From Moments of Sum to Moments of Product

Raymond Kan *

Joseph L. Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario, Canada M5S 3E6

Abstract

We provide an identity that relates the moment of a product of random variables to the moments of different linear combinations of the random variables. Applying this identity, we obtain new formulae for the expectation of the product of normally distributed random variables and the product of quadratic forms in normally distributed random variables. In addition, we generalize the formulae to the case of multivariate elliptically distributed random variables. Unlike existing formulae in the literature, our new formulae are extremely efficient for computational purposes.

MSC 2000 subject classifications: 60E10; 62H10

Key words: Moments of product, Product of random variables, Product of quadratic forms

1 Introduction

Let $z = [z_1, \ldots, z_n]' \sim N(\mu, \Sigma)$ be a normal random vector, where $\Sigma = (\sigma_{ij})$ is a positive semidefinite matrix. For nonnegative integers s_i, we are interested in obtaining analytical and computationally efficient expressions for (1) the expectation of a product of the elements of z,

$$\mu_{s_1,s_2,\ldots,s_n} = E[z_1^{s_1} z_2^{s_2} \cdots z_n^{s_n}],$$

* Corresponding author.
Email address: kan@chass.utoronto.ca (Raymond Kan).

and (2) the expectation of a product of quadratic forms in z,

$$Q_{s_1,s_2,\ldots,s_p} = E[(z'A_1z)^{s_1}(z'A_2z)^{s_2}\cdots(z'A_pz)^{s_p}],$$

where A_1 to A_p are symmetric matrices.

Explicit expressions of these expectations have long been available in the statistics literature. For the expectation of a product of central normal random variables, its formula is available since Isserlis [8]. In physics literature, Isserlis's formula is often written as the hafnian of Σ and it is known as the Wick's formula. However, for even $s = s_1 + \cdots + s_n$, this formula requires summing up $(s-1)!! = 1\times 3\times\cdots\times(s-1)$ terms of product of $s/2$ elements of the Σ matrix. Even for moderately large s, the number of calculations is astronomical. For example, if one wishes to calculate $E[(z_1z_2z_3z_4z_5)^4]$, then one would need to sum up $19!! = 654{,}729{,}075$ number of terms to obtain the answer, which is clearly impractical. Holmquist [6] provides an elegant formula for all the higher order product moments of z. However, his formula requires one to construct a large symmetric matrix of dimension $n^s \times n^s$. For $n = 5$ and $s = 20$ as in the previous example, one needs to construct a matrix with 9.0949×10^{27} elements, which is infeasible even with the use of sparse matrix. Schott [19] also provides a compact formula for the moment matrices of the normal distribution. Similar to Holmquist's formula, Schott's formula requires the construction of a permutation matrix of size $n^{\frac{s}{2}} \times n^{\frac{s}{2}}$, which is also impractical for computational purpose. Recently, Blacher [3] provides another formula for $\mu_{s_1,s_2,\ldots,s_n}$. His formula requires one only to sum up $([s_1/2]+1)\times\cdots\times([s_n/2]+1)$ terms, but within each of the term, one needs to enumerate all the possible symmetric $n \times n$ matrices with zero diagonal elements that satisfies n constraints. Unless n is very small, it is very time consuming to enumerate all these symmetric matrices. For the case that $s_1 = \cdots = s_n = 1$, Blacher's formula is in fact the same as Isserlis's formula, so again it requires summing up the same large number of terms.

The situation for the product of quadratic forms is even worse. Most of the existing work express $Q_{s_1,\ldots,s_p}$ as a sum of various products of the traces of $s = s_1 + \ldots + s_p$ matrices related to A_i's and Σ (see Kumar [11], Magnus [12,13], Don [4], Tracy and Jindasa [21], Jindasa and Tracy [9], Tracy and Sultan [22] for the development of this type of formulae, see also Mathai and Provost [16] for an excellent review of quadratic forms in random variables). While explicit expressions of $Q_{s_1,\ldots,s_p}$ are available when s is small ($s \leq 4$), current methods are impractical for computing $Q_{s_1,\ldots,s_p}$ even for moderately large s. This is because the number of terms grows exponentially as s increases. For $s = 4$, we only have 17 terms but even for $s = 12$, there are 171,453,343 terms (see Don [4] and Magnus [12] for a method of counting the number of terms). For $s = 20$, there are 6.6337×10^{17} terms, so computing something like $E[(z'A_1z)^4(z'A_2z)^4(z'A_3z)^4(z'A_4z)^4(z'A_5z)^4]$ is simply impossible. Holmquist [7] provides a compact expression of $Q_{s_1,\ldots,s_p}$ but again his formula requires the construction of an $n^s \times n^s$ symmetric matrix, which is impractical even for fairly small n and s.

In this paper, we take a different approach in dealing with these problems. At the heart of our derivation is an identity that is motivated by a lemma in Magnus [12]. This identity allows us to express a product of random variables as a polynomial of various sums of the random variables. In many cases, moments of a sum of random variables are readily available but expectation of a product of random variables is relatively difficult to obtain, especially when the random variables are not independent of each other. With the help of this identity, we present new analytical expressions for $\mu_{s_1,\ldots,s_n}$ and $Q_{s_1,\ldots,s_p}$ that are computationally far more efficient than existing formulae.

The rest of the paper is organized as follows. Section 2 presents the key identity that allows us to compute the expectation of a product of random variables by using the moments of various sums of the random variables. Section 3 applies our identity to obtain an explicit formula for the expectation of the product of normal random variables. It also generalizes the results to the case of multivariate elliptical distribution. Section 4 presents an expression for the product of quadratic forms in multivariate elliptical random variables. Section 5 concludes the paper.

