A New Breed of Copulas for Risk and Portfolio Management

Attilio Meucci[1]

published[2]: September 1 2011
this revision: August 26 2011
latest revision and code: http://symmys.com/node/335

Abstract

We introduce the copula-marginal algorithm (CMA), a commercially viable technique to generate and manipulate a much wider variety of copulas than those commonly used by practitioners.

CMA consists of two steps: separation, to decompose *arbitrary* joint distributions into their copula and marginals; and combination, to glue *arbitrary* copulas and marginals into new joint distributions.

Unlike traditional copula techniques, CMA a) is not restricted to few parametric copulas such as elliptical or Archimedean; b) never requires the explicit computation of marginal cdf's or quantile functions; c) does not assume equal probabilities for all the scenarios, and thus allows for advanced techniques such as importance sampling or entropy pooling; d) allows for arbitrary transformations of copulas.

Furthermore, the implementation of CMA is also computationally very efficient in arbitrary large dimensions.

To illustrate benefits and applications of CMA, we propose two case studies: stress-testing with a panic copula which hits non-symmetrically the downside and displays non-equal, risk-premium adjusted probabilities; and arbitrary rotations of the panic copula.

Documented code for the general algorithm and for the applications of CMA is available at http://symmys.com/node/335.

JEL Classification: C1, G11

Keywords: panic copula, copula transformations, Archimedean, elliptical, Student t, non-parametric, scenarios-probabilities, empirical distribution, entropy pooling, importance sampling, grade, unit cube.

[1] The author is grateful to Garli Beibi
[2] This article appears as Meucci, A. (2011) "A New Breed of Copulas for Risk and Portfolio Management" *Risk*, **24**, 9, 122-126

1 Introduction

The multivariate distribution of a set of risk factors such as stocks returns, interest rates or volatility surfaces is fully specified by the separate marginal distributions of the factors and by their copula or, loosely speaking, the correlations among the factors.

Modeling the marginals and the copula separately provides greater flexibility for the practitioner to model randomness. As a result, copulas have been used extensively in finance, both on the sell-side to price derivatives, see e.g. Li (2000), and on the buy-side to model portfolio risk, see e.g. Meucci, Gan, Lazanas, and Phelps (2007).

In practice, a large variety of marginal distributions can be modeled by parametric or non-parametric specifications. However, unlike for marginal distributions, despite the wealth of theoretical results on copulas, only a few parametric families of copulas, such as elliptical or Archimedean, are used in practice in commercial applications.

Here we introduce a technique, which we call "copula-marginal algorithm" (CMA) to generate and use in practice new, extremely flexible copulas. CMA enables us to extract the copulas and the marginals from *arbitrary* joint distributions; to perform *arbitrary* transformations of those extracted copulas; and then to glue those transformed copulas back with another set of *arbitrary* marginal distributions.

This flexibility follows from the fact that, unlike traditional approaches to copulas implementation, CMA does not require the explicit computation of marginal cdf's and their inverses. As a result, CMA can generate scenarios for many more copulas than the few parametric families used in the traditional approach. For instance, it includes large-dimensional, downside-only panic copulas which can be coupled with, say, extreme value theory marginals for portfolio stress-testing.

An additional benefit of CMA is that it does not assume that all the scenarios have equal probabilities.

Finally, CMA is computationally very efficient even in large markets, as can be verified in the code available for download.

We summarize in the table below the main differences between CMA and the traditional approach to apply the theory of copulas in practice

	Copula	Marginals	Probabilities
Traditional	parametric	flexible	equal
CMA	flexible	flexible	flexible

(1)

In Section 2 we review the basics of copula theory. In Section 3 we discuss the traditional approaches to copula implementation. In Section 4 we introduce CMA in full generality. In Section 5 we present a first application of CMA: we create a panic copula for stress-testing that hits non-symmetrically the downside and is probability-adjusted for risk premium. In Section 6 we discuss a second application of CMA, namely how to perform arbitrary transformations of copulas.

Documented code for the general algorithm and for the applications of CMA is available at http://symmys.com/node/335.

2 Review of copula theory

The two-step theory of copulas is simple and powerful. For much more review on the subject, the reader is referred to articles such as Embrechts, A., and Straumann (2000), Durrleman, Nikeghbali, and Roncalli (2000), Embrechts, Lindskog, and McNeil (2003), or monographs such as to Nelsen (1999), Cherubini, Luciano, and Vecchiato (2004), Brigo, Pallavicini, and Torresetti (2010) and Jaworski, Durante, Haerdle, and Rychlik (2010). For a concise, visual primer with all the main results and proofs see Meucci (2011).

Consider a set of N joint random variables $\mathbf{X} \equiv (X_1, \ldots, X_N)'$ with a given joint distribution which we represent in terms of the cdf

$$F_{\mathbf{X}}(x_1, \ldots, x_N) \equiv \mathbb{P}\{X_1 \leq x_1, \ldots, X_N \leq x_N\}. \tag{2}$$

We call the first step "separation". This step separates the distribution $F_{\mathbf{X}}$ into the pure "individual" information contained in each variable X_n, i.e. the marginals F_{X_n}, and the pure "joint" information of all the entries of $\mathbf{X}$, i.e. the copula $F_{\mathbf{U}}$. The copula is the joint distribution of the grades, i.e. the random variables $\mathbf{U} \equiv (U_1, \ldots, U_N)'$ defined by feeding the original variables X_n into their respective marginal cdf

$$U_n \equiv F_{X_n}(X_n), \qquad n = 1, \ldots, N. \tag{3}$$

Each grade U_n has a uniform distribution on the interval $[0, 1]$ and thus it can be interpreted as a "non-linear z-score" of the original variables X_n which lost all the "individual" information of the distribution of X_n and only preserved its joint information with other X_m's. To summarize, the separation step $\mathcal{S}$ proceeds as follows

$$\mathcal{S}: \qquad \begin{pmatrix} X_1 \\ \vdots \\ X_N \end{pmatrix} \sim F_{\mathbf{X}} \quad \mapsto \quad \begin{cases} F_{X_1}, \ldots, F_{X_N} \\ \begin{pmatrix} U_1 \\ \vdots \\ U_N \end{pmatrix} \sim F_{\mathbf{U}} \end{cases} \tag{4}$$

