

HOKKAIDO UNIVERSITY

Title	Geodesics of multivariate normal distributions and a Toda lattice type Lax pair
Author(s)	Kobayashi, Shimpei
Citation	Physica Scripta, 98(11) https://doi.org/10.1088/1402-4896/ad0087
Issue Date	2023-11-01
Doc URL	https://hdl.handle.net/2115/93713
Rights	This is the Accepted Manuscript version of an article accepted for publication in Physica Scripta. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The Version of Record is available online at http://dx.doi.org/10.1088/1402-4896/ad0087
Rights(URL)	https://creativecommons.org/licenses/by-nc-nd/4.0/
Type	journal article
File Information	2023-September-19-revised-Geodesics-normal-dist.pdf



Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

GEODESICS OF MULTIVARIATE NORMAL DISTRIBUTIONS AND A TODA LATTICE TYPE LAX PAIR

SHIMPEI KOBAYASHI

Abstract. We study geodesics of multivariate normal distributions with respect to the Fisher metric. First it will be shown that a computational formula for geodesics can be understood using the block Cholesky decomposition and a natural Riemannian submersion. Next a mid point algorithm for geodesics will be obtained. And finally a new Toda lattice type Lax pair will be derived from the geodesic and the block Cholesky decomposition.

1. Introduction and the main results

Due to the invariance property under sufficient statistics [1], it is important to consider a parametric statistical model, that is a family of probability distributions with a finite number of parameters, as a Riemannian manifold with specific metric, the so-called *Fisher metric*:

$$g(X,Y) = E_\theta[(X \log p(x,\theta))(Y \log p(x,\theta))], \quad E_\theta[f] = \int_{\mathbb{R}^n} f(x)p(x,\theta)\,dx,$$

where $p(x,\theta)$ denotes a family of probability density functions of $x \in \mathbb{R}^n$, and X, Y are tangent vectors at a point θ of the manifold. Fundamental parametric models are given by the model of normal distributions and its multivariate generalization, the *multivariate normal distributions model*, that is, the probability density function is given by

$$p\left(x,(\Sigma,\mu)\right) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right),$$

where $x \in \mathbb{R}^n$, and $\mu \in \mathbb{R}^n$ is called the *mean vector* and Σ is called the *covariance matrix*, which is in order n positive definite matrices, $\mathrm{Sym}^+(n,\mathbb{R})$. In this paper, the $\frac{1}{2}n(n+3)$-dimensional manifold $\mathcal{N} = \left\{(\Sigma,\mu) \mid \Sigma \in \mathrm{Sym}^+(n,\mathbb{R}),\ \mu \in \mathbb{R}^n\right\}$ with the Fisher metric g will be called the *multivariate normal manifold* $(\mathcal{N}, g)$. See [18] for Riemannian geometry of $(\mathcal{N}, g)$, and note that it has been shown [12, 5, 6] that $(\mathcal{N}, g)$ has a solvable Lie group structure.

It is a fundamental problem to compute geodesics of a Riemannian manifold. For the multivariate normal manifold $(\mathcal{N}, g)$, it was first given by Eriksen [4] as follows: First we realize

Date: September 23, 2023.

2020 *Mathematics Subject Classification.* Primary 53C35, Seondary 62B10, 17B80.

Key words and phrases. Multivariate normal distributions; geodesics; symmetric spaces; Riemannian submersions; Toda lattice.

The author is partially supported by Kakenhi 22K03304.

the multivariate normal manifold $\mathcal{N}$ as

$$\mathcal{N} = \left\{ \begin{pmatrix} \Theta & \delta \\ \delta^T & 1+\|\delta\|^2_{\Theta^{-1}} \end{pmatrix} \;\middle|\; \begin{array}{c} \Theta = \Sigma^{-1} \in \mathrm{Sym}^+(n,\mathbb{R}), \\ \delta = \Sigma^{-1}\mu \in \mathbb{R}^n,\ \|\delta\|^2_{\Theta^{-1}} = \delta^T\Theta^{-1}\delta \end{array} \right\} \subset \mathrm{Sym}^+(n+1,\mathbb{R}), \tag{1.1}$$

and the geodesic equation (integrated once) can be formulated as

$$\frac{d}{dt}\begin{pmatrix} \Theta & \delta \end{pmatrix} = \begin{pmatrix} -A_0 & a_0 \end{pmatrix} \begin{pmatrix} \Theta & \delta \\ \delta^T & 1+\|\delta\|^2_{\Theta^{-1}} \end{pmatrix},$$

where $A_0 \in \mathrm{Sym}(n,\mathbb{R})$ and $a_0 \in \mathbb{R}^n$ are some constant. Second, to construct a solution of the geodesic equation, let

$$G(t) = \exp(tV) \quad \text{with} \quad V = \begin{pmatrix} -A_0 & a_0 & 0 \\ a_0^T & 0 & -a_0^T \\ 0 & -a_0 & A_0 \end{pmatrix}, \quad \text{and} \quad J = \begin{pmatrix} & & \mathrm{id}_n \\ & \mathrm{id}_1 & \\ \mathrm{id}_n & & \end{pmatrix}, \tag{1.2}$$

that is, $V \in \mathrm{Sym}(2n+1,\mathbb{R})$ with $\operatorname{tr} V = 0$. Since G is positive definite with determinant 1 and satisfies a symmetry relation $JG^{-1}J = G$, the block Cholesky decomposition of $G = MDM^T$, see Lemma 2.1, is given by

$$M = \begin{pmatrix} \mathrm{id}_n & 0 & 0 \\ \delta^T\Theta^{-1} & 1 & 0 \\ * & -\Theta^{-1}\delta & \mathrm{id}_n \end{pmatrix}, \quad D = \begin{pmatrix} \Theta & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \Theta^{-1} \end{pmatrix}, \tag{1.3}$$

where id_n is the identity matrix of order n, $\delta : \mathbb{R} \to \mathbb{R}^n$ and $\Theta : \mathbb{R} \to \mathrm{Sym}^+(n,\mathbb{R})$ and $*$ denote some $n \times n$ matrix. Finally from the form of $G(t)$ given by (1.3), the sub-matrix of $G(t)$ by the first $(n+1)$-rows and the first $(n+1)$-columns defines a curve in $(\mathcal{N}, g)$ and it is in fact a geodesic in $\mathcal{N}$. Note that in [4] the formula was derived by a direct computation without using the block Cholesky decomposition, and a more explicit formula of the geodesic was derived in [2]. In Section 2.1, we will review the construction using the block Cholesky decomposition in details.

