TTIC 31150/CMSC 31150 Mathematical Toolkit (Fall 2024)

Avrim Blum

Lecture 2: Vector Space Applications and Linear Transformations

Recap

- Fields (like $\mathbb{R}, \mathbb{Q}, \mathbb{F}_p$)
- Vector spaces (like $\mathbb{R}^n, \mathbb{F}_p^n$)
- Linear dependence / independence
- Span(S)
- Basis of V
- Steinitz Exchange Principle
- Dimension of finitely-generated vector space

Existence of bases in general vector spaces

- Any finitely-generated vector space (∃ finite set T s.t. $Span(T) = V$) has a basis.
- Turns out also true for general vector spaces (even infinite-dimensional).
  - Example of such vector space? Polynomials $R\,[X]$ over $\mathbb{R}$, or $\mathbb{R}$ over $\mathbb{Q}$
    - $f(n) = x^n$, for $n = 0,1,2,\cdots$
  - We define span using finite linear combination (Hamel Basis)
  - Generic vector space may not have notion of distance, closeness and convergence
- Proving it uses "Zorn's lemma" which is equivalent to axiom of choice.
- Won't get into here.

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?

- Dimension is n. Standard basis is $\{1, x, x^2, \ldots, x^{n-1}\}$.

Lagrange Interpolation (contd)

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 $p(a_i) = b_i \; \forall i \in [n]$.

- Why unique?
  - If there were two, say p_1, p_2, then $p_1 - p_2$ would have at least n roots. But a nonzero polynomial of degree at most $n-1$ can have at most $n-1$ roots.

Lagrange Interpolation (contd)

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

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

Lagrange Interpolation (contd)

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 $c_1, \ldots, c_n \in \mathbb{R}$.

Lagrange Interpolation (contd)

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

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

Because all the other terms evaluate to 0

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

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

Lagrange Interpolation (contd)

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 $p(a_i) = b_i \; \forall i \in [n]$.

- Argument works if replace $\mathbb{R}$ with any field $\mathbb{F}$ having at least n distinct points.

Secret Sharing

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.

- E.g., password, (decryption key for) sensitive data, etc.

Shamir's secret sharing

7 languages

Article Talk | Read Edit View history

From Wikipedia, the free encyclopedia

> This article includes a list of general references, but **it lacks sufficient corresponding inline citations**. Please help to improve this article by introducing more precise citations. *(February 2019) (Learn how and when to remove this template message)*

Shamir's secret sharing (SSS) is an efficient secret sharing algorithm for distributing private information (the "secret") among a group so that the secret cannot be revealed unless a quorum of the group acts together to pool their knowledge. To achieve this, the secret is mathematically divided into parts (the "shares") from which the secret can be reassembled only when a sufficient number of shares are combined. SSS has the property of information-theoretic security, meaning that even if an attacker steals some shares, it is impossible for the attacker to reconstruct the secret unless they have stolen the quorum number of shares.

Shamir's secret sharing is used in some applications to share the access keys to a master secret.

Secret Sharing

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

One direction: If any d get together, can uniquely determine Q by Lagrange interpolation, recover secret by evaluating Q at 0.

Secret Sharing

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 $(i, Q(i))$.

Other direction:

- If $d - 1$ get together, for any secret s', exists a consistent polynomial Q'. In fact, exactly one.
- Because Q chosen randomly from p^{d-1} polynomials consistent with secret, this means any two secrets have the same probability of producing the observed $d - 1$ shares.

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

- *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* $\varphi(f)(x) = f(x^2)$.

Important properties

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 $v \in B$. ■

Important properties

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 $\varphi(v) = \alpha(v)\ \forall v \in B$.*

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:*

- $\ker(\varphi) := \{v \in V \mid \varphi(v) = 0_W\}$. [Kernel also called “nullspace”]
- $\operatorname{im}(\varphi) = \{\varphi(v) \mid v \in V\}$.

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

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

What is rank of φ_A for $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 2 \end{bmatrix}$?

Rank is 2

Nullspace just 0_V since columns are independent

What is rank of φ_B for B= $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}$?

Rank is 2

How about nullspace? All multiples of $\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$

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

- $\ker(\varphi) := \{v \in V \mid \varphi(v) = 0_W\}$. [Kernel also called “nullspace”]
- $\operatorname{im}(\varphi) = \{\varphi(v) \mid v \in V\}$.

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

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

How about $A = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$?

(Mapping $\mathbb{R}^7$ to $\mathbb{R}^3$)

Rank is 3

Nullity is 4

Nullspace spanned by $\begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$

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

$$\dim(\ker(\varphi)) + \dim(\mathrm{im}(\varphi)) \;=\; \dim(V)\,.$$

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

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

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 $v_1, ..., v_k$. ■