The above separation step can be reverted by a second step, which we call "combination". We start from arbitrary marginal distributions $\bar{F}_{X_n}$, in general different from the above F_{X_n}, and grades $\mathbf{U} \equiv (U_1, \ldots, U_N)'$ distributed according to a chosen arbitrary copula $\bar{F}_{\mathbf{U}}$, which can, but does not need to, be obtained by separation as the above $F_{\mathbf{U}}$. Then we combine the marginals $\bar{F}_{X_n}$ and the copula $\bar{F}_{\mathbf{U}}$ into a new joint distribution $\bar{F}_{\mathbf{X}}$ for $\mathbf{X}$. To do so, for each marginal $\bar{F}_{X_n}$ we first compute the inverse cdf $\bar{F}_{X_n}^{-1}$, or quantile, and then we apply the inverse cdf to the respective grade from the copula

$$X_n \equiv \bar{F}_{X_n}^{-1}(U_n), \qquad n = 1, \ldots, N. \tag{5}$$

To summarize, the combination step $\mathcal{C}$ proceeds as follows

$$\mathcal{C}: \qquad \left. \begin{array}{c} \bar{F}_{X_1}, \ldots, \bar{F}_{X_N} \\ \\ \begin{pmatrix} U_1 \\ \vdots \\ U_N \end{pmatrix} \sim \bar{F}_{\mathbf{U}} \end{array} \right\} \quad \mapsto \quad \begin{pmatrix} X_1 \\ \vdots \\ X_N \end{pmatrix} \sim \bar{F}_{\mathbf{X}} \tag{6}$$

3 Traditional copula implementation

In general, the separation step (4) and the combination step (6) cannot be performed analytically. Therefore, in practice, one resorts to Monte Carlo scenarios.

In the traditional implementation of the separation step (4), first of all we select a parametric N-variate joint distribution $F^{\theta}_{\mathbf{X}}$ to model $\mathbf{X} \equiv (X_1, \ldots, X_N)'$, whose marginal distributions $F^{\theta}_{X_n}$ can be represented *analytically*. Then we draw J joint Monte Carlo scenarios $\{x_{1,j}, \ldots, x_{N,j}\}_{j=1,\ldots,J}$ from $F^{\theta}_{\mathbf{X}}$. Next, we compute the marginal cdf's $F^{\theta}_{X_n}$ from their analytical representation. Then, the joint scenarios for X are mapped as in (3) into joint grade scenarios by means of the respective marginal cdf's

$$u_{n,j} \equiv F^{\theta}_{X_n}(x_{n,j}), \qquad n = 1, \ldots, N; \quad j = 1, \ldots, J. \tag{7}$$

The grades scenarios $\{u_{1,j}, \ldots, u_{N,j}\}_{j=1,\ldots,J}$ now represent simulations from the parametric copula $F^{\theta}_{\mathbf{U}}$ of X.

To illustrate the traditional implementation of the separation, $F^{\theta}_{\mathbf{X}}$ can be normal, and the scenarios $x_{n,j}$ can be simulated by twisting N independent standard normal draws by the Cholesky decomposition of the covariance and adding the expectations. The marginals of the joint normal distribution are normal, and the normal cdf's $F^{\theta}_{X_n}$ are computed by quadratures of the normal pdf. Then the scenarios for the normal copula follow from (7).

We can summarize the traditional implementation of the separation step as follows

$$\mathcal{S}: \qquad \left\{ \begin{pmatrix} x_{1,j} \\ \vdots \\ x_{N,j} \end{pmatrix} \right\} \sim F^{\theta}_{\mathbf{X}} \quad \mapsto \quad \left\{ \begin{array}{c} F^{\theta}_{X_1}, \ldots, F^{\theta}_{X_N} \\ \\ \left\{ \begin{pmatrix} u_{1,j} \\ \vdots \\ u_{N,j} \end{pmatrix} \right\} \sim F^{\theta}_{\mathbf{U}} \end{array} \right. \tag{8}$$

where for brevity we dropped the subscript $j = 1, \ldots, J$ from the curly brackets, displaying only the generic j-th joint N-dimensional scenario.

In the traditional implementation of the combination (6), we first generate scenarios from the desired copula $\bar{F}^{\theta}_{\mathbf{U}}$, typically obtained via a parametric separation step, i.e. $\bar{F}^{\theta}_{\mathbf{U}} \equiv F^{\theta}_{\mathbf{U}}$ and thus $\bar{u}_{n,j} \equiv u_{n,j}$. Then we specify the desired marginal distributions, typically parametrically $\bar{F}^{\theta}_{X_n}$, and we compute analytically or by quadratures the inverse cdf's $\bar{F}^{\theta^{-1}}_{X_n}$. Then we feed as in (5) each grade scenario $\bar{u}_{n,,j}$ into the respective quantiles

$$\bar{x}_{n,j} \equiv \bar{F}^{\theta^{-1}}_{X_n}(\bar{u}_{n,j}), \qquad n = 1, \ldots, N; \quad j = 1, \ldots, J \tag{9}$$

The joint scenarios $\{\bar{x}_{1,j}, \ldots, \bar{x}_{N,j}\}_{j=1,\ldots,J}$ display the desired copula $\bar{F}^{\theta}_{\mathbf{U}}$ and marginals $\bar{F}^{\theta}_{X_n}$.

To illustrate the traditional implementation of the combination, we can use the previously obtained normal copula scenarios and combine them with, say, chi-square marginals with different degrees of freedom, giving rise to a multivariate correlated chi-square distribution.