We now explain the above result more geometrically: The idea is that the curve $G(t)$ in (1.2) seems to define a geodesic in some homogeneous space, and the projection to the sub-matrix could be understood geometrically. First the curve $G(t)$ is in $\mathcal{L} = \mathrm{Sym}^+_*(2n+1,\mathbb{R})$, where $\mathrm{Sym}^+_*(2n+1,\mathbb{R})$ denotes the set of positive definite matrices with determinant 1 and $\mathcal{L}$ is a Riemannian symmetric space of type AI:

$$\mathcal{L} = \mathrm{SL}(2n+1,\mathbb{R})/\,\mathrm{SO}(2n+1),$$

and thus $G(t)$ is a homogeneous geodesic in $\mathcal{L}$. Moreover, there is a non-compact Riemannian symmetric space of type BDI:

$$\mathcal{M} = \mathrm{SO}(n+1,n)/\,\mathrm{S}(\mathrm{O}(n+1)\times\mathrm{O}(n)) \subset \mathcal{L} = \mathrm{Sym}^+_*(2n+1,\mathbb{R}), \tag{1.4}$$

which is a totally geodesic submanifold of $\mathcal{L}$. Note that $\mathcal{M}$ is in fact a reflective submanifold (a special totally geodesic submanifold) of $\mathcal{L}$, see [14]. It is easy to see that a point $m \in \mathcal{L} = \mathrm{Sym}^+_*(2n+1,\mathbb{R})$ belongs to $\mathcal{M}$ if and only if m satisfies the symmetry $Jm^{-1}J = m$, and the geodesic $G(t)$ in $\mathcal{L} = \mathrm{Sym}^+_*(2n+1,\mathbb{R})$ satisfies the symmetry and thus it is a geodesic of $\mathcal{M}$. Since $\mathcal{N}$ can be realized in $\mathrm{Sym}^+(n+1,\mathbb{R})$ as in (1.1), the projection from $\mathcal{M}$ to the sub-matrix by the first $(n+1)$-rows and the first $(n+1)$-columns can be understood as the Riemannian submersion $\pi : \mathcal{M} \to \mathcal{N}$. Finally, the geodesic $G(t)$ in $\mathcal{M}$ is in fact

horizontal, that is, a tangent vector of $G(t)$ at any point $p \in \mathcal{M}$ is orthogonal to the fiber $\pi^{-1}(\pi(p))$, and a standard argument about Riemannian submersions shows the following geometric characterization:

Theorem A. *Any geodesic in the multivariate normal manifold $(\mathcal{N}, g)$ can be obtained by the Riemannian submersion of a horizontal geodesic of the Riemannian symmetric space $\mathcal{M}$ in* (1.4)*.*

The proof of Theorem A will be given in Section 2.2 below.

Remark 1.1. A part of the result in Theorem A had been obtained by a direct computation in [10, Chapter 4]. In this paper we give a geometric proof using the fundamental result of Riemannian submersions [17, 8].

For the above characterization, it is important that two Riemannian symmetric spaces $\mathcal{L} = \mathrm{Sym}^+_*(2n+1,\mathbb{R}) = \mathrm{SL}(2n+1,\mathbb{R})/\mathrm{SO}(2n+1)$ and $\mathcal{M} = \mathrm{SO}(n+1,n)/\mathrm{S}(\mathrm{O}(n+1)\times\mathrm{O}(n))$ are given in the space of positive definite matrices with determinant 1 $\mathrm{Sym}^+_*(2n+1,\mathbb{R})$. Moreover since $\mathcal{M}$ is embedded totally geodesically in $\mathcal{L}$, the horizontal geodesic $G(t)$ can be thought as a geodesic in $\mathrm{Sym}^+_*(2n+1,\mathbb{R})$. On the one hand, Nakamura [16] had given an algorithm to obtain the mid point of a geodesic segment connecting two points p and q in the space of positive definite matrices, the so-called *AHM algorithm* (Arithmetic and harmonic mean algorithm). Thus applying the AHM algorithm to the horizontal geodesic of $\mathcal{M} \subset \mathcal{L}$ and by using the projection $\pi : \mathcal{M} \to \mathcal{N}$, we obtain the following result:

Theorem B. *Let $p, q \in \mathcal{N}$ and γ be distinct points and the geodesic segment connecting p and q, respectively. Moreover, let r be the midpoint on γ. Then there exist two sequences $\{p_n \in \mathcal{N}\}_{n=0,1,2,\dots}$ and $\{q_n \in \mathcal{N}\}_{n=0,1,2,\dots}$ with $p_0 = p$ and $q_0 = q$ such that $\{p_n\}_{n=0,1,2,\dots}$ and $\{q_n\}_{n=0,1,2,\dots}$ both converge to the mid point r of γ.*

The proof of Theorem B will be given in Section 2.3 below.

Remark 1.2.

(1) By Theorem A and a simple computation, the geodesic segment $\gamma \subset \mathcal{N}$ can be obtained by a geodesic segment in $\mathcal{M}$

$$\tilde{\gamma}(t) = P^{1/2}\left(P^{-1/2}QP^{-1/2}\right)^t P^{1/2}, \quad (t \in [0,1])$$

such that $\gamma = \pi \circ \tilde{\gamma}$ and $\pi P = p, \pi Q = q$. Note that lifting points P and Q in $\mathcal{M}$ are not unique, but one of them is chosen, then the other point is uniquely determined. For an example, if $p = \mathrm{id}$ then $P = \mathrm{id}$ is a natural choice and then Q will be uniquely determined.

