Math 115AH

Max Smiley

Winter 2021

This is an honors, introductory course to abstract linear algebra that I took at UCLA. There were three broad categories of the course: Vector Spaces, Linear Transformations, and Inner Product Spaces. In each section, I'll go over definitions, theorems, and some important proofs and corollaries in chronological order of the course. I'll also attempt to give extra commentary and applications to things I find interesting. This is an informal set of notes, mainly intended to be referenced by myself down the road.

Vector Spaces

Def: a *field* is a set F with two maps $+ : F \times F \to F$ and $\cdot : F \times F \to F$ satisfying the following properties

- $\forall\ a, b, c \in F,\ (a + b) + c = a + (b + c)$
- $\forall\ a, b \in F,\ a + b = b + a$
- $\exists!\ 0 \in F,\ such\ that\ a + 0 = 0 + a\quad \forall\ a \in F$
- $\forall\ a \in F,\ \exists\ b \in F,\ such\ that\ a + b = b + a = 0.\ We\ write\ a = -b$
- $\forall\ a, b, c \in F,\ (a \cdot b) \cdot c = a \cdot (b \cdot c)$
- $\forall\ a, b \in F,\ a \cdot b = b \cdot a$
- $\exists!\ 1 \in F,\ such\ that\ 1 \cdot a = a \cdot 1 = a\quad \forall\ a \in F$
- $\forall\ a \neq 0 \in F,\ \exists!\ b \in F\ such\ that\ a \cdot b = b \cdot a = 1.\ We\ write\ a = b^{-1}$
- $\forall\ a, b, c \in F,\ a \cdot (b + c) = a \cdot b + a \cdot c$
- $\forall\ a, b, c \in F,\ a \cdot (b + c) \in F$

This defines two operators on a set, addition and multiplication. Essentially, these rules break down as follows - associativity, commutativity, existence (and uniqueness) of inverses, distributivity and closure of addition and multiplication.

A few examples of fields are the set of all rational numbers $\mathbb{Q}$, the set of all real numbers $\mathbb{R}$, and the set of all complex numbers $\mathbb{C}$. Note that the set of all integers $\mathbb{Z}$ does not form a field, as their multiplicative inverses are *not* in the field.

Note that we also made no mention of what these operations actually have to represent, only that they must follow those four rules. This allows for a greater degree of generality - the multiplication and addition operators don't necessarily have to be what we're used to them being. Defining the notion of a field is also useful, as if we can prove a statement is true about a field, we can apply it to *any* set that fulfills this description. Namely, when we define vector spaces, we define them over a field, which allows us to prove theorems for the reals, complexes, or rationals to name a few.

Though all of the fields I named in the previous paragraph were infinite sets, fields can be finite. Consider the following set, and tables of operations.

$$F = \{0, 1\}$$

+	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

This satisfies all of the field axioms.

Def: a *vector space* over a field F is a set V with two operations $+ : V \times V \to V$ and $\cdot : F \times V \to V$ (referred to as vector addition and scalar multiplication) satisfying the following properties

- $\forall\ v_1, v_2, v_3 \in V,\ (v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$

- $\forall\ v_1, v_2 \in V,\ v_1 + v_2 = v_2 + v_1$

- $\exists!\ 0 \in V,\ such\ that\ v + 0 = 0 + v \quad \forall\ v \in V$

- $\forall\ v \in V,\ \exists\ w \in V,\ such\ that\ v + w = w + v = 0.\ We\ write\ v = -w$

- $\forall\ v \in V,\ v \cdot 1 = 1 \cdot v = v$

- $\forall\ v \in V,\ and\ \alpha, \beta \in F,\ \alpha \cdot (\beta \cdot v) = (\alpha \cdot \beta) \cdot v$

- $\forall\ v \in V,\ and\ \alpha, \beta \in F,\ (\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$

- $\forall\ v_1, v_2 \in V,\ and\ \alpha \in F,\ \alpha \cdot (v_1 + v_2) = \alpha \cdot v_1 + \alpha \cdot v_2$

- $\forall\ v_1, v_2 \in V,\ and\ \alpha \in F,\ \alpha \cdot v_1 + v_2 \in F$

These axioms looks suspiciously similar to the field axioms. In fact, *any field is a vector space over itself.* By this construction, the vector space axioms are inherited by the field axioms. For example, the field $\mathbb{R}$ is a vector space over $\mathbb{R}$. To give an example of a vector space that's a bit more interesting, the set of 3 element vectors over the field $\mathbb{R}$ forms a vector space, denoted as $\mathbb{R}^3$. In general, the set of all n element vectors over the a field F form the vector space denoted as F^n.

Def: a *subspace* of a vector space V is a subset $W \subseteq V$ such that W is also a vector space.

