Correlation matrices with average constraints

Citation for published version:
Tuitman, J, Vanduffel, S & Yao, J 2020, 'Correlation matrices with average constraints', *Statistics and Probability Letters*, vol. 165, 108868. https://doi.org/10.1016/j.spl.2020.108868

Digital Object Identifier (DOI):
10.1016/j.spl.2020.108868

Published In:
Statistics and Probability Letters

Correlation Matrices with Average Constraints

Jan Tuitman* Steven Vanduffel† Jing Yao‡

Abstract

We develop an algorithm that makes it possible to generate all correlation matrices satisfying a constraint on their average value. We extend the results to the case of multiple constraints. These results can be used to assess the extent to which methodologies driven by correlation matrices are robust to misspecification thereof.

1 Introduction

Correlation matrices are a crucial model component in many disciplines including statistics, risk management, finance and insurance. For instance, they appear in portfolio optimization models (e.g., Markowitz' framework), models for setting capital requirements (e.g., Solvency II directive), and credit risk portfolio management (e.g., KMV's credit risk portfolio framework). Clearly, it is often difficult to accurately estimate all pairwise correlations, i.e., full information is rarely available. For instance, in a credit risk context the pairwise default correlation between the default events of two obligors is essentially equivalent to the probability they default together. As defaults are rare events such joint probability is very hard to estimate with a reasonable degree of precision. Simulating correlation matrices that are consistent with all available, yet incomplete, information is then useful to provide insight into the sensitivity of the model used. Indeed, while pairwise correlations might not be (all) known, information might be available at a more aggregate level. Specifically, the average correlation might be known with sufficient degree of accuracy. In this paper, we aim to describe the set of all correlation matrices that are consistent with one or more average constraints. Specifically, we provide an algorithm to randomly simulate from this set.

We are not the first to describe sets of correlation matrices that satisfy some constraints. Marsaglia and Olkin (1984) discuss how to generate correlation matrices with given eigenvalues. Joe (2006) proposes a method to generate correlation matrices based on partial correlations, which was extended in Lewandowski,

*KU Leuven, Belgium

†Vrije Universiteit Brussel, Belgium.

Please contact me when you wish to obtain the R-code for the algorithms we provide in this letter.

‡Maxwell Institute for Mathematical Sciences and School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, U.K.

Kurowicka and Joe (2009) using a vine method. To account for the uncertainty on the input correlations, Hardin, Garcia and Golan (2014) propose to add in a controlled manner a noisy matrix to the input correlation matrix. Our paper is closest to the one of Hüttner and Mai (2019) who propose an algorithm to generate correlation matrices with Perron-Frobenius property. As this property is known to often hold in financial data series such algorithm is highly relevant for applications in finance. In this letter, we propose an algorithm to generate random correlation matrices with given average correlation. We extend the results to the case of multiple average constraints. To the best of our knowledge, we are the first to study this problem.

The rest of the letter is organized as follows. In Section 2, we develop an algorithm to generate correlation matrices with given average. Section 3 provides the extension to further account for multiple average constraints. We provide pseudo-code and numerical examples showing that our algorithms work well. Section 4 concludes the paper.

2 Correlation Matrices with a Given Average

2.1 Problem formulation

Let $n \geq 2$ be an integer. We denote by $M_{n\times n}(\mathbb{R})$ the set consisting of $n \times n$ matrices in $\mathbb{R}$. A correlation matrix $C \in M_{n\times n}(\mathbb{R})$ is a positive semidefinite matrix that is symmetric and which has ones on the diagonal. We denote by $\langle \cdot, \cdot \rangle$ the standard inner product on $\mathbb{R}^n$, i.e., for all $x, y \in \mathbb{R}^n$, $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$. Furthermore, $\|\cdot\|$ denotes the corresponding (Euclidean) norm, i.e., $\|x\| = \sqrt{\langle x, x \rangle}$. In the sequel, we will also refer to the norm of a vector as its length. For $x, y \in \mathbb{R}^n$, the so-called triangle inequalities are well-known:

$$| \, \|x\| - \|y\| \, | \leq \|x - y\| \leq \|x\| + \|y\|.$$

Marsaglia and Olkin (1984) prove the following theorem, which provides a manner to generate correlation matrices matrix $C \in M_{n\times n}(\mathbb{R})$.

Theorem 1 (Marsaglia and Olkin (1984)) *A matrix $C \in M_{n\times n}(\mathbb{R})$ is a correlation matrix if and only if there exists a matrix $T \in M_{n\times n}(\mathbb{R})$ such that $C = TT^t$ where the rows of T considered as vectors in $\mathbb{R}^n$ have length 1. Moreover, one can assume the matrix T to be lower triangular, i.e. $T_{ij} = 0$ for all $i < j$, in which case T is unique.*

Theorem 1 is also known as Cholesky decomposition; see e.g. Gentle (2012).

