GENERATING VALID $4 \times 4$ CORRELATION MATRICES

MARK BUDDEN, PAUL HADAVAS, LORRIE HOFFMAN, AND CHRISTOPHER PRETZ

Abstract. In this article, we provide an algorithm for generating valid $4 \times 4$ correlation matrices by creating bounds for the remaining three correlations that insure positive semidefiniteness once correlations between one variable and the other three are supplied. This is achieved by attending to the constraints placed on the leading principal minor determinants. Prior work in this area is restricted to $3 \times 3$ matrices or provides a means to investigate whether a $4 \times 4$ matrix is valid but does not offer a method of construction. We do offer such a method and supply a computer program to deliver a $4 \times 4$ positive semidefinite correlation matrix.

1. Introduction

The level of sophistication and complexity of statistical inquiry has placed demands on researchers to increasingly employ simulation techniques. Frequently, distributions that reflect the essence of an application are assumed to be multivariate normal with a general prior knowledge of the correlation structure. More theoretical investigations such as those in the area of Bayesian analysis rely on Markov Chain Monte Carlo methods to investigate posterior distributions. Liechty, Liechty and Muller [4] encounter a challenge when conducting their simulations due to the ‘awkward manner in which $r_{ij}$ is embedded in the likelihood’ (p. 6) noting that the requirement of positive semidefiniteness is the constraint that imposes an analysis dealing with truncated distributions. This leads them to develop a procedure that generates normalizing constants. This positive semidefinite matrix complication is encountered throughout the applied literature and sometimes leads to erroneous outcomes when simulation results are aggregated while using an invalid correlation matrix.

The need to forecast demand for a group of products in order to realize savings by properly managing inventories is one such application requiring use of correlation matrices. Xu and Evers [10] offer a proof that partial pooling of customers can not lead to fewer resource requirements than complete pooling as was indicated in Tyagi and Das [9]. They further elucidate their position by showing that the error in the examples offered by these authors centered on the use of infeasible correlation matrices. Xu and Evers reveal that the matrices are not positive semidefinite and do so in the three variable case by revisiting boundary conditions offered by Marsaglia and Olkin [5] and via a computer program they developed to produce eigenvalues for the four variable case. Xu and Evers indicate that such an oversight...
is understandable because of the complex nature and interrelationships inherent in correlation structures particularly for 'non-statisticians like us ([10], p. 301)'. We contend that even statisticians and mathematicians find the problem of developing feasible correlation matrices to be challenging.

The work done by Xu and Evers is helpful in that it prevents a researcher from offering an infeasible example or conducting a meaningless simulation due to non-positive semidefinite correlation matrices. Their approach, though, is not constructive, as they say 'a correlation matrix is checked, (but) the determination of boundaries necessary for constructing feasible matrices is not performed ([10], p. 305)'. We offer such a method in the four variable case.

The paper proceeds as follows. In the next section, previous results for $3 \times 3$ correlation matrices are explained. The generation of valid $4 \times 4$ correlation matrices is detailed in section 3. Finally, in section 4, we illustrate the utility of this constructive approach with an example from Xu and Evers [10]. Additionally, in section 4, other examples of valid $4 \times 4$ correlation matrices are provided along with directions for further research.

2. Review of $3 \times 3$ correlation matrices

The method described in this section for generating $3 \times 3$ correlation matrices has been noted by many sources. Although the bounds described here were known earlier, Stanley and Wang [8] gave the first proof of the bounds in 1969. Other (geometric) proofs were subsequently given by Glass and Collins [1] and Leung and Lam [5]. Olkin [6] investigated the general question of how correlations are restricted when a submatrix of a correlation matrix is fixed. More recently, Rousseeuw and Molenberghs [7] used these bounds to investigate the volume of the set of valid $3 \times 3$ correlation matrices.

Let $x_1$, $x_2$, and $x_3$ be random variables and let $r_{ij}$ denote the correlation coefficient for the variables $x_i$ and $x_j$ where $i, j \in \{1, 2, 3\}$. A $3 \times 3$ correlation matrix is a positive semidefinite matrix of the form

$$A = \begin{pmatrix}
1 & r_{12} & r_{13} \\
r_{12} & 1 & r_{23} \\
r_{13} & r_{23} & 1
\end{pmatrix}.$$

