Signal Processing 89 (2009) 1704–1714

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Conjugate gradient algorithm for optimization under unitary matrix constraint

Traian Abrudan [*,1,2], Jan Eriksson [2], Visa Koivunen

SMARAD CoE, Department of Signal Processing and Acoustics, Helsinki University of Technology, FIN-02015 HUT, Finland

ARTICLE INFO

Article history:
Received 8 August 2008
Received in revised form
27 November 2008
Accepted 9 March 2009
Available online 24 March 2009

Keywords:
Optimization
Unitary matrix constraint
Array processing
Subspace estimation
Source separation

ABSTRACT

In this paper we introduce a Riemannian algorithm for minimizing (or maximizing) a real-valued function $\mathcal{J}$ of complex-valued matrix argument $\mathbf{W}$ under the constraint that $\mathbf{W}$ is an $n \times n$ unitary matrix. This type of constrained optimization problem arises in many array and multi-channel signal processing applications.

We propose a conjugate gradient (CG) algorithm on the Lie group of unitary matrices $U(n)$. The algorithm fully exploits the group properties in order to reduce the computational cost. Two novel geodesic search methods exploiting the almost periodic nature of the cost function along geodesics on $U(n)$ are introduced. We demonstrate the performance of the proposed CG algorithm in a blind signal separation application. Computer simulations show that the proposed algorithm outperforms other existing algorithms in terms of convergence speed and computational complexity.

1. Introduction

Constrained optimization problems arise in many signal processing applications. In particular, we are addressing the problem of optimization under unitary matrix constraint. Such problems may be found in communications and array signal processing, for example, blind and constrained beamforming, high-resolution direction finding, and generally in all subspace-based methods. Another important class of applications is source separation and independent component analysis (ICA). This type of optimization problems occur also in multiple-input multiple-output (MIMO) communication systems. See [2,1,23] for recent reviews.

Commonly, optimization under unitary matrix constraint is viewed as a constrained optimization problem on the Euclidean space. Classical gradient algorithms are combined with different techniques for imposing the constraint, for example, orthogonalization, approaches stemming from the Lagrange multipliers method, or some stabilization procedures. If the unitary criterion is enforced by using such techniques, one may experience slow convergence or departures from the unitary constraint as shown in [2].

A constrained optimization problem may be converted into an unconstrained one on a different parameter space determined by the constrained set. The unitary matrix constraint considered in this paper determines a parameter space which is the Lie group of $n \times n$ unitary matrices $U(n)$. This parameter space is a Riemannian manifold [6] and a matrix group [19] under the standard matrix multiplication at the same time. By using modern tools of Riemannian geometry, we take full benefit of the nice geometrical properties of $U(n)$ in order to solve the optimization problem efficiently and satisfy the constraint with high fidelity at the same time.

Pioneering work in Riemannian optimization may be found in [21,14,28,3]. The optimization with orthogonality constraints is considered in detail in [7]. Steepest descent

* Corresponding author. Tel.: +358 9451 2934.
E-mail addresses: tabrudan@signal.hut.fi (T. Abrudan), jamaer@signal.hut.fi (J. Eriksson), visa@signal.hut.fi (V. Koivunen).

[1] Funded partly by GETA Graduate School in Electronics, Telecommunications and Automation.

[2] This work was supported by the Academy of Finland.

0165-1684/$ - see front matter
doi:10.1016/j.sigpro.2009.03.015

(SD), conjugate gradient (CG), and Newton algorithms on Stiefel and Grassmann manifolds are derived. A CG algorithm on the Grassmann manifold has also been proposed recently in [18]. A non-Riemannian approach, which is a general framework for optimization, is introduced in [22]. Modified SD and Newton algorithms on Stiefel and Grassmann manifolds are derived. SD algorithms operating on orthogonal group are considered recently in [24,8,26] and on the unitary group in [1,2]. Riemannian SD algorithms on some specific subgroups of the unitary Lie group $U(n)$, such as the unitary unimodular matrix group $SU(2)$ and the special unitary group $SU(3)$ have been recently considered by Fiori in [11,10], respectively. Fixed point (or leap-frog) type of algorithms that do not stem from a from first-order or second-order (CG or Newton-like) cost function optimization have been recently proposed for neural learning on $U(n)$ [9]. A CG on the special linear group is proposed in [30]. Algorithms in the existing literature [21,14,28,29,7,22] are, however, more general in the sense that they can be applied on more general manifolds than $U(n)$. For this reason when applied to $U(n)$, they do not take full benefit of the special properties arising from the Lie group structure of the manifold [2].