We aim to design an algorithm that makes it possible to generate all correlation matrices that have a certain average correlation value, i.e, for given weights $\sigma_i > 0$ and a given (admissible) value for ρ we wish to generate all correlation matrices $C \in M_{n\times n}(\mathbb{R})$ such that

$$\frac{\sum_{i<j} \sigma_i \sigma_j C_{ij}}{\sum_{i<j} \sigma_i \sigma_j} = \rho \tag{1}$$

Clearly, fixing ρ is equivalent to fixing $\sum_{i,j=1}^{n} \sigma_i \sigma_j C_{ij}$, which can be seen as the variance of a random sum $X_1 + ... + X_n$ in which the components X_i are obtained from independent standard normals N_i via the transformation.

$$X_i = \sigma_i \sum_{j=1}^{n} T_{ij} N_j, i = 1, ..., n. \tag{2}$$

We formulate the following problem:

Problem I *Given an integer $n \geq 2$, positive real numbers $\sigma_1, \ldots, \sigma_n$ and a non-negative real number s, generate correlation matrices C satisfying the variance constraint*

$$s^2 = \sum_{i,j=1}^{n} C_{ij} \sigma_i \sigma_j.$$

In what follows we will call $\sum_{i,j=1}^{n} C_{ij} \sigma_i \sigma_j$ the total variance of the matrix C. Note that if C is the solution to Problem I, i.e., when C has total variance equal to s^2 then C has average correlation ρ_s given as

$$\rho_s = \frac{s^2 - \sum_{i=1}^{n} \sigma_i^2}{2 \sum_{i<j} \sigma_i \sigma_j}. \tag{3}$$

The problem we deal with can thus be interpreted as finding matrices T that linearly transform (via (2)) standard independent normal random variables N_i into jointly normal variables X_i that have zero mean, standard deviation σ_i and satisfy the constraint $var(X_1 + X_2 + ... + X_n) = s^2$. If we denote the rows of T by $t_i \in \mathbb{R}^n$, we can write $Cov(X_i, X_j) = \sigma_i \sigma_j \langle t_i, t_j \rangle$ and we obtain that

$$s^2 \quad = \quad \|\sigma_1 t_1 + \sigma_2 t_2 + ... + \sigma_n t_n\|^2$$

must hold. In other words, to solve Problem I we only need to generate n vectors $t_i \in \mathbb{R}^n$ (rows of T) having length one and satisfying $s^2 = \|\sigma_1 t_1 + \sigma_2 t_2 + ... + \sigma_i t_n\|^2$. Clearly, Problem I may not be well posed, for instance when s is larger than $\sum_i \sigma_i$. We characterize the admissible values of s in Lemma 1.

Lemma 1 (Well-posedness) *Problem I has a solution if and only if*

$$\max\{\sigma_{i_{max}} - \sum_{i \neq i_{max}} \sigma_i, 0\} \leq s \leq \sum_{i=1}^{n} \sigma_i,$$

where $1 \leq i_{max} \leq n$ denotes an index for which $\sigma_{i_{max}}$ is maximal among all σ_i.

Proof.

Suppose there exists a solution, then $s = \|\sum_{i=1}^{n} \sigma_i t_i\|$ is clearly non-negative and by the triangle inequalities satisfies

$$\sigma_{i_{max}} - \sum_{i \neq i_{max}} \sigma_i \leq s \leq \sum_{i=1}^{n} \sigma_i.$$

Conversely, it is clear that for all s satisfying these conditions a solution exists. ■

2.2 The solution (algorithm) to Problem I

To solve Problem 1 we iteratively construct $t_i \in \mathbb{R}^n$ $(i = 1, ..., n)$ such that the subsequent partial sums $\sigma_1 t_1 + \ldots + \sigma_i t_i$ have an admissible length in that $\|\sigma_1 t_1 + \ldots + \sigma_n t_n\| = s$ can be attained. We formally define

$$l_i = \|\sigma_1 t_1 + \ldots + \sigma_i t_i\|, i = 1, 2, ..., n. \tag{4}$$

and note that the l_i can also be seen as the variances of the partial sums $S_i = X_1 + ... + X_i$. Furthermore,

$$l_1 = \sigma_1 \text{ and } l_n = s \tag{5}$$

must hold and because of the triangle inequalities we also have that

$$|l_i - l_{i-1}| \leq \sigma_i \leq l_i + l_{i-1}, 2 \leq i \leq n. \tag{6}$$

We proceed in two steps. First, we explain how to generate a series of admissible lengths $l_1, ..., l_n$ in a random way. Using the $l_1, \ldots, l_n$ as input, we are then able to iteratively generate vectors $t_1, \ldots, t_n$ in $\mathbb{R}^n$ in a second step such that $\|t_i\| = 1$ and $\|\sigma_1 t_1 + \ldots + \sigma_i t_i\| = l_i$ for all $1 \leq i \leq n$. In particular, $\|\sigma_1 t_1 + \ldots + \sigma_i t_n\| = s$, i.e., $C = TT^t$ in which T is a matrix with rows t_i solves problem I.

