Quarternion gaussian integral and its application to a geometry of quarternion gaussian distributions

Tatsuo Suzuki\textsuperscript{1, a}
\textsuperscript{1}Department of Mathematical Sciences, Shibaura Institute of Technology, Japan
E-mail: \textsuperscript{a}suzukita@shibaura-it.ac.jp

Abstract. We give an explicit formula of a quarternion Gaussian integral in terms of a quarternionic determinant. Then, we obtain a simple form of a probability density function of multivariate quarternion Gaussian distributions. Moreover, as an application to a geometry, we obtain the Fisher metric of quarternion Gaussian distributions by using the potential function of them.

1. Introduction
The statistical model of real multivariate Gaussian distributions (or normal distributions) plays an important role in the information geometry [1]. On the other hand, in the quaternion valued statistical signal processing and related areas, quaternion analysis have been developed, for instance, see [5], [7] and references within. Therefore, we investigate geometry of the quarternion Gaussian distributions.

1.1. Exponential family
A statistical model $S = \{p_\theta | \theta \in \Xi\}$ is called an exponential family if, for functions $\{C, F_1, \cdots, F_n\}$ on $\chi$ and a function $\psi$ on $\Xi$, $p_\theta$ is written as

$$p_\theta(x) = p(x; \theta) = \exp \left\{ C(x) + \sum_{i=1}^{n} \theta^i F_i(x) - \psi(\theta) \right\}. \quad (1)$$

1.2. Geometry of univariate Gaussian distributions
Let $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_{>0}$. A probability density function of univariate Gaussian distribution is given by

$$p(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\} \quad (-\infty < x < \infty) \quad (2)$$

$$= \exp \left\{ \frac{\mu^2}{2\sigma^2} x - \frac{1}{2\sigma^2} x^2 - \left( \frac{\mu^2}{2\sigma^2} + \log(\sqrt{2\pi}\sigma) \right) \right\}$$

$$= \exp \left\{ \theta^1 F_1(x) + \theta^2 F_2(x) - \psi(\theta) \right\}. \quad (3)$$
This is an exponential family with $C(x) = 0$, $\theta^1 = \frac{x}{2\sigma^2}$, $\theta^2 = \frac{1}{2\sigma^2}$, $F_1(x) = x$, $F_2(x) = -x^2$. The potential function $\psi$ is given by

$$\psi(\theta) = \frac{\mu^2}{2\sigma^2} + \log(\sqrt{2\pi}\sigma) = \left(\theta^1\right)^2 - \frac{1}{2} \log(\theta^2) + \frac{1}{2} \log \pi$$

The derivatives of log-likelihood function $l_\theta = \log p_\theta$ with respective to parameters are called score functions. Then, the Fisher metric (the Fisher information matrix) is defined as follows:

$$g_{ij} = E_\theta \left[ (\partial_i l_\theta)(\partial_j l_\theta) \right]$$

where $\partial_i := \frac{\partial}{\partial \theta^i}$ and $E_\theta$ denotes the expectation with respect to the distribution $p_\theta$. If $S$ is an exponential family, it is well-known that the Fisher metric is the Hessian of the potential function [1]. In the case of univariate Gaussian distributions, the Fisher metric is

$$g = \sum_{i,j=1}^2 \frac{\partial^2 \psi}{\partial \theta^i \partial \theta^j} d\theta^i d\theta^j = \frac{1}{\sigma^2} (d\mu^2 + 2d\sigma^2).$$

2. Quaternion Gaussian integral

2.1. Real and complex Gaussian integral

For a positive-definite real symmetric matrix $S \in \text{Symm}(N, \mathbb{R})_+$, it is well-known that

$$\int_{-\infty}^{\infty} e^{-\langle x, Sx \rangle} dx = \frac{\pi^{N/2}}{|S|^{1/2}}$$

Similarly, for a positive-definite complex Hermitian matrix $A \in \text{Herm}(N, \mathbb{C})_+$, we have [3]

$$\int_{-\infty}^{\infty} e^{-\langle z, Az \rangle} dxdy = \frac{\pi^N}{|A|^1},$$

where $(z, Az) = z^TAz$ and $z = x + yi$ ($x, y \in \mathbb{R}^N$).

2.2. Quaternion Gaussian integral

We denote $\mathbb{H}^N$ by a set of $N$-component quaternion vectors. We write $x = x_0 + x_1i + x_2j + x_3k = z_1 + z_2j \in \mathbb{H}^N$, where $x_0, x_1, x_2, x_3 \in \mathbb{R}^N$ and $z_1, z_2 \in \mathbb{C}^N$.