2 An Identity

The identity in this section is motivated by Lemma 4.1 of Magnus [12]. His lemma expresses the product of n real numbers as a polynomial of different linear combinations of these n numbers. Our identity serves the same purpose as Magnus's lemma but with fewer terms.

Let a, b, and c be real variables. After observing the following relations,

$$4ab = (a+b)^2 - (a-b)^2$$

and

$$24abc = (a+b+c)^3 - (a+b-c)^3 - (a-b+c)^3 + (a-b-c)^3,$$

it is natural to conjecture that a product of n real numbers $x_1 x_2 \cdots x_n$ can be expressed as the sum of 2^{n-1} terms, with each term taking the form of $(x_1 \pm x_2 \pm x_3 \cdots \pm x_n)^n$. The following lemma presents a more general version of this identity by allowing the variables in the product to be raised to a different power.

Lemma 1. *Let x_1 to x_n be real numbers, s_1 to s_n be nonnegative integers, and $s = s_1 + s_2 + \cdots + s_n$. Then,*

$$x_1^{s_1} x_2^{s_2} \cdots x_n^{s_n} = \frac{1}{s!} \sum_{\nu_1=0}^{s_1} \cdots \sum_{\nu_n=0}^{s_n} (-1)^{\sum_{i=1}^n \nu_i} \binom{s_1}{\nu_1} \cdots \binom{s_n}{\nu_n} \left(\sum_{i=1}^n h_i x_i \right)^s, \tag{1}$$

where $h_i = s_i/2 - \nu_i$.

The proof of Lemma 1 is given in Appendix A. For actual calculations, half of the terms on the right hand side of (1) are repeated, so we only need to sum up half of the terms and multiply the result by two. To illustrate this, we first assume at least one of the s_i is odd. Without loss of generality, we assume s_1 is odd. Then, it is easy to verify that for $\nu_1 = 0$ to $(s_1-1)/2$, we have

$$\begin{aligned}
&\sum_{\nu_2=0}^{s_2}\cdots\sum_{\nu_n=0}^{s_n}(-1)^{\nu_1+\sum_{i=2}^n \nu_i}\binom{s_1}{\nu_1}\binom{s_2}{\nu_2}\cdots\binom{s_n}{\nu_n}\left[\left(\frac{s_1}{2}-\nu_1\right)x_1+\sum_{i=2}^n h_i x_i\right]^s \\
=&\sum_{\nu_2=0}^{s_2}\cdots\sum_{\nu_n=0}^{s_n}(-1)^{(s_1-\nu_1)+\sum_{i=2}^n \nu_i}\binom{s_1}{s_1-\nu_1}\binom{s_2}{\nu_2}\cdots\binom{s_n}{\nu_n}\left[\left(\nu_1-\frac{s_1}{2}\right)x_1+\sum_{i=2}^n h_i x_i\right]^s .
\end{aligned}$$

The equality is obtained by replacing ν_i with $s_i - \nu_i'$ for $i = 2,\ldots,n$ and using the fact that $\binom{s}{s-\nu} = \binom{s}{\nu}$. This implies that we can obtain $x_1^{s_1}x_2^{s_2}\cdots x_n^{s_n}$ using

$$x_1^{s_1}\cdots x_n^{s_n} = \frac{2}{s!}\sum_{\nu_1=0}^{(s_1-1)/2}\sum_{\nu_2=0}^{s_2}\cdots\sum_{\nu_n=0}^{s_n}(-1)^{\sum_{i=1}^n \nu_i}\binom{s_1}{\nu_1}\cdots\binom{s_n}{\nu_n}\left(\sum_{i=1}^n h_i x_i\right)^s,$$

which involves only $(s_1+1)(s_2+1)\cdots(s_n+1)/2$ terms. For the case when all the s_i's are even, we can repeat the above exercise for ν_2 to ν_n to obtain

$$\begin{aligned}
x_1^{s_1}\cdots x_n^{s_n} =& \frac{2}{s!}\sum_{\nu_1=0}^{\frac{s_1}{2}-1}\sum_{\nu_2=0}^{s_2}\cdots\sum_{\nu_n=0}^{s_n}(-1)^{\sum_{i=1}^n \nu_i}\binom{s_1}{\nu_1}\cdots\binom{s_n}{\nu_n}\left(\sum_{i=1}^n h_i x_i\right)^s \\
&+(-1)^{\frac{s_1}{2}}\binom{s_1}{\frac{s_1}{2}}\frac{1}{s!}\sum_{\nu_2=0}^{s_2}\cdots\sum_{\nu_n=0}^{s_n}(-1)^{\sum_{i=2}^n \nu_i}\binom{s_2}{\nu_2}\cdots\binom{s_n}{\nu_n}\left(\sum_{i=2}^n h_i x_i\right)^s \\
=& \frac{2}{s!}\sum_{\nu_1=0}^{\frac{s_1}{2}-1}\sum_{\nu_2=0}^{s_2}\cdots\sum_{\nu_n=0}^{s_n}(-1)^{\sum_{i=1}^n \nu_i}\binom{s_1}{\nu_1}\cdots\binom{s_n}{\nu_n}\left(\sum_{i=1}^n h_i x_i\right)^s \\
&+(-1)^{\frac{s_1}{2}}\binom{s_1}{\frac{s_1}{2}}\frac{2}{s!}\sum_{\nu_2=0}^{\frac{s_2}{2}-1}\sum_{\nu_3=0}^{s_3}\cdots\sum_{\nu_n=0}^{s_n}(-1)^{\sum_{i=2}^n \nu_i}\binom{s_2}{\nu_2}\cdots\binom{s_n}{\nu_n}\left(\sum_{i=2}^n h_i x_i\right)^s+\cdots \\
&+(-1)^{\frac{s_1+s_2+s_{n-1}}{2}}\binom{s_1}{\frac{s_1}{2}}\binom{s_2}{\frac{s_2}{2}}\binom{s_{n-1}}{\frac{s_{n-1}}{2}}\frac{2}{s!}\sum_{\nu_n=0}^{\frac{s_n}{2}-1}(-1)^{\nu_n}\binom{s_n}{\nu_n}(h_n x_n)^s,
\end{aligned}$$

which has only $[(s_1+1)(s_2+1)\cdots(s_n+1)-1]/2$ terms. Therefore, for both cases, we only need to sum up $[(s_1+1)(s_2+1)\cdots(s_n+1)/2]$ terms to obtain the right hand side of (1).