(2) It is not known a closed formula for $\tilde{\gamma}(t)$, but for $t = 1/2$, that is the *geometric mean* of two matrices P and Q, the AHM algorithm can be applied (see Section 2.3), and for example it is faster than Heron's method [16, p.171]. Using Theorem B repeatedly, any point on the geodesic segment γ can be efficiently computed.

In the above construction, the block Cholesky decomposition has a key role. As in (1.3), for any $G \in \mathrm{Sym}^+_*(2n+1,\mathbb{R})$ there exist unique matrices M and D such that $G = MDM^T$

holds, where M is a block lower triangular matrix with identity block diagonal matrices and D is a block diagonal matrix, respectively. Setting $G_1 = MD$ and $G_2 = M^T$, we have

$$G = G_1 G_2 \quad \text{for any } G \in \mathrm{Sym}_*^+(2n+1, \mathbb{R}), \tag{1.5}$$

where G_1 is the unique block lower triangular matrix and G_2 is the unique block upper triangular matrix with identity block diagonal matrices, respectively, see Lemma 2.1 for the decomposition $G = MDM^T$ in detail.

Since we have a horizontal geodesic $G(t) = \exp(tV)$ and the block Cholesky dedomposition, it is natural to expect some integrable system (Lax pair) associate to them, see [7, Chapter 8]. Then the main result of the paper is the following:

Theorem C. *Let $G(t) = \exp(tV)$ be the horizontal geodesic in (1.2) of the symmetric space $\mathcal{M}$ through the identity and set $G(t) = G_1(t)G_2(t)$ to be the block Cholesky decomposition as in (1.5). Then the adjoint orbit*

$$L(t) = \mathrm{Ad}(G_1(t)^{-1})V : \mathbb{R} \to \mathfrak{so}(n+1, n), \tag{1.6}$$

defines the following system of ODEs (Lax equation) in the Lie algebra $\mathfrak{so}(n+1, n)$:

$$\dot{L} = [L, M], \quad L = \begin{pmatrix} -Q & r & 0 \\ a_0^T & 0 & -r^T \\ 0 & -a_0 & Q^T \end{pmatrix}, \quad M = \begin{pmatrix} -Q & 0 & 0 \\ a_0^T & 0 & 0 \\ 0 & -a_0 & Q^T \end{pmatrix}, \tag{1.7}$$

where $r : \mathbb{R} \to \mathbb{R}^n, Q : \mathbb{R} \to \mathrm{M}(n \times n, \mathbb{R})$, and $a_0 \in \mathbb{R}^n$ is the constant vector defined in (1.2). More explicitly (1.7) is equivalent to the following system of ODEs for r and Q:

$$\dot{Q} = -ra_0^T, \quad \dot{r} = Qr, \tag{1.8}$$

and it is further equivalent to

$$2\dot{Q} = Q^2 - A_0^2 - 2a_0 a_0^T, \quad \dot{r} = Qr, \tag{1.9}$$

with the initial condition $Q(0) = A_0$ and $r(0) = a_0$.

Conversely, any solution of (1.7) with the initial condition $L(0) = V$ given in (1.2) can be obtained by (1.6).

Remark 1.3. For $n = 1$, the Lax equation becomes the finite non-periodic Toda lattice [7, Chapters 5, 8], and thus (1.7) gives a Toda lattice type Lax pair. The construction of the Toda lattice like Lax pair is similar to the Kostant-Toda lattices construction given in [13]. However, the author would not know a precise relation.

The proof of Theorem C will be given in Section 3 below.

2. Geodesics and AHM algorithm

2.1. **An explicit construction of geodesics.** It is known [18, Theorem 6.1] that the Riemannian geodesic equations for $\mathcal{N} = (\Sigma, \mu)$ can be formulated as

$$\ddot{\Sigma} + \dot{\mu}\dot{\mu}^T - \dot{\Sigma}\Sigma^{-1}\dot{\Sigma} = 0, \quad \ddot{\mu} + \dot{\Sigma}\Sigma^{-1}\dot{\mu} = 0, \tag{2.1}$$

where dot denotes derivative with respect to a parameter t. From the second equation, $\frac{d}{dt}(\Sigma^{-1}\dot{\mu}) = 0$ holds, that is, there exists some constant $a_0 \in \mathbb{R}^n$ such that $\Sigma^{-1}\dot{\mu} = a_0$ holds.

Moreover from the first equation, $\frac{d}{dt}(\Sigma^{-1}\dot{\Sigma} + a_0\mu^T) = 0$ holds, that is, there exists some constant $A_0 \in \mathrm{Sym}(n,\mathbb{R})$ such that $\Sigma^{-1}\dot{\Sigma} + a_0\mu^T = A_0$ holds. Thus the system (2.1) can be integrated by an initial point $(\mathrm{id}_p, 0)$ and an initial direction $(A_0, a_0) \in \mathrm{Sym}(n,\mathbb{R}) \times \mathbb{R}^n$. Introducing

$$\Theta = \Sigma^{-1} \in \mathrm{Sym}^+(n,\mathbb{R}) \quad \text{and} \quad \delta = \Sigma^{-1}\mu \in \mathbb{R}^n,$$

the geodesic equations in (2.1) can be formulated as

$$\frac{d}{dt}\begin{pmatrix}\Theta & \delta\end{pmatrix} = \begin{pmatrix}-A_0 & a_0\end{pmatrix}\begin{pmatrix}\Theta & \delta \\ \delta^T & 1+\|\delta\|^2_{\Theta^{-1}}\end{pmatrix}, \quad \Theta(0) = \mathrm{id},\ \delta(0) = 0, \tag{2.2}$$

