# Remarks on geodesics for multivariate normal models

Imai, Takuro
Graduate School of Mathematics, Kyushu University

Takaesu, Akira
NEC Communication Systems, Ltd. | Graduate School of Mathematics, Kyushu University

Wakayama, Masato
Institute of Mathematics for Industry, Kyushu University



出版情報：Journal of Math-for-Industry (JMI). 3 (B), pp.125-130, 2011-10-12. Faculty of Mathematics, Kyushu University
バージョン：
権利関係：

Journal of Math-for-Industry, Vol. 3 (2011B-6), pp. 125–130

# Remarks on geodesics for multivariate normal models

Takuro Imai, Akira Takaesu and Masato Wakayama

Received on May 10, 2011 / Revised on September 30, 2011

**Abstract.** A complete description of the geodesic curves on the Riemann manifold of multivariate normal distributions equipped with the Fisher information metric has been accomplished by Eriksen in 1987, and later by Calvo and Oller in 1991 but in a different manner. The former describes geodesic curves in terms of an exponential map in somewhat mysterious way and the latter obtains a solution of the differential equation of a geodesic curve explicitly by solving much general system of differential equations. The method what Erikson had taken seems to have a group theoretic nature while it is still unclear. The purposes of this short note are to derive the explicit formula of the geodesic curve from the result obtained by Eriksen and to clarify why such exponential map may give geodesic curves for the one dimensional normal distribution case.

*Keywords.* geodesics, multivariate normal distribution, statistical manifold, Fisher's information metric, Riemann symmetric spaces, Lorentz group.

## 1. Introduction

The advantage of introducing geometric concepts and tools in statistics are quite well-understood now. One of the main emphasis on the investigation of this direction has been put on determining the geodesic curves and the distance between two statistical distributions, initiating with the work by Rao [6].

Among several basic works on this subject, the geodesic curves on the Riemann manifold $M$ of multivariate normal distributions equipped with the Fisher information metric have been studied by many people, and first determined by Eriksen [3] in 1987 and Calvo-Oller [2] later in 1991 in a very different manner. In fact, the former [3] describes geodesic curves in terms of an exponential map in somewhat mysterious way and the latter [2] obtains an explicit solution of the differential equation of a geodesic curve. The method what Erikson had taken and the result obtained there are still not so clear. Actually, he has integrated the geodesic equations via using the exponential matrix function of a given compound matrix (made of the initial data) but did not provide an explicit form of the solution. In contrast with this, the latter [2] solved much general system of differential equations and then obtained an explicit description of the solution from it.

The purpose of this note is to make two remarks on the work [3]. Actually, we first shows that one can derive the solutions of the geodesic differential equation explicitly from the result obtained in [3]. For the second, we clarify why such exponential map can provide nicely a description of the geodesic curves when the model (= the statistical manifold of the one dimensional normal distributions) is being isomorphic to the Poincaré upper half plane. The first result is considered to be another proof of the result in [2].

## 2. Preliminaries

Consider the $p$th multivariate normal distribution $(\mu, \Sigma)$, where $\mu \in \mathbb{R}^p$ is the mean vector and a positive definite symmetric matrix $\Sigma \in GL(p, \mathbb{R})$ is the covariance matrix, where $GL(p, \mathbb{R})$ is the general linear group of order $p$ over $\mathbb{R}$. Let $N_p := \{(\mu, \Sigma)\}$ be the statistical Riemann manifold defined by the collection of the $p$th multivariate normal distributions $(\mu, \Sigma)$ equipped with the Fisher information metric, which is computed as

$$ds^2 = ({}^t d\mu)\Sigma^{-1}(d\mu) + \frac{1}{2}\mathrm{tr}((\Sigma^{-1}d\Sigma)^2). \tag{1}$$

(See, e.g. [8], for the detailed Riemann structure of $N_p$ and also for a brief introduction to the development of studies on geometric structure of statistical models. For further details, see [1].)

A geodesic curve $(\mu(t), \Sigma(t))$ $(t \in \mathbb{R})$ on $N_p$ satisfies

$$\begin{cases} \ddot{\Sigma} + \dot{\mu}{}^t\dot{\mu} - \dot{\Sigma}\Sigma^{-1}\dot{\Sigma} = 0, \\ \ddot{\mu} - \dot{\Sigma}\Sigma^{-1}\dot{\mu} = 0, \end{cases}$$

where $\dot{}$ denotes the differentiation with respect to the variable $t$. This system of differential equations can be obtained by the Euler-Lagrange equation for the Lagrangian

$$\mathcal{L} = {}^t\dot{\mu}\Sigma^{-1}\dot{\mu} + \frac{1}{2}\mathrm{tr}((\Sigma^{-1}\dot{\Sigma})^2).$$

See [4] for a general derivation of geodesic equations and [8] for this special case.