Suppose x_1 to x_n are random variables. Then taking expectation of both sides of (1) gives us

$$E\left[\prod_{i=1}^n x_i^{s_i}\right] = \frac{1}{s!}\sum_{\nu_1=0}^{s_1}\cdots\sum_{\nu_n=0}^{s_n}(-1)^{\sum_{i=1}^n \nu_i}\binom{s_1}{\nu_1}\cdots\binom{s_n}{\nu_n}E\left[\left(\sum_{i=1}^n h_i x_i\right)^s\right], \tag{2}$$

provided that the expectation exists. At first glance, this identity is of little use because instead of computing one expectation, we now need to compute $[(s_1+1)(s_2+1)\cdots(s_n+1)/2]$ expectations, with each term being the s-th moment of a linear combination of x_i's. However, it is often the case that direct evaluation of the expectation of a product of random variables is difficult but yet simple analytical expression for the s-th moment of the linear combination of the random variables is readily available. In such cases, (2) allows us to provide a significant simplification for the computation of the moment of the product.[1] We illustrate the use of this identity with different applications in the rest of the paper.

3 Expectation of a Product of Normal Random Variables

We focus on the central normal case first by assuming $\mu = 0$. The noncentral normal case will be dealt with later. For $n = 1$ or 2, all moments of the normal distribution can be easily computed. For $n = 1$, it is well known that

$$\mu_s = \begin{cases} 0 & \text{if } s \text{ is odd,} \\ (s-1)!!\sigma^s & \text{if } s \text{ is even.} \end{cases}$$

For the case that $n = 2$ (i.e., z is bivariate normal), a simple closed-form solution of μ_{s_1,s_2} is available from Isserlis [8] and Kendall and Stuart [10] (p.94), which is given by

$$\mu_{s_1,s_2} = \begin{cases} 0 & \text{if } s_1+s_2 \text{ is odd,} \\ \sigma_1^{s_1}\sigma_2^{s_2}\dfrac{s_1!s_2!}{2^{s/2}}\displaystyle\sum_{j=0}^{t}\frac{(2\rho)^{2j}}{\left(\frac{s_1}{2}-j\right)!\left(\frac{s_2}{2}-j\right)!(2j)!} & \text{if } s_1 \text{ and } s_2 \text{ are even,} \\ \sigma_1^{s_1}\sigma_2^{s_2}\dfrac{s_1!s_2!}{2^{s/2}}\displaystyle\sum_{j=0}^{t}\frac{(2\rho)^{2j+1}}{\left(\frac{s_1-1}{2}-j\right)!\left(\frac{s_2-1}{2}-j\right)!(2j+1)!} & \text{if } s_1 \text{ and } s_2 \text{ are odd,} \end{cases}$$

where ρ is the correlation coefficient between z_1 and z_2 and $t = [\min[s_1, s_2]/2]$.

For $n > 2$, the most popular expression for $\mu_{s_1,\ldots,s_n}$ is due to Isserlis [8]. In the physics literature, this formula is known as the Wick's formula. In order to use this formula to compute $\mu_{s_1,\ldots,s_n}$, we let $z^* = [z_1 1'_{s_1},\ z_2 1'_{s_2}\ldots, z_n 1'_{s_n}]'$ be an expanded s-vector formed from the original normal random vector z and denote Σ^* as the covariance matrix of z^*. Then the Wick's formula suggests that

$$\mu_{s_1,\ldots,s_n} = \begin{cases} 0 & \text{if } s \text{ is odd,} \\ \mathrm{Haf}(\Sigma^*) & \text{if } s \text{ is even,} \end{cases}$$

[1] Magnus appears to be unaware of the power of his lemma. Instead of computing the s-th moment of the sum, he proceeds to expand it into many terms which leads to great computational complexity.

where $\mathrm{Haf}(\Sigma^*)$ is the hafnian of $\Sigma^* = (\sigma^*_{ij})$, which is defined as

$$\mathrm{Haf}(\Sigma^*) = \sum_{p\in\Pi_s} \prod_{i=1}^{\frac{s}{2}} \sigma^*_{p_{2i-1},p_{2i}},$$

and Π_s is the set of all permutations p of $\{1, 2, \ldots, s\}$ satisfying the property $p_1 < p_3 < p_5 < \cdots < p_{s-1}$ and $p_1 < p_2$, $p_3 < p_4$, $\ldots$, $p_{s-1} < p_s$. This formula is quite simple and it is also very easy to program. All it requires is to enumerate all the p's in Π_s and compute the product of $s/2$ elements of Σ^* and then sum them up. However, the set Π_s has $(s-1)!!$ elements in it, so the number of elements in Π_s quickly goes up as s increases, making this formula impractical to use even when s is moderately large.