Subspace Theorem: *Let V be a vector space over the field F, and let $W \subseteq V$ be nonempty. Then the following are equivalent.*

- W is a subspace of V

- W is closed under addition and scalar multiplication

- $\forall\, w_1, w_2 \in W$ and $\alpha \in F,\ \alpha \cdot w_1 + w_2 \in W$

Proof: 1) $\Rightarrow$ 2) by vector space axioms. 2) $\Rightarrow$ 3) by definition. Now assume 3). Take $\alpha = 1$ to get $w_1 + w_2 \in W \ \forall \ w_1, w_2 \in W$. Take $\alpha = -1$ to get $w_2 - w_1 \in W \ \forall \ w_1, w_2 \in W$. Finally, take $w_2 = 0_V \in W$ to get $\alpha \cdot w_1 \in W \ \forall \ \alpha \in F$ and $w_1 \in W$. Thus, W is closed under addition and scalar multiplication, so it is a subspace, and so 3) $\Rightarrow$ 1).

Def: let V be a vector space over a field F. A *linear combination* of elements in V is a sum of multiples of elements of V. For example, $\alpha v_1 + \beta v_2 + \gamma v_3$ is a linear combination for $v_1, v_2, v_3 \in V$ and $\alpha, \beta, \gamma \in F$. A vector $v \in V$ is said to be a *linear combination* of vectors $v_i \in V$ if $\exists$ scalars $\alpha_i \in F$ such that $v = \sum_i \alpha_i v_i$.

Def: let V be a vector space over a field F, and $\beta = \{v_1 \ldots v_n\}$ a set of n vectors in V. The *span* of β is the set of all linear combinations of elements of β with scalars from F.

$$span(v_1, \ldots v_n) = \sum_{i=1}^{n} \alpha_i v_i \quad \forall\, \alpha_i \in F$$

Note that if $v_1, \ldots v_n \in V$, then $span(v_1, \ldots v_n)$ is a subspace of V. It is a subset, and is closed under addition and scalar multiplication.

Def: let $W_1, W_2 \subseteq V$ be subspaces. We define the sum

$$W_1 + W_2 = \{w_1 + w_2\ :\ w_1 \in W_1,\ w_2 \in W_2\} = span(W_1 \cup W_2)$$

More generally, if $W_i \subseteq V$ are subspaces indexed by I, we define

$$\sum_I W_i = span\left(\bigcup_I W_i\right)$$

This is a subspace, due to the way in which we defined span. Note that if we had

defined $+$ by taking the union of the two sets, this would not form a subspace, as the unions of two subspaces are not necessarily closed under addition between the two elements.

Def: let $W_1, W_2 \subseteq V$ be subspaces. We define the direct sum $W_1 \oplus W_2$ as $W_1 + W_2$ if $W_1 \cap W_2 = \{0\}$ More generally, if $W_i \subseteq V$ are subspaces indexed by I, we write

$$\bigoplus_I W_i = \sum_I W_i \quad if \quad W_j \cap \sum_{I \setminus \{j\}} W_i = \{0\} \ \forall j \in I$$

Def: a set of vectors in a vector space V is *linearly independent* if

$$\sum_{i=1}^n \alpha_i v_i = 0 \Rightarrow \alpha_i = 0 \ \forall \alpha_i \in F, \ \forall v_i \in V \ for \ not \ all \ v_i = 0$$

Likewise, a set of vectors is *linearly dependent* if

$$\exists \ \alpha_i \neq 0 \ such \ that \ \sum_{i=1}^n \alpha_i v_i = 0 \ is \ true \ for \ not \ all \ v_i = 0$$

Linear (in)dependence is a big deal in linear algebra. Without diving into the mathematics of it (yet), linearly dependent vectors mean that some vector in the set can be expressed by other vectors in the set. In a sense, there is some repeated information. In contrast, each newly added linearly independent vector to a set adds new information.

Toss In Theorem: *Let V be a vector space over F, and $\varnothing \neq S \subseteq V$ be a linearly independent subset of V. Suppose $v \in V$ is not in the span of S. Then $S \cup \{v\}$ is linearly independent.*

Proof: Suppose $S \cup \{v\}$ is linearly dependent. Then $\exists \ v_1, \ldots v_n \in S$ and $\alpha_i, \ldots \alpha_n \in F$ not all zero such that

$$\alpha_1 v_1 + \ldots + \alpha_n v_n + \alpha v = 0$$

Suppose $\alpha = 0$. Then our equation reads $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$ for not all $\alpha_i = 0$. However, this contradicts the linear independence of S.

Now suppose $\alpha \neq 0$. Then α^{-1} exists, so we can multiply both sides and rearrange to arrive at

$$v = -\alpha^{-1}\alpha_1 v_1 - \ldots - \alpha^{-1}\alpha_n v_n$$

The above is simply a linear combination of elements of S, so it is in the span of S. However, we initially assumed $v \notin span(S)$, and so we have arrived at a contradiction. Since all scenarios lead to a contradiction, we can conclude what we had assumed is false, proving that $S \cup \{v\}$ is linearly independent.

