Bedford and Cooke (2001) and Bedford and Cooke (2002) gave a probabilistic construction of multivariate distributions based on simple building blocks called pair-copulas. See also Kurowicka and Cooke (2006).
- Extends work by Joe (1996).
- Overseen so far by statisticians and econometricians.

Introduction

- Bedford and Cooke (2001) and Bedford and Cooke (2002) gave a probabilistic construction of multivariate distributions based on simple building blocks called pair-copulas. See also Kurowicka and Cooke (2006).
- Extends work by Joe (1996).
- Overseen so far by statisticians and econometricians.
- We believe that they provide a central tool in model building, when conditional independence is not natural
- Maintains the logic of building complexity by simple elementary blocks.

Multivariate modeling

- Parametric multivariate distribution
  - Not appropriate when all variables do not have the same distribution.

Multivariate modeling

- Parametric multivariate distribution
  - Not appropriate when all variables do not have the same distribution.
- Marginal distribution + copula
  - Not appropriate when all pairs of variables do not have the same dependency structure.

Multivariate modeling

- Parametric multivariate distribution
  - Not appropriate when all variables do not have the same distribution.
- Marginal distribution + copula
  - Not appropriate when all pairs of variables do not have the same dependency structure.
- Marginal distributions + pair-copula construction
  - Very flexible structure

Multivariate distributions

Consider n random variables $\mathbf{X} = (X_1, \ldots, X_n)$ with

- joint density $f(x_1, \ldots, x_n)$ and marginal densities $f_i(x_i), i = 1, \ldots, n$
- joint cdf $F(x_1, \ldots, x_n)$ and marginal cdf's $F_i(x_i), i = 1, \ldots, n$

Multivariate distributions

Consider n random variables $\mathbf{X} = (X_1, \ldots, X_n)$ with

- joint density $f(x_1, \ldots, x_n)$ and marginal densities $f_i(x_i), i = 1, \ldots, n$
- joint cdf $F(x_1, \ldots, x_n)$ and marginal cdf's $F_i(x_i), i = 1, \ldots, n$
- factorization

$$f(x_1, \ldots, x_n) = \\ f_n(x_n) \cdot f(x_{n-1}|x_n) \cdot f(x_{n-2}|x_{n-1}, x_n) \ldots \cdot f(x_1|x_2, \ldots, x_n) \tag{1}$$

Copula

A copula is a multivariate distribution on $[0,1]^n$ with uniformly distributed marginals.

- copula cdf $C(u_1, \ldots, u_n)$
- copula density $c(u_1, \ldots, u_n)$

Copula

A copula is a multivariate distribution on $[0,1]^n$ with uniformly distributed marginals.

- copula cdf $C(u_1, \ldots, u_n)$
- copula density $c(u_1, \ldots, u_n)$

Using Sklar's Theorem (1959) we have for absolutely continuous distributions with continuous marginal cdf's

$$f(x_1, \ldots, x_n) = c_{12\cdots n}(F_1(x_1), \ldots F_n(x_n)) \cdot f_1(x_1) \cdots f_n(x_n) \qquad (2)$$

for some n-variate copula density $c_{12\cdots n}(\cdot)$.

Pair-copula constructions (PCC)

Copula decompositions (1)

- $n = 2$

$$
\begin{aligned}
&\text{joint density} \\
f(x_1, x_2) &= c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2) \\
&\text{conditional densities} \\
f_{1|2}(x_1|x_2) &= c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \\
f_{2|1}(x_2|x_1) &= c_{12}(F_1(x_1), F_2(x_2)) \cdot f_2(x_2)
\end{aligned}
$$

Copula decompositions (2)

- $n = 3$
  - first decomposition:

$$
\begin{aligned}
f_{1|23}(x_1|x_2, x_3) &= \frac{f_{12|3}(x_1, x_2|x_3)}{f_{2|3}(x_2|x_3)} \\
&= c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)) \cdot f_{1|3}(x_1|x_3)
\end{aligned}
$$

Copula decompositions (2)

- $n = 3$
  - first decomposition:

$$\begin{aligned} f_{1|23}(x_1|x_2, x_3) &= \frac{f_{12|3}(x_1, x_2|x_3)}{f_{2|3}(x_2|x_3)} \\ &= c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)) \cdot f_{1|3}(x_1|x_3) \\ &= c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)) \\ &\quad \times c_{13}(F_1(x_1), F_3(x_3)) \cdot f_1(x_1) \end{aligned}$$

Copula decompositions (2)

- $n = 3$
  - first decomposition:

$$\begin{aligned} f_{1|23}(x_1|x_2,x_3) &= \frac{f_{12|3}(x_1,x_2|x_3)}{f_{2|3}(x_2|x_3)} \\ &= c_{12|3}(F_{1|3}(x_1|x_3),F_{2|3}(x_2|x_3)) \cdot f_{1|3}(x_1|x_3) \\ &= c_{12|3}(F_{1|3}(x_1|x_3),F_{2|3}(x_2|x_3)) \\ &\quad \times c_{13}(F_1(x_1),F_3(x_3)) \cdot f_1(x_1) \end{aligned}$$

  - Second decomposition:

$$\begin{aligned} f(x_1|x_2,x_3) &= c_{13|2}(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2)) \cdot f(x_1|x_2) \\ &= c_{13|2}(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2)) \\ &\quad \times c_{12}(F_1(x_1),F_2(x_2)) \cdot f_1(x_1) \end{aligned}$$

Copula decompositions (3)

- general n
  - For $\mathbf{v} = (v_1, \ldots, v_d)$ and any $j = 1, \ldots, d$

$$
\begin{aligned}
f(x|\mathbf{v}) &= c_{xv_j|\mathbf{v}_{-j}}(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j})) \cdot f(x|\mathbf{v}_{-j}) \\
\mathbf{v}_{-j} &= (v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_d) \\
c_{xv_j|\mathbf{v}_{-j}}(\cdot) &= \text{bivariate copula density}
\end{aligned}
$$

Copula decompositions (3)

- general n
  - For $\mathbf{v} = (v_1, \dots, v_d)$ and any $j = 1, \dots, d$

$$\begin{aligned} f(x|\mathbf{v}) &= c_{xv_j|\mathbf{v}_{-j}}(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j})) \cdot f(x|\mathbf{v}_{-j}) \\ \mathbf{v}_{-j} &= (v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_d) \\ c_{xv_j|\mathbf{v}_{-j}}(\cdot) &= \text{bivariate copula density} \end{aligned}$$

  - Combining this decomposition of the conditional distribution with the factorization (1), we derive at a decomposition of $f(x_1, \dots, x_n)$ that only consist of pair-copulae.
  - We call this a pair-copula construction.

