Math 115A: Linear Algebra

Michael Andrews
UCLA Mathematics Department

March 20, 2018

Contents

1 January 8: a little about sets

Definition 1.1. A *set* is any collection of mathematical things. The members of a set are often called its *elements*. We write $x \in X$ to mean that "x is an element of a set X."

Notation 1.2. Curly brackets (braces) are often used to show sets. The set whose elements are $a_1, a_2, a_3, \ldots, a_n$ is written $\{a_1, a_2, a_3, \ldots, a_n\}$. Similarly, the set whose elements are those of an infinite sequence $a_1, a_2, a_3, \ldots$ of objects is denoted $\{a_1, a_2, a_3, \ldots\}$.

Definition 1.3. A common way to define a set is as "those elements with a certain property." This can be problematic (Russell's paradox), but usually it is okay. We write

$$\{x : x \text{ has property } P\}$$

for the set consisting of elements x with a property P, or

$$\{x \in X : x \text{ has property } P\}$$

for the elements of a previously defined set X with a property P.

The colons should be read as "such that."

Notation 1.4. Here is some notation for familiar sets. ":=" means "is defined to be equal to."

1. the *natural numbers* $\mathbb{N} := \{1, 2, 3, \ldots\}$,
2. the *integers* $\mathbb{Z} := \{0, -1, 1, -2, 2, -3, 3, \ldots\}$,
3. the *rationals* $\mathbb{Q} := \{\frac{m}{n} : m \in \mathbb{Z},\ n \in \mathbb{N}\}$,
4. the *reals* $\mathbb{R}$,
5. the *complex numbers* $\mathbb{C} := \{x + iy : x, y \in \mathbb{R}\}$.

Definition 1.5. Given two sets X and Y, we say that X is a *subset* of Y and write $X \subseteq Y$ **iff** whenever $z \in X$, we have $z \in Y$.

Remark 1.6. In this definition "iff" stands for "if and only if." This means that saying "X is a subset of Y" is exactly the same as saying "whenever $z \in X$, we have $z \in Y$."

Often, in definitions, mathematicians say "if" even though the meaning is "iff." I normally do this, and I feel a little silly for writing "iff," but I decided that it's the least confusing thing I can do. To make this "iff" feel different from a non-definitional "iff" I have used bold.

Definition 1.7. The *empty set*, written $\emptyset$, is the set with no elements.

Definition 1.8. Suppose X and Y are sets. We say that $X = Y$ **iff** ($X \subseteq Y$ and $Y \subseteq X$), i.e. **iff** $z \in X$ if and only if $z \in Y$.

Example 1.9.

1. $\{0, 1\} = \{1, 0\}$.
2. $\{2, 2\} = \{2\}$.
3. $\{\emptyset\} \neq \emptyset$. This is because $\emptyset \in \{\emptyset\}$, but $\emptyset$ is not an element of $\emptyset$, so $\{\emptyset\}$ is not a subset of $\emptyset$.
4. $\{n \in \mathbb{N} :\ n \text{ divides } 12\} = \{1, 2, 3, 4, 6, 12\}$.

2 January 9 (discussion)

2.1 Some definitions: union, intersection, set complements

Definition 2.1.1. Suppose X and Y are sets.

We write $X \cup Y$ for the set

$$\{z : z \in X \text{ or } z \in Y\}.$$

$X \cup Y$ is read as "X *union* Y" or "the *union* of X and Y."

We write $X \cap Y$ for the set

$$\{z : z \in X \text{ and } z \in Y\}.$$

$X \cup Y$ is read as "X *intersect* Y" or "the *intersection* of X and Y."

We write $X \setminus Y$ for the set

$$\{z : z \in X \text{ and } z \notin Y\}.$$

$X \setminus Y$ is read as "X *takeaway* Y." $x \notin Y$ means, and is read as "x is not an element of Y."

Example 2.1.2. The irrationals are the set $\mathbb{R} \setminus \mathbb{Q}$.

Example 2.1.3. For any sets X and Y we have the following *inclusions*.

1. $X \subseteq X \cup Y$, $Y \subseteq X \cup Y$.
2. $X \cap Y \subseteq X$, $X \cap Y \subseteq Y$.
3. $X \cap Y \subseteq X \cup Y$.

2.2 Using definitions: $\emptyset$ is always a subset, De Morgan's Laws

Here's a slightly silly example.

Theorem 2.2.1. *Suppose X is a set. Then $\emptyset \subseteq X$.*

Proof. Let X be a set. We must check that whenever $z \in \emptyset$, we have $z \in X$. However, since $\emptyset$ has no elements, there is nothing to check. $\square$

The next proof is more useful for your learning. Here are some things to try and learn from the following proof.

- Introducing the sets you are going to be using in your proof.
- Verifying set equality: you have to check $\subseteq$ and $\supseteq$.
- Stating the if-then sentence you are going to verify.
- How to prove an if-then statement directly: you suppose the premise is true, and must verify the conclusion to be true.
- Using definitions of union, intersection, etc.

- Although you might think the following result is "obvious," there is something to say in order to make the result feel clear to the reader. You should avoid the word "obvious" at all costs. I manage to break the proof into stages where in each line I am either stating what we need to check, stating an assumption, using a definition, making a deduction, using the line before, or summarizing what has happened before.

Theorem 2.2.2 (De Morgan's Laws)**.** *Suppose X, A, and B are sets. Then*

$$X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B),$$

$$X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B).$$

Proof. Let X, A, and B be sets.

1. Homework.

2. By definition of set equality, we have to show that

$$z \in X \setminus (A \cap B) \text{ if and only if } z \in (X \setminus A) \cup (X \setminus B).$$

   (a) First, we show that if $z \in X \setminus (A \cap B)$, then $z \in (X \setminus A) \cup (X \setminus B)$.
   Suppose $z \in X \setminus (A \cap B)$. By definition of $\setminus$, this means $z \in X$ and $z \notin A \cap B$.
   We cannot have both $x \in A$ and $x \in B$ since, by definition of $\cap$, this would tell us that $z \in A \cap B$.

      i. Case 1: $z \notin A$. Since $z \in X$, the definition of $\setminus$ tells us that $z \in X \setminus A$.
      Thus, by definition of $\cup$, $z \in (X \setminus A) \cup (X \setminus B)$.

      ii. Case 1: $z \notin B$. Since $z \in X$, the definition of $\setminus$ tells us that $z \in X \setminus B$.
      Thus, by definition of $\cup$, $z \in (X \setminus A) \cup (X \setminus B)$.

   In either case we have shown that $z \in (X \setminus A) \cup (X \setminus B)$.

   (b) Next, we show that if $z \in (X \setminus A) \cup (X \setminus B)$, then $z \in X \setminus (A \cap B)$.
   Suppose $z \in (X \setminus A) \cup (X \setminus B)$. By definition of $\cup$, this means $z \in X \setminus A$ or $z \in X \setminus B$.

      i. Case 1: $z \in X \setminus A$. By definition of $\setminus$, this means that $z \in X$ and $z \notin A$.
      We cannot have $z \in A \cap B$, since otherwise, by the definition of $\cap$, we'd have $z \in A$.
      So $z \in X$ and $z \notin A \cap B$, and the definition of $\setminus$ gives $z \in X \setminus (A \cap B)$.

      ii. Case 1: $z \in X \setminus B$. By definition of $\setminus$, this means that $z \in X$ and $z \notin B$.
      We cannot have $z \in A \cap B$, since otherwise, by the definition of $\cap$, we'd have $z \in B$.
      So $z \in X$ and $z \notin A \cap B$, and the definition of $\setminus$ gives $z \in X \setminus (A \cap B)$.

   In either case we have shown that $z \in X \setminus (A \cap B)$.

$\square$

Remark 2.2.3. In the proof of the second result above, both directions broke into cases. In both instances, the cases were almost identical with A and B swapping roles. Because of the symmetry of the situation it might be reasonable to expect the reader to see that both cases are proved similarly. Often in such cases the proof-writer might say, "case 2 is similar" or they might say "without loss of generality we only need to consider the first case." I'd hesistate to do such things unless you have written out the proof and seen that it really is identical. Sometimes things can appear symmetric, but after careful consideration, you might realize it's not that easy!

2.3 Cartesian product and the definition of a function

Definition 2.3.1. Suppose X and Y are sets. We write $X \times Y$ for the set

$$\{(x, y) : \ x \in X, \ y \in Y\},$$

that is the set of ordered pairs where one coordinate has its value in X and the other has its value in Y. $X \times Y$ is called the *Cartesian product* of X and Y.

Example 2.3.2.

1. $\{0, 1\} \times \{5, 6, 7\} = \{(0, 5), \ (0, 6), \ (0, 7), \ (1, 5), \ (1, 6), \ (1, 7)\}$.
2. $\mathbb{R} \times \mathbb{R}$ is the Cartesian plane.
3. $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$ is the home of 3D calculus.

Definition 2.3.3. A *function* $f : X \longrightarrow Y$ consists of:

- a set X called the *domain* of f;
- a set Y called the *codomain* of f;
- a subset G_f of the Cartesian product $X \times Y$ with certain properties.

When $(x, y) \in G_f$, we write $f(x) = y$.

Remark 2.3.4. This is the formal definition of a function $f : X \longrightarrow Y$. Normally we don't specify a subset of $X \times Y$ explicitly and instead just give a formula for $f(x)$.

The reason for the notation G_f is that it stands for "graph of f."

The omitted properties regarding $G_f \subseteq X \times Y$ say that for every x-value there is a corresponding y value, and that for each x-value there can only be one y-value.

Notation 2.3.5. We often use the notation $f : X \longrightarrow Y$, $x \longmapsto f(x)$.

3 Some more on sets and functions

I am not sure when or even if we will cover this material explicitly in class yet. We may not have time. It is certainly useful for you to read and understand.

First, we complete the definition of a function in case you're interested.
Recall that a function $f : X \longrightarrow Y$ consists of:

- a set X called the *domain* of f;
- a set Y called the *codomain* of f;
- a subset G_f of the Cartesian product $X \times Y$.

