Introduction to vine copulas

Nicole Krämer & Ulf Schepsmeier
Technische Universität München
[kraemer, schepsmeier]@ma.tum.de

NIPS Workshop, Granada, December 18, 2011
1 Motivation and background

2 Pair-copula construction (PCC) of vine distribution

3 Model selection and estimation

4 Applications and extensions

5 Summary and Outlook
Motivation

- Copulas model marginal and common dependencies separately.
- There is a wide range of parametric copula families:

  - **Gauss**
  - **Frank**
  - **Clayton**

- **But:** Standard multivariate copulas
  - can become inflexible in high dimensions.
  - do not allow for different dependency structures between pairs of variables.

⇒ **Vine copulas** for higher-dimensional data.
Overview Vines

Vine pair-copulas

- **Bivariate copulas** are building blocks for higher-dimensional distributions.
- The dependency structure is determined by the bivariate copulas and a **nested set of trees**.

→ Vine approach is more flexible, as we can select bivariate copulas from a wide range of (parametric) families.

Model estimation

1. **graph theory** to determine the dependency structure of the data
2. **statistical inference** (maximum-likelihood, Bayesian approach ...) to fit bivariate copulas.
Background - Bivariate Copulas

Bivariate Copula

A bivariate copula function

\[ C : [0, 1]^2 \rightarrow \mathbb{R} \]

is a distribution on \([0, 1]^2\) with uniform marginals.

Let \( F \) be a bivariate distribution with marginal distributions \( F_1, F_2 \).

Sklar’s Theorem (1959)

There exists a two dimensional copula \( C(\cdot, \cdot) \), such that

\[ \forall (x_1, x_2) \in \mathbb{R}^2 : \quad F(x_1, x_2) = C(F_1(x_1), F_2(x_2)) . \]

If \( F_1 \) and \( F_2 \) are continuous, the copula \( C \) is unique.
Copula densities

Copula density (2-dimensional)

\[ c_{12}(u_1, u_2) = \frac{\partial^2 C_{12}(u_1, u_2)}{\partial u_1 \partial u_2} \]

This implies

- **joint density**

\[ f(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2) \]

- **conditional density**

\[ f(x_2|x_1) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_2(x_2) \]
Important: pair-copula constructions

We can represent a density \( f(x_1, \ldots, x_d) \) as a product of pair copula densities and marginal densities!

Example: \( d = 3 \) dimensions. One possible decomposition of \( f(x_1, x_2, x_3) \) is:

\[
f(x_1, x_2, x_3) = f_{3|12}(x_3|x_1, x_2)f_{2|1}(x_2|x_1)f_1(x_1)
\]

\[
f_{2|1}(x_2|x_1) = c_{12}(F_1(x_1), F_2(x_2))f_2(x_2)
\]

\[
f_{3|12}(x_3|x_1, x_2) = c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2))f_{3|2}(x_3|x_2)
\]

\[
f_{3|2}(x_3|x_2) = c_{23}(F_2(x_2), F_3(x_3))f_3(x_3)
\]

\[
f(x_1, x_2, x_3) = f_3(x_3)f_2(x_2)f_1(x_1) \text{ (marginals)} \times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \text{ (unconditional pairs)} \times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \text{ (conditional pair)}
\]
Important: pair-copula constructions

We can represent a density \( f(x_1, \ldots, x_d) \) as a product of pair copula densities and marginal densities!

Example: \( d = 3 \) dimensions. One possible decomposition of \( f(x_1, x_2, x_3) \) is:

\[
f(x_1, x_2, x_3) = f_{3|12}(x_3|x_1, x_2)f_{2|1}(x_2|x_1)f_1(x_1)
\]

\[
f_{2|1}(x_2|x_1) = c_{12}(F_1(x_1), F_2(x_2))f_2(x_2)
\]
\[
f_{3|12}(x_3|x_1, x_2) = c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2))f_{3|2}(x_3|x_2)
\]
\[
f_{3|2}(x_3|x_2) = c_{23}(F_2(x_2), F_3(x_3))f_3(x_3)
\]

\[
f(x_1, x_2, x_3) = f_3(x_3)f_2(x_2)f_1(x_1) \text{ (marginals)}
\times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \text{ (unconditional pairs)}
\times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \text{ (conditional pair)}
\]
Important: pair-copula constructions

We can represent a density \( f(x_1, \ldots, x_d) \) as a product of \textit{pair} copula densities and marginal densities!