Conditional cdf's

- Univariate v: Since $f(x|v) = c_{xv}(F_x(x), F_v(v))f_x(x)$ we have

$$F(x|v) = \int_{-\infty}^{x} c_{xv}(F_x(u), F_v(v))f_x(u)du$$

Conditional cdf's

- Univariate v: Since $f(x|v) = c_{xv}(F_x(x), F_v(v))f_x(x)$ we have

$$
\begin{aligned}
F(x|v) &= \int_{-\infty}^{x} c_{xv}(F_x(u), F_v(v))f_x(u)du \\
&= \int_{-\infty}^{x} \frac{\partial^2 C_{xv}(F_x(u), F_v(v))}{\partial F_x(u)\partial F_v(v)} f_x(u)du
\end{aligned}
$$

Conditional cdf's

- Univariate v: Since $f(x|v) = c_{xv}(F_x(x), F_v(v))f_x(x)$ we have

$$
\begin{aligned}
F(x|v) &= \int_{-\infty}^{x} c_{xv}(F_x(u), F_v(v))f_x(u)du \\
&= \int_{-\infty}^{x} \frac{\partial^2 C_{xv}(F_x(u), F_v(v))}{\partial F_x(u)\partial F_v(v)} f_x(u)du \\
&= \frac{\partial}{\partial F_v(v)} \int_{-\infty}^{x} \frac{\partial C_{xv}(F_x(u), F_v(v))}{\partial F_x(u)} \frac{dF_x(u)}{du} du
\end{aligned}
$$

Conditional cdf's

- Univariate v: Since $f(x|v) = c_{xv}(F_x(x), F_v(v))f_x(x)$ we have

$$
\begin{aligned}
F(x|v) &= \int_{-\infty}^{x} c_{xv}(F_x(u), F_v(v))f_x(u)du \\
&= \int_{-\infty}^{x} \frac{\partial^2 C_{xv}(F_x(u), F_v(v))}{\partial F_x(u)\partial F_v(v)} f_x(u)du \\
&= \frac{\partial}{\partial F_v(v)} \int_{-\infty}^{x} \frac{\partial C_{xv}(F_x(u), F_v(v))}{\partial F_x(u)} \frac{dF_x(u)}{du} du \\
&= \frac{\partial\, C_{xv}(F_x(x), F_v(v))}{\partial F_v(v)}
\end{aligned}
$$

Conditional cdf's (2)

- In the case of uniform marginal cdf's, i.e. $F_x(x) = x$ and $F_v(v) = v$ we have for parametrized $C_{xv}(x, v, \Theta)$

$$h(x, v, \Theta) := F(x|v) = \frac{\partial\, C_{xv}(x, v, \Theta)}{\partial v}.$$

These h functions will become essential for simulation and inference.

Conditional cdf's (2)

- In the case of uniform marginal cdf's, i.e. $F_x(x) = x$ and $F_v(v) = v$ we have for parametrized $C_{xv}(x, v, \Theta)$

$$h(x, v, \Theta) := F(x|v) = \frac{\partial\, C_{xv}(x, v, \Theta)}{\partial v}.$$

These h functions will become essential for simulation and inference.

- Under regularity conditions Joe (1996) showed that

$$F(x|\mathbf{v}) = \frac{\partial\, C_{x,v_j|\mathbf{v}_{-j}}(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j}))}{\partial F(v_j|\mathbf{v}_{-j})},$$

Examples of PCC's

- $n = 3$

$$f(x_1, x_2, x_3) = f_3(x_3) f(x_2|x_3) f(x_1|x_2, x_3)$$

Examples of PCC's

- $n = 3$

$$
\begin{aligned}
f(x_1, x_2, x_3) = & \\
& f_3(x_3) f(x_2|x_3) f(x_1|x_2, x_3) \\
= \quad & f_3(x_3)[c_{23}(F_2(x_2), F_3(x_3)) f_2(x_2)] \\
& \times [c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)) f(x_1|x_3)]
\end{aligned}
$$

Examples of PCC's

- $n = 3$

$$
\begin{aligned}
f(x_1, x_2, x_3) = & \\
& f_3(x_3) f(x_2|x_3) f(x_1|x_2, x_3) \\
= & f_3(x_3)[c_{23}(F_2(x_2), F_3(x_3)) f_2(x_2)] \\
& \times [c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)) f(x_1|x_3)] \\
= & f_3(x_3)[c_{23}(F_2(x_2), F_3(x_3)) f_2(x_2)] \\
& \times [c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3))] \\
& \times [c_{13}(F_1(x_1), F_3(x_3)) f_1(x_1)]
\end{aligned}
$$

Examples of PCC's

- $n = 3$

$$
\begin{aligned}
f(x_1, x_2, x_3) = \\
& f_3(x_3) f(x_2|x_3) f(x_1|x_2, x_3) \\
= \; & f_3(x_3)[c_{23}(F_2(x_2), F_3(x_3)) f_2(x_2)] \\
& \times [c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)) f(x_1|x_3)] \\
= \; & f_3(x_3)[c_{23}(F_2(x_2), F_3(x_3)) f_2(x_2)] \\
& \times [c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3))] \\
& \times [c_{13}(F_1(x_1), F_3(x_3)) f_1(x_1)] \\
= \; & f_1(x_1) f_2(x_2) f_3(x_3) \times c_{23}(F_2(x_2), F_3(x_3)) \\
& \times c_{13}(F_1(x_1), F_3(x_3)) c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3))
\end{aligned}
$$

Examples of PCC's

- $n = 3$

$$
\begin{aligned}
f(x_1, x_2, x_3) = & \\
& f_3(x_3) f(x_2|x_3) f(x_1|x_2, x_3) \\
= & \ f_3(x_3)[c_{23}(F_2(x_2), F_3(x_3)) f_2(x_2)] \\
& \times [c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)) f(x_1|x_3)] \\
= & \ f_3(x_3)[c_{23}(F_2(x_2), F_3(x_3)) f_2(x_2)] \\
& \times [c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3))] \\
& \times [c_{13}(F_1(x_1), F_3(x_3)) f_1(x_1)] \\
= & \ f_1(x_1) f_2(x_2) f_3(x_3) \times c_{23}(F_2(x_2), F_3(x_3)) \\
& \times c_{13}(F_1(x_1), F_3(x_3)) c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3))
\end{aligned}
$$

If X_1 ind. X_2 given X_3 $\Rightarrow c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)) = 1$

- $n = 5$

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4, x_5) = \\
& f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \cdot f_5(x_5) \\
\cdot \quad & c_{12}\left(F_1(x_1), F_2(x_2)\right) \cdot c_{13}\left(F_1(x_1), F_3(x_3)\right) \cdot c_{14}\left(F_1(x_1), F_4(x_4)\right) \\
\cdot \quad & c_{15}\left(F_1(x_1), F_5(x_5)\right) \cdot c_{23|1}\left(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)\right) \\
\cdot \quad & c_{24|1}\left(F_{2|1}(x_2|x_1), F_{4|1}(x_4|x_1)\right) \cdot c_{25|1}\left(F_{2|1}(x_2|x_1), F_{5|1}(x_5|x_1)\right) \\
\cdot \quad & c_{34|12}\left(F_{3|12}(x_3|x_1, x_2), F_{4|12}(x_4|x_1, x_2)\right) \\
\cdot \quad & c_{35|12}\left(F_{3|12}(x_3|x_1, x_2), F_{5|12}(x_5|x_1, x_2)\right) \\
\cdot \quad & c_{45|123}\left(F_{4|123}(x_4|x_1, x_2, x_3), F_{5|123}(x_5|x_1, x_2, x_3)\right),
\end{aligned}
$$

There are 240 different such constructions for $n = 5$

Vines

- Hence, for high-dimensional distributions there are many possible pair-copula constructions.
- Bedford and Cooke (2001) introduced a graphical model called regular vine to help organize them.
- The class of regular vines is large and embraces a large number of possible PCC's.

Vines

- Hence, for high-dimensional distributions there are many possible pair-copula constructions.
- Bedford and Cooke (2001) introduced a graphical model called regular vine to help organize them.
- The class of regular vines is large and embraces a large number of possible PCC's.
- We concentrate on two special cases (Kurowicka and Cooke 2004):
  - D-vine
  - Canonical Vine

Vines(2)

- An *n*-dimensional vine is represented by n-1 trees.

Vines(2)

- An n-dimensional vine is represented by n-1 trees.
- Tree j has $n+1-j$ nodes and $n-j$ edges.

Vines(2)

- An *n*-dimensional vine is represented by n-1 trees.
- Tree j has $n+1-j$ nodes and $n-j$ edges.
- Each edge corresponds to a pair-copula density.

Vines(2)

- An *n-dimensional vine* is represented by n-1 trees.
- Tree j has $n+1-j$ nodes and $n-j$ edges.
- Each edge corresponds to a pair-copula density.
- Edges in tree j become nodes in tree $j+1$.

