Dimensional Argument

Yanheng Wang

Suppose we want to study a collection $\mathcal{A}$ defined via certain constraints. The *dimensional argument* could help us bound the size $|\mathcal{A}|$, a first indicator of the behaviour of $\mathcal{A}$. To apply the argument, we injectively map $\mathcal{A} \ni a \mapsto \sigma_a \in V$ where V is a linear space. Then we exploit the constraints to show that $\{\sigma_a\}_{a \in \mathcal{A}}$ are linearly independent, thus concluding $|\mathcal{A}| = |\{\sigma_a\}_{a \in \mathcal{A}}| \leqslant \dim(V)$.

The key to dimensional argument is designing a suitable injection $\sigma \colon \mathcal{A} \to V$. Typically we work backwards:

- Explore the constraints and model them algebraically;
- Design the σ_a's so that we *could* prove their linear independence by the algebraic properties we collected earlier.
- Bound the dimension of the linear space $V := \operatorname{span}\{\sigma_a : a \in \mathcal{A}\}$.

1 Town theorems

Odd-even town

Let $\mathcal{A} \subseteq 2^{[n]}$ be a collection of subsets of $[n]$. For all distinct $A, B \in \mathcal{A}$, we require that $|A|$ is odd while $|A \cap B|$ is even. How large can $m := |\mathcal{A}|$ be?

- We may encode a set $A \subseteq [n]$ as a binary vector (termed the *charactersitic vector*) $\chi_A \in \{0,1\}^n$ by putting a 1 at coordinate i iff $i \in A$. Then the constraints simply say $\langle \chi_A, \chi_A \rangle = 1 \pmod 2$ and $\langle \chi_A, \chi_B \rangle = 0 \ (\mathrm{mod}\ 2)$.
- If we interpret the characteristic vectors as vectors in $\mathbb{F}_2^n$ then they are orthonormal! From basic linear algebra we know orthonormal vectors are linear independent.
- With this in mind, we simply design $\sigma \colon \mathcal{A} \to \mathbb{F}_2^n, A \mapsto \chi_A$ and conclude $m \leqslant \dim(\mathbb{F}_2^n) = n$.

Even-odd town

In fact, we may switch the parities (i.e. constraining $|A|$ even and $|A \cap B|$ odd) and derive the same bound. One lazy proof is to append a dummy element 0 to all the sets, thus alternating the parities back. (A small modification is necessary in the argument: the characteristic vectors are now living in $\{1\} \times \mathbb{F}_2^n$, a linear space of dimension n still.)

But it is instructive to present a direct proof. Again we map sets to characteristic vectors in $\mathbb{F}_2^n$ and find $\langle \chi_A, \chi_A \rangle = 0$ and $\langle \chi_A, \chi_B \rangle = 1$. This time, showing linear independence is less straightforward. To do so, we place the vectors in a matrix $M \in \mathbb{F}_2^{n \times m}$ and try to show $\operatorname{rank}(M) = m$. Note

$$M^{\mathrm{T}} M = \begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 0 \end{pmatrix} = J - I.$$

Recall that J has eigenvalues $m, 0, \ldots, 0$, so $J - I$ has eigenvalues $m-1, -1, \ldots, -1$. Hence

$$\operatorname{rank}(M) = \operatorname{rank}(M^{\mathrm{T}} M) = \operatorname{rank}(J - I) = m.$$

Therefore the vectors are indeed independent, implying $m \leqslant n$. (Alternatively, we could compute the $\det(M^{\mathrm{T}} M) \neq 0$ to conclude the same.)

Even-even and odd-odd towns?

What if we require both $|A|$ and $|A \cap B|$ even (or both odd)? Then our previous argument breaks since the charactersitic vectors are deeply dependent. Actually in this scenario, m could be as large as $2^{n/2}$ – consider partitioning $[n]$ into $n/2$ pairs, say $p_i := \{2i-1, 2i\}$ for $i \in [n/2]$, and letting $\mathcal{A} := \{\bigcup_{i \in I} p_i : I \subseteq [n/2]\}$.

Via very basic properties of orthogonal complements, one can show a tight upper bound $m \leqslant 2^{n/2}$. The proof is easy but off-topic, so we do not pursue it here.

