ESTIMATES FOR MOMENTS OF RANDOM VECTORS

AUTHOR’S ABSTRACT OF THE PHD DISSERTATION

MARTA STRZELECKA

My dissertation is devoted to estimates of norms of some natural classes of random vectors in $\mathbb{R}^n$. Dimension-free bounds are of most interest, since they may be generalised to infinitely-dimensional spaces. However, if the dependence on the dimension is nontrivial (especially if an estimate depends only on the logarithm of the dimension), a bound is useful too and gives us a better understanding of the behaviour of the class of random vectors we investigate. Let us describe three types of estimates we are dealing with in this thesis.

1. Comparison of weak and strong moments

In convex geometry the class of log-concave vectors is often investigated. One of the fundamental property of this class is the Paouris inequality from [15], which in a version from [1] states that for a log-concave vector $X$ in $\mathbb{R}^n$,

\begin{equation}
(\mathbb{E}\|X\|^p)^{1/p} \leq C_1 \left( (\mathbb{E}\|X\|^2)^{1/2} + \sigma_X(p) \right) \quad \text{for } p \geq 1,
\end{equation}

where

\[ \sigma_X(p) := \sup_{\|t\|_2 \leq 1} \left( \mathbb{E} \left( \sum_{i=1}^n t_i X_i \right)^p \right)^{1/p} \]

is the Euclidean weak $p$-th moment of $X$. We call the quantity $(\mathbb{E}\|X\|^p)^{1/p}$ the $p$-th strong moment of $X$ (with respect to the Euclidean norm). Since a bound reverse to (1.1) holds trivially, the Paouris inequality states in fact, that weak and strong moments of the Euclidean norm of a log-concave vector are comparable.

It is natural to ask whether inequality (1.1) may be generalized to non-Euclidean norms. In [8] Latała formulated and discussed the following conjecture.

Conjecture 1.1. There exists a universal constant $C$ such that for any log-concave vector $X$ with values in a finite dimensional normed space $(F, \|\|)$,

\begin{equation}
(\mathbb{E}\|X\|^p)^{1/p} \leq C \left( \mathbb{E}\|X\| + \sup_{\varphi \in F^*, \|\varphi\| \leq 1} (\mathbb{E}|\varphi(X)|^p)^{1/p} \right) \quad \text{for } p \geq 1.
\end{equation}

Today we only know that Conjecture 1.1 holds in some special cases, and we do not know any possible counterexample. Therefore even some partial results in this matter deepen our understanding of log-concave vectors.

It is also interesting to find more general assumptions than log-concavity under which (1.2) holds in some special cases. Latała and Tkocz proved in [12, Theorem 2.3] that for vectors with independent coordinates we may indeed assume less then the log-concavity
for (1.2) to hold. This weaker assumption is the $\alpha$-regularity of growth of moments of coordinates of $X$ (then the constant $C$ depends on $\alpha$). However, in the case of dependent coordinates the $\alpha$-regularity of growth of moments of $\langle t, X \rangle$ (for all $t \in \mathbb{R}^n$) does not imply (1.2) even for the Euclidean norm.

1.1. Comparison of moments for $\ell_r$-norms. The first main result of the dissertation states that an analogue of the Paouris inequality holds with the $\ell_r$-norm of any log-concave vector, with a constant depending linearly on $r$. It comes from the joint work with Rafał Latała [10]. This result may be easily generalised to the analogue estimate for spaces that may be isometrically embedded in $\ell_r$ for some $r \geq 1$.

**Theorem 1.2.** Let $X$ be a log-concave vector with values in a normed space $(F, \|\|)$ which may be isometrically embedded in $\ell_r$ for some $r \in [1, \infty)$. Then for $p \geq 1$,

$$(\mathbb{E}\|X\|^p)^{1/p} \leq Cr \left( \mathbb{E}\|X\| + \sup_{\varphi \in F^*, \|\varphi\|_\ast \leq 1} \mathbb{E}|\varphi(X)|^p \right)^{1/p}. $$

This theorem implies the following deviation inequality for $\|X\|$.

**Corollary 1.3.** Let $X$ and $F$ be as above. Then

$$\mathbb{P}(\|X\| \geq 2eCr t \mathbb{E}\|X\|) \leq \exp \left( -\sigma^{-1}_{\|\|, X}(t \mathbb{E}\|X\|) \right) \quad \text{for } t \geq 1. $$

We may take $C$ as in Theorem 1.2.

To show the above theorem we follow the approach from [9] and establish the following cut version of the above inequality.