Vines(2)

- An *n*-dimensional vine is represented by n-1 trees.
- Tree j has $n+1-j$ nodes and $n-j$ edges.
- Each edge corresponds to a pair-copula density.
- Edges in tree j become nodes in tree $j+1$.
- Two nodes in tree $j+1$ are joined by an edge if the corresponding edges in tree j share a node.

Vines(2)

- An n-dimensional vine is represented by n-1 trees.
- Tree j has $n+1-j$ nodes and $n-j$ edges.
- Each edge corresponds to a pair-copula density.
- Edges in tree j become nodes in tree $j+1$.
- Two nodes in tree $j+1$ are joined by an edge if the corresponding edges in tree j share a node.
- The complete decomposition is defined by the $\frac{n(n-1)}{2}$ edges (i.e. pair copula densities) and the marginal densities.

Canonical and D-vines

- Canonical vine:
  A regular vine for which each tree has a unique node that is connected to $n - j$ edges.

Canonical and D-vines

- Canonical vine:
  A regular vine for which each tree has a unique node that is connected to $n-j$ edges.
- D-vine:
  A regular vine for which no node in any tree is connected to more than two edges

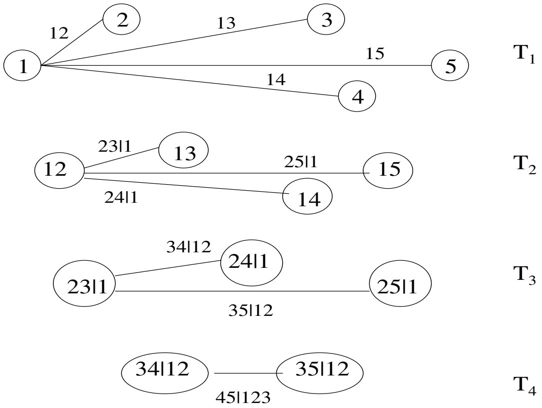
Canonical and D-vines

- Canonical vine:
  A regular vine for which each tree has a unique node that is connected to $n-j$ edges.
- D-vine:
  A regular vine for which no node in any tree is connected to more than two edges
- Abbreviations:

$$
\begin{aligned}
c_{ij|\mathbf{v}} &:= c_{ij|\mathbf{v}}(F_{i|\mathbf{v}}(x_i|\mathbf{x_v}), F_{j|\mathbf{v}}(x_j|\mathbf{x_v})) \\
f_j &:= f_j(x_j) \\
f_{\mathbf{v}} &:= f_{\mathbf{v}}(x_{\mathbf{v}})
\end{aligned}
$$

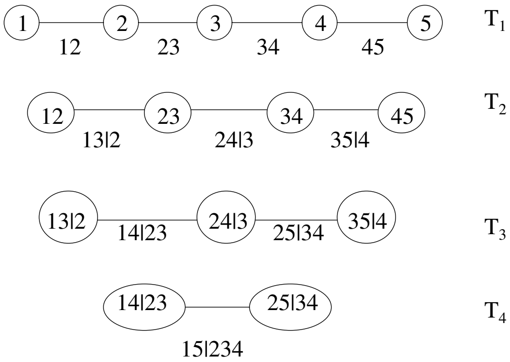
Five dimensional canonical vine

$$f_{12345} = f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdot f_5 \cdot c_{12} \cdot c_{13} \cdot c_{14} \cdot c_{15} \cdot c_{23|1} \cdot c_{24|1} \cdot c_{25|1} \cdot c_{34|12} \cdot c_{35|12} \cdot c_{45|123}$$



Five dimensional D-vine

$$f_{12345} = f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdot f_5 \cdot c_{12} \cdot c_{23} \cdot c_{34} \cdot c_{45} \cdot c_{13|2} \cdot c_{24|3} \cdot c_{35|4} \cdot c_{14|23} \cdot c_{25|34} \cdot c_{15|234}$$



General density expressions

- Canonical vine density

$$\prod_{k=1}^{n} f_k \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{j,j+i|1,\dots,j-1}$$

General density expressions

- Canonical vine density

$$\prod_{k=1}^{n} f_k \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{j,j+i|1,\ldots,j-1}$$

- D-vine density

$$\prod_{k=1}^{n} f_k \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{i,i+j|i+1,\ldots,i+j-1}$$

where index j identifies the trees, while i runs over the edges in each tree.

Simulation of Pair-copula constructions

For simplicity we assume that the marginal distributions are uniform, i.e. $f_k(x_k) = 1$ and $F_k(x_k) = x_k$ for $k = 1, \ldots, n$. Both canonical and D-vines are simulated as follows:

For simplicity we assume that the marginal distributions are uniform, i.e. $f_k(x_k) = 1$ and $F_k(x_k) = x_k$ for $k = 1, \ldots, n$. Both canonical and D-vines are simulated as follows:

- Sample $w_i, i = 1, \ldots, n$ independent uniform on [0,1].
- Set

$$
\begin{aligned}
x_1 &= w_1 \\
x_2 &= F^{-1}_{2|1}(w_2|x_1) \\
x_3 &= F^{-1}_{3|1,2}(w_3|x_1, x_2) \\
.. &= .... \\
x_n &= F^{-1}_{n|1,2,\ldots,n-1}(w_n|x_1, \ldots, x_{n-1}).
\end{aligned}
$$

The simulation procedure for the canonical and D-vine differs in how $F(x_j|x_1, x_2, \ldots, x_{j-1})$ is computed.

The simulation procedure for the canonical and D-vine differs in how $F(x_j|x_1, x_2, \ldots, x_{j-1})$ is computed.

- For the canonical vine $F(x_j|x_1, x_2, \ldots, x_{j-1})$ is computed as

$$\frac{\partial\, C_{j,j-1|1,\ldots j-2}(F(x_j|x_1, \ldots, x_{j-2}), F(x_{j-1}|x_1, \ldots, x_{j-2}))}{\partial F(x_{j-1}|x_1, \ldots, x_{j-2})}.$$

The simulation procedure for the canonical and D-vine differs in how $F(x_j|x_1, x_2, \ldots, x_{j-1})$ is computed.

- For the canonical vine $F(x_j|x_1, x_2, \ldots, x_{j-1})$ is computed as

$$\frac{\partial\, C_{j,j-1|1,\ldots j-2}(F(x_j|x_1, \ldots, x_{j-2}), F(x_{j-1}|x_1, \ldots, x_{j-2}))}{\partial F(x_{j-1}|x_1, \ldots, x_{j-2})}.$$

- For the D-vine $F(x_j|x_1, x_2, \ldots, x_{j-1})$ is computed as

$$\frac{\partial\, C_{j,1|2,\ldots j-1}(F(x_j|x_2, \ldots, x_{j-1}), F(x_1|x_2, \ldots, x_{j-1}))}{\partial F(x_1|x_2, \ldots, x_{j-1})}.$$

Simulation for $n = 3$

For $n = 3$ we have canonical = D-vine

- Sample $w_i, i = 1, \ldots, 3$ i.i.d $U[0, 1]$
- Set $x_1 = w_1$

Simulation for $n = 3$

For $n = 3$ we have canonical = D-vine

- Sample $w_i, i = 1, \ldots, 3$ i.i.d $U[0, 1]$
- Set $x_1 = w_1$
- Set $x_2 = h^{-1}(w_2, x_1, \Theta_{12})$, since $F_{2|1}(x_2|x_1) = h(x_2, x_1, \Theta_{12})$

Simulation for $n = 3$

For $n = 3$ we have canonical = D-vine

- Sample $w_i, i = 1, \ldots, 3$ i.i.d $U[0,1]$
- Set $x_1 = w_1$
- Set $x_2 = h^{-1}(w_2, x_1, \Theta_{12})$, since $F_{2|1}(x_2|x_1) = h(x_2, x_1, \Theta_{12})$
- $x_3 = h^{-1}(h^{-1}(w_3, h(x_2, x_1, \Theta_{12}), \Theta_{23|1}), x_1, \Theta_{13})$ since

$$
\begin{aligned}
F(x_3|x_1, x_2) &= h(F_{3|1}(x_3|x_1), F_{2|1}(x_2|x_1), \Theta_{23|1}) \\
&= h(h(x_3, x_1, \Theta_{13}), h(x_2, x_1, \Theta_{12})), \Theta_{23|1})
\end{aligned}
$$

Simulation algorithms for general n can be found in Aas, Czado, Frigessi, and Bakken (2007).

