On the Best Constants in the Khintchine Inequality

Johannes Wissel and Jϕseph Turian
Mentor: Arteom Zvavitch Supervisor: Gideon Schechtman
Department of Theoretical Mathematics, Weizmann Institute of Science

Abstract

We show a self-contained new proof for the best B_p constant in the Khintchine Inequality for $p > 3$ using only elementary calculus.

Introduction

Consider the Khintchine Inequality:

$$A_p \cdot (\sum_{i=1}^{n} \alpha_i^2)^{\frac{1}{2}} \leq (E|\sum_{i=1}^{n} \varepsilon_i \cdot \alpha_i|^p)^{\frac{1}{p}} \leq B_p \cdot (\sum_{i=1}^{n} \alpha_i^2)^{\frac{1}{2}},$$

where $p \in I\!\!R, p > 0$ and, $\alpha_i \in I\!\!R$, for all i such that $1 \leq i \leq n$ and $\varepsilon_1, \cdots, \varepsilon_n$ are independent random variables that each assume the values ± 1 with probabilities $\frac{1}{2}$.

The actual inequality (in which B_p and A_p are not necessarily the best) was proved by Khintchine. But we are looking for the best constants A_p and B_p that satisfy the inequality for all natural numbers n. For all p, a solution and proof was given in [1]. Here we present another proof, for $p > 3$, which uses only elementary calculus.

Proof

Because of the homogeneity of the inequality, we may assume that $\sum_{i=1}^{n} \alpha_i^2 = 1$. Therefore, we can write the inequality in the following way:

$$A_p \leq (E|\sum_{i=1}^{n} \alpha_i \cdot \varepsilon_i|^p)^{\frac{1}{p}} \leq B_p.$$

Using the definition of mathematical expectation, we arrive to the following formula:

$$A_p^p \leq \sum_{\substack{\varepsilon_i = \pm 1, \\ 1 \leq i \leq n}} \frac{1}{2^n} |\sum_{i=1}^{n} \alpha_i \cdot \varepsilon_i|^p \leq B_p^p.$$

And finally:

$$A_p^p \cdot 2^n \leq \sum_{\substack{\varepsilon_i = \pm 1, \\ 1 \leq i \leq n}} |\sum_{i=1}^{n} \alpha_i \cdot \varepsilon_i|^p \leq B_p^p \cdot 2^n.$$

As changing the signs of the α_i does not change the sum, we may assume that every α is non-negative. For $n = 1$, we obtain $\alpha_1 = 1$. If $n \geq 2$ and $a_k = 0$ for some k, without loss of generality let $k = n$. Then note that:

$$\sum_{\substack{\varepsilon_i=\pm 1,\\ 1\leq i\leq n}} |\sum_{i=1}^{n} \alpha_i \cdot \varepsilon_i|^p = 2 \cdot \sum_{\substack{\varepsilon_i=\pm 1,\\ 1\leq i\leq n-1}} |\sum_{i=1}^{n-1} \alpha_i \cdot \varepsilon_i|^p,$$

and that the inequality is therefore

$$A_p^p \cdot 2^{n-1} \leq \sum_{\substack{\varepsilon_i=\pm 1,\\ 1\leq i\leq n-1}} |\sum_{i=1}^{n-1} \alpha_i \cdot \varepsilon_i|^p \leq B_p^p \cdot 2^{n-1}.$$

Thus, the equality for n reduces to the inequality for $n - 1$.

We would like to show that to find the best value of B_p, it suffices to find the upper bound of

$$\frac{1}{2^n} \cdot \sum_{\varepsilon_i=\pm 1} |\sum_{i=1}^{n} \frac{\varepsilon_i}{\sqrt{n}}|^p$$

for all valid ns.

We now define the following function f:

$$f(\alpha_1, \ldots, \alpha_n) = \sum_{\substack{\varepsilon_i=\pm 1,\\ 1\leq i\leq n}} |\sum_{i=1}^{n} \alpha_i \cdot \varepsilon_i|^p,$$

with $\sum_{i=1}^{n} \alpha_i^2 = 1$.

