Integer sequences and semidefinite programming

Dedicated to Kálmán Győry
on the occasion of his 60th birthday.

László Lovász
Microsoft Research
One Microsoft Way, Redmond, WA 98052
lovasz@microsoft.com

May 2000

Technical Report MSR-TR-2000-48

Abstract

We show that Roth's theorem on the discrepancy of the family of arithmetic progressions can be derived using rather standard arguments in semidefinite optimization.

1 Roth's Theorem

Let $\mathcal{F}$ be a family of subsets of $\{0,1,\ldots,n\}$. We want to find a sequence $x = (x_0, x_1, \ldots, x_{n-1})$ of ± 1's so that each member of $\mathcal{F}$ contains about as many 1's as -1's. More exactly, we define the *discrepancy of the sequence* x by

$$\max_{A \in \mathcal{F}} \left| \sum_{i \in A} x_i \right|,$$

and the *discrepancy of the family* $\mathcal{F}$ by

$$\min_{x \in \{-1,1\}^n} \max_{A \in \mathcal{F}} \left| \sum_{i \in A} x_i \right|.$$

The following basic theorem in discrepancy theory was proved by Roth [4]:

Theorem 1 *The discrepancy of the family of arithmetic progressions is* $\Omega(n^{1/4})$.

One way of looking at this result is to think of the x_i in the definition of discrepancy as the output of a pseudorandom number generator, and of the discrepancy, as a randomness test. If the x_i are truly random, we expect this discrepancy to be about $n^{1/2}$. Most "bad" sequences one encounters fail by producing a larger discrepancy. Roth's Theorem shows that the discrepancy cannot be arbitrarily small, but it allows sequences to have substantially smaller discrepancy than a random sequence. One might expect that the lower bound in the theorem can be strengthened to about $\Omega(n^{1/2}$, but it was shown by Beck [2] that Roth's estimate is sharp up to a logarithmic factor. Recently, even this logarithmic factor was removed by Matoušek and Spencer [5].

Let us state Roth's Theorem with the constant we are going to prove.

Theorem 2

$$\max_{A} \left| \sum_{i \in A} x_i \right| > \frac{1}{7} n^{1/4}.$$

The aim of this note is to show that Roth's estimate can be obtained by a rather standard argument based on a quite different field, namely semidefinite optimization (see [1] and [6] for surveys of this field). It is beyond the scope of this note to give an introduction to semidefinite programming, but the arguments will be self-contained.

It seems that all proofs of this theorem establish more. First, one shows that this quantity is large even if we maximize over the following subfamily of arithmetic progressions. Let $k = \lfloor \sqrt{n/8} \rfloor$. Consider arithmetic progressions with difference at most $8k$ and length exactly k. We consider arithmetic progressions modulo n, *i.e.*, we

let them wrap around. (Of course, in this case it may happen that the progression with the large discrepancy is wrapped; but since $(k-1)(8k) < n$, it wraps over n at most once, and so it is the union of two unwrapped arithmetic progressions, one of which has discrepancy at least half the original.) Let $\mathcal{H}$ denote the family of such arithmetic progressions. Clearly $|\mathcal{H}| = 8kn$.

Roth proves that for every set S the discrepancy of arithmetic progressions in $\mathcal{H}$ is large even on the average:

Theorem 3

$$\frac{1}{8kn} \sum_{A \in \mathcal{H}} \left| \sum_{i \in A} x_i \right|^2 > \frac{1}{49} n^{1/2}. \tag{1}$$

We prove this theorem in the next section.

2 Semidefinite relaxation

Let $q(x)$ denote the quadratic form on the left hand side of (1). We want to show that the minimum of $q(x)$ over all $x_i = \pm 1$ is at least $(1/49)n^{1/2}$. The condition that $x_i = \pm 1$ can be written as a quadratic equation, and so we get the quadratic program

$$\begin{aligned} \text{minimize} \quad & q(x) && (2) \\ \text{subject to} \quad & x_0^2 = \ldots = x_{n-1}^2 = 1. && (3) \end{aligned}$$

The technique of semidefinite optimization we apply here is that we introduce new variables $y_{ij} = x_i x_j$, and consider the matrix $Y = (y_{ij})$. In these terms, $q(x) = \ell(Y)$ is a linear function of the entries of Y, and so the objective function and the constraints become linear:

$$\begin{aligned} \text{minimize} \quad & \ell(Y) && (4) \\ \text{subject to} \quad & y_{00} = \ldots = y_{n-1,n-1} = 1. && (5) \end{aligned}$$

In addition, we can note that

$$Y \text{ is positive semidefinite,} \tag{6}$$

and

$$Y \text{ has rank 1.} \tag{7}$$

It is easy to see that the minimum in (4) subject to (5), (6) and (7) is the same as the minimum in (2) subject to (3): if Y is positive semidefinite and has rank 1, then we can write $y_{ij} = x_i x_j$, and the vector x defined this way is a solution of (2).

