NORTH-HOLLAND

Dualistic Differential Geometry of Positive Definite Matrices and Its Applications to Related Problems*

Atsumi Ohara† and Nobuhide Suda
Department of Systems Engineering
Osaka University
1-3, Machikane-yama, Toyanaka
Osaka 560, Japan

and

Shun-ichi Amari
Department of Mathematical Engineering and Information Physics
University of Tokyo
7-3-1, Hongo, Bunkyo-ku
Tokyo 113, Japan

Submitted by T. Ando

ABSTRACT

Analysis on the set of positive definite matrices is involved in many engineering problems. We develop a differential geometric theory of the set of positive definite matrices by means of a specific class of connections introduced into it. Consequently, various dualistic aspects of the set are elucidated. Next, using the theory, we derive new results and interesting geometrical interpretations for matrix approximation, positive definite matrix completion, and linear matrix inequality problems.

* Some of the results in this paper were given in [16] without proofs.

† E-mail: ohara@sys.es.osaka-u.ac.jp.

LINEAR ALGEBRA AND ITS APPLICATIONS 247:31–53 (1996)

655 Avenue of the Americas, New York, NY 10010
0024-3795/96/$15.00
SSDI 0024-3795(94)00348-H

0. NOTATION

$\mathrm{Sym}(n)$	Set of n by n real symmetric matrices
$\mathrm{PD}(n)$	Set of n by n real positive definite symmetric matrices
$\mathrm{GL}(n)$	Set of n by n real nonsingular matrices
$\mathbf{R}$	Set of real numbers
$T_P\mathcal{M}$	Tangent vector space of the manifold $\mathcal{M}$ at a point $P \in \mathcal{M}$
I	Identity matrix of appropriate size
A^T	Transpose of the matrix A
$\lambda_{\min}(A)$	Smallest eigenvalue of the symmetric matrix A
$\mathrm{tr}(A)$	Trace of the matrix A
δ_i^j	Kronecker's delta
N	Dimension of $\mathrm{PD}(n)$ and $\mathrm{Sym}(n)$ [$= n(n+1)/2$]
m	Dimension of submanifold $\mathcal{S} \subset \mathrm{PD}(n)$ defined in Section 4
$\mathrm{argmin}_{A \in \mathcal{S}} f(A)$	Matrix A in the set $\mathcal{S}$ that minimizes the function $f(A)$

We adopt Einstein's summation convention for the indices which appear twice as sub and superscripts, e.g., $c^k = a_{ij}b^{ijk}$ automatically means $c^k = \sum_i \sum_j a_{ij}b^{ijk}$. We use i, j, k, and l to run from 1 to N; α, β, γ, and δ from 1 to m; and κ and μ from $m+1$ to N.

---

1. INTRODUCTION

It is widely known that positive definite matrices often play important roles in various aspects of mathematical analysis for signal processing, control theory, numerical analysis, and operator theory via the Lyapunov stability theorem, the Riccati equation, covariance analysis, and so on (see the references). Hence, studying properties of the set of positive definite matrices, which itself has mathematical interest, gives important insights into the above fields.

The set of positive definite matrices has been studied as not only a convex cone [11, 2], but also a Riemannian manifold [10, 8, 16] in the context of differential geometry. While these differential geometric analyses give many fundamental and fruitful results, the main tool used there has been limited to Riemannian geometry, where metric-preserving (Levi-Civita) connection [13] is essential. Hence, by introducing other specific connections, we have a possibility to investigate more abundant properties and geometric structures of the manifold, which are deeply related to several applications in engineering.

In this paper, we first define and exploit a new geometry of the manifold of positive definite matrices by means of *dual connections*, which have been introduced by Amari [1] in statistical analysis. As a result, we feature the dualistic nature of the manifold and its information theoretic implications. Specifically, *dual coordinate systems* and *divergence* functions, introduced here, are found to be fundamental tools of geometrical analysis of the manifold. Next, we discuss the relations or applications of the theory to matrix approximation (MA) [6, 5], positive definite matrix completion (PDMC) [7, 12, 9], and linear matrix inequality (LMI) [3, 4] problems. Consequently, we can give new results and geometric interpretations of existing results for these problems from the point of view of the dualistic differential geometry.

As a related work, we should refer to the interesting paper [8], which discusses the Kahler structure of the set of positive definite matrices. The point of view of this paper is quite different, but its relation to our present results, which is not clear now, will make the geometry of the set more attractive.

2. RIEMANNIAN METRIC AND DUAL CONNECTIONS ON PD(n)

2.1. *Riemannian Metric*

Let $\{E_i\}$, $i = 1, \ldots, N$, be linearly independent basis matrices of Sym(n), then $P \in$ PD(n) can be represented as

$$P = P(\theta) := \theta^i E_i \in \mathrm{PD}(n). \tag{1}$$

Hence, we can regard $\theta = (\theta^i)$ as a global coordinate system for PD(n) and $\partial_i := \partial/\partial\theta^i$ as a tangent vector field on PD(n).

Denote the tangent vector space of PD(n) at a point P by T_pPD(n). Since T_pPD(n) is isomorphic to Sym(n) for each P, we will identify, E_i with the natural basis $(\partial_i)_P := (\partial/\partial\theta^i)_P$ of T_PPD(n). Then tangent vectors X_P in T_PPD(n) can be represented by matrices X in Sym(n):

$$T_P\mathrm{PD}(n) \ni X_P := a^i(\partial_i)_P \equiv X := a^i E_i \in \mathrm{Sym}(n), \qquad (a^i) \in R^N. \tag{2}$$

Here, the symbol $\equiv$ denotes the identification.

Now, we will consider the following Riemannian metric g on PD(n) [16] defined by

$$g_{ij}(\theta) = g_P(\partial_i, \partial_j) := \mathrm{tr}\left(P^{-1}E_i P^{-1}E_j\right). \tag{3}$$

The above Riemannian metric g can be proved invariant under the inverse transformation $\iota : P \mapsto P^{-1}$ and the congruent transformations $\tau_T : P \mapsto TPT^T$ for all $T \in \mathrm{GL}(n)$, i.e., the two diffeomorphism ι and τ_T are isometries on $(\mathrm{PD}(n), g)$ [16]. This follows from the fact that the differential of ι

$$\iota_* : T_P\mathrm{PD}(n) \ni X_P \equiv X \mapsto -P^{-1}XP^{-1} \equiv \iota_*(X_P) \in T_{\iota(P)}\mathrm{PD}(n) \quad (4)$$

and the differential of τ_T

$$\tau_{T*} : T_P\mathrm{PD}(n) \ni X_P \equiv X \mapsto TXT^T \equiv \tau_{T*}(X_P) \in T_{\tau_T(P)}\mathrm{PD}(n), \quad (5)$$

which are described in the matrix representation (2), satisfy

$$g_P(X_P, Y_P) = g_{\iota(P)}(\iota_*(X_P), \iota_*(Y_P)),$$
$$g_P(X_P, Y_P) = g_{\tau_T(P)}(\tau_{T*}(X_P), \tau_{T*}(Y_P)). \quad (6)$$

2.2. *Dual Connections*

For any smooth curve $c : P(t)$ on $\mathrm{PD}(n)$ where t belongs to a certain interval $I_0 \subset \mathbf{R}$, *parallel displacement* [14, p. 70] along the curve c defines linear mappings from $T_{P(t_1)}\mathrm{PD}(n)$ to $T_{P(t_2)}\mathrm{PD}(n)$ for all $t_1, t_2 \in I_0$. Using the matrix representation (2) again, we introduce the following two specific parallel displacements.

$$\Pi_c : X \equiv X_{P(t_1)} \mapsto X_{P(t_2)} := \Pi_c X_{P(t_1)} \equiv X, \quad (7)$$

$$\Pi_c^* : X \equiv X_{P(t_1)} \mapsto X_{P(t_2)} := \Pi_c^* X_{P(t_1)} \equiv P(t_2)P(t_1)^{-1}XP(t_1)^{-1}P(t_2). \quad (8)$$

Note that Π_c and Π_c^* do not depend on the curve c they follow, but only on $P(t_1)$ and $P(t_2)$.