We note that the Riemann manifold $N_p$ is a homogeneous Riemann space. In fact, the positive affine motion group $GA^+(p) := \{g = (\delta, P) \in \mathbb{R}^p \times GL(p,\mathbb{R}) \mid \det P > 0\}$ acts $N_p$ as

$$g.(\mu,\Sigma) := (P\mu + \delta, P\Sigma{}^tP) \in N_p\,, (g = (\delta, P) \in GA^+(p)).$$

This action, obviously, leaves the Fisher metric on $N_p$ invariant and is transitive. Therefore, since the special orthogonal group $SO(p) := \{g \in GL(p,\mathbb{R}) \mid {}^tgg = \mathrm{I}_p, \det g = 1\}$ is the stabilizer subgroup of the point $(0, \mathrm{I}_p)$, one obtains the following isomorphism:

$$\begin{array}{ccc} GA^+(p)/SO(p) & \overset{\sim}{\to} & N_p \\ g = (\delta, P) & \mapsto & g.(0,\mathrm{I}_p) = (\delta, P{}^tP) \end{array}.$$

Hence, we may restrict ourselves to describe the geodesic through the origin $(0, \mathrm{I}_p)$ of $N_p$ by translation.

The system of the equations for the geodesic curve can be partially integrated as

$$\dot{\mu} = \Sigma x, \quad \dot{\Sigma} = \Sigma(B - x{}^t\mu),$$

where $x$ ($p$th vector) and $B$ ($p$th square matrix) are integration constants. Set

$$\Delta(t) = \Sigma(t)^{-1} \quad \text{and} \quad \delta(t) = \Sigma(t)^{-1}\mu(t).$$

Then it is known that in [3] (see, e.g. [2] for the derivation) the system of differential equations above becomes

$$\begin{cases} \dot{\Delta} = -B\Delta + x{}^t\delta, \\ \dot{\delta} = -B\delta + (1 + {}^t\delta\Delta^{-1}\delta)x, \\ \Delta(0) = \mathrm{I}_p,\ \delta(0) = 0. \end{cases}$$

We may also take the initial direction of the curves as

$$\begin{cases} \dot{\Delta}(0) = -B & (B \in \mathrm{Sym}_p(\mathbb{R})), \\ \dot{\delta}(0) = x & (x \in \mathbb{R}^p), \end{cases}$$

where $\mathrm{Sym}_p(\mathbb{R})$ denotes the set of all real symmetric matrices of degree $p$.

Let $\mathrm{Mat}_p = \mathrm{Mat}_p(\mathbb{R})$ be the set of all real square matrices of degree $p$. Put

$$A = \begin{pmatrix} -B & x & 0 \\ {}^tx & 0 & -{}^tx \\ 0 & -x & B \end{pmatrix} \in \mathrm{Mat}_{2p+1}.$$

Eriksen proved the following result (Theorem in [3]).

**Proposition 1.** *For $t \in \mathbb{R}$, put*

$$\Lambda(t) := \exp(tA) = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!} =: \begin{pmatrix} \Delta & \delta & \Phi \\ {}^t\delta & \epsilon & {}^t\gamma \\ {}^t\Phi & \gamma & \Gamma \end{pmatrix}.$$

*Then the geodesics curve through the origin $(0, \mathrm{I}_p)$ of $N_p$ with tangent $(x, -B)$ is given by $(\delta(t), \Delta(t))$.*

## 3. Calculation of the exponential

In this section, we calculate the exponential of the matrix $A$ in Proposition 1 and obtain explicit expressions of the geodesic curves on $N_p = \{(\mu, \Sigma)\}$.

We first define two matrices of degree $2p + 1$ by

$$H = \begin{pmatrix} -\frac{1}{2}B & x & -\frac{1}{2}B \\ 0 & 0 & 0 \\ \frac{1}{2}B & -x & \frac{1}{2}B \end{pmatrix}, \ L = \frac{1}{2}\begin{pmatrix} \mathrm{I}_p & 0 & -\mathrm{I}_p \\ 0 & 0 & 0 \\ -\mathrm{I}_p & 0 & \mathrm{I}_p \end{pmatrix}.$$

Then it is immediate to see that

$$A = H + {}^tH,\ H^2 = {}^tH^2 = 0$$

and

$$\begin{aligned} {}^tHH &= \begin{pmatrix} \frac{1}{2}B^2 & -Bx & \frac{1}{2}B^2 \\ -{}^txB & 2{}^txx & -{}^txB \\ \frac{1}{2}B^2 & -Bx & \frac{1}{2}B^2 \end{pmatrix}, \\ H{}^tH &= \begin{pmatrix} \frac{1}{2}B^2 + x{}^tx & 0 & -\frac{1}{2}B^2 - x{}^tx \\ 0 & 0 & 0 \\ -\frac{1}{2}B^2 - x{}^tx & 0 & \frac{1}{2}B^2 + x{}^tx \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} G^2 & 0 & -G^2 \\ 0 & 0 & 0 \\ -G^2 & 0 & G^2 \end{pmatrix}, \end{aligned}$$

where $G$ is a positive semi-definite symmetric matrix defined by $G^2 = B^2 + 2x{}^tx$. Also, we notice that

$$LH = H \text{ and } L^2 = L = {}^tL.$$

**Lemma 1.** *Put $M = H{}^tH$. Then one has*

$$\begin{aligned} A^{2n+1} &= M^nH + {}^tHM^n, \\ A^{2n+2} &= M^{n+1} + {}^tHM^nH. \end{aligned}$$

