On Moments of Folded and Truncated Multivariate Normal Distributions

Raymond Kan*
Department of Finance, University of Toronto
105 St. George Street, Toronto, Ontario M5S 3E6, Canada
(E-mail: kan@chass.utoronto.ca)
and
Cesare Robotti
Department of Finance, Imperial College London
South Kensington Campus, London SW7 2AZ, United Kingdom
(E-mail: c.robotti@imperial.ac.uk)

March 12, 2018

Abstract

Recurrence relations for integrals that involve the density of multivariate normal distributions are developed. These recursions allow fast computation of the moments of folded and truncated multivariate normal distributions. Besides being numerically efficient, the proposed recursions also allow us to obtain explicit expressions of low order moments of folded and truncated multivariate normal distributions.

Keywords: Multivariate normal distribution; Folded normal distribution; Truncated normal distribution.

*The first author gratefully acknowledges financial support from National Bank Financial of Canada.

1 Introduction

Suppose $\boldsymbol{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ follows a multivariate normal distribution with mean $\boldsymbol{\mu}$ and positive definite covariance matrix $\boldsymbol{\Sigma}$. We are interested in computing $E(|X_1^{k_1} \cdots X_n^{k_n}|)$ and $E(X_1^{k_1} \cdots X_n^{k_n} \mid a_i < X_i < b_i,\ i = 1, \ldots, n)$ for $k_i \geq 0,\ i = 1, \ldots, n$. The first expression is the moment of a folded multivariate normal distribution $|\boldsymbol{X}| = (|X_1|, \ldots, |X_n|)^{\mathrm{T}}$. The second expression is the moment of a truncated multivariate normal distribution, with X_i truncated at the lower limit a_i and upper limit b_i. In the second expression, some of the a_i's can be $-\infty$ and some of the b_i's can be ∞. When all the b_i's are ∞, the distribution is called the lower truncated multivariate normal, and when all the a_i's are $-\infty$, the distribution is called the upper truncated multivariate normal.

The folded univariate normal distribution was first introduced by Leone et al. (1961), and Elandt (1961) provides expressions for its moments. Psarakis and Panaretos (2001) generalize the folded distribution to the bivariate normal case and provide the moment generating function when $\boldsymbol{\mu} = \mathbf{0}$. Recently, Chakraborty and Chatterjee (2013) introduce the folded multivariate normal distribution. They present the joint density, the moment generating function, and the mean and covariance matrix of $|\boldsymbol{X}|$. Unfortunately, as pointed out by Murthy (2015), the moment generating function as well as the mean and covariance matrix expressions given in Chakraborty and Chatterjee (2013) are incorrect. The moments of the folded multivariate normal distribution are simply the absolute moments of the multivariate normal distribution. When $\boldsymbol{\mu} = \mathbf{0}$, there is a literature that provides explicit formulae for these absolute moments. Nabeya (1951) derives an explicit expression of the absolute moments for the bivariate normal case. Nabeya (1952) presents explicit expressions of the absolute moments for the trivariate normal case (up to 12th order, see also related results in Kamat (1953)). For the 4-variate case, Nabeya (1961) provides explicit expressions of some low order absolute moments. However, the computation of higher order absolute moments has been a challenge for $n > 2$ even when $\boldsymbol{\mu} = \mathbf{0}$. When $\boldsymbol{\mu} \neq \mathbf{0}$, we are unaware of any result that enables us to compute arbitrary order absolute moments of a multivariate normal distribution (except when $n = 1$).

There is a long and rich literature on truncated normal distributions. For $n = 1$, Cohen (1991) provides a comprehensive review of the literature. For the lower truncated univariate normal, Cohen (1951a) proposes a recursive formula for its moments. In addition, Cohen (1951b) derives a recursive formula for the moments of the doubly truncated univariate normal. For $n = 2$, Rosenbaum (1961) provides the first two moments for the singly truncated case, and Khatri and Jaiswal (1963) provide a recurrence relation to obtain all the bivariate moments for the lower truncated case. For the doubly truncated case, Shah and Parikh (1964) and Dyer (1973) propose recurrence formulae for the bivariate moments. For the n-dimensional case, Birnbaum and Meyer (1953) derive a recursive formula for the bivariate moments in the lower truncated case. Gupta and Tracy (1976) provide a recurrence relation between different product moments of a doubly truncated multivariate normal. Unfortunately, since their recurrence relation does not express the product moments in terms of lower order product moments, it has been of little practical use besides the case of bivariate moments. Lee (1983) also presents a recurrence relation between product moments of a doubly truncated multivariate normal, but his relation requires the powers of all but one of the variables to be equal to one. Therefore, his formula cannot be used when all the variables have powers greater than one. The moment generating function of the lower truncated multivariate normal distribution is available in Tallis (1961) and, in principle, it can be used to compute all the product moments for the lower truncated multivariate normal. Tallis (1961) provides explicit expressions of some low order moments for the $n = 2$ and 3 cases. However, differentiating this moment generating function to obtain higher order moments involves tedious calculations. Recently, Arismendi (2013) overcomes this difficulty by providing explicit expressions for computing arbitrary order product moments. However, the required calculations for this approach can be quite time consuming. For example, when $n = 6$, computing all the fourth order moments, i.e., $k_1 + \cdots + k_6 = 4$, for the lower truncated multivariate normal distribution requires more than 5.4 hours on a PC with an Intel i7-4790K.[1] In contrast, the Matlab program based on our algorithm computes all the product moments with $0 \leq k_i \leq 4$ $(i = 1, \ldots, 6)$ in less than 29 seconds.

Instead of differentiating the moment generating function, we approach the problem by directly computing the moments of folded and truncated multivariate normal distributions, which require evaluating n-dimensional integrals that involve the multivariate normal density. We develop simple and efficient recurrence formulae for these multivariate integrals. In the most general case, the recurrence formula involves $3n + 1$ terms, but in many cases the number of terms can be reduced to $n + 1$. Besides giving us a very efficient approach for computing the product moments of folded and truncated multivariate normal distributions, our recurrence formula may also be applicable to other similar problems. The rest of the paper is organized as follows. Section 2 presents a recurrence formula for an integral that is essential for the evaluation of moments of folded and truncated multivariate normal distributions. Section 3 presents the results for the folded multivariate normal distribution. Section 4 presents the results for the truncated multivariate normal distribution. Besides providing the numerical algorithm for computing these moments, we also present explicit expressions for some low order moments of folded and truncated multivariate normal distributions. Section 5 discusses possible extensions.

[1]We thank Juan Arismendi for kindly sharing his Matlab programs with us.

2 A Recurrence Relation for a Multivariate Integral

Suppose $\boldsymbol{X} = (X_1, \ldots, X_n)^{\mathrm{T}} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^{\mathrm{T}}$ is the mean of $\boldsymbol{X}$, $\boldsymbol{\Sigma} = (\sigma_{ij})$ is the covariance matrix of $\boldsymbol{X}$, and $\sigma_i^2 \equiv \sigma_{ii}$ stands for the variance of X_i. The joint density function of $\boldsymbol{X}$ is

$$\phi_n(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}.$$

The cumulative distribution function of $\boldsymbol{X}$ is denoted as

$$\Phi_n(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \int_{-\boldsymbol{\infty}}^{\boldsymbol{x}} \phi_n(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \mathrm{d}\boldsymbol{y},$$

where we make use of the short-hand notation

$$\int_{\boldsymbol{a}}^{\boldsymbol{b}} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \equiv \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n) \mathrm{d}x_n \cdots \mathrm{d}x_1.$$

When $\boldsymbol{\mu} = \boldsymbol{0}$, we suppress the argument $\boldsymbol{\mu}$ and simply write $\phi_n(\boldsymbol{x}; \boldsymbol{\Sigma})$ and $\Phi_n(\boldsymbol{x}; \boldsymbol{\Sigma})$. In addition, let

$$L_n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \int_{\boldsymbol{a}}^{\boldsymbol{b}} \phi_n(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \mathrm{d}\boldsymbol{y}.$$

Based on the inclusion-exclusion principle, this probability can be written as a linear combination of 2^n different values of $\Phi_n(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$, that is,

$$L_n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{i_1, \ldots, i_n \in \{0,1\}} (-1)^{n - \sum_{j=1}^{n} i_j} \Phi_n((y_{i_1}, \ldots, y_{i_n})^{\mathrm{T}}; \boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

where $y_{i_j} = a_j$ if $i_j = 0$ and $y_{i_j} = b_j$ if $i_j = 1$.

For the special case of univariate standard normal (i.e., $n = 1$, $\mu = 0$, $\sigma = 1$), we use $\phi(x)$ and $\Phi(x)$ to denote its density and cumulative distribution functions, respectively. In addition, for the standard bivariate normal (i.e., $n = 2$, $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$), we let $\phi_2(x_1, x_2; \rho)$ stand for $\phi_2(\boldsymbol{x}; \boldsymbol{\Sigma})$ and $\Phi_2(x_1, x_2; \rho)$ stand for $\Phi_2(\boldsymbol{x}; \boldsymbol{\Sigma})$, where ρ is the correlation coefficient between X_1 and X_2.

For two n-vectors $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{T}}$ and $\boldsymbol{\kappa} = (k_1, \ldots, k_n)^{\mathrm{T}}$, let $\boldsymbol{x}^{\boldsymbol{\kappa}}$ stand for $x_1^{k_1} \cdots x_n^{k_n}$. By $\boldsymbol{a}_{(i)}$ we denote a vector $\boldsymbol{a}$ with its ith element removed. For a matrix $\boldsymbol{A}$, we let $\boldsymbol{A}_{i,(j)}$ stand for the ith row of $\boldsymbol{A}$ with its jth element removed. Similarly, $\boldsymbol{A}_{(i),(j)}$ stands for the matrix $\boldsymbol{A}$ with its ith row and jth column removed. Finally, we let $\boldsymbol{e}_i$ denote an n-vector with its ith element equal to one and zero otherwise.

The integral that we are interested in evaluating is

$$F_{\boldsymbol{\kappa}}^n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \equiv \int_{\boldsymbol{a}}^{\boldsymbol{b}} \boldsymbol{x}^{\boldsymbol{\kappa}} \phi_n(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \mathrm{d}\boldsymbol{x}.$$

The boundary condition is obviously $F_{\boldsymbol{0}}^n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = L_n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$. When $n = 1$, it is straightforward to use integration by parts to show that (with the arguments of F_k^1 suppressed)

$$\begin{aligned} F_0^1 &= \Phi(\beta) - \Phi(\alpha), \\ F_{k+1}^1 &= \mu F_k^1 + k\sigma^2 F_{k-1}^1 + \sigma\{a^k \phi(\alpha) - b^k \phi(\beta)\} \quad (k \geq 1), \end{aligned}$$

where $\alpha = (a - \mu)/\sigma$ and $\beta = (b - \mu)/\sigma$. When $n > 1$, we need a similar recurrence relation in order to compute $F_{\boldsymbol{\kappa}}(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$. The following theorem presents the required result. Lee (1983) also presents a similar recursive relation but his result can only be applied when $\boldsymbol{\kappa}$ is in the form of $(1, \ldots, 1)^{\mathrm{T}} + k_i \boldsymbol{e}_i$, whereas our result allows for an arbitrary $\boldsymbol{\kappa} > \boldsymbol{0}$.