Let ∇ and ∇^* denote connections [13, p. 63] corresponding to the above parallel displacements, respectively. The components of each connections represented in the θ coordinate system are as follows (see Appendix A):

$$\Gamma_{ijk}(\theta) := g(\nabla_{\partial_i}\partial_j, \partial_k) = 0, \quad (9)$$

$$\Gamma_{ijk}^*(\theta) := g(\nabla_{\partial_i}^*\partial_j, \partial_k) = -2\,\mathrm{tr}(P^{-1}E_iP^{-1}E_jP^{-1}E_k). \quad (10)$$

While these connections do not preserve the Riemannian metric g (non-metric connection [13, p. 158]), they have the following properties:

THEOREM 1.

(i) *The connections* ∇ *and* ∇^* *satisfy*

$$Ag(B,C) = g(\nabla_A B, C) + g(B, \nabla_A^* C) \tag{11}$$

for any vector fields A, B, C *on* $\mathrm{PD}(n)$.

(i) *The torsion and curvature with respect to the connection* ∇ *are zero.*

(iii) *The torsion and curvative with respect to the connection* ∇^* *are zero.*

Proof. First, note that since all the matrices P, E_i belong to $\mathrm{Sym}(n)$, then we have

$$\begin{aligned}\mathrm{tr}\left(P^{-1}E_iP^{-1}E_jP^{-1}E_k\right) &= \mathrm{tr}\left(P^{-1}E_iP^{-1}E_kP^{-1}E_j\right)\\ &= \mathrm{tr}\left(P^{-1}E_jP^{-1}E_iP^{-1}E_k\right).\end{aligned} \tag{12}$$

Here, we have used a formula $\mathrm{tr}(FG) = \mathrm{tr}(GF) = \mathrm{tr}(F^TG^T)$ for the first equality with $F := P^{-1}E_iP^{-1}$ with $G := E_jP^{-1}E_k$, and for the second equality with $F := P^{-1}E_iP^{-1}E_k$ and $G := P^{-1}E_j$. For any constant matrices F, we also have

$$\partial_i \,\mathrm{tr}(P^{-1}F) = -\mathrm{tr}\left(P^{-1}E_iP^{-1}F\right) \qquad \forall F \in \mathrm{Sym}(n). \tag{13}$$

using the differentiation of the matrix inverse.

(i): By direct differentiation of (3), we get

$$\partial_i g_{jk}(\theta) = -2\,\mathrm{tr}\left(P^{-1}E_iP^{-1}E_jP^{-1}E_k\right). \tag{14}$$

Then from (9) and (10) we have

$$\partial_i g_{jk} = \Gamma_{ijk} + \Gamma^*_{ikj}. \tag{15}$$

This implies (11).

(ii): Obvious from $\Gamma_{ijk} = 0$.

(iii): Let $T^*_{ijk}(\theta)$ be the components of torsion tensor with respect to ∇^*; then

$$T^*_{ijk} := g\left(\nabla^*_{\partial_i}\partial_j - \nabla^*_{\partial_j}\partial_i, \partial_k\right) = \Gamma^*_{ijk} - \Gamma^*_{jik} = 0. \tag{16}$$

Further, when $P(t_1) = P(t_2)$, (8) implies the parallel displacement Π_c^* does not change tangent vectors for any closed curve c. Hence, the curvature tensor with respect to ∇^* vanishes. ■

A pair of connections ∇ and ∇^* that satisfy (11) are called *mutually dual* [1]. By virtue of statement (ii) [(iii)], we shall say PD(n) is ∇-*flat* [∇^*-*flat*].

For general Riemannian manifold $(\mathscr{M}, g)$, Amari has shown the following results:

LEMMA 1 [1, pp. 80-82]. *If the Riemannian manifold $(\mathscr{M}, g)$ is ∇- and ∇^*-flat with respect to a certain pair of dual connections ∇ and ∇^*, then*:

(i) *There exist two specific coordinate systems $\theta = (\theta^i)$ and $\eta = (\eta_i)$ on $\mathscr{M}$ which satisfy*

$$g(\partial_i, \partial^j) = \delta_i^j, \qquad \text{where} \quad \partial^j := \partial/\partial\eta_j. \tag{17}$$

(*We will call such a pair of coordinate systems* mutually dual.) *The coordinate system θ [η] is* affine *for the connection ∇ [∇^*], i.e.*, $\Gamma_{ijk}(\theta) = g(\nabla_{\partial_i}\partial_j, \partial_k) = 0$ [$\Gamma^{*ijk}(\eta) = g(\nabla^*_{\partial^i}\partial^j, \partial^k) = 0$].

(ii) *There exist potential functions $\psi(\theta)$ and $\phi(\eta)$ on $\mathscr{M}$ such that*

$$\eta_i = \eta_i(\theta) = \partial_i\psi(\theta), \qquad \theta^i = \theta^i(\eta) = \partial^i\phi(\eta), \tag{18}$$

$$g_{ij}(\theta) = \partial_i\,\partial_j\psi(\theta) = \partial\eta_j/\partial\theta^i, \qquad g^{ij}(\eta) = \partial^i\partial^j\phi(\eta) = \partial\theta^j/\partial\eta_i, \tag{19}$$

$$g^{ij}(\eta)g_{jk}(\theta) = \delta_k^i, \qquad \textit{at any point} \quad p(\theta) = p(\eta) \in \mathscr{M}. \tag{20}$$

Further, they are related to each other via Legendre transformation (*note the relation* (18)):

$$\psi(\theta) + \phi(\eta) - \theta^i\eta_i = 0 \qquad \textit{at any point} \quad p = p(\theta) = p(\eta) \in \mathscr{M}. \tag{21}$$

Here, the quantities $g^{ij}(\eta)$ and $\Gamma^{*ijk}(\eta)$ denote the components of the Riemannian metric g and the connection ∇^* represented in the η coordinate system.

Since $(\mathrm{PD}(n), g)$ is both ∇- and ∇^*-flat by Theorem 1, we can apply the above lemma. Actually, the dual coordinate system η and the potential functions $\phi(\theta)$ and $\psi(\eta)$ satisfying the lemma can be defined as follows:

Let $\{E^i\}$ be another basis matrices for $\mathrm{Sym}(n)$ biorthogonal to $\{E_i\}$, i.e.,

$$-\mathrm{tr}\left(E^i E_j\right) = \delta_j^i, \tag{22}$$

and represent P using $\{E^i\}$ as

$$P = P(\eta) := \left(\eta_i E^i\right)^{-1}. \tag{23}$$

Then $\eta = (\eta_i)$ can be also regarded as a global coordinate system of $\mathrm{PD}(n)$. As a result, we find the obtained coordinate system η is dual to θ defined by (1) and satisfies $\Gamma^{*ijk}(\eta) = 0$ (cf. Appendix B).

