A simpler SDP approach to coloring discrepancy

Aryan Hemmati*

July 2023

Abstract

Coloring discrepancy is the problem of coloring a ground set V of vertices with two colors such that the deviation of this coloring from a totally balanced state is minimized. We study different variations of this problem, along with its importance and applications in rounding methods. And finally present a new approach to the problem that might get to some newer results on the problem.

Complexity of the problem

Coloring discrepancy problem (in its most general form) consists of a ground set $V = \{1, \cdots, n\}$ and a family $\mathcal{F} \in 2^V$ and intends to find a coloring $\chi \in \{\pm 1\}^V$ such that $\text{disc}(\chi) = \max\limits_{F_i \in \mathcal{F}} |\sum\limits_{v_j \in F_i} \chi(v_j)|$ is minimized and the value $\min\limits_{\chi \in \{\pm 1\}^V} \text{disc}(\chi)$ which indicates the minimum deviation from the balanced coloring amongst all possible colorings, is referred to as the coloring discrepancy of the set system $(V, \mathcal{F})$. For studying computational complexity of this problem, one might check the related decision problems in the first place:

1. Given a set system $(V, \mathcal{F})$ and $k \in \mathbb{N}$, can one determine, in polynomial time, if $\text{disc}(V, \mathcal{F}) \geq k$?

2. Given a set system $(V, \mathcal{F})$ and $k \in \mathbb{N}$, can one determine, in polynomial time, if $\text{disc}(V, \mathcal{F}) = k$?

3. Given a set system $(V, \mathcal{F})$, can one determine, in polynomial time, whether or not $\text{disc}(V, \mathcal{F}) = 0$?

First of all note that for any set system $(V, \mathcal{F})$, one can obviously deduce that $\text{disc}(V, \mathcal{F}) \leq n$. Now assuming there is an algorithm for solving (1), one can run the algorithm for values $k, k+1$. If either $\text{disc}(V, \mathcal{F}) \geq k, \text{disc}(V, \mathcal{F}) \geq k+1$, or $\text{disc}(V, \mathcal{F}) < k$, then the answer to the decision problem (2) is negative. But if $\text{disc}(V, \mathcal{F}) \geq k, \text{disc}(V, \mathcal{F}) < k+1$, then the answer is affirmative. On the other hand, having been provided with an algorithm for solving (2), one can execute the algorithm for values $k = 0, \cdots, k_0 - 1$ (note that since $\text{disc}(V, \mathcal{F}) \leq n$, one can return a negative answer for any entry $k_0 > n$. Hence, we'll assume that $k_0 \leq n$). If any of the executions return an affirmative answer, the answer to (1) is negative and if none of the executions return an affirmative answer, the answer to (1) will be affirmative. Hence, one can deduce that decision problems (1),(2) can be reduced to each other. Also it's pretty clear that (3) can be reduced to (2). Hence, (3) can also be reduced to (1). Now to solve (3), one may take the linear algebraic approach. Let $\mathbf{A} \in \mathbf{M_{m\times n}}(\mathbb{Z}_\mathbf{2})$ be the incidence matrix of the set system $\text{disc}(V, \mathcal{F})$. Then one can deduce that (3) is equivalent to the following decision problems:

4. Given a matrix $\mathbf{A} \in \mathbf{M_{m\times n}}(\mathbb{Z}_\mathbf{2})$, does there exist $\mathbf{x} \in \{\pm 1\}^n$ such that $\mathbf{Ax} = \mathbf{0}$.

5. Given a matrix $\mathbf{A} \in \mathbf{M_{m\times n}}(\mathbb{Z}_2), \mathbf{b} \in \{\mathbb{Z}_\mathbf{2}\}^\mathbf{n}$, does there exist $\mathbf{x} \in \{\mathbb{Z}_2\}^n$ such that $\mathbf{Ax} = \mathbf{b}$.

