A PSEUDO-RIEMANNIAN-GRADIENT APPROACH TO THE LEAST-SQUARES PROBLEM ON THE REAL SYMPLECTIC GROUP

Simone Fiori

Dipartimento di Ingegneria Biomedica, Elettronica e Telecomunicazioni (DiBET)
Facoltà di Ingegneria, Università Politecnica delle Marche
Via Brecce Bianche, Ancona I-60131, Italy
eMail: s.fiori@univpm.it
URL: http://web.dibet.univpm.it/fiori

ABSTRACT

The present paper discusses the problem of geodesic least-squares over the real symplectic group of matrices $\text{Sp}(2n, \mathbb{R})$. As the space $\text{Sp}(2n, \mathbb{R})$ is a non-compact Lie group, it is convenient to endow it with a pseudo-Riemannian geometry instead of a Riemannian one. Indeed, a pseudo-Riemannian metric allows the computation of geodesic arcs and geodesic distances in closed form.


1. INTRODUCTION

The standard ‘least squares’ method is used to approximately solve overdetermined systems, i.e., systems of equations in which there are more equations than unknowns. Least-squares problems fall into two categories: linear and non-linear. Linear least-squares problems admit a closed-form solution, while non-linear problems do not and are usually solved by iterative refinement. The standard least-squares method may be extended and generalized in order to accommodate a broader range of applications.

On a Riemannian manifold $M$, the least-square problem may be formulated via the criterion function $f : M \rightarrow \mathbb{R}^+_0$:

$$f(x) \stackrel{\text{def}}{=} \sum_i d^2(x, \tau_i),$$

where $\tau_i \in M$ denote target matrices and the function $d(\cdot, \cdot)$ denotes a geodesic distance on the manifold $M$ corresponding to the metric that the manifold is endowed with.

On non-compact Riemannian Lie groups, the formulation of a least-squares criterion and its optimization are rather involved problems, because it might be hard to compute geodesic distances in closed form. However, a non-compact Lie group may be treated as a pseudo-Riemannian manifold with a bi-invariant pseudo-Riemannian metric [10]. In the pseudo-Riemannian context, geodesic arcs, geodesic distances and pseudo-Riemannian gradient may be calculated in closed form, therefore, Riemannian-gradient-based optimization of a least-squares criterion may be amenable.

One of such non-compact Lie groups is the real symplectic group, denoted as $\text{Sp}(2n, \mathbb{R})$. Real symplectic matrices play an important role in applied fields. Some applications are: Coding theory [3], time-series prediction [1], control of beam systems in particle accelerators [4], quantum computing [7] and vibration analysis [11].

In the present manuscript, we discuss the problem of least-squares on the real symplectic group. After a review of results known from literature about optimization on the real symplectic manifold, an optimization method based on endowing it with a pseudo-Riemannian geometry will be discussed.

2. OPTIMIZATION ON SMOOTH MANIFOLDS

For a reference on differential geometry, see [12]. Let $M$ be a Riemannian manifold. The tangent space at $x \in M$ to the manifold is denoted by $T_x M$. A metric on $M$ is a non-degenerate, smooth, symmetric, bilinear map which assigns a real number to pairs of tangent vectors at each tangent space of the manifold $M$. Let us denote the metric by $\langle \cdot, \cdot \rangle_x : T_x M \times T_x M \rightarrow \mathbb{R}$. The metric $\langle \cdot, \cdot \rangle_x$ also defines the squared norm $\|v\|^2_x \stackrel{\text{def}}{=} \langle v, v \rangle_x$ for $v \in T_x M$.

The geodesic curve connecting two points $x_1, x_2 \in M$ is the curve $G(t) \in M$, parameterized by $t \in [0, 1]$, that minimizes the energy integral:

$$\int_0^1 \langle \dot{x}, \dot{x} \rangle_x dt, \text{ such that } \langle \dot{G}, \dot{G} \rangle_G \text{ is constant}. \quad (2)$$

By the calculus of variation on manifold, the solution of the geodesic equation may be written in normal form as:

$$\ddot{x} + \Gamma_x(\dot{x}, \dot{x}) = 0. \quad (3)$$

In the above expressions, the over-dot and the double over-dot denote first-order and second-order derivation with respect to
parameter \( t \), respectively, while symbol \( \Gamma(\cdot,\cdot,\cdot) \) denotes the Christoffel operator. The solution of the geodesic equation may be written in terms of two known quantities that play a role in the boundary conditions for the second-order geodesic differential equation. When the values \( x(0) = x \in M \) and \( \dot{x}(0) = v \in T_x M \) are specified, we shall denote the solution as \( G_{x,v}(t) \).