Generalisations

We list several generalisations of town theorems and give hints on how to approach them:

- For prime p, require $|A| \neq 0 \pmod{p}$ and $|A \cap B| = 0 \pmod{p}$.
  - Result: $m \leqslant n$.
  - Method: work in $\mathbb{F}_p^n$.
- For $q = p^k$ where p is prime, require $|A| \neq 0 \pmod{q}$ and $|A \cap B| = 0 \pmod{q}$.
  - Result: $m \leqslant n$.
  - Method: work in $\mathbb{Q}^n$ and use some number theory to derive linear independence. We cannot work in $\mathbb{F}_q^n$ because it is not $\mathbb{Z}_q$! More to the point, $\operatorname{char}(\mathbb{F}_q^n) = p$, so a zero sum only implies "mod p" rather than "mod q".
- For $q = \prod_{i=1}^{r} p_i^{k_i}$ where p_i's are distinct primes, require $|A| \neq 0 \pmod{q}$ and $|A \cap B| = 0 \pmod{q}$.
  - Result: $m \leqslant rn$.
  - Method: by induction on r; partition $\mathcal{A}$ depending on $|A| = 0 \pmod{p_r^{k_r}}$ or not; then apply induction hypothesis and the previous part.
- We may also generalise the number of intersecting sets. For example, require $|A|$ odd but $|A \cap B \cap C|$ even.
  - Result: $m \leqslant n\,(n+1)$.
  - Method: build a graph $G := (\mathcal{A}, \{\{A, B\} : |A \cap B| \text{ odd}\})$. Using known results and the property, show $\Delta(G) \leqslant n$. Then $\chi(G) \leqslant n+1$, and $\alpha(G) \geqslant \frac{m}{n+1}$. But on the other hand any independent set corresponds to odd-even town, hence $\alpha(G) \leqslant n$.

2 Fisher's inequality

Again consider a collection $\mathcal{A}=\{A_1,\ldots,A_m\}\subseteq 2^{[n]}$. This time we require that $|A_i\cap A_j|=\lambda$ for all distinct i,j.

First we rule out an uninteresting case. If there is a set A_i of size exactly λ, then all other sets must be supersets of A_i, and their pairwise intersections give A_i. The picture looks like a "sunflower" where A_i locates at its core and the "petals" are disjoint. So the total number of sets, m, is at most n.

Now we may assume $|A_i|\geqslant\lambda+1$ for all i. We copy the argument in previous section. Here we work in $\mathbb{R}^n$ and arrange the characteristic vectors into matrix $M\in\mathbb{R}^{m\times m}$ such that

$$M^{\mathrm{T}}M=\begin{pmatrix} |A_1| & \lambda & \cdots & \lambda \\ \lambda & |A_2| & \cdots & \lambda \\ \vdots & \vdots & \ddots & \vdots \\ \lambda & \lambda & \cdots & |A_m| \end{pmatrix}.$$

We calculate the determinant and find it $\left(1+\sum_{i=1}^m \frac{\lambda}{|A_i|-\lambda}\right)\prod_{i=1}^m (|A_i|-\lambda)\neq 0$, thus proving $m\leqslant n$.

Next we present another approach which maps sets to polynomials rather than characteristic vectors. For each $A\in\mathcal{A}$ we associate a (linear) polynomial

$$p_A(x):=\langle x,\chi_A\rangle-\lambda.$$

Note that $p_A(\chi_A)=|A|-\lambda\geqslant 1$ while $p_A(\chi_B)=0$ for all $B\in\mathcal{A}\setminus\{A\}$. Therefore the polynomials are linearly independent. (Prove by definition of linear independence; or decompose the polynomial to a standard basis and compute rank.) On the other hand, the dimension of the space is at most n, since every such polynomial is a linear combination of monomials $x_1,\ldots,x_n$. This implies $m\leqslant n$.

3 k-distance set

Let $P\subseteq\mathbb{R}^d$ be a point set in Euclidean space. We call it a k-distance set if every pair $\{p,q\}\in\binom{P}{2}$ satisfies $\|p-q\|^2\in\{\delta_1,\ldots,\delta_k\}$, where $\delta_1,\ldots,\delta_k$ are distinct positive numbers. How large can P be?

For the case $k=1$, it is quite intuitive that $|P|\leqslant d+1$, with the extremal example being the regular simplex. One argument goes as follows. Without loss of generality assume $\mathbf{0}\in P$ and $\delta_1:=1$. We consider the remaining points $P':=P\setminus\{\mathbf{0}\}$. By equidistant property we see $\|p\|=1$ and $1=\|p-q\|^2=2-2\langle p,q\rangle$ for all $p,q\in P'$. So we could again show that the points are linearly independent, so $|P'|\leqslant n$. (Remark: in this proof we use the trivial mapping $\sigma:=\mathrm{id}$.)

