Bayesian Inference for Pair-copula Constructions of Multiple Dependence

Claudia Czado and Aleksey Min
Technische Universität München
cczado@ma.tum.de, aleksmin@ma.tum.de

December 7, 2007
## Overview

1. Introduction
2. Pair-copula constructions (PCC)
3. Bayesian Analysis of PCC’s
5. Bayesian Model Selection among PCC’s
6. Euro Swap Rates
7. Summary and Outlook

Extends work by Joe (1996).

Aas, Czado, Frigessi, and Bakken (2007) used the PCC construction to construct flexible multivariate copulas based on pair-copulas such as bivariate Gaussian, t-, Gumbel and Clayton copulas.
Parameters are estimated by maximum likelihood (ML) and PCC’s are sucessfully applied to financial return data.

Berg and Aas (2007) give further example of PCC’s and developed an R-package for fitting PCC’s using MLE.

The Fisher information is difficult to handle analytically, so therefore estimated standard errors for parameter estimates are so far absent.

Here we follow a Bayesian approach where such interval estimates are easy to establish.
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

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

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

for some $n$-variante copula density $c_{12\ldots n}(\cdot)$. 
Pair-copula constructions (PCC)

- **n = 3**
  \[ f_{1|23}(x_1|x_2, x_3) = \frac{f_{12|3}(x_1, x_2|x_3)}{f_{2|3}(x_2|x_3)} = c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)) \cdot f_{1|3}(x_1|x_3) \]

- **general n** For \( \mathbf{v} = (v_1, \ldots, v_d) \) and any \( j = 1, \ldots, d \)
  \[ f(x|\mathbf{v}) = c_{xv_j|v_{-j}}(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j})) \cdot f(x|\mathbf{v}_{-j}) \]
  \[ \mathbf{v}_{-j} = (v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_d) \]
  \[ c_{xv_j|v_{-j}}(\cdot) = \text{bivariate copula density} \]

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

We call this a pair-copula construction (PCC).
Conditional cdf’s

**Univariate** $\nu$: Since $f(x|\nu) = c_{x\nu}(F_x(x), F_\nu(\nu))f_x(x)$ we have

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

$$= \frac{\partial C_{x\nu}(F_x(x), F_\nu(\nu))}{\partial F_\nu(\nu)}$$

**General** $\nu$: Under regularity conditions Joe (1996) showed that

$$F(x|\nu) = \frac{\partial C_{x,v_j|v_{-j}}(F(x|v_{-j}), F(v_j|v_{-j}))}{\partial F(v_j|v_{-j})}. $$
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
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.

- **D-vine:**
  A regular vine for which no node in any tree is connected to more than two edges

**Abbreviations:**

\[
\begin{align*}
c_{ij|v} & := c_{ij|v}(F_{i|v}(x_i|x_v), F_{j|v}(x_j|x_v)) \\
f_j & := f_j(x_j) \\
f_v & := f_v(x_v)
\end{align*}
\]
Pair-copula constructions (PCC)

Five dimensional canonical vine

\[ f_{12345} = f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdot f_5 \cdot c_{12} \cdot c_{13} \cdot c_{14} \cdot c_{15} \cdot c_{23|1} \cdot c_{24|1} \cdot c_{25|1} \cdot c_{34|12} \cdot c_{35|12} \cdot c_{45|123} \]
Pair-copula constructions (PCC)

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 

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

- ** D-vine density 

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

where index \( j \) identifies the trees, while \( i \) runs over the edges in each tree.
Assume bivariate t-copula for each pair copula term, i.e.

\[ \theta = (\rho, \nu) \quad \rho = \text{correlation} \ , \nu = \text{df} \]

\[ c(u_1, u_2 | \theta) = \frac{\Gamma(\frac{\nu+2}{2})/\Gamma(\frac{\nu}{2})}{\nu \pi t_\nu(x_1) t_\nu(x_2) \sqrt{1 - \rho^2}} \left(1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu (1 - \rho^2)}\right)^{-\frac{\nu+2}{2}}, \]

where

\[ x_1 = t_\nu^{-1}(u_1) \]
\[ x_2 = t_\nu^{-1}(u_2) \]
\[ t_\nu(\cdot) = \text{pdf of univariate } t_\nu \text{ distribution} \]
\[ t_\nu^{-1}(\cdot) = \text{quantile function of } t_\nu \text{ distribution} \]
Bayesian Analysis of PCC’s with t-copula pairs

- For the PCC’s we assume a regular vine structure with edge indices denoted by $s := ij|v$ for $s \in S$.
- Let $\theta_s = (\rho_s, \nu_s)$ the corresponding parameters of the pair copula associated with edge $s$.
- Let $\theta = \{\theta_s, s \in S\}$ denote the vector of all parameters of the PCC with t-copula pairs.
- We assume uniform priors for $\rho_s$ on $(-1, 1)$ and for $\nu_s$ on $(1, U)$ for each $s$ and prior independence among all parameters.
- Inference is based on the posterior distribution given by

$$p(\theta|data) = \frac{f(data|\theta) \times p(\theta)}{\int f(data|\theta) \times p(\theta)d\theta}$$

where $f(data|\theta)$ is the likelihood of the PCC and $p(\theta)$ the prior for $\theta$. 