For $f : X \longrightarrow Y$ to be a function the subset $G_f \subseteq X \times Y$ must have the following properties:

- Suppose $x \in X$. Then there exists a $y \in Y$ such that $(x, y) \in G_f$.
- Suppose $x \in X$ and $y_1, y_2 \in Y$. If $(x, y_1) \in G_f$ and $(x, y_2) \in G_f$, then $y_1 = y_2$.

When $(x, y) \in G_f$, we write $f(x) = y$.

Definition 3.1. Suppose X and Y are sets and that $f : X \longrightarrow Y$ is a function.

1. We say f is *injective* **iff** whenever $x_1, x_2 \in X$, $f(x_1) = f(x_2)$ implies $x_1 = x_2$.
2. We say f is *surjective* **iff** whenever $y \in Y$, we can find an $x \in X$ such that $f(x) = y$.
3. We say f is *bijective* **iff** f is injective and surjective.

Remark 3.2. We may also use "one-to-one" for injective, and "onto" for surjective.

There are noun forms of the words in the previous definition too. We speak of an injection, a surjection, and a bijection.

Definition 3.3. Suppose X, Y, and Z are sets, and that $f : X \to Y$ and $g : Y \to Z$ are functions. The *composition of f and g* is the function $g \circ f : X \longrightarrow Z$ defined by $(g \circ f)(x) = g(f(x))$.

Theorem 3.4.

1. *Suppose that $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ are injective functions. Then $g \circ f$ is injective.*
2. *Suppose that $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ are surjective functions. Then $g \circ f$ is surjective.*

Proof. Let $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ be functions.

1. Suppose f and g are injective. We must check $g \circ f$ is injective.

   Let $x_1, x_2 \in X$. Suppose that $(g \circ f)(x_1) = (g \circ f)(x_2)$. This means that $g(f(x_1)) = g(f(x_2))$. Since g is injective, this gives $f(x_1) = f(x_2)$. Since f is injective, this gives $x_1 = x_2$.

2. Suppose f and g are surjective. We must check $g \circ f$ is surjective.

   Let $z \in Z$. Since g is surjective, we can choose a $y \in Y$ such that $g(y) = z$. Since f is surjective, we can choose an $x \in X$ such that $f(x) = y$. We now have $(g \circ f)(x) = g(f(x)) = g(y) = z$.

$\square$

Corollary 3.5. *Suppose $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ are bijections. Then $g \circ f$ is a bijection.*

Definition 3.6. If X is a set, the *identity* function is the function

$$1_X : X \longrightarrow X, \ x \longmapsto x.$$

Definition 3.7. Suppose X and Y are sets and $f : X \longrightarrow Y$ is a function.

f is said to be *invertible* **iff** there is a function $g : Y \longrightarrow X$ such that

$$g \circ f = 1_X \text{ and } f \circ g = 1_Y.$$

Theorem 3.8. *A function $f : X \longrightarrow Y$ is invertible if and only if it is bijective.*

4 January 10

4.1 Basic matrix operations

Recall that an $m \times n$ matrix is a grid of numbers with m rows and n colums:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}.$$

We can add two $m \times n$ matrices entrywise, so if $A = (a_{ij})$ and $B = (b_{ij})$, then $A + B = (a_{ij} + b_{ij})$. We can also multiply by a scalar, so if $A = (a_{ij})$, then $\lambda A = (\lambda a_{ij})$.

Given an $m \times n$ matrix A, and an $n \times p$ matrix B, where

$$A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}, \; B = (b_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le p}},$$

we can multiply A and B to obtain an $m \times p$ matrix AB. If

$$AB = (c_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le p}},$$

then there is a concise formula for c_{ij}, the (i, j)-entry of AB, in terms of the entries of A and B:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + a_{i3} b_{3j} + \ldots + a_{in} b_{nj}.$$

One way to think of this is as forming the dot product between the i-th row of A, and the j-th column of B. Thinking about things this way will allow you to multiply two matrices correctly. It gives NO insight into what on earth is going on.

4.2 Thinking about matrix-vector multiplication the right way

A special case of matrix multiplication is multiplying an $m \times n$ matrix and an $n \times 1$ matrix. An $n \times 1$ matrix is a vector. Recall $\mathbb{R}^n$, the set of n-dimensional vectors

$$\left\{ \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} : \; \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R} \right\}.$$

Since these n-dimensional vectors are just $n \times 1$ matrices, you can add them, and multiply them by scalars as described above.

Here is the PURPOSE of an $m \times n$ matrix: multiplication by an $m \times n$ matrix takes a vector in $\mathbb{R}^n$ to a vector in $\mathbb{R}^m$. Said another way, if A is an $m \times n$ matrix, then we can define a function

$$A \cdot (\;) : \mathbb{R}^n \longrightarrow \mathbb{R}^m, \; x \longmapsto Ax.$$

So a matrix does something! It takes a load of vectors to a load of other vectors. In fact, it does this in a special way. It *acts linearly.* Don't worry if you don't know what this means yet.

Consider the vectors $e_1, e_2, \dots, e_n \in \mathbb{R}^n$ defined as follows:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \; e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \; \dots, \; e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

There are a lot of numbers in a matrix and they seem like a mess. Here is the MOST IMPORTANT FACT:

$$Ae_j = \text{the } j\text{-th column of the matrix } A.$$

Summarizing...

- An $m \times n$ matrix A does something: it takes vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$.
- The j-th column of A tells you where e_j goes.

For this reason, when thinking about matrix-vector multiplication, the columns of the matrix are the most important. Write an $m \times n$ matrix A as its column vectors next to each other:

$$A = \left(v_1 \middle| v_2 \middle| \cdots \middle| v_n \right), \quad v_1, v_2, \dots, v_n \in \mathbb{R}^m.$$

Then matrix-vector multiplication is described as follows:

$$A \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

This is the RIGHT way to think about matrix-vector multiplication. Given an $m \times n$ matrix A, and a vector $x \in \mathbb{R}^n$, Ax is a *linear combination* of the columns of the matrix. The components of the vector tell you what scalar multiple of each column of the matrix to take.

4.3 Thinking about matrix multiplication the right way

As long as we're thinking about $m \times n$ matrices as functions $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ (and this is how we'll most often think about them), then there is a RIGHT way to think about matrix multiplication too.

Given an $m \times n$ matrix A, and an $n \times p$ matrix B, write B as its column vectors next to each other

$$B = \left(w_1 \middle| w_2 \middle| \cdots \middle| w_p \right), \quad w_1, w_2, \dots, w_p \in \mathbb{R}^n.$$

Then

$$AB = \left(Aw_1 \middle| Aw_2 \middle| \cdots \middle| Aw_p \right),$$

and each of $Aw_1, Aw_2, \dots, Aw_p \in \mathbb{R}^m$ can be calculated as just described in the previous subsection.

4.4 Vector spaces over a field

Definition 4.4.1. A field is a set $\mathbb{F}$ together with operations

- $+ \; : \mathbb{F} \times \mathbb{F} \longrightarrow \mathbb{F}, \; (\lambda, \mu) \longmapsto \lambda + \mu$ (addition)
- $\cdot \; : \mathbb{F} \times \mathbb{F} \longrightarrow \mathbb{F}, \; (\lambda, \mu) \longmapsto \lambda\mu$ (multiplication)

which satisfy various axioms. We will not concern ourselves with the details of the axioms. Among other things they say that we can subtract, and that we can divide by non-zero numbers.

Example 4.4.2.

1. $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are fields with the notion of addition and multiplication that you know.
2. $\mathbb{F}_2 = \{0, 1\}$ is a field with addition and multiplication defined as follows:

$$\begin{aligned} 0 + 0 = 0, &\;\; 0 \cdot 0 = 0 \\ 0 + 1 = 1, &\;\; 0 \cdot 1 = 0 \\ 1 + 0 = 1, &\;\; 1 \cdot 0 = 0 \\ 1 + 1 = 0, &\;\; 1 \cdot 1 = 1. \end{aligned}$$

   A good way to think about these operations is $0 =$ "even" and $1 =$ "odd".

Definition 4.4.3. A *vector space over a field* $\mathbb{F}$ is a set V together with operations

- $+ \; : V \times V \longrightarrow V, \; (v, w) \longmapsto v + w$ (addition)
- $\cdot \; : \mathbb{F} \times V \longrightarrow V, \; (\lambda, v) \longmapsto \lambda v$ (scalar multiplication)

which satisfy the following axioms ("$\forall$" = "for all", "$\exists$" = "there exists", ":" = "such that"):

1. $\forall u \in V, \; \forall v \in V, \; u + v = v + u$
   (vector space addition is commutative).
2. $\forall u \in V, \; \forall v \in V, \; \forall w \in V, \; (u + v) + w = u + (v + w)$
   (vector space addition is associative).
3. There exists an element of V, which we call 0, with the property that $\forall v \in V, \; v + 0 = v$
   (there is an identity element for vector space addition).
4. $\forall u \in V, \; \exists v \in V : \; u + v = 0$
   (additive inverses exist for vector space addition).
5. $\forall v \in V, \; 1v = v$
   (the multiplicative identity element of the field acts sensibly under scalar multiplication).
6. $\forall \lambda \in \mathbb{F}, \; \forall \mu \in \mathbb{F}, \; \forall v \in V, \; (\lambda\mu)v = \lambda(\mu v)$
   (the interaction of field multiplication and scalar multiplication is sensible).

7. $\forall \lambda \in \mathbb{F},\ \forall u \in V,\ \forall v \in V,\ \lambda(u+v) = \lambda u + \lambda v$
   (the interaction of scalar multiplication and vector space addition is sensible).

8. $\forall \lambda \in \mathbb{F},\ \forall \mu \in \mathbb{F},\ \forall v \in V,\ (\lambda+\mu)v = \lambda v + \mu v$
   (the interaction of field addition and scalar multiplication is sensible).

5 January 12: vector spaces over a field

For the most of this class we'll take $\mathbb{F} = \mathbb{R}$, the real numbers.

Example 5.1.

