Dealing With Vine's Issues

P.A. Maugis[1] and D. Guegan[2]

[1]ENS, CES
University Paris I Pantheon-Sorbonne
[2]CNRS, CES
University Paris I Pantheon-Sorbonne

3rd Financial Risks International Forum

Overview

General Copula Definition

Definition: Copulas

- Let $X = [X_1, X_2, ..., X_n]$ be a vector of random variables:
  - of distribution function F,
  - of joint density f,
  - of marginal distribution functions $F_1, F_2, ..., F_n$,
  - of marginal densities $f_1, f_2, ..., f_n$.
- Sklar (1959)'s theorem insures the existence of two functions c and C such that:

$$\begin{aligned} F(x) &= C(F_1(x_1), F_2(x_2), ..., F_n(x_n)) \\ f(x) &= \prod_{i=1}^{n} f_i(x_i) \cdot c(F_1(x_1), F_2(x_2), ..., F_n(x_n)) \end{aligned}$$

Vine Copula Definition

Definition: Vine Copulas

A vine copula is a n-variate parametric copula built by decomposing the multivariate density in a product of bivariate copulas.

- Origins: Joe (1997) and Bedford and Cooke (2002, 2001)
- Applications: Aas et al. (2009), Berg and Aas (2009), Czado et al. (2009) and Guégan and Maugis (2009).

Decomposition Example In Dimension 3

$$
\begin{aligned}
f(x_1, x_2, x_3) &= f(x_1) \cdot f(x_2) \cdot f(x_3) \\
&\cdot c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{13}(F_1(x_1), F_3(x_3)) \\
&\cdot c_{23|1}(\partial_1 C_{12}(F_1(x_1), F_2(x_2)), \partial_1 C_{13}(F_1(x_1), F_2(x_2)))
\end{aligned}
$$

Trivariate Vine Copula

$$
\begin{aligned}
c_{123} &= c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{13}(F_1(x_1), F_3(x_3)) \\
&\cdot c_{23|1}(\partial_1 C_{12}(F_1(x_1), F_2(x_2)), \partial_1 C_{13}(F_1(x_1), F_2(x_2))) \\
c_{123} &= c_{12} \cdot c_{13} \cdot c_{23|1}
\end{aligned}
$$

Selection Issue

Issues:

- Number of vine copulas:

$$\frac{n!}{2}\prod_{i=1}^{n-3} i!$$

- Copula Selection: Chen et al. (2004) and Chen and Fan (2006) and Kiefer and Choi (2006).

Lattice Selection:

Gabriel (1969), Andersson (1988), Edwards and Havranek (1987) and Antoch and Hanousek (2000)

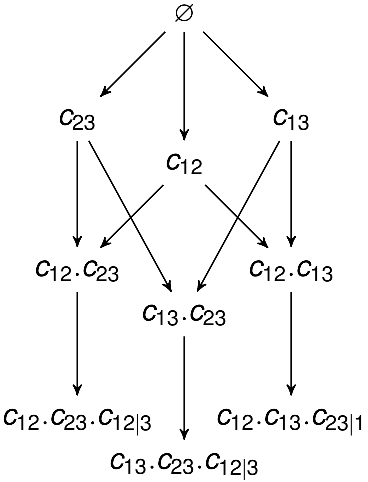
Specification & Generalization: Lattice structure

Among models, some can be considered as specifications of others, and some as generalizations of others.

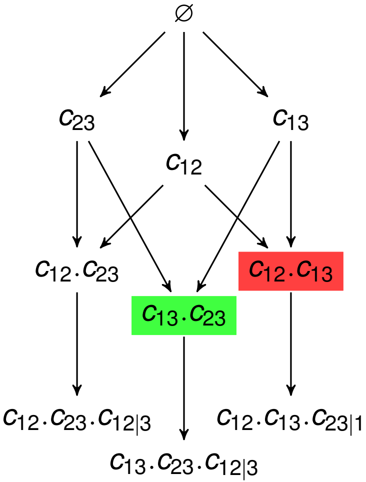
Coherence Principle:

Specifications of a false model are false, and generalizations of valid models are valid.