Def: a *basis* for a vector space V is a set of linearly independent vectors in V that spans V.

Def: for a given basis $\beta = \{v_1, \ldots v_n\}$ of a vector space V, the *coordinates* refer to the scalars associated with the basis vectors. For example, an arbitrary $v \in V$ has representation

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n \quad \forall \, \alpha_i \in F$$

as β spans V. α_i is referred to as the "i'th" coordinate of β, associated with the basis vector v_i.

Coordinate Theorem: *for any given basis $\beta = \{v_1, \ldots v_n\}$ the coordinates associated with a given element are unique*

Proof: Suppose we have two equivalent vectors with representations

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n \text{ and } v = \beta_1 v_1 + \ldots + \beta_n v_n \Rightarrow (\alpha_1 - \beta_1)v_1 + \ldots + (\alpha_n - \beta_n)v_n = 0$$

However, β is linearly independent by definition of a basis, implying $\alpha_i - \beta_i = 0 \, \forall i$ $\Rightarrow \alpha_i = \beta_i \, \forall i$ and so the coordinates are unique.

Toss Out Theorem: *Suppose $v \in span(v_1, \ldots v_n)$. Then $span(v, v_1, \ldots v_n) = span(v_1 \ldots v_n)$.* In plain text, we can "toss out" a linearly dependent vector from a set to maintain the same span.

Proof: Let $W = span(v, v_1, \ldots v_n)$ and $W' = span(v_1, \ldots v_n)$. Clearly, $W' \subseteq W$. Now suppose $w \in W$. Then $w = \alpha v + \alpha_1 v_1 + \ldots \alpha_n v_n$ for some $\alpha, \alpha_1 \ldots \alpha_n \in F$. But

since $v \in W'$, $v = \beta_1 v_1 + \ldots + \beta_n v_n$ for some $\beta_1, \ldots \beta_n \in F$

$$\Rightarrow w = \alpha(\beta_1 v_1 + \ldots + \beta_n v_n) + \alpha_1 v_1 + \ldots + \alpha_n v_n$$

$$\Rightarrow w = (\alpha\beta_1 + \alpha_1)v_1 + \ldots + (\alpha\beta_n + \alpha_n)v_n$$

Since w can be represented purely as a linear combination of elements in W',

$$\Rightarrow w \in W' \Rightarrow W \subseteq W' \Rightarrow W = W'$$

Replacement Theorem: *Let V be a vector space over a field F with a given basis $\beta = \{v_1, \ldots v_n\}$. Suppose that $v \in V$ is a linear combination of the basis vectors, with some coordinate $\alpha_i \neq 0$. Then $\beta' = \{v_1, \ldots v_{i-1}, v, v_{i+1}, \ldots v_n\}$ is also a basis for V.* In plain text, if we find some vector with a nonzero component along the i'th basis vector, we can form a new basis by replacing that basis vector with our new one.

Proof: Without loss of generality, assume $i = 1$. The α_i^{-1} exists, and so for an arbitrary $v \in V$, $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$, we can rewrite this to conclude

$$v_1 = \alpha_1^{-1} v - \alpha_1^{-1}\alpha_2 v_2 - \ldots - \alpha_1^{-1}\alpha_n v_n \Rightarrow v_1 \in span(v, v_2, \ldots v_n)$$

Since v is simply a linear combination of $v_1, \ldots v_n$, it is evident that $span(v_1, \ldots v_n) = span(v, v_1, \ldots v_n)$. However, since $v_1 \in span(v, v_2, \ldots v_n)$, by the Toss Out Theorem, $span(v, v_1, \ldots v_n) = span(v, v_2, \ldots, v_n)$. Hence, $span(\beta) = span(\beta')$, and so β' spans V. Thus, it suffices to show β' is a basis if it is linearly independent. By way of contradiction, assume β' is linearly dependent.
$\Rightarrow \exists \, \beta, \beta_2, \ldots \beta_n$ not all zero such that $\beta v + \beta_2 v_2 + \ldots + \beta_n v_n = 0$.

Case 1: $\beta = 0$
Then this implies $\beta_2 v_2 + \ldots + \beta_n v_n = 0$ for $\beta_2 \ldots \beta_n$ not all equal to zero. However, this contradicts their linear independence, as $\beta = \{v_1 \ldots v_n\}$ is linearly independent.

Case 2: $\beta \neq 0$
Then β^{-1} exists, so we can multiply both sides by this, solve for v, and set it equal to our original equation for v.

$$v = 0 \cdot v_1 - \beta^{-1}\beta_2 v_2 - \ldots - \beta^{-1}\beta_n v_n = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

By the coordinate theorem, $\alpha_1 = 0$. However, this is exactly the α_i we had assumed to be nonzero, so we have arrived at a contradiction. In either case, the assumption that β' is linearly dependent leads to a contradiction, and so β' is linearly independent.