1. Supppose $n \in \mathbb{N}$. Then the set of n-tuples $\mathbb{R}^n$ is a vector space over $\mathbb{R}$ under coordinatewise addition and scalar multiplication.
2. Suppose $m, n \in \mathbb{N}$. Then the set of real $m \times n$ matrices, $M_{m\times n}(\mathbb{R})$ is a vector space over $\mathbb{R}$ under matrix addition and scalar multiplication.
3. Suppose X is a non-empty set. The set of real-valued functions from X, $\{f : X \longrightarrow \mathbb{R}\}$ is a vector space over $\mathbb{R}$ under pointwise addition and scalar multiplication.
4. Suppose $n \in \mathbb{N}$. The set of degree n real-valued polynomials $\mathcal{P}_n(\mathbb{R})$ is a vector space over $\mathbb{R}$ under coefficientwise addition and scalar multiplication.
5. The set of real-valued polynomials $\mathcal{P}(\mathbb{R}) = \bigcup_{n\in\mathbb{N}} \mathcal{P}_n(\mathbb{R})$ is a vector space over $\mathbb{R}$ under coefficientwise addition and scalar multiplication.

Example 5.2. I made a number of errors while doing these examples in class - sorry about that. Everything is correct here!

1. Define an unusual addition and scalar multiplication on $\mathbb{R}^2$ by

   $$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 - y_2),\ \ \lambda(x_1, x_2) = (\lambda x_1, \lambda x_2).$$

   Axioms 3, 4, 5, 6, 7 all hold; checking 3 and 4 is a good exercise.

   Axioms 1, 2, and 8 fail.

   Let's consider axiom 1: $\forall (x_1, x_2) \in \mathbb{R}^2,\ \forall (y_1, y_2) \in \mathbb{R}^2,\ (x_1, x_2) + (y_1, y_2) = (y_1, y_2) + (x_1, x_2)$. This says $\forall (x_1, x_2) \in \mathbb{R}^2,\ \forall (y_1, y_2) \in \mathbb{R}^2,\ (x_1 + y_1, x_2 - y_2) = (y_1 + x_1, y_2 - x_2)$.

   You might say that this is false because $x_2 - y_2 \neq y_2 - x_2$, but this is *sometimes* true. The best way to demonstrate the falseness of a "for all" statement is to give a very explicit example of its failure. In this case, I would say axiom 1 fails because

   $$(0, 1) + (0, 0) = (0, 1) \neq (0, -1) = (0, 0) + (0, 1).$$

   Similarly, axiom 2 fails because

   $$((0, 0) + (0, 0)) + (0, 1) = (0, 0) + (0, 1) = (0, -1),$$
   $$(0, 0) + ((0, 0) + (0, 1)) = (0, 0) + (0, -1) = (0, 1),$$

   and $(0, -1) \neq (0, 1)$.

Thank you to Jiachen for pointing out that axiom 8 fails. This is because

$$(0+1)(0,1) = 1(0,1) = (0,1) \neq (0,-1) = (0,0) + (0,1) = 0(0,1) + 1(0,1)$$

There's a lesson to be learned here... My laziness in checking the axiom with pen and paper cost me correctness, so don't take shortcuts even if you're the professor!

2. Define an unusual addition and scalar multiplication on $\mathbb{R}^2$ by

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, 0), \;\; \lambda(x_1, x_2) = (\lambda x_1, 0).$$

Axioms 1, 2, 6, 7, 8 all hold.

However, axiom 3 fails. To see this, suppose for contradiction that there exists an element 0 with the property that

$$\forall (x_1, x_2) \in \mathbb{R}^2, (x_1, x_2) + 0 = (x_1, x_2).$$

In particular, by taking $(x_1, x_2) = (0, 1)$, we see that $(0,1)+0 = (0,1)$. However, by definition of addition, the second coordinate of $(0,1) + 0$ is 0, not 1, and this contradicts the previous equation. Thus, there cannot be such an element 0.

Remark 5.3. If you've never seen proof by contradiction before, here's how it works. You want to prove a statement P. "Suppose for contradiction, that P is false." You then make a load of arguments based on this assumption. As soon as you get a contradiction, you should explain what the contradiction is. After that, you can conclude that P is true. In the previous argument, the statement P was "there is not an element 0 with the property that..."

Because axiom 3 fails, axiom 4 does not even make sense.

Axiom 5 also fails. Its negation

$$\exists (x_1, x_2) \in \mathbb{R}^2 : \;\; 1(x_1, x_2) \neq (x_1, x_2)$$

is true because $1(0,1) = (0,0) \neq (0,1)$.

6 January 17: vector spaces over a field

The next theorem proves facts which we will take for granted whenever dealing with vector spaces. I'm not a big fan of axiomatic proofs, but there are some things you should learn from the following proof:

- Introducing elements that you are going to talk about in your argument.
- How to prove an if-then statement directly: you suppose the premise is true, and must verify the conclusion to be true.
- Laying out assumptions clearly, stating what it is necessary to show, and perhaps explaining your idea to the reader.
- Referencing and using axioms clearly.

- What is means to be the unique element with a property, and using uniqueness.
- Reducing a proof to a more easily obtained goal using a previously proven fact, and referencing previously obtained facts.
- Leaving things to the reader and the fact that it is frustrating, so you can see why we deduct points when not everything is properly explained.

Theorem 6.1. *Suppose V is a vector space over a field $\mathbb{F}$.*

1. *Let $u, v, w \in V$. If $u + w = v + w$, then $u = v$.*
2. *The vector $0 \in V$ is the* unique *vector with the property that $\forall v \in V,\ v + 0 = v$.*
3. *Suppose $u \in V$. There is a* unique *vector with the property that $u + v = 0$. We call it $(-u)$.*
4. *Let $\lambda \in \mathbb{F}$ and 0 be the identity additive element of V. Then $\lambda 0 = 0$.*
5. *Let $v \in V$. Then $0v = 0$.*
6. *Let $v \in V$. Then $(-1)v = -v$.*
7. *Let $\lambda \in \mathbb{F}$ and $v \in V$. Then $-(\lambda v) = (-\lambda)v = \lambda(-v)$.*

Proof. Suppose V is a vector space over a field $\mathbb{F}$.

1. Let $u, v, w \in V$. Suppose that $u + w = v + w$. We need to show that $u = v$. The idea is to add an element to both sides to cancel w.

   By axiom 4, we know there exists an element w' such that $w + w' = 0$. We add w' to both sides, and we can summarize our calculation in one line: by using axiom 3, $0 = w + w'$, axiom 2, $u + w = v + w$, axiom 2, $w + w' = 0$, axiom 3, in exactly that order, we find that

   $$u = u + 0 = u + (w + w') = (u + w) + w' = (v + w) + w' = v + (w + w') = v + 0 = v.$$

2. To show $0 \in V$ is unique, we suppose that there exists another such element $0' \in V$ with the property that $\forall v \in V,\ v + 0' = v$. We need to show that $0 = 0'$. This is true because

   $$0 = 0 + 0' = 0' + 0 = 0',$$

   where the first equality uses the property of $0'$, the second equality uses commutativity of addition, and the last equality uses the property of 0.

3. Suppose $u \in V$. Axiom 4 tells us that there is a $v \in V$ such that $u + v = 0$. We must now address uniqueness. Suppose that there is another $v' \in V$ such that $u + v' = 0$. Then we have $u + v = u + v'$. Because vector space addition is commutative, this tells us that $v + u = v' + u$. Using part 1 to cancel the u's, we obtain $v = v'$.

4. Let $\lambda \in \mathbb{F}$ and 0 be the identity additive element of V. We wish to show $\lambda 0 = 0$. By part 1, it is enough to show $\lambda 0 + \lambda 0 = 0 + \lambda 0$:

   $$\lambda 0 + \lambda 0 \overset{7.}{=} \lambda(0 + 0) \overset{3.}{=} \lambda 0 \overset{3.}{=} \lambda 0 + 0 \overset{1.}{=} 0 + \lambda 0.$$

5. Let $v \in V$. We wish to show $0v = 0$. By part 1, it is enough to show that $0v + 0v = 0 + 0v$. In the field $\mathbb{F}$, we have $0 + 0 = 0$, and so

$$0v + 0v \overset{8.}{=} (0 + 0)v = 0v \overset{3.}{=} 0v + 0 \overset{1.}{=} 0 + 0v.$$

6. Let $v \in V$. Then

$$v + (-1)v \overset{5.}{=} 1v + (-1)v \overset{8.}{=} (1 + (-1))v = 0v.$$

By part 5, this is 0, and so $v + (-1)v = 0$. By part 3, $-v$ is the unique element such that $v + (-v) = 0$, so $(-1)v = -v$.

7. Let $\lambda \in \mathbb{F}$ and $v \in V$. Then we have

$$-(\lambda v) = (-1)(\lambda v) \overset{6.}{=} ((-1)\lambda)v = (-\lambda)v = (\lambda(-1))v \overset{6.}{=} \lambda((-1)v) = \lambda(-v),$$

where the unmarked inequalities come from part 6 or field properties (I'll leave it to you to figure out which is which).

□

7 January 19: Subspaces

Definition 7.1. Suppose V is a vector space over $\mathbb{F}$ with operations $+ : V \times V \longrightarrow V$ and $\cdot : \mathbb{F} \times V \longrightarrow V$, and that W is a set.

We call W a *subspace* of V **iff**

- $W \subseteq V$;
- The operations $+$ and $\cdot$ restrict to operations

$$+ : W \times W \longrightarrow W, \;\cdot\; : \mathbb{F} \times W \longrightarrow W;$$

- With these operations W is a vector space over $\mathbb{F}$.

Example 7.2. Suppose that V is a vector space over $\mathbb{F}$ with zero element 0. Then $\{0\}$ and V are subspaces of V.

With the definition above it seems laborious to check that something is a subspace: we have to check it is a vector space in its own right, and we've seen that checking the axioms is tedious. This is why the following theorem, often called the subspace test, is useful.