For a positive-definite quaternion Hermitian matrix $H = A + Bj \in \text{Herm}(N, \mathbb{H})_+$, where $A^T = A$, $B^T = -B$, we have the following theorem.

**Theorem 1.**

$$\int_{-\infty}^{\infty} e^{-\langle x, Hx \rangle} dx_0 dx_1 dx_2 dx_3 = \frac{\pi^{2N}}{S\text{det}(H)}$$

where $S\text{det}(H)$ denotes the Study determinant of a quaternion matrix $H$ [2];

$$S\text{det}(H) := \left| \begin{array}{cc} A & B \\ -B & A \end{array} \right| = |A||\tilde{A} + BA^{-1}B|$$
Proof.

\[(x, Hx) = \overline{x}^THx = \text{Re}(\overline{x}^THx) = (z_1, Az_1 - B\bar{z}_2) + (\bar{z}_2, \bar{A}\bar{z}_2 + Bz_1)
\]

\[(z_1 - A^{-1}B\bar{z}_2, Az_1 - B\bar{z}_2) + (A^{-1}B\bar{z}_2, A\bar{z}_1) - (A^{-1}B\bar{z}_2, B\bar{z}_2)
\]

\[= (z_1 - A^{-1}B\bar{z}_2, A(z_1 - A^{-1}B\bar{z}_2)) + (B\bar{z}_2, z_1) + (\bar{z}_2, BA^{-1}B\bar{z}_2)
\]

\[= (z_1 - A^{-1}B\bar{z}_2, A(z_1 - A^{-1}B\bar{z}_2)) + (\bar{z}_2, (\bar{A} + BA^{-1}B)\bar{z}_2)
\]

By the formula of the complex case (8), we have

\[
\int_{-\infty}^{\infty} e^{-(x, Hx)}dx_0dx_1dx_2dx_3 = \pi^N |A| \frac{\pi^N}{|A + BA^{-1}B|} = \frac{\pi^N}{|\text{Sdet}(H)|}.
\]

We remark that if

\[A = A_1 + iA_2, B = B_1 + iB_2, \quad A, B \in \mathbb{C}^{N \times N}, \quad A_j, B_j \in \mathbb{R}^{N \times N},\]

we put

\[
\begin{pmatrix}
A & B \\
-B & B
\end{pmatrix} = \begin{pmatrix}
A_1 & B_1 \\
-B_1 & A_1
\end{pmatrix} + i \begin{pmatrix}
A_2 & B_2 \\
B_2 & -A_2
\end{pmatrix} \equiv X + iY.
\]

Then we have

\[
\text{det} \begin{pmatrix}
A & B \\
-B & B
\end{pmatrix} = \left\{ \text{det} \begin{pmatrix}
X & -Y \\
Y & X
\end{pmatrix} \right\}^{\frac{1}{2}}. \quad (9)
\]

For more detail, see Appendix 1. Hence, in consideration to degrees of components of a matrix, for \(H \in \text{Herm}(N, \mathbb{H})_+\), we denote

\[
\text{det} H = (\text{Sdet}(H))^{\frac{1}{2}}. \quad (10)
\]

Remark 1. We can summarize as follows; For \(K = \mathbb{R}, \mathbb{C}, \mathbb{H}\), \(d = \dim_{\mathbb{R}} K = 1, 2, 4\) and \(A \in \text{Herm}(N, K)_+\), we have

\[
\int_{-\infty}^{\infty} e^{-(x, Ax)}dx = \frac{\pi^d N}{(\text{det} A)^{\frac{d}{2}}}. \quad (11)
\]

3. Definition of the quaternion Gaussian distribution

By using theorem 1, we can determine the normalization constant of quaternion Gaussian distribution.

Definition 2. For \((\mu, \Sigma) \in \mathbb{H}^N \times \text{Herm}(N, \mathbb{H})_+\), we define a probability density function of a quaternion Gaussian distribution of a \(N\)-dimensional quaternionic random vector;

\[
p(x; \mu, \Sigma) = \frac{1}{(2\pi)^{2N}(\text{det} \Sigma)^2} \exp\left\{-\frac{1}{2}(\bar{x} - \bar{\mu})^T \Sigma^{-1} (x - \mu)\right\} \quad (x \in \mathbb{H}^N \cong \mathbb{R}^{4N}). \quad (12)
\]
We remark that (12) is a simpler expression than in the previous paper [4].

