Mathematical Toolkit **Autumn 2021**

Lecture 2: September 30, 2021

Lecturer: Madhur Tulsiani

1 Span and Bases

Recall the definition of the span of a set, discussed in the previous lecture.

Definition 1.1 *Given a set $S \subseteq V$, we define its* span *as*

$$\text{Span}(S) = \left\{ \sum_{i=1}^{n} a_i \cdot v_i \mid a_1, \ldots, a_n \in \mathbb{F}, v_1, \ldots, v_n \in S, n \in \mathbb{N} \right\}.$$

Note that we only include finite *linear combinations. Also, since linear combinations of vectors are still in V, we have* $\text{Span}(S) \subseteq V$. *In fact, you can check that* $\text{Span}(S)$ *is also a vector space. Such a subset of V, which is also a vector space, is called a* subspace *of V.*

Remark 1.2 *Note that the definition above and the previous definitions of linear dependence and independence, all involve only finite linear combinations of the elements. Infinite sums cannot be said to be equal to a given element of the vector space without a notion of convergence or distance, which is not necessarily present in an abstract vector space.*

Definition 1.3 *A set B is said to be a* basis *for the vector space V if B is linearly independent and* $\text{Span}(B) = V$.

We will say that a set $B \subseteq V$ is a maximal linearly independent set if B is linearly independent and for all $v \in V \setminus B$, $B \cup \{v\}$ is linearly dependent. We also discussed that $B \subseteq V$ is a basis for V if and only if B is a maximal linearly independent set.

We will now discuss a tool that'll be very helpful in arguing about bases of vector spaces.

Proposition 1.4 (Steinitz exchange principle) *Let $\{v_1, \ldots, v_k\}$ be linearly independent and $\{v_1, \ldots, v_k\} \subseteq \text{Span}(\{w_1, \ldots, w_n\})$. Then $\forall i \in [k]$ $\exists j \in [n]$ such that $w_j \notin \{v_1, \ldots, v_k\} \setminus \{v_i\}$ and $\{v_1, \ldots, v_k\} \setminus \{v_i\} \cup \{w_j\}$ is linearly independent.*

Proof: Assume not. Then, there exists $i \in [k]$ such that for all $w_j \notin \{v_1, \ldots, v_k\} \setminus \{v_i\}$, $\{v_1, \ldots, v_k\} \setminus \{v_i\} \cup \{w_j\}$ is linearly dependent. Note that this means we cannot have $v_i \in \{w_1, \ldots, w_n\}$ (why?)

The above gives that for all $j \in [n]$, $w_j \in \text{Span}(\{v_1, \ldots, v_k\} \setminus \{v_i\})$. However, this implies

$$\{v_1, \ldots, v_k\} \subseteq \text{Span}(\{w_1, \ldots, w_n\}) \subseteq \text{Span}(\{v_1, \ldots, v_k\} \setminus \{v_i\}),$$

which is a contradiction. ■

The following is an easy corollary of the Steinitz exchange principle.

Corollary 1.5 *Let $B_1 = \{v_1, \ldots, v_k\}$ and $B_2 = \{w_1, \ldots, w_n\}$ be two bases of a vector space V. Then, they must have the same size i.e., $k = n$.*

Note that we are already assuming in the above statement that the bases are finite, which may not necessarily be the case for all vector spaces. The above is just saying that *if* there happen to be two finite bases, then they must be of equal sizes.

Proof Sketch: Use the exchange principle to successively replace elements from B_1 by those from B_2. Since we need to replace k elements and no element of B_2 can be used twice (why?) we must have $k \leq n$. By symmetry, we must also have $n \leq k$.

Note that the above argument just needs that we can remove elements from B_1 and replace them by new elements from B_2. While it is true that the intermediate sets we will construct will also be bases, we don't need this to argue $k \leq n$. □

1.1 Finitely generated spaces

A vector space V is said to be finitely generated if there exists a finite set T such that $\text{Span}(T) = V$. Note that Corollary 1.5 proves that all bases of a finitely generated vector space (if they exist!) have the same size. It is easy to see that a similar argument can also be used to prove that a basis must always exist.

Exercise 1.6 *Prove that a finitely generated vector space with a generating set T has a basis (which is a subset of the generating set T).*

The above argument can also be used to prove a stronger statement.

Exercise 1.7 *Let V be a finitely generated vector space and let $S \subseteq V$ be any linearly independent set. Then S can be "extended" to a basis of V i.e., there exists a basis B such that $S \subseteq B$.*

The size of all bases of a vector space is called the dimension of the vector space, denoted as $\dim(V)$. Using the above arguments, it is also easy to check that *any* linearly independent set of the right size must be a basis.

