Louis Meunier

Algebra 2 — MATH251

Course Outline:

Based on Lectures from Winter, 2024 by Prof. Anush Tserunyan.

Contents

1 Introduction

Remark 1.1. *This course is about vector spaces and linear transformations between them; a vector space involves multiplication by* scalars, *where the scalars come from some field. We recall first examples of fields, then vector spaces, as a motivation, before presenting a formal definition.*

1.1 Vector Spaces

Remark 1.2. *Much of this is recall from Algebra 1.*

⊛ **Example 1.1: Examples of Fields**

1. $\mathbb{Q}$; the field of rational numbers.
2. $\mathbb{R}$; the field of real numbers; $\mathbb{Q} \subseteq \mathbb{R}$.
3. $\mathbb{C}$; the field of complex numbers; $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$.
4. $\mathbb{F}_p \equiv \mathbb{Z}/p\mathbb{Z} \equiv \{0, 1, \ldots, p-1\}; the (unique) field of \mathrm{p} elements, where \mathrm{p} prime.$[a]
   (a) $p = 2$; $\mathbb{F}_2 \equiv \{0, 1\}$.
   (b) $p = 3$; $\mathbb{F}_3 \equiv \{0, 1, 2\}$.
   (c) $\cdots$

[a] where $a +_p b :=$ remainder of $\frac{a+b}{p}$, $a \cdot_p b :=$ remainder of $\frac{a \cdot b}{p}$.

Remark 1.3. *Throughout the course, we will denote an abstract field as* $\mathbb{F}$.

⊛ **Example 1.2: Examples of Vector Spaces**

1. $\mathbb{R}^3 := \{(x, y, z) : x, y, z \in \mathbb{R}\}$. We can add elements in $\mathbb{R}^3$, and multiply them by real scalars.
2. $\mathbb{F}^n := \underbrace{\mathbb{F} \times \mathbb{F} \times \cdots \mathbb{F}}_{n \text{ times}} := \{(a_1, a_2, \ldots, a_n) : a_i \in \mathbb{F}\}$, where $n \in \mathbb{N}$[1]; this is a generalization of the previous example, where we took $n = 3$, $\mathbb{F} = \mathbb{R}$. Operations follow identically; addition:

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) := (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n)$$

and, taking a scalar $\lambda \in \mathbb{F}$, multiplication:

$$\lambda \cdot (a_1, a_2, \ldots, a_n) := (\lambda \cdot a_1, \lambda \cdot a_2, \ldots, \lambda \cdot a_n).$$

We refer to these elements $(a_1, \cdots, a_n)$ as *vectors* in $\mathbb{F}^n$; the vector for which $a_i = 0 \,\forall\, i$ is the 0 *vector*, and is the additive identity, making $\mathbb{F}^n$ an abelian group under addition, that admits multiplication by scalars from $\mathbb{F}$.

3. $C(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} : f \text{ continuous}\}$. Here, we have the constant zero function as our additive identity ($x \mapsto 0 \,\forall\, x$), and addition/scalar multiplication of two continuous real functions are continuous.

4. $\mathbb{F}[t] := \{a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n : a_i \in \mathbb{F} \,\forall\, i, n \in \mathbb{N}\}$, ie, the set of all polynomials in t with coefficients from $\mathbb{F}$. Here, we can add two polynomials;

$$(a_0 + a_1 t + \cdots + a_n t^n) + (b_0 + b_1 t + \cdots + b_m t^m) := \sum_{i=0}^{\max\{n,m\}} (a_i + b_i) t^i,$$

(where we "take" undefined a_i/b_i's as 0; that is, if $m > n$, then $a_{m-n}, a_{m-n+1}, \ldots, a_m$ are taken to be 0). Scalar multiplication is defined

$$\lambda \cdot (a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n) := \lambda a_0 + \lambda a_1 t + \lambda a_2 t^2 + \cdots + \lambda a_n t^n.$$

Here, the zero polynomial is simply 0 (that is, $a_i = 0 \,\forall\, i$).

↪ Definition 1.1: Vector Space

A *vector space* V over a field $\mathbb{F}$ is an *abelian group* with an operation denoted $+$ (or $+_V$) and identity element[2] denoted 0_V, equipped with *scalar multiplication* for each scalar $\lambda \in \mathbb{F}$ satisfying the following axioms:

1. $1 \cdot v = v$ for $1 \in \mathbb{F}$, $\forall\, v \in V$.
2. $\alpha \cdot (\beta \cdot v) = (\alpha \cdot \beta) v$, $\forall\, \alpha, \beta \in \mathbb{F}, v \in V$.
3. $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$, $\forall\, \alpha, \beta \in \mathbb{F}, v \in V$.
4. $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$, $\forall\, \alpha \in \mathbb{F}, u, v \in V$.

We refer to elements $v \in V$ as *vectors*.

↪ Proposition 1.1

For a vector space V over a field $\mathbb{F}$, the following holds:

1. $0 \cdot v = 0_V$, $\forall\, v \in V$ (where $0 := 0_{\mathbb{F}}$)
2. $-1 \cdot v = -v$, $\forall\, v \in V$ (where $1 := 1_{\mathbb{F}}$)[3]
3. $\alpha \cdot 0_V = 0_V$, $\forall\, \alpha \in \mathbb{F}$

Proof. 1. $0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v \implies 0 \cdot v = 0_V$ (by "cancelling" one of the $0 \cdot v$ terms on each side).

[1] Where we take $0 \in \mathbb{N}$, for sake of consistency. Moreover, by convention, we define $\mathbb{F}^0$ (that is, when $n = 0$) to be $\{0\}$; the trivial vector space.
[2] The "zero vector".
[3] NB: "additive inverse"

2. $v + (-1 \cdot v) = (1 \cdot v + (-1) \cdot v) = (1 - 1) \cdot v = 0 \cdot v = 0_V \implies (-1 \cdot v) = -v$.
3. $\alpha \cdot 0_V = \alpha \cdot (0_V + 0_V) = \alpha \cdot 0_V + \alpha \cdot 0_V \implies \alpha \cdot 0_V = 0_V$ (by, again, cancelling a term on each side).

