CS 761 : Randomized Algorithms, Spring 2023, Waterloo

Lecture 9 : Discrepancy Minimization

We study the random walk algorithm by Lovett and Meka for discrepancy minimization whose analysis uses martingales.

Spencer's Theorem

Discrepancy minimization is a classical topic in discrete mathematics with connections to different areas.

A standard setting is when we are given a set system $S_1, S_2, \ldots, S_m$ of the ground set $[n]$, and the goal is to find a coloring $\chi : [n] \to \{1,-1\}$ that minimizes the maximum discrepancy $\max_i |\sum_{j \in S_i} \chi(j)|$.

It is not difficult to prove that a random coloring has discrepancy $O(\sqrt{n \log m})$ with high probability.

A celebrated result by Spencer beats the random coloring bound.

Theorem (Spencer) For any set system with n elements and m sets, there is a coloring with discrepancy at most $k \cdot \sqrt{n \cdot \log_2(m/n)}$ where k is a universal constant.

In the special case when $m=n$, Spencer bounded the discrepancy by $6\sqrt{n}$, and this is known as "the six standard derivation suffice" paper.

The proof of Spencer is non-constructive, based on some entropy and pigeonhole principle argument (see e.g. Alon-Spencer).

It was even conjectured that such an efficient algorithm does not exist.

In a breakthrough work in 2010, Bansal gave a randomized polynomial time algorithm that finds a coloring with discrepancy $O(\sqrt{n} \cdot \log_2(m/n))$, thus matching Spencer's bound when $m = O(n)$.

Bansal's algorithm is based on semidefinite programming and also on Spencer's original proof.

In a subsequent work, Lovett and Meka gave a randomized polynomial time algorithm that finds a coloring with discrepancy $O(\sqrt{n \cdot \log_2(m/n)})$, thus matching Spencer's bound for all m and n.

Lovett-Meka's algorithm is only based on linear programming and the proof is self-contained.

Both the algorithm and the analysis are elegant and clean. This is what we study in this lecture.

Geometric Approach

Spencer's result uses the partial coloring method (similar to the work of Beck on local lemma), which finds a partial coloring $\chi : [n] \to \{-1, 0, 1\}$ such that the maximum discrepancy is at most $O(\sqrt{n \cdot \log(m/n)})$ and there are at least a constant fraction of elements that are set to $\{-1, 1\}$.

Given such a partial coloring, one can then recurse on the remaining variables (those that are assigned 0) and there are at least a constant fraction of elements that are set to $\{-1, 1\}$.

Given such a partial coloring, one can then recurse on the remaining variables (those that are assigned 0) until every variable is assigned -1 or 1, and the total discrepancy is still at most $O(\sqrt{n \cdot \log(m/n)})$ as the sum is a decreasing geometric sequence. We will see a similar argument below.

Lovett and Meka also use this partial coloring method, but the approach to find a partial coloring is more geometrical than combinatorial.

The main statement is about finding a point in a polytope with many "close-to $\{\pm 1\}$ coordinates".

<u>Theorem</u> (Lovett-Meka) Let $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$ be vectors, and $x_0 \in [-1,1]^n$ be a starting point. Let $c_1, \ldots, c_m \geq 0$ be thresholds such that $\sum_{j=1}^{m} e^{-c_j^2/16} \leq n/16$. Let δ be a small approximation parameter. Then there is a randomized polynomial time algorithm, with constant probability, finds a point $x \in [-1,1]^n$ s.t.

(i) $|\langle x - x_0, v_j \rangle| \leq c_j \cdot \|v_j\|_2$, and (ii) $|x_i| \geq 1-\delta$ for at least $n/2$ indices $i \in [n]$.

The main difference with Spencer's result is that $x \in [-1,1]^n$ rather that $x \in \{-1, 0, 1\}^n$.

Also, this result gives us additional flexibility by controlling how much a constraint is violated by adjusting the thresholds $c_1, \ldots, c_m$, which is a useful feature in subsequent works in designing approximation algorithms.

To apply to discrepancy minimization, we set $v_i = \chi_{S_i}$ where $\chi_{S_i} \in \{0,1\}^n$ is the indicator vector of the set S_i so that $\|v_i\|_2 = \sqrt{|S_i|}$, set $c_i = \Delta_{S_i} / \sqrt{|S_i|}$ for some parameter Δ_{S_i} $\forall i$, set $x_0 = \vec{0}$ and $\delta = \frac{1}{\text{poly}(n)}$.