We drop constraint (7) (which is non-convex), and show that the bound claimed holds for the solution of (4) subject to (5) and (6). This is clearly stronger than Theorem 3.

The next step is to notice that both (5) and (6) define convex sets in the space of matrices, and the objective function is linear. Moreover, all of them are invariant under the cyclic shift of indices. Hence by averaging, we get an optimum solution Y which itself is invariant under the cyclic shift of indices, *i.e.*, it satisfies

$$y_{i+1,j+1} = y_{ij} \tag{8}$$

(where the addition in the subscript is modulo n).

Now since Y is positive semidefinite, we can write $Y_{ij} = u_i^T u_j$, where $u_i \in \mathbb{R}^d$ for some $d \leq n$. We may assume that the u_i span $\mathbb{R}^d$. The objective function (4) becomes

$$\frac{1}{8kn} \sum_{A \in \mathcal{H}} \left| \sum_{i \in A} u_i \right|^2 . \tag{9}$$

Equation (5) implies that the u_i are unit vectors, and (8) says that $u_i^T u_j = u_{i+1}^T u_{j+1}$. In other words, the cyclic shift $u_0 \mapsto u_1 \mapsto \ldots u_{n-1} \mapsto u_0$ preserves the length of the u_i and all the angles between them, and hence there is an orthogonal matrix M such that $u_{i+1} = M u_i$, and hence $u_t = M^t u_0$ for $t = 0, 1, \ldots$. Thus $M^n u_i = u_i$ for all i, and hence $M^n = I$ (the identity matrix).

Up to this point, our arguments have been standard in semidefinite optimization. The next trick is to allow complex numbers and choose a basis so that M has a diagonal form

$$M = \begin{pmatrix} \varepsilon_1 & 0 & \ldots & 0 \\ 0 & \varepsilon_2 & \ldots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \ldots & \varepsilon_d \end{pmatrix}$$

where each ε_t is an n-th root of unity. The the objective function (9) can be written as

$$\frac{1}{8kn} \sum_{t=1}^{d} \sum_{A \in \mathcal{H}} \left| \sum_{j \in A} \varepsilon_t^j \right|^2 |u_{0t}|^2. \tag{10}$$

Now take any coordinate t, and let $\varepsilon_t = e^{2\pi a/n}$. By Dirichlet's Theorem, there is are integers $1 \leq q \leq 8k$ and p such that $|q(a/n) - p| < 1/(8k)$. This implies that for every

arithmetic progression A of difference q and length k, the complex numbers ε_t^j $(j \in A)$ point in almost the same direction: the maximum angle between them is less than $(k-1)(2\pi/(8k)) < \pi/4$. Hence

$$\left|\sum_{j\in A} \varepsilon_t^j\right|^2 > \frac{k^2}{2}.$$

Since there are n arithmetic progressions in $\mathcal{H}$ with this difference, we get

$$\sum_{A\in\mathcal{H}} \left|\sum_{j\in A} \varepsilon_t^j\right|^2 > \frac{k^2 n}{2},$$

and thus

$$\frac{1}{8kn}\sum_{t=1}^{d}\sum_{A\in\mathcal{H}}\left|\sum_{j\in A}\varepsilon_t^j\right|^2 |u_{0t}|^2 > \frac{k^2 n}{2}\frac{1}{8kn}\sum_{t=0}^{n-1}|u_{0t}|^2 = \frac{k}{16} > \frac{n^{1/2}}{49}$$

as claimed. □

References

[1] F. Alizadeh, Interior point methods in semidefinite programming with applications to combinatorial optimization, *SIAM J. Optim.* **5** (1995), 13–51.

[2] J. Beck: Roth's estimate on the discrepancy of integer sequences is nearly sharp, *Combinatorica* **1** (1981) 327–335.

[3] J. Beck and V.T. Sós: Discrepancy Theory, in: *Handbook of Combinatorics* (eds. R.L. Graham, M. Grötschel, L. Lovász), North-Holland, Amsterdam (1995), 1405–1446.

[4] K.F. Roth: Remark concerning integer sequences, *Acta Arith.* **35**, 257–260.

[5] J. Matoušek and J. Spencer, Discrepancy in arithmetic progressions, *J. Amer. Math. Soc.* **9** (1996) 195–204.

[6] L. Vandenberghe and S. Boyd: Semidefinite programming, in: *Math. Programming: State of the Art* (ed. J. R. Birge and K. G. Murty), Univ. of Michigan, 1994; revised verion *SIAM Rev.* **38** (1996), 49–95.