Since $\theta$ is high dimensional the posterior is not analytically tractable.

Posterior distribution is estimated using Markov Chain Monte Carlo (MCMC) methods.

For each univariate parameter Metropolis-Hastings (MH) updates with symmetric normal random walk proposals are used.

Proposal variances are determined by pilot runs to achieve acceptance rates between 20% and 80%.

Posterior mean, mode, median and density are estimated using the MCMC iterates.

Credible intervals are estimated by empirical quantiles of the MCMC iterates.
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

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

\[
x_{i,t} = c_i + \alpha_i x_{i,t-1} + \sigma_{i,t} z_{i,t},
\]

\[
E[z_{t,i}] = 0 \text{ and } \text{Var}[z_{t,i}] = 1,
\]

\[
\sigma_{i,t}^2 = a_{i,0} + a_i \epsilon_{i,t-1}^2 + b_i \sigma_{i,t-1}^2 \text{ where } \epsilon_{i,t-1} = \sigma_{i,t} z_{i,t}
\]

are found to be independent and are transformed to uniform margins
**Data set**

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

<table>
<thead>
<tr>
<th>Between</th>
<th>M</th>
<th>T</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>4.21</td>
<td>34.16</td>
<td>14.47</td>
</tr>
<tr>
<td>M</td>
<td>16.65</td>
<td></td>
<td>15.48</td>
</tr>
<tr>
<td>T</td>
<td></td>
<td>12.60</td>
<td></td>
</tr>
</tbody>
</table>
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.
Bayesian Inference of the Norwegian return data

For each pair copula we assume a bivariate t-copula with correlation $\rho$ and $\nu$ degree of freedom.

For each $\rho$ we assume uniform(-1,1) prior and for $\nu$ a uniform(1,1000) prior. Further prior independence between all parameters are assumed.

For the likelihood we assume the following PCC

$$c(x_S, x_M, x_T, x_B) = c_{SM}(F(x_S), F(x_M)) \cdot c_{MT}(F(x_M), F(x_T))$$

$$\cdot c_{TB}(F(x_T), F(x_B)) \cdot c_{ST|M}(F(x_S|x_M), F(x_T|x_M))$$

$$\cdot c_{MB|T}(F(x_M|x_T), F(x_B|x_T))$$

$$\cdot c_{SB|MT}(F(x_S|x_M, x_T), F(x_B|x_M, x_T))$$
Bayesian Inference of the Norwegian return data

- For the **MH updates** in the MCMC algorithm we used
  - symmetric random walk proposals
  - normal proposals with variances determined by pilot runs to achieve acceptance rates between 25% and 80%
- **10000** MCMC iterations were run
- MLE estimates were used as **starting values**
- Implemented using **Daniel Berg’s R-package**
Trace Plots of MCMC iterations

- $\rho_{SM}$
- $\nu_{SM}$
- $\rho_{MT}$
- $\nu_{MT}$
- $\rho_{TB}$
- $\nu_{TB}$
- $\rho_{ST|M}$
- $\nu_{ST|M}$
- $\rho_{MB|T}$
- $\nu_{MB|T}$
- $\rho_{SB|MT}$
- $\nu_{SB|MT}$
Autocorrelations of the MCMC iterations

For the further analysis we used burnin 1000 iteration and only every 20th iteration.
Bayesian Inference for Pair-copula Constructions of Multiple Dependence
Application: Financial Returns

Estimated Posterior Densities

- $\rho_{SM}$
- $\nu_{SM}$
- $\rho_{MT}$
- $\nu_{MT}$
- $\rho_{TB}$
- $\nu_{TB}$
- $\rho_{ST|M}$
- $\nu_{ST|M}$
- $\rho_{MB|T}$
- $\nu_{MB|T}$
- $\rho_{SB|MT}$
- $\nu_{SB|MT}$
Summary statistics for thinned MCMC