To produce a valid correlation matrix, we begin by randomly picking the correlation coefficients $-1 < r_{12} < 1$ and $-1 < r_{13} < 1$. Here, we exclude the possibilities of having correlations of $\pm 1$ since such correlations indicate the redundancy of one variable where it is a perfect linear combination of another and can thus be eliminated from any analysis. Once these coefficients are selected, the possible range of values for $r_{23}$ can be described by considering the determinant

$$\det A = -r_{23}^2 + (2r_{12}r_{13})r_{23} + (1 - r_{12}^2 - r_{13}^2) \geq 0.$$ 

This determinant is equal to 0 when

$$r_{23} = r_{12}r_{13} \pm \sqrt{(1 - r_{12}^2)(1 - r_{13}^2)}$$

and it obtains its maximum value (for $r_{12}$ and $r_{13}$ fixed) of

$$(1 - r_{12}^2)(1 - r_{13}^2) \leq 1.$$
when \( r_{23} = r_{12}r_{13} \). Thus, \( A \) is a correlation matrix exactly when \( r_{23} \) satisfies
\[
(1) \quad r_{12}r_{13} - \sqrt{(1 - r_{12}^2)(1 - r_{13}^2)} \leq r_{23} \leq r_{12}r_{13} + \sqrt{(1 - r_{12}^2)(1 - r_{13}^2)}.
\]
Again, we do not allow \( r_{23} = \pm 1 \) for the reason described above.

A special case to notice is when \( r_{12}r_{13} = 0 \). Without loss of generality, assume that \( r_{12} = 0 \). The range of possible values of \( r_{23} \) is given by
\[
-\sqrt{1 - r_{13}^2} \leq r_{23} \leq \sqrt{1 - r_{13}^2}.
\]
From this one sees that if \( r_{12} = 0 = r_{13} \), then \( r_{23} \) can take on any possible correlation.

3. Algorithmic Generation of 4 × 4 Correlation Matrices

We now consider a 4 × 4 matrix of the form
\[
A = \begin{pmatrix}
1 & r_{12} & r_{13} & r_{14} \\
r_{12} & 1 & r_{23} & r_{24} \\
r_{13} & r_{23} & 1 & r_{34} \\
r_{14} & r_{24} & r_{34} & 1
\end{pmatrix}.
\]
The correlations \( r_{12}, r_{13}, \) and \( r_{14} \) can be randomly picked from the interval \((-1, 1)\).

Here, we exclude the possibilities of have correlations of \( \pm 1 \) for the reason noted in the previous section. We note here that selections from intervals created herein should also avoid \( \pm 1 \).

To find suitable ranges for the other correlations, we use the fact that a symmetric matrix \( C \) is positive semidefinite if and only if all of its symmetric submatrices (including itself) have a nonnegative determinant (see [2]). In other words, \( A \) is a correlation matrix if and only if \( \det A \geq 0 \) and every matrix of the form
\[
A_{ijk} = \begin{pmatrix}
1 & r_{ij} & r_{ik} \\
r_{ij} & 1 & r_{jk} \\
r_{ik} & r_{jk} & 1
\end{pmatrix}
\]
is a correlation matrix for \( i, j, k \in \{1, 2, 3, 4\} \) (no two of \( i, j, \) and \( k \) are equal).

To find a range of possible values for \( r_{23} \), we will need to use the bounds described in (1). However, if we were to choose any value of \( r_{23} \) within this range, there is no guarantee that there will be any possible values of \( r_{24} \) and \( r_{34} \) for which there exists a valid correlation matrix. The reason for this is that the range given in (1) only takes into consideration the determinant of
\[
\begin{pmatrix}
1 & r_{12} & r_{13} \\
r_{12} & 1 & r_{23} \\
r_{13} & r_{23} & 1
\end{pmatrix}
\]
and neglects the determinants of \( A \) and
\[
\begin{pmatrix}
1 & r_{23} & r_{24} \\
r_{23} & 1 & r_{34} \\
r_{24} & r_{34} & 1
\end{pmatrix}.
\]

Taking this information into account, we describe the bounds of \( r_{23} \) as follows.
3.1. The Range of $r_{23}$. To find a lower bound $L_{23}$ for $r_{23}$, we need to take into account three possible limitations. We have already noted by (1) that $L_{23}^{(1)} \leq r_{23}$ where

$$L_{23}^{(1)} := r_{12} r_{13} - \sqrt{(1 - r_{12}^2)(1 - r_{13}^2)}.$$ 

We must also consider the determinant of

$$
\begin{pmatrix}
1 & r_{23} & r_{24} \\
r_{23} & 1 & r_{34} \\
r_{24} & r_{34} & 1
\end{pmatrix}
$$

which gives the lower bound $L_{23}^{(2)} \leq r_{23}$ where

$$L_{23}^{(2)} := r_{24} r_{34} - \sqrt{(1 - r_{24}^2)(1 - r_{34}^2)}$$

is minimized over all possible values of $r_{24}$ and $r_{34}$ within the ranges

(2) $r_{12} r_{14} - \sqrt{(1 - r_{12}^2)(1 - r_{14}^2)} \leq r_{24} \leq r_{12} r_{14} + \sqrt{(1 - r_{12}^2)(1 - r_{14}^2)}$

and

(3) $r_{13} r_{14} - \sqrt{(1 - r_{13}^2)(1 - r_{14}^2)} \leq r_{34} \leq r_{13} r_{14} + \sqrt{(1 - r_{13}^2)(1 - r_{14}^2)}.