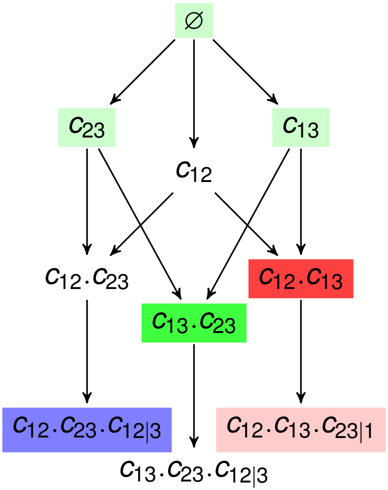
Example


∅
C_{23}
C_{12}
C_{13}
$C_{12}.C_{23}$
$C_{13}.C_{23}$
$C_{12}.C_{13}$
$C_{12}.C_{23}.C_{12|3}$
$C_{13}.C_{23}.C_{12|3}$
$C_{12}.C_{13}.C_{23|1}$

Example

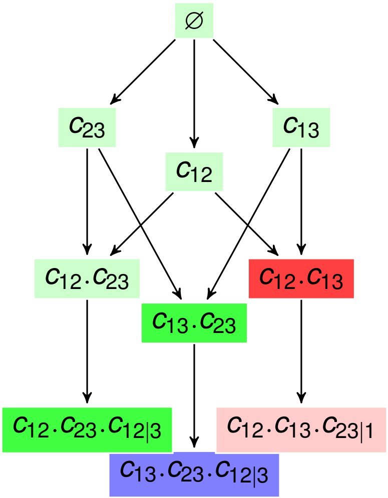

∅
c_{23}
c_{13}
c_{12}
$c_{12} \cdot c_{23}$
$c_{12} \cdot c_{13}$
$c_{13} \cdot c_{23}$
$c_{12} \cdot c_{23} \cdot c_{12|3}$
$c_{12} \cdot c_{13} \cdot c_{23|1}$
$c_{13} \cdot c_{23} \cdot c_{12|3}$

Example

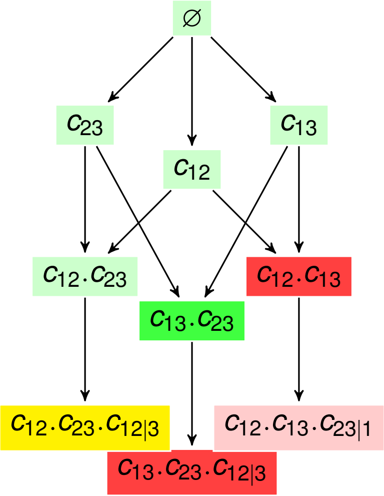


∅
c_{23}
c_{12}
c_{13}
$c_{12} \cdot c_{23}$
$c_{13} \cdot c_{23}$
$c_{12} \cdot c_{13}$
$c_{12} \cdot c_{23} \cdot c_{12|3}$
$c_{13} \cdot c_{23} \cdot c_{12|3}$
$c_{12} \cdot c_{13} \cdot c_{23|1}$

Example


∅
C_{23}
C_{13}
C_{12}
$C_{12}.C_{23}$
$C_{12}.C_{13}$
$C_{13}.C_{23}$
$C_{12}.C_{23}.C_{12|3}$
$C_{12}.C_{13}.C_{23|1}$
$C_{13}.C_{23}.C_{12|3}$