**Theorem 1.4.** Suppose that $r \in [1, \infty)$ and $X$ is a log-concave $n$-dimensional random vector. Let

$$d_i := (\mathbb{E}X_i^2)^{1/2}, \quad d := \left( \sum_{i=1}^n d_i^r \right)^{1/r}. $$

Then for $p \geq r$,

$$\mathbb{E}\left( \sum_{i=1}^n |X_i|^r 1_{\{|X_i| \geq td_i^r\}} \right)^{p/r} \leq (C_2r \sigma_{r, X}(p))^p \quad \text{for } t \geq C_3r \log \left( \frac{d}{\sigma_{r, X}(p)} \right), $$

where

$$\sigma_{r, X}(p) := \sup_{\|t\|_r \leq 1} \left( \mathbb{E}\left( \sum_{i=1}^n t_i X_i^p \right)^{1/p} \right). $$

**Remark 1.5.** Any finite dimensional space embeds isometrically in $\ell_\infty$, so to show Conjecture 1.1 it is enough to establish Theorem 1.2 (with a universal constant in place of $Cr$) for $r = \infty$. It is known that such an estimate holds for isotropic log-concave vectors. However a linear image of an isotropic vector does not have to be isotropic, so to establish the conjecture we need to consider either isotropic vectors and an arbitrary norm or vectors with a general covariance structure and the standard $\ell_\infty$-norm.
Remark 1.6. An \( n \)-dimensional space embeds isometrically in \( \ell_\infty^N \), where \( N \sim e^n \). Moreover, in \( \mathbb{R}^N \) we have \( e^{-1} \| \cdot \|_{\log N} \leq \| \cdot \|_{\infty} \leq \| \cdot \|_{\log N} \). Therefore Theorem 1.2 implies (1.2) with \( C \sim \log N \sim n \). If Theorem 1.2 held with \( Cr^\gamma \) instead of \( Cr \), then (1.2) would hold with \( C \sim n^\gamma \), which is unknown for any \( \gamma \leq \frac{1}{2} \).

1.2. Comparison of moments in the independent case. Let us now present results obtained in another joint work with Rafał Latała [11]. We may look at the comparison of moments in a slightly different way than the one presented before. For an \( n \)-dimensional random vector \( X \) instead of taking the moments of norms of \( X \) we may considering the moments of \( \sup_{t \in T} | \sum_{i=1}^n t_i X_i | \) – if \( T \) is a unit ball of the dual norm of \( \| \cdot \| \), then this quantity coincides with \( \| X \| \). This approach is useful in the proof of our second main result concerning the comparison of weak and strong moments, which generalise the aforementioned result of [12, Theorem 2.3] for vectors with independent regular coordinates.

**Theorem 1.7.** Let \( X_1, \ldots, X_n \) be independent mean zero random variables with finite moments such that

\[
\| X_i \|_p \leq \alpha \| X_i \|_p \quad \text{for every } p \geq 2 \text{ and } i = 1, \ldots, n,
\]

where \( \alpha \) is a finite positive constant. Then for every \( p \geq 1 \) and every nonempty set \( T \subset \mathbb{R}^n \) we have

\[
\left( \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^n t_i X_i \right|^p \right)^{1/p} \leq C(\alpha) \left[ \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^n t_i X_i \right| + \sup_{t \in T} \left( \mathbb{E} \left| \sum_{i=1}^n t_i X_i \right|^p \right)^{1/p} \right],
\]

where \( C(\alpha) \) is a constant which depends only on \( \alpha \).

It turns out that Theorem 1.7 may be reversed in the i.i.d. case (see the theorem below). Therefore one cannot weaken assumption 1.5 in Theorem 1.7.

**Theorem 1.8.** Let \( X_1, X_2, \ldots \) be i.i.d. random variables. Assume that there exists a constant \( L \) such that for every \( p \geq 1 \), every \( n \) and every nonempty set \( T \subset \mathbb{R}^n \) we have

\[
\left( \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^n t_i X_i \right|^p \right)^{1/p} \leq L \left[ \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^n t_i X_i \right| + \sup_{t \in T} \left( \mathbb{E} \left| \sum_{i=1}^n t_i X_i \right|^p \right)^{1/p} \right].
\]

Then

\[
\| X_1 \|_p \leq \alpha(L) \| X_1 \|_p \quad \text{for } p \geq 2,
\]

where \( \alpha(L) \) is a constant which depends only on \( L \geq 1 \).

