World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:4, No:4, 2010

EML-Estimation of Multivariate *t* Copulas with Heuristic Optimization

Jin Zhang and Wing Lon Ng

Open Science Index, Computer and Information Engineering Vol:4, No:4, 2010 publications.waset.org/3163.pdf

Abstract—In recent years, copulas have become very popular in financial research and actuarial science as they are more flexible in modelling the co-movements and relationships of risk factors as compared to the conventional linear correlation coefficient by Pearson. However, a precise estimation of the copula parameters is vital in order to correctly capture the (possibly nonlinear) dependence structure and joint tail events. In this study, we employ two optimization heuristics, namely Differential Evolution and Threshold Accepting to tackle the parameter estimation of multivariate *t* distribution models in the EML approach. Since the evolutionary optimizer does not rely on gradient search, the EML approach can be applied to estimation of more complicated copula models such as high-dimensional copulas. Our experimental study shows that the proposed method provides more robust and more accurate estimates as compared to the IFM approach.

Keywords—Copula Models, Student *t* Copula, Parameter Inference, Differential Evolution, Threshold Accepting.

I. INTRODUCTION

NOWADAYS, copulas have been widely applied by market practitioners to model the dependence structure of financial risk factors, such as equity and exchange rate returns. The popularity of copulas is mainly due to their flexibility as they can be used to model both the linear and non-linear dependence structure of a multivariate distribution. The linear correlation by Pearson is not only insufficient in describing the dependence of risk factors when moving away from elliptical distributions, but also inconsistent under nonlinear strictly increasing transformations of risk factors (see McNeil *et al.* [6]). Therefore, using copula-based dependence measures will be more accurate in capturing the dependence structure than calculating the linear correlation.

However, a precise estimation of parameters in copula models is crucial to dependence modelling. In the literature, several ways based on the statistical inference theory are developed to estimate parametric and non-parametric copula models (Joe [3]). These approaches can be mainly classified into three types: (parametric) maximum likelihood estimation, semi-parametric estimation and non-parametric methods. The maximum likelihood estimations usually include the exact maximum likelihood method (EML) and the inference for margins method (IFM). The EML method could be computationally very intensive when using traditional optimization methods if the dimension of risk factors turns out to be high. Furthermore, since the EML jointly estimates the marginal distribution parameters and the dependence structure (copula) parameters, traditional optimization approaches tend to converge to local optima. Thus, Joe [3] proposed the IFM approach, a computationally simpler approach that first estimates the margin parameters and then the copula parameters. However, the estimators from the IFM method do not hold with the ones in EML in general. Due to this reason, the former set of estimates are usually used as a starting guess for the latter, leading a cumbersome procedure, i.e. a ‘two-step’ maximum likelihood.

This study compares two evolutionary methods, namely Differential Evolution (DE) and Threshold Accepting (TA) in tackling the parameter estimation of multivariate *t* copula models under the EML framework. The approach is a one-step estimation procedure, and it does not require any starting guess of the decision variables. It employs a derivative-free optimization method to overcome the curse of dimensionality problem. Therefore, the proposed approach is particularly suitable for the inference of large and complicated copula models by using EML, whereas traditional optimization procedures tend to stop at local optima in such cases. in particular, we find that DE outperforms TA in terms of stability and accuracy of estimates.

The structure of the paper is organized as follows: Section 2 introduces the copula model and the parameter inference problem. Section 3 presents the optimization problem, and the evolutionary methods which are used to solve the problem. Section 4 reports the experiment results and Section 5 concludes.

II. THE COPULA MODEL

Copulas have become an important tool in finance with various applications, e.g., risk management, derivatives pricing, portfolio management, etc. In fact, copulas were initially introduced by Sklar [10]. Let H denote a joint distribution of function with margins $F_1, ..., F_d$, then there exists an unique copula C

$$H(x_1, ...x_d) = C(F_1(x_1), ..., F_d(x_d)), \qquad (1)$$

if $F_1, ..., F_d$ are continuous. The copula model interprets multivariate distributions by coupling the marginal distribution function $F_{x_1}(x_1), ..., F_{x_d}(x_d)$ with the dependence structure C (Nelsen [8]). In other words, the joint distribution can be expressed by combining the marginal distributions with the dependence structure, yielding

$$C(u_1, ..., u_d) = H(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)), \qquad (2)$$

Jin Zhang is a COMISEF researcher appointed at the Centre for Computational Finance and Economic Agents, School of Computer Science and Electronic Engineering, University of Essex, United Kingdom. Email: jzhangf@essex.ac.uk.