Theorem 7.3. *Suppose V is a vector space over $\mathbb{F}$ with operations $+ : V \times V \to V$, $\cdot : \mathbb{F} \times V \to V$, zero element 0, and that W is a subset of V. W is a subspace of V if and only if the following three conditions hold:*

1. *$0 \in W$.*
2. *If $w \in W$ and $w' \in W$, then $w + w' \in W$.*
3. *If $\lambda \in \mathbb{F}$ and $w \in W$, then $\lambda w \in W$.*

Proof. Suppose V is a vector space over $\mathbb{F}$ with operations $+ : V \times V \longrightarrow V$ and $\cdot\ : \mathbb{F} \times V \longrightarrow V$, and that W is a subset of V.

First, we show the "only if" direction of the theorem statement. So suppose W is a subspace of V. By definition of what it means to be a subspace, the operations $+$ and $\cdot$ restrict to operations

$$+ : W \times W \longrightarrow W, \ \cdot\ : \mathbb{F} \times W \longrightarrow W.$$

This is exactly the same as saying 2 and 3. So we just have to think about why 1 is true. Because W is a vector space, it has a unique zero element. Right now, for all we know, this zero element could be different to the unique zero element in V, so, for clarity, call the zero element of V, 0_V, and the zero element of W, 0_W. We show that they're not different by showing $0_V = 0_W$. For this, we note that following equalities in V:

$$0_V + 0_W = 0_W + 0_V = 0_W = 0_W + 0_W.$$

The first equality is by commutativity of addition in V. The second equality is because 0_V is the zero element of V. The third equality is because $0_W \in W$, the addition in W coincides with that in V, and 0_W is the zero element of W. By cancellation, we conclude that $0_V = 0_W$, as required.

Next, we show the "if" direction. So suppose that statements 1, 2, and 3 hold. The first part of the definition of a subspace holds because $W \subseteq V$. Assumptions 2 and 3 tell us the operations in V restrict to operations in W. We just have to show W is a vector space over $\mathbb{F}$, i.e. that the axioms hold. Axioms 1, 2, 5, 6, 7, 8 all hold because they hold in V. We just have to think about axioms 3 and 4. By assumption 1, $0 \in W$, and so axiom 3 holds. For axiom 4, suppose $w \in W$. Because V is a vector space, we have the multiplicative inverse $-w \in V$. By theorem 6.1 part 6, we know that $-w = (-1)w$ and by assumption 3, this shows $-w \in W$. $\square$

Remark 7.4. Having proved this theorem, we should never worry EVER AGAIN about the 0 of a subspace being different to the 0 in the containing vector space.

8 January 22

8.1 Subspaces

Example 8.1.1.

1. Suppose $n \in \mathbb{N}$, and $a_1, a_2, \ldots, a_n \in \mathbb{R}$. Then

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n : \ a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0 \right\}$$

is a subsapce of $\mathbb{R}^n$ (with the usual addition and scalar mulitplication). In fact, this is called a *hyperplane*.

2. Suppose $n \in \mathbb{N}$. The set of *symmetric* $n \times n$ matrices

$$\mathcal{S} = \{A \in M_{n\times n}(\mathbb{R}) : \ A^T = A\}$$

is a subspace of $M_{n\times n}(\mathbb{R})$.

3. Suppose $n \in \mathbb{N}$. The set of *skew-symmetric* $n \times n$ matrices

$$\mathfrak{o}(n) = \{A \in M_{n\times n}(\mathbb{R}) : \; A^T = -A\}$$

is a subspace of $M_{n\times n}(\mathbb{R})$.

4. Suppose $n \in \mathbb{N}$. The set of *traceless* $n \times n$ matrices

$$\mathfrak{sl}(n, \mathbb{R}) = \{A \in M_{n\times n}(\mathbb{R}) : \; \mathrm{tr}(A) = 0\}$$

is a subspace of $M_{n\times n}(\mathbb{R})$.

5. Suppose $n \in \mathbb{N}$. The set of *diagonal* $n \times n$ matrices

$$\left\{ \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix} : \; \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R} \right\}$$

is a subspace of $M_{n\times n}(\mathbb{R})$.

6. The set of even functions

$$\{f : \mathbb{R} \longrightarrow \mathbb{R} : \text{ for all } t \in \mathbb{R}, \; f(-t) = f(t)\}$$

is a subspace of $\{f : \mathbb{R} \longrightarrow \mathbb{R}\}$.

7. The set of odd functions

$$\{f : \mathbb{R} \longrightarrow \mathbb{R} : \text{ for all } t \in \mathbb{R}, \; f(-t) = -f(t)\}$$

is a subspace of $\{f : \mathbb{R} \longrightarrow \mathbb{R}\}$.

8. Suppose $n, m \in \mathbb{N}$ and $n < m$. $\mathcal{P}_n(\mathbb{R})$ is a subspace of $\mathcal{P}_m(R)$.

9. Suppose $n \in \mathbb{N}$. $\mathcal{P}_n(\mathbb{R})$ is a subspace of $\mathcal{P}(R)$.

10. You might want to try and say that $\mathcal{P}(R)$ is a subspace of $\{f : \mathbb{R} \longrightarrow \mathbb{R}\}$. However, there is an issue here that needs to be considered carefully. A polynomial is NOT, by definition, a function; it is a formal sum of terms like $a_n x^n$. By allowing yourself to plug in real numbers for the variable, you obtain a function. But what if two different polynomials end up defining the same function? This cannot happen over $\mathbb{R}$, but it can happen if you consider polynomials over $\mathbb{F}_2$, and so this definitely requires justification. Can you prove that the function

$$i : \mathcal{P}(\mathbb{R}) \longrightarrow \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$$

defined by $i(p(x))(t) = p(t)$, is injective? Once the following sentence makes sense, your best strategy is to show that i is linear and that its kernel is $\{0\}$.

Theorem 8.1.2. *Suppose V is a vector space over a field $\mathbb{F}$ (with some operations). Suppose that*

$$\{W_i : i \in I\}$$

is some collection of subspaces. Then the intersection

$$\bigcap_{i\in I} W_i := \{w \in V : \ \forall i \in I, \ w \in W_i\}$$

is a subspace of V.

Proof. Suppose V is a vector space over a field $\mathbb{F}$ (with some operations). Suppose that $\{W_i : i \in I\}$ is some collection of subspaces and let $W = \bigcap_{i\in I} W_i$. We wish to show W is a subspace of V.

First, we have to show $0 \in W$. Since each W_i is a subspace, we have $0 \in W_i$ for all $i \in I$, and so, by definition of the intersection, $0 \in W$.

Next suppose $w \in W$, $w' \in W$, and $\lambda \in \mathbb{F}$. We have to show that $w + w' \in W$ and $\lambda w \in W$. Since each W_i is a subspace, we have $w + w' \in W_i$ and $\lambda w \in W_i$ for all $i \in I$, and so, by definition of the intersection, $w + w' \in W$ and $\lambda w \in W$. $\square$

Remark 8.1.3. In the above, I is an indexing set. For example, if $I = \{1, 2, 3\}$, then

$$\bigcap_{i\in I} W_i = W_1 \cap W_2 \cap W_3.$$

However, I allowed to be infinite in this notation.

8.2 Linear combinations

Definition 8.2.1. Suppose V is a vector space over a field $\mathbb{F}$ and that $v_1, v_2, \ldots, v_n$ are vectors in V. A vector $v \in V$ is said to be *a linear combination of the vectors* $v_1, v_2, \ldots, v_n$ **iff** there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$ such that

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

In this case, we call $\lambda_1, \lambda_2, \ldots, \lambda_n$ the *coefficients* of the linear combination.

Example 8.2.2. Suppose V is a vector space over a field $\mathbb{F}$, and that $v_1, v_2, \ldots, v_n \in V$. Then 0 is a linear combination of $v_1, v_2, \ldots, v_n$. This is because

$$0 = 0v_1 + 0v_2 + \ldots + 0v_n.$$

Example 8.2.3.

1. In $\mathbb{R}^3$, $\begin{pmatrix}1\\8\\1\end{pmatrix}$ is a linear combination of the vectors $\begin{pmatrix}1\\2\\1\end{pmatrix}$, $\begin{pmatrix}0\\\pi\\0\end{pmatrix}$.

2. In $\mathbb{R}^3$, $\begin{pmatrix}0\\0\\1\end{pmatrix}$ is not a linear combination of vectors $\begin{pmatrix}1\\1\\0\end{pmatrix}$, $\begin{pmatrix}0\\1\\1\end{pmatrix}$.

Definition 8.2.4. Suppose V is a vector space over a field $\mathbb{F}$ and that $\emptyset \neq S \subseteq V$. A vector $v \in V$ is said to be *a linear combination of vectors in S* **iff** there exist vectors $v_1, v_2, \ldots, v_n \in S$ such that v is a linear combination of the vectors $v_1, v_2, \ldots, v_n$.

Example 8.2.5. Suppose V is a vector space over a field $\mathbb{F}$, and that $\emptyset \neq S \subseteq V$. Then 0 is a linear combination vectors in S. This is because for any $v \in S$, we have

$$0 = 0v.$$

Moreover, you might view 0 as a sum of no vectors, so that 0 is a linear combination vectors in $\emptyset$.

Definition 8.2.6. Suppose V is a vector space over a field $\mathbb{F}$, and that $\emptyset \neq S \subseteq V$. The *span of S*, denoted $\mathrm{span}(S)$, is the set consisting of all linear combinations of the vectors in S.

$$\mathrm{span}(S) := \{v \in V : v \text{ is a linear combination of vectors in } S\}.$$

It is convenient to define $\mathrm{span}(\emptyset) := \{0\}$.

Remark 8.2.7. The definitions above apply both in the case when S is finite and when S is infinite. In the case that S is nonempty and finite, we can describe $\mathrm{span}(S)$ a little more easily:

$$\mathrm{span}(\{v_1, v_2, \ldots, v_n\}) = \{\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n : \ \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}\}.$$

Theorem 8.2.8. *Suppose V is a vector space over a field $\mathbb{F}$ and $S \subseteq V$. Then span(S) is a subspace of V.*

Proof. Suppose V is a vector space over a field $\mathbb{F}$ and $S \subseteq V$. We wish to show that $\mathrm{span}(S)$ is a subspace of V.