*aryanhemmati1382@gmail.com

Note that $(\exists \mathbf{x} \in \{\pm 1\}^n : \mathbf{A}\mathbf{x} = \mathbf{0}) \iff (\exists \mathbf{x} \in \{\pm 1\}^n : \mathbf{A}(\mathbf{x} + \mathbf{J_n}) = \mathbf{A}\mathbf{J_n} = \mathbf{b} \in \mathbb{Z}^{\mathbf{n}}) \iff \left(\exists \mathbf{x}' \in \{\mathbb{Z}_2\}^n : \mathbf{A}\mathbf{x}' = \frac{\mathbf{1}}{\mathbf{2}}\mathbf{b}\right)$. From this result, it's clear that (5) can indeed be reduced to (4). And since (4),(3) are equivalent, this means that (5) can also be reduced to (1). But it's obvious that (5) is the linear algebraic formulation of the exact cover problem which is well known to be NP-complete. Therefore the problem of evaluating coloring discrepancy of a set system is NP. In fact, in [1] it's proved that it is NP-hard to distinguish between set systems of discrepancy 0 and those with discrepancy of $\Omega(\sqrt{n})$. Therefore, it's also relatively hard to even approximate coloring discrepancy.

Background of the problem, variations and related results

One of the basic results on discrepancy is by random colorings. Color any of the vertices with a $\frac{1}{2}$ probability. Then for any $S \in \mathcal{F}$, $\chi(S)$ would be a random variable with mean 0 and variance $|S|$. Then, by Chernoff bounds:

$$\mathbf{Pr}\left(|\chi(S)| \geq c\sqrt{|S|}\right) \leq 2\exp\left(\frac{-c^2}{2}\right)$$

Letting $c = 2\sqrt{\log m}$, one can deduce that $\mathbf{Pr}\left(|\chi(S)| \geq c\sqrt{|S|}\right) \leq \frac{2}{m^2}$. Hence by Boole's inequality, the coloring will satisfy $|\chi(S)| \leq 2\sqrt{\log m|S|} \leq 2\sqrt{n\log m}$ for all $S \in \mathcal{F}$ with probability $1 - |\mathcal{F}|\frac{2}{m^2} = 1 - \frac{2}{m}$. This result can be improved to $\mathcal{O}(\sqrt{n\log\frac{m}{n}})$.[2]

One of the more famous settings for this problem is the Beck-Fiala setting, where along with the set system, we're given $t \in \mathbb{N}$ and guaranteed that $\forall v \in V : |\{S \in \mathcal{F} \mid v \in S\}| < t$. Noting the random coloring result, we can achieve a better result simply bounding $|S|$ by t in the last inequality. So random coloring will ensure a non-constructive $\mathcal{O}(\sqrt{t\log m})$ bound in the Beck-Fiala setting.

Since disc itself is very hard to approximate, it's commonly easier to bound it with relevant, more well-behaved values. Along with all the things about complexity of evaluating exact or approximate value for discrepancy, a property that makes it so hard to even approximate is it's not well-behaved at all and is in a sense, fragile. For understanding this disadvantage, let $(V, \mathcal{F})$ be an arbitrary set system with very high discrepancy. Let $(V', \mathcal{F}')$ be a copy of $(V, \mathcal{F})$ on the new ground set V'. ($\mathcal{F}'$ is constructed accordingly) Now let $\mathcal{T} = \{S_i \cup S'_i \mid i = 1, \cdots, m\}$ where S_i, S'_i are corresponding subsets of the ground set V, V'. Now consider the set system $(V \cup V', \mathcal{T})$ which is of zero discrepancy since you can color all of V with the first color and all of V' with the second color. But restricting the ground set of this system to V, we would get to a system of $\text{disc}(V, \mathcal{F})$ discrepancy, which is a relatively huge value. For fixing this issue, we define a more well-behaved notion of coloring discrepancy:

$$\text{herdisc}(V, \mathcal{F}) = \max_{V' \subseteq v} \text{disc}(V', \mathcal{F}\mid_{V'}) \quad \text{where} \quad \mathcal{F}\mid_{V'} = \{S \cap V' \mid \forall S \in \mathcal{F})$$

However, Bansal proved (using special random walks) a very close relationship between disc, herdisc of a set system:

There exists a randomized polynomial-time algorithm that "computes"
a coloring χ with $\text{disc}(\chi) = \mathcal{O}(H\log(mn))$ where $H > \text{herdisc}(V, \mathcal{F})$

