Geometry of log-concave Ensembles of random matrices and approximate reconstruction

Radosław ADAMCZAK [a,1,4], Rafał LATAŁA [a,1,4], Alexander E. LITVAK [b,4]
Alain PAJOR [c,2,4] Nicole TOMCZAK-JAEGERMANN [b,3,4]

[a] *Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warszawa, Poland*
[b] *Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1*
[c] *Equipe d'Analyse et Mathématiques Appliquées, Université Paris Est, 5 boulevard Descartes, Champs sur Marne, 77454 Marne-la-Vallee, Cedex 2, France*

Abstract

We study the Restricted Isometry Property of a random matrix Γ with independent isotropic log-concave rows. To this end, we introduce a parameter $\Gamma_{k,m}$ that controls uniformly the operator norm of sub-matrices with k rows and m columns. This parameter is estimated by means of new tail estimates of order statistics and deviation inequalities for norms of projections of an isotropic log-concave vector.

On étudie la propriété d'isométrie restreinte d'une matrice aléatoire Γ dont les lignes sont des vecteurs aléatoires indépendants isotropes log-concave. Pour cela on introduit un paramètre $\Gamma_{k,m}$ qui contrôle uniformément les normes d'opérateurs des sous-matrices de k lignes et m colonnes. Ce paramètre est estimé à l'aide de nouvelles inégalités de queue des statistiques d'ordre et d'inégalités de déviation des normes de projections d'un vecteur aléatoire log-concave.

Introduction. Let $T \subset \mathbb{R}^N$ and Γ be an $n \times N$ matrix. Consider the problem of reconstructing any vector $x \in T$ from the data $\Gamma x \in \mathbb{R}^n$, with a fast algorithm. Clearly one needs some a priori hypothesis on the subset T and of course, the matrix Γ should be suitably chosen. The common and useful hypothesis is that T consists of sparse vectors, that is vectors with short support. In that setting, Compressed Sensing provides a way of reconstructing the original signal x from its compression Γx with $n \ll N$ by the so-called ℓ_1-minimization method. The problem of reconstruction can be reformulated after D. Donoho [10]

Email addresses: `R.Adamczak@mimuw.edu.pl` (Radosław ADAMCZAK), `rlatala@mimuw.edu.pl` (Rafał LATAŁA), `alexandr@math.ualberta.ca` (Alexander E. LITVAK), `alain.pajor@univ-mlv.fr` (Alain PAJOR), `nicole@ellpspace.math.ualberta.ca` (Nicole TOMCZAK-JAEGERMANN).

1 . Research partially supported by MNiSW Grant no. N N201 397437 and the Foundation for Polish Science.
2 . Research partially supported by the ANR project ANR-08-BLAN-0311-01.
3 . This author holds the Canada Research Chair in Geometric Analysis.
4 . The research was partially conducted while the authors participated in the Thematic Program on Asymptotic Geometric Analysis at the Fields Institute in Toronto in Fall 2010.

in a language of high dimensional geometry, namely, in terms of neighborliness of polytopes obtained by taking the convex hull of the columns of Γ. In this spirit, the sensing matrix is described by its columns. From another point of view, the matrix Γ may be also determined by measurements, e.g. by its rows.

Let $0 \le m \le N$. Denote by U_m the subset of unit vectors in $\mathbb{R}^N$, which are m-sparse, i.e. have at most m non-zero coordinates. The natural scalar product, the Euclidean norm and the unit sphere are denoted by $\langle \cdot, \cdot \rangle$, $|\cdot|$ and S^{N-1}. We also denote by the same notation $|\cdot|$ the cardinality of a set. For any $x = (x_i) \in \mathbb{R}^n$ we let $\|x\|_\infty = \max_i |x_i|$. By C, C_1, c, etc. we will denote absolute positive constants.

Let $\delta_m = \delta_m(\Gamma) = \sup_{x \in U_m} \left| |\Gamma x|^2 - \mathbb{E}|\Gamma x|^2 \right|$ be the Restricted Isometry Property (RIP) parameter of order m. This concept was introduced by E. Candes and T. Tao in [9] and its important feature is that if δ_{2m} is appropriately small then every m-sparse vector x can be reconstructed from its compression Γx by the ℓ_1-minimization method. The goal now is to check this property for certain models of matrices.