Theorem 1

$$f(\alpha_1, \ldots, \alpha_n) = \sum_{\varepsilon_i=\pm 1} |\sum_{i=1}^{n} \varepsilon_i \alpha_i|^p,$$

under the constraints $p > 3$ and $\sum_{i=1}^{n} \alpha_i^2 = 1$, is not maximal for the given α sequence if there exists an α_j and α_k such that $\alpha_k > \alpha_j > 0$.

It follows from **Theorem 1** that f is only maximal when the non-zero αs are equal.

Proof of Theorem 1:

We prove **Theorem 1** by contradiction. Let us suppose that f is maximal for some $\alpha_1, \cdots, \alpha_n$, such that, without loss of generality, $\alpha_2 > \alpha_1 > 0$.

We define following function:

$$h(\theta) = f(k(\theta), l(\theta), \alpha_3, \ldots, \alpha_n) = \sum_{\varepsilon_i=\pm 1} |\varepsilon_1 k(\theta) + \varepsilon_2 l(\theta) + A_\varepsilon|^p,$$

with A_ε as a constant dependent upon the ε sequence, $\{\varepsilon_i\}_{i=3}^{n}$:

$$A_\varepsilon = \sum_{i=3}^{n} \varepsilon_i \alpha_, i$$

and with

$$\begin{aligned} k(\theta) &= \sqrt{(1-\theta)\alpha_1^2+\theta\alpha_2^2} = \sqrt{\theta(\alpha_2^2-\alpha_1^2)+\alpha_1^2}, \\ l(\theta) &= \sqrt{(1-\theta)\alpha_2^2+\theta\alpha_1^2} = \sqrt{\theta(\alpha_1^2-\alpha_2^2)+\alpha_2^2}, \end{aligned}$$

and $\theta \in [0,1]$. Note that $k(0) = \alpha_1$, $l(0) = \alpha_2$, and $k(\theta)^2 + l(\theta)^2 = \alpha_1^2 + \alpha_2^2$, for all $\theta \in [0,1]$. Because $\alpha_2 \neq 0$, $k(\theta)$ and $l(\theta)$ are not zero-functions.

If we can show that $h(\theta)$ increases at $\theta = 0$, then we have a contradiction because f increases with $h(\theta)$ and is therefore not maximal for $\alpha_1, \cdots, \alpha_n$ as assumed.

Lemma 1 *$h(\theta)$ is increasing at $\theta = 0$.*

Proof of Lemma 1:

$$\begin{aligned} k'(\theta) &= \frac{\alpha_2^2-\alpha_1^2}{2\sqrt{\theta(\alpha_2^2-\alpha_1^2)+\alpha_1^2}} \\ l'(\theta) &= \frac{\alpha_1^2-\alpha_2^2}{2\sqrt{\theta(\alpha_1^2-\alpha_2^2)+\alpha_2^2}} \end{aligned}$$

and therefore $k'(0) = \frac{a_2^2-a_1^2}{2\cdot a_1}$ and $l'(0) = \frac{a_1^2-a_2^2}{2\cdot a_2}$.

Note that $k(\theta)k'(\theta) + l(\theta)l'(\theta) = 0$ for any θ.

Define the following constant: $c = \frac{\alpha_2^2-\alpha_1^2}{2\cdot\alpha_1\cdot\alpha_2}$. Thus, $c > 0$.

Note that

$$k'(0) - l'(0) = \tfrac{\alpha_2^2-\alpha_1^2}{2\cdot\alpha_1\cdot\alpha_2}\cdot(l(0)+k(0)) \quad = c\cdot(l(0)+k(0))$$

and

$$k'(0) + l'(0) = \tfrac{\alpha_2^2-\alpha_1^2}{2\cdot\alpha_1\cdot\alpha_2}\cdot(l(0)-k(0)) \quad = c\cdot(l(0)-k(0))$$