Theorem 1 *For $n > 1$,*

$$F_{\boldsymbol{\kappa}+\boldsymbol{e}_i}^n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \mu_i F_{\boldsymbol{\kappa}}^n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) + \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{\Sigma} \boldsymbol{c}_{\boldsymbol{\kappa}} \quad (i = 1, \ldots, n), \tag{1}$$

where $\boldsymbol{c}_{\boldsymbol{\kappa}}$ is an n-vector with jth element

$$c_{\boldsymbol{\kappa},j} = k_j F_{\boldsymbol{\kappa}-\boldsymbol{e}_j}^n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) + a_j^{k_j} \phi_1(a_j; \mu_j, \sigma_j^2) F_{\boldsymbol{\kappa}_{(j)}}^{n-1}(\boldsymbol{a}_{(j)}, \boldsymbol{b}_{(j)}; \tilde{\boldsymbol{\mu}}_j^{\boldsymbol{a}}, \tilde{\boldsymbol{\Sigma}}_j)$$

$$- b_j^{k_j}\phi_1(b_j;\mu_j,\sigma_j^2)F^{n-1}_{\boldsymbol{\kappa}_{(j)}}(\boldsymbol{a}_{(j)},\boldsymbol{b}_{(j)};\tilde{\boldsymbol{\mu}}_j^{\boldsymbol{b}},\tilde{\boldsymbol{\Sigma}}_j), \tag{2}$$

and

$$\begin{aligned}
\tilde{\boldsymbol{\mu}}_j^{\boldsymbol{a}} &= \boldsymbol{\mu}_{(j)} + \boldsymbol{\Sigma}_{(j),j}\frac{a_j-\mu_j}{\sigma_j^2},\\
\tilde{\boldsymbol{\mu}}_j^{\boldsymbol{b}} &= \boldsymbol{\mu}_{(j)} + \boldsymbol{\Sigma}_{(j),j}\frac{b_j-\mu_j}{\sigma_j^2},\\
\tilde{\boldsymbol{\Sigma}}_j &= \boldsymbol{\Sigma}_{(j),(j)} - \frac{1}{\sigma_j^2}\boldsymbol{\Sigma}_{(j),j}\boldsymbol{\Sigma}_{j,(j)}.
\end{aligned}$$

When $k_j = 0$, the first term in (2) vanishes. When $a_j = -\infty$, the second term vanishes, and when $b_j = \infty$, the third term vanishes.

Proof: Taking the derivative of the multivariate normal density, we have

$$-\frac{\partial \phi_n(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial \boldsymbol{x}} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\phi_n(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma}).$$

Multiplying each element on both sides by $\boldsymbol{x}^{\boldsymbol{\kappa}}$ and integrating $\boldsymbol{x}$ from $\boldsymbol{a}$ to $\boldsymbol{b}$, we have (after suppressing the arguments of $F^n_{\boldsymbol{\kappa}}$)

$$\boldsymbol{c}_{\boldsymbol{\kappa}} = \boldsymbol{\Sigma}^{-1}\begin{bmatrix} F^n_{\boldsymbol{\kappa}+\boldsymbol{e}_1} - \mu_1 F^n_{\boldsymbol{\kappa}} \\ F^n_{\boldsymbol{\kappa}+\boldsymbol{e}_2} - \mu_2 F^n_{\boldsymbol{\kappa}} \\ \vdots \\ F^n_{\boldsymbol{\kappa}+\boldsymbol{e}_n} - \mu_n F^n_{\boldsymbol{\kappa}} \end{bmatrix}, \tag{3}$$

where the jth element of the left hand side is

$$\begin{aligned}
c_{\boldsymbol{\kappa},j} &= -\int_{\boldsymbol{a}}^{\boldsymbol{b}} \boldsymbol{x}^{\boldsymbol{\kappa}}\frac{\partial \phi_n(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma})}{\partial x_j}\mathrm{d}\boldsymbol{x}\\
&= -\int_{\boldsymbol{a}_{(j)}}^{\boldsymbol{b}_{(j)}} \boldsymbol{x}^{\boldsymbol{\kappa}}\phi_n(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma})\Big|_{x_j=a_j}^{b_j}\mathrm{d}\boldsymbol{x}_{(j)} + \int_{\boldsymbol{a}}^{\boldsymbol{b}} k_j \boldsymbol{x}^{\boldsymbol{\kappa}-\boldsymbol{e}_j}\phi_n(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma})\mathrm{d}\boldsymbol{x} \qquad (4)
\end{aligned}$$

by using integration by parts. Using the fact that

$$\begin{aligned}\phi_n(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma})\big|_{x_j=a_j} &= \phi_1(a_j;\mu_j,\sigma_j^2)\phi_{n-1}(\boldsymbol{x}_{(j)};\tilde{\boldsymbol{\mu}}_j^{\boldsymbol{a}},\tilde{\boldsymbol{\Sigma}}_j),\\ \phi_n(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma})\big|_{x_j=b_j} &= \phi_1(b_j;\mu_j,\sigma_j^2)\phi_{n-1}(\boldsymbol{x}_{(j)};\tilde{\boldsymbol{\mu}}_j^{\boldsymbol{b}},\tilde{\boldsymbol{\Sigma}}_j),\end{aligned}$$

we obtain

$$\begin{aligned}c_{\boldsymbol{\kappa},j} = k_j F^n_{\boldsymbol{\kappa}-\boldsymbol{e}_j}(\boldsymbol{a},\boldsymbol{b};\boldsymbol{\mu},\boldsymbol{\Sigma}) &+ a_j^{k_j}\phi_1(a_j;\mu_j,\sigma_j^2)F^{n-1}_{\boldsymbol{\kappa}_{(j)}}(\boldsymbol{a}_{(j)},\boldsymbol{b}_{(j)};\tilde{\boldsymbol{\mu}}_j^{\boldsymbol{a}},\tilde{\boldsymbol{\Sigma}}_j)\\ &- b_j^{k_j}\phi_1(b_j;\mu_j,\sigma_j^2)F^{n-1}_{\boldsymbol{\kappa}_{(j)}}(\boldsymbol{a}_{(j)},\boldsymbol{b}_{(j)};\tilde{\boldsymbol{\mu}}_j^{\boldsymbol{b}},\tilde{\boldsymbol{\Sigma}}_j).\end{aligned}$$

When $k_j = 0$, the last integral in (4) is equal to zero, and the first term of $c_{\boldsymbol{\kappa},j}$ drops out. When $a_j \to -\infty$, $a_j^{k_j}\phi_1(a_j;\mu_j,\sigma_j^2) \to 0$, so the second term of $c_{\boldsymbol{\kappa},j}$ drops out. Similarly, when $b_j \to \infty$, the third term of $c_{\boldsymbol{\kappa},j}$ drops out. Finally, multiplying both sides of (3) by $\boldsymbol{\Sigma}$, we obtain (1). This concludes the proof of Theorem 1.

It should be emphasized that Gupta and Tracy (1976) present a similar recurrence relation for the moments of a doubly truncated multivariate normal distribution. Besides the fact that they are dealing with the special case of $\boldsymbol{a} = \boldsymbol{0}$, the main difference is that their recurrence relation is essentially stated as

$$c_{\boldsymbol{\kappa},j} = \boldsymbol{e}_j^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\begin{bmatrix} F^n_{\boldsymbol{\kappa}+\boldsymbol{e}_1} - \mu_1 F^n_{\boldsymbol{\kappa}} \\ F^n_{\boldsymbol{\kappa}+\boldsymbol{e}_2} - \mu_2 F^n_{\boldsymbol{\kappa}} \\ \vdots \\ F^n_{\boldsymbol{\kappa}+\boldsymbol{e}_n} - \mu_n F^n_{\boldsymbol{\kappa}} \end{bmatrix} \qquad (j = 1,\dots,n).$$

In this form, one cannot compute $F^n_{\boldsymbol{\kappa}}$ by using only lower order terms, and it is difficult to use this recursion in practice.[2] Due to this unfortunate situation, no attempts have

[2]For example, Manjunath and Wilhelm (2012) comment that "But since except for the mean there are fewer equations than parameters, these recurrent conditions do not uniquely identify moments of order $\geq$ 2 and are therefore not sufficient for the computation of the variance and other higher order moments."

been made to use this recurrence relation to compute higher order moments of a truncated multivariate normal for $n \geq 3$. We overcome this problem in Theorem 1 by multiplying both sides of (3) by $\boldsymbol{\Sigma}$. This delivers a simple way to compute $F^n_{\boldsymbol{\kappa}}(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ based on at most $3n + 1$ lower order terms, with $n + 1$ of them being n-dimensional integrals and the rest being $(n - 1)$-dimensional integrals.

Although Theorem 1 is stated as a recurrence relation, it is better to avoid using a recursive function to compute $F^n_{\boldsymbol{\kappa}}(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$. For speed gains, it is much more efficient to first compute all the necessary $(n - 1)$-dimensional integrals ($F^{n-1}_{\boldsymbol{\nu}}(\boldsymbol{a}_{(j)}, \boldsymbol{b}_{(j)}, \tilde{\boldsymbol{\mu}}^{\boldsymbol{a}}_j, \tilde{\boldsymbol{\Sigma}}_j)$ and $F^{n-1}_{\boldsymbol{\nu}}(\boldsymbol{a}_{(j)}, \boldsymbol{b}_{(j)}, \tilde{\boldsymbol{\mu}}^{\boldsymbol{b}}_j, \tilde{\boldsymbol{\Sigma}}_j)$ for $\boldsymbol{0} \leq \boldsymbol{\nu} \leq \boldsymbol{\kappa}_{(j)}$, $j = 1, \ldots, n$) and then build up the entire table of $F^n_{\boldsymbol{\nu}}(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ for $\boldsymbol{0} \leq \boldsymbol{\nu} \leq \boldsymbol{\kappa}$.[3]

When all the a_i's are $-\infty$ or all the b_i's are ∞, the length of the recurrence relation is reduced to $2n + 1$. When all the a_i's are $-\infty$ and all the b_i's are ∞, we have

$$F^n_{\boldsymbol{\kappa}}(-\boldsymbol{\infty}, \boldsymbol{\infty}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = E(\boldsymbol{X}^{\boldsymbol{\kappa}}),$$

which is the product moments of $\boldsymbol{X}$. In this case, the recurrence relation is

$$E(\boldsymbol{X}^{\boldsymbol{\kappa}+\boldsymbol{e}_i}) = \mu_i E(\boldsymbol{X}^{\boldsymbol{\kappa}}) + \sum_{j=1}^{n} \sigma_{ij} k_j E(\boldsymbol{X}^{\boldsymbol{\kappa}-\boldsymbol{e}_j}) \quad (i = 1, \ldots, n),$$

and it is of length $n+1$. This recurrence relation was obtained by Takemura and Takeuchi (1988) and Willink (2005).

