Institut canadien des dérivés
Canadian Derivatives Institute

L'Institut bénéficie du soutien financier de l'Autorité des marchés financiers ainsi que du ministère des Finances du Québec

Document de recherche

DR 18-07

Joint Distribution of a Random Vector with Known Margins of Linear Combinations

Mai, 2018

Ce document de recherche a été rédigée par :

Carole Bernard, Grenoble École de Management
Oleg Bondarenko, Université de l'Illinois
Steven Vanduffel, Vrije Université de Bruxelles

Joint Distribution of a Random Vector with Known Margins of Linear Combinations

Carole Bernard,* Oleg Bondarenko† and Steven Vanduffel‡§

March 4, 2018

Abstract

We provide an algorithm for obtaining a joint distribution of a random vector $(X_1, \ldots, X_d)$ when the marginal distribution functions of all its components as well as some linear combinations thereof are known. These results can be useful to build a non-parametric approach for multivariate option pricing.

Key-words: Rearrangement Algorithm, Model-free dependence, Model uncertainty, Fréchet-Hoefding bounds, Risk-Neutral Density.

*Carole Bernard, Department of Accounting, Law and Finance, Grenoble Ecole de Management (GEM) (email: carole.bernard@grenoble-em.com). Department of Economics and Political Sciences at Vrije Universiteit Brussel (VUB).

†Oleg Bondarenko, Department of Finance, University of Illinois at Chicago (UIC). (e-mail: olegb@uic.edu).

‡Steven Vanduffel, Department of Economics and Political Sciences at Vrije Universiteit Brussel (VUB). (e-mail: steven.vanduffel@vub.ac.be).

§The authors gratefully acknowledge funding from IFSID and the Global Risk Institute (GRI) for the related project "Inferring tail correlations from option prices".

1 Introduction

Let $X_1, \ldots, X_d$ be d real valued random variables (rvs) that are defined on an atomless probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We assume that the distribution functions F_j of the X_j as well as the distribution functions, F_{α^k}, of given linear combinations $\sum_{j=1}^d \alpha_j^k X_j$, $\alpha_j^k \in \mathbb{R}$ $(k = 1, \ldots, K)$ are all known. Each F_j describes a marginal risk X_j, $j = 1, \ldots, d$, whereas every F_{α^k}, $k = 1, \ldots, K$, yields information on their interdependence. In fact, if we only use the marginal distributions F_j as source of information (i.e., when $K = 0$), then for each d-dimensional copula C the map from $\mathbb{R}^d$ to $[0, 1]$

$$(x_1, \ldots, x_d) \mapsto C(F_1(x_1), \ldots, F_d(x_d))$$

defines a d-dimensional joint distribution function F with univariate marginal distributions F_1,...,F_d. In this case every dependence (copula) is thus admissible to build a joint distribution for the random vector $(X_1, \ldots, X_d)$. When adding the dependence information that is contained in the distributions F_{α^k}, $k = 1, \ldots, K$, only a subset of copulas will still be admissible. However, it is typically no longer clear how to specify them. In this regard, note that the distribution of a random vector is uniquely specified if and only if the distribution of all linear combinations of the variables are known. Here, we only specify a finite number of them and hence the distribution of the random vector is no longer unique in general.

The contribution of this paper is to provide a constructive algorithm that yields a joint distribution that is consistent with the imposed distributional constraints. To illustrate that this problem is not easy, let us mention that a closely related problem concerns finding a joint distribution (or equivalently, a copula) for given marginal distributions and a given covariance matrix. For this problem, Cario and Nelson (1997) develop the so-called Normal to anything method (NORTA) to obtain an admissible Gaussian dependence. Ghosh and Henderson (2003) show that one cannot expect that this approach works well in higher dimensions (see also Kurowicka and Cooke (2006)) and develop an involved numerical approach based on checkerboard copulas. If the distributions of all the pairwise sums (thus, also all covariances) are known, then we can use our algorithm to obtain a consistent joint distribution.