We can summarize the traditional implementation of the combination step as follows

$$\mathcal{C}: \quad \left. \begin{array}{c} \bar{F}^{\theta}_{X_1}, \ldots, \bar{F}^{\theta}_{X_N} \\ \\ \{\begin{pmatrix} \bar{u}_{1,j} \\ \vdots \\ \bar{u}_{N,j} \end{pmatrix}\} \sim \bar{F}^{\theta}_{\mathbf{U}} \end{array} \right\} \mapsto \quad \{\begin{pmatrix} \bar{x}_{1,j} \\ \vdots \\ \bar{x}_{N,j} \end{pmatrix}\} \sim \bar{F}^{\theta}_{\mathbf{X}}. \tag{10}$$

In practice, only a handful of parametric joint distributions is used to obtain the copula scenarios that appear in (8) and (10), because in general it is impossible to compute the marginal cdf's and thus perform the transformations $u_{n,j} \equiv F_{X_n}(x_{n,j})$ in (3) and (7). As a result, practitioners resort to elliptical distributions such as normal or Student t, or a few isolated tractable distributions for which the cdf's are known, such as in Daul, De Giorgi, Lindskog, and McNeil (2003).

An alternative approach proposed to broaden the choice of copulas involves simulating the grades scenarios $\bar{u}_{n,j}$ in (10) directly from a parametric copula $\bar{F}^{\theta}_{\mathbf{U}}$, without obtaining them from a separation step (8). However, the parametric specifications that allow for direct simulation are limited to the Archimedean family, see Genest and Rivest (1993), and few other extensions. Furthermore, the parameters of the Archimedean family are not immediate to interpret. Finally, simulating the grades scenarios $\bar{u}_{n,j}$ from the Archimedean family when the dimension N is large is computationally challenging.

To summarize, only a restrictive set of parametric copulas is used in practice, whether they stem from parametric joint distributions or they are simulated directly from parametric copula specifications. CMA intends to greatly extends the set of copulas that can be used in practice.

4 The copula-marginal algorithm (CMA)

Unlike the traditional approach, CMA does *not* require the analytical representation of the marginals that appear in theory in (3) and in practice in (7). Instead, we construct these cdf's non-parametrically from the joint scenarios for $F_{\mathbf{X}}$. Then, it becomes easy to extract the copula. This allows us to start from arbitrary parametric or non parametric joint distributions $F_{\mathbf{X}}$ and thus achieve much higher flexibility.

Even better, CMA allows us to extract both the marginal cdf's and the copula from distributions that are represented by joint scenarios with fully general, non-equal probabilities. Therefore, we can include distributions F_X obtained from advanced Monte Carlo techniques such as importance sampling, see Glasserman (2004); or from posterior probabilities driven by the Entropy Pooling approach, see Meucci (2008); or from "Fully Flexible Probabilities" as in Meucci (2010).

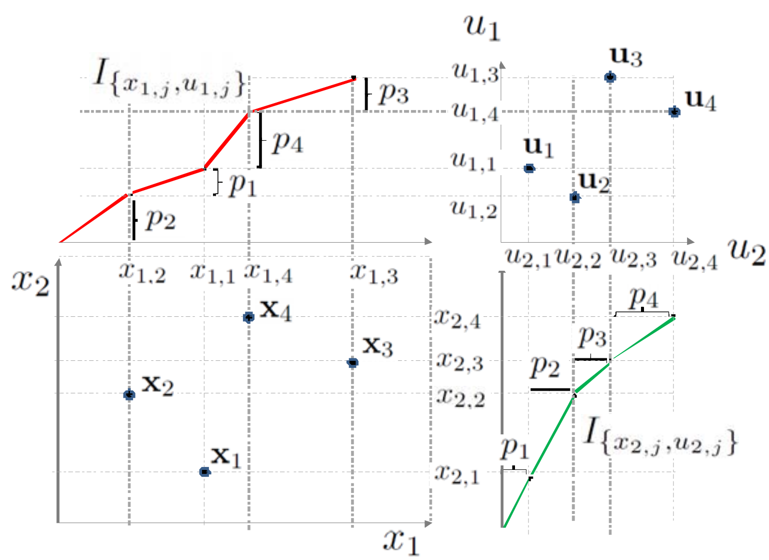

Figure 1: Copula-Marginal Algorithm: separation

Let us first discuss the separation step $\mathcal{S}$. For this step, CMA takes as input the scenarios-probabilities representation $\{x_{1,j},\ldots,x_{N,j};p_j\}_{j=1,\ldots,J}$ of a fully general distribution $F_{\mathbf{X}}$, see Figure 1, where we display a $N=2$-variate distribution with $J=4$ scenarios. With this input, CMA computes the grades scenarios $u_{n,j}$ as the probability-weighted empirical grades

$$u_{n,j} \equiv \sum_{i=1}^{J} p_i 1_{x_{n,i}\leq x_{n,j}}, \qquad n=1,\ldots,N; \quad j=1,\ldots,J, \tag{11}$$

where 1_A denotes the indicator function for the generic statement A, which is equal to 1 if A is true and 0 otherwise, refer again to Figure 1.

With the grades scenarios (11) we are ready to separate both the copula and the marginals in the distribution $F_{\mathbf{X}}$.

For the copula, we associate the probabilities p_j of the original scenarios $x_{n,j}$ with the grade scenarios $u_{n,j}$. As it turns out, the copula $F_{\mathbf{U}}$, i.e. the joint distribution of the grades, is given by the scenarios-probabilities $\{u_{1,j},\ldots,u_{N,j};p_j\}_{j=1,\ldots,J}$.

For the marginal distributions, as in Meucci (2006) CMA creates from each grid of scenarios pairs $\{x_{n,j}, u_{n,j}\}$ a function $I_{\{x_{n,j},u_{n,j}\}}$ that inter/extra-polates those values, see Figure 1. This function is now the cdf of the generic n-th variable

$$F_{X_n}(x) \equiv I_{\{x_{n,j},u_{n,j}\}}(x), \qquad n = 1, \ldots, N. \tag{12}$$

To summarize, the CMA separation step attains from the distribution $F_{\mathbf{X}}$ the scenarios-probabilities representation of the copula $F_{\mathbf{U}}$ and the inter/extra-polation representation of the marginal cdf's $F_{X_n} \equiv I_{\{x_{n,j},u_{n,j}\}}$ as follows

$$\mathcal{S}_{\text{CMA}}: \qquad \{(\begin{matrix} x_{1,j} \\ \vdots \\ x_{N,j} \end{matrix}); p_j\} \sim F_{\mathbf{X}} \quad \mapsto \quad \begin{cases} I_{\{x_{1,j},u_{1,j}\}}, \ldots, I_{\{x_{N,j},u_{N,j}\}} \\ \\ \{(\begin{matrix} u_{1,j} \\ \vdots \\ u_{N,j} \end{matrix}); p_j\} \sim F_{\mathbf{U}} \end{cases} \tag{13}$$

Notice that CMA avoids the parametric cdf's $F_{X_n}^{\boldsymbol{\theta}}$ that appear in (7).

