On the Normal Orthant Probability with a Tri-Diagonal Correlation Matrix

Sun, Hong-Jie
Department of Information Systems, Kyushu University

Asano, Chooichiro
Department of Information Systems, Kyushu University

https://doi.org/10.15017/17181
ON THE NORMAL ORTHANT PROBABILITY WITH A
TRI-DIAGONAL CORRELATION MATRIX

Hong-Jie SUN* and Chooichiro ASANO*
(Received January 16, 1990)

We propose a new method to reduce the normal orthant probability with a positive definite tri-diagonal correlation matrix. We also investigate some properties of this method. This method is simple and is suitable to use computer to evaluate the normal orthant probability. For practical purpose, we give the reduced formulas for dimensions from 5 to 9.

1. Introduction

Let $X_1, X_2, \cdots, X_n$ be random variables with zero means, unit variances and a joint normal distribution with a positive definite tri-diagonal correlation matrix $R_n = \{|p_{ij}|\}$. The orthant probability is defined as all the $x_i$ are simultaneously positive. This probability has many applications especially in the theory of order restricted inference (Barlow, et al., 1972), but its evaluation is not simple for $n \geq 4$. So far as the present authors are aware of, there is no closed expression for $n \geq 4$. The evaluation problem of this probability has been considered by many researchers (see, for example, Martynov, 1981). Vaart (1953, 1955) obtained that this probability can be represented as a linear combination of some multiple integrals with order no more than $[n/2]$. Moran (1985) investigated two methods, one is using iterative integration, the other is using Tetrachoric series in its convergence region. These two methods also have the accuracy and computing time problems in practice.

It is given in Sun (1988) that for arbitrary positive definite correlation matrix $\Lambda_n$, the orthant probability can be represented as a linear combination of some multiple integrals with order no more than $([n/2] - 1)$. Along this line, we consider the evaluation of $P_n$ with $R_n$ being a tri-diagonal matrix. Our main purpose is to give a method for practical evaluation of this probability. In section 2, we give the method and investigate some properties of it. By using a computer to evaluate the probability, our method has obviously advantages over Vaart’s, especially when $n$ is large. We also give simplified formulas for $n = 5, 6, \cdots, 9$ in section 3.

2. A reduction of $P_n$ with a tri-diagonal correlation matrix

Without loss of generality, let $p_n$ denote a standardized multivariate normal density function and $P_n$ the correspondent orthant probability.

*Department of Information Systems
\[ P_n = \int_0^\infty p_n(x) \, dx, \] (1)

where \( x \) denotes a column vector of \( n \) components and \( x^T \) its transpose, and

\[ p_n(x) = (2\pi)^{-n/2} | R_n |^{-1/2} \exp \left( -\frac{1}{2} x^T R_n^{-1} x \right), \] (2)

\( R_n, R_n^{-1}, \) and \( | R_n | \) denote the positive definite correlation matrix \( |\rho_{ij}| \) which has a tri-diagonal form, its inverse and its determinant, respectively.

For \( n = 2k \), let

\[ I_n(\Lambda_n) = \frac{1}{(-2\pi)^k} \int_{-\infty}^\infty \prod_{i=1}^n \frac{1}{\omega_i} \exp \left( -\frac{1}{2} \omega^T \Lambda_n \omega \right) d^n \omega, \] (3)

where \( \Lambda_n = |\lambda_{ij}| \) is a covariance matrix of \( n \) variates. By using results of Nabeya (1961), Childs (1967), we have

\[ P_n = \frac{1}{2^n} + \frac{1}{2^{n-1}} \sum_{i=1}^{n-1} \sin^{-1} \rho_{ii+1} + \sum_{j=2}^{[n/2]} \frac{1}{2^{n-j} \pi^j} \sum_{i_1 < i_2 < \ldots < i_j = 1} I_{2j}(R_n^{(i_1i_2 \ldots i_j)}) \] (4)

where \( R_n^{(i_1i_2 \ldots i_j)} \) denotes the submatrix consisting of the rows and columns of \( R_n \) corresponding to \((i_1i_2 \ldots i_j)\) variates. For simplifying our representation, we introduce the following notation.

