2. Fundamental methods of projection theory

Thursday Feb 6

In this lecture, we introduce two fundamental methods for proving estimates in projection theory: the double counting method and the Fourier method.

These methods are cleanest in the setting of finite fields, so we begin with that case.

We write $\mathbb{F}_q$ for the finite field with q elements. Our projections will be a set of linear maps $\mathbb{F}_q^2 \to \mathbb{F}_q$. For each $\theta \in \mathbb{F}_q$, we define $\pi_\theta : \mathbb{F}_q^2 \to \mathbb{F}_q$ by

$$\pi_\theta(x_1, x_2) = x_1 + \theta x_2 \tag{1}$$

Consider the following setup.

Setup.

$$X \subset \mathbb{F}_q^2$$

$$D \subset \mathbb{F}_q \text{ (set of directions)}$$

$$S = S(X, D) := \max_{\theta \in D} |\pi_\theta(X)|.$$

The first example of a set which has many small directions is an integer grid.

Example 1. (Integer grid example) For simplicity suppose that $q = p$ is prime. Write $[N]$ for $\{1, ..., N\}$. For some $N \leq p$, define

$$X = \{(x_1, x_2) : x_1, x_2 \in [N]\}$$

For some $A \leq p$, define

$$D = \{a_1/a_2 : a_1, a_2 \in [A]\}$$

If $\theta \in D$, and $(x_1, x_2) \in X$, we have

$$\pi_\theta(x_1, x_2) = \frac{a_2 x_1 + a_1 x_2}{a_2}.$$

Therefore, $|\pi_\theta(X)| \lesssim AN$. So we get

$$S(X, D) \sim \max(AN, p).$$

The configuration is interesting when $S \leq p/2$. In this case, we have $S \sim AN$ and so

(2) $$|D| \sim \frac{S^2}{|X|}$$

This example generalizes to any finite field $\mathbb{F}_q$ (or any field). But when $q = p^r$ with $r > 1$, there is also a more dramatic example based on the subfields of $\mathbb{F}_q$. We illustrate this in the case $q = p^2$.

Example 2. (Subfield example) Suppose that $q = p^2$ with p prime. Define

$$X = \mathbb{F}_p^2 \subset \mathbb{F}_q^2$$

$$D = \mathbb{F}_p \subset \mathbb{F}_q$$

If $\theta \in D$, and $(x_1, x_2) \in X$, then we have $\pi_\theta(x_1, x_2) = x_1 + \theta x_2 \in \mathbb{F}_p$. So $|\pi_\theta(X)| \leq p$.

So $|X| = p^2 = q$, $|D| = p = q^{1/2}$, and $S = S(X, D) = p = q^{1/2}$.

Comparing with Example 1, we see that $|D|$ is much larger than $\frac{S^2}{|X|}$.

Over $\mathbb{F}_p$, there is no known example more dramatic than the integer grid example. In fact, all known examples with many small projections are small variations of the integer grid example. This leads to the following conjecture.

Conjecture 2.1. *Suppose $X \subset \mathbb{F}_p^2$, $D \subset \mathbb{F}_p$, and $S = \max_{\theta \in D} |\pi_\theta(X)|$. If $S \leq p/2$, then*

$$|D| \lesssim \frac{S^2}{|X|}$$

Here we need $S \leq p/2$ because for any sets X, D, we always have $S \leq p$. If $S = p$, then we cannot get any information about $|D|, |X|$. For fields $\mathbb{F}_q$, I have not seen a conjecture written down anywhere, but informally it is expected that the extreme examples are minor variations on Examples 1 and 2.

We will prove two fundamental estimates about projection theory in $\mathbb{F}_q^2$. The proofs of these results introduce two main techniques that we will use repeatedly: double counting and the orthogonality / Fourier method.

Theorem 2.2. *(Double counting) Suppose $X \subset \mathbb{F}_q^2$, $D \subset \mathbb{F}_q$, and $S = \max_{\theta \in D} |\pi_\theta(X)|$. If $S \leq |X|/2$, then*

$$|D| \lesssim S$$

Theorem 2.3. *(Orthogonality/ Fourier) Suppose $X \subset \mathbb{F}_q^2$, $D \subset \mathbb{F}_q$, and $S = \max_{\theta \in D} |\pi_\theta(X)|$. If $S \leq q/2$, then*

$$|D| \lesssim \frac{Sq}{|X|}.$$

Remark. When $S = q/2$, or when $S \sim q$, Theorem 2.3 matches the grid example and it is sharp. Theorem 2.3 is also sharp for the subfield example. If $q = p$, then whenever S is much less than q, Theorem 2.3 does not appear to be sharp. And even if $q = p^2$, there are many values of S, $|X|$ where Theorem 2.3 does not appear to be sharp.

