Appl. Comput. Harmon. Anal. 40 (2016) 396–416

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Letter to the Editor

A perturbation inequality for concave functions of singular values and its applications in low-rank matrix recovery

Man-Chung Yue [a], Anthony Man-Cho So [b,*]

[a] *Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong*
[b] *Department of Systems Engineering and Engineering Management, and, by courtesy, CUHK-BGI Innovation Institute of Trans-Omics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong*

A R T I C L E   I N F O

Article history:
Received 27 June 2014
Received in revised form 12 March 2015
Accepted 19 June 2015
Available online 26 June 2015
Communicated by Thomas Strohmer

Keywords:
Singular value perturbation inequality
Schatten quasi-norm
Low-rank matrix recovery
Exact and robust recovery

A B S T R A C T

In this paper, we establish the following perturbation result concerning the singular values of a matrix: Let $A, B \in \mathbb{R}^{m\times n}$ be given matrices, and let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be a concave function satisfying $f(0) = 0$. Then, we have

$$\sum_{i=1}^{\min\{m,n\}} \left| f(\sigma_i(A)) - f(\sigma_i(B)) \right| \le \sum_{i=1}^{\min\{m,n\}} f(\sigma_i(A-B)),$$

where $\sigma_i(\cdot)$ denotes the i-th largest singular value of a matrix. This answers an open question that is of interest to both the compressive sensing and linear algebra communities. In particular, by taking $f(\cdot) = (\cdot)^p$ for any $p \in (0, 1]$, we obtain a perturbation inequality for the so-called Schatten p-quasi-norm, which allows us to confirm the validity of a number of previously conjectured conditions for the recovery of low-rank matrices via the popular Schatten p-quasi-norm heuristic. We believe that our result will find further applications, especially in the study of low-rank matrix recovery.

1. Introduction

The problem of low-rank matrix recovery, with its many applications in computer vision [12,20], trace regression [31,23], network localization [19,21], etc., has been attracting intense research interest in recent years. In a basic version of the problem, the goal is to reconstruct a low-rank matrix from a set of possibly noisy linear measurements. To achieve this, one immediate idea is to formulate the recovery problem as a rank minimization problem:

* Corresponding author.
E-mail addresses: mcyue@se.cuhk.edu.hk (M.-C. Yue), manchoso@se.cuhk.edu.hk (A.M.-C. So).

http://dx.doi.org/10.1016/j.acha.2015.06.006

$$\begin{array}{ll} \text{minimize} & \text{rank}(X) \\ \text{subject to} & \|\mathcal{A}(X) - y\|_2 \le \eta, \ \ X \in \mathbb{R}^{m\times n}, \end{array} \tag{1}$$

where the linear measurement map $\mathcal{A} : \mathbb{R}^{m\times n} \to \mathbb{R}^l$, the vector of measurements $y \in \mathbb{R}^l$, and the noise level $\eta \ge 0$ are given. However, Problem (1) is NP-hard in general, as it includes the NP-hard vector cardinality minimization problem [30] as a special case. Moreover, since the rank function is discontinuous, Problem (1) can be challenging from a computational point-of-view. To circumvent this intractability, a popular approach is to replace the objective of (1) with the so-called Schatten (quasi)-norm of X. Specifically, given a matrix $X \in \mathbb{R}^{m\times n}$ and a number $p \in (0, 1]$, let $\sigma_i(X)$ denote the i-th largest singular value of X and define the *Schatten p-quasi-norm* of X by

$$\|X\|_p = \left(\sum_{i=1}^{\min\{m,n\}} \sigma_i^p(X) \right)^{1/p}.$$

One can then consider the following *Schatten p-quasi-norm heuristic* for low-rank matrix recovery:

$$\begin{array}{ll} \text{minimize} & \|X\|_p^p \\ \text{subject to} & \|\mathcal{A}(X) - y\|_2 \le \eta, \quad X \in \mathbb{R}^{m\times n}. \end{array} \tag{2}$$

Note that the function $X \mapsto \|X\|_p^p$ is continuous for each $p \in (0, 1]$. Thus, algorithmic techniques for continuous optimization can be used to tackle Problem (2). The Schatten quasi-norm heuristic is motivated by the observation that $\|X\|_p^p \to \text{rank}(X)$ as $p \searrow 0$. In particular, when $p = 1$, the function $X \mapsto \|X\|_1$ defines a norm—known as the *nuclear norm*—on the set of $m \times n$ matrices, and we obtain the well-known *nuclear norm heuristic* [13]. In this case, Problem (2) is a convex optimization problem that can be solved efficiently by various algorithms; see, e.g., [18] and the references therein. On the other hand, when $p \in (0, 1)$, the function $X \mapsto \|X\|_p$ only defines a quasi-norm. In this case, Problem (2) is a non-convex optimization problem and is NP-hard in general; cf. [16]. Nevertheless, a number of numerical algorithms implementing the Schatten p-quasi-norm heuristic (where $p \in (0, 1)$) have been developed (see, e.g., [28,32,21,26] and the references therein), and they generally have better empirical recovery performance than the (convex) nuclear norm heuristic.

From a theoretical perspective, a natural and fundamental question concerning the aforementioned heuristics is about their recovery properties. Roughly speaking, this entails determining the conditions under which a given heuristic can recover, either exactly or approximately, a solution to Problem (1). A first study in this direction was done by Recht, Fazel, and Parrilo [35], who showed that techniques used to analyze the ℓ_1 heuristic for sparse vector recovery (see [40] for an overview and further pointers to the literature) can be extended to analyze the nuclear norm heuristic. Since then, recovery conditions based on the restricted isometry property (RIP) and various nullspace properties have been established for the nuclear norm heuristic; see, e.g., [33,7,6,22] for some recent results. In fact, many recovery conditions for the nuclear norm heuristic can be derived in a rather simple fashion from their counterparts for the ℓ_1 heuristic by utilizing a perturbation inequality for the nuclear norm [33].

Compared with the nuclear norm heuristic, recovery properties of the Schatten p-quasi-norm heuristic are much less understood, even though the corresponding heuristic for sparse vector recovery, namely the ℓ_p heuristic with $p \in (0, 1)$, has been extensively studied; see, e.g., [42,44,34,43] and the references therein. As first pointed out in [33] and later further elaborated in [25], the difficulty seems to center around the following question, which concerns the validity of certain perturbation inequality for the Schatten p-quasi-norm:

Question (Q). *Given a number $p \in (0, 1)$ and matrices $A, B \in \mathbb{R}^{m\times n}$, does the inequality*