Let us now address the combination step $\mathcal{C}$. The two inputs are an arbitrary copula $\overline{F}_{\mathbf{U}}$ and arbitrary marginal distributions, represented by the cdf's $\overline{F}_{X_n}$. For the copula, we take any copula $\bar{F}_{\mathbf{U}}$ obtained with the separation step, i.e. a set of scenarios-probabilities $\{\bar{u}_{1,j}, \ldots, \bar{u}_{N,j}; \bar{p}_j\}$. For the marginals, we take any parametric or non-parametric specification of the cdf's $\bar{F}_{X_n}$. Then for each n we construct, in one of a few ways discussed in the appendix, a grid of significant points $\{\tilde{x}_{n,k}, \tilde{u}_{n,k}\}_{i=1,\ldots,K_n}$, where $\tilde{u}_{n,k} \equiv \bar{F}_{X_n}(\tilde{x}_{n,k})$. Then, CMA takes each grade scenario for the copula $\bar{u}_{n,j}$ and maps it into the desired combined scenarios $\bar{x}_{n,j}$ by inter/extra-polation of the copula scenarios $\bar{u}_{n,j}$ on the grid in a manner similar to (12), but reversing the axes

$$\bar{x}_{n,j} \equiv I_{\{\tilde{u}_{n,k},\tilde{x}_{n,k}\}}(\bar{u}_{n,j}), \qquad n = 1, \ldots, N; \quad j = 1, \ldots, J. \tag{14}$$

Notice that the inter/extra-polation (14) replaces the computation of the inverse cdf $\bar{F}_{X_n}^{-1}$ that appears in (5) and in (9).

To summarize, the CMA combination step achieves the scenarios-probabilities representation of the joint distribution $\bar{F}_{\mathbf{X}}$ that glues the copula $\bar{F}_{\mathbf{U}}$ with the marginals $\bar{F}_{X_n}$ as follows

$$\mathcal{C}_{\text{CMA}}: \qquad \left.\begin{matrix} \bar{F}_{X_1}, \ldots, \bar{F}_{X_N} \\ \\ \{(\begin{matrix} \bar{u}_{1,j} \\ \vdots \\ \bar{u}_{N,j} \end{matrix}); \bar{p}_j\} \sim \bar{F}_{\mathbf{U}} \end{matrix}\right\} \quad \mapsto \quad \{(\begin{matrix} \bar{x}_{1,j} \\ \vdots \\ \bar{x}_{N,j} \end{matrix}); p_j\} \sim \bar{F}_{\mathbf{X}} \tag{15}$$

From a computational perspective, both the separation step (13) and the combination step (15) are extremely efficient, as they run in fractions of a second even in large markets with very large numbers of scenarios. Please refer to the code available at http://symmys.com/node/335 and the appendix for more details.

5 Case study: panic copula

Here we apply CMA to generate a large dimensional panic copula for stress-testing and portfolio optimization. The code for this case study is available at http://symmys.com/node/335.

Consider a N-dimensional vector of financial random variables $\mathbf{X} \equiv (X_1, \ldots, X_N)'$, such as the yet to be realized returns of the $N = 500$ stocks in the S&P 500. Our aim is to construct a panic stress-test joint distribution $\bar{F}_{\mathbf{X}}$ for $\mathbf{X}$.

To do so, we first introduce a distribution $F_{\mathbf{X}}$ which is driven by two separate sets of random variables $\mathbf{X}^{(c)}$ and $\mathbf{X}^{(p)}$, representing the calm market and the panic-stricken market. From $F_{\mathbf{X}}$ we will extract the panic copula, which we will then glue with marginal distributions fitted to empirical data.

Accordingly, we first define the joint distribution $F_{\mathbf{X}}$ with a few components, as follows

$$\mathbf{X} \stackrel{d}{\equiv} (\mathbf{1}_N - \mathbf{B}) \circ \mathbf{X}^{(c)} + \mathbf{B} \circ \mathbf{X}^{(p)}, \tag{16}$$

where $\mathbf{1}_N$ is a N-dimensional vector of ones and the operation $\circ$ multiplies vectors term-by-term. The first component, $\mathbf{X}^{(c)} \equiv (X_1^{(c)}, \ldots, X_N^{(c)})'$ are the calm-market drivers, which are normally distributed with expectation a N-dimensional vector of zeros 0_N and correlation matrix $\boldsymbol{\rho}$

$$X^{(c)} \sim \mathrm{N}\left(0_N, \boldsymbol{\rho}\right). \tag{17}$$

The second component, $\mathbf{X}^{(p)} \equiv (X_1^{(p)}, \ldots, X_N^{(p)})'$ are panic-market drivers independent of $\mathbf{X}^{(c)}$, with high homogeneous correlations r amongst each other

$$\mathbf{X}^{(p)} \sim \mathrm{N}\left(\left(\begin{array}{c} 0 \\ \vdots \\ 0 \end{array}\right), \left(\begin{array}{ccc} 1 & r & \ddots \\ r & 1 & r \\ \ddots & r & 1 \end{array}\right)\right). \tag{18}$$

The variable $\mathbf{B} \equiv (B_1, \ldots, B_N)'$ triggers panic. More precisely, $\mathbf{B}$ selects the panic downside endogenously a-la Merton (1974)

$$B_n \equiv \begin{cases} 1 & \text{if } X_n^{(p)} < \Phi^{-1}(b) \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

where Φ is the standard normal cdf and b is a low threshold probability.

