Some Useful Expected Values with Multivariate Normal Distribution and Uniform Distribution on Sphere

Liyuan Cao

February 15, 2020

1 Useful values with multivariate normally distributed variables

Let $\epsilon \in \mathbb{R}^n$ be a set of variables following multivariate normal distribution: $\epsilon \sim \mathcal{N}(0_{n\times 1}, I_{n\times n})$. Let a be an arbitrary constant vector in $\mathbb{R}^n$. We will derive the following expectations.

$$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}(\epsilon^\intercal\epsilon)^k = \frac{(n+2k-2)!!}{(n-2)!!} = n(n+2)(n+4)\cdots(n+2k-2) \tag{1a}$$

$$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}(\epsilon^\intercal\epsilon)^k\epsilon\epsilon^\intercal = \frac{(n+2k)!!}{n!!}I = (n+2)(n+4)\cdots(n+2k)I \tag{1b}$$

$$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}(a^\intercal\epsilon)^{2k} = (a^\intercal a)^k \tag{1c}$$

$$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}(a^\intercal\epsilon)^{2k_1}(\epsilon^\intercal\epsilon)^{k_2} = (2k_1-1)!!\frac{(n+2k_1+2k_2-2)!!}{(n+2k_1-2)!!}(a^\intercal a)^{k_1} \tag{1d}$$

There is also an attempted derivation of

$$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}\mathrm{E}_{\bar\epsilon\sim\mathcal{N}(0,I)}(a^\intercal\epsilon)^{2k_1}(a^\intercal\bar\epsilon)^{2k_2}(\epsilon^\intercal\bar\epsilon)^{k_3}$$

with k_1, k_2, k_3 being natural numbers, but I was unable to get the result. Instead, the direct approach was use to calculate the following two expectations used in the paper:

$$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}\mathrm{E}_{\bar\epsilon\sim\mathcal{N}(0,I)}(a^\intercal\epsilon)^2(a^\intercal\bar\epsilon)^2(\epsilon\bar\epsilon)^2$$
$$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}\mathrm{E}_{\bar\epsilon\sim\mathcal{N}(0,I)}(a^\intercal\epsilon)^3(a^\intercal\bar\epsilon)(\epsilon^\intercal\bar\epsilon).$$

To calculate the above values, we need to first notice ϵ_i and ϵ_j, the ith and jth element of ϵ, each follows standard normal distribution independently for any $i \neq j$. In addition, any odd moment of a normally distributed variable is 0, i.e., for any $k \geq 0$ that is odd

$$\mathrm{E}_{\epsilon_i\sim\mathcal{N}(0,1)}\epsilon_i^k = 0. \tag{2}$$

$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}(\epsilon^\intercal\epsilon)^k = n(n+2)(n+4)\cdots(n+2k-2)$

$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}(\epsilon^\intercal\epsilon)^k\epsilon\epsilon^\intercal = (n+2)(n+4)\cdots(n+2k)I$

Notice that $\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}(\epsilon^\intercal\epsilon)^k$ is the kth moment of a Chi-square distributed random variable, so we can get this expectation by simply looking it up in Wikipedia.

The expectation $\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}(\epsilon^\intercal\epsilon)^k\epsilon\epsilon^\intercal$ is an $n\times n$ matrix. Let's denote this matrix as A, and the element on its ith row and jth column as A_{ij}. We can see

$$A_{ij} = (\epsilon^\intercal\epsilon)^k\epsilon_i\epsilon_j = \left(\sum_{t=1}^{n}\epsilon_t^2\right)^k\epsilon_i\epsilon_j.$$

By expanding $\left(\sum_{t=1}^{n}\epsilon_t^2\right)^k$, A_{ij} can be written as the sum of n^k terms. Let $\epsilon_{t_1}^2\epsilon_{t_2}^2\ldots\epsilon_{t_k}^2\epsilon_i\epsilon_j$ be one of those terms, where $(t_1,t_2,\ldots,t_k)\in\{1,2,\ldots,n\}^k$. Considering (2), the only way $\mathbb{E}\epsilon_{t_1}^2\epsilon_{t_2}^2\ldots\epsilon_{t_k}^2\epsilon_i\epsilon_j$ is not zero is when $i\neq j$. Thus $A_{ij}=0$ for all (i,j) such that $i\neq j$.