There is one last limitation to $L_{23}$ that we must consider. We must be sure that for all values of $r_{23}$ within our range, there exists valid $r_{24}$ and $r_{34}$ within their respective ranges (2) and (3) such that the determinant of $A$ is nonnegative. Note that the determinant has the form

$$
\det A = (r_{14}^2 - 1)r_{23}^2 + (2r_{24} r_{34} + 2r_{12} r_{13} - 2r_{12} r_{14} r_{34} - 2r_{13} r_{14} r_{24}) r_{23} \\
+ (1 - r_{12}^2 - r_{13}^2 - r_{14}^2 - r_{24}^2 - r_{34}^2 + 2r_{12} r_{14} r_{24} + 2r_{13} r_{14} r_{34} \\
- 2r_{12} r_{13} r_{24} r_{34} + r_{12}^2 r_{34}^2 + r_{13}^2 r_{24}^2),
$$

which is a quadratic in $r_{23}$. The left-most root of this quadratic (minimized over the ranges of $r_{24}$ and $r_{34}$ given in (2) and (3)) gives us our third lower bound:

$$L_{23}^{(3)} := \frac{r_{24} r_{34} + r_{12} r_{13} - r_{12} r_{14} r_{34} - r_{13} r_{14} r_{24} - \sqrt{\det(A_{124}) \cdot \det(A_{134})}}{(1 - r_{14}^2)}.$$

In order for there to exist a valid correlation matrix, $r_{23}$ must be chosen to be greater than or equal to all three of the above described bounds. So we set

$$L_{23} := \max \left\{ L_{23}^{(1)}, L_{23}^{(2)}, L_{23}^{(3)} \right\}.$$

Finding the upper bound $U_{23}$ of $r_{23}$ is similar to the method used to find the lower bound. We first note that the correlation $r_{23}$ is subject to the constraint $r_{23} \leq U_{23}^{(1)}$ where

$$U_{23}^{(1)} := r_{12} r_{13} + \sqrt{(1 - r_{12}^2)(1 - r_{13}^2)}.$$

The second constraint to the upper bound of $r_{23}$ is that $r_{23} \leq U_{23}^{(2)}$ where

$$U_{23}^{(2)} := r_{24} r_{34} + \sqrt{(1 - r_{24}^2)(1 - r_{34}^2)}$$

is maximized over the ranges of $r_{24}$ and $r_{34}$ described in (2) and (3).
From the determinant of $A$ that was described in section (3.1), we also have $r_{23} \leq U_{23}^{(3)}$ where

$$U_{23}^{(3)} := \frac{r_{24} r_{34} + r_{12} r_{13} - r_{12} r_{14} r_{34} - r_{13} r_{14} r_{24} + \sqrt{\det(A_{124}) \cdot \det(A_{134})}}{(1 - r_{14}^2)},$$

which is maximized over the ranges described in (2) and (3).

Since $r_{23}$ must satisfy all of the above inequalities, we have that

$$U_{23} := \min \left\{ U_{23}^{(1)}, U_{23}^{(2)}, U_{23}^{(3)} \right\}.$$

As long as we pick a value of $r_{23}$ within the range

$$L_{23} \leq r_{23} \leq U_{23},$$

we are guaranteed that there exists a valid correlation matrix with the chosen values of $r_{12}$, $r_{13}$, $r_{14}$, and $r_{23}$.

3.2. The Range of $r_{24}$. In this section, we will describe the range of possible values of the correlation $r_{24}$, once $r_{12}$, $r_{13}$, $r_{14}$, and $r_{23}$ have been fixed using the method described above. Although the construction of the bounds is identical to that of $r_{23}$, we now have the added benefit that our optimization is over single variable functions.

