A Goemans-Williamson-inspired approach

Thomas Zhang

1 Introduction to vector discrepancy problem

In Nikolov's paper [1], he presents a two-fold relaxation to the set a.k.a. combinatorial discrepancy, problem, called the vector discrepancy problem (frankly, a pretty confounding name but oh well). My interpretation of the set discrepancy problem is the following: we have n types of items and m boxes, each box containing some collection of items from the n types. We then want to assign either a positive or negative charge on each type of item such that net charge on each box is "minimized" in the min-max sense, i.e. the net charge of the box with the largest net charge by magnitude is minimized. We can conveniently translate this into matrix-vector language.

We can treat each box as a $\{0,1\}^n$ row vector, where 1 indicates that the corresponding item is in the box and 0 otherwise. We can then line up these m row vectors to form what is called the "incidence" matrix. Observe that we can encode the assignment of charges into a $\{-1,1\}^n$ vector, call it x, such that the vector $Ax \in \mathbb{Z}^m$ encodes as its element the net charge of each box. For a given set system, we can construct its corresponding incidence matrix A and pose the following objective:

$$\begin{aligned} &\min_{x\in\{-1,+1\}^n} \max_{i\in\{1,\dots,m\}} |(Ax)_i| \\ \iff &\min_{x\in\{-1,+1\}^n} \|Ax\|_\infty . \end{aligned}$$

Nikolov studies a version of the set discrepancy problem that is relaxed in two ways. Firstly, instead of having A as a 0/1 matrix, we can relax the conditions such that the columns of A are bounded in the ℓ_2 norm−w.l.o.g. at most norm 1. This is clearly a relaxation because if A were a 0/1 matrix, its columns are clearly bounded in ℓ^2 norm by m. Secondly, instead of considering $x \in \{-1,+1\}^n$, which is analagous to assigning a sign to each column of A, we instead consider assigning a "vector" to each column of A. Formalizing this notion: we observe that in the set discrepancy case, $Ax = \sum_{j=1}^n A_{*,j}x(j)$, where $A_{*,j}$ is the j-th column of matrix A and $x(j)$ is the j-th element of x, i.e. the "sign" of $A_{*,j}$. Observe that $\|Ax\|_\infty = \max_{i=1,\dots,m} \left|A_{i,*}^\top x\right| = \max_{i=1,\dots,m} \left|\sum_{j=1}^n A_{ij}x(j)\right|$. In the relaxation, we replace $x(j)$ with a ℓ^2 unit vector u_j and take $\left\|\sum_{j=1}^n A_{ij}u_j\right\|_2$ to get the value analogous to $\sum_{j=1}^n A_{ij}x(j)$.

We then define the "vector" discrepancy:

$$\begin{aligned} & \operatorname{disc}(A) := \min_{x(1),\ldots,x(n)\in\{-1,+1\}} \max_{i=1,\ldots,m} \left| \sum_{j=1}^{n} A_{ij}x(j) \right| \\ \sim \quad & \operatorname{vecdisc}(A) := \min_{u_1,\ldots,u_n,\, \|u_j\|_2=1} \max_{i=1,\ldots,m} \left\| \sum_{j=1}^{n} A_{ij}u_j \right\|_2 . \end{aligned}$$

To see why this opaque mess is actually ends up being a relaxation, we see every instance of $x \in \{-1,+1\}^n$ can be translated to a collection of unit vectors $\{u_j\}_{j=1}^n$: simply set $u_j = x(j)v$, where v is any ℓ^2 unit vector. Given $x \in \{-1,+1\}^n$ that attains $\operatorname{disc}(A)$,

$$\begin{aligned} \operatorname{disc}(A) = \|Ax\|_\infty &= \max_{i=1,\ldots,m} \left| \sum_{j=1}^{n} A_{ij}x(j) \right| \\ &= \max_{i=1,\ldots,m} \left| \sum_{j=1}^{n} A_{ij}x(j) \right| \|v\|_2 \\ &= \max_{i=1,\ldots,m} \left\| \sum_{j=1}^{n} A_{ij}x(j)v \right\|_2 \\ &= \max_{i=1,\ldots,m} \left\| \sum_{j=1}^{n} A_{ij}u_j \right\|_2 , \quad \|u_j\|_2 = 1 \\ \implies \operatorname{disc}(A) \geq \operatorname{vecdisc}(A). \end{aligned}$$

The first relaxation is not unique to this paper; in fact, it is essentially the version of the problem as initially given by Dan, i.e. trying to add and subtract a collection of vectors (read: columns of the incidence matrix) to get as small of a resulting vector as possible. However, the second relaxation may seem to come out of nowhere, and I certainly don't claim to have a geometric or combinatorial explanation of it. To understand the motivation behind the second relaxation, we have to first introduce the Komlós conjecture.