It is clear from the proof of Theorem 1.8 that it suffices to assume (1.7) for \( T = \{ \pm e_j : j \in \{1, \ldots, n\} \} \) only, where \( \{e_1, \ldots, e_n\} \) is the canonical basis of \( \mathbb{R}^n \).

The comparison of weak and strong moments (1.6) yields also a deviation inequality for \( \sup_{t \in T} | \sum_{i=1}^n t_i X_i | \).

**Corollary 1.9.** Assume \( X_1, X_2, \ldots \) satisfy the assumptions of Theorem 1.7. Then for any \( u \geq 0 \) and any nonempty set \( T \) in \( \mathbb{R}^n \),

\[
P \left( \sup_{t \in T} \left| \sum_{i=1}^n t_i X_i \right| \geq C_1(\alpha) \left[ u + \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^n t_i X_i \right| \right] \right) \leq C_2(\alpha) \sup_{t \in T} P \left( \left| \sum_{i=1}^n t_i X_i \right| \geq u \right),
\]
where constants $C_1(\alpha)$ and $C_2(\alpha)$ depend only on the constant $\alpha$ in (1.5).

Another consequence of Theorem 1.8 is the following Khintchine-Kahane type inequality.

**Corollary 1.10.** Assume $X_i, 1 \leq i \leq n$ satisfy the assumptions of Theorem 1.7. Then for any $p \geq q \geq 2$ and any nonempty set $T$ in $\mathbb{R}^n$ we have,

$$\left( \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^{n} t_i X_i \right|^p \right)^{1/p} \leq C(\alpha) \left( \frac{p}{q} \right)^{\max\{1/2, \log_2 \alpha\}} \left( \mathbb{E} \sup_{t \in T} \left| \sum_{i=1}^{n} t_i X_i \right|^q \right)^{1/q}$$

where a constant $C(\alpha)$ depends only on the constant $\alpha$ in (1.5). Moreover, the exponent $\max\{1/2, \log_2 \alpha\}$ is optimal.

2. **Convex infimum convolution inequality**

The results presented in this section come from the joint work with Michał Strzelecki and Tomasz Tkocz [19].

Let $X$ be a random vector with values in $\mathbb{R}^n$ and let $\varphi : \mathbb{R}^n \to [0, \infty]$ be a measurable function. We say that the pair $(X, \varphi)$ satisfies the infimum convolution inequality (ICI for short) if for every bounded measurable function $f : \mathbb{R}^n \to \mathbb{R}$,

$$E e^{f \square \varphi(X)} E e^{-f(X)} \leq 1,$$

(2.1)

where $f \square \varphi$ denotes the infimum convolution of $f$ and $\varphi$, defined as $f \square \varphi(x) = \inf\{f(y) + \varphi(x-y) : y \in \mathbb{R}^n\}$ for $x \in \mathbb{R}^n$. The function $\varphi$ is called a cost function and $f$ is called a test function. We also say that the pair $(X, \varphi)$ satisfies the convex infimum convolution inequality if (2.1) holds for every convex function $f : \mathbb{R}^n \to \mathbb{R}$ bounded from below.

The recent works [5] and [4] enable to view the ICI from a different perspective. In [5] Gozlan, Roberto, Samson, and Tetali introduce weak transport-entropy inequalities and establish their dual formulations. The dual formulations are exactly the convex ICIs. In [4] Gozlan, Roberto, Samson, Shu and Tetali investigate extensively the weak transport cost inequalities on the real line, obtaining a characterisation for arbitrary cost functions which are convex and quadratic near zero, thus providing a tool for studying the convex ICI. Around the same time, the convex ICI for the quadratic-linear cost function was fully understood by Feldheim, Marsiglietti, and Nayar in [3].

Using the aforementioned novel tools from [4], we show that product measures with symmetric marginals having log-concave tails satisfy the optimal convex ICI, which complements Latała and Wojtaszczyk’s result about log-concave product measures. This has applications to concentration and moment comparison. We also offer an example showing that the assumption of log-concave tails cannot be weakened substantially.

Let us explain what the optimal convex ICI is. For a random vector $X$ in $\mathbb{R}^n$ we define

$$\Lambda_X^*(x) := \mathcal{L} \Lambda_X(x) := \sup_{y \in \mathbb{R}^n} \{ \langle x, y \rangle - \ln E e^{\langle y, X \rangle} \},$$

which is the Legendre transform of the cumulant-generating function

$$\Lambda_X(x) := \ln E e^{\langle x, X \rangle}, \quad x \in \mathbb{R}^n.$$

If $X$ is symmetric and the pair $(X, \varphi)$ satisfies the convex ICI, then $\varphi(x) \leq \Lambda_X^*(x)$ for every $x \in \mathbb{R}^n$. In other words, $\Lambda_X^*$ is the optimal cost function $\varphi$ for which the convex
ICI can hold. We say that $X$ satisfies (convex) IC($\beta$) if the pair $(X, \Lambda_X(\cdot/\beta))$ satisfies the (convex) ICI.

