The Matrix Angular Central Gaussian Distribution

YASUKO CHIKUSE

Kagawa University, Kagawa 760, Japan

Communicated by the Editors

The Riemann space whose elements are \( m \times k \) (\( m \geq k \)) matrices \( X \) such that \( X'X = I_k \) is called the Stiefel manifold and denoted by \( V_{k,m} \). Some distributions on \( V_{k,m} \), e.g., the matrix Langevin (or von Mises-Fisher) and Bingham distributions and the uniform distribution, have been defined and discussed in the literature. In this paper, we present methods to construct new kinds of distributions on \( V_{k,m} \) and discuss some properties of these distributions. We investigate distributions of the "orientation" \( H_Z = Z(Z'Z)^{-1/2} (\in V_{k,m}) \) of an \( m \times k \) random matrix \( Z \). The general integral form of the density of \( H_Z \) reduces to a simple mathematical form, when \( Z \) has the matrix-variate central normal distribution with parameter \( \Sigma \), an \( m \times m \) positive definite matrix. We may call this distribution the matrix angular central Gaussian distribution with parameter \( \Sigma \), denoted by the MACG (\( \Sigma \)) distribution. The MACG distribution reduces to the angular central Gaussian distribution on the hypersphere for \( k = 1 \), which has been already known. Then, we are concerned with distributions of the orientation \( H_Y \) of a linear transformation \( Y = BZ \) of \( Z \), where \( B \) is an \( m \times m \) matrix such that \( |B| \neq 0 \). Utilizing properties of these distributions, we propose a general family of distributions of \( Z \) such that \( H_Z \) has the MACG (\( \Sigma \)) distribution.

1. INTRODUCTION

The Riemann space whose elements are \( m \times k \) (\( m \geq k \)) matrices \( X \) such that \( X'X = I_k \), the \( k \times k \) identity matrix, is called the Stiefel manifold, and denoted by \( V_{k,m} \). For \( k = m \), the Stiefel manifold is the orthogonal group \( O(m) \). Practical examples of data on \( V_{k,m} \) are illustrated by Downs [5] in vectorcardiography and by Jupp and Mardia [11] in astronomy.

Received February 16, 1989; revised August 21, 1989.

AMS 1980 subject classifications: primary 62H10, 62E10; secondary 58C35.

Key words and phrases: Stiefel manifolds, orientation of a random matrix, matrix angular Gaussian distributions, matrix-variate normal distributions, matrix elliptically symmetric distributions.
An invariant measure on $V_{k,m}$ is given by the differential form
\[
(X' dX) = \bigwedge_{i<j}^k x_j' \, dx_j \bigwedge_{j=1}^{m-k} x_{k+j} \, dx_j,
\]
in terms of the exterior products ($\bigwedge$), where we choose an $m \times (m-k)$ matrix $X_1$ such that $(X: X_1) = (x_1, \ldots, x_k; x_{k+1}, \ldots, x_m) \in O(m)$ and $dx$ is an
$m \times 1$ vector of differentials. See, e.g., Muirhead [13] for the use of exterior products. The volume of $V_{k,m}$ is
\[
w(k, m) = \int_{V_{k,m}} (X' dX) = 2^k \pi^{km/2} / \Gamma_k(m/2),
\]
where $\Gamma_k(a) = \pi^{k(k-1)/4} \prod_{i=1}^k \Gamma(a - (i-1)/2)$. Let $[dX]$ denote the normalized invariant measure ($(X' dX)/w(k, m)$) of unit mass on $V_{k,m}$. A detailed discussion of manifolds and their invariant measures may be found in James [9] and Farrell [6, Chaps. 6–8].

Some distributions on $V_{k,m}$ have been defined and discussed in the literature. The matrix Langevin (or von Mises–Fisher) distribution has the density proportional to $etr (F'X)$, with $F$ an $m \times k$ matrix, with respect to $[dX]$ (Downs [5]), and is a uni-modal distribution with a modal orientation around which the distribution is rotational symmetric. See Chikuse [2], Downs [5], Jupp and Mardia [11], and Khatri and Mardia [12] for further discussions of distribution theory and inference of this distribution. The matrix Bingham distribution has the density proportional to $etr (X'AX)$, with $A$ an $m \times m$ symmetric matrix, with respect to $[dX]$, and is an antipodally symmetric distribution. Here, since $tr(XX') = tr(X'X) = k$, a restriction is imposed on $A$ to ensure the identifiability of $A$, e.g., $tr A = 0$. See Bingham [1] (for $k=1$), Jupp and Mardia [11], and Prentice [14] for a detailed discussion of this distribution. The case $F=0$ or $A=0$ in the above distributions reduces to the uniform distribution $[dX]$ on $V_{k,m}$.