There are two cases. First, when $S = \emptyset$, $\mathrm{span}(S) = \{0\}$ and this is a subspace of V.

We are left with the case when $S \neq \emptyset$ and for this we appeal to the subspace test (thm 7.3).

First, $0 \in \mathrm{span}(S)$: choose any $v \in S$; then we have $0 = 0v$.

Next, suppose that $w, w' \in \mathrm{span}(S)$, and $\mu \in \mathbb{F}$. We will show that $w + w', \mu w \in \mathrm{span}(S)$. By definition of $\mathrm{span}(S)$, there are vectors $u_1, u_2, \ldots, u_n \in S$, and scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$ such that

$$w = \lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_n u_n,$$

and $u'_1, u'_2, \ldots, u'_{n'} \in S$, $\lambda'_1, \lambda'_2, \ldots, \lambda'_{n'} \in \mathbb{F}$ such that

$$w' = \lambda'_1 u'_1 + \lambda'_2 u'_2 + \ldots + \lambda'_{n'} u'_{n'}.$$

The following equations show that $w + w' \in \mathrm{span}(S)$ and $\mu w \in \mathrm{span}(S)$, respectively:

$$w + w' = \lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_n u_n + \lambda'_1 u'_1 + \lambda'_2 u'_2 + \ldots + \lambda'_{n'} u'_{n'},$$

$$\mu w = (\mu\lambda_1) u_1 + (\mu\lambda_2) u_2 + \ldots + (\mu\lambda_n) u_n.$$

□

9 January 24

9.1 Linear combinations

Theorem 9.1.1. *Suppose V is a vector space over a field $\mathbb{F}$ and that $S \subseteq V$. Then span(S) is the smallest subspace of V which contains S.*

Proof. Suppose V is a vector space over a field $\mathbb{F}$ and that $S \subseteq V$. We wish to show $\text{span}(S)$ is the smallest subspace of V which contains S.

There are two cases. First, when $S = \emptyset$, $\text{span}(S) = \{0\}$ and so we are done because $\{0\}$ is the smallest subspace of V, and $\emptyset \subseteq \{0\}$.

We are left with the case when $S \neq \emptyset$. We have already shown that $\text{span}(S)$ is a subspace of V. To see that $\text{span}(S)$ contains S, that is $S \subseteq \text{span}(S)$, note that whenever $v \in S$, we have $v = 1v \in \text{span}(S)$. To show that $\text{span}(S)$ is the smallest such subspace, suppose we have a subspace W of V with $S \subseteq W$; we must must show that $\text{span}(S) \subseteq W$, so let $v \in \text{span}(S)$. By definition, we can write $v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n$, where $v_1, v_2, \ldots, v_n \in S$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$. Since $S \subseteq W$, we have $v_1, v_2, \ldots, v_n \in W$. Since W is a subspace of V, we have $\lambda_1 v_1, \lambda_2 v_2, \ldots, \lambda_n v_n \in W$ and $v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n \in W$. $\square$

Definition 9.1.2. Suppose V is a vector space over a field $\mathbb{F}$, and that $S \subseteq V$. We say that S *spans* V **iff** $\text{span}(S) = V$.

Definition 9.1.3. Suppose V is a vector space over a field $\mathbb{F}$ and that $v_1, v_2, \ldots, v_n$ are vectors in V. We say that $v_1, v_2, \ldots, v_n$ *span* V **iff** $\{v_1, v_2, \ldots, v_n\}$ spans V, i.e. $\text{span}(\{v_1, v_2, \ldots, v_n\}) = V$.

Example 9.1.4.

1. Let $n \in \mathbb{N}$. The set $\{e_1, \ldots, e_n\}$ spans $\mathbb{R}^n$. We could also say that the vectors $e_1, \ldots, e_n$ span $\mathbb{R}^n$.
2. The set $\{e_1, e_1 + e_2, e_2\}$ spans $\mathbb{R}^2$.
3. The matrices $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$ span $\mathfrak{o}(3)$.
4. The matrices $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ span $\mathfrak{sl}(2, \mathbb{R})$.
5. The set $\{1, x, x^2, x^3, \ldots\}$ spans $\mathcal{P}(\mathbb{R})$.

9.2 Linear dependence and linear independence

Definition 9.2.1. Suppose V is a vector space over a field $\mathbb{F}$ and that $v_1, v_2, \ldots, v_n \in V$. The n-tuple $(v_1, v_2, \ldots, v_n)$ is said to be *linearly dependent* **iff** there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{F}$, not all zero, such that

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0. \tag{9.2.2}$$

Remark 9.2.3. Equation (9.2.2) is called a *dependency relation for* $v_1, v_2, \ldots, v_n$. Given vectors $v_1, v_2, \ldots, v_n$ in a vector space V, we always have the *trivial dependency relation*

$$0v_1 + 0v_2 + \ldots + 0v_n = 0.$$

Thus, an n-tuple $(v_1, v_2, \ldots, v_n)$ is linearly dependent precisely when there is a *nontrivial* dependency relation for $v_1, v_2, \ldots, v_n$.

Example 9.2.4. Suppose V is a vector space over a field $\mathbb{F}$ and $v \in V$. The 1-tuple (v) is linearly dependent if and only if $v = 0$.

Proof. Suppose V is a vector space over a field $\mathbb{F}$ and that $v \in V$. First, assume that $v = 0$. Then $1v = 0$ is a nontrivial dependency relation, so the 1-tuple (v) is linearly dependent. Conversely, assume that the 1-tuple (v) is linearly dependent. Then there exists a nonzero scalar $\lambda \in \mathbb{F}$ such that $\lambda v = 0$. Thus, $v = \lambda^{-1}(\lambda v) = \lambda^{-1}0 = 0$. □

Example 9.2.5. Suppose V is a vector space over a field $\mathbb{F}$ and $v \in V$. The 2-tuple (v, v) is linearly dependent. This is becase $1v + (-1)v = 0$ is a non-trivial dependency relation.

Definition 9.2.6. Suppose V is a vector space over a field $\mathbb{F}$ and that $v_1, v_2, \ldots, v_n \in V$. The n-tuple $(v_1, v_2, \ldots, v_n)$ is said to be *linearly independent* **iff** it is not linearly dependent, i.e. the only dependency relation for $v_1, v_2, \ldots, v_n$ is the trivial dependency relation.

By convention, the 0-tuple () is regarded as linearly independent and not linearly dependent.

Example 9.2.7. Suppose V is a vector space over a field $\mathbb{F}$ and $v \in V$. The 1-tuple (v) is linearly independent if and only if $v \neq 0$.

Remark 9.2.8. My definition differs a little from that of the textbook. The authors of the textbook talk about a *set* of vectors being linearly (in)dependent, whereas I talk about an n-tuple of vectors being linearly (in)dependent. Later on in the book, they will want to say that an $n \times n$ matrix is invertible if and only if its column vectors are linearly independent. At this moment, they should realize, but they don't :(, that their definition sucks! The issue is that their definition says the *set*

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$$

is linearly independent, but my definition says that the 2-tuple

$$\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)$$

is linearly dependent. In the end, the concept of a multiset is the best-suited to talking about linear dependence. Even though they are not particularly difficult, I don't expect you to know what they are, so we'll stick with n-tuples (which are elements of a Cartesian product).

Notice that my proofs of theorem 10.1.2 and theorem 11.2 are easier than the corresponding proofs in the textbook, those of theorems 1.7 and 1.9, largely because of these choices of definitions.

Example 9.2.9. Further examples were given in lecture and can be found in homework 3.

10 January 26

10.1 Linear dependence and linear independence

Theorem 10.1.1. *Suppose V is a vector space over a field $\mathbb{F}$, that $n, m \in \mathbb{N} \cup \{0\}$, and that*

$$v_1, \ldots, v_n, v_{n+1}, \ldots, v_{n+m} \in V.$$

If $(v_1, \ldots, v_n)$ is linearly dependent, then $(v_1, \ldots, v_{n+m})$ is linearly dependent; if $(v_1, \ldots, v_{n+m})$ is linearly independent, then $(v_1, \ldots, v_n)$ is linearly independent.

Theorem 10.1.2. *Suppose V is a vector space over a field $\mathbb{F}$, that $n \in \mathbb{N} \cup \{0\}$, that $(v_1, \ldots, v_n)$ is linearly independent in V, and that $v_{n+1} \in V$.*

Then $(v_1, \ldots, v_n, v_{n+1})$ is linearly dependent if and only if $v_{n+1} \in span(\{v_1, \ldots, v_n\})$.

Proof. Suppose V is a vector space over a field $\mathbb{F}$, that $n \in \mathbb{N} \cup \{0\}$, that $(v_1, \ldots, v_n)$ is linearly independent in V, and that $v_{n+1} \in V$.

First, suppose that $v_{n+1} \in \operatorname{span}(\{v_1, \ldots, v_n\})$. By definition of span, there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ such that

$$v_{n+1} = \lambda_1 v_1 + \ldots \lambda_n v_n.$$

$(v_1, \ldots, v_n, v_{n+1})$ is linearly dependent because the following dependency relation is nontrivial:

$$(-\lambda_1)v_1 + \ldots (-\lambda_n)v_n + 1 \cdot v_{n+1} = 0.$$

Now suppose that $(v_1, \ldots, v_n, v_{n+1})$ is linearly dependent. By definition, there is a nontrivial dependency relation:

$$\lambda_1 v_1 + \ldots + \lambda_n v_n + \lambda_{n+1} v_{n+1} = 0.$$

We cannot have $\lambda_{n+1} = 0$, since otherwise we'd have a nontrivial dependency relation

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0,$$

contradicting the fact that $(v_1, \ldots, v_n)$ is linearly independent. Thus, we have

$$v_{n+1} = \lambda_{n+1}^{-1}(\lambda_{n+1} v_{n+1}) = (-\lambda_1 \lambda_{n+1}^{-1})v_1 + \ldots + (-\lambda_n \lambda_{n+1}^{-1})v_n \in \operatorname{span}(\{v_1, \ldots, v_n\}).$$

This proof reads well when $n \in \mathbb{N}$. When $n = 0$, you have to be a bit more careful about what some expressions mean. With the correct reading it is just example 9.2.4 all over again, but you might prefer to divide the proof into two cases. (The textbook fails to point this out.) □

10.2 Bases and dimension

Definition 10.2.1. Suppose V is a vector space over a field $\mathbb{F}$. An n-tuple $(v_1, v_2, \ldots, v_n)$ of vectors in V is said to be a *basis for V* **iff**:

1. $\{v_1, v_2, \ldots, v_n\}$ spans V;
2. $(v_1, v_2, \ldots, v_n)$ is linearly independent.