Let a positive matrix \( \Sigma \) be defined by \( \sigma_{11} = t^{-2}, \sigma_{12} = \sigma_{21} = \lambda_{12}/\sqrt{\lambda_{11}} \) and \( \sigma_{ij} = \lambda_{ij} \). We partition \( \Sigma \) into a \( 2 \times 2 \) principal matrix \( \Sigma_{11}, \) a \( 2 \times (n-2) \) matrix \( \Sigma_{12}, \) and an \( (n-2) \times (n-2) \) matrix \( \Sigma_{22}, \) where the diagonal entries of \( \Sigma_{11} \) are \( t^{-2} \) and \( \lambda_{22} \). It is obvious that if \( \Sigma \) is a tri-diagonal matrix, also the matrix \( \Lambda_{n-2} = \Sigma_{22} - \Sigma_{12}^T \Sigma_{11}^{-1} \Sigma_{12} \). This property plays an important role in the following theorems.

**Theorem 1.** It is well known that \( I_2(\Lambda_2) = \sin^{-1} (\lambda_{12}/\sqrt{\lambda_{11} \lambda_{22}}) \). For \( n = 2k \) with \( k > 1 \), we have that

\[ I_n(\Lambda_n) = \int_0^1 \sqrt{\frac{\lambda_{12}}{\lambda_{11} \lambda_{22} - \lambda_{12}^2 t^2}} I_{n-2}(\Lambda_{n-2}) \, dt. \] (5)

**Proof.** By making \( \lambda_{1i} = 0, i > 2, \) in the theorem of Sun (1988), the formula (5) follows.

This theorem says that we can obtain a reduced expression by using it successively and the \( n \)-variate orthant probability can be computed by multiple integrals with order no more than \( ([n/2] - 1) \). This is different from the result of Vaart (1953, 1955) in which the integral is of order \( [n/2] \). By using a computer to evaluate \( P_n \), our method has the advantage of less time over Vaart's especially for larger \( n \).
**Theorem 2.** For arbitrary $n$, $2 \leq j \leq [n/2]$, the expression

$$
\sum_{i_1 < i_2 < \cdots < i_j \leq n-1} I_{j; i_1, i_2, \cdots, i_j} (R_n^{(i_1 i_2 \cdots i_j)})
$$

in (4) can be computed at most by $(n-2j+1)$ integrals with order $(j-1)$ and $(n-j-2j+1)$ integrals with order less than $(j-1)$.

**Proof.** By the definition of $R_n^{(i_1 i_2 \cdots i_j)}$, there are a total number of $\binom{n}{2j}$ $(2j) \times (2j)$ submatrices. In all these submatrices, there are $(n-2j+1)$ submatrices with the off-diagonal terms being non-zero, and each of these can be calculated by a multiple integral with order $(j-1)$ according to theorem 1. The other $\binom{n}{2j} - (n-2j+1)$ submatrices must contain zero in the off-diagonal terms. For normal variables, the submatrix with the off-diagonal terms containing zero means that it can be partitioned into at least two independent blocks with the number of dimension less than $2j$. For each of these blocks, according to theorem 1, it can be calculated by some multiple integrals with order less than $(j-1)$. This completes the proof.

Theorem 2 gives the outline of the consisting elements of (6). It can be used to check the correctness of a reduced expression.

### 3. Expressions for $4 \leq n \leq 9$

In the section, we illustrate the case $n = 5$ and give reduced expressions for $n = 6, 7, \ldots, 9$. For $n$ greater than 9, one can use the theorems in section 2 to obtain the reduced expression.

It is well known in Childs [2] that for $n = 4$,

$$
P_4 = \frac{1}{2^4} + \frac{1}{2^3 \pi} \sum_{i=1}^{3} \sin^{-1} \rho_{i,i+1} + \frac{1}{2^2 \pi^2} I_4 (R_4)
$$

$$
I_4 (R_4) = \int_{0}^{1} \frac{\rho_{12}}{\sqrt{1 - \rho_{12}^2 t^2}} \sin^{-1} \left( \frac{a_{12}}{\sqrt{a_{11} a_{22}}} \right) \, dt,
$$

$$
a_{11} (t) = 1 - \rho_{23}^2 - \rho_{12}^2 t^2,
$$

$$
a_{22} (t) = 1 - \rho_{12}^2 t^2,
$$

$$
a_{12} (t) = \rho_{34} (1 - \rho_{12}^2 t^2)
$$

For $n = 5$, according to the formula (4), the correlation matrix $R_5$ can be partitioned into five $4 \times 4$ submatrices, which are

$$
A = \begin{bmatrix}
1 & \rho_{12} & 0 & 0 \\
\rho_{12} & 1 & \rho_{23} & 0 \\
0 & \rho_{23} & 1 & \rho_{34} \\
0 & 0 & \rho_{34} & 1
\end{bmatrix},
$$

