Discrepancy and Combinatorial Optimization
Lecture 1 - IPCO summer school

Nikhil Bansal

May 20, 2019

Abstract

Discrepancy theory is an area of mathematics that studies how well continuous objects can be approximated by discrete ones, and it has various connections to problems in combinatorics, optimization and computer science. E.g. how well a fractional solution to a linear program can be rounded to an integral one. In these notes, we look at some results and techniques that have been useful in combinatorial optimization, such as rounding techniques based on random walks and linear algebra and tools from convex geometry to show existence of and find integer points in convex bodies.

1 Introduction

Combinatorial discrepancy deals with the following question. Given a universe of elements $U = \{1, \ldots, n\}$ and some collection $\mathcal{S} = \{S_1, \ldots, S_m\}$ of subsets S_i of U. Find a red-blue coloring of points in U such that each set in $\mathcal{S}$ is colored as evenly as possible. More formally, given a $-1, +1$ coloring $x = (x_1, \ldots, x_n)$ of points in U, the discrepancy of coloring x is

$$\mathrm{disc}(\mathcal{S}, x) = \max_{j \in m} |\sum_{i \in S_j} x_i|$$

The discrepancy of $\mathcal{S}$ is defined as the minimum discrepancy over all colorings

$$\mathrm{disc}(\mathcal{S}) = \min_{x \in \{-1,1\}^n} \mathrm{disc}(\mathcal{S}, x).$$

Letting A denote the $m \times n$ incidence matrix of $(U, \mathcal{S})$, we have

$$\mathrm{disc}(\mathcal{S}) = \mathrm{disc}(A) := \min_{x \in \{-1,1\}^n} \|Ax\|_\infty$$

This also defines discrepancy for arbitrary real matrices A. Typically results in discrepancy are stated for set systems, but most of the techniques extend to general matrices.

2 Applications

We now discuss some motivating applications. For much more we refer to [12, 6].

2.1 Rounding

Typically in approximation algorithms, one takes a fractional solution to a problem, and tries to round it suitably to an integral one. The following relates the rounding error to discrepancy.

Theorem 2.1 (Lovász, Spencer, and Vesztergombi [11]). *For any $x \in \mathbb{R}^n$ satisfying $Ax = b$, there is a $\tilde{x} \in \mathbb{Z}^n$ with $\|\tilde{x} - x\|_\infty < 1$, such that $\|A(x - \tilde{x})\|_\infty \leq \text{herdisc}(A)$.*

Here herdisc(A) is the hereditary discrepancy of A, defined as follows.

Definition 1. *For a subset of columns $S \subseteq [n]$, let $A_{|S}$ be A restricted to S. Then*

$$\text{herdisc}A = \max_{S \subset [n]} \text{disc}(A_{|S}) = \max_{S \subset [n]} \min_{x \in \{-1,1\}^n} \|Ax\|_\infty$$

is the maximum discrepancy over all column restrictions $A_{|S}$ of A.

It is a more robust version of discrepancy that is monotone under taking subset of columns. But for almost all problems, any technique for bounding discrepancy implies the same result for hereditary discrepancy.

The main idea behind Theorem 2.1 is this. Suppose x is 1/2-integral (i.e. each x_i has fractional part 0 or 1/2), and let S be the subset of variables with fractional part 1/2, then the low discrepancy coloring of elements of S can be used to round the elements up or down.

Formally, let y be ± 1 coloring of S with disc$(A_{|S})$. Then, observe that $x' = x + y/2$ is integral, and the rounding error satisfies

$$\|Ax' - Ax\|_\infty = \|A(y/2)\|_\infty = \frac{1}{2}\text{disc}(A_{|S}) \leq \frac{1}{2}\text{herdisc}(A),$$

Proof. (Theorem 2.1) Fix an integer k, and consider the first k bits of the fractional part of each x_i. Applying the idea described above to the least significant bit, makes this bit 0 for each x_i, and introduces at most 2^{-k} herdisc(A) error.