The algorithm that we propose is a generalization of the so-called Rearrangement Algorithm (RA) of Puccetti and Rüschendorf (2012) and Embrechts, Puccetti, and Rüschendorf (2013); see also Rüschendorf (1983) for an early version that applies in special cases as well as Gaffke and Rüschendorf (1981), Wang and Wang (2011), Wang, Peng, and Yang (2013), Puccetti and Wang (2015) and Wang and Wang (2016) for important contributions on the strongly related mathematical concept of mixability of random variables. Specifically, Puccetti and Rüschendorf (2012) and Embrechts, Puccetti, and Rüschendorf (2013) aim at finding the worst-case dependence among variables with given marginal distributions in that a suitable risk measure (typically the Value-at-Risk) becomes maximum. They show that this problem is in essence closely related to the problem of rearranging a matrix (in which the columns represent samples of the marginal distributions)

such that the row sums become as constant as possible and develop the RA as a powerful heuristic to deal with it. The RA and the mathematical study of the concept of mixability have greatly contributed to the development of the literature on model risk assessment.[1] Moreover, the RA has also shown to be a useful tool for dealing with problems in operations research (Haus (2015), Boudt, Jakobsons, and Vanduffel (2017), Jakobsons and Wang (2017)). Further properties, performance analysis and technical generalizations of the RA have been studied in Bernard and McLeish (2016), Boudt, Jakobsons, and Vanduffel (2017), Bernard, Rüschendorf, and Vanduffel (2017), Bernard, Bondarenko, and Vanduffel (2018), Bernard, Rüschendorf, Vanduffel, and Yao (2017) and Hofert, Memartoluie, Saunders, and Wirjanto (2017).

The RA can be seen as a constructive way to obtain an approximation of the joint distribution of a random vector under the linear constraint that the sum of components has a degenerate distribution (i.e., is constant almost surely). We generalize this algorithm in that we show how to obtain a joint distribution under general distributional constraints on an arbitrary number of linear combinations of the components. We label the algorithm that we propose *Constrained Block Rearrangement Algorithm (CBRA)*. We also briefly discuss an application; we show how this algorithm can be applied to build a non-parametric model for pricing any multivariate option when prices of call and put options on single assets and linear combinations thereof (basket and index options) are available.

2 Inferring a Joint Distribution

2.1 Problem Formulation and Discretization

The information that is available on the marginal distribution functions F_j $(j = 1, \dots, d)$ and F_{α^k} $(k = 1, \dots, K)$ is generally not sufficient to identify a unique joint distribution.[2] In order to find a consistent distribution, we effectively need to specify a joint dependence of a vector $(U_1, \dots, U_d)$ of uniformly distributed variables, i.e., a copula, having the property that for all $\mathrm{x} \in \mathbb{R}$ and for all $k = 1, \dots, K$

$$\mathbb{P}\left(\sum_{j=1}^{d} \alpha_j^k F_j^{-1}(U_j) \leqslant x\right) = F_{\alpha^k}(x). \tag{1}$$

[1] See for instance Embrechts, Wang, and Wang (2015), Lux and Papapantoleon (2016), Bernard, Rüschendorf, and Vanduffel (2017), Bernard, Rüschendorf, Vanduffel, and Yao (2017), Puccetti, Rüschendorf, Small, and Vanduffel (2017), and Rüschendorf and Witting (2017).

[2] When F_α is available for *any* possible choice of $\alpha \in \mathbb{R}$, then the moment generating function of $(X_1, X_2, \dots, X_d)$ is also known, from which the unique joint distribution function can be obtained. When there is only a limited number of linear combinations with known distributions, then several consistent joint distributions will typically exist, but it is no longer clear *a priori* how to find them.

Indeed, in this case a consistent joint distribution F is obtained as

$$F(x_1, \ldots, x_d) = \mathbb{P}\left(F_1^{-1}(U_1) \leqslant x_1, \ldots, F_d^{-1}(U_d) \leqslant x_d\right), \qquad (x_1, \ldots, x_d) \in \mathbb{R}^d.$$

Next, we observe that in order to obtain a consistent distribution, it is equivalent to look for rearrangements f_i of F_i^{-1} such that for a uniformly distributed variable U and all $k = 1, \ldots, K$ it holds that

$$\mathbb{P}\left(\sum_{j=1}^{d} \alpha_j^k f_j(U) \leqslant x\right) = F_{\alpha^k}(x), \tag{2}$$

since then $F(x_1, \ldots, x_d)$ can be taken as

$$F(x_1, \ldots, x_d) = \mathbb{P}\left(f_1(U) \leqslant x_1, \ldots, f_d(U) \leqslant x_d\right).$$