$$
B = \begin{bmatrix}
1 & \rho_{23} & 0 & 0 \\
\rho_{23} & 1 & \rho_{34} & 0 \\
0 & \rho_{34} & 1 & \rho_{45} \\
0 & 0 & \rho_{45} & 1
\end{bmatrix},
$$

$$
C = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & \rho_{23} & 0 \\
0 & \rho_{23} & 1 & \rho_{34} \\
0 & 0 & \rho_{34} & 1
\end{bmatrix}.
$$
\[
D = \begin{bmatrix}
1 & \rho_{12} & 0 & 0 \\
\rho_{12} & 1 & 0 & 0 \\
0 & 0 & 1 & \rho_{34} \\
0 & 0 & \rho_{34} & 1
\end{bmatrix}, \quad E = \begin{bmatrix}
1 & \rho_{23} & 0 & 0 \\
\rho_{23} & 1 & \rho_{34} & 0 \\
0 & \rho_{34} & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}.
\]

It is now obvious that \( A \) and \( B \) can be computed by \( I_4 \) in (7). The integrals \( I_4(C) \) and \( I_4(E) \) are equal to zero, and \( I_4(D) = \sin^{-1}\rho_{12}\sin^{-1}\rho_{34} \). These results coincide with Theorem 2. So we have

\[
P_5 = \frac{1}{2^5} + \frac{1}{2^4} \pi \sum_{i=1}^{4} \sin^{-1}\rho_{i,i+1} + \frac{1}{2^3} \pi \sum_{i=1}^{5} \sum_{j=i+3}^{5} \sin^{-1}\rho_{i,i+1}\sin^{-1}\rho_{j,j+1} + I_4(A) + I_4(B)
\]

The followings are expressions for \( n = 6, 7, \ldots, 9 \), where the notation \( A_k(i) \) is used for a tri-diagonal correlaiton matrix with the number of dimension \( k, k \leq n \); and the off-diagonal terms starting from the \( i \)-th term of \( R_n \), \( 1 \leq i \leq n \).

\[
P_6 = \frac{1}{2^6} + \frac{1}{2^5} \pi \sum_{i=1}^{5} \sin^{-1}\rho_{i,i+1} + \frac{1}{2^4} \pi \sum_{i=1}^{6} \sum_{j=i+3}^{6} \sin^{-1}\rho_{i,i+1}\sin^{-1}\rho_{j,j+1} + I_4(A)
\]

\[
I_6(R_6) = \int_0^1 \frac{\rho_{12}}{\sqrt{1 - \rho_{12}^2}} \int_0^1 \frac{a_{12}}{\sqrt{a_{11}a_{22} - a_{12}^2u^2}} \sin^{-1}\left( \frac{b_{12}}{\sqrt{b_{11}b_{22}}} \right) \, du \, dt,
\]

\[
a_{11}(i) = 1 - \rho_{23}^2 - \rho_{12}^2
\]

\[
a_{ii}(i) = 1 - \rho_{12}^2 t^2 \quad i = 2, 3, 4.
\]

\[
a_{i,i+1}(i) = \rho_{i+2,i+3}(1 - \rho_{12}^2 t^2) \quad i = 1, 2, 3.
\]

\[
b_{11}(i, u) = a_{33}(a_{11}a_{22} - a_{12}^2 u^2) - a_{11}a_{23}^2
\]

\[
b_{22}(i, u) = a_{44}(a_{11}a_{22} - a_{12}^2 u^2)
\]

\[
b_{12}(i, u) = a_{34}(a_{11}a_{22} - a_{12}^2 u^2).
\]

\[
P_7 = \frac{1}{2^7} + \frac{1}{2^6} \pi \sum_{i=1}^{6} \sin^{-1}\rho_{i,i+1} + \frac{1}{2^5} \pi \sum_{i=1}^{5} \sum_{j=i+3}^{5} \sin^{-1}\rho_{i,i+1}\sin^{-1}\rho_{j,j+1} + \sum_{i=1}^{4} I_4(A)
\]

\[
I_7(R_7) = \int_0^1 \frac{\rho_{12}}{\sqrt{1 - \rho_{12}^2}} \int_0^1 \frac{a_{12}}{\sqrt{a_{11}a_{22} - a_{12}^2u^2}} \sin^{-1}\left( \frac{b_{12}}{\sqrt{b_{11}b_{22}}} \right) \, du \, dt.
\]

