Nonparametric Estimation of Copulas for Time Series

PEDRO A. MORETTIN*

Institute of Mathematics and Statistics
University of São Paulo
05311-970 São Paulo, Brazil
pam@ime.usp.br
http://www.ime.usp.br/~ pam
*Joint work with Clélia M.C. Toloi, Chang Chiann and José C.S. de Miranda

March 13-16, 2007

1. Copulas

- Copulas provide convenient tool for describing dependence between variables.
- Applications in Finance, Insurance, Risk Management, EVT.
- Software: SPlus(S+FinMetrics), R, MatLab, Mathematica etc.
- Copula: multivariate distribution whose marginal distributions are uniform on $(0, 1)$.
- Copulas allow to model marginal distributions and dependence struture of multivariate distribution separately.
- Some critical remarks (Mikosch, 2005):
  - not all problems related with stochastic dependences and multivariate distributions can be solved via copulas;
  - "curse of dimensionality";
  - why do we transform marginals to uniform?
  - do copulas fit in stochastic processes and time series?

2. Fitting Copulas

Estimation of copulas: complicated for multivariate time series in finance, given the presence of time dependences, serial autocorrelation, heteroskedasticity in asset returns, interest rates and exchange rates.

Copula estimates were developed basically in the context of iid samples.

(i) if the copula belongs to some parametric family consistent and asymptotically normal estimators of parameters can be obtained by maximum likelihod (ML)

Procedure in two stages:

- estimate marginal distributions;
- estimate copula.

The choice of marginals determines the copula, hence the dependence structure. Therefore, different instruments to fit marginals lead to different dependences.

2. Fitting Copulas

Three possibilities to fit marginals:

a) fit parametric distributions to the marginals;

b) use the edf $F_{i,n}$ to estimate F_i;

c) same as in (b), but add some tail, for exemple GPD tails.

Any of these approaches can go wrong, eg, if we want to estimate VaR.

Another problem: choice of copula: Archimedian, Gaussian, t, elliptical, extreme value etc.

Genest et al. (1995), Shi and Louis (1995).

2. Fitting Copulas

(ii) use empirical copulas. Deheuvels (1978, 1981).

(iii) nonparametric estimation: obtains smooth estimators, without assuming that the copula belong to some parametric family; based on kernels, splines, wavelets.

Fermanian & Scaillet (2003), Morettin et al. (2006a,b).

General references: Joe (1997), Nelsen (1999), Cherubini et al. (2004).

3. Definition

An $n-$ *dimensional copula* is a function C from $[0,1]^n$ to the interval $[0,1]$, with the properties:

(i) C is *grounded*: for every $\mathbf{u} = (u_1, \dots, u_n) \in [0,1]^n$, $C(\mathbf{u}) = 0$ if at least one coordinate $u_i = 0$, $i = 1, \dots, n$;

(ii) C is *n-increasing*: for every $\mathbf{u}$ and $\mathbf{v}$ in $[0,1]^n$, with $\mathbf{u} \leq \mathbf{v}$, the C-volume $V_C([\mathbf{u}, \mathbf{v}])$ of the box $[\mathbf{u}, \mathbf{v}]$ is non-negative;

(iii) $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$, for all $u_i \in [0,1]$.

See Nelsen (1999) for the definition of C-volume and further details on copulas. The following important theorem links the definition of copula with an n-dimensional distribution function and its marginal distributions.

4. Theorem (Sklar).

Let F be an n-dimensional distribution function with margins $F_1, \ldots, F_n$. Then there exists an n- copula C such that for all $\mathbf{x} = (x_1, \ldots, x_n) \in [-\infty, \infty]^n$, we have

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)). \tag{1}$$

Conversely, if C is an n-copula and $F_1, \ldots, F_n$ are distribution functions, the function F defined by (1) is an n-dimensional distribution function with margins $F_1, \ldots, F_n$. Moreover, if the margins are all continuous, then C is unique. Otherwise, C is uniquely determined on $RanF_1 \times \ldots \times RanF_n$.

Therefore, given Sklar's theorem, it is easy to construct the corresponding copula, namely

$$C(u_1, \ldots, u_n) = F(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)), \tag{2}$$

where $F_i^{-1}(u_i) = \inf\{x_i \mid F_i(x_i) \geq u_i\}$, $i = 1 \ldots, n$.