We now repeat this for the $k-1$-th bits[1] and so on, until all the k bits of the fractional part are 0. This gives overall error $\sum_{\ell=k}^{1} 2^{-\ell}$ herdisc$(A) \leq$ herdisc(A), and the result follows by letting k go to infinity. □

2.2 Ordering with small prefix sums

Let $v_1, \ldots, v_n$ be vectors in $\mathbb{R}^d$ with $\|v_i\| \leq 1$ for each $i \in [n]$, and $\sum_i v_i = 0$. We wish to find an ordering of these vectors so that norm of each prefix sum is small. The is called the Steinitz problem and has a fascinating history, see [3] for a nice survey, with many surprising applications in scheduling and optimization. The following is a classic result.

Theorem 2.2. *(Steinitz). For any $v_1, \ldots, v_n$ satisfying the properties above, there is a permutation π with $m(\pi) = O(d)$, where $m(\pi) = \max_k \|\sum_{i=1}^{k} v_{\pi(i)}\|$.*

[1] these may have changed from those in x, due to carry over from rounding the k-th bit.

Remark 2.3. *This result actually holds for any norm* $\|\cdot\|$.

The idea will be to start from an arbitrary ordering, and iteratively improve it using a low discrepancy coloring. Formally, consider following discrepancy problem.

Definition 2 (Discrepancy of prefix sums.)**.** *Given an ordered list of vectors* $w_1, \ldots, w_n \in \mathbb{R}^d$ *with norm* $\|w_i\| \le 1$*, find* $x \in \{-1, +1\}^n$ *that minimizes* $\max_{k \in [n]} \|\sum_{i=1}^k s_i w_i\|$.

We have the following.

Theorem 2.4. *An* $f(d)$ *bound for discrepancy of prefix sums implies an* $f(d)$ *bound for the Steinitz problem.*

Proof. Let $v_1, \ldots, v_n$ be vectors satisfying the condition of Steinitz lemma. We will give a procedure below that given any ordering π of the v_i, finds another ordering π' with $m(\pi') \le (m(\pi) + f(d))/2$. So as long as $m(\pi) > f(d)$, $m(\pi')$ reduces by at least 1, and repeating this eventually gives an ordering $\tilde{\pi}$ with $m(\tilde{\pi}) \le f(d)$. □

The Procedure. Let π be some ordering of the vectors, and for this ordering, let $s \in \{-1, 1\}^n$ be a coloring with discrepancy of prefix sums at most $f(d)$. Let P denote the set of indices i with $s_i = 1$, and N be those with $s_i = -1$. The permutation π' is defined by first listing the indices in P (in the order of π) and then listing the indices in N (in the reverse order of π).

Example. Let π be the ordering $v_1\ v_2\ v_3\ v_4\ v_5\ v_6\ v_7\ v_8$. Suppose in the signed sum v_1, v_4, v_6, v_7 are colored $+1$ and v_2, v_3, v_5, v_8 are colored -1. Then the ordering by π' is $v_1\ v_4\ v_6\ v_7\ v_8\ v_5\ v_3\ v_2$.

Claim 1. $m(\pi') \le (m(\pi) + f(d))/2$.

Proof. The key point is that for any $k \in [n]$, $\sum_{i=1}^k v_i + \sum_{i=1}^k s_i v_i = 2\sum_{i \in P \cap [k]} v_i$.

As $\|\sum_{i=1}^k v_i + \sum_{i=1}^k s_i v_i\| \le \|\sum_{i=1}^k v_i\| + \|\sum_{i=1}^k s_i v_i\| \le m(\pi) + f(d)$, by the triangle inequality, this gives that

$$\|\sum_{i \in P \cap [k]} v_i\| \le (m(\pi) + f(d))/2,$$

i.e. every prefix sum in π' (ending at an index in P) has norm $(m(\pi) + f(d))/2$.

Similarly, as $\sum_{i=1}^k v_i - \sum_{i=1}^k s_i v_i = 2\sum_{i \in N \cap [k]} v_i$, for each $k \in [n]$

$$\|\sum_{i \in N \cap [k]} v_i\| \le (m(\pi) + f(d))/2.$$

As the order of the indices in N is reversed in π', this gives taht any suffix of π' (starting at an index in N) has norm $(m(\pi) + f(d))/2$. But as $\sum_i v_i = 0$, any prefix $[k]$ of π' has the same norm as the suffix $[n] \setminus [k]$. □

2.3 Sparsification

An important use of discrepancy is to replace a universe of elements by a much smaller one, that scales well for the sets we care about.