Example: \( d = 3 \) dimensions. One possible decomposition of \( f(x_1, x_2, x_3) \) is:

\[
f(x_1, x_2, x_3) = f_{3|12}(x_3|x_1, x_2)f_{2|1}(x_2|x_1)f_1(x_1)
\]

\[
f_{2|1}(x_2|x_1) = c_{12}(F_1(x_1), F_2(x_2))f_2(x_2)
\]
\[
f_{3|12}(x_3|x_1, x_2) = c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2))f_{3|2}(x_3|x_2)
\]
\[
f_{3|2}(x_3|x_2) = c_{23}(F_2(x_2), F_3(x_3))f_3(x_3)
\]

\[
f(x_1, x_2, x_3) = f_3(x_3)f_2(x_2)f_1(x_1) \text{ (marginals)}
\times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \text{ (unconditional pairs)}
\times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \text{ (conditional pair)}
\]
Important: pair-copula constructions

We can represent a density \( f(x_1, \ldots, x_d) \) as a product of pair copula densities and marginal densities!

Example: \( d = 3 \) dimensions. One possible decomposition of \( f(x_1, x_2, x_3) \) is:

\[
f(x_1, x_2, x_3) = f_{3|12}(x_3|x_1, x_2)f_{2|1}(x_2|x_1)f_1(x_1)
\]

\[
f_{2|1}(x_2|x_1) = c_{12}(F_1(x_1), F_2(x_2))f_2(x_2)
\]

\[
f_{3|12}(x_3|x_1, x_2) = c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2))f_{3|2}(x_3|x_2)
\]

\[
f_{3|2}(x_3|x_2) = c_{23}(F_2(x_2), F_3(x_3))f_3(x_3)
\]

\[
f(x_1, x_2, x_3) = f_3(x_3)f_2(x_2)f_1(x_1) \text{ (marginals)}
\times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \text{ (unconditional pairs)}
\times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \text{ (conditional pair)}
Pair-copula construction (PCC) in $d$ dimensions


\[
f(x_1, \ldots, x_d) = \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,(i+j)|(i+1),\ldots,(i+j-1)} \cdot \prod_{k=1}^{d} f_k(x_k)
\]

pair copula densities
imperical densities

with

\[
c_{i,j|i_1,\ldots,i_k} := c_{i,j|i_1,\ldots,i_k}(F(x_i|x_{i_1}, \ldots, x_{i_k}), (F(x_j|x_{i_1}, \ldots, x_{i_k}))
\]

for $i, j, i_1, \ldots, i_k$ with $i < j$ and $i_1 < \cdots < i_k$.

Remarks:

- The decomposition is not unique.
- Bedford and Cooke (2001) introduced a graphical structure called regular vine structure to help organize them.
Important: regular vine structure

Example: $d = 3$ dimensions

\[
  f(x_1, x_2, x_3) = f_3(x_3)f_2(x_2)f_1(x_1) \quad \text{(marginals)}
\]

\[
  \times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \quad \text{(unconditional pairs)}
\]

\[
  \times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \quad \text{(conditional pair)}
\]
Important: regular vine structure

Example: $d = 3$ dimensions

\[
f(x_1, x_2, x_3) = f_3(x_3)f_2(x_2)f_1(x_1) \text{ (marginals)} \\
\times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \text{ (unconditional pairs)} \\
\times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \text{ (conditional pair)}
\]
Important: regular vine structure

Example: \( d = 3 \) dimensions

\[
f(x_1, x_2, x_3) = f_3(x_3)f_2(x_2)f_1(x_1) \quad \text{(marginals)}
\]
\[
\times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \quad \text{(unconditional pairs)}
\]
\[
\times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \quad \text{(conditional pair)}
\]
Important: regular vine structure

Example: $d = 3$ dimensions

$$f(x_1, x_2, x_3) = f_3(x_3)f_2(x_2)f_1(x_1) \quad \text{(marginals)}$$
$$\times c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \quad \text{(unconditional pairs)}$$
$$\times c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \quad \text{(conditional pair)}$$
Important: regular vine structure

Example: $d = 3$ dimensions

$$f(x_1, x_2, x_3) = f_3(x_3)f_2(x_2)f_1(x_1) \ (\text{marginals})$$
$$\times \ c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \ (\text{unconditional pairs})$$
$$\times \ c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \ (\text{conditional pair})$$
R-vine structure \((d = 5)\)

Pair-copulas:

1. \(C_{12}, C_{13}, C_{34}, C_{34}, C_{15}\)
2. proximity condition If two nodes are joined by an edge in tree \(j + 1\), the corresponding edges in tree \(j\) share a node.
3. \(C_{23|1}, C_{14|3}, C_{35|1}\)
4. \(C_{24|13}, C_{45|13}\)
5. \(C_{25|134}\)
R-vine structure \((d = 5)\)

Pair-copulas:

1. \(c_{12}, c_{13}, c_{34}, c_{34}, c_{15}\)
2. proximity condition If two nodes are joined by an edge in tree \(j + 1\), the corresponding edges in tree \(j\) share a node.
3. \(c_{23|1}, c_{14|3}, c_{35|1}\)
4. \(c_{24|13}, c_{45|13}\)
5. \(c_{25|134}\)
R-vine structure \((d = 5)\)