When some of the s_i's are not equal to one, we can speed up the computation of $\mu_{s_1,\ldots,s_n}$ without actually having to enumerate all the $(s-1)!!$ terms. For the example $E[(z_1z_2z_3z_4z_5)^4]$ that we give in the introduction, we only have $n = 5$ and a large number of the $(s-1)!!$ terms are repeated. In order to cut down the computation time, we can use the following recursive relation

$$\mu_{s_1,\ldots,s_n} = (s_1 - 1)\sigma_{11}\mu_{s_1-2,\ldots,s_n} + \sum_{i=2}^{n} s_i\sigma_{1i}\mu_{s_1-1,s_2,\ldots,s_i-1,\ldots,s_n},$$

with the first term vanishes when $s_1 = 1$.[2] Then together with the closed-form solutions for $n = 1$ and $n = 2$ as the boundary conditions, we can numerically solve for $\mu_{s_1,\ldots,s_n}$. However, for $s_1 = s_2 = \cdots = s_n = 1$, there does not appear to be a faster way other than to enumerate all the elements of Π_s.

Using our identity in Section 2, the following Proposition presents a new formula for $\mu_{s_1,\ldots,s_n}$.

Proposition 1 *Suppose $z = [z_1, \ldots, z_n]' \sim N(0, \Sigma)$, where Σ is an $n\times n$ positive semidefinite matrix. For nonnegative integers s_1 to s_n, we have*

$$E\left[\prod_{i=1}^{n} z_i^{s_i}\right] = \begin{cases} 0 & \text{if } s \text{ is odd,} \\ \frac{1}{(\frac{s}{2})!}\sum_{\nu_1=0}^{s_1}\cdots\sum_{\nu_n=0}^{s_n}(-1)^{\sum_{i=1}^n \nu_i}\binom{s_1}{\nu_1}\cdots\binom{s_n}{\nu_n}\left(\frac{h'\Sigma h}{2}\right)^{\frac{s}{2}} & \text{if } s \text{ is even,}\end{cases}$$

where $s = s_1 + \cdots + s_n$ and $h = \left[\frac{s_1}{2} - \nu_1, \ldots, \frac{s_n}{2} - \nu_n\right]'$.

[2] This recursive relation can be easily proven by using an induction method as in Triantafyllopoulos [23]. Our experience suggests that rearranging z_i such that s_i is in ascending order also significantly reduces the computation time.

Proof: Putting $x_i = z_i$ in (2), we have

$$E\left[\prod_{i=1}^n z_i^{s_i}\right] = \frac{1}{s!}\sum_{\nu_1=0}^{s_1}\cdots\sum_{\nu_n=0}^{s_n}(-1)^{\sum_{i=1}^n \nu_i}\binom{s_1}{\nu_1}\cdots\binom{s_n}{\nu_n}E\left[\left(\sum_{i=1}^n h_i z_i\right)^s\right].$$

Note that $Y = \sum_{i=1}^n h_i z_i \sim N(0, h'\Sigma h)$ and the s-th moment of Y is given by $E[Y^s] = (s-1)!!(h'\Sigma h)^{\frac{s}{2}}$ if s is even and 0 if s is odd. Using the fact that $s!/(s-1)!! = 2^{\frac{s}{2}}\left(\frac{s}{2}\right)!$, we obtain our expression. □

In using our expression to compute $\mu_{s_1,\ldots,s_n}$, we need to sum up $[(s_1+1)(s_2+1)\cdots(s_n+1)/2]$ terms and each term only involves computing the variance of a normal random variable.[3] For the example of $E[(z_1z_2z_3z_4z_5)^4]$, we only need to compute $[5^5/2] = 1{,}562$ terms, which is far less than the $19!! = 654{,}729{,}075$ terms required under the Isserlis's formula. For a given value of s, the worst case scenario for our method is when $s_1 = \ldots = s_n = 1$ which requires us to compute 2^{s-1} terms. However, even in this worst case scenario, it only requires computing 524,288 terms for $s = 20$, which is quite manageable with today's computers.[4]

A small modification of Proposition 1 enables us to compute $\mu_{s_1,\ldots,s_n}$ for $z \sim N(\mu, \Sigma)$, where μ is not a zero vector. This is because $Y = h'z$ has a distribution of $N(h'\mu, h'\Sigma h)$ and its s-th moment is given by

$$\begin{aligned}
E[Y^s] &= E[(Y - h'\mu + h'\mu)^s] \\
&= \sum_{r=0}^{[s/2]}\binom{s}{2r}(2r-1)!!(h'\Sigma h)^r(h'\mu)^{s-2r} \\
&= \sum_{r=0}^{[s/2]}\frac{s!}{r!(s-2r)!}\left(\frac{h'\Sigma h}{2}\right)^r(h'\mu)^{s-2r}.
\end{aligned}$$

Therefore, we only need to do one more summation to obtain $\mu_{s_1,\ldots,s_n}$ for the noncentral normal case. The results are summarized in the following Proposition.

Proposition 2 *Suppose $z = [z_1, \ldots, z_n]' \sim N(\mu, \Sigma)$, where Σ is an $n \times n$ positive semidefinite matrix. For nonnegative integers s_1 to s_n, we have*

$$E\left[\prod_{i=1}^n z_i^{s_i}\right] = \sum_{\nu_1=0}^{s_1}\cdots\sum_{\nu_n=0}^{s_n}\sum_{r=0}^{[s/2]}(-1)^{\sum_{i=1}^n \nu_i}\binom{s_1}{\nu_1}\cdots\binom{s_n}{\nu_n}\frac{\left(\frac{h'\Sigma h}{2}\right)^r(h'\mu)^{s-2r}}{r!(s-2r)!},$$

where $s = s_1 + \cdots + s_n$ and $h = \left[\frac{s_1}{2} - \nu_1, \ldots, \frac{s_n}{2} - \nu_n\right]'$.

[3] For programming purpose, we only need to update a small part of $h'\Sigma h$ as we loop through ν_1 to ν_n, and there is no need to compute $h'\Sigma h$ from scratch each time.