In the sequel, we use the latter formulation for describing a joint distribution and we thus merely aim at finding rearrangements f_j of the F_j^{-1} such that the K constraints in (2) are all satisfied. To this end, we first discretize the problem. Specifically, since each F_j can be approximated to any degree of accuracy by a discrete distribution function, we assume F_j to be discrete, i.e., we consider

$$F_j(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{(x_{ij}, \infty)}(x), \quad j = 1, \ldots, d.$$

in which the values x_{ij} are chosen as

$$x_{ij} = F_j^{-1}\left(\frac{i - 0.5}{n}\right), \quad i = 1, \ldots, n.$$

In other words, we sample n equiprobable values from each distribution F_j to obtain an $n \times d$ matrix

$$\mathbf{X} = (x_{ij}) = \begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1d} \\ x_{21} & x_{22} & \ldots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{nd} \end{bmatrix}.$$

At this point the matrix $\mathbf{X}$ represents the support of the comonotonic random vector

$$(F_1^{-1}(U), \ldots, F_d^{-1}(U)).$$

Specifically, the rows of $\mathbf{X}$ depict all joint realizations, each occurring with probability $1/n$ (note that there might be duplicates), and its columns represent the random variables $F_j^{-1}(U)$, $j =$

$1, \ldots, d$. This observation motivates why in the sequel we apply probabilistic operators (such as the variance) onto columns of matrices. We will use the notation X_j for depicting both the random variable X_j (which takes the n equiprobable values x_{ij} for $i = 1, \ldots, n$), as well as the j-th column of the matrix $\mathbf{X}$.

We follow a similar approach to discretize the distribution functions F_{α^k} for the K linear constraints and obtain an $n \times K$ matrix

$$\mathbf{S} = (s_{ik}) = \begin{bmatrix} s_{11} & s_{12} & \ldots & s_{1K} \\ s_{21} & s_{22} & \ldots & s_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \ldots & s_{nK} \end{bmatrix},$$

where the column vector S_k contains the sampled values from F_{α^k}, i.e.,

$$s_{ik} = F^{-1}_{\alpha^k}\left(\frac{i - 0.5}{n}\right), \quad i = 1, \ldots, n.$$

Next we let $\mathbf{Y} = [\mathbf{X}; \mathbf{S}]$ denote an $n \times (d+K)$ matrix which combines the columns for the d marginal distributions and K linear constraints. By rearranging elements within each column of matrix $\mathbf{Y}$ we can obtain a new matrix $\mathbf{Y}^\pi$:

$$\mathbf{Y}^\pi \in P(\mathbf{Y}) := \left\{ (y^\pi_{ij}) : \quad y^\pi_{ij} := y_{\pi_j(i)j}, \quad i = 1, \ldots, n; \quad j = 1, \ldots, d+K \right\},$$

in which the $\pi_1, \ldots, \pi_{d+K}$ denote permutations of $\{1, \ldots, n\}$. In what follows we will call the vector of permutations $\pi := (\pi_1, \ldots, \pi_{d+K})$ a *rearrangement* and $\mathbf{Y}^\pi$ a *rearranged matrix*. We aim to find a rearranged matrix $\mathbf{Y}^\pi \in P(\mathbf{Y})$ which satisfies constraints in (2) are all satisfied. This can be expressed as

$$L^\pi_k := \sum_{j=1}^{d+K} \widetilde{\alpha}^k_j Y^\pi_j = 0, \tag{3}$$

where

$$\widetilde{\alpha}^k_j = \begin{cases} \alpha^k_j, & \text{for } j = 1, \ldots, d \\ -1, & \text{for } j = d + k \\ 0, & \text{otherwise} \end{cases}$$

