Miskolc Mathematical Notes
Vol. 25 (2024), No. 1, pp. 287–300

HU e-ISSN 1787-2413
DOI: 10.18514/MMN.2024.4255

NOTE ON THE MOMENT GENERATING FUNCTION OF THE MULTIVARIATE NORMAL DISTRIBUTION

KENICHI HIROSE

To the memory of Michael Cichon

Received 21 May, 2022

Abstract. We present a streamlined proof of a formula for the derivatives of the moment generating function of the multivariate normal distribution. We formulate it in terms of the summation of the contractions by pairings, which encodes a combinatorial computation procedure. We give two applications. First, we provide a simple proof of Isserlis' theorem and derive a formula for the moments of the multivariate normal distribution.Second, we calculate the moments of the product of a finite number of correlated normally and lognormally distributed random variables.

2010 *Mathematics Subject Classification:* 60E05

Keywords: moment generating function, multivariate normal distribution, Isserlis' theorem, normal-lognormal product distribution

1. Introduction

In this note, we present a streamlined proof of a formula for arbitrary derivatives of the moment generating function of the multivariate normal distribution. As applications of the formula, we provide a simple proof of Isserlis' theorem and derive a formula for the moments of the multivariate normal distribution. Then we calculate the moments of the product of a finite number of correlated normally and lognormally distributed random variables.

Throughout this note, we let $\mathbb{N} = \{0, 1, 2, \dots\}$ denote the set of nonnegative integers, $\mathbb{R}$ the set of real numbers, and $\mathbb{R}_+ = \{x \in \mathbb{R} \,|\, x > 0\}$ the set of positive real numbers. Also, we let N denote a positive integer.

2. Derivatives of the moment generating function of the multivariate normal distribution

2.1. *Preliminaries*

Let $\boldsymbol{X} = (X_1, \ldots, X_N)^T$ denote an N-dimensional normally distributed random vector with mean vector $\boldsymbol{m} = (m_i) \in \mathbb{R}^N$ and covariance matrix $C = (c_{ij}) \in \mathbb{R}^{N\times N}$. We have $\mathbb{E}[X_i] = m_i, \mathrm{Cov}(X_i, X_j) = c_{ij}$, and in particular $\mathrm{Var}[X_i] = \sigma_i{}^2 = c_{ii}$. With these notations, the probability density function of $\boldsymbol{X}$ is given by

$$f_N(\boldsymbol{x};\boldsymbol{m},C) = (2\pi)^{-\frac{n}{2}}(\det C)^{-\frac{1}{2}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})^T C^{-1}(\boldsymbol{x}-\boldsymbol{m})\right) \quad \text{for } \boldsymbol{x} \in \mathbb{R}^N,$$

and the moment generating function of $\boldsymbol{X}$ is given by

$$M_{\boldsymbol{X}}(\boldsymbol{t}) = \mathbb{E}[e^{\boldsymbol{t}\cdot\boldsymbol{X}}] = \exp\left(\boldsymbol{m}^T\boldsymbol{t} + \frac{1}{2}\boldsymbol{t}^T C\boldsymbol{t}\right) \quad \text{for } \boldsymbol{t} = (t_1, \ldots, t_N)^T$$

where $\boldsymbol{t}\cdot\boldsymbol{X} = \boldsymbol{t}^T\boldsymbol{X}$ (see [5, Chapter 45]).

By differentiating $M_{\boldsymbol{X}}(\boldsymbol{t})$, we have the following expectation value

$$\mathbb{E}[X_1{}^{\alpha_1}\cdots X_N{}^{\alpha_N}e^{\boldsymbol{t}\cdot\boldsymbol{X}}] = \left(\frac{\partial}{\partial t_1}\right)^{\alpha_1}\cdots\left(\frac{\partial}{\partial t_N}\right)^{\alpha_N} M_{\boldsymbol{X}}(\boldsymbol{t}).$$

Put

$$Q(\boldsymbol{t}) = \boldsymbol{m}^T\boldsymbol{t} + \frac{1}{2}\boldsymbol{t}^T C\boldsymbol{t} = \sum_{i=1}^N m_i t_i + \frac{1}{2}\sum_{i,j=1}^N c_{ij}t_i t_j,$$

so that $M_{\boldsymbol{X}}(\boldsymbol{t}) = e^{Q(\boldsymbol{t})}$. Also put $\boldsymbol{\xi}(\boldsymbol{t}) = \boldsymbol{m} + C\boldsymbol{t}$, namely, $\boldsymbol{\xi}(\boldsymbol{t}) = (\xi_1(\boldsymbol{t}), \ldots, \xi_N(\boldsymbol{t}))^T$ where $\xi_k(\boldsymbol{t}) = m_k + \sum_{j=1}^N c_{kj}t_j$ $(1 \le k \le N)$. Then, for $k, i \in \{1, \ldots, N\}$ we have

$$\frac{\partial Q(\boldsymbol{t})}{\partial t_k} = \xi_k(\boldsymbol{t}), \tag{2.1}$$

$$\frac{\partial \xi_k(\boldsymbol{t})}{\partial t_i} = c_{ki}. \tag{2.2}$$

In other words, $\boldsymbol{\xi}(\boldsymbol{t})$ is the gradient vector of $Q(\boldsymbol{t})$ and C is the Hessian matrix of $Q(\boldsymbol{t})$.

Given any differentiable function $F(\boldsymbol{t})$, we apply the product rule and Eq.(2.1) to get

$$\frac{\partial}{\partial t_k}\left(e^{Q(\boldsymbol{t})}F(\boldsymbol{t})\right) = e^{Q(\boldsymbol{t})}\left(\frac{\partial Q(\boldsymbol{t})}{\partial t_k}F(\boldsymbol{t}) + \frac{\partial F(\boldsymbol{t})}{\partial t_k}\right) = e^{Q(\boldsymbol{t})}\left(\xi_k(\boldsymbol{t})F(\boldsymbol{t}) + \frac{\partial F(\boldsymbol{t})}{\partial t_k}\right). \tag{2.3}$$

By successively applying Eq.(2.3) and Eq.(2.2), one can calculate the derivatives of $M_{\boldsymbol{X}}(\boldsymbol{t}) = e^{Q(\boldsymbol{t})}$.

Example 1. For brevity, we write $Q = Q(\boldsymbol{t})$ and $\xi_k = \xi_k(\boldsymbol{t})$.

(i) $\mathbb{E}[X_1 e^{\boldsymbol{t}\cdot\boldsymbol{X}}] = \frac{\partial}{\partial t_1}M_{\boldsymbol{X}}(\boldsymbol{t}) = e^Q\xi_1$.

(ii) $\mathbb{E}[X_1X_2e^{\boldsymbol{t}\cdot\boldsymbol{X}}] = \frac{\partial}{\partial t_1}\frac{\partial}{\partial t_2}M_{\boldsymbol{X}}(\boldsymbol{t}) = e^Q(\xi_1\xi_2 + c_{12})$. By setting $X_2 = X_1$, we have $\mathbb{E}[X_1{}^2e^{\boldsymbol{t}\cdot\boldsymbol{X}}] = e^Q(\xi_1{}^2 + \sigma_1{}^2)$.

(iii) $\mathbb{E}[X_1X_2X_3e^{\boldsymbol{t}\cdot\boldsymbol{X}}] = \frac{\partial}{\partial t_1}\frac{\partial}{\partial t_2}\frac{\partial}{\partial t_3}M_{\boldsymbol{X}}(\boldsymbol{t}) = e^Q(\xi_1\xi_2\xi_3 + c_{12}\xi_3 + c_{13}\xi_2 + c_{23}\xi_1)$. By setting $X_3 = X_1$, we have $\mathbb{E}[X_1{}^2X_2e^{\boldsymbol{t}\cdot\boldsymbol{X}}] = e^Q(\xi_1{}^2\xi_2 + 2c_{12}\xi_1 + \sigma_1{}^2\xi_2)$. Further, by setting $X_2 = X_1$, we have $\mathbb{E}[X_1{}^3e^{\boldsymbol{t}\cdot\boldsymbol{X}}] = e^Q(\xi_1{}^3 + 3\sigma_1{}^2\xi_1)$.