[4] Using the Windows version of Matlab running on an Opteron 165 processor, it takes less than five seconds to do this calculation.

Finally, Propositions 2 can be extended to the case that z is multivariate elliptically distributed with parameters μ and V (note that μ is the mean of z but V is in general not the covariance matrix of z). We assume z has a characteristic function $\exp(i\mu' t)\psi(t'Vt)$ for some function ψ. When z is multivariate elliptically distributed, $Y = h'z$ is also elliptically distributed. From Berkane and Bentler [2], we know the s-th central moment of Y (when it exists) is given by

$$E[(Y - h'\mu)^s] = \begin{cases} 0 & \text{if } s \text{ is odd,} \\ \left[1 + \kappa\left(\frac{s}{2}\right)\right](s-1)!!(h'\Sigma h)^{\frac{s}{2}} & \text{if } s \text{ is even,} \end{cases}$$

where $\Sigma = -2\psi'(0)V$ is the covariance matrix of z and $\kappa(r)$ is a moment parameter of the elliptical distribution, defined as [5]

$$\kappa(r) = \frac{\psi^{(r)}(0)}{\psi'(0)^r} - 1. \tag{3}$$

Therefore, one only needs to multiply our results for the normal case by some moment parameters to deal with the elliptical distribution case. The results are summarized in the following Proposition.

Proposition 3 *Suppose $z = [z_1, \ldots, z_n]'$ has a multivariate elliptical distribution with mean μ and covariance matrix Σ, where Σ is an $n \times n$ positive semidefinite matrix. Let $s = s_1 + \ldots + s_n$, where s_1 to s_n are nonnegative integers. Assuming the s-th moment of z exists and its moment parameters are $\kappa(r)$, $r = 1, \ldots, [s/2]$, (with $\kappa(0)$ defined as 0), the expectation of a product of s elements of z is given by*

$$E\left[\prod_{i=1}^n z_i^{s_i}\right] = \sum_{\nu_1=0}^{s_1} \cdots \sum_{\nu_n=0}^{s_n} \sum_{r=0}^{[s/2]} (-1)^{\sum_{i=1}^n \nu_i} \binom{s_1}{\nu_1} \cdots \binom{s_n}{\nu_n} \frac{[1+\kappa(r)]\left(\frac{h'\Sigma h}{2}\right)^r (h'\mu)^{s-2r}}{r!(s-2r)!},$$

where $h = \left[\frac{s_1}{2} - \nu_1, \ldots, \frac{s_n}{2} - \nu_n\right]'$.

4 Expectation of a Product of Quadratic Forms in Normal Random Variables

In time series analysis, a lot of estimators take the form of quadratic form or ratio of quadratic forms. Examples of these include the sample autocovariances and the sample autocorrelation coefficients are discussed in Anderson [1]. In obtaining the joint moments of these estimators, one needs to come up with a method to compute $Q_{s_1,\ldots,s_p}$.

[5] For example, under the multivariate t-distribution with k degrees of freedom, we have $\kappa(r) = \left(\frac{k}{2} - 1\right)^r / \left[\left(\frac{k}{2} - 1\right) \cdots \left(\frac{k}{2} - r\right)\right] - 1$ for $r < k/2$.

For $p = 1$, all moments of $z'Az$ can be easily computed. For example, Lemma 2 of Magnus [14] provides the following expression of Q_s

$$Q_s(A) = E[(z'Az)^s] = 2^s s! \sum_{\lambda} \prod_{j=1}^{s} \frac{[\mathrm{tr}(A\Sigma)^j + j\mu'(A\Sigma)^{j-1}A\mu]^{\lambda_j}}{\lambda_j!(2j)^{\lambda_j}}, \tag{4}$$

where the summation is over all s-vector $\lambda = (\lambda_1, \ldots, \lambda_s)$ whose elements λ_j are nonnegative integers satisfying $\sum_{j=1}^{s} j\lambda_j = s$, and tr stands for the trace operator.

Two computational remarks on this equation are in order. The first remark is that this equation requires an algorithm to enumerate all partitions of the integer s. Such an algorithm is readily available and can be found in Nijenhuis and Wilf [17]. The second remark is that when s is moderately large, it is often more efficient to compute the numerator of (4) by first performing an eigenvalue decomposition. When Σ is positive definite, we let L be a lower triangular matrix such that $LL' = \Sigma$. Suppose $L'AL = PDP'$, where $D = \mathrm{Diag}(d_1, \ldots, d_m)$ is a diagonal matrix of the $m \leq n$ nonzero eigenvalues of $L'AL$, and P is an $n \times m$ matrix of the corresponding eigenvectors. Then, denoting $\eta = P'L^{-1}\mu$, we can write

$$\mathrm{tr}(A\Sigma)^j + j\mu'(A\Sigma)^{j-1}A\mu = \mathrm{tr}D^j + j\eta' D^j \eta = \sum_{i=1}^{m} (1 + j\eta_i^2) d_i^j.$$

As it turns out, there exists a recursive expression for Q_s that is computationally far more efficient than (4). Based on the recursive relation between moments and cumulants, we can compute Q_s using (see, for example, Ruben [18] and Mathai and Provost [16] (Eq.3.2b.8))

$$Q_s(A) = s!2^s d_s(A), \tag{5}$$

where $d_s(A)$ is obtained using the following recursive relation

$$d_s(A) = \frac{1}{2s} \sum_{i=1}^{s} [\mathrm{tr}(A\Sigma)^i + i\mu'(A\Sigma)^{i-1}A\mu] d_{s-i}(A), \qquad d_0(A) = 1. \tag{6}$$

Although (5) is not an explicit expression, it is easier to program than (4) and it also takes much less time to compute.