For $k\geqslant 2$ the previous proof does not generalise. The natural step is to try some stronger and more flexible object, the polynomials. To this end, we design, for each $p\in P$, a polynomial

$$f_p(x):=\prod_{i=1}^k (\|x-p\|^2-\delta_i)$$

which nicely captures the k-distant property. Namely, $f_p(p) \neq 0$ yet $f_p(q) = 0$ for any distinct $p, q \in P$. Hence the polynomials are linearly independent. It remains to bound the dimension of this polynomial space. As a very crude estimate, the dimension is at most $(d+1)^{2k}$ because $\deg(f_p) \leqslant 2k$ and there are at most that many (in fact, $\binom{d+1+2k}{2k}$ to be precise) monomials to choose.

But we could be more refined. We expand the definition by

$$f_p(x) = \prod_{i=1}^{k} \left(\|x\|^2 - 2\sum_{i=1}^{d} p_j x_j + (\|p\|^2 - \delta_i) \right)$$

Observe that the polynomials are spanned by the monomials

$$\left\{ x_1^{\beta_1} \cdots x_d^{\beta_d} \|x\|^{2\beta_0} \,\middle|\, \sum_{j=0}^{d} \beta_j \leqslant k \right\}.$$

Basically, the types of monomials are quite restricted. It remains to count the number, but this is a variant of "balls and bins" model. We introduce a slack variable to make the "$\leqslant$" a "$=$". Then we count the number of possibilities to distribute k balls into $d+2$ bins – which can be realised by choosing $d+1$ positions for delimiters from $d+1+k$ possible slots. So the count is $\binom{d+1+k}{d+1} = \binom{d+1+k}{k}$.

4 L Family

Suppose $\mathcal{A} \subseteq 2^{[n]}$ and let p be a prime.

- We call it an L-mod-p family if $|A \cap B| \in L \pmod{p}$ and $|A| \notin L \pmod{p}$ for all distinct $A, B \in \mathcal{A}$.
- We call it an L family if $|A \cap B| \in L$ for all distinct $A, B \in \mathcal{A}$.
- We call it r-uniform if every set $A \in \mathcal{A}$ has the same cardinality $|A| = r$.

Given n and $k := |L|$, how large can an L (resp. L-mod-p) family be?

This is a general framework which models both town theorem and Fisher's inequality. Their conditions can be rephrased as "$\{0\}$-mod-2 family" and "$\{\lambda\}$ family", respectively.

Let's look at L-mod-p family first. It's no longer a good idea to map to characteristic vectors because we don't have enough information for proving independence. So we try mapping to polynomials in a way that they are easily seen independent. For each A define polynomial $f_A \colon \mathbb{Z}_p^n \to \mathbb{Z}_p$ by

$$f_A(x) := \prod_{l \in L} (\langle \chi_A, x \rangle - l).$$

Observe that $f_A(\chi_A) \neq 0$ and $f_A(\chi_B) = 0$ for all $B \in \mathcal{A} \setminus \{A\}$. So the polynomials are linearly independent. It remains to bound the dimension of this polynomial space.

To this end, we note that the polynomials have degrees at most $k = |L|$, so they are of course spanned by all $\binom{n+k}{k}$ monomials. This gives a (quite trivial) upper bound on the dimension. But we can do better by exploiting the simple fact $1^t = 1$ and $0^t = 0$. Namely, we could replace any high-order term in $p_A(x)$ by a multilinear term, for instance replacing $x_1^3 x_4^2 x_5$ with $x_1 x_4 x_5$. This operation will change the polynomial, but it preserves the values on characteristic vectors! Hence the resulting polynomials are still linearly independent; furthermore they are spanned by all *multilinear* polynomials on n variables of degree at most k. Now we derive a bound $\sum_{i=0}^{k} \binom{n}{i}$, tighter than our previous $\binom{n+k}{k}$ especially when k is large.