Equivalently, we want to find a rearranged matrix $\mathbf{Y}^\pi$ which minimizes the objective

$$V := V(\mathbf{Y}^\pi) = \sum_{k=1}^{K} \mathrm{var}(L^\pi_k) = 0. \tag{4}$$

Verifying this condition for all possible rearranged matrices is not practical as there are $(n!)^{d+K-1}$

possibilities to consider. It turns out, however, that a particular kind of rearrangements, called *block rearrangements*, are suitable to the problem we consider. A block rearrangement consists in swapping entire rows within a given block of columns (and not just elements within a given column). Specifically, following the same terminology as in Boudt, Jakobsons, and Vanduffel (2017), let $\delta \in \{0,1\}^{d+K}$ indicate by $\delta_j = 1$ the columns j to be rearranged, $j = 1, \ldots, d+K$. A rearrangement π of a matrix $\mathbf{Y}$ into $\mathbf{Y}^\pi$ is called a block rearrangement if there exists $\delta \in \{0,1\}^{d+K}$ and a permutation $\pi^\delta : \{1, \ldots, n\} \to \{1, \ldots, n\}$ such that $\pi_j(\cdot) = \pi^\delta(\cdot)$ for $\delta_j = 1$ and $\pi_j(\cdot) = (\cdot)$ for $\delta_j = 0$ (which means that for all x, $\pi_j(x) = x$). As a special case, if the vectors $\delta \in \{0,1\}^{d+K}$ satisfy $\sum_{j=1}^{d+K} \delta_j = 1$, i.e., when the block only concerns a single column, then the block rearrangement reduces to rearranging this particular column.[3]

2.2 Constrained Rearrangement Algorithm

We propose an algorithm for rearranging $\mathbf{Y}$ into $\mathbf{Y}^\pi \in P(\mathbf{Y})$ such that condition (4) is approximately met. The algorithm builds on the fact that two random variables W and Z with given distribution have minimum correlation if and only if they are antimonotonic. In the sequel, we use the notation

$$W \,\updownarrow\, Z$$

to reflect that W and Z are antimonotonic. We rely on the following lemma.

Lemma 2.1 (Minimum Covariance)**.** *Let W and $Z_1, Z_2, \ldots, Z_m$ be random variables. Assume that $\widetilde{W}$ has the same distribution as W and is anti-monotonic with $\sum_{i=1}^m Z_i$. Then*

$$\operatorname{cov}\left(W, \sum_{i=1}^m Z_i\right) \geqslant \operatorname{cov}\left(\widetilde{W}, \sum_{i=1}^m Z_i\right).$$

When W is not anti-monotonic with $\sum_{i=1}^m Z_i$, the inequality is strict.

In order to construct a rearranged matrix that approximately meets condition (4), let us note that for any $\delta \in \{0,1\}^{d+K}$ with corresponding sets I_δ and I_{δ^c} satisfying $I_\delta \cap I_{\delta^c} = \varnothing$ and $I_\delta \cup I_{\delta^c} =$

[3] In the original Rearrangement Algorithm (RA) of Puccetti and Rüschendorf (2012) and Embrechts, Puccetti, and Rüschendorf (2013) single columns are rearranged. It was noted in Bernard, Rüschendorf, and Vanduffel (2017) and Bernard and McLeish (2016) that their algorithm could be improved by block rearrangements, which has led to the development of Block RA (BRA); see also Puccetti and Wang (2015) for the related mathematical concept of Σ-countermonotonic matrices. BRA exploits the fact that a necessary condition for minimum variance of a sum is that each (partial) sub-sum has minimum correlation with the complementary sub-sum, whereas the original RA only considers sub-sums that contain one component (column) (Remark 4.1 in Bernard, Rüschendorf, and Vanduffel (2017), Proposition 2.1 of Bernard and McLeish (2016) as well as Section 3).

$\{1, \ldots, d+K\}$, it holds that

$$\sum_{k=1}^{K} \operatorname{var}(L_k) = C + \sum_{k=1}^{K} \operatorname{cov}\left(\sum_{j \in I_\delta} \widetilde{\alpha}_j^k Y_j, \sum_{j \in I_\delta^c} \widetilde{\alpha}_j^k Y_j \right)$$

where $j \in I_\delta$ if and only if $\delta_j = 1$, and

$$C = \sum_{k=1}^{K} \left(\operatorname{var}\left(\sum_{j \in I_\delta} \widetilde{\alpha}_j^k Y_j \right) + \operatorname{var}\left(\sum_{j \in I_\delta^c} \widetilde{\alpha}_j^k Y_j \right) \right).$$