*Proof.* One finds that

$$A^2 = H{}^tH + {}^tHH,$$

whence it follows that

$$\begin{aligned} A^{2n+3} &= A^{2n+1}A^2 = (M^nH + {}^tHM^n)(H{}^tH + {}^tHH) \\ &= {}^tHM^{n+1} + M^{n+1}H, \\ A^{2n+2} &= A^{2n+1}A = (M^nH + {}^tHM^n)(H + {}^tH) \\ &= M^nH{}^tH + {}^tHM^nH. \end{aligned}$$

By induction, the claim follows immediately. □

From the lemma above we have

$$\begin{aligned} \Lambda(t) &= \mathrm{I}_{2p+1} + \sum_{n=1}^{\infty} \frac{t^{2n}}{(2n)!}A^{2n} + \sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!}A^{2n+1} \\ &= \mathrm{I}_{2p+1} + \sum_{n=1}^{\infty} \frac{t^{2n}}{(2n)!}M^n + {}^tH\sum_{n=1}^{\infty} \frac{t^{2n}}{(2n)!}M^{n-1}H \\ &\quad + \sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!}M^nH + {}^tH\sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!}M^n. \end{aligned} \tag{2}$$

We recall here the notion of a g-inverse of a square matrix.

**Definition 1** (g-inverse)**.** Let $G$ be a square matrix. If the square matrix $G^-$ satisfies

$$GG^-G = G,$$

then $G^-$ is called a generalized inverse (simply, g-inverse) of $G$.

The following lemmas are known (see e.g. [7]).

**Lemma 2.** *If $G^-$ be the g-inverse of a symmetric matrix $G$, then*

$$G^-GG^- = G^-,\ G^-G = GG^-.$$

**Lemma 3.** *If $X = X_1 + X_2$, with $X_1$ and $X_2$ are positive semi-definite symmetric, then there exist positive semi-definite symmetric matrices $G$ and $G^-$, where $G^-$ is a g-inverse of $G$, such that the following conditions hold.*

**(a)** $G^2 = X$,

**(b)** $X_iGG^- = GG^-X_i = X_i$.

*Moreover, if $X_i = R_i^2$, where $R_i$ is a positive semi-definite symmetric matrix, then*

$$R_iGG^- = GG^-R_i = R_i.$$

Let $X = G^2 = B^2 + 2x{}^tx$. Then, by Lemma 3, there exists a positive semi-definite symmetric g-inverse $G^-$ of $G$ such that the following equations hold.

$$BGG^- = GG^-B = B,$$
$$GG^-x{}^tx = x{}^txGG^- = x{}^tx.$$

Since $x$ is an eigenvector of the matrix $x{}^tx$, using the latter relation above, we observe

$$({}^txx)GG^-x = GG^-x{}^txx = x{}^txx = ({}^txx)x.$$

Hence, it follows that

$$GG^-x = x.$$

To make our subsequent discussion simpler, we introduce the following convention. For $X \in \mathrm{Mat}_p$ and $K \in \mathrm{Mat}_{2p+1}$ of the form

$$K = \begin{pmatrix} A & x & B \\ 0 & 0 & 0 \\ C & y & D \end{pmatrix} \quad (A, B, C, D \in \mathrm{Mat}_p, x, y \in \mathbb{R}^p),$$

we define a matrix $X \circ_R K \in \mathrm{Mat}_{2p+1}$ by

$$X \circ_R K = \begin{pmatrix} XA & Xx & XB \\ 0 & 0 & 0 \\ XC & Xy & XD \end{pmatrix}.$$

Then for $X, X' \in \mathrm{Mat}_p$ and for such matrices $K, K' \in \mathrm{Mat}_{2p+1}$, it is easy to see that

$$(XX') \circ_R K = X \circ_R (X' \circ_R K),$$
$$X \circ_R (KK') = (X \circ_R K)K'.$$

Similarly, for $X \in \mathrm{Mat}_p$ and

$$J = \begin{pmatrix} A & 0 & B \\ {}^tx & 0 & {}^ty \\ C & 0 & D \end{pmatrix} \in \mathrm{Mat}_{2p+1},$$

we define

$$J \circ_L X = \begin{pmatrix} AX & 0 & BX \\ {}^txX & 0 & {}^tyX \\ CX & 0 & DX \end{pmatrix} \in \mathrm{Mat}_{2p+1}.$$