Now we move on to L family. The idea is the same but with some twist because it is now possible that $|A| \in L$. Still, we may enforce linear independence by "upper triangular form" instead of the "diagonal form". To be specific, define $f_A \colon \mathbb{R}^n \to \mathbb{R}$ by

$$f_A(x) := \prod_{\substack{l \in L \\ l < |A|}} (\langle \chi_A, x \rangle - l).$$

Still $f_A(\chi_A) \neq 0$. Besides, $f_A(\chi_B) = 0$ whenever $|A \cap B| < |A|$, or equivalently $A \not\subseteq B$. We claim that the polynomials are linearly independent. Suppose for the sake of contradiction that $\sum_{A \in \mathcal{A}} \alpha_A \cdot f_A(x) \equiv 0$ for some non-trivial coefficients. Then let $B \in \mathcal{A}$ be the inclusion-minimal set with $\alpha_A \neq 0$. We evaluate the equation at $x := \chi_A$ and find everything except $\alpha_A \cdot f_A$ go away (since all $A \subset B$ have zero coefficient, and all $A \not\subseteq B$ have zero f_A value). So we must conclude $\alpha_A = 0$, a contradiction.

By the same line of argument as in previous proof, we derive an upper bound $\sum_{i=0}^{k} \binom{n}{i}$.

Next, we examine r-uniform L family and see how the uniformity condition sharpens our bound. The key ingredient is to "squeeze in" more polynomials besides $f_A(x)$ yet maintaining linear independence and the dimension. We define, for each $I \subseteq [n] \colon |I| \leqslant k - 1$, a polynomial $h_I \colon \mathbb{R}^n \to \mathbb{R}$ by

$$h_I(x) := (\langle x, \mathbf{1} \rangle - r) \cdot \prod_{i \in I} x_i.$$

It is clear from definition that $h_I(\chi_A) = 0$ for all $A \in \mathcal{A}$. The product term serves as a "unique identifier" for this coined-up polynomial; it will come handy when showing linear independence.

Suppose $\sum_A \alpha_A \cdot f_A(x) + \sum_I \beta_I \cdot h_I(x) \equiv 0$. First assume that the α_A's are non-trivial. Then we may repeat our previous argument without difficulties because the h_I's disappear when substituting in characteristic vectors of $A, B \in \mathcal{A}$. This would lead to a contradiction, so we must conclude the α_A's are all zero.

Hence $\sum_I \beta_I \cdot h_I(x) \equiv 0$. Assume that the t_I's are non-trivial. Then let J be the inlusion-minimal set such that $\beta_I \neq 0$. We evaluate the equation at $x := \chi_J$ and all terms except $\beta_J \cdot h_J$ vanish. (For $I \subset J$ the coefficient is zero. For $I \not\subseteq J$, there exists $i \in I \setminus J$, and in particular $\chi_J(i) = 0$ cancels the product term $\prod_{i \in I} x_i$.) Therefore we derive $\beta_J = 0$, a contradiction.

Finally, we remark that the same bound applies to uniform L-mod-p family as well, but the proof requires Möbius inversion formula.

5 Application Highlights

Explicit construction in Ramsey theory

We will construct explicitly a graph $G = (V, E)$ on n vertices with $\alpha(G), \omega(G) \leqslant k$. This would imply that the Ramsey number $R(k+1)$ is at least $n+1$.

Our vertex set is taken to be $V := \binom{[N]}{p^2-1}$ where N is a parameter we shall fix later. We join two sets $A, B \in V$ by an edge iff $|A \cap B| = p-1 \pmod{p}$. Now observe:

- Every independent set of the graph is a (p^2-1)-uniform $\{0, \ldots, p-2\}$-mod-p family. So its size is at most $\binom{N}{p-1}$.
- Every clique of the graph is a (p^2-1)-uniform $\{p-1\}$-mod-p family. So the size of any pairwise intersection can only be $p-1, 2p-1, \ldots, (p-1)p-1$. Therefore it is also a (p^2-1)-uniform $\{p-1, 2p-1, \ldots, (p-1)p-1\}$ family. So its size is at most $\binom{N}{p-1}$.

Hence we constructed a graph with $n := \binom{N}{p^2-1}$ and $\alpha, \omega \leqslant k := \binom{N}{p-1}$. One may tune $N := p^3$ to get optimal asymptotics.

Chromatic number of $\mathbb{R}^d$

We are asked to paint the Euclidean space $\mathbb{R}^d$ using as least colour as possible. The requirement is that any pair of points $x, y \in \mathbb{R}^d : \|x - y\| \leqslant 1$ are assigned different colours.

A natural approach is to tile hypercubes in $\mathbb{R}^d$. With a more refined method, one may prove that 9^d colours suffice.