Another case of special interest occurs when $a_i = 0$ and $b_i = \infty$, $i = 1, \ldots, n$. For this scenario, let

$$I^n_{\boldsymbol{\kappa}}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \equiv F^n_{\boldsymbol{\kappa}}(\boldsymbol{0}, \boldsymbol{\infty}; \boldsymbol{\mu}, \boldsymbol{\Sigma}).$$

[3] A set of Matlab programs to evaluate $F^n_{\boldsymbol{\kappa}}(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ and other expressions in the paper is available at `http://www-2.rotman.utoronto.ca/~kan/research.htm`.

The recurrence relation for $I^n_{\boldsymbol{\kappa}}$ can be written as

$$I^n_{\boldsymbol{\kappa}+\boldsymbol{e}_i}(\boldsymbol{\mu},\boldsymbol{\Sigma}) = \mu_i I^n_{\boldsymbol{\kappa}}(\boldsymbol{\mu},\boldsymbol{\Sigma}) + \sum_{j=1}^n \sigma_{ij} d_{\boldsymbol{\kappa},j} \quad (i=1,\ldots,n), \tag{5}$$

where

$$d_{\boldsymbol{\kappa},j} = \begin{cases} k_j I^n_{\boldsymbol{\kappa}-\boldsymbol{e}_j}(\boldsymbol{\mu},\boldsymbol{\Sigma}) & (k_j>0), \\ \phi_1(\mu_j;\sigma_j^2) I^{n-1}_{\boldsymbol{\kappa}_{(j)}}(\tilde{\boldsymbol{\mu}}_j,\tilde{\boldsymbol{\Sigma}}_j) & (k_j=0), \end{cases}$$

with $\tilde{\boldsymbol{\mu}}_j = \boldsymbol{\mu}_{(j)} - \boldsymbol{\Sigma}_{(j),j}\mu_j/\sigma_j^2$. The length of this recursion is only $n+1$. For $n=1$, our $I^1_k(\mu,\sigma^2)$ function is closely related to the I_k function of Fisher (1931), which is defined as

$$I_k(\xi) = \frac{1}{(2\pi)^{\frac{1}{2}}}\int_0^\infty \frac{x^k}{k!} e^{-\frac{(x+\xi)^2}{2}} \mathrm{d}x.$$

It can be readily seen that $I_k(\xi) = I^1_k(-\xi,1)/k!$, and it satisfies the recurrence relation

$$(k+1)I_{k+1}(\xi) = -\xi I_k(\xi) + I_{k-1}(\xi) \quad (k \geq 1).$$

3 Folded Multivariate Normal

The folded multivariate normal distribution is simply the distribution of $|\boldsymbol{X}|$, where $\boldsymbol{X} \sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$. In this section, we present the correct expression of the moment generating function of $|\boldsymbol{X}|$ as well as our approach for computing arbitrary order moments of $|\boldsymbol{X}|$. In addition, we present some explicit expressions of some low order moments of $|\boldsymbol{X}|$, including the mean and covariance matrix of $|\boldsymbol{X}|$.

Following Chakraborty and Chatterjee (2013), let

$$\mathcal{S}(n) = \{\boldsymbol{s} : \boldsymbol{s} = (s_1,\ldots,s_n),\ \text{with } s_i = \pm 1,\ i=1,\ldots,n\}$$

be a set of different combinations of n positive and negative signs. By defining $\boldsymbol{\Lambda}_{\boldsymbol{s}} = \mathrm{Diag}(s_1, \ldots, s_n)$, Chakraborty and Chatterjee (2013) show that the joint density of $\boldsymbol{Y} = |\boldsymbol{X}|$ is

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \sum_{\boldsymbol{s} \in \mathcal{S}(n)} \phi_n(\boldsymbol{y}; \boldsymbol{\mu}_{\boldsymbol{s}}, \boldsymbol{\Sigma}_{\boldsymbol{s}}) \qquad (\boldsymbol{y} \geq \mathbf{0}),$$

where $\boldsymbol{\mu}_{\boldsymbol{s}} = \boldsymbol{\Lambda}_{\boldsymbol{s}}\boldsymbol{\mu}$, $\boldsymbol{\Sigma}_{\boldsymbol{s}} = \boldsymbol{\Lambda}_{\boldsymbol{s}}\boldsymbol{\Sigma}\boldsymbol{\Lambda}_{\boldsymbol{s}}$, and the cumulative distribution function of $\boldsymbol{Y}$ is simply

$$F_{\boldsymbol{Y}}(\boldsymbol{y}) = \Pr[-\boldsymbol{y} \leq \boldsymbol{X} \leq \boldsymbol{y}] = L_n(-\boldsymbol{y}, \boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \qquad (\boldsymbol{y} \geq \mathbf{0}).$$

Using the same derivations as in Tallis (1961), it is easy to show that

$$\begin{aligned} \int_{\mathbf{0}}^{\infty} e^{\boldsymbol{t}^{\mathrm{T}}\boldsymbol{y}} \phi_n(\boldsymbol{y}; \boldsymbol{\mu}_{\boldsymbol{s}}, \boldsymbol{\Sigma}_{\boldsymbol{s}}) \mathrm{d}\boldsymbol{y} &= \int_{\mathbf{0}}^{\infty} \frac{1}{(2\pi)^{\frac{n}{2}} |\boldsymbol{\Sigma}_{\boldsymbol{s}}|^{\frac{1}{2}}} e^{\boldsymbol{t}^{\mathrm{T}}\boldsymbol{y} - \frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}_{\boldsymbol{s}})^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{s}}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_{\boldsymbol{s}})} \mathrm{d}\boldsymbol{y} \\ &= e^{\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{s}} + \frac{\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{s}}\boldsymbol{t}}{2}} \int_{\mathbf{0}}^{\infty} \frac{1}{(2\pi)^{\frac{n}{2}} |\boldsymbol{\Sigma}_{\boldsymbol{s}}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}_{\boldsymbol{s}}-\boldsymbol{\Sigma}_{\boldsymbol{s}}\boldsymbol{t})^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{s}}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_{\boldsymbol{s}}-\boldsymbol{\Sigma}_{\boldsymbol{s}}\boldsymbol{t})} \mathrm{d}\boldsymbol{y} \\ &= e^{\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{s}} + \frac{\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{s}}\boldsymbol{t}}{2}} \Phi_n(\boldsymbol{\mu}_{\boldsymbol{s}} + \boldsymbol{\Sigma}_{\boldsymbol{s}}\boldsymbol{t}; \boldsymbol{\Sigma}_{\boldsymbol{s}}). \end{aligned}$$

It follows that the moment generating function of $\boldsymbol{Y}$ is

$$m_{\boldsymbol{Y}}(\boldsymbol{t}) = E(e^{\boldsymbol{t}^{\mathrm{T}}\boldsymbol{y}}) = \sum_{\boldsymbol{s} \in \mathcal{S}(n)} e^{\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\mu}_{\boldsymbol{s}} + \frac{\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{s}}\boldsymbol{t}}{2}} \Phi_n(\boldsymbol{\mu}_{\boldsymbol{s}} + \boldsymbol{\Sigma}_{\boldsymbol{s}}\boldsymbol{t}; \boldsymbol{\Sigma}_{\boldsymbol{s}}).$$

While it is possible to differentiate $m_{\boldsymbol{Y}}(\boldsymbol{t})$ to obtain the product moments of $\boldsymbol{Y}$, it is much easier to compute the product moments of $\boldsymbol{Y}$ using our $I_{\boldsymbol{\kappa}}^n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ function. Specifically, we have

$$E(\boldsymbol{Y}^{\boldsymbol{\kappa}}) = \sum_{\boldsymbol{s} \in \mathcal{S}(n)} \int_{\mathbf{0}}^{\infty} \boldsymbol{y}^{\boldsymbol{\kappa}} \phi_n(\boldsymbol{y}; \boldsymbol{\mu}_{\boldsymbol{s}}, \boldsymbol{\Sigma}_{\boldsymbol{s}}) \mathrm{d}\boldsymbol{y} = \sum_{\boldsymbol{s} \in \mathcal{S}(n)} I_{\boldsymbol{\kappa}}^n(\boldsymbol{\mu}_{\boldsymbol{s}}, \boldsymbol{\Sigma}_{\boldsymbol{s}}).$$

All we need is to evaluate 2^n different $I_{\boldsymbol{\kappa}}^n(\boldsymbol{\mu}_{\boldsymbol{s}}, \boldsymbol{\Sigma}_{\boldsymbol{s}})$ to obtain $E(\boldsymbol{Y}^{\boldsymbol{\kappa}})$. Using our recurrence relation in (5), these calculations are very fast even for moderately large n. For example, when running our Matlab program on a PC with an Intel i7-4790K CPU, it takes 3.7 seconds to compute $E(\boldsymbol{Y}^{\boldsymbol{\nu}})$ for $\mathbf{0} \leq \boldsymbol{\nu} \leq (5,5,5,5)^{\mathrm{T}}$ when $n = 4$, and 45.2 seconds to compute $E(\boldsymbol{Y}^{\boldsymbol{\nu}})$ for $\mathbf{0} \leq \boldsymbol{\nu} \leq (5,5,5,5,5)^{\mathrm{T}}$ when $n = 5$.

3.1 Explicit Expressions for Low Order Moments

The recurrence relation for $I^n_{\boldsymbol{\kappa}}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ can be used to obtain explicit expressions for the product moments of $\boldsymbol{Y}$. In the following, we provide explicit expressions for $E(\boldsymbol{Y}^{\boldsymbol{\kappa}})$ up to the fourth order, i.e., $\sum_{i=1}^n k_i \leq 4$. In our expressions, we assume $\sigma_1 = \cdots = \sigma_n = 1$. This implies that $\boldsymbol{\Sigma} = \boldsymbol{R}$, where $\boldsymbol{R} = (\rho_{ij})$ is the correlation matrix of $\boldsymbol{X}$, with $\rho_{ij} = \sigma_{ij}/(\sigma_i\sigma_j)$. For the general $\boldsymbol{\Sigma}$ case, we just need to replace μ_i in our expressions with μ_i/σ_i, and then multiply the result by $\sigma_1^{k_1} \cdots \sigma_n^{k_n}$.

