LINEAR AND HEREDITARY DISCREPANCY

BENJAMIN DOERR

ABSTRACT. Let A be a $m \times n$–matrix and $q := \lfloor \log_2(m) \rfloor + 1$. In this article we are going to improve the well–known bound $\operatorname{lindisc}(A) \leq 2 \operatorname{herdisc}(A)$ and show

$$\operatorname{lindisc}(A) \leq 2\left(1 - 2^{-q}\right) \operatorname{herdisc}(A) \quad \left(\leq 2(1 - \frac{1}{2m}) \operatorname{herdisc}(A)\right).$$

Like the previous proofs on this problem, ours is constructive as well. We'll give a so-called "on–line algorithm" and analyze it using game theory.

INTRODUCTION

Discrepancy theory is dealing with questions of the kind "How far does the optimal solution of a problem deviate from an ideal solution?" In combinatorial discrepancy theory our problem is to color the points of a hypergraph with two colors in such a way that all hyperedges are about balancedly colored. So an ideal solution would be one where every hyperedge has the same number of points in one color as in the other. To be more precise:

Let $\mathcal{H} = (X, \mathcal{E})$ denote a finite *hypergraph*, i. e. X is a finite set (of "*vertices*") and $\mathcal{E}$ is a family of subsets of X (called *hyperedges*). A *coloring* of $\mathcal{H}$ is simply a mapping $\varepsilon : X \to \{-1, +1\}$ (we call -1 and $+1$ *colors*). For a hyperedge $E \in \mathcal{E}$ define $\varepsilon(E) = \sum_{x \in E} \varepsilon(x)$. The *discrepancy* of $\mathcal{H}$ is defined by

$$\operatorname{disc}(\mathcal{H}) = \min_{\varepsilon : X \to \{-1,+1\}} \max_{E \in \mathcal{E}} |\varepsilon(E)|.$$

This concept may be generalized to matrices in a natural way. Let $A = (a_{ij})$ be any $m \times n$–matrix and set

$$\operatorname{disc}(A) := \min_{\varepsilon \in \{-1,1\}^n} \|A\varepsilon\|,$$

Date: January 4, 1999.

1991 *Mathematics Subject Classification.* Primary 11K38; Secondary 90D05.

Key words and phrases. linear discrepancy, hereditary discrepancy, on-line algorithm, game theory.

supported by the graduate school 'Effiziente Algorithmen und Multiskalenmethoden', Deutsche Forschungsgemeinschaft.

where the norm shall always denote the maximum–norm.

If A is the incidence–matrix of $\mathcal{H}$, we have $\operatorname{disc}(A) = \operatorname{disc}(\mathcal{H})$, so discrepancy of matrices is indeed a more general concept (if, of course, we restrict ourselves to zero–one–matrices both concepts are equivalent). For our specific problem the matrix–concept provides no extra difficulties, so we'll continue in the language of matrices.

There are two related notions: The *linear discrepancy* of A is defined as

$$\operatorname{lindisc}(A) := \max_{p \in [-1,1]^n} \min_{\varepsilon \in \{-1,1\}^n} \|A(p - \varepsilon)\|.$$

Linear discrepancy can be seen as a measure of how well a fractional solution can be rounded to an integer solution. The *hereditary discrepancy* is

$$\operatorname{herdisc}(A) := \max_{J \subseteq [n]} \operatorname{disc}((a_{ij})_{i \in [m], j \in J}).$$

If A is the incidence matrix of a hypergraph, $\operatorname{herdisc}(A)$ is just the maximum discrepancy of the induced subgraphs. As a general reference we'd like to recommend Beck and Sós[1].

It's easy to see that we have both $\operatorname{disc}(A) \le \operatorname{lindisc}(A)$ and $\operatorname{disc}(A) \le \operatorname{herdisc}(A)$ for any A. More interesting is the relation between $\operatorname{lindisc}(A)$ and $\operatorname{herdisc}(A)$. Results of Beck and Spencer[2] and Lovász, Spencer and Vesztergombi[3] show that there is a non–trivial relation between the two concepts: For any A we have $\operatorname{lindisc}(A) \le 2 \operatorname{herdisc}(A)$. In [4] Spencer gives an improvement in terms of n: $\operatorname{lindisc}(A) \le 2(1 - 2^{-2^n}) \operatorname{herdisc}(A)$. He also gives an example of an A satisfying $\operatorname{lindisc}(A) = 2(1 - \frac{1}{n+1}) \operatorname{herdisc}(A)$ and introduces the problem to close this gap.