Example



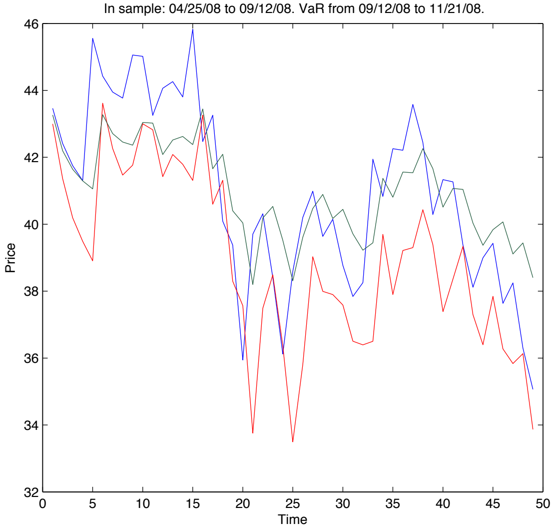
$\varnothing$
c_{23}
c_{12}
c_{13}
$c_{12}.c_{23}$
$c_{13}.c_{23}$
$c_{12}.c_{13}$
$c_{12}.c_{23}.c_{12|3}$
$c_{12}.c_{13}.c_{23|1}$
$c_{13}.c_{23}.c_{12|3}$

CAC40

Procedure:

- Five main assets composing the *CAC*40 from 25/04/08 to 21/11/08 (Datastream).
- In-sample: 04/25/08 to 09/12/08.
  Out-Sample: 09/12/08 to 11/21/08.
- Fully automated estimation and search.
  - Univariate process: *GARCH* using *AIC*.
  - Multivariate estimation: Maximum Likelihood.
  - Test: Modified Anderson Darling (Chen et al., 2004).

VaR



In sample: 04/25/08 to 09/12/08. VaR from 09/12/08 to 11/21/08.
Price
Time
46
44
42
40
38
36
34
32
0
5
10
15
20
25
30
35
40
45
50

Summary

- Vine copulas are parametric copulas built using bivariate copulas.
- Vine copulas can be selected using a lattice based search algorithm.
- By completely modeling the multivariate distribution, Vine copulas permits to build better risk measures.

Issues & Prospects

Issues:

Heavily Computational.

Prospects:

Better bivariate copulas.

Bibliography I

Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44, 182–198.

Andersson, S., 1988. The Lattice Structure Of Orthogonal Linear Models And Orthogonal Variance Component Models. University of Washington Technical Report.

Antoch, J., Hanousek, J., 2000. Model Selection and Simplification Using Lattices. CERGE-EI Working Paper.

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

Bedford, T., Cooke, R., 2002. Vines: A New Graphical Model for Dependent Random Variables. The Annals of Statistics 30 (4), 1031–1068.

Berg, D., Aas, K., 2009. Models for construction of multivariate dependence. Forthcoming in The Europen Journal of Finance.

Chen, X., Fan, Y., 2006. Estimation of copula-based semi-parametric time series models. Journal of Econometrics 130, 307–335.

Bibliography II

Chen, X., Fan, Y., Patton, A., 2004. Simple Tests for Models of Dependence Between Multiple Financial Time Series, with Applications to U.S. Equity Returns and Exchange Rates. London Economics Financial Markets Group Working Paper (483), london, UK.

Czado, C., Gartner, F., Min, A., 2009. Analysis of Australian electricity loads using joint Bayesian inference of D-Vines with autoregressive margins. Zentrum Mathematik Technische Universitat Munchen: Working Paper Munchen, Germany.

Edwards, D., Havranek, T., 1987. A Fast Model Selection Procedure for Large Families of Models. Journal of the American Statistical Association 82 (397), 205–213.

Gabriel, K. R., 1969. Simultaneous Test Procedures - Some Theory of Multiple Comparisons. Annals of Mathematics Statistics 40 (1), 224–250.

Guégan, D., Maugis, P. A., 2009. Prospects On Vines. Working Paper.

Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman & Hall, London, UK.

Bibliography III

Kiefer and Choi, 2006. Robust Model Selection in Dynamic Models with an Application to Comparing Predictive Accuracy. CAE Working Paper.

Sklar, A., 1959. Fonctions de répartition à *n* dimensions et leurs marges. Publications de l'Institut de Statistique de L' Université de Paris 8, 229–231.