The articles [1,2,3,4,5] considered random matrices with independent *columns*, and investigated high dimensional geometric properties of the convex hull of the columns and the RIP for various models of matrices, including the log-concave Ensemble build with independent isotropic log-concave columns. It was shown that various properties of random vectors can be efficiently studied via operator norms and the parameter $\Gamma_{n,m}$ recalled below. In order to control this parameter an efficient technique of chaining was developed in [3] and [4].

In [14], the authors studied the RIP and more generally the parameter $\delta_T = \sup_{x \in T} \left| |\Gamma x|^2 - \mathbb{E}|\Gamma x|^2 \right|$ for random matrices with independent isotropic subgaussian *rows*. It is natural to ask whether random matrices with independent isotropic log-concave *rows* also have the RIP.

Fix integers $n, N \ge 1$. Let $Y_1, \ldots, Y_n$ be independent random vectors in $\mathbb{R}^N$ and let Γ be the $n \times N$ random matrix with rows Y_i. Let $T \subset S^{N-1}$ and $1 \le k \le n$ and define the parameter $\Gamma_k(T)$ by

$$\Gamma_k(T)^2 = \sup_{y \in T} \sup_{\substack{I \subset \{1,\ldots,n\} \\ |I| = k}} \sum_{i \in I} |\langle Y_i, y \rangle|^2. \tag{1}$$

We also denote $\Gamma_{k,m} = \Gamma_k(U_m)$. The role of this parameter with respect to the RIP is revealed by the following lemma which reduces a concentration inequality to a deviation inequality.

Lemma 1 *Let $Y_1, \ldots, Y_n$ be independent isotropic random vectors in $\mathbb{R}^N$. Let $T \subset S^{N-1}$ be a finite set. Let $0 < \theta < 1$ and $B \ge 1$. Then with probability at least $1 - |T| \exp\left(-3\theta^2 n / 8B^2\right)$ one has*

$$\sup_{y \in T} \left| \frac{1}{n} \sum_{i=1}^n (|\langle Y_i, y \rangle|^2 - \mathbb{E}|\langle Y_i, y \rangle|^2) \right| \le \theta + \frac{1}{n} \left(\Gamma_k(T)^2 + \mathbb{E}\Gamma_k(T)^2 \right),$$

where $k \le n$ is the largest integer satisfying $k \le (\Gamma_k(T)/B)^2$.

In this note we focus on the compressed sensing setting where T is the set of sparse vectors. Lemma 1 shows that after a suitable discretization, checking the RIP reduces to estimating $\Gamma_{k,m}$. The idea of such an approach, when $k = n$, originated from the work of J. Bourgain [8] on the empirical covariance matrix. It was developed in [3] and [5] (with $T = U_m$), where the estimate of $\Gamma_{n,m}$ played a central role for solving the Kannan-Lovász-Simonovits conjecture related to complexity of computing high-dimensional volumes ([11]); and it was studied in [13], where $\Gamma_k(T)$ was estimated by means of Talagrand γ-functionals.

Using Lemma 1 it can be shown (cf., [5] for a similar argument) that if $0 < \theta < 1$, $B \ge 1$, and $m \le N$ satisfies $m \log(CN/m) \le 3\theta^2 n / 16B^2$, then with probability at least $1 - \exp\left(-3\theta^2 n / 16B^2\right)$ one has

$$\delta_m(\Gamma/\sqrt{n}) = \sup_{y \in U_m} \left| \frac{1}{n} \sum_{i=1}^n (|\langle Y_i, y \rangle|^2 - \mathbb{E}|\langle Y_i, y \rangle|^2) \right| \le C\theta + \frac{C}{n} \left(\Gamma_{k,m}^2 + \mathbb{E}\Gamma_{k,m}^2 \right), \tag{2}$$

where $k \le n$ is the largest integer satisfying $k \le (\Gamma_{k,m}/B)^2$ (note that k is a random variable).

We consider the log-concave Ensemble of $n \times N$ matrices with independent isotropic log-concave rows. Recall that a random vector is isotropic log-concave if it is centered, its covariance matrix is the identity and its distribution has a log-concave density. Our goal is to bound $\Gamma_{k,m}$ for this Ensemble. This leads to questions that require a deeper understanding of some geometric parameters of log-concave measures, such as tail estimates for order statistics and deviation inequalities for norms of projections. Proofs and related results will be presented in [6].