Model selection

Inference for a pair-copula constructions should in principle consider

1. The specific pair-copula construction

Model selection

Inference for a pair-copula constructions should in principle consider

1. The specific pair-copula construction
2. The choice of the pair-copula type such as
   - Gaussian (no tail dependence)
   - Student's t (upper and lower tail dependence)
   - Clayton (lower tail dependence)
   - Gumbel (upper tail dependence)

Model selection

Inference for a pair-copula constructions should in principle consider

1. The specific pair-copula construction
2. The choice of the pair-copula type such as
   - Gaussian (no tail dependence)
   - Student's t (upper and lower tail dependence)
   - Clayton (lower tail dependence)
   - Gumbel (upper tail dependence)
3. The estimation of the parameters of the chosen pair-copulae

Number of PCC's and some smart choices

- There are $\frac{n!}{2}$ different canonical vines and $\frac{n!}{2}$ different D-vines
- For $n > 3$ canonical and D-vines are always different
- There are many regular vine decompositions that are neither canonical nor D-vines

Number of PCC's and some smart choices

- There are $\frac{n!}{2}$ different canonical vines and $\frac{n!}{2}$ different D-vines
- For $n > 3$ canonical and D-vines are always different
- There are many regular vine decompositions that are neither canonical nor D-vines
- For $n \leq 4$ one can estimate parameters for all decompositions if only one type of pair-copulae is used.
- For higher dimensions, one can first estimate most important bivariate relationships and use them to determine which decomposition to use.
- In D-vines one can select more freely which pairs to model than in the canonical vine.

Inference

- We assume that data $x_t = (x_{1t}, \cdots, x_{nt})$ for $t = 1, \cdots, T$ has uniform margins.

Inference

- We assume that data $x_t = (x_{1t}, \cdots, x_{nt})$ for $t = 1, \cdots, T$ has uniform margins.
- For general n and a specific vine decomposition the log likelihood $L(\Theta|x)$ can be computed using the density expressions for canonical and D-vines. Here x denotes the complete data and Θ the vector of all pairwise copula parameters.

Inference

- We assume that data $x_t = (x_{1t}, \cdots, x_{nt})$ for $t = 1, \cdots, T$ has uniform margins.
- For general n and a specific vine decomposition the log likelihood $L(\Theta|x)$ can be computed using the density expressions for canonical and D-vines. Here x denotes the complete data and Θ the vector of all pairwise copula parameters.
- Algorithmic expressions are given in (Aas et al. 2007).

Inference for $n = 3$

$$
\begin{aligned}
L(\Theta|x) &= \sum_{t=1}^{T} [\log c_{13}(x_{1t}, x_{3t}|\Theta_{13}) + \log c_{23}(x_{2t}, x_{3t}|\Theta_{23}) \\
&+ \log c_{12|3}(v_{1t}, v_{2t}|\Theta_{1|23})]
\end{aligned}
$$

where

$$
\begin{aligned}
v_{1t} &:= F(x_{1t}|x_{3t}) = h(x_{1t}, x_{3t}, \Theta_{13}) \\
v_{2t} &:= F(x_{2t}|x_{3t}) = h(x_{2t}, x_{3t}, \Theta_{23}) \\
\Theta &:= (\Theta_{13}, \Theta_{23}, \Theta_{12|3})
\end{aligned}
$$

Starting values

- Estimate Θ_{13} from data $(x_{1,t}, x_{3,t}, t = 1, \cdots, T)$
- Estimate Θ_{23} from data $(x_{2,t}, x_{3,t}, t = 1, \cdots, T)$

Starting values

- Estimate Θ_{13} from data $(x_{1,t}, x_{3,t}, t = 1, \cdots, T)$
- Estimate Θ_{23} from data $(x_{2,t}, x_{3,t}, t = 1, \cdots, T)$
- Use these estimates to define

$$\begin{aligned} \hat{v}_{1t} &:= h(x_{1t}, x_{3t}, \hat{\Theta}_{13}) \\ \hat{v}_{2t} &:= h(x_{2t}, x_{3t}, \hat{\Theta}_{23}) \end{aligned}$$

Starting values

- Estimate Θ_{13} from data $(x_{1,t}, x_{3,t}, t = 1, \cdots, T)$
- Estimate Θ_{23} from data $(x_{2,t}, x_{3,t}, t = 1, \cdots, T)$
- Use these estimates to define

$$\begin{aligned}
\hat{v}_{1t} &:= h(x_{1t}, x_{3t}, \hat{\Theta}_{13}) \\
\hat{v}_{2t} &:= h(x_{2t}, x_{3t}, \hat{\Theta}_{23})
\end{aligned}$$

- Estimate $\Theta_{12|3}$ from data $(\hat{v}_{1,t}, \hat{v}_{2,t}, t = 1, \cdots, T)$

Starting values

- Estimate Θ_{13} from data $(x_{1,t}, x_{3,t}, t = 1, \cdots, T)$
- Estimate Θ_{23} from data $(x_{2,t}, x_{3,t}, t = 1, \cdots, T)$
- Use these estimates to define

$$
\begin{aligned}
\hat{v}_{1t} &:= h(x_{1t}, x_{3t}, \hat{\Theta}_{13}) \\
\hat{v}_{2t} &:= h(x_{2t}, x_{3t}, \hat{\Theta}_{23})
\end{aligned}
$$

- Estimate $\Theta_{12|3}$ from data $(\hat{v}_{1,t}, \hat{v}_{2,t}, t = 1, \cdots, T)$
- For Gauss and t-copula an estimate of Kendall's tau is used to estimate the correlation parameter, since this is independent of margins.
- Kendall's tau can also be used for the Clayton copula

Necessary Expressions

- For each pair-copula in the decomposition, three expressions are necessary:
  - bivariate density
  - h-function $h(x, v, \Theta) = F(x|v) = \frac{C_{x,v}(x,v,\Theta)}{\partial v}$
  - inverse of h function wrt to x

Necessary Expressions

- For each pair-copula in the decomposition, three expressions are necessary:
  - bivariate density
  - h-function $h(x, v, \Theta) = F(x|v) = \frac{C_{x,v}(x,v,\Theta)}{\partial v}$
  - inverse of h function wrt to x
- For Clayton this is available in closed form.
- For the Gumbel, Gaussian, Student's the inverse of the h-function must be obtained numerically.

Application: Financial Returns

- Tail dependence properties are important in finance.

Application: Financial Returns

- Tail dependence properties are important in finance.
- n-dimensional Student's t-copula has been often used for modeling financial returns, however it only has a single parameter for tail dependence.

Application: Financial Returns

- Tail dependence properties are important in finance.
- *n*-dimensional Student's t-copula has been often used for modeling financial returns, however it only has a single parameter for tail dependence.
- Pair-copula constructions allows for multiple parameters for modeling tail dependence.