It is noted here that, throughout this paper, the density of a random matrix $X$ on $V_{k,m}$ is expressed with respect to the normalized measure $[dX]$, while the density of a $q \times k$ random matrix $Y=(y_{ij})$ is expressed with respect to the measure $(dY)$. Here we denote
\[
(dY) \equiv \bigwedge_{j=1}^q \bigwedge_{i=1}^k dy_{ij} \\
\equiv \bigwedge_{1 \leq i \leq j \leq k} dy_{ij}, \quad \text{if } Y \text{ is } k \times k \text{ symmetric}.
\]

In this paper, we present methods to construct new kinds of distributions on $V_{k,m}$ and discuss some properties of these distributions. For an $m \times k$
nonrandom matrix $A$ of rank $k$ ($m \geq k$), we define the unique polar decomposition of $A$ as

$$A = H_A T_A^{1/2}, \quad \text{with} \quad H_A = A(\check{A}'\check{A})^{-1/2} \quad \text{and} \quad T_A = \check{A}'\check{A},$$

so that $H_A \in V_{k,m}$, where $T_A^{1/2}$ denotes the unique positive definite square root of the $k \times k$ positive definite matrix $T_A$. Let $Z$ be an $m \times k$ random matrix. Then, $Z$ is of rank $k$ almost everywhere, and hence, extending the above definition to random matrices, we may write the "unique" polar decomposition of $Z$ as

$$Z = H_Z T_Z^{1/2}, \quad \text{with} \quad H_Z = Z(\check{Z}'\check{Z})^{-1/2} \quad \text{and} \quad T_Z = \check{Z}'\check{Z}. \quad (1.1)$$

$H_Z$ indicates the "orientation" of the random matrix $Z$, and we are interested in the distribution of $H_Z$. The orientation of a random matrix reduces to the direction of a random vector for $k = 1$. The product matrix $T_Z = Z'Z$ indicates the inner products of the columns of $Z$ and may also be of importance in statistical inference of orientational data. In particular, when $Z$ is the sum of a random sample of size $n$, $T_Z$ plays an important role. Although our main interest of this paper is to investigate distributions of $H_Z$, we may give some distributional results on $T_Z$ when they are worth being noted (see also Chikuse [2]).

In Section 2, we investigate distributions of the orientation $H_Z$. The general integral form of the density of $H_Z$ is presented. When $Z$ has the matrix-variate central normal distribution with parameter $\Sigma$, an $m \times m$ positive definite matrix, the density of $H_Z$ is written as (2.5). We may call this distribution the "matrix angular central Gaussian distribution" with parameter $\Sigma$, denoted by the MACG ($\Sigma$) distribution. The MACG distribution reduces to the angular central Gaussian distribution of a directional random vector on the hypersphere in $R^n$ for $k = 1$. The MACG distributions have statistical properties similar to those for $k = 1$. It is noted that a further discussion on the MACG distributions, including limit theorems, may be found in Chikuse [3].

In connection with the discussion of distributions of the orientation of a random matrix, we present a theorem which may be useful in multivariate distribution theory. The well-known characterization of the matrix-variate standard normal distribution (see James [9, Section 8]) is a special case of the theorem.

Section 3 is concerned with distributions of the orientation $H_Y$ of a linear transformation $Y = BZ$ of $Z$, where $B$ is an $m \times m$ matrix such that $|B| \neq 0$. The general form of the density of $H_Y$ is presented in terms of the density function of $H_Z$. It is shown that the property for $H_Z$ to be MACG is maintained for $H_Y$ after the linear transformation. Then, we propose a general family of distributions of a random matrix $Z$ such that $H_Z$ has the MACG
268 YASUKO CHIKUSE

(Σ) distribution. It is seen that the matrix elliptically symmetric distribution and hence the matrix-variate central normal distribution belong to this family of distributions.

