Geometry of log-concave Ensembles of random matrices

Nicole Tomczak-Jaegermann

Joint work with
Radosław Adamczak, Rafał Latała, Alexander Litvak, Alain Pajor

Cortona, June 2011

Basic definitions and Notation

Let $X = (X(1), \dots, X(N))$ be a random vector in $\mathbb{R}^N$ with full dimensional support. We say that the distribution of X is

- *log-concave*, if X has density of the form $e^{-h(x)}$ with $h\colon \mathbb{R}^N \to (-\infty, \infty]$ convex;
- *isotropic*, if $\mathbb{E}X_i = 0$ and $\mathbb{E}X_iX_j = \delta_{i,j}$.

For $x, y \in \mathbb{R}^N$ we put

- $|x| = \|x\|_2 = \left(\sum_{i=1}^N x_i^2 \right)^{1/2}$
- $\langle x, y \rangle$ is the inner product
- For $\emptyset \neq I \subset \{1, \dots, N\}$, by P_I we denote the coordinate projection onto $\{y \in \mathbb{R}^N : \ \mathrm{supp}(y) \subset I\}$.

Examples: log-concave random vectors

1. Let $K \subset \mathbb{R}^n$ be a convex body ($=$ compact convex, with non-empty interior) (symmetric means $-K = K$).
X a random vector uniformly distributed in K. Then the corresponding probability measure on $\mathbb{R}^N$

$$\mu_K(A) = \frac{|K \cap A|}{|K|}$$

is log-concave (by **Brunn-Minkowski**).
Moreover, for every convex body K there exists an affine map T such that μ_{TK} is isotropic.

2. The Gaussian vector $G = (g_1, ..., g_n)$, where g_i's have $\mathcal{N}(0, 1)$ distribution, is isotropic and log-concave.

3. Similarly the vector $X = (\xi_1, ..., \xi_n)$, where ξ_i's are independent with the exponential distribution (i.e., with density $f(t) = \frac{1}{\sqrt{2}} \exp(-\sqrt{2}|t|)$, for $t \in \mathbb{R}$) is isotropic and log-concave.

Matrices

Let n, N be integers; $X \in \mathbb{R}^N$ a random vector, $X_1, \dots, X_n$ independent, distributed as X.
A is an $n \times N$ matrix which has X_i's as rows

$$A = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_n \end{pmatrix} = \begin{pmatrix} \dots\dots\dots\dots \\ \dots\dots\dots\dots \\ \dots\dots\dots\dots \\ \dots\dots\dots\dots \end{pmatrix}$$

For $J \subset \{1, \dots, n\}$, $I \subset \{1, \dots, N\}$, let $A(J, I)$ be the sub-matrix with rows from J and columns from I. For $k \leqslant n$, $m \leqslant N$, let

$$A_{k,m} = \sup_{J,I} \|A(J, I)\|,$$

over all $|J| \leqslant k$, $|I| \leqslant m$. This is maximum of norms of sub-matrices of k columns and m rows (the operator norms from ℓ_2^N to ℓ_2^n).
Question: Upper bounds for $A_{k,m}$ with high probability

Example: Norms of sub-matrices, uniform version of P.

Earlier case: $m = N$, i.e., $A_{k,N}$. We select k rows and take *all* columns. Solved by ALPT (JAMS 2010); connected with length of good approximations of the covariance matrices

Case $k = 1$**:** For any J with $|J| = 1$, and I with $|I| \leqslant m$, sub-matrix of one row has a form

$$A(J, I) = \Big(P_I X \Big), \qquad \text{so} \quad A_{1,m} = \sup_{|I| \leqslant m} |P_I X|.$$

A bound for $A_{1,m}$ will imply a uniform bound for $P_I X$, for $t \geqslant 1$.

$$\sup_{|I|=m} |P_I X| \leqslant \text{ bound for } \ A_{1,m}, \qquad (*)$$

with high probability. Compare with Paouris' theorem for a *fixed* I with $|I| \leqslant m$: for $s \geqslant 1$,

$$\mathbb{P}\Big(|P_I X| \leqslant C s \sqrt{N}\Big) \geqslant 1 - \exp\Big(-s\sqrt{N}\Big).$$

However the complexity of the family of subsets is $\binom{N}{m} \gg \exp(c\sqrt{N})$. To prove $(*)$ we need more advanced technique.