(iv) $\mathbb{E}[X_1X_2X_3X_4e^{\boldsymbol{t}\cdot\boldsymbol{X}}] = \frac{\partial}{\partial t_1}\frac{\partial}{\partial t_2}\frac{\partial}{\partial t_3}\frac{\partial}{\partial t_4}M_{\boldsymbol{X}}(\boldsymbol{t}) = e^Q(\xi_1\xi_2\xi_3\xi_4 + c_{12}\xi_3\xi_4 + c_{13}\xi_2\xi_4 + c_{14}\xi_2\xi_3 + c_{23}\xi_1\xi_4 + c_{24}\xi_1\xi_3 + c_{34}\xi_1\xi_2 + c_{12}c_{34} + c_{13}c_{24} + c_{14}c_{23})$. By setting $X_3 = X_1$ and $X_4 = X_2$, we have $\mathbb{E}[X_1{}^2X_2{}^2e^{\boldsymbol{t}\cdot\boldsymbol{X}}] = e^Q(\xi_1{}^2\xi_2{}^2 + \sigma_1{}^2\xi_2{}^2 + \sigma_2{}^2\xi_1{}^2 + 4c_{12}\xi_1\xi_2 + 2c_{12}{}^2 + \sigma_1{}^2\sigma_2{}^2)$. Further, by setting $X_2 = X_1$, we have $\mathbb{E}[X_1{}^4e^{\boldsymbol{t}\cdot\boldsymbol{X}}] = e^Q(\xi_1{}^4 + 6\sigma_1{}^2\xi_1{}^2 + 3\sigma_1{}^4)$.

2.2. *The key formula*

By generalizing Example 1, we prove a formula for arbitrary derivatives of the moment generating function of the multivariate normal distribution. Our proof is adapted from [1, Lemma 5.2.], but we slightly extend it to treat high-order derivatives and introduce combinatorial notions to describe the computation procedure. As we shall see later, these small improvements turn out to be useful in the applications.

For an N-dimensional multi-index $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{N}^N$, put $|\alpha| = \sum_{i=1}^N \alpha_i$ and $\alpha! = \prod_{i=1}^N \alpha_i!$. For $\boldsymbol{t} = (t_1, \dots, t_N)^T$, write $\boldsymbol{t}^\alpha = t_1{}^{\alpha_1} \dots t_N{}^{\alpha_N}$ and denote the differential operator of index α with respect to $\boldsymbol{t}$ by $\partial^\alpha = \left(\frac{\partial}{\partial t_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial t_N}\right)^{\alpha_N}$. For $\alpha = (\alpha_i), \beta = (\beta_i) \in \mathbb{N}^N$, write $\beta \le \alpha$ if $\beta_i \le \alpha_i$ for all i. If $\beta \le \alpha$, then $\alpha - \beta = (\alpha_i - \beta_i) \in \mathbb{N}^N$.

For $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{N}^N$, define its corresponding multiset by $V_\alpha = \{(i,j) \mid 1 \le i \le N, 1 \le j \le \alpha_i\} \subset \{1, \dots, N\} \times \mathbb{N}$. For $1 \le i \le N$, write $V_\alpha(i) = \{(i,j) \mid 1 \le j \le \alpha_i\} \subset V_\alpha$. Then $\bigcup_{i=1}^N V_\alpha(i) = V_\alpha$. Note that $\#V_\alpha(i) = \alpha_i$ and $\#V_\alpha = |\alpha|$. Conversely, for a finite subset E of $\{1, \dots, N\} \times \mathbb{N}$, define $\iota(E) = (\nu_1, \dots, \nu_N) \in \mathbb{N}^N$ by $\nu_i = \#(\pi^{-1}(i) \cap E) \in \mathbb{N}$. Here π denotes the projection from $\{1, \dots, N\} \times \mathbb{N}$ onto $\{1, \dots, N\}$, namely $\pi(i,j) = i$ for $(i,j) \in \{1, \dots, N\} \times \mathbb{N}$. Then we have $\iota(V_\alpha) = \alpha$ and $\iota(W) \le \alpha$ for any subset W of V_α.

In general, given a finite set E and an integer $k \ge 0$, we call a k-pairing of E a set of k (unordered) pairs formed by $2k$ distinct elements of E. We denote the set of k-pairings of V_α by $\mathcal{P}_k(\alpha)$, namely

$$\mathcal{P}_k(\alpha) = \left\{ \sigma = \{W_i\}_{i=1}^k \;\middle|\; W_i\text{'s are distinct 2-element subsets of } V_\alpha \right\}.$$

Note that $\mathcal{P}_k(\alpha)$ is nonempty only if $0 \le k \le \lfloor |\alpha|/2 \rfloor$, where the symbol $\lfloor x \rfloor$ denotes the greatest integer which does not exceed $x \in \mathbb{R}$. In fact, if $0 \le k \le \lfloor |\alpha|/2 \rfloor$, then we have

$$\#\mathcal{P}_k(\alpha) = \frac{1}{k!}\prod_{j=0}^{k-1}\binom{|\alpha| - 2j}{2} = \frac{|\alpha|!}{2^k k!(|\alpha| - 2k)!}. \tag{2.4}$$

Put $\mathcal{P}(\alpha) = \bigcup_{k=0}^{\lfloor|\alpha|/2\rfloor} \mathcal{P}_k(\alpha)$, the set of all possible pairings of V_α.

Given a vector $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_N)^T \in \mathbb{R}^N$ and a k-pairing $\sigma = \{W_i\}_{i=1}^k \in \mathcal{P}_k(\alpha)$, we define the contraction of $\boldsymbol{\eta}$ by σ, denoted by $\widehat{\boldsymbol{\eta}}_\sigma$, by

$$\widehat{\boldsymbol{\eta}}_\sigma = c_\sigma \boldsymbol{\eta}^{\iota(V_\alpha \setminus \bigcup\sigma)} \in \mathbb{R}$$

where $c_\sigma = \prod_{i=1}^k c_{\pi(W_i)} = \prod_{i=1}^k c_{\pi(a_i),\pi(b_i)}, W_i = \{a_i, b_i\} \subset V_\alpha, a_i \neq b_i, \bigcup\sigma = \bigcup_{i=1}^k W_i$. Note that c_σ is well-defined because the covariance matrix $C = (c_{ij})$ is symmetric. For $\sigma \in \mathcal{P}_k(\alpha)$, write $\bar{\sigma} = \iota(\bigcup\sigma) \in \mathbb{N}^N$. Then we have $\bar{\sigma} \leq \alpha$, $|\bar{\sigma}| = 2k$ and $\iota(V_\alpha \setminus \bigcup\sigma) = \alpha - \bar{\sigma}$. We can thus write $\widehat{\boldsymbol{\eta}}_\sigma = c_\sigma \boldsymbol{\eta}^{\alpha - \bar{\sigma}}$. Note also that for $\sigma \in \mathcal{P}_k(\alpha)$, $\tau \in \mathcal{P}_l(\alpha)$ such that $\bigcup\sigma \cap \bigcup\tau = \varnothing$, we have $\sigma \cup \tau \in \mathcal{P}_{k+l}(\alpha)$ and $\widehat{\boldsymbol{\eta}}_{\sigma\cup\tau} = c_\sigma c_\tau \boldsymbol{\eta}^{\iota(V_\alpha \setminus \bigcup(\sigma\cup\tau))}$.