Unlike the case of $p = 1$ where we can compute all the moments of $z'Az$ with relative ease, the case of $p > 1$ is far more challenging. For small s ($s \leq 4$), explicit expressions for $Q_{s_1,\ldots,s_p}$ are widely available (see, for example, Kumar [11], Magnus [12] and Holmquist [6]). For general s, Magnus [12,13] and Don [4] present an algorithm for computing $Q_{s_1,\ldots,s_p}$ when $\mu = 0$.[6] They express $Q_{s_1,\ldots,s_p}$ as an $A(s)$ polynomial, which is a polynomial with terms equal to various combinations of traces of s matrices. Explicit algorithm for enumerating the terms

[6] It is possible to generalize their results to the case of $\mu \neq 0$.

and computing the coefficients in this $A(s)$ polynomial was given by Magnus [12,13] and Don [4]. With some effort, their algorithm can be programmed to enumerate all the terms in the $A(s)$ polynomial. The main difficulty in applying this algorithm is the large number of terms involved in the $A(s)$ polynomial. Although there are only 17 terms for $s = 4$, the number of terms increases to 73 and 388 for $s = 5$ and 6. This probably explains why we never see explicit expression of $Q_{s_1,\dots,s_p}$ for $s > 4$. The number of terms in the $A(s)$ polynomial grows very quickly as s increases, so it is infeasible even for a computer to perform this exercise. For example, when $s = 12$, the $A(s)$ polynomial has 171,453,343 terms and it is a formidable task even for the fastest computer.

As we remarked earlier, computation of $E[(z'Az)^s]$ is relatively easy, so it is an ideal case to apply our identity on $(z'A_1z)^{s_1}\cdots(z'A_pz)^{s_p}$. The following Proposition presents the resulting new formula for $Q_{s_1,\dots,s_p}$.

Proposition 4 *Suppose $z = [z_1,\dots,z_n]' \sim N(\mu,\Sigma)$, where Σ is an $n\times n$ positive semidefinite matrix. For symmetric matrices A_1 to A_p, we have*

$$E\left[\prod_{i=1}^{p}(z'A_iz)^{s_i}\right] = \frac{1}{s!}\sum_{\nu_1=0}^{s_1}\cdots\sum_{\nu_p=0}^{s_p}(-1)^{\sum_{i=1}^{p}\nu_i}\binom{s_1}{\nu_1}\cdots\binom{s_p}{\nu_p}Q_s(B_\nu),$$

where $s = s_1+\cdots+s_p$, $B_\nu = \left(\frac{s_1}{2}-\nu_1\right)A_1+\cdots+\left(\frac{s_p}{2}-\nu_p\right)A_p$, and $Q_s(B_\nu)$ is given by the recursive equation in (5)–(6).

Proof: Putting $x_i = z'A_iz$ in (2), we have

$$E\left[\prod_{i=1}^{p}(z'A_iz)^{s_i}\right] = \frac{1}{s!}\sum_{\nu_1=0}^{s_1}\cdots\sum_{\nu_p=0}^{s_p}(-1)^{\sum_{i=1}^{p}\nu_i}\binom{s_1}{\nu_1}\cdots\binom{s_p}{\nu_p}E\left[\left(\sum_{i=1}^{p}h_iz'A_iz\right)^s\right].$$

As

$$E\left[\left(\sum_{i=1}^{p}h_iz'A_iz\right)^s\right] = E\left[\left(z'\left[\sum_{i=1}^{p}h_iA_i\right]z\right)^s\right] = E\left[(z'B_\nu z)^s\right],$$

we can use (5)–(6) to compute its expectation. □

Similar to the case of the products of normal random variables, there are $[(s_1+1)(s_2+1)\cdots(s_p+1)/2]$ terms in the expression for $Q_{s_1,\dots,s_p}$, and each term involves the computation of the s-th moment of a quadratic form in normal random variables. The worst case scenario occurs when $s_1 = s_2 = \cdots = s_p = 1$, in which case we need to evaluate $2^s - 1$ terms. As the computation of $Q_s(B_\nu)$ is quite efficient, we can now compute $Q_{s_1,\dots,s_p}$ for reasonably large s. As an example, for $n = 60$, $E[\prod_{i=1}^{5}(z'A_iz)^4]$ takes less than three seconds to compute using the Windows version of Matlab running on an Opteron 165 processor. Even for the worst case scenario of $E[\prod_{i=1}^{20}(z'A_iz)]$, the answer can be obtained in less than 13 minutes.

Extension of Proposition 4 to the case of multivariate elliptical distribution is possible.

We start out with the case of $\mu = 0$. For the multivariate normal case, all the terms in $(z'A_1z)^{s_1}(z'A_1z)^{s_2}\cdots(z'A_pz)^{s_p}$ are just products of $2s$ central normal random variables. Using Lemma 2 of Maruyama and Seo [15], it is easy to show that $Q_{s_1,\ldots,s_p}$ for the multivariate elliptical distribution case is simply $1+\kappa(s)$ times the value of $Q_{s_1,\ldots,s_p}$ for the multivariate normal case, where $\kappa(s)$ is a moment parameter of the multivariate elliptical distribution defined in (3). For the case that $\mu \neq 0$, it takes a bit more work. The key is to obtain an expression for $E[(z'Az)^s]$ where z follows a multivariate elliptical distribution with parameters μ and V. This can be accomplished as follows. Let $\varepsilon = z - \mu$, then ε has the same multivariate elliptical distribution as z but with zero mean. Also define $u \sim N(0,V)$ as a vector of normal random variables that has the same covariance as z but has zero mean. Expanding $(z'Az)^s$ using the binomial theorem, we have

