ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 17, No. 2, 2022, pp.154-160

New Kantorovich's theorems for Newton's method on Lie groups for mappings and matrix optimization problems

Li Nan[1]
School of Mathematics and Statistics, Nanjing University of information Science and Technology, Nanjing 210044, China.

(Received August 10, 2022, accepted October 11, 2022)

Abstract: We propose a new Kantorovich theorem for Newton's method on Lie groups for mappings and matrix low-rank optimization problems, which arises from many applications. Under the classical hypothesis of f, we establish the convergence criteria of Newton's method from Lie group to its Lie algebra with weakened conditions, which improves the corresponding results in [20].

keywords: Lie group; Newton's method; Trace function; Lipschitz condition.

1. Introduction

In recent years, more and more attentions have been focused on studying numerical algorithms on manifolds. Classical optimization problems on manifolds are given by symmetric eigenvalue problems, low-rank nearest correlation matrix estimation, invariant subspace computations, optimization problems with equality constraints (see [7][9][21]). In this paper we focus on optimization problems on Lie groups. Consider the following problem:

$$\min_{x\in M} \phi(x), \tag{1.1}$$

where M is a Riemannian manifold and ϕ is a real-valued function on M. We will explore the optimization problem when ϕ is matrix trace function. It is essentially a kind of constrained matrix optimization problem. Many scholars have studied the problem. In [20], $\phi: G \to \mathbb{R}$ in (1.1) be given by

$$\phi(x) = -\mathrm{tr}(x^{\mathrm{T}}CxQ) \text{ for each } x \in G, \tag{1.2}$$

where $G = \mathrm{SO}(n,\mathbb{R}) := \{x \in \mathbb{R}^{n\times n} | x^{\mathrm{T}}x = \mathrm{I}_n, \det x = 1\}$, C is a fixed symmetric matrix and $Q = \mathrm{diag}(0_{n-\varsigma,n-\varsigma}, Q_\varsigma)$ with $Q_\varsigma = \mathrm{diag}(q_1,\cdots,q_\varsigma), 0 < q_1 \le q_2 \le \cdots \le q_\varsigma$, solved a kind of matrix trace function optimization problem with orthogonal constraints. Xu solved a generalized singular value of a Grassmann matrix pair or a real matrix pair. If $Q_\varsigma = \mathrm{diag}(\mathrm{I}_\varsigma, 0_{n-\varsigma,n-\varsigma})$ for $1 \le \varsigma \le n$, Xu solved this case by Riemannian inexact Newton-CG method [21]. Different from method in [21], We consider Newton's method on Lie group to solve this problem.

Brockett studied the optimization problem when

$$\phi(x) = -\mathrm{tr}(x^{\mathrm{T}}QxD) \text{ for each } x \in G \tag{1.3}$$

in (1.1), where Q is a fixed symmetric matrix and D with the following structure

$$D = \mathrm{diag}(1,2,\dots,n),$$

showed that the minimum $x^* \in G$ occurs when ${x^*}^{\mathrm{T}}Qx^*$ is a diagonal matrix with diagonal entries (eigenvalues

[1] Corresponding Author E-mail: linan3749@126.com.

of Q) in ascending order [3][4]. Sato & Iwai studied the maximum value of the following functions on Riemannian manifolds:

$$\mathrm{tr}(U^{\mathrm{T}}AVN),$$

where $U \in \mathbb{R}^{m\times p}, V \in \mathbb{R}^{n\times p}$ and $U^{\mathrm{T}}U + V^{\mathrm{T}}V = \mathrm{I}_p, N \in \mathbb{R}^{p\times p}$ is a diagonal matrix, $A \in \mathbb{R}^{m\times n}$. The global optimal solution of this problem provides a set of left and right singular vectors, and transforms the problem of matrix trace function into finding singular values and singular vectors of A [16]. Mahony developed the Newton method with a single-parameter subgroup on the Lie group, and proved the local convergence [12]; Xu designed the Newton-CG method for the Grassmannian manifold problem to solve the singular value of the matrix pair [21].

Lie groups were originally used to solve differential equations. For solving ordinary differential equations on Lie groups, Owren and Welfert used the implicit Euler method for Lie groups [14]. Newton method is an effective method for solving approximate solutions of equations, and is widely used in large-scale optimal control problems, constrained smooth and non-smooth problems (see [13][15]). In Banach space, Kantorovich's theorem (see [10]) is an important result on Newton's method. It ensures the quadratic convergence of Newton's method, the existence and local uniqueness of the solution under very mild assumptions that the second Fréchet derivative of f is bounded on a proper open metric ball of the initial point x_0. Smith studied Newton's method in Riemannian manifolds [17][18], Ferreira and Svaiter generalized Kantorovich's theorem of Newton's method in Riemannian manifolds [5]. Li introduced the concept of the γ condition of the map f and established the γ condition of the Newton's method of the map f, extending and developing Smale's α-theory and γ-theory [11]. Wang established Kantorovich's theorem for Newton's method on Lie groups, under the classical assumption of the map f, they proved the convergence criterion of Newton's method to the zeros of the map f, and obtained the estimation of the convergence domains [20]. He established the unique ball of a zero of a map on Lie group and an estimation of the radius of convergence ball by Newton's method on a Lie group [6]. Argyros presented the local convergence analysis of Newton's method, obtained a larger convergence ball and a more precise distance error bound [1]. Argyros demonstrated semi-local convergence of Newton's method with sufficiently weak convergence criteria and tighter distance error bounds [2].

In this paper, we propose New Kantorovich's theorems for Newton method on Lie groups for mappings and matrix low-rank optimization problems. Under the classical assumption of f, we establish the convergence criterion of Newton's method from Lie group to its Lie algebra with weakened conditions. The rest of this paper is organized as follows. In Section 2 some useful notations, and lemma are given. In Section 3 we will give some theorems and an algorithm. Finally, in Section 4 concluding remarks are drawn.

1. Notions and preliminaries

Most of the notions and notation that are used in the present paper are standard. $\mathbb{R}$ and $\mathbb{R}^{n\times n}$ denote the sets of real numbers and $n \times n$ matrices with entries in $\mathbb{R}$. The symbols I_n and $0_{m\times n}$ represent the n-order identity matrix and the $m \times n$ zeros matrix, $\mathrm{tr}(\cdot)$ denote the trace function. A Lie group $(G,\cdot)$ is both a manifold and a topological group, and its group multiplication map and inverse map are both C^∞. We assume that the Lie group G is n-dimensional. The symbol e denotes the identity element of G. The tangent space T_eG of G at e is the Lie algebra of the Lie group G, and is also the set of all left-invariant vector fields of G, denoted as $\mathcal{G}$, equipped with the Lie bracket $[\cdot,\cdot]: \mathcal{G} \times \mathcal{G} \longrightarrow \mathcal{G}$. For any element x in the Lie group G, T_xG represents the tangent space of x.

Next, we will introduce some definitions that will be used. We define for each $y \in G$ the left translation $L_y: G \rightarrow G$ by

$$L_y(z) = y \cdot z \quad \text{for each} \quad z \in G. \tag{2.1}$$

The differential of L_y at z is denoted by $(L_y{}')z$, which determines a linear isomorphism from the tangent space T_zG to $T_{(y\cdot z)}G$. The exponential map exp

$$\begin{aligned} \exp: \quad & \mathcal{G} \rightarrow G \\ & u \mapsto \exp(u) \end{aligned}$$