The parameters $(\boldsymbol{\rho}, r, b)$ that specify the joint distribution (16) have an intuitive interpretation. The correlation matrix $\boldsymbol{\rho}$ characterizes the dependence structure of the market in times of regular activity. This matrix can be obtained by fitting a normal copula to the realizations of $\mathbf{X}$ that occurred in non-extreme regimes, as filtered by the minimum volume ellipsoid, see e.g. Meucci (2005) for a review and the code. The homogeneous correlation level r determines the dependence structure of the market in the panic regime. The probability b determines the likelihood of a high-correlation crash event. Therefore, r and b steer the effect of a non-symmetric panic correlation structure of an otherwise calm-market correlation ρ and are set as stress-test parameters.

The highly non-symmetrical joint distribution $F_{\mathbf{X}}$ defined by (16) is not analytically tractable. Nonetheless, we can easily generate a large number J of equal-probability joint scenarios $\{x_{1,j}, \ldots, x_{N,j}\}_{j=1,\ldots,J}$ from this distribution, and for enhanced precision impose as in Meucci (2009) that the first two moments of the simulations match the theoretical distribution.

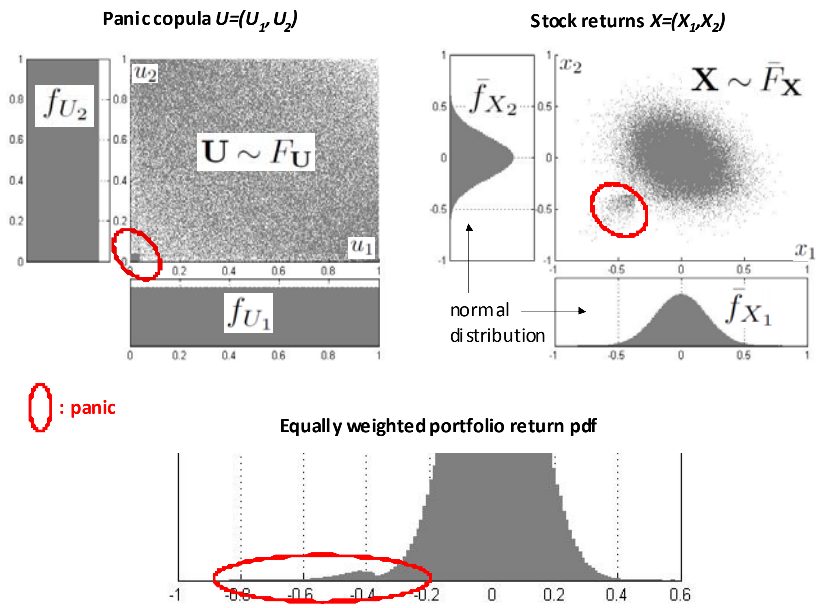

Figure 2: Panic copula, normal marginals, and skewed portfolio of normal returns

Due to the non-symmetrical nature of the panic triggers (19), this distribution has negative expectations, i.e. $\frac{1}{J}\sum_j x_{n,j} < 0$. Now we perform a second step to create a more realistic distribution that compensates for the exposure to downside risk. To this purpose, we use the Entropy Pooling approach as in Meucci (2008). Accordingly, we twist the probabilities $\bar{p}_j$ of the Monte Carlo scenarios in such a way that they display the least distortion with respect to the original probabilities $p_j \equiv 1/J$, and yet they give rise to non-negative expectations for the market $\mathbf{X}$, i.e. $\sum_j p_j x_{n,j} \geq 0$. In practice this amount to solving the following minimization

$$
\begin{aligned}
\{\bar{p}_j\} \quad &\equiv \quad \underset{\{p_j\}}{\operatorname{argmin}} \textstyle\sum_j p_j \ln\left(J p_j\right) \\
&\qquad \text{such that } \textstyle\sum_j p_j x_{n,j} \geq 0, \quad \sum_j p_j \equiv 1, \quad p_j \geq 0,
\end{aligned}
\tag{20}
$$

see Meucci (2008) for more details. Now the scenarios-probabilities $\{x_{1,j}, \ldots, x_{N,j}; \bar{p}_j\}$ represent a panic distribution $F_{\mathbf{X}}$ adjusted for risk-premium.

Using the separation step of CMA (13), we produce the scenario-probability representation $\{u_{1,j}, \ldots, u_{N,j}; \bar{p}_j\}$ of the panic copula $F_{\mathbf{U}}$. Then, using the combination step of CMA (15), we glue the panic copula $F_{\mathbf{U}}$ with marginals $\bar{F}_{X_n}$ fitted to the empirical observations of $\mathbf{X}$, creating the scenarios-probabilities $\{\bar{x}_{1,j}, \ldots, \bar{x}_{N,j}; \bar{p}_j\}$ for the panic distribution $\bar{F}_{\mathbf{X}}$, which fits the empirical data. The distribution $\bar{F}_{\mathbf{X}}$ can be used for stress-testing, or it can be fed into an optimizer to select an optimal panic-aware portfolio allocation.

To illustrate the panic copula, we show in the top-left portion of Figure 2 the scenarios of this copula with panic correlations $r \equiv 90\%$ and with very low panic probability $b \equiv 2\%$, for two stock returns. In the circle we highlighted the non-symmetrical downside panic scenarios.

For the marginals, a possible choice are Student t fits, as in Meucci, Gan, Lazanas, and Phelps (2007). Alternatively, we can construct the marginals as the kernel-smoothed empirical distributions of the returns, with tails fitted using extreme value theory, see Embrechts, Klueppelberg, and Mikosch (1997).

However, for didactical purposes, in the top-right portion Figure 2 we combine the panic copula with normal marginals fitted to the empirical data. This way we obtain a deceptively tame joint market distribution $\bar{F}_{\mathbf{X}}$ of normal returns. Nevertheless, even with perfectly normal marginal returns, and even with a very unlikely panic probability $b \equiv 2\%$, the market is dangerously skewed toward less favorable outcomes: portfolios of normal securities are not necessarily normal! In the bottom portion of Figure 2 we can visualize this effect for the equally weighted portfolio.