5. Wavelets

- Z: the set of all integers
- ${}_a\mathrm{Z}$: the set of all integers z with $z \geq a$
- ${}_a\mathrm{Z}_b$: the set of all integers z such that $a \leq z \leq b$
- $\mathrm{Z}(0) = \mathrm{Z} \cup {}_0\mathrm{Z} \times \mathrm{Z}$
- $\mathrm{Z}(0, J) = \mathrm{Z} \cup {}_0\mathrm{Z}_J \times \mathrm{Z}$.

When we write $\psi_\eta, \ \eta \in \mathrm{Z}(0)$ we mean that $\psi_\eta = \phi_{0,\eta}$, if $\eta \in \mathrm{Z}$, and $\psi_\eta = \psi_{(j,k)} = \psi_{j,k}$, if $\eta \in {}_0\mathrm{Z} \times \mathrm{Z}$.

Given an unknown function f we consider its wavelet expansion

$$f = \sum_{\eta \in \mathrm{Z}(0)} \beta_\eta \psi_\eta, \tag{3}$$

where $\{\psi_\eta, \eta \in \mathrm{Z}(0)\}$ is a compactly supported wavelet basis.

The wavelet coefficients are given by

$$\beta_\eta = \int f \psi_\eta d\ell. \tag{4}$$

Interest: $f(\mathbf{x})$ is a density, supposed to belong to $L_2(\mathbb{R}^n)$, where $\mathbf{x} = (x_1, \ldots, x_n)'$. A wavelet expansion similar to (4) for $f(\mathbf{x})$ will hold, where the wavelets are obtained as products of one-dimensional wavelets.

We illustrate here the case $n = 2$.

One possibility: basis with a single scale:

$$f(x_1, x_2) = c_{0,\mathbf{0}} + \sum_{j=0}^{\infty} \sum_{\mathbf{k}} \sum_{\mu=h,v,d} d_{j,\mathbf{k}}^{\mu} \Psi_{j,\mathbf{k}}^{\mu}(x_1, x_2), \tag{5}$$

with the wavelet coefficients given by

$$d_{j,\mathbf{k}}^{\mu} = \int f(x_1, x_2) \Psi_{j,\mathbf{k}}^{\mu}(x_1, x_2) dx_1 dx_2. \tag{6}$$

Another possibility: basis as the tensor product of two one-dimensional bases with different scales for each dimension. Here

$$f(x_1, x_2) = \sum_I d_I \mu_I(x_1, x_2), \tag{7}$$

where $I = (j_1, j_2, k_1, k_2)$ and $\mu_I(x_1, x_2) = \psi_{j_1,k_1}(x_1)\psi_{j_2,k_2}(x_2)$. The $\psi_{j,k}(\cdots)$'s here include also father wavelets. The two bases imply different tilings of the time-scale plane.

6. Wavelet estimators

Consider an n-dimensional process $\{\mathbf{X}_t, t \in \mathbb{Z}\}$, such that for all t, $\mathbf{X}_t = (X_{1t}, \ldots, X_{nt})'$ has a density $f_t(\mathbf{x}) = f(\mathbf{x})$ and distribution function $F_t(\mathbf{x}) = F(\mathbf{x})$, with $\mathbf{x} = (x_1, \ldots, x_n)'$. Stationarity is not assumed.

Observe the process at T time points, obtaining $\mathbf{X}_1, \ldots, \mathbf{X}_T$.

f_i and F_i, $i = 1, \ldots, n$: pdf and cdf of each X_{it}, respectively.
Our purpose: estimate F, F_i and C in (2) using wavelet methods.

Notation:

- $J = (J_1, \ldots, J_n)'$, where each $J_i = J_i(T)$.
- $\eta = (\eta_1, \ldots, \eta_n)'$.
- $\eta_\ell = (j_\ell, k_\ell)$, or $\eta_\ell = k_\ell \in \mathbb{Z}, \ell = 1, \ldots, n$.
- $\psi_\eta = \psi_{\eta_1} \otimes \cdots \otimes \psi_{\eta_n}$.