The squared geodesic distance between the geodesic’s endpoints is defined as:

\[
d^2(x_1, x_2) \overset{\text{def}}{=} \left( \int_0^1 \sqrt{\langle G(t), G(t) \rangle} \, dt \right)^2 = \langle G(1), G(1) \rangle_{G}^{t=0}. \tag{4}
\]

Note that if the geodesic curve is expressed as \( G_{x,v}(t) \), then the squared geodesic distance equals \( \|v\|^2_x \).

The Riemannian gradient \( \nabla_x f \in T_x M \) of a regular criterion function \( f : M \to \mathbb{R} \) in a point \( x \in M \) may be defined as:

\[
\nabla_x f \overset{\text{def}}{=} (df)_x^\sharp,
\]

where symbol \( ^\sharp \) denotes the ‘sharp’ isomorphism and symbol \( df \) denotes differential.

A pseudo-Riemannian manifold is a manifold endowed with a metric that is not positive-definite. On a pseudo-Riemannian manifold \( M \), the quantity \( \|v\|^2_x \) may be positive, negative or null even for \( 0 \neq v \in T_x M \).

The basic idea to cope with pseudo-Riemannian manifolds is to partition each tangent space \( T_x M \) as follows:

\[
\begin{align*}
T^+_x M &\overset{\text{def}}{=} \{ v \in T_x M \text{ such that } \|v\|_x^2 > 0 \}, \\
T^0_x M &\overset{\text{def}}{=} \{ v \in T_x M \text{ such that } \|v\|_x^2 = 0 \}, \\
T^-_x M &\overset{\text{def}}{=} \{ v \in T_x M \text{ such that } \|v\|_x^2 < 0 \}.
\end{align*}
\]

The notion of geodesics may be defined on a pseudo-Riemannian manifold by the calculus of variation on the energy integral (2). A geodesic arc \( G_{x,v}(t) \) will be the solution of the differential equation (3). On a pseudo-Riemannian manifold \( M \), however, it holds \( \|v\|_x^2 < 0 \) for \( v \in T^+_x M \), therefore the notion of squared pseudo-Riemannian geodesic distance may be defined as:

\[
d^2(x_1, x_2) \overset{\text{def}}{=} \|v\|_{x_2}, \tag{7}
\]

where \( x \in M \) and \( v \in T_x M \) are the parameters of the geodesic arc connecting points \( x_1 \) and \( x_2 \) on \( M \).

The definition (5) holds even in the case of pseudo-Riemannian gradient. The gradient steepest descent optimization equation on a pseudo-Riemannian manifold reads:

\[
\dot{x} = \begin{cases} 
-\nabla_x f & \text{if } \nabla_x f \in T^+_x M \cup T^0_x M, \\
\nabla_x f & \text{if } \nabla_x f \in T^-_x M.
\end{cases} \tag{8}
\]

Note that it differs from the gradient descent equation on a Riemannian manifold (see, e.g., [5]). It induces the dynamics:

\[
\dot{f} = \begin{cases} 
-\|\nabla_x f\|_x^2 & \text{if } \nabla_x f \in T^+_x M \cup T^0_x M, \\
\|\nabla_x f\|_x^2 & \text{if } \nabla_x f \in T^-_x M
\end{cases} \leq 0. \tag{9}
\]

The equation (8) may be solved by the numerical optimization algorithm:

\[
x_{k+1} = \begin{cases} 
G_{x,v,-\nabla_x f}(\eta) \text{ if } \nabla_x f \in T^+_x M \cup T^0_x M, \\
G_{x,v,\nabla_x f}(\eta) \text{ if } \nabla_x f \in T^-_x M,
\end{cases} \tag{10}
\]

where symbol \( G_{x,v}(t) \) denotes a pseudo-Riemannian geodesic arc departing from the point \( x \in M \) with initial direction \( v \in T_x M \) and parameter \( t \in [0,1] \). The parameter \( \eta > 0 \) plays the role of optimization stepsize.