<table>
<thead>
<tr>
<th></th>
<th>2.5% Quantile</th>
<th>50% Quantile</th>
<th>97.5% Quantile</th>
<th>Est. Post. Mean</th>
<th>Est. Post. Mode</th>
<th>MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{SM}$</td>
<td>-0.316</td>
<td>-0.254</td>
<td>-0.184</td>
<td>-0.253</td>
<td>-0.25</td>
<td>-0.25</td>
</tr>
<tr>
<td>$\nu_{SM}$</td>
<td>3.483</td>
<td>4.658</td>
<td>7.263</td>
<td>4.849</td>
<td>4.40</td>
<td>4.34</td>
</tr>
<tr>
<td>$\rho_{MT}$</td>
<td>0.422</td>
<td>0.466</td>
<td>0.508</td>
<td>0.465</td>
<td>0.47</td>
<td>0.47</td>
</tr>
<tr>
<td>$\nu_{MT}$</td>
<td>13.304</td>
<td>228.006</td>
<td>948.340</td>
<td>326.388</td>
<td>108.61</td>
<td>16.26</td>
</tr>
<tr>
<td>$\rho_{TB}$</td>
<td>-0.224</td>
<td>-0.170</td>
<td>-0.106</td>
<td>-0.168</td>
<td>-0.17</td>
<td>-0.17</td>
</tr>
<tr>
<td>$\nu_{TB}$</td>
<td>12.829</td>
<td>321.556</td>
<td>973.567</td>
<td>383.984</td>
<td>148.08</td>
<td>13.17</td>
</tr>
<tr>
<td>$\rho_{ST</td>
<td>M}$</td>
<td>-0.163</td>
<td>-0.104</td>
<td>-0.047</td>
<td>-0.103</td>
<td>-0.11</td>
</tr>
<tr>
<td>$\nu_{ST</td>
<td>M}$</td>
<td>106.098</td>
<td>560.460</td>
<td>964.869</td>
<td>550.005</td>
<td>672.33</td>
</tr>
<tr>
<td>$\rho_{MB</td>
<td>T}$</td>
<td>-0.033</td>
<td>0.031</td>
<td>0.090</td>
<td>0.029</td>
<td>0.03</td>
</tr>
<tr>
<td>$\nu_{MB</td>
<td>T}$</td>
<td>45.736</td>
<td>514.904</td>
<td>972.857</td>
<td>510.722</td>
<td>632.56</td>
</tr>
<tr>
<td>$\rho_{SB</td>
<td>MT}$</td>
<td>0.226</td>
<td>0.281</td>
<td>0.337</td>
<td>0.282</td>
<td>0.28</td>
</tr>
<tr>
<td>$\nu_{SB</td>
<td>MT}$</td>
<td>13.557</td>
<td>285.269</td>
<td>958.333</td>
<td>366.153</td>
<td>127.36</td>
</tr>
</tbody>
</table>

red corresponds to uncorrelatedness
magenta corresponds to near normality
Kendall's $\tau$ measures dependence and is defined by

$$\tau := P((X_1 - X_2)(Y_1 - Y_2) > 0) - P((X_1 - X_2)(Y_1 - Y_2) < 0)$$

where $(X_1, Y_1)$ and $(X_2, Y_2)$ are i.i.d.

- $\tau$ is invariant under monotone transformations
- Relationship between $\tau$ and $\rho$ for bivariate t-distributions

$$\tau = \frac{2}{\pi} \arcsin(\rho)$$

Estimated posterior summaries of $\tau$:

<table>
<thead>
<tr>
<th></th>
<th>2.5% Quantile</th>
<th>50% Quantile</th>
<th>97.5% Quantile</th>
<th>Est. Post. Mean</th>
<th>Est. Post Mode</th>
<th>Empirical $\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{SM}$</td>
<td>-0.204</td>
<td>-0.164</td>
<td>-0.1179</td>
<td>-0.163</td>
<td>-0.164</td>
<td>-0.158</td>
</tr>
<tr>
<td>$\tau_{MT}$</td>
<td>0.277</td>
<td>0.308</td>
<td>0.3393</td>
<td>0.308</td>
<td>0.309</td>
<td>0.313</td>
</tr>
<tr>
<td>$\tau_{TB}$</td>
<td>-0.144</td>
<td>-0.109</td>
<td>-0.0677</td>
<td>-0.108</td>
<td>-0.110</td>
<td>-0.110</td>
</tr>
</tbody>
</table>
Model fit: Tail dependence

- Lower and upper tail dependence:

\[
\lambda_l := \lim_{\nu \to 0} P(X \leq F_X^{-1}(\nu) | Y \leq F_Y^{-1}(\nu))
\]

\[
\lambda_u := \lim_{\nu \to 1} P(X \geq F_X^{-1}(\nu) | Y \geq F_Y^{-1}(\nu))
\]

- For t distributions we have

\[
\lambda := \lambda_l = \lambda_u = 2 t_{\nu+1} \left( -\sqrt{\nu + 1} \sqrt{\frac{1 - \rho}{1 + \rho}} \right)
\]

- Estimated posterior summaries of \( \lambda \)