where $\|\delta\|^2_{\Theta^{-1}} = \delta^T\Theta^{-1}\delta$. From the form of (2.2) it is natural to expect that it could be solved by exponential of (A_0, a_0). Unfortunately, the simplest matrix exponential $\exp\left\{t\begin{pmatrix}-A_0 & a_0 \\ a_0 & 0\end{pmatrix}\right\}$ is not a solution of (2.2) by the term $1+\|\delta\|^2_{\Theta^{-1}}$ of the right hand side of (2.2). To understand the term $1 + \|\delta\|^2_{\Theta^{-1}}$, let us consider the block Cholesky decomposition of the matrix in the right hand side of (2.2):

$$\begin{pmatrix}\Theta & \delta \\ \delta^T & 1+\|\delta\|^2_{\Theta^{-1}}\end{pmatrix} = \begin{pmatrix}\mathrm{id} & 0 \\ \delta^T\Theta^{-1} & 1\end{pmatrix}\begin{pmatrix}\Theta & 0 \\ 0 & 1\end{pmatrix}\begin{pmatrix}\mathrm{id} & \Theta^{-1}\delta \\ 0 & 1\end{pmatrix}.$$

Then the second term in the right-hand side has the special property, that is, the $(n+1, n+1)$-entry is just 1. Thus we would like to find a exponential matrix which has such a middle term property in a larger matrix: Let

$$J = \begin{pmatrix} & & \mathrm{id}_n \\ & \mathrm{id}_1 & \\ \mathrm{id}_n & & \end{pmatrix}, \tag{2.3}$$

where id_j denotes the identity matrix of degree j. In fact the following lemma gives an answer in order $2n+1$ positive definite matrices with determinant 1, $\mathrm{Sym}^+_*(2n+1,\mathbb{R})$:

Lemma 2.1 (Lemma 7 in [11]). *Let $G \in \mathrm{Sym}^+_*(2n+1,\mathbb{R})$ with symmetry $JG^{-1}J = G$, where J is in* (2.3) *and consider the block Cholesky decomposition of*

$$G = MDM^T, \quad M \in \mathcal{L}, D \in \mathcal{A}$$

where

$$\mathcal{L} = \left\{ \begin{pmatrix}\mathrm{id}_n & 0 & 0 \\ L_{21} & 1 & 0 \\ L_{31} & L_{32} & \mathrm{id}_n\end{pmatrix} \in \mathrm{GL}(2n+1,\mathbb{R}) \right\},$$

$$\mathcal{A} = \left\{ \begin{pmatrix}A_1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & A_1^{-1}\end{pmatrix} \,\middle|\, A_1 \in \mathrm{Sym}^+(n,\mathbb{R}) \right\}.$$

Then it can be computed as

$$M = \begin{pmatrix}\mathrm{id}_n & 0 & 0 \\ \tilde{\delta}^T\tilde{\Theta}^{-1} & 1 & 0 \\ * & -\tilde{\Theta}^{-1}\tilde{\delta} & \mathrm{id}_n\end{pmatrix}, \quad D = \begin{pmatrix}\tilde{\Theta} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \tilde{\Theta}^{-1}\end{pmatrix}, \tag{2.4}$$

where $\tilde{\Theta} \in \mathrm{Sym}^+(n,\mathbb{R})$ and $\tilde{\delta} \in \mathbb{R}^n$ are matrix valued functions of t, and $*$ denotes some $n \times n$ matrix.

Proof. The symmetry $JG^{-1}J = G$ implies the forms of L and D in (2.4). □

Remark 2.2. A proof of existence of the block Cholesky decomposition can be found in [11, Lemma 5].

Since the matrix V in (1.2) has a symmetry $-JVJ = V$, $G(t) = \exp(tV)$ in (1.2) has the symmetry $JG^{-1}J = G$. Therefore the block lower triangular matrix M and the block diagonal matrix D have the forms as in (2.4). Let H be a sub-matrix of G given by first $n+1$ rows and first $n+1$ columns. From the forms of M and D in (2.4), it can be computed as

$$H = \pi(G) = \begin{pmatrix} \tilde{\Theta} & \tilde{\delta} \\ \tilde{\delta}^T & 1 + \|\tilde{\delta}\|^2_{\tilde{\Theta}^{-1}} \end{pmatrix}, \tag{2.5}$$

and it is easy to see that H is a Riemannian geodesic in $\mathcal{N}$.

Theorem 2.3 ([4])**.** *The sub-matrix $H(t)$ of $G(t)$ defines a Riemannian geodesic through the initial point $(\mathrm{id}_p, 0)$ and the direction (A_0, a_0).*

Proof. The matrix valued function G satisfies the following system of ODEs:

$$\dot{G} = VG, \quad G(0) = \mathrm{id}\,.$$

where V is given in (1.2). Thus the sub-matrix H in (2.5) satisfies the same system of the geodesic equation (2.2) for (Θ, δ) and the uniqueness of ODEs implies that $\Theta = \tilde{\Theta}$ and $\delta = \tilde{\delta}$. □

Remark 2.4. The embedding of $\mathcal{N}$ into $\mathrm{Sym}^+(n+1, \mathbb{R})$ as in (1.1) had been considered in [3, 15]. Note that in there $(\mu, \Sigma) \in (\mathbb{R}^n, \mathrm{Sym}^+(n, \mathbb{R}))$ had been embedded as

$$\begin{pmatrix} \Sigma + \mu\mu^T & -\mu \\ -\mu^T & 1 \end{pmatrix} \subset \mathrm{Sym}^+(n+1, \mathbb{R}).$$

Our embedding is the inverse of it, that is,

$$\begin{pmatrix} \Sigma + \mu\mu^T & -\mu \\ -\mu^T & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \Theta & \delta \\ \delta^T & 1 + \|\delta\|^2_{\Theta^{-1}} \end{pmatrix}.$$

Geodesics of the multivariate normal manifold had been first computed by [19] for $n = 1$. In fact for $n = 1$, the multivariate normal manifold becomes the two-dimensional hyperbolic space $\mathbb{H}^2$.