These theorems give interesting bounds but they don't give a complete picture of projection theory over $\mathbb{F}_q^2$. In part, this is because the techniques that we study today don't distinguish prime fields from non-prime fields, but the optimal projection estimates do depend on whether the field is prime. It is fairly difficult to prove bounds going beyond these two theorems, and we will return to that later in the course.

2.1. **Double Counting.**

Proof of Theorem 1. We will apply double counting to the set

$$(*) := \{\theta \in D, x_1 \neq x_2 \in X : \pi_\theta(x_1) = \pi_\theta(x_2)\}$$

(Note on notation: here x_1, x_2 are points in X, not components of a vector.)

We call $(*)$ the set of coincidences. The idea of the proof is as follows. If there are many directions θ where $\pi_\theta(X)$ is small, then there must be a lot of coincidences. But for any $x_1 \neq x_2 \in X$, there is only one direction θ so that $\pi_\theta(x_1) = \pi_\theta(x_2)$, and so there can't be that many coincidences.

If $\theta \in D$, then we have $|\pi_\theta(X)| \leq S \leq |X|/2$. Therefore, using Cauchy-Schwarz, we get

$$\#\{x_1 \neq x_2 \in X : \pi_\theta(x_1) = \pi_\theta(x_2)\} \gtrsim S \left(\frac{|X|}{S}\right)^2.$$

(Details of this argument are on the first problem set.). And so

$$(*) \gtrsim |X|^2 S^{-1} |D|.$$

On the other hand, for each $x_1 \neq x_2 \in X$, there is only one direction θ so that $\pi_\theta(x_1) = \pi_\theta(x_2)$, and so

$$(*) \leq |X|^2.$$

All together we have

$$|X|^2 S^{-1} |D| \lesssim (*) \lesssim |X|^2,$$

and so $|D| \lesssim S$. □

2.2. **Orthogonality / Fourier method.**

Proof of Theorem 2.3. The fibers of the map π_θ are parallel lines in $\mathbb{F}_q^2$. So if $|\pi_\theta(X)| \leq S$, then we can cover X using at most L lines coming from fibers of π_θ.

Recall that for each $\theta \in D$, $|\pi_\theta(X)| \leq S$. Let $\mathbb{L}_\theta$ be a set of S fibers of π_θ which covers X. Let $\mathbb{L} = \cup_{\theta \in D} \mathbb{L}_\theta$. Note that

$$|\mathbb{L}| = |D| S.$$

If L is a line in $\mathbb{F}_q^2$, we write $L(x)$ for the characteristic function of L. We define

$$f(x) = \sum_{L \in \mathbb{L}} L(x)$$

Notice that for each $x \in X$,

$$f(x) = |D|.$$

We will estimate the function f using orthogonality. To do that, we first break up each function L as a constant function plus a mean zero part:

$$L(x) = \underbrace{\frac{1}{q}}_{L_0(x)} + \underbrace{L(x) - \frac{1}{q}}_{L_h(x)} \tag{3}$$

Here $L_0(x) = 1/q$ is the mean value of $L(x)$, and so $L_h(x)$ has mean zero. (The mean value of a function $g : \mathbb{F}_q^d \to \mathbb{C}$ is $\frac{1}{q^d} \sum_{x \in \mathbb{F}_q^d} g(x)$.) We can break up f in a similar way:

$$f(x) = \sum_{L \in \mathbb{L}} L(x) = \underbrace{\frac{|\mathbb{L}|}{q}}_{f_0(x)} + \underbrace{\sum_{L \in \mathbb{L}} L_h(x)}_{f_h(x)} \tag{4}$$

The constant function f_0 is very simple to understand. Since $|\mathbb{L}| = SD$, and since we assumed $S \leq q/2$, we have $f_0(x) \leq |D|/2$. Now for every $x \in X$, $f(x) = |D|$, and so

$$|f_h(x)| \geq |D|/2 \text{ for all } x \in X$$

The key point is that the functions $L_h(x)$ are essentially orthogonal, and we can use this to estimate the function f_h. We state the orthogonality as a lemma.