Further, define two potential functions,

$$\begin{aligned} \psi(\theta) &= \psi(P(\theta)) := -\log\det P(\theta) - c_\psi, \\ \phi(\eta) &= \phi(P(\eta)) := -\log\det P(\eta)^{-1} - c_\phi, \end{aligned} \tag{24}$$

where c_ψ and c_ϕ are any constants that satisfy $c_\psi + c_\phi = n$. Then the statements in Lemma 1(ii) are verified (cf. Appendix B).

REMARK 1. Using the pair of dual connections (∇, ∇^*), we can define the one parameter family of connections

$$\overset{a}{\nabla} := \frac{1-a}{2}\nabla + \frac{1+a}{2}\nabla^*, \qquad a \in \mathbf{R}. \tag{25}$$

The pair of connections $(\overset{a}{\nabla}, \overset{-a}{\nabla})$ are found to be mutually dual. In particular, $\overset{0}{\nabla}$ is self-dual, i.e., coincides with the Riemannian (Levi-Civita) connection on $(\mathrm{PD}(n), g)$.

3. GEODESICS, SUBMANIFOLDS, AND DIVERGENCES

For use in the following sections, we introduce some important concepts or tools derived from the dualistic geometric structures of $\mathrm{PD}(n)$ as simply as possible. For more details, consult [1, Chapter 3].

First, since connections are given to (PD(n), g), we can define a special curve called the *geodesic*, whose tangent vectors are parallel with respect to the connection in question. We call geodesics with respect to the connections ∇ and ∇^*, respectively, ∇- and ∇^*-*geodesic*. In the θ coordinate system, ∇- and ∇^*-geodesics are characterized, respectively, as the solutions of the following differential equations with proper initial conditions [13, p. 146]:

$$\ddot{\theta}^i(t) = 0, \tag{26}$$

$$g_{ij}(\theta(t))\ddot{\theta}^j(t) + \Gamma^*_{ijk}(\theta(t))\dot{\theta}^j(t)\dot{\theta}^k(t) = 0. \tag{27}$$

On the other hand, in the η coordinate system, they are characterized as the solutions of the following differential equations:

$$g^{ij}(\eta(t))\ddot{\eta}_j(t) + \Gamma^{ijk}(\eta(t))\dot{\eta}_j(t)\dot{\eta}_k(t) = 0, \tag{28}$$

$$\ddot{\eta}_i(t) = 0. \tag{29}$$

The equations (26) and (29) follow from the facts that $\Gamma_{ijk}(\theta) = 0$ and $\Gamma^{*ijk}(\eta) = 0$, respectively.

We should note that from Lemma 1(i) or (26) and (29), the coordinate curves θ^i and η_i are examples of ∇- and ∇^*-geodesics on PD(n), respectively.

Next, let $\mathscr{Q}$ be an m-dimensional smooth submanifold in PD(n). We denote by $\tilde{X}_\alpha, \tilde{Y}_\beta$, $\alpha, \beta = 1, \dots, m$, tangent vector fields whose vectors at each $P \in \mathscr{Q}$ form a basis for $T_P\mathscr{Q}$; and by $\tilde{Z}_\mu$, $\mu = m+1, \dots, N$, tangent vector fields whose vectors at each $P \in \mathscr{Q}$ form a basis for $T_P\mathscr{Q}^\perp$ [the orthogonal complement of $T_P\mathscr{Q}$ in T_PPD(n)].

The *Euler-Schouten imbedding curvature tensors* (or *second fundamental forms*) of the submanifold $\mathscr{Q}$ with respect to ∇ and ∇^* are defined by

$$H_{\alpha\beta\mu} := g\left(\nabla_{\tilde{X}_\alpha}\tilde{Y}_\beta, \tilde{Z}_\mu\right), \qquad H^*_{\alpha\beta\mu} := g\left(\nabla^*_{\tilde{X}_\alpha}\tilde{Y}_\beta, \tilde{Z}_\mu\right).$$

These quantities show how curved the submanifold $\mathscr{Q}$ is in PD(n) in the sense of the connection ∇ or ∇^*. When $H_{\alpha\beta\mu}$ ($H^*_{\alpha\beta\mu}$) is zero, the submanifold $\mathscr{Q}$ is said to be ∇-*autoparallel* (∇^*-*autoparallel*). In particular, a one-dimensional ∇- (∇^*-)autoparallel manifold is a ∇- (∇^*-)geodesic.

Finally, consider the following functions $\mathrm{PD}(n) \times \mathrm{PD}(n) \to \mathbf{R}$ defined via potential functions.

$$\begin{aligned} D(P_1, P_2) &:= \psi(P_1) + \phi(P_2) - \theta_1^i \eta_{2i} \\ &= \log\det P_2 - \log\det P_1 + \mathrm{tr}\left(P_2^{-1} P_1\right) - n, \end{aligned} \tag{30}$$

$$\begin{aligned} D^*(P_1, P_2) &:= \phi(P_1) + \psi(P_2) - \eta_{1i}\theta_2^i \\ &= \log\det P_1 - \log\det P_2 + \mathrm{tr}\left(P_1^{-1} P_2\right) - n \\ &= D(P_2, P_1), \end{aligned} \tag{31}$$

where (θ_1^i) and (η_{1i}) [(θ_2^i) and (η_{2i})] are the θ- and η-coordinates of P_1 [P_2]. These functions are called ∇- and ∇^*-*divergences*. In particular, $D(P_1, P_2)$ coincides with the well-known *Kullback-Leibler information* (or *relative entropy*) between two zero-mean Gaussian distributions whose covariances are P_1 and P_2, respectively.

While the divergences do not satisfy the definition of a distance function, they play the role of a measure for closeness in $\mathrm{PD}(n)$ because of (i), (ii) and (iii) in the following lemma:

LEMMA 2 (Properties of divergences [1, pp. 84-93]).

(i) *For all P_1 and P_2, $D(P_1, P_2) \geqslant 0$. The equality holds only when $P_1 = P_2$.*

(ii) *If $P_1 = (\theta^i) = (\eta_i)$ and $P_2 = (\theta^i + d\theta^i) = (\eta_i + d\eta_i)$, i.e., infinitesimal displacement of P_1, then*

$$D(P_1, P_2) = \frac{g_{ij}\, d\theta^i\, d\theta^j}{2} + O\left(|d\theta^i|^3\right) = \frac{g^{ij}\, d\eta_i\, d\eta_j}{2} + O\left(|d\eta_i|^3\right). \tag{32}$$

(iii) (*Pythagorean theorem*) *Let c_+ and c_- be the ∇-geodesic connecting P_1 with P_2 and ∇^*-geodesic connecting with P_2 with P_3, respectively. Then $D(P_1, P_3) - D(P_1, P_2) - D(P_2, P_3)$ is positive, zero, or negative according as the angle between tangent vectors of c_+ and c_- at P_2 is greater than, equal to, or less than $\pi/2$.*

(iv) *For a given smooth submanifold $\mathscr{Q}$ and point P_0 in* $\mathrm{PD}(n)$, *let $\mathscr{P}$ be a set of points $P \in \mathscr{Q}$ such that ∇-geodesics connecting P_0 with P are orthogonal to $T_P\mathscr{Q}$ (such points are called* ∇-projections *of P_0 to $\mathscr{Q}$). Then a point $\hat{P} :=$* $\arg\min_{Q \in \mathscr{Q}}$ *$D(P_0, Q)$ is included in $\mathscr{P}$. In particular if $\mathscr{Q}$ is*

∇^**-autoparallel, the* ∇*-projection of* P_0 *to* $\mathscr{Q}$ *uniquely exists (and hence* $\hat{P}$ *does).*

REMARK 2. For $D^*(P_1, P_2)$, the above statements also hold if we swap ∇ and ∇^*. Statement (iv) implies that a local property (orthogonality) determines a global one (minimality of divergences) when $\mathscr{Q}$ is ∇- or ∇^*-autoparallel. This fact is found useful for constructing practical algorithms to obtain optimal approximation of P to $\mathscr{Q}$ via optimization methods. More generally, when $\mathscr{Q}$ is ∇^*- (∇-)convex, the uniqueness of ∇- (∇^*-)projection can be assured via (iii) [1, p. 91].