Data set

- Daily data from Jan. 4, 1999 until July 8, 2003 for

T	=	Norwegian stock index (TOTX)
M	=	MSCI world stock index
B	=	Norwegian bond index (BRIX)
S	=	SSBWG hedged bond index

Data set

- Daily data from Jan. 4, 1999 until July 8, 2003 for

  T = Norwegian stock index (TOTX)
  M = MSCI world stock index
  B = Norwegian bond index (BRIX)
  S = SSBWG hedged bond index

- Standardized residuals of a AR(1)-GARCH(1,1) model for the log return $x_{i,t}$

$$\begin{aligned} x_{i,t} &= c_i + \alpha_i\, x_{i,t-1} + \sigma_{i,t}\, z_{i,t}, \\ \mathrm{E}[z_{t,i}] &= 0 \text{ and } \mathrm{Var}[z_{t,i}] = 1, \\ \sigma^2_{i,t} &= a_{i,0} + a_i\, \epsilon^2_{i,t-1} + b_i\, \sigma^2_{i,t-1} \text{ where } \epsilon_{i,t-1} = \sigma_{i,t}\, z_{i,t} \end{aligned}$$

  are found to be independent and are transformed to uniform margins

Data set

- Fitted degree of freedom for a t-copula to each pair

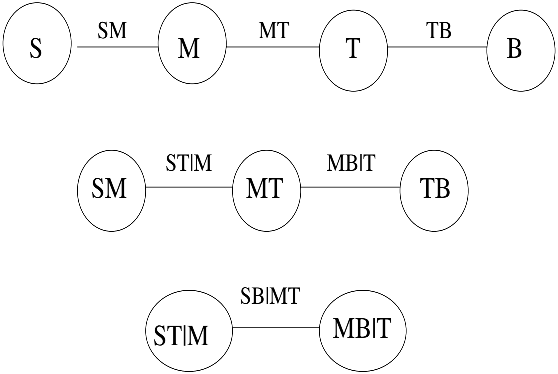
Between	M	T	B
S	4.21	34.16	14.47
M		8.03	15.48
T			12.60

Which PCC?

Strongest dependence between S and M, M and T and T and B, which will be used in the top tree of a D-vine

Which PCC?

Strongest dependence between S and M, M and T and T and B, which will be used in the top tree of a D-vine



Parameter estimates

Param	*Start*	*Final*
ρ_{SM}	-0.25	-0.25
ρ_{MT}	0.47	0.47
ρ_{TB}	-0.17	-0.17
$\rho_{ST\|M}$	-0.11	-0.11
$\rho_{MB\|T}$	0.02	0.03
$\rho_{SB\|MT}$	0.29	0.28
ν_{SM}	4.21	4.34
ν_{MT}	16.65	16.26
ν_{TB}	12.60	13.17
$\nu_{ST\|M}$	300.00	300.00
$\nu_{MB\|T}$	130.33	45.59
$\nu_{SB\|MT}$	15.58	15.04
log.likelih.	267.86	268.17

Parameter estimates

Param	*Start*	*Final*
ρ_{SM}	-0.25	-0.25
ρ_{MT}	0.47	0.47
ρ_{TB}	-0.17	-0.17
$\rho_{ST\|M}$	-0.11	-0.11
$\rho_{MB\|T}$	0.02	0.03
$\rho_{SB\|MT}$	0.29	0.28
ν_{SM}	4.21	4.34
ν_{MT}	16.65	16.26
ν_{TB}	12.60	13.17
$\nu_{ST\|M}$	300.00	300.00
$\nu_{MB\|T}$	130.33	45.59
$\nu_{SB\|MT}$	15.58	15.04
log.likelih.	267.86	268.17

- BRIX and MSCI are nearly independent given TOTX

Parameter estimates

Param	*Start*	*Final*
ρ_{SM}	-0.25	-0.25
ρ_{MT}	0.47	0.47
ρ_{TB}	-0.17	-0.17
$\rho_{ST\|M}$	-0.11	-0.11
$\rho_{MB\|T}$	0.02	0.03
$\rho_{SB\|MT}$	0.29	0.28
ν_{SM}	4.21	4.34
ν_{MT}	16.65	16.26
ν_{TB}	12.60	13.17
$\nu_{ST\|M}$	300.00	300.00
$\nu_{MB\|T}$	130.33	45.59
$\nu_{SB\|MT}$	15.58	15.04
log.likelih.	267.86	268.17

- BRIX and MSCI are nearly independent given TOTX
- SSBWG and BRIX are not independent given TOTX and MSCI

Parameter estimates

Param	*Start*	*Final*
ρ_{SM}	-0.25	-0.25
ρ_{MT}	0.47	0.47
ρ_{TB}	-0.17	-0.17
$\rho_{ST\|M}$	-0.11	-0.11
$\rho_{MB\|T}$	0.02	0.03
$\rho_{SB\|MT}$	0.29	0.28
ν_{SM}	4.21	4.34
ν_{MT}	16.65	16.26
ν_{TB}	12.60	13.17
$\nu_{ST\|M}$	300.00	300.00
$\nu_{MB\|T}$	130.33	45.59
$\nu_{SB\|MT}$	15.58	15.04
log.likelih.	267.86	268.17

- BRIX and MSCI are nearly independent given TOTX
- SSBWG and BRIX are not independent given TOTX and MSCI
- cond. distribution of SSBWG and TOTX given MSCI is normal

Parameter estimates

Param	*Start*	*Final*
ρ_{SM}	-0.25	-0.25
ρ_{MT}	0.47	0.47
ρ_{TB}	-0.17	-0.17
$\rho_{ST\|M}$	-0.11	-0.11
$\rho_{MB\|T}$	0.02	0.03
$\rho_{SB\|MT}$	0.29	0.28
ν_{SM}	4.21	4.34
ν_{MT}	16.65	16.26
ν_{TB}	12.60	13.17
$\nu_{ST\|M}$	300.00	300.00
$\nu_{MB\|T}$	130.33	45.59
$\nu_{SB\|MT}$	15.58	15.04
log.likelih.	267.86	268.17

- BRIX and MSCI are nearly independent given TOTX
- SSBWG and BRIX are not independent given TOTX and MSCI
- cond. distribution of SSBWG and TOTX given MSCI is normal
- cond. distribution of MSCI and BRIX given TOTX is close to normal

Comparison with multivariate Student's t-copula

- AIC

4 dim. Student's t-copula	-512.33
4 dim. Student's t pair copula decomposition	-487.42

[1] Can be used approximately since the 4 dim. Student's t-copula is nested within the 4 dim. Student's pair copula decomposition for large df's

Comparison with multivariate Student's t-copula

- AIC

4 dim. Student's t-copula	-512.33
4 dim. Student's t pair copula decomposition	-487.42

- Likelihood ratio test statistic [1]
  - 2 * Likelihood difference is 34.92 with 5 df
  - P-value is $1.5610^{-6} \Rightarrow$ 4 dim. Student's t-copula is rejected in favor of the PCC.

[1] Can be used approximately since the 4 dim. Student's t-copula is nested within the 4 dim. Student's pair copula decomposition for large df's

Tail dependence

- Upper and lower tail dependence coefficient for bivariate t-copula

$$\lambda_l(X,Y) = \lambda_u(X,Y) = 2\, t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$$

Tail dependence

- Upper and lower tail dependence coefficient for bivariate t-copula

$$\lambda_l(X,Y) = \lambda_u(X,Y) = 2\, t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$$

- Estimated tail dependence coefficients

Margin	*PCC*	*Student's t-copula*
SM	0.0279	0.0001
MT	0.0229	0.0317
TB	0.0005	0.0003

The probability of a joint large portfolio loss in SSBWG and BRIX stocks is much larger in the 4 dim. Student's t PCC compared to a 4 dim. Student's t-copula.