<table>
<thead>
<tr>
<th></th>
<th>2.5% Quantile</th>
<th>50% Quantile</th>
<th>97.5% Quantile</th>
<th>Est. Post. Mean</th>
<th>Est. Post. Mode</th>
<th>Empirical ( \lambda_l )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_{SM} )</td>
<td>0.0048</td>
<td>0.02384</td>
<td>0.04656</td>
<td>0.02426</td>
<td>0.02312</td>
<td>0.02941</td>
</tr>
<tr>
<td>( \lambda_{MT} )</td>
<td>0.0000</td>
<td>0.00000</td>
<td>0.04378</td>
<td>0.00397</td>
<td>0.00000</td>
<td>0.02778</td>
</tr>
<tr>
<td>( \lambda_{TB} )</td>
<td>0.0000</td>
<td>0.00000</td>
<td>0.00051</td>
<td>0.00009</td>
<td>0.00000</td>
<td>0.12500</td>
</tr>
</tbody>
</table>
The λ function is defined as a *shifted copula distribution function* given by

\[
\lambda(z, \theta) := z - K(z, \theta)
\]

\[
K(z, \theta) := P(C(u_1, u_2, \theta) \leq z)
\]

- Here \( \theta \) denotes the parameters of the bivariate copula pair.
- For the conditional pairs \( ST|M, MB|T \) and \( SB|MT \), an empirical estimate of \( \lambda(z, \theta) \) is based on transformed data using posterior mode estimates.
Estimated pointwise posterior mode $\lambda(z, \theta)$

- **Pair ST**
  - Red = 95% CI, Black = est. posterior mode, Green = empirical

- **Pair MT**

- **Pair TB**

- **Pair ST|M**

- **Pair MB|T**

- **Pair SB|MT**
Want to select among

- $K$ models $M_1, \cdots, M_K$,

where Model $M_k$ has parameter $\theta_k$

and compare them on the basis of posterior model probabilities given by

$$P(\text{Model } M_k|\text{data}), k = 1, \cdots, K$$

We distinguish two situations

- **$K$ is small** and it is feasible timewise to fit all models (Congdon (2006) and Scott (2002))

- **$K$ is large** and we want only to fit models which are probable using reversible jump MCMC (RJMCMC) (Green (1995))
Bayesian Model Selection : K small

Congdon (2006) made the following assumptions

- The distribution of the data is independent of \( \{\theta_j \neq k\} \) given \( M_k \)
- Independence among \( \theta_k \)’s given Model M

and showed that

- Posterior distributions of \( \theta_k \) are independent given \( M = M_k \) and can be sampled individually

and used that

\[
P(M = M_k | \text{data, } \theta) \propto P(\text{data} | \theta_k, M = M_k) P(\theta | M = M_k) P(M = M_k) \tag{3}
\]
Bayesian Model Selection: K small

Assume that $K$ independently MCMC runs result in

$$M_1 : \theta_1^{(r)}, \ r = 1, \cdots, R \quad p(\theta_1|\text{data})$$

$$\vdots \quad \text{approximating} \quad \vdots$$

$$M_K : \theta_K^{(r)}, \ r = 1, \cdots, R \quad p(\theta_K|\text{data})$$

Using $\{\theta^{(r)} := (\theta_1^{(r)}, \cdots, \theta_K^{(r)}), \ r = 1, \cdots, R\}$

$$P(M|\text{data}) = \int P(M|\theta, \text{data})p(\theta|\text{data})d\theta$$

is approximated by $P(\widehat{M}|\text{data}) := \frac{1}{R} \sum_{r=1}^{R} P(M|\theta^{(r)}, \text{data})$. 
Bayesian Model Selection : K small

Using (3) we can estimate \( P(M = M_k | \text{data, } \theta^{(r)}) \) by

\[
w_k^{(r)} := \frac{G_k^{(r)}}{\sum_{j=1}^{K} G_j^{(r)}}, \text{ where }
\]

\[
G_k^{(r)} := \exp(L_k^{(r)} - L_{\text{max}}^{(r)})
\]

\[
L_k^{(r)} := \log \left( P(\text{data} | \theta_k^{(r)}, M = M_k) P(\theta_k^{(r)} | M = M_k) P(M = M_k) \right)
\]

\[
L_{\text{max}}^{(r)} := \max_{k=1,\ldots,K} L_k^{(r)}
\]

