Prague Summer School: Discrepancy Theory Summer 2018

Lecture 3

Aleksandar Nikolov

1 The Beck-Fiala Theorem

So far we saw a general upper bound on the combinatorial discrepancy of set systems consisting of m sets over a universe of size n. While this bound is tight in general, already in the exercises we indicated that it can be improved when the set system has special structure. In this lecture we study "sparse" set systems, i.e. set systems in which no element appears in too many sets.

Definition 1. *Let* $(\mathcal{U}, \mathcal{S})$ *be a finite set system. The* degree *of* $(\mathcal{U}, \mathcal{S})$ *is*

$$d(\mathcal{S}) = \max_{j \in \mathcal{U}} |\{S \in \mathcal{S} : j \in S\}|.$$

Beck and Fiala [4] proved the following remarkable theorem, showing that the discrepancy of a set system $\mathcal{S}$ can be bounded in terms of its degree.

Theorem 2. *For any set system* $(\mathcal{S}, \mathcal{U})$*, we have* $\operatorname{disc}(\mathcal{S}) \le 2d(\mathcal{S}) - 1$.

Proof. For brevity, let $d = d(\mathcal{S})$. Without loss of generality, we will assume that $\mathcal{U} = [n]$. We provide an "iterative rounding" procedure to construct a colouring χ satisfying $\operatorname{disc}(\chi, \mathcal{S}) \le 2d - 1$. To be more precise, we construct a sequence $x(0), x(1), \ldots, x(T)$ so that for all $j \in [n]$:

1. $x_j(0) = 0$,
2. for all $t = 0, 1, \ldots, T$, $x_j(t) \in [-1, +1]$,
3. $x_j(T) \in \{-1, +1\}$.

This is similar to the Lovett-Meka proof of Spencer's theorem from the last lecture, except that at the last step we have $x(T) \in \{-1, 1\}^n$, which defines a coloring $\chi(j) = x_j(T)$. We will show that this coloring has discrepancy $\operatorname{disc}(\chi, \mathcal{S}) \le 2d - 1$.

At each time $t = 1, 2, \ldots, T$ we specify two sets: first, the *fixed* indices at t are

$$\mathcal{V}_t := \{j : x_j(t) \in \{-1, +1\}\},$$

and second, the *dangerous* sets at t are

$$\mathcal{D}_t = \{S \in \mathcal{S} : |S \setminus \mathcal{V}_t| > d\}.$$

Our goal at time $t + 1$ is to construct a $x(t + 1)$ so that

1. $\forall S \in \mathcal{D}_t$, $\sum_{j \in S} x_j(t + 1) = 0$;

2. $\forall j \in \mathcal{V}_t$, $x_j(t+1) = x_j(t)$;

3. $\mathcal{V}_t \subsetneq \mathcal{V}_{t+1}$.

First we prove the theorem, assuming that we can construct such a sequence. It is clear that $T \leq n$, since the set of fixed indices grows in cardinality by at least one in every time step. Now, choose an arbitrary $S \in \mathcal{S}$ and suppose that S becomes safe (i.e. not dangerous) at time t. Then by the triangle inequality

$$|\chi(S)| = \left|\sum_{j \in S} x_j(T)\right| \leq \left|\sum_{j \in S} x_j(t)\right| + \left|\sum_{j \in S} (x_j(T) - x_j(t))\right|.$$

Since S became safe at iteration t we have $\sum_{j \in S} x_j(t) = 0$. Moreover, for any $j \in \mathcal{V}_t$, $x_j(T) = x_j(t)$. Applying these observations, and using the definition of the dangerous sets $\mathcal{D}_t$ we have

$$|\chi(S)| \leq \left|\sum_{j \in S \setminus \mathcal{V}_t} (x_j(T) - x_j(t))\right| < 2|S \setminus \mathcal{V}_t| \leq 2d,$$

where the strict inequality follows since $(x_j(T) - x_j(t)) \in (-1, 1)$ for any j which was not fixed (and, thus, was not -1 or $+1$) at time t. But, the discrepancy must be an integer, so we must have $|\chi(S)| \leq 2d - 1$.