Wing Lon Ng is a lecture with the Centre for Computational Finance and Economic Agents, School of Computer Science and Electronic Engineering, University of Essex, United Kingdom. Email: wlng@essex.ac.uk.

Open Science Index, Computer and Information Engineering Vol:4, No:4, 2010 publications.waset.org/3163.pdf

with $\boldsymbol{u} \in [0,1]^d$, and $F_i^{-1}(.)$ denoting the inverse of the marginal distribution $F_i(.)$. In this study, we consider the general Student *t* distribution and Student *t* copula to model the marginal distribution $F_i(.)$ and the dependence structure $C(.)$, respectively.

Particularly in finance and risk management, the Student *t* distribution has been used instead of the normal distribution, because of its fat tail behavior, which can be applied to capture financial extreme events (Bollerslev [1]). The current work mainly focuses on the parameter estimation of multivariate *t* distributions. The marginal distributions of a multivariate *t* distribution are univariate Student *t* distributions. The probability density function $f(.)$ of general Student *t* distributions can be written as

$$f^t_{\nu_m,\mu_m,\sigma_m}(x) = \frac{\Gamma(\frac{\nu_m+1}{2})}{\Gamma(\frac{\nu_m}{2})} \frac{1}{\sqrt{\nu_m \pi \sigma_m^2}} \left(1 + \frac{1}{\nu_m} \frac{(x-\mu_m)^2}{\sigma_m^2}\right)^{-\frac{\nu_m+1}{2}}, \tag{3}$$

where $\Gamma(.)$ is the gamma function, ν_m denotes the marginal degrees of freedom (DOF), μ_m and σ_m represent the location and dispersion of the marginal distribution (see Meucci [7]).

According to Sklar [10], the *t* copula of the random vector **u** can be expressed as

$$C^t_{\nu,\boldsymbol{\rho}}(\mathbf{u}) = t^d_{\nu,\rho}(t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_d)), \tag{4}$$

where $\rho_{i,j} = \Sigma_{i,j}/\sqrt{\Sigma_{i,i}\Sigma_{j,j}}$, with $i,j \in 1,...,d$. $t^d_{\nu,\rho}(.)$ denotes the distribution function $H(.)$ of $\sqrt{\nu}/\sqrt{S}\cdot\mathbf{Z}$; and $t_\nu^{-1}(.)$ represents the inverse of the marginal *t* distribution function $F_i^{-1}(.)$. The corresponding *t*-copula density $c(u_1,...,u_d) = \frac{\partial^h C(u_1,...,u_d)}{\partial u_1 ... \partial u_d}$ can be written as

$$c^t_{\nu,\rho}(u_1,...u_d) = \frac{1}{\sqrt{|\rho|}} \frac{\Gamma(\frac{\nu+d}{2})\Gamma(\frac{\nu}{2})^{d-1}}{\Gamma(\frac{\nu+1}{2})^d} \frac{\prod_{k=1}^d (1+\frac{y_k^2}{\nu})^{\frac{\nu+1}{2}}}{(1+\frac{y'\rho^{-1}y}{\nu})^{\frac{\nu+d}{2}}}. \tag{5}$$

One should note that, if the DOF ν_m of the marginal distribution of Eq. (3) is consistent with the DOF ν in the copula function in Eq. (5), the multivariate distribution is referred to as a multivariate *t* distribution (see Mcneil *et al.* [6]).