Therefore

\[
\frac{1}{R} \sum_{r=1}^{R} w_k^{(r)} \text{ estimates } P(M = M_k | \text{data}).
\]
Since zero in 95% credible interval for $\rho_{MB|T}$ and the posterior mode of $\rho_{ST|M}$ and $\rho_{MB|T}$ small we want to use Congdon (2006)’s method if these copula pairs needed

| Model | PCC | Formula | $P(M_k|\text{data})$ |
|-------|-----|---------|---------------------|
| $M_1$: | with all pairs | $c_{SM}c_{MT}c_{TB}c_{ST|M}c_{MB|T}c_{SB|MT}$ | 0.00071 |
| $M_2$: | without $c_{MB|T}$ | $c_{SM}c_{MT}c_{TB}c_{ST|M}c_{SB|MT}$ | 0.25900 |
| $M_3$: | without $c_{MB|T}$ and $c_{ST|M}$ | $c_{SM}c_{MT}c_{TB}c_{SB|MT}$ | 0.68807 |
| $M_4$: | without $c_{MB|T},c_{ST|M}$ and $c_{TB}$ | $c_{SM}c_{MT}c_{SB|MT}$ | 0.05220 |
Many of the posterior estimates of the degree of freedom are large, so want to check if we can use for these pairs a Gaussian copula.

| Model | PCC | $P(M_k | \text{data})$ |
|-------|-----|-------------------------|
| $M_1$: | All pairs are $t-$copulas | $1.74 \cdot 10^{-13}$ |
| $M_2$: | First pair $c_{SM}$ is $t-$copula, all other pairs are Gaussian copulas | $0.99999924$ |
| $M_3$: | All pairs are Gaussian copulas | $7.62 \cdot 10^{-6}$ |

Only for $c_{SM}$ a $t$-copula is needed.
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.

Identifiability: Use only one decomposition

\[ n = 3: \]

| Full PCC’s:          | \( c_{12}c_{23}c_{13|2} \) |
|----------------------|-------------------------------|
| Reduced by           |                               |
| 1 pair-copula:       | \( c_{12}c_{23}, c_{12}c_{13|2}, c_{23}c_{13|2} \) |
| Reduced by           |                               |
| 2 pair copulas:      | \( c_{12}, c_{23}, c_{13|2} \) |
Bayesian framework of PCC’s with $n = 3$

- Each of the 7 different PCC’s with $n \leq 3$ is identified by a model index $\mathbf{m} = (m_1, m_2, m_3) = (i_1j_1, i_2j_2, i_3j_3|k_3)$.  
  Example: $c_{12} c_{23|1}$ corresponds to $\mathbf{m} = (12, 00, 23|1)$.

- Each PCC model $\mathbf{m}$ has parameter vector $\theta_\mathbf{m} = (\theta_{m_1}, \theta_{m_2}, \theta_{m_3})$.  
  Example: Model $\mathbf{m} = (12, 23, 00|0)$ has parameter vector $\theta_\mathbf{m} = (\theta_{12}, \theta_{23}, \theta_{00|0}) = (\theta_{12}, \theta_{23})$.

- **Goal** is to estimate the best fitting model $\mathbf{m}$ and the corresponding parameter vector $\theta_\mathbf{m}$ using a Bayesian approach, i.e. the model $\mathbf{m}$ and $\theta_\mathbf{m}$ are considered random quantities.

- Inference about $\mathbf{m}$ and $\theta_\mathbf{m}$ is done via the joint posterior distribution of $(\mathbf{m}, \theta_\mathbf{m})$. 
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 $\theta_m$.

Requirement: Need to accommodate varying model dimension, i.e. construct MCMC iterates $\theta_{\mathbf{m}^r} = (\theta_{m_1^r}, \theta_{m_2^r}, \theta_{m_3^r})$ 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).
Bayesian Inference for Pair-copula Constructions of Multiple Dependence
Bayesian Model Selection among PCC's

General RJ MCMC (2)

Model
- $M_1$
- $M_2$

Parameter
- $\theta^{(1)} \in \mathbb{R}^{d_1}$
- $\theta^{(2)} \in \mathbb{R}^{d_2}$
- $d_1 < d_2$

Proposal
- $\eta^{(1)} \sim \varphi_1(\cdot)$
- $\eta^{(2)} \sim \varphi_2(\cdot)$
- $\eta^{(1)} \in \mathbb{R}^{a_1}$
- $\eta^{(2)} \in \mathbb{R}^{a_2}$

Dimension matching
- $d_1 + a_1 = d_2 + a_2$

Bijection
- $\left( \begin{array}{c} \theta^{(1)} \\ \eta^{(1)} \end{array} \right) \leftrightarrow \left( \begin{array}{c} \theta^{(2)} \\ \eta^{(2)} \end{array} \right)$
Acceptance probability for MH step from $M_1$ to $M_2$

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

$$A := \frac{p(2, \theta^{(2)}|y)}{p(1, \theta^{(1)}|y)} \cdot \frac{p_{2\rightarrow1}}{p_{1\rightarrow2}} \cdot \frac{\varphi_2(\eta^{(2)})}{\varphi_1(\eta^{(1)})} \cdot \left| \frac{\partial (\theta^{(2)}, \eta^{(2)})}{\partial (\theta^{(1)}, \eta^{(1)})} \right|$$