Example 10.2.2. The 0-tuple () is a basis for $\{0\}$: we have $\operatorname{span}(\emptyset) = \{0\}$ and, by convention, () is linearly independent.

Theorem 10.2.3. *Suppose V is a vector space over a field $\mathbb{F}$. An n-tuple $(v_1, v_2, \ldots, v_n)$ of vectors in V is a basis for V if and only if for all $v \in V$, there are unique $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ such that*

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

11 January 29: Bases and dimension

We would like to make the following definition.

Definition 11.1. Suppose V is a vector space over a field $\mathbb{F}$. We say V is *finite-dimensional* **iff** V is the span of a finite set; otherwise, we say V is *infinite-dimensional.* When V is finite dimensional, the *dimension* of V, written $\dim V$, is the number of elements in a basis for V.

At the present moment, there are two problems with this definition. First, if V is spanned by finitely many elements, does V even have a basis? Second, what if V is finite-dimensional and there are two bases for V with different sizes; then doesn't dimension become ambiguous?

The answer to the first question is "yes," and the answer to the second question is "this cannot happen." We need to prove these answers!

Theorem 11.2. *Suppose V is a vector space over a field $\mathbb{F}$, that $(v_1, v_2, \ldots, v_n)$ is an n-tuple of vectors in V, and that $\{v_1, v_2, \ldots, v_n\}$ spans V. Then some sub-tuple $(v_{m_1}, v_{m_2}, \ldots, v_{m_r})$ is a basis for V. In particular, V has a basis.*

Proof. Suppose V is a vector space over a field $\mathbb{F}$, $(v_1, v_2, \ldots, v_n)$ is an n-tuple of vectors in V, and that $\{v_1, v_2, \ldots, v_n\}$ spans V.

First, we note the two extreme cases. If $(v_1, v_2, \ldots, v_n)$ is linearly independent, then $(v_1, \ldots, v_n)$ is a basis for V and there is nothing to do: we can simply take $m_k = k$ for $k = 1, \ldots, n$. Also, if every $v_i = 0$, then $V = \{0\}$, and the 0-tuple $()$ is a basis for V: we never even have to pick an m_1. Notice that when $n = 0$, these extreme cases coincide.

Suppose that not every $v_i = 0$. First, we shall construct a sub-tuple $(v_{m_1}, \ldots, v_{m_r})$ of $(v_1, \ldots, v_n)$ which is linearly independent and maximal with this property, i.e. adding any vector of $(v_1, \ldots, v_n)$ results in it becoming linearly dependent. Choose m_1 to be the smallest number such that $v_{m_1} \neq 0$. By example 9.2.7, this is the same as choosing m_1 to be the smallest number such that the 1-tuple (v_{m_1}) is linearly independent. While possible, repeat this process, choosing m_2 to be the smallest number such that (v_{m_1}, v_{m_2}) is linearly independent, and m_3 to be the smallest number such that $(v_{m_1}, v_{m_2}, v_{m_3})$ is linearly independent... This process cannot continue forever. This is because we'd find ourselves adding a vector of $(v_1, \ldots, v_n)$ to the tuple twice, and this results in a tuple which is linearly dependent (example 9.2.5). Thus, we obtain $m_1, m_2, \ldots, m_r$ such that $(v_{m_1}, v_{m_2}, \ldots, v_{m_r})$ is linearly independent and adding any vector of $(v_1, \ldots, v_n)$ results in it becoming linearly dependent.

We claim that $(v_{m_1}, v_{m_2}, \ldots, v_{m_r})$ is a basis for V.

First, by construction, we have that $(v_{m_1}, \ldots, v_{m_r})$ is linearly independent. So we just have to show that $\{v_{m_1}, \ldots, v_{m_r}\}$ spans V. i.e. $\operatorname{span}(\{v_{m_1}, \ldots, v_{m_r}\}) = V$. Since $\operatorname{span}(\{v_1, v_2, \ldots, v_n\}) = V$, theorem 9.1.1 tells us that V is the smallest subspace containing $v_1, v_2, \ldots, v_n$. Thus, it is enough to show that $\{v_1, v_2, \ldots, v_n\} \subseteq \operatorname{span}(\{v_{m_1}, \ldots, v_{m_r}\})$. With this goal, let $i \in \{1, \ldots, n\}$. We must show that $v_i \in \operatorname{span}(\{v_{m_1}, \ldots, v_{m_r}\})$. By theorem 10.1.2, it is enough for us to know that $(v_{m_1}, \ldots, v_{m_r}, v_i)$ is linearly dependent, and this is true by construction of $(v_{m_1}, \ldots, v_{m_r})$. □

12 January 31: Bases and dimension

Theorem 12.1 (Replacement theorem)**.** *Let V be a vector space over a field $\mathbb{F}$, and $m, n \in \mathbb{N} \cup \{0\}$.*

Suppose that $v_1, v_2, \ldots, v_m \in V$ and $w_1, w_2, \ldots, w_n \in V$.

Moreover, suppose that $(v_1, v_2, \ldots, v_m)$ is linearly independent and $\{w_1, w_2, \ldots, w_n\}$ spans V.

Then $m \leq n$, and we can pick $n - m$ vectors from $w_1, w_2, \ldots, w_n$, say $w_{k_1}, w_{k_2}, \ldots, w_{k_{n-m}}$ such that $\{v_1, v_2, \ldots, v_m, w_{k_1}, w_{k_2}, \ldots, w_{k_{n-m}}\}$ spans V.

Proof. Let V be a vector space over a field $\mathbb{F}$.

First, we note that if $m = 0$ and $n \in \mathbb{N} \cup \{0\}$, the statement of the theorem is true: $0 \leq n$ and we just pick all the vectors $w_1, w_2, \ldots, w_n$.

Now, suppose that statement of the theorem is true for $m, n \in \mathbb{N} \cup \{0\}$. We would like to show that the result is true for $m + 1, n$. So suppose that $v_1, v_2, \ldots, v_m, v_{m+1} \in V$, $w_1, w_2, \ldots, w_n \in V$, $(v_1, v_2, \ldots, v_m, v_{m+1})$ is linearly independent, and $\{w_1, w_2, \ldots, w_n\}$ spans V.

Since $(v_1, v_2, \ldots, v_m)$ is linearly independent (theorem 10.1.1), we have, by the m, n result, that $m \leq n$, and we can pick $n - m$ vectors from $w_1, w_2, \ldots, w_n$, say $w_{k_1}, w_{k_2}, \ldots, w_{k_{n-m}}$ such that

$$S = \{v_1, v_2, \ldots, v_m, w_{k_1}, w_{k_2}, \ldots, w_{k_{n-m-1}}, w_{k_{n-m}}\}$$

spans V. Therefore, there exists scalars $\lambda_1, \lambda_2, \ldots, \lambda_m, \mu_1, \mu_2, \ldots, \mu_{n-m} \in \mathbb{F}$ such that

$$v_{m+1} = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_m v_m + \mu_1 w_{k_1} + \mu_2 w_{k_2} + \ldots + \mu_{n-m-1} w_{k_{n-m-1}} + \mu_{n-m} w_{k_{n-m}}.$$

Since $(v_1, v_2, \ldots, v_m, v_{m+1})$ is linearly independent, theorem 10.1.2 tells us that

$$v_{m+1} \notin \operatorname{span}(\{v_1, v_2, \ldots, v_m\}).$$

Thus, we must have $n - m > 0$, and at least one μ_i must be nonzero. This gives $n \geq m + 1$. Also, by reordering the w_i's and μ_i's we may as well assume that $\mu_{n-m} \neq 0$. We claim that

$$S' = \{v_1, v_2, \ldots, v_m, v_{m+1}, w_{k_1}, w_{k_2}, \ldots, w_{k_{n-m-1}}\}$$

spans V. This claim would complete our goal of showing that the $m + 1, n$ result is true. Notice that we're *replacing* $w_{k_{n-m}}$ with v_{m+1} to go from S to S' which is why the theorem has its name.

First, we'll show that $S \subseteq \operatorname{span}(S')$. We immediately have

$$\begin{aligned} v_1, v_2, \ldots, v_m, w_{k_1}, w_{k_2}, \ldots, w_{k_{n-m-1}} &\in \operatorname{span}(\{v_1, v_2, \ldots, v_m, v_{m+1}, w_{k_1}, w_{k_2}, \ldots, w_{k_{n-m-1}}\}) \\ &= \operatorname{span}(S'). \end{aligned}$$

Also, because

$$\begin{aligned} w_{k_{n-m}} = &(-\lambda_1 \mu_{n-m}^{-1}) v_1 + (-\lambda_2 \mu_{n-m}^{-1}) v_2 + \ldots + (-\lambda_m \mu_{n-m}^{-1}) v_m + \mu_{n-m}^{-1} v_{m+1} + \\ &(-\mu_1 \mu_{n-m}^{-1}) w_{k_1} + (-\mu_2 \mu_{n-m}^{-1}) w_{k_2} + \ldots + (-\mu_{n-m-1} \mu_{n-m}^{-1}) w_{k_{n-m-1}}. \end{aligned}$$

we have $w_{k_{n-m}} \in \operatorname{span}(\{v_1, v_2, \ldots, v_m, v_{m+1}, w_{k_1}, w_{k_2}, \ldots, w_{k_{n-m-1}}\}) = \operatorname{span}(S')$.