Assume now that $\delta \in \{0,1\}^{d+K}$ is chosen such that for every constraint $k \in \{1, \ldots, K\}$, the coefficients $\widetilde{\alpha}_j^k$ do not depend on j, i.e., $\widetilde{\alpha}_j^k := \widetilde{\alpha}^k$, $j \in I_\delta$, then C is constant since its value does not change when applying a *block rearrangement* π_δ. We refer to such δ as *admissible*. An admissible δ always exists; the stated condition on $\widetilde{\alpha}_j^k$ in particular holds for all $\delta \in \{0,1\}^{d+K}$ satisfying $\sum_{j=1}^{d+K} \delta_j = 1$ (case of singleton "blocks", i.e., columns). The point is then that, for any such δ, one can identify a corresponding block rearrangement π^δ that decreases the value of the objective V for the rearranged matrix $\mathbf{Y}^{\pi^\delta}$. Indeed, it follows from Lemma 2.1 that to lower V, π^δ should be taken such that

$$\sum_{j \in I_\delta} \widetilde{\alpha}^k Y_j^{\pi^\delta} \,\text{⇅}\, \left(\sum_{j \in I_\delta^c} \widetilde{\alpha}_j^k Y_j \right).$$

These considerations lead to the following Constrained Block Rearrangement Algorithm (CBRA).

Constrained Block Rearrangement Algorithm (CBRA)

1. Consider the initial matrix $\mathbf{Y}$ and select an admissible δ.
2. Apply the block rearrangement π^δ such that

$$\sum_{j \in I_\delta} \widetilde{\alpha}^k Y_j^{\pi^\delta} \,\text{⇅}\, \left(\sum_{j \in I_\delta^c} \widetilde{\alpha}_j^k Y_j \right).$$

3. Compute $V := \sum_{k=1}^{K} \operatorname{var}\left(L_k^{\pi^\delta} \right)$.
4. If there is no improvement in V, output the current matrix $\mathbf{Y}^{\pi^\delta}$, otherwise, return to Step 1.

From the output matrix, one can then simply isolate the first d columns (corresponding to rearrangments of the X_j), which yields the solution of the problem we consider.

The algorithm guarantees that at each step the objective V decreases and, thus, the algorithm converges in finite time to a local optimum. In the ideal situation, after the algorithm stops, the rearranged matrix is such that all $L_k = 0$, $k = 1, \ldots, K$ (see (3)). In practice, however, due to discretization errors and the fact that the algorithm is a heuristic, the final L_k will deviate from the zero vector. We can denote these deviations by column vectors E_k, $k = 1, \ldots, K$ (one column for each constraint). It is possible to test whether these deviations (errors) are significant, in that one could adjust S_k for the noise E_k and assess whether the perturbed sum $S_k + E_k$ still statistically follows the given distribution F_{S_k}.

Remark 1 (Block RA). When $K = 1$ and the constraint is that the total sum $X_1 + X_2 + ... + X_d$ is constant (a degenerate distribution), our CBRA reduces to the Block RA (BRA) extensively studied in Bernard, Bondarenko, and Vanduffel (2018).

Remark 2 (Convergence Issues). It is known that (B)RA performs greatly in dealing with rearrangement problems (Embrechts, Puccetti, and Rüschendorf (2013), Bernard and McLeish (2016), Boudt, Jakobsons, and Vanduffel (2017), Hofert, Memartoluie, Saunders, and Wirjanto (2017),Bernard, Bondarenko, and Vanduffel (2018)), and similar performance can be expected from CBRA. There is, however, no proof that these algorithms *always* converge in that the deviations (errors) always vanish. In order to reduce the likelihood of non-convergence, it is a good practice to randomize in Step 1 of the CBRA the initial input matrix (i.e., to randomly change the order of the elements in each column), as well as to randomize the selection of admissible δ. Note that doing so makes it also possible to obtain different solutions.

Remark 3 (Different Objective). The objective V in (3) effectively assigns equal weights to all errors $\text{var}(L_k^{\pi^\delta})$, $k = 1, \ldots, K$. It might sometimes be preferable to use varying weights, for example to account for the different magnitudes the errors have. In practice, the difficulty of minimizing the error for the k-th constraint ($\text{var}(L_k^\pi)$) is related to the number of non-zero components used in the constraint and their variabilities. To accommodate such a possibility, we can re-write the objective as:

$$V := V(\mathbf{Y}^\pi) = \sum_{k=1}^{K} \gamma_k \, \text{var}(L_k^\pi) = 0, \tag{5}$$

for some positive constants γ_k, $k = 1, \ldots, K$. This problem can easily be solved by the above algorithm except the objective will now be computed as in (5) and Step 2 will change to