To obtain the potential function of the quaternion Gaussian distribution, we rewrite (12) as follows:

\[
p(x; \mu, \Sigma) = \frac{1}{(2\pi)^{2N}(\det \Sigma)^{2}} \exp\left\{-\frac{1}{2}(\bar{x} - \bar{\mu})^{T}\Sigma^{-1}(x - \mu)\right\}
\]

\[
= \exp\left\{-\frac{1}{2}\bar{x}^{T}\Sigma^{-1}x + \frac{1}{2}\bar{x}^{T}\Sigma^{-1}\mu + \frac{1}{2}\bar{\mu}^{T}\Sigma^{-1}\mu
\right.
\]

\[
- \left(\frac{1}{2}\bar{\mu}^{T}\Sigma^{-1}\mu + \log((\det \Sigma)^{2}) + 2N \log(2\pi)\right)\}
\]

(13)

\[
= \exp\left\{-\frac{1}{2}\bar{x}^{T}\Sigma^{-1}x - \frac{1}{2}\bar{x}^{T}\Sigma^{-1}\mu + \frac{1}{2}\bar{\mu}^{T}\Sigma^{-1}\mu
\right. 
\]

\[
- \left(\frac{1}{2}\bar{\mu}^{T}\Sigma^{-1}\mu + \log((\det \Sigma)^{2}) + 2N \log(2\pi)\right)\}
\]

(14)

4. The Fisher metric of quaternion Gaussian distributions

For \((\mu, \Sigma) \in H^{N} \times \text{Herm}(N, H)\), let \(M_{N} := \{(\mu, \Sigma)\}\) be the statistical model defined by the collection of the quaternion Gaussian distributions.

According to Yoshizawa-Tanabe [8], we define canonical parameters as follows:

\[
\theta_{H} := \frac{1}{\sqrt{2}}\Sigma^{-1}\mu, \quad \Theta_{H} := \Sigma^{-1}
\]

(15)

We remark that

\[
\frac{1}{2}\bar{\mu}^{T}\Sigma^{-1}\mu + \log((\det \Sigma)^{2}) = \bar{\theta}_{H}^{T}\Theta_{H}^{-1}\theta_{H} - \log((\det \Theta_{H})^{2}).
\]

(16)

Therefore, by (14), the potential function is given by

\[
\psi(\theta_{H}, \Theta_{H}) := \bar{\theta}_{H}^{T}\Theta_{H}^{-1}\theta_{H} - \log((\det \Theta_{H})^{2}) + 2N \log(2\pi).
\]

(17)

Next, we use a quaternion (matrix) derivatives (see Appendix 2). For more detail, see [7], [6].

**Theorem 3.** The Fisher metric of quaternion Gaussian distributions is the Hessian of the potential function (17) with respect to \((\theta_{H}, \Theta_{H})\).

**Proof.** For simplicity, we prove \(N = 1\) case. We put \(\Sigma = \sigma^{2}\) and \(\theta^{1} = \theta_{H}, \theta^{2} = \Theta_{H}\). Then

\[
l_{\theta} = \log p_{\theta} = \frac{1}{2\sigma^{2}}(\bar{x}\mu + \bar{\mu}x) - \frac{1}{2\sigma^{2}}\bar{x}x - \psi(\theta^{1}, \theta^{2}) = \bar{F}_{1}(x)\theta^{1} + \bar{\theta}^{1}F_{1}(x) + \theta^{2}F_{2}(x) - \psi(\theta^{1}, \theta^{2}).
\]

(18)

(19)

By using the quaternion derivatives, we have

\[
\frac{\partial}{\partial \theta^{i}}l_{\theta} = \frac{1}{2}\bar{F}_{1}(x) - \frac{\partial \psi}{\partial \theta^{i}}, \quad \frac{\partial}{\partial \theta^{i}}l_{\theta} = \frac{1}{2}F_{1}(x) - \frac{\partial \psi}{\partial \theta^{i}}, \quad \frac{\partial}{\partial \theta^{2}}l_{\theta} = F_{2}(x) - \frac{\partial \psi}{\partial \theta^{2}},
\]

\[
\frac{\partial^{2}l_{\theta}}{\partial \theta^{i} \partial \theta^{j}} = -\frac{\partial^{2} \psi}{\partial \theta^{i} \partial \theta^{j}}, \quad (i, j = 1, 1, 2).
\]

(20)

(21)

Therefore, we obtain

\[
g_{ij}(\theta) = E_{\theta}\left[\left(\frac{\partial}{\partial \theta^{i}}l_{\theta}\right)\left(\frac{\partial}{\partial \theta^{j}}l_{\theta}\right)\right] = -E_{\theta}\left[\frac{\partial^{2}l_{\theta}}{\partial \theta^{i} \partial \theta^{j}}\right] = \frac{\partial^{2} \psi}{\partial \theta^{i} \partial \theta^{j}}.
\]

(22)
By direct calculations of the Hessian of the potential function (17) with the quaternion derivatives and using the parameters \((\mu, \Sigma)\) again, we obtain the following theorem.