2. DISTRIBUTIONS OF THE ORIENTATION OF A RANDOM MATRIX

Let Z be an \( m \times k \) random matrix having the density \( f_Z(Z) \). Let us write the unique polar decomposition of Z as in (1.1) with the orientation \( H_Z \in V_{k,m} \) and the product matrix \( T_Z \). We shall investigate the “marginal” distribution of \( H_Z \) (and also that of \( T_Z \)).

The following lemma, which is essentially due to Herz [8, Lemma 1.4] (see also James [9, identity (8.19)], and Muirhead [13, Theorem 2.1.14]), is useful for the derivation of the distribution of \( H_Z \) and \( T_Z \).

**LEMMA 2.1.** The measure \((dZ)\) is decomposed as

\[
(dZ) = c \mid T_Z \mid^{(m-k-1)/2} [dH_Z](dT_Z),
\]

where

\[
c = \pi^{km/2}/\Gamma_k(m/2).
\]

Utilizing Lemma 2.1, we establish

**THEOREM 2.1.** We have the density \( f_{H_Z}(H_Z) \) of \( H_Z \),

\[
f_{H_Z}(H_Z) = c \int_{T>0} f_Z(H_Z T^{1/2}) \mid T \mid^{(m-k-1)/2} (dT),
\]

and the density \( f_{T_Z}(T_Z) \) of \( T_Z \),

\[
f_{T_Z}(T_Z) = c \mid T_Z \mid^{(m-k-1)/2} \int_{V_{k,m}} f_Z(H T_Z^{1/2})[dH],
\]

where \( c \) is given by (2.2) and the integral in (2.3) is over the space of \( k \times k \) positive definite matrices.

Now, let us consider the case when \( Z \) is normally distributed. In this case, we can evaluate the integrals in (2.3) and (2.4) in compact mathematical forms.

**THEOREM 2.2.** Suppose that \( Z \) has the \( m \times k \) matrix-variate central normal distribution with parameter \( \Sigma \), whose density is

\[
f_Z(Z) = (2\pi)^{-km/2} \mid \Sigma \mid^{-k/2} \text{etr}(-\frac{1}{2}Z^T \Sigma^{-1}Z),
\]
where $\Sigma$ is an $m \times m$ positive definite matrix. Then, we have the density of $H_Z$

$$f_{H_Z}(H_Z) = |\Sigma|^{-k/2} |H'_Z \Sigma^{-1} H_Z|^{-m/2},$$  

(2.5)

and the density of $T_Z$

$$f_{T_Z}(T_Z) = [2^{km/2} \Gamma_k(m/2)|\Sigma|^{k/2}]^{-1} e_0^{(k)}(\Sigma^{-1}, -\frac{1}{2} T_Z) |T_Z|^{(m-k-1)/2}. \quad (2.6)$$

Here, $e_0^{(r)}(S_1, S_2)$ is the hypergeometric function with $p_1 \times p_1$ and $p_2 \times p_2$ symmetric matrix arguments $S_1$ and $S_2$, respectively, which has a series expansion

$$\sum_{l=0}^{\infty} \sum_{\lambda} C_{\lambda}(S_1) C_{\lambda}(S_2)/l! C_\lambda(I_p),$$

in terms of the zonal polynomials $C_{\lambda}(S_1)$ and $C_{\lambda}(S_2)$, where $\lambda$ runs over the ordered partitions of the integer $l$ into not more than $r$ parts ($r \leq \min(p_1, p_2)$, $p = \max(p_1, p_2)$). The zonal polynomial $C_{\lambda}(S)$ is a homogeneous symmetric polynomial of degree $l$ in the latent roots of a symmetric matrix $S$. See, e.g., James [10] and Constantine [4] for a detailed discussion of the zonal polynomials and hypergeometric function with matrix arguments.