To illustrate the procedure of contraction, let $N \geq 3$ and $\alpha = (3,3,2,0,\ldots,0)$. Figure 1 depicts an illustrative 2-pairing of V_α consisting of $\{(1,1),(1,2)\}$ and $\{(2,2),(3,1)\}$. The contraction of $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_N)^T$ by this 2-pairing equals $c_{11}c_{23}\eta_1\eta_2{}^2\eta_3$.

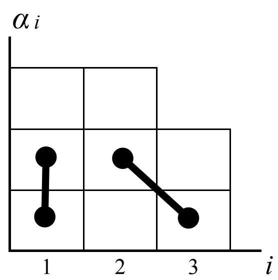

FIGURE 1. An illustrative contraction by a 2-pairing

Finally, for $\boldsymbol{\eta} \in \mathbb{R}^N$ and $\alpha \in \mathbb{N}^N$, we introduce the following notation for the sum of the contractions of $\boldsymbol{\eta}$ over all possible pairings of V_α, namely

$$R_\alpha(\boldsymbol{\eta}) = \sum_{\sigma \in \mathcal{P}(\alpha)} \widehat{\boldsymbol{\eta}}_\sigma = \sum_{\sigma \in \mathcal{P}(\alpha)} c_\sigma \boldsymbol{\eta}^{\iota(V_\alpha \setminus \bigcup\sigma)} \in \mathbb{R}.$$

The following key formula will serve as our main tool.

Lemma 1. *Let $\boldsymbol{X}$ denote an N-dimensional normally distributed random vector with mean vector $\boldsymbol{m} = (m_i) \in \mathbb{R}^N$ and covariance matrix $C = (c_{ij}) \in \mathbb{R}^{N\times N}$, and let $M_{\boldsymbol{X}}(\boldsymbol{t})$ denote its moment generating function. Then, for $\alpha = (\alpha_i) \in \mathbb{N}^N$, we have*

$$\mathbb{E}[\boldsymbol{X}^\alpha e^{\boldsymbol{t}\cdot\boldsymbol{X}}] = \partial^\alpha M_{\boldsymbol{X}}(\boldsymbol{t}) = e^{Q(\boldsymbol{t})} R_\alpha(\boldsymbol{\xi}(\boldsymbol{t})) \tag{2.5}$$

where $Q(\boldsymbol{t}) = \boldsymbol{m}^T\boldsymbol{t} + (1/2)\boldsymbol{t}^T C\boldsymbol{t}$ and $\boldsymbol{\xi}(\boldsymbol{t}) = \boldsymbol{m} + C\boldsymbol{t}$.

Proof. We need to prove the second equality in Eq.(2.5). We proceed by induction on α. First, when $\alpha = 0 = (0,\ldots,0)$, then $\mathcal{P}(\alpha)$ consists only of the 0-pairing, namely $\mathcal{P}(0) = \{\varnothing\}$. Thus, we have $R_0(\boldsymbol{\xi}(\boldsymbol{t})) = \boldsymbol{\xi}(\boldsymbol{t})^0 = 1$. Hence, Eq.(2.5) holds trivially in this case.

Next, assuming that the second equality in Eq.(2.5) is true for $\alpha = (\alpha_i) \in \mathbb{N}^N$, we show that it holds for $\widetilde{\alpha} = (\alpha_1, \ldots, \alpha_j + 1, \ldots, \alpha_N)$ for $1 \le j \le N$. Differentiating Eq.(2.5) with respect to t_j and using Eq.(2.3), we have

$$\partial^{\widetilde{\alpha}} M_{\boldsymbol{X}}(\boldsymbol{t}) = \frac{\partial}{\partial t_j}\left(e^{Q(\boldsymbol{t})} R_\alpha(\boldsymbol{\xi}(\boldsymbol{t}))\right) = e^{Q(\boldsymbol{t})}\left(\xi_j(\boldsymbol{t}) R_\alpha(\boldsymbol{\xi}(\boldsymbol{t})) + \frac{\partial R_\alpha(\boldsymbol{\xi}(\boldsymbol{t}))}{\partial t_j}\right).$$

Put $S(\boldsymbol{t}) = \xi_j(\boldsymbol{t}) R_\alpha(\boldsymbol{\xi}(\boldsymbol{t})) + \frac{\partial R_\alpha(\boldsymbol{\xi}(\boldsymbol{t}))}{\partial t_j}$. Also, put $\omega = (j, \alpha_j + 1)$ so that $V_{\widetilde{\alpha}} = V_\alpha \cup \{\omega\}$. We want to show $S(\boldsymbol{t}) = R_{\widetilde{\alpha}}(\boldsymbol{\xi}(\boldsymbol{t}))$.

Concerning the first term of $S(\boldsymbol{t})$, we have

$$\xi_j(\boldsymbol{t}) R_\alpha(\boldsymbol{\xi}(\boldsymbol{t})) = \sum_{\sigma \in \mathcal{P}(\alpha)} c_\sigma \boldsymbol{\xi}(\boldsymbol{t})^{\iota(V_{\widetilde{\alpha}} \setminus \bigcup \sigma)}.$$

If we regard $\sigma \in \mathcal{P}(\alpha)$ as an element of $\mathcal{P}(\widetilde{\alpha})$, then $\xi_j(\boldsymbol{t}) R_\alpha(\boldsymbol{\xi}(\boldsymbol{t}))$ is the sum of the contractions of $\boldsymbol{\xi}(\boldsymbol{t})$ over all the pairings of $V_{\widetilde{\alpha}}$ that do not involve ω.

Next we calculate the second term of $S(\boldsymbol{t})$. For $\sigma \in \mathcal{P}_k(\alpha)$, let $\iota(V_\alpha \setminus \bigcup \sigma) = (\nu_1, \ldots, \nu_N)$. Then, by Eq.(2.2) we have

$$\frac{\partial}{\partial t_j} \boldsymbol{\xi}(\boldsymbol{t})^{\iota(V_\alpha \setminus \bigcup \sigma)} = \frac{\partial}{\partial t_j} \boldsymbol{\xi}(\boldsymbol{t})^{(\nu_1, \ldots, \nu_N)} = \sum_{i : \nu_i > 0} \nu_i c_{ij} \boldsymbol{\xi}(\boldsymbol{t})^{(\nu_1, \ldots, \nu_i - 1, \ldots, \nu_N)}.$$

Observe that the rightmost term is the sum of the contractions of $\boldsymbol{\xi}(\boldsymbol{t})^{\iota(V_\alpha \setminus \bigcup \sigma)}$ over all the 1-pairings of $V_{\widetilde{\alpha}}$ matching ω with each $s \in V_\alpha \setminus \bigcup \sigma$. Note that ω and s are distinct since $\omega \notin V_\alpha$. For $\sigma \in \mathcal{P}_k(\alpha)$ and $s \in V_\alpha \setminus \bigcup \sigma$, define $\sigma_s \in \mathcal{P}_{k+1}(\widetilde{\alpha})$ by $\sigma_s = \sigma \cup \{s, \omega\}$. Then we have

$$\frac{\partial}{\partial t_j} \widehat{\boldsymbol{\xi}(\boldsymbol{t})}_\sigma = \sum_{s \in V_\alpha \setminus \bigcup \sigma} \widehat{\boldsymbol{\xi}(\boldsymbol{t})}_{\sigma_s}.$$

Hence $\frac{\partial R_\alpha(\boldsymbol{\xi}(\boldsymbol{t}))}{\partial t_j}$ is the sum of the contractions of $\boldsymbol{\xi}(\boldsymbol{t})$ over all the pairings of $V_{\widetilde{\alpha}}$ involving ω.