4. SUBMANIFOLD OF LINEARLY CONSTRAINED MATRICES IN PD(n)

For given $E_0, \dots, E_m \in \mathrm{Sym}(n)$, where $\{E_\alpha\}$, $\alpha = 1, \dots, m < N$, are linearly independent, consider an m-dimensional submanifold $\mathscr{S}$ in PD(n) defined by

$$\mathscr{S} := \{P(x) \mid P(x) = E_0 + x^\alpha E_\alpha \in \mathrm{PD}(n)\}. \tag{33}$$

We can regard $x = (x^\alpha)$ as a coordinate system of the submanifold $\mathscr{S}$.

For this submanifold $\mathscr{S}$, we can formulate the following problems:

(a) Find a certain point $P \in \mathscr{S}$ (or equivalently, check whether $\mathscr{S}$ is nonempty).

(b) Find an approximate point $\hat{P} \in \mathscr{S}$ for $P \notin \mathscr{S}$ with some distance measures.

(c) Parametrize $\mathscr{S}$.

Problem (a) is referred to as a *linear matrix inequality* [4, 3] or *positive definite matrix completion* [7, 9, 12] problem, and problem (b) as a *matrix approximation* [5, 6] problem. Many mathematical and engineering problems, such as approximation by Toeplitz or Hankel type matrices in signal processing or the LMI approach for control system synthesis, can be cast into the above forms (see the above references). In the rest of the paper, we will exploit geometric properties of $\mathscr{S}$ using the dualistic geometry introduced in the previous sections to apply to the above problems. We treat mainly problems (a) and (b). The consideration of problem (c) via the dualistic geometry will be given in another place.

In this section we first consider the case $\mathscr{L} \neq \emptyset$ to treat problem (b).

Now we define the Riemannian metric and dual connections on $\mathscr{L}$ *induced* from PD(n). Represent E_0 and each E_α, $\alpha = 1, \ldots, m$, as a linear combination of E_i:

$$E_0 := \theta_0^i E_i, \qquad E_\alpha := B_\alpha^i E_i, \text{ where } B_\alpha^i \text{ is a constant.}$$

Then the θ-coordinate of $P(x)$ and the tangent vector field $\partial_\alpha := \partial / \partial x^\alpha$ on $\mathscr{L}$ are represented as

$$\theta^i(x) = \theta_0^i + B_\alpha^i x^\alpha, \qquad \partial_\alpha = B_\alpha^i \, \partial_i. \tag{34}$$

Hence, the components of the *induced* Riemannian metric and dual connections on $\mathscr{L}$ are calculated from those of PD(n) as

$$g_{\alpha\beta}(x) = B_\alpha^i B_\beta^j g_{ij}(\theta(x)), \tag{35}$$

$$\Gamma_{\alpha\beta\gamma}(x) = B_\alpha^i B_\beta^j B_\gamma^k \Gamma_{ijk}(\theta(x)) + \left(\partial_\alpha B_\beta^j\right) B_\gamma^k g_{jk}(\theta(x)) = 0. \tag{36}$$

$$\begin{aligned} \Gamma^*_{\alpha\beta\gamma}(x) &= B_\alpha^i B_\beta^j B_\gamma^k \Gamma^*_{ijk}(\theta(x)) + \left(\partial_\alpha B_\beta^j\right) B_\gamma^k g_{jk}(\theta(x)) \\ &= B_\alpha^i B_\beta^j B_\gamma^k \Gamma^*_{iji}(\theta(x)). \end{aligned} \tag{37}$$

Here, we use $\Gamma_{ijk}(\theta) = 0$ and $\partial_\alpha B_\beta^j = 0$. These are fundamental quantities that determine the dualistic structures of $\mathscr{L}$.

Let $\partial_\kappa, \partial_\mu$, $\kappa, \mu = m + 1, \ldots, N$, be tangent vector fields whose vectors at each $P(x) \in \mathscr{L}$ form a basis for $T_P \mathscr{L}^\perp$. The following theorem is the key result for geometric analysis of $\mathscr{L}$:

THEOREM 2. *The submanifold $\mathscr{L}$ is ∇-autoparallel in* PD(n).

Proof. Let $\tilde{B}_\mu^k(x) := \partial_\mu \theta^k \mid_{P(x)}$. Then the components of the Euler-Schouten imbedding curvature of $\mathscr{L}$ are calculated by

$$\begin{aligned} H_{\alpha\beta\mu}(x) = g\left(\nabla_{\partial_\alpha} \beta_\beta, \partial_\mu\right) &= B_\alpha^i B_\beta^j \tilde{B}_\mu^k(x) \Gamma_{ijk}(\theta(x)) \\ &\quad + \left(\partial_\alpha B_\beta^j\right) \tilde{B}_\mu^k(x) g_{jk}(\theta(x)) = 0, \end{aligned}$$

because $\Gamma_{ijk}(\theta) = 0$ and $\partial_\alpha B_\beta^j = 0$. ∎

From this theorem, the following result follows:

COROLLARY 1. *The submanifold $\mathscr{L}$ is itself ∇- and ∇^*-flat.*

Proof. From (36), we immediately find $\mathscr{L}$ is itself ∇-flat.

It is easy from (37) to see $\mathscr{L}$ is torsion-free with respect to ∇^*. Further, it is known the components of Riemannian curvature tensor of $\mathscr{L}$ with respect to ∇^*, denoted by $R^*_{\alpha\beta\gamma\delta}$, satisfy the following equation [19, 1]:

$$R^*_{\alpha\beta\gamma\delta}(x) = B^i_\alpha B^j_\beta B^k_\gamma B^l_\delta R^*_{ijkl}(\theta(x)) + \left[H_{\alpha\delta\kappa}(x) H^*_{\beta\gamma\mu}(x) - H^*_{\alpha\gamma\kappa}(x) H_{\beta\delta\mu}(x)\right] g^{\kappa\mu}(\theta(x)),$$

where $g^{\kappa\mu}(\theta)$ is the inverse of $g_{\kappa\mu} = g(\partial_\kappa, \partial_\mu)$, and $R_{ijkl}{}^*(\theta)$ are the components of the Riemannian curvature tensor of PD(n) with respect to ∇^*. Since PD(n) is ∇^*-flat, i.e., $R^*_{ijkl}(\theta) = 0$ and $\mathscr{L}$ is ∇-autoparallel, i.e., $H_{\alpha\beta\kappa}(x) = 0$ we obtain $R^*_{\alpha\beta\gamma\delta}(x) = 0$. Thus, $\mathscr{L}$ is itself ∇^*-flat. ■

For problem (b), we have to choose a distance measure for approximation. The most popular measure to approximate P_1 by P_2 is the Frobenius norm

$$\|P_1 - P_2\|_F := \left\{\mathrm{tr}\left[(P_1 - P_2)^2\right]\right\}^{1/2}, \qquad P_1, P_2 \in \mathrm{PD}(n).$$

In many applications to engineering, the matrices P_1 and P_2 represent quadratic physical quantities, e.g., correlations. Hence, measures for approximations are often required to satisfy the invariance under the congruent transformations τ_T, which physically corresponds to scalings or change of units for the quantities. Since the Frobenius norm does not generally satisfy it, i.e.,

$$\|P_1 - P_2\|_F \neq \|\tau_T(P_1) - \tau_T(P_2)\|_F,$$

we have to use the other measures for this purpose.