Then it is obvious that

$${}^t(X \circ_R K) = {}^tK \circ_L {}^tX.$$

Using this convention, we have

$$M = H{}^tH = G^2 \circ_R L = L \circ_L G^2.$$

In particular, since $L^2 = L$ it follows that

$$M^n = G^{2n} \circ_R L.$$

We now calculate the exponential $\Lambda(t) = \exp(tA)$ explicitly according to the four parts decomposition of the sum (2). Since

$$\begin{aligned} M^nH &= (G^{2n} \circ_R L)H = G^{2n} \circ_R (LH) = G^{2n} \circ_R H \\ &= (G^{2n+1}G^-) \circ_R H = G^{2n+1} \circ_R (G^- \circ_R H), \end{aligned}$$

it is not hard to calculate the first term as

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!} M^nH &= \sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!} G^{2n+1} \circ_R (G^- \circ_R H) \\ &= \sinh(tG)G^- \circ_R H. \end{aligned}$$

Similarly, we have

$$\begin{aligned} {}^tH \sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!} M^n &= {}^t(\sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!} M^nH) \\ &= {}^tH \circ_L \sinh(tG)G^-, \\ \sum_{n=1}^{\infty} \frac{t^{2n}}{(2n)!} M^n &= \sum_{n=1}^{\infty} \frac{t^{2n}}{(2n)!} G^{2n} \circ_R L \\ &= (\cosh(tG) - \mathrm{I}_p) \circ_R L, \\ {}^tH \sum_{n=1}^{\infty} \frac{t^{2n}}{(2n)!} M^{n-1}H &= {}^tH \sum_{n=1}^{\infty} \frac{t^{2n}}{(2n)!} G^{2n-2} \circ_R (LH) \\ &= {}^tH \sum_{n=1}^{\infty} \frac{t^{2n}}{(2n)!} G^{2n}(G^-)^2 \circ_R H \\ &= {}^tH(\cosh(tG) - \mathrm{I}_p)(G^-)^2 \circ_R H. \end{aligned}$$

A small manipulation hence yields the following.

**Theorem 1.** *Retain the notation and assumption in the*

*Preliminaries. Then, one has*

$$\begin{aligned}\Delta(t) =& \mathrm{I}_p + \frac{1}{2}(\cosh(tG) - \mathrm{I}_p) + \frac{1}{2}B(\cosh(tG) - \mathrm{I}_p)(G^-)^2 B \\ &- \frac{1}{2}\sinh(tG)G^- B - \frac{1}{2}B\sinh(tG)G^-, \\ \delta(t) =& - B(\cosh(tG) - \mathrm{I}_p)(G^-)^2 x + \sinh(tG)G^- x, \\ \Phi(t) =& - \frac{1}{2}(\cosh(tG) - \mathrm{I}_p) + \frac{1}{2}B(\cosh(tG) - \mathrm{I}_p)(G^-)^2 B \\ &- \frac{1}{2}\sinh(tG)G^- B + \frac{1}{2}B\sinh(tG)G^-\end{aligned}$$

*and*

$$\Gamma(t) = \Delta(-t), \quad \gamma(t) = \delta(-t).$$

Remark that the last two relations in the theorem can be also seen from the fact $\Lambda(t)^{-1} = \Lambda(-t)$. Further, since $\dot{\Lambda}(t) = A\Lambda(t)$, one observes easily that $\dot{\delta} = -B\delta + \epsilon x$. Hence, noticing (a part of) the geodesic equation $\dot{\delta} = -B\delta + (1 + {}^t\delta\Delta^{-1}\delta)x$, we have

$$\epsilon(t) = 1 + {}^t\delta(t)\Delta(t)^{-1}\delta(t).$$

From the first two equations in Theorem 1, one can easily derive the explicit formulas of the geodesic curves for the multivariate normal model obtained in [2].

**Corollary 1.** *When $n = 1$, one has*

$$\begin{aligned}\Delta(t) =& 1 + \frac{1}{2}(\cosh(tg) - 1)(1 + b^2 g^{-2}) - bg^{-1}\sinh(tg), \\ \delta(t) =& - b(\cosh(tg) - 1)g^{-2}x + \sinh(tg)g^{-1}x\end{aligned}$$

*and*

$$\Phi(t) = \frac{1}{2}(\cosh(tg) - 1)(b^2 g^{-2} - 1).$$

We make here a small remark. Consider the submanifold $N_p(0) := \{(0, \Sigma)\}$ of $N_p$ (that is, with the 0 mean value). Then, since $x = 0$ we have $G = B$. It follows immediately from the theorem above that $\Sigma(t) = \Delta(t)^{-1} = \exp(tB)$. In other words, the geodesic curves on $N_p(0)$ are given by the exponential map. This meets the fact that the metric $ds^2 = \frac{1}{2}\mathrm{tr}((\Sigma^{-1}d\Sigma)^2)$ on $N_p(0)$ is equal to that of the space of positive definite symmetric matrices of order $p$, i.e. the Riemann symmetric space $GL(p, \mathbb{R})/O(p)$, where $O(p)$ is the orthogonal group of order $p$.