2.2. **Riemannian submersion.** We would like to understand geometrically the construction of the previous section. First it is clear that $\mathrm{SL}(2n+1, \mathbb{R})$ acts on $\mathrm{Sym}^+_*(2n+1, \mathbb{R})$ by

$$g \mapsto gPg^T, \quad g \in \mathrm{SL}(2n+1, \mathbb{R}), \quad P \in \mathrm{Sym}^+_*(2n+1, \mathbb{R}). \tag{2.6}$$

Then the stabilizer at identity is given by

$$\mathrm{Fix}_\tau(\mathrm{SL}(2n+1, \mathbb{R})) = \{g \in \mathrm{SL}(2n+1, \mathbb{R}) \mid \tau(g) = g\} = \mathrm{SO}(2n+1),$$

where the involution τ is given by

$$\tau(g) = (g^T)^{-1}, \quad g \in \mathrm{SL}(2n+1, \mathbb{R}). \tag{2.7}$$

Thus $\mathrm{Sym}^+_*(2n+1,\mathbb{R})$ can be realized as a non-compact Riemannian symmetric space of type AI:

$$\mathcal{L} = \mathrm{SL}(2n+1,\mathbb{R})/\,\mathrm{SO}(2n+1).$$

By using the involution in $Jg^{-1}J$ for $\mathrm{Sym}^+_*(2n+1,\mathbb{R})$ as in Lemma 2.1 and the action in (2.6) one can induce an involution of $\mathrm{SL}(2n+1,\mathbb{R})$ as follows:

$$\sigma(g) = Jg^{-T}J, \quad g \in \mathrm{SL}(2n+1,\mathbb{R}), \tag{2.8}$$

and the fixed point set of the involution σ can be computed as

$$\mathrm{Fix}_\sigma(\mathrm{SL}(2n+1,\mathbb{R})) = \{g \in \mathrm{SL}(2n+1,\mathbb{R}) \mid \sigma(g) = g\} = \mathrm{SO}(n+1,n),$$

see (2.9) for the Lie algebra computation. Since σ in (2.8) and τ in (2.7) commute, one can consider a fixed point set of τ on $\mathrm{Fix}_\sigma(\mathrm{SL}(2n+1,\mathbb{R})) = \mathrm{SO}(n+1,n)$, which is $\mathrm{S}(\mathrm{O}(n+1)\times \mathrm{O}(n))$, and the corresponding homogeneous space is given by

$$\mathcal{M} = \mathrm{SO}(n+1,n)/\,\mathrm{S}(\mathrm{O}(n+1)\times\mathrm{O}(n)) \subset \mathcal{L} = \mathrm{Sym}^+_*(2n+1,\mathbb{R}),$$

which is an another non-compact Riemannian symmetric space type BDI. Note that a standard reference for symmetric spaces is [9].

On the Lie algebra level, $\mathfrak{g} = \mathfrak{so}(n+1,n)$ can be realized in $\mathfrak{sl}(2n+1,\mathbb{R})$ as follows: Let σ be an involution on $\mathfrak{sl}(2n+1,\mathbb{R})$ given by

$$\sigma(X) = -JX^TJ, \quad X \in \mathfrak{sl}(2n+1,\mathbb{R}), \tag{2.9}$$

where J is defined in (2.3). Note that σ is a derivative of the involution for $\mathrm{SL}(2n+1,\mathbb{R})$ in (2.3) and $J^2 = \mathrm{id}_{2n+1}$ holds. The eigenvalues of σ are 1 with multiplicity $n+1$ and -1 with multiplicity n. Thus the fixed point subalgebra $\mathfrak{g} = \mathrm{Fix}_\sigma(\mathfrak{sl}(2n+1,\mathbb{R}))$ of σ is isomorphic to $\mathfrak{so}(n+1,n)$:

$$\begin{aligned}\mathfrak{g} &= \left\{ \begin{pmatrix} -Q & r & R \\ t^T & 0 & -r^T \\ S & -t & Q^T \end{pmatrix} \mid Q \in \mathrm{M}(n\times n,\mathbb{R}),\ R,S \in \mathrm{Skew}(n,\mathbb{R}),\ r,t \in \mathbb{R}^n \right\} \\ &\cong \mathfrak{so}(n+1,n).\end{aligned} \tag{2.10}$$

Moreover, let τ be an involution on $\mathfrak{g} = \mathfrak{so}(n+1,n)$:

$$\tau(X) = -X^T, \quad X \in \mathfrak{so}(n+1,n).$$

Then the fixed point subalgebra $\mathfrak{k} = \mathrm{Fix}_\tau(\mathfrak{g})$ can be computed as

$$\begin{aligned}\mathfrak{k} &= \left\{ \begin{pmatrix} -Q & r & R \\ -r^T & 0 & -r^T \\ R & r & Q^T \end{pmatrix} \mid Q,R \in \mathrm{Skew}(n,\mathbb{R}),\ r \in \mathbb{R}^n \right\} \\ &\cong \mathfrak{o}(n+1)\times\mathfrak{o}(n).\end{aligned} \tag{2.11}$$

The compliment subspace $\mathfrak{m}$ can be computed as

$$\mathfrak{m} = \left\{ \begin{pmatrix} -Q & r & R \\ r^T & 0 & -r^T \\ -R & -r & Q \end{pmatrix} \mid Q \in \mathrm{Sym}(n,\mathbb{R}),\ R \in \mathrm{Skew}(n,\mathbb{R}),\ r \in \mathbb{R}^n \right\}. \tag{2.12}$$

Note that $\mathfrak{m} \subset \mathrm{Sym}(2n+1,\mathbb{R})$. Thus by (2.10), (2.11) and (2.12), we have

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{m},$$

and $[\mathfrak{k},\mathfrak{k}] \subset \mathfrak{k}$, $[\mathfrak{m},\mathfrak{m}] \subset \mathfrak{k}$ and $[\mathfrak{m},\mathfrak{k}] \subset \mathfrak{m}$ hold. It is evident that the tangent space of $\mathcal{M}$ at the identity can be represented by $\mathfrak{m}$ in (2.12), see [9] in details.