\[
P_8 = \frac{1}{2^8} + \frac{1}{2^7} \pi \sum_{i=1}^{7} \sin^{-1}\rho_{i,i+1} + \frac{1}{2^6} \pi \sum_{i=1}^{6} \sum_{j=i+3}^{6} \sin^{-1}\rho_{i,i+1}\sin^{-1}\rho_{j,j+1} + \sum_{i=1}^{5} I_4(A)
\]
平成元年
九州大学総合理工学研究科報告 第12巻 第1号 — 57 —

\[
\frac{1}{2^5 \pi^3} \left\{ \sum_{i=1}^{2} I_4(A_4(i)) \sum_{j=i+1}^{7} \sin^{-1} \rho_{j,i+1} + \sum_{i=4}^{5} I_4(A_4(i)) \sum_{j=i+1}^{8} \sin^{-1} \rho_{j,i+1} \right. \\
\left. + \sin^{-1} \rho_{12} \sin^{-1} \rho_{45} \sin^{-1} \rho_{78} + \sum_{j=1}^{3} I_6(A_6(i)) \right\} + \frac{1}{2^4 \pi^4} I_8(R_8)
\]

\[
I_8(R_8) = \int_0^1 \frac{\rho_{12}}{\sqrt{1-\rho_{12}^2}} \int_0^1 \frac{a_{12}}{\sqrt{a_{11}a_{22} - a_{12}^2 u^2}} \int_0^1 \frac{b_{12}}{\sqrt{b_{11}b_{22} - b_{12}^2 v^2}} \sin^{-1} \left( \frac{c_{12}}{\sqrt{c_{11}c_{22}}} \right) dvdu dt.
\]

- \[
a_{11}(i) = 1 - \rho_{23}^2 - \rho_{12}^2 t^2
\]
- \[
a_{ii}(i) = 1 - \rho_{12}^2 t^2 \quad i = 2, 3, \ldots, 6.
\]
- \[
a_{i,i+1}(i) = \rho_{i+2,i+3}(1 - \rho_{12}^2 t^2) \quad i = 1, 2, \ldots, 5.
\]
- \[
b_{11}(t, u) = a_{33}(a_{11}a_{22} - a_{12}^2 u^2) - a_{11}a_{23}^2
\]
- \[
b_{ii}(t, u) = a_{44}(a_{11}a_{22} - a_{12}^2 u^2) \quad i = 2, 3, 4.
\]
- \[
b_{i,i+1}(t, u) = a_{i+2,i+3}(a_{11}a_{22} - a_{12}^2 u^2) \quad i = 1, 2, 3.
\]
- \[
c_{11}(t, u, v) = b_{33}(b_{11}b_{22} - b_{12}^2 v^2) - b_{11}b_{23}^2
\]
- \[
c_{ii}(t, u, v) = b_{44}(b_{11}b_{22} - b_{12}^2 v^2)
\]
- \[
c_{i,i+1}(t, u, v) = b_{i+2,i+3}(b_{11}b_{22} - b_{12}^2 v^2) \quad i = 1, 2, 3.
\]

- \[
P_9 = \frac{1}{2^9} + \frac{1}{2^8 \pi} \sum_{i=1}^{8} \sin^{-1} \rho_{i,i+1} + \frac{1}{2^7 \pi^2} \sum_{i=1}^{8} \sin^{-1} \rho_{j,i+1} + \sum_{i=1}^{8} I_4(A_4(i))
\]
- \[
+ \frac{1}{2^8 \pi^3} \left\{ \sum_{i=1}^{3} I_4(A_4(i)) \sum_{j=i+1}^{8} \sin^{-1} \rho_{j,i+1} + \sum_{i=4}^{5} I_4(A_4(i)) \sum_{j=i+1}^{8} \sin^{-1} \rho_{j,i+1} \right. \\
\left. + \sum_{i=1}^{2} \sin^{-1} \rho_{j,i+1} \sum_{j=i+1}^{5} \sin^{-1} \rho_{j,j+1} + \sum_{i=4}^{9} \sum_{k=j+3}^{4} I_6(A_6(i)) \right\}
\]
- \[
+ \frac{1}{2^5 \pi^4} \left\{ \sin^{-1} \rho_{12} I_6(A_6(4)) + \sin^{-1} \rho_{89} I_6(A_6(1)) + I_4(A_4(1)) I_4(A_4(6)) \right. \\
\left. + \sum_{i=1}^{2} I_8(A_8(i)) \right\}.
\]
References


