LECTURE NOTES ON APPLICATIONS OF GROTHENDIECK'S INEQUALITY

THE INEQUALITY

JOP BRIËT

Abstract. In this lecture we cover Grothendieck's inequality via an extremely elegant proof of Krivine. In addition, we cover a useful lemma of Grothendieck's that yields a "factorization" version of the inequality.

1. Grothendieck's inequality

Let $\langle \cdot, \cdot \rangle$ be the standard inner product on $\mathbb{R}^d$. Denote the set of unit vectors in $\mathbb{R}^d$ by $S^{d-1} = \{x \in \mathbb{R}^d : \langle x, x \rangle = 1\}$. Grothendieck's inequality can be given in terms of the following two quantities of a matrix $A \in \mathbb{R}^{n\times n}$:

$$\|A\|_{\infty\to 1} = \max\Big\{ \sum_{i,j=1}^n A_{ij} a_i b_j \,:\, a, b \in \{-1,1\}^n \Big\}$$

$$\|A\|_G = \sup\Big\{ \sum_{i,j=1}^n A_{ij} \langle x_i, y_j \rangle \,:\, d \in \mathbb{N},\, x_i, y_j \in S^{d-1} \Big\}.$$

The notation suggests that these quantities are norms and it turns out that they are, but we shall not use this fact here.

Theorem 1.1 (Grothendieck's inequality)**.** *There exists an absolute constant $K_G \in (1,2)$ such that the following holds. For any positive integer n and matrix $A \in \mathbb{R}^{n\times n}$, we have*

$$\|A\|_G \le K_G \|A\|_{\infty\to 1}. \tag{1}$$

Here we give arguably the most elegant proof of this theorem, which is due to Krivine [Kri79].

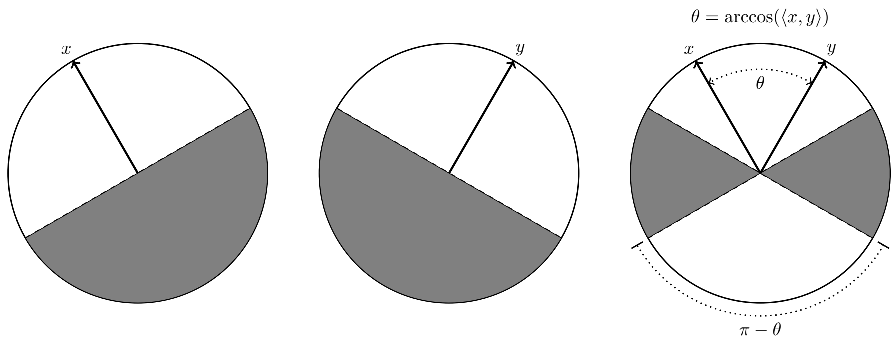

Figure 1. Grothendieck's identity in two dimensions.

2. Krivine's proof of Grothendieck's inequality

The first ingredient of the proof is the following simple lemma.

Lemma 2.1 (Grothendieck's identity)**.** *Let x, y be n-dimensional real unit vectors and let $g = (g_1, \dots, g_n) \sim N(0, I_n)$ be an n-dimensional standard Gaussian vector. Then,*

$$\mathbb{E}\big[\,\mathrm{sign}(\langle x, g\rangle)\,\mathrm{sign}(\langle y, g\rangle)\big] = \frac{2}{\pi}\arcsin(\langle x, y\rangle). \tag{2}$$

Proof sketch: If $x = y$ or $x = -y$ then there is nothing to prove, so assume that $x \neq y$. We begin by projecting to the two-dimensional subspace spanned by the pair x, y. The projection of g onto this subspace is still a two-dimensional standard Gaussian. Observe that $\mathrm{sign}(\langle x, g\rangle)\,\mathrm{sign}(\langle y, g\rangle)$ is positive if and only if g lies above or below both of the half-planes orthogonal to x and y respectively (Figure 1). Since the direction of g is uniform on the unit circle, it follows that this happens with probability

$$\frac{2}{2\pi}\big(\pi - \arccos(\langle x, y\rangle).$$

Hence, the expectation in (2) equals

$$\frac{1}{\pi}\big(\pi - \arccos(\langle x, y\rangle) - \frac{1}{\pi}\big(\,\arccos(\langle x, y\rangle) = 1 - \frac{2}{\pi}\arccos(\langle x, y\rangle),$$

which equals the right-hand side of (2). □

We will use the Taylor expansions of the sine and hyperbolic sine functions, given by

$$\sin(t) = \sum_{k=1}^{\infty} \alpha_k t^k$$
$$\sinh(t) = \sum_{k=1}^{\infty} |\alpha_k| t^k,$$

for some sequence of Taylor coefficients $(\alpha_k)_k \subseteq \mathbb{R}$. Note the fact that the Taylor coefficients of the hyperbolic sine functions are given by the absolute values of those of the sine function. Both of these Taylor series converge on the interval $[-1, 1]$.