The sum of the both terms of $S(\boldsymbol{t})$ exhausts the contractions of $\boldsymbol{\xi}(\boldsymbol{t})$ by all pairings of $V_{\widetilde{\alpha}}$. Hence, the proof is complete. □

3. MOMENTS OF THE MULTIVARIATE NORMAL DISTRIBUTION

As an application of the key formula (2.5), we provide a simple proof of Isserlis' theorem and derive a formula for the moments of the multivariate normal distribution (see [9]). In fact, evaluating formula (2.5) at $\boldsymbol{t} = \boldsymbol{0}$, we have

$$\mathbb{E}\left[\prod_{i=1}^N X_i^{\alpha_i}\right] = e^{Q(\boldsymbol{0})} R_\alpha(\boldsymbol{\xi}(\boldsymbol{0})) = R_\alpha(\boldsymbol{m}) = \sum_{\sigma \in \mathcal{P}(\alpha)} c_\sigma \boldsymbol{m}^{\iota(V_\alpha \setminus \bigcup \sigma)}. \tag{3.1}$$

This immediately implies Isserlis' theorem ([4]) as follows.

Proposition 1. *Let $\boldsymbol{X} = (X_1,\ldots,X_N)^T$ denote an N-dimensional normally distributed random vector with zero mean. Then we have*

$$\mathbb{E}\left[\prod_{i=1}^{N} X_i\right] = \begin{cases} \sum_{\sigma\in\mathcal{P}_{N/2}(N)} \prod_{\{i,j\}\in\sigma} \mathrm{Cov}(X_i,X_j) & \text{if } N \text{ is even,} \\ 0 & \text{if } N \text{ is odd,} \end{cases}$$

where $\mathcal{P}_{N/2}(N)$ denotes the set of $(N/2)$-pairings of $\{1,\ldots,N\}$.

Proof. In Eq.(3.1) choose $\alpha = (1,\ldots,1)$ and identify V_α with $\{1,\ldots,N\}$. When $\boldsymbol{m} = \boldsymbol{0}$, the term under summation in Eq.(3.1) is zero except when $V_\alpha \setminus \bigcup\sigma = \varnothing$, namely $\bigcup\sigma = \{1,\ldots,N\}$. This condition holds only when N is even; and, in this case we have

$$\mathbb{E}\left[\prod_{i=1}^{N} X_i\right] = \sum_{\sigma\in\mathcal{P}_{N/2}(\alpha)} c_\sigma = \sum_{\sigma\in\mathcal{P}_{N/2}(N)} \prod_{\{i,j\}\in\sigma} c_{ij}.$$

In the last step, we identified $\mathcal{P}_{N/2}(\alpha) = \mathcal{P}_{N/2}(N)$ under the above identification $V_\alpha = \{1,\ldots,N\}$. This completes the proof. □

Next we want to express $\mathbb{E}\left[\prod_{i=1}^{N} X_i^{\alpha_i}\right]$ as a polynomial in c_{ij}'s and m_i's. From Eq.(3.1), we have

$$\mathbb{E}\left[\prod_{i=1}^{N} X_i^{\alpha_i}\right] = R_\alpha(\boldsymbol{m}) = \sum_{k=0}^{\lfloor|\alpha|/2\rfloor} \sum_{\sigma\in\mathcal{P}_k(\alpha)} c_\sigma \boldsymbol{m}^{\alpha-\bar{\sigma}}. \tag{3.2}$$

Here $\bar{\sigma} = (\bar{\sigma}_i) = \iota(\bigcup\sigma) \in \mathbb{N}^N$. Note that $|\bar{\sigma}| = 2k$ and $\bar{\sigma} \leq \alpha$ for $\sigma \in \mathcal{P}_k(\alpha)$. Let $\sigma = \{W_s\}_{s=1}^{k} \in \mathcal{P}_k(\alpha)$. Then we can write

$$c_\sigma = \prod_{s=1}^{k} c_{\pi(W_s)} = \prod_{i\leq j} c_{ij}{}^{n_{ij}} = C^{\Delta(\sigma)}$$

where $\Delta(\sigma) = (n_{ij}) \in \mathbb{N}^{N\times N}$ is determined by

$$n_{ij} = \begin{cases} \#\{s \,|\, \pi(W_s) = \{i,j\}, 1\leq s\leq k\} & \text{for } i\leq j, \\ 0 & \text{for } i > j. \end{cases}$$

Here we used the notation $A^B = \prod_{i,j} a_{ij}{}^{b_{ij}} \in \mathbb{R}$ for two real matrices $A = (a_{ij}), B = (b_{ij})$ of the same size (let $0^0 = 1$ by convention). These n_{ij}'s satisfy $\sum_{i,j=1}^{N} n_{ij} = k$ and $\sum_{i=1}^{N} n_{ih} + \sum_{j=1}^{N} n_{hj} = \bar{\sigma}_h \leq \alpha_h$ for $1 \leq h \leq N$.

Conversely, for $\alpha \in \mathbb{N}^N$ and $0 \leq k \leq \lfloor|\alpha|/2\rfloor$, consider an upper triangular, non-negative integer matrix $L = (l_{ij}) \in \mathbb{N}^{N\times N}$ satisfying the following set of conditions, which we call $\mathcal{C}(\alpha,k)$:

$$\begin{cases} l_{ij} = 0 & \text{for } i > j, \\ \sum_{i,j=1}^{N} l_{ij} = k, & \\ \sum_{i=1}^{N} l_{ih} + \sum_{j=1}^{N} l_{hj} \leq \alpha_h & \text{for } 1 \leq h \leq N. \end{cases} \tag{3.3}$$

For such L, put $\lambda(L) = (\lambda(L)_h) \in \mathbb{N}^N$ where $\lambda(L)_h = \sum_{i=1}^N l_{ih} + \sum_{j=1}^N l_{hj}$ for $1 \le h \le N$. Namely, put $\lambda(L) = (|\tilde{l}_1|, \dots, |\tilde{l}_N|)$ where $\tilde{l}_1, \dots, \tilde{l}_N \in \mathbb{N}^N$ are defined in column vector notation by $(\tilde{l}_1, \dots, \tilde{l}_N) = L + L^T$. Then the second and third conditions in (3.3) are equivalent to $|\lambda(L)| = 2k$ and $\lambda(L) \le \alpha$, respectively. Note that if $\sigma \in \mathcal{P}_k(\alpha)$ satisfies $\Delta(\sigma) = L$, then $\bar{\sigma} = \iota(\bigcup \sigma) = \lambda(L)$.

Given a matrix $L = (l_{ij}) \in \mathbb{N}^{N\times N}$ satisfying the condition $\mathcal{C}(\alpha, k)$ for some $\alpha \in \mathbb{N}^N$ and $0 \le k \le \lfloor |\alpha|/2 \rfloor$, a k-pairing $\sigma \in \mathcal{P}_k(\alpha)$ such that $\Delta(\sigma) = L$ consists of l_{ij} distinct pairs matching between the elements of $V_\alpha(i)$ and $V_\alpha(j)$ for each $i, j \in \{1, \dots, N\}$. Let $d_L(\alpha, k) \in \mathbb{N}$ denote the number of k-pairings $\sigma \in \mathcal{P}_k(\alpha)$ such that $\Delta(\sigma) = L$. Then we have

$$d_L(\alpha,k) = \frac{1}{2^{\sum_{i=1}^N l_{ii}}} \frac{1}{\prod_{i,j=1}^N l_{ij}!} \frac{\prod_{i=1}^N \alpha_i!}{\prod_{i=1}^N (\alpha_i - \lambda(L)_i)!} = \frac{1}{2^{\mathrm{Tr}L}} \frac{1}{L!} \frac{\alpha!}{(\alpha - \lambda(L))!} \tag{3.4}$$

where $\mathrm{Tr}L = \sum_{i=1}^N l_{ii}$ and $L! = \prod_{i,j=1}^N l_{ij}!$.