The complete copula model consists of marginal cumulative distributions $F(.)$ and a joint cumulative distribution $H(.)$, distribution parameters on the two levels are ideally estimated jointly in the exact maximum likelihood (EML) method. The log-likelihood function ℓ^m of the Student *t* marginal distribution $f(.)$ can be written as

$$\ell^m = -n_o \cdot [\, \log(\sigma_j) + \log(\sqrt{\nu_j}) + \log(\sqrt{\pi}) + \log\left(\Gamma\left(\frac{\nu_j}{2}\right)\right) + \log\left(\Gamma\left(\frac{1+\nu_j}{2}\right)\right)] \tag{6}$$

$$-\left(\frac{\nu_j+1}{2}\right) \cdot \sum_{i=1}^{n_o} \log\left(1 + \frac{(x_{j,i}-\mu_j)^2}{\sigma_j^2 \cdot \nu_j}\right), \tag{7}$$

where n_o is the observation number; and μ_j, σ_j, ν_j denote the location, dispersion and degrees of freedom of the j-th marginal distribution, respectively. The log-likelihood function ℓ^C of the *t* copula density in Eq. (5) can be written as

$$\begin{aligned} \ell^C = n_o \cdot [\, & -\frac{1}{2} \cdot \log(|\boldsymbol{\rho}|) - 2 \cdot \log\left(\Gamma\left(\frac{\nu+1}{2}\right)\right) \\ & + \log\left(\Gamma\left(\frac{\nu+2}{2}\right)\right) + \log\left(\Gamma\left(\frac{\nu}{2}\right)\right)] \\ & + \sum_{j=1}^{d} \sum_{i=1}^{n_o} \frac{\nu+1}{2} \cdot \log\left(1 + \frac{y_{j,i}^2}{\nu}\right) \\ & - \frac{\nu+2}{2} \cdot \sum_{i=1}^{n_o} \log\left[1 + \frac{1}{\nu} \boldsymbol{y_i}' \cdot \boldsymbol{\rho}^{-1} \cdot \boldsymbol{y_i}\right], \end{aligned} \tag{8}$$

where d denotes the dimension of the risk factors; $y_{j,i}$ represents the inverse transform of Student *t* with ν degrees of freedom for the j-th risk factor's observations after a strictly increasing transform, i.e., the Student *t* cumulative distribution function.

Since the EML method can be computationally burdensome if the common Newton-Raphson algorithm is applied to optimize the objective function, the literature suggests the inference for margins (IFM) approach as this two-step procedure can obtain the estimates faster – but at the cost of higher bias. Basically, the IMF approach estimates the parameters of the marginal distributions Eq. (7) first. The variables $x_{j,i}$ are transferred into $y_{j,i}$ by using the estimates, then the inference of the copula parameters in Eq. (8) is performed. IFM approach is a two-step procedure and it is implemented by using the Newton-Raphson algorithm. However, the approach cannot guarantee the parameter ν_j in Eq. (7) and the ν in Eq. (8) being consistent. EML overcomes the barrier since it estimates the marginal distributions and the copula jointly. The objective function used in the EML approach is then written as

$$\begin{aligned} \ell = -n_o \cdot \sum_{j=1}^{d} [\, & \log(\sigma_j) + \log(\sqrt{\nu}) + \log(\sqrt{\pi}) \\ & + \log\left(\Gamma\left(\frac{\nu}{2}\right)\right) + \log\left(\Gamma\left(\frac{1+\nu}{2}\right)\right)] \\ & - \sum_{j=1}^{d} \left(\frac{\nu+1}{2}\right) \cdot \sum_{i=1}^{n_o} \log\left(1 + \frac{(x_{j,i}-\mu_j)^2}{\sigma_j^2 \cdot \nu}\right) \\ & + n_o \cdot [\, -\frac{1}{2} \cdot \log(|\boldsymbol{\rho}|) - 2 \cdot \log\left(\Gamma\left(\frac{\nu+1}{2}\right)\right) \\ & + \log\left(\Gamma\left(\frac{\nu+2}{2}\right)\right) + \log\left(\Gamma\left(\frac{\nu}{2}\right)\right)] \\ & + \sum_{j=1}^{d} \sum_{i=1}^{n_o} \frac{\nu+1}{2} \cdot \log\left(1 + \frac{y_{j,i}^2}{\nu}\right) \\ & - \frac{\nu+2}{2} \cdot \sum_{i=1}^{n_o} \log\left[1 + \frac{1}{\nu} \boldsymbol{y_i}' \cdot \boldsymbol{\rho}^{-1} \cdot \boldsymbol{y_i}\right]. \end{aligned} \tag{9}$$

