Algebraic Methods in Computer Science	5.11.14
Lecture: 2	
Lecturer: *Amir Shpilka*	Scribe: *Roi Tabach*

In this and the next lecture we will see several examples where we map a combinatorial problem to a question about polynomials, which is solved using dimension arguments. In particular we will see the following examples.

- Graham-Pollak Thm.
- Points with 2-distances
- Ray-Chahudhury-Wilson type theorems.
- Applications of R-W theorem.

Graham-Pollak Theorem

Let K_n be the complete graph on n vertices. We wish to cover all edges of K_n by disjoint, complete bi-partite graphs. I.e. every edge belongs to exactly one such complete bi-partite graph. The question we are interested in is what is the minimal number of bi-partite graphs in such a cover.

The trivial solution is to view each edge as a complete bi-partite graph. This gives a cover using $\binom{n}{2}$ bi-partite graphs. A better solution is to cover K_n with $n-1$ graphs, where the k'th graph contains all edges connecting the k'th vertex to vertices $\{k_1, \ldots, n\}$. Clearly this gives a cover with $n-1$ complete bi-partite graphs. The Graham-Pollak theorem says that any cover must contain at least that many complete bi-partite graphs. We note that the disjointness requirement is essential. Without it we could cover K_n with $\log n$ complete bi-partite graphs. We interpret each $k \in [n]$ as a $\log n$ long binary string. The i'th graph connects all vertices whose i'th coordinate is 0 with those whose i'th coordinate is 1.

Theorem 1 (Graham-Pollak)**.** *The minimal number of disjoint bi-partite graphs needed to cover K_n is $n-1$.*

Proof. Let $B_1, \ldots, B_m$ be disjoint complete bi-partite graphs that cover K_n. Denote the set of vertices of B_i with (L_i, R_i). It follows that the adjacency matrix of B_i satisfies $B_{i,j} = 1$ if and only if $i \in L_i$ and $j \in R_i$, or vice versa. I.e., it is composed of two rank-1 matrices and its rank equals 2. Because of the disjointness property, we get that summation of all matrices equals the adjacency matrix of K_n. Using rank arguments this immediately gives a lower bound of $n/2$ on m.

We shall now see how to improve that bound. We associate a polynomial with each B_i:

$$B_i \mapsto P_i(x_1, ..., x_n) \triangleq \left(\sum_{k \in L_i} x_k\right) \cdot \left(\sum_{j \in R_i} x_j\right) = \sum_{k \sim j \in E(B_i)} x_k x_j. \tag{1}$$

Hence,

$$\sum_{i=1}^{m} P_{B_i} = \sum_{j<k} x_j x_k = \frac{1}{2}\vec{x}^t \cdot \left(\sum_{i=1}^{m} B_i\right) \cdot \vec{x} = \frac{1}{2}\vec{x}^t \cdot A_{K_n} \cdot \vec{x}, \tag{2}$$

where $\vec{x} = (x_1, ..., x_n)$ and A_{K_n} is the adjacency matrix of K_n. We note that $A_{K_n} = J - I$, where J is the all 1 matrix and I is the identity matrix. Therefore,

$$\frac{1}{2}\vec{x}^t \cdot A_{K_n} \cdot \vec{x} = \frac{1}{2}\left(\left(\sum_{i=1}^{n} x_i\right)^2 - \sum_{i=1}^{n} x_i^2\right). \tag{3}$$

Combining equations (1),(2) and (3) we get

$$2\sum_{i=1}^{m}\left(\sum_{k \in L_i} x_k\right)\left(\sum_{j \in R_i} x_j\right) = \left(\sum_{i=1}^{n} x_i\right)^2 - \sum_{i=1}^{n} x_i^2.$$

Thus,

$$\sum_{i=1}^{n} x_i^2 = -2\sum_{i=1}^{m}\left(\sum_{k \in L_i} x_k\right)\left(\sum_{j \in R_i} x_j\right) - \left(\sum_{i=1}^{n} x_i\right)^2.$$

In other words, we found a way to express $\sum_{i=1}^{n} x_i^2$ as a sum of $m+1$ many products of linear funcitons $\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{m+1} \ell_i \cdot \ell_i'$, where ℓ_i, ℓ_i' are linear forms in $(x_1, ..., x_n)$. It is clear that for every $\vec{x} \neq \vec{0}$, the LHS is greater then zero. on the other hand, if $m+1 < n$ then we can find $\vec{x} \neq \vec{0}$ such that for every i, $\ell_i(\vec{x}) = 0$. Indeed, this is possible since we have $m+1 < n$ linear forms in n variables. In particular, for such a nonzero solution, $\sum_{i=1}^{m+1} \ell_i \cdot \ell_i' = 0$, and hence the RHS vanishes, in contradiction. It therefore follows that $m+1 \geq n$ and thus, $m \geq n-1$ as claimed. □

Points with 2-distances:

Let $v_1, \ldots, v_m$ be some m points in R^n. We are interested in an upper bound for m, given certain limitations on $\{v_i\}_i$.