■

↪ *Wed Jan 10 14:16:29 EST 2024*

1.2 Creating Spaces from Other Spaces

↪ **Definition 1.2: Product/Direct Sum of Vector Spaces**

For vector spaces U, V over the same field $\mathbb{F}$, we define their *product* (or *direct sum*) as the set

$$U \times V = \{(u, v) : u \in U, v \in V\},$$

with the operations:

$$(u_1, v_1) + (u_2, v_2) := (u_1 + u_2, v_1 + v_2)$$
$$\lambda \cdot (u, v) := (\lambda \cdot u, \lambda \cdot v)$$

⊛ **Example 1.3:** $\mathbb{F}$

$\mathbb{F}^2 = \mathbb{F} \times \mathbb{F}$, where $\mathbb{F}$ is considered as the vector space over $\mathbb{F}$ (itself).

↪ **Definition 1.3: Subspace**

For a vector space V over a field $\mathbb{F}$, a *subspace* of V is a subset $W \subseteq V$ s.t.

1. $0_V \in W$[4]
2. $u + v \in W \, \forall \, u, v \in W$ (closed under addition)
3. $\alpha \cdot u \in W \, \forall \, u \in W, \alpha \in \mathbb{F}$[5]

Then, W is a vector space in its own right.

⊛ **Example 1.4: Examples of Subspaces**

1. Let $V := \mathbb{F}^n$.

   - $W := \{(x_1, x_2, \ldots, x_n) \in \mathbb{F}^n : x_1 = 0\} = \{(0, x_2, x_3, \ldots, x_n) : x_i \in \mathbb{F}\}$.
   - $W := \{(x_1, x_2, \ldots, x_n) \in \mathbb{F}^n : x_1 + 2 \cdot x_2 = 0\}$

     Proof. Let $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in W$. Then, $x + y = (x_1 + y_1, \ldots, x_n + y_n)$, and $x_1 + y_1 + 2 \cdot (x_2 + y_2) = x_1 + 2 \cdot x_2 + y_1 + 2 \cdot y_2 = 0 + 0 = 0 \implies x + y \in W$. Similar logic follows for axioms 2., 3. ■

   - (More generally)

     $$W := \{(x_1, \ldots, x_n) \in \mathbb{F}^n : \begin{matrix} a_{11}x_1 + \cdots a_{1n}x_n = 0 \\ a_{21}x_1 + \cdots + a_{2n}x_n = 0 \\ \ddots \\ a_{k1}x_1 + \cdots a_{kn}x_n = 0 \end{matrix}\},$$

     that is, a linear combination of homogenous "conditions" on each term.
   - $W^* := \{(x_1, \ldots, x_n) : x_1 + x_2 = 1\}$ is *not* a subspace; it is not closed under addition, nor under scalar multiplication.

2. Let $\mathbb{F}[t]_n := \{a_0 + a_1 t + \cdots + a_n t^n : a_i \in \mathbb{F}\}$. Then, $\mathbb{F}[t]_n$ is a subspace of $\mathbb{F}[t]$, the more general polynomial space. *However*, the set of all polynomials of degree *exactly* n (all axioms fail, in fact) is not a subspace of $\mathbb{F}[t]_n$.

   - $W := \{p(t) \in \mathbb{F}[t]_n : p(1) = 0\}$.
   - $W := \{p(t) \in \mathbb{F}[t]_n : p''(t) + p'(t) + 2p(t) = 0\}$.

3. Let $V := C(\mathbb{R})$ be the space of continuous function $\mathbb{R} \to \mathbb{R}$.

[4] This is equivalent to requiring that $W \neq \varnothing$; stated this way, axiom 3. would necessitate that $0 \cdot w = 0_V \in W$.

[5] Note that these axioms are equivalent to saying that W is a subgroup of V with respect to vector addition; 2. ensures closed under addition, and 3. ensures the existence of additive inverses (as per $-1 \cdot v = -v$).

- $W := \{f \in C(\mathbb{R}) : f(\pi) + 7f(\sqrt{2}) = 0\}$.
- $W := C^1(\mathbb{R}) :=$ everywhere differentiable functions.
- $W := \{f \in C(\mathbb{R}) : \int_0^1 f \, \mathrm{d}x = 0\}$.

↪ **Proposition 1.2**

Let W_1, W_2 be subspaces of a vector space V over $\mathbb{F}$. Then, define the following:

1. $W_1 + W_2 := \{w_1 + w_2 : w_1 \in W_1, w_2 \in W_2\}$
2. $W_1 \cap W_2 := \{w \in V : w \in W_1 \wedge w \in W_2\}$

These are both subspaces of V.

Proof.

1. (a) $0_V \in W_1$ and $0_V \in W_2 \implies 0_V = 0_V + 0_V \in W_1 + W_2$.
   (b) $(u_1 + u_2) + (v_1 + v_2) = (u_1 + v_1) + (u_2 + v_2) \in W_1 + W_2$.
   (c) $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v \in W_1 + W_2$
2. (a) $0_V \in W_1$ and $0_V \in W_2 \implies 0_V = 0_V + 0_V \in W_1 \cap W_2$.
   (b) $u, v \in W_1 \cap W_2 \implies u + v \in W_1 \wedge u + v \in W_2 \implies u + v \in W_1 \cap W_2$.
   (c) $\alpha \cdot u \in W_1 \wedge \alpha \cdot u \in W_2 \implies \alpha \cdot u \in W_1 \cap W_2$.

■

1.3 Linear Combinations and Span

↪ **Definition 1.4: Linear Combination**

Let V be a vector space over a field $\mathbb{F}$. For finitely many vectors $v_1, v_2, \ldots, v_n$, their *linear combination* is a sum of the form

$$\sum_{i=1}^{n} a_i v_i = a_1 \cdot v_1 + \cdots + a_n \cdot v_n,$$

where $a_i \in \mathbb{F} \, \forall \, i$.