III. THE OPTIMIZATION PROBLEM

The estimation of the parameters is based on the maximization of an objective function, i.e. the log-likelihood function from the copula density function. Ideally, the parameters of the cumulative distribution function should be estimated jointly

Open Science Index, Computer and Information Engineering Vol:4, No:4, 2010 publications.waset.org/3163.pdf

with parameters of the copula model. The fitness of the final objective function is defined as the sum of log-likelihood values of both the marginal and copula density functions in Eq. (9). In the multivariate *t* copula model, the marginal distributions and copula model have a same DoF ν. The fitness value of the objective function $\mathcal{O}$ depends on μ_m, σ_m^2, ρ and ν. The optimization problem can be briefly described as:

$$\max_{\mu_{\mathbf{m}},\sigma_{\mathbf{m}},\rho,\nu} \mathcal{O} = \ell$$

subject to

$$1 > \rho > -1, \quad \nu > 3. \tag{10}$$

To estimate the copula model, we adopt a population based evolutionary method to optimize parameters of copula models while taking the marginal distribution and dependent structure into account simultaneously.

A. Differential Evolution

Heuristic methods provide ways of tackling combinatorial optimization problems. Differential Evolution (DE) which was originally proposed by Storn and Price [11], is a population based heuristic method for solving the optimization problems with continuous space. The approach generates new solutions by linear combination and cross-over based on current solutions. The resulting solution would replace the current best solution if the new solution has a higher fitness value. For each current solution $\boldsymbol{\imath_p}$, a new solution $\boldsymbol{\imath_c}$ is generated from the following procedure: randomly selecting three different members from the current population ($p_1 \neq p_2 \neq p_3 \neq p$); linearly combining the solution vectors at probability π_1, or inheriting the original p-th solution otherwise. We use the 'Ditter' and 'Jitter' version of the standard DE (see Price *et al.* [9]), which considers adding normally distributed random numbers to the weighting factor K, and the difference of two solution vectors respectively. Vectors $\boldsymbol{z_1}$ and $\boldsymbol{z_2}$ represent the extra noise in the algorithm; they contain random numbers being zero with probability $\boldsymbol{\pi_2}$ and $\boldsymbol{\pi_3}$ respectively, or independently follow the normal distributions $N(0, \mathfrak{d}_1^2)$ and $N(0, \mathfrak{d}_2^2)$. The DE algorithm is described by the pseudo code in Algorithm 1. π_1 is the cross-over probability. After the linear combination and cross-over, DE updates the population. More specifically, if the fitness value of $\boldsymbol{\imath_c}$ is higher than the one of $\boldsymbol{\imath_p}$, the solution $\boldsymbol{\imath_p}$ is replaced by $\boldsymbol{\imath_c}$, and the updated $\boldsymbol{\imath_p}$ exists in the current population; otherwise $\boldsymbol{\imath_p}$ survives.

The technical parameters of DE algorithm used in this study are listed as follows. Population size and iteration number were set at 50 and 500; the value of K was set at a value 0.5; and the cross-over probability π_1 was at 60%. The parameters used for the extra noise were: $\pi_2 = 50\%$, $\pi_3 = 10\%$, $\mathfrak{d}_1^2 = 0.1$ and $\mathfrak{d}_2^2 = 0.1$. We use the dotted parameters to represent the chromosomes of the population. Repair functions are used to translate the chromosomes to the solutions which meet the constraints:

$$\nu = 3 + |\dot{\nu}| \tag{11}$$

Algorithm 1 Differential Evolution.

```
randomly initialize population of vectors $\boldsymbol{\imath_p}$, $p$= 1,...,$P$
**while** the halting criterion is not met **do**
    **for** all current solutions $\boldsymbol{\imath_p}$, $p$=1,...,$P$ **do**
        randomly pick three different solutions, i.e. $p_1 \neq p_2 \neq p_3 \neq p$
        $\boldsymbol{\imath_c}[i] \leftarrow \imath_{p_1}[i] + (K + z_1[i])(\imath_{p_2}[i] - \imath_{p_3}[i] + z_2[i])$ with probability $\pi_1$, or $\imath_c[i] \leftarrow \imath_p[i]$ otherwise
        compute the fitness value of $\boldsymbol{\imath_p}$, i.e. the sum of log-likelihood value of the marginal and copula density functions
    **end for**
    **for** the current solution $\boldsymbol{\imath_p}$, $p$ = 1,...,$P$ **do**
        **if** Fitness($\boldsymbol{\imath_c}$) $>$ Fitness($\boldsymbol{\imath_p}$) **then** $\boldsymbol{\imath_p} \leftarrow \boldsymbol{\imath_c}$ **end if**
    **end for**
**end while**
```