Lemma 2.4. *If L_1, L_2 are two different lines in $\mathbb{F}_q^2$, then*

$$\sum_{x \in \mathbb{F}_q^2} L_{1,h}(x) L_{2,h}(x) \leq 0.$$

Using Lemma 2.4, we can bound the L^2 norm of f_h:

$$\sum_{x \in \mathbb{F}_q^2} |f_h(x)|^2 = \sum_{L_1, L_2 \in \mathbb{L}} \sum_{x \in \mathbb{F}_q^2} L_{1,h}(x) L_{2,h}(x) \leq \sum_{L \in \mathbb{L}} \sum_{x \in \mathbb{F}_q^2} |L_h(x)|^2.$$

For each line L, we can compute $\sum_{x \in \mathbb{F}_q^2} |L_h(x)|^2$ by hand. It is slightly smaller than $\sum_{x \in \mathbb{F}_q^2} L(x) = q$. So all together we get the L^2 bound

$$\sum_{x \in \mathbb{F}_q^2} |f_h(x)|^2 \leq |\mathbb{L}| q \tag{5}$$

Combining everything we have done so far, we see that

$$|X||D|^2 \lesssim \sum_{x \in \mathbb{F}_q^2} |f_h(x)|^2 \leq |\mathbb{L}| q = |D| S q$$

Rearranging gives $|D| \lesssim \frac{Sq}{|X|}$.

□

Before we prove Lemma 2.4, we make some comments about the proof. Our bounds here are interesting when $|\mathbb{L}|$ is much larger than q. The key input is the L^2 estimate for f_h in (5. When $|\mathbb{L}|$ is much bigger than q, then this estimate shows that $\sum_{x \in \mathbb{F}_q^2} |f_0(x)|^2$ is much bigger than $\sum_{x \in \mathbb{F}_q^2} |f_h(x)|^2$. So $f(x)$ is equal to a constant function f_0 plus a perturbation f_h, and for most x, $|f_h(x)|$ is much smaller than $|f_0(x)|$. Informally, we could say that the function $f(x)$ is almost constant.

Looking back at the proof of our L^2 estimate (5), the argument applies to any set of lines $\mathbb{L}$. The crux of the matter is that if $|\mathbb{L}|$ is much bigger than q, and if $f(x) = \sum_{L \in \mathbb{L}} L(x)$, then $f = f_0 + f_h$ where f_0 is a constant function, and f_h has small L^2 norm.

The key to the L^2 estimate is the orthogonality in Lemma 2.4. Now we discuss the proof of Lemma 2.4. One simple proof is just to compute $\sum_{x \in \mathbb{F}_q^2} L_{1,h}(x) L_{2,h}(x)$. Recall that

$$L_{1,h}(x) = \begin{cases} 1 - 1/q & x \in L_1 \\ -1/q & x \notin L_1 \end{cases}$$

We can now compute $\sum_{x \in \mathbb{F}_q^2} L_{1,h}(x) L_{2,h}(x)$. With a little algebra, we find

$$\sum_{x \in \mathbb{F}_q^2} L_{1,h}(x) L_{2,h}(x) \begin{cases} = 0 & \text{if } L_1, L_2 \text{ are not parallel} \\ < 0 & \text{if } L_1, L_2 \text{ are parallel} \end{cases}$$

The main case is when L_1, L_2 are not parallel. In this case something interesting is happening that causes the sum to be zero, and we should look for a conceptual explanation. One explanation comes from independence. After a change of coordinates, we can assume that L_1 is the vertical axis and L_2 is the horizontal axis. In these coordinates, $L_{1,h}$ only depends on x_1 and $L_{2,h}$ only depends on x_2, and so $L_{1,h}$ and $L_{2,h}$ are independent. Therefore,

$$\sum_{x \in \mathbb{F}_q^2} L_{1,h}(x) L_{2,h}(x) = \left(\sum_{x \in \mathbb{F}_q}^{2} L_{1,h}(x) \right) \left(\sum_{x \in \mathbb{F}_q}^{2} L_{2,h}(x) \right) = 0 \cdot 0 = 0.$$

Another conceptual explanation comes from Fourier analysis. We now pause to review the Fourier transform over finite fields, and then we use Fourier analysis to explain why $L_{1,h}$ and $L_{2,h}$ are orthogonal when L_1, L_2 are not parallel.

Suppose that $e : \mathbb{F}_q \to \mathbb{C}^*$ is a non-trivial homomorphism from the group $\mathbb{F}_q^+$ to the group $\mathbb{C}^*$. If $q = p$ is prime, then we can take $e(x) = e^{2\pi i \frac{x}{p}}$.