For univariate moments, Winkelbauer (2012) shows that

$$E(Y_i^k) = \frac{2^{\frac{k}{2}}\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi}} {}_1F_1\left(-\frac{k}{2}; \frac{1}{2}; -\frac{\mu_i^2}{2}\right),$$

where ${}_1F_1(a; b; z)$ is the confluent hypergeometric function. It follows that the first four moments of Y_i are

$$\begin{aligned}
E(Y_i) &= \mu_i \operatorname{erf}\left(\frac{\mu_i}{\sqrt{2}}\right) + 2\phi(\mu_i), \\
E(Y_i^2) &= 1 + \mu_i^2, \\
E(Y_i^3) &= \mu_i(3 + \mu_i^2)\operatorname{erf}\left(\frac{\mu_i}{\sqrt{2}}\right) + (4 + 2\mu_i^2)\phi(\mu_i), \\
E(Y_i^4) &= 3 + 6\mu_i^2 + \mu_i^4,
\end{aligned}$$

where $\operatorname{erf}(\mu_i/\sqrt{2}) = \Phi(\mu_i) - \Phi(-\mu_i)$ is the error function. Using the recurrence relation of confluent hypergeometric functions, higher order moments of Y_i can be obtained using[4]

$$E(Y_i^k) = (\mu_i^2 + 2k - 3)E(Y_i^{k-2}) - (k-2)(k-3)E(Y_i^{k-4}) \qquad (k \geq 4).$$

[4]Elandt (1961) expresses the higher order moments of Y_i in terms of Fisher's I_k functions, and her expression (Eq. 8) is less efficient than ours.

For bivariate moments, we define $z_{i\cdot j} = (\mu_i - \rho_{ij}\mu_j)/(1-\rho_{ij}^2)^{1/2}$ and use (5) repeatedly to obtain

$$
\begin{aligned}
I^2_{(1,1)}((\mu_i,\mu_j)^{\mathrm{T}},\rho_{ij}) &= (\mu_i\mu_j + \rho_{ij})\Phi_2(\mu_i,\mu_j;\rho_{ij}) + \mu_i\phi(\mu_j)\Phi(z_{i\cdot j}) \\
&\quad + \mu_j\phi(\mu_i)\Phi(z_{j\cdot i}) + (1-\rho_{ij}^2)\phi_2(\mu_i,\mu_j;\rho_{ij}), \\
I^2_{(2,1)}((\mu_i,\mu_j)^{\mathrm{T}},\rho_{ij}) &= \left\{(1+\mu_i^2)\mu_j + 2\rho_{ij}\mu_i\right\}\Phi_2(\mu_i,\mu_j;\rho_{ij}) + (\mu_i\mu_j + 2\rho_{ij})\phi(\mu_i)\Phi(z_{j\cdot i}) \\
&\quad + (1+\mu_i^2+\rho_{ij}^2)\phi(\mu_j)\Phi(z_{i\cdot j}) + \mu_i(1-\rho_{ij}^2)\phi_2(\mu_i,\mu_j;\rho_{ij}), \\
I^2_{(3,1)}((\mu_i,\mu_j)^{\mathrm{T}},\rho_{ij}) &= \left\{\mu_i\mu_j(3+\mu_i^2) + 3\rho_{ij}(1+\mu_i^2)\right\}\Phi_2(\mu_i,\mu_j;\rho_{ij}) \\
&\quad + \left\{(2+\mu_i^2)\mu_j + 3\mu_i\rho_{ij}\right\}\phi(\mu_i)\Phi(z_{j\cdot i}) \\
&\quad + \left\{\mu_i^3 + 3\mu_i(1+\rho_{ij}^2) - \mu_j\rho_{ij}^3\right\}\phi(\mu_j)\Phi(z_{i\cdot j}) \\
&\quad + (2+\mu_i^2+\rho_{ij}^2)(1-\rho_{ij}^2)\phi_2(\mu_i,\mu_j;\rho_{ij})
\end{aligned}
$$

for $i \neq j$. Summing up these terms for the four quadrants, i.e., with (μ_i,μ_j,ρ_{ij}) in the above expressions replaced by $(\mu_i,-\mu_j,-\rho_{ij})$, $(-\mu_i,\mu_j,-\rho_{ij})$, and $(-\mu_i,-\mu_j,\rho_{ij})$ in the other three quadrants, we obtain for $i \neq j$

$$
\begin{aligned}
E(Y_iY_j) &= (\mu_i\mu_j + \rho_{ij})p_2(\mu_i,\mu_j;\rho_{ij}) + 2\mu_i\phi(\mu_j)\,\mathrm{erf}\left(\frac{z_{i\cdot j}}{\sqrt{2}}\right) \\
&\quad + 2\mu_j\phi(\mu_i)\,\mathrm{erf}\left(\frac{z_{j\cdot i}}{\sqrt{2}}\right) + 4(1-\rho_{ij}^2)\phi_2(\mu_i,\mu_j;\rho_{ij}), \\
E(Y_i^2Y_j) &= \left\{(1+\mu_i^2)\mu_j + 2\rho_{ij}\mu_i\right\}\mathrm{erf}\left(\frac{\mu_j}{\sqrt{2}}\right) + 2(1+\mu_i^2+\rho_{ij}^2)\phi(\mu_j) \\
E(Y_i^3Y_j) &= \left\{\mu_i\mu_j(3+\mu_i^2) + 3\rho_{ij}(1+\mu_i^2)\right\}p_2(\mu_i,\mu_j;\rho_{ij}) \\
&\quad + 2\left\{(2+\mu_i^2)\mu_j + 3\rho_{ij}\mu_i\right\}\phi(\mu_i)\,\mathrm{erf}\left(\frac{z_{j\cdot i}}{\sqrt{2}}\right) \\
&\quad + 2\left\{\mu_i^3 + 3\mu_i(1+\rho_{ij}^2) - \mu_j\rho_{ij}^3\right\}\phi(\mu_j)\,\mathrm{erf}\left(\frac{z_{i\cdot j}}{\sqrt{2}}\right) \\
&\quad + 4(2+\mu_i^2+\rho_{ij}^2)(1-\rho_{ij}^2)\phi_2(\mu_i,\mu_j;\rho_{ij}),
\end{aligned}
$$

$$E(Y_i^2Y_j^2) = E(X_i^2X_j^2) = (1+\mu_i^2)(1+\mu_j^2) + 4\mu_i\mu_j\rho_{ij} + 2\rho_{ij}^2,$$

where

$$p_2(\mu_i, \mu_j; \rho_{ij}) = 4\Phi_2(\mu_i, \mu_j; \rho_{ij}) - 2\Phi(\mu_i) - 2\Phi(\mu_j) + 1.$$

With the univariate and bivariate moments of $\boldsymbol{Y}$ available, the expected value and covariance matrix of $\boldsymbol{Y}$ for the general $\boldsymbol{\Sigma}$ case are

$$\begin{aligned}
E(Y_i) &= \mu_i \operatorname{erf}\left(\frac{\tilde{\mu}_i}{\sqrt{2}}\right) + 2\sigma_i\phi(\tilde{\mu}_i), \\
\operatorname{var}(Y_i) &= \mu_i^2 + \sigma_i^2 - E(Y_i)^2, \\
\operatorname{cov}(Y_i, Y_j) &= (\mu_i\mu_j + \sigma_{ij})\{4\Phi_2(\tilde{\mu}_i, \tilde{\mu}_j; \rho_{ij}) - 2\Phi(\tilde{\mu}_i) - 2\Phi(\tilde{\mu}_j) + 1\} \\
&\quad + 2\mu_i\sigma_j\phi(\tilde{\mu}_j)\operatorname{erf}\left(\frac{\tilde{\mu}_i - \rho_{ij}\tilde{\mu}_j}{\sqrt{2}(1-\rho_{ij}^2)^{\frac{1}{2}}}\right) + 2\mu_j\sigma_i\phi(\tilde{\mu}_i)\operatorname{erf}\left(\frac{\tilde{\mu}_j - \rho_{ij}\tilde{\mu}_i}{\sqrt{2}(1-\rho_{ij}^2)^{\frac{1}{2}}}\right) \\
&\quad + 4\sigma_i\sigma_j(1-\rho_{ij}^2)\phi_2(\tilde{\mu}_i, \tilde{\mu}_j; \rho_{ij}) - E(Y_i)E(Y_j),
\end{aligned}$$

where $\tilde{\mu}_i = \mu_i/\sigma_i$.

For trivariate moments, we define $z_{i\cdot jk} = (z_{i\cdot k} - \rho_{ij\cdot k}z_{j\cdot k})/(1-\rho_{ij\cdot k}^2)^{1/2}$, where $\rho_{ij\cdot k} = (\rho_{ij} - \rho_{ik}\rho_{jk})/\{(1-\rho_{ik}^2)(1-\rho_{jk}^2)\}^{1/2}$. Let $\tilde{\boldsymbol{\mu}} = (\mu_i, \mu_j, \mu_k)^{\mathrm{T}}$ and $\tilde{\boldsymbol{R}}$ be a 3 by 3 submatrix of $\boldsymbol{R}$ that consists of the (i,j,k)th rows and columns of $\boldsymbol{R}$. Applying (5) repeatedly, we obtain

$$\begin{aligned}
I^3_{(1,1,1)}(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{R}}) &= (\mu_i\mu_j\mu_k + \mu_i\rho_{jk} + \mu_j\rho_{ik} + \mu_k\rho_{ij})\Phi_3(\tilde{\boldsymbol{\mu}}; \tilde{\boldsymbol{R}}) \\
&\quad + (\mu_j\mu_k + \rho_{ij}\rho_{ik} + \rho_{jk})\phi(\mu_i)\Phi_2(z_{j\cdot i}, z_{k\cdot i}; \rho_{jk\cdot i}) \\
&\quad + (\mu_i\mu_k + \rho_{ij}\rho_{jk} + \rho_{ik})\phi(\mu_j)\Phi_2(z_{i\cdot j}, z_{k\cdot j}; \rho_{ik\cdot j}) \\
&\quad + (\mu_i\mu_j + \rho_{ik}\rho_{jk} + \rho_{ij})\phi(\mu_k)\Phi_2(z_{i\cdot k}, z_{j\cdot k}; \rho_{ij\cdot k}) \\
&\quad + \mu_i(1-\rho_{jk}^2)\phi_2(\mu_j, \mu_k; \rho_{jk})\Phi(z_{i\cdot jk}) \\
&\quad + \mu_j(1-\rho_{ik}^2)\phi_2(\mu_i, \mu_k; \rho_{jk})\Phi(z_{j\cdot ik})
\end{aligned}$$