2. Apply the block rearrangement π^δ such that

$$\sum_{j \in I_\delta} \widetilde{\alpha}_j^k Y_j^{\pi^\delta} \,\text{⇅}\, \left(\sum_{j \in I_\delta^c} \gamma_k \left(\widetilde{\alpha}_j^k Y_j \right) \right).$$

Such variation of CBRA allows for example to adaptively select weights γ_k to ensure that the terms $\gamma_k \operatorname{var}(L_k^\pi)$ for $k = 1, \ldots, K$ are roughly the same magnitude.

Remark 4 (Generalization). The above algorithm can be generalized to the case in which an arbitrarily number of distributions of linear combinations is known, but not necessarily the marginal distributions for all of the d individual variables X_j $(j = 1, \ldots, d)$ are available, as we assume in this paper.[4] We are grateful to Ruodu Wang for this suggestion. To simplify the exposition, we do not present the algorithm in its full generality here and leave this generalization for future research.

Example 2.2 (Normal Distributions)**.** Consider six components, which are standard normally distributed, i.e., $X_j \sim N(0,1)$, for $j = 1, \ldots, 6$. Furthermore, assume that there are three linear combinations (sums) with known normal distribution. Specifically,

$$\begin{aligned} X_1 + X_2 + X_3 + X_4 &\sim N(0,10) := F_{\alpha^1}, \qquad \alpha^1 = [1,1,1,1,0,0], \\ X_3 + X_4 + X_5 + X_6 &\sim N(0,10) := F_{\alpha^2}, \qquad \alpha^2 = [0,0,1,1,1,1], \\ X_1 + X_2 + X_3 + X_4 + X_5 + X_6 &\sim N(0,24) := F_{\alpha^3}, \qquad \alpha^3 = [1,1,1,1,1,1]. \end{aligned}$$

Let $J_1 = \{1,2,3,4\}$, $J_2 = \{3,4,5,6\}$, and $J_3 = \{1,2,3,4,5,6\}$ denote sets containing the indexes of those components that appear in the first sum, the second sum and third sum, respectively. For a given set J, let $\bar{\rho}(J)$ denote the average pairwise correlation of all componenents with indexes in set I. That is,

$$\bar{\rho}(J) = \frac{2}{|J|(|J|-1)} \sum_{i,j \in J,\, i>j} \rho_{ij},$$

where ρ_{ij} denotes the correlation between X_i and X_j. It is easy to check that the above constraints on the three sums are equivalent to the following restrictions on the average correlations:

$$\bar{\rho}(J_1) = 0.5, \qquad \bar{\rho}(J_2) = 0.5, \qquad \bar{\rho}(J_3) = 0.6. \tag{6}$$

To infer a possible joint distribution for $(X_1, \ldots, X_6)$, we first discretize the continuous distributions using $n = 10,000$ equiprobable values and then run CBRA $M = 1,000$ times. To minimize the objective function, CBRA rearranges along all one-element blocks, as well as three additional two-element admissible blocks: $\{1,2\}$, $\{3,4\}$, and $\{5,6\}$. Depending on initial randomization, each run of CBRA produces a different solution for the joint distribution, but they all satisfy the three constraints in (6) very closely, with the averages across M runs being 0.5006, 0.5006, and 0.5993, respectively. The table below reports the correlation matrix for the six components averaged across

[4] Clearly, the marginal distribution for a single variable X_j may be viewed as a special case of a linear constraint.

all M runs:

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	1.000	0.854	0.501	0.501	0.782	0.782
X_2	0.854	1.000	0.501	0.501	0.782	0.782
X_3	0.501	0.501	1.000	0.146	0.500	0.500
X_4	0.501	0.501	0.146	1.000	0.500	0.500
X_5	0.782	0.782	0.500	0.500	1.000	0.858
X_6	0.782	0.782	0.500	0.500	0.858	1.000

Note also that the components X_1 and X_2 enter the constraints symmetrically (as well as have the same marginal distributions). The average pairwise correlations in the above table reflect this symmetry. The same also applies to the component pairs (X_3, X_4) and (X_5, X_6).