Main Theorem: *Suppose V is a vector space over a field F with $V = span(v_1, \ldots v_n)$. Then any linearly independent subset in V has at most n elements.*

Proof: by the Toss Out Theorem, we can toss out linearly dependent vectors from this set without affecting the span, until we have arrived at a linearly independent set. Since this set spans V and is linearly independent, it is a basis for V. Assume $\beta = \{v_1, \ldots, v_n\}$ is a basis for V. Now suppose $\{w_1, \ldots w_m\} \subseteq V$ is a linearly independent set. We must show $m \leq n$. We claim it suffices to show that for each $k \leq n$, $\{w_1, \ldots w_k, v_{k+1}, \ldots v_n\}$ is a basis. If $k < n$, then $m < n$, and so the theorem is proved. If $k = n$, then $\{w_1, \ldots w_n\}$ is a basis, as any w_{k+c} for $c > 0$ implies w_{k+c} can be written as a linear combination of $\{w_1, \ldots, w_n\}$, i.e. the w's are not linearly independent, contradicting the claim that the w's form a basis.

Base Case: $k = 1$.
Then $w_1 = \alpha_1 v_1 + \ldots + \alpha_n v_n$ for some $\alpha_i \in F$. Since w_1 is a basis element, some α_i is nonzero, so without loss of generality, assume $\alpha_1 \neq 0$. Rearranging, we can express v_1 as a linear combination of $\{w_1, v_2, \ldots, v_n\}$, and so by the Replacement Theorem, $\{w_1, v_2, \ldots, v_n\}$ is a basis for V.

Inductive Step: *does $P(n-1) \Rightarrow P(n)$?*
Assume $P(n-1)$, i.e. for some $k < n$, $\{w_1, \ldots w_k, v_{k+1}, \ldots v_n\}$ is a basis. We want to show that $\{w_1, \ldots w_{k+1}, v_{k+2}, \ldots v_n\}$ is a basis. w_{k+1} has representation $w_{k+1} = \alpha_1 w_1 + \ldots + \alpha_k w_k + \beta_{k+1} v_{k+1} + \ldots \beta_n v_n$. Note $\exists \beta_i \neq 0$, as if all β_i were 0, then $w_{k+1} \in span(w_1, \ldots w_m)$, contradicting the linear independence of the w's. Without loss of generality, assume $\beta_{k+1} \neq 0$. Hence, $\{w_1, \ldots w_{k+1}, v_{k+2}, \ldots v_n\}$ is a basis by the Replacement Theorem.

Def: the *dimension* of a vector space V is the cardinality of a basis for V. This may be finite or infinite dimensional.

Many people, upon first coming to grips with the notion of an infinite dimensional

vector space, are rightly confused. *Aren't most vector spaces infinite sets?* This is true, but the dimension of a vector space represents something different than the mere size of the vector space itself. It represents *how many basis vectors we need* in order to describe an arbitrary object in V through linear combination.

For example, consider $\mathbb{R}^2$, the set of all (x, y) such that $x \in \mathbb{R}$ and $y \in \mathbb{R}$, or equivalently, $\mathbb{R} \times \mathbb{R}$. This is a finite dimensional vector space, as there exists a basis with only two elements, $(1, 0)$ and $(0, 1)$ to name an example.

In contrast, consider the set of all polynomials with coefficients over a field, denoted as $F[t]$, with an arbitrary element having representation $\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \ldots + \alpha_n t^n$ for $\alpha_0, \ldots \alpha_n \in F$. This is an infinite dimensional vector space, since the monomials generated by t^n form basis for $F[t]$, but since $n \in \mathbb{N}$, n is not bounded by any finite number (the span of t^n is evident, and their linear independence is a result from differential equations). It is important to make the distinction between infinite and finite dimensional vector spaces, as some theorems only hold in the finite dimensional case.

Corollary: *Let V be a finite-dimensional vector space over a field F of dimension $n > 0$, and $\varnothing \neq S \subseteq V$ a subset. Then*

- If $|S| > n$, then S is linearly dependent

- If $|S| < n$, then S does not span V

The theorem states that a maximally linearly independent set in V is a basis. By the Toss Out Theorem, a minimal spanning set of V is a basis.

Corollary: *Let V be a finite-dimensional vector space over a field F with bases β_1 and β_2. Then $|\beta_1| = |\beta_2|$.*

By definition, a basis β must span V and be linearly independent. Then, by the previous corollary, $|\beta| \geq dim(V)$, and $|\beta| \leq dim(V)$, or $|\beta| = dim(V)$. Thus, for a given V, all bases have the same cardinality, and so $|\beta_1| = |\beta_2|$

This is what we expected, after all it would be disastrous if the dimension of a vector space depended on what basis we chose. Since we've also concluded that a basis is simply a minimal spanning set, we can conclude that every vector space has a basis. Evidently, there is an intimate relationship between bases and vector spaces. In

fact, when we talk about linear transformations, much of our work will be describing what the linear transformations does to the basis elements, to give an example of how useful describing things in terms of their bases are.