To illustrate the correspondence between matrices satisfying $\mathcal{C}(\alpha, k)$ and pairings of V_α, let $N = 2$, $\alpha = (3, 1)$ and $L = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. Then $\lambda(L) = (3, 1)$ and $|\lambda(L)| = 4$. Thus L satisfies the condition $\mathcal{C}(\alpha, 2)$. Hence, $d_L(\alpha, 2) = 3$. Figure 2 depicts the three 2-pairings σ of V_α such that $\Delta(\sigma) = L$.

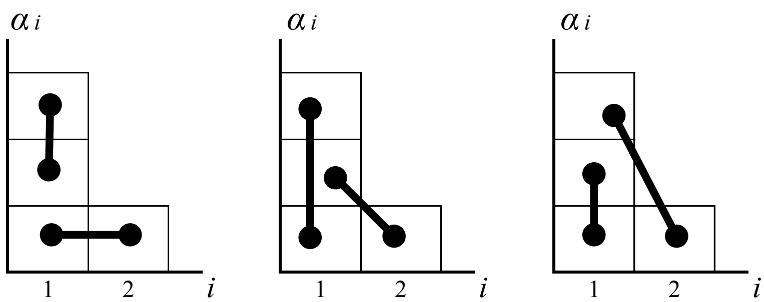

FIGURE 2. Three 2-pairings $\sigma \in \mathcal{P}_2((3,1))$ such that $\Delta(\sigma) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$

Theorem 1. *Let $\boldsymbol{X} = (X_1, \dots, X_N)^T$ denote an N-dimensional normally distributed random vector with mean vector $\boldsymbol{m} = (m_i) \in \mathbb{R}^N$ and covariance matrix $C = (c_{ij}) \in \mathbb{R}^{N\times N}$. Then, for $\alpha = (\alpha_i) \in \mathbb{N}^N$, we have*

$$\mathbb{E}\left[\prod_{i=1}^N X_i^{\alpha_i}\right] = \sum_{k=0}^{\lfloor|\alpha|/2\rfloor} \sum_{L:\, \mathcal{C}(\alpha,k)} \frac{\alpha!}{2^{\mathrm{Tr}L} L! (\alpha - \lambda(L))!} C^L \boldsymbol{m}^{\alpha - \lambda(L)}$$

where the second summation is taken over $L = (l_{ij}) \in \mathbb{N}^{N\times N}$ satisfying the condition $\mathcal{C}(\alpha, k)$ given in (3.3).

Proof. From Eq.(3.2) and the observations made in the above, we have

$$\mathbb{E}\left[\prod_{i=1}^{N} X_i^{\alpha_i}\right] = \sum_{k=0}^{\lfloor|\alpha|/2\rfloor} \sum_{\sigma\in\mathcal{P}_k(\alpha)} c_\sigma \boldsymbol{m}^{\alpha-\bar{\sigma}} = \sum_{k=0}^{\lfloor|\alpha|/2\rfloor} \sum_{L} d_L(\alpha,k) C^L \boldsymbol{m}^{\alpha-\lambda(L)}$$

where the second summation in the last term is taken over $L \in \mathbb{N}^{N\times N}$ such that $L = \Delta(\sigma)$ for some $\sigma \in \mathcal{P}_k(\alpha)$. Hence, from Eq.(3.4) we have the desired formula. This completes the proof. □

A result equivalent to Theorem 1 is stated in [7] and a proof based on mathematical induction is presented in [6]. Our combinatorial proof gives an alternative proof.

Corollary 1. *For $\alpha \in \mathbb{N}^N$ and $0 \le k \le \lfloor|\alpha|/2\rfloor$, we have*

$$\sum_{L:\,C(\alpha,k)} \frac{\alpha!}{2^{\mathrm{Tr}L} L!(\alpha-\lambda(L))!} = \frac{|\alpha|!}{2^k k!(|\alpha|-2k)!}$$

where the summation over $L \in \mathbb{N}^{N\times N}$ is taken in the same manner as in Theorem 1.

Proof. Indeed, $\#\mathcal{P}_k(\alpha)$ is the sum of $d_L(\alpha,k)$ over L. The above relation thus follows from Eq.(2.4) and Eq.(3.4). This completes the proof. □

4. The normal-lognormal product distribution and its moments

Consider a random variable expressed as the product of a finite number of correlated normally and lognormally distributed random variables. As another application of the key formula (2.5), we calculate the moments of this random variable.

4.1. *The normal-lognormal product distribution*

This subsection summarizes basic properties of the multivariate distribution composed of normal and lognormal distributions, as well as the univariate distribution defined as the product of normal and lognormal distributions. Note that these distributions were considered independently by researchers in various fields ([2],[3],[10]).

Let $N = n + l$ where $n, l \in \mathbb{N}$. We say that an N-dimensional random vector $\boldsymbol{Y} = (Y_1,\ldots,Y_N)^T$ follows joint normal-lognormal distribution of index (n,l) if $\widetilde{\boldsymbol{Y}} = (Y_1,\ldots,Y_n,\log Y_{n+1},\ldots,\log Y_N)^T$ is normally distributed. This definition is equivalent to say that $\boldsymbol{Y}$ is of the form $\boldsymbol{Y} = \left(X_1,\ldots,X_n,e^{X_{n+1}},\ldots,e^{X_N}\right)^T$ where $\boldsymbol{X} = (X_1,\ldots,X_N)^T$ is normally distributed. We say that a random variable Z follows normal-lognormal product distribution of index (n,l) if Z is of the form $Z = \prod_{i=1}^N Y_i$ where $\boldsymbol{Y} = (Y_1,\ldots,Y_N)^T$ follows joint normal-lognormal distribution of index (n,l). Equivalently, Z is of the form $Z = (\prod_{i=1}^n X_i)e^{\sum_{j=1}^l X_{n+j}}$ where $\boldsymbol{X} = (X_1,\ldots,X_N)^T$ is normally distributed. We call $\boldsymbol{X}$ the underlying normal distribution of $\boldsymbol{Y}$ and Z.

Proposition 2. *Let $\boldsymbol{Y} = (Y_1,\ldots,Y_N)^T$ denote an N-dimensional random vector following joint normal-lognormal distribution of index (n,l) with $N = n+l$, and* *let* $\boldsymbol{X} = (X_1, \ldots, X_N)^T$ *denote its underlying normal distribution with mean vector* $\boldsymbol{m} = (m_i) \in \mathbb{R}^N$ *and covariance matrix* $C = (c_{ij}) \in \mathbb{R}^{N\times N}$. *Let* $I = \{1, \ldots, n\}$ *and* $J = \{n+1, \ldots, N\}$.

(1) *The probability density function of* $\boldsymbol{Y}$ *is given by*

$$g_{n,l}(\boldsymbol{y};\boldsymbol{m},C) = (2\pi)^{-\frac{n}{2}}(\det C)^{-\frac{1}{2}}(\prod_{i\in J} y_i)^{-1}\exp\left(-\frac{1}{2}(\widetilde{\boldsymbol{y}}-\boldsymbol{m})^T C^{-1}(\widetilde{\boldsymbol{y}}-\boldsymbol{m})\right)$$

where $\boldsymbol{y} = (y_1, \ldots, y_n, y_{n+1}, \ldots, y_N)^T \in \mathbb{R}^n \times \mathbb{R}^l_+$ *and* $\widetilde{\boldsymbol{y}} = (y_1, \ldots, y_n, \log y_{n+1}, \ldots, \log y_N)^T \in \mathbb{R}^N$.