In the table below we report relevant risk statistics for the equally weighted portfolio in our panic market $\bar{F}_{\mathbf{X}}$. We also report the same statistics in a perfectly normal market, which follows by setting $b \equiv 0$ in (19)

Risk	Panic copula	Normal copula
CVaR 95%	-29%	-24%
Exp. value	0	0
St. deviation	12%	12%
Skewness	-0.4	0
Kurtosis	4.4	3

(21)

For more details, documented code is available at http://symmys.com/node/335.

6 Case study: copula transformations

Here we use CMA to perform arbitrary operations on arbitrary copulas. The documented code for this case study is available at http://symmys.com/node/335.

By construction, a generic copula $F_{\mathbf{U}}$ lives on the unit cube because each grade is normalized to have a uniform distribution on the unit interval. At times, when we need to modify the copula, the unit-interval, uniform normalization is impractical. For instance, one might need to reshuffle the dependence structure of the $N \times 1$ vector of the grades $\mathbf{U}$ by means of a linear transformation

$$T_{\boldsymbol{\gamma}} : \mathbf{U} \mapsto \boldsymbol{\gamma}\mathbf{U}, \tag{22}$$

where $\boldsymbol{\gamma}$ is a $N \times N$ matrix. Unfortunately, the transformed entries of $\boldsymbol{\gamma}\mathbf{U}$ are *not* the grades of a copula. This is easily verified because in general (22) transforms the copula domain, which is the unit cube, into a parallelotope that trespasses the boundaries of the unit cube. Furthermore, the marginal distribution of the transformed variable $\boldsymbol{\gamma}\mathbf{U}$ are not uniform.

In order to perform transformations on copulas, we propose to simply use alternative, not necessarily uniform, normalizations for the copulas, operate the transformation on the normalized variables, and then map the result back in the unit cube.

To be concrete, let us focus on the linear transformation (22). First, we normalize each grade to have a standard normal distribution, instead of uniform, i.e. we define the following random variables

$$Z_n \equiv \Phi^{-1}(U_n) \sim \mathrm{N}(0,1). \tag{23}$$

This is a special case of a combination step (6), where $\bar{F}_{X_n} \equiv \Phi$. Then we operate the linear transformation (22) of the normalized variables

$$\tilde{T}_{\boldsymbol{\gamma}} : \mathbf{Z} \mapsto \tilde{\mathbf{Z}} \equiv \boldsymbol{\gamma}\mathbf{Z}. \tag{24}$$

Finally, we map the transformed variables $\tilde{\mathbf{Z}}$ back into the unit cube space of the copula by means of the marginal cdf's of $\tilde{\mathbf{Z}}$.

$$\tilde{U}_n \equiv F_{\widetilde{Z}_n}(\widetilde{Z}_n) \sim \mathrm{U}([0,1]). \tag{25}$$

This step entails performing a separation step (4) and then only retaining the copula. This way we obtain the distribution of the grades $\tilde{F}_{\mathbf{U}} \equiv F_{\tilde{\mathbf{U}}}$

We summarize the copula transformation process in the following diagram

$$\begin{array}{ccc} F_{\mathbf{U}} & \overset{T}{\rightsquigarrow} & \tilde{F}_{\mathbf{U}} \\ \downarrow\mathcal{C} & & \uparrow\mathcal{S} \\ F_{\mathbf{Z}} & \xrightarrow{\tilde{T}} & F_{\tilde{\mathbf{Z}}} \end{array} \tag{26}$$

It is trivial to generalize the above steps and diagram to arbitrary non-linear transformations T. It is also possible to consider non-normal standardizations of the grades in the combination step (23), which can be tailored to the desired transformation T. The theory of the most suitable standardization for a given transformation is the subject of a separate publication.

In rare cases, the above copula transformations can be implemented analytically. However, the family of copulas that can be transformed analytically is extremely small, and depends on the specific transformation. For instance, for linear transformations we can only rely on elliptical copulas.

Instead, to implement copula transformations in practice, we rely on CMA, which allows us to perform arbitrary combination steps and separation steps, which are suitable for fully general transformations of arbitrary copulas.

To illustrate how to transform a copula using CMA, we perform a special case of the linear transformation $\boldsymbol{\gamma}$ in (24), namely a rotation on the panic

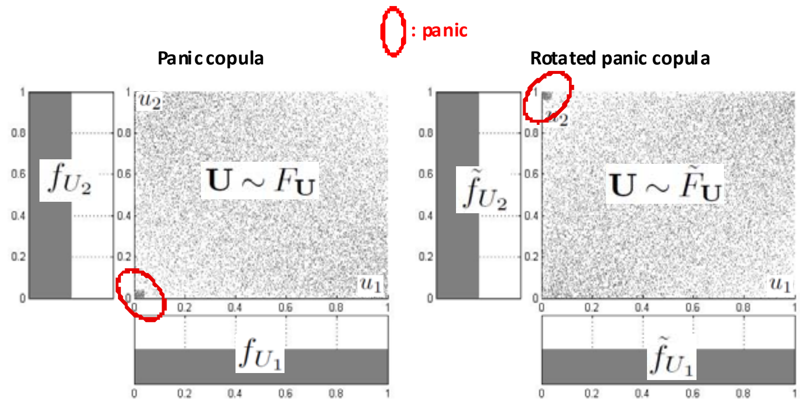

Figure 3: Copula-Marginal Algorithm for copula transformations: rotation of panic copula

copula introduced in Section 5. In the bivariate case we can parametrize the rotations by an angle θ as follows

$$\boldsymbol{\gamma} \equiv \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{27}$$

In Figure 3 we display the result for $\theta \equiv \pi/2$: the non-symmetric panic scenarios now affect positively the second security. For more details, documented code is available at http://symmys.com/node/335.

7 Conclusions

We introduced CMA, or "copula-marginal algorithm", a technique to generate new flexible copulas for risk management and portfolio management.

CMA generates flexible copulas and glues them with arbitrary marginals using the scenarios-probabilities representation of a distribution. CMA generates many more copulas than the few parametric families used in traditional copula implementations. For instance, with CMA we can generate large-dimensional, downside-only panic copulas. CMA also allows us to perform arbitrary transformations of copulas, despite the fact that copulas are only defined on the unit cube. Finally, unlike in traditional approaches to copula implementation, the probabilities of the scenarios are not assumed equal. Therefore CMA allows us to leverage techniques such as importance sampling and Entropy Pooling.