Extension Theorem: *Let V be a finite-dimensional vector space over a field F, and $W \subseteq V$ a subspace of V. Then any linearly independent subset of W is finite, and part of a basis for W.* It also follows that any linearly independent subset of V may be extended to a basis for V.

Proof: By the main theorem, any linearly independent subset $\beta \subseteq W$ has at most $dim(W)$ linearly independent vectors. In particular, $dim(span(\beta)) \leq dim(W)$. If $dim(span(\beta)) = dim(W)$, we are done, as any linearly independent subset of a vector space with cardinality equal to the dimension of the given vector space forms a basis for that vector space. If $dim(span(\beta)) < dim(W)$, by the Toss In Theorem, we can add in some vector $v \in V$ where $v \notin span(\beta)$ such that $\beta \cup \{v\}$ is linearly independent, and our new set β' has dimension $|\beta| + 1$. Since V is finite dimensional, this process will eventually terminate when $dim(\beta') = dim(W)$, in which case β' will be a basis for W by the Main Theorem.

Counting Theorem: *Let V be a vector space over F, with $W_1, W_2 \subseteq V$ finite-dimensional subsets of V. Then*

- $W_1 \cap W_2$ is a finite-dimensional vector space
- $W_1 + W_2$ is a finite-dimensional vector space
- $dim(W_1) + dim(W_2) = dim(W_1 + W_2) + dim(W_1 \cap W_2)$

Proof: $W_1 \cap W_2 \subseteq W$, so it is finite-dimensional since W is finite-dimensional. Moreover, it is a vector space, as $\forall\, x, y \in W_1 \cap W_2$ and $\forall\, c \in F$

$$\Rightarrow x \in W_1, W_2 \text{ and } y \in W_1, W_2 \Rightarrow cx + y \in W_1, W_2 \text{ since } W_1 \text{ and } W_2 \text{ are subspaces}$$

If β_1, β_2 are bases for W_1, W_2 respectively, then by definition, $W_1 + W_2 = span(\beta_1 \cup \beta_2)$. Since β_1, β_2 are finite, then $W_1 + W_2$ is finite dimensional. Moreover, $\beta_1 \cup \beta_2$ can be reduces to a basis by the Toss Out Theorem, and so $W_1 + W_2$ is also a vector space.

Now suppose β is a basis for $W_1 \cup W_2$. By the Extension Theorem, we can extend this to a basis for W_1 and W_2, denoted as β_1, β_2 respectively. We want to claim that $dim(W_1 + W_2) = dim(span(\beta_1 \cup \beta_2))$. To do so, we simply need to show $\beta_1 \cup \beta_2$ is

linearly independent. By way of contradiction, assume $\beta_1 \cup \beta_2$ is linearly dependent. Note that given a general representation of a vector in V for our basis elements, it equaling zero implies not all coefficients associated with β and β_2 are zero, as this would violate the linear independence of β_1. Similarly, not all coefficients regarding β and β_2 are equal to zero. Then, if our basis vectors are

$$\beta = \{v_1, \ldots v_n\}, \beta_1 = \{y_1, \ldots y_m\}, \beta_2 = \{z_1, \ldots z_r\} \; with \; scalars \; \alpha_1, \ldots \alpha_n, \beta_1, \ldots \beta_m, \gamma_1, \ldots \gamma_r \in F$$

$$\Rightarrow v = \alpha_1 v_1 + \ldots + \alpha_n v_n + \beta_1 y_1 + \ldots + \beta_m y_m + \gamma_1 z_1 + \ldots + \gamma_r z_r = 0$$

Without loss of generality, assume $\gamma_1 \neq 0$. Note our representations of z are

$$0 \neq z = \gamma_1 z_1 + \ldots + \gamma_r z_r \in span(\beta_2) = W_2$$

$$0 \neq z = -\alpha_1 v_1 - \ldots - \alpha_n v_n - \beta_1 y_1 - \ldots - \beta_m y_m \in span(\beta_2) = W_2$$

$$\Rightarrow \gamma_1 z_1 + \ldots + \gamma_r z_r = -\alpha_1 v_1 - \ldots - \alpha_n v_n - \beta_1 y_1 - \ldots - \beta_m y_m$$

$$\Rightarrow \exists \; \delta_1, \ldots \delta_n \; such \; that \; 0 \neq z = \delta_1 v_1 + \ldots + \delta_n v_n$$

$$\Rightarrow \delta_1 v_1 + \ldots + \delta_n v_n + 0 \cdot z_1 + \ldots + 0 \cdot z_r = \gamma_1 z_1 + \ldots \gamma_r z_r$$