Now we know $A=\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}(\epsilon^\intercal\epsilon)^k\epsilon\epsilon^\intercal$ is a diagonal matrix. Also notice

$$\sum_{i=1}^{n}A_{ii} = \sum_{i=1}^{n}\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}(\epsilon^\intercal\epsilon)^k\epsilon_i^2 = \mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}(\epsilon^\intercal\epsilon)^{k+1}.$$

We should also have $A_{ii}=A_{jj}$ for all $(i,j)\in\{1,2,\ldots,n\}$ by summetry, so

$$A = \frac{1}{n}\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}(\epsilon^\intercal\epsilon)^{k+1}I = (n+2)(n+4)\cdots(n+2k)I.$$

$\mathrm{E}_{\epsilon \sim \mathcal{N}(0,I)}(a^\intercal \epsilon)^{2k} = (a^\intercal a)^k$

The moment generating function of $a^\intercal \epsilon$ is

$$\begin{aligned} M_{a^\intercal \epsilon}(t) &= \mathbb{E}\exp\left(t a^\intercal \epsilon\right) = \mathbb{E}\prod_{i=1}^{n} \exp\left(t a_i \epsilon_i\right) \\ &= \prod_{i=1}^{n} \mathbb{E}\exp\left(t a_i \epsilon_i\right) = \prod_{i=1}^{n} \exp\left(\frac{1}{2} t^2 a_i^2\right) = \exp\left(\frac{1}{2} t^2 a^\intercal a\right) \\ &= \sum_{i=1}^{\infty} \frac{1}{i!}\left(\frac{1}{2} t^2 a^\intercal a\right)^i . \end{aligned}$$

Thus we have

$$\begin{aligned} &\mathrm{E}_{\epsilon \sim \mathcal{N}(0,I)}(a^\intercal \epsilon)^{2k} \\ &= \frac{\mathrm{d}^{2k}}{\mathrm{d}t^{2k}} M_{a^\intercal \epsilon}(t) = \frac{\mathrm{d}^{2k}}{\mathrm{d}t^{2k}} \frac{1}{k!}\left(\frac{1}{2} t^2 a^\intercal a\right)^k \\ &= \frac{(2k)!}{2^k k!}(a^\intercal a)^k = (2k-1)!!(a^\intercal a)^k \end{aligned}$$

$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}(a^\intercal\epsilon)^{2k_1}(\epsilon^\intercal\epsilon)^{k_2} = (2k_1-1)!!\frac{(n+2k_1+2k_2-2)!!}{(n+2k_1-2)!!}(a^\intercal a)^{k_1}$

The moment generating function for $(a^\intercal\epsilon)(\epsilon^\intercal\epsilon)$ is

$$
\begin{aligned}
&M_{a^\intercal\epsilon,\epsilon^\intercal\epsilon}(t_1,t_2)\\
&=\mathbb{E}\exp\left(t_1(a^\intercal\epsilon)\right)\exp\left(t_2(\epsilon^\intercal\epsilon)\right)=\prod_{i=1}^n\mathbb{E}\exp\left(t_1a_i\epsilon_i+t_2\epsilon_i^2\right)\\
&=\prod_{i=1}^n\int_{\mathbb{R}^n}\exp\left(t_1a_iu_i+t_2u_i^2\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{u_i^2}{2}\right)\mathrm{d}u_i\\
&=\prod_{i=1}^n\frac{1}{\sqrt{1-2t_2}}\exp\left(\frac{(t_1a_i)^2}{2(1-2t_2)}\right)\int_{\mathbb{R}^n}\frac{\sqrt{1-2t_2}}{\sqrt{2\pi}}\exp\left(-\frac{1-2t_2}{2}\left(u_i-\frac{t_1a_i}{1-2t_2}\right)^2\right)\mathrm{d}u_i\\
&=\prod_{i=1}^n\frac{1}{\sqrt{1-2t_2}}\exp\left(\frac{(t_1a_i)^2}{2(1-2t_2)}\right)=(1-2t_2)^{-\frac{n}{2}}\exp\left(\frac{t_1^2a^\intercal a}{2(1-2t_2)}\right)\\
&=(1-2t_2)^{-\frac{n}{2}}\sum_{i=1}^\infty\frac{1}{i!}\left(\frac{t_1^2a^\intercal a}{2(1-2t_2)}\right)^i.
\end{aligned}
$$