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

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

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

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

For moves from $M_2$ to $M_1$ we use $A^{-1}$. 
<table>
<thead>
<tr>
<th>Notation</th>
<th>Meaning</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>B0D0</td>
<td>stay move</td>
<td>$c_{12} \rightarrow c_{12}$</td>
</tr>
<tr>
<td>B1D0</td>
<td>birth of 1 factor</td>
<td>$c_{12} \rightarrow c_{12}c_{23}$</td>
</tr>
<tr>
<td>D1B0</td>
<td>death of 1 factor</td>
<td>$c_{12}c_{23} \rightarrow c_{12}$</td>
</tr>
</tbody>
</table>
Graph of all PCC’s for $n = 3$

With probabilities $a$ or $b$ the chain moves to other model or stays in the current model.
Acceptance probability for B1D0 move

Actual state \( \theta^o := \theta^o_{m^o} \) \( m^o = (m_1^o, m_2^o, m_3^o) \)

Proposed state \( \theta^p := \theta^p_{m^p} \) \( m^p = (m_1^p, m_2^p, m_3^p) \)

Birth index \( \{m^p_s\} := m^p \setminus m^o \)

Stay indices \( \{m^o_v, m^o_w\} := m^p \cap m^o \)

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

Bijection:

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

Here \( \eta^o_{m^p_s} \) is bivariate normal centered at \( \theta^p_{m^p_s, last} \) with specified covariance matrix.
Is the correlation $\rho_{ST|M} = -0.11$ significantly negligible?
Would Model $M_2$ now be preferred to Model $M_1$?

<table>
<thead>
<tr>
<th>Model notation</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{SM}c_{MT}c_{ST</td>
<td>M}$</td>
</tr>
<tr>
<td>$c_{SM}c_{MT}$</td>
<td>$M_2$</td>
</tr>
<tr>
<td>$c_{SM}c_{ST</td>
<td>M}$</td>
</tr>
<tr>
<td>$c_{MT}c_{ST</td>
<td>M}$</td>
</tr>
<tr>
<td>$c_{SM}$</td>
<td>$M_5$</td>
</tr>
<tr>
<td>$c_{MT}$</td>
<td>$M_6$</td>
</tr>
<tr>
<td>$c_{ST</td>
<td>M}$</td>
</tr>
</tbody>
</table>
Financial returns: Statistics of visited models

Model $M_1$ was visited 2872 times out of 10000. Model $M_2$ was visited 7128 times out of 10000.
## Summary statistics for visited Models

**Model $M_2$: $c_{SM} c_{MT}$**

<table>
<thead>
<tr>
<th></th>
<th>2.5% Quantile</th>
<th>50% Quantile</th>
<th>97.5% Quantile</th>
<th>Est. Post. Mean</th>
<th>Est. Post. Mode</th>
<th>MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{SM}$</td>
<td>-0.31</td>
<td>-0.25</td>
<td>-0.18</td>
<td>-0.25</td>
<td>-0.25</td>
<td>-0.25</td>
</tr>
<tr>
<td>$\nu_{SM}$</td>
<td>3.39</td>
<td>4.51</td>
<td>6.86</td>
<td>4.65</td>
<td>4.39</td>
<td>4.21</td>
</tr>
<tr>
<td>$\rho_{MT}$</td>
<td>0.42</td>
<td>0.47</td>
<td>0.50</td>
<td>0.47</td>
<td>0.47</td>
<td>0.47</td>
</tr>
<tr>
<td>$\nu_{MT}$</td>
<td>12.64</td>
<td>270.89</td>
<td>944.71</td>
<td>343.05</td>
<td>122.94</td>
<td>16.65</td>
</tr>
</tbody>
</table>

**Model $M_1$: $c_{SM} c_{MT} c_{ST|\lambda}$**

<table>
<thead>
<tr>
<th></th>
<th>2.5% Quantile</th>
<th>50% Quantile</th>
<th>97.5% Quantile</th>
<th>Est. Post. Mean</th>
<th>Est. Post. Mode</th>
<th>MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{SM}$</td>
<td>-0.32</td>
<td>-0.25</td>
<td>-0.18</td>
<td>-0.25</td>
<td>-0.25</td>
<td>-0.25</td>
</tr>
<tr>
<td>$\nu_{SM}$</td>
<td>3.31</td>
<td>4.61</td>
<td>7.08</td>
<td>4.71</td>
<td>4.49</td>
<td>4.24</td>
</tr>
<tr>
<td>$\rho_{MT}$</td>
<td>0.43</td>
<td>0.47</td>
<td>0.52</td>
<td>0.47</td>
<td>0.47</td>
<td>0.47</td>
</tr>
<tr>
<td>$\nu_{MT}$</td>
<td>16.24</td>
<td>264.85</td>
<td>861.86</td>
<td>305.85</td>
<td>156.72</td>
<td>16.45</td>
</tr>
<tr>
<td>$\rho_{ST</td>
<td>\lambda}$</td>
<td>-0.17</td>
<td>-0.11</td>
<td>-0.04</td>
<td>-0.11</td>
<td>-0.11</td>
</tr>
<tr>
<td>$\nu_{ST</td>
<td>\lambda}$</td>
<td>126.20</td>
<td>498.14</td>
<td>958.40</td>
<td>527.08</td>
<td>420.38</td>
</tr>
</tbody>
</table>
## Posterior model probabilities