But, by the coordinate theorem we get $\gamma_1 = 0$, when we had assumed $\gamma_1 \neq 0$! Thus, we have arrived at a contradiction, and so $\beta_1 \cup \beta_2$ is linearly independent. Thus, $dim(W_1 + W_2) = dim(span(\beta_1 \cup \beta_2))$

$$\Rightarrow dim(W_1 + W_2) = |\beta_1| + |\beta_2| - |\beta_1 \cap \beta_2|$$

$$\Rightarrow dim(W_1 + W_2) = dim(W_1) + dim(W_2) - dim(W_1 \cap W_2)$$

At this point, we have fleshed out the concepts we need to describe vector spaces, and proved important theorems that we'll build upon in the next section. We defined fields, then built upon them with vector spaces. We related the notion of linear independence, span, basis, and dimension to each other, and rigorously proved ideas we may have taken for granted in our first pass at linear algebra.

Linear Transformations

Def: a *function* $f : X \to Y$ is a map that assigns a unique $y \in Y$ for each $x \in X$. The set X is called the *domain* of f, while the set Y is called the *codomain* of f. The *range* or *image* of f denotes the set of all $f(x) \in Y \ \forall x \in X$, which is not strictly the same as the codomain.

Def: a *linear transformation* is a function $T : V \to W$ such that $\forall v_1, v_2 \in V$ and $c \in F$, then $T(cv_1 + v_2) = cT(v_1) + T(v_2)$.

Essentially, every function introduced in linear algebra will be a linear transformation. Linear functions are everywhere in mathematics, and so the study of them will be enriching in many areas. For example, derivatives and integrals are linear functions, as $\frac{\partial}{\partial t}(cf + g) = c\frac{\partial f}{\partial t} + \frac{\partial g}{\partial t}$, and $\int(cf + g) = c\int f + \int g$, for some functions f and g, and some constant c. Note our definition of a linear transformation implies $T(0_V) = 0_W$, as for some $v \in V$, $T(0_V) = T(v - v) = T(v) - T(v) = 0_W$.

Def: a *linear operator* is a linear transformation $T : V \to V$ such that its domain is its codomain.

It's useful to distinguish this case, as we're not always interested in mapping elements from different sets. For example, a rotation transformation in $\mathbb{R}^2$ takes vectors from $\mathbb{R}^2$ to $\mathbb{R}^2$.

Def: the *kernel* of a linear transformation is the set of all $v \in V$ such that $T(v) = 0$.

We claim both the image and kernel of a linear transformation $T : V \to W$ are subspaces. $im(T) := \{T(v) \in W \mid v \in V\} \Rightarrow im(T) \subseteq W$. Now suppose we have $w_1, w_2 \in im(T)$ and $c \in F$. Then $\exists v_1, v_2 \in V$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$. By linearity, $T(cv_1 + v_2) = cT(v_1) + T(v_2) = cw_1 + w_2$. Thus, $cw_1 + w_2 \in im(T)$, and so $im(T)$ is a subspace of W. $ker(T) := \{v \in V \mid T(v) = 0\} \Rightarrow ker(T) \subseteq V$. Now suppose we have $v_1, v_2 \in ker(T)$, and $c \in F$. By linearity, $T(cv_1 + v_2) = cT(v_1) + T(v_2)$. But, since $v_1, v_2 \in ker(T)$, $cT(v_1) + T(v_2) = c \cdot 0 + 0 = 0 \Rightarrow cv_1 + v_2 \in ker(T)$. Thus, $ker(T)$ is a subspace of V.

Def: an *injection* is a function $f : X \to Y$ such that $f(x_1) = f(x_2) \Rightarrow x_1 = x_2 \; \forall \; x_1, x_2 \in X$. It is also called a *one to one function*, as it implies no two x's in X map to the same y in Y.

An example of an injective function is $f : \mathbb{R} \to \mathbb{R}$ by $x \mapsto e^x$, as since $\frac{\partial}{\partial x} e^x > 0 \; \forall x \in \mathbb{R}$, e^x is always increasing, and so two distinct x's in $\mathbb{R}$ will never map to the same $f(x) \in \mathbb{R}$. An example of a non-injective function is $f : \mathbb{R} \to \mathbb{R}$ by $x \mapsto x^2$, as both $x = -1$ and $x = 1$ map to $f(x) = 1$.

Def: a *surjection* is a function $f : X \to Y$ such that $\forall y \in Y, \; \exists x \in X$ such that $f(x) = y$. It is also called *onto*, as it implies every value in the codomain is mapped to. Equivalently, the range or image is the same set as the codomain.

