Linear Algebra and its Applications 504 (2016) 108–123

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Controlling the least eigenvalue of a random Gram matrix

Pavel Yaskov [a,b,*,1]

[a] Steklov Mathematical Institute of RAS, Russia
[b] National University of Science and Technology MISIS, Russia

ARTICLE INFO

Article history:
Received 14 August 2015
Accepted 28 March 2016
Available online 12 April 2016
Submitted by H. Rauhut

MSC:
60B20

Keywords:
Covariance matrices
Gram matrices
Random matrices

ABSTRACT

Consider a $n \times p$ random matrix X with i.i.d. rows. We show that the least eigenvalue of $n^{-1}X^\top X$ is bounded away from zero with high probability when $p/n \leqslant y$ for some fixed y in $(0, 1)$ and normalized orthogonal projections of rows are not too close to zero. The principal difference from the previous results is that y can be arbitrarily close to one. Our results cover many cases of interest in high-dimensional statistics and random matrix theory.

1. Introduction

Lower bounds on the least eigenvalue of a random Gram matrix play an important role in the least squares problems in high-dimensional statistics and compressed sensing (e.g., see [5,8], and [11]). In particular, these bounds control the rate of convergence of the least squares estimator in the ordinary linear regression model.

* Correspondence to: Steklov Mathematical Institute of RAS, Russia.
E-mail address: yaskov@mi.ras.ru.
1 The research is supported by the Russian Science Foundation via the grant 14-21-00162.

http://dx.doi.org/10.1016/j.laa.2016.03.048

For precise statements, we need to introduce some notation. For a random vector X_0 in $\mathbb{R}^p$, consider a random $n \times p$ matrix X with i.i.d. rows $\{X_k^\top\}_{k=1}^n$ distributed as $X_0^\top$ and the Gram matrix $X^\top X$ of X's columns. Note also that $n^{-1}X^\top X = n^{-1}\sum_{k=1}^n X_k X_k^\top$ is the sample covariance matrix. We are interested in $\lambda_{\min}(X^\top X)$, the least eigenvalue of $X^\top X$.

Recently, a great deal of attention has been paid to bounds implying that $\lambda_{\min}(n^{-1}X^\top X)$ stays away from zero with large probability for large enough p, n with $p/n \leqslant y$ for some fixed $y \in (0,1)$ (see [4,5,7,8,14,15,17,18]). In [8], Lecué and Mendelson obtained the most general sufficient condition for these bounds, i.e., *the small-ball condition* stating that

$$\inf_{v\in\mathbb{S}^{p-1}} \mathbb{P}(|X_0^\top v| > a) \geqslant b \text{ for some } a > 0 \text{ and } b \in (0,1]. \tag{1}$$

Here $\mathbb{S}^{p-1}$ is the unit sphere of $\mathbb{R}^p$ w.r.t. the Euclidean norm $|v| = (v^\top v)^{1/2}$, $v \in \mathbb{R}^p$. In particular, it was proven in [8] that

> there are absolute constants $c, C > 0$ such that if (1) holds for some $(a,b) \in \mathbb{R}_+ \times (0,1]$ and $p/n \leqslant y \in (0, Cb^2]$, then
>
> $$\mathbb{P}(\lambda_{\min}(n^{-1}X^\top X) > a^2 b/2) \geqslant 1 - \exp\{-cnb^2\}.$$

However, this general result doesn't cover the case where y can be arbitrarily close to one. Allowing any y in $(0,1)$, Tikhomirov [15] derived a similar result under a more restrictive assumption that the entries of $X_0 - \mathbb{E}X_0$ are i.i.d.

In the present paper, we show that $\lambda_{\min}(n^{-1}X^\top X)$ is bounded away from zero with high probability when p, n are large, $p/n \leqslant y$ for fixed $y \in (0,1)$, and normalized orthogonal projections of X_0 are not too close to zero. In the proofs we modify the approach of Srivastava and Vershynin [14] that is a randomization of the Batson–Spielman–Srivastava sparsification method [2]. Another modification of this approach has been recently used by Chafaï and Tikhomirov in [4]. We compare their result with ours in Section 3.

The paper is structured as follows. Section 2 contains the main result. Section 3 deals with applications. The proofs are given in Section 4. Some auxiliary lemmas are proved in Appendices.

2. Main result

Introduce two main quantities

$$c(X_0) = \inf \mathbb{E}|X_0^\top v| \text{ and } L(\delta, \varepsilon) = L(\delta, \varepsilon; X_0) = \sup \mathbb{P}(|\Pi X_0|^2 \leqslant \delta \operatorname{rank}(\Pi)),$$