The *tensor product* of vectors $x \in \mathbb{R}^{d_1}$ and $y \in \mathbb{R}^{d_1}$ is the $d_1 d_2$-dimensional vector $x \otimes y \in \mathbb{R}^{d_1 d_2}$ given by

$$x \otimes y = (x_i y_j)_{(i,j) \in [d_1] \times [d_2]}.$$

For a positive integer k, denote by $x^{\otimes k}$ the k-fold iterated tensor product of x with itself.

Proposition 2.2. *For any $x, y \in \mathbb{R}^d$, we have $\langle x^{\otimes k}, y^{\otimes k} \rangle = \langle x, y \rangle^k$.*

Proof of Theorem 1.1: Let

$$c = \sinh^{-1}(1) = \ln(1 + \sqrt{2}).$$

Fix a positive integer d and two sets of d-dimensional unit vectors $x_1, \ldots, x_n, y_1, \ldots, y_n \in S^{d-1}$. We show that there exist $\{-1, 1\}$-valued random variables $a_1, \ldots, a_n$, $b_1, \ldots, b_n$ such that

$$\mathbb{E}[a_i b_j] = \frac{2c}{\pi} \langle x_i, y_j \rangle$$

holds for all $i, j \in [n]$. To see why this suffices to prove the theorem, observe that by linearity of expectation,

$$\frac{2c}{\pi} \sum_{i,j=1}^{n} A_{ij} \langle x_i, y_j \rangle = \mathbb{E}\Big[\sum_{i,j=1}^{n} A_{ij} a_i b_j \Big] \leq \max_{a_i, b_j \in \{-1,1\}} \sum_{i,j=1}^{n} A_{ij} a_i b_j.$$

This shows that $K_G \leq \pi/(2c)$.

To obtain the random signs, define two new sequences of vectors:

$$u_i = \bigoplus_{k=1}^{\infty} \sqrt{|\alpha_k|} c^{k/2}\, x_i^{\otimes k}$$

$$v_j = \bigoplus_{k=1}^{\infty} \operatorname{sign}(\alpha_k)\, \sqrt{|\alpha_k|} c^{k/2}\, y_j^{\otimes k}.$$

Using the Taylor expansions of sin and sinh and Proposition 2.2 it is easy to verify that these are unit vectors and that

$$\langle u_i, v_j \rangle = \sin(c\langle x_i, v_j \rangle).$$

Observe that these are infinite-dimensional. However, since there are only $2n$ of them, they span a space of dimension at most $2n$, and it follows that there exist unit vectors $u_1', \dots, u_n',\ v_1', \dots, v_n' \in S^{2n-1}$ such that $\langle u_i', v_j' \rangle = \langle u_i, v_j \rangle$ for all $i, j \in [n]$. Let $g = (g_1, \dots, g_{2n})$ be a random vector of independent standard normal random variables. Define

$$a_i = \operatorname{sign}(\langle u_i', g \rangle) \quad \text{and} \quad b_j = \operatorname{sign}(\langle v_j', g \rangle).$$

Then, by Lemma 2.1,

$$\begin{aligned} \frac{\pi}{2}\mathbb{E}[a_i b_j] &= \arcsin(\langle u_i', v_j' \rangle) \\ &= \arcsin(\sin(c\langle x_i, y_j \rangle)) \\ &= c\langle x_i, y_j \rangle. \end{aligned}$$

This proves the theorem. □

The *Grothendieck constant* K_G is the smallest real number for which Theorem 1.1 holds true. The problem of determining its exact value, posed in [Gro53], remains open. The best lower and upper bounds $1.6769 \dots \le K_G < 1.7822 \dots$ were proved by Davie and Reeds [Dav84, Ree91], and Braverman et al. [BMMN13], resp.

