Rounding Techniques in Approximation Algorithms

Lecture 12: Iterative Relaxation for Discrepancy

Lecturer: Nathan Klein

1 Discrepancy

In this lecture, we will discuss a famous theorem of Beck and Fiala from the 80s [BF81] about the discrepancy of bounded degree hypergraphs. We will interpret their proof as an iterative algorithm that uses facts about the structure of extreme point solutions to an LP.

Given a hypergraph $H = (V, E)$, the **discrepancy** of H, $\text{disc}(H)$, is the minimum value of $\max_{e \in E} |x(e)|$ over all colorings $x : V \to \{-1, 1\}$, where $x(e) = \sum_{v \in e} x_v$. Equivalently, given a matrix $A \in \{0,1\}^{m \times n}$ we want to find $x \in \{-1, 1\}^n$ that minimizes $\|Ax\|_\infty$.

We will prove the following theorem:

Theorem 1.1. *Let H be a hypergraph in which the degree of every vertex is bounded by t. Then, $disc(H) \leq 2t$.*

In the homework, you are asked to improve this to either $2t-1$ for all t or to $2t-3$ when $t \geq 3$. The following example shows that the latter bound is tight for $t = 3$.

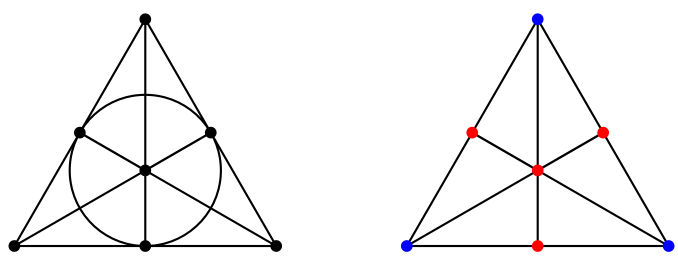

Figure 1: On the left is the Fano plane. The sets consist of all the lines, all joining three points. $t = 3$ and the discrepancy is, perhaps surprisingly, 3. On the right is a modified instance with discrepancy 1 and a coloring demonstrating this.

The proof of this theorem actually derives from a polyhedral view of discrepancy, which at first glance seems pretty useless:

$$P_{\text{disc}} = \begin{cases} x(e) = 0 & \forall e \in E \\ -1 \leq x_v \leq 1 & \forall v \in V \\ x_v = -1 & \forall v \in R \\ x_v = 1 & \forall v \in B \end{cases}$$

Why useless? Well, because we can just set $x_v = 0$ for all vertices, at least supposing R and B are empty (you'll see why they are useful later). So there's no point in even solving it – or so it seems at first glance! It turns out that the vertices of this polytope are *not* just all 0 (at least, under certain conditions).

To prove this, we begin with a very innocent-sounding fact, where F is the set of "fractional" variables between -1 and 1, i.e. $F = V \setminus (R \cup B)$.

Fact 1.2. *Suppose every vertex is in at most t edges. Furthermore, suppose every edge contains at least $t+1$ vertices in F. Then, there are fewer edges than vertices in F.*

Proof. There are at most $|F|t$ vertex-edge incidences from vertices in F, but also at least $|E|(t+1)$. The lemma follows. □

Corollary 1.3. *In any vertex of P_{disc} in which every edge contains at least $t+1$ vertices, there is a variable in F set to -1 or 1.*

Proof. After ignoring R and B, which are irrelevant, there are $|F|$ variables, and fewer than $|F|$ non-trivial constraints by the above fact. By the rank lemma there must be a variable set to -1 or 1. □

These two simple facts lead us to the proof of our main theorem.

Theorem 1.1. *Let H be a hypergraph in which the degree of every vertex is bounded by t. Then, $disc(H) \leq 2t$.*

Proof. Until all elements are in R or B, do the following. Delete all edges with at most t vertices in F. Let x be a vertex of P_{disc} with the current set of edges. By Corollary 1.3, there must be a vertex $v \in F$ with $x_v \in \{-1, 1\}$. If $x_v = -1$, add it to R and otherwise to B.

This algorithm must terminate, since a new color is set every iteration, and clearly the LP remains feasible since we are only deleting constraints and fixing variables to values they had in a feasible point. So it remains to prove the guarantee on the discrepancy.

Before an edge is deleted, it has $x(e) = 0$. At the moment of deletion, it has at most t fractional coordinates. Therefore, $x(e)$ can change by at most $2t$, since each variable can change by at most 2. This proves the theorem. □

2 Iterative Relaxation

This falls into a general framework known as iterative relaxation.

Iterative Relaxation

Consider any linear program and an extreme point solution x. Fix all integer coordinates of x, delete one of the constraints (in some problem-specific manner), and re-solve. Iterate until all coordinates are integral.

Fact 2.1. *In every iteration, the cost of x can only decrease. So, the cost of the resulting integer solution is no more than integer OPT. Of course, it may not obey all the same constraints as integer OPT since we deleted some.*

In the next lecture, we will use this to prove the following theorem of Singh and Lau [SL15].

Theorem 2.2. *Let $G = (V, E)$ be a weighted graph and $k \in \mathbb{N}$. Let T^* be the cheapest tree with maximum degree k. Then, there exists a polynomial time algorithm which outputs a tree of cost at most $c(T^*)$ and maximum degree $k+1$.*

This allows us to consider a different metric for approximation: instead of losing on cost, we can output a solution that has slightly weaker properties than the OPT we compare against.

Notice that the above theorem is tight (unless P=NP). Setting $k = 2$, this is the Hamiltonian path problem, so it cannot be solved without losing something on maximum degree (even without costs).

References

[BF81] József Beck and Tibor Fiala. ""Integer-making" theorems". In: *Discrete Applied Mathematics* 3.1 (1981), pp. 1–8 (cit. on p. 1).

[SL15] Mohit Singh and Lap Chi Lau. "Approximating minimum bounded degree spanning trees to within one of optimal". In: *Journal of the ACM* 62.1 (2015). DOI: 10.1145/2629366 (cit. on p. 2).