## 4. A group theoretic interpretation for the one dimensional model

The aim of this section is to make a group theoretical interpretation of Proposition 1 when $p = 1$.

Let $SL(2, \mathbb{R}) = \{g \in GL(2, \mathbb{R}) \mid \det g = 1\}$ be the special linear group of order 2. The corresponding Lie algebra is $\mathfrak{sl}(2, \mathbb{R}) := \{X \in \mathrm{Mat}_2(\mathbb{R}) \mid \mathrm{tr}X = 0\}$ and the Lie bracket is given by the matrices commutator $[X, Y] := XY - YX$.

The stage we consider is the complex upper plane $H := \{z = x + iy \in \mathbb{C} \mid y > 0\}$ equipped with the Poincaré metric $ds^2 = \frac{2}{y^2}(dx^2 + dy^2)$. The group $SL(2, \mathbb{R})$ acts on $H$ by the linear fractional transform:

$$g.z = \frac{az + b}{cz + d} \quad (g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{R}),\ z \in H).$$

Note that the action leaves invariant the Poincaré metric. Moreover, it is well known (see, e.g. [5]) that every $g \in SL(2, \mathbb{R})$ can be uniquely written as $g = n_x a_y k$ (the Iwasawa decomposition), where $n_x := \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$ $(x \in \mathbb{R})$, $a_y := \begin{pmatrix} y^{\frac{1}{2}} & 0 \\ 0 & y^{-\frac{1}{2}} \end{pmatrix}$ $(y > 0)$ and $k \in SO(2)$. This shows that the stabilizer group of the point $i \in H$ is given by $K := SO(2)$. Hence the (isometric) map $SL(2, \mathbb{R})/SO(2) \ni gK \mapsto g.i = x + iy \in H$ induces a structure of the Riemann symmetric space on $H$. We sometimes identify the coset $gK$ with the point $z = g.i$.

For $g \in SL(2, \mathbb{R})$ and $X, Y \in \mathfrak{sl}(2, \mathbb{R})$, define the adjoint representations of $SL(2, \mathbb{R})$ and $\mathfrak{sl}(2, \mathbb{R})$ respectively by

$$\mathrm{Ad}(g)X := gXg^{-1} \quad \text{and} \quad \mathrm{ad}(X)Y := [X, Y].$$

Also, the Killing form $B$ of $\mathfrak{sl}(2, \mathbb{R})$ is defined by

$$B(X, Y) := \mathrm{tr}(\mathrm{ad}(X)\mathrm{ad}(Y)) = 4\mathrm{tr}XY.$$

It defines a non-degenerated quadratic form of signature $(2, 1)$ on the space $\mathfrak{sl}(2, \mathbb{R})$. We put $B_o(X, Y) = \frac{1}{2}\mathrm{tr}XY$.

Let $e_1 = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, e_2 = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ and $e_3 = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$. Then $(e_1, e_2, e_3)$ gives an orthogonal basis of $\mathfrak{sl}(2, \mathbb{R})$ with respect to $B_o$. We will fix this basis in the sequel. Then, since

$$\begin{aligned}\mathrm{ad}\begin{pmatrix} b & x \\ x & -b \end{pmatrix} e_1 =& 2be_3, \\ \mathrm{ad}\begin{pmatrix} b & x \\ x & -b \end{pmatrix} e_2 =& - 2xe_3, \\ \mathrm{ad}\begin{pmatrix} b & x \\ x & -b \end{pmatrix} e_3 =& 2be_1 - 2xe_2,\end{aligned}$$

we have

$$\mathrm{ad}\begin{pmatrix} b & x \\ x & -b \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2b \\ 0 & 0 & -2x \\ 2b & -2x & 0 \end{pmatrix}.$$

Similarly, for $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{R})$, we have

$$\begin{aligned}&\mathrm{Ad}(g) \\ =& \begin{pmatrix} \frac{1}{2}(a^2 - b^2 - c^2 + d^2) & -ab + cd & \frac{1}{2}(a^2 + b^2 - c^2 - d^2) \\ bd - ac & ad + bc & -(ac + bd) \\ \frac{1}{2}(a^2 - b^2 + c^2 - d^2) & -(ab + cd) & \frac{1}{2}(a^2 + b^2 + c^2 + d^2) \end{pmatrix}.\end{aligned}$$