Taking the $2k_1$th derivative of $M_{a^\intercal\epsilon,\epsilon^\intercal\epsilon}(t_1,t_2)$ with respect to t_1 yields

$$
\begin{aligned}
&\frac{\partial^{2k_1}}{\partial t_1^{2k_1}}M_{a^\intercal\epsilon,\epsilon^\intercal\epsilon}(t_1,t_2)\\
&=(1-2t_2)^{-\frac{n}{2}}\frac{1}{k_1!}\left(\frac{a^\intercal a}{2(1-2t_2)}\right)^{k_1}(2k_1)!\\
&=(1-2t_2)^{-\frac{n+2k_1}{2}}(2k_1-1)!!(a^\intercal a)^{k_1},
\end{aligned}
$$

and further take the k_2th derivative with respect to t_2 yields

$$
\begin{aligned}
&\frac{\partial^{k_2}}{\partial t_2^{k_2}}\frac{\partial^{2k_1}}{\partial t_1^{2k_1}}M_{a^\intercal\epsilon,\epsilon^\intercal\epsilon}(t_1,t_2)\\
&=\frac{\partial^{k_2}}{\partial t_2^{k_2}}(1-2t_2)^{-\frac{n+2k_1}{2}}(2k_1-1)!!(a^\intercal a)^{k_1}\\
&=\frac{(n+2k_1+2k_2-2)!!}{(n+2k_1-2)!!}(2k_1-1)!!(a^\intercal a)^{k_1},
\end{aligned}
$$

where the last equality can be easily shown with induction.

$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}\mathrm{E}_{\bar\epsilon\sim\mathcal{N}(0,I)}(a^\intercal\epsilon)^{2k_1}(a^\intercal\bar\epsilon)^{2k_2}(\epsilon^\intercal\bar\epsilon)^{k_3}$

The moment generating function:

$$\begin{aligned}
&\mathbb{E}\exp\left(t_1 a^\intercal\epsilon\right)\exp\left(t_2 a^\intercal\bar\epsilon\right)\exp\left(t_3\epsilon^\intercal\bar\epsilon\right)\\
=&\mathbb{E}\exp\left(\sum_{i=1}^n t_1 a_i\epsilon_i + t_2 a_i\bar\epsilon_i + t_3\epsilon_i\bar\epsilon_i\right)\\
=&\prod_{i=1}^n \mathbb{E}\exp\left(t_1 a_i\epsilon_i + t_2 a_i\bar\epsilon_i + t_3\epsilon_i\bar\epsilon_i\right)\\
=&\prod_{i=1}^n \int_{\mathbb{R}^2}\exp\left(t_1 a_i u + t_2 a_i v + t_3 uv\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{u^2}{2}\right)\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{v^2}{2}\right)\mathrm{d}(u,v)\\
=&\prod_{i=1}^n \int_{\mathbb{R}^2}\frac{1}{2\pi}\int_{\mathbb{R}^2}\exp\left(-\frac{1}{2}u^2-\frac{1}{2}v^2+t_1 a_i u + t_2 a_i v + t_3 uv\right)\mathrm{d}(u,v);
\end{aligned}$$

by doing the replacement

$$\begin{bmatrix}u\\v\end{bmatrix}=\begin{bmatrix}\sqrt{2}/2 & \sqrt{2}/2\\ -\sqrt{2}/2 & \sqrt{2}/2\end{bmatrix}\begin{bmatrix}u'\\v'\end{bmatrix}$$

$$\begin{aligned}
=&\prod_{i=1}^n \int_{\mathbb{R}^2}\frac{1}{2\pi}\int_{\mathbb{R}^2}\exp\left(-\frac{1+t_3}{2}u'^2-\frac{1-t_3}{2}v'^2+\frac{\sqrt{2}(t_1-t_2)}{2}a_i u'+\frac{\sqrt{2}(t_1+t_2)}{2}a_i v'\right)\mathrm{d}(u',v')\\
=&\prod_{i=1}^n\left[\begin{array}{l}\exp\left(\frac{a_i^2(t_1-t_2)^2}{4(1+t_3)}\right)\frac{1}{\sqrt{1+t_3}}\int_{\mathbb{R}}\frac{\sqrt{1+t_3}}{\sqrt{2\pi}}\exp\left(-\frac{1+t_3}{2}\left(u-\frac{\sqrt{2}a_i(t_1-t_2)}{2(t_3+1)}\right)^2\right)\\ \cdot\exp\left(\frac{a_i^2(t_1+t_2)^2}{4(1-t_3)}\right)\frac{1}{\sqrt{1-t_3}}\int_{\mathbb{R}}\frac{\sqrt{1-t_3}}{\sqrt{2\pi}}\exp\left(-\frac{1-t_3}{2}\left(v+\frac{\sqrt{2}a_i(t_1+t_2)}{2(t_3-1)}\right)^2\right)\end{array}\right]\\
=&\prod_{i=1}^n\exp\left(\frac{a_i^2(t_1-t_2)^2}{4(1+t_3)}\right)\frac{1}{\sqrt{1+t_3}}\exp\left(\frac{a_i^2(t_1+t_2)^2}{4(1-t_3)}\right)\frac{1}{\sqrt{1-t_3}}\\
=&(1-t_3^2)^{-\frac{n}{2}}\exp\left(\left[\frac{(t_1-t_2)^2}{4(1+t_3)}+\frac{(t_1+t_2)^2}{4(1-t_3)}\right]a^\intercal a\right)
\end{aligned}$$