Also note that

$$k(0)l'(0) + k'(0)l(0) = \frac{(\alpha_2^2-\alpha_1^2)^2}{2\cdot\alpha_1\cdot\alpha_2} = c\cdot(\alpha_2^2-\alpha_1^2).$$

$$h'(\theta) = \quad p\cdot\sum_{\varepsilon_i=\pm1}|\varepsilon_1 k(\theta)+\varepsilon_2 l(\theta)+A_\varepsilon|^{p-2}(\varepsilon_1 k(\theta)+\varepsilon_2 l(\theta)+A_\varepsilon)(\varepsilon_1 k'(\theta)+\varepsilon_2 l'(\theta))$$

From this point forward, k (and k') shall be understood as $k(0)$ (and $k'(0)$) and l (and l') as $l(0)$ (and $l'(0)$).

$$\begin{aligned} h'(0) = &\quad p\cdot\sum_{\varepsilon_i=\pm1}|\varepsilon_1 k+\varepsilon_2 l+A_\varepsilon|^{p-2}(\varepsilon_1 k+\varepsilon_2 l+A_\varepsilon)(\varepsilon_1 k'+\varepsilon_2 l') \\ = &\quad p\cdot\sum_{\varepsilon_i=\pm1}|\varepsilon_1 k+\varepsilon_2 l+A_\varepsilon|^{p-2}(k\cdot k'+l\cdot l'+\varepsilon_1\cdot\varepsilon_2(k\cdot l'+k'\cdot l)+ \\ + &\quad A_\varepsilon(\varepsilon_1 k'+\varepsilon_2 l')) \\ = &\quad p\cdot\sum_{\varepsilon_i=\pm1}|\varepsilon_1 k+\varepsilon_2 l+A_\varepsilon|^{p-2}(\varepsilon_1\cdot\varepsilon_2\cdot c\cdot(\alpha_2^2-\alpha_1^2)+A_\varepsilon(\varepsilon_1 k'+\varepsilon_2 l')) \end{aligned}$$

Replacing ε_1 and ε_2 with explicit values, we obtain the following equation:

$$\begin{aligned} h'(0) = p \cdot \sum_{\substack{\varepsilon_i=\pm 1,\\ 3\le i\le n}} & |k+l+A_\varepsilon|^{p-2}(c\cdot(\alpha_2^2-\alpha_1^2)+A_\varepsilon(k'+l')) - \\ & - |k-l-A_\varepsilon|^{p-2}(c\cdot(\alpha_2^2-\alpha_1^2)+A_\varepsilon(k'-l')) - \\ & - |k-l+A_\varepsilon|^{p-2}(c\cdot(\alpha_2^2-\alpha_1^2)-A_\varepsilon(k'-l')) + \\ & + |k+l-A_\varepsilon|^{p-2}(c\cdot(\alpha_2^2-\alpha_1^2)-A_\varepsilon(k'+l')) \\ = p \cdot \sum_{\substack{\varepsilon_i=\pm 1,\\ 3\le i\le n}} & |k+l+A_\varepsilon|^{p-2}(c\cdot(\alpha_2^2-\alpha_1^2)+A_\varepsilon\cdot c\cdot(l-k) - \\ & - |k-l-A_\varepsilon|^{p-2}(c\cdot(\alpha_2^2-\alpha_1^2)+A_\varepsilon\cdot c\cdot(l+k) - \\ & - |k-l+A_\varepsilon|^{p-2}(c\cdot(\alpha_2^2-\alpha_1^2)-A_\varepsilon\cdot c\cdot(l+k) + \\ & + |k+l-A_\varepsilon|^{p-2}(c\cdot(\alpha_2^2-\alpha_1^2)-A_\varepsilon\cdot c\cdot(l-k) \end{aligned}$$