Order Statistics

For an N–dimensional random vector X by $X_1^* \geqslant X_2^* \geqslant \ldots \geqslant X_N^*$ we denote the nonincreasing rearrangement of $|X(1)|, \ldots, |X(N)|$ (in particular $X_1^* = \max\{|X(1)|, \ldots, |X(N)|\}$ and $X_N^* = \min\{|X(1)|, \ldots, |X(N)|\}$). Random variables X_k^*, $1 \leqslant k \leqslant N$, are called order statistics of X.

Problem: Find an upper bound for $\mathbb{P}(X_k^* \geqslant t)$.

If coordinates of X_i are independent symmetric exponential r.v. with variance 1 then $\mathrm{Med}(X_k^*) \sim \log(eN/k)$ for $k \leqslant N/2$.

Order Statistics for isotropic log-concave vectors

Theorem

Let X be N-dim. log-concave isotropic vector. Then for all $t \geqslant C \log\left(\frac{eN}{k}\right)$,

$$\mathbb{P}\left(X_k^* \geqslant t\right) \leqslant \exp\left(-\frac{1}{C}\sqrt{k}t\right).$$

Actually we need a stronger theorem that uses the following important "weak parameter"

Weak parameter

For a vector X in $\mathbb{R}^N$ we define

$$\sigma_X(p) := \sup_{t \in S^{N-1}} (\mathbb{E}|\langle t, X\rangle|^p)^{1/p} \quad p \geqslant 2.$$

We also let σ_X^{-1} to be the left-inverse function.

$$\sigma_X^{-1}(s) := \sup \{t \colon \sigma_X(t) \leqslant s\}$$

Examples

- For isotropic log-concave vectors X, $\sigma_X(p) \leqslant p/\sqrt{2}$.
- For subgaussian vectors X, $\sigma_X(p) \leqslant C\sqrt{p}$.
- We say that an isotropic vector X is ψ_α if $\sigma_X(p) \leqslant Cp^{1/\alpha}$ (uniform distributions on suitable normalized B_p^N balls are ψ_α with $\alpha = \min(p, 2)$)

Order Statistics with weak parameter

Theorem

For any N*-dim. log-concave isotropic vector* X, $\ell \geqslant 1$ *and* $t \geqslant C \log\left(\frac{eN}{\ell}\right)$,

$$\mathbb{P}\Big(X_\ell^* \geqslant t\Big) \leqslant \exp\Big(-\sigma_X^{-1}\Big(\frac{1}{C} t\sqrt{\ell}\Big)\Big)$$

This theorem implies uniform estimates for norms of projections.

Uniform bound for norms of projections

Theorem

Let X be an isotropic log-concave vector in $\mathbb{R}^N$, and $m \leqslant N$. Then for any $t \geqslant 1$,

$$\mathbb{P}\left(\sup_{|I|=m} |P_I X| \geqslant Ct\sqrt{m}\log\left(\frac{eN}{m}\right)\right)$$

is less than or equal to

$$\exp\left(-\sigma_X^{-1}\left(\frac{t\sqrt{m}\log\left(\frac{eN}{m}\right)}{\sqrt{\log(em)}}\right)\right) \leqslant \exp\left(-t\frac{\sqrt{m}}{\sqrt{\log(em)}}\log\left(\frac{eN}{m}\right)\right).$$

The bound is sharp, except for $\sqrt{\log}$ in the probability estimate.
We conjecture this factor is not needed.

Uniform bound for norms of projections, idea

Idea of the proof. We want to estimate

$$\mathbb{P}\left(\sup_{|I|=m} |P_I X| \geqslant Ct\sqrt{m} \log\left(\frac{eN}{m}\right)\right).$$

We have

$$\sup_{|I|=m} |P_I X| = \Big(\sum_{k=1}^{m} |X_k^*|^2 \Big)^{1/2} \leqslant 2 \Big(\sum_{i=0}^{s-1} 2^i |X_{2^i}^*|^2 \Big)^{1/2},$$

where $s = \lceil \log_2 m \rceil$.
We then use the estimates for order statistics.

Estimates for order statistics

Our approach to estimates for order statistics is based on the suitable estimate of moments of the process $N_X(t)$, where for $t \geqslant 0$,

$$N_X(t) := \sum_{i=1}^{N} \mathbb{1}_{\{X_i \geqslant t\}}$$

That is, $N_X(t)$ is equal to the number of coordinates of X larger than or equal to t.