(2) *The mean of* Y_i *is given by*

$$\mathbb{E}[Y_i] = \begin{cases} m_i & \text{for } i \in I, \\ e^{m_i + \frac{\sigma_i^2}{2}} & \text{for } i \in J. \end{cases}$$

(3) *The covariance between* Y_i *and* Y_j *is given by*

$$\mathrm{Cov}(Y_i, Y_j) = \begin{cases} c_{ij} & \text{for } i, j \in I, \\ e^{m_j + \frac{\sigma_j^2}{2}} c_{ij} & \text{for } i \in I, j \in J, \\ e^{m_i + m_j + \frac{\sigma_i^2 + \sigma_j^2}{2}}(e^{c_{ij}} - 1) & \text{for } i, j \in J. \end{cases}$$

From this it follows that $\mathrm{Cov}(Y_i,Y_j) = 0$ *if and only if* $\mathrm{Cov}(X_i,X_j) = 0$ *for* $i, j = 1, \ldots, N$.

Proof.

(1) Let $f_N(x_1, \ldots, x_N; \boldsymbol{m}, C)$ be the probability density function of $\boldsymbol{X}$. By change of variables $(X_1, \ldots, X_N) = (Y_1, \ldots, Y_n, \log Y_{n+1}, \ldots, \log Y_N)$, we have

$$\begin{aligned} g_{n,l}(y_1, \ldots, y_N; \boldsymbol{m}, C) &= f_N(x_1, \ldots, x_N; \boldsymbol{m}, C)\frac{\partial(x_1, \ldots, x_N)}{\partial(y_1, \ldots, y_N)} \\ &= f_N(y_1, \ldots, y_n, \log y_{n+1}, \ldots, \log y_N; \boldsymbol{m}, C)\prod_{i=1}^{l}\frac{1}{y_{n+i}}. \end{aligned}$$

(2) For $i \in I$, we have $\mathbb{E}[Y_i] = \mathbb{E}[X_i] = m_i$. For $i \in J$, we have $\mathbb{E}[Y_i] = \mathbb{E}[e^{X_i}] = M_{X_i}(1) = e^{m_i + \frac{\sigma_i^2}{2}}$.

(3) For $i, j \in I$, we have $\mathrm{Cov}(Y_i,Y_j) = \mathrm{Cov}(X_i,X_j) = c_{ij}$.

For $i \in I, j \in J$, recall Stein's lemma ([8]. See also [3, Chapter 9, Appendix 2].) which asserts that $\mathrm{Cov}(X, \varphi(Y)) = \mathbb{E}[\varphi'(Y)]\mathrm{Cov}(X,Y)$ for a bivariate normal distribution (X,Y) and a differentiable function $\varphi(x)$ such that (i) $\lim_{x\to\pm\infty}\varphi(x)e^{-ax^2} = 0$ for any $a > 0$, and (ii) $\mathbb{E}[\varphi'(Y)]$ exists. Noting that (X_i,X_j) is normally distributed, we have $\mathrm{Cov}(Y_i,Y_j) = \mathrm{Cov}(X_i, e^{X_j}) = \mathbb{E}[e^{X_j}]\mathrm{Cov}(X_i,X_j) = e^{m_j + \frac{\sigma_j^2}{2}} c_{ij}$. An alternative proof is given in Example 2 (i) below.

For $i,j\in J$, we have

$$\mathrm{Cov}(Y_i,Y_j)=\mathrm{Cov}(e^{X_i},e^{X_j})=\mathbb{E}[e^{X_i+X_j}]-\mathbb{E}[e^{X_i}]\mathbb{E}[e^{X_j}].$$

Without loss of generality, we may fix $i=1, j=2$. Thus, we have

$$\mathbb{E}[e^{t_1X_1+t_2X_2}]=\exp\left(m_1t_1+m_2t_2+\frac{1}{2}(c_{11}{t_1}^2+2c_{12}t_1t_2+c_{22}{t_2}^2)\right).$$

Evaluating this function at $(t_1,t_2)=(1,1)$, we have

$$\mathbb{E}[e^{X_1+X_2}]=e^{m_1+m_2+\frac{1}{2}({\sigma_1}^2+2c_{12}+{\sigma_2}^2)}.$$

Thus

$$\begin{aligned}\mathrm{Cov}(Y_1,Y_2)&=e^{m_1+m_2+\frac{1}{2}({\sigma_1}^2+2c_{12}+{\sigma_2}^2)}-e^{m_1+\frac{{\sigma_1}^2}{2}}e^{m_2+\frac{{\sigma_2}^2}{2}}\\&=e^{m_1+m_2+\frac{{\sigma_1}^2+{\sigma_2}^2}{2}}(e^{c_{12}}-1).\end{aligned}$$

□

When $n=1$, we have the following integral representation of the probability density function of the normal-lognormal product distribution of index $(1,l)$.

Proposition 3. *Let Z denote a random variable following normal-lognormal product distribution of index $(1,l)$ derived by the joint normal-lognormal distribution $\boldsymbol{Y}$ of the same index. Let $\boldsymbol{m}=(m_i)\in\mathbb{R}^{l+1}$ and $C=(c_{ij})\in\mathbb{R}^{(l+1)\times(l+1)}$ denote, respectively, the mean vector and the covariance matrix of the underlying normal distribution. Then, the probability density function of Z is given by*

$$p_{1,l}(z)=\int_{\mathbb{R}^l}g_{1,l}(ze^{-w_1-w_2-\cdots-w_l},e^{w_1},\ldots,e^{w_l};\boldsymbol{m},C)dw_1\ldots dw_l$$

where $g_{1,l}(y_1,\ldots,y_{l+1};\boldsymbol{m},C)$ is the probability density function of $\boldsymbol{Y}$ given in Proposition 2 (1).

Proof. Since $Z=Y_1Y_2\ldots Y_{l+1}$ and $Y_2\ldots Y_{l+1}>0$, we have

$$p_{1,l}(z)=\int_{\mathbb{R}_+^l}g_{1,l}\left(\frac{z}{y_2\ldots y_{l+1}},y_2,\ldots,y_{l+1};\boldsymbol{m},C\right)\frac{dy_2}{y_2}\cdots\frac{dy_{l+1}}{y_{l+1}}.$$

By putting $w_i=\log y_{i+1} (1\le i\le l)$, we have the desired expression. □

4.2. *Moments of the normal-lognormal product distribution*

Theorem 2. *Let Z denote a random variable following normal-lognormal product distribution of index (n,l) with $N=n+l$, and let $\boldsymbol{X}=(X_1,\ldots,X_N)^T$ denote its underlying normal distribution with mean vector $\boldsymbol{m}=(m_i)\in\mathbb{R}^N$ and covariance matrix $C=(c_{ij})\in\mathbb{R}^{N\times N}$. Then the d-th raw moment of Z is given by*

$$\mathbb{E}[Z^d]=\mathbb{E}\left[\left(X_1\cdots X_ne^{X_{n+1}+\cdots+X_N}\right)^d\right]=e^{Q(\boldsymbol{d}_{0,l})}R_{d_{n,0}}(\boldsymbol{\zeta})\qquad(4.1)$$

where $d_{n,0} = (\overbrace{d,\dots,d}^{n},0,\dots,0)$, $\boldsymbol{d}_{0,l} = (0,\dots,0,\overbrace{d,\dots,d}^{l})^T$ *and* $\boldsymbol{\zeta} = (\zeta_i) = \boldsymbol{\xi}(\boldsymbol{d}_{0,l})$. *Explicitly,* $Q(\boldsymbol{d}_{0,l}) = d\sum_{j=1}^{l} m_{n+j} + \frac{d^2}{2}\sum_{i,j=1}^{l} c_{n+i,n+j}$ *and* $\zeta_i = m_i + d\sum_{j=1}^{l} c_{i,n+j}$ $(1 \le i \le N)$.

