ARTICLE IN PRESS
YACHA:1156

Appl. Comput. Harmon. Anal. • • • (• • • •) • • •–• • •

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Second-order matrix concentration inequalities

Joel A. Tropp

ARTICLE INFO

Article history:
Received 25 April 2015
Received in revised form 11 July 2016
Accepted 24 July 2016
Available online xxxx
Communicated by Charles K. Chui

MSC:
primary: 60B20
secondary: 60F10
60G50
60G42

Keywords:
Concentration inequality
Moment inequality
Random matrix

ABSTRACT

Matrix concentration inequalities give bounds for the spectral-norm deviation of a random matrix from its expected value. These results have a weak dimensional dependence that is sometimes, but not always, necessary. This paper identifies one of the sources of the dimensional term and exploits this insight to develop sharper matrix concentration inequalities. In particular, this analysis delivers two refinements of the matrix Khintchine inequality that use information beyond the matrix variance to improve the dimensional dependence.

1. Motivation

Matrix concentration inequalities provide spectral information about a random matrix that depends smoothly on many independent random variables. In recent years, these results have become a dominant tool in applied random matrix theory. There are several reasons for the success of this approach:

- **Flexibility.** Matrix concentration applies to a wide range of random matrix models. In particular, we can obtain bounds for the spectral norm of a sum of independent random matrices in terms of the properties of the summands.
- **Ease of use.** For many applications, matrix concentration tools require only a small amount of matrix analysis. No expertise in random matrix theory is required to invoke the results.
- **Power.** For a large class of examples, including independent sums, matrix concentration bounds are provably close to optimal.

See the monograph [22] for an overview of this theory and a comprehensive bibliography.

E-mail address: jtropp@cms.caltech.edu.

http://dx.doi.org/10.1016/j.acha.2016.07.005
1063-5203/

Please cite this article in press as: J.A. Tropp, Second-order matrix concentration inequalities, Appl. Comput. Harmon. Anal. (2016), http://dx.doi.org/10.1016/j.acha.2016.07.005

The matrix concentration inequalities in the literature are suboptimal for certain examples because of a weak dependence on the dimension of the random matrix. Removing this dimensional term is difficult because there are many situations where it is necessary. The purpose of this paper is to identify one of the sources of the dimensional factor. Using this insight, we will develop some new matrix concentration inequalities that are qualitatively better than the current generation of results, although they sacrifice some of our desiderata. Ultimately, we hope that this line of research will lead to general tools for applied random matrix theory that are flexible, easy to use, and that give sharp results in most cases.

2. The matrix Khintchine inequality

To set the stage, we present and discuss the primordial matrix concentration result, the *matrix Khintchine inequality*, which describes the behavior of a special random matrix model, called a *matrix Gaussian series*. This result already exhibits the key features of more sophisticated matrix concentration inequalities, and it can be used to derive concentration bounds for more general models. As such, the matrix Khintchine inequality serves as a natural starting point for deeper investigations.

2.1. Matrix Gaussian series

In this work, we focus on an important class of random matrices that has a lot of modeling power but still supports an interesting theory.

Definition 2.1 *(Matrix Gaussian series).* Consider fixed Hermitian matrices $\boldsymbol{H}_1, \dots, \boldsymbol{H}_n$ with common dimension d, and let $\{\gamma_1, \dots, \gamma_n\}$ be an independent family of standard normal random variables. Construct the random matrix

$$\boldsymbol{X} := \sum\nolimits_{i=1}^{n} \gamma_i \boldsymbol{H}_i. \tag{2.1}$$

We refer to a random matrix with this form as a matrix Gaussian series with Hermitian coefficients or, for brevity, an *Hermitian matrix Gaussian series*.

Matrix Gaussian series enjoy a surprising amount of modeling power. It is easy to see that we can express any random Hermitian matrix with jointly Gaussian entries in the form (2.1). More generally, we can use matrix Gaussian series to analyze a sum of independent, zero-mean, random, Hermitian matrices $\boldsymbol{Y}_1, \dots, \boldsymbol{Y}_n$. Indeed, for any norm $|||\cdot|||$ on matrices,

$$\mathbb{E}\left|\left|\left|\sum\nolimits_{i=1}^{n} \boldsymbol{Y}_i\right|\right|\right| \le \sqrt{2\pi} \cdot \mathbb{E}\left[\mathbb{E}\left[\left|\left|\left|\sum\nolimits_{i=1}^{n} \gamma_i \boldsymbol{Y}_i\right|\right|\right| \,\middle|\, \boldsymbol{Y}_1, \dots, \boldsymbol{Y}_n\right]\right]. \tag{2.2}$$

The process of passing from an independent sum to a conditional Gaussian series is called *symmetrization*. See [13, Lem. 6.3 and Eqn. (4.8)] for details about this calculation. Furthermore, some techniques for Gaussian series can be adapted to study independent sums directly without the artifice of symmetrization.

Note that our restriction to Hermitian matrices is not really a limitation. We can also analyze a rectangular matrix $\boldsymbol{Z}$ with jointly Gaussian entries by working with the Hermitian dilation of $\boldsymbol{Z}$, sometimes known as the Jordan–Wielandt matrix. See [22, Sec. 2.1.16] for more information on this approach.

2.2. The matrix variance

Many matrix concentration inequalities are expressed most naturally in terms of a matrix extension of the variance.

Please cite this article in press as: J.A. Tropp, Second-order matrix concentration inequalities, Appl. Comput. Harmon. Anal. (2016), http://dx.doi.org/10.1016/j.acha.2016.07.005