Bayesian model selection among PCC's

joint work with Dr. A. Min

- If a single pair copula type is used in a PCC of dimension n, then all other possible PCC's of dimension n provide a factorization of the same joint density

Bayesian model selection among PCC's

joint work with Dr. A. Min

- If a single pair copula type is used in a PCC of dimension n, then all other possible PCC's of dimension n provide a factorization of the same joint density
- So we want to find reduced PCC's which best fits the data, i.e. let the data discover conditional independence conditions.

Bayesian model selection among PCC's

joint work with Dr. A. Min

- If a single pair copula type is used in a PCC of dimension n, then all other possible PCC's of dimension n provide a factorization of the same joint density
- So we want to find reduced PCC's which best fits the data, i.e. let the data discover conditional independence conditions.
- $n = 3$:

Full PCC's:	$c_{12}c_{23}c_{13\|2} = c_{13}c_{23}c_{12\|3} = c_{12}c_{13}c_{23\|1}$
Reduced by 1 pair-copula:	$c_{12}c_{23},\ c_{12}c_{13\|2}$ $c_{13}c_{23},\ c_{23}c_{12\|3},\ c_{13}c_{12},\ c_{13}c_{23\|1}$
Reduced by 2 pair copulas:	$c_{12},\ c_{13},\ c_{23}$ $c_{13\|2},\ c_{12\|3},\ c_{23\|1}$

- Identifiability: Note that $c_{12}c_{23|1} = c_{23}c_{12|3}$ since $c_{12}c_{23|1} = c_{12}c_{13}c_{23|1} = c_{13}c_{23}c_{12|3} = c_{23}c_{12|3}$ when $c_{13} = 1$.

Bayesian framework of PCC's with $n \leq 3$

- Each of the 13 different PCC's with $n \leq 3$ is identified by a model index $\mathbf{m} = (m_1, m_2, m_3) = (i_1 j_1, i_2, j_2, i_3 j_3 | k_3)$.
  Example: $c_{12} c_{23|1}$ corresponds to $\mathbf{m} = (12, 00, 23|1)$.
- Each PCC model $\mathbf{m}$ has parameter vector $\boldsymbol{\theta}_m = (\boldsymbol{\theta}_{m_1}, \boldsymbol{\theta}_{m_2}, \boldsymbol{\theta}_{m_3})$.
  Example: Model $\mathbf{m} = (12, 23, 00|0)$ has parameter vector $\boldsymbol{\theta}_m = (\boldsymbol{\theta}_{12}, \boldsymbol{\theta}_{23}, \boldsymbol{\theta}_{00|0}) = (\boldsymbol{\theta}_{12}, \boldsymbol{\theta}_{23})$.

Bayesian framework of PCC's with $n \leq 3$

- Each of the 13 different PCC's with $n \leq 3$ is identified by a model index $\mathbf{m} = (m_1, m_2, m_3) = (i_1 j_1, i_2, j_2, i_3 j_3 | k_3)$.
  Example: $c_{12} c_{23|1}$ corresponds to $\mathbf{m} = (12, 00, 23|1)$.
- Each PCC model $\mathbf{m}$ has parameter vector $\boldsymbol{\theta}_m = (\boldsymbol{\theta}_{m_1}, \boldsymbol{\theta}_{m_2}, \boldsymbol{\theta}_{m_3})$.
  Example: Model $\mathbf{m} = (12, 23, 00|0)$ has parameter vector $\boldsymbol{\theta}_m = (\boldsymbol{\theta}_{12}, \boldsymbol{\theta}_{23}, \boldsymbol{\theta}_{00|0}) = (\boldsymbol{\theta}_{12}, \boldsymbol{\theta}_{23})$.
- Goal is to estimate the best fitting model $\mathbf{m}$ and the corresponding parameter vector $\boldsymbol{\theta}_m$ using a Bayesian approach, i.e. the model $\mathbf{m}$ and θ_m are considered random quantities.
- Inference about $\mathbf{m}$ and θ_m is done via the joint posterior distribution of $(\mathbf{m}, \boldsymbol{\theta}_m)$.

MCMC algorithm for PCC's with $n \leq 3$

- Problems:
  - Want to facilitate estimation without having to fit all models

MCMC algorithm for PCC's with $n \leq 3$

- Problems:
  - Want to facilitate estimation without having to fit all models
  - Joint posterior is not analytically tractable

MCMC algorithm for PCC's with $n \leq 3$

- Problems:
  - Want to facilitate estimation without having to fit all models
  - Joint posterior is not analytically tractable
- Approach: Construct a Markov Chain Monte Carlo (MCMC) algorithm which simultaneously estimates $\mathbf{m}$ and $\boldsymbol{\theta}_m$.

MCMC algorithm for PCC's with $n \leq 3$

- Problems:
  - Want to facilitate estimation without having to fit all models
  - Joint posterior is not analytically tractable
- Approach: Construct a Markov Chain Monte Carlo (MCMC) algorithm which simultaneously estimates $\mathbf{m}$ and $\boldsymbol{\theta}_m$.
- Requirement: Need to accommodate varying model dimension, i.e. construct MCMC iterates $\boldsymbol{\theta}_{\mathbf{m}^r} = (\boldsymbol{\theta}^r_{m^r_1}, \boldsymbol{\theta}^r_{m^r_2}, \boldsymbol{\theta}^r_{m^r_3})$ where $\mathbf{m}^r$ is the current model at iteration r.

MCMC algorithm for PCC's with $n \leq 3$

- Problems:
  - Want to facilitate estimation without having to fit all models
  - Joint posterior is not analytically tractable
- Approach: Construct a Markov Chain Monte Carlo (MCMC) algorithm which simultaneously estimates $\mathbf{m}$ and $\boldsymbol{\theta}_m$.
- Requirement: Need to accommodate varying model dimension, i.e. construct MCMC iterates $\boldsymbol{\theta}_{\mathbf{m}^r} = (\boldsymbol{\theta}^r_{m^r_1}, \boldsymbol{\theta}^r_{m^r_2}, \boldsymbol{\theta}^r_{m^r_3})$ where $\mathbf{m}^r$ is the current model at iteration r.
- Solution: Reversible jump (RJ) MCMC proposed by (Green 1995)

General RJ MCMC (1)

- algorithm stays in current model using a Metropolis Hastings (MH) step
- algorithm moves to a larger model using a MH step (birth)
- algorithm moves to a smaller model using a MH step (death)

General RJ MCMC (2)

Model	M_1		M_2
Parameter	$\boldsymbol{\theta}^{(1)} \in \mathbf{R}^{d_1}$		$\boldsymbol{\theta}^{(2)} \in \mathbf{R}^{d_2}$
	d_1	$<$	d_2
Proposal	$\boldsymbol{\eta}^{(1)} \sim \varphi_1(\cdot)$		$\boldsymbol{\eta}^{(2)} \sim \varphi_2(\cdot)$
	$\boldsymbol{\eta}^{(1)} \in \mathbf{R}^{a_1}$		$\boldsymbol{\eta}^{(2)} \in \mathbf{R}^{a_2}$
Dimension matching	$d_1 + a_1$	$=$	$d_2 + a_2$
Bijection	$\begin{pmatrix} \boldsymbol{\theta}^{(1)} \\ \boldsymbol{\eta}^{(1)} \end{pmatrix}$	$\leftrightarrow$	$\begin{pmatrix} \boldsymbol{\theta}^{(2)} \\ \boldsymbol{\eta}^{(2)} \end{pmatrix}$

Acceptance probability for MH step from M_1 to M_2

$$\alpha(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}) = \min\{1, A\}, \text{ where}$$

$$A := \frac{p(2, \boldsymbol{\theta}^{(2)}|\mathbf{y})}{p(1, \boldsymbol{\theta}^{(1)}|\mathbf{y})} \cdot \frac{p_{2\to1}}{p_{1\to2}} \cdot \frac{\varphi_2(\boldsymbol{\eta}^{(2)})}{\varphi_1(\boldsymbol{\eta}^{(1)})} \cdot \left|\frac{\partial\left(\boldsymbol{\theta}^{(2)}, \boldsymbol{\eta}^{(2)}\right)}{\partial\left(\boldsymbol{\theta}^{(1)}, \boldsymbol{\eta}^{(1)}\right)}\right|$$