If $x, \xi \in \mathbb{F}_q^d$, we define the dot product $x \cdot \xi$ by

$$x \cdot \xi = x_1 \xi_1 + ... + x_d \xi_d.$$

If $f : \mathbb{F}_q^d \to \mathbb{C}$, then we define its Fourier transform $\hat{f} : \mathbb{F}_q^d \to \mathbb{C}$ by

$$\hat{f}(\xi) := \sum_{x \in \mathbb{F}_q^d} f(x) e(-x \cdot \xi) \tag{6}$$

With this setup, we can write down the two fundamental theorems in Fourier analysis: Fourier inversion and Plancherel.

Theorem 2.5. *If $f : \mathbb{F}_q^d \to \mathbb{C}$, then*

$$f(x) = \frac{1}{q^d} \sum_{\xi \in \mathbb{F}_q^d} \hat{f}(\xi) e(x \cdot \xi) = \underbrace{\frac{1}{q^d} \hat{f}(0)}_{f_0(x)} + \underbrace{\frac{1}{q^d} \sum_{\xi \neq 0} \hat{f}(\xi) e(x \cdot \xi)}_{f_h(x)}$$

Theorem 2.6. *If $f, g : \mathbb{F}_q^d \to \mathbb{C}$, then*

$$\sum_{x \in \mathbb{F}_q^d} f(x) \overline{g(x)} = \frac{1}{q^d} \sum_{\xi \in \mathbb{F}_q^d} \hat{f}(\xi) \overline{\hat{g}(\xi)}$$

Let us now revisit how we broke up a function f as $f_0 + f_h$. Starting with Fourier inversion, we can write f as

$$f(x) = \frac{1}{q^d} \sum_{\xi \in \mathbb{F}_q^d} \hat{f}(\xi) e(x \cdot \xi) = \underbrace{\frac{1}{q^d} \hat{f}(0)}_{f_0(x)} + \underbrace{\frac{1}{q^d} \sum_{\xi \neq 0} \hat{f}(\xi) e(x \cdot \xi)}_{f_h(x)}$$

Since $\hat{f}(0) = \sum_{q \in \mathbb{F}_q^d} f(x)$, we see that f_0 is just the mean value of $f(x)$. So this decomposition is the same one we used above in the proof of Theorem 2.3. We can think of f_0 as the contribution of the zero frequency, and we think of f_h as the contribution of the non-zero frequencies. The letter h stands for 'high', and we think of f_h as the 'high-frequency' part of f. In general, for any function f, we can define f_h as above, and we have

$$\hat{f}_h(\xi) = \begin{cases} \hat{f}(\xi) & \xi \neq 0 \\ 0 & \xi = 0 \end{cases}$$

The Fourier transform interacts in a nice way with lines, and more generally with affine subspaces. Suppose that $P \subset \mathbb{F}_q^d$ is an affine k-plane. We write $P(x)$ for the characteristic function of P. We define $P^\perp$ as

$$P^\perp = \{\xi \in \mathbb{F}_q^d : (x_1 - x_2) \cdot \xi = 0 \text{ for all } x_1, x_2 \in P\}.$$

(Here the vector $x_1 - x_2$ is tangent to P, and so $P^\perp$ is the set of vectors perpendicular to P. Note that P is affine, so it may not contain 0, whereas $P^\perp$ is a subspace, and it does contain 0.)

Lemma 2.7. *If $P(x)$ is the characteristic function of an affine k-plane in $\mathbb{F}_q^d$, then*

$$|\hat{P}(\xi)| = \begin{cases} q^k & \xi \in P^\perp \\ 0 & \xi \notin P^\perp \end{cases}$$

The proof of Lemma 2.7 is on the first problem set. The main point is that when P is an affine plane, then $\hat{P}(\xi) = \sum_{x \in P} e(-x \cdot \xi)$ is a geometric series, and so we can sum it exactly. For most ξ, the geometric series sums to zero because of symmetry.

Using Fourier analysis, we can now give another proof that when L_1, L_2 are not parallel, then $L_{1,h}$ and $L_{2,h}$ are orthogonal. By Plancherel, we have

$$\sum_{x \in \mathbb{F}_q^2} L_{1,h}(x) L_{2,h}(x) = \frac{1}{q^d} \sum_{\xi \in \mathbb{F}_q^2} \hat{L}_{1,h}(\xi) \overline{\hat{L}_{2,h}(\xi)} = \frac{1}{q^d} \sum_{\xi \neq 0} \hat{L}_1(\xi) \overline{\hat{L}_2(\xi)}$$

But by Lemma 2.7, the support of $\hat{L}_1$ is $L_1^\perp$ and the support of $\hat{L}_2$ is $L_2^\perp$. Since the two supports intersect only at $\xi = 0$, our last sum is zero.