As before, we first note that $L_{24}^{(1)} \leq r_{24}$ where

$$L_{24}^{(1)} := r_{12} r_{14} - \sqrt{(1 - r_{12}^2)(1 - r_{14}^2)}.$$

The second constraint to the lower bound of $r_{24}$ is that $L_{24}^{(2)} \leq r_{24}$ where

$$L_{24}^{(2)} := r_{23} r_{34} - \sqrt{(1 - r_{23}^2)(1 - r_{34}^2)}$$

is minimized over the range of $r_{34}$ described in (3).

Next, we consider the determinant of $A$, written as a quadratic in $r_{24}$:

$$\det A = (r_{13}^2 - 1) r_{24}^2 + (2 r_{23} r_{34} + 2 r_{12} r_{14} - 2 r_{12} r_{13} r_{34} - 2 r_{13} r_{14} r_{23}) r_{24}$$

$$+ (1 - r_{12}^2 - r_{13}^2 - r_{14}^2 - r_{23}^2 - r_{34}^2 + 2 r_{12} r_{13} r_{23} + 2 r_{13} r_{14} r_{34}$$

$$- 2 r_{12} r_{14} r_{23} r_{34} + r_{12}^2 r_{34}^2 + r_{13}^2 r_{23}^2 + r_{14}^2 r_{23}^2).$$

The left-most root of this quadratic gives us the lower bound

$$L_{24}^{(3)} := \frac{r_{23} r_{34} + r_{12} r_{14} - r_{13} r_{14} r_{23} - r_{12} r_{13} r_{34} - \sqrt{\det(A_{123}) \cdot \det(A_{134})}}{(1 - r_{13}^2)},$$

which is minimized over the range of $r_{34}$ given in (3). For the lower bound of $r_{24}$, we set

$$L_{24} := \max \left\{ L_{24}^{(1)}, L_{24}^{(2)}, L_{24}^{(3)} \right\}.$$

Similarly, for the upper bound of $r_{24}$, we define

$$U_{24}^{(1)} := r_{12} r_{14} + \sqrt{(1 - r_{12}^2)(1 - r_{14}^2)},$$

$$U_{24}^{(2)} := r_{23} r_{34} + \sqrt{(1 - r_{23}^2)(1 - r_{34}^2)},$$

and

$$U_{24}^{(3)} := \frac{r_{23} r_{34} + r_{12} r_{14} - r_{13} r_{14} r_{23} - r_{12} r_{13} r_{34} + \sqrt{\det(A_{123}) \cdot \det(A_{134})}}{(1 - r_{13}^2)}.$$
where both $U_{24}^{(2)}$ and $U_{24}^{(3)}$ are maximized over the values of $r_{34}$ given in (3). Setting

$$U_{24} := \min \left\{ U_{24}^{(1)}, U_{24}^{(2)}, U_{24}^{(3)} \right\},$$

we are guaranteed the existence of a valid correlation matrix with $r_{24}$ chosen from the range

$$L_{24} \leq r_{24} \leq U_{24}.$$

3.3. The Range of $r_{34}$. Finally, we assume that all of the correlations except $r_{34}$ have been chosen by the above method. Our arguments have shown that there will exist at least one value of $r_{34}$ so that the matrix $A$ is a valid correlation matrix. To describe the range of values that $r_{34}$ can have, we follow our previous arguments, without any optimization.

First, we let

$$L_{34}^{(1)} := r_{13}r_{14} - \sqrt{(1 - r_{13}^2)(1 - r_{14}^2)},$$

and

$$L_{34}^{(2)} := r_{23}r_{24} - \sqrt{(1 - r_{23}^2)(1 - r_{24}^2)}.$$

Noting that the determinant of $A$ can be written as

$$\det A = (r_{12} - 1)r_{34}^2 + (2r_{23}r_{24} + 2r_{13}r_{14} - 2r_{12}r_{13}r_{24} - 2r_{12}r_{14}r_{23})r_{34}
+ (1 - r_{12}^2 - r_{13}^2 - r_{14}^2 - r_{23}^2 - r_{24}^2 + 2r_{12}r_{14}r_{24} + 2r_{12}r_{13}r_{23}
- 2r_{13}r_{14}r_{23}r_{24} + r_{13}^2r_{24}^2 + r_{14}^2r_{23}^2)\det(A_{123})\cdot\det(A_{124})
- (1 - r_{12}^2)$$

we set

$$L_{34}^{(3)} := r_{23}r_{24} + r_{13}r_{14} - r_{12}r_{14}r_{23} - r_{12}r_{13}r_{24} - \sqrt{\det(A_{123})\cdot\det(A_{124})}$$