The Riemannian metric of $\mathcal{M}$ at identity can be induced from a Killing metric on $\mathcal{L} = \mathrm{Sym}_*^+(2n+1,\mathbb{R}) = \mathrm{SL}(2n+1,\mathbb{R})/\mathrm{SO}(2n+1)$:

$$\langle A, B\rangle = \mathrm{Tr}(AB) \quad \text{for } A, B \in \mathrm{Sym}(2n+1,\mathbb{R}).$$

Moreover, any geodesic in $\mathcal{M}$ through identity is given by $\tilde{G}(t) = \exp(t\tilde{V})$ for some $\tilde{V} \in \mathfrak{m}$, and thus $G(t)$ in (1.2) is a geodesic in $\mathcal{M}$. From Lemma 2.1, we know that an element $m \in \mathcal{M} \subset \mathrm{Sym}^+(2n+1,\mathbb{R})$ can be represented by

$$m = \begin{pmatrix} \Theta & \delta & G_{13} \\ \delta^T & 1+\|\delta\|^2_{\Theta^{-1}} & G_{23} \\ G_{13}^T & G_{23}^T & G_{33} \end{pmatrix} \in \mathcal{M} \subset \mathrm{Sym}_*^+(2n+1,\mathbb{R}), \tag{2.13}$$

where $\Theta \in \mathrm{Sym}^+(n,\mathbb{R}), \delta \in \mathbb{R}^n, G_{13} \in \mathrm{M}(n\times n,\mathbb{R}), G_{23} \in \mathrm{M}(n\times 1,\mathbb{R}), G_{33} \in \mathrm{Sym}(n,\mathbb{R})$. Since the multivariate normal manifold $\mathcal{N} = \left\{(\Sigma,\mu) \mid \Sigma \in \mathrm{Sym}^+(n,\mathbb{R}),\ \mu \in \mathbb{R}^n\right\}$ can be identified by (1.1), thus $\pi : \mathcal{M} \to \mathcal{N}$ is naturally defined for $m \in \mathcal{M}$ in (2.13) by

$$\pi(m) = \begin{pmatrix} \Theta & \delta \\ \delta^T & 1+\|\delta\|^2_{\Theta^{-1}} \end{pmatrix} \in \mathcal{N},$$

and it is a Riemannian submersion: $d\pi|_{\mathrm{id}} : T_{\mathrm{id}}\mathcal{M} \to T_{\mathrm{id}}\mathcal{N}$ with

$$T_{\mathrm{id}}\mathcal{M} \ni \begin{pmatrix} -Q & r & R \\ r^T & 0 & -r^T \\ -R & -r & Q \end{pmatrix} \mapsto \begin{pmatrix} -Q & r \\ r^T & 0 \end{pmatrix} \in T_{\mathrm{id}}\mathcal{N}.$$

Note that the Riemannian metric of $\mathcal{N}$ at identity can be computed as

$$\langle A, B\rangle = 2\,\mathrm{Tr}(AB) \quad \text{for } A, B \in T_{\mathrm{id}}\mathcal{N} \subset \mathrm{Sym}(n+1,\mathbb{R}),$$

see [15, 6]. For a Riemannian submersion $\pi : \mathcal{M} \to \mathcal{N}$, a tangent vector to $\mathcal{N}$ at p is *horizontal* if it is orthogonal to the fiber $\pi^{-1}(\pi(p))$. Then the horizontal subspace $\mathfrak{h} \subset \mathfrak{m}$ can be identified with

$$\mathfrak{h} = \left\{ X = \begin{pmatrix} -Q & r & 0 \\ r^T & 0 & -r^T \\ 0 & -r & Q \end{pmatrix} \in \mathfrak{m} \right\},$$

and thus the geodesic $G(t)$ in $\mathcal{M}$ given in (1.2) is in fact horizontal at identity, that is, $V \in \mathfrak{h}$. Then the following fact about Riemannian submersions is fundamental:

Theorem 2.5 (Corollary 2 in [17])**.** *Let $\pi : \mathcal{M} \to \mathcal{N}$ be a Riemannian submersion. If a geodesic γ of $\mathcal{M}$ that is horizontal at some one point, γ is always horizontal (hence $\pi \circ \gamma$ is a geodesics of $\mathcal{N}$).*

Remark 2.6. The result first was proved by Hermann [8] by using length-minimizing properties and O'Neill [17] gave a simpler proof.

The geodesic γ in Theorem 2.5 is called the *horizontal geodesic*, and thus the curve $G(t)$ in (1.2) is the horizontal geodesic in $\mathcal{M}$.

Proof of Theorem A. From Theorem 2.5, we understand the curve $G(t)$ is a horizontal geodesic in $\mathcal{M}$ and any geodesic in $\mathcal{N}$ can be obtained by the Riemannian submersion of a horizontal geodesic $G(t)$. This completes the proof. □

2.3. **Arithmetic harmonic mean algorithm.** It is known that $\mathcal{M}$ is a totally geodesic submanifold in $\mathcal{L} = \mathrm{Sym}^+_*(2n+1,\mathbb{R})$, see [14]. Since geodesics in $\mathcal{N}$ are obtained by the projection of horizontal geodesics in $\mathcal{M}$, therefore one can think $G(t)$ as a geodesic in $\mathcal{L} = \mathrm{Sym}^+_*(2n+1,\mathbb{R})$ and use it to compute geodesics in $\mathcal{N}$. In [16, Theorem 10], Nakamura gave an algorithm to obtain the midpoint R_0 of the Riemannian geodesic from P_0 and Q_0 in the space of positive definite matrices. First note that the midpoint R_0 of P_0 and Q_0 is given by

$$R_0 = P_0^{1/2}\left(P_0^{-1/2}Q_0P_0^{-1/2}\right)^{1/2}P_0^{1/2},$$

and in particular if $P_0 = \mathrm{id}_{2n+1}$, then $R_0 = Q_0^{1/2}$ holds. For $n = 0, 1, 2, \ldots$, define

$$P_{n+1} = \frac{1}{2}(P_n + Q_n), \tag{2.14}$$

$$Q_{n+1} = 2(P_n^{-1} + Q_n^{-1})^{-1}. \tag{2.15}$$

Then the sequences of matrices $\{P_n\}_{n=0,1,2,\ldots}$ and $\{Q_n\}_{n=0,1,2,\ldots}$ defined by (2.14) and (2.15) are all positive definite.