To summarize, $\hat{L}_{1,h}$ and $\hat{L}_{2,h}$ have disjoint supports, and so $L_{1,h}$ and $L_{2,h}$ are orthogonal.

Remark. We don't necessarily need Fourier analysis to prove Theorem 2.3, but in some further developments the Fourier analysis is helpful. For instance, if we want to generalize Theorem 2.3 to higher dimensions, the Fourier analysis point of view is important. You will explore this on the first problem set.

2.3. **Projection theory for balls in Euclidean space.** Next we will start to study projection theory in Euclidean space. We will consider the projections of a set of unit balls in Euclidean space, and we will adapt our two fundamental methods to that setting. There is a new issue that appears for balls in Euclidean space, which has to do with how the balls are clustered. In this lecture, we start to set up our problems in the context of balls in Euclidean space, and we see how the clustering comes into play.

In this section, for a set $X \subset \mathbb{R}^d$, we write $|X|$ for the d-dimensional measure of X.

Setup

Suppose that X is a set of disjoint unit balls in $B_R \subset \mathbb{R}^2$.

Suppose that D is a finite set in S^1, which is $1/R$-separated.

Define $S(X, D) = \max_{\theta \in D} |\pi_\theta(X)|$, the maximal 1-dimensional measure of $\pi_\theta(X)$.

Here we suppose that the directions in D are $1/R$-separated because otherwise the projections would be essentially equivalent.

Next we can consider some examples. There is an integer grid example which is analogous to the one we mentioned in finite fields.

Example 1. (Widely spaced integer grid example)

We let X be an $N \times N$ grid of unit balls in B_2^R, spaced as widely as possible. The centers of the balls lie on the lattice $\frac{R}{N}\mathbb{Z} \times \frac{R}{N}$.

We choose a parameter $A \leq R$, and we let D be the set of directions with slope in the set $\{a_1/a_2 : a_1, a_2 \in [A]\}$.

By a similar analysis to the one in finite fields, we see that

$$S(X, D) \sim \max(AN, R).$$

The configuration is interesting when $S \leq R/2$. In this case, we have $S \sim AN$ and so

$$|D| \sim \frac{S^2}{|X|} \tag{7}$$

Notice that the numerology of Example 1 in the setting of balls exactly matches the numerology for the integer grid in $\mathbb{F}_p^2$.

We recall that for projection theory in $\mathbb{F}_q^2$, there were interesting examples related to subfields of $\mathbb{F}_q$. The field $\mathbb{R}$ does have subfields, such as the field of rational numbers. However, these subfields do not lead to interesting sets of unit balls in B_R. I think that the issue is that $\mathbb{Q}$ is not closed. To get a set of unit balls, we might take the 1-neighborhood of $\mathbb{Q} \times \mathbb{Q}$, but that is all of $\mathbb{R}^2$.

But there is a new phenomenon for projection theory of balls in Euclidean space which has to do with clustering. As a second example, we consider a tightly clustered set of balls.

Example 2. (Clustered example)

For some $N \leq R$, we let X be a set of $\sim N^2$ disjoint unit balls in $B_N \subset B_R$. We have $|X| \sim N^2$

Now for every direction θ, we have $|\pi_\theta(X)| \lesssim N$.

So we can let D be a maximal set of $1/R$ separated directions, so $|D| \sim R$, and we can take $S = 2N$.

Plugging in, we find that $|D|$ is much larger than $\frac{S^2}{|X|} \sim 1$. And so this example is more extreme than Example 1.

The new theme in this setting is that projection estimates depend on how much X is clustered. It turns out that it is important to consider both how X is clustered and how D is clustered. We can quantify the clustering of X and D with the following definitions.

We write $B(c, r)$ for the ball with center c and radius r. For any $1 \leq r \leq R$, we define

$$N_X(r) = \max_{c \in B_R} |X \cap B(c, r)| \tag{8}$$

We write γ for an arc of S^1, and $|\gamma|$ for its length. For any $\rho \in [1/R, 1]$, we define

$$N_D(\rho) = \max_{|\gamma| = \rho} \#(D \cap \gamma) \tag{9}$$

Our goal will be to prove projection estimates that depend on the functions $N_X(r)$ and $N_D(\rho)$.

In the next lecture we will work out analogues of Theorem 2.2 and Theorem 2.3 for balls in $\mathbb{R}^2$. The main idea will be to adapt the methods we used today in order to take account of clustering information from $N_X(r)$ and $N_D(\rho)$.