Thus, $S \subseteq \operatorname{span}(S')$. Since $\operatorname{span}(S')$ is a subspace of V, and $\operatorname{span}(S)$ is the *smallest* subspace containing S, we obtain $\operatorname{span}(S) \subseteq \operatorname{span}(S')$. Because $V = \operatorname{span}(S)$, this tells us that $\operatorname{span}(S') = V$, as we claimed. Thus, we have demonstrated that the $m + 1, n$ result follows from the m, n result.

The theorem is true by mathematical induction on n. $\square$

Corollary 12.2. *Let V be a vector space over a field $\mathbb{F}$, and $m, n \in \mathbb{N} \cup \{0\}$.*

Suppose that $(v_1, v_2, \ldots, v_m)$ and $(w_1, w_2, \ldots, w_n)$ are bases for V. Then $m = n$.

Proof. Let V be a vector space over a field $\mathbb{F}$, and $m, n \in \mathbb{N} \cup \{0\}$.

Suppose that $(v_1, v_2, \ldots, v_m)$ and $(w_1, w_2, \ldots, w_n)$ are bases for V. We use the previous theorem twice: since $(v_1, v_2, \ldots, v_m)$ is linearly independent and $\{w_1, w_2, \ldots, w_n\}$ spans V, we have $m \leq n$; since $(w_1, w_2, \ldots, w_n)$ is linearly independent and $\{v_1, v_2, \ldots, v_m\}$ spans V, we have $n \leq m$. Thus, $m = n$. □

13 February 2: Bases and dimension

Theorem 13.1. *Definition 11.1 makes sense.*

Proof. There was no doubt that the definition of "finite-dimensional" and "infinite-dimensional" made sense. Suppose V is a finite-dimensional vector space over $\mathbb{F}$. By definition, V has a finite spanning set. Thus, we can find a tuple $(v_1, v_2, \ldots, v_n)$ such that $\{v_1, v_2, \ldots, v_n\}$ spans V. Theorem 11.2 shows that V has a basis. Moreover, the last corollary shows that any bases have the same number of elements. Thus, $\dim V$ is well-defined. □

Remark 13.2. Definition 11.1 together with theorem 10.2.3 tell us that to describe a vector v in an n-dimensional vector space V, we need n scalars.

Example 13.3.

1. Let $\mathbb{F}$ be a field. The vector space $\{0\}$ has dimension 0.

   Moreover, it is the only vector space over $\mathbb{F}$ of dimension 0.

2. Let $n \in \mathbb{N}$. The real vector space $\mathbb{R}^n$ has dimension n: $(e_1, \ldots, e_n)$ is a basis.

3. Let $m, n \in \mathbb{N}$. The real vector space $M_{m\times n}(\mathbb{R})$ has dimension mn:

   $$(E_{11}, E_{21}, E_{31}, \ldots, E_{m1}, E_{12}, E_{22}, E_{32}, \ldots, E_{m2}, \ldots, E_{1n}, E_{2n}, E_{3n}, \ldots, E_{mn})$$

   is a basis; here

   $$E_{pq} = (\delta_{p,i}\delta_{q,j})_{\substack{1\leq i\leq m,\\ 1\leq j\leq n}} \text{ where } \delta_{k,l} = 1 \text{ if } k = l, \text{ and 0 otherwise.}$$

4. Let $n \in \mathbb{N}$. The real vector space $\mathcal{P}_n(\mathbb{R})$ has dimension $n + 1$: $\{1, x, \ldots, x^n\}$ is a basis.

5. $\mathbb{C}$ can be viewed as a vector space over itself, and as a vector space over $\mathbb{R}$. Over $\mathbb{C}$, (1) is a basis, and so it is 1-dimensional. Over $\mathbb{R}$, $(1, i)$ is a basis, and so it is 2-dimensional.

Theorem 11.2 and the replacement theorem have further implications.

Theorem 13.4. *Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and suppose* $\dim V = n$.

1. *Suppose $\{v_1, v_2, \ldots, v_m\}$ spans V. Then $m \geq n$.*

2. *Suppose $\{v_1, v_2, \ldots, v_n\}$ spans V. Then $(v_1, v_2, \ldots, v_n)$ is a basis for V.*

3. *Suppose $(v_1, \ldots, v_m)$ is linearly independent in V. Then $m \leq n$. Moreover, there are vectors $v_{m+1}, \ldots, v_n \in V$ such that $(v_1, \ldots, v_m, v_{m+1}, \ldots v_n)$ is a basis for V.*

4. *Suppose $(v_1, v_2, \ldots, v_n)$ is linearly independent in V. Then $(v_1, v_2, \ldots, v_n)$ is a basis for V.*

Proof. Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and suppose $\dim V = n$.

1. Suppose $v_1, v_2, \ldots, v_m$ span V. By theorem 11.2, we know that some sub-tuple of $(v_1, \ldots, v_m)$ is a basis of V. Because $\dim V = n$, this sub-tuple has n elements, and so $m \geq n$.

2. Suppose $v_1, v_2, \ldots, v_n$ span V. By theorem 11.2 we know that some sub-tuple of $(v_1, \ldots, v_n)$ is a basis of V. On the other hand, a basis of V must have n elements, so the sub-tuple must be the whole tuple, and $(v_1, \ldots, v_n)$ is a basis.

3. Suppose $(v_1, v_2, \ldots, v_m)$ is linearly independent in V, and let $(b_1, b_2, \ldots, b_n)$ be a basis for V. The replacement theorem tells us that $m \leq n$, and that it is possible for us to pick $n - m$ vectors from $b_1, b_2, \ldots, b_n$, say $b_{k_1}, b_{k_2}, \ldots, b_{k_{n-m}}$ such that $\{v_1, v_2, \ldots, v_m, b_{k_1}, b_{k_2}, \ldots, b_{k_{n-m}}\}$ spans V. Let $v_{m+i} = b_{k_i}$ for $i \in \{1, \ldots, n-m\}$. We know $v_1, \ldots, v_m, v_{m+1}, \ldots v_n$ span V. By part 2, $(v_1, \ldots, v_m, v_{m+1}, \ldots v_n)$ is a basis for V.

4. This is the $m = n$ version of part 3.

□

Parts 2, 4 of the previous theorem are useful when we know the dimension of a finite-dimensional vector space V, and we wish to check if a given tuple of vectors is a basis. As long as it's a $\dim V$-tuple, we only have to check whether its vectors span, or whether it is linearly independent.

Example 13.5. $(1, 8+2x, 3-2x+6x^2, 20+5x-8x^2+9x^3, 88+4000x-31x^2+100000x^3+5x^4)$ is linearly independent in $\mathcal{P}_4(\mathbb{R})$ because of homework 3, question 8.(b). Thus, it is a basis.

Theorem 13.6. *Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and W be a subspace of V. Then W is finite-dimensional and $\dim W \leq \dim V$. Moreover, if $\dim W = \dim V$, then $W = V$.*

Proof. Let V be a finite-dimensional vector space over a field $\mathbb{F}$, and W be a subspace of V. First, we note what happens when $W = \{0\}$. Then W is finite-dimensional because it is spanned by $\emptyset$, and $\dim W = 0 \leq \dim V$. Moreover, if $\dim V = \dim W$, then $\dim V = 0$, and so $V = \{0\} = W$.

Now suppose $W \neq \{0\}$ and let $n = \dim V$. We can choose $w_1 \in W$, such that $w_1 \neq 0$. This is the same as choosing $w_1 \in W$ such that (w_1) is linearly independent. While possible, repeat this process, choosing $w_2 \in W$ such that (w_1, w_2) is linearly independent, and $w_3 \in W$ such that (w_1, w_2, w_3) is linearly independent… Since we cannot have $n+1$ linearly independent vectors (theorem 13.4, part 3), this process must stop. In this way, we contruct a tuple $(w_1, w_2, \ldots, w_m)$, with $m \leq n$, which is linearly independent, and such that for all $w \in W$, $(w_1, w_2, \ldots, w_m, w)$ is linearly dependent. Theorem 10.1.2 and the last statement show that $\{w_1, w_2, \ldots, w_m\}$ spans W, so $(w_1, w_2, \ldots, w_m)$ is a basis for W, W is finite-dimensional, and $\dim W = m \leq n$. If $m = n$, then theorem 13.4, part 2 or 4, shows $(w_1, w_2, \ldots, w_m)$ is a basis for V, so $W = V$. □

Corollary 13.7. *Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and W be a subspace of V. Then any basis for W can be extended to a basis for V.*

Proof. Let $(w_1, w_2, \ldots, w_m)$ be a basis for W. Because $(w_1, w_2, \ldots, w_m)$ is linearly independent in V, theorem 13.4, part 3, says we can extend it to a basis of V. □

14 February 5: Linear transformations

Definition 14.1. Let V and W be vector spaces over a field $\mathbb{F}$. A function $T : V \longrightarrow W$ is said to be a *linear transformation* **iff**

1. $\forall v_1 \in V, \forall v_2 \in V, T(v_1 + v_2) = T(v_1) + T(v_2)$.
2. $\forall \lambda \in \mathbb{F}, \forall v \in V, T(\lambda v) = \lambda T(v)$.

Often, we say just "T is linear."

Lemma 14.2. *Let V and W be vector spaces over a field $\mathbb{F}$ and $T : V \longrightarrow W$ be a function.*

1. *If T is linear, then $T(0) = 0$.*
2. *T is linear if and only if*

$$\forall \lambda \in \mathbb{F},\ \forall v_1 \in V,\ \forall v_2 \in V,\ T(\lambda v_1 + v_2) = \lambda T(v_1) + T(v_2).$$

3. *If T is linear, $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$, and $v_1, \ldots, v_n \in V$, then*

$$T\left(\sum_{k=1}^{n} \lambda_k v_k\right) = \sum_{k=1}^{n} \lambda_k T(v_k).$$

Example 14.3. Suppose $m, n \in \mathbb{N}$ and $A \in M_{m\times n}(\mathbb{R})$. Then $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ defined by $T(x) = Ax$ is a linear transformation. This follows from facts about matrix-vector multiplication.