Theorem 2 tells us when a partial sequence $l_1, \ldots, l_k$ $(k < n)$ can be extended to obtain a complete sequence $l_1, ..., l_n$ and provides a manner to generate C that solves Problem I.

Theorem 2 *Suppose that $\sigma_1 \leq \sigma_2 \leq \ldots \leq \sigma_n$. Let $1 \leq k \leq n-2$ be an integer and $l_1, \ldots, l_k$ non-negative real numbers that satisfy $l_1 = \sigma_1$ and $|l_i - l_{i-1}| \leq \sigma_i \leq l_i + l_{i-1}$ for all $2 \leq i \leq k$. Then the following two are equivalent:*

1. *There exist non-negative real numbers $l_{k+1}, \ldots, l_n$ with $l_n = s$ such that $|l_i - l_{i-1}| \leq \sigma_i \leq l_i + l_{i-1}$ for all $2 \leq i \leq n$.*
2. *$|s - l_k| \leq \sigma_{k+1} + \ldots + \sigma_n$ and $s + l_k \geq \sigma_n - (\sigma_{k+1} + \ldots + \sigma_{n-1})$.*

Proof.
(1. ⇒ 2.) If $l_{k+1}, \ldots, l_n$ satisfying the conditions exist, then

$$|s - l_k| = |l_n - l_k| \leq |l_{k+1} - l_k| + \ldots + |l_n - l_{n-1}| \leq \sigma_{k+1} + \ldots + \sigma_n,$$
$$s + l_k = (l_n + l_{n-1}) + (l_k - l_{k+1}) + \ldots + (l_{n-2} - l_{n-1}) \geq \sigma_n - (\sigma_{k+1} + \ldots + \sigma_{n-1}).$$

(2. ⇒ 1.) We construct the l_i inductively for $k+1 \leq i \leq n-1$, by choosing

$$l_i \in \left[\max\{|l_{i-1} - \sigma_i|, s - \sum_{j=i+1}^{n} \sigma_j, \sigma_n - \sum_{j=i+1}^{n-1} \sigma_j - s\}, \min\{l_{i-1} + \sigma_i, s + \sum_{j=i+1}^{n} \sigma_j\}\right]$$

arbitrarily. Note that this closed interval has been obtained by imposing all conditions from 1. and 2. on l_i. From the assumptions

$$|s - l_{i-1}| \leq \sum_{j=i}^{n} \sigma_j \text{ and } s + l_{i-1} \geq \sigma_n - \sum_{j=i}^{n-1} \sigma_j$$

it follows that this interval is non-empty. Here we are also using that the σ_i are increasing. It follows from the definition of l_i that

1. $|l_i - l_{i-1}| \leq \sigma_i \leq l_i + l_{i-1}$,
2. $|s - l_i| \quad \leq \sigma_{i+1} + \ldots + \sigma_n$,
3. $s + l_i \quad \geq \sigma_n - (\sigma_{i+1} + \ldots + \sigma_{n-1})$.

Finally, using the definition of l_{n-1} it is easily checked that $l_n = s$ satisfies $|l_n - l_{n-1}| \leq \sigma_n \leq l_n + l_{n-1}$. ■

Note that in Theorem 2 the assumption of increasing σ_i is made without loss of generality.

We use Theorem 2 to provide an explicit algorithm for generating feasible lengths $l_1, \ldots, l_n$.

Algorithm 1a : Step 1 - Generation of series of admissible lengths

given $\sigma_1, \ldots, \sigma_n$ increasing and s such that
$\sigma_n - (\sigma_1 + \ldots + \sigma_{n-1}) \leq s \leq \sigma_1 + \ldots + \sigma_n$
generate $l_1, \ldots, l_n$
$l_1 := 1$
for $i := 2$ to $n - 1$ do
$l_i :=$ random in $[\max\{s - \sum_{j=i+1}^{n} \sigma_j, \sigma_n - \sum_{j=i+1}^{n-1} \sigma_j - s, |l_{i-1} - \sigma_i|\},$
$\min\{s + \sum_{j=i+1}^{n} \sigma_j, l_{i-1} + \sigma_i\}]$
end for
$l_n := s$
return $l_1, \ldots, l_n$ =0

Note that when implementing this algorithm, one still needs to choose a distribution of the l_i on the intervals of allowed values for l_i. A natural first choice could be to take the uniform distribution on these intervals. However, other choices could be envisaged as well.