Proof. For $\alpha = (\alpha_i) \in \mathbb{N}^N$ where $N = n+l$, put $\alpha_{n,0} = (\alpha_1,\dots,\alpha_n,0,\dots,0)$ and $\boldsymbol{\alpha}_{0,l} = (0,\dots,0,\alpha_{n+1},\dots,\alpha_N)^T$. Substituting $\alpha = \alpha_{n,0}$ in the key formula (2.5) and evaluating it at $\boldsymbol{t} = \boldsymbol{\alpha}_{0,l}$, we have

$$\mathbb{E}\left[{X_1}^{\alpha_1}\cdots{X_n}^{\alpha_n}e^{\alpha_{n+1}X_{n+1}+\cdots+\alpha_N X_N}\right] = \partial^{\alpha_{n,0}}M_{\boldsymbol{X}}(\boldsymbol{\alpha}_{0,l}) = e^{Q(\boldsymbol{\alpha}_{0,l})}R_{\alpha_{n,0}}(\boldsymbol{\xi}(\boldsymbol{\alpha}_{0,l})). \quad (4.2)$$

Setting $\alpha_1 = \cdots = \alpha_N = d \in \mathbb{N}$ in Eq.(4.2), we have Eq.(4.1). This completes the proof. □

When $d = 1$, we have the mean of the normal-longnormal product distribution of index (n,l).

Corollary 2. *With the notation as above, when* $d = 1$*, we have*

$$\mathbb{E}[Z] = \mathbb{E}\left[X_1\cdots X_n e^{X_{n+1}+\cdots+X_N}\right] = e^{Q_0}\sum_{k=0}^{\lfloor n/2\rfloor}\sum_{\sigma\in\mathcal{P}_k(n)}\left(\prod_{\{i,j\}\in\sigma}c_{ij}\prod_{i\in\{1,\dots,n\}\setminus\bigcup\sigma}\zeta_i\right)$$

where $Q_0 = \sum_{j=1}^{l} m_{n+j} + \frac{1}{2}\sum_{i,j=1}^{l} c_{n+i,n+j}$, $\zeta_i = m_i + \sum_{j=1}^{l} c_{i,n+j}$ $(1 \le i \le n)$ *and* $\mathcal{P}_k(n)$ *denotes the set of k-pairings of* $\{1,\dots,n\}$.

Proof. When $d = 1$, $Q(\boldsymbol{d}_{0,l}) = Q_0$. For notational simplicity put $\nu = 1_{n,0} = (\overbrace{1,\dots,1}^{n},$ $0,\dots,0)$. Then we have

$$R_\nu(\boldsymbol{\zeta}) = \sum_{\sigma\in\mathcal{P}(\nu)} c_\sigma \boldsymbol{\zeta}^{\nu-\bar{\sigma}} = \sum_{k=0}^{\lfloor n/2\rfloor}\sum_{\sigma\in\mathcal{P}_k(\nu)} c_\sigma \boldsymbol{\zeta}^{\nu-\bar{\sigma}}$$
$$= \sum_{k=0}^{\lfloor n/2\rfloor}\sum_{\sigma\in\mathcal{P}_k(n)}\left(\prod_{\{i,j\}\in\sigma}c_{ij}\prod_{i\in\{1,\dots,n\}\setminus\bigcup\sigma}\zeta_i\right).$$

In the last step, we identified $\mathcal{P}_k(\nu)$ with $\mathcal{P}_k(n)$ by identifying $V_\nu = \{1,\dots,n\}$. Thus, from Eq.(4.1) the corollary follows. □

When $n = 1$, we have the following explicit formula for arbitrary moments of the normal-longnormal product distribution of index $(1,l)$.

Corollary 3. *With the notation as above, when* $n = 1$*, we have*

$$\mathbb{E}[Z^d] = \mathbb{E}\left[\left(X_1 e^{X_2+\cdots+X_{l+1}}\right)^d\right] = e^{Q_1}\sum_{k=0}^{\lfloor d/2\rfloor}\frac{d!}{2^k k!(d-2k)!}{\sigma_1}^{2k}{\zeta_1}^{d-2k}$$

where ${\sigma_1}^2 = c_{11}$, $Q_1 = d\sum_{j=1}^{l} m_{j+1} + \frac{d^2}{2}\sum_{i,j=1}^{l} c_{i+1,j+1}$ *and* $\zeta_1 = m_1 + d\sum_{j=1}^{l} c_{1,j+1}$.

Proof. When $n=1$, $Q(\boldsymbol{d}_{0,l})=Q_1$. By definition, for $d_{1,0}=(d,0,\dots,0)\in\mathbb{N}^{l+1}$ and $\boldsymbol{\zeta}=(\zeta_i)\in\mathbb{R}^{l+1}$, we have

$$R_{d_{1,0}}(\boldsymbol{\zeta})=\sum_{k=0}^{\lfloor d/2\rfloor}\sum_{\sigma\in\mathcal{P}_k(d_{1,0})}c_\sigma\boldsymbol{\zeta}^{d_{1,0}-\bar{\sigma}}=\sum_{k=0}^{\lfloor d/2\rfloor}\#\mathcal{P}_k(d_{1,0}){c_{11}}^k{\zeta_1}^{d-2k}$$
$$=\sum_{k=0}^{\lfloor d/2\rfloor}\frac{d!}{2^k k!(d-2k)!}{\sigma_1}^{2k}{\zeta_1}^{d-2k}.$$

Here we used Eq.(2.4) in the last step. Therefore, from Eq.(4.1) the corollary follows. □

By applying the idea of the proof of Theorem 2 to the results of Example 1, we can directly calculate the first few moments of the normal-lognormal product distribution for $n,l\in\{1,2\}$ and $d\le 8/(n+l)$.

Example 2. For subsequent use, we evaluate $\xi_k(\boldsymbol{t})$ and $Q(\boldsymbol{t})$ at $\boldsymbol{t}=a\boldsymbol{e}_i+b\boldsymbol{e}_j$ (namely, $t_i=a, t_j=b$, and $t_h=0$ for $h\ne i,j$):

$$\xi_k(a\boldsymbol{e}_i+b\boldsymbol{e}_j)=m_k+ac_{ki}+bc_{kj},$$
$$Q(a\boldsymbol{e}_i+b\boldsymbol{e}_j)=am_i+bm_j+\frac{1}{2}(a^2{\sigma_i}^2+2abc_{ij}+b^2{\sigma_j}^2).$$

(i) Index $(1,1)$ case: Evaluating $\mathbb{E}[X_1e^{\boldsymbol{t}\cdot\boldsymbol{X}}]$ at $\boldsymbol{t}=\boldsymbol{e}_2$, we have

$$\mathbb{E}[Z]=\mathbb{E}[X_1e^{X_2}]=e^{m_2+\frac{{\sigma_2}^2}{2}}(m_1+c_{12}).$$

Evaluating $\mathbb{E}[{X_1}^2e^{\boldsymbol{t}\cdot\boldsymbol{X}}]$ at $\boldsymbol{t}=2\boldsymbol{e}_2$, we have

$$\mathbb{E}[Z^2]=\mathbb{E}[{X_1}^2e^{2X_2}]=e^{2m_2+2{\sigma_2}^2}((m_1+2c_{12})^2+{\sigma_1}^2).$$

Hence, we have

$$\mathrm{Var}[Z]=\mathbb{E}[Z^2]-\mathbb{E}[Z]^2=e^{2m_2+{\sigma_2}^2}\left[e^{{\sigma_2}^2}((m_1+2c_{12})^2+{\sigma_1}^2)-(m_1+c_{12})^2\right].$$

Also, from the above results, we have

$$\mathrm{Cov}(X_1,e^{X_2})=\mathbb{E}[X_1e^{X_2}]-\mathbb{E}[X_1]\mathbb{E}[e^{X_2}]=e^{m_2+\frac{{\sigma_2}^2}{2}}c_{12}.$$