An example of a surjective function is $f : \mathbb{R} \to \mathbb{R}$ by $x \mapsto 2x$, as for any $y \in Y$, we can find an $x \in X$ such that $f(x) = y$, namely $x = \frac{y}{2}$ in this case, which is in $\mathbb{R}$, as $\mathbb{R}$ is closed under division. An example of a non-surjective function is $f : \mathbb{Z} \to \mathbb{Z}$ by $x \mapsto 2x$, as, for example, $y = 3$ is never mapped to, as it would require $x = \frac{3}{2}$, but $\frac{3}{2} \notin \mathbb{Z}$.

Def: a *bijection* is a function $f : X \to Y$ such that f is an injection and a surjection. If f is a *bijection*, then each $x \in X$ is associated with a unique $y \in Y$, and each $y \in Y$ is associated with a unique $x \in X$. Thus, $|X| = |Y|$. Additionally, we can define a new map $f^{-1} : Y \to X$ by $y \mapsto x$ if $f(x) = y$.

Dimension Theorem: *Let V and W be finite dimensional vector spaces over a field F. Let $T : V \to W$ be a linear transformation. Then $dim(V) = dim(im(T)) + dim(ker(T))$.*

Proof: Let $dim(V) = n$. Suppose $\beta_0 = \{v_1, \ldots v_m\}$ is a basis for $ker(T)$. Since $ker(T)$ is a subspace of V, by the Extension Theorem, we can extend it to a basis $\beta = \{v_1, \ldots v_n\}$ for V (for $m \leq n$). If we can show $\beta' = \{Tv_{m+1}, \ldots Tv_n\}$ is a basis for $im(T)$, the theorem will have been proved, as it implies $|\beta| + |\beta'| = n \Rightarrow dim(im(T) + dim(ker(T)) = dim(V)$.

To show β' is a basis for $im(T)$, we must show span and linear independence. Suppose $w \in im(T)$ is an arbitrary element of the image. Then $\exists v \in V$ such that $T(v) = w$.

Since β is a basis for V, we have $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ for some $\alpha_1, \ldots \alpha_n \in F$. Then $w = T(\alpha_1 + v_1 + \ldots + \alpha_n) = \alpha_1 T(v_1) + \ldots + \alpha_n T(v_n)$. However, since the v_i's up to v_m are in $ker(T)$ (since they are in β_0), then $w = \alpha_{m+1} T(v_{m+1}) + \ldots + \alpha_n T(v_n)$, and thus β' spans $im(T)$. Now consider $\alpha_{m+1} T(v_{m+1}) + \ldots + \alpha_n T(v_n) = 0$.

$$\Rightarrow T(\alpha_{m+1} v_{m+1} + \ldots + \alpha_n v_n) = 0$$

$$\Rightarrow \alpha_{m+1} v_{m+1} + \ldots + \alpha_n v_n \in ker(T)$$

But, since β_0 is a basis for $ker(T)$, then this vector also has the representation

$$\beta_1 v_1 + \ldots + \beta_m v_m \ for \ some \ \beta_1, \ldots \beta_m \in F$$

$$\Rightarrow \alpha_{m+1} v_{m+1} + \ldots + \alpha_n v_n = \beta_1 v_1 + \ldots + \beta_m v_m$$

$$\Rightarrow -\beta_1 v_1 - \ldots - \beta_m v_m + \alpha_{m+1} v_{m+1} + \ldots + \alpha_n v_n = 0$$

Remember $\beta = \{v_1, \ldots v_n\}$ is a basis. Thus, $\{v_1, \ldots v_n\}$ are linearly independent, and so the above equation implies $\alpha_i, \beta_j = 0 \ \forall i, j$. Thus, $\alpha_{m+1} T(v_{m+1}) + \ldots + \alpha_n T(v_n) = 0$ is only true for $\alpha_{m+1} = \ldots = \alpha_n = 0$, and so $\{T(v_{m+1}), \ldots T(v_n)\}$ is linearly independent.

Def: a *monomorphism* is an injective linear transformation.

Def: an *epimorphism* is a surjective linear transformation.

Def: an *isomorphism* is a bijective linear transformation.

We say two vector spaces V and W are *isomorphic* if there exists an isomorphism between the two, denoted as $V \cong W$. In general, an isomorphism is an extension of the notion of equality. For example, the x and y axes of the $\mathbb{R}^2$ plane are not strictly equal, as they contain different elements, but they share a certain mathematical structure, in that we can construct a mapping between the two sets, matching vectors to vectors such that operations in one set will preserve relationships between elements in the other. An isomorphism is an equivalence relation.

As our isomorphism $T : V \rightarrow W$ is a bijection, there exists an inverse transformation $T^{-1} : W \rightarrow V$. I claim this transformation is also linear. Let $w_1, w_2 \in W$ and $c \in F$.