### 3. THE REAL SYMPLECTIC GROUP

The present section aims at recalling the definition of the real symplectic group and its properties, along with some recent results about optimization on it.

The real symplectic group is defined as follows:

\[
\text{Sp}(2n, \mathbb{R}) \overset{\text{def}}{=} \{ x \in \mathbb{R}^{2n \times 2n} | x^T q x = q \}, \tag{11}
\]

\[
q \overset{\text{def}}{=} \begin{bmatrix} 0_n & e_n \\
-e_n & 0_n \end{bmatrix}, \tag{12}
\]

where symbol \( e_n \) denotes the \( n \times n \) identity matrix, symbol \( 0_n \) denotes a whole-zero \( n \times n \) matrix and superscript \( T \) denotes matrix transpose.

The space \( \text{Sp}(2n, \mathbb{R}) \) is a curved smooth manifold of dimension \( n(2n + 1) \) that may also be endowed with smooth algebraic-group structure (namely, group multiplication and group inverse and possesses a identity element) in a manner that is compatible with the manifold structure. Therefore, the space \( \text{Sp}(2n, \mathbb{R}) \) has the structure of a Lie group. In particular, standard matrix multiplication and inverse work as algebraic group operations.

The tangent space \( T_x \text{Sp}(2n, \mathbb{R}) \) has structure:

\[
T_x \text{Sp}(2n, \mathbb{R}) = \{ v \in \mathbb{R}^{2n \times 2n} | v^T q x + x^T q v = 0_{2n} \}. \tag{13}
\]

The tangent space at the identity of the Lie group, namely the Lie algebra \( \text{sp}(2n, \mathbb{R}) \), has structure:

\[
\text{sp}(2n, \mathbb{R}) = \{ h \in \mathbb{R}^{2n \times 2n} | h^T q + q h = 0 \}. \tag{14}
\]

To the best of the present author knowledge, the most advanced results about optimization on the manifold \( \text{Sp}(2n, \mathbb{R}) \) find in the contribution [2]. Let \( \sigma : \text{sp}(2n, \mathbb{R}) \to \text{sp}(2n, \mathbb{R}) \) be a symmetric positive-definite operator with respect to the Euclidean inner product \( \langle \cdot, \cdot \rangle_E \) on the space \( \text{sp}(2n, \mathbb{R}) \) given by \( \langle h_1, h_2 \rangle_E = \text{tr}(h_1^T h_2) \) for every \( h_1, h_2 \in \text{sp}(2n, \mathbb{R}) \). The minimizing curve of the integral:

\[
\int_{t_1}^{t_2} \langle h, \sigma(h) \rangle_E dt
\]

over all curves \( x(t) \in \text{Sp}(2n, \mathbb{R}) \) with \( t \in [t_1, t_2] \) and with fixed endpoints \( x(t_1) = x_1 \in \text{Sp}(2n, \mathbb{R}) \) and \( x(t_2) = x_2 \in \text{Sp}(2n, \mathbb{R}) \).
Sp(2n, R), and where h is defined by \( \dot{x} = xh \), so that \( h \in \text{sp}(2n, \mathbb{R}) \), is the solution of the system:

\[
\begin{cases}
\dot{x} = xh, \\
\dot{m} = \sigma^T(h)m - m\sigma^T(h), \\
h = \sigma^{-1}(m),
\end{cases}
\]

where symbol \( \sigma^{-1} \) denotes the inverse of the operator \( \sigma \).

The simplest choice for the symmetric positive-definite operator \( \sigma \) is \( \sigma(h) = h \). The above choice for the operator \( \sigma \) implies that \( m = h \) and that the corresponding geodesic equation on the real symplectic group satisfies the equations:

\[
\dot{x} - \dot{x}x^{-1}\dot{x} + x\dot{x}^Txqqx^{-1}\dot{x} - \dot{x}\dot{x}^Tqq = 0.
\]

(16)

Closed-form solutions of the above equations are unknown to the authors of [2] and to the present author.

