Mathematical Toolkit **Fall 2024**

Lecture 2: October 2, 2024

Lecturer: Avrim Blum (notes based on notes from Madhur Tulsiani)

1 Applications of our development so far

1.1 Lagrange interpolation

Lagrange interpolation is used to find the unique polynomial of degree at most $n-1$, taking given values at n distinct points. We can derive the formula for such a polynomial using basic linear algebra.

Recall that the space of polynomials of degree at most $n-1$ with real coefficients, denoted by $\mathbb{R}^{\leq n-1}[x]$, is a vector space. What is the dimension of this space? What would be a simple example of a basis?

Let $a_1, \ldots, a_n \in \mathbb{R}$ be distinct. Say we want to find the unique polynomial p of degree at most $n-1$ satisfying $p(a_i) = b_i \; \forall i \in [n]$. Recall from the last lecture that if we define $g(x)$ as $\prod_{i=1}^{n}(x-a_i)$, the degree $n-1$ polynomials defined as

$$f_i(x) \;=\; \frac{g(x)}{x-a_i} \;=\; \prod_{j \neq i}^{n}(x-a_j)\,,$$

are n linearly independent polynomials in $\mathbb{R}^{\leq n-1}[x]$. Thus, they must form a basis for $\mathbb{R}^{\leq n-1}[x]$ and we can write the required polynomial, say p as

$$p \;=\; \sum_{i=1}^{n} c_i \cdot f_i\,,$$

for some $c_1, \ldots, c_n \in \mathbb{R}$. Evaluating both sides at a_i gives $p(a_i) = b_i = c_i \cdot f_i(a_i)$. Thus, we get

$$p(x) \;=\; \sum_{i=1}^{n} \frac{b_i}{f_i(a_i)} \cdot f_i(x)\,.$$

1.2 Secret Sharing

Note that the only property of the field $\mathbb{R}$ we used in Lagrange interpolation, was the fact that $\mathbb{R}$ is large enough to contain n distinct points $a_1, \ldots, a_n$. Check that the same argument

can be used to find a polynomial of degree at most $n-1$ in the space $\mathbb{F}[x]$ for any field $\mathbb{F}$ such that $|\mathbb{F}| \geq n$. This can then be used to develop another nice application as below.

Consider the problem of sharing a secret s, which is an integer in a known range $[0, M]$ with a group of n people, such that if any d of them get together, they are able to learn the secret message. However, if fewer than d of them are together, they do not get any information about the secret. We can then proceed as follows:

- Choose a finite field $\mathbb{F}_p$, with $p > \max(n, M)$.
- Choose $d-1$ random values $b_1, \ldots, b_{d-1}$ in $\{0, ..., p-1\}$, and let $Q \in \mathbb{F}_p^{\leq d-1}[x]$ be the polynomial
$$Q = s + b_1 x + b_2 x^2 + ... + b_{d-1} x^{d-1}.$$
Note that the secret is $Q(0)$.
- For $i = 1, ..., n$, give person i the pair $(i, Q(i))$.

Note that if any group of d or more people get together, they can uniquely determine the polynomial Q by Lagrange interpolation. They can then recover the secret by evaluating Q at 0. However, if $d-1$ of them gather, then there is always a polynomial consistent with the values they hold, and any possible value at 0. To precisely say that they learn nothing about the secret, we we use the fact that there is *exactly one* polynomial consistent with the values they hold and any given value at 0. Since for any given secret s there are exactly p^{d-1} polynomials with $Q(0) = s$, and we chose the polynomial at random conditioned on the secret, this means that any two secrets have the same probability of producing the observed $(d-1)$-tuple of shares. We will talk in more depth about arguments like this when we discuss probability in the second half of the course.

2 Existence of bases in general vector spaces

We proved that any finitely-generated vector space must have a basis. Recall that a vector space V is said to be finitely generated if there exists a finite set T such that $\text{Span}(T) = V$.

It turns out that general vector spaces (including infinite-dimensional ones) have bases too. Proving this uses a fact called "Zorn's lemma" which states that any partially-ordered set (think of the set of linearly-independent sets of vectors, ordered by inclusion) having the property that all totally-ordered subsets are bounded (meaning that there is some element that is greater than or equal to all of them) must have a maximal element (an element with no other element greater than or equal to it). In the case of linearly-independent sets of vectors ordered by inclusion, the maximal element will be a basis. I won't go through the details of the argument here but feel free to think about it. Zorn's lemma turns out to be equivalent to the axiom of choice. (So the statement about existence of bases in general vector spaces depends on the axiom of choice.)