3. Grothendieck factorization

For a vector $x \in \mathbb{R}^n$, denote by $\operatorname{Diag}(x) \in \mathbb{R}^{n \times n}$ the matrix whose diagonal is x and whose off-diagonals are all zero. The *operator norm* of a matrix $A \in \mathbb{R}^{n \times n}$ is defined by

$$\|A\| = \sup\left\{ |\langle Ax, y \rangle| :\ x, y \in S^{n-1} \right\}.$$

Lemma 3.1 (Grothendieck)**.** *For any matrix $A \in \mathbb{R}^{n\times n}$, there exist positive unit vectors $u, v \in \mathbb{R}_{>0} \cap S^{n-1}$ such that for any $x, y \in \mathbb{R}^n$,*

$$|\langle Ax, y\rangle| \le \|A\|_G \, \|x \circ u\|_2 \|y \circ v\|_2, \tag{3}$$

where $\circ$ denotes the entry-wise product.

Corollary 3.2. *For any matrix $A \in \mathbb{R}^{n\times n}$, there exist positive unit vectors $u, v \in \mathbb{R}^n_{>0} \cap S^{n-1}$ such that the matrix*

$$B = \frac{1}{K_G} \operatorname{Diag}(u)^{-1} A \operatorname{Diag}(v)^{-1} \tag{4}$$

satisfies $\|B\| \le \|A\|_{\ell_\infty \to \ell_1}$.

The proof of Lemma 3.1 is based on a clever application of the Hahn–Banach separation theorem [Rud91, Theorem 3.4]. Recall that a set $K \subseteq \mathbb{R}^d$ is *convex* if for any $x, y \in K$ and $\lambda \in [0, 1]$, we have that $\lambda x + (1 - \lambda_y \in K$. Moreover, K is a *cone* if for any $x \in K$ and $\lambda > 0$, we have that $\lambda x \in K$.

Theorem 3.3 (Hahn–Banach separation theorem)**.** *Let $C, D \subseteq \mathbb{R}^n$ be convex sets and let C be algebraically open. Then, the sets C and D are disjoint if and only if there exists a vector $\lambda \in \mathbb{R}^n$ and an $\alpha \in \mathbb{R}$ such that $\langle \lambda, c\rangle < \alpha$ for every $c \in C$ and $\langle \lambda, d\rangle \ge \alpha$ for every $d \in D$. Morever, if C and D are convex cones we may take $\alpha = 0$.*

Proof of Lemma 3.1: Let $M = A/\|A\|_G$, so that $\|M\|_G \le 1$. Then, for arbitrary vectors x_i, y_j, we have

$$\sum_{i,j=1}^{n} M_{ij} \langle x_i, y_j\rangle \le \max_{i,j\in[n]} \|x_i\|\|y_j\| \le \frac{1}{2} \max_{i,j\in[n]} (\|x_i\|^2 + \|y_j\|^2), \tag{5}$$

where the second inequality is by AMGM inequality. Define the set $K \subseteq \mathbb{R}^{n\times n}$ by

$$K = \left\{ \left(\|x_i\|^2 + \|y_j\|^2 - 2 \sum_{k,\ell=1}^{n} M_{k\ell} \langle x_k, y_\ell\rangle \right)_{i,j=1}^{n} : \, d \in \mathbb{N}, \;\, x_i, y_j \in \mathbb{R}^d \right\}.$$

We show that K is a convex cone. For every $t \in \mathbb{R}_+$ and matrix $Q \in K$ given by vectors x_i, y_j, the vectors $x'_i = \sqrt{t} x_i$ and $y'_j = \sqrt{t} y_j$ similarly define tQ, and so K is a cone. We now show K is a convex set. Let $Q, Q' \in K$ be specified by x_i, y_j and x'_i, y'_j respectively. Then, for any $\lambda \in [0, 1]$, the convex combination $\lambda Q + (1-\lambda)Q'$ also belongs to K, as it can be specified by the vectors $(\sqrt{\lambda} x_i, \sqrt{1-\lambda} x'_i), (\sqrt{\lambda} y_j, \sqrt{1-\lambda} y'_j)$.

Additionally, it follows from (5) that K is disjoint from the open convex cone $\mathbb{R}^{n\times n}_{<0}$ of matrices with strictly negative entries. By Theorem 3.3 (the Hahn–Banach separation theorem), we conclude that there exists a nonzero matrix $L \in \mathbb{R}^{n\times n}$ such that $\langle L, Q\rangle \geq 0$ for every $Q \in K$ and $\langle L, N\rangle < 0$ for every $N \in \mathbb{R}^{n\times n}_{<0}$. In particular, the second inequality implies that $L \in \mathbb{R}^{n\times n}_{+}$. Let $P = L/\sum_{ij} L_{ij}$, so that $\{P_{ij}\}_{i,j=1}^n$ defines a probability distribution over $[n]^2$. Then, for any $Q \in K$,