The next step should be calculating

$$\frac{\partial^{2k_1}}{\partial t_1^{2k_1}}\frac{\partial^{2k_2}}{\partial t_2^{2k_2}}\frac{\partial^{k_3}}{\partial t_3^{k_3}}\mathbb{E}\exp\left(t_1 a^\intercal\epsilon\right)\exp\left(t_2 a^\intercal\bar\epsilon\right)\exp\left(t_3\epsilon^\intercal\bar\epsilon\right).$$

However I could not do it.

$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}\mathrm{E}_{\bar{\epsilon}\sim\mathcal{N}(0,I)}(a^\intercal\epsilon)^2(a^\intercal\bar{\epsilon})^2(\epsilon\bar{\epsilon})^2 = (n+8)(a^\intercal a)^2$

$$
\begin{aligned}
&\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}\mathrm{E}_{\bar{\epsilon}\sim\mathcal{N}(0,I)}(a^\intercal\epsilon)^2(a^\intercal\bar{\epsilon})^2(\epsilon\bar{\epsilon})^2 \\
= &\sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{t=1}^{n}\sum_{w=1}^{n}(a_p u_p)(a_q u_q)(a_r\bar{u}_r)(a_s\bar{u}_s)(u_t\bar{u}_t)(u_w\bar{u}_w)
\end{aligned}
$$

Now there are the following scenarios:

1. $p=q=r=s=t=w$

$$
\begin{aligned}
&\mathbb{E}(a_p u_p)(a_q u_q)(a_r\bar{u}_r)(a_s\bar{u}_s)(u_t\bar{u}_t)(u_w\bar{u}_w) \\
= &\mathbb{E}(a_p u_p)^2(a_p\bar{u}_p)^2(u_p\bar{u}_p)^2 \\
= &9a_p^4;
\end{aligned}
$$

2. $p=q=r=s$, $t=w$ (and $s\neq t$ is implied here and similarly below)

$$
\begin{aligned}
&\mathbb{E}(a_p u_p)(a_q u_q)(a_r\bar{u}_r)(a_s\bar{u}_s)(u_t\bar{u}_t)(u_w\bar{u}_w) \\
= &\mathbb{E}(a_p u_p)^2(a_p\bar{u}_p)^2(u_t\bar{u}_t)^2 \\
= &a_p^4;
\end{aligned}
$$

3. $p=q$, $r=s$, $t=w$

$$
\begin{aligned}
&\mathbb{E}(a_p u_p)(a_q u_q)(a_r\bar{u}_r)(a_s\bar{u}_s)(u_t\bar{u}_t)(u_w\bar{u}_w) \\
= &\mathbb{E}(a_p u_p)^2(a_r\bar{u}_r)^2(u_t\bar{u}_t)^2 \\
= &a_p^2a_r^2;
\end{aligned}
$$

4. $p=q=t=w$, $r=s$ (or $r=s=t=w$, $p=q$)

$$
\begin{aligned}
&\mathbb{E}(a_p u_p)(a_q u_q)(a_r\bar{u}_r)(a_s\bar{u}_s)(u_t\bar{u}_t)(u_w\bar{u}_w) \\
= &\mathbb{E}(a_p u_p)^2(a_r\bar{u}_r)^2(u_p\bar{u}_p)^2 \\
= &3a_p^2a_r^2;
\end{aligned}
$$

5. $p=r=t$, $q=s=w$ (or $p=s=t$, $q=r=w$ or $q=r=t$, $p=s=w$ or $q=s=t$, $p=r=w$)

$$
\begin{aligned}
&\mathbb{E}(a_p u_p)(a_q u_q)(a_r\bar{u}_r)(a_s\bar{u}_s)(u_t\bar{u}_t)(u_w\bar{u}_w) \\
= &\mathbb{E}(a_p u_p)(a_p\bar{u}_p)(u_p\bar{u}_p)(a_q u_q)(a_q\bar{u}_q)(u_q\bar{u}_q) \\
= &a_p^2a_q^2.
\end{aligned}
$$