Pair-copulas:

1. \(c_{12}, c_{13}, c_{34}, c_{34}, c_{15}\)
2. Proximity condition: If two nodes are joined by an edge in tree \(j + 1\), the corresponding edges in tree \(j\) share a node.
3. \(c_{23|1}, c_{14|3}, c_{35|1}\)
4. \(c_{24|13}, c_{45|13}\)
5. \(c_{25|134}\)
R-vine structure \((d = 5)\) formal definition

Pair-copulas:

1. \(c_{12}, c_{13}, c_{34}, c_{34}, c_{15}\)

2. **proximity condition** If two nodes are joined by an edge in tree \(j + 1\), the corresponding edges in tree \(j\) share a node.

3. \(c_{23|1}, c_{14|3}, c_{35|1}\)

4. \(c_{24|13}, c_{45|13}\)

5. \(c_{25|134}\)
R-vine structure \((d = 5)\) formal definition

Pair-copulas:

1. \(C_{12}, C_{13}, C_{34}, C_{34}, C_{15}\)
2. **proximity condition** If two nodes are joined by an edge in tree \(j + 1\), the corresponding edges in tree \(j\) share a node.
3. \(C_{23|1}, C_{14|3}, C_{35|1}\)
4. \(C_{24|13}, C_{45|13}\)
5. \(C_{25|134}\)
R-vine structure \((d = 5)\) ▶ formal definition

Pair-copulas:

1. \(c_{12}, c_{13}, c_{34}, c_{34}, c_{15}\)

2. **proximity condition** If two nodes are joined by an edge in tree \(j + 1\), the corresponding edges in tree \(j\) share a node.

3. \(c_{23|1}, c_{14|3}, c_{35|1}\)

4. \(c_{24|13}, c_{45|13}\)

5. \(c_{25|134}\)
C-anonical vines
Each tree has a unique node that is connected to all other nodes.

\[ f_{1234} = f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdot c_{12} \cdot c_{13} \cdot c_{14} \cdot c_{23\mid1} \cdot c_{24\mid1} \cdot c_{34\mid12} \]

- Edges in \( T_1 \): nodes in \( T_1 \)
- Edges in \( T_2 \): nodes in \( T_2 \)
- Edge in \( T_3 \): nodes in \( T_3 \)
D-vines

Each tree is a path.

\[ f_{1234} = f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdot c_{12} \cdot c_{23} \cdot c_{34} \cdot c_{13|2} \cdot c_{24|3} \cdot c_{14|23} \]

- nodes in \( T_1 \)
- edges in \( T_1 \)
- nodes in \( T_2 \)
- edges in \( T_2 \)
- nodes in \( T_3 \)
- edge in \( T_3 \)
Preliminary summary: pair-copula decomposition

So far

Given a $d$-dimensional density, we can

- decompose it into products of marginal densities and bivariate copula densities.
- represent this decomposition with nested set of trees that fulfill a proximity condition.

Question

Given data, how can we estimate a pair-copula decomposition?
Preliminary summary: pair-copula decomposition

So far

Given a $d$-dimensional density, we can
- decompose it into products of marginal densities and bivariate copula densities.
- represent this decomposition with nested set of trees that fulfill a proximity condition.

Question

Given data, how can we estimate a pair-copula decomposition?
Model selection and parameter estimation

Model = structure (trees) + copula families + copula parameters

Use our software package CDVine!
(Brechmann and Schepsmeier (2011))
Model selection and parameter estimation

Model = structure (trees) + copula families + copula parameters

Data

Use our software package **CDVine**!
(Brechmann and Schepsmeier (2011))
Model selection and parameter estimation

Model = \textit{structure \ (trees)} + copula families + copula parameters

Use our software package \textbf{CDVine}!
(Brechmann and Schepsmeier (2011))
Model selection and parameter estimation

Model = structure (trees) + copula families + copula parameters

Use our software package **CDVine**!
(Brechmann and Schepsmeier (2011))
Model selection and parameter estimation

Model = structure (trees) + copula families + copula parameters

Use our software package **CDVine**! (Brechmann and Schepsmeier (2011))
Model selection and parameter estimation

Model = structure (trees) + copula families + copula parameters

Normal, $\rho = 0.5$

Clayton, $\theta = 2.5$

Gumbel, $\theta = 1.7$

Use our software package CDVine!
(Brechmann and Schepsmeier (2011))
Model = structure (trees) + copula families + copula parameters

Problem:
- Huge number of possible vines  $\rightarrow$ structure selection
- $\frac{d(d-1)}{2}$ pair-copulas  $\rightarrow$ copula selection
  $\rightarrow$ parameter estimation