We have the following group isomorphism:

$$\begin{array}{ccc} SL(2, \mathbb{R})/\{\pm \mathrm{I}_2\} \ni g & \mapsto & \tilde{g} = \mathrm{Ad}(g) \in SO_o(2, 1) \\ \cup & & \cup \\ SO(2)/\{\pm \mathrm{I}_2\} \ni k & \mapsto & \tilde{k} = \mathrm{Ad}(k) \in S(O(2) \times O(1)), \end{array}$$

where $SO_o(2, 1) := SO_o(B_o)$ is the connected component of the special Lorentz group $SO(B_o)(\cong SO(2, 1)) := \{g \in$

$GL(3,\mathbb{R}) \mid B_o(gX, gY) = B_o(X,Y), \det g = 1\}$ of type $(2,1)$. It is elementary to verify that $B_o(X,Y) = x_1y_1 + x_2y_2 - x_3y_3$ for $X = (x_1,x_2,x_3)$, $Y = (y_1,y_2,y_3) \in \mathbb{R}^3$ with respect to our fixed basis $(e_1,e_2,e_3)$. Since the adjoint action $\mathrm{Ad}(g)$ $(g \in SL(2,\mathbb{R}))$ leaves the quadratic form $B_o(X,Y)$ invariant, as Riemann symmetric spaces, we have the isomorphisms:

$$H \cong SL(2,\mathbb{R})/SO(2) \cong SO_o(2,1)/S(O(2)\times O(1)).$$

Now, in order to make our story meet the picture of Proposition 1 (and Theorem 1), we take a realization of the Lie algebra $\mathfrak{so}(2,1)$ as

$$\mathfrak{so}(2,1) := \{Z \in \mathrm{Mat}_3(\mathbb{R}) \mid {}^tZJ + JZ = 0,\ \mathrm{tr}(Z) = 0\},$$

where $J = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$. Remark that the corresponding realization of the Lorentz group is given by

$$SO(2,1) := \{g \in GL(3,\mathbb{R}) \mid {}^tgJg = J,\ \det g = 1\}.$$

It is the conjugate image of the group $SO(B_o)$ above by $C = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{pmatrix}$. Obviously, $C^{-1} = C$.

Let $X = \begin{pmatrix} -\frac{b}{2} & -\frac{x}{\sqrt{2}} \\ -\frac{x}{\sqrt{2}} & \frac{b}{2} \end{pmatrix} \in \mathfrak{sl}(s,\mathbb{R})$. Then we see that

$$C\mathrm{ad}(X)C^{-1} = \begin{pmatrix} -b & x & 0 \\ x & 0 & -x \\ 0 & -x & b \end{pmatrix} \in \mathfrak{so}(2,1).$$

Since $\mathrm{Ad}(\exp tX) = \exp(t\mathrm{ad}(X))$, we observe

$$\begin{aligned} &C\mathrm{Ad}(\exp tX)C^{-1} \\ &= C\exp(t\mathrm{ad}(X))C^{-1} = \exp(tC\mathrm{ad}(X)C^{-1}) \\ &= \exp t\begin{pmatrix} -b & x & 0 \\ x & 0 & -x \\ 0 & -x & b \end{pmatrix} =: \begin{pmatrix} \Delta & \delta & \Phi \\ \delta & \epsilon & \gamma \\ \Phi & \gamma & \Gamma \end{pmatrix}. \end{aligned}$$

It follows that

$$\begin{aligned} &\mathrm{Ad}(\exp tX) = C^{-1}\begin{pmatrix} \Delta(t) & \delta(t) & \Phi(t) \\ \delta(t) & \epsilon(t) & \gamma(t) \\ \Phi(t) & \gamma(t) & \Gamma(t) \end{pmatrix} C \\ &= \frac{1}{2}\begin{pmatrix} \Delta(t)+2\Phi(t)+\Gamma(t) & \sqrt{2}(\delta(t)+\gamma(t)) & \Delta(t)-\Gamma(t) \\ \sqrt{2}(\delta(t)+\gamma(t)) & 2\epsilon(t) & \sqrt{2}(\delta(t)-\gamma(t)) \\ \Delta(t)-\Gamma(t) & \sqrt{2}(\delta(t)-\gamma(t)) & \Delta(t)-2\Phi(t)+\Gamma(t) \end{pmatrix}. \end{aligned}$$

On the other hand, one finds that $X^2 = \frac{g^2}{4}\mathrm{I}_2$, where $g^2 = b^2 + 2x^2$. It follows that

$$X^{2n} = (\frac{g}{2})^{2n}\mathrm{I}_2 \quad \text{and} \quad X^{2n+1} = (\frac{g}{2})^{2n}X.$$