Remark 1.1 (Komlós Conjecture) *Given matrix A whose columns have ℓ^2 norm at most 1, (set) disc$(A) \leq K$ for some absolute constant K independent of the number of rows or columns of A. In other words,*

$$\|A_{*,j}\|_2 \leq 1\, \forall j \implies \min_{x\in\{-1,+1\}^n} \|Ax\|_\infty \leq K.$$

Nikolov's paper essentially refutes the following line of thought: can we *disprove* the Komlós conjecture by relaxing the constraints of the set discrepancy problem and then showing the discrepancy of the relaxed problem is lower bounded by some non-constant factor (say $K\log(n)$)? Nikolov shows that this line of thought does not work, at least with the aforementioned relaxations to the problem.

2 Improvement on proof of vecdisc(A) ≤ 1

Let us consolidate the statement of the "vector discrepancy" problem: given $A \in M_{m,n}$, we want to find

$$vecdisc(A) := \min_{u_1,\ldots,u_n,\, \|u_j\|_2=1} \max_{i\in\{1,\ldots,m\}} \left\| \sum_{j=1}^{n} A_{ij} u_j \right\|_2 . \tag{1}$$

For this problem we have a natural analog to the Komlós conjecture that Nikolov proves in his paper

Theorem 2.1 (Vector Komlós) *For matrix $A \in M_{m,n}$ that has columns with ℓ^2 norm at most 1, $vecdisc(A) \leq 1$.*

Observe that this defeats any hope of using this relaxed version of the set discrepancy problem to disprove the Komlós conjecture. I won't go into Nikolov's proof for the above theorem, but here's the quick rundown:

1. Formulate a semi-definite program that is equivalent to the vector discrepancy problem
2. Derive the dual program that is feasible only for D such that vecdisc$(A) \geq D$ (strong duality holds here)
3. Use a bunch of matrix/linear algebra lemmas to create an ellipsoid argument that contradicts the dual program being feasible for any $D = 1 + \epsilon \implies$ optimal value of primal program is upper bounded by 1, and it is tight by trivial examples.

Just by making some observations about the primal SDP and its dual that I will derive, I think I have come up with a simpler proof for actually a slightly more general result that implies the above theorem.

Observe that we can re-phrase problem (1) into a minimization problem:

$$\begin{aligned}
&\min \quad t \\
&\text{subject to} \\
&t \geq \left\| \sum_{j=1}^{n} A_{ij} u_j \right\|_2^2 = \sum_{j=1}^{n} \sum_{k=1}^{n} A_{ij} A_{ik} u_j^\top u_k,\ i = 1, \ldots, m \\
&\|u_j\|_2 = 1,\ j = 1, \ldots, n.
\end{aligned}$$

where t^* is vecdisc$(A)^2$. We'll see later that no assumption needs to be made of A. Conveniently, we can embed this problem into a semi-definite program:

$$\begin{aligned}
&\min \quad t \\
&\text{subject to} \\
&t \geq A_{i,*}^\top X A_{i,*},\ i = 1, \ldots, m \\
&X_{jj} = 1,\ j = 1, \ldots, n \\
&X \succeq 0.
\end{aligned}$$

We appeal to Cholesky factorization to establish the equivalence between the former and latter programs: given a list of u_j such that $\|u_j\| = 1$, we look at the matrix $X = U^\top U = \begin{bmatrix} u_1 & \dots & u_n \end{bmatrix}^\top \begin{bmatrix} u_1 & \dots & u_n \end{bmatrix}$ and immediately observe that X is psd and $X_j j = \|u_j\|_2^2 = 1$. On the other hand, given a psd matrix X whose diagonal entries are all 1, we can use Cholesky factorization to get $X = U^\top U$, where U has n columns, and each column is a ℓ^2 unit vector.