and obtain the lower bound

$$L_{34} := \max \left\{ L_{34}^{(1)}, L_{34}^{(2)}, L_{34}^{(3)} \right\}.$$

For the upper bound, we set

$$U_{34}^{(1)} := r_{13}r_{14} + \sqrt{(1 - r_{13}^2)(1 - r_{14}^2)},$$

$$U_{34}^{(2)} := r_{23}r_{24} + \sqrt{(1 - r_{23}^2)(1 - r_{24}^2)},$$

and

$$U_{34}^{(3)} := r_{23}r_{24} + r_{13}r_{14} - r_{12}r_{14}r_{23} - r_{12}r_{13}r_{24} + \sqrt{\det(A_{123})\cdot\det(A_{124})}
- (1 - r_{12}^2)$$

and obtain the lower bound

$$L_{34} := \max \left\{ L_{34}^{(1)}, L_{34}^{(2)}, L_{34}^{(3)} \right\}.$$

For

$$U_{34} := \min \left\{ U_{34}^{(1)}, U_{34}^{(2)}, U_{34}^{(3)} \right\},$$

$r_{34}$ can be chosen within the range

$$L_{34} \leq r_{34} \leq U_{34}.$$

The constructive approach of this algorithm guarantees that any matrix formed in this way will be a valid correlation matrix. Furthermore, the ranges of possible values of each of the correlations $r_{23}$, $r_{24}$, and $r_{34}$, are the largest possible ranges.
so that there will exist some valid correlation matrix with any chosen value. In other words, any value chosen outside of the determined range will fail one of the necessary conditions for being valid. Thus, our algorithm provides both necessary and sufficient conditions for producing valid $4 \times 4$ correlation matrices.

4. Applications

Having a method to create valid correlation matrices, we can now discuss the application presented by the resource allocation problem and present other examples that illustrate the difficulty of being able to guess proper correlation values. In offering a proof related to resource allocation, Xu and Evers [10] refute a prior claim by Tyagi and Das [9]. Xu and Evers explain that the error in logic was fostered by errors in the examples that inadvertently used invalid correlation matrices - ones that failed to be positive semidefinite. In the three variable case, Xu and Evers rely on boundary conditions offered by Marsaglia and Olkin [5] to establish that the matrix used by Tyagi and Das was not positive semidefinite. In the four variable case, they employ their own computer program developed to produce eigenvalues.

The previous section contains a constructive algorithm that generates valid $4 \times 4$ correlation matrices. We have developed a computer program called “validcor” (2005) to accomplish this task. Our program delivers the remaining correlations after specifying those relating one variable to the other three. The program can be found in the general archive at Carnegie Mellon University’s Statlab.

http://lib.stat.cmu.edu

To illustrate the utility of this constructive approach we use Tyagi and Das’ matrix

$$
\begin{pmatrix}
1 & -0.5 & 0.5 & -0.5 \\
-0.5 & 1 & -0.5 & 0.5 \\
0.5 & -0.5 & 1 & 0.5 \\
-0.5 & 0.5 & 0.5 & 1
\end{pmatrix}
$$

and show that it would not be generated by our program. For this example we set $r_{12} = -0.5$, $r_{13} = 0.5$, and $r_{14} = -0.5$. Feasible bounds found by our program for $r_{23}$ are $(-1, 0.5]$. We chose $r_{23} = -0.5$. Then, feasible bounds found for $r_{24}$ are $[-0.5, 1)$. We chose $r_{24} = 0.5$. Finally, feasible bounds found for $r_{34}$ are $(-1, 3.333]$. The $r_{34}$ value selected for the Tyagi and Das matrix was $0.5$ and is outside the range that assures a positive semidefinite matrix.

In addition to the use of the program showing how certain matrices are not valid correlation matrices, we provide the table below where each of the five rows correspond to generations of valid correlation matrices. The first three columns give the input parameters. The following columns give the bounds for the remaining correlations and the choices for $r_{23}$ and $r_{24}$ that led to the successive feasible bounds found by the computer program. No choice for $r_{34}$ is given as any value within the bounds will complete a valid matrix.

One interesting outcome to note from the table are the final bounds for the $r_{34}$ value. When both selections for $r_{23}$ and $r_{24}$ occur near their respective bounds, the interval for $r_{34}$ is severely limited. However, in the last example, a selection of $r_{23}$ not close to either bound allows $r_{34}$ to take on a wide range of possible values. This phenomenon along with its potential application to finding the hypervolume of possible $4 \times 4$ correlation matrices in 6-space will be explored in future research.
Another direction will include the generalization of the algorithm to $n \times n$ correlation matrices where $n$ is greater than 4.

**REFERENCES**