We then compute the exponential as

$$\begin{aligned} &\exp tX \\ &= \sum_{n=0}^{\infty}\frac{t^{2n}}{(2n)!}(\frac{g}{2})^{2n}\mathrm{I}_2 + \sum_{n=0}^{\infty}\frac{t^{2n+1}}{(2n+1)!}(\frac{g}{2})^{2n}X \\ &= \cosh(\frac{tg}{2})\mathrm{I}_2 + 2g^{-1}\sinh(\frac{tg}{2})X \\ &= \begin{pmatrix} \cosh(\frac{tg}{2}) - g^{-1}\sinh(\frac{tg}{2})b & -\sqrt{2}g^{-1}\sinh(\frac{tg}{2})x \\ -\sqrt{2}g^{-1}\sinh(\frac{tg}{2})x & \cosh(\frac{tg}{2}) + g^{-1}\sinh(\frac{tg}{2})b \end{pmatrix}. \end{aligned}$$

Set $\exp tX = \begin{pmatrix} \alpha(t) & \beta(t) \\ \beta(t) & \gamma(t) \end{pmatrix}$. Then, since

$$\begin{aligned} &\mathrm{Ad}(\exp tX) \\ &= \begin{pmatrix} \frac{1}{2}(\alpha^2+\gamma^2-2\beta^2) & -\alpha\beta+\beta\gamma & \frac{1}{2}(\alpha^2-\gamma^2) \\ \beta\gamma-\beta\alpha & \alpha\gamma+\beta^2 & -(\alpha\beta+\beta\gamma) \\ \frac{1}{2}(\alpha^2-\gamma^2) & -(\alpha\beta+\beta\gamma) & \frac{1}{2}(\alpha^2+2\beta^2+\gamma^2) \end{pmatrix}, \end{aligned}$$

comparing the elements of two expressions of $\mathrm{Ad}(\exp tX)$ above, i.e. by taking the sum of (1.1), (1.3), (3.1) and (3.3)-components, and of (1.2) and (2.3)-components respectively, we conclude that

$$\Delta = \alpha^2 \quad \text{and} \quad \delta = -\sqrt{2}\alpha\beta. \tag{3}$$

By Proposition 1, taking the exponential $\Lambda(t) = \exp tA$ of the matrix $A = \mathrm{ad}(X) \in \mathrm{Mat}_3(\mathbb{R})$, we find that the equation (3) should give a geodesic on $N_1$. (In fact, one can easily obtain the explicit formulas in Corollary 1 from (3).)

We now give a brief explanation why the equation (3) can give a geodesic on $N_1$ directly from the general theory of geodesics on Riemann symmetric spaces. Notice first that the Fisher information metric given by the equation (1) for the one dimensional normal model $N_1 = \{(\mu,\sigma^2)\}$ is

$$ds^2 = \sigma^{-2}(d\mu^2 + 2d\sigma^2).$$

It is hence easy to check that the following map defines an isomorphism between $N_1$ and the Poincaré upper half plane $H$:

$$\begin{array}{ccc} N_1 & \xrightarrow{\sim} & H \cong SL(2,\mathbb{R})/SO(2) \\ (\mu,\sigma^2) & \mapsto & \frac{1}{\sqrt{2}}\mu + i\sigma \end{array}.$$

Hence the manifold $N_1$ is a Riemann symmetric space. In particular, one may identify $(0,1) \in N_1$ with $i \in H$ (the origin). It is well known that the geodesic curve through the origin $eK$ with a tangent vector $X$ on a Riemann symmetric space $G/K$ is given by the exponential map $(\exp tX)K \in G/K$ (see [4]). Therefore, what we have to show that is the curve $(\mu(t),\sigma(t)^2) = (\Delta(t)^{-1}, \Delta(t)^{-1}\delta(t))$ can be obtained by $\exp(tY).i$ for some $Y \in \mathfrak{sl}(2,\mathbb{R})$.

Let $X = \begin{pmatrix} -\frac{b}{2} & -\frac{x}{\sqrt{2}} \\ -\frac{x}{\sqrt{2}} & \frac{b}{2} \end{pmatrix} \in \mathfrak{sl}(2,\mathbb{R})$ as before. Define the real valued functions $\tilde{\mu}(t)$ and $\tilde{\sigma}(t)$ by

$$\frac{1}{\sqrt{2}}\tilde{\mu}(t) + i\tilde{\sigma}(t) := \exp(-\frac{t}{2}X).i \in H \quad (t \in \mathbb{R}).$$

Then we observe

$$\begin{aligned} \tilde{\sigma}(-2t) &= \mathrm{Im}\{\exp tX.i\} = \frac{1}{\beta(t)^2+\gamma(t)^2} \\ &= \frac{1}{\cosh(tg) + g^{-1}b\sinh(tg)} = \frac{1}{\alpha(-2t)}. \end{aligned}$$