To see the idea, suppose there is a zero-discrepancy coloring. Let U' be points colored $+$ (or only $-$). Then $|U'| = |U|/2$ (we can ensure this by adding the set $S = U$), and each set S restricted to U' has $S/2$ elements. In general, small discrepancy allows one to iterated this process to get very interesting results. For example,

1. Given an undirected graph $G = (V, E)$, the set system with edges as elements and sets all cuts woud correspond to graph sparsification. One can decompose the complete graph into $\Omega(n)$ expanders with degree $O(1)$.

2. For set systems with VC dimension d, one can find a $O(d/\epsilon \log(1/\epsilon))$ universe that hits each set of size at least ϵn.

3. In numerical integration for a region R (say of volume 1), we want to replace R by a set of n points P that approximate R for some test sets S, in the sense that $\mathrm{vol}(R \cap S)$ is close to $\frac{1}{n}|P \cap S|$.

3 Linear Algebraic Methods

In the following sections we look at various methods to upper bound discrepancy. The first is based on a simple linear algebraic idea, but it often gives surprisingly powerful results. More generally, this technique is called iterated rounding, and has several applications in combinatorial optimization, a great reference is [10].

The method. We start with the coloring $x(0) = 0^n$, which is then updated over several iterations $1, \ldots, T$ until the final coloring $x(T)$ has all variables $-1, +1$. During the intermediate steps t, the variables $x_i(t)$ can take (fractional) values in the range $[-1, 1]$. At the beginning of time t, we call a variable i floating if $x_i(t-1) \in (-1, 1)$.

The update at time t. Pick some subset $S(t)$ of the floating variables at time t, and consider any linear system

$$B(t)y = 0, \quad y \in \mathbb{R}^n, \quad y_i = 0 \text{ for } i \in [n] \setminus S(t).$$

The idea is that if $B(t)$ has fewer than $|S(t)|$ rows, then this system has a non-zero solution $y(t)$, and we can update the solution $x(t-1)$ in the *direction* $y(t)$ until some variable reaches -1 or $+1$ to get $x(t)$, i.e.

$$x(t+1) = x(t) + \delta y(t)$$

where $\delta > 0$ is the largest real that ensures that $x(t+1) \in [-1, 1]^n$.

All the ingenuity in this method lies in choosing $B(t)$, at each time t. Note that once a variable reaches -1 or 1, it is fixed and never updated anymore. The algorithm takes at most n steps as at least one floating variable reaches ± 1 in each step. Let us consider a few examples to make this more concrete.

3.1 The Beck-Fiala Problem

Let $\mathcal{S}$ be a set system on n elements where each elements lies in at most k sets. One of the most important open problems in discrepancy is the following.

Beck Fiala Conjecture. $\mathrm{disc}(\mathcal{S}) = O(k^{1/2})$.

We show a $2k-1$ bound using the linear algebraic method. The best known bound independent of n is $2k - \Omega(\log^* k)$ [5].

Theorem 3.1 ([4])**.** $\mathrm{disc}(\mathcal{S}) \leq 2k - 1$.

Proof. Let A denote the incidence matrix of $\mathcal{S}$, where the rows correspond to sets and columns to elements. By our assumption, each column has at most k ones.

To apply the linear algebraic method, consider some iteration t and let $F(t)$ be the floating variables. We pick $S(t) = F(t)$, and $B(t)$ to be the rows with more than k elements. As each element in $S(t)$ lies in at most k rows, counting the number of 1's in the submatrix $B(t) \times S(t)$, we have $|B(t) < |S(t)|$.

As long as a set has more than k floating elements, its discrepancy remains 0. But, once a set has at most k floating elements, no matter how these variables are rounded in subsequent iterations, the additional discrepancy will be less than $2k$ (e.g. if all floating variables are -0.99 but get rounded to 1). As discrepancy of a set system is integral, we get the bound $2k - 1$. □

3.2 Discrepancy of prefix sums (Steinitz problem)

Recall the discrepancy of prefix sums problem. We are given $v_1, \ldots, v_n \in \mathbb{R}^d$ with $\|v_i\| \leq 1$ for $i \in [n]$. The goal is to find $s \in \{-1, 1\}^n$ with minimum $\max_{k \in [n]} \| \sum_{i=1}^k s_i v_i \|$.