**Proof.** The proof of (2.5) is straightforward by utilizing the integral definition of $\Gamma_k(a)$. Substituting the normal density of $Z$ into (2.4), and then making the transformation $H \to Q_1 H, Q_1 \in O(m)$ and integrating over $Q_1 \in O(m)$ yields

$$f_{T_Z}(T_Z) = c(2\pi)^{-km/2} |\Sigma|^{-k/2}$$

$$\times |T_Z|^{(m-k-1)/2} \int_{V_{k,m}} \int_{O(m)} \text{etr}(-\frac{1}{2} H'Q_1 \Sigma^{-1} Q_1 HT_Z)[dQ_1] [dH]$$

$$= [2^{km/2} \Gamma_k(m/2)|\Sigma|^{k/2}]^{-1}$$

$$\times |T_Z|^{(m-k-1)/2} \int_{V_{k,m}} e_0^{(m)}(\Sigma^{-1}, -\frac{1}{2} HT_Z H') [dH]$$

$$= [2^{km/2} \Gamma_k(m/2)|\Sigma|^{k/2}]^{-1}$$

$$\times |T_Z|^{(m-k-1)/2} e_0^{(k)}(\Sigma^{-1}, -\frac{1}{2} T_Z).$$

It is noted that the result (2.6) has been obtained as the distribution of a quadratic form, by a slightly different approach, though (e.g., Hayakawa [7, Theorem 1]).
We note that, when $\Sigma = I_m$ in Theorem 2.2, $H_Z$ is uniformly distributed on $V_{k,m}$ and $T_Z$ has the central Wishart distribution $W_k(m, I_k)$.

The special case $k = 1$ of (2.5) has been noticed by several authors (e.g., Watson [16, Section 3.6]), and has been called the angular central Gaussian distribution on the hypersphere $V_{1,m}$ in $R^m$. This distribution may be denoted by the ACG ($\Sigma$) distribution for the rest of this paper. The ACG distribution is an alternative to the Bingham distribution (Bingham [11]) for modeling antipodally symmetric directional data on $V_{1,m}$, and its statistical theory has been discussed by Tyler [15].

We may call the distribution having the density of the form (2.5) the "matrix angular central Gaussian distribution" with parameter $\Sigma$ on the Stiefel manifold $V_{k,m}$, which may be denoted by the MACG ($\Sigma$) distribution. This distribution has statistical properties similar to those of the ACG ($\Sigma$) distribution for $k = 1$. The case $\Sigma = I_m$ gives the uniform distribution on $V_{k,m}$. There is an indeterminancy in the parameter matrix $\Sigma$ by multiplication by a positive scalar. The distribution (2.5) is invariant under the transformation $H_Z \rightarrow H_Z Q_2$, for any $Q_2 \in O(k)$.

Let $X_1, \ldots, X_n$ be a random sample of size $n$ from the MACG ($\Sigma$) distribution. Then, it is shown by a method, essentially due to Tyler [15], utilizing the Taylor's expansion in matrix arguments that the maximum likelihood estimate $\hat{\Sigma}$ of $\Sigma$ is the solution of the equation

$$\hat{\Sigma} = \left[\frac{m}{kn}\right] \sum_{i=1}^{n} X_i (X_i' \hat{\Sigma}^{-1} X_i)^{-1} X_i'.$$

Before closing this section, we present a theorem which may be useful in multivariate distribution theory.

**Theorem 2.3.** An $m \times k$ random matrix $Z$ has the density of the form $g(Z'Z)$, a function of $Z'Z$, if and only if the following three conditions are satisfied:

(i) $H_Z$ and $T_Z$ are independent,

(ii) $H_Z$ is uniformly distributed on $V_{k,m}$, and

(iii) $T_Z$ has the density of the form

$$f_{T_Z}(T_Z) = c g(T_Z) |T_Z|^{(m-k-1)/2}.$$

where $c$ is given by (2.2).

**Proof.** The proof is straightforward by utilizing Lemma 2.1.

Theorem 2.3 may be useful to characterize multivariate distributions of matrix variates by specifying certain forms of the density (2.8) of $T_Z$. The well-known characterization of the matrix variate standard normal dis-
distribution (see James [9, Section 8]) is given by the above conditions (i), (ii), and (iii) with \( f_{T_{Z}}(T_{Z}) \) being the density of the Wishart distribution \( W_{k}(m, I_{k}) \), i.e., \( g(T_{Z}) = (2\pi)^{-km/2} \text{etr}(-\frac{1}{2}T_{Z}) \).