6. Wavelet estimators

As an estimator of $f_i(x_i)$ take

$$\hat{f}_{i,J_i}(x_i) = \sum_{\eta_i} \hat{\beta}_{\eta_i} \psi_{\eta_i}, \tag{8}$$

then estimate the d.f. $F_i(x_i)$ by

$$\hat{F}_{i,J_i}(x_i) = \int_{-\infty}^{x_i} \hat{f}_{i,J_i}(y)dy. \tag{9}$$

As an estimator of $f(\mathbf{x})$ take

$$\hat{f}_J(\mathbf{x}) = \sum_{\eta} \hat{\beta}_{\eta} \psi_{\eta}(\mathbf{x}). \tag{10}$$

6. Wavelet estimators

We may consider nonlinear estimators by replacing $\hat{\beta}_\eta$ in (10), by $\delta(\hat{\beta}_\eta, \lambda)$, where $\delta(\cdot, \lambda)$ is a threshold and λ is a threshold parameter which can be specified in a number of ways.

Examples of thresholds are

$$\delta_s(x, \lambda) = sgnx(|x| - \lambda)_+,$$

$$\delta_h(x, \lambda) = xI\{|x| \geq \lambda\},$$

the so-called soft and hard thresholds, respectivelly.

Instances of the choice of λ are the universal, SURE, minimax and Bayesian procedures.

6. Wavelet estimators

The empirical wavelet coefficients are given by

$$\hat{\beta}_\eta = \frac{1}{T} \sum_{t=1}^{T} \psi_\eta(\mathbf{X}_t). \qquad (11)$$

If

$$F_T(\mathbf{x}) = \frac{1}{T} \sum_{t=1}^{T} I\{\mathbf{X}_t \le \mathbf{x}\}$$

is the empirical distribution function, then the estimator (11) can be written as

$$\hat{\beta}_\eta = \int_{\mathbb{R}^n} \psi_\eta dF_T(\mathbf{x}).$$

It follows that these are unbiased estimators of the corresponding coefficients.

As an estimator of the distribution function (df) F of $\mathbf{X}_t$ at $\mathbf{x}$ we take

$$\hat{F}_J(\mathbf{x}) = \int_{-\infty}^{\mathbf{x}} \hat{f}_J(\mathbf{y})d\mathbf{y}. \tag{12}$$

Write $\hat{F}_J = (\hat{F}_{1,J_1}, \ldots, \hat{F}_{n,J_n})'$.

By (2), to estimate the copula at some point $\mathbf{u} = (u_1, \ldots, u_n)'$, we propose

$$\hat{C}_J(\mathbf{u}) = \hat{F}_J(\hat{\xi}), \tag{13}$$

$\hat{\xi} = (\hat{\xi}_1, \ldots, \hat{\xi}_n)'$, $\hat{\xi}_i = \inf\{x_i \in \mathbb{R} : \hat{F}_{i,J_i}(x_i) \geq u_i\}$, $i = 1, \ldots, n$, from which it follows that $\hat{\xi}_i$ is the wavelet estimator of the quantile of X_{it} with probability u_i, $i = 1, \ldots, n$.

7. Properties of the estimators

We restrict attention to the case where the densities belong to $L_2(\mathbb{R})$ and $L_2(\mathbb{R}^n) \cap L_\infty(\mathbb{R}^n)$, respectively, for the marginals and joint distribution.

Let $f_{t,s}(\mathbf{x}, \mathbf{y})$ be the joint density of $\mathbf{X}_t$ and $\mathbf{X}_s$ and $f(\mathbf{x})$ the density of $\mathbf{X}_t$, for every t, s. Define, for $t \neq s$,

$$q_{t,s}(\mathbf{x}, \mathbf{y}) = f_{t,s}(\mathbf{x}, \mathbf{y}) - f(\mathbf{x})f(\mathbf{y}), \tag{14}$$

and let the wavelet expansion of $q_{t,s}$ be given by

$$q_{t,s}(\mathbf{x}, \mathbf{y}) = \sum_{\mu} \sum_{\rho} \gamma_{\mu,\rho}^{(t,s)} \psi_\mu \otimes \psi_\rho(\mathbf{x}, \mathbf{y}). \tag{15}$$

Assumptions. Mixing conditions in terms of the coefficients $\gamma_{\mu,\rho}^{(t,s)}$ of the expansion (13) and in terms of the behavior of the scale $J_i(T)$, as $T \to \infty$, e.g.,

$\sum_{t\neq s} |\gamma_{\eta,\xi}^{(t,s)}| < \infty$, for all η, ξ,

$\left(2^{\sum J_i}\right) / T \to 0$, as $T \to \infty$

Theorem 1.