Theorem 3.2. *The discrepancy of prefix sums is at most* $2d$*.*

Proof. We apply the linear algebraic method. Consider some iteration t, and let $S(t) = \{i_1, i_2, \ldots, i_{d+1}\}$ be the subset of the floating variables with the $d+1$ smallest indices (if there are d or fewer floating variables, we round them arbitrarily and terminate the algorithm). Let B be $d \times n$ matrix with columns v_i.

For any $k \in [n]$, consider the prefix sum $\sum_{i=1}^k x_i(t) v_i$. The key observation is that as long as $S(t)$ is contained in $[k]$, the update at time t satisfies

$$\sum_{i=1}^k y_i(t) v_i = \sum_{i \in [k] \cap S(t)} y_i(t) v_i = \sum_{i \in [n] \cap S(t)} y_i(t) v_i = By(t) = \mathbf{0},$$

which ensures that the discrepancy of the prefix remains 0.

At the first time when $S(t)$ is not contained in $[k]$, there can be at most d floating variables with indices in $[k]$, and no matter how they are rounded in subsequent iterations, the discrepancy incurred can be at most $2d$. □

4 Partial Coloring Technique

The partial coloring technique is one of the most powerful and widely applicable technique in discrepancy, and gives a very non-trivial way to go beyond the probabilistic method and union bound. Let us first consider what a random coloring gives us, for a set system $U = [n]$ and $\mathcal{S} = \{S_1, \ldots, S_m\}$.

Random Coloring. Suppose each element $i \in [n]$ is colored independently -1 or $+1$ with probability $1/2$. By standard Chernoff bounds, for any set S,

$$\Pr[\text{disc}(S) \geq \Delta_S] \leq 2 \exp\left(-\frac{\Delta_S^2}{2|S|}\right)$$

A union bound argument directly gives the following.

Theorem 4.1 (Union Bound.)**.** *Let $\Delta_j > 0$ be given for each set S_j, $j \in [m]$. If the Δ_j satisfy the condition*

$$\sum_j 2 \exp\left(-\frac{\Delta_j^2}{2|S_j|}\right) < 1 \tag{1}$$

then there is a coloring with discrepancy at most Δ_j for each set S_j.

If the right hand side of (1) is say 0.9 (or $1 - n^{-O(1)}$), this gives a polynomial time algorithm to find such a coloring.

Partial Coloring Lemma. The partial coloring lemma below looks similar, but can give substantially better discrepancy bounds than Theorem 4.1.

Theorem 4.2. *Let $\Delta_j > 0$ be given for each set S_j, $j \in [m]$. Suppose the Δ_j satisfy the condition*

$$\sum_{j \in [m]} g\left(\frac{\Delta_j}{\sqrt{|S_j|}}\right) \leq \frac{n}{5} \tag{2}$$

where

$$g(\lambda) = \begin{cases} Ke^{-\lambda^2/9} & \text{if } \lambda > 0.1 \\ K \ln(\lambda^{-1}) & \text{if } \lambda \leq 0.1 \end{cases}$$

and K is some absolute constant. Then there is a partial coloring that assigns ± 1 to at least $n/10$ variables (and 0 to the rest), with discrepancy at most Δ_j for each $j \in [m]$.

A key difference from Theorem 4.1 is that the condition (2) has $n/5$ on the right side, while (1) has 1. This gives substantially more power. For example, in Theorem 4.1 we cannot set $\Delta_j < \sqrt{S_j}$ even for a single set, while Theorem 4.2 allows us to set $\Delta_j < 1$ for $O(n/\log n)$ sets (which give a partial coloring with 0 discrepancy for all of those sets!).

The original proofs of Theorem 4.2 we all non-algorithmic, and based on the pigeonhole method and counting. More recently, several algorithmic variants have been discovered, and we shall consider these later.

4.1 Applications

Spencer's six standard deviations result. The partial coloring method was developed in the form above by Spencer [14], to show the following result, solving a problem proposed by Erdös.

Theorem 4.3. *Any set system on n elements and n sets has discrepancy $O(\sqrt{n})$.*

A union bound argument gives the bound $O(\sqrt{n \log n})$, which is actually tight for a random coloring. So, Spencer's result removes the $\sqrt{\log n}$ factor. This is the best possible up to constants, as there exist set systems with discrepancy at least $\Omega(\sqrt{n})$.