Note that $\forall A_\varepsilon \ge 0$, there is an $A_{-\varepsilon} \le 0$ such that $|A_\varepsilon| = |A_{-\varepsilon}|$. The inverse is also true. Therefore, the equation can be condensed:

$$\begin{aligned} h'(0) = 2\cdot p \cdot \sum_{\substack{\varepsilon_i=\pm 1,\\ 3\le i\le n}} & |k+l+A_\varepsilon|^{p-2}(c\cdot(\alpha_2^2-\alpha_1^2)+A_\varepsilon\cdot c\cdot(l-k) - \\ & - |k-l+A_\varepsilon|^{p-2}(c\cdot(\alpha_2^2-\alpha_1^2)-A_\varepsilon\cdot c\cdot(l+k) \\ = 2\cdot p\cdot c \cdot \sum_{\substack{\varepsilon_i=\pm 1,\\ 3\le i\le n}} & |k+l+A_\varepsilon|^{p-2}(\alpha_2^2-\alpha_1^2+A_\varepsilon\cdot(l-k) - \\ & - |k-l+A_\varepsilon|^{p-2}(\alpha_2^2-\alpha_1^2-A_\varepsilon\cdot(l+k) \end{aligned}$$

Now we compare every non-negative A_ε with its corresponding $A_{-\varepsilon}$.

$$\begin{aligned} h'(0) = 2\cdot p\cdot c \cdot \sum_{\substack{\varepsilon_i=\pm 1,\\ A_\varepsilon\ge 0}} & |k+l+A_\varepsilon|^{p-2}((\alpha_2^2-\alpha_1^2)+A_\varepsilon\cdot(l-k) + \\ & + |k+l-A_\varepsilon|^{p-2}((\alpha_2^2-\alpha_1^2)-A_\varepsilon\cdot(l-k) + \\ & - |k-l+A_\varepsilon|^{p-2}((\alpha_2^2-\alpha_1^2)-A_\varepsilon\cdot(l+k) - \\ & - |k-l-A_\varepsilon|^{p-2}((\alpha_2^2-\alpha_1^2)+A_\varepsilon\cdot(l+k). \end{aligned}$$

Note that when the summation constraint is $A_\varepsilon \ge 0$, the sum is greater than or equal to the sum when the summation constraint is $A_\varepsilon > 0$, that is

$$\begin{aligned} h'(0) \ge 2\cdot p\cdot c \cdot \sum_{\substack{\varepsilon_i=\pm 1,\\ A_\varepsilon > 0}} & |k+l+A_\varepsilon|^{p-2}((\alpha_2^2-\alpha_1^2)+A_\varepsilon\cdot(l-k) + \\ & + |k+l-A_\varepsilon|^{p-2}((\alpha_2^2-\alpha_1^2)-A_\varepsilon\cdot(l-k) + \\ & - |k-l+A_\varepsilon|^{p-2}((\alpha_2^2-\alpha_1^2)-A_\varepsilon\cdot(l+k) - \\ & - |k-l-A_\varepsilon|^{p-2}((\alpha_2^2-\alpha_1^2)+A_\varepsilon\cdot(l+k). \end{aligned}$$

Let us prove that each summand is greater than 0, which will prove that $h'(0) > 0$. For each summand, define the constants: $x_\varepsilon = \frac{l+k}{A_\varepsilon}$ and $y_\varepsilon = \frac{l-k}{A_\varepsilon}$. Note that $x_\varepsilon > y_\varepsilon > 0$. Then:

$$\begin{aligned} h'(0) \geq 2 \cdot p \cdot c \cdot \sum_{\substack{\varepsilon_i = \pm 1, \\ A_\varepsilon > 0}} A_\varepsilon^p \cdot (\quad & |x_\varepsilon + 1|^{p-2}(x_\varepsilon \cdot y_\varepsilon + y_\varepsilon) + \\ + \quad & |x_\varepsilon - 1|^{p-2}(x_\varepsilon \cdot y_\varepsilon - y_\varepsilon) - \\ - \quad & |y_\varepsilon - 1|^{p-2}(x_\varepsilon \cdot y_\varepsilon - x_\varepsilon) - \\ - \quad & |y_\varepsilon + 1|^{p-2}(x_\varepsilon \cdot y_\varepsilon + x_\varepsilon)) \end{aligned}$$