<u>Corollary</u> For any set system with n elements $[n]$ and m sets $S_1, S_2, \ldots S_m$, if $\Delta_{S_1}, \ldots, \Delta_{S_m}$ satisfies $\sum_{j=1}^{m} e^{-\Delta_{S_j}^2/16|S_j|} \leq n/16$, then there exists $x \in [-1,+1]^n$ with $|\{i \mid |x_i| = 1\}| \geq n/2$ and $|\sum_{i \in S_j} x_i| \leq \Delta_S + 1/\text{poly}(n)$ for every set S_j.

Moreover, there is a randomized polynomial time algorithm to find such a coloring x.

In particular, if we set $\Delta_{S_j} = 8\sqrt{n \cdot \log(m/n)}$, then there is a partial coloring with discrepency at most $8\sqrt{n \cdot \log(m/n)}$ for each set S_j.

<u>From partial coloring to full coloring</u>

Then we can apply the theorem recursively with the partial coloring x as the starting point x_1 in the next iteration, on the remaining variables and the set system restricted on them.

Since there are at most $n/2$ remaining variables, we can set $\Delta_{S_j} = 8\sqrt{(n/2) \cdot \log(m/(n/2))}$.

Then we take the partial coloring of the current iteration as the starting point of the next iteration, and repeat until we get a full coloring x_T.

Then we take the partial coloring of the current iteration as the starting point of the next iteration, and repeat until we get a full coloring x_T.

The total discrepancy of a set S_j in the full coloring is

$$|\langle x_T, \chi_{S_j}\rangle| \leq \sum_{t=1}^{T} |\langle x_t - x_{t-1}, \chi_{S_j}\rangle| \leq \sum_{t=1}^{T} \Delta_{S_j}^{(t)} = 8\sqrt{n}\sum_{t=1}^{T}\sqrt{\frac{\log(m\cdot 2^t/n)}{2^t}} = O\left(\sqrt{n\cdot\log(m/n)}\right).$$

Henceforth, we focus on finding a partial coloring guaranteed by the Lovett-Meka theorem.

Partial Coloring by Random Walks in the Polytope

The algorithm to find a partial coloring is surprising simple: just do a random walk in the polytope!

To ensure that the random walk stays in the polytope, the walk is restricted to the subspace orthogonal to the tight variable constraints (i.e. $|x_i| \geq 1-\delta$) and the tight discrepancy constraints (i.e. $\langle x_t, v_j\rangle \geq c_j - \delta$).

The intuition is that the condition $\sum_i e^{-c_i^2/16} \leq n/16$ implies that the discrepancy constraints are farther from the origin on average than the variable constraints, and so a random walk will hit more variable constraints than the discrepancy constraints, thus obtaining a partial coloring.

Algorithm

Let $\varepsilon > 0$ be a small step size so that $\delta = O(\varepsilon\sqrt{\log(nm/\varepsilon)})$.

Let $T = K_1/\varepsilon_2$ where $K_1 = 16/3$.

The algorithm will produce $X_0 = x_0, X_1, \ldots, X_T \in \mathbb{R}^n$ using the following random process.

We assume without loss of generality that $\|v_j\|_2 = 1$ for $1 \leq j \leq m$.

For $t = 1$ to T do

Let $C_t^{var} := \{ i \in [n] : |(X_{t-1})_i| \geq 1-\delta \}$ be the set of (nearly) tight variable constraints.

Let $C_t^{disc} := \{ j \in [m] : |\langle X_{t-1} - x_0, v_j\rangle| \geq c_j - \delta \}$ be the set of (nearly) tight discrepancy constraints.

Let $V_t := \{ u \in \mathbb{R}^n : u_i = 0 \ \forall i \in C_t^{var} \text{ and } \langle u, v_j\rangle = 0 \ \forall j \in C_t^{disc} \}$ be the linear subspace orthogonal to the (nearly) tight constraints.

Set $X_t := X_{t+1} + \varepsilon U_t$ where $U_t \sim N(V_t)$ is a random direction in V_t generated as follows:

Construct an orthonormal basis $\{w_1, w_2, \ldots, w_d\}$ of V_t and sample independent Gaussian random variables $g_1, g_2, \ldots, g_d \in N(0,1)$ and set $U_t := g_1 w_1 + g_2 w_2 + \ldots + g_d w_d$.

Preliminaries

To analyze the random walk algorithm, we will use some basic properties of Gaussian random variables, and a martingale concentration inequality for Gaussian random variables.

Claim 1: Let $U_t \sim N(V_t)$. For any $y \in \mathbb{R}^n$, $\langle U_t, y\rangle \sim N(0, \sigma^2)$ where $\sigma^2 \le \|y\|_2^2$.