$$
\begin{aligned}
&+ \mu_k(1-\rho_{ij}^2)\phi_2(\mu_i,\mu_j;\rho_{ij})\Phi(z_{k\cdot ij}) + |\tilde{\boldsymbol{R}}|\phi_3(\tilde{\boldsymbol{\mu}};\tilde{\boldsymbol{R}}),\\
I^3_{(2,1,1)}(\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{R}}) = &\{(1+\mu_i^2)(\mu_j\mu_k+\rho_{jk}) + 2\mu_i(\mu_j\rho_{ik}+2\mu_k\rho_{ij}) + 2\rho_{ij}\rho_{ik}\}\Phi_3(\tilde{\boldsymbol{\mu}};\tilde{\boldsymbol{R}})\\
&+ \{\mu_i(\mu_j\mu_k+\rho_{jk}) + 2\mu_j\rho_{ik} + 2\mu_k\rho_{ij}\}\phi(\mu_i)\Phi_2(z_{j\cdot i}, z_{k\cdot i};\rho_{jk\cdot i})\\
&+ \{2\mu_i(\rho_{ik}+\rho_{ij}\rho_{jk}) + \mu_k(1+\mu_i^2+\rho_{ij}^2) - \mu_j\rho_{ij}^2\rho_{jk})\}\phi(\mu_j)\Phi_2(z_{i\cdot j}, z_{k\cdot j};\rho_{ik\cdot j})\\
&+ \{2\mu_i(\rho_{ij}+\rho_{ik}\rho_{jk}) + \mu_j(1+\mu_i^2+\rho_{ik}^2) - \mu_k\rho_{ik}^2\rho_{jk})\}\phi(\mu_k)\Phi_2(z_{i\cdot k}, z_{j\cdot k};\rho_{ij\cdot k})\\
&+ (\mu_i\mu_k + 2\rho_{ik} + \rho_{ij}\rho_{jk})(1-\rho_{ij}^2)\phi_2(\mu_i,\mu_j;\rho_{ij})\Phi(z_{k\cdot ij})\\
&+ (\mu_i\mu_j + 2\rho_{ij} + \rho_{ik}\rho_{jk})(1-\rho_{ik}^2)\phi_2(\mu_i,\mu_k;\rho_{ik})\Phi(z_{j\cdot ik})\\
&+ (1+\mu_i^2+\rho_{ij}^2+\rho_{ik}^2)(1-\rho_{jk}^2)\phi_2(\mu_j,\mu_k;\rho_{jk})\Phi(z_{k\cdot ij})\\
&+ \mu_i|\tilde{\boldsymbol{R}}|\phi_3(\tilde{\boldsymbol{\mu}};\tilde{\boldsymbol{R}}),
\end{aligned}
$$

where i, j, k are distinct positive integers. Summing up these terms over 8 different values of $(\boldsymbol{\mu_s}, \boldsymbol{\Sigma_s})$ and after simplification, we obtain for distinct positive integers i, j, k,

$$
\begin{aligned}
E(Y_iY_jY_k) = &(\mu_i\mu_j\mu_k + \mu_i\rho_{jk} + \mu_j\rho_{ik} + \mu_k\rho_{ij})p_3(\mu_i,\mu_j,\mu_k;\rho_{ij},\rho_{ik},\rho_{jk})\\
&+ 2(\mu_j\mu_k + \rho_{ij}\rho_{ik} + \rho_{jk})\phi(\mu_i)p_2(z_{j\cdot i}, z_{k\cdot i};\rho_{jk\cdot i})\\
&+ 2(\mu_i\mu_k + \rho_{ij}\rho_{jk} + \rho_{ik})\phi(\mu_j)p_2(z_{i\cdot j}, z_{k\cdot j};\rho_{ik\cdot j})\\
&+ 2(\mu_i\mu_j + \rho_{ik}\rho_{jk} + \rho_{ij})\phi(\mu_k)p_2(z_{i\cdot k}, z_{j\cdot k};\rho_{ij\cdot k})\\
&+ 4\mu_i(1-\rho_{jk}^2)\phi_2(\mu_j,\mu_k;\rho_{jk})\,\mathrm{erf}\left(\frac{z_{i\cdot jk}}{\sqrt{2}}\right)\\
&+ 4\mu_j(1-\rho_{ik}^2)\phi_2(\mu_i,\mu_k;\rho_{ik})\,\mathrm{erf}\left(\frac{z_{j\cdot ik}}{\sqrt{2}}\right)\\
&+ 4\mu_k(1-\rho_{ij}^2)\phi_2(\mu_i,\mu_j;\rho_{ij})\,\mathrm{erf}\left(\frac{z_{k\cdot ij}}{\sqrt{2}}\right) + 8|\tilde{\boldsymbol{R}}|\phi_3(\tilde{\boldsymbol{\mu}};\tilde{\boldsymbol{R}}),\\
E(Y_i^2Y_jY_k) = &\{(1+\mu_i^2)(\mu_j\mu_k+\rho_{jk})\\
&+ 2\mu_i(\mu_j\rho_{ik}+\mu_k\rho_{ij}) + 2\rho_{ij}\rho_{ik}\}p_2(\mu_j,\mu_k;\rho_{jk})
\end{aligned}
$$

$$+ 2\{2\mu_i(\rho_{ik} + \rho_{ij}\rho_{jk}) + \mu_k(1 + \mu_i^2 + \rho_{ij}^2) - \mu_j\rho_{ij}^2\rho_{jk}\}\phi(\mu_j)\,\mathrm{erf}\left(\frac{z_{k\cdot j}}{\sqrt{2}}\right)$$

$$+ 2\{2\mu_i(\rho_{ij} + \rho_{ik}\rho_{jk}) + \mu_j(1 + \mu_i^2 + \rho_{ik}^2) - \mu_k\rho_{ik}^2\rho_{jk}\}\phi(\mu_k)\,\mathrm{erf}\left(\frac{z_{j\cdot k}}{\sqrt{2}}\right)$$

$$+ 4(1 + \mu_i^2 + \rho_{ij}^2 + \rho_{ik}^2)(1 - \rho_{jk}^2)\phi_2(\mu_j, \mu_k; \rho_{jk}),$$

where

$$p_3(\mu_i, \mu_j, \mu_k; \rho_{ij}, \rho_{ik}, \rho_{jk})$$

$$= 8\Phi_3\left(\begin{bmatrix}\mu_i\\ \mu_j\\ \mu_k\end{bmatrix}; \begin{bmatrix}1 & \rho_{ij} & \rho_{ik}\\ \rho_{ij} & 1 & \rho_{jk}\\ \rho_{ik} & \rho_{jk} & 1\end{bmatrix}\right) - 4\Phi_2(\mu_i, \mu_j; \rho_{ij}) - 4\Phi_2(\mu_i, \mu_k; \rho_{ik})$$

$$- 4\Phi_2(\mu_j, \mu_k; \rho_{jk}) + 2\Phi(\mu_i) + 2\Phi(\mu_j) + 2\Phi(\mu_k) - 1.$$

For the fourth order product moment, we define $z_{i\cdot jkl} = (z_{i\cdot kl} - \rho_{ij\cdot kl}z_{j\cdot kl})/(1 - \rho_{ij\cdot kl}^2)^{1/2}$, where $\rho_{ij\cdot kl} = (\rho_{ij\cdot l} - \rho_{ik\cdot l}\rho_{jk\cdot l})/\{(1 - \rho_{ik\cdot l}^2)(1 - \rho_{jk\cdot l}^2)\}^{1/2}$. Let $\hat{\boldsymbol{\mu}} = (\mu_i, \mu_j, \mu_k, \mu_l)^{\mathrm{T}}$ and $\hat{\boldsymbol{R}}$ be a 4 by 4 submatrix of $\boldsymbol{R}$ that consists of the (i, j, k, l)th rows and columns of $\boldsymbol{R}$. Applying (5) repeatedly to obtain $I^4_{(1,1,1,1)}(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{R}})$ and then summing up the expression for 16 different values of $(\boldsymbol{\mu_s}, \boldsymbol{\Sigma_s})$, we obtain for distinct positive integers i, j, k, l,

$$E(Y_iY_jY_kY_l) = (\mu_i\mu_j\mu_k\mu_l + \mu_i\mu_j\rho_{kl} + \mu_i\mu_k\rho_{jl} + \mu_i\mu_l\rho_{jk} + \mu_j\mu_k\rho_{il} + \mu_j\mu_l\rho_{ik} + \mu_k\mu_l\rho_{ij}$$

$$+ \rho_{ij}\rho_{kl} + \rho_{ik}\rho_{jl} + \rho_{il}\rho_{jk})p_4(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{R}})$$

$$+ 2\{\mu_j\mu_k\mu_l + \mu_j(\rho_{ik}\rho_{il} + \rho_{kl}) + \mu_k(\rho_{ij}\rho_{il} + \rho_{jl}) + \mu_l(\rho_{ij}\rho_{ik} + \rho_{jk})$$

$$- \mu_i\rho_{ij}\rho_{ik}\rho_{il}\}\phi(\mu_i)p_3(z_{j\cdot i}, z_{k\cdot i}, z_{l\cdot i}; \rho_{jk\cdot i}, \rho_{jl\cdot i}, \rho_{kl\cdot i})$$

$$+ 2\{\mu_i\mu_k\mu_l + \mu_i(\rho_{jk}\rho_{jl} + \rho_{kl}) + \mu_k(\rho_{ij}\rho_{jl} + \rho_{il}) + \mu_l(\rho_{ij}\rho_{jk} + \rho_{ik})$$

$$- \mu_j\rho_{ij}\rho_{jk}\rho_{jl}\}\phi(\mu_j)p_3(z_{i\cdot j}, z_{k\cdot j}, z_{l\cdot j}; \rho_{ik\cdot j}, \rho_{il\cdot j}, \rho_{kl\cdot j})$$

$$+ 2\{\mu_i\mu_j\mu_l + \mu_i(\rho_{jk}\rho_{kl} + \rho_{jl}) + \mu_j(\rho_{ik}\rho_{kl} + \rho_{il}) + \mu_l(\rho_{ik}\rho_{jk} + \rho_{ij})$$