The following lemma tells us how to generate a vector t_k based on lengths $l_1, \ldots, l_k$ and vectors $t_1, \ldots, t_{k-1}$.

Lemma 2 *Assume that $t_1, \ldots, t_{k-1}$ are vectors in $\mathbb{R}^n$ such that $\|t_i\| = 1$ and $\|u_i\| = l_i$ for all $1 \leq i \leq k-1$ in which $u_i := \sigma_1 t_1 + \ldots + \sigma_i t_i$. Then we have the following equivalence:*

$$\|u_k\| = l_k \Leftrightarrow \langle t_k, u_{k-1} \rangle = \frac{l_k^2 - l_{k-1}^2 - \sigma_k^2}{2\sigma_k} \tag{7}$$

Proof. This follows easily from:

$$\begin{aligned}\|u_k\|^2 &= \|u_{k-1} + \sigma_k t_k\|^2 \\ &= \|u_{k-1}\|^2 + \|\sigma_k t_k\|^2 + 2\langle \sigma_k t_k, u_{k-1}\rangle \\ &= l_{k-1}^2 + \sigma_k^2 + 2\sigma_k \langle t_k, u_{k-1}\rangle\end{aligned}$$

■

We now provide the pseudo-code for implementing the second step of our approach, i.e., the generation of row vectors t_i that compose a matrix T such that $C = TT^t$ solves Problem I.

Algorithm 1b : Step 2 - Generation of row vectors t_i

```
generate t_1, ..., t_n, given l_1, ..., l_n
t_1 := random in R^n of length 1
u_1 := σ_1 t_1
for i := 2 to n do
x := random in R^n of length 1
y := x − u_{i−1}⟨x, u_{i−1}⟩/‖u_{i−1}‖^2
z := (u_{i−1}/‖u_{i−1}‖^2)(l_i^2 − l_{i−1}^2 − σ_i^2)/(2σ_i)
t_i := z + (sqrt(1 − ‖z‖^2)/‖y‖)y
u_i := u_{i−1} + σ_i t_i
end for
return t_1, ..., t_n =0
```

Proof of correctness. Note that in the loop y is orthogonal to u_{i-1} which is parallel to z and $\langle z, u_{i-1}\rangle = (l_i^2 - l_{i-1}^2 - \sigma_i^2)/(2\sigma_i)$. Since t_i differs from z by a multiple of y, we have $\langle t_i, u_{i-1}\rangle = \langle z, u_{i-1}\rangle$. Finally, one easily checks that $\langle t_i, t_i\rangle = 1$. ■

The matrix T is then composed by the row vectors $t_1, \ldots, t_n$ and the correlation matrix $C = TT^t$ is then a solution to Problem I.

We present an example.

Example 1 *We provide an example of how to compute a random 6×6 correlation matrix C with average correlation $\rho = 0.2$ using Algorithms 1a and 1b. We use uniform distributions when simulating the lengths l_i. This gives as a particular solution,*

$$C = \begin{pmatrix} 1.0000000 & -0.7669058 & 0.1442050 & 0.1372040 & 0.3284773 & 0.6463319 \\ -0.7669058 & 1.0000000 & -0.6059352 & -0.2850269 & -0.3895216 & -0.6645325 \\ 0.1442050 & -0.6059352 & 1.0000000 & 0.7252801 & 0.7406292 & 0.6270028 \\ 0.1372040 & -0.2850269 & 0.7252801 & 1.0000000 & 0.9604118 & 0.6341328 \\ 0.3284773 & -0.3895216 & 0.7406292 & 0.9604118 & 1.0000000 & 0.7682472 \\ 0.6463319 & -0.6645325 & 0.6270028 & 0.6341328 & 0.7682472 & 1.0000000 \end{pmatrix}$$

3 Multiple Average Constraints

We now want to generate random correlation matrices with some blocks (submatrices) each satisfying an average constraint. To make this more precise, suppose that $n = n_1 + \ldots + n_k$ are all positive integers and that we are given an $n_i \times n_i$ correlation matrix C_i for each $1 \leq i \leq k$, where each C_i satisfies an average constraint, as discussed in the previous section.

In this section, we specifically denote by $T_i \in M_{n_i \times n_i}(\mathbb{R})$ the lower triangular matrix such that $C_i = T_i T_i^t$. The rows of each T_i are denoted as vectors $t_{ij} \in \mathbb{R}^{n_i}$, $j = 1, 2, ..., n_i$. Finally, we define $v_i \in \mathbb{R}^{n_i}$ as

$$v_i = \sigma_{i1} t_{i1} + \ldots + \sigma_{in_i} t_{in_i}, \tag{8}$$

and denote

$$\|v_i\| = s_i, 1 \leq i \leq k \tag{9}$$

i.e., s_i is the total variance of the correlation matrix C_i.