Our Results

The main result of this article is

Theorem. *Let A be any $m \times n$–matrix. Set $q := \lfloor \log_2(m) \rfloor + 1$. Then*

$$\operatorname{lindisc}(A) \le 2\left(1 - 2^{-q}\right) \operatorname{herdisc}(A).$$

If A is the incidence matrix of a hypergraph (i. e. $A \in \{0,1\}^{m \times n}$), then we may assume $m \le 2^n - 1$, as each two rows can be assumed different and different from $(0, 0, \dots, 0)$. (In the language of hypergraphs this just means that no edge occurs twice and that the empty edge can be ignored). This yields $2\,(1 - 2^{-q}) \le 2\,(1 - 2^{-n})$, giving

Corollary. *For $A \in \{0,1\}^{m \times n}$ we have*

$$\operatorname{lindisc}(A) \le 2(1 - 2^{-n}) \operatorname{herdisc}(A).$$

So for hypergraphs in special and matrices with $m \ll 2^{2^n}$ in general our result gives considerable improvement to the known results; the improvement being better the sparser the matrix is. Our result is quite close to the optimal: The example of [4] has

$$\operatorname{lindisc}(A) = 2\left(1-\frac{1}{m}\right)\operatorname{herdisc}(A),$$

the theorem gives

$$\operatorname{lindisc}(A) \leq 2\left(1-\frac{1}{m+1}\right)\operatorname{herdisc}(A)$$

for $m = 2^l - 1$.

The Proof

Our proof uses the original proof, which we state here for convenience:

Original Proof: Let $p \in [-1,1]^n$. We'll construct an $\varepsilon \in \{-1,1\}^n$ such that $\|A(p-\varepsilon)\|$ is small. Set $a_j^{(0)} := \frac{1}{2}(p_j+1) \in [0,1]$ for all $j \in [n]$. As

$$p \mapsto \min_{\varepsilon \in \{-1,1\}^n} \|A(p-\varepsilon)\|$$

is a continuous function and $\{\sum_{i=1}^n x_i 2^{-i} | n \in \mathbb{N}, x \in \{0,1\}^n\}$ is dense in $[0,1]$, we may assume that there is $k \in \mathbb{N}$ such that $a_j^{(0)} 2^k \in \mathbb{Z}$ for all $j \in [n]$. We are going to round the $a^{(0)}$ successively to a vector of shorter binary expansion until we have a 0–1–vector.

Suppose that for a $l \in \{0,\ldots,k-1\}$ the $(a_j^{(l)})_{j=1,\ldots,n}$ are already defined and fulfill $a_j^{(l)} 2^{k-l} \in \mathbb{Z}$ for all $j \in [n]$. Set $X := \{j \in [n] | a_j^{(l)} 2^{k-l} \text{ odd}\}$, the set of all j such that the binary expansion of $a_j^{(l)} 2^{k-l}$ ends on 1 (these are the components of $a^{(l)}$ that need to be rounded). Find $\varepsilon^{(l)}: X \to \{-1,+1\}$ such that

$$d_i^{(l)} := \sum_{j \in X} \varepsilon_j^{(l)} a_{ij} \in [-\operatorname{herdisc}(A), \operatorname{herdisc}(A)]$$

for all $i \in [m]$.

Define

$$a_j^{(l+1)} := \begin{cases} a_j^{(l)} + 2^{-(k-l)}\varepsilon_j^{(l)} & \text{if } j \in X \\ a_j^{(l)} & \text{otherwise.} \end{cases}$$

Then we have $a_j^{(l+1)} 2^{k-l-1} \in \mathbb{Z}$ for all $j \in [n]$ and

$$\sum_{j\in[n]} a_{ij}(a_j^{(l+1)} - a_j^{(l)}) = 2^{-(k-l)} d_i^{(l)}$$

for all $i \in [m]$, that means we rounded the $a^{(l)}$ to $a^{(l+1)}$ in a way that $\|A(a^{(l+1)} - a^{(l)})\|$ is small.