Here is where we diverge from Nikolov's work. We explicitly write out the dual of the above semi-definite program by essentially adapting the concept of Lagrangian dual for convex programming (there's probably a formal name for that technique but I don't know it):

$$\begin{aligned} &\max \quad \sum_{j=1}^{n} \lambda_j \\ &\text{subject to} \\ &\sum_{i=1}^{m} \mu_i A_{i,*} A_{i,*}^\top \succeq \operatorname{diag}(\lambda_1, \dots, \lambda_n) \\ &\sum_{i=1}^{m} \mu_i = 1, \mu_i \geq 0, \ i = 1, \dots, m. \end{aligned}$$

At this point, we might have to consider the duality gap in such a problem. However, we will show that primal-dual Slater's condition for strong duality holds here. In a geometric sense, we want to show that the feasible region for *both* the primal and dual problems has a non-empty interior. What this boils down to in our case that we have to show that we can attain an example of *strict* inequalities for all our constraints that have inequalities (a helpful picture would be to consider the Slater condition in the linear programming case: do our inequalities form a polytope in which one can find a point that is not touching any of the faces?). For our primal program, this is rather simple: given matrix A, we find n linearly independent u_j such that $X = U^\top U$ has diagonals all 1 and is strictly pd. Since A has finitely many rows, we can pick any t such that it exceeds all $A_{i,*}^\top X A_{i,*}$. For our dual program, given a list of $\mu_i \geq 0$ such that $\sum_{i=1}^m \mu_i = 1$, we observe that $\sum_{i=1}^m \mu_i A_{i,*} A_{i,*}^\top$ is at least a psd matrix. Since λ_j are unconstrained, we simply find "negative enough" λ_j such that

$$\sum_{i=1}^{m} \mu_i A_{i,*} A_{i,*}^\top - \operatorname{diag}(\lambda_1, \dots, \lambda_n) \succ 0.$$

Therefore, by showing that we can find feasible solutions to both the primal and dual programs that don't hug any equalities in the inequality constraints, we have shown that the feasible regions of both have nonempty interiors. Slater's condition then tells us that strong duality holds for this primal-dual pair, i.e. $t^* = \sum_{i=1}^m \lambda_j^*$.

At this point we claim the main theorem of this section

Theorem 2.2 *Given* $A \in M_{m,n}$, $A = \begin{bmatrix} a_1^\top \\ \vdots \\ a_m^\top \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$,

$$\textit{vecdisc}(A) \leq \min\left\{ \max_{i=1,\ldots,m} \|a_i\|_2^2 \, , \; \max_{j=1,\ldots,n} \|v_j\|_2^2 \right\}.$$

This immediately shows as a corollary the main result of Nikolov's paper:

Corollary 2.3 *Given* $A \in M_{m,n}$ *such that* $\|v_j\|_2 \leq 1$ *for* $j = 1, \ldots, n$, *then*

$$\textit{vecdisc}(A) \leq 1.$$

To prove the theorem, we turn our focus to the dual program in particular. Let $(\vec{\lambda}^*, \vec{\mu}^*)$ be an optimal solution to the dual program. We will prove that $\text{vecdisc}(A) \leq \max_{i=1,\ldots,m} \|a_i\|_2^2$. Just by plugging our optimal values into the constraints of the dual program, we get:

$$\begin{aligned}
\sum_{i=1}^{m} \mu_i^* a_i a_i^\top &\succeq \text{diag}(\lambda_1^*, \ldots, \lambda_n^*) \\
\implies \text{tr}\left(\sum_{i=1}^{m} \mu_i^* a_i a_i^\top \right) &\geq \sum_{j=1}^{n} \lambda_j^* = \text{vecdisc}(A) \\
\text{tr}\left(\sum_{i=1}^{m} \mu_i^* a_i a_i^\top \right) &= \sum_{i=1}^{m} \text{tr}\left(\mu_i^* a_i a_i^\top\right) \\
&= \sum_{i=1}^{m} \text{tr}\left(\mu_i^* a_i^\top a_i\right) \\
&= \sum_{i=1}^{m} \mu_i^* \max_{i=1,\ldots,m} \|a_i\|_2^2 \\
&= \|a_i\|_2^2 \\
\implies \text{vecdisc}(A) &\leq \max_{i=1,\ldots,m} \|a_i\|_2^2 \, .
\end{aligned}$$

