Linear Algebra and its Applications 524 (2017) 159–181

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Disentangling orthogonal matrices

Teng Zhang [a,*], Amit Singer [b]

[a] Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA
[b] Department of Mathematics and The Program in Applied and Computational Mathematics (PACM), Princeton University, Princeton, NJ 08544, USA

ARTICLE INFO

Article history:
Received 1 May 2016
Accepted 5 March 2017
Available online 9 March 2017
Submitted by V. Mehrmann

MSC:
15A24
90C22

Keywords:
SDP relaxation
Orthogonal Procrustes problem
Cryo-EM

ABSTRACT

Motivated by a certain molecular reconstruction methodology in cryo-electron microscopy, we consider the problem of solving a linear system with two unknown orthogonal matrices, which is a generalization of the well-known orthogonal Procrustes problem. We propose an algorithm based on a semi-definite programming (SDP) relaxation, and give a theoretical guarantee for its performance. Both theoretically and empirically, the proposed algorithm performs better than the naïve approach of solving the linear system directly without the orthogonal constraints. We also consider the generalization to linear systems with more than two unknown orthogonal matrices.

1. Introduction

In this paper, we consider the following problem: given known matrices $\boldsymbol{X}_1, \boldsymbol{X}_2 \in \mathbb{R}^{N\times D}$ and unknown orthogonal matrices $\boldsymbol{V}_1, \boldsymbol{V}_2 \in O(D)$, recover $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ from $\boldsymbol{X}_3 \in \mathbb{R}^{N\times D}$ defined by

* Corresponding author.
E-mail addresses: teng.zhang@ucf.edu (T. Zhang), amits@math.princeton.edu (A. Singer).

http://dx.doi.org/10.1016/j.laa.2017.03.002
0024-3795/

$$\boldsymbol{X}_3 = \boldsymbol{X}_1 \boldsymbol{V}_1 + \boldsymbol{X}_2 \boldsymbol{V}_2. \tag{1}$$

A naïve approach would be solving (1) while dropping the constraints of orthogonality on $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$. This linear system has ND linear constraints and $2D^2$ unknown variables, therefore, this approach can recover $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ when $N \geq 2D$. The question is, can we develop an algorithm that takes the constraints of orthogonality into consideration, so that it is able to recover $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ when $N < 2D$, and more stably when the observation $\boldsymbol{X}_3$ is contaminated by noise?

The associated least squares problem

$$\min_{\boldsymbol{V}_1, \boldsymbol{V}_2 \in O(D)} \|\boldsymbol{X}_1 \boldsymbol{V}_1 + \boldsymbol{X}_2 \boldsymbol{V}_2 - \boldsymbol{X}_3\|_F^2 \tag{2}$$

can be considered as a generalization of the well-known orthogonal Procrustes problem [1]:

$$\min_{\boldsymbol{V} \in O(D)} \|\boldsymbol{X}_1 \boldsymbol{V} - \boldsymbol{X}_2\|_F^2, \tag{3}$$

with the main difference being that the minimization in (2) is over two orthogonal matrices instead of just one as in (3). Although the orthogonal Procrustes problem has a closed form solution using the singular value decomposition, problem (2) does not enjoy this property.

Still, (2) can be reformulated so that it belongs to a wider class of problems called the little Grothendieck problem [2], which again belongs to QO-OC (Quadratic Optimization under Orthogonality Constraints) considered by Nemirovski [3]. QO-OCs have been well studied and include many important problems as special cases, such as Max-Cut [4] and generalized orthogonal Procrustes [5–7]

$$\min_{\boldsymbol{V}_1, \ldots, \boldsymbol{V}_n \in O(D)} \sum_{1 \leq i,j \leq n} \|\boldsymbol{X}_i \boldsymbol{V}_i - \boldsymbol{X}_j \boldsymbol{V}_j\|_F^2,$$

which has applications to areas such as psychometrics, image and shape analysis and biometric identification.

The non-commutative little Grothendieck problem [8] is defined by:

$$\min_{\boldsymbol{V}_1, \ldots, \boldsymbol{V}_n \in O(D)} \sum_{i,j=1}^{n} \operatorname{tr}(\boldsymbol{C}_{ij} \boldsymbol{V}_i \boldsymbol{V}_j^\top). \tag{4}$$

Problem (2) can be considered as a special case of (4) with $n = 3$. The argument is as follows. For convenience, we homogenize (1) by introducing a slack unitary variable $\boldsymbol{V}_3 \in O(D)$ and consider the augmented linear system

$$\boldsymbol{X}_1 \boldsymbol{V}_1 + \boldsymbol{X}_2 \boldsymbol{V}_2 + \boldsymbol{X}_3 \boldsymbol{V}_3 = \boldsymbol{0} \tag{5}$$