A linear combination is called *trivial* if $a_i = 0 \, \forall \, i$, that is, all coefficients are 0.

If $n = 0$ (ie, we are "summing up" 0 vectors), we define the sum as the zero vector; $\sum_{i=1}^{0} a_i v_i := 0_V$.

↪ *Tue Jan 23 15:04:02 EST 2024*

↪ **Definition 1.5: A More General Definition of Linear Combination**

For a (possibly infinite) set S of vectors from V, a *linear combination* of vectors in S is a linear combination of

$a_1v_1 + \cdots a_nv_n$ for some finite subset $\{v_1, \ldots, v_n\} \subseteq S$.[6]

↪ **Definition 1.6: Span**

For a subset $S \subseteq V$, we define its *span* as

$$\operatorname{Span}(S) := \text{ set of all linear combinations of } S := \{a_1v_1 + \cdots a_nv_n : a_i \in \mathbb{F}, v_i \in S\}.$$

By convention, we set $\operatorname{Span}(\varnothing) = \{0_V\}$.

⊛ **Example 1.5**

Let $S := \{(1, 0, -1), (0, 1, -1), (1, 1, -2)\} \subseteq \mathbb{R}^3$. Then,

$$0_{\mathbb{R}^3} = (0, 0, 0) = 1 \cdot (1, 0, -1) + 1 \cdot (0, 1, -1) + -1 \cdot (1, 1, -2).$$

We claim, moreover, that $\operatorname{Span}(S) = U := \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$ (a plane through the origin).

Proof. Note that $S \subseteq U$, hence $S \subseteq \operatorname{Span} S \subseteq U$. OTOH, if $(x, y, z) \in U$, we have $z = -x - y$, and so

$$(x, y, z) = (x, y, -x - y) = x \cdot (1, 0, -1) + y \cdot (0, 1, -1) \in \operatorname{Span}(S)$$

hence $U \subseteq \operatorname{Span}(S)$ and thus $\operatorname{Span}(S) = U$. ■

Remark 1.4. *We implicitly used the following claim in the proof above; we prove it more generally.*

↪ **Proposition 1.3**

Let V be a vector space over $\mathbb{F}$ and let $S \subseteq V$. Then, $\operatorname{Span}(S)$ is always a subspace. Moreover, it is the smallest (minimal) subspace containing S (that is, for any subspace $U \supseteq S$, we have that $U \supseteq \operatorname{Span} S$).

Proof. Because adding/scalar multiplying linear combinations of elements of S again results in a linear combination of elements of S, and $0_V \in \operatorname{Span}(S)$ by definition, we have that $\operatorname{Span}(S)$ is indeed a subspace.

If $U \supset S$ is a subspace of V containing S, then by definition U is closed under addition, that is, taking linear combinations of its elements (in particular, of elements of S); hence, $U \supset \operatorname{Span}(S)$. ■

↪ **Lemma 1.1**

For $S \subseteq V$ and $v \in V$, $v \in \operatorname{Span}(S) \iff \operatorname{Span}(S \cup \{v\}) = \operatorname{Span}(S)$.

Proof. ($\implies$) Let $v \in \operatorname{Span}(S) \implies v = a_1v_1 + \cdots a_nv_n, a_i \in \mathbb{F}, v_i \in V$. Then, for any linear combination

$$b_1u_1 + \cdots b_mu_m + b \cdot v = b_1u_1 + \cdots b_mu_m + b(a_1v_1 + \cdots + a_nv_n)$$

[6]That is, we do not allow infinite sums.

is a linear combination of vectors in $S \cup \{v\}$ (first equality) or equivalently, a combination of vectors in S (second equality) and thus $\operatorname{Span}(S \cup \{v\}) \subseteq \operatorname{Span} S$. The reverse inclusion follows trivially.

$(\Longleftarrow)$ $\operatorname{Span}(S \cup \{v\}) = \operatorname{Span} S \implies v \in \operatorname{Span}(S)$. ■

⊛ **Example 1.6**

(From the above example) We have

$$\operatorname{Span}(\{(1,0,-1),(0,1,-1)\} \cup \{(1,1,-2)\}) = \operatorname{Span}(\{(1,0,-1),(0,1,-1)\}),$$

since $(1,1,-2) \in \operatorname{Span}(\{(1,0,-1),(0,1,-1)\})$ (it was redundant, as it could be generated by the other two vectors).

↪ **Definition 1.7: Spanning Set**

Let V be a vector space over a field $\mathbb{F}$. We call $S \subseteq V$ a *spanning set* for V if $\operatorname{Span}(S) = V$. We call such a spanning set *minimal* if no proper subset of S is a spanning set ($\nexists v \in S$ s.t. $S \setminus \{v\}$ spanning).

Remark 1.5. *Note that any $S \subseteq V$ is a spanning for $\operatorname{Span}(S)$. But, S may not be minimal; indeed, consider the previous example. We were able to remove a vector from S while having the same span.*

⊛ **Example 1.7**

For $\mathbb{F}^n$ as a vector space over $\mathbb{F}$, the *standard spanning set*

$$\underset{n}{\mathrm{St}} := \{\underbrace{(1,\ldots,0)}_{:=e_1}, \underbrace{(0,1,0,\ldots,0)}_{:=e_2}, \ldots, \underbrace{(0,\ldots,1)}_{e_n}\}.$$

Given any $x := (x_1, \ldots, x_n) \in \mathbb{F}^n$, we can write

$$x = x_1 \cdot e_1 + \cdots x_n \cdot e_n.$$

This is clearly minimal; removing any e_i would then result in a 0 in the ith "coordinate" of a vector, hence $\mathrm{St} \setminus \{e_i\}$ would span only vectors whose ith coordinate is 0.

↪ **Definition 1.8: Linear Dependence**

Let V be a vector space over a field $\mathbb{F}$. A set $S \subseteq V$ is said to be *linearly dependent* if there is a nontrivial linear combination of vectors in S that is equal to 0_V.