Similarly, noticing $\alpha\gamma - \beta^2 = 1$, we have

$$\begin{aligned} \frac{1}{\sqrt{2}}\tilde{\mu}(-2t) &= \mathrm{Re}\{\exp tX.i\} = \frac{(\alpha(t)+\gamma(t))\beta(t)}{\beta(t)^2+\gamma(t)^2} \\ &= \frac{-\sqrt{2}g^{-1}x\sinh(tg)}{\alpha(-2t)} = -\frac{\beta(-2t)}{\alpha(-2t)} \\ &= -\Delta(-2t)^{-1}\alpha(-2t)\beta(-2t). \end{aligned}$$

Hence, it follows from (3) that

$$\begin{cases} \tilde{\sigma}(t)^2 = \Delta(t)^{-1} = \sigma(t)^2, \\ \tilde{\mu}(t) = \Delta(t)^{-1}\delta(t) = \mu(t). \end{cases}$$

This explains group theoretically the legitimacy of the result in [3] for $p = 1$.

## 5. Concluding discussions

The matrix $A$ taken in the Preliminaries can be regarded as an element of the Lie algebra $\mathfrak{g} := \mathfrak{so}(p+1, p) = \{X \in \mathrm{Mat}_{2p+1}(\mathbb{R}) \,|\, {}^tXJ_{2p+1} + J_{2p+1}X = 0\}$ of the special Lorentz group $G = SO_o(p+1, p)$ as Eriksen [3] pointed out. Here the matrix $J_{2p+1}$, which defines a non-degenerate $(p+1,\, p)$ quadratic form (invariant under the action of $G$, i.e. ${}^tgJ_{2p+1}g = J_{2p+1}$ for $g \in G$), is given by

$$J_{2p+1} = \begin{pmatrix} 0 & 0 & \mathrm{I}_p \\ 0 & 1 & 0 \\ \mathrm{I}_p & 0 & 0 \end{pmatrix}.$$

More precisely, the matrix $A$ is regarded as an element of the $\mathfrak{p}$-part of the Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, where $\mathfrak{k}$ is the Lie algebra of a maximal compact subgroup $K = S(O(p+1) \times O(p))$ of $G$. It is well known (see [4]) that its exponential $\exp(tA)K$ defines a geodesic curve on the Riemann symmetric space $G/K$.

When $p = 1$, there exists an isomorphism of Riemann manifolds $N_1 \cong SO_o(2,1)/S(O(2) \times O(1))$. Thus, as we have shown that $\exp(tA)$ for $A = \mathrm{ad}(X) \in \mathfrak{so}(2,1)$ indeed describes a geodesic curve on $N_1$. If $p > 1$, however, the Riemann structure defined by the Fisher information metric on the space $N_p$ is not equivalent to the one induced from the Killing form of $\mathfrak{g}$ on $G/K$. Although $N_p$ can be embedded into $G/K$ as a differentiable manifold, this fact implies that $\exp(tA)K$ is not necessarily defines a geodesic on the normal model $N_p$ when $p > 1$. It is the future work to study this point by clarifying the geometric meaning of each component of the matrix $\exp(tA)$.

## Acknowledgments

The authors would like to thank the referee for useful comments, which helped them to improving the first version of this paper.

This work was partially supported by Grant-in Aid for Scientific Research (B) No. 21340011 from the Japan Society for the Promotion of Science.

## References

[1] Amari, S., and Nagaoka, H.: *Methods of Information Geometry.* Trans. Math. Monographs **191**, Amer. Math. Soc. and Oxford Univ. Press, 2001.

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

[3] Eriksen, P.S.: Geodesics connected with the Fisher metric on the multivariate normal manifold, *Proceedings of the GST Workshop, Lancaster* (1987), 225–229.

[4] Helgason, S.: *Differential Geometry, Lie Groups, and Symmetric Spaces.* Academic Press, New York, 1978.

[5] Lang, S.: $SL_2(\mathbb{R})$ Addison-Wesley, 1975.

[6] Rao, C.R.: Information and the accuracy attainable in the estimation of statistical parameters, *Bull. Calcutta Math. Soc.* **37**(1945), 81-91.

[7] Rao, C.R. and Mitra, S.K.: *Generalized inverse of matrices and its applications.* John Wiley & Sons, New York, 1971.

[8] Skovgaard, L.T.: A Riemannian geometry of the multivariate normal model, *Scand. J. Statist.* **11** (1984), 211–223.

Takuro Imai and Akira Takaesu
Graduate School of Mathematics, Kyushu University, 744, Motooka, Nishi-ku, Fukuoka, 819-0395, Japan
E-mail: t-imai(at)math.kyushu-u.ac.jp
takaesu.ak(at)gmail.com

Akira Takaesu
NEC Communication Systems, Ltd., Tokyo 108-0073, Japan (*current address*)
E-mail: takaesu.ak(at)gmail.com

Masato Wakayama
Institute of Mathematics for Industry, Kyushu University, 744, Motooka, Nishi-ku, Fukuoka, 819-0395, Japan
E-mail: wakayama(at)imi.kyushu-u.ac.jp