Exercise 1.8 *Let V be a finitely generated vector space and let S be a linearly independent set with $|S| = \dim(V)$. Prove that S must be a basis of V.*

1.2 What if V is not finitely generated?

Of course, it need not always be the case that the vector space we are dealing with is finitely generated. For example, the vector space $\mathbb{R}[x]$ of polynomials has no finite generating set (since the maximum degree in any finite set T generating set will be an upper bound on the degree of polynomials in Span (T).) However, it is still the case that every vector space has a (possibly infinite) basis, such that all elements can be expressed as *finite* linear combinations of the basis elements. Such a basis is known as a Hamel basis. We will present the argument for existence of a Hamel basis below, in case you are interested. However, this will not be incuded in tests or homeworks for the class.

To prove the existence of a basis for every vector space, we will need Zorn's Lemma (which is equivalent to the axiom of choice). We first define the concepts needed to state and apply the lemma.

Definition 1.9 *Let X be a non-empty set. A relation $\preceq$ between elements of X is called a* partial order

- $x \preceq x$ *for all* $x \in X$.
- $x \preceq y, y \preceq x \Rightarrow x = y$.
- $x \preceq y, y \preceq z \Rightarrow x \preceq z$.

The relation is called a partial order since not all the elements of X may be related. A subset $Z \subseteq X$ is called totally ordered *if for every $x, y \in Z$ we have $x \preceq y$ or $y \preceq x$. A set $Z \subseteq X$ is called* bounded *if there exists $x_0 \in X$ such that $z \preceq x_0$ for all $z \in Z$. An element $x_0 \in X$ is maximal if there does not exist any other $y \in X \in \{x_0\}$ such that $x_0 \preceq x$.*

Proposition 1.10 (Zorn's Lemma) *Let X be a partially ordered set such that every totally ordered subset of X is bounded. Then X contains a maximal element.*

We can use Zorn's Lemma to in fact prove a stronger statement than the existence of a basis (which we already saw for finitely generated vector spaces).

Proposition 1.11 *Let V be a vector space over a field $\mathbb{F}$ and let S be a linearly independent subset. Then there exists a basis B of V containing the set S.*

Proof: Let X be the set of all linearly independent subsets of V that contain S. For $T_1, T_2 \in X$, we say that $T_1 \preceq T_2$ if $T_1 \subseteq T_2$. Let Z be a totally ordered subset of X. Define T^* as

$$T^* := \bigcup_{T \in Z} T = \{v \in V \mid \exists T \in Z \text{ such that } v \in T\} .$$

Then we claim that T^* is linearly independent and is hence in X. It is clear that $T \preceq T^*$ for all $T \in Z$ and this will prove that Z is bounded by T^*. By Zorn's Lemma this shows that X contains a maximal element (say) B, which must be a basis containing S.

To show that T^* is linearly independent, note that we only need to show that no *finite* subset of T^* is linearly dependent. Indeed, let $\{v_1, \ldots, v_n\}$ be a finite subset of T^*. By the definition of T^*, there exists a $T \in X$ such that $\{v_1, \ldots, v_n\} \subseteq T$. Thus, $\{v_1, \ldots, v_n\}$ must be linearly independent. This proves the claim. ∎

1.3 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.

Let $a_1, \ldots, a_n \in \mathbb{R}$ be distinct. Say we want to find the unique (why?) polynomial p of degree at most $n-1$ satisfying $p(a_i) = b_i \; \forall i \in [n]$. 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. Also, 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

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

Exercise 1.12 *Check that the above 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 $|\mathbb{F}| \geq n$.*

1.4 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 $(i, Q(i))$.

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 Linear Transformations

Definition 2.1 *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 2.2 *The following are all linear transformations:*

- *A matrix $A \in \mathbb{R}^{m \times n}$ defines a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.*
- *$\varphi : C([0,1], \mathbb{R}) \to C([0,2], \mathbb{R})$ defined by $\varphi(f)(x) = f(x/2)$.*

- $\varphi : C([0,1],\mathbb{R}) \to C([0,1],\mathbb{R})$ *defined by* $\varphi(f)(x) = f(x^2)$.
- $\varphi : C([0,1],\mathbb{R}) \to C([0,1],\mathbb{R})$ *defined by* $\varphi(f)(x) = f(1-x)$.
- $\varphi_{\text{left}} : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ *defined by* $\varphi_{\text{left}}(f)(n) = f(n+1)$.
- *The derivative operator acting on* $\mathbb{R}[x]$.

Proposition 2.3 *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 $\varphi(v) = \alpha(v)\ \forall v \in B$.*

Definition 2.4 *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 2.5 $\ker(\varphi)$ *is a subspace of V and* $\text{im}(\varphi)$ *is a subspace of W.*

Proposition 2.6 (rank-nullity theorem) *If V is a finite dimensional vector space and $\varphi : V \to W$ is a linear transformation, then*

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

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

Example 2.7 *Consider the matrix A which defines a linear transformation from $\mathbb{F}_2^7$ to $\mathbb{F}_2^3$:*

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

- $\dim(\text{im}(\varphi)) = 3$.
- $\dim(\ker(\varphi)) = 4$.
- *Check that* $\ker(\varphi)$ *is a* code *which can recover from one bit of error.*
- *Check that 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.