To prove that $\text{vecdisc}(A) \leq \max_{j=1,\ldots,n} \|v_j\|_2^2$, we once again plug in $(\vec{\lambda}^*, \vec{\mu}^*)$ to the dual program:

$$
\begin{aligned}
\sum_{i=1}^{m} \mu_i^* a_i a_i^\top &\succeq \operatorname{diag}(\lambda_1^*, \ldots, \lambda_n^*) \\
\sum_{i=1}^{m} \mu_i^* a_i a_i^\top &= A^\top \operatorname{diag}(\mu_1^*, \ldots, \mu_m^*) \\
&= \begin{bmatrix} v_1^\top \\ \vdots \\ v_n^\top \end{bmatrix} \operatorname{diag}(\vec{\mu}^*) \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \\
&= \begin{bmatrix} v_1^\top \operatorname{diag}(\vec{\mu}^*) v_1 & & * \\ & \ddots & \\ * & & v_n^\top \operatorname{diag}(\vec{\mu}^*) v_n \end{bmatrix}.
\end{aligned}
$$

Since we have that $\operatorname{tr}\left(\sum_{i=1}^{m} \mu_i^* a_i a_i^\top \right) \geq \sum_{j=1}^{n} \lambda_j^*$, we then have

$$
\begin{aligned}
\operatorname{tr} \begin{bmatrix} v_1^\top \operatorname{diag}(\vec{\mu}^*) v_1 & & * \\ & \ddots & \\ * & & v_n^\top \operatorname{diag}(\vec{\mu}^*) v_n \end{bmatrix} &\geq \sum_{j=1}^{n} \lambda_j^* \\
\iff \sum_{j=1}^{n} v_j^\top \operatorname{diag}(\vec{\mu}^*) v_j &\geq \sum_{j=1}^{n} \lambda_j^*.
\end{aligned}
$$

To upper bound the value on the left hand side, we consider the following linear program:

$$
\begin{aligned}
&\max \quad \sum_{j=1}^{n} v_j^\top \operatorname{diag}(\mu_1, \ldots, \mu_m) v_j \\
&\text{subject to} \\
&\sum_{i=1}^{m} \mu_i = 1, \ \mu_i \geq 0, i = 1, \ldots, m.
\end{aligned}
$$

Observe that the optimal value to this program is an upper bound for $\sum_{j=1}^{n} v_j^\top \operatorname{diag}(\vec{\mu}^*) v_j$. Since this is a linear (hence convex) maximization over a bounded polytope, its optimal value has a vertex solution, i.e. some $\mu_k = 1$ and $\mu_j = 0$ for other indices. Quick observation tells us the μ_k we should grant the value of 1 is the index k such that $v_k^\top v_k = \|v_k\|_2^2 = \max_{j=1,\ldots,n} \|v_j\|_2^2$. Therefore,

$$
\max_{j=1,\ldots,n} \|v_j\|_2^2 \geq \sum_{j=1}^{n} v_j^\top \operatorname{diag}(\vec{\mu}^*) v_j \geq \sum_{j=1}^{n} \lambda_j^* \geq \operatorname{vecdisc}(A).
$$

This completes the proof of the claim.

3 Randomized algorithm using Goemans-Williamson machinery

Disclaimer: this part is mostly me toying around with things I don't understand fully from Goemans and Williamson's approximation scheme of `Max 2-cut`, and the bounds I get are pretty bad right now, but I think there is room for improvement, and even if there isn't, it's been a great learning experience.