$$
\begin{aligned}
&- \mu_k\rho_{ik}\rho_{jk}\rho_{kl}\}\phi(\mu_k)p_3(z_{i\cdot k}, z_{j\cdot k}, z_{l\cdot k}; \rho_{ij\cdot k}, \rho_{il\cdot k}, \rho_{jl\cdot k}) \\
&+ 2\{\mu_i\mu_j\mu_k + \mu_i(\rho_{jl}\rho_{kl} + \rho_{jk}) + \mu_j(\rho_{il}\rho_{kl} + \rho_{ik}) + \mu_k(\rho_{il}\rho_{jl} + \rho_{ij}) \\
&- \mu_l\rho_{il}\rho_{jl}\rho_{kl}\}\phi(\mu_l)p_3(z_{i\cdot l}, z_{j\cdot l}, z_{k\cdot l}; \rho_{ij\cdot l}, \rho_{ik\cdot l}, \rho_{jk\cdot l}) \\
&+ 4(1-\rho_{ij}^2)(\mu_k\mu_l + \rho_{ik}\rho_{il} + \rho_{jk}\rho_{jl} + \rho_{kl})\phi_2(\mu_i, \mu_j; \rho_{ij})p_2(z_{k\cdot ij}, z_{l\cdot ij}; \rho_{kl\cdot ij}) \\
&+ 4(1-\rho_{ik}^2)(\mu_j\mu_l + \rho_{ij}\rho_{il} + \rho_{jk}\rho_{kl} + \rho_{jl})\phi_2(\mu_i, \mu_k; \rho_{ik})p_2(z_{j\cdot ik}, z_{l\cdot ik}; \rho_{jl\cdot ik}) \\
&+ 4(1-\rho_{il}^2)(\mu_j\mu_k + \rho_{ij}\rho_{ik} + \rho_{jl}\rho_{kl} + \rho_{jk})\phi_2(\mu_i, \mu_l; \rho_{il})p_2(z_{j\cdot il}, z_{k\cdot il}; \rho_{jk\cdot il}) \\
&+ 4(1-\rho_{jk}^2)(\mu_i\mu_l + \rho_{ij}\rho_{jl} + \rho_{ik}\rho_{kl} + \rho_{il})\phi_2(\mu_j, \mu_k; \rho_{jk})p_2(z_{i\cdot jk}, z_{l\cdot jk}; \rho_{il\cdot jk}) \\
&+ 4(1-\rho_{jl}^2)(\mu_i\mu_k + \rho_{ij}\rho_{jk} + \rho_{il}\rho_{kl} + \rho_{ik})\phi_2(\mu_j, \mu_l; \rho_{jl})p_2(z_{i\cdot jl}, z_{k\cdot jl}; \rho_{ik\cdot jl}) \\
&+ 4(1-\rho_{kl}^2)(\mu_i\mu_j + \rho_{ik}\rho_{jk} + \rho_{il}\rho_{jl} + \rho_{ij})\phi_2(\mu_k, \mu_l; \rho_{kl})p_2(z_{i\cdot kl}, z_{j\cdot kl}; \rho_{ij\cdot kl}) \\
&+ 8\mu_i|\hat{\boldsymbol{R}}_{(1),(1)}|\phi_3(\hat{\boldsymbol{\mu}}_{(1)}; \hat{\boldsymbol{R}}_{(1),(1)})\,\mathrm{erf}\left(\frac{z_{i\cdot jkl}}{\sqrt{2}}\right) \\
&+ 8\mu_j|\hat{\boldsymbol{R}}_{(2),(2)}|\phi_3(\hat{\boldsymbol{\mu}}_{(2)}; \hat{\boldsymbol{R}}_{(2),(2)})\,\mathrm{erf}\left(\frac{z_{j\cdot ikl}}{\sqrt{2}}\right) \\
&+ 8\mu_k|\hat{\boldsymbol{R}}_{(3),(3)}|\phi_3(\hat{\boldsymbol{\mu}}_{(3)}; \hat{\boldsymbol{R}}_{(3),(3)})\,\mathrm{erf}\left(\frac{z_{k\cdot ijl}}{\sqrt{2}}\right) \\
&+ 8\mu_l|\hat{\boldsymbol{R}}_{(4),(4)}|\phi_3(\hat{\boldsymbol{\mu}}_{(4)}; \hat{\boldsymbol{R}}_{(4),(4)})\,\mathrm{erf}\left(\frac{z_{l\cdot ijk}}{\sqrt{2}}\right) + 16|\hat{\boldsymbol{R}}|\phi_4(\hat{\boldsymbol{\mu}}; \hat{\boldsymbol{R}}),
\end{aligned}
$$

where

$$
\begin{aligned}
p_4(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{R}}) = {} & 16\Phi_4(\hat{\boldsymbol{\mu}}; \hat{\boldsymbol{R}}) - 8\Phi_3(\hat{\boldsymbol{\mu}}_{(1)}; \hat{\boldsymbol{R}}_{(1),(1)}) - 8\Phi_3(\hat{\boldsymbol{\mu}}_{(2)}; \hat{\boldsymbol{R}}_{(2),(2)}) \\
& - 8\Phi_3(\hat{\boldsymbol{\mu}}_{(3)}; \hat{\boldsymbol{R}}_{(3),(3)}) - 8\Phi_3(\hat{\boldsymbol{\mu}}_{(4)}; \hat{\boldsymbol{R}}_{(4),(4)}) + 4\Phi_2(\mu_i, \mu_j; \rho_{ij}) + 4\Phi_2(\mu_i, \mu_k; \rho_{ik}) \\
& + 4\Phi_2(\mu_i, \mu_l; \rho_{il}) + 4\Phi_2(\mu_j, \mu_k; \rho_{jk}) + 4\Phi_2(\mu_j, \mu_l; \rho_{jl}) + 4\Phi_2(\mu_k, \mu_l; \rho_{kl}) \\
& - 2\Phi(\mu_i) - 2\Phi(\mu_j) - 2\Phi(\mu_k) - 2\Phi(\mu_l) + 1.
\end{aligned}
$$

4 Truncated Multivariate Normal

The doubly truncated multivariate normal distribution is obtained by conditioning on $\boldsymbol{a} \leq \boldsymbol{X} \leq \boldsymbol{b}$, where $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Let $\boldsymbol{Z}$ be the resulting truncated normal random vector with density function

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = \frac{\phi_n(\boldsymbol{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma})}{L_n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})} \quad (\boldsymbol{a} \leq \boldsymbol{z} \leq \boldsymbol{b}).$$

The cumulative distribution function of $\boldsymbol{Z}$ is

$$F_{\boldsymbol{Z}}(\boldsymbol{z}) = \frac{1}{L_n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})} \int_{\boldsymbol{a}}^{\boldsymbol{z}} \phi_n(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \mathrm{d}\boldsymbol{x} = \frac{L_n(\boldsymbol{a}, \boldsymbol{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma})}{L_n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})} \quad (\boldsymbol{a} \leq \boldsymbol{z} \leq \boldsymbol{b}).$$

Generalizing the results in Tallis (1961), it is easy to show that the moment generating function of $\boldsymbol{Z}$ is

$$m_{\boldsymbol{Z}}(\boldsymbol{t}) = E(e^{\boldsymbol{t}^{\mathrm{T}}\boldsymbol{z}}) = \frac{1}{L_n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})} e^{\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\mu} + \frac{\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{t}}{2}} L_n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu} + \boldsymbol{\Sigma}\boldsymbol{t}, \boldsymbol{\Sigma}).$$

In principle, one could differentiate this moment generating function to obtain $E(\boldsymbol{Z}^{\boldsymbol{\kappa}}) = E(\boldsymbol{X}^{\boldsymbol{\kappa}} \mid \boldsymbol{a} \leq \boldsymbol{X} \leq \boldsymbol{b})$, but for higher order moments, these calculations are extremely tedious, and the resulting expressions are not computationally efficient. Instead, we express $E(\boldsymbol{Z}^{\boldsymbol{\kappa}})$ in terms of our $F_{\boldsymbol{\kappa}}^n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ in Section 2 as

$$E(\boldsymbol{Z}^{\boldsymbol{\kappa}}) = \frac{1}{L_n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})} \int_{\boldsymbol{a}}^{\boldsymbol{b}} \boldsymbol{z}^{\boldsymbol{\kappa}} \phi_n(\boldsymbol{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \mathrm{d}\boldsymbol{z} = \frac{F_{\boldsymbol{\kappa}}^n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})}{L_n(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})}.$$

Using our recurrence relation in Theorem 1, the computation of $E(\boldsymbol{Z}^{\boldsymbol{\kappa}})$ is very fast even for moderately large n. For example, when running our Matlab program on a PC with an Intel i7-4790K CPU, it takes 0.97 second to compute $E(\boldsymbol{Z}^{\boldsymbol{\nu}})$ for $\mathbf{0} \leq \boldsymbol{\nu} \leq (5, 5, 5, 5)^{\mathrm{T}}$ when $n = 4$, and 10.1 seconds to compute $E(\boldsymbol{Z}^{\boldsymbol{\nu}})$ for $\mathbf{0} \leq \boldsymbol{\nu} \leq (5, 5, 5, 5, 5)^{\mathrm{T}}$ when $n = 5$.

Our algorithm allows for the possibility that $a_i = -\infty$, $b_i = \infty$, or both $a_i = -\infty$ and $b_i = \infty$, i.e., no truncation on X_i. When all the a_i's are $-\infty$ (b_i's are ∞), we have the upper (lower) truncated multivariate normal distributions. For these special cases, we can express $E(\boldsymbol{Z}^{\boldsymbol{\kappa}})$ in terms of the $I^n_{\boldsymbol{\kappa}}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ function, which can be computed with a shorter recursion. We first provide an illustration of this method for the lower truncated multivariate normal distribution. In this scenario, we can write $E(\boldsymbol{Z}^{\boldsymbol{\kappa}})$ as

$$\begin{aligned} E(\boldsymbol{Z}^{\boldsymbol{\kappa}}) &= \frac{1}{L_n(\boldsymbol{a}, \boldsymbol{\infty}; \boldsymbol{\mu}, \boldsymbol{\Sigma})} \int_{\boldsymbol{a}}^{\boldsymbol{\infty}} \boldsymbol{z}^{\boldsymbol{\kappa}} \phi_n(\boldsymbol{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \mathrm{d}\boldsymbol{z} \\ &= \frac{1}{\Phi_n(\boldsymbol{\mu} - \boldsymbol{a}; \boldsymbol{\Sigma})} \int_{\boldsymbol{0}}^{\boldsymbol{\infty}} (\boldsymbol{y} + \boldsymbol{a})^{\boldsymbol{\kappa}} \phi_n(\boldsymbol{y}; \boldsymbol{\mu} - \boldsymbol{a}, \boldsymbol{\Sigma}) \mathrm{d}\boldsymbol{y} \\ &= \frac{1}{\Phi_n(\boldsymbol{\mu} - \boldsymbol{a}; \boldsymbol{\Sigma})} \sum_{\boldsymbol{0} \le \boldsymbol{\nu} \le \boldsymbol{\kappa}} \binom{\boldsymbol{\kappa}}{\boldsymbol{\nu}} \boldsymbol{a}^{\boldsymbol{\kappa} - \boldsymbol{\nu}} I^n_{\boldsymbol{\nu}}(\boldsymbol{\mu} - \boldsymbol{a}, \boldsymbol{\Sigma}), \end{aligned} \tag{6}$$

where $\boldsymbol{\nu} = (\nu_1, \dots, \nu_n)^{\mathrm{T}}$ and

$$\binom{\boldsymbol{\kappa}}{\boldsymbol{\nu}} = \prod_{i=1}^{n} \frac{k_i!}{\nu_i!(k_i - \nu_i)!}.$$