Main results. Our main result, Theorem 6, provides upper estimates for $\Gamma_{k,m}$ valid with large probability for matrices from the log-concave Ensemble. To achieve this we need some intermediate steps also of a major importance. For a random vector X and $p>0$, we define the following natural parameter

$$\sigma_X(p) = \sup_{t\in S^{N-1}} (\mathbb{E}|\langle t, X\rangle|^p)^{1/p}.$$

It is known that $\sigma_X(p) \le p$ for isotropic log-concave X and $p \ge 2$. Paouris' Theorem ([15]) states

$$(\mathbb{E}|X|^p)^{1/p} \le C((\mathbb{E}|X|^2)^{1/2} + \sigma_X(p)). \tag{3}$$

It is a consequence of Theorem 8.2 combined with Lemma 3.9 in [15], note that Lemma 3.9 holds not only for convex bodies but for log-concave measures as well.

We extend the Paouris Theorem to the following bound on deviations of norm of projections of an isotropic log-concave vector, uniform over all coordinate projections P_I of a fixed rank.

Theorem 2 *Let $m \le N$ and X be an isotropic log-concave vector in $\mathbb{R}^N$. Then for every $t \ge 1$ one has*

$$\mathbb{P}\left(\sup_{\substack{I\subset\{1,\ldots,N\}\\ |I|=m}} |P_I X| \ge Ct\sqrt{m}\log\left(\frac{eN}{m}\right)\right) \le \exp\left(-t\frac{\sqrt{m}}{\sqrt{\log(em)}}\log\left(\frac{eN}{m}\right)\right).$$

This theorem is sharp up to $\sqrt{\log(em)}$ in the probability estimate as the case of a vector with independent exponential coordinates shows. Actually our further applications require a stronger result in which the bound for probability is improved by involving the parameter σ_X and its inverse σ_X^{-1}, namely

Theorem 3 *Let $m \le N$ and X be an isotropic log-concave vector in $\mathbb{R}^N$. Then for any $t \ge 1$,*

$$\mathbb{P}\left(\sup_{\substack{I\subset\{1,\ldots,N\}\\ |I|=m}} |P_I X| \ge Ct\sqrt{m}\log\left(\frac{eN}{m}\right)\right) \le \exp\left(-\sigma_X^{-1}\left(\frac{t\sqrt{m}\log\left(\frac{eN}{m}\right)}{\sqrt{\log(em/m_0)}}\right)\right),$$

where $m_0 = m_0(X,t) = \sup\left\{k \le m:\ k\log(eN/k) \le \sigma_X^{-1}\left(t\sqrt{m}\log(eN/m)\right)\right\}$.

Theorem 3 is based on tail estimates for order statistics of isotropic log-concave vectors. By $(X^*(i))_i$ we denote the non-increasing rearrangement of $(|X(i)|)_i$. Combining (3) with methods of [12] we obtain

Theorem 4 *Let X be an N-dimensional isotropic log-concave vector. Then for every $t \ge C\log(eN/\ell)$,*

$$\mathbb{P}(X^*(\ell) \ge t) \le \exp(-\sigma_X^{-1}(C^{-1}t\sqrt{\ell})).$$

Introduction of the parameter σ_X enables us to obtain new inequalities for convolutions of log-concave measures. Let $X_1,\ldots,X_n$ be independent isotropic log-concave random vectors in $\mathbb{R}^N$. We will consider weighted sums of the vectors X_i of the form $Y=\sum_{i=1}^n x_i X_i$, where $x=(x_1,\ldots,x_n)\in\mathbb{R}^n$. Bernstein's inequality and ψ_1 estimate for isotropic log-concave random vectors give $\sigma_Y(p) \le C(\sqrt{p}|x| + p\|x\|_\infty)$ for $p \ge 1$. Together with Theorem 3 this yields the following

Corollary 5 *Assume that $|x| \le 1$ and $1 \ge b \ge \max(\|x\|_\infty, 1/\sqrt{m})$. Then for any $t \ge 1$,*

$$\mathbb{P}\left(\sup_{\substack{I\subset\{1,\ldots,N\}\\ |I|=m}} |P_I Y| \ge Ct\sqrt{m}\log\left(\frac{eN}{m}\right)\right) \le \exp\left(-\frac{t\sqrt{m}\log\left(\frac{eN}{m}\right)}{b\sqrt{\log(e^2b^2m)}}\right).$$

We now pass to bounds on deviation of $\Gamma_{k,m}$. To get a slightly simplified formula we assume that $N \ge n$.