Conversely, S is called *linearly independent* if there is no nontrivial linear combination of vectors in S that is equal to 0_V; all linear combinations of vectors in S that equal 0_V are trivial.

↪ *Tue Jan 23 15:04:36 EST 2024*

⊛ **Example 1.8**

1. The empty set $\varnothing$ is linearly independent; there are no non-trivial linear combinations that equal 0_V (there are no linear combinations at all).

2. For $v \in V$, the set $\{v\}$ is linearly dependent iff $v = 0_V$.

3. $S := \{(1,0,-1),(0,1,-1),(1,1,-2)\} := \{v_1, v_2, v_3\}$; S is linearly dependent ($v_1 + v_2 - v_3 = (0,0,0)$).

4. $V := \mathbb{F}^3$; $S := \{(1,0,-1),(0,1,-1),(0,0,1)\} = \{v_1, v_2, v_3\}$ is linearly independent.

   Proof. Suppose

$$\begin{aligned}
a_1v_1 + a_2v_2 + a_3v_3 = 0_V& \\
&\implies a_1 = 0 \wedge a_2 = 0 \wedge -a_1 - a_2 + a_3 = 0 \implies a_3 = 0 \\
&\implies a_1 = a_2 = a_3 = 0
\end{aligned}$$

   Hence only a trivial linear combination is possible. ■

5. St_n is linearly independent.

   Proof.

$$\sum_{i=1}^{n} a_i e_i = 0_{\mathbb{F}^n} \implies a_i = 0 \,\forall\, i$$

■

↪ **Lemma 1.2**

Let V be a vector space over a field $\mathbb{F}$, and $S \subseteq V$ (possibly infinite).

1. S is linearly dependent $\iff$ there is a finite subset $S_0 \subseteq S$ that is linearly dependent.

2. S is linearly independent $\iff$ all finite subsets of S are linearly independent.

Proof. 2. follows from the negation of 1.

($\impliedby$) Trivial.

($\implies$) Suppose S linearly dependent. Then, $0_V =$ some nontrivial linear combination of vectors $v_1, \ldots, v_n$ in S. Let $S_0 = \{v_1, \ldots, v_n\}$, then, S_0 is linearly dependent itself. ■

1.4 Linear Dependence and Span

↪ **Proposition 1.4**

Let V be a vector space over a field $\mathbb{F}$and $S \subseteq V$.

1. S linearly dependent $\iff \exists v \in \operatorname{Span}(S \setminus \{v\})$.
2. S linearly independent $\iff$ there is no $v \in \operatorname{Span}(S \setminus \{v\})$.

Proof. 2. follows from the negation of 1.

($\implies$) Suppose S linearly dependent. Then, $0_V = \sum_{i=1}^n a_i v_i$ for some nontrivial linear combination of distinct vectors S. At least one of $a_i \neq 0$; we can assume wlog (reindexing) $a_1 \neq 0$. Then,

$$a_1 v_1 = -\sum_{i=2}^n a_i v_i \implies v_1 = (-a_1^{-1}) \sum_{i=2}^n a_i v_i = \sum_{i=2}^n (-a_1^{-1} a_i) v_i,$$

hence, $v_1 \in \operatorname{Span}(\{v_2, \ldots, v_n\}) \subseteq \operatorname{Span}(S \setminus \{v\})$

($\impliedby$) Suppose $v \in \operatorname{Span}(S \setminus \{v\})$, then $v = a_1 v_1 + \cdots + a_n v_n$, with $v_1, \ldots, v_n \in S \setminus \{v\}$, thus

$$0_V = a_1 v_1 + \cdots a_n v_n - v,$$

which is not a trivial combination (-1 on the v; v cannot "merge" with the other vectors), hence S is linearly dependent. ■

↪ **Corollary 1.1**

$S \subseteq V$ is linearly independent $\iff$ S a minimal spanning set of $\operatorname{Span} S$.

Proof. Follows from proposition 1.4, 2. ■

↪ **Definition 1.9: Maximally Independent**

Let V be a vector space over a field $\mathbb{F}$. A set $S \subseteq V$ is called *maximally independent* if S is linearly independent and $\nexists v \in V \setminus S$ s.t. $S \cup \{v\}$ is still linearly independent.

In other words, there is no proper supset $\tilde{S} \supsetneq S$ that is still independent.

↪ **Lemma 1.3**

If $S \subseteq V$ maximally independent, then S is spanning for V.

Proof. Let $S \subseteq V$ be maximally independent. Let $v \in V$; supposing $v \notin S$ (in the case that $v \in S$, then $v \in \operatorname{Span}(S)$ trivially). By maximality, $S \cup \{v\}$ is linearly dependent, hence there exists a nontrivial linear combination that equals

0_V. Since S independent, this combination must include v, with a nonzero coefficient. We can write

$$
\begin{aligned}
& av + \sum_{i=1}^{n} a_i v_i = 0_V \quad a \neq 0, v_i \in S \\
\implies & v = \sum_{i=1}^{n} (-a^{-1} a_i) v_i \in \operatorname{Span} S.
\end{aligned}
$$

■

↪ **Theorem 1.1**

Let V be a vector space over a field $\mathbb{F}$and let $S \subseteq V$. TFAE:

1. S is a minimal spanning set;
2. S is linearly independent and spanning;
3. S is a maximally linearly independent set;
4. Every vector in V is equal to *unique* linear combination of vectors in S.

↪ *Mon Jan 15 13:44:34 EST 2024*

Proof. (1. $\implies$ 2.) Suppose S is spanning for V and is minimal. Then, by corollary 1.1, we have that S is linearly independent, and is thus both linearly independent and spanning.

(2. $\implies$ 3.) Suppose S is linearly independent and spanning. Let $v \in V \setminus S$; S is spanning, hence $v \in \operatorname{Span} S$, that is, there exists a linear combination of vectors in S that is equal to v:

$$
v = a_1 v_1 + \cdots + a_n v_n, a_i \in \mathbb{F}, v_i \in S.
$$

Thus, $0_V = a_1 v_1 + \cdots + a_n v_n - v$, thus $S \cup \{v\}$ is linearly dependent, and so S is maximally linearly independent.