We first explain how to generate a random $n \times n$ correlation matrix C which contains the C_i as consecutive blocks along the diagonal. Next, we explain how to do this in the presence of an additional variance constraint (equivalently, an average correlation constraint) on C.

3.1 Block structures with given averages

Our problem reads as follows.

Problem II *Generate random correlation matrices C containing the C_i as consecutive blocks along the diagonal.*

To solve this problem we will use orthogonal matrices. We first introduce some notation and definitions.

Definition 1 *Let $m < n$ be positive integers. We will always consider $\mathbb{R}^m$ to be a vector subspace of $\mathbb{R}^n$ by embedding it onto the first m components, i.e. by sending the vector $(x_1, \ldots, x_m)$ to the vector $(x_1, \ldots, x_m, 0, \ldots, 0)$. Note that this choice of embedding preserves the standard inner product, by which we mean that $\langle x, y \rangle$ does not depend on whether we consider $x, y \in \mathbb{R}^m$ to be elements of $\mathbb{R}^m$ or $\mathbb{R}^n$.*

Definition 2 *For a positive integer n, we let $O(n)$ denote the orthogonal group in dimension n, i.e. a matrix $M_{n \times n} \in O(n)$ if and only if it preserves the standard inner product on $\mathbb{R}^n$, by which $\langle Mv, Mw \rangle = \langle v, w \rangle$ for all vectors $v, w \in \mathbb{R}^n$.*

The following lemma is crucial in solving Problem II.

Lemma 3 *Let m, n be positive integers. Suppose that $v_1, \ldots, v_m$ and $w_1, \ldots, w_m$ are vectors in $\mathbb{R}^n$ such that $\langle v_i, v_j \rangle = \langle w_i, w_j \rangle$ for all $1 \leq i, j \leq m$. Then there exists an orthogonal matrix $M \in O(n)$ such that $Mv_i = w_i$ for all $1 \leq i \leq m$.*

Proof. Let V be the vector space spanned by $v_1, \ldots, v_m$ and W the vector space spanned by $w_1, \ldots, w_m$. Note that if $c_1v_1 + \ldots + c_mv_m = 0$ for some $c_1, \ldots, c_m \in R$ then $c_1w_1 + \ldots + c_mw_m = 0$ as well. This follows from the fact that a vector $u \in \mathbb{R}^n$ is zero if and only if $\langle u, u \rangle = 0$. In particular this implies that $\dim(V) = \dim(W)$. Thus, there exists a unique linear transformation $L : V \to W$ such that $L(v_i) = w_i$ for all $1 \leq i \leq m$. Let $e_1, \ldots, e_l$ be an orthogonal basis for the orthogonal complement of V in $\mathbb{R}^n$ and $f_1, \ldots, f_l$ an orthogonal basis for the orthogonal complement of W in $\mathbb{R}^n$. We extend L to all of $\mathbb{R}^n$ by setting $L(e_i) = f_i$ for all $1 \leq i \leq l$. By construction, L preserves all inner products on $\mathbb{R}^n$. Therefore, its matrix M is indeed an element of $O(n)$. ■

Now we are ready to present the algorithm that solves Problem II.

Algorithm 2

generate C with blocks $C_1, \ldots, C_k$ along the diagonal
for $i := 1$ **to** k **do**
Let T_i be the lower triangular matrix (with rows t_{ij}) such that $C_i = T_iT_i^t$
$M_i :=$ random in $O(n)$
$\tau_{ij} := M_it_{ij}$
end for
$T :=$ matrix with rows τ_{ij}
$C := TT^t$
return C =0

Remark 1 *Note that we have not specified the distribution of a random element of $O(n)$. However, in this case (unlike in the previous section) there is a very natural choice of distribution. Since $O(n)$ is a compact topological group, it carries a unique (left) translation invariant probability measure called the Haar measure. Note that there are good algorithms available to generate random orthogonal matrices according to this distribution (Stewart (1980)).*

Example 2 *Suppose we are looking for a 6×6 correlation matrix C such that the 3×3 upper left corner has average correlation 0.2 and the 3×3 lower right corner average correlation 0. Using Algorithms 1a and 1b that we developed in Section 1 we obtain for the upper left block*

$$C_1 = \begin{pmatrix} 1.0000000 & 0.1307564 & 0.8562048 \\ 0.1307564 & 1.0000000 & -0.3869612 \\ 0.8562048 & -0.3869612 & 1.0000000 \end{pmatrix}$$

and for the lower right block

$$C_2 = \begin{pmatrix} 1.00000000 & 0.03949805 & 0.4043501 \\ 0.03949805 & 1.00000000 & -0.4438482 \\ 0.40435012 & -0.44384816 & 1.0000000 \end{pmatrix}$$