3. DISTRIBUTIONS OF THE ORIENTATION OF A LINEAR TRANSFORMATION

It may be interesting to see the relationship between the distributions of the orientation of a random matrix \( Z \) and that of a linear transformation of \( Z \).

**Theorem 3.1.** Let \( Z \) be an \( m \times k \) random matrix with the density \( f_{Z}(Z) \) which is assumed to be invariant under the transformation \( Z \rightarrow ZQ_{2} \), for any \( Q_{2} \in O(k) \). We consider a new \( m \times k \) random matrix \( Y = BZ \), where \( B \) is an \( m \times m \) matrix such that \( |B| \neq 0 \). Let us write

\[
Z = H_{Z} T_{Z}^{1/2}, \quad \text{with} \quad H_{Z} = Z(Z'Z)^{-1/2} \quad \text{and} \quad T_{Z} = Z'Z,
\]

and

\[
Y = H_{Y} T_{Y}^{1/2}, \quad \text{with} \quad H_{Y} = Y(Y'Y)^{-1/2} \quad \text{and} \quad T_{Y} = Y'Y,
\]

and let \( f_{H_{Z}}(H_{Z}) \) be the density of \( H_{Z} \) (which is determined from \( f_{Z}(Z) \) by Theorem 2.1.). Then, the density \( f_{H_{Y}}(H_{Y}) \) of \( H_{Y} \), the orientation of the new random matrix \( Y \), is given by

\[
f_{H_{Y}}(H_{Y}) = |B|^{-k} |W^{'W}|^{-m/2} f_{H_{Z}}(H_{w}), \quad (3.1)
\]

where \( W = B^{-1}H_{Y} \) and \( H_{w} = W(W^{'W})^{-1/2} \).

**Proof.** Since the Jacobian of \( Z \rightarrow BZ \) is \( |B|^{k} \), \( Y \) has the density

\[
f_{Y}(Y) = |B|^{-k} f_{Z}(B^{-1}Y). \quad (3.2)
\]

Substituting (3.2) into (2.3) gives the density \( f_{H_{Y}}(H_{Y}) \) of \( H_{Y} \)

\[
f_{H_{Y}}(H_{Y}) = c |B|^{-k} \int_{T > 0} f_{Z}(B^{-1}H_{Y} T^{1/2}) |T|^{(m - k - 1)/2} (dT). \quad (3.3)
\]

Now we make the transformation, in (3.3),

\[
T = (W^{'W})^{-1/2} S(W^{'W})^{-1/2}, \quad \text{with} \quad W = B^{-1}H_{Y}. \quad (3.4)
\]

Then, the Jacobian of \( T \rightarrow (W^{'W})^{-1/2} S(W^{'W})^{-1/2} \) is \( |W^{'W}|^{-(k + 1)/2} \). From (3.4) and the assumption of the invariance of the density \( f_{Z}(Z) \), we have

\[
f_{Z}(WT^{1/2}) = f_{Z}(H_{w} S^{1/2}).
\]
Hence, (3.3) can be written as

\[ f_{H_Y}(H_Y) = c \left| B \right|^{-k} \left| W'W \right|^{-m/2} \int_{S > 0} f_Z(H_WS^{1/2}) \left| S \right|^{(m-k-1)/2} (dS). \]  (3.5)

In view of Theorem 2.1, (3.5) leads to our desired result (3.1).

It is noted that, for \( k = 1 \), (3.1) reduces to Watson [16, Eq. (3.6.4)].

It is seen that the property for the orientation of a random matrix \( Z \) to be MACG is maintained for that of the linear transformation of \( Z \). In more detail, we have

**Corollary 3.1.1.** We assume the condition of Theorem 3.1.

1. If \( H_Z \) has the MACG (\( \Sigma \)) distribution, then \( H_Y \), with \( Y = BZ \), has the MACG (\( B\Sigma B' \)) distribution.
2. In particular, if \( H_Z \) is uniformly distributed on \( V_{k,m} \) (i.e., \( H_Z \) has the MACG (\( I_m \)) distribution), then \( H_Y \) has the MACG (\( BB' \)) distribution.
3. If \( H_Z \) has the MACG (\( \Sigma \)) distribution and \( B \) is chosen such that \( B'B = \Sigma^{-1} \), then \( H_Y \) is uniformly distributed on \( V_{k,m} \).