$p(k, \boldsymbol{\theta}^{(k)}|y)$ = joint posterior density of M_k and $\boldsymbol{\theta}^{(k)}$

$p_{i\to j}$ = prior switching probability from M_i to M_j

$\varphi_k(\boldsymbol{\eta}^{(k)})$ = proposal distribution of $\boldsymbol{\eta}^{(k)}$ (suitably chosen)

$\left|\frac{\partial\left(\boldsymbol{\theta}^{(2)},\boldsymbol{\eta}^{(2)}\right)}{\partial\left(\boldsymbol{\theta}^{(1)},\boldsymbol{\eta}^{(1)}\right)}\right|$ = Jacobian of bijection

Acceptance probability for MH step from M_1 to M_2

$$\alpha(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}) = \min\{1, A\}, \text{ where}$$

$$A := \frac{p(2, \boldsymbol{\theta}^{(2)}|\mathbf{y})}{p(1, \boldsymbol{\theta}^{(1)}|\mathbf{y})} \cdot \frac{p_{2\to 1}}{p_{1\to 2}} \cdot \frac{\varphi_2(\boldsymbol{\eta}^{(2)})}{\varphi_1(\boldsymbol{\eta}^{(1)})} \cdot \left| \frac{\partial\left(\boldsymbol{\theta}^{(2)}, \boldsymbol{\eta}^{(2)}\right)}{\partial\left(\boldsymbol{\theta}^{(1)}, \boldsymbol{\eta}^{(1)}\right)} \right|$$

$p(k, \boldsymbol{\theta}^{(k)}|y)$ = joint posterior density of M_k and $\boldsymbol{\theta}^{(k)}$

$p_{i\to j}$ = prior switching probability from M_i to M_j

$\varphi_k(\boldsymbol{\eta}^{(k)})$ = proposal distribution of $\boldsymbol{\eta}^{(k)}$ (suitably chosen)

$\left| \frac{\partial\left(\boldsymbol{\theta}^{(2)}, \boldsymbol{\eta}^{(2)}\right)}{\partial\left(\boldsymbol{\theta}^{(1)}, \boldsymbol{\eta}^{(1)}\right)} \right|$ = Jacobian of bijection

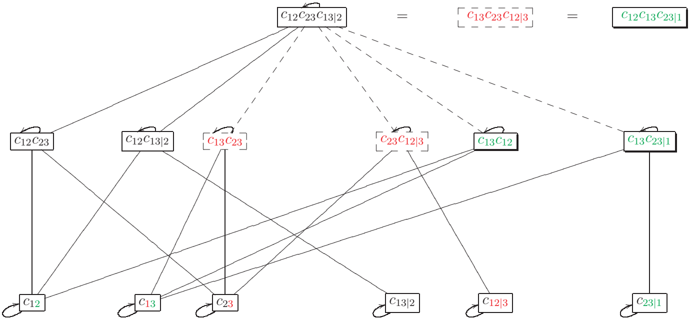
For moves from M_2 to M_1 we use A^{-1}.

RJ MCMC moves for selecting PCC's (n=3)

Notation	Meaning	Example
B0D0	stay move	$c_{12} \longrightarrow c_{12}$
B1D0	birth of 1 factor	$c_{12} \longrightarrow c_{12}\,c_{23}$
D1B0	death of 1 factor	$c_{12}\,c_{23} \longrightarrow c_{12}$
B2D1	birth of 2 factors death of 1 factor	$c_{13}\,c_{23} \longrightarrow c_{12}\,c_{23}\,c_{13\|2}$
D2B1	death of 2 factors birth of 1 factor	$c_{12}\,c_{23}\,c_{13\|2} \longrightarrow c_{13}\,c_{23}$
B3D2	birth of 3 factors death of 2 factors	$c_{13}\,c_{23\|1} \longrightarrow c_{12}\,c_{23}\,c_{13\|2}$
D3B2	death of 3 factors birth of 2 factors	$c_{12}\,c_{23}\,c_{13\|2} \longrightarrow c_{13}\,c_{23\|1}$

Graph of all PCC's with $n \leq 3$

Solid lines: B0D0, B1D0 and D1B0 moves
Dotted lines: B2D1, D2B1, B3D2 and D3B2 moves



Priors and proposal distributions

- Each model is equally likely a priori

Priors and proposal distributions

- Each model is equally likely a priori
- We assume equal prior switching probabilities for staying, moving to a smaller and moving to a larger model

Priors and proposal distributions

- Each model is equally likely a priori
- We assume equal prior switching probabilities for staying, moving to a smaller and moving to a larger model
- Independent normal distributions are assumed for each η component centered at corresponding previous parameter component with specified variance

Likelihood for PCC's with $n \leq 3$

Assume bivariate t-copula for each pair copula term, i.e.

$$
\begin{aligned}
\boldsymbol{\theta} &= (\rho, \nu) \quad \rho = \textbf{correlation} \ , \nu = \ \textbf{df} \\
c(u_1, u_2 | \boldsymbol{\theta}) &= \frac{\Gamma(\frac{\nu+2}{2})/\Gamma(\frac{\nu}{2})}{\nu\pi t_\nu(x_1) t_\nu(x_2)\sqrt{1-\rho^2}} \left(1 + \frac{x_1^2 + x_2^2 - 2\rho\, x_1\, x_2}{\nu(1-\rho^2)}\right)^{-\frac{\nu+1}{2}},
\end{aligned}
$$

where

$$
\begin{aligned}
x_1 &= t_\nu^{-1}(u_1) \\
x_2 &= t_\nu^{-1}(u_2) \\
t_\nu(\cdot) &= \text{pdf of univariate } t_\nu \text{ distribution} \\
t_\nu^{-1}(\cdot) &= \text{quantile function of } t_\nu \text{ distribution}
\end{aligned}
$$

Acceptance probability for B2D1 move

Actual state	$\theta^o := \theta^o_{\mathbf{m}^o}$	$\mathbf{m}^o = (m^o_1, m^o_2, m^o_3)$
Proposed state	$\theta^p := \theta^p_{\mathbf{m}^p}$	$\mathbf{m}^p = (m^p_1, m^p_2, m^p_3)$

Acceptance probability for B2D1 move

Actual state	$\theta^o := \theta^o_{\mathbf{m}^o}$	$\mathbf{m}^o = (m^o_1, m^o_2, m^o_3)$
Proposed state	$\theta^p := \theta^p_{\mathbf{m}^p}$	$\mathbf{m}^p = (m^p_1, m^p_2, m^p_3)$
birth indices	$\{m^p_s, m^p_t\} :=$	$\mathbf{m}^p \backslash \mathbf{m}^o$
death index	$m^o_v :=$	$\mathbf{m}^o \bigcap \mathbf{m}^p$
stay index	$m^o_w :=$	$\mathbf{m}^o \backslash \mathbf{m}^p$

Example: $(\theta^o_{13}, \theta^o_{23}) \rightarrow (\theta^p_{12}, \theta^p_{23}, \theta^p_{13|2})$

Acceptance probability for B2D1 move

Actual state	$\theta^o := \theta^o_{\mathbf{m}^o}$	$\mathbf{m}^o = (m^o_1, m^o_2, m^o_3)$
Proposed state	$\theta^p := \theta^p_{\mathbf{m}^p}$	$\mathbf{m}^p = (m^p_1, m^p_2, m^p_3)$
birth indices	$\{m^p_s, m^p_t\} :=$	$\mathbf{m}^p \backslash \mathbf{m}^o$
death index	$m^o_v :=$	$\mathbf{m}^o \bigcap \mathbf{m}^p$
stay index	$m^o_w :=$	$\mathbf{m}^o \backslash \mathbf{m}^p$