Having defined $a_j^{(l)}$ for all $j \in [n], l \in \{0, \dots, k\}$ we set $\varepsilon_j := 2a_j^{(k)} - 1$ (this is in $\{-1, 1\}$) and have

$$\begin{aligned} \|A(\varepsilon - p)\| &= 2\left\|A(a^{(k)} - a^{(0)})\right\| \\ &= 2\left\|\sum_{l\in[k]} A(a^{(l)} - a^{(l-1)})\right\| \\ &\le \left\|\sum_{l\in[k]} 2^{-k+l} d^{(l)}\right\| \\ &\le 2\,\mathrm{herdisc}(A). \end{aligned}$$

□

The key idea now is based on the following simple observation: If we replace $\varepsilon^{(l)}$ by $-\varepsilon^{(l)}$, we get $-d^{(l)}$ instead of $d^{(l)}$. By choosing signs for the $\varepsilon^{(l)}$, $l \in [k]$ in a clever way and not using the triangle–inequality, we are going to improve the above result.

Note that if we change the sign of one $\varepsilon^{(l)}$, this leads to a different $a^{(l+1)}$ and thus may change all the subsequently determined variables. So we have to decide the signs "online". This might be described best in the language of games. Consider the following two–player game:

The Game: At the start of the game the vector $v \in \mathbb{R}^m$ is zero. One round of the game is: Player A gives a vector $w \in [-1, 1]^m$, Player B chooses a sign $\delta \in \{-1, 1\}$. v is then updated to $v := 0.5v + \delta w$. The game is played for a fixed number of k rounds. Player A aims to maximize $\|v\|$ while B wants to keep the norm down. What is the maximum number c that A can reach?

This is the game we are playing (as Player B against the algorithm as Player A) when deciding on the signs of the $\varepsilon^{(l)}$, $l \in \{0, \dots, k-1\}$. In the game we normed the w to be in $[-1, 1]^m$ (while above we have $d_i^{(l)} \in \{-\mathrm{herdisc}(A), \dots, \mathrm{herdisc}(A)\}$), but it is clear that this just changes c to $c\,\mathrm{herdisc}(A)$ as an upper bound. So we have the following general result:

Lemma 1. *If c is the maximum value Player A can reach in the above described game, then* $\operatorname{lindisc}(A) \le c \operatorname{herdisc}(A)$.

We complete the proof by determining this constant c. We show

Lemma 2. *The maximum value Player A can reach is* $c = 2\,(1 - 2^{-q})$.

Proof. We investigate the following strategy for Player B: Whatever vectors $w^{(1)}, \dots, w^{(k-q)}$ Player A chooses in the first $k-q$ rounds, pick $\delta^{(1)}, \dots, \delta^{(k-q)} := 1$ (any other choice would do, too). Set $w := \sum_{j=1}^{k-q} 2^{-k+j} w^{(j)}$. Choose the next sign $\delta^{(k-q+1)}$ in such a way, that the number of components $i \in [m] =: X_1$ such that $\operatorname{sgn}(w_i)$ and $\operatorname{sgn}(\delta^{(k-q+1)} w_i^{(k-q+1)})$ are different, is maximal. Set $X_2 := \{i \in [m] |\operatorname{sgn}(w_i) = \operatorname{sgn}(\delta^{(k-q+1)} w_i^{(k-q+1)})\}$. Next choose $\delta^{(k-q+2)} \in \{-1, 1\}$ such that the number of components $i \in X_2$ such that $\operatorname{sgn}(w_i)$ and $\operatorname{sgn}(\delta^{(k-q+2)} w_i^{(k-q+2)})$ are different, is maximal. Set $X_3 := \{i \in X_2 |\operatorname{sgn}(w_i) = \operatorname{sgn}(\delta^{(k-q+2)} w_i^{(k-q+2)})\}$. Continue in this fashion until $\delta^{(k)}$ and X_q are determined.