Also, $\mathbb{E}[\|U_t\|_2^2] = \dim(V_t)$.

The proof uses the orthonormal property of the basis $\{v_1, \dots, v_d\}$ and is left as an exercise (or see Lovett-Meka).

The following is a standard tail bound for Gaussian variables that we've seen in dimension reduction in L05.

Claim 2 Let $g \sim N(0,1)$. Then $\Pr[|g| \ge \lambda] \le 2e^{-\lambda^2/2}$ for any $\lambda > 0$.

The following martingale concentration inequality for Gaussian random variables will be used for discrepancy constraints.

Lemma Let $X_1, \dots, X_T$ be random variables and $Y_1, \dots, Y_T$ be random variables where each Y_i is a function of X_i.

Suppose that for all $1 \le i \le T$, any $x_1, \dots, x_{i-1} \in \mathbb{R}$, the random variable $Y_i \mid (X_1 = x_1, \dots, X_{i-1} = x_{i-1})$ is a Gaussian random variable with mean zero and variance at most one.

Then, $\Pr[|Y_1 + \dots + Y_T| \ge \lambda\sqrt{T}] \le 2e^{-\lambda^2/2}$ for any $\lambda > 0$.

Analysis

To prove Lovett-Meka's theorem about partial coloring, we need to prove that

1. $X_0, X_1, \dots, X_T$ are in the polytope. i.e. they all satisfy all the constraints.
2. $|(X_T)_i| \ge 1-\delta$ for at least $n/2$ indices $i \in [n]$.

Claim A For $\varepsilon \le \delta / \sqrt{c \cdot \log(mn/\varepsilon)}$ for a constant $c \ge 4$, then $X_0, \dots, X_T$ are all in the polytope with probability at least $1 - 1/(mn)^{\frac{c}{2}-1}$.

Proof This is almost by construction, as when the constraints are nearly tight, then the walk would be orthogonal to them.

The only possibility that a constraint is violated is when $|(X_t)_i| < 1-\delta$ but $|(X_t)_i| > 1$ for the variable constraints, or when $\langle X_t, v_j\rangle < c_j - \delta$ but $\langle X_{t+1}, v_j\rangle > c_j$.

In either case, it is because the Gaussian step is too large in some direction.

More precisely, there exists some unit vector w (recall that we assume wlog that the v_j's are unit vectors) such that $\langle X_{t+1} - X_t, w\rangle > \delta \iff \langle U_t, w\rangle > \delta/\varepsilon$.

By claim 1, $\langle U_t, w\rangle$ is a Gaussian random variable with mean 0 and variance at most 1 as $\|w\|_2 = 1$.

Thus, by claim 2, $\Pr[|\langle U_t, w\rangle| > \delta/\varepsilon] \le 2e^{-(\delta/\varepsilon)^2/2} = O((\varepsilon/mn)^{c/2})$ by the choice $\delta/\varepsilon = \sqrt{c \cdot \log(mn/\varepsilon)}$.

Taking a union bound over all $T = O(1/\varepsilon^2)$ time steps and over all $n+m$ constraints,

$\Pr[\text{some } X_t \text{ is not in the polytope}] \lesssim T \cdot (n+m) \cdot (\varepsilon/mn)^{c/2} \le 1/(mn)^{c/2-1}$ when $c \ge 4$. □

The interesting part is to prove ②.

First, we use the martingale concentration inequality to prove that only a small fraction of discrepancy constraints are nearly tight.

Clearly, by construction, $C_t^{disc} \subseteq C_{t+1}^{disc}$ for any t, so we just need to bound the size of $|C_T^{disc}|$.

Claim B $\mathbb{E}[|C_T^{disc}|] \leq n/4$.

Proof We divide the discrepancy constraints into two groups.

Let $J := \{ j \mid c_j \leq 10\delta \}$ be the set of "close" constraints.

We argue that there are not too many of them, because

$$n/16 \geq \sum_{j \in J} e^{-c_j^2/16} \geq |J| \cdot e^{-100\delta^2/16} \geq |J| \cdot e^{-1/16} > 9|J|/10 \quad \text{where we assume } \delta < 0.1.$$

For any "far" constraint $j \notin J$, if $j \in C_T^{disc}$, then $|\langle X_T - x_0, v_j \rangle| \geq c_j - \delta \geq 0.9 c_j$.

We will bound the probability that such $j \in C_T^{disc}$ by the martingale concentration inequality.

Recall that $X_T = x_0 + \varepsilon(U_1 + \ldots + U_T)$ and define $Y_i = \langle U_i, v_j \rangle$.