(i) $E(\hat{\beta}_\eta) = \beta_\eta$.

(ii) $\mathrm{Cov}(\hat{\beta}_\eta, \hat{\beta}_\xi) = \frac{1}{T}(\int \psi_\eta(\mathbf{x})\psi_\xi(\mathbf{x})f(\mathbf{x})d\mathbf{x} - \beta_\eta\beta_\xi) + \frac{1}{T^2}\sum_{t\neq s}\gamma_{\eta,\xi}^{(t,s)}$.

(iii) Under the assumptions, the empirical wavelet coefficients are consistent and asymptotically uncorrelated.

Theorem 2.

(i) The covariance structure of $\hat{F}_J(\mathbf{x})$ is given by:

$$\mathrm{Cov}(\hat{F}_J(\mathbf{x}), \hat{F}_J(\mathbf{y})) = \sum_{\eta} \sum_{\xi} \int_{-\infty}^{\mathbf{x}} \int_{-\infty}^{\mathbf{y}} \psi_\eta \otimes \psi_\xi \, d\ell^2 \mathrm{Cov}(\hat{\beta}_\eta, \hat{\beta}_\xi),$$

where the sums are for all scales up to and including scale J

(ii) Under the assumptions, the estimators $\hat{F}_J(\cdot)$ are consistent and asymptotically non-correlated, no matter how $J \to \infty$, as $T \to \infty$.

7. Properties of the estimators

Theorem 3. Under the assumptions, we have that

$$\sup_{\mathbf{x}} |\hat{F}_J(\mathbf{x}) - F(\mathbf{x})| \xrightarrow{P} 0, \quad T \to \infty. \tag{16}$$

Theorem 4. Under the assumptions, we have that

$$\sup_{\mathbf{0}<\mathbf{u}<\mathbf{1}} |\hat{C}_J(\mathbf{u}) - C(\mathbf{u})| \xrightarrow{P} 0, \quad T \to \infty. \tag{17}$$

(1) Bivariate VAR(1):

$$\mathbf{X}_t = \mathbf{A} + \mathbf{B}\mathbf{X}_{t-1} + \epsilon_t, \tag{18}$$

$\mathbf{X}_t = (X_{1t}, X_{2t})$, with independent components : $C(u_1, u_2) = u_1 u_2$
$\epsilon_t \sim N(0, \Sigma)$, $\quad \mathbf{A} = (1, 1)^{'}$,
$\mathrm{vec}(\mathbf{B}) = (0.25, 0, 0, 0.75)^{'}$,
$\mathrm{vec}(\Sigma) = (0.75, 0, 0, 1.25)^{'}$.

Number of Monte Carlo replications : 500, $T = 2^{10} = 1024$.

8. Some simulations

Table 1: *Bias and MSE of Haar wavelet estimator: independent case*

$\times 10^{-4}$	C(.01,.01)	C(.05,.05)	C(.25,.25)	C(.50,.50)	C(.75,.75)	C(.95,.95)	C(.99,.99)
True	1.00	25.00	625.00	2500.00	5625.00	9025.00	9801.00
Bias	0.05103	0.71621	9.06150	25.32195	28.57155	13.56050	5.32317
MSE	5.08978e-7	9.3428e-5	0.015	0.113	0.151	0.032	0.0055

Table 2: *Bias and MSE of kernel estimator: independent case*

$\times 10^{-4}$	C(.01,.01)	C(.05,.05)	C(.25,.25)	C(.50,.50)	C(.75,.75)	C(.95,.95)	C(.99,.99)
True	1.00	25.00	625.00	2500.00	5625.00	9025.00	9801.00
Bias	-0.09	-0.08	0.40	1.12	-0.90	-0.04	4.66
MSE	0.00	0.01	0.25	0.48	0.25	0.01	0.05