Several distance measures derived from the Riemannian metric g satisfy the requirement because g is invariant for τ_T. Among them, the Riemannian distance function [16]

$$\mathrm{dist}(P_1, P_2) := \left\{\mathrm{tr}\left[\left(\log P_1^{-1/2} P_2 P_1^{-1/2}\right)^2\right]\right\}^{1/2}$$

may be the most natural, because it satisfies the definition of a distance function. While the primal divergence $D(P_1, P_2)$ does not satisfy the definition of distance function, it may be also suitable for special purposes because it coincides with the Kullback-Leibler information, which implies the approximation with D has a statistical meaning: *maximum likelihood estimation* or minimum relative entropy approximation. However, to find the best approximation in $\mathscr{S}$ with these distance measures dist$(\bullet, \bullet)$ or $D(\bullet, \bullet)$ is difficult in general because the extreme points are not always unique and do not always constitute the global minimum.

To the contrary, from Theorem 2 and Remark 2, we find the dual divergence $D^*(P_1, P_2)$ has the following nice properties for treating problem (b):

COROLLARY 2. *For given $P \notin \mathscr{S}$, the best approximation point with D^*, i.e., $\hat{P} := \arg\min_{Q \in \mathscr{S}} D^*(P, Q)$, uniquely exists and is characterized as the ∇^*-projection of P to $\mathscr{S}$.*

REMARK 3. As in the previous section, the equation of the ∇^*-geodesic is represented by $\ddot{\eta}_i(t) = 0$ in the η coordinate system. Hence, the ∇^*-geodesic connecting a given point $P = (\theta^i) = (\eta_i)$ and any point $P(x) = (\theta^i(x)) = (\eta_i(x)) \in \mathscr{S}$ is

$$\eta_i(t) := \eta_i + t\big[\eta_i(x) - \eta_i\big], \qquad t \in \mathbf{R}, \tag{38}$$

and its tangent vectors are $[\eta_i(x) - \eta_i]\partial^i$. To obtain the ∇^*-projection, we have to solve the orthogonality condition at $P(x)$ for x:

$$\begin{aligned} g\big(\partial_\alpha, [\eta_i(x) - \eta_i]\partial^i\big) &= g\big(B_\alpha^j \partial_j, [\eta_i(x) - \eta_i]\partial^i\big) \\ &= B_\alpha^i\big[\eta_i(x) - \eta_i\big] = 0. \end{aligned} \tag{39}$$

As the above corollary states, this is equivalent to obtaining the minimizing point of $D^*(P, P(x))$. Actually, the extreme condition $\partial_\alpha D^*(P, P(\hat{x})) = 0$ can be modified to the orthogonality condition (39) using (18), (31):

$$\begin{aligned} \partial_\alpha\big[\psi(\theta(x)) - \eta_j\theta^j(x)\big] &= B_\alpha^i\, \partial_i\big[\psi(\theta(x)) - \eta_j\theta^j(x)\big] \\ &= B_\alpha^i\big[\eta_i(x) - \eta_i\big] = 0. \end{aligned}$$

Further, $D^*(P, P(x))$ is a convex function on $\mathscr{S}$ because

$$\partial_\alpha\, \partial_\beta D^*(P, P(x)) = B_\alpha^i B_\beta^j\, \partial_i\, \partial_j \psi(\theta(x)) = B_\alpha^i B_\beta^j g_{ij}(\theta(x)) = g_{\alpha\beta}(x)$$

is positive definite. Thus, we find the extreme point is the global minimum. Consequently, one practical way to obtain the best approximation with D^* is to use suitable convex programming algorithms. The condition (39) can be used to check the convergence.

5. ANALYTIC CENTER FOR $\mathscr{S}$ AND ANCILLARY SUBMANIFOLD

To develop the theory further and apply it to problem (a), we introduce a special point called the *analytic center* (AC) [3] *maximum entropy completion* (MEC) [7, 9, 12], or *ancillary submanifold* [1].

First, we extend the coordinate system x to all of PD(n). Let $\{E_\kappa\}$, $\kappa = m + 1, \dots, N$, be complementary basis matrices for $\{E_\alpha\}$, i.e., $\{E_\alpha; E_\kappa\}$ form a basis of Sym(n), and x_0^κ and x_0^κ be constants that satisfy

$$E_0 = x_0^\alpha E_\alpha + x_0^\kappa E_\kappa .$$

Since any $P \in \mathrm{PD}(n)$ can be represented as

$$P = P(x^\alpha; x^\kappa) := E_0 + x^\alpha E_\alpha + x^\kappa E_\kappa = (x^\alpha + x_0^\alpha) E_\alpha + (x^\kappa + x_0^\kappa) E_\kappa,$$

we can regard $\tilde{x} := (x^\alpha; x^\kappa)$ as a global coordinate system of PD(n), and $(x^\alpha; 0)$ specifies points on $\mathscr{S}$.

As we have seen in Section 2.2, the coordinate system $\tilde{y} = (y_\alpha; y_\kappa)$ that is dual to $\tilde{x}$ can be defined by

$$y_\alpha(\tilde{x}) := \partial_\alpha \psi(P(\tilde{x})), \quad y_\kappa(\tilde{x}) := \partial_\kappa \psi(P(\tilde{x})),$$
$$\text{where} \quad \partial_\alpha := \frac{\partial}{\partial x^\alpha}, \quad \partial_\kappa := \frac{\partial}{\partial x^\kappa}. \tag{40}$$

Conversely, we find that $\tilde{x}$ is represented in $\tilde{y}$ using x_0^α and x_0^κ due to the definition of $\tilde{x}$:

$$x^\alpha(\tilde{y}) = \partial^\alpha \phi(P(\tilde{y})) - x_0^\alpha, \quad x^\kappa(\tilde{y}) = \partial^\kappa \phi(P(\tilde{y})) - x_0^\kappa,$$
$$\text{where} \quad \partial^\alpha := = \frac{\partial}{\partial y_\alpha}, \quad \partial^\kappa := \frac{\partial}{\partial y_\kappa}. \tag{41}$$

We should note that vector fields $\partial^\kappa := \partial/\partial y_\kappa$ are orthogonal to $T_P\mathscr{S}$ at each $P \in \mathscr{S}$ because of Lemma 1(i).

Now let $\mathscr{L}$ be bounded. The AC for $\mathscr{L}$, denoted by P_{AC}, is defined as the point that maximizes the function log det $P(x)$ on $\mathscr{L}$ and plays an important role in an LMI or PDMC problem.

Some results about the matrix form of P_{AC}^{-1} (what is called the *inverse pattern*) have been obtained and discussed for the case of band matrices [7] or matrices whose diagonal entries are specified [9]. We find each of them to be a special case of $\mathscr{L}$ by choosing appropriate basis matrices E_α in (33). We first generalize these results to the case of $\mathscr{L}$, using the geometric approach in the previous sections.