Proof. The coloring is constructed in phases $i = 0, 1, \ldots$, where in each phase i we apply the partial coloring lemma to the system restricted to the n_i remaining uncolored elements. So, $n_0 = n$ and $n_{i+1} \leq (0.9)n_i$ and hence $n_{i+1} \leq (0.9)^n$.

As there are n sets and n_i elements in phase i, setting $\Delta_S = c(n_i \log(2n/n_i))^{1/2}$ for each set S, with c large enough constant, satisfies (2) as,

$$ng(\lambda) \leq nK \exp\left(-\frac{\lambda^2}{9}\right) = Kn \exp\left(-\frac{c^2}{9} \log\left(\frac{2n}{n_i}\right)\right) \leq Kn \left(\frac{n_i}{2n}\right)^{c^2/9} \leq \frac{n_i}{5},$$

and so there is some partial coloring with discrepancy $c(n_i \log(2n/n_i))^{1/2}$ for each set in phase i. Summing up over the phases, the total discrepancy is at most

$$\sum_i c(n_i \log(2n/n_i))^{1/2} \leq \sum_i c \left(n(0.9)^i \log\left(\frac{2n}{n(0.9)^i}\right)\right)^{1/2} = O(n^{1/2}) \qquad \square$$

More generally for $m \geq n$, the method gives the bound $O(\sqrt{n \log(m/n)}$, which is also tight up to $O(1)$ factors.

Beck Fiala problem. A set system has degree k if each element lies in at most k sets.

Theorem 4.4. *Any set system with degree k has discrepancy $O(\sqrt{k} \log n)$.*

Proof. We will show that there exists a partial coloring with discrepancy $O(\sqrt{k})$. As there are $O(\log n)$ partial coloring phases this will give the result. So we set $\Delta_S = c\sqrt{k}$ for each set S for some $c = O(1)$, and show that the condition (2) holds.

This follows by noting that there are most $nk/2^\ell$ sets of size in $[2^\ell, 2^{\ell+1})$, as each element lies in at most k sets. Using that a sets of size $s \leq k$ contributes $e^{-\Omega(k/s)}$ to the left hand side of (2), and $O(\ln(s/k))$ if $s > k$, a simple but slightly tedious computation (that we can leave as an exercise) gives the result. $\square$

Remark 4.5. *Unlike in the Spencer's result, here the $O(\sqrt{k})$ error does not decrease over the partial coloring phases (as k does not reduce over the phases).*

Steinitz Problem. We show an $O(\sqrt{d}\log n)$ for discrepancy of prefix sums in the ℓ_∞ norm. For much more on other norms see [8, 2, 3]. It suffices to show the following.

Theorem 4.6. *There is a partial coloring with discrepancy $O(d^{1/2})$.*

Remark 4.7. *While Theorem 4.2 only considers set systems, the same result holds with $|S_j|$ replaced by $\|a_j\|_2$ (the ℓ_2 norm of the j-th row of A).*

We leave this as a guided exercise.

4.2 Proof of Partial Coloring Lemma

A nice exposition of Spencer's original proof, based on the pigeonhole principle and the entropy method is in [1]. Here we describe the independent proof due to Gluskin based on convex geometry.

Geometric View of Discrepancy. The following is a simple but very useful observation that connects geometry to discrepancy.

Observation 4.8. *The discrepancy of an $m \times n$ matrix A with rows a_i is Δ if and only if the polytope $P = \{x : |a_i x| \leq \Delta, i \in [m]\}$, contains some point in $\{-1, 1\}^n$.*

So proving Theorem 4.2 is equivalent to showing that the polytope $P = \{x : |a_i x| \leq \Delta_i, i \in [m]\}$, contains some point in $\{-1, 0, 1\}^n$ with at least $n/10$ non-zero coordinates, if Δ_i for $i \in [m]$ satisfy (2).

Gluskin's Theorem. A convex body K is symmetric if $x \in K$ implies $-x \in K$. The standard gaussian measure on the real line is given by the probability density $\gamma(x) = (1/\sqrt{2\pi})e^{-x^2/2}$. The n-dimensional gaussian measure is given by the density

$$\gamma_n(x_1, \ldots, x_n) = \prod_{i=1}^{n} \gamma(x_i) = \frac{1}{(2\pi)^{n/2}} e^{-\sum_i x_i^2/2}.$$

The gaussian measure γ_n has several interesting properties. In particular, it is rotationally invariant, and only depends on the euclidean distance $\|x\|_2$.