(3. $\implies$ 1.) Suppose S is maximally linearly independent. By lemma 1.3, S is spanning, and since S is linearly independent, by corollary 1.1, S is minimally spanning for $\operatorname{Span} S$.

(2. $\implies$ 4.) Suppose S is linearly independent and spans V, and let $v \in V$. We have that $v \in \operatorname{Span} S$ and hence is equal to a linear combination of vectors in S. This gives existence; we now need to prove uniqueness.

Suppose there exist two linear combinations that equal v,

$$
v = a_1 v_1 + \cdots + a_n v_n = b_1 u_1 + \cdots + b_m u_m,
$$

$a_i, b_j \in \mathbb{F}$, $v_i, u_j \in S$. With appropriate reindexing/relabelling and allowing certain scalars to equal 0, we can assume that the combinations use the same vectors (with potentially different coefficients), that is,

$$
v = a_1 w_1 + \cdots + a_k w_k = b_1 w_1 + \cdots + a_k w_k.
$$

This implies, then,

$$
(a_1 - b_1) w_1 + \cdots + (a_k - b_k) w_k = 0_V,
$$

and by the assumed linear independent of S, each coefficient $(a_i - b_i) = 0 \,\forall\, i \implies a_i = b_i \,\forall\, i$, hence, these are indeed the same representations, and thus this representation is unique.

(4. $\implies$ 2.) Suppose every vector in V admits a unique linear combination of vectors in S. Clearly, then, S is spanning. It remains to show S is linearly independent. Suppose

$$0_V = a_1 v_1 + \cdots + a_n v_n$$

for $v_i \in S$. But we have that every vector has a unique representation, and we know that $a_i = 0 \,\forall\, i$ is a (valid) linear combination that gives 0_V; hence, this must be the unique combination, $a_i = 0 \,\forall\, i$, and the linear combination above is trivial. Hence, S is linearly independent and spanning. ∎

↪ **Definition 1.10: Basis**

If any (hence all) of the above statements hold, we call S a *basis* for V.

In the words of 4., we call the unique linear combination of vectors in S that is equal to v the *unique representation of v in S*. Its coefficients are called the *Fourier coefficients of v in S*.

⊛ **Example 1.9**

1. $\mathrm{St}_n = \{e_i : 1 \leq i \leq n\}$ is a basis for $\mathbb{F}^n$.
2. In $\mathbb{F}^3$, the set
$$\{(1, 0, -1), (0, 1, -1), (0, 0, 1)\}$$
is a basis; it is linearly independent and spanning.
3. For $\mathbb{F}[t]_n$, the standard basis is
$$\{1, t, t^2, \dots, t^n\}.$$
4. For $\mathbb{F}[t]$, the standard basis is
$$S := \{1, t, t^2, \dots\} = \{t^n : n \in \mathbb{N}\}.$$
5. Let $\mathbb{F}[\![t]\!]$ denote the space of all formal power series $\sum_{n\in\mathbb{N}} a_n t^n$; polynomials are an example, but with only finite nonzero coefficients. Note that, then, the set S defined above is not a basis for this "extended" set. We *can* in fact find a basis for this set; we need more tools first.

↪ **Theorem 1.2**

Every vector space has a basis.

Remark 1.6. *This theorem relies on assuming the Axiom of Choice.*

↪ *Wed Jan 17 13:37:26 EST 2024*

Proof (Attempt). (Of theorem 1.2) We will try to "inductively" build a maximally independent set, as follows:

Begin with an empty set $S_0 := \varnothing$, and iteratively add more vectors to it. Let $v_0 \in V$ be a non-zero vector, and let $S_1 := \{v_0\}$.

If S_1 is maximal, then we are done. Otherwise, there exists a new vector $v_1 \in V \setminus S_1$ s.t. $S_2 := \{v_0, v_1\}$ is still independent.

If S_2 is maximal, then we are done. Otherwise, there exists a new vector $v_2 \in V \setminus S_2$ s.t. $S_3 := \{v_0, v_1, v_2\}$ is still independent.

Continue in this manner; this would take arbitrarily many finite, or even infinite, steps; we would need some "choice function" that would "allow" us to choose any particular ith vector v_i.

We can make this construction precise via the Axiom of Choice and transfinite induction (on ordinals); alternatively, we will prove a statement equivalent to the Axiom of Choice, Zorn's Lemma. ■

Remark 1.7. *Before stating Zorn's Lemma, we introduce the following terminology.*

↪ **Axiom 1.1: Axiom of Choice**

Let X be a set of nonempty sets. Then, there exists a choice function f defined on X that maps each set of X to an element of that set.

↪ **Definition 1.11: Inclusion-Maximal Element**

A *inclusion-maximal* element of I is a set $S \in I$ s.t. there is no strict super set $S' \supsetneq S$ s.t. $S' \in I$.

↪ **Definition 1.12: Chain**

Let X a set. Call a collection $\mathcal{C} \subseteq \mathcal{P}(X)$ a *chain* if any two $A, B \in \mathcal{C}$ are comparable, ie, $A \subseteq B$ or $B \subseteq A$.

↪ **Definition 1.13: Upper Bound**

An *upper bound* of a collection $\tau \subseteq \mathcal{P}(X)$ is a set $U \subseteq X$ s.t. $U \supseteq J \, \forall \, J \in \tau$; U contains the union of all sets in J.

⊛ **Example 1.10: Of The Previous Definitions**

Let $X := \mathbb{N}, I := \{\varnothing, \{0\}, \{1, 2\}, \{1, 2, 3\}\} \subseteq \mathcal{P}(\mathbb{N})$.

The maximal elements of I would be $\{0\}$ and $\{1, 2, 3\}$.

Chains would include $\mathcal{C}_0 := \{\varnothing, \{1, 2\}, \{1, 2, 3\}\}, \mathcal{C}_1 := \{\varnothing, \{0\}\}, \mathcal{C}_2 := \{\varnothing\}$ (or any set containing a single element).