THEOREM 3:. *Let* $\{E^\kappa\}$, $\kappa = m+1, \dots, N$, *be the bi orthogonal basis for* $\{E_\alpha; E_\kappa\}$ *such that* $-\mathrm{tr}(E^\kappa E_\alpha) = 0$ *and* $-\mathrm{tr}(E^\kappa E_\mu) = \delta_\mu^\kappa$. *Then the AC for* $\mathscr{L}$ *is characterized by its inverse as*

$$P_{\mathrm{AC}}^{-1} = c_\kappa E^\kappa,$$

where c_κ *is the solution of equations for* y_κ:

$$\partial^\kappa \phi(P(\tilde{y}))\mid_{\tilde{y}=(0;\,y_\kappa)} - x_0^\kappa = 0.$$

Proof. Since P_{AC} is characterized by the potential function $\psi(P)$:

$$P_{\mathrm{AC}} := \arg\max_{P\in\mathscr{L}} \log\det P = \arg\min_{P\in\mathscr{L}} \psi(P),$$

we find P_{AC} for $\mathscr{L}$ uniquely exists because $\psi(P(x))$ is a strictly convex function on $\mathscr{L}$ [or equivalently, $\partial_\alpha\,\partial_\beta\,\psi(P(\tilde{x})) = g_{\alpha\beta}(\tilde{x})$ is positive definite] and $\mathscr{L}$ is bounded. Hence, P_{AC} is the extreme point that satisfies

$$y_\alpha = \partial_\alpha \psi(P(\tilde{x})) = 0,$$

i.e., the dual coordinate $\tilde{y}$ of P_{AC} is of the form $(0; c_\kappa)$, where each c_κ is constant. These c_κ can be determined by the condition that P_{AC} belongs to $\mathscr{L}$. This implies, from (41), solving $N - m$ equations

$$x^\kappa = \partial^\kappa \phi(\tilde{y})\mid_{\tilde{y}=(0;\,y_\kappa)} - x_0^\kappa = 0$$

for y_κ. ∎

Finally, we will discuss the relation with the dualistic geometry and the convex algorithm ([8], [5], etc.) to solve LMI or PDMC problems, or check the nonemptiness of $\mathscr{L}$. Let us introduce two new subsets in PD(n).

First, we define the set $\mathscr{L}_\lambda$ for some fixed $\lambda \in \mathbf{R}$ by

$$\mathscr{L}_\lambda := \{P \mid P = P(x) - \lambda I = E_0 - \lambda I + x^\alpha E_\alpha \in \mathrm{PD}(n)\}.$$

If λ is smaller than $\lambda_{\min}(E_0)$, then $\mathscr{L}_\lambda$ is necessarily nonempty.

We should note that $\mathscr{L} \neq \emptyset$ iff $\lambda^* := \sup\{\lambda \mid \mathscr{L}_\lambda \neq \emptyset\} > 0$, and in this case $\lambda^* = \sup_{P \in \mathscr{L}} \lambda_{\min}(P)$. Hence, we will concentrate on getting λ^* to check the nonemptiness of $\mathscr{L}$. One of the solutions for LMI can be obtained through this process.

Next, assume $\mathscr{L}_\lambda \neq \emptyset$, and let $\mathscr{L}_\lambda^\perp(P)$ denote the family of ∇^*-geodesics that pass through $P \in \mathscr{L}_\lambda$ and are orthogonal to $\mathscr{L}_\lambda$ at P. According to Lemma 2 (iv), $\mathscr{L}_\lambda^\perp(P)$ is just the set of all the points in PD(n) whose ∇^*-projections to $\mathscr{L}_\lambda$ are identical to P.

An $(N-m)$-dimensional submanifold $\mathscr{L}_\lambda^\perp(P)$ has the following properties:

THEOREM 4. *Assume $\mathscr{L}_\lambda \neq \emptyset$. Then*
(i) *for each $P \in \mathscr{L}_\lambda$, $\mathscr{L}_\lambda^\perp(P)$ is ∇^*-autoparallel;*
(ii) *each $\mathscr{L}_\lambda^\perp(P)$ defines a* foliation *structure in* PD(n), *i.e.*,

$$\mathrm{PD}(n) = \bigcup_{P \in \mathscr{L}_\lambda} \mathscr{L}_\lambda^\perp(P), \qquad \mathscr{L}_\lambda^\perp(P) \cap \mathscr{L}_\lambda^\perp(P') = \emptyset \quad \text{if } P \neq P'.$$

Proof. Statement (i) follows similarly to Theorem 2. To prove (ii), we follow [13, p. 10]. Let $\mathscr{E}$ be the *distribution* generated by vector fields $\{\partial^\kappa\}$, $\kappa = m+1, \dots, N$. Using (i), i.e., $g(\nabla^*_{\partial^\kappa}\partial^\mu, \partial_\alpha) = 0$, we find

$$\nabla^*_X Y \in \mathscr{E}, \qquad \forall X, Y \in \mathscr{E}.$$

Since ∇^* is torsion-free,

$$[X, Y] = \nabla^*_X Y - \nabla^*_Y X \in \mathscr{E}.$$

Thus, $\mathscr{E}$ is *involutive*. It is easy to see its *maximal integrable manifold* is $\mathscr{L}_\lambda^\perp(P)$. Hence, statement (ii) follows. ■

The above theorem implies that if we assume the $\tilde{y}$ coordinate for P is $(p_\alpha; p_\kappa)$, then every point $Q(y_\kappa) \in \mathscr{L}_\lambda^\perp(P)$ is specified by $(p_\alpha; y_\kappa)$ in the $\tilde{y}$ coordinate, and any tangent vector of $\mathscr{L}_\lambda^\perp(P)$ at $Q(y_\kappa)$ is a linear combination of $(\partial^\kappa)_Q$, $\kappa = m+1, \dots, N$.

For each λ where $\lambda < \lambda^*$, we can define the AC for $\mathscr{L}_\lambda$, denoted by $P_{\mathrm{AC}}(\lambda)$. The curve $P_{\mathrm{AC}}(\lambda)$ in PD(n) for the parameter $\lambda < \lambda^*$ is called the *path* of AC for $\mathscr{L}_\lambda$. As shown in the proof of Theorem 3, the $\tilde{y}$ coordinate of $P_{\mathrm{AC}}(\lambda)$ is just $(0; c_\kappa(\lambda))$. Hence, if we simply denote $\mathscr{L}_\lambda^\perp(P_{\mathrm{AC}}(\lambda))$ by $\mathscr{L}^\perp$, which is actually independent of λ from Theorem 4, then we find $P_{\mathrm{AC}}(\lambda)$ is contained in $\mathscr{L}^\perp$ for all $\lambda < \lambda^*$.

Moreover, note that $\mathscr{L}_\lambda$ is ∇-autoparallel (which is proved in the same way as Theorem 2) and that $\mathscr{L}^\perp$ is ∇^*-autoparallel, then we obtain the following important characterizations for $P_{\mathrm{AC}}(\lambda)$:

COROLLARY 3 [Characterizations of $P_{\mathrm{AC}}(\lambda)$]. *For each $\lambda < \lambda^*$, let $Q_1 \in \mathscr{L}_\lambda$ and $Q_2 \in \mathscr{L}^\perp$. Then*

(i) $P_{\mathrm{AC}}(\lambda) = \mathscr{L}_\lambda \cap \mathscr{L}^\perp$;

(ii) $P_{\mathrm{AC}}(\lambda)$ *is the ∇-projection of Q_1 to $\mathscr{L}^\perp$*;

(iii) $P_{\mathrm{AC}}(\lambda)$ *is the ∇^*-projection of Q_2 to $\mathscr{L}_\lambda$.*

Since $\mathscr{L}_\lambda \neq \varnothing$ iff $P_{\mathrm{AC}}(\lambda)$ exists, to obtain λ^* we only have to follow the path of AC in the direction of λ increasing from λ_0. This is the basic idea of the *path following method* in general convex programming problems [15].