To find a correlation matrix C with C_1 and C_2 as blocks along the diagonal, we *now apply Algorithm 2 and we obtain*

$$C = \begin{pmatrix} 1.0000000 & 0.1307564 & 0.85620478 & 0.73725058 & 0.51956587 & 0.0912944 \\ 0.1307564 & 1.0000000 & -0.38696122 & -0.43638596 & 0.84869870 & -0.7474391 \\ 0.8562048 & -0.3869612 & 1.00000000 & 0.87710025 & 0.06164776 & 0.5356547 \\ 0.7372506 & -0.4363860 & 0.87710025 & 1.00000000 & 0.03949805 & 0.4043501 \\ 0.5195659 & 0.8486987 & 0.06164776 & 0.03949805 & 1.00000000 & -0.4438482 \\ 0.0912944 & -0.7474391 & 0.53565467 & 0.40435012 & -0.44384816 & 1.0000000 \end{pmatrix}$$

3.2 Adding a global average constraint

As before, let $n = n_1 + \ldots + n_k$ be positive integers. we want to generate a random $n \times n$ correlation matrix C which contains $n_i \times n_i$ correlation matrices C_i as consecutive blocks and which also satisfies a global average constraint. Our problem can be formulated as follows:

Problem III *Generate random correlation matrices C containing the C_i, $i = 1, 2, ..., k$ as consecutive blocks along the diagonal and satisfying the variance constraint s^2, i.e., for given weights σ_{ij}, $i = 1, 2, ..., k; j = 1, 2, ..., n_i$ and given $s > 0$, C must also satisfy*

$$s^2 = \sum_{i,j=1}^{n} C_{ij}\sigma_i\sigma_j. \tag{10}$$

Let $C = TT^t$ solve Problem III and denote by $\tau_{ij} \in \mathbb{R}^n$ the rows of T (first ordered by $1 \le i \le k$ then $1 \le j \le n_i$). In what follows, we also specifically denote the weights by $\sigma_{ij}(1 <= i <= k, 1 <= j <= n_i)$; that is, σ_{ij} is equal to σ_l where $l = n_1 + ...n_{i-1} + j$ with the convention that $n_1 + n_0 = 0$. Let $w_i = \sigma_{i1}\tau_{i1} + \ldots + \sigma_{in_i}\tau_{in_i}$, $1 \le i \le k$. As C solves Problem II, it holds that $\|w_1 + \ldots + w_k\| = s$ and $\|w_i\| = s_i$ for all $1 \le i \le k$. The following theorem shows how to derive from the given matrices T_i a matrix T with these properties. In particular, $C = TT^t$ then solves Problem III.

Theorem 3 *Let $w_1, \ldots, w_k \in \mathbb{R}^n$ be vectors such that $\|w_1 + \ldots + w_k\| = s$ and $\|w_i\| = s_i$ for all $1 \le i \le k$. Then for all $1 \le i \le k$ there exists an orthogonal matrix $M_i \in O(n)$ such that $w_i = M_i v_i$. For all $1 \le i \le k$ and $1 \le j \le n_i$ define $\tau_{ij} \in \mathbb{R}^n$ by $\tau_{ij} = M_i t_{ij}$. Let T be the matrix with rows τ_{ij} (first ordered by i then j). Then the matrix $C = TT^t$ is a solution to Problem III.*

Conversely, suppose that the matrix C is a solution to Problem III and let T be a matrix such that $C = TT^t$. Let $\tau_{ij} \in \mathbb{R}^n$ be the rows of T (first ordered by $1 \le i \le k$ then $1 \le j \le n_i$). For all $1 \le i \le k$, let $w_i \in \mathbb{R}^n$ denote the vector $w_i = \sigma_{i1}\tau_{i1} + \ldots + \sigma_{in_i}\tau_{in_i}$. Then $\|w_1 + \ldots + w_k\| = s$ and $\|w_i\| = s_i$ for all $1 \le i \le k$. Moreover, for each $1 \le i \le k$ there exists an orthogonal matrix $M_i \in O(n)$ such that $M_i t_{ij} = \tau_{ij}$ for all $1 \le j \le n_i$.

Proof. If $\|w_i\| = s_i$, then $\langle v_i, v_i\rangle = \langle w_i, w_i\rangle$. Therefore, by Lemma 3 for all $1 \le i \le k$ there exists $M_i \in O(n)$ such that $M_i v_i = w_i$. Since M_i is orthogonal and $\tau_{ij} = M_i t_{ij}$ the inner products between τ_{ij} are the same as between t_{ij} for any fixed value of $1 \leq i \leq k$. Therefore, the matrix C contains C_i as consecutive blocks along the diagonal and the total variance of C is equal to

$$\| \sum_{\substack{1 \leq i \leq k \\ 1 \leq j \leq n_i}} \sigma_{ij} \tau_{ij} \|^2 = \| \sum_{i=1}^{k} w_k \|^2 = s^2,$$

so that the matrix C is indeed a solution to Problem III.