Algorithm 2 Threshold Accepting.

```
Initialize threshold sequence $\mathfrak{T}_i$, with $i = 1, ..., n_s$ and
Randomly generate an initial solution $\boldsymbol{\imath^c}$
**for** $i = 1$ to $n_s$ **do**
    Choose $\boldsymbol{\imath^n} \in \mathcal{N}(\boldsymbol{\imath^c})$
    **if** $\mathcal{O}(\boldsymbol{\imath^n}) - \mathcal{O}(\boldsymbol{\imath^c}) < \mathfrak{T}_i$, **then**
    $\boldsymbol{\imath^c} \leftarrow \boldsymbol{\imath^n}$
    **end if**
**end for**
```

$$\rho = \begin{cases} \exp \frac{-0.15}{|\dot{\rho}|}, & \text{if } \dot{\rho} > 0 \\ -\exp \frac{-0.15}{|\dot{\rho}|}, & \text{if } \dot{\rho} < 0 \end{cases} \tag{12}$$

The idea of the above repair mechanism is proposed by Maringer and Oyewumi [5], who apply a similar idea to translate solutions from DE to valid asset weights for index tracking problems.

B. Threshold Accepting

Dueck and Scheuer [2] introduced threshold accepting (TA) which is a heuristics approach analog to simulated annealing (SA) (see Kirkpatrick *et al.* [4]). Winker [12] gave a comprehensive introduction of TA and its applications in economics. TA is a refined version of the standard local search procedure, mainly it differs from the standard approaches in its acceptance criterion. Given a minimization problem, let $\boldsymbol{\imath^c}$ denote an initial element and $\boldsymbol{\imath^n}$ represent a new element in the neighborhood $\mathcal{N}(\boldsymbol{\imath^c})$, TA will accept $\boldsymbol{\imath^n}$ as a new current solution if and only if the solution is better than $\boldsymbol{\imath^c}$ in terms of objective function, i.e. $\mathcal{O}(\boldsymbol{\imath^n}) - \mathcal{O}(\boldsymbol{\imath^n}) < \mathfrak{T}$ for some preassigned non-negative threshold value $\mathfrak{T}$, and the threshold $\mathfrak{T}$ is decreased gradually and reaches the value of zero after a given number of steps n_s.

To generate a new solution in the neighborhood of $\boldsymbol{\imath^c}$, we add normally distributed randomness from $N(0, \mathfrak{d}_3^2)$ to each molecule in the chromosome of the $\boldsymbol{\imath^c}$ at a probability of π_4, otherwise the molecule of the new solution $\boldsymbol{\imath^n}$ inherits the one of $\boldsymbol{\imath^c}$. The parameter $\mathfrak{d}_3$ and π_4 are assigned with values of 0.1 and 0.5. The sequence of threshold $\mathfrak{F}_i, i = 1, ..., n_s$ is decided by using a data-driven approach, which is suggested by Winker [12] as a standard approach in deciding the sequence. The empirical distribution of the distance between a number of

Open Science Index, Computer and Information Engineering Vol:4, No:4, 2010 publications.waset.org/3163.pdf

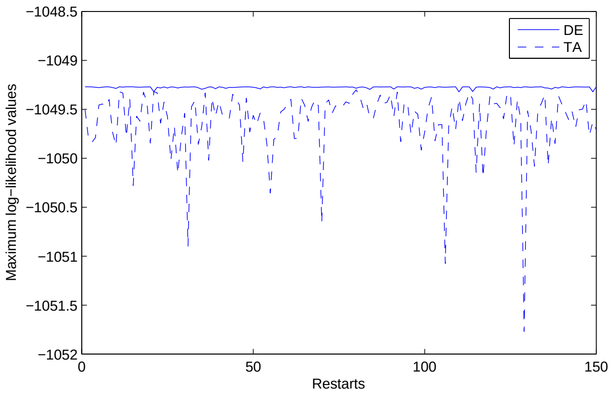

Fig. 1. Log-Likelihood values of DE and TA

randomly chosen solutions and their corresponding neighborhood is computed, and the distance is defined as the absolute difference in the values of the objective function which are computed by using the solution and its neighborhood. The empirical distance then is sorted in decreasing order, and the i-th quantile of the sorted distance are use as the threshold for the i-th step. The mapping functions described in the last subsection are used to satisfy the constraints. The TA approach is described by Algorithm 2.