All remain cases have expectation 0 as they will fall into the situation of (2). Thus the final expectation is

$$
\begin{aligned}
&\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}\mathrm{E}_{\bar{\epsilon}\sim\mathcal{N}(0,I)}(a^\intercal\epsilon)^2(a^\intercal\bar{\epsilon})^2(\epsilon\bar{\epsilon})^2 \\
&= 9\sum_{p=1}^{n} a_p^4 + \sum_{p=1}^{n}\sum_{t\neq p} a_p^4 + \sum_{p=1}^{n}\sum_{r\neq p}\sum_{t\neq p, t\neq r} a_p^2 a_r^2 + 2\sum_{p=1}^{n}\sum_{r\neq p} 3a_p^2 a_r^2 + 4\sum_{p=1}^{n}\sum_{q\neq p} a_p^2 a_q^2 \\
&= (n+8)\sum_{p=1}^{n} a_p^4 + (n+8)\sum_{p=1}^{n}\sum_{r\neq p} a_p^2 a_r^2 \\
&= (n+8)(a^\intercal a)^2.
\end{aligned}
$$

Similar technique can be used to derive the following expectations:

$$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}\mathrm{E}_{\bar{\epsilon}\sim\mathcal{N}(0,I)}(a^\intercal\epsilon)(a^\intercal\bar{\epsilon})(\epsilon^\intercal\bar{\epsilon}) = a^\intercal a$$

$$\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}\mathrm{E}_{\bar{\epsilon}\sim\mathcal{N}(0,I)}(a^\intercal\epsilon)^3(a^\intercal\bar{\epsilon})(\epsilon^\intercal\bar{\epsilon}) = 3(a^\intercal a)^2$$

2 Useful values with variables uniformly distributed on unit sphere

Uniform distribution on unit sphere is closely related to multivariate normal distribution because for $\epsilon \sim \mathcal{N}(0, I)$, $\epsilon/\|\epsilon\| \sim U(S(0,1))$. Here $U(\cdot)$ stands for uniform distribution on a set, and $S(0,1)$ stands for a sphere centered at the all zero vector and has radius of 1.

One important thing worth noticing is that the angle and the magnitude of an $\epsilon \sim \mathcal{N}(0, I)$ are independent, so we can have

$$\begin{aligned}\mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}\left\{(\epsilon^\intercal\epsilon)^2\epsilon\epsilon^\intercal\right\} &= \mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}\left\{\frac{(\epsilon^\intercal\epsilon)^2\epsilon\epsilon^\intercal}{\|\epsilon\|^6}\|\epsilon\|^6\right\} \\ &= \mathrm{E}_{\epsilon\sim U(S(0,1))}\left\{(\epsilon^\intercal\epsilon)^2\epsilon\epsilon^\intercal\right\}\mathrm{E}_{X\sim\chi^2_n}\left\{X^3\right\}\end{aligned}$$

with ϵ and X being independent from each other.

Then we have, for example,

$$\begin{aligned}\mathrm{E}_{\epsilon\sim U(S(0,1))}\left\{(\epsilon^\intercal\epsilon)^2\epsilon\epsilon^\intercal\right\} &= \mathrm{E}_{\epsilon\sim\mathcal{N}(0,I)}\left\{(\epsilon^\intercal\epsilon)^2\epsilon\epsilon^\intercal\right\}/\mathrm{E}_{X\sim\chi^2_n}\left\{X^3\right\} \\ &= \frac{(n+2)(n+4)}{n(n+2)(n+4)}I \\ &= \frac{1}{n}I.\end{aligned}$$

Similarly,

$$\mathrm{E}_{\epsilon\sim U(S(0,1))}(a^\intercal\epsilon)^2\epsilon\epsilon^\intercal = \frac{a^\intercal a I + 2aa^\intercal}{n(n+2)} \tag{3a}$$

$$\mathrm{E}_{\epsilon\sim U(S(0,1))}(\epsilon^\intercal\epsilon)\epsilon\epsilon^\intercal = \frac{(n+2)I}{n(n+2)} = \frac{1}{n}I \tag{3b}$$

$$\mathrm{E}_{\epsilon\sim U(S(0,1))}(\epsilon^\intercal\epsilon)^2\epsilon\epsilon^\intercal = \frac{(n+2)(n+4)I}{n(n+2)(n+4)} = \frac{1}{n}I \tag{3c}$$

$$\mathrm{E}_{\epsilon\sim U(S(0,1))}(\epsilon^\intercal\epsilon)^3\epsilon\epsilon^\intercal = \frac{(n+2)(n+4)(n+6)I}{n(n+2)(n+4)(n+6)} = \frac{1}{n}I, \tag{3d}$$

where a is any vector in $\mathbb{R}^n$ independent from ϵ.