$$
\begin{aligned}
E[(z'Az)^s] &= E[((\mu+\varepsilon)'A(\mu+\varepsilon))^s] \\
&= E[(\mu'A\mu+\varepsilon'A\varepsilon+2\mu'A\varepsilon)^s] \\
&= \sum_{i=0}^{s}\binom{s}{i}E[(\mu'A\mu+\varepsilon'A\varepsilon)^{s-i}(2\mu'A\varepsilon)^i] \\
&= \sum_{r=0}^{[s/2]}\binom{s}{2r}E[(\mu'A\mu+\varepsilon'A\varepsilon)^{s-2r}(2\mu'A\varepsilon)^{2r}] \\
&= \sum_{r=0}^{[s/2]}\binom{s}{2r}\sum_{q=0}^{s-2r}\binom{s-2r}{q}(\mu'A\mu)^{s-2r-q}E[(\varepsilon'A\varepsilon)^q(2\mu'A\varepsilon)^{2r}] \\
&= \sum_{r=0}^{[s/2]}\binom{s}{2r}\sum_{q=0}^{s-2r}\binom{s-2r}{q}(\mu'A\mu)^{s-2r-q}[1+\kappa(r+q)]E[(u'Au)^q(2\mu'Au)^{2r}].
\end{aligned}
$$

The fourth equality follows because the expectation vanishes when i is odd due to symmetry of the distribution of $\mu'A\varepsilon$. The last equality is obtained by applying Lemma 2 of Maruyama and Seo [15] and noting that the terms in the expectation are products of $2(r+q)$ random variables from a central elliptical distribution. With this expression, we can then compute $Q_s(A)$ for the case of elliptical distribution and $Q_{s_1,\ldots,s_p}$ follows naturally.

5 Conclusion

Expectations of products of random variables are generally difficult to obtain. Even when explicit expressions are available, they are often impractical for computational purpose when the number of random variables in the product is moderately large. In contrast, expressions of moments of sums of random variables are often much simpler and easier to compute, regardless of the number of random variables in the sum. In this paper, we present an identity that allows us to evaluate the expectation of a product of random variables by computing the moments of various sums of the same random variables. To illustrate the value of this identity, we apply it to derive expressions for the expectation of a product of normal random variables and also for the expectation of a product of quadratic forms in normal random variables. The resulting formulae are computationally far more efficient than existing formulae.[7] The results are easily generalized to the case that the random variables have a multivariate elliptical distribution. These results are potentially useful in many applications. As an example, we can model the residuals in a time series to have a multivariate Student t distribution. Applying our results, we can then easily compute the moments and cross-moments of statistical estimators that are quadratic forms of the residuals. Finally, we anticipate our identity can also be used to provide efficient formulae for computing expectation of products of other nonnormal random variables.

Acknowledgement

I thank Jan Magnus, Serge Provost, Kostas Triantafyllopoulos and an anonymous referee for helpful comments, and Xiaolu Wang for research assistance. Financial support from the National Bank Financial of Canada is gratefully acknowledged.

Appendix A. Proof of Lemma 1

Define $f_i(t) = (e^{tx_i/2} - e^{-tx_i/2})^{s_i}$ and $f(t) = \prod_{i=1}^{n} f_i(t)$. Note that

$$\frac{\mathrm{d}^p f_i(t)}{\mathrm{d}t^p} = \frac{\mathrm{d}^p}{\mathrm{d}t^p} \sum_{r=0}^{s_i} (-1)^{s_i - r} \binom{s_i}{r} e^{tx_i\left(r - \frac{s_i}{2}\right)} = \sum_{r=0}^{s_i} (-1)^{s_i - r} \binom{s_i}{r} e^{tx_i\left(r - \frac{s_i}{2}\right)} \left[x_i \left(r - \frac{s_i}{2}\right)\right]^p,$$

so we have

$$\left.\frac{\mathrm{d}^p f_i(t)}{\mathrm{d}t^p}\right|_{t=0} = x_i^p \sum_{r=0}^{s_i} (-1)^{s_i - r} \binom{s_i}{r} \left(r - \frac{s_i}{2}\right)^p.$$

Using the following identity (see, for example, Feller [5] (p.65))

$$\sum_{r=0}^{s_i} (-1)^{s_i - r} \binom{s_i}{r} r^p = \begin{cases} 0 & \text{if } p < s_i, \\ s_i! & \text{if } p = s_i, \end{cases}$$

and the fact that $(r - \frac{s_i}{2})^p$ is a polynomial in r with the leading term as r^p, we have

$$\left.\frac{\mathrm{d}^p f_i(t)}{\mathrm{d}t^p}\right|_{t=0} = \begin{cases} 0 & \text{if } p < s_i, \\ x_i^{s_i} s_i! & \text{if } p = s_i. \end{cases} \tag{7}$$

[7] A set of Matlab programs to implement our formulae is available upon request.

Differentiating $f(t)$ s times and using the generalized Leibniz rule,[8] we have

$$\frac{\mathrm{d}^s f(t)}{\mathrm{d}t^s} = \sum_{\substack{\nu_1+\cdots+\nu_n=s\\ \nu_i\geq 0}} \frac{s!}{\nu_1!\cdots\nu_n!}\prod_{i=1}^n \frac{\mathrm{d}^{\nu_i} f_i(t)}{\mathrm{d}t^{\nu_i}}. \tag{8}$$

When we evaluate the derivative at $t=0$, we know from (7) that the right hand side of (8) is only nonzero when $\nu_i = s_i$, $i=1,\ldots,n$. It follows that

$$\left.\frac{\mathrm{d}^s f(t)}{\mathrm{d}t^s}\right|_{t=0} = \frac{s!}{s_1!\cdots s_n!}\prod_{i=1}^n s_i! x_i^{s_i} = s! x_1^{s_1}\cdots x_n^{s_n}. \tag{9}$$