Next we move on to constructing the sequence $x(0), x(1), \ldots, x(T)$. Fix any time $t \in \{0, 1, 2, \ldots, T\}$, and consider the system of equations

$$\forall S \in \mathcal{D}_t : \sum_{j \in S \setminus \mathcal{V}_t} \Delta x_j = 0 \tag{1}$$

where there is one variable Δx_j for each active $j \notin \mathcal{V}_t$. Note that

$$|\mathcal{D}_t| \cdot d < |\{(j, S) : j \in S \setminus \mathcal{V}_t, S \in \mathcal{D}_t\}| \leq (n - |\mathcal{V}_t|)d,$$

where the first inequality follows since each set $S \in \mathcal{D}_t$ is dangerous and so $|S \setminus \mathcal{V}_t| > d$, and the second inequality follows since each element j is contained in at most d sets. This means that $|\mathcal{D}_t| < n - |\mathcal{V}_t|$, and thus the above system of equations has more variables than constraints, and, therefore, a nonzero solution Δx. Then we can define

$$x_j(t+1) = \begin{cases} x_j(t) & \text{if } j \in \mathcal{V}_t \\ x_j(t) + \gamma \Delta x_j & \text{otherwise} \end{cases}$$

where γ is the largest real so that $x(t) \in [-1, +1]^n$. There must be at least one j that becomes $+1$ or -1 after this update, and it follows that $\mathcal{V}_t \subsetneq \mathcal{V}_{t+1}$. The other two requirements follow directly from the construction. $\square$

Exercise 1. *Modify the proof of the Beck-Fiala theorem to show the improved bound* $\operatorname{disc}(\mathcal{S}) \leq \max\{2d(\mathcal{S}) - 3, d\}$.

Exercise 2. *Let A be an $m \times n$ matrix such that for every $j \in [n]$ we have*

$$\sum_{i:a_{ij}>0} a_{ij} \le 1,$$
$$-\sum_{i:a_{ij}<0} a_{ij} \le 1.$$

Prove that for any $x(0) \in [0,1]^n$ there exists a $x \in \{0,1\}^n$ so that

$$\forall i \in [m] : \sum_{j=1}^{n} a_{ij}(x_j - x_j(0)) < 1.$$

Hint*: For any set $S \subseteq [n]$, any i, and any $x \in \{0,1\}^n$,*

$$\sum_{j \in S} a_{ij}(x_j - x_j(0)) \le \sum_{j:a_{ij}>0} a_{ij}(1 - x_j(0)) - \sum_{j:a_{ij}<0} a_{ij}(x_j(0)).$$

2 Towards the Beck-Fiala Conjecture

In the same paper, Beck and Fiala made the following conjecture, which remains open today:

Open Problem: Prove or disprove the upper bound $\mathrm{disc}(\mathcal{S}) \le O(\sqrt{d(\mathcal{S})})$.

This is perhaps the most prominent open problem in combinatorial discrepancy theory. The best known bound is $2d - \log^*(d)$ [5], where $\log^*$ is the iterated logarithm function (the inverse of the tower of 2's function). On the other hand, we can prove bounds which come close to the conjecture, at the cost of a mild dependendce on m, the number of sets in the set system. The best one is due to Banaszczyk [1]:

Theorem 3. *For any set system $(\mathcal{S},\mathcal{U})$,* $\mathrm{disc}(\mathcal{S}) = O(\sqrt{d \log |\mathcal{S}|})$.