Use our software package **CDVine**!
(Brechmann and Schepsmeier (2011))
Structure selection

Possible edge weights

- Kendall’s $\tau$
- Spearman’s $\rho$
- $p$-values of Goodness-of-Fit tests
- distances

Model selection

is done tree by tree via

- optimal C-vines structure selection (Czado et al. (2011))
- Traveling Salesman Problem for D-vines
- Maximum Spanning Tree for R-vines (Dissmann et al. (2011))
- Bayesian approaches (Reversible Jump MCMC)
Structure selection

Possible edge weights

- Kendall’s $\tau$
- Spearman’s $\rho$
- p-values of Goodness-of-Fit tests
- distances

Model selection

is done tree by tree via

- optimal C-vines structure selection (Czado et al. (2011))
- Traveling Salesman Problem for D-vines
- Maximum Spanning Tree for R-vines (Dissmann et al. (2011))
- Bayesian approaches (Reversible Jump MCMC)
Copula selection

Copula selection can be done via
- Goodness-of-fit tests
- Independence test
- AIC/BIC
- Graphical tools like contour plots, $\lambda$-function, ...

Possible copula families
- Elliptical copulas (Gauss, t-)
- One-parametric Archimedean copulas (Clayton, Gumbel, Frank, Joe,...)
- Two-parametric Archimedean copulas (BB1, BB7,...)
- Rotated versions of the Archimedean for neg. dependencies
- ...

Krämer & Schepsmeier (TUM)
Parameter estimation

Estimation approaches:

- **Maximum likelihood estimation**
- **Sequential estimation:**
  - Parameters are estimated sequentially starting from the top tree.
  - Parameter estimates can be used to define pseudo observations for the next tree.
  - Parameter estimation via $\theta = f(\tau)$ or bivariate MLE.
  - Sequential estimates can be used as starting values for maximum likelihood.
- **Bayesian estimation**
Parameter estimation

Estimation approaches:

- **Maximum likelihood estimation**
- **Sequential estimation:**
  - Parameters are estimated sequentially starting from the top tree.
  - Parameter estimates can be used to define pseudo observations for the next tree.
  - Parameter estimation via $\theta = f(\tau)$ or bivariate MLE
  - Sequential estimates can be used as **starting values** for maximum likelihood
- **Bayesian estimation**
Parameter estimation

Estimation approaches:

- **Maximum likelihood estimation**
- **Sequential estimation**: Parameters are estimated sequentially starting from the top tree. Parameter estimates can be used to define pseudo observations for the next tree. Parameter estimation via $\theta = f(\tau)$ or bivariate MLE. Sequential estimates can be used as starting values for maximum likelihood.
- **Bayesian estimation**
Applications

Dimensionality of applications

- **Gaussian** vines in arbitrary dimensions (Kurowicka and Cooke 2006)
- First non Gaussian D-vine models using joint maximum likelihood in 4 dimensions
- Bayesian D-vines with credible intervals in 7 and 12 dimensions
- Joint maximum likelihood now feasible in 50 dimensions for R-vines
- Sequential estimation of R-vines in 100 dimensions
- Sequential estimation for \( d \gg 100 \) dimensions with truncation (i.e. higher order trees only contain independent copulas)
- Heinen and Valdesogo (2009) sequentially fit a C-vine autoregressive model in 100 dimensions

Application areas:

- finance
- insurance
- genetics
- health
- images
- ...
Extensions (Projects of our research group)

**Special vine models:**
- vine copulas with *time varying* parameters
- regime switching vine models
- non parametric vine pair copulas
- Non Gaussian directed acyclic graphical (DAG) models based on PCC’s
- discrete vine copulas
- truncated and simplified R-vines
- spatial vines
- copula discriminant analysis
Summary and outlook

- PCC’s such as C-, D- and R-vines allow for very flexible class of multivariate distributions
- Efficient parameter estimation methods are available for dimensions up to 50
- Model selection of vine tree structures and pair copula types for regular vines still needs further work
- Efficient distance measures between vine distributions would be useful


Reading material, software and current projects: [http://www-m4.ma.tum.de/en/research/vine-copula-models](http://www-m4.ma.tum.de/en/research/vine-copula-models)
An $d$-dimensional regular vine is a sequence of $d$-1 trees

1. **tree 1**
   - $d$ nodes: $X_1, \ldots, X_d$
   - $d - 1$ edges: pair-copula densities between nodes $X_1, \ldots, X_d$

2. **tree $j$**
   - $d + 1 - j$ nodes: edges of tree $j - 1$
   - $d - j$ edges: conditional pair-copula densities

**Proximity condition:** If two nodes in tree $j + 1$ are joined by an edge, the corresponding edges in tree $j$ share a node.