The goal here is to come up with a randomized algorithm to come up with $x \in \{-1, +1\}^n$ for the set discrepancy problem. We will later see at multiple points that the naive implementation of machinery from Goemans-Williamson makes it so that only a positive matrix A gets sensical bounds. Fortunately, this is fine for the original set discrepancy problem, where A is a 0/1 matrix. Perhaps the core to this whole section is the simple rounding scheme that Goemans and Williamson use: say that in solving the primal semi-definite program from above to optimality we get matrix X^* as an optimal solution. We then treat it as a covariance matrix and sample

$$\xi \sim N(0, X^*).$$

We then round ξ in the following manner to extract a ± 1 vector $x(\xi)$:

$$x(\xi)_i = \begin{cases} +1 & \text{if } \xi(i) > 0 \\ -1 & \text{if } \xi(i) < 0. \end{cases}$$

Sheppard proved in 1899 that

$$\mathbb{E}[x(\xi)_i x(\xi)_j] = \frac{2}{\pi} \arcsin(X^*_{ij}).$$

We now introduce three lemmas, two of which are elementary, and the third coming from Goemans and Williamson.

Lemma 3.1 (Hoeffding's Lemma) *Suppose $Z \in [a, b]$. Then*

$$\Pr[Z \geq \mathbb{E}[Z] + t] \leq e^{\frac{-2t^2}{(b-a)^2}}.$$

Lemma 3.2 (Union (intersection) Bound) *Given random variables $Z_1, \ldots, Z_n$,*

$$\Pr\left[\bigcap_{i=1}^{n} Z_i\right] \geq 1 - \sum_{i=1}^{n} \Pr[Z_i^c],$$

where we intentionally denote Z_i^c as the complement of Z_i.

Lemma 3.3 *Suppose $x \in [-1, +1)$, then*

$$\frac{2}{\pi} \arcsin(x) \leq 1 - \alpha + \alpha x,$$

where $\alpha = 0.878\ldots$

Let us now define $Z_i = \left(a_i^\top x(\xi)\right)^2$, where $a_i^\top$ is the i-th row of the incidence matrix A. Observe that since $x(\xi) \in \{-1,+1\}^n$, Z_i is bounded: $Z_i \in [0, \|a_i\|_1^2]$. Furthermore, if we compute the expectation of Z_i, we get something interesting

$$\begin{aligned}
\mathbb{E}[Z_i] &= \mathbb{E}[a_i^\top x(\xi)x(\xi)^\top a_i] \\
&= a_i^\top \mathbb{E}[x(\xi)x(\xi)^\top]a_i \\
&= a_i^\top \begin{bmatrix} \frac{2}{\pi}\arcsin(X_{11}^*) & \cdots & \frac{2}{\pi}\arcsin(X_{1n}^*) \\ \vdots & \ddots & \vdots \\ \frac{2}{\pi}\arcsin(X_{n1}^*) & \cdots & \frac{2}{\pi}\arcsin(X_{nn}^*) \end{bmatrix} a_i \\
&= \frac{2}{\pi} a_i^\top \arcsin(X^*)a_i,
\end{aligned}$$

where $\arcsin(X^*)$ denotes arcsin applied element-wise to the matrix X^*.

Now applying Lemma 3.1, we get

$$\Pr[Z_i \geq \mathbb{E}[Z_i] + t] = \Pr[Z_i \geq \frac{2}{\pi}a_i^\top \arcsin(X^*)a_i + t] \leq e^{\frac{-2t^2}{\|a_i\|_1^4}}.$$

Applying Lemma 3.3 element-wise with respect to X^*, we get

$$\frac{2}{\pi}\arcsin(X^*) \leq (1-\alpha)\mathbb{1}\mathbb{1}^\top + \alpha X^*,$$

which implies

$$\frac{2}{\pi}a_i^\top \arcsin(X^*)a_i \leq (1-\alpha)(a_i^\top \mathbb{1})^2 + \alpha a_i^\top X^* a_i.$$

Here is where we make the assumption that A is a **positive matrix**, in which case $(a_i^\top \mathbb{1})^2 = \|a_i\|_1^2$. Returning to the probability bound, we now have

$$\begin{aligned}
&\Pr[Z_i \geq (1-\alpha)\|a_i\|_1^2 + \alpha a_i^\top X^* a_i + t] \\
\leq\ &\Pr[Z_i \geq \frac{2}{\pi}a_i^\top \arcsin(X^*)a_i + t] \\
\leq\ &e^{\frac{-2t^2}{\|a_i\|_1^4}}.
\end{aligned}$$