IV. EXPERIMENTAL RESULTS

First, we compare the log-likelihood values and the estimates of the copula model using the DE and the TA in the optimisation procedure. We independently restart the two algorithms 150 times to obtain the distributions of the fitness value and the parameter estimates using a same data sample. The fitness values i.e., the log-likelihood values defined in Eq. (9) are provided in Figure 1. The figure shows that the DE yields better and more stable fitness values than the TA. As Winker [12] pointed out that the performance of TA highly depends on (a) the construction of local structure, (b) the threshold sequence. In order to have a successful implementation of TA, the solution space of the specified problem should be explored in finding a proper local structure and a suitable threshold sequence. Regarding DE, the algorithm usually only requires a fine tune of the parameters K and π_1.

It has been shown that DE works better than TA for the above problem. Therefore, we employ DE as the EML optimizer in the following Monte-Carlo simulation experiments. To assess the performance of the proposed EML estimation, we first simulate a set of 200×2 random variables with bivariate student t distribution at a total iteration number of $I_n = 5,000$. The true distribution parameters are set as $\mu_1 = 0$, $\mu_2 = 0, \sigma_1 = 0.2548, \sigma_2 = 0.2250, \rho = 0.43$ and $\nu = 6$. After that, we estimate the parameters of the bivariate t-copula with t margins by using the proposed approach and a traditional approach respectively. Since the standard hill-climbing algorithm such as the Newton-Raphson approach for the EML method did not generate any results but only for the IFM framework, we only compare the latter one with the DE procedure applied on EML (see also Joe (1997)). Table I shows the numerical results with the standard descriptive statistics of the estimated parameters in the N bootstrap samples. As expected, the EML estimators, obtained by maximizing the log-likelihood function with the DE, are often (a) closer to the true values, (b) less biased, (c) less skewed and (d) less kurtotic as compared to the IFM alternatives.

TABLE I
COMPARISON OF THE SAMPLE MOMENTS OF THE BOOTSTRAPPED PARAMETERS

		EML				
θ	True Value		Mean$(\hat{\theta})$	St.dev.$(\hat{\theta})$	Skewness$(\hat{\theta})$	Kurtosis$(\hat{\theta})$
μ_1	0.0000		0.0000	0.0105	0.0390	2.9275
μ_2	0.0000		0.0000	0.0092	0.0051	3.0494
σ_1	0.2548		0.2551	0.0095	0.0080	3.0637
σ_2	0.2252		0.2255	0.0086	0.0863	3.0170
ρ	0.4300		0.4374	0.0330	-0.0950	3.0309
ν	6.0000		6.2589	1.1123	1.2754	6.2565
		IMF				
θ	True Value		Mean$(\hat{\theta})$	St.dev.$(\hat{\theta})$	Skewness$(\hat{\theta})$	Kurtosis$(\hat{\theta})$
μ_1	0.0000		0.0000	0.0106	0.0397	2.9311
μ_2	0.0000		0.0000	0.0093	0.0103	3.0495
σ_1	0.2548		0.2554	0.0111	0.0453	3.0506
σ_2	0.2252		0.2257	0.0100	0.0966	3.0502
ρ	0.4300		0.4375	0.0336	-0.0620	3.0299
ν	6.0000		6.6798	2.7531	4.8531	59.7094