Gluskin proved the following more general theorem, which as we show below, will imply Theorem 4.2 (with different constants).

Theorem 4.9 (Gluskin [9])**.** *Any symmetric convex body K in $\mathbb{R}^n$ with $\gamma_n \geq 2^{-n/5}$, contains a $\{-1, 0, 1\}^n$ with at least $n/10$ non-zero coordinates.*

The idea of the proof is simple. For $x \in \mathbb{R}^n$, let $K_x := K + x$ denote K shifted by x. Consider the 2^n copies K_x for $x \in \{-1/2, 1/2\}^n$, then there is some point $y \in \mathbb{R}^n$ where $2^{\Omega(n)}$ copies overlap. So there exist $x, x' \in \{-1/2, 1/2\}^n$ differing in $\Omega(n)$ coordinates such that $y \in K_x$ and $y \in K_{x'}$. We then show that the point $x - x' \in \{-1, 0, 1\}^n$ satisfies the required conditions.

Before proving Theorem 4.9, we need some simple facts.

Lemma 4.10. *For any $x \in \mathbb{R}^n$ and any symmetric body K, $\gamma_n(K+x) \geq e^{-\|x\|^2}\gamma_n(K)$.*

Proof. As K is symmetric, let us see how the average density of y and $-y$ changes on shifting by x.

$$\begin{aligned} & \frac{1}{2}(\gamma_n(y+x)+\gamma_n(x-y)) \geq (\gamma_n(y-x)\gamma_n(y+x))^{1/2} \qquad \text{(AM-GM)} \\ = \quad & (2\pi)^{-n/2}e^{-(\|x-y\|^2+\|x+y\|^2)/4} = e^{-\|x\|^2/2}\gamma_n(y) \end{aligned}$$

where the last step uses that $\|x-y\|^2 = \|x\|^2+\|y\|^2-2\langle x,y\rangle$ and similarly for $\|x+y\|^2$. □

Lemma 4.11. *Any subset of $\{0,1\}^n$ of size $2^{n/2}$ contains some vector with*

Let $h(\alpha) = \alpha\log_2\alpha + (1-\alpha)\log_2(1-\alpha)$ be the binary entropy function.

Lemma 4.12. *Let $X = \{x : x \in \{0,1\}^n, \|x\|_1 \leq \alpha n\}$. Then $|X| < 2^{h(\alpha)n}$. In particular, for $\alpha = 1/10$, $|X| \leq 2^{n/2}$.*

Proof. Let us sample a random vector $Y = (Y_1,\ldots,Y_n)$ in X. The entropy of Y is $H(Y) = \log|X|$. By sub-additivity of entropy

$$H(Y) \leq \sum_i H(Y_i) = nh(Y_1) \leq nh(\alpha)$$

as each coordinate Y_i has value 1 with probability at most α. □

Proof. (Theorem 4.9) For $x \in \mathbb{R}^n$, let $K_x = \{k+x : k \in K\}$ denote the body K shifted by x. Consider the 2^n shifted bodies K_x for $x \in \{-1/2,1/2\}^n$. By Lemma 4.10, $\gamma_n(K_x) \geq e^{-n/8}\gamma_n(K)$ for each such x. So the total gaussian measure of these 2^n bodies is at least $2^n e^{-n/8}\gamma_n(K) \geq 2^{n/2}$.

As $\gamma_n(\mathbb{R}^n) = 1$, there must be some point y that lies in at least $2^{n/2}$ different K_x. By Lemma 4.12, there exist some x, x' with hamming distance at least $n/10$ such that y lies in both $K_{x'}$ and K_x. We claim that $z = x - x')$ satisfies the conditions.