This alternative expression shows that by using a binomial expansion, we can write $E(\boldsymbol{Z}^{\boldsymbol{\kappa}})$ as a linear combination of $\prod_{i=1}^n (k_i + 1)$ different $I^n_{\boldsymbol{\nu}}(\boldsymbol{\mu} - \boldsymbol{a}, \boldsymbol{\Sigma})$. In computing $I^n_{\boldsymbol{\kappa}}(\boldsymbol{\mu} - \boldsymbol{a}, \boldsymbol{\Sigma})$, all the $I^n_{\boldsymbol{\nu}}(\boldsymbol{\mu} - \boldsymbol{a}, \boldsymbol{\Sigma})$ with $\boldsymbol{0} \le \boldsymbol{\nu} \le \boldsymbol{\kappa}$ have already been computed. Therefore, no additional work is required besides summing up these terms.

Similarly, for the upper truncated case, we can write

$$E(\boldsymbol{Z}^{\boldsymbol{\kappa}}) = \frac{1}{\Phi_n(\boldsymbol{b} - \boldsymbol{\mu}; \boldsymbol{\Sigma})} \sum_{\boldsymbol{0} \le \boldsymbol{\nu} \le \boldsymbol{\kappa}} \binom{\boldsymbol{\kappa}}{\boldsymbol{\nu}} \boldsymbol{b}^{\boldsymbol{\kappa} - \boldsymbol{\nu}} (-1)^{\sum_{i=1}^n \nu_i} I^n_{\boldsymbol{\nu}}(\boldsymbol{b} - \boldsymbol{\mu}, \boldsymbol{\Sigma}).$$

4.1 Explicit Expressions for Low Order Moments

Using our recurrence relation for $F^n_{\boldsymbol{\kappa}}(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$, we present some low order product moments for the lower truncated multivariate normal. For the upper truncated multivariate normal,

$$E(\boldsymbol{Z}^{\boldsymbol{\kappa}}) = (-1)^{\sum_{i=1}^n k_i} E((-\boldsymbol{Z})^{\boldsymbol{\kappa}}) = (-1)^{\sum_{i=1}^n k_i} E((-\boldsymbol{X})^{\boldsymbol{\kappa}} \mid -\boldsymbol{X} > -\boldsymbol{b}).$$

Since $-\boldsymbol{X} \sim N(-\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we just need to replace $\boldsymbol{\mu}$ with $-\boldsymbol{\mu}$ and $\boldsymbol{a}$ with $-\boldsymbol{b}$ in the expression for the product moment of a lower truncated multivariate normal, and then multiply the result by $(-1)^{\sum_{i=1}^n k_i}$ to obtain the product moment of an upper truncated multivariate normal.[5] In our derivations, we assume $\sigma_1 = \cdots = \sigma_n = 1$, i.e., $\boldsymbol{\Sigma} = \boldsymbol{R}$. The result for the general $\boldsymbol{\Sigma}$ case can be obtained by replacing a_i with a_i/σ_i, μ_i with μ_i/σ_i, and multiplying the result by $\sigma_1^{k_1} \cdots \sigma_n^{k_n}$.

Let $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)^{\mathrm{T}} = \boldsymbol{\mu} - \boldsymbol{a}$. When $n = 1$, Cohen (1951a) expresses $E(Z^k)$ using Fisher's I_k functions, which is essentially equivalent to (6) for the case of $n = 1$. However, we can also use the recursion for $F_k^1(a, \infty; \mu, 1)$ to obtain the more efficient recursion

$$E(Z^{k+1}) = \mu E(Z^k) + kE(Z^{k-1}) + \frac{a^k \phi(\eta)}{\Phi(\eta)} \quad (k \geq 1),$$

with the boundary condition $E(Z^0) = 1$. Using this recurrence relation, we obtain the first three moments of Z as

$$\begin{aligned} E(Z) &= \mu + \frac{\phi(\eta)}{\Phi(\eta)}, \\ E(Z^2) &= 1 + \mu^2 + \frac{(\mu + a)\phi(\eta)}{\Phi(\eta)}, \\ E(Z^3) &= 3\mu + \mu^3 + \frac{(\mu^2 + a\mu + a^2 + 2)\phi(\eta)}{\Phi(\eta)}. \end{aligned}$$

When $n = 2$, we use (5) and (6) to obtain $E(Z_1^{k_1} Z_2^{k_2})$ for $1 \leq k_1 + k_2 \leq 3$. Specifically, we have

$$\begin{aligned} E(Z_1) &= \mu_1 + \frac{\phi(\eta_1)\Phi(w_{2\cdot 1}) + \rho_{12}\phi(\eta_2)\Phi(w_{1\cdot 2})}{\Phi_2(\eta_1, \eta_2; \rho_{12})}, \\ E(Z_1^2) &= 1 + \mu_1^2 + \frac{(\mu_1 + a_1)\phi(\eta_1)\Phi(w_{2\cdot 1})}{\Phi_2(\eta_1, \eta_2; \rho_{12})} \end{aligned}$$

[5]Although analytically attainable, we do not report the results for the doubly truncated multivariate normal distribution because the expressions of the product moments can be very lengthy.

$$+\frac{\rho_{12}\{(2\mu_1-\rho_{12}\eta_2)\phi(\eta_2)\Phi(w_{1\cdot 2})+(1-\rho_{12}^2)\phi_2(\eta_1,\eta_2;\rho_{12})\}}{\Phi_2(\eta_1,\eta_2;\rho_{12})},$$

$$E(Z_1Z_2)=\mu_1\mu_2+\rho_{12}+\frac{(\mu_2+\rho_{12}a_1)\phi(\eta_1)\Phi(w_{2\cdot 1})+(\mu_1+\rho_{12}a_2)\phi(\eta_2)\Phi(w_{1\cdot 2})}{\Phi_2(\eta_1,\eta_2;\rho_{12})}$$
$$+\frac{(1-\rho_{12}^2)\phi_2(\eta_1,\eta_2;\rho_{12})}{\Phi_2(\eta_1,\eta_2;\rho_{12})},$$

$$E(Z_1^3)=3\mu_1+\mu_1^3+\frac{(\eta_1^2+3a_1\mu_1+2)\phi(\eta_1)\Phi(w_{2\cdot 1})}{\Phi_2(\eta_1,\eta_2;\rho_{12})}$$
$$+\frac{\rho_{12}\{3+3\mu_1^2-3\rho_{12}\mu_1\eta_2+\rho_{12}^2(\eta_2^2-1)\}\phi(\eta_2)\Phi(w_{1\cdot 2})}{\Phi_2(\eta_1,\eta_2;\rho_{12})}$$
$$+\frac{\rho_{12}(1-\rho_{12}^2)(2\mu_1+a_1-\rho_{12}\eta_2)\phi_2(\eta_1,\eta_2;\rho_{12})}{\Phi_2(\eta_1,\eta_2;\rho_{12})},$$

$$E(Z_1^2Z_2)=(1+\mu_1^2)\mu_2+2\rho_{12}\mu_1+\frac{\{(\mu_1+a_1)\mu_2+\rho_{12}(2+a_1^2)\}\phi(\eta_1)\Phi(w_{2\cdot 1})}{\Phi_2(\eta_1,\eta_2;\rho_{12})}$$
$$+\frac{\{1+\mu_1^2+2\rho_{12}a_2\mu_1+\rho_{12}^2(1-a_2\eta_2)\}\phi(\eta_2)\Phi(w_{1\cdot 2})}{\Phi_2(\eta_1,\eta_2;\rho_{12})}$$
$$+\frac{(1-\rho_{12}^2)(\mu_1+a_1+\rho_{12}a_2)\phi_2(\eta_1,\eta_2;\rho_{12})}{\Phi_2(\eta_1,\eta_2;\rho_{12})},$$

where $w_{i\cdot j}=(\eta_i-\rho_{ij}\eta_j)/(1-\rho_{ij}^2)^{1/2}$.

When $n=3$, we again use (5) and (6) to obtain $E(Z_1^{k_1}Z_2^{k_2}Z_3^{k_3})$ for $1\le k_1+k_2+k_3\le 3$. Specifically, we have

$$E(Z_1)=\mu_1+q_1+\rho_{12}q_2+\rho_{13}q_3,$$

$$E(Z_1^2)=1+\mu_1^2+(\mu_1+a_1)q_1+\rho_{12}(2\mu_1-\rho_{12}\eta_2)q_2+\rho_{13}(2\mu_1-\rho_{13}\eta_3)q_3$$
$$+\rho_{12}(1-\rho_{12}^2)h_1+\rho_{13}(1-\rho_{13}^2)h_2+gh_3,$$

$$E(Z_1Z_2)=\mu_1\mu_2+\rho_{12}+(\mu_2+\rho_{12}a_1)q_1+(\mu_1+\rho_{12}a_2)q_2$$
$$+(\rho_{23}\mu_1+\rho_{13}\mu_2-\rho_{13}\rho_{23}\eta_3)q_3+(1-\rho_{12}^2)h_1+\rho_{23}(1-\rho_{13}^2)h_2$$
$$+\rho_{13}(1-\rho_{23}^2)h_3,$$

$$E(Z_1^3)=3\mu_1+\mu_1^3+(\eta_1^2+3a_1\mu_1+2)q_1+\rho_{12}\{3+3\mu_1^2-3\rho_{12}\mu_1\eta_2+\rho_{12}^2(\eta_2^2-1)\}q_2$$