Therefore, it is enough to prove that for all $x > y > 0$

$$\begin{aligned} \gamma_x(y) = \quad & (|x_\varepsilon + 1|^{p-2}(x_\varepsilon + 1) + |x_\varepsilon - 1|^{p-2}(x_\varepsilon - 1)) \cdot y - \\ - \quad & (|y_\varepsilon - 1|^{p-2}(y_\varepsilon - 1) + |y_\varepsilon + 1|^{p-2}(y_\varepsilon + 1)) \cdot x > 0 \end{aligned}$$

Note that $\gamma_x(0) = 0$ and $\gamma_x(x) = 0$. Thus, if we can prove that $\gamma_x''(y)$ is less than 0, we will prove that $\gamma_x(y)$ is greater than 0 for valid values of x and y.

$\gamma_x'(y) = (|x+1|^{p-2} \cdot (x+1) + |x-1|^{p-2} \cdot (x-1)) - x \cdot (p-1) \cdot (|y-1|^{p-2} + |y+1|^{p-2})$.

Case 1: $y - 1 = 0$

The question that arises is if $\gamma_x'(y)$ is differentiable, or, more specifically, if $\delta(y) = |y-1|^{p-2}$ is differentiable at $y = 1$.
$\delta'(y) = (p-2) \cdot (y-1)^{p-3}$ for $y - 1 > 0$
$\delta'(y) = -(p-2) \cdot (1-y)^{p-3}$ for $y - 1 < 0$
By the definition of the derivative,

$$\delta'(y_0) = \lim_{y \to y_0} \frac{|y-1|^{p-2} - |y_0 - 1|^{p-2}}{y - y_0},$$

(if the limit exists).

To prove that the limit exists, we take the limit from both sides.

$$\begin{aligned} \lim_{y \to 1^+} \frac{|y-1|^{p-2}}{y-1} &= \lim_{y \to 1^+} |y-1|^{p-3} = 0 \\ \lim_{y \to 1^-} \frac{|y-1|^{p-2}}{y-1} &= -\lim_{y \to 1^-} \frac{|y-1|^{p-2}}{|y-1|} = -\lim_{y \to y^-} |y-1|^{p-3} = 0 \end{aligned}$$

So thus $\delta'(1) = 0$ and $\gamma_x''(y) < 0$.

Otherwise, when $y - 1 \neq 0$,

$$\gamma_x''(y) = -x \cdot (p-1) \cdot (p-2) \cdot (|y-1|^{p-4} \cdot (y-1) + |y+1|^{p-4} \cdot (y+1)).$$

Case 2: $y - 1 > 0$

$$\gamma_x''(y) = -x \cdot (p-1) \cdot (p-2) \cdot (|1+y|^{p-3} + |1-y|^{p-3}) < 0.$$

Case 3: $y - 1 < 0$

$$\gamma_x''(y) = -x \cdot (p-1) \cdot (p-2) \cdot (|1+y|^{p-3} - |1-y|^{p-3}) < 0.$$

Therefore, $\gamma_x(y)$ is greater than 0 for $x > y > 0$. Note that this inequality does not hold for $p \leq 3$
Thus, $h'(0) > 0$ and **Lemma 1** has been proven.

Thus, **Theorem 1** has been proven.

Q.E.D.

Conclusion

Thus, using elementary calculus, the best B_p constant for the Khintchine Inequality has been derived for $p > 3$.

Suggestions for Future Research

It could be possible to create a proof for $2 < p < 3$ using a similar methodology to that of this project. If we can show that $h'(0) = 0$ implies $h''(0) > 0$, then we could also get the contradiction with which we proved **Theorem 1**.

Acknowledgments

We would like to thank our mentor, Arteom Zvavitch, for all his support and inspiration throughout this project.
We would also like to the coordinators and staff of the 29th Annual Bessie F. Lawrence Summer Science Institute and the American Committee for the Weizmann Institute of Science for giving us the opportunity to "transport the sun."

References

[1] HAAGERUP U. (1981). *The best constants in the Khintchine inequality*, Studia Mathematica **70**, pp. 232-283.