Now, applying the binomial theorem to $f_i(t)$, we can also write $f(t)$ as

$$\begin{aligned} f(t) &= (e^{tx_1/2} - e^{-tx_1/2})^{s_1}(e^{tx_2/2} - e^{-tx_2/2})^{s_2}\cdots(e^{tx_n/2} - e^{-tx_n/2})^{s_n} \\ &= \left[\sum_{\nu_1=0}^{s_1}\binom{s_1}{\nu_1}(-1)^{\nu_1} e^{t\left(\frac{s_1}{2}-\nu_1\right)x_1}\right]\cdots\left[\sum_{\nu_n=0}^{s_n}\binom{s_n}{\nu_n}(-1)^{\nu_n} e^{t\left(\frac{s_n}{2}-\nu_n\right)x_n}\right] \\ &= \sum_{\nu_1=0}^{s_1}\cdots\sum_{\nu_n=0}^{s_n}(-1)^{\nu_1+\cdots+\nu_n}\binom{s_1}{\nu_1}\cdots\binom{s_n}{\nu_n} e^{\left(\sum_{i=1}^n h_i x_i\right)t}. \end{aligned}$$

Differentiating it s times, we have

$$\frac{\mathrm{d}^s f(t)}{\mathrm{d}t^s} = \sum_{\nu_1=0}^{s_1}\cdots\sum_{\nu_n=0}^{s_n}(-1)^{\sum_{i=1}^n \nu_i}\binom{s_1}{\nu_1}\cdots\binom{s_n}{\nu_n} e^{\left(\sum_{i=1}^n h_i x_i\right)t}\left(\sum_{i=1}^n h_i x_i\right)^s.$$

Evaluating the derivative at $t=0$, we obtain

$$\left.\frac{\mathrm{d}^s f(t)}{\mathrm{d}t^s}\right|_{t=0} = \sum_{\nu_1=0}^{s_1}\cdots\sum_{\nu_n=0}^{s_n}(-1)^{\sum_{i=1}^n \nu_i}\binom{s_1}{\nu_1}\cdots\binom{s_n}{\nu_n}\left(\sum_{i=1}^n h_i x_i\right)^s. \tag{10}$$

Equating (9) with (10), we obtain our identity. □

References

[1] T. W. Anderson, The Statistical Analysis of Time Series, Wiley, New York, 1971.

[2] M. Berkane, and P. M. Bentler, Moments of elliptically distributed random variates, Statistics and Probability Letters 4 (1986) 333–335.

[8] See Thaheen and Laradji [20] for a simple proof of the generalized Leibniz rule.

[3] R. Blacher, Multivariate quadratic forms of random vectors, Journal of Multivariate Analysis 87 (2003) 2–23.

[4] F. J. H. Don, The expecation of products of quadratic forms in normal variables, Statistica Neerlandica 33 (1979) 73–79.

[5] W. Feller, An Introduction to Probability Theory and its Applications, Vol. I, 3rd edition, Wiley, New York, 1968.

[6] B. Holmquist, Moments and cumulants of the multivariate normal distribution, Stochastic Analysis and Applications 6 (3) (1988) 273–278.

[7] B. Holmquist, Expectations of products of quadratic forms in normal variables, Stochastic Analysis and Applications 14 (2) (1996) 149–164.

[8] L. Isserlis, On a fomula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables, Biometrika 12 (1 and 2) (1918) 134–139.

[9] K. Jinadasa, D. S. Tracy, Higher order moments of random vectors using matrix derivatives, Stochastic Analysis and Applications 4 (1986) 399–407.

[10] M. G. Kendall, A. Stuart, The Advanced Theory of Statistics, Vol. 1, 4th edition, Charles Griffin and Co., London, 1977.

[11] A. Kumar, Expectations of products of quadratic forms, Sankhyā Series B 35 (1973) 359–362.

[12] J. Magnus, The moment of products of quadratic forms in normal variables, Statistica Neerlandica 32 (1978) 201–210.

[13] J. Magnus, The expectation of products of quadratic forms in normal variables: the practice, Statistica Neerlandica 33 (1979) 131–136.

[14] J. Magnus, The exact moments of a ratio of quadratic forms in normal variables, Annales d'Économie et de Statistique 4 (1986) 95–109.

[15] Y. Maruyama, T. Seo, Estimation of moment parameter in elliptical distributions, Journal of Japan Statistical Society 33 (2) (2003) 215–229.

[16] A. M. Mathai, S. B. Provost, Quadratic Forms in Random Variables: Thoery and Applications, Marcel Dekker, New York, 1992.

[17] A. Nijenhus, H. Wilf, Combinatorial Algorithms, 2nd edition, Academic Press, New York, 1978.

[18] H. Ruben, Probability content of regions under spherical normal distributions, IV: The distribution of homogeneous and non-homogeneous quadratic functions of normal variables, Annals of Mathematical Statistics 33 (1962) 542–570.

[19] J. R. Schott, Kronecker product permutation matrics and their application to moment matrices of the normal distribution, Journal of Multivariate Analysis 87 (2003) 177–190.

[20] A. B. Thaheem, A. Laradji, A generalization of Leibniz rule for higher derivatives, International Journal of Mathematical Education in Science and Technology 34 (6) (2003) 905–951.

[21] D. S. Tracy, K. G. Jinadasa, Expectations of products of random quadratic forms, Stochastic Analysis and Applications 4 (1986) 111–116.

[22] D. S. Tracy, S. A. Sultan, Higher order moments of multivariate normal distribution using matrix derivatives, Stochastic Analysis and Applications 11 (1993) 337–348.

[23] K. Triantafyllopoulos, On the central moments of the multidimensional Guassian distribution, Mathematical Scientist 28 (2003) 125–128.