Annex 1

$$
\begin{aligned}
f(x_1, x_2, x_3) &= f(x_1, x_2) \cdot \frac{f(x_2, x_3 | x_1)}{f(x_2 | x_1)} \\
f(x_1, x_2, x_3) &= f(x_1) \cdot f(x_2) \cdot c_{12}(F_1(x_1), F_2(x_2)) \\
&\quad \cdot \frac{f(x_2|x_1) \cdot f(x_3|x_1)}{f(x_2|x_1)} \cdot c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)) \\
f(x_1, x_2, x_3) &= f(x_1) \cdot f(x_2) \cdot c_{12}(F_1(x_1), F_2(x_2)) \\
&\quad \cdot \frac{f(x_1, x_3)}{f(x_1)} \cdot c_{23|1}(\partial_1 C_{12}(F_1(x_1), F_2(x_2)), \partial_1 C_{13}(F_1(x_1), F_2(x_2))) \\
f(x_1, x_2, x_3) &= f(x_1) \cdot f(x_2) \cdot f(x_3) \\
&\quad \cdot c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{13}(F_1(x_1), F_3(x_3)) \\
&\quad \cdot c_{23|1}(\partial_1 C_{12}(F_1(x_1), F_2(x_2)), \partial_1 C_{13}(F_1(x_1), F_2(x_2)))
\end{aligned}
$$

Annex 2 I

The test is the following: we are testing the null hypothesis H_0 against the alternative hypothesis H_1:

$$H_0 : Pr(C(U_1, \dots, U_n) = C_0(U_1, \dots, U_n)) = 1$$

$$H_1 : Pr(C(U_1, \dots, U_n) = C_0(U_1, \dots, U_n)) < 1$$

where C is the true copula and C_0 the estimated copula. We define then $\{Z_i\}_{i<n}$ and W as:

$$Z_i = F(U_i | U_1, \dots, U_{i-1}) \text{ and } W = \sum_1^n [\Phi^{-1}(Z_j)]^2,$$

where Φ is the standard normal distribution function. Then W follows a χ^2_d distribution; our test will be based on this result. For this purpose we will use this univariate boundary kernel $K_h(x, y)$

$$K_h(x, y) = \begin{cases} k(\frac{x-y}{h}) / \int_{-\frac{x}{h}}^{1} k(u)du & \text{if } x \in [0, h) \\ k(\frac{x-y}{h}) & \text{if } x \in [h, 1-h] \\ k(\frac{x-y}{h}) / \int_{-\frac{x}{h}}^{1} k(u)du & \text{if } x \in (1-h, 1] \end{cases}$$

Annex 2 II

in which $k(\cdot)$ is the quartic kernel: $k(u) = \frac{15}{16}(1-u^2)^2\mathbf{1}_{|u|<1}$. Then we define g_W as a kernel estimation of its density function, where h is fixed using the rule of thumb ($h = \sqrt{Var(W)}n^{-1/5}$):

$$g_W(\omega) = \frac{1}{n * h}\sum_t^T K_h(\omega, F_{\chi^2_d}(W_t))$$

The alternative test is based on:

$$J_n = \int_0^1 [g_W(\omega) - 1]^2 \, d\omega$$

Then under classical hypothesis we have that:

$$Stat_n = \frac{(nh^{1/2}J_n - r_n)}{\sigma} \rightarrow N(0,1) \text{ in distribution.}$$

Annex 2 III

Where:

$$r_n = h1/2\left[(h^{-1} - 2)\int_{-1}^{1} k^2(\omega)d\omega + 2\int_{0}^{1}\int_{-1}^{z} k_z^2(y)dydz\right]$$

$$\sigma^2 = 2\int_{-1}^{1}\left[\int_{-1}^{1} k(u+v)k(v)dv\right]^2 du$$

$$k_z(y) = k(y)/\int_{-1}^{z} k(u)du$$

For further details on the test, including Monte-Carlo estimation of the power of the test, see Chen et al. (2004).