Example: $(\theta^o_{13}, \theta^o_{23}) \rightarrow (\theta^p_{12}, \theta^p_{23}, \theta^p_{13|2})$ **Bijection:**

$$\begin{pmatrix} \theta^p \\ \eta^p_{m^o_w} \end{pmatrix} := \begin{pmatrix} \theta^p_{m^p_s} \\ \theta^p_{m^p_t} \\ \theta^p_{m^o_v} \\ \eta^p_{m^o_w} \end{pmatrix} = \begin{pmatrix} \eta^o_{m^p_s} \\ \eta^o_{m^p_t} \\ \theta^o_{m^o_v} \\ \theta^o_{m^o_w} \end{pmatrix} =: \begin{pmatrix} \eta^o_{m^p_s} \\ \eta^o_{m^p_t} \\ \theta^o \end{pmatrix}$$

Acceptance probability for B2D1 move

Actual state	$\theta^o := \theta^o_{\mathbf{m}^o}$	$\mathbf{m}^o = (m_1^o, m_2^o, m_3^o)$
Proposed state	$\theta^p := \theta^p_{\mathbf{m}^p}$	$\mathbf{m}^p = (m_1^p, m_2^p, m_3^p)$
birth indices	$\{m_s^p, m_t^p\} :=$	$\mathbf{m}^p \backslash \mathbf{m}^o$
death index	$m_v^o :=$	$\mathbf{m}^o \bigcap \mathbf{m}^p$
stay index	$m_w^o :=$	$\mathbf{m}^o \backslash \mathbf{m}^p$

Example: $(\theta^o_{13}, \theta^o_{23}) \to (\theta^p_{12}, \theta^p_{23}, \theta^p_{13|2})$ **Bijection:**

$$\begin{pmatrix} \theta^p \\ \eta^p_{m_w^o} \end{pmatrix} := \begin{pmatrix} \theta^p_{m_s^p} \\ \theta^p_{m_t^p} \\ \theta^p_{m_v^o} \\ \eta^p_{m_w^o} \end{pmatrix} = \begin{pmatrix} \eta^o_{m_s^p} \\ \eta^o_{m_t^p} \\ \theta^o_{m_v^o} \\ \theta^o_{m_w^o} \end{pmatrix} =: \begin{pmatrix} \eta^o_{m_s^p} \\ \eta^o_{m_t^p} \\ \theta^o \end{pmatrix}$$

Here $\eta^o_{m_s^p}$ and $\eta^o_{m_t^p}$ are bivariate normal centered at $\theta_{m_s^p, last}$, and $\theta_{m_t^p, last}$ with specified covariance matrix, while we set $\eta^p_{m_w^o} = \theta^o_{m_w^o}$

First results for selecting between $c_{12}c_{23}c_{13|2}$ and $c_{12}c_{23}$

5000 data points generated from $c_{12}c_{23}c_{13|2}$ with every pair copula term a bivariate t-copula
RJ MCMC visited 4749 the full model $c_{12}c_{23}c_{13|2}$ and 251 the reduced model $c_{12}c_{23}$

First results for selecting between $c_{12}c_{23}c_{13|2}$ and $c_{12}c_{23}$

5000 data points generated from $c_{12}c_{23}c_{13|2}$ with every pair copula term a bivariate t-copula

RJ MCMC visited 4749 the full model $c_{12}c_{23}c_{13|2}$ and 251 the reduced model $c_{12}c_{23}$

	Full model				Reduced model			
Parameter	**mean**	**median**	**5%**	**95%**	**mean**	**median**	**5%**	**95%**
$\rho_{12} = .2$	0.19	0.19	0.17	0.21	0.19	0.19	0.16	0.22
$\nu_{12} = 5$	5.07	4.97	4.39	6.17	5.23	5.19	4.47	6.04
$\rho_{23} = 0.6$	0.60	0.60	0.58	0.61	0.60	0.60	0.59	0.62
$\nu_{23} = 7$	7.58	7.44	6.38	9.40	8.01	7.85	6.51	10.06
$\rho_{13\|2} = 0.2$	0.19	0.19	0.13	0.23				
$\nu_{13\|2} = 40$	13.35	13.20	11.88	15.62				

Conclusions and Outlook

- PCC's such as canonical and D-Vines allow for very flexible class of multivariate distributions.

Conclusions and Outlook

- PCC's such as canonical and D-Vines allow for very flexible class of multivariate distributions.
- Canonical and D-Vines can be simulated fast and the likelihood can be calculated fast as well.

Conclusions and Outlook

- PCC's such as canonical and D-Vines allow for very flexible class of multivariate distributions.
- Canonical and D-Vines can be simulated fast and the likelihood can be calculated fast as well.
- Computer intensive methods such as RJ MCMC algorithms are needed to estimate model type (full or reduced) and corresponding parameters.

Conclusions and Outlook

- PCC's such as canonical and D-Vines allow for very flexible class of multivariate distributions.
- Canonical and D-Vines can be simulated fast and the likelihood can be calculated fast as well.
- Computer intensive methods such as RJ MCMC algorithms are needed to estimate model type (full or reduced) and corresponding parameters.
- Incorporate choice of different copula types
- Efficient model selection criteria are needed to compare nested and non-nested models.

Conclusions and Outlook

- PCC's such as canonical and D-Vines allow for very flexible class of multivariate distributions.
- Canonical and D-Vines can be simulated fast and the likelihood can be calculated fast as well.
- Computer intensive methods such as RJ MCMC algorithms are needed to estimate model type (full or reduced) and corresponding parameters.
- Incorporate choice of different copula types
- Efficient model selection criteria are needed to compare nested and non-nested models.
- Time series structures for margins can be incorporated by first fitting marginal time series models and then considering standardized residuals. Joint estimation ?

Conclusions and Outlook

- PCC's such as canonical and D-Vines allow for very flexible class of multivariate distributions.
- Canonical and D-Vines can be simulated fast and the likelihood can be calculated fast as well.
- Computer intensive methods such as RJ MCMC algorithms are needed to estimate model type (full or reduced) and corresponding parameters.
- Incorporate choice of different copula types
- Efficient model selection criteria are needed to compare nested and non-nested models.
- Time series structures for margins can be incorporated by first fitting marginal time series models and then considering standardized residuals. Joint estimation ?

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2007).
Pair-copula constructions of multiple dependence.
Insurance, Mathematics and Economics, DOI 10.1016/j.insmatheco.2007.02.001, preprint available at SFB 386 Statistical Analysis of Diskrete Structures, www.stat.uni–muenchen.de/sfb386/.

Bedford, T. and R. M. Cooke (2001).
Probability density decomposition for conditionally dependent random variables modeled by vines.
Annals of Mathematics and Artificial Intelligence 32, 245–268.

Bedford, T. and R. M. Cooke (2002).
Vines - a new graphical model for dependent random variables.
Annals of Statistics 30(4), 1031–1068.

Green, P. (1995).
Reversible jump markov chain monte carlo computation and bayesian model determination.

Biometrika 82, 711–732.

Joe, H. (1996).
Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters.
In L. Rüschendorf and B. Schweizer and M. D. Taylor (Ed.), *Distributions with Fixed Marginals and Related Topics*.

Kurowicka, D. and R. Cooke (2006).
Uncertainty analysis with high dimensional dependence modelling.
Chichester: Wiley.

Kurowicka, D. and R. M. Cooke (2004).
Distribution - free continuous bayesian belief nets.
In *Fourth International Conference on Mathematical Methods in Reliability Methodology and Practice, Santa Fe, New Mexico*.