Estimate for N_X

Theorem

For any isotropic log-concave vector X *and* $p \geqslant 2$ *we have*

$$\mathbb{E}(t^2 N_X(t))^p \leqslant (Cp)^{2p} \quad \textit{for } t \geqslant C \log\left(\frac{Nt^2}{p^2}\right).$$

$$\mathbb{E}(t^2 N_X(t))^p \leqslant (C\sigma_X(p))^{2p} \quad \textit{for } t \geqslant C \log\left(\frac{Nt^2}{\sigma_X^2(p)}\right).$$

Estimate for N_X

Theorem

For any isotropic log-concave vector X *and* $p \geqslant 2$ *we have*

$$\mathbb{E}(t^2 N_X(t))^p \leqslant (Cp)^{2p} \quad \textit{for } t \geqslant C \log\left(\frac{Nt^2}{p^2}\right).$$

$$\mathbb{E}(t^2 N_X(t))^p \leqslant (C\sigma_X(p))^{2p} \quad \textit{for } t \geqslant C \log\left(\frac{Nt^2}{\sigma_X^2(p)}\right).$$

To get estimate for order statistics we observe that $X_k^* \geqslant t$ implies that $N_X(t) \geqslant k/2$ or $N_{-X}(t) \geqslant k/2$ and vector $-X$ is also isotropic and log-concave, and $\sigma_X = \sigma_{-X}$. Estimates for N_X and Chebyshev's inequality give

$$\mathbb{P}(X_k^* \geqslant t) \leqslant \left(\frac{2}{k}\right)^p \left(\mathbb{E}N_X(t)^p + \mathbb{E}N_{-X}(t)^p\right) \leqslant \left(\frac{C\sigma_X(p)}{t\sqrt{k}}\right)^{2p}$$

provided that $t \geqslant C\log(Nt^2/\sigma_X^2(p))$. Set $p = \sigma_X^{-1}\left(\frac{1}{eC} t\sqrt{k}\right)$ and notice that the restriction on t follows by the assumption that $t \geqslant C\log(eN/k)$.

Convolutions of measures

Let $X_1, \dots, X_n$ be independent isotropic log-concave vectors in $\mathbb{R}^N$, and let $x = (x_i) \in \mathbb{R}^n$. Consider the vector

$$Y = \sum_{i=1}^{n} x_i X_i \in \mathbb{R}^N.$$

Probability for bounds of the process

$$\sup_{|I|=m} |P_I Y|$$

depends on the vector x. Specifically, it is convenient to assume the normalization $|x| \leqslant 1$ and $\|x\|_\infty \leqslant b \leqslant 1$.

Uniform bound for projections of convolutions

Theorem

Let $Y = \sum_{i=1}^{n} x_i X_i$, *where* $X_1, \ldots, X_n$ *are independent isotropic* N*-dimensional log-concave vectors. Assume that* $|x| \leqslant 1$ *and* $\|x\|_\infty \leqslant b \leqslant 1$*.*
i) If $b \geqslant \frac{1}{\sqrt{m}}$, *then for any* $t \geqslant 1$,

$$\mathbb{P}\left(\sup_{|I|=m} |P_I Y| \geqslant Ct\sqrt{m}\log\left(\frac{eN}{m}\right)\right) \leqslant \exp\left(- \frac{t\sqrt{m}\log\left(\frac{eN}{m}\right)}{b\sqrt{\log(e^2 b^2 m)}}\right).$$

ii) If $b \leqslant \frac{1}{\sqrt{m}}$ *then for any* $t \geqslant 1$,

$$\mathbb{P}\left(\sup_{|I|=m} |P_I Y| \geqslant Ct\sqrt{m}\log\left(\frac{eN}{m}\right)\right)$$
$$\leqslant \exp\left(- \min\left\{t^2 m \log^2\left(\frac{eN}{m}\right), \frac{t}{b}\sqrt{m}\log\left(\frac{eN}{m}\right)\right\}\right).$$

Uniform bound for norms of submatrices

Let A be an $n \times N$ random matrix with independent log-concave isotropic rows $X_1, \ldots, X_n \in \mathbb{R}^N$. For $k \leqslant n, m \leqslant N$ we let

$$A_{k,m} = \text{the maximal operator norm of a } k \times m \text{ submatrix of } A.$$

For simplicity assume $n \leqslant N$.

