MAT 307: Combinatorics

Lecture 15: Applications of linear algebra

Instructor: Jacob Fox

1 Linear algebra in combinatorics

After seeing how probability and topology can be useful in combinatorics, we are going to exploit an even more basic area of mathematics - linear algebra. While the probabilistic method is usually useful to construct examples and prove lower bounds, a common application of linear algebra is to prove an upper bound, where we show that a collection of objects satisfying certain properties cannot be too large. A typical argument to prove this is that we replace the objects by vectors in a linear space of a certain dimension, and we show that the respective vectors are linearly independent. Hence, there cannot be more of them than the dimension of the space.

2 Even and odd towns

We start with the following classical example. Suppose there is a town where residents love forming different clubs. To limit the number of possible clubs, the town council establishes the following rules:

Even town.

- Every club must have an even number of members.
- Two clubs must not have exactly the same members.
- Every two clubs must share an even number of members.

How many clubs can be formed in such a town? We leave it as an exercise to the reader that there can be as many as $2^{n/2}$ clubs (for an even number of residents n). Thus, the town council reconvened and invited a mathematician to help with this problem. The mathematician suggested the following modified rules.

Odd/even town.

- Every club must have an odd number of members.
- Every two clubs must share an even number of members.

The residents soon found out that they were able to form only n clubs under these rules, for example by each resident forming a separate club. In fact, the mathematician was able to prove that more than n clubs are impossible to form.

Theorem 1. *Let $\mathcal{F} \subset 2^{[n]}$ be such that $|A|$ is odd for every $A \in \mathcal{F}$ and $|A \cap B|$ is even for every distinct $A, B \in \mathcal{F}$. Then $|\mathcal{F}| \leq n$.*

Proof. Consider the vector space Z_2^n, where $Z_2 = \{0, 1\}$ is a finite field with operations modulo 2. Represent each club $A \in \mathcal{F}$ by its *incidence vector* $\mathbf{1}_A \in Z_2^n$, where a component i is equal to 1 exactly if $i \in A$. We claim that these vectors are linearly independent.

Suppose that $z = \sum_{A \in \mathcal{F}} \alpha_A \mathbf{1}_A = 0$. Fix any $B \in \mathcal{F}$. We consider the inner product $z \cdot \mathbf{1}_B = 0$. By the linearity of the inner product and the odd-town rules,

$$0 = z \cdot \mathbf{1}_B = \sum_{A \in \mathcal{F}} \alpha_A (\mathbf{1}_A \cdot \mathbf{1}_B) = \alpha_B,$$

all operations over Z_2. We conclude that $\alpha_B = 0$ for all $B \in \mathcal{F}$. Therefore, the vectors $\{\mathbf{1}_A : A \in \mathcal{F}\}$ are linearly independent and their number cannot be more than n, the dimension of Z_2^n. $\square$

An alternative variant is an even/odd town, where the rules are reversed.

Even/odd town.

- Every club must have an even number of members.
- Every two clubs must share an odd number of members.

Exercise. By a simple reduction, any even/odd town with n residents and m clubs can be converted to an odd/even town with $n + 1$ residents and m clubs. This shows that there is no even/odd town with n residents and $n + 2$ clubs.

Theorem 2. *Let $\mathcal{F} \subset 2^{[n]}$ be such that $|A|$ is even for every $A \in \mathcal{F}$ and $|A \cap B|$ is odd for every distinct $A, B \in \mathcal{F}$. Then $|\mathcal{F}| \leq n$.*

Proof. Assume that $|\mathcal{F}| = n + 1$. All calculations in the following are taken mod 2. The $n + 1$ vectors $\{\mathbf{1}_A : A \in \mathcal{F}\}$ must be linearly dependent, i.e. $\sum_{A \in \mathcal{F}} \alpha_A \mathbf{1}_A = 0$ for some non-trivial linear combination. Note that $\mathbf{1}_A \cdot \mathbf{1}_B = 1$ for distinct $A, B \in \mathcal{F}$ and $\mathbf{1}_A \cdot \mathbf{1}_A = 0$ for any $A \in \mathcal{F}$. Therefore,

$$\mathbf{1}_B \cdot \sum_{A \in \mathcal{F}} \alpha_A \mathbf{1}_A = \sum_{A \in \mathcal{F} : A \neq B} \alpha_A = 0.$$

By subtracting these expressions for $B, B' \in \mathcal{F}$, we get $\alpha_B = \alpha_{B'}$. This means that all the coefficients α_B are equal and in fact equal to 1 (otherwise the linear combination is trivial).

We have proved that for any even/odd town with $n + 1$ clubs, $\sum_{A \in \mathcal{F}} \mathbf{1}_A = 0$. Moreover, for any $B \in \mathcal{F}$, $0 = \mathbf{1}_B \cdot \sum_{A \in \mathcal{F}} \mathbf{1}_A = |\mathcal{F}| - 1 = n$ which means that $|\mathcal{F}|$ is odd and n is even.