Banaszczyk's proof uses deep facts from convex geometry and proves a more general statement, which we will discuss in the next lecture. In the remainder of this lecture we will sketch an algorithmic proof of the theorem which is a refinement of the Beck and Fiala proof, and is due to Bansal, Dadush, and Garg [2]. The main idea is to randomize the steps in the Beck-Fiala argument: rather than walk in an arbitrary direction, we pick the direction at random. Then the hope is that once a set becomes safe, its discrepancy behaves like that of a random coloring on a set of size at most d. I.e. we hope that the discrepancy of a safe set behaves like a Gaussian with variance bounded by $O(d)$. If that's the case, then a standard tail bound argument would give the $O(\sqrt{d \log |\mathcal{S}|})$ discrepancy bound. Unfortunately, because we are restricted to the subspace on which the system of equations (1) are satisfied, the values $x_j(t)$ for different j may be correlated so that they don't behave like a random coloring. The key insight in the proof is to that it is possible to make the random choices so that this correlation has minimal effect.

The following main lemma shows that we can "randomize" Beck and Fiala's proof so that the discrepancy of every set behaves as if we chose a random coloring for a set of size $O(d)$.

Lemma 4. *There exists a constant C such that for any set system $\mathcal{S}$ with degree $d = d(\mathcal{S})$, there exists a sequence of random vectors $x(0) = 0, x(1), \ldots, x(T)$ so that $x(T) \in \{-1, 1\}^n$, and for any set $S \in \mathcal{S}$ and any $0 \leq \eta \leq \frac{1}{2C}$, we have*

$$\max\{\mathbb{E}e^{\eta \sum_{j\in S} x_j(T)}, \mathbb{E}e^{-\eta \sum_{j\in S} x_j(T)}\} \leq e^{C\eta^2 d}. \tag{2}$$

Let us first finish the proof of Theorem 3 assuming Lemma 4.

Proof of Theorem 3. The proof follows along the lines of a standard Chernoff bound. Observe first that if $\sqrt{d \log |\mathcal{S}|} = \Omega(d)$, then the theorem follows from the Beck-Fiala theorem. Assume otherwise. By Lemma 4 and Markov's inequality, we have that for any $t, \eta \geq 0$,

$$\Pr\left(\sum_{j\in S} x_j(T) > t\right) = \Pr\left(e^{\eta \sum_{j\in S} x_j(T)} > e^{\eta t}\right) < \frac{\mathbb{E}e^{\eta \sum_{j\in S} x_j(T)}}{e^{\eta t}} \leq e^{C\eta^2 d - \eta t}.$$

The choice of η that minimizes the right hand side is $\eta = \frac{t}{2Cd}$, which is at most $\frac{1}{2C}$ as long as $t \leq d$. Then, we get $e^{-t^2/4Cd}$ on the right. The same argument works for $-\sum_{j\in S} x_j(T)$, and we have that, for any $t \leq d$,

$$\Pr\left(\big|\sum_{j\in S} x_j(T)\big| > t\right) < 2e^{-t^2/4Cd}.$$

Setting $t = \sqrt{4Cd \log 2|\mathcal{S}|}$ and taking a union bound shows that the discrepancy of the coloring χ defined by $\chi(j) = x_j(T)$ is $O(\sqrt{d \log |\mathcal{S}|})$ with positive probability. ☐

The proof of Lemma 4 relies on two technical lemmas which we will not prove in this lecture. The first lemma is a a form of a martingale concentration inequality, which is essentially due to Freedman [6] (see also [3]).

Lemma 5. *Let $y_0, y_1, \ldots, y_T$ and $\Delta y_1, \ldots, \Delta y_T$ be random variables such that*

1. *y_0 is deterministic;*
2. *$y_t - y_{t-1} \leq \Delta y_t$ for all t with probability 1;*
3. *$\Delta y_t \leq 1$ for all t with probability 1;*
4. *$\mathbb{E}[\Delta y_t + \Delta y_t^2 \mid y_0, \ldots, y_{t-1}] \leq 0$ for all t with probability 1.*

Then $\mathbb{E}e^{y_T} \leq e^{y_0}$.

The second lemma shows that in any subspace of sufficiently large dimension, we can find a non-trivial distribution on random vectors with essentially uncorrelated coordinates.