Theorem

For any integers $n \leqslant N$, $k \leqslant n$, $m \leqslant N$ and any $t \geqslant 1$, we have

$$\mathbb{P}\Big(A_{k,m} \geqslant Ct\lambda_{mk}\Big) \leqslant \exp\left(-\frac{t\lambda_{mk}}{\sqrt{\log(3m)}}\right),$$

where

$$\lambda_{mk} = \sqrt{\log\log(3m)}\sqrt{m}\log\left(\frac{eN}{m}\right) + \sqrt{k}\log\left(\frac{en}{k}\right).$$

Uniform bound for norms of submatrices, cont.

The threshold value of λ_{km} is optimal, up to the factor $\sqrt{\log\log(3m)}$. Under the additional assumption of unconditionality of the rows we can remove this factor and get the sharp estimate. For example, for expectations:

Theorem

Let A be an $n \times N$ matrix whose rows are independent isotropic log-concave unconditional random vectors. Then

$$\mathbb{E}A_{km} \leqslant C\left(\sqrt{m}\,\log\frac{3N}{m} + \sqrt{k}\,\log\frac{3n}{k}\right).$$

Reconstruction

Let n, N be integers; $X \in \mathbb{R}^N$ a random vector, $X_1, \dots, X_n$ independent, distributed as X. A is an $n \times N$ matrix which has X_i's as rows

$$A = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_n \end{pmatrix} = \begin{pmatrix} \dots\dots\dots\dots \\ \dots\dots\dots\dots \\ \dots\dots\dots\dots \\ \dots\dots\dots\dots \end{pmatrix}$$

We can treat $A : \mathbb{R}^N \to \mathbb{R}^n$ given by $x \leftarrow \big(\langle X_i, x\rangle\big) \in \mathbb{R}^n$.

Let $m \leqslant N$. A vector $x \in \mathbb{R}^N$ is m-sparse if $|\operatorname{supp} x| \leqslant m$.
Problem from Compressed Sensing theory: Reconstruct any m-sparse vector $x \in \mathbb{R}^N$ from the data $Ax \in \mathbb{R}^n$, with a fast algorithm.

Given Ax, find x, knowing that it is sparse. Note that of course A is not-invertible.

RIP and Geometry pf Polytopes

Define $\delta_m = \delta_m(A)$ as the infimum of $\delta > 0$ such that

$$\left| |Ax|^2 - \mathbb{E}|Ax|^2 \right| \leqslant \delta \, n \, |x|^2$$

holds for all m-sparse vectors $x \in R^N$.
δ_m is the Restricted Isometry Property (RIP) parameter of order m.

Candes and Tao (2006): if δ_{2m} is sufficiently small then
$(*)$ whenever $y = Ax$ has a m-sparse solution x, then x is the unique solution of the ℓ_1-minimization program: min $\|t\|_{\ell_1}$ with the min over all t such that $At = y$.

Geometry of Polytopes: By Donoho (2005), $(*)$ is equivalent to the condition that the centrally symmetric polytope $A(B_1^N)$ is m-centrally neighborly (i.e., any set of less than m vertices containing no opposite pairs, is a vertex set of a face).

Lemma

Let $X_1, \dots, X_n$ *be independent isotropic random vectors in* $\mathbb{R}^N$*. Let* $0 < \theta < 1$ *and* $B \geqslant 1$*. Then with probability at least*

$$1 - \binom{N}{m} \exp\left(-3\theta^2 n / 8B^2\right)$$

one has

$$\delta_m \leqslant \theta + \frac{1}{n}\left(A_{k,m}^2 + \mathbb{E}A_{k,m}^2\right),$$

where $k \leqslant n$ *is the largest integer satisfying* $k \leqslant (A_{k,m}/B)^2$*;*

RIP theorem for matrices with independent rows:

Theorem

Let $n \leqslant N$ *and* $0 < \theta < 1$. *Let* A *be an* $n \times N$ *matrix, whose rows are independent isotropic log-concave random vectors* X_i, $i \leqslant n$.
There exists an absolute constant $c > 0$, *such that if* $m \leqslant N$ *satisfies*

$$m \log\log(3m) \left(\log \frac{3N}{m}\right)^2 \leqslant c \left(\frac{\theta}{\log(3/\theta)}\right)^2 n$$

then

$$\delta_m \leqslant \theta$$

with overwhelming probability.

Optimal up to a log log factor.
For unconditional distributions can be removed