3 Model-free Approach for Pricing Multi-Asset Derivatives

An important application for our algorithm is pricing multi-asset derivatives. Consider a financial market with $d \geqslant 2$ risky non-dividend paying stocks, modeled as a $\mathbb{R}^d_+$-valued random process $(X_1(t), \ldots, X_d(t))_{t\in[0,T]}$ on a stochastic basis $(\Omega, \mathcal{A}, \mathbb{P})$. There also exists a risk-free asset with interest rate r.

We are interested in finding time-0 prices of various *path-independent* derivatives, i.e., a derivative whose payoff is a function of $(X_1(T), \ldots, X_d(T))$. From now on, we omit the reference to time in notation, i.e., we write X_j instead of $X_j(T)$. Let $C_j(L)$ denote the time-0 price of the European-style call option with strike L and maturity T on the underlying stock X_j. Under the standard assumptions, the call price is equal to the expected value of its payoff under a suitably chosen *risk-neutral* probability measure $\mathbb{Q}$:

$$C_j(L) = e^{-rT} E_t^{\mathbb{Q}} \left[(X_j - L)^+ \right] = e^{-rT} \int_0^\infty (x - L)^+ \, h_j(x) \, dx,$$

where $h_j(x)$ denotes the *risk-neutral density* (RND). The RND satisfies the relationship first discovered by Ross (1976), Breeden and Litzenberger (1978), Banz and Miller (1978):

$$h_j(x) = e^{rT} \left. \frac{\partial^2 C_j(L)}{\partial L^2} \right|_{L=x} \tag{7}$$

Although not directly observable, the RND can be recovered using the relationship in (7), provided that prices of call options with all strikes $L \in \mathbb{R}$ are available. The assumption that traded options are available for a continuum of strikes is evidently violated in reality, but a number of interpolation approaches have been proposed in the literature that allow to circumvent this shortcoming and to retrieve the implied risk-neutral distribution function when call prices are only

available for a finite number of strikes; see, for example, Jackwerth and Rubinstein (1996), Aït-Sahalia and Lo (2000), and Bondarenko (2003). We therefore assume that the option market offers sufficiently enough strikes that the risk-neutral distributions $F_1, \ldots, F_d$ for the stocks $X_1, \ldots, X_d$ can be accurately estimated.

Besides the individual stocks, we assume that market participants can also trade various indexes, as well as their options. Specifically, suppose that there are K market indexes (or, baskets) available, where

$$S_k = \sum_{k=1}^{K} \alpha_i^k X_i, \qquad k = 1, \ldots, K.$$

For example, Microsoft Corp (MSFT) can be a constituent of several popular indexes, including Dow Jones Industrial Average (DJIA), S&P 500, NASDAQ 100, S&P Technology Select Sector, and others. These indexes differ in the number of constituents and their weights (some indexes are price-weighted, while others are market capitalization-weighted). Furthermore, for many of these market indexes, liquid options are traded for a wide range of strikes. Thus, using the relationship in (7), the risk-neutral distributions F_{α_k}, $(k = 1, \ldots, K)$ can also be obtained.

Armed with the risk-neutral distributions for the d individual stocks as well as for K indexes, we can use CBRA to find a candidate joint distribution consistent with prices of existing derivative securities. The traditional approaches rely on strong parametric assumptions to model a multi-dimensional joint distribution and are prone to misspecification errors. In contrast, the results of our paper make it possible to develop a completely model-free method to retrieve a joint model that is entirely consistent with prices of traded options on both single assets and their various indexes. The estimated candidate joint distribution then can be used to price other path-independent derivatives written on the d stocks, including basket options, exchange options, correlation swaps, etc. To the best of our knowledge no other model-free approach for pricing multi-asset derivatives is available in the literature.

References

Aït-Sahalia, Y., and A. W. Lo (2000): "Nonparametric risk management and implied risk aversion," *Journal of econometrics*, 94(1), 9–51.

Banz, R. W., and M. H. Miller (1978): "Prices for state-contingent claims: Some estimates and applications," *Journal of Business*, 51(4), 653–672.

Bernard, C., O. Bondarenko, and S. Vanduffel (2018): "Rearrangement Algorithm and Maximum Entropy," *Annals of Operations Research*, 261(1-2), 107–134.