Then, for $j \notin J$, $\Pr[j \in C_T^{disc}] = \Pr[\langle \varepsilon(U_1 + \ldots + U_T), v_j \rangle \geq c_j - \delta] \leq \Pr[|Y_1 + \ldots + Y_T| \geq 0.9 c_j / \varepsilon]$.

Check that the conditions of the martingale concentration inequality are satisfied with $U_1, \ldots, U_T$ and $Y_1, \ldots, Y_T$,

so it follows that $\Pr[j \in C_T^{disc}] \leq 2\exp\left(-(0.9c_j)^2 / 2\varepsilon^2 T\right) < 2e^{-c_j^2/16}$ if $T \leq 16/3\varepsilon^2$.

Therefore, $\mathbb{E}[|C_T^{disc}|] \leq |J| + \sum_{j \notin J} \Pr[j \in C_T^{disc}] \leq 5n/72 + \sum_{j \notin J} 2e^{-c_j^2/16} \leq 5n/72 + 2n/16 < n/4$,

where the second last inequality is by the assumption $\sum_j e^{-c_j^2/16} \leq n/16$. □

Given that the number of tight discrepancy constraints is at most $n/4$, if the number of tight variable constraints is also small, then the dimension of the subspace V_t is large and so we expect $\|U_t\|_2^2$ to be large, and thus $\|X_t\|_2^2$ will increase significantly and this would lead to many tight variables.

Claim B $\mathbb{E}[\|X_T\|_2^2] \leq n$.

Proof This is a simple claim that we omit the proof (see Lovett-Meka), basically it follows from

$\mathbb{E}[(X_T)_i^2] \leq 1 - \delta + \varepsilon \mathbb{E}[|(U_T)_i|^2] \leq 1$ as $(U_T)_i$ is a Gaussian with variance ≤ 1 by Claim 1. □

The interesting claim is to prove that there are many tight variable constraints using the above plan.

Claim C $\mathbb{E}[|C_T^{var}|] \geq 0.65n$.

Proof By Claim 1, $\mathbb{E}[U_t \mid X_{t-1}] = 0$ and $\mathbb{E}[\|U_t\|_2^2 \mid X_{t-1}] = \dim(V_t)$, hence

$$\mathbb{E}[\|X_t\|_2^2] = \mathbb{E}[\|X_{t-1} + \varepsilon U_t\|_2^2] = \mathbb{E}[\|X_{t-1}\|_2^2] + \varepsilon^2 \mathbb{E}[\|U_t\|_2^2] = \mathbb{E}[\|X_{t-1}\|_2^2] + \varepsilon^2 \cdot \mathbb{E}[\dim(V_t)].$$

So, by Claim B, assuming $T = K_1/\varepsilon^2$ with $K_1 = 16/3$, it follows that

$$n \geq \mathbb{E}[\|X_T\|_2^2] \geq \varepsilon^2 \cdot \sum_{t=1}^{T} \mathbb{E}[\dim(V_t)] \geq \varepsilon^2 \cdot T \cdot \mathbb{E}[\dim(V_T)] \geq K_1 \cdot \mathbb{E}[\dim(V_T)] = K_1 \cdot \mathbb{E}[n - |C_T^{var}| - |C_T^{disc}|].$$

Therefore, by Claim A, $\mathbb{E}[|C_T^{var}|] \geq n(1-\frac{1}{10}) - \mathbb{E}[|C_T^{disc}|] \geq n(1-\frac{1}{K_1}-\frac{1}{4}) = 0.56n$. □

We are ready to finish the proof of Lovett-Meka theorem.

By Claim C and the upper bound that $|C_T^{var}|$, a Markov-type argument shows that $\Pr[|C_T^{var}| \geq \frac{n}{2}] \geq 0.12$.

Combining with Claim A, both ① and ② are satisfied with probability $\geq 0.12 - \frac{1}{\text{poly}(mn)} \geq 0.1$.

This completes the proof of the partial coloring theorem.

Apply it recursively as explained before. this provides a self-contained and constructive proof of Spencer's theorem.

References

- Bansal. Constructive algorithms for discrepancy minimization, 2010.
- Lovett. Meka. Constructive discrepancy minimization by walking on the edges, 2012.

These two works have sparked a lot of interests in discrepancy minimization, and lead to more constructive results for more general non-constructive theorems, including a recent one in matrix discrepancy.

A major open problem remained unresolved is the Beck-Fiala Conjecture, which states that there is a coloring with discrepancy $O(\sqrt{d})$ if every element is contained in at most d sets.

The best known result is a coloring with discrepancy $O(\sqrt{d \cdot \log n})$.