8. Some simulations

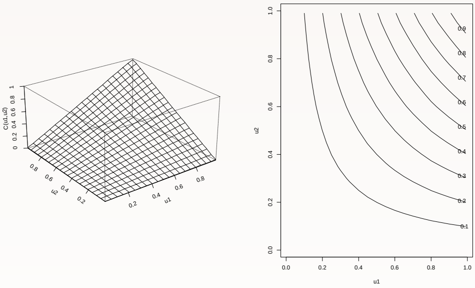

Figure 1: *Wavelet estimator for Haar wavelet: independent case.*

8. Some simulations

(2) Components of $\mathbf{X}_t$ are dependent processes, with
$\mathbf{A} = (1,1)'$, $\text{vec}(\mathbf{B}) = (0.25, 0.2, 0.2, 0.75)'$
$\text{vec}(\Sigma) = (0.75, 0.5, 0.5, 1.25)'$.
X_{1t} and X_{2t} are positively dependent, $C(u_1, u_2) > u_1 u_2$.
500 Monte Carlo replications, the data length $T = 1024$.

Table 3: *Bias and MSE of Haar wavelet estimator: dependent case*

$\times 10^{-4}$	C(.01,.01)	C(.05,.05)	C(.25,.25)	C(.50,.50)	C(.75,.75)	C(.95,.95)	C(.99,.99)
True	27.08	197.95	1511.74	3747.68	6511.74	9197.95	9827.08
Bias	0.6024	-1.6883	-21.3189	-32.9333	-13.2733	9.2234	16.7489
MSE	0.019	0.11	0.46	0.73	0.49	0.11	0.05

8. Some simulations

Table 4: *Bias and MSE of kernel estimator: dependent case*

$\times 10^{-4}$	C(.01,.01)	C(.05,.05)	C(.25,.25)	C(.50,.50)	C(.75,.75)	C(.95,.95)	C(.99,.99)
True	27.08	197.95	1511.74	3747.68	6511.74	9197.95	9827.08
Bias	-7.474	-34.88	-130.32	-172.28	-130.53	-35.25	-7.65
MSE	0.01	0.18	1.98	3.36	1.99	0.18	0.01

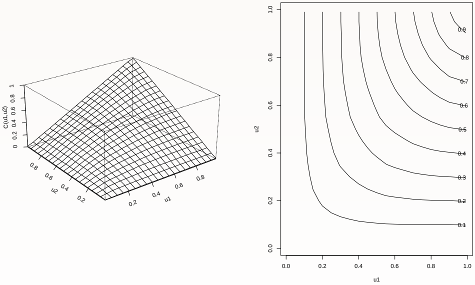

Figure 2: *Wavelet estimator for Haar wavelet: dependent case.*

9. An empirical application

Daily returns of the São Paulo Stock Exchange index (Ibovespa) and of prices of stocks of the Brazilian oil company, Petrobras, from January 2, 1995 to February 3, 1999 ($T = 1024$ observations).

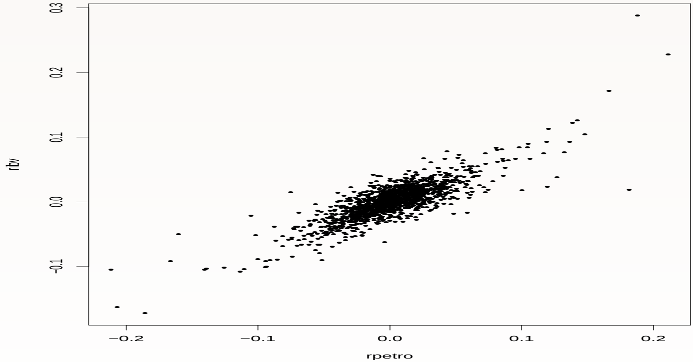

Figure 3: *Returns for Ibovespa and Petrobras.*

9. An empirical application

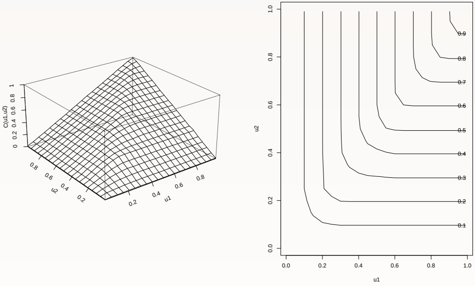

Figure 4: *Copula estimate for Ibovespa and Petrobras using Haar wavelet.*

Assume that the copula $C(u,v) \in L_\infty([0,1]^2)$ and consider its wavelet expansion