The investigation into lower bounds is much more difficult. However, using results from set systems, we can prove a surprisingly nice lower bound 1.1^d without much effort.

Let $d := 4p$ and $m := \binom{4p}{p-1} + 1$. We construct a set $S \subseteq \mathbb{R}^d$ that contains all ± 1-vectors with $2p$ many $+1$'s and the same amount of -1's, and additionally with first coordinate $+1$. These vectors naturally correspond to sets in $\binom{[4p]}{2p}$ containing the first element. Clearly $|S| = \binom{4p}{2p}/2$.

Eventually we will show that S is "full of conflicts". That is, any m-subset of S shall contain a unit-distant pair. Then it's impossible to paint S with only $|S|/m$ colours (otherwise there exists a colour class containing at least m points, hence getting a internal conflict). This certifies that the chromatic number of the entire space is at least

$$\frac{|S|}{m} = \frac{\binom{4p}{2p}}{2\left(\binom{4p}{p-1} + 1\right)} > 1.1^d.$$

Towards our goal, we will do something slightly different:

Lemma. Every m-subset of S contains an orthogonal pair.

Proof. Let $x, y \in S$; assume A, B are their corresponding sets. Note that $x = 2\chi_A - \mathbf{1}$ and similarly $y = 2\chi_B - \mathbf{1}$. Therefore

$$\begin{aligned}\langle x, y\rangle &= 4\langle \chi_A, \chi_B\rangle - 2\langle \chi_A, \mathbf{1}\rangle - 2\langle \chi_B, \mathbf{1}\rangle + \langle \mathbf{1}, \mathbf{1}\rangle \\ &= 4|A \cap B| - 4p - 4p + 4p \\ &= 4|A \cap B| - 4p,\end{aligned}$$

and $x \perp y \Longleftrightarrow |A \cap B| = p$.

Now suppose $T \subseteq S$ does not contain orthogonal pairs, then the corresponding set system of T is a p-uniform $\{1, \ldots, p-1\}$-mod-p family. (0 is invalid because $A \cap B \neq \emptyset$, $|A \cap B| < 2p$ and $|A \cap B| \neq p$.) Therefore $|T| \leqslant \binom{4p}{p-1} < m$. ☐

Our goal almost follows immediately from the lemma. Note that all vectors in S have length $2\sqrt{p}$. We rescale them to obtain length $\frac{1}{\sqrt{2}}$ each. Orthogonality is not harmed, clearly. By the lemma, every m-subset of (the rescaled) S contains an orthogonal pair x, y, and consequently $\|x - y\| = \sqrt{\|x\|^2 + \|y\|^2} = 1$ as desired.

Counterexample of Borsuk's conjecture

Another exciting manifestation of algebraic tools is constructing a counterexample of Borsuk's conjecture:

> Any set $D \subseteq \mathbb{R}^d$ of diameter 1 can be decomposed into $d+1$ parts of diameters strictly less than 1.

What we will show is quite the contrary, in a drastic sense:

> There exists $S_0 \subseteq \mathbb{R}^d$ of diameter 1 that satisfies the following. No matter how we decompose it into $< 1.1^{\sqrt{d}}$ parts, there is always a part of diameter 1.

In fact this counterexample is not far-reaching. It is directly related to our previous set S of ± 1 vectors. We want to reduce the detection of "diameter=1" to the detection of orthogonality. In other words, we wish the existence of orthogonal pair certifies the fact that the diameter is 1.

For our wish to become true, we must ensure that the vectors have angles at most $\pi/2$, which is not quite the case for S.

The solution uses a *tensorisation* trick, which transforms S to a set of the desired property, yet orthogonality is preserved. Define the tensor product of two vectors to be $x \otimes y := x\, y^{\mathrm{T}}$. It is easy to check by linear algebra that

$$\langle x \otimes y, u \otimes v\rangle = \mathrm{tr}((x\, y^{\mathrm{T}})^{\mathrm{T}} u\, v^{\mathrm{T}}) = \mathrm{tr}(y(x^{\mathrm{T}} u) v) = \langle x, u\rangle \langle y, v\rangle.$$

(Here we identify $d \times d$ matrix with d^2-dimensional vector). Now we generate

$$S_0 := \{x \otimes x \in \mathbb{R}^{d \times d} : x \in S\}.$$

Then

$$\langle x \otimes x, y \otimes y\rangle = \langle x, y\rangle^2 \geqslant 0$$

with equality iff $x \perp y$. This is exactly what we are seeking.