For the converse, let C be a solution to Problem III and let T be a matrix such that $C = TT^t$. Moreover, let τ_{ij} be the rows of T (first ordered by $1 \leq i \leq k$ then $1 \leq j \leq n_i$). Since the matrix C contains C_i as consecutive blocks along the diagonal, the inner products between τ_{ij} are the same as between t_{ij} for any fixed value of $1 \leq i \leq k$. Therefore, by Lemma 3 for all $1 \leq i \leq k$ there exists an orthogonal matrix $M_i \in O(n)$ such that $M_i t_{ij} = \tau_{ij}$ for all $1 \leq j \leq n_i$. In particular, we have that $\|w_i\| = \|v_i\| = s_i$ for all $1 \leq i \leq k$. The equality $\|w_1 + \ldots + w_k\| = s$ follows directly from the fact that the matrix C has total variance s^2. ■

According to Theorem 3, Problem III can be solved in two steps:

1. Generate vectors $w_1, \ldots, w_k \in \mathbb{R}^n$ such that $\|w_1 + \ldots + w_k\| = s$ and $\|w_i\| = s_i$ for all $1 \leq i \leq k$.

2. Generate $M_1, \ldots, M_k \in O(n)$ such that $M_i v_i = w_i$ (this is always possible). Let T be the matrix with rows $\tau_{ij} = M_i t_{ij}$ (first ordered by $1 \leq i \leq k$ then $1 \leq j \leq n_i$) and define $C = TT^t$.

Step 1 is similar to approach we used to solve Problem I in Section 1. There we generated $t_1, \ldots, t_n \in \mathbb{R}^n$ such that $\|\sigma_1 t_1 + \ldots + \sigma_n t_n\| = s$ and $\|t_i\| = 1$ for all $1 \leq i \leq n$. Here we want to generate $w_1, \ldots, w_k \in \mathbb{R}^n$ with $k \leq n$ such that $\|w_1 + \ldots + w_k\| = S$ and $\|w_i\| = S_i$. Taking $n = k$, $\sigma_i = S_i$ and $t_i = w_i / S_i$ this is equivalent, apart from the fact that $w_1, \ldots, w_k$ are to be contained in $\mathbb{R}^n$ and not $\mathbb{R}^k$. However, this last problem is easily solved by generating $w_1, \ldots, w_k \in \mathbb{R}^k$, embedding them into $\mathbb{R}^n$ as in Definition 1 and multiplying all by a single random orthogonal matrix $M \in O(n)$.

Moreover, in a similar way as in Theorem 1 in Section 2, s_i and s have to be admissible to generate the corresponding correlation matrices. Thus, we have the following theorem.

Theorem 4 *Problem III has a solution if and only if*

$$s_{i_{max}} - \sum_{i \neq i_{max}} s_i \leq s \leq \sum_{i=1}^{k} s_i,$$

where i_{max} is an index for which s_i is maximal.

Proof. We know from (the proof of) Theorem 3 that Problem III has a solution if and only if there exist $w_1, \ldots, w_k \in \mathbb{R}^n$ such that $\|w_1 + \ldots + w_k\| = s$

and $\|w_i\| = s_i$ for all $1 \leq i \leq k$. Therefore, the result follows from Theorem 1. ■

Algorithm 3

---

generate C with blocks $C_1, \ldots, C_k$ along the diagonal and total variance s^2
generate $t_1, \ldots, t_k \in \mathbb{R}^k$ with $\|t_i\| = 1$ such that $\|s_1 t_1 + \ldots + s_k t_k\| = s$ (as in Section 2) and set $w_i := s_i t_i$ for all $1 \leq i \leq k$
$M :=$ random in $O(n)$.
for $i := 1$ to k do
$w_i := M w_i$ // consider w_i as an element of $\mathbb{R}^n$
Compute T_i (with rows t_{ij}) from C_i
$v_i := \sigma_{i1} t_{i1} + \ldots + \sigma_{in_i} t_{in_i}$
$M_i :=$ random in $O(n)$ such that $M_i v_i = w_i$
for $j := 1$ to n_i do
$\tau_{ij} := M_i t_{ij}$
end for
end for
$T :=$ matrix with rows τ_{ij}
$C := TT^t$
return C =0

We present numerical examples with block structure.