Documented code for CMA is available at http://symmys.com/node/335.

References

Brigo, D., A. Pallavicini, and R. Torresetti, 2010, *Credit Models and the Crisis: A Journey Into CDOs, Copulas, Correlations and Dynamic Models* (Wiley).

Cherubini, U., E. Luciano, and W. Vecchiato, 2004, *Copula Methods in Finance* (Wiley).

Daul, S., E. De Giorgi, F. Lindskog, and A. McNeil, 2003, The grouped t-copula with an application to credit risk, *Working Paper* Available at http://ssrn.com/abstract=1358956.

Durrleman, V., A. Nikeghbali, and T. Roncalli, 2000, Which copula is the right one?, *Working Paper*.

Embrechts, P., McNeil A., and D. Straumann, 2000, Correlation: Pitfalls and alternatives, *Working Paper*.

Embrechts, P., C. Klueppelberg, and T. Mikosch, 1997, *Modelling Extremal Events* (Springer).

Embrechts, P., F. Lindskog, and A. J. McNeil, 2003, Modelling dependence with copulas and applications to risk management, *Handbook of heavy tailed distributions in finance*.

Genest, C., and R. Rivest, 1993, Statistical inference procedures for bivariate Archimedean copulas, *Journal of the American Statistical Association* 88, 1034– 1043.

Glasserman, P., 2004, *Monte Carlo Methods in Financial Engineering* (Springer).

Jaworski, P., F. Durante, W. Haerdle, and T. (Editors) Rychlik, 2010, *Copula Theory and its Applications* (Springer, Lecture Notes in Statistics - Proceedings).

Li, D. X., 2000, On default correlation: A copula function approach, *Journal of Fixed Income* 9, 43–54.

Merton, R. C., 1974, On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29, 449–470.

Meucci, A., 2005, *Risk and Asset Allocation* (Springer) Available at http://symmys.com.

———, 2006, Beyond Black-Litterman in practice: A five-step recipe to input views on non-normal markets, *Risk* 19, 114–119 Article and code available at http://symmys.com/node/157.

———, 2008, Fully Flexible Views: Theory and practice, *Risk* 21, 97–102 Article and code available at http://symmys.com/node/158.

——— , 2009, Simulations with exact means and covariances, *Risk* 22, 89–91 Article and code available at http://symmys.com/node/162.

——— , 2010, Historical scenarios with Fully Flexible Probabilities, *GARP Risk Professional - "The Quant Classroom by Attilio Meucci"* December, 40–43 Article and code available at http://symmys.com/node/150.

——— , 2011, A short, comprehensive, practical guide to copulas, *GARP Risk Professional - "The Quant Classroom by Attilio Meucci"* October, 54–59 Article and code available at http://symmys.com/node/351.

——— , Y Gan, A. Lazanas, and B. Phelps, 2007, A portfolio mangers guide to Lehman Brothers tail risk model, *Lehman Brothers Publications*.

Nelsen, R. B., 1999, *An Introduction to Copulas* (Springer).

A Appendix

The appendix below can be skipped at first reading.

A.1 Separation: from joint to copula/marginal

Let us first introduce some terminology. Consider a set of J numbers $\{z_j\}$ and the set of all permutations of the first J integers $\pi(\{1,\ldots,J\})$.

Sorting $\{z_j\}$ is equivalent to defining a permutation a of the first J integers

$$\{z_j\} \mapsto a \in \pi(\{1,\ldots,J\}), \tag{28}$$

such that

$$z_{a_j} = z_{j:J}. \tag{29}$$

where $z_{j:J}$ is the j-th smallest number among $\{z_j\}$.

Sorting $\{z_j\}$ is also equivalent to defining the permutation b of the first J integers which is the inverse of the permutation a

$$\{z_j\} \mapsto b \in \pi(\{1,\ldots,J\}). \tag{30}$$

It is easy to verify that the permutation n represents the ranking of the entries of $\{z_j\}$.

To illustrate, consider an example with $J \equiv 4$ entries

$$\{z_j\} \equiv \begin{pmatrix} 5.2 \\ 7.4 \\ 2.3 \\ 1.7 \end{pmatrix} \mapsto a \equiv \begin{pmatrix} a_1 \equiv 4 \\ a_2 \equiv 3 \\ a_3 \equiv 1 \\ a_4 \equiv 2 \end{pmatrix}, \quad b \equiv \begin{pmatrix} b_1 \equiv 3 \\ b_2 \equiv 4 \\ b_3 \equiv 2 \\ b_4 \equiv 1 \end{pmatrix}. \tag{31}$$

Then

$$\begin{array}{ll} z_{a_1} = z_4 = 1.7 = z_{1:4} & z_1 \equiv 5.2 \text{ is the } b_1 \equiv 3\text{-rd smallest} \\ z_{a_2} = z_3 = 2.3 = z_{2:4} & z_2 \equiv 7.4 \text{ is the } b_2 \equiv 4\text{-th smallest} \\ z_{a_3} = z_1 = 5.2 = z_{3:4} & z_3 \equiv 2.3 \text{ is the } b_3 \equiv 2\text{-nd smallest} \\ z_{a_4} = z_2 = 7.4 = z_{4:4} & z_4 \equiv 1.7 \text{ is the } b_4 \equiv 1\text{-st smallest} \end{array} \tag{32}$$

Now consider a set of scenarios-probabilities $\{x_{1,j},\ldots,x_{N,j};p_j\}$ that represents the fully general joint distribution f_X of the random variable $X \equiv (X_1,\ldots,X_N)$. For the generic n-th entry, first we sort, generating the sorting/ranking permutations (28) and (30)

$$\{x_{n,j}\}_{j=1,\ldots,J} \mapsto a, b. \tag{33}$$

Then we define the sorted entries $\overline{x}_{n,j}$ and the respective sorted cumulative probabilities $\overline{u}_{j,n}$ using the permutation a in (33) as follows

$$\overline{x}_{n,j} \equiv x_{n,a_j}, \quad \overline{u}_{n,j} \equiv \textstyle\sum_{s=1}^{j} p_{a_s}. \tag{34}$$