Bernard, C., and D. McLeish (2016): "Algorithms for Finding Copulas Minimizing Convex Functions of Sums," *Asia-Pacific Journal of Operational Research*, 33(5), 1650040.

Bernard, C., L. Rüschendorf, and S. Vanduffel (2017): "Value-at-Risk bounds with variance constraints," *Journal of Risk and Insurance*, 84(3), 1539–6975.

Bernard, C., L. Rüschendorf, S. Vanduffel, and J. Yao (2017): "How robust is the Value-at-Risk of credit risk portfolios?," *European Journal of Finance*, 23(6), 507–534.

Bondarenko, O. (2003): "Estimation of risk-neutral densities using positive convolution approximation," *Journal of Econometrics*, 116(1), 85–112.

Boudt, K., E. Jakobsons, and S. Vanduffel (2017): "Block rearranging elements within matrix columns to minimize the variability of the row sums," *4OR - A Quarterly Journal of Operations Research, forthcoming*, pp. 1–20, doi.org/10.1007/s10288–017–0344–4.

Breeden, D. T., and R. H. Litzenberger (1978): "Prices of state-contingent claims implicit in option prices," *Journal of business*, 51(4), 621–651.

Cario, M. C., and B. L. Nelson (1997): "Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix," Discussion paper, Technical Report, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois.

Embrechts, P., G. Puccetti, and L. Rüschendorf (2013): "Model uncertainty and VaR aggregation," *Journal of Banking & Finance*, 37(8), 2750–2764.

Embrechts, P., B. Wang, and R. Wang (2015): "Aggregation-robustness and model uncertainty of regulatory risk measures," *Finance and Stochastics*, 19(4), 763–790.

Gaffke, N., and L. Rüschendorf (1981): "On a class of extremal problems in statistics," *Optimization*, 12(1), 123–135.

Ghosh, S., and S. G. Henderson (2003): "Behavior of the NORTA method for correlated random vector generation as the dimension increases," *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 13(3), 276–294.

Haus, U.-U. (2015): "Bounding stochastic dependence, joint mixability of matrices, and multidimensional bottleneck assignment problems," *Operations Research Letters*, 43(1), 74–79.

Hofert, M., A. Memartoluie, D. Saunders, and T. Wirjanto (2017): "Improved algorithms for computing worst Value-at-Risk," *Statistics & Risk Modeling*, 34(1-2), 13–31.

Jackwerth, J., and M. Rubinstein (1996): "Recovering probability distributions from option prices," *Journal of Finance*, 51, 1611–1631.

Jakobsons, E., and R. Wang (2017): "Negative dependence in matrix arrangement problems," *Forthcoming in Annals of Operations Research*.

Kurowicka, D., and R. M. Cooke (2006): *Uncertainty analysis with high dimensional dependence modelling*. John Wiley & Sons.

Lux, T., and A. Papapantoleon (2016): "Model-free bounds on Value-at-Risk using partial dependence information," *arXiv preprint arXiv:1610.09734*.

Puccetti, G., and L. Rüschendorf (2012): "Computation of sharp bounds on the distribution of a function of dependent risks," *Journal of Computational and Applied Mathematics*, 236(7), 1833–1840.

Puccetti, G., L. Rüschendorf, D. Small, and S. Vanduffel (2017): "Reduction of Value-at-Risk bounds via independence and variance information," *Scandinavian Actuarial Journal*, 2017(3), 245–266.

Puccetti, G., and R. Wang (2015): "Extremal Dependence Concepts," *Statistical Science*, 30(4), 485–517.

Ross, S. A. (1976): "Options and efficiency," *Quarterly Journal of Economics*, 90(1), 75–89.

Rüschendorf, L. (1983): "On the multidimensional assignment problem," *Zeitschrift für Operations Research*, Serie A(47), 107–113.

Rüschendorf, L., and J. Witting (2017): "VaR bounds in models with partial dependence information on subgroups," *Dependence Modeling*, 5(1), 59–74.

Wang, B., and R. Wang (2011): "The complete mixability and convex minimization problems with monotone marginal densities," *Journal of Multivariate Analysis*, 102(10), 1344–1360.

——— (2016): "Joint mixability," *Mathematics of Operations Research*, 41(3), 808–826.

Wang, R., L. Peng, and J. Yang (2013): "Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities," *Finance and Stochastics*, 17(2), 395–417.