Remark 2 *The complexity of all our algorithms is driven by the complexity of the computation of the matrix product $C = T * T^t$. This complexity is at least $O(n^2)$ and, by the latest results on fast matrix multiplication, it is at most $O(n^{2.3728639})$; see, Le Gall (2014).*

Example 3 *Generate a correlation matrix $C_{6\times 6}$ with upper 3×3 sub-average correlation 0.2, lower 3×3 sub-average correlation 0.4 and total average correlation $\rho = 0.25$. We show 2 samples of simulation.*

$$\begin{pmatrix} 1 & 0.44953864 & 0.4383173 & 0.6524330 & 0.2731052 & 0.60398060 \\ 0.44953864 & 1 & -0.2878560 & -0.1556668 & -0.1022461 & 0.08723503 \\ 0.4383173 & -0.2878560 & 1 & 0.3080049 & 0.3916332 & -0.10847894 \\ 0.6524330 & -0.1556668 & 0.3080049 & 1 & 0.6190522 & 0.68384052 \\ 0.2731052 & -0.1022461 & 0.3916332 & 0.6190522 & 1 & -0.10289275 \\ 0.60398060 & 0.08723503 & -0.10847894 & 0.68384052 & -0.10289275 & 1 \end{pmatrix};$$

$$\begin{pmatrix} 1 & 0.34104551 & 0.29730654 & 0.0995247 & -0.1172317 & 0.1399298 \\ 0.34104551 & 1 & -0.03835204 & 0.6223703 & 0.4797049 & 0.4089176 \\ 0.29730654 & -0.03835204 & 1 & -0.5257222 & 0.4260202 & 0.4164864 \\ 0.0995247 & 0.6223703 & -0.5257222 & 1 & 0.2634622 & 0.2275388 \\ -0.1172317 & 0.4797049 & 0.4260202 & 0.2634622 & 1 & 0.7089991 \\ 0.1399298 & 0.4089176 & 0.4164864 & 0.2275388 & 0.7089991 & 1 \end{pmatrix}.$$

4 Conclusion

In this letter, we propose an algorithm to generate correlation matrices subject to a given average correlation. We extend the algorithm to account for correlation matrices with block structures. Our algorithms are useful for stress-testing models. Specific applications to model risk assessment are considered in further research. Finally, we wish to point out that the problem of finding correlation matrices with given average that we consider in this paper is strongly related to the so-called problem of joint mixability, which has been studied in a series of papers initiated by Wang and Wang (2011) and Wang et al. (2013) with application to the assessment of worst-case portfolio models under (partial) dependence uncertainty; see e.g., Bernard et al. (2017a), Bernard et al. (2017b), Embrechts et al. (2013), Puccetti et al. (2017) and Bernard, Rüschendorf and Vanduffel (2017). Specifically, A tuple of univariate distributions is said to be jointly mixable if there exist random variables, with respective distributions, such that their sum is a constant. The results in this paper make it possible to describe all these joint mixes when the marginal distributions are normally distributed and their dependence is required to be gaussian.

Acknowledgement Jing Yao acknowledges the National Natural Science Foundation of China (No.11971506).

References

[1] Bernard, C., Rüschendorf, L., Vanduffel, S., Yao, J. (2017). How robust is the value-at-risk of credit risk portfolios? The European Journal of Finance, 23(6), 507-534.

[2] Bernard, C., Rüschendorf, L., Vanduffel, S. (2017a). Value-at-risk bounds with variance constraints. Journal of Risk and Insurance, 84(3), 923-959.

[3] Bernard, C., Rüschendorf, L., Vanduffel, S., Wang, R. (2017b). Risk bounds for factor models. Finance and Stochastics, 21(3), 631-659.

[4] Embrechts, P., Puccetti, G., Rüschendorf, L. (2013). Model uncertainty and VaR aggregation. Journal of Banking Finance, 37(8), 2750-2764.

[5] Gentle, James E. (2012). Numerical linear algebra for applications in statistics. Springer Science Business Media, 2012.

[6] Hardin, J., Garcia, S. R. & Golan, D. (2013). A method for generating realistic correlation matrices. The Annals of Applied Statistics, 1733-1762.

[7] Hüttner, A., Mai, J. F. (2019). Simulating realistic correlation matrices for financial applications: correlation matrices with the PerronCFrobenius property. Journal of Statistical Computation and Simulation, 89(2), 315-336.

[8] Joe, H. (2006). Generating random correlation matrices based on partial correlations. Journal of Multivariate Analysis, 97(10), 2177-2189.

[9] Le Gall, F. (2014, July). Powers of tensors and fast matrix multiplication. In Proceedings of the 39th international symposium on symbolic and algebraic computation (pp. 296-303).

[10] Lewandowski, D., Kurowicka, D. & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. Journal of multivariate analysis, 100(9), 1989-2001.

[11] Marsaglia, G. and Olkin, I. (1984). Generating correlation matrices. SIAM J. Sci. Statist. Comput., 5(2):470–475.

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

[13] Stewart, G.W. (1980). The efficient generation of random orthogonal matrices with an application to condition estimators. SIAM J. Numer. Anal., 17(3):403–409.

[14] Wang, B., Wang, R. (2011). The complete mixability and convex minimization problems with monotone marginal densities. Journal of Multivariate Analysis, 102(10), 1344-1360.

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