Then the n-th empirical cdf $F_{X_n} \equiv \sum_{j=1}^{J} p_j 1_{[x_{n,j},\infty)}$ satisfies

$$F_{X_n}(\overline{x}_{n,j}) = \overline{u}_{n,j}. \tag{35}$$

Therefore the marginal cdf's at arbitrary points are represented by the inter/extrapolation of a grid

$$F_{X_n}(x) \equiv I_{\{\overline{x}_{n,j},\overline{u}_{n,j}\}}(x). \tag{36}$$

As for the n-th grade $u_{n,j} \equiv F_{X_n}(x_{n,j})$ it satisfies $u_{n,j} = \sum_{x_{n,s}\leq x_{n,j}} p_s = \overline{u}_{n,b_j}$, where j is the permutation defined in (33).

Repeating for all the entries $n = 1,\ldots,N$ of the random variable $\mathbf{X} \equiv (X_1,\ldots,X_N)'$ we obtain scenarios-probabilities representation of the copula

$$f_{\mathbf{U}} \iff \{\overline{u}_{n,b_j}, p_j\}.. \tag{37}$$

To illustrate, consider a $N \equiv 2$-variate distribution

$$f_{\mathbf{X}} \quad \Leftrightarrow \quad \{x_{n,j}\} \equiv \begin{pmatrix} 5.2 & 3.6 \\ 7.4 & 2.5 \\ 2.3 & 1.9 \\ 1.7 & 6.4 \end{pmatrix}, \quad \{p_j\} \equiv \begin{pmatrix} 0.2 \\ 0.4 \\ 0.3 \\ 0.1 \end{pmatrix} \tag{38}$$

Then for the marginal

$$\{\overline{x}_{n,j}\} \equiv \begin{pmatrix} 1.7 & 1.9 \\ 2.3 & 2.5 \\ 5.2 & 3.6 \\ 7.4 & 6.4 \end{pmatrix}, \quad \{\overline{u}_{n,j}\} \equiv \begin{pmatrix} 0.1 & 0.3 \\ 0.4 & 0.7 \\ 0.6 & 0.9 \\ 1.0 & 1.0 \end{pmatrix} \tag{39}$$

Therefore (35) becomes

$$F_1(1.7) = 0.1, \quad F_1(2.3) = 0.4, \quad F_1(5.2) = 0.6, \quad F_1(7.4) = 1.0 \tag{40}$$
$$F_2(1.9) = 0.3, \quad F_2(2.5) = 0.7, \quad F_2(3.6) = 0.9, \quad F_2(6.4) = 1.0 \tag{41}$$

As for the copula, (37) becomes

$$f_{\mathbf{U}} \quad \Leftrightarrow \quad \{u_{n,j}\} \equiv \begin{pmatrix} 0.6 & 0.9 \\ 1.0 & 0.7 \\ 0.4 & 0.3 \\ 0.1 & 1.0 \end{pmatrix}, \quad \{p_j\} \equiv \begin{pmatrix} 0.2 \\ 0.4 \\ 0.3 \\ 0.1 \end{pmatrix} \tag{42}$$

Commented code for the separation algorithm is available at http://symmys.com/node/335.

A.2 Combination: from copula/marginal to joint

Consider a generic scenario-probabilities representation $\{u_{n,j}, p_j\}$ for the copula and a set of marginal distributions $\{F_{Y_n}\}$.

For each marginal F_{Y_n}, we select a significant grid of values $\{\tilde{y}_{n,k}\}_{k=1,\ldots,K_n}$, where the number of elements K_n in the grid can, but need not be, equal to the number of scenarios J. To select such scenarios several approaches are possible. For instance, if the quantile function is readily available, one can choose a grid equally-spaced in probability

$$\begin{aligned} \tilde{y}_{n,1} &\equiv F_{Y_n}^{-1}(\epsilon) \\ &\vdots \\ \tilde{y}_{n,k} &\equiv F_{Y_n}^{-1}(\epsilon + (k-1)\Delta) \\ &\vdots \\ \tilde{y}_{n,K_n} &\equiv F_{Y_n}^{-1}(1-\epsilon), \end{aligned} \tag{43}$$

where $\epsilon \ll 1$ and Δ is set to solve $1 - \epsilon \equiv \epsilon + (K-1)\Delta$.

Alternatively, an equally-spaced grid between lower and upper quantile

$$\begin{aligned} \tilde{y}_{n,1} &\equiv F_{Y_n}^{-1}(\epsilon) \\ &\vdots \\ \tilde{y}_{n,k} &\equiv \tilde{y}_{n,1} + (k-1)\Delta_n \\ &\vdots \\ \tilde{y}_{n,K_n} &\equiv F_{Y_n}^{-1}(1-\epsilon), \end{aligned} \tag{44}$$

where again $\epsilon \ll 1$, the lower and upper quantile are computed by quadrature and Δ is set to solve $F_{Y_n}^{-1}(1-\epsilon) = F_{Y_n}^{-1}(\epsilon) + (K-1)\Delta_n$.

Alternatively, $\{\tilde{y}_{n,k}\}$ can be Monte Carlo or empirical scenarios from the marginal distribution F_{Y_n}.

Then we compute the grid $\{\tilde{u}_{n,k}, \tilde{y}_{n,k}\}_{k=1,\ldots,K_n}$, where $\tilde{u}_{n,k} \equiv F_{Y_n}(\tilde{y}_{n,k})$.

Next, we define the joint scenarios by inter/extra-polation

$$y_{n,j} \equiv I_{\{\tilde{u}_{n,k}, \tilde{y}_{n,k}\}}(u_{n,j}). \tag{45}$$

The probabilities $\{p_j\}$ are unaffected. Therefore, we obtain the scenario-probabilities representation $\{x_{n,j}, p_j\}$ for the copula. Commented code for the combination algorithm is available at http://symmys.com/node/335.