MAT 307: Combinatorics

Lecture 16: Linear algebra - continued

Instructor: Jacob Fox

1 Spaces of polynomials

In the previous lecture, we considered objects represented by 0/1 vectors. A vector $\mathbf{1}_A$ corresponding to a set A can be also viewed as a *linear form* $f(\vec{x}) = \sum_{i \in A} x_i$. (All our proofs could be written equivalently in the language of linearly independent linear forms.) More generally, however, we can represent objects by *polynomials* $f(x)$. Polynomials of a certain degree form a vector space and we can still apply the same arguments about dimension and linear independence. This gives us more flexibility and power compared to the linear case.

2 Two-distance sets

Consider a set of points $A \subset R^n$. If all the pairwise distances between points in A are equal, then these are the vertices of a simplex. The number of such points can be at most $n+1$.

What if we relax the condition and require that there are *two possible distances* c, d, so that any pairwise distance is either c or d? Such a set is called a two-distance set.

Exercise. Construct a two-distance set in R^n with $\binom{n}{2}$ points.

Theorem 1. *Any two-distance set in R^n has at most $\frac{1}{2}(n+1)(n+4)$ points.*

Proof. Let $A \subset R^n$ be a two-distance set. For each point $a \in A$, we define a polynomial on R^n,

$$f_a(x) = \left(||x-a||^2 - c^2\right)\left(||x-a||^2 - d^2\right).$$

Here, $||x||^2 = \sum_i x_i^2$ denotes the square of the euclidean norm. Let's prove that the polynomials $f_a(x)$ are linearly independent. Suppose that $\sum_{a \in A} \alpha_a f_a(x)$ is identically zero. Then plug in $x = b$ for some point $b \in A$. We have $f_a(b) = 0$ for any $b \neq a$, because $||a-b||$ is either c or d. So we have $0 = \sum_{a \in A} \alpha_a f_a(b) = \alpha_b f_b(b) = \alpha_b c^2 d^2$. Since $cd \neq 0$, this implies $\alpha_b = 0$, for any $b \in A$. This shows that the polynomials $f_a(x)$ are linearly independent.

Finally, we want to bound the dimension of the vector space containing our polynomials. By expanding the euclidean norms, it can be seen that each $f_a(x)$ can be expressed as a linear combination of the following polynomials:

$$V = \left\{ (\sum_{i=1}^n x_i^2)^2, x_j \sum_{i=1}^n x_i^2, x_i x_j, x_i, 1 \mid i, j \in [n] \right\}.$$

The number of generators here is $1 + n + \frac{1}{2}n(n+1) + n + 1 = \frac{1}{2}(n+1)(n+4)$. Therefore, the polynomials $f_i(x)$ reside in a vector space of dimension $\frac{1}{2}(n+1)(n+4)$. □

3 Sets with few possible intersection sizes

Here we discuss a generalization of Fisher's inequality. Consider a family of sets $\mathcal{F} \subseteq 2^{[n]}$ and let $L \subset \{0, 1, \ldots, n\}$. We say that $\mathcal{F}$ is L-intersecting if $|A \cap B| \in L$ for any distinct $A, B \in \mathcal{F}$. Fisher's inequality says that if $|L| = 1$ then $|\mathcal{F}| \leq n$. Frankl and Wilson proved the following generalization in 1981.

Theorem 2. *If $\mathcal{F}$ is an L-intersecting family of subsets of $[n]$, then*

$$|\mathcal{F}| \leq \sum_{k=0}^{|L|} \binom{n}{k}.$$

Note that the family of all subsets of size at most ℓ is L-intersecting, for $L = \{0, 1, \ldots, \ell - 1\}$, so this bound is best possible.

Proof. Let $\mathcal{F} \subset 2^{[n]}$ and $|L| = s$. For any $A \neq B \in \mathcal{F}$, $|A \cap B| \in L$. We define a polynomial on R^n for each $A \in \mathcal{F}$:

$$f_A(x) = \prod_{\ell \in L : \ell < |A|} \left(\sum_{e \in A} x_e - \ell \right).$$

Observe that for any $B \in \mathcal{F}, A \not\subset B$, if we plug in the indicator vector $\mathbf{1}_B$, we get

$$f_A(\mathbf{1}_B) = \prod_{\ell \in L : \ell < |A|} (|A \cap B| - \ell) = 0$$

because $|A \cap B| = \ell < |A|$ for some $\ell \in L$. On the other hand,

$$f_A(\mathbf{1}_A) = \prod_{\ell \in L : \ell < |A|} (|A| - \ell) > 0.$$

By an argument similar to the one we used before, the polynomials $\{f_A(x) : A \in L\}$ are independent.

It remains to compute the dimension of the space containing all these polynomials. A trick that helps reduce the dimension is that we are only using 0/1 vectors here. Thus, we can replace all higher powers x_i^k by x_i itself; this does not change the linear independence property. Then, the polynomials are generated by all *monomials* $\prod_{i \in I} x_i$, where $|I| \leq s$. The number of such monomials is exactly $\sum_{k=0}^{s} \binom{n}{k}$, as required. □

Using essentially the same argument, we can also prove the following modular version of the theorem.

Theorem 3. *Let p be prime and $L \subset Z_p$. Assume $\mathcal{F} \subset 2^{[n]}$ is a family of sets such that*

- $|A| \notin L \pmod{p}$ *for any $A \in \mathcal{F}$.*
- $|A \cap B| \in L \pmod{p}$ *for any distinct $A, B \in \mathcal{F}$.*

Then

$$|\mathcal{F}| \leq \sum_{k=0}^{|L|} \binom{n}{k}.$$

Proof. Let $\mathcal{F} \subset 2^{[n]}$ and $L \subset Z_p$. In the following, all operations are mod p. For any $A, B \in \mathcal{F}$ distinct, $|A \cap B| \in L$. We define a polynomial on Z_p^n for each $A \in \mathcal{F}$:

$$f_A(x) = \prod_{\ell \in L} \left(\sum_{e \in A} x_e - \ell \right).$$

Observe that for any $B \in \mathcal{F}, B \neq A$, if we plug in the indicator vector $\mathbf{1}_B$, we get

$$f_A(\mathbf{1}_B) = \prod_{\ell \in L} (|A \cap B| - \ell) = 0$$

because $|A \cap B| \in L$. On the other hand,

$$f_A(\mathbf{1}_A) = \prod_{\ell \in L} (|A| - \ell) \neq 0.$$

Again, we replace each $f_A(x)$ by $\tilde{f}_A(x)$ where each factor x_i^k is replaced by x_i. Since we are only substituting 0/1 values, this does not affect the properties above. Hence, the polynomials $\tilde{f}_A(x)$ are linearly independent. They are generated by all monomials $\prod_{i \in I} x_i$, where $|I| \leq |L|$. The number of such monomials is exactly $\sum_{k=0}^{|L|} \binom{n}{k}$, as required. □