It is known for \( k = 1 \) (e.g., Tyler [15]) that, if an \( m \times 1 \) random vector \( Z \) has the elliptically symmetric distribution whose density is of the form \( \left| \Sigma \right|^{-1/2} g(Z'\Sigma^{-1}Z) \), with \( \Sigma \) an \( m \times m \) positive definite matrix, then the distribution of its direction \( (Z'Z)^{-1/2} Z \) is ACG (\( \Sigma \)) on the hypersphere in \( R^m \). This fact can be readily extended to our general case \( k \geq 1 \) on the Stiefel manifold. That is, if an \( m \times k \) random matrix \( Z \) has the matrix elliptically symmetric distribution whose density is of the form \( \left| \Sigma \right|^{-k/2} g(\text{tr} Z'\Sigma^{-1}Z) \), then its orientation \( H_Z \) has the MACG (\( \Sigma \)) distribution. It is now to be furthermore generalized. The following theorem proposes a more general family of distributions of a random matrix such that its orientation has the MACG (\( \Sigma \)) distribution.

**Theorem 3.2.** Suppose that an \( m \times k \) random matrix \( Z \) has the density of the form

\[ f_Z(Z) = \left| \Sigma \right|^{-k/2} g(Z'\Sigma^{-1}Z) \]  (3.6)

which is invariant under the transformation \( Z \to ZQ_2 \), for any \( Q_2 \in O(k) \), where \( \Sigma \) is an \( m \times m \) positive definite matrix. Then its orientation \( H_Z \) has the MACG (\( \Sigma \)) distribution.

**Proof.** Let us choose an \( m \times k \) matrix \( B \) such that \( BB' = \Sigma \) and \( |B| \neq 0 \), and put \( U = B^{-1}Z \). Then, the density of \( U \) is \( g(U'U) \), which is also invariant under the transformation \( U \to UQ_2 \), for any \( Q_2 \in O(k) \). From
Theorem 2.3, the orientation $H_U$ of $U$ is uniformly distributed on $V_{k,m}$, and hence, in view of (ii) of Corollary 3.1.1, $H_Z$ has the MACG ($\Sigma (=BB')$) distribution.

It is noticed that the $m \times k$ matrix-variate central normal distribution (see Theorem 2.2) and, in more general terms, the matrix elliptically symmetric distribution belong to the general family of distributions in Theorem 3.2. The densities of distributions in this family would be, in general, certain functions of the latent roots of $Z'\Sigma^{-1}Z$.

We have seen that the condition (3.6) with the invariance under the transformation $Z \rightarrow ZQ_2$, for any $Q_2 \in O(k)$, is a sufficient condition on a random matrix $Z$ for $H_Z$ to have the MACG($\Sigma$) distribution. The same condition is, in fact, necessary and sufficient for a set of conditions including that $H_Z$ has the MACG ($\Sigma$) distribution. That is, we have

**Corollary 3.2.1.** We assume that an $m \times k$ random matrix $Z$ has the density $f_Z(Z)$ which is invariant under the transformation $Z \rightarrow ZQ_2$, for any $Q_2 \in O(k)$. Then $f_Z(Z)$ is of the form

$$f_Z(Z) = |\Sigma|^{-k/2} g(Z'\Sigma^{-1}Z),$$

with $\Sigma$ an $m \times m$ positive definite matrix, if and only if the following three conditions are satisfied:

1. $H_Z$ has the MACG ($\Sigma$) distribution,
2. $H_U$ and $T_U$ are independent, and
3. the density $f_{T_U}(T_U)$ of $T_U$ is of the form

$$f_{T_U}(T_U) = c |T_U|^{m-k-1/2} g(T_U),$$

where $c$ is given by (2.2). Here, we put $U = B^{-1}Z$ for an $m \times m$ matrix $B$ such that $BB' = \Sigma$ and write $U = H_UT_U^{1/2}$, with $H_U = U(U'U)^{-1/2}$ and $T_U = U'U$.

**Proof.** The proof is straightforward from that of Theorem 3.2 by utilizing Theorem 2.3 and Corollary 3.1.1.

**Acknowledgment**

The author is very grateful to the referee for valuable comments and suggestions which led to an improved revision of this paper.
REFERENCES