Note that $|X_{j+1}| \le \left\lfloor \frac{|X_j|}{2} \right\rfloor$ for all $j \in [q-1]$, which gives $|X_q| < 1$, i. e. $X_q = \emptyset$. So for every component i there is a $j \in \{k-q+1, \dots, k\}$ such that w_i and $\delta^{(q)} w_i^{(q)}$ have different signs. The worst–case is the one, where all $\delta^{(j)} w_i^{(j)}$, $j \in \{k-q+1, \dots, k\}$ are 1 (or -1) and w_i is zero. This gives us

$$\begin{aligned} \left\| \sum_{j=1}^{k} 2^{-k+j} \delta^{(j)} w^{(j)} \right\| &= \left\| \sum_{j=k-q+1}^{k} 2^{-k+j} \delta^{(j)} w^{(j)} + w \right\| \\ &\le \sum_{z=0}^{q-1} 2^{-z} \\ &= 2\left(1 - 2^{-q}\right). \end{aligned}$$

Player B can't do any better, as the following strategy for A reveals: Let r denote the biggest power of 2 that is less than or equal to m (so $r = 2^{q-1}$). Choose the first $k-q$ vectors as zero. The last q vectors choose like this: Components greater than r are always set zero (or what ever you like). For an index $i = 1 + \sum_{j=0}^{q-2} x_j 2^j \le r$, $x : \{0, \dots, q-2\} \to \{0, 1\}$ and a $p \in \{0, \dots, q-1\}$ set $w_i^{(k-p)} := 2x_p - 1$.

Here's an example for m=5:

$$\begin{aligned} w^{(j)} &= (0,0,0,0,0) \text{ for } i \in [k-q] \\ w^{(k-2)} &= (-1,-1,-1,-1,0) \\ w^{(k-1)} &= (-1,-1,+1,+1,0) \\ w^{(k)} &= (-1,+1,-1,+1,0). \end{aligned}$$

Whatever signs $\delta^{(j)}, j \in [k]$ you choose, there'll always be a component $i \in [r]$ such that $w_i^{(k-q+1)} = \ldots = w_i^{(k)}$ and thus

$$\sum_{j=1}^{k} 2^{-k+j} w_i^{(j)} = w_i^{(k)} \sum_{j=0}^{q-1} 2^{-j} = w_i^{(k)} 2 \left(1 - 2^{-q}\right).$$

This proves Lemma 2 and thus the theorem. □

Discussion

All of the above is constructive in the following sense: Let A and p be given as above. Assume that one has a coloring for every submatrix of discrepancy not greater than herdisc(A) (discrepancy is an NP–complete problem, so we can't skip this assumption). Then we have an algorithm rounding p to ε in time at most $\mathcal{O}(k \max\{n, m\})$ such that

$$\|A(p-\varepsilon)\| \leq 2^{-k}\|A\| + 2\left(1 - 2^{-q}\right) \operatorname{herdisc}(A),$$

where $\|A\| := \sup_{\|x\|=1} \|Ax\|$ and k is the maximum binary length we use to express the p_i in our algorithm.

There are some ideas which we didn't know how to use (or didn't want to). Remember that the $d_i^{(l)}$ are discrepancies of hyperedges of induced subgraphs under an optimal coloring. We assumed all $d_i^{(l)}$ to take the worst possible value in $[-\operatorname{herdisc}(A), \operatorname{herdisc}(A)]$, but in general an optimal coloring just creates a few badly colored hyperedges. Furthermore, in general not all induced subgraphs do have herdisc(A) as discrepancy.

This last point might be used also in a different ε–choosing–strategy: Choose the $\varepsilon^{(l)}$ in such a way that the resulting $a^{(l+1)}$ represents a hypergraph with small discrepancy.

Acknowledgments

I would like to thank J. Spencer for pointing out a mistake in Lemma 2 and some discussion on this topic. He is also the one who drew my

attention on this problem at the Discrepancy Workshop in Kiel this summer.

References

[1] J. Beck and V. Sós, *Discrepancy Theory* in R. Graham, M. Grötschel and L. Lovász, *Handbook of Combinatorics*, 1995, Chapter 26

[2] J. Beck and J. Spencer, *Integral approximation sequences*, Math. Programming **30** (1984), 88–98

[3] L. Lovász, J. Spencer and K. Vesztergombi, *Discrepancies of set–systems and matrices*, European J. Combin. **7** (1986), 151–160

[4] J. Spencer, *Ten Lectures on the Probabilistic Method*, 1987, SIAM, Philadelphia

Benjamin Doerr, Institut für Informatik und Praktische Mathematik, Mathematisches Seminar II, Christian–Albrechts–Universität zu Kiel, Hermann-Rodewald–Str. 3/1, 24098 Kiel, Germany

E-mail address: bed@numerik.uni-kiel.de