Theorem 2.7 (Theorem 10 in [16])**.** *The sequences $\{P_n\}_{n=0,1,2,\ldots}$ and $\{Q_n\}_{n=0,1,2,\ldots}$ tend to the midpoint $R_0 = P_0^{1/2}(P_0^{-1/2}Q_0P_0^{-1/2})^{1/2}P_0^{1/2}$ of the geodesic segment connecting P_0 and Q_0 in a quadratic order.*

Remark 2.8. Since $\{P_n\}_{n=0,1,2,\ldots}$ and $\{Q_n\}_{n=0,1,2,\ldots}$ in Theorem 2.7 satisfy

$$Q_{n+1} - P_{n+1} = \frac{1}{2}(Q_n - P_n)^2(Q_n + P_n)^{-1},$$

see [16, p.171], thus they converge to R_0 in a "quadratic order".

We now prepare the following Lemma.

Lemma 2.9. *Let $Q_0, P_0 \in \mathrm{Sym}^+_*(2n+1,\mathbb{R})$. Then the sequences of matrices $\{P_n\}_{n=0,1,2,\ldots}$ and $\{Q_n\}_{n=0,1,2,\ldots}$ given by* (2.14) *and* (2.15) *also take values in* $\mathrm{Sym}^+_*(2n+1,\mathbb{R})$.

Proof. By the constructions in (2.14) and (2.15), it is clear that Q_n and P_n take values in $\mathrm{Sym}^+(2n+1,\mathbb{R})$. We show that Q_n and P_n have determinant 1.

For the case of $Q_0 = \mathrm{id}$: By Lemma 7 in [16], $Q_nP_n = P_nQ_n$ for $n = 0, 1, 2, \ldots$ and therefore Q_n and P_n can be simultaneously diagonalized and if they have determinant 1 then Q_{n+1} and P_{n+1} also have determinant 1.

For the general case of Q_0: By (30) in [16], the the sequences of matrices $\{P'_n\}_{n=0,1,2,\ldots}$ and $\{Q'_n\}_{n=0,1,2,\ldots}$ are introduced by

$$P'_n = Q_0^{-1/2}P_nQ_0^{1/2}, \quad Q'_n = Q_0^{-1/2}Q_nQ_0^{1/2}, \quad n = 0, 1, 2, \ldots,$$

and clearly $Q'_0 = \mathrm{id}$. Thus Q'_n and P'_n are determinant 1 as before, and thus Q_n and P_n are determinant 1 as well. □

The algorithm defined in (2.14) and (2.15) together with Lemma 2.9 will give an algorithm for the mid point of geodesic segment in $\mathcal{N}$:

The proof of Theorem B. Let $p, q \in \mathcal{N}$ be distinct points and the geodesic segment γ connecting p and q. Let P_0 and Q_0 be distinct points and the horizontal geodesic $G(t)$ connecting P_0 and Q_0 in $\mathcal{M}$ such that $\pi(P_0) = p, \pi(Q_0) = q$ and $\pi(G(t)) = \gamma$. Since $\mathcal{M} \subset \mathcal{L} = \mathrm{Sym}^+_*(2n+1, \mathbb{R})$ is totally geodesic, the P_0, Q_0 and G can be points and a geodesic in $\mathrm{Sym}^+_*(2n+1, \mathbb{R})$. Now applying Theorem 2.7, one obtains sequences of matrices $\{P_n\}_{n=0,1,2,\dots}$ and $\{Q_n\}_{n=0,1,2,\dots}$ such that they converge to the midpoint

$$R_0 = P_0^{1/2}(P_0^{-1/2} Q_0 P_0^{-1/2})^{1/2} P_0^{1/2}$$

of $G(t)$ in a quadratic order. Moreover by Lemma 2.9, $\{P_n\}_{n=0,1,2,\dots}$ and $\{Q_n\}_{n=0,1,2,\dots}$ take values in $\mathrm{Sym}^+_*(2n+1, \mathbb{R})$. Then the Riemannian submersion $\pi : \mathcal{M} \to \mathcal{N}$ gives a corresponding sequences of points $\{p_n = \pi(P_n) \in \mathcal{N}\}_{n=0,1,2,\dots}$ and $\{q_n = \pi(Q_n) \in \mathcal{N}\}_{n=0,1,2,\dots}$ such that they converge to the midpoint $r = \pi(R_0)$ of the geodesic segment $\gamma \subset \mathcal{N}$ connecting $p, q \in \mathcal{N}$. □

3. Toda lattice type Lax pair

Let us rephrase the block Cholesky decomposition of $G \in \mathrm{SO}(n+1,n) \subset \mathrm{Sym}^+_*(2n+1, \mathbb{R})$ in Lemma 2.1 as follows:

$$G = G_1 G_2, \quad G_1 = MD, \quad G_2 = M^T, \tag{3.1}$$

where M is the block lower triangular matrix with trivial diagonal and D is the block diagonal matrix. More explicitly,

$$M = \begin{pmatrix} \mathrm{id}_n & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & -m_{21}^T & \mathrm{id}_n \end{pmatrix}, \quad D = \begin{pmatrix} d_{11} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & d_{11}^{-1} \end{pmatrix},$$

where m_{21} is a row vector with length n, d_{11} is a order n positive definite symmetric matrix and m_{31} is a $n \times n$ matrix. Note that from the symmetry $JG^{-1}J = G$, m_{31} and m_{21} satisfy a symmetry relation

$$m_{31} + m_{31}^T = -m_{21} m_{21}^T.$$

Let us denote the projection to the first factor of the above decomposition by π_1. Accordingly, the Lie algebra $\mathfrak{g} = \mathfrak{so}(n+1, n)$ of $\mathrm{SO}(n+1,n)$ can be decomposed as