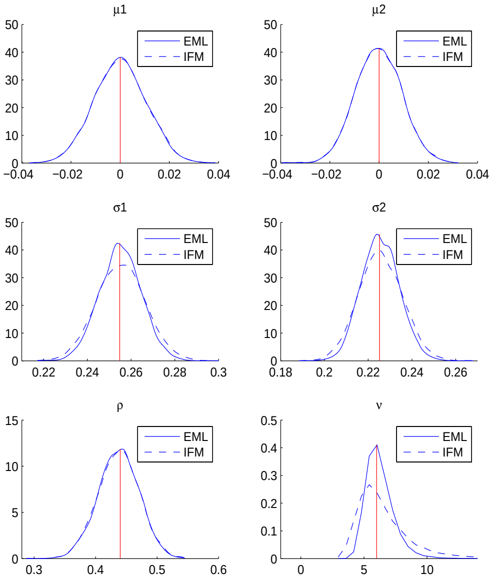

Fig. 2. Kernel densities of the estimated copula parameters from the EML and IFM

Figure 2 compares the kernel densities of the estimated distribution parameters for both estimation procedures. As discernible, the differences between the distribution of estimators for μ_1, μ_2 (top panels) and ρ (bottom left panel) are negligible. More interestingly, the middle panels reveal that the dispersion parameters σ_1 and σ_2 can be more accurately and more efficiently estimated with the EML approach as their kernel densities are higher in the centered region and lower in the tail regions. Finally, the bottom right panel shows that the estimate for the degree of freedom ν that controls the probability

Open Science Index, Computer and Information Engineering Vol:4, No:4, 2010 publications.waset.org/3163.pdf

of tail events in the distribution is even biased in the IFM method as the peak of its kernel density is not localized at the true parameters position (as indicated by the blue line). In total, it can be seen that the parameters responsible centered moments of the distribution, i.e. σ_1, σ_2 and ν can be better estimated with the EML approach. These results are indeed essential as they reveal that the IFM approach often preferred in the financial literature is more likely to provide less reliable estimators of the underlying joint distribution and, hence, less able to correctly capture the dependence structure and, more importantly, the tail dependence of risk factors (i.e., the extreme losses).

V. CONCLUSION

This paper suggests implementing an evolutionary algorithm in the exact maximum likelihood estimation of multivariate copula models as standard hill-climbing procedure. Usually, the standard Newton-Raphson algorithm fails to optimize the objective function when the number of dimensions turns out to be high, while a derivative-free optimizer can overcome this problem. We employ two evolutionary approaches, namely Differential Evolution and Threshold Accepting in the EML approach, and we find that the former works better. Through a simple Monte-Carlo simulation study, we show that the proposed methodology already provide reasonably good results in a straightforward 2-dimensional setting with a bivariate Student t-copula. As expected, the estimates obtained by the EML approach enhanced with the heuristics approaches are often closer to the true values as compared to the IFM alternatives. Furthermore, the Differential Evolution is competent for the EML inference of more complicated copula models than the bivariate Student t copula studied.

ACKNOWLEDGMENT

The authors gratefully acknowledge financial support from CSEE University of Essex and the EU Commission through MRTN-CT-2006-034270 COMISEF.

REFERENCES

[1] T. Bollerslev. A conditional heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics*, 69:542–547, 1987.

[2] G. Dueck and T. Scheuer. Threshold Accepting: a general purpose optimization algorithm appearing superior to Simulated annealing. *Journal of Computational Physics*, 90:161–175, 1990.

[3] Harry Joe. *Multivariate Models and Dependence Concepts*. Chapman & Hall/CRC, 1997.

[4] S. Kirkpatrick, C. Gelatt, and M. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[5] Dietmar Maringer and Olufemi Oyewumi. Index tracking with constrained portfolios. *Intelligent Systems in Accounting and Finance Management*, 15(1):51–71, 2007.

[6] A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management*. Princeton Series in Finance. Princeton University Press, 2005.

[7] Attilio Meucci. *Risk and Asset Allocation*. Springer, 2005.

[8] Roger B Nelsen. *An Introduction to Copulas*. Springer, 1998.

[9] Kenneth V. Price, Rainer M. Storn, and Jouni A. Lampinen. *Differential Evolution: A Practical Approach to Global Optimization*. Springer, 1998.

[10] A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de 1'institut de statistique de 1'Université de Paris*, 8:229–231, 1959.

[11] Rainer Storn and Kenneth Price. Differential Evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359, 1997.

[12] Peter Winker. *Optimization Heuristics In Econometrics: Applications of Threshold Accepting*. John Wiley & Sons, 2001.