The proof is based on Semi-Definite programming rounding and sticky random walks. On the other hand, a classic result of Beck and Fiala [3] implies that $\text{herdisc}(V, \mathcal{F})$ is less than the maximum l_1-norm of all columns of the incidence matrix of the set system. Since the incidence matrix is positive, this is equal to $||\mathbf{A1}||_\infty$. However, the Beck-Fiala setting can also be translated into the same thing, adding a constraint that $\mathbf{A1} \leq t\mathbf{1} \implies ||\mathbf{A1}||_\infty < t$. Therefore, this implies that Bansal's algorithm gives a constructive bound of $\mathcal{O}(\sqrt{t}\log(mn))$. We'd proceed to exhibit a relatively simpler approach to the sticky walk of Bansal's algorithm:

SDP Approach

First of all note that the discrepancy problem can be modeled as the following program: small

$$\begin{aligned} \text{minimize} \quad & D \\ \text{subjectto}: \quad & \Big|\sum_{v_i \in S_j} x_i\Big| \leq D, \quad \forall S_j \in \mathcal{F} \\ & x \in \{\pm 1\}^n \end{aligned}$$

Applying Semi-Definite relaxation: small

$$\begin{aligned} \text{minimize} \quad & D \\ \text{subjectto}: \quad & \Big|\Big|\sum_{v_i \in S_j} \mathbf{u_i}\Big|\Big| \leq D, \quad \forall S_j \in \mathcal{F} \\ & ||u_i|| = 1, \quad \forall i \in [n] \qquad\qquad U \succeq 0 \end{aligned}$$

The optimal solution to this SDP is referred to as the vector discrepancy. (verdisc) Clearly, since this is a relaxation of the previous program, $\text{verdisc}(V, \mathcal{F}) \leq \text{disc}(V, \mathcal{F})$. It is well-known that a SDP can be solved within a given error bound in polynomial time. (that is; if the SDP is well-behaved, meaning that the coefficients of the program are all rational and that the feasible solutions are bounded in some sense, since we mostly use interior-point methodology to find a approximate feasible solution for a SDP) Knowing this, the algorithm proceeds with a randomized iterative rounding, that with high probability reaches a valid integer solution. (and hence a valid coloring) Consider a hypercube $[-1, 1]^n$ where each axis represents an element of the ground set and each vertex represents a valid coloring. We'd from $\mathbf{x} = \mathbf{0}$ which represents a pseudo-coloring (since it has non-integer coordinates) with discrepancy 0. (yet useless for our rounding) and take small random steps inside the cube through each iteration. In the i-th iteration of algorithm, we generate an increment vector δ_i in the following manner: Fix $j \in [n]$, we'd take a random (equidistributively) unit vector $z_{t,j}$ and an n-dimensional Gaussian λ_t. (independent of $z_{t,k}$) Finally let $C_t = \sigma {\lambda_t}^T z_{t,j}$ where σ should be a relatively small value. Yet too small values of σ would result in increasing the run-time of the algorithm. Therefore a constant of $\sigma = \frac{1}{Cn \log n}$ would work efficiently, where C is a sufficiently huge value. It's quite easy to see that C_t's are independent random variables with one-dimensional normal distribution of $N(0, \sigma^2)$. Now initialize $\delta_t = 0$ and then set $(\delta_t)_j = C_t$ and do this in each step for all $j \in [n]$ if $(x_t)_j \in (-1, 1)$. In each iteration let x_t be equal to $x_{t-1} + \delta_t$, after it's pruned; where pruning is defined as if we set each coordinate to its closest extremum value (± 1) if the current position is out of bounds of this coordinate. One can see that if the point reaches the extremum value of any of the coordinates, the increments are set in a way that it won's move off the relevant bound face of the hypercube. The walk would then be proceeding for $B\sigma^{-2}\log n$ steps. If not ending on a vertex, the algorithm would start over. Finally it can be proved by the Boole's inequality that in each execution of the algorithm, there is a possibility of at least $1-\frac{1}{n}$ that the algorithm terminates. (landing on a vertex) which completes the algorithm.

References

[1] M. Charikar, A. Newman, and A. Nikolov. Tight hardness results for minimizing discrepancy. In Proc. 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), pages 1607–1614, 2011

[2] J. Spencer. Ten Lectures on the Probabilistic Method. CBMS-NSF. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, USA, 1987

[3] J. Beck and T. Fiala, 'Integer-making' theorems, Discrete Appl. Math. 3, 1981