The sets $\{0, 1, 2, 3\}$ and $\{0, 1, 2, 3, 4, 5\}$ are upper bounds for I, while neither is an element of I. The set $\{1, 2, 3\}$ is an upper bound for $\mathcal{C}_0$. A chain $\{\varnothing, \{0\}, \{0, 1\}, \{0, 1, 2\}, \ldots\}$ has an upper bound of $\mathbb{N}$.

↪ **Lemma 1.4: Zorn's Lemma**

Let X be an ambient set and $I \subseteq \mathcal{P}(X)$ be a nonempty collection of subsets of X. If every chain $\mathcal{C} \subseteq I$ has an upper bound in I, then I has a maximal element.

"Proof". This is equivalent to the Axiom of Choice; proving it is beyond the scope of this course :(. ■

Proof of theorem 1.2, cnt'd. We obtain a maximal independent set using Zorn's Lemma.

Let I be the collection of all linearly independent subsets of V. I is nonempty; $\varnothing \in I$, as is $\{v\} \in I$ for any nonzero $v \in V$. To apply Zorn's, we need to show that every chain $\mathcal{C}$ if sets in I has an upper bound in I; that is, every linearly independent set has an upper bound that itself is linearly independent.

Let $\mathcal{C}$ be a chain in I. Let $S := \bigcup \mathcal{C}$ be the union of all sets in $\mathcal{C}$. To show S is linearly independent, it suffices to show that every finite subset $\{v_1, \dots, v_n\} \subseteq S$ is linearly independent. Let $S_i \in \mathcal{C}$ be s.t. $v_i \in S_i$ for each i. Because $\mathcal{C}$ a chain, for each i, j we have either $S_i \subseteq S_j$ or $S_j \subseteq S_i$, and so we can order $S_1, \dots, S_n$ in increasing order w.r.t $\subseteq$. This implies, then, there is a maximal S_{i_0} s.t. $S_{i_0} \supseteq S_i \, \forall \, i \in \{1, \dots, n\}$. Moreover, we have that $\{v_1, \dots, v_n\} \in S_{i_0}$, and that S_{i_0} is linearly independent and thus $\{v_1, v_2, \dots, v_n\}$ is also linearly independent.

Thus, as we can apply Zorn's Lemma, we conclude that I has a maximal element, ie, there is a maximal independent set, and thus a V indeed has a basis. ■

↪ *Fri Jan 19 13:36:58 EST 2024*

↪ **Theorem 1.3**

For every vector space V over a field $\mathbb{F}$, any two bases $\mathcal{B}_1, \mathcal{B}_2$ are equinumerous/of equal size/cardinality, ie, there is a bijection between $\mathcal{B}_1$ and $\mathcal{B}_2$.

Remark 1.8. *We will only prove this for vector spaces that admit a finite basis.*

↪ **Lemma 1.5: Steinitz Substitution**

Let V be a vector space over a field $\mathbb{F}$. Let $Y \subseteq V$ be a (possibly infinite) linearly independent set and let $Z \subseteq V$ be a finite spanning set. Then:

1. $k := |Y| \leq |Z| =: n$
2. There is $Z' \subseteq Z$ of size $n - k$ s.t. $Y \cup Z'$ is still spanning.

Proof. We prove by induction on k.

$k = 0$ gives that $Y = \varnothing$, and so $Z' = Z$ itself works ($Z' \cup Y = Z$) as a spanning set.

Suppose the statement holds for some $k \geq 0$. Let Y be an independent set such that $|Y| = k + 1$, ie

$$Y := \{y_1, y_2, \dots, y_k, y_{k+1}\}, \quad y \in V.$$

By our inductive assumption, we can consider $Y' := \{y_1, \dots, y_k\} \subseteq Y$ of size k, to obtain a set

$$Z' = \{z_1, z_2, \dots, z_{n-k}\} \subseteq Z, \text{ s.t. } Y' \cup Z' = \{y_1, \dots, y_k, z_1, \dots, z_{n-k}\}$$

is spanning. As this is spanning, we can write y_{k+1} as a linear combination of vectors in $Y' \cup Z'$, ie

$$y_{k+1} = a_1 y_1 + \cdots + a_k y_k + b_1 z_1 + \cdots + b_{n-k} z_{n-k}, \quad a_i, b_j \in \mathbb{F}.$$

It must be that at least one of b_j's must be nonzero; if they were all zero, then y_{k+1} would simply be a linear combination of vector y_i giving that y_{k+1} linearly dependent, contradicting our construction of Y linearly independent.

Assume, wlog, $b_{n-k} \neq 0$. Then, we can write

$$z_{n-k} = b_{n-k}^{-1} y_{k+1} - b_{n-k}^{-1} a_1 y_1 - \cdots - b_{n-k}^{-1} a_k y_k - b_{n-k}^{-1} b_1 z_1 - \cdots - b_{n-k}^{-1} b_{n-k-1} z_{n-k-1},$$

and hence

$$z_{n-k} \in \operatorname{Span}\{y_1, \ldots, y_{k+1}, z_1, \ldots, z_{n-k-1}\} = \operatorname{Span}\left(\underbrace{\{y_1, \ldots, y_{k+1}\}}_{Y} \cup \underbrace{\{z_1, \ldots, z_{n-k-1}\}}_{:=Z''}\right).$$

We had that $Y' \cup Z'$ was spanning, and $(Y' \cup Z') \setminus (Y \cup Z'') = \{z_{n-k}\} \subseteq \operatorname{Span}(Y \cup Z'')$, and we thus have that $Y \cup Z''$ is also spanning. ■

↪ **Corollary 1.2: Finite Basis Case for theorem 1.3**

Let V be a vector space that admits a finite basis. Then, any two bases of V are equinumerous.

Proof. Let Y, Z be two finite bases for V. Then, Y is independent and Z is spanning, so by Steinitz Substitution, $|Y| \leq |Z|$. OTOH, Z is independent, and Y is spanning, so by Steinitz Substitution, $|Z| \leq |Y|$, and we conclude that $|Y| = |Z|$. Let $n := |Y|$.