### RJMCMC

| Model | PCC                        | Formula                 | $P(M_k|data)$ |
|-------|----------------------------|-------------------------|---------------|
| $M_1$ | with all pairs             | $c_{SM}c_{MT}c_{ST|M}$  | 0.29          |
| $M_2$ | without $c_{ST|M}$         | $c_{SM}c_{MT}$          | 0.71          |
| $M_3$ | without $c_{MT}$           | $c_{SM}c_{ST|M}$        | 0             |
| $M_4$ | without $c_{SM}$           | $c_{MT}c_{ST|M}$        | 0             |
| $M_5$ | without $c_{MT}$ and $c_{ST|M}$ | $c_{SM}$               | 0             |
| $M_6$ | without $c_{SM}$ and $c_{ST|M}$ | $c_{MT}$               | 0             |
| $M_7$ | without $c_{SM}$ and $c_{MT}$ | $c_{ST|M}$             | 0             |

### Congdon

| Model | PCC                        | Formula                 | $P(M_k|data)$ |
|-------|----------------------------|-------------------------|---------------|
| $M_1$:| with all pairs             | $c_{SM}c_{MT}c_{ST|M}$  | 0.22          |
| $M_2$:| without $c_{ST|M}$         | $c_{SM}c_{MT}$          | 0.78          |
Euro Swap Rates

- daily swap rates with 2, 3, 5, 7 and 10 year maturity quoted in Euro’s
- considered period is Dec 7, 1988 until June 21, 2001
- serial marginal correlation removed with an ARMA(1,1)-GARCH(1,1) model and standardized residuals transformed empirically to data with uniform margins
- Notation: Swap2, Swap3, Swap5, Swap7 and Swap10 or abbreviated $S_1, ..., S_5$
Bayesian Inference for Pair-copula Constructions of Multiple Dependence
Euro Swap Rates

Dependence structure of Euro Swap Rates
Estimated df’s of bivariate t-copula margins for swap rates

<table>
<thead>
<tr>
<th></th>
<th>Swap2</th>
<th>Swap3</th>
<th>Swap5</th>
<th>Swap7</th>
<th>Swap10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swap3</td>
<td>2.87</td>
<td>3.62</td>
<td>2.86</td>
<td>4.53</td>
<td>5.23</td>
</tr>
<tr>
<td>Swap5</td>
<td>2.07</td>
<td>3.00</td>
<td>3.56</td>
<td>4.93</td>
<td></td>
</tr>
<tr>
<td>Swap7</td>
<td>5.23</td>
<td>4.93</td>
<td>3.56</td>
<td>2.07</td>
<td></td>
</tr>
</tbody>
</table>
Chosen D-vine structure for swap rates
Autocorrelations among the MCMC iterations for the swap rates

For the further analysis we used burnin 1000 iteration and only every 20th iteration.
Estimated posterior densities of D-vine parameters
## Summary statistics for thinned MCMC