**Theorem 4.** The Fisher metric of \(M_N\) is given by

\[
ds^2 = \frac{1}{4} d\bar{\mu}^T \Sigma^{-1} d\mu + \frac{1}{2} \text{Retr}(\Sigma^{-1} d\Sigma)^2
\]

where \(\text{Retr}(H)\) is the real part of trace of \(H\) and

\[
d\mu = d\mu_0 + id\mu_1 + jd\mu_2 + kd\mu_3, \quad d\bar{\mu}^T = d\mu_0^T - id\mu_1^T - jd\mu_2^T - kd\mu_3^T
\]

In the case of \(N = 1\), we have

\[
ds^2 = \frac{|d\mu|^2 + 8d\sigma^2}{4\sigma^2}.
\]

5. Conclusion and comments

We obtain a unified formula of a Gaussian integral in the case of real, complex and quaternion. By using this, we can express a probability density function of multivariate Gaussian distributions in a simple form even if in the quaternion case. Finally, as an application to a geometry, the Fisher metric of quaternion Gaussian distributions is obtained by using the potential function.

Based on these results described in this paper, we would like to construct a quaternionic information geometry. We will report some results in another paper.

Acknowledgments
The author would like to thank Shintaro Yoshizawa for his helpful comments and suggestions.

Appendix 1

If

\[
A = A_1 + iA_2, \quad B = B_1 + iB_2, \quad A, B \in \mathbb{C}^{N \times N}, \quad A_j, B_j \in \mathbb{R}^{N \times N},
\]

we put

\[
\begin{pmatrix}
A & B \\
-B & B
\end{pmatrix} = \begin{pmatrix}
A_1 & B_1 \\
-B_1 & A_1
\end{pmatrix} + i \begin{pmatrix}
A_2 & B_2 \\
B_2 & -A_2
\end{pmatrix} \equiv X + iY.
\]

Then we have

\[
det \begin{pmatrix}
X & -Y \\
Y & X
\end{pmatrix} = det \begin{pmatrix}
E & iE \\
O & E
\end{pmatrix} \begin{pmatrix}
X & -Y \\
Y & X
\end{pmatrix} \begin{pmatrix}
E & -iE \\
O & E
\end{pmatrix} \quad (E: \text{unit matrix})
\]

\[
= det \begin{pmatrix}
X + iy & O \\
Y & X - iy
\end{pmatrix}
\]

\[
= det(X + iy) \det(X - iy)
\]

\[
= |det(X + iy)|^2 = \left| \det \begin{pmatrix}
A & B \\
-B & B
\end{pmatrix} \right|^2.
\]

In the case of \(H = A + Bj\) is a positive-definite quaternion Hermitian matrix, \(\det \begin{pmatrix}
A & B \\
-B & B
\end{pmatrix}\) is a positive real number. Therefore we obtain

\[
det \begin{pmatrix}
A & B \\
-B & B
\end{pmatrix} = \left( \det \begin{pmatrix}
X & -Y \\
Y & X
\end{pmatrix} \right)^{\frac{1}{2}} = \left\{ \det \begin{pmatrix}
A_1 & B_1 & -A_2 & -B_2 \\
-A_1 & B_1 & A_2 & -B_2 \\
A_2 & B_2 & A_1 & B_2 \\
B_2 & -A_2 & -B_1 & A_1
\end{pmatrix} \right\}^{\frac{1}{2}}.
\]
Appendix 2
We introduce quaternion derivatives in [7] within the minimum essentials. For quaternion matrix derivatives, the see [6].

For quaternions $\mathbf{H} = \{ q = q_a + q_i + q_j + q_k| q_a, q_i, q_j, q_k \in \mathbb{R} \}$, the quaternion conjugate $q^*$ and the quaternion rotation $q^\mu$ are defined by

$$ q^* := q_a - q_i - q_j - q_k $$

$$ q^\mu := \mu q \mu^{-1} \quad \mu \neq 0 $$

“HR calculus” is introduced in Mandic et al. [5](2011) as an hypercomplex extension of well-known CR calculus. Moreover, “GHR calculus” (generalized HR calculus) is developed.