$$
\begin{aligned}
&+ \rho_{13}\{3 + 3\mu_1^2 - 3\rho_{13}\mu_1\eta_3 + \rho_{13}^2(\eta_3^2 - 1)\}q_3 \\
&+ \rho_{12}(1-\rho_{12}^2)(2\mu_1 + a_1 - \rho_{12}\eta_2)h_1 + \rho_{13}(1-\rho_{13}^2)(2\mu_1 + a_1 - \rho_{13}\eta_3)h_2 \\
&+ \Big\{3g\mu_1 + \rho_{23}(\rho_{12}^3\eta_2 + \rho_{13}^3\eta_3) - \frac{\rho_{12}^2(3\rho_{13} - \rho_{12}\rho_{23})w_{2\cdot 3}}{(1-\rho_{23}^2)^{\frac{1}{2}}} \\
&- \frac{\rho_{13}^2(3\rho_{12} - \rho_{13}\rho_{23})w_{3\cdot 2}}{(1-\rho_{23}^2)^{\frac{1}{2}}}\Big\}h_3 + \frac{g|\boldsymbol{R}|\phi_3(\boldsymbol{\eta};\boldsymbol{R})}{(1-\rho_{23}^2)\Phi_3(\boldsymbol{\eta};\boldsymbol{R})},
\end{aligned}
$$

$$
\begin{aligned}
E(Z_1^2 Z_2) = {}& (1+\mu_1^2)\mu_2 + 2\rho_{12}\mu_1 + \{(\mu_1 + a_1)\mu_2 + \rho_{12}(2 + a_1^2)\}q_1 \\
&+ \{1 + \mu_1^2 + 2\rho_{12}a_2\mu_1 + \rho_{12}^2(1 - a_2\eta_2)\}q_2 \\
&+ [\rho_{13}\{2\rho_{12} + \mu_2(2\mu_1 - \rho_{13}\eta_3)\} + \rho_{23}\{1 - \rho_{13}^2 + (\mu_1 - \rho_{13}\eta_3)^2\}]q_3 \\
&+ (1-\rho_{12}^2)(\mu_1 + a_1 + \rho_{12}a_2)h_1 + (1-\rho_{13}^2)\{\mu_2\rho_{13} + \rho_{23}(\mu_1 + a_1 - \rho_{13}\eta_3)\}h_2 \\
&+ \{ga_2 + \rho_{13}(1-\rho_{23}^2)(2\mu_1 - \rho_{13}\eta_3)\}h_3 + \frac{\rho_{13}|\boldsymbol{R}|\phi_3(\boldsymbol{\eta};\boldsymbol{R})}{\Phi_3(\boldsymbol{\eta};\boldsymbol{R})},
\end{aligned}
$$

$$
\begin{aligned}
E(Z_1 Z_2 Z_3) = {}& \mu_1\mu_2\mu_3 + \rho_{23}\mu_1 + \rho_{13}\mu_2 + \rho_{12}\mu_3 \\
&+ \{\mu_2\mu_3 + \rho_{12}\rho_{13} + \rho_{23} + a_1(\mu_2\rho_{13} + \mu_3\rho_{12} - \eta_1\rho_{12}\rho_{13})\}q_1 \\
&+ \{\mu_1\mu_3 + \rho_{12}\rho_{23} + \rho_{13} + a_2(\mu_1\rho_{23} + \mu_3\rho_{12} - \eta_2\rho_{12}\rho_{23})\}q_2 \\
&+ \{\mu_1\mu_2 + \rho_{13}\rho_{23} + \rho_{12} + a_3(\mu_1\rho_{23} + \mu_2\rho_{13} - \eta_3\rho_{13}\rho_{23})\}q_3 \\
&+ (1-\rho_{12}^2)(\mu_3 + a_1\rho_{13} + a_2\rho_{23})h_1 + (1-\rho_{13}^2)(\mu_2 + a_1\rho_{12} + a_3\rho_{23})h_2 \\
&+ (1-\rho_{23}^2)(\mu_1 + a_2\rho_{12} + a_3\rho_{13})h_3 + \frac{|\boldsymbol{R}|\phi_3(\boldsymbol{\eta};\boldsymbol{R})}{\Phi_3(\boldsymbol{\eta};\boldsymbol{R})},
\end{aligned}
$$

where

$$
\begin{aligned}
g &= 2\rho_{12}\rho_{13} - \rho_{23}(\rho_{12}^2 + \rho_{13}^2), \\
q_1 &= \phi(\eta_1)\Phi_2(w_{2\cdot 1}, w_{3\cdot 1}; \rho_{23\cdot 1})/\Phi_3(\boldsymbol{\eta};\boldsymbol{R}), \\
q_2 &= \phi(\eta_2)\Phi_2(w_{1\cdot 2}, w_{3\cdot 2}; \rho_{13\cdot 2})/\Phi_3(\boldsymbol{\eta};\boldsymbol{R}), \\
q_3 &= \phi(\eta_3)\Phi_2(w_{1\cdot 3}, w_{2\cdot 3}; \rho_{12\cdot 3})/\Phi_3(\boldsymbol{\eta};\boldsymbol{R}),
\end{aligned}
$$

$$h_1 = \phi_2(\eta_1, \eta_2; \rho_{12})\Phi(w_{3\cdot 12})/\Phi_3(\boldsymbol{\eta}; \boldsymbol{R}),$$
$$h_2 = \phi_2(\eta_1, \eta_3; \rho_{13})\Phi(w_{2\cdot 13})/\Phi_3(\boldsymbol{\eta}; \boldsymbol{R}),$$
$$h_3 = \phi_2(\eta_2, \eta_3; \rho_{23})\Phi(w_{1\cdot 23})/\Phi_3(\boldsymbol{\eta}; \boldsymbol{R}),$$

and $w_{i\cdot jk} = (w_{i\cdot k} - \rho_{ij\cdot k} w_{j\cdot k})/(1 - \rho_{ij\cdot k}^2)^{1/2}$.

5 Conclusion

The results in this paper can be easily generalized to the case of multivariate normal mixtures. Generalizing the results to multivariate elliptical distributions requires a lot more work. Although the product moments of multivariate elliptical distributions can be obtained from the product moments of multivariate normal distributions (see, for example, Berkane and Bentler (1986) and Maruyama and Seo (2003)), it is not clear how to obtain product moments of folded and truncated multivariate elliptical distributions. We leave this topic for future research.

SUPPLEMENTARY MATERIAL

Matlab-package: Matlab-package `ftnorm` contains a set of programs to compute the moment expressions given in the article.

References

Arismendi, J. C. (2013), "Multivariate Truncated Moments," *Journal of Multivariate Analysis*, 117, 41–75.

Berkane, M., and Bentler, P. M. (1986), "Moments of Elliptically Distributed Random Variates," *Statistics & Probability Letters*, 4, 333–335.

Birnbaum, Z., and Meyer, P. L. (1953), "On the Effect of Truncation in Some or All Coordinates of a Multinormal Population," *Journal of the Indian Society of Agricultural Statistics*, 5, 18–28.

Chakraborty, A. K., and Chatterjee, M. (2013), "On Multivariate Folded Normal Distribution," *Sankhyā*, Series B, 75, 1–15.

Cohen, A. C. (1951a), "On Estimating the Mean and Variance of Singly Truncated Normal Frequency Distributions from the First Three Sample Moments," *Annals of the Institute of Statistical Mathematics*, 3, 37–44.

Cohen, A. C. (1951b), "Estimation of Parameters in Truncated Pearson Frequency Distributions," *Annals of Mathematical Statistics*, 22, 256–265.

Cohen, A. C. (1991), *Truncated and Censored Samples: Theory and Applications*, New York: Marcel Dekker.

Dyer, D. D. (1973), "On Moments Estimation of the Parameters of a Truncated Bivariate Normal Distribution," *Journal of the Royal Statistical Society*, Series C, 22, 287–291.

Elandt, R. C. (1961), "The Folded Normal Distribution: Two Methods of Estimating Parameters from Moments," *Technometrics*, 3, 551–562.

Fisher, R. A. (1931), "Properties and Applications of Hh Functions," *Introduction to British A.A.S. Mathematical Tables*, 1, 26–35.

Gupta A. K., and Tracy, D. S. (1976), "Recurrence Relations for the Moments of Truncated Multinormal Distribution," *Communications in Statistics – Theory and Methods*, 5, 855–865.

Kamat, A. R. (1953), "Incomplete and Absolute Moments of the Multivariate Normal Distribution with Some Applications," *Biometrika*, 40, 20–34.

Khatri, C. G., and Jaiswal, M. C. (1963), "Estimation of Parameters of a Truncated Bivariate Normal Distribution," *Journal of the American Statistical Association*, 58, 519–526.

Lee, L. F. (1983), "The Determination of Moments of the Doubly Truncated Multivariate Normal Tobit Model," *Economics Letters*, 11, 245–250.

Leone, F. C., Nelson, L. S., and Nottingham, R. B. (1961), "The Folded Normal Distribution," *Technometrics*, 3, 543–550.

Manjunath, B. G., and Wilhelm, S. (2012), "Moments Calculation for the Doubly Truncated Multivariate Normal Density," arXiv:1206.5387v1.

Maruyama Y., and Seo, T. (2003), "Estimation of Moment Parameter in Elliptical Distributions," *Journal of the Japan Statistical Society*, 33, 215–229.

Murthy, G. S. R. (2015), "A Note on Multivariate Folded Normal Distribution," *Sankhyā*, Series B, 77, 108–113.

Nabeya, S. (1951), "Absolute Moments in 2-Dimensional Normal Distribution," *Annals of the Institute of Statistical Mathematics*, 3, 2–6.

Nabeya, S. (1952), "Absolute Moments in 3-Dimensional Normal Distribution," *Annals of the Institute of Statistical Mathematics*, 4, 15–30.

Nabeya, S. (1961), "Absolute and Incomplete Moments of the Multivariate Normal Distribution," *Biometrika*, 48, 77–84.

Psarakis, S., and Panaretos, J. (2001), "On Some Bivariate Extensions of the Folded Normal and the Folded t Distributions," *Journal of Applied Statistical Science*, 10, 119–136.

Rosenbaum, S. (1961), "Moments of a Truncated Bivariate Normal Distribution," *Journal of the Royal Statistical Society*, Series B, 23, 405–408.

Shah, S. M., and Parikh, N. T. (1964), "Moments of Single and Doubly Truncated Standard Bivariate Normal Distribution," *Vidya*, 7, 82–91.

Tallis, G. M. (1961), "The Moment Generating Function of the Truncated Multi-Normal Distribution," *Journal of the Royal Statistical Society*, Series B, 23, 223–229.

Takemura, A., and Takeuchi, K. (1988), "Some Results on Univariate and Multivariate Cornish-Fisher Expansion: Algebraic Properties and Validity," *Sankhyā*, Series A, 50, 111–136.

Willink, R. (2005), "Normal Moments and Hermite Polynomials," *Statistics & Probability Letters*, 73, 271–275.

Winkelbauer, A. (2012), "Moments and Absolute Moments of the Normal Distribution," arXiv:1209.4340v2.