<table>
<thead>
<tr>
<th></th>
<th>2.5% Quantile</th>
<th>50% Quantile</th>
<th>97.5% Quantile</th>
<th>Est. Post. Mean</th>
<th>Est. Post. Mode</th>
<th>MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{S_1 S_2}$</td>
<td>0.93</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>$\nu_{S_1 S_2}$</td>
<td>2.33</td>
<td>2.73</td>
<td>3.26</td>
<td>2.74</td>
<td>2.72</td>
<td>2.49</td>
</tr>
<tr>
<td>$\rho_{S_2 S_3}$</td>
<td>0.93</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>$\nu_{S_2 S_3}$</td>
<td>2.62</td>
<td>3.03</td>
<td>3.54</td>
<td>3.04</td>
<td>3.00</td>
<td>3.08</td>
</tr>
<tr>
<td>$\rho_{S_3 S_4}$</td>
<td>0.92</td>
<td>0.92</td>
<td>0.93</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
</tr>
<tr>
<td>$\nu_{S_3 S_4}$</td>
<td>2.62</td>
<td>3.03</td>
<td>3.48</td>
<td>3.04</td>
<td>3.02</td>
<td>2.87</td>
</tr>
<tr>
<td>$\rho_{S_4 S_5}$</td>
<td>0.95</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>$\nu_{S_4 S_5}$</td>
<td>1.86</td>
<td>2.12</td>
<td>2.50</td>
<td>2.14</td>
<td>2.11</td>
<td>2.27</td>
</tr>
<tr>
<td>$\rho_{S_1 S_3</td>
<td>S_2}$</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.07</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>$\nu_{S_1 S_3</td>
<td>S_2}$</td>
<td>4.19</td>
<td>5.32</td>
<td>7.00</td>
<td>5.40</td>
<td>5.27</td>
</tr>
<tr>
<td>$\rho_{S_2 S_4</td>
<td>S_3}$</td>
<td>0.04</td>
<td>0.08</td>
<td>0.12</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>$\nu_{S_2 S_4</td>
<td>S_3}$</td>
<td>4.63</td>
<td>5.63</td>
<td>7.36</td>
<td>5.70</td>
<td>5.53</td>
</tr>
<tr>
<td>$\rho_{S_3 S_5</td>
<td>S_4}$</td>
<td>0.08</td>
<td>0.12</td>
<td>0.15</td>
<td>0.12</td>
<td>0.11</td>
</tr>
<tr>
<td>$\nu_{S_3 S_5</td>
<td>S_4}$</td>
<td>4.64</td>
<td>5.62</td>
<td>7.22</td>
<td>5.67</td>
<td>5.55</td>
</tr>
<tr>
<td>$\rho_{S_1 S_4</td>
<td>S_2 S_3}$</td>
<td>-0.10</td>
<td>-0.07</td>
<td>-0.03</td>
<td>-0.07</td>
<td>-0.07</td>
</tr>
<tr>
<td>$\nu_{S_1 S_4</td>
<td>S_2 S_3}$</td>
<td>10.84</td>
<td>18.12</td>
<td>55.27</td>
<td>22.21</td>
<td>16.48</td>
</tr>
<tr>
<td>$\rho_{S_2 S_5</td>
<td>S_3 S_4}$</td>
<td>-0.11</td>
<td>-0.08</td>
<td>-0.04</td>
<td>-0.08</td>
<td>-0.08</td>
</tr>
<tr>
<td>$\nu_{S_2 S_5</td>
<td>S_3 S_4}$</td>
<td>7.10</td>
<td>9.51</td>
<td>14.58</td>
<td>9.90</td>
<td>9.20</td>
</tr>
<tr>
<td>$\rho_{S_1 S_5</td>
<td>S_2 S_3 S_4}$</td>
<td>-0.06</td>
<td>-0.03</td>
<td>0.01</td>
<td>-0.03</td>
<td>-0.03</td>
</tr>
<tr>
<td>$\nu_{S_1 S_5</td>
<td>S_2 S_3 S_4}$</td>
<td>15.37</td>
<td>64.51</td>
<td>954.21</td>
<td>223.08</td>
<td>55.84</td>
</tr>
</tbody>
</table>
### Parameter values

<table>
<thead>
<tr>
<th>Model Description</th>
<th>Loglikelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivariate t with common df</td>
<td>13937.15</td>
</tr>
<tr>
<td>MLE</td>
<td>14036.83</td>
</tr>
<tr>
<td>Posterior Mode</td>
<td>14037.96</td>
</tr>
</tbody>
</table>

- **Likelihood Ratio Test** rejects a multivariate t copula with common df decisively, estimated common df is 3.36.
- Since posterior mode log likelihood is close to MLE likelihood, prior choice for $\rho$ and $\nu$ are close to being noninformative in this data set.
## Model selection for swap rate data

| Model | PCC                        | $P(M_k|\text{data})$  |
|-------|----------------------------|------------------------|
| $M_1$: | with all pairs             | 0.19                   |
| $M_2$: | without $c_{s_1s_5|s_2s_3s_4}$ | 0.81                   |
| $M_3$: | without $c_{s_1s_5|s_2s_3s_4}$ and $c_{s_2s_5|s_3s_4}$ | $2.5 \times 10^{-8}$ |
| $M_4$: | without $c_{s_1s_5|s_2s_3s_4}$, $c_{s_2s_5|s_3s_4}$ and $c_{s_1s_4|s_2s_3}$ | $5.2 \times 10^{-11}$ |
Summary and Conclusion

- PCC’s such as canonical and D-Vines allow for very flexible class of multivariate distributions.
- The Bayesian approach can solve estimation as well as model selection problems. It also gives credible intervals for parameters of interest.
- The proposed RJMCMC Bayesian algorithm for $n > 3$ is under implementation. Incorporation of pair-copulas from different parametric families is under consideration.
- The proposed RJMCMC algorithm is an alternative to DIC and other Bayesian model choice algorithms.
- Congdon’s method can be used to make model comparisons for a small number of models.