Question. *Given m points in R, where the distance between each two points is the same, how large can m be?*

The question will be discussed in the next HW. We will only state here that there exists a construction where: $m \geq n+1$, and that $n+1$ is an upper bound. The construction is based on the triangle in R^2 or the Tetrahedron in R^3.

Now we will look into the more complex problem: 2-distance sets. Given m points in R so that for every pair, it's distance is either δ_1 or δ_2, we would like to know what is the upper bound on m.

First we will try to make a construction. Let $H : \{0,1\}^n \mapsto N$ be the Hamming-weight function, meaning $H(x) = \sum_i x_i$. If we look in the set of vectors in R^n who's Hamming weight is 2 (Let it be denoted S), we will get a 2-distance set. The reason it is a 2-distance set is that for every pair $x, y \in S$, either x and y share one or zero indices where they aren't zero. On the former case, we get something of the form:

$$\|x - y\| = \left\| \begin{bmatrix} \pm 1 \\ \mp 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\|$$

And on the latter case we get something of the form:

$$\|x - y\| = \left\| \begin{bmatrix} \pm 1 \\ \pm 1 \\ \mp 1 \\ \mp 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\|$$

All together, we got that S is a 2-distance set.

We have shown a construction of a 2-distance set with $\|S\| = \binom{n}{2}$. Now we will try giving an upper bound for m.

Theorem 2 (Size of a 2-distance set)**.** *Let S be a 2-distances set. So $|S| \leq \binom{n}{2} + (n+1)$*

Proof. Our strategy will again be finding a mapping to a polynomial space, then finding upper bounds on it's dimension. We would like to map each point $v_i \in S$ to a polynom $p_i \in R[x_1, \ldots, x_n]$, so that each p_i will vanish on every other point in S. This can be achieved by choosing

$$p_i(\vec{x}) := (\|\vec{x} - \vec{v_i}\| - \delta_1) \cdot (\|\vec{x} - \vec{v_i}\| - \delta_2) \tag{4}$$

And indeed for the polynom we definded it holds that

$$p_i(v_j) = \begin{cases} \delta_1 \cdot \delta_2 & i = j \\ 0 & i \neq j \end{cases} \tag{5}$$

as wanted. We will show now that these polynoms are linear independent, and that they are all in the span of a small set of polynoms. This will give us an upper bound on the size of S.

Linear independence follows from the fact that if $j \neq i$ then $p_i(v_j) = 0$. If we look at a linear combination of the form

$$f := \sum_{i=1}^{m} \alpha_i \cdot p_i$$

where not all α_is are zero, and $f \equiv 0$, it is easy to show that, for example, $\alpha_1 = 0$:

$$0 = f(v_1) = \sum_i \alpha_i \cdot p_i(v_1) \overset{(4)}{=} \alpha_1 \cdot p_1(v_1) \overset{(5)}{=} \alpha_1 \cdot (\delta_1 \cdot \delta_2)$$

$$\Rightarrow \alpha_1 = 0$$

Now we will look further into (4).

$$p_i(x_1, \ldots, x_n) = \left(\left(\sum_{j=1}^{n} (x_j - v_{ij})^2 \right) - \delta_1 \right) \cdot \left(\left(\sum_{j=1}^{n} (x_j - v_{ij})^2 \right) - \delta_2 \right) =$$

$$= \left[\left(\sum_j {x_i}^2 \right) - 2 \left(\sum_j x_j \cdot v_{ij} \right) + \left(\sum_j v_{ij}^2 - \delta_1 \right) \right] \cdot$$

$$\cdot \left[\left(\sum_j {x_i}^2 \right) - 2 \left(\sum_j x_j \cdot v_{ij} \right) + \left(\sum_j v_{ij}^2 - \delta_2 \right) \right] =$$

$$= \left(\sum x_i^2 \right)^2 + \sum_{i=1}^{n} c_i x_i \left(\sum_j x_j^2 \right) + \sum d_{ij} x_i x_j + \sum e_i x_i + f$$

for some c_i, d_{ij}, e_i, f. We got that for all of our p_is:

$$p_i \in \text{Span} \left(\left\{ \left(\sum_i x_i^2 \right)^2 \right\} \cup \left\{ x_i \cdot \left(\sum_j x_j^2 \right) \right\}_{i=1}^{n} \cup \{x_i \cdot x_j\}_{i,j=1}^{n} \cup \{x_i\}_{i=1}^{n} \cup \{1\} \right)$$

Which we will denote V. And since V is defined as the Span of $1 + n + \binom{n+1}{2} + n + 1$ vectors, we get

$$m \leq \dim(V) \leq 1 + n + \binom{n+1}{2} + n + 1 = \binom{n+2}{2} + (n+1)$$

□