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2,$$

that is,

$$\pi_1 X = \begin{pmatrix} -Q & 0 & 0 \\ t^T & 0 & 0 \\ S & t & Q^T \end{pmatrix} \quad \text{for} \quad X = \begin{pmatrix} -Q & r & R \\ t^T & 0 & -r^T \\ S & -t & Q^T \end{pmatrix} \in \mathfrak{g} = \mathfrak{so}(n+1, n).$$

Let V be the matrix in (1.2) and the horizontal geodesic $G(t) = \exp(tV)$, and consider the block Cholesky decomposition $G(t) = G_1(t) G_2(t)$ by (3.1). Note that since $G(t)$ takes values in $\mathrm{Sym}^+_*(2n+1, \mathbb{R})$, the decomposition is defined for all $t \in \mathbb{R}$. Then consider an adjoint orbit

$$L(t) = \mathrm{Ad}(G_1(t)^{-1}) V : \mathbb{R} \to \mathfrak{g} = \mathfrak{so}(n+1, n).$$

We are now ready to prove the main result.

Proof of Theorem C. We first show that $L(t)$ has the form in (1.7). Since $G_1(t)$ is a block lower triangular matrix, it is easy to see that the upper right part, that is, the sub-matrix given by the first n-rows and the last n-columns, is zero. On the other hand, $L(t)$ can be rephrased as

$$L(t) = \mathrm{Ad}(G_1(t)^{-1})V = \mathrm{Ad}(G_2(t)G(t)^{-1})V = \mathrm{Ad}(G_2(t))V,$$

since $G(t) = \exp(tV)$ commutes with V. Obviously, $G_2(t)$ is the block upper triangular and has a trivial block diagonal part, and therefore the lower left part of L, that is, the sub-matrix given by the last n-rows and the first n-columns, is zero. Moreover a straightforward computation shows that the $(n+1, n+1)$ entry is zero and the sub-matrices given by $n+1$-th row and $n+1$-th column are constants a_0^T and $-a_0$, respectively.

Finally explicit systems of ODEs in (1.8) and (1.9) are given by straightforward computations. □

The author states that there is no conflict of interest. Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Acknowledgements. We would like to express our gratitude for the anonymous referee's comments on the manuscript, particularly for their valuable input in the substantial reformulation of Theorem B.

References

[1] N. N. Čencov. Statistical Decision Rules and Optimal Inference. *Translations of mathematical monographs*, American Mathematical Society **53**, 1981.

[2] M. Calvo, J. M. Oller. An explicit solution of information geodesic equations for the multivariate model. *Statistics & Decisions* **9**:119–138, 1991.

[3] M. Calvo, J. M. Oller. A distance between multivariate normal distributions based in embedding into the Siegel group. *J. Multivariate Analysis* **35**:223–242, 1990.

[4] P. S. Eriksen. Geodesics connected with the Fisher metric on the multivariate normal manifold. *Geometrization of statistical theory (Lancaster, 1987)*, 225–229, 1987.

[5] H. Furuhata, J. Inoguchi, and S.-P. Kobayashi. A characterization of the alpha-connections on the statistical manifold of normal distributions. *Inf. Geom.* **4**(1):1–12, 2020.

[6] W. Globke, R. Quiroga-Barranco. Information geometry and asymptotic geodesics on the space of normal distributions. *Inf. Geom.* **4**(1):131–153, 2021.

[7] M. Guest. *Harmonic maps, loop groups, and integrable systems.* London Mathematical Society Student Texts, **38**, 1997.

[8] R. Hermann. The differential geometry of foliations. *Ann. of Math.* **72**:445–457, 1960.

[9] S. Helgason. *Differential geometry, Lie groups, and symmetric spaces.* Pure and Applied Mathematics, **80**. Academic Press, Inc. 1978.

[10] H. Inoue. A symmetric subdomain contained in a fiber of the Siegel upper half space. PhD thesis, Kyushu University, 2016.

[11] H. Inoue. Matrix-valued Bratu equation and the exact solution of its initial value problem. *Int. J. Math. Ind.* **12**, 1950007, 8pp, 2020.

[12] S.-P. Kobayashi, Y. Ohno. A characterization of the alpha-connections on the statistical manifold of multivariate normal distributions. *Preprint*, 2023.

[13] B. Kostant. The solution to a generalized Toda lattice and representation theory. *Adv. Math.* **34**(3):195–338, 1979.

[14] D. S. P. Leung. On the classification of reflective submanifolds of Riemannian symmetric spaces. *Indiana Univ. Math. J.* **24**:327–339, 1974/75.
[15] M. Lovrić, M. Min-Oo, E. Ruh. Multivariate normal distributions parametrized as a Riemannian symmetric space. *J. Multivariate Anal.* **74**(1):36–48, 2000.
[16] Y. Nakamura. Algorithms associated with arithmetic, geometric and harmonic means and integrable systems. *J. Comput. Appl. Math.* **131**(1–2):161–174, 2001.
[17] B. O'Neill. Submersions and geodesics. *Duke Math. J.* **34**:363–373, 1967.
[18] L. T. Skovgaard. A Riemannian geometry of the multivariate normal model. Technical Report No.167, Stanford University, 1981.
[19] T. Yoshizawa. A geometry of parameter space and its statistical interpretation. *Surikaiseki-kenkyusho Kokyuroku* **167**, 103–131, 1972.

Department of Mathematics, Hokkaido University, Sapporo, 060-0810, Japan

Email address: shimpei@math.sci.hokudai.ac.jp