We are ready to present our first main result.

**Theorem 2.1.** There exists a universal constant $\beta \leq 1680e$ such that every symmetric random variable with log-concave tails satisfies convex IC($\beta$).

The convex ICI tensorises, thus we have the following corollary.

**Corollary 2.2.** Let $X$ be a symmetric random vector with values in $\mathbb{R}^n$ and independent coordinates with log-concave tails. Then $X$ satisfies convex IC($\beta$) with a universal constant $\beta \leq 1680e$.

Note that the class of distributions from Theorem 2.1 is wider than the class of symmetric log-concave product distributions considered by Latała and Wojtaszczyk in [14]. Among others, it contains measures which do not have a connected support, e.g. a symmetric Bernoulli random variable.

Recall that variables with log-concave tails are 1-regular. However, the assumption of log-concave tails in Theorem 2.1 cannot be replaced by a weaker one of $\alpha$-regularity of moments (an example is given both in the dissertation and in [19]). Thus it seems that the assumptions of Theorem 2.1 are not far from necessary conditions for the convex ICI to hold with an optimal cost function (random variables with moments growing regularly are akin to random variables with log-concave tails as the former can essentially be sandwiched between the latter, see (4.6) in [12]).

Another corollary to Theorem 2.1 is the comparison of weak and strong moments.

**Corollary 2.3.** Let $X$ be a symmetric random vector with values in $\mathbb{R}^n$ and with independent coordinates which have log-concave tails. Then for every norm $\| \cdot \|$ on $\mathbb{R}^n$ and every $p \geq 2$ we have

$$\left( \mathbb{E} \| X \|^p \right)^{1/p} \leq \mathbb{E} \| X \| + D\sigma_{\| \cdot \|,X}(p),$$

where $D$ is a universal constant (one can take $D = 6720\sqrt{2}e^2 < 70223$).

Note that the constant standing at $\mathbb{E} \| X \|$ is equal to 1. If we only assume that the coordinates of $X$ are independent and their moments grow $\alpha$-regularly, then (2.2) does not always hold, so also in the corollary the assumption about log-concave tails is not far from the optimal one.

3. Estimates of norms of log-concave matrices

A special type of norms are operator norms of matrices (an $(mn)$-dimensional vector may be treated as an $m \times n$ matrix). We are interested in estimating the expected value of the operator norm from $\ell_p^n$ to $\ell_q^m$ of random matrices. Most results concerning this quantity deal with the spectral norm only (i.e. the operator norm from $\ell_2^n$ to $\ell_2^m$). Moreover, in all known results one has to assume the independence of entries of the matrix. The part of the thesis devoted to estimates of norms of random matrices comes from a work in progress [18] by the author.

A classical result regarding spectra of random matrices is Wigner’s Semicircle Law, which describes the limit of empirical spectral measures of a random matrix with independent centred entries with equal variance. Theorems of this type say nothing about
the largest eigenvalue (i.e. the operator norm from $\ell_2$ to $\ell_2$). However, Seginer proved in [17] that for a random matrix $X$ with i.i.d. symmetric entries $\mathbb{E}\|X\|_{2,2}^1$ is of the same order as the expectation of the maximum Euclidean norm of rows and columns of $X$. The same holds true for the structured Gaussian matrices (i.e. when $X_{ij} = a_{ij}g_{ij}$ and $g_{ij}$ are i.i.d. standard Gaussian variables), as was recently shown in [13], and up to a logarithmic factor for any $X$ with independent centred entries, see [16]. The advance of the two latest results is that they do not require that the entries of $X$ are equally distributed.

Another upper bound for $\mathbb{E}\|X\|_{2,2}$ also does not require equal distributions but only the independence of entries: by [7] we know that

$$\mathbb{E}\|X\|_{2,2} \lesssim \max_i \sqrt{\sum_j \mathbb{E}X_{ij}^2} + \max_j \sqrt{\sum_i \mathbb{E}X_{ij}^2} + \sqrt{\sum_{i,j} \mathbb{E}X_{ij}^4}.$$ 

This bound is dimension free, but in some cases is worse than the one from [16].