It remains to show that there exist no infinite bases for V; it suffices to show that there is no independent set of size $n+1$. To this end, let $I \subseteq V$ such that $|I| = n+1$ be an independent set. Y is still spanning, hence, by the substitution lemma, $n+1 \leq n$, a contradiction. Hence, I as defined cannot exist and so any basis of V must be of size n. ■

↪ **Definition 1.14: Dimension**

Let V be a vector space over a field $\mathbb{F}$. The *dimension* of V, denote

$$\dim(V)$$

as the cardinality/size of any basis for V. We call V *finite dimensional* if $\dim(V)$ is a natural number, i.e. V admits a finite basis. Otherwise, we say V is infinite dimensional.

↪ **Corollary 1.3: of Steinitz Substitution**

Let V be a finite dimensional vector space over $\mathbb{F}$ and denote $n := \dim(V)$. Then:

1. Every linearly independent subset $I \subseteq V$ has size $\leq n$;

2. Every spanning set $S \subseteq V$ for V has size $\geq n$;

3. Every independent set I can be completed to a basis to V, ie, there exists a basis B for V s.t. $I \subseteq B$.

Proof. Fix a basis B for V, $|B| =: n$.

1. If I is a independent set, then because B spanning, Steinitz Substitution gives $|I| \leq |B|$.
2. If S spanning for V, then because B is linearly independent, Steinitz Substitution gives $|B| \leq |S|$.
3. Let I be an independent set. Then, because B is spanning, Steinitz Substitution gives $B' \subseteq B$ of size $n - |I|$ s.t. $I \cup B'$ is spanning. Moreover, $|I \cup B'| \leq n$, and by 2. it must have size $\geq n$, and thus has size precisely n and is thus a minimally spanning set and thus a basis.

■

↪ **Corollary 1.4: Monotonicity of Dimension**

Let V be a vector space over a field $\mathbb{F}$. For any subspace $W \subseteq$, $\dim W \leq \dim V$, and

$$\dim W = \dim V \iff W = V.$$

Proof. Let $B \subseteq W$ be a basis for W. Because B is independent, $|B| \leq \dim(V)$ by 1. of corollary 1.3, so $\dim(W) = |B| \leq \dim(V)$.

If $|B| = \dim(V)$, then B is a basis for V again by 1. of corollary 1.3, so $W = \operatorname{Span}(B) = V$. ■

↪ *Mon Jan 22 13:43:44 EST 2024*

2 Linear Transformations

2.1 Definitions

↪ **Definition 2.1: Linear Transformation**

Let V, W be vector spaces over a field $\mathbb{F}$. A function $T : V \to W$ is called a *linear transformation* if it preserves the vector space structures, that is,

1. $T(v_0 + v_1) = T(v_0) + T(v_1)$, $\forall\, v_0, v_1 \in V$;
2. $T(\alpha \cdot v) = \alpha \cdot T(v)$, $\forall\, \alpha \in \mathbb{F}, v \in V$;
3. $T(0_V) = 0_W$.

Remark 2.1. *Note that 3. is redundant, implied by 2., but included for emphasis:*

$$T(0_V) = T(0_{\mathbb{F}} \cdot 0_V) = 0_{\mathbb{F}} \cdot T(0_V) = 0_W.$$

⊛ **Example 2.1: Linear Transformations**

1. $T : \mathbb{F}^2 \to \mathbb{F}^2, T(a_1, a_2) := (a_1 + 2a_2, a_1)$.
2. Let $\theta \in \mathbb{R}$, and let $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be the rotation by θ. The linearity of this is perhaps most obvious in polar coordinates, ie $v \in \mathbb{R}^2, v = r(\cos\alpha, \sin\alpha)$ for appropriate r, α, and $T_\theta(v) = r(\cos(\alpha+\theta), \sin(\alpha+\theta))$.
3. $T : \mathbb{R}^2 \to \mathbb{R}^2$, a reflection about the x-axis, ie, $T(x, y) = (x, -y)$.
4. Projections, $T : \mathbb{F}^n \to \mathbb{F}^n$.
5. The transpose on $M_n(\mathbb{F})$, ie, $T : M_n(\mathbb{F}) \to M_n(\mathbb{F})$, where $A \mapsto A^t$.
6. The derivative on space of polynomials of degree leq n, $D : \mathbb{F}[t]_{n+1} \to \mathbb{F}[t]_n, p(t) \mapsto p'(t)$.

↪ **Theorem 2.1**

Linear transformations are completely determined by their values on a basis.

That is, let $\mathcal{B} := \{v_1, \dots, v_n\}$ be a basis for a vector space V over $\mathbb{F}$. Let W also be a vector space over $\mathbb{F}$ and let $w_1, \dots, w_n \in W$ be arbitrary vectors. Then, there is a unique linear transformation $T : V \to W$ s.t. $T(v_i) = w_i \, \forall \, i = 1, \dots, n$.

Proof. We aim to define $T(v)$ for arbitrary $v \in V$. We can write

$$v = a_1 v_1 + \cdots + a_n v_n$$

as the unique representation of v in terms of the basis $\mathcal{B}$. Then, we simply define

$$T(v) := a_1 w_1 + \cdots + a_n w_n,$$

for our given w_i's. Then, $T(v_i) = 1 \cdot w_i = w_i$, as desired, and T is linear;

1. Let $u, v \in V; u := \sum_n a_i v_i, v := \sum_n b_i v_i$. Then,

$$T(u + v) = T(\sum_n a_i v_i + \sum_n b_i v_i) = T(\sum (a_i + b_i) v_i) = (a_i + b_i) \sum_n w_i = \sum_n a_i w_i + \sum_n b_i w_i.$$

2. Scalar multiplication follows similarly.

To show uniqueness, suppose T_0, T_1 are two linear transformations satisfying $T_0(v_i) = w_i = T_1(v_i)$. Let $v \in V$, and write $v = \sum_n a_i v_i$. By linearity,

$$T_k(v) = T_k(\sum_n a_i v_i) = \sum_n a_i T(v_i) = \sum_n a_i w_i,$$

for $k = 0, 1$, hence, $T_1(v) = T_0(v)$ for arbitrary v, hence the transformations are equivalent. ■

↪ **Definition 2.2: Some Important Transformations**

We denote $T_0 : V \to W$ by $T_0(v) := 0_W \,\forall\, v \in V$ the *zero transformation*. We denote $I_V : V \to V$, $I_V(v) := v \,\forall\, v \in V$, as the *identity transformation*.