The algorithm on the basis of the above idea is as follows: For the sake of simplicity, we assume that a constant λ_0 such that $\mathscr{L}_{\lambda_0} \neq \varnothing$ and an AC for $\mathscr{L}_{\lambda_0}$ denoted by P_0 are given.

ALGORITHM. As initial conditions, substitute $\lambda^{(0)} := \lambda_0$ and $P_{\mathrm{AC}}(\lambda^{(0)}) := P_0$, and let ϵ_1 and ϵ_2 be sufficiently small positive constants.

Step 1. Let $\lambda^{(l)} := \lambda^{(l-1)} + \lambda_{\min}(P_{\mathrm{AC}}(\lambda^{(\lambda-1)})) - \epsilon_1$.

Step 2. Find the AC for $\mathscr{L}_{\lambda^{(l)}}$, and let $P_{\mathrm{AC}}(\lambda^{(l)})$ denote it.

Step 3. If $\lambda_{\min}(P_{\mathrm{AC}}(\lambda^{(l)})) < \epsilon_2$, then stop and let $\lambda^* := \lambda^{(l)}$. Otherwise, continue.

Step 4. Return to step 1.

REMARK 5. If $\lambda^{(l)} > 0$, a point $P \in \mathscr{L}$ is obtained as $P := P_{\mathrm{AC}}(\lambda^{(l)}) + \lambda^{(l)} I$.

The most important part in the above algorithm is step 2. According to Corollary 3, we can find $P_{AC}(\lambda^{(l)})$ by the following two methods:

(I) Let $Q_2^{(l)}$ be any point in $\mathscr{L}^{\perp}$ [e.g., $Q_2^{(l)} = P_{AC}(\lambda^{(l-1)})$], and find the ∇^*-projection of $Q_2^{(l)}$ to $\mathscr{L}_{\lambda^{(l)}}$. Then it is $P_{AC}(\lambda^{(l)})$.

(II) Let $Q_1^{(l)}$ be any point in $\mathscr{L}_{\lambda^{(l)}}$ [e.g., $Q_1^{(l)} = P_{AC}(\lambda^{(l-1)}) - (\lambda^{(l)} - \lambda^{(l-1)})I$], and find the ∇-projection of $Q_1^{(l)}$ to $\mathscr{L}^{\perp}$. Then it is $P_{AC}(\lambda^{(l)})$. (See Figure 1.)

Method (I) is based on Corollary 3 (iii). As described at the end of Section 4, finding each $P_{AC}(\lambda^{(l)})$ as the ∇^*-projection is cast to a convex programming problem to find P that minimizes $D^*(Q_2^{(l)}, P)$ on $\mathscr{L}_{\lambda^{(l)}}$ using the primal coordinate x^{α}, $\alpha = 1, \dots, m$. The algorithm with method (I) is essentially equivalent to that proposed by [3]. On the contrary, method (II) corresponds to Corollary 3(ii). From (30) and (19),

$$\partial^{\kappa}\partial^{\mu}D(Q_1, P(\tilde{y})) = \partial^{\kappa}\partial^{\mu}\phi(P(\tilde{y})) = g^{\kappa\mu}(\tilde{y})$$

is positive definite for given $Q_1 \in \mathrm{PD}(n)$. Hence, finding each $P_{AC}(\lambda^{(l)})$ as the ∇-projection, which is equivalent to finding P that minimizes $D(Q_1^{(l)}, P)$ on $\mathscr{L}^{\perp}$, is also a convex programming problem using the dual coordinate y_{κ},

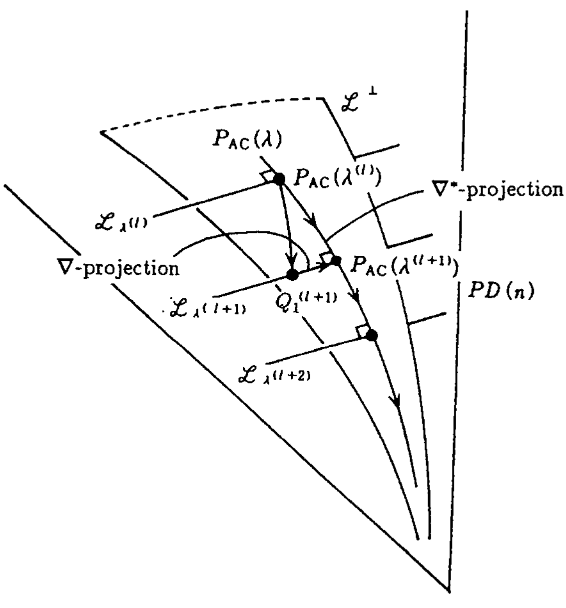

FIG. 1. Following the path of the analytic center by ∇- and ∇^*-projections.

$\kappa = m + 1, \ldots, N$. The algorithm with method (II) using this dual coordinate system $\tilde{y}$ may be new. Considering the dimension of the variables, we can easily guess that the algorithm with method (I) [method (II)] is faster when $m \ll N/2$ [$m \gg N/2$].

Thus, by means of the dualistic geometry, we can solve problem (a) and give geometric interpretations of the above two algorithms. The remarkable point is that the Hesse matrices of the objective function in the above two convex programming problems coincide with the Riemannian metrics $g_{\alpha\beta}$ and $g^{\kappa\mu}$, respectively.

6. CONCLUSIONS

We have exploited the theory of dualistic differential geometry for the set of positive definite matrices. We find that several kinds of duality exist and play essential roles in the theory.

In addition, we have elucidated relations between the theory and some mathematical problems, i.e., MA, PDMC, and LMI. For these problems, the theory can provide new results and nice geometrical insight. Since positive definite matrices appear in various fields of the engineering, there is possibility of its further applications to practical problems.

We are grateful to anonymous reviewers and the corresponding editor for their constructive suggestions.

APPENDIX A. THE COMPONENTS OF THE CONNECTIONS ∇ AND ∇^*

Let $c : P(t)$ be any smooth curve which passes through a given point $P = P(0)$, and $\Pi_c(t)$ and $\Pi_c^*(t)$ denote the parallel displacements from $T_p\mathrm{PD}(n)$ to $T_{p(t)}\mathrm{PD}(n)$ along the curve c defined in (7) and (8). Then *covariant derivative* of a vector field X to the direction $\dot{P}(0)$ is derived from parallel displacement as

$$\nabla_{\dot{P}(0)} X = \lim_{t \to 0} \frac{1}{t} \left\{ \Pi_c(t)^{-1} X_{p(t)} - X_P \right\}.$$

Since the parallel displacement $\Pi_c(t)$ does not change tangent vectors, we get $\nabla \partial_i \, \partial_j = 0$. Hence, (9) follows.