In this paper we derive a CG algorithm operating on the Lie group of unitary matrices $U(n)$. The proposed CG algorithm provides faster convergence compared to the existing SD algorithms [2,22] at even lower complexity. There are two main contributions in this paper. First, a computationally efficient CG algorithm on the Lie group of unitary matrices $U(n)$ is proposed. The algorithm fully exploits the Lie group properties such as simple expressions for the geodesics and tangent vectors.

The second main contribution in this paper is that we propose Riemannian optimization algorithms which exploit the *almost periodic* property [12] of the cost function along geodesics on $U(n)$. Based on this property we derive novel high accuracy *line search methods* [28] that facilitate fast convergence and selection of suitable step size parameter. Many of the existing geometric optimization algorithms do not include practical line search methods [7,24], or if they do, they are too complex when applied to optimization on $U(n)$ [22,8]. In some cases, the line search methods are either valid only for specific cost functions [28], or the resulting search is not highly accurate [22,8,18,2,26]. Because the CG algorithm assumes exact search along geodesics, the line search method is crucial for the performance of the resulting algorithm. An accurate line search method exploiting the periodicity of a cost function which appears on the limited case of special orthogonal group $SO(n)$ for $n\leqslant 3$, is proposed in [26] for non-negative ICA. The method can also be applied for $n>3$, but the accuracy decreases, since the periodicity of the cost function is lost. Univariate descent (UVD) methods on homogeneous manifolds have been recently proposed by Celledoni and Fiori [5]. The main advantage is that the multi-dimensional optimization is split into several one-dimensional optimization problems which are easier to solve.

The high-accuracy line search methods proposed in this paper use a different approach. They perform optimization of the cost function along geodesics by exploiting the almost periodic behavior of the cost function on $U(n)$, due to the exponential map. Their complexity is lower compared to well-known efficient methods [2] such as the Armijo method [27]. The proposed line search methods are used together with the proposed CG algorithm. To our best knowledge the proposed CG algorithm is the first ready-to-implement CG algorithm on the Lie group of unitary matrices $U(n)$. It is also valid for the special orthogonal group $SO(n)$.

This paper is organized as follows. In Section 2 we approach the problem of optimization under unitary constraint by using tools from Riemannian geometry. We show how the geometric properties may be exploited in order to solve the optimization problem in an efficient way. Two novel line search methods are introduced in Section 3. The practical CG algorithm for optimization under unitary matrix constraint is given in Section 4. Simulation results and applications are presented in Section 5. Finally, Section 6 concludes the paper.

2. Optimization on the unitary group $U(n)$

In this section we show how the problem of optimization under unitary matrix constraint can be solved efficiently and in an elegant manner by using tools of Riemannian geometry. In Section 2.1 we review some important properties of the unitary group $U(n)$ which are needed later in our derivation. Few important properties of $U(n)$ that are very beneficial in optimization are pointed out in Section 2.2. The difference in behavior of the SD and CG algorithm on Riemannian manifolds in explained in Section 2.3. A generic CG algorithm on $U(n)$ is proposed in Section 2.4.

2.1. Some key geometrical features of $U(n)$

This subsection describes briefly some Riemannian geometry concepts related to the Lie group of unitary matrices $U(n)$ and show how they can be exploited in optimization algorithms. Consider a real-valued function $\mathscr{J}$ of an $n\times n$ complex matrix $\mathbf{W}$, i.e., $\mathscr{J}:\mathbb{C}^{n\times n}\rightarrow\mathbb{R}$. Our goal is to minimize (or maximize) the function $\mathscr{J}=\mathscr{J}(\mathbf{W})$ under the constraint that $\mathbf{W}\mathbf{W}^H=\mathbf{W}^H\mathbf{W}=\mathbf{I}$, i.e., $\mathbf{W}$ is unitary. This constrained optimization problem on $\mathbb{C}^{n\times n}$ may be converted into an unconstrained one on the space determined by the unitary constraint, i.e., the Lie group of unitary matrices. We view our cost function $\mathscr{J}$ as a function defined on $U(n)$. The space $U(n)$ is a *real* differentiable manifold [6]. Moreover, the unitary matrices are closed under the standard matrix multiplication, i.e., they form a Lie group [19]. The additional properties arising from the Lie group structure may be exploited to reduce the complexity of the optimization algorithms.

2.1.1. Tangent vectors and tangent spaces

The tangent space $T_{\mathbf{W}}U(n)$ is an n^2-dimensional real vector space attached to every point $\mathbf{W}\in U(n)$. At the group identity $\mathbf{I}$, the tangent space is the *real Lie algebra* of skew-Hermitian matrices $\mathfrak{u}(n)\triangleq T_{\mathbf{I}}U(n)=\{\mathbf{S}\in\mathbb{C}^{n\times n}|\mathbf{S}=$