$$\begin{aligned}
0 \leq \langle P, Q\rangle &= \sum_{i,j=1}^n P_{ij}(\|x_i\|^2 + \|y_j\|^2) - 2\sum_{k,\ell=1}^n M_{k\ell}\langle x_k, y_\ell\rangle \\
&= \sum_{i=1}^n \sigma_i\|x_i\|^2 + \sum_{j=1}^n \mu_j\|y_j\|^2 - 2\sum_{k,\ell=1}^n M_{k\ell}\langle x_k, y_\ell\rangle,
\end{aligned}$$

where $\sigma_i = P_{i1} + \cdots + P_{in}$ and $\mu_j = P_{1j} + \cdots + P_{nj}$. Observe that σ_i, μ_j are strictly positive because $P_{ij} > 0$. Rearranging the inequality above and using bi-linearity, it follows that for every $\lambda > 0$, we have

$$\begin{aligned}
2\sum_{k,\ell=1}^n M_{k\ell}\langle x_k, y_\ell\rangle &= 2\sum_{k,\ell=1}^n M_{k\ell}\langle \lambda x_k, \lambda^{-1} y_\ell\rangle \\
&\leq \lambda^2 \sum_{i=1}^n \sigma_i\|x_i\|_2^2 + \lambda^{-2}\sum_{j=1}^n \mu_j\|y_j\|_2^2. \qquad (6)
\end{aligned}$$

Setting

$$\lambda = \left(\frac{\sum_{j=1}^n \mu_j\|y_j\|_2^2}{\sum_{i=1}^n \sigma_i\|x_i\|_2^2}\right)^{1/4}$$

in Eq. (6), we find that

$$2\sum_{k,\ell=1}^n M_{k\ell}\langle x_k, y_\ell\rangle \leq \Big(\sum_{i=1}^n \sigma_i\|x_i\|_2^2\Big)^{1/2}\Big(\sum_{j=1}^n \mu_j\|y_j\|_2^2\Big)^{1/2}.$$

In particular, for the case where $x_k, y_\ell \in \mathbb{R}$, we have

$$x^{\mathsf{T}} M y \leq \|\operatorname{diag}(\sigma)^{1/2} x\|_2 \|\operatorname{diag}(\mu)^{1/2} y\|_2.$$

This implies

$$x^{\mathsf{T}}\Big(\operatorname{Diag}(\sigma)^{-1/2} M \operatorname{Diag}(\mu)^{-1/2}\Big) y \leq \|x\|_2 \cdot \|y\|_2,$$

which in particular implies that $\|\operatorname{Diag}(\sigma)^{-1/2} M \operatorname{Diag}(\mu)^{-1/2}\| \leq 1$. The result follows by letting $u_i = \sqrt{\sigma_i}$, $v_i = \sqrt{\mu_i}$ for every $i \in [n]$. $\square$

4. Exercises

Exercise 4.1. Let $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$. Show that $\|A\|_G \geq \sqrt{2}\|A\|_{\infty\to 1}$. [Hint: For the lower bound on $\|A\|_G$, two-dimensional vectors suffice.]

Exercise 4.2. Let $A \in \mathbb{R}^{n\times n}$ and $u, v \in \mathbb{R}_{>0} \cap S^{n-1}$ be such that for any $x, y \in \mathbb{R}^n$, we have $|\langle Ax, y\rangle| \leq \|x \circ y\|_2 \|y \circ v\|_2$. Prove that $\|A\|_G \leq 1$. Conclude that equality holds in Lemma 3.1.

References

[BMMN13] Mark Braverman, Konstantin Makarychev, Yury Makarychev, and Assaf Naor. The Grothendieck constant is strictly smaller than Krivine's bound. *Forum Math. Pi*, 1:e4, 42, 2013. Preliminary version in FOCS'11.

[Dav84] Alexander Davie. Lower bound for K_G. Unpublished, 1984.

[Gro53] Alexander Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques (French). *Bol. Soc. Mat. São Paulo*, 8:1–79, 1953.

[Kri79] J. L. Krivine. Constantes de Grothendieck et fonctions de type positif sur les sphères. *Adv. Math.*, 31:16–30, 1979.

[Ree91] James Reeds. A new lower bound on the real Grothendieck constant. Manuscript (`http://www.dtc.umn.edu/~reedsj/bound2.dvi`), 1991.

[Rud91] W. Rudin. *Functional analysis*. McGraw-Hill Science, 1991.