↪ *Mon Jan 22 14:26:24 EST 2024*

2.2 Isomorphisms, Kernel, Image

↪ **Definition 2.3: Isomorphism**

Let V, W be vector spaces over $\mathbb{F}$. An *isomorphism* from V to W is a linear transformation $T : V \to W$ (a homomorphism for vector spaces) which admits an inverse T^{-1} that is also linear.

If such an isomorphism exists, we say V and W are *isomorphic*.

↪ **Proposition 2.1**

$T : V \to W$ is an isomorphism $\iff$ T is linear and bijective.

Proof. The direction $\implies$ is trivial.

Suppose $T : V \to W$ is linear and bijective, ie T^{-1} exists. We need to show that T^{-1} is linear. Let $w_1, w_2 \in W, a_1, a_2 \in \mathbb{F}$. Then:

$$\begin{aligned} T^{-1}(a_1w_1 + a_2w_2) &= T^{-1}(a_1T(T^{-1}(w_1)) + a_2T(T^{-1}(w_2))) \\ \text{(by linearity of } T\text{)} \quad &= T^{-1}(T(a_1T^{-1}(w_1) + a_2T^{-1}(w_2))) \\ &= a_1T^{-1}(w_1) + a_2T^{-1}(w_2). \end{aligned}$$

■

Remark 2.2. *This proposition holds for all structures that only have operations; it does not for those with relations, such as graphs, orders, etc..*

↪ **Theorem 2.2**

For $n \in \mathbb{N}$, every n-dimensional vector space V over $\mathbb{F}$ is isomorphic to $\mathbb{F}^n$. In particular, all n-dim vector spaces over $\mathbb{F}$ are isomorphic.

Proof. Fix a basis $\mathcal{B} := \{v_1, \ldots, v_n\}$ for V, and let $T : V \to \mathbb{F}^n$ be the unique linear transformation determined by $\mathcal{B}$ with $T(v_i) = e_i$, where $\{e_1, \ldots, e_n\}$ is the standard basis for $\mathbb{F}^n$. We show that T is a bijection.

(Injective) Suppose $T(x) = T(y), x, y \in V$. Write $x = a_1v_1 + \cdots + a_nv_n, y = b_1v_1 + \cdots + b_nv_n$, the unique representation of x, y in the basis $\mathcal{B}$. We have:

$$a_1e_1 + \cdots + a_ne_n = a_1T(v_1) + \cdots + a_nT(v_n) = T(a_1v_1 + \cdots + a_nv_n) = T(x) = T(y) = \cdots = b_1e_1 + \cdots + b_ne_n,$$

but by the uniqueness of representation in a basis, it follows that each $a_i = b_i$, hence, $x = y$.

(Surjective) Let $w \in \mathbb{F}^n$. Then, $w = a_1 e_1 + \cdots + a_n e_n$ (uniquely). But then,

$$w = a_1 T(v_1) + \cdots + a_n T(v_n) = T(a_1 v_1 + \cdots + a_n v_n),$$

where $a_1 v_1 + \cdots + a_n v_n \in V$, hence T indeed surjective. ■

Remark 2.3. *Replacing $\mathbb{F}^n$ with an arbitrary n-dim vector space W over $\mathbb{F}$ yields the following.*

↪ **Theorem 2.3: Freeness of Vector Space**

Let W, V be vector spaces over $\mathbb{F}$ and let β, γ be bases for V, W respectively. Every bijection $T : \beta \to \gamma$ can be extended to an isomorphism $\hat{T} : V \to W$.

In particular, all vector spaces over $\mathbb{F}$ with equinumerous bases are isomorphic.

Remark 2.4. *The proof follows very similarly to the previous theorem, but extended to arbitrary, possible infinite, spaces.*

Proof. ■

↪ **Definition 2.4: Image/Kernel**

For a linear transformation $T : V \to W$, where V, W are vector spaces over $\mathbb{F}$, we define the *image*

$$\mathrm{Im}(T) := T(v),$$

and its *kernel*

$$\ker(T) = T^{-1}(\{0_W\}).$$

↪ **Proposition 2.2**

$\ker(T)$ and $\mathrm{Im}\, T$ are subspaces of V, W resp.

Proof. ($\ker(T)$) Let $v_0, v_1 \in \ker T$ and $a_0, a_1 \in \mathbb{F}$, then

$$T(a_0 v_0 + a_1 v_1) = a_0 T(v_0) + a_1 T(v_1) = 0_W \implies a_0 v_0 + a_1 v_1 \in \ker T.$$

($\mathrm{Im}(T)$) Let $w_0, w_1 \in \mathrm{Im}\, T$, $a_0, a_1 \in \mathbb{F}$. Then $w_i = T(v_i), v_i \in V$, and so

$$a_0 w_0 + a_1 w_1 = a_0 T(v_0) + a_1 T(v_1) = T(a_0 v_0 + a_1 v_1) \implies a_0 w_0 + a_1 w_1 \in \mathrm{Im}\, T.$$

■

↪ **Proposition 2.3**

Let $T : V \to W$ be a linear transformation, where V, W vector spaces over $\mathbb{F}$. Let β be a (possibly infinite) basis for V. Then, $T(B)$ spans $\operatorname{Im}(T)$.

In particular, T is surjective iff $T(\beta)$ spans W.

$\hookrightarrow$ *Wed Jan 24 14:26:13 EST 2024*