Upper bounds for the expectation of other operator norms were investigated in [2] in the case of independent centred entries bounded by 1. For $q \geq 2$ and $m \times n$ matrices the authors proved that $\mathbb{E}\|X\|_{p,q} \lesssim \max\{m^{1/q}\sqrt{n}\}$. In [6] Guédon, Hinrichs, Litvak, and Prochno proved that for a structured Gaussian matrix $X = (a_{ij}X_{ij})_{i \leq m, j \leq n}$ and $p, q \geq 2$,

$$\mathbb{E}\|X\|_{p',q} \leq C(p, q) \left[ (\log m)^{1/q} \max_{1 \leq i \leq m} \left( \sum_{j=1}^n |a_{ij}|^p \right)^{1/p} + \max_{1 \leq j \leq n} \left( \sum_{i=1}^m |a_{ij}|^q \right)^{1/q} \right] + (\log m)^{1/q} \mathbb{E} \max_{1 \leq i \leq m} |X_{ij}|.$$ 

It is easy to see that this estimate is optimal up to logarithmic terms. Note that in the case $(p, q) \neq (2, 2)$ moment methods fails in estimating $\mathbb{E}\|X\|_{p',q}$ (as they give information only on the spectrum of $X$).

All the mentioned results require the independence of entries of $X$. We generalise the main result of [6] to a wide class of random matrices with independent log-concave rows, following the scheme of proof of the original theorem from [6]. Our estimate is optimal (for fixed $p, q \geq 2$) up to a factor depending logarithmically on the dimension. Let us stress that we do not require the rows of $X$ to have independent, but only uncorrelated coordinates (and to be log-concave). In the proof we use results from other parts of the dissertation.

To make the notation more clear, if $A = (A_{ij})_{i \leq m, j \leq n}$ is an $m \times n$ matrix, we denote by $A_i \in \mathbb{R}^n$ its $i$-th row and by $A^{(j)} \in \mathbb{R}^m$ we denote its $j$-th column.

**Theorem 3.1.** Let $m \geq 2$, let $Y_1, \ldots, Y_m$ be i.i.d. isotropic log-concave vectors in $\mathbb{R}^n$, and let $A = (A_{ij})$ be an $m \times n$ (deterministic) matrix. Consider a random matrix $X$ with entries $X_{ij} = A_{ij}Y_{ij}$ for $i \leq m, j \leq n$, where $Y_{ij}$ is the $j$'th coordinate of $Y_i$. Then for

\[1\| \cdot \|_{p,q} \text{ stands for the operator norm from } \ell_p \text{ to } \ell_q.\]
every $p, q \geq 2$ we have

\begin{align}
E \|X\|_{p', q} & \leq C(p, q) \left[ (\log m)^{1/q} \max_{1 \leq i \leq m} \|A_i\|_p + \max_{1 \leq j \leq n} \|A^{(j)}\|_q + (\log m)^{1/q + 1} E \max_{1 \leq i \leq m} |X_{ij}| \right],
\end{align}

where $C(p, q)$ depends only on $p$ and $q$.

The next corollary is a version of Theorem 3.1 in the spirit of the aforementioned results from [17, 13, 16].

**Corollary 3.2.** Under the assumptions of Theorem 3.1 we have

\begin{align}
E \|X\|_{p', q} & \leq C(p, q)(\log m)^{1+1/q} \left( E \max_{1 \leq i \leq m} \left( \sum_{j=1}^{n} |X_{ij}|^p \right)^{1/p} + E \max_{1 \leq j \leq n} \left( \sum_{i=1}^{m} |X_{ij}|^q \right)^{1/q} \right)
\end{align}

We may also use the main theorem to get an analogue bound for random matrices, which rows are Gaussian mixtures.

**Corollary 3.3.** Let $m, n \geq 2$, and let $G = (G_{ij})_{i \leq m, j \leq n}$ be a matrix which entries are i.i.d. standard Gaussian variables. Let $X_{ij} = R_{ij} B_{ij} G_{ij}$, where $R$ is a log-concave and isotropic random matrix \(^2\) independent of $G$. Then for every $p, q \geq 2$ we have

\begin{align}
E \|X\|_{p', q} & \leq C(p, q) \left( (\log m)^{1/q + 1} \left[ \max_{1 \leq i \leq m} \|B_i\|_p + E \max_{1 \leq i \leq m} |X_{ij}| \right] + \log n \max_{1 \leq j \leq n} \|B^{(j)}\|_q \right).
\end{align}

**References**


\(^2\)This means that $R$ treated as an $(mn)$-dimensional vector is log-concave and isotropic.