$$C(u,v) = c_{0,\mathbf{0}} + \sum_{j=0}^{\infty} \sum_{\mathbf{k}} \sum_{\mu=h,v,d} d_{j,\mathbf{k}}^{\mu} \Psi_{j,\mathbf{k}}^{\mu}(u,v), \tag{19}$$

with the wavelet coefficients given by

$$c_{0,\mathbf{0}} = \int C(u,v)dudv, \quad d_{j,\mathbf{k}}^{\mu} = \int C(u,v)\Psi_{j,\mathbf{k}}^{\mu}(u,v)dudv. \tag{20}$$

We take as the *empirical wavelet coefficients*,

$$\hat{d}^{\mu}_{j,\mathbf{k}} = \int C_n(u,v)\Psi^{\mu}_{j,\mathbf{k}}(u,v)dudv, \tag{21}$$

and a similar expression for $c_{0,\mathbf{0}}$, where $C_n(u,v)$ is the *empirical copula*.

We have that the corresponding estimator for $C(u,v)$ is then

$$\hat{C}(u,v) = \hat{c}_{0,\mathbf{0}} + \sum_{j,\mathbf{k}} \sum_{\mu} \delta(\hat{d}^{\mu}_{j,\mathbf{k}}, \lambda)\Psi^{\mu}_{j,\mathbf{k}}(u,v), \tag{22}$$

where $\delta(\cdot,\lambda)$ is a threshold.

Morettin et al. (2006b)

References

- Bouyé, E., Durrleman, V., Nikeghbali, A., Riboulet, G. and Roncalli, T. (2000). *Copulas for Finance: A Reading Guide and Some Applications*. Groupe de Recherche Opérationelle, Credit Lyonnais, Paris.
- Cherubini, U. and Luciano, E. (2000). Value-at-risk trade-off and capital allocation with copulas. *Economic Notes*, **30**.
- Daubechies, I. (1992) *Ten Lectures on Wavelets*. Philadelphia: SIAM.
- Deheulves, P. (1979). La fonction de dépendance empirique et ses propriétés- Un test non paramétrique d'independence. *Académie Royale de Belgique-Bulletin de la Classe des Sciences-5e Série*, **65**, 274–292.
- Deheulves,P. (1981) A non parametric test for independence. *Publications de l'Institut de Statistique de l'Université de Paris*, **26**, 29–50.

References

- Donoho, D.L., Johnstone, I., Kerkyacharian, G. and Picard, D. (1995). Wavelet shrinkage: Asymptopia? *Journal of the Royal Statistical Society*, Series B, **57**, 301–369.
- Donoho, D.L., Johnstone, I., Kerkyacharian, G. and Picard, D. (1996). Density estimation by wavelet thresholding. *Annals of Statistics*, **24**, 508–539.
- Fermanian, J.-D. and Scaillet, O. (2003). Nonparametric estimation of copulas for time series. *Journal of Risk*, **5**, 25–54.
- Genest, C., Ghoudi, K. and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, **82**, 543–552.
- Joe, H. (1997). *Multivariate Models and Dependence Concepts.* New York: Chapman and Hall.
- Mikosch, T. (2005). Copulas: Tales and Facts. To appear, *Extremes*.

References

- Morettin, P.A., Toloi, C.M.C., Chiann, C. and de Miranda, J.C.S. (2006a). Wavelet estimation of copulas for time series. Preprint.
- Morettin, P.A., Toloi, C.M.C., Chiann, C. and de Miranda, J.C.S. (2006b). Some remarks on copula estimators and a new proposal. Preprint.
- Nelsen, R. (1999). *An Introduction to Copulas*. Lecture Notes in Statistics. New York: Springer.
- Shi, J. and Louis, T. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrika*, **51**, 1384-1399.
- Sklar, A.(1959). Fonctions de répartition à *n* dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, **8**, 229–231.