We now appeal to Lemma 3.2 to bound the intersection of probabilities across rows $a_i^\top$:

$$\begin{aligned}
&\Pr\left[\bigcap_{i=1}^m \left(Z_i \leq (1-\alpha)\|a_i\|_1^2 + \alpha a_i^\top X^* a_i + t\right)\right] \\
\geq\ &1 - \sum_{i=1}^m \Pr[Z_i \geq (1-\alpha)\|a_i\|_1^2 + \alpha a_i^\top X^* a_i + t] \\
\geq\ &1 - \sum_{i=1}^m e^{\frac{-2t^2}{\|a_i\|_1^4}}
\end{aligned}$$

Let us set $c = \max_{i=1,\dots,m} \|a_i\|_1$ and $t = \frac{\sqrt{2}}{2}c^2\sqrt{d}$:

$$\begin{aligned}\Pr\left[\bigcap_{i=1}^{m}\left(Z_i \le (1-\alpha)\|a_i\|_1^2 + \alpha a_i^\top X^* a_i + \frac{\sqrt{2}}{2}c^2\sqrt{d}\right)\right] &\ge 1 - \sum_{i=1}^{m} e^{\frac{-c^4 d}{c^4}} \\ &= 1 - me^{-d}.\end{aligned}$$

For the final flourish, we set $d = \log(2m)$ to get

$$\Pr\left[\bigcap_{i=1}^{m}\left(Z_i \le (1-\alpha)\|a_i\|_1^2 + \alpha a_i^\top X^* a_i + \frac{\sqrt{2}}{2}c^2\sqrt{d}\right)\right] \ge 1/2.$$

Recall that $Z_i := (a_i^\top x(\xi))^2$, i.e. the squared net "charge" of the i-th row. Furthermore, recall the primal semi-definite program

$$\begin{aligned}&\min \quad t \\ &\text{subject to} \\ &t \ge a_i^\top X a_i,\ i = 1,\dots,m \\ &X_{jj} = 1,\ j = 1,\dots,n \\ &X \succeq 0.\end{aligned}$$

Observe that by the constraints of the program, each $a_i^\top X^* a_i \le \text{vecdisc}(A)$. Therefore, for positive matrix A, the randomized algorithm will return $x(\xi)$ that attains a discrepancy no more than

$$\begin{aligned}&\sqrt{(1-\alpha)\max_{i=1,\dots,m}\|a_i\|_1^2 + \alpha\,\text{vecdisc}(A)^2 + \frac{\sqrt{2}}{2}\max_{i=1,\dots,m}\|a_i\|_1^2\sqrt{\log(2m)}} \\ &= \sqrt{\left(1-\alpha+\frac{\sqrt{2}}{2}\sqrt{\log(2m)}\right)\max_{i=1,\dots,m}\|a_i\|_1^2 + \alpha\,\text{vecdisc}(A)^2}\end{aligned}$$

with probability at least 1/2.

To finish off this write-up, we make some remarks on some ramifications of this result and areas of possible improvement.

Remark 3.4 *In the process of analyzing this randomized algorithm, some of the bounds were notably loose. For example, if we were to significantly improve the bounds on the random variable Z_i for the application of Hoeffding's Lemma , we might get a more palatable end result, as having $Z_i \in [0, \|a_i\|_1^2]$ is the sole reason we have $\max_{i=1,\dots,m}\|a_i\|_1^2$ in the final bound.*

Remark 3.5 *As Chris pointed out to me, since the algorithm is designed for instances of the original set discrepancy problem (positive A matrix and returns $x(\xi) \in \{-1,+1\}^n$) yet the guarantee involves vecdisc(A), which hints that perhaps a similar approach might help us bound the integrality gap between disc(A) and its relaxed value vecdisc(A).*

References

[1] Aleksandar Nikolov. The Komlós conjecture holds for vector colorings, 2013. Pre-print.

[2] Yinyu Ye. MS&E314/CME336: Conic Linear Optimization, Chapter 4: Semi-definite Programming Approximation (lecture notes), 2014. Retrieved from: http://web.stanford.edu/class/msande314/Chapter4.pdf