Now we use the following duality. Replace each set $A \in \mathcal{F}$ by its complement $\bar{A}$. Since the total number of elements n is even, we get $|\bar{A}|$ even and $|\bar{A} \cap \bar{B}|$ odd for any distinct $A, B \in \mathcal{F}$. This means that the $n + 1$ complementary clubs $\bar{A}$ should also form an even/odd town and therefore again, we should have $\sum_{A \in \mathcal{F}} \mathbf{1}_{\bar{A}} = 0$. But then,

$$0 = \sum_{A \in \mathcal{F}} \mathbf{1}_A + \sum_{A \in \mathcal{F}} \mathbf{1}_{\bar{A}} = |\mathcal{F}| \mathbf{1}$$

where $\mathbf{1}$ is the all-ones vector. This implies that $|\mathcal{F}|$ is even, contradicting our previous conclusion that $|\mathcal{F}|$ is odd. $\square$

3 Fisher's inequality

A slight modification of the odd-town rules is that every two clubs share a fixed number of members k (there is no condition here on the size of each club). We get a similar result here, which is known as Fisher's inequality.

Theorem 3 (Fisher's inequality)**.** *Suppose that $\mathcal{F} \subset 2^{[n]}$ is a family of nonempty clubs such that for some fixed k, $|A \cap B| = k$ for every distinct $A, B \in \mathcal{F}$. Then $|\mathcal{F}| \le n$.*

Proof. Again, we consider the incidence vectors $\{\mathbf{1}_A : A \in \mathcal{F}\}$, this time as vectors in the real vector space R^n. We have $\mathbf{1}_A \cdot \mathbf{1}_B = k$ for all $A \neq B$ in $\mathcal{F}$. Suppose that $\sum_{A\in\mathcal{F}} \alpha_A \mathbf{1}_A = 0$. Then

$$\begin{aligned} 0 &= ||\sum_{A\in\mathcal{F}} \alpha_A \mathbf{1}_A||^2 = \left(\sum_{A\in\mathcal{F}} \alpha_A \mathbf{1}_A\right) \cdot \left(\sum_{B\in\mathcal{F}} \alpha_B \mathbf{1}_B\right) \\ &= \sum_{A\in\mathcal{F}} \alpha_A^2 |A| + \sum_{A\neq B\in\mathcal{F}} \alpha_A \alpha_B k = k\left(\sum_{A\in\mathcal{F}} \alpha_A\right)^2 + \sum_{A\in\mathcal{F}} \alpha_A^2(|A| - k). \end{aligned}$$

Note that $|A| \ge k$, and at most one set A^* can actually have size k. Therefore, the contributions to the last expression are all nonnegative and $\alpha_A = 0$ except for $|A^*| = k$. But then, $\sum_{A\in\mathcal{F}} \alpha_A = \alpha_{A^*}$ and this must be zero as well.

We have proved that the vectors $\{\mathbf{1}_A : A \in \mathcal{F}\}$ are linearly independent in R^n and hence their number can be at most n. $\square$

Fisher's inequality is related to the study of *designs*, set systems with special intersection patterns. We show here how such a system can be used to construct a graph on n vertices, which does not have any clique or independent set of size $\omega(n^{1/3})$. Recall that in a random graph, there are no cliques or independent sets significantly larger than $\log n$; so this explicit construction is very weak in comparison.

Lemma 1. *For a fixed k, let G be a graph whose vertices are triples $T \in \binom{[k]}{3}$ and $\{A, B\}$ is an edge if $|A \cap B| = 1$. Then G does not contain any clique or independent set of size more than k.*

Proof. Suppose Q is a clique in G. This means we have a set of triples on $[k]$ where each pair intersects in exactly one element. By Fisher's inequality, the number of such triples can be at most k.

Suppose S is an independent set in G. This is a set of triples on $[k]$ where each pair intersects in an even number of elements, either 0 or 2. By the odd-town theorem, the number of such triples is again at most k. $\square$

Another application of Fisher's inequality is the following.

Lemma 2. *Suppose P is a set of n points in the plane, not all on one line. Then pairs of points from P define at least n distinct lines.*

Proof. Let L be the set of lines defined by pairs of points from P. For each point $x_i \in P$, let $A_i \subseteq L$ be the set of lines containing x_i. We have $|A_i| \ge 2$, otherwise all points lie on the same line. Also, A_i is different for each point; the same set of at least 2 lines would define the same point. Moreover, any two points share exactly one line, i.e. $|A_i \cap A_{i'}| = 1$ for any $i \neq i'$. By Fisher's inequality, we get $|P| \le |L|$. $\square$