This gives an alternative proof of the second case of Proposition 2 (3).
Evaluating $\mathbb{E}[{X_1}^3e^{\boldsymbol{t}\cdot\boldsymbol{X}}]$ at $\boldsymbol{t}=3\boldsymbol{e}_2$, and $\mathbb{E}[{X_1}^4e^{\boldsymbol{t}\cdot\boldsymbol{X}}]$ at $\boldsymbol{t}=4\boldsymbol{e}_2$, we have

$$\mathbb{E}[Z^3]=\mathbb{E}[{X_1}^3e^{3X_2}]=e^{3m_2+(9/2){\sigma_2}^2}((m_1+3c_{12})^3+3{\sigma_1}^2(m_1+3c_{12})),$$
$$\mathbb{E}[Z^4]=\mathbb{E}[{X_1}^4e^{4X_2}]=e^{4m_2+4{\sigma_2}^2}((m_1+4c_{12})^4+6{\sigma_1}^2(m_1+4c_{12})^2+3{\sigma_1}^4).$$

(ii) Index $(1,2)$ case: Evaluating $\mathbb{E}[X_1e^{\boldsymbol{t}\cdot\boldsymbol{X}}]$ at $\boldsymbol{t}=\boldsymbol{e}_2+\boldsymbol{e}_3$, we have

$$\mathbb{E}[Z]=\mathbb{E}[X_1e^{X_2+X_3}]=e^{m_2+m_3+\frac{1}{2}({\sigma_2}^2+2c_{23}+{\sigma_3}^2)}(m_1+c_{12}+c_{13}).$$

Evaluating $\mathbb{E}[{X_1}^2 e^{\boldsymbol{t}\cdot\boldsymbol{X}}]$ at $\boldsymbol{t} = 2\boldsymbol{e}_2 + 2\boldsymbol{e}_3$, we have

$$\mathbb{E}[Z^2] = \mathbb{E}[{X_1}^2 e^{2X_2+2X_3}] = e^{2(m_2+m_3)+2({\sigma_2}^2+2c_{23}+{\sigma_3}^2)}((m_1+2c_{12}+2c_{13})^2+{\sigma_1}^2).$$

(iii) Index $(2,1)$ case: Evaluating $\mathbb{E}[X_1X_2 e^{\boldsymbol{t}\cdot\boldsymbol{X}}]$ at $\boldsymbol{t} = \boldsymbol{e}_3$, we have

$$\mathbb{E}[Z] = \mathbb{E}[X_1X_2e^{X_3}] = e^{m_3+\frac{{\sigma_3}^2}{2}}((m_1+c_{13})(m_2+c_{23})+c_{12}).$$

Evaluating $\mathbb{E}[{X_1}^2{X_2}^2 e^{\boldsymbol{t}\cdot\boldsymbol{X}}]$ at $\boldsymbol{t} = 2\boldsymbol{e}_3$, we have

$$\begin{aligned}\mathbb{E}[Z^2] &= \mathbb{E}[{X_1}^2{X_2}^2e^{2X_3}]\\ &= e^{2m_3+2{\sigma_3}^2}\big[(m_1+2c_{13})^2(m_2+2c_{23})^2+{\sigma_2}^2(m_1+2c_{13})^2\\ &\quad+{\sigma_1}^2(m_2+2c_{23})^2+4c_{12}(m_1+2c_{13})(m_2+2c_{23})+2{c_{12}}^2+{\sigma_1}^2{\sigma_2}^2\big].\end{aligned}$$

(iv) Index $(2,2)$ case: Evaluating $\mathbb{E}[X_1X_2 e^{\boldsymbol{t}\cdot\boldsymbol{X}}]$ at $\boldsymbol{t} = \boldsymbol{e}_3+\boldsymbol{e}_4$, we have

$$\begin{aligned}\mathbb{E}[Z] &= \mathbb{E}[X_1X_2e^{X_3+X_4}]\\ &= e^{m_3+m_4+\frac{1}{2}({\sigma_3}^2+2c_{34}+{\sigma_4}^2)}((m_1+c_{13}+c_{14})(m_2+c_{23}+c_{24})+c_{12}).\end{aligned}$$

Evaluating $\mathbb{E}[{X_1}^2{X_2}^2 e^{\boldsymbol{t}\cdot\boldsymbol{X}}]$ at $\boldsymbol{t} = 2\boldsymbol{e}_3+2\boldsymbol{e}_4$, we have

$$\begin{aligned}\mathbb{E}[Z^2] &= \mathbb{E}[{X_1}^2{X_2}^2e^{2X_3+2X_4}]\\ &= e^{2(m_3+m_4)+2({\sigma_3}^2+2c_{34}+{\sigma_4}^2)}\big[(m_1+2c_{13}+2c_{14})^2(m_2+2c_{23}+2c_{24})^2\\ &\quad+{\sigma_2}^2(m_1+2c_{13}+2c_{14})^2+{\sigma_1}^2(m_2+2c_{23}+2c_{24})^2\\ &\quad+4c_{12}(m_1+2c_{13}+2c_{14})(m_2+2c_{23}+2c_{24})+2{c_{12}}^2+{\sigma_1}^2{\sigma_2}^2\big].\end{aligned}$$

ACKNOWLEDGEMENTS

I would like to acknowledge Mr. Subramaniam Iyer for calling my attention to what we call normal-lognormal product distributions in this note.

REFERENCES

[1] T. Alberts and D. Khoshnevisan, *Calculus on Gauss Space: An Introduction to Gaussian Analysis.*, 2018, book in preparation available at http://www.math.utah.edu/ davar/math7880/F18/GaussianAnalysis.pdf.

[2] S. J. Fletcher and M. Zupanski, "A hybrid multivariate normal and lognormal distribution for data assimilation." *Atmos. Sci. Let.*, vol. 7, pp. 43–46, 2006, doi: 10.1002/asl.128.

[3] K. Hirose, *Topics in quantitative analysis of social protection systems.*, 1999, Issues in Social Protection, Discussion Paper No. 6. International Labour Office, Social Security Department, Geneva.

[4] L. Isserlis, "On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables." *Biometrika*, vol. 12, no. 1-2, pp. 134–139, 1918, doi: 10.1093/biomet/12.1-2.134.

[5] S. Kotz, N. Balakrishnan, and N. L. Johnson, *Continuous Multivariate Distributions: Volume 1: Models and Applications. 2nd Edition.* New York: John Wiley and Sons, Inc., 2000. doi: 10.1002/0471722065.

[6] I. Song, “A proof of the explicit foumula for product moments of multivariate gaussian random variables.” 2017, preprint, available at arXiv:1705.00163 [math.PR].
[7] I. Song and S. Lee, “Explicit foumulae for product moments of multivariate gaussian random variables.” *Stat. Probab. Lett.*, vol. 100, pp. 27–34, 2015, doi: 10.1016/j.spl.2015.01.030.
[8] C. Stein, “Estimation of the mean of a multivariate normal distribution.” *Ann. Statist.*, vol. 9, no. 6, pp. 1135–1151, 1981, doi: 10.1214/aos/1176345632.
[9] C. S. Withers, “The moments of the multivariate normal.” *Bull. Aust. Math. Soc.*, vol. 32, no. 1, pp. 103–107, 1985, doi: 10.1017/S000497270000976X.
[10] M. Yang, “Normal log-normal mixture, leptokurtosis and skewness.” *Appl. Econ, Lett.*, vol. 15, no. 9, pp. 737–742, 2008, doi: 10.1080/13504850600749073.

Author’s address

Kenichi Hirose
10-17 Moto-machi, Ono City, Fukui 912-0081, Japan
E-mail address: kenichihirose@gmail.com