Suppose $T^{-1}(w_1) = v_1$ and $T^{-1}(w_2) = v_2$, for $v_1, v_2 \in V$. Then $T(v_1) = w_1$ and $T(v_2) = w_2$.

$$T^{-1}(cw_1 + w_2) = T^{-1}(cT(v_1) + T(v_2)) = T^{-1}(T(cv_1 + v_2))$$

$$T^{-1}(T(cv_1 + v_2)) = cv_1 + v_2 = cT^{-1}(v_1) + T^{-1}(w_2)$$

Since $T^{-1}(cw_1 + w_2) = cT^{-1}(w_1) + T^{-1}(w_2)$, T^{-1} is linear.

Monomorphism Theorem: *Let $T : V \to W$ be a linear transformation, for V and W vector spaces over a field F. Then the following are equivalent.*

- T is a monomorphism
- T takes linearly independent sets to linearly independent sets
- $ker(T) = \{0\}$
- $dim(ker(T)) = 0$

Proof: Let T be a monomorphism. Let $\{v_1, \ldots v_n\}$ be an arbitrary, linearly independent set in V. Suppose, by way of contradiction, that $\{T(v_1), \ldots T(v_n)\}$ is linearly dependent. Then $\exists \alpha_i \neq 0$ such that $\alpha_1 T(v_1) + \ldots + \alpha_n T(v_n) = 0$ for $\alpha_1, \ldots \alpha_n \in F$. By linearity, $\alpha_1 T(v_1) + \ldots + \alpha_n + T(v_n) = 0 \Rightarrow T(\alpha_1 v_1 + \ldots + \alpha_n v_n) = 0$. Since T is a monomorphism, $\alpha_1 v_1 + \ldots + \alpha_n v_n$ is the unique element that maps to 0. Moreover, as T is a linear transformation, $T(0) = 0 \Rightarrow \alpha_1 v_1 + \ldots + \alpha_n v_n = 0$. However, since $\{v_1, \ldots v_n\}$ is linearly independent, it implies $\alpha_i = 0 \; \forall i$, which violates our assumption. Thus, $\{T(v_1), \ldots T(v_n)\}$ is linearly independent, and so 1) $\Rightarrow$ 2) is proved.

Now assume 2). Suppose, by way of contradiction that there exists some nonzero vector $v \in ker(T)$. Since $\{v\}$ is linearly independent, $\{T(v)\}$ must be too. However, $T(v) = 0$, which is linearly dependent. With this contradiction, we have shown 2) $\Rightarrow$ 3).

3) $\Leftrightarrow$ 4) is immediate.

Suppose $ker(T) = \{0\}$. Let $v_1, v_2 \in V$, such that $T(v_1) = T(v_2)$. Rearranging, we arrive at $T(v_1 - v_2) = 0 \Rightarrow v_1 - v_2 = 0 \Rightarrow v_1 = v_2$. Thus, T is an injection, ergo T is a monomorphism. Hence, 3) $\Rightarrow$ 1) is proved.

Isomorphism Theorem: *Let $T : V \to W$ be a linear transformation, for V and W finite dimensional vector spaces over a field F with $n = dim(V) = dim(W)$. Then the following are equivalent*

- T is an isomorphism

- T is a monomorphism

- T is an epimorphism

- If $\beta = \{v_1, \ldots v_n\}$ is a basis for V, then $\beta' = \{T(v_1), \ldots T(v_n)\}$ is a basis for W

- $\exists$ a basis for V such that T takes it to a basis for W

1) $\Rightarrow$ 2) By definition, an isomorphism is a monomorphism.

2) $\Rightarrow$ 3) By the Dimension Theorem, $dim(V) = dim(im(T)) + dim(ker(T))$. Since T is a monomorphism, by the Monomorphism Theorem, $dim(ker(T)) = 0$, and so $dim(im(T)) = dim(V) = dim(W)$. Thus, T is an epimorphism.

3) $\Rightarrow$ 2) As T is an epimorphism, $dim(V) = dim(W)$. By the Dimension Theorem, $dim(im(T)) + dim(ker(T)) = dim(V) \Rightarrow dim(ker(T)) = 0$. By the Monomorphism Theorem, T is a monomorphism.

3) $\Rightarrow$ 1) If T is an epimorphism, we have shown this implies it is a monomorphism. Thus, it is an isomorphism.

2) $\Rightarrow$ 4) By the Monomorphism Theorem, β' is linearly independent. By the Main Theorem, any n sized set of linearly independent vectors in an n dimensional vector space forms a basis for that vector space, thus β' is a basis for W.

4) $\Rightarrow$ 5) By the Main Theorem, every vector space has a basis, so 5) immediately follows from 4).

5) $\Rightarrow$ 3) Let $\beta = \{v_1, \ldots v_n\}$ be a basis for V such that $\beta' = \{T(v_1), \ldots T(v_n)\}$ is a basis for W. Thus, β' spans W, and therefore T is an epimorphism.

It is a strong condition that V and W must be finite dimensional. The theorem does not hold unless this is true.

Classification Theorem: *Let V and W be finite dimensional vector spaces over a field F. Then $V \cong W \Leftrightarrow dim(V) = dim(W)$.*