Lemma 6. *Let W be a linear subspace of $\mathbb{R}^n$ of dimension at least $\frac{n}{2}$. Then there exists a probability distribution D_W supported on W such that $z \sim D_W$ satisfies*

$$\forall j \in [n] : \mathbb{E} z_j = 0$$
$$\mathbb{E} \sum_{j=1}^{n} z_j^2 = n,$$
$$\forall u \in \mathbb{R}^n : \mathbb{E} \left(\sum_{j=1}^{n} u_j z_j \right)^2 \leq 2\mathbb{E} \sum_{j=1}^{n} u_j^2 z_j^2,$$

and $|z_j| \leq n$ with probability 1.

In fact the distribution has the form $\sum_{i=1}^{k} \sigma_i w_i$, where $\sigma_1, \ldots, \sigma_k \in \{-1, 1\}$ are chosen independently and uniformly, and $w_1, \ldots, w_k \in \mathbb{R}^n$ are efficiently computable vectors. The lemma reduces to a statement about the existence of positive semidefinite matrices with particular properties.

Proof of Lemma 4. The proof proceeds along the lines of the proof of Theorem 2 but using the distribution in Lemma 6 to randomize each step, and using Lemma 5 to analyze the random process.

We define the set of fixed variables $\mathcal{V}_t$ at time t to contain, as before, all j for which $x_j(t) < -1 + \delta$ or $x_j(t) > 1 - \delta$. Here δ is very small, for instance $\delta = 1/n$. At the end of the algorithm we will round all coordinates in $\mathcal{V}_t$ to the nearest integer, which does not change the estimate significantly.

We slightly modify the definition of dangerous sets to $\mathcal{D}_t = \{S \in \mathcal{S} : |S \setminus \mathcal{V}_t| > 2d\}$. Now an argument analogous to the one before shows that, at any time step t, the subspace $W \subseteq \mathbb{R}^{[n] \setminus \mathcal{V}_t}$ of solutions to the system of equations (1) has dimension at least $(n - \mathcal{V}_t)/2$. Let D_W be the distribution supported on W from Lemma 6, and let $\Delta x(t)$ be sampled from D_W. Then we define

$$x_j(t+1) = \begin{cases} x_j(t) & \text{if } j \in \mathcal{V}_t \\ x_j(t) + \gamma \Delta x_j(t) & \text{otherwise} \end{cases},$$

where γ is a tiny constant, to be chosen later. The final time step T is the one in which $\mathcal{V}_T = [n]$

Exercise 3. *Using a variant of the analysis in the proof of Spencer's theorem from the previous lecture, show that $\mathbb{E}[T] = O\left(\frac{\log n}{\gamma^2}\right)$.*

We analyze this construction using Lemma 5. Let $S \in \mathcal{S}$ be an arbitrary set in the set system, and let t be the last time when S is dangerous, i.e. the largest t such that $S \in \mathcal{D}_t$. Define $y_0 = C\eta^2 d$ for a constant C to be chosen later, and for any $s \in \{1, \ldots, T - t\}$ define

$$y_s = \eta \sum_{j \in S} x_j(t+s) + C\eta^2 \sum_{j \in S} (1 - x_j(t+s)^2).$$

Intuitively, the first term is the (positive) discrepancy of S, and the second term is how much progress the elements of S have made towards becoming fixed. In particular, at time T the second term vanishes, and we have $y_{T-t} = \eta \sum_{j \in S} x_j(t+s)$.