3 Linear Transformations

Definition 3.1 *Let V and W be vector spaces over the same field $\mathbb{F}$. A map $\varphi : V \to W$ is called a* linear transformation *if*

- $\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) \quad \forall v_1, v_2 \in V.$
- $\varphi(c \cdot v) = c \cdot \varphi(v) \quad \forall v \in V.$

Example 3.2 *The following are all linear transformations:*

- *A matrix $A \in \mathbb{R}^{m \times n}$ (m rows, n columns) defines a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.*

  Note that we are using $\varphi_A(v) = Av$, where we are viewing the elements of $\mathbb{R}^m$ and $\mathbb{R}^n$ as column vectors.
- *$\varphi : C([0,1], \mathbb{R}) \to C([0,2], \mathbb{R})$ defined by $\varphi(f)(x) = f(x/2)$. Recall that $C([a,b], \mathbb{R}) = \{f : [a,b] \to \mathbb{R} \mid f \text{ is continuous}\}$.*
- *$\varphi : C([0,1], \mathbb{R}) \to C([0,1], \mathbb{R})$ defined by $\varphi(f)(x) = f(x^2)$.*
- *$\varphi : C([0,1], \mathbb{R}) \to C([0,1], \mathbb{R})$ defined by $\varphi(f)(x) = f(1-x)$.*
- *The derivative operator acting on $\mathbb{R}[x]$. (Polynomials in x with real-valued coefficients)*

Proposition 3.3 *Let V, W be vector spaces over $\mathbb{F}$ and let B be a basis for V. Let $\alpha : B \to W$ be an arbitrary map. Then there exists a unique linear transformation $\varphi : V \to W$ satisfying $\varphi(v) = \alpha(v)$ $\forall v \in B$.*

Proof: Since B is a basis, any $u \in V$ can be written in a unique way as a sum $\sum_{v \in B} a_v v$, where the values a_v are in $\mathbb{F}$ and only finitely many are nonzero. By the two properties of a linear transformation, we must then have $\varphi(u) = \sum_{v \in B} a_v \varphi(v)$. Since the values $\varphi(v)$ are fixed for all $v \in B$, this gives the unique solution of $\varphi(u) = \sum_{v \in B} a_v \alpha(v)$. Moreover, this φ indeed satisfies the property that $\varphi(v) = \alpha(v)$ for all $v \in B$. ∎

Proposition 3.3 solidifies the connection between linear transformations and matrices. We saw that a matrix $A \in \mathbb{F}^{m \times n}$ corresponds to a linear transformation φ_A from $\mathbb{F}^n$ to $\mathbb{F}^m$ defined as $\varphi_A(v) = Av$. But we can also go the other way as well. Given a linear transformation $\varphi : \mathbb{F}^n \to \mathbb{F}^m$, consider the standard basis $B = \{e_1, ..., e_n\}$ for $\mathbb{F}^n$, where e_i has 1 in its ith coordinate and 0 in all other coordinates. By Proposition 3.3, φ is uniquely defined by its effect on B, and so can be represented by the matrix $A \in \mathbb{F}^{m \times n}$ with $\varphi(e_i)$ as its ith column.

Definition 3.4 *Let $\varphi : V \to W$ be a linear transformation. We define its* kernel *and* image *as:*

- $\ker(\varphi) := \{v \in V \mid \varphi(v) = 0_W\}$.
- $\text{im}(\varphi) = \{\varphi(v) \mid v \in V\}$.

Proposition 3.5 $\ker(\varphi)$ *is a subspace of V and* $\text{im}(\varphi)$ *is a subspace of W.*

Definition 3.6 $\dim(\text{im}(\varphi))$ *is called the rank and* $\dim(\ker(\varphi))$ *is called the nullity of φ.*

Proposition 3.7 (rank-nullity theorem) *If V is a finite dimensional vector space and $\varphi : V \to W$ is a linear transformation, then*

$$\dim(\ker(\varphi)) + \dim(\text{im}(\varphi)) = \dim(V).$$

Proof: Let $n = \dim(V)$ and let $k = \dim(\ker(\varphi))$. Choose a basis $v_1, ..., v_k$ for the kernel and then extend this to a basis B for V with linearly independent vectors $v_{k+1}, ..., v_n$ (which we can always do, as we saw in the last class). We know that

$$\text{im}(\varphi) = \text{Span}(\{\varphi(v_1), ..., \varphi(v_n)\}) = \text{Span}(\{\varphi(v_{k+1}), ..., \varphi(v_n)\}).$$

So, to show that the rank is $n - k$, all that remains is to show that $\varphi(v_{k+1}), ..., \varphi(v_n)$ are linearly independent. This follows from the definition of linear transformation: if some linear combination of $\varphi(v_{k+1}), ..., \varphi(v_n)$ equals 0 then so does φ of the same linear combination of $v_{k+1}, ..., v_n$, meaning that this linear combination of $v_{k+1}, ..., v_n$ lies in the kernel. This contradicts the fact that they were all linearly independent of $v_1, ..., v_k$. ■

Example 3.8 *Consider the matrix A which defines a linear transformation from $\mathbb{F}_2^7$ to $\mathbb{F}_2^3$:*

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

- $\dim(\text{im}(\varphi)) = 3$.
- $\dim(\ker(\varphi)) = 4$.

Remark. *In the above example,* $\ker(\varphi)$ *is a* code *which can recover from one bit of error. More generally, this is also true for the $(2^k - 1) \times k$ matrix A_k where the i^{th} column is the number i written in binary (with the most significant bit at the top).*

This code is known as the Hamming Code and the matrix A is called the parity-check matrix of the code.