To obtain $\Gamma^*_{ijk}(\theta)$, consider a specific curve $\bar{c} : P(t)$ defined by

$$P(t) := P^{1/2} \exp(Xt) \, P^{1/2}.$$

The curve $\bar{c}$ is found to satisfy

$$P(0) = P \quad \text{and} \quad \dot{P}(0) \equiv P^{1/2} X P^{1/2}. \tag{42}$$

Hence, to calculate the covariant derivative $\nabla^*_{\partial_i} \partial_j$ using the curve $\bar{c}$, we shall set

$$X := P^{-1/2} E_i P^{-1/2}.$$

Then $\dot{P}(0) = (\partial_i)_P \equiv E_i$ and from the definition of covariant derivative, we can obtain

$$\left(\nabla^*_{\dot{P}(0)} \partial_j\right)_P = \left(\nabla^*_{\partial_i} \partial_j\right)_P \equiv \lim_{t \to 0} \frac{1}{t} \left\{ \Pi^*_{\bar{c}}(t)^{-1} \left(\partial_j\right)_{P(t)} - \left(\partial_j\right)_P \right\}. \tag{43}$$

Using the matrix representation of $\Pi^*_{\bar{c}}(t)$ in (8), its inverse $\Pi^*_{\bar{c}}(t)^{-1} : T_{P(t)}\mathrm{PD}(n) \to T_P \mathrm{PD}(n)$ is

$$\begin{aligned} \Pi^{*-1}_{\bar{c}(t)} \left(\partial_j\right)_{P(t)} &\equiv P P(t)^{-1} E_j P(t)^{-1} P \\ &= P^{1/2} \exp(-Xt) \, P^{-1/2} E_j P^{-1/2} \exp(-Xt) \, P^{1/2}. \end{aligned}$$

Hence, we substitute this expression in (43) to get

$$\begin{aligned} \left(\nabla^*_{\partial_i} \partial_j\right)_p &\equiv \frac{d}{dt} \left\{ P^{1/2} \exp(-Xt) \, P^{-1/2} E_j P^{-1/2} \exp(-Xt) \, P^{1/2} \right\} \Big|_{t=0} \\ &= -P^{1/2} X P^{-1/2} E_j - E_j P^{-1/2} X P^{1/2} \\ &= -E_i P^{-1} E_j - E_j P^{-1} E_i . \end{aligned}$$

Thus, (10) follows from (12).

APPENDIX A. VERIFICATION OF THE PROPERTIES IN LEMMA 1 FOR THE DUAL COORDINATE SYSTEM η AND THE POTENTIAL FUNCTIONS $\phi(\theta)$ AND $\psi(\eta)$ DEFINED IN (23) AND (24)

By means of the representation (2), the basis tangent vector ∂^i at P^{-1} is identified as

$$(\partial^i)_{P^{-1}} \in T_{P^{-1}}\mathrm{PD}(n) \equiv E^i \in \mathrm{Sym}(n).$$

If we pull back $(\partial^i)_{p^{-1}} \in T_{p^{-1}}\mathrm{PD}(n)$ to $T_p\mathrm{PD}(n)$ via the differential ι_*, then

$$(\partial^i)_P \equiv (\iota_*)^{-1} E^i = -PE^iP.$$

Hence,

$$g(\partial_i, \partial^j) = \mathrm{tr}\{P^{-1}E_i P^{-1}(-PE^jP)\} = \delta_i^j.$$

Using ι_*, similarly, we find (8) implies $\Gamma^{*ijk}(\eta) = 0$.

Next, note the relations $\eta_i = -\mathrm{tr}(P^{-1}E_i)$ and $\theta^i = -\mathrm{tr}(PE^i)$ hold. Then (18) and (19) are derived from Taylor expansion of $\psi(P)$ for $dP = E_i\,d\theta^i$:

$$\begin{aligned}
\psi(P + dP) - \psi(P) &= -\log\frac{\det(P + dP)}{\det P} = -\log\det(I + P^{-1}\,dP)\\
&= -\mathrm{tr}\log(I + P^{-1}\,dP)\\
&= \mathrm{tr}\sum_{k=1}^{\infty} k^{-1}(-1)^k(P^{-1}\,dP)^k\\
&= \eta_i\,d\theta^i + \tfrac{1}{2}g_{ij}\,d\theta^i\,d\theta^j + \frac{1}{3!}\Gamma^*_{ijk}\,d\theta^i\,d\theta^j\,d\theta^k + \cdots
\end{aligned}$$

and that of $\phi(P)$ for $dP^{-1} = E^i\,d\eta_i$ similarly. The equation (20) is obvious from (19).

Finally, at any point $P = P(\theta) = P(\eta)$, the equation (21) follows from the identity

$$\operatorname{tr} P(\theta) P^{-1}(\eta) = \operatorname{tr} I = n$$

$$= \operatorname{tr}\left[\left(\theta^i E_i\right)\left(\eta_j E^j\right)\right] = \theta^i \eta_j \operatorname{tr}\left(E_i E^j\right) = -\theta^i \eta_i.$$

REFERENCES

1 S. Amari, *Differential-Geometrical Methods in Statistics*, Springer-Verlag, 1985.
2 G. P. Barker, Theory of cones, *Linear Algebra Appl.* 39:263–291 (1981).
3 S. Boyd and L. El Ghaoui, Method of centers for minimizing generalized eigenvalues, *Linear Algebra Appl.* 188, 189:63–111 (1993).
4 S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, 1994.
5 J. P. Burg, D. G. Luengberger, and D. L. Wenger, Estimation of structured covariance matrices, *Proc. IEEE* 70(9):963–974 (1982).
6 B. De Moor, Structured total least squares and L_2 approximation problems, *Linear Algebra Appl.* 188, 189:163–205 (1993).
7 H. Dym and I. Gohberg, Extension of band matrices with band inverses, *Linear Algebra Appl.* 36:1–24 (1981).
8 L. Faybusovich, A Hamiltonian structure for generalized affine-scaling vector fields, *J. Nonlinear Sci.*, to appear.
9 R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkovicz, Positive definite completions of partial Hermitian matrices, *Linear Algebra Appl.* 58:109–124 (1984).
10 S. Helgason, *Differential Geometry and Symmetric Spaces*, Academic, New York, 1962.
11 R. D. Hill, On the cone of positive semidefinite matrices, *Linear Algebra Appl.* 90:81–88 (1987).
12 C. R. Johnson, Positive definite completions: A guide to selected literature, in *Signal Processing Part I: Signal Processing Theory* (L. Auslander et al., Eds.), Springer-Verlag, 1990, pp. 169–188.
13 S. Kobayashi and K. Nomizu, *Foundation of Differential Geometry I*, Wiley, 1963.
14 S. L. Kullback, *Information Theory and Statistics*, Wiley, New York, 1959.
15 Yu. Nesterov and A. Nemirovsky, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, 1994.
16 A. Ohara and T. Kitamori, Distances on Gramians and their applications, in *Recent Advances in Mathematical Theory of Systems, Control, Networks and Signal Processing I, Proceedings of MTNS*-91 (H. Kimura and S. Kodama, Eds.), MIT Press, 1992, pp. 449–454.

17 A. Ohara, N. Suda, and S. Amari, Dualistic differential geometry of positive definite matrices and its application to matrix approximation and completion problems, in *Proceedings of the MTNS*-93, Regensburg, Germany, Aug. 1993.
18 E. B. Vinberg, The theory of convex homogeneous cones, *Trans. Moscow Math. Soc.* 12:340–403 (1963).
19 P. W. Vos, Fundamental equations for statistical submanifolds with applications to the Bartlett correction, *Ann. Inst. Statist. Math.* 41(3):429–450 (1989).

Received 18 July 1994; final manuscript accepted 1 December 1994