Definition 5 (The GHR Derivatives). The left GHR derivatives of a real-differentiable function $f : \mathbf{H} \to \mathbf{H}$ wrt $q^\mu$ and $q^{\mu\alpha}$ ($\mu \neq 0, \mu \in \mathbf{H}$) are defined as

$$ \frac{\partial f}{\partial q^\mu} := \frac{1}{4} \left( \frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b} i^\mu - \frac{\partial f}{\partial q_c} j^\mu - \frac{\partial f}{\partial q_d} k^\mu \right) $$

$$ \frac{\partial f}{\partial q^{\mu\alpha}} := \frac{1}{4} \left( \frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b} i^\mu + \frac{\partial f}{\partial q_c} j^\mu + \frac{\partial f}{\partial q_d} k^\mu \right) $$

Especially, in the case of $\mu = 1$,

$$ \frac{\partial f}{\partial q} = \frac{1}{4} \left( \frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b} i - \frac{\partial f}{\partial q_c} j - \frac{\partial f}{\partial q_d} k \right) $$

$$ \frac{\partial f}{\partial q^*} = \frac{1}{4} \left( \frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b} i + \frac{\partial f}{\partial q_c} j + \frac{\partial f}{\partial q_d} k \right) $$

Then, we find

$$ \frac{\partial q}{\partial q} = \frac{1}{4} \left( \frac{\partial q_a}{\partial q_a} - i \frac{\partial q_b}{\partial q_b} - j \frac{\partial q_c}{\partial q_c} - k \frac{\partial q_d}{\partial q_d} \right) = \frac{1}{4} \left( 1 - i \cdot i - j \cdot j - k \cdot k \right) = 1 $$

We remark a following important relation in $\mathbf{H}$. This is a big difference to the complex case.

$$ q^* = -\frac{1}{2} (q + iqj + jqi + kqk) $$

Then, we have

$$ \frac{\partial q^*}{\partial q} = \frac{1}{4} \left( \frac{\partial q_a^*}{\partial q_a} - i \frac{\partial q_b^*}{\partial q_b} - j \frac{\partial q_c^*}{\partial q_c} - k \frac{\partial q_d^*}{\partial q_d} \right) = \frac{1}{4} \left( 1 - i \cdot (-i) - j \cdot (-j) - k \cdot (-k) \right) = -\frac{1}{2} $$

$$ f(q) = |q|^2 = q^* q, \quad \frac{\partial |q|^2}{\partial q} = \frac{1}{4} \left( 2q_a - i \cdot 2q_b - j \cdot 2q_c - k \cdot 2q_d \right) = \frac{1}{2} q^*. $$

We need the GHR calculus because of the noncommutativity of the quaternions.

$$ \frac{\partial q_{a}^{\beta}}{\partial q_{b}^{\beta}} = \frac{1}{4} \left( \frac{\partial q_{a}^{\beta}}{\partial q_{a}^{\beta}} - \frac{\partial q_{a}^{\beta}}{\partial q_{b}^{\beta}} - \frac{\partial q_{a}^{\beta}}{\partial q_{c}^{\beta}} - \frac{\partial q_{a}^{\beta}}{\partial q_{d}^{\beta}} \right) $$

$$ = \frac{1}{4} \left( 1 \cdot \beta - (-i) \cdot \beta i - (-j) \cdot \beta j - (-k) \cdot \beta k \right) = -\frac{1}{2} \beta^*. $$

Lemma 6. For constants $\alpha, \beta$, we have

$$ \frac{\partial (\alpha f \beta)}{\partial q^\mu} = \alpha \frac{\partial f}{\partial q^\mu} \beta. $$

Proposition 7. We have a following Leibniz formula:

$$ \frac{\partial (fg)}{\partial q} = f \frac{\partial g}{\partial q} + \frac{\partial (fg)}{\partial q} \bigg|_{g=const} = f \frac{\partial g}{\partial q} + \frac{\partial f}{\partial q^\mu} g. $$
References
[1] Amari S and Nagaoka H 2000 Methods of Information Geometry (New York: American Mathematical Society and Oxford University Press)
[2] Aslaksen H 1996 Quaternionic determinants The Mathematical Intelligencer 18 57-65
[3] Bellman R 1959 Representation theorem and inequalities for hermitian matrices Duke. Math. J. 26 485-490
[4] Loots M 2010 The development of the quaternion normal distribution Thesis at Univ. of Pretoria
[5] Mandic D P et al 2011 A quaternion gradient operator and its applications IEEE Signal Process. Let. 18 47-50
[6] Xu D and Mandic D P 2015 The theory of quaternion matrix derivatives IEEE trans. Sig. Proc. 63 1543-1556
[7] Xu D et al 2015 Enabling quaternion derivatives: the generalized HR calculus Royal Soc. Open Sci. 2 150255
[8] Yoshizawa S and Tanabe K 1999 Dual differential geometry associated with the Kullback-Leibler information on the Gaussian distributions and its 2-parameter deformations SUT J. of Math. 35 113-137