Theorem 6 *Let $1 \le n \le N$, and let Γ be an $n \times N$ random matrix with independent isotropic log-concave rows. For any integers $k \le n$, $m \le N$ and any $t \ge 1$, we have*
$$\mathbb{P}(\Gamma_{k,m} \ge Ct\lambda) \le \exp(-t\lambda/\sqrt{\log(3m)}),$$
where $\lambda = \sqrt{\log\log(3m)}\sqrt{m}\log(eN/m) + \sqrt{k}\log(en/k)$.

The threshold value λ in Theorem 6 is optimal, up to the factor of $\sqrt{\log\log(3m)}$. Assuming additionally unconditionality of the distributions of the rows, we can remove this factor and get a sharp estimate ([7]).

The proof of the above theorem is composed of two parts, depending on the relation between k and the quantity $k' = \inf\{\ell \ge 1\colon m\log(eN/m) \le \ell\log(en/\ell)\}$. First, we adjust the chaining argument from [3] to reduce the problem to the case $k \le k'$. This step also involves Theorem 2. Next, we use Corollary 5 combined with another chaining to complete the argument.

Theorem 6 together with (2) allows us to prove the RIP result for matrices Γ with independent isotropic log-concave rows. The result is optimal, up to the factor $\log\log 3m$, as shown in [4]. As for Theorem 6, assuming unconditionality of the distributions of the rows, we can remove this factor ([7]).

Theorem 7 *Let $0 < \theta < 1$, $1 \le n \le N$. Let Γ be an $n \times N$ random matrix with independent isotropic log-concave rows. There exists $c(\theta) > 0$ such that $\delta_m(\Gamma/\sqrt{n}) \le \theta$ with overwhelming probability whenever*
$$m\log^2(2N/m)\log\log 3m \le c(\theta)n.$$

References

[1] R. Adamczak, O. Guédon, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, Condition number of a square matrix with i.i.d. columns drawn from a convex body, *Proc. Amer. Math. Soc.*, to appear.

[2] R. Adamczak, O. Guédon, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, Smallest singular value of random matrices with independent columns, *C. R. Math. Acad. Sci. Paris*, **346** (2008), 853–856.

[3] R. Adamczak, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, Quantitative estimates of the convergence of the empirical covariance matrix in log-concave Ensembles, *Journal of AMS*, **234** (2010), 535–561.

[4] R. Adamczak, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, Restricted isometry property of matrices with independent columns and neighborly polytopes by random sampling, *Constructive Approximation*, **34** (2011), 61–88.

[5] R. Adamczak, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, Sharp bounds on the rate of convergence of empirical covariance matrix, *C.R. Math. Acad. Sci. Paris*, **349** (2011), 195–200.

[6] R. Adamczak, R. Latała, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, Tail estimates for norms of sums of log-concave random vectors, preprint.

[7] R. Adamczak, R. Latała, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, Chevet type inequality and norms of submatrices, preprint.

[8] J. Bourgain, Random points in isotropic convex sets. *In:* "Convex geometric analysis, Berkeley, CA, 1996", Math. Sci. Res. Inst. Publ., Vol. 34, 53–58, Cambridge Univ. Press, Cambridge (1999).

[9] E.J. Candés and T. Tao, Decoding by linear programming, *IEEE Trans. Inform. Theory*, **51** (2005), 4203–4215.

[10] D.L. Donoho, Neighborly Polytopes and Sparse solutions of underdetermined linear equations, Department of Statistics, Stanford University, 2005.

[11] R. Kannan, L. Lovász and M. Simonovits, Random walks and $O^*(n^5)$ volume algorithm for convex bodies, *Random structures and algorithms*, **2** (1997), 1–50.

[12] R. Latała, Order statistics and concentration of l_r norms for log-concave vectors, *J. Funct. Anal.*, **261** (2011), 681–696.

[13] S. Mendelson, Empirical Processes with a bounded ψ_1 diameter, *Geom. Funct. Anal.*, **20** (2010), 988–1027.

[14] S. Mendelson, A. Pajor and N. Tomczak-Jaegermann, Reconstruction and subgaussian operators in asymptotic geometric analysis, *Geom. Funct. Anal.*, **17** (2007), 1248–1282.

[15] G. Paouris, Concentration of mass on convex bodies. *Geom. Funct. Anal.*, **16** (2006), 1021–1049.