Let us define, for any $s \in \{0, \ldots, T-t-1\}$,

$$\Delta y_{s+1} = \eta\gamma \sum_{j \in S} \Delta x_j(t+s) - C\eta^2 \sum_{j \in S} (2\gamma x_j(t+s)\Delta x_j(t+s) + \gamma^2 \Delta x_j(t+s)^2).$$

It is easy to see that conditions 1 and 2 of Lemma 5 are satisfied, and condition 3 is satisfied if γ is chosen to be small enough. The main challenge is to verify the fourth condition of the lemma. In the rest of the proof, all expectations will be conditional on $y_1, \ldots, y_s$. For convenience we will denote $\Delta y = \Delta y_{s+1}$, and also $\Delta x = \Delta x(t+s)$ and $x = x(t+s)$. We calculate,

$$\mathbb{E}\Delta y = -C\eta^2\gamma^2 \mathbb{E} \sum_{j \in S} \Delta x_j^2, \tag{3}$$

and

$$\begin{aligned}
\mathbb{E}\Delta y^2 &= 2\eta^2\gamma^2 \mathbb{E}\left(\sum_{j \in S} \Delta x_j\right)^2 + 8C^2\eta^4\gamma^2 \mathbb{E}\left(\sum_{j \in S} x_j \Delta x_j\right)^2 + O(\gamma^3) \\
&\le 4\eta^2\gamma^2 \mathbb{E} \sum_{j \in S} \Delta x_j^2 + 16C^2\eta^4\gamma^2 \mathbb{E} \sum_{j \in S} x_j^2 \Delta x_j^2 + O(\gamma^3) \\
&\le (4 + 16C^2\eta^2)\eta^2\gamma^2 \mathbb{E} \sum_{j \in S} \Delta x_j^2 + O(\gamma^3),
\end{aligned}$$

where the first inequality follows from Cauchy-Schwarz, the second from Lemma 6, and the last inequality because $x_j^2 \le 1$. Here the notation $O(\gamma^3)$ hides all terms multiplied by a power of γ greater than 2. Because the algorithm makes at most $O(\gamma^{-2}\log n)$ steps (with constant, and also with high probability), all such terms can be made negligible by choosing γ small enough. To finish verifying the fourth condition of Lemma 5, we just need to set C such that $4 + 16C^2\eta^2 < C$, which holds whenever $\eta \le \frac{1}{2C}$ and $C > 8$. Then, Lemma 5 implies that $\mathbb{E}e^{y_{T-t}} \le e^{y_0}$, which, by the definition of y_s is equivalent to

$$\mathbb{E}e^{\eta \sum_{j \in S} x_j(T)} \le e^{C\eta^2 d}.$$

The bound for $\mathbb{E}e^{-\eta \sum_{j \in S} x_j(T)}$ is proved analogously. □

Exercise 4. *Let A be an $m \times n$ matrix each of whose columns has Euclidean norm at most* 1*. Modify the proof above to show that* $\mathrm{disc}(A) = O(\sqrt{\log m})$*.*

A famous conjecture by Komlós posits that the bound in the exercise above can be improved to $\mathrm{disc}(A) = O(1)$. This conjecture implies the Beck-Fiala conjecture.

References

[1] Wojciech Banaszczyk. Balancing vectors and Gaussian measures of n-dimensional convex bodies. *Random Structures Algorithms*, 12(4):351–360, 1998.

[2] Nikhil Bansal, Daniel Dadush, and Shashwat Garg. An algorithm for Komlós conjecture matching Banaszczyk's bound. In *57th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2016*, pages 788–799. IEEE Computer Soc., Los Alamitos, CA, 2016.

[3] Nikhil Bansal, Daniel Dadush, Shashwat Garg, and Shachar Lovett. The gram-schmidt walk: a cure for the banaszczyk blues. In Ilias Diakonikolas, David Kempe, and Monika Henzinger, editors, *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 587–597. ACM, 2018.

[4] József Beck and Tibor Fiala. “Integer-making” theorems. *Discrete Appl. Math.*, 3(1):1–8, 1981.

[5] Boris Bukh. An improvement of the Beck-Fiala theorem. *Combin. Probab. Comput.*, 25(3):380–398, 2016.

[6] David A. Freedman. On tail probabilities for martingales. *Ann. Probability*, 3:100–118, 1975.