First, z has all coordinates in $\{-1,0,1\}$ at least $n/10$ are non-zero. Second, as $y = x+k$ and $y = x'+k'$ for some $k,k' \in K$, we have $x-x' = k-k' \leq 2K$ (as $-k' \in K$ by symmetry of K) and hence $z \in K$. □

Recovering the bound for Polytopes. Define a strip $S_{v,\lambda}$ as $\langle x,v\rangle \leq \lambda$. The polytope P is the intersection of strips $\cap_{i\in[m]}S_{a_i,\Delta_i}$. The lemma below lower bounds the Gaussian measure of a convex body by a strip.

Lemma 4.13. *(Sidak-Khatri) For any symmetric convex body K and slab S*

$$\gamma_n(K \cap S) \geq \gamma_n(K)\gamma_n(S)$$

The proof is quite simple can be found here [8].

This gives that

$$\gamma_n(P) \geq \prod_i \gamma_n(S_{a_i,\Delta_i}).$$

Estimating Volume of P. We first give an estimate on the volume of a slab $\gamma_n(S_{a_i,\Delta_i})$.

By rotational symmetry of the Gaussian, we can assume that $a_i = \|a_i\|_2 e_1$. So, for $\Delta_i = \lambda_i \|a_i\|_2$, the volume is about $1 - e^{-\lambda_i^2/2}$. Using $1 - \varepsilon \approx e^{-\varepsilon}$. For large λ_i, this is roughly $\exp(\exp(-\lambda_i^2/2)$.

On the other hand, for $\lambda_i \ll 1$, as the one dimensional gaussian measure roughly uniformly distributed in an interval or size 1 around the origin, the volume of the slice is roughly $1/\lambda_i$.

So the condition, we need for $\gamma_n(P) \geq 2^{-n/5}$, upon taking logarithm precisely becomes (2).

5 Rothvoss' Algorithm

Let K be the convex body with $\gamma(K) \geq 2^{-\varepsilon n}$, for $\epsilon > 0$ small enough. Let $C = [-1,1]^n$ be the cube.

Algorithm. Sample a random point according to γ_n. Find the closest point $y \in K \cap C$ to x. Output y.

Solving for y is a convex optimization problem, and so this gives an efficient algorithm. Rothvoss showed that y gives the desired point.

Theorem 5.1 (Rothvoss [13])**.** *With high probability y has at least C_ϵ coordinates ± 1.*

The proof is a very elegant application of concentration of measure and Lemma 4.13. It is very readable and we refer the reader to [13].

Another algorithm. A related algorithm due to [7] is the following:

Pick a random direction c and optimize

$$\max c \cdot y, \quad y \in K \cap C$$

References

[1] Noga Alon and Joel H. Spencer. The probabilistic method, 2000.

[2] Wojciech Banaszczyk. On series of signed vectors and their rearrangements. *Random Structures & Algorithms*, 40(3):301–316, 2012.

[3] Imre Bárány. On the power of linear dependencies. In *Building bridges*, pages 31–45. Springer, 2008.

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

[5] Boris Bukh. An improvement of the beck-fiala theorem. *Combinatorics, Probability & Computing*, 25(3):380–398, 2016.

[6] Bernard Chazelle. *The discrepancy method: randomness and complexity.* Cambridge University Press, 2000.

[7] Ronen Eldan and Mohit Singh. Efficient algorithms for discrepancy minimization in convex sets. *CoRR*, abs/1409.2913, 2014.

[8] Apostolos Giannopoulos. On some vector balancing problems. *Studia Mathematica*, 122(3):225–234, 1997.

[9] E. D. Gluskin. Extremal Properties of Orthogonal Parallelepipeds and Their Applications to the Geometry of Banach Spaces. *Sbornik: Mathematics*, 64:85–96, February 1989.

[10] Lap-Chi Lau, R. Ravi, and Mohit Singh. *Iterative Methods in Combinatorial Optimization.* Cambridge University Press, 2011.

[11] László Lovász, Joel Spencer, and Katalin Vesztergombi. Discrepancy of set-systems and matrices. *European Journal of Combinatorics*, 7(2):151–160, 1986.

[12] Jiri Matousek. *Geometric discrepancy: An illustrated guide.* Springer Science, 2009.

[13] Thomas Rothvoss. Constructive discrepancy minimization for convex sets. In *IEEE Symposium on Foundations of Computer Science, FOCS*, pages 140–145, 2014.

[14] Joel Spencer. Six standard deviations suffice. *Transactions of the American Mathematical Society*, 289(2):679–706, 1985.