Let us consider the following pseudo-Riemannian metric on the Lie group \( \text{Gl}(n) \) [10]:

\[
\langle u, v \rangle_x \overset{\text{def}}{=} \text{tr}(x^{-1}ux^{-1}v), \quad \forall u, v \in T_x \text{Gl}(n).
\]

(17)

Under the above pseudo-Riemannian metric, it is indeed possible to solve the geodesic equation in closed form. The energy integral in this case reads:

\[
\int_0^1 \text{tr}((x^{-1}\dot{x})^2)dt,
\]

(18)

and the corresponding geodesic curve and the squared geodesic distance have the following expressions:

\[
G_{x,v}(t) = x \exp(tx^{-1}v),
\]

\[
d^2(x_1, x_2) = \text{tr}(|\log^2(x_1^{-1}x_2)|).
\]

(19)

(20)

The pseudo-Riemannian gradient of a regular function \( f : \text{Sp}(2n, \mathbb{R}) \rightarrow \mathbb{R} \) is the solution of the compatibility condition:

\[
\text{tr}(\partial_x^Tvf) = \text{tr}(x^{-1}\nabla_xfx^{-1}v),
\]

that reads:

\[
\nabla_xf = \frac{1}{2} \left( q\partial_xfq + x(\partial_xf)^Tx \right).
\]

(21)

(22)

Having chosen a pseudo-Riemannian metric for the real symplectic group, the expression of the geodesic arc as well as the expression of the geodesic distance may be computed in closed form. The least-squares problem (1) may be thus set up, where target matrices \( \tau_i \) belong to \( \text{Sp}(2n, \mathbb{R}) \). Moreover, the numerical scheme (10) may be effectively implemented.

4. NUMERICAL TESTS

As a numerical test, consider the computation of the mean of a set of given real symplectic matrices. An application is described in [8].

On a metrizable manifold \( M \), the ‘intrinsic mean’ may be defined as [6, 9]:

\[
\mu \overset{\text{def}}{=} \arg \min_{x \in M} \sum_i d^2(x, \tau_i),
\]

(23)

where the matrices \( \tau_i \in M \) are distributed around a center-of-mass to be estimated by \( \mu \in M \). Setting \( M = \text{Sp}(2n, \mathbb{R}) \), the problem (23) is a least-squares problem on \( \text{Sp}(2n, \mathbb{R}) \).

The Figure 1 shows a result obtained with the iterative algorithm (10) for \( n = 5 \). The picture shows the value of the criterion function in (23) as well as the value of the squared norm of its pseudo-Riemannian gradient during iteration. The Figure also shows the squared distances \( d^2(x, \tau_i) \) before iteration (with initial guess chosen as \( \epsilon_{10} \)) and after the iteration. The Figure shows that the algorithm converges steadily to-

![Fig. 1. Optimization over the real symplectic group Sp(10, R).](image-url)

ward the minimal criterion value (in fact, the distances from the found center of mass are much smaller than the distances from the initial guess).

A close-up of the numerical behavior of the least-squares optimization algorithm (10) comes from the examination of the case \( n = 1 \). The group \( \text{Sp}(2, \mathbb{R}) \) is a 3-dimensional manifold, therefore the following parametrization may be taken advantage of:

\[
\mathbb{R}^3 \ni (a, b, c) \rightarrow \begin{bmatrix} a & b \\ c & z \end{bmatrix} \in \text{Sp}(2, \mathbb{R}),
\]

(24)

(z determined by condition \( az - bc = 0 \)).

Hence, the elements of the group \( \text{Sp}(2, \mathbb{R}) \) may be rendered on a 3-dimensional figure.
The Figure 2 shows a result obtained with the iterative algorithm (10) for \( n = 1 \). The Figure shows the location of the target matrices \( \tau_i \) (circles), the location of the center-of-mass (cross), the trajectory of the optimization algorithm over the search space (solid-dotted line) and the location of the final point computed by the algorithm (diamond). The Figure shows that the algorithm is convergent toward the center of mass. (Because of finite sample size, the empirical center of mass differs from the actual one.)

5. CONCLUSION

The standard least-squares problem may be extended and generalized to smooth curved parameter spaces. The present paper discussed the problem of geodesic least-squares over the real symplectic group of matrices.

We suggested to regard the real symplectic group as a pseudo-Riemannian manifold and chose a metric that allows for the computation of closed-forms for the geodesic arcs and hence for the geodesic distance. On the basis of these findings, the geodesic least-squared problem may be properly set up and the geodesic-based numerical stepping method may be properly implemented.

Numerical tests have been performed with reference to the computation of the intrinsic mean of a collection of symplectic matrices. Numerical results show that the developed optimization algorithm is suitable.

6. REFERENCES