Example 14.4. Suppose that X is a nonempty set. Recall that $\mathcal{F}(X) = \{f : X \longrightarrow \mathbb{R}\}$ is a vector space over $\mathbb{R}$. Given $x_0 \in X$, $\mathrm{ev}_{x_0} : \mathcal{F}(X) \longrightarrow \mathbb{R}$, $f \longmapsto f(x_0)$ is linear.

Example 14.5. Suppose V is a vector space over a field $\mathbb{F}$. The identity function

$$1_V : V \longrightarrow V,\ v \longmapsto v$$

is a linear transformation.

Example 14.6. Suppose V and W are vector spaces over a field $\mathbb{F}$. The zero function

$$0_{V,W} : V \longrightarrow W,\ v \longmapsto 0$$

is a linear transformation.

Definition 14.7. Suppose U, V, and W are vector spaces over a field $\mathbb{F}$, and that $S : V \longrightarrow W$ and $T : U \longrightarrow V$ are linear transformations. Then $ST : U \longrightarrow W$ is defined by

$$(ST)(u) := S(T(u)),$$

i.e. ST is the composite of $S \circ T$.

Lemma 14.8. *Suppose U, V, and W are vector spaces over a field $\mathbb{F}$, and that $S : V \longrightarrow W$ and $T : U \longrightarrow V$ are linear transformations. Then $ST : U \longrightarrow W$ is a linear transformation.*

Example 14.9. Suppose $a, b \in \mathbb{R}$ with $a < b$. Let $[a,b] = \{x \in \mathbb{R} : \ a \leq x \leq b\}$. Recall that

$$\mathcal{F}([a,b]) = \{f : [a,b] \longrightarrow \mathbb{R}\}$$

is a vector space over $\mathbb{R}$. Let

$$\begin{aligned} \mathcal{F}_{\text{cont}}([a,b]) &= \{f : [a,b] \longrightarrow \mathbb{R} : \ f \text{ is continuous}\}, \\ \mathcal{F}_{\text{diff}}([a,b]) &= \{f : [a,b] \longrightarrow \mathbb{R} : \ f \text{ is differentiable}\}. \end{aligned}$$

In 131A, you prove that $\mathcal{F}_{\text{cont}}([a,b])$ and $\mathcal{F}_{\text{diff}}([a,b])$ are subspaces of $\mathcal{F}([a,b])$ and that

$$\begin{aligned} &\mathcal{D} : \mathcal{F}_{\text{diff}}([a,b]) \longrightarrow \mathcal{F}([a,b]), \quad f \longmapsto f', \\ &\mathcal{I} : \mathcal{F}_{\text{cont}}([a,b]) \longrightarrow \mathbb{R}, \quad f \longmapsto \int_a^b f(t)\ dt, \\ &\mathcal{A} : \mathcal{F}_{\text{cont}}([a,b]) \longrightarrow \mathcal{F}_{\text{diff}}([a,b]), \quad f \longmapsto \left(x \longmapsto \int_a^x f(t)\ dt\right) \end{aligned}$$

are linear transformations. Moreover, the fundamental theorem of calculus part 2 says that

$$\mathcal{D}\mathcal{A} : \mathcal{F}_{\text{cont}}([a,b]) \longrightarrow \mathcal{F}([a,b]), \ f \longmapsto f$$

is the natural inclusion.

Definition 14.10. Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be a linear transformation. We say that T is an *isomorphism* **iff** T is a bijection.

Remark 14.11. It is a theorem in set theory that a function $f : X \longrightarrow Y$ is a bijection if and only if it has an inverse function $g : Y \longrightarrow X$ (this means that $g \circ f = 1_X$ and $f \circ g = 1_Y$).

Theorem 14.12. *Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be an isomorphism. Let $S : W \longrightarrow V$ be the inverse function of T. Then S is linear.*

Proof. Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be an isomorphism. Let $S : W \longrightarrow V$ be the inverse function of T. We wish to show that:

1. $\forall w_1 \in V$, $\forall w_2 \in V$, $S(w_1 + w_2) = S(w_1) + S(w_2)$.
2. $\forall \lambda \in \mathbb{F}$, $\forall w \in W$, $S(\lambda w) = \lambda S(w)$.

So let $w_1, w_2 \in W$. Because T is injective, to show $S(w_1 + w_2) = S(w_1) + S(w_2)$, it is enough to show that $T(S(w_1 + w_2)) = T(S(w_1) + S(w_2))$. Since S is the inverse function to T, the LHS is $w_1 + w_2$. Since T is linear, and S is the inverse function to T,

$$T(S(w_1) + S(w_2)) = T(S(w_1)) + T(S(w_2)) = w_1 + w_2$$

so the RHS is also equal to $w_1 + w_2$. The second condition is verified similarly. $\square$

Corollary 14.13. *Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be a linear transformation. T is an isomorphism if and only if there exists a linear transformation $S : W \longrightarrow V$ such that*

$$ST = 1_V, \ TS = 1_W.$$

15 February 7: Kernels and Images

Definition 15.1. Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be a linear transformation.

The *kernel* of T is defined to be the set

$$\ker T := \{v \in V : \; T(v) = 0\}.$$

The *image* of T is defined to be the set

$$\operatorname{im} T := \{w \in W : \text{ there exists } v \in V \text{ such that } Tv = w\}.$$

Remark 15.2. Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be a linear transformation. By definition, we have $\ker T \subseteq V$ and $\operatorname{im} T \subseteq W$.

Theorem 15.3. *Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be a linear transformation. Then $\ker T$ is a subspace of V, and $\operatorname{im} T$ is a subspace of W.*

Theorem 15.4. *Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be a linear transformation. Then T is injective if and only if $\ker T = \{0\}$.*

Remark 15.5. Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be a linear transformation. Then T is surjective if and only if $\operatorname{im} T = W$.

Corollary 15.6. *Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be a linear transformation. Then T is an isomorphism if and only if $\ker T = \{0\}$ and $\operatorname{im} T = W$.*

Theorem 15.7. *Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be a linear transformation. Suppose that $\{v_1, \ldots, v_n\}$ spans V. Then*

$$\operatorname{im} T = \mathit{span}(\{T(v_1), \ldots, T(v_n)\}).$$

In particular, if V is finite-dimensional, then $\operatorname{im} T$ is finite-dimensional.

Definition 15.8. Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be a linear transformation. Suppose that V is finite-dimensional.

The *nullity* of T, written $\mathit{null}(T)$, is defined to be $\dim(\ker T)$.

The *rank* of T, written $\mathit{rank}(T)$, is defined to be $\dim(\operatorname{im} T)$.

16 February 9: Rank-Nullity

Theorem 16.1 (Rank-Nullity Theorem)**.** *Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be a linear transformation. Suppose that V is finite-dimensional. Then*

$$\mathit{rank}(T) + \mathit{null}(T) = \dim V.$$

Proof. Let V and W be vector spaces over a field $\mathbb{F}$ and let $T : V \longrightarrow W$ be a linear transformation. Suppose that V is finite-dimensional, let $m = \dim V$, and $n = null(T)$. Recall that $n = \dim(\ker T)$ by definition of nullity.

Choose a basis $(v_1, \ldots, v_n)$ for $\ker T$. Extend it to a basis $(v_1, \ldots, v_n, v_{n+1}, \ldots, v_m)$ of V. We claim that $(T(v_{n+1}), \ldots, T(v_m))$ is a basis for $\mathrm{im}(T)$. As long as we prove the claim, we will have $rank(T) = \dim(\mathrm{im}\, T) = m - n = \dim V - null(T)$, and the theorem will follow.

First, we show that $\{T(v_{n+1}), \ldots, T(v_m)\}$ spans W. Since $\{v_1, \ldots, v_n, v_{n+1}, \ldots, v_m\}$ spans V, theorem 15.7 tells us $\{T(v_1), \ldots, T(v_n), T(v_{n+1}), \ldots, T(v_m)\}$ spans W. Since $v_1, \ldots, v_n \in \ker T$, we have $T(v_1) = \ldots = T(v_n) = 0$. So

$$\mathrm{span}(\{T(v_{n+1}), \ldots, T(v_m)\}) = \mathrm{span}(\{T(v_1), \ldots, T(v_n), T(v_{n+1}), \ldots, T(v_m)\}) = W.$$

Now we show that $(T(v_{n+1}), \ldots, T(v_m))$ is linearly independent. Suppose that $\lambda_{n+1}, \ldots, \lambda_m \in \mathbb{F}$ and

$$\lambda_{n+1}T(v_{n+1}) + \ldots + \lambda_m T(v_m) = 0.$$

By linearity of T, this gives $T(\lambda_{n+1}v_{n+1} + \ldots + \lambda_m v_m) = 0$, so that $\lambda_{n+1}v_{n+1} + \ldots + \lambda_m v_m \in \ker T$. Thus, $\lambda_{n+1}v_{n+1} + \ldots + \lambda_m v_m$ is a linear combination of $v_1, \ldots, v_n$, i.e. there exists $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ such that

$$\lambda_1 v_1 + \ldots \lambda_n v_n = \lambda_{n+1}v_{n+1} + \ldots + \lambda_m v_m.$$

Equivalently, $(-\lambda_1)v_1 + \ldots (-\lambda_n)v_n + \lambda_{n+1}v_{n+1} + \ldots + \lambda_m v_m = 0$. Because $(v_1, \ldots, v_n, v_{n+1}, \ldots, v_m)$ is linearly independent, we see that $\lambda_{n+1} = \ldots = \lambda_m = 0$. □

Theorem 16.2. *Let $n \in \mathbb{N} \cup \{0\}$, let V and W be finite-dimensional vector spaces over a field $\mathbb{F}$ each with dimension n, and let $T : V \longrightarrow W$ be a linear transformation. The following conditions are equivalent:*

1. *T is injective;*
2. $\ker T = \{0\}$*;*
3. *$null(T) = 0$;*
4. *T is surjective;*
5. $\mathrm{im}\, T = W$*;*
6. *$rank(T) = n$.*

Proof. Let $n \in \mathbb{N} \cup \{0\}$, let V and W be finite-dimensional vector spaces over a field $\mathbb{F}$ each with dimension n, and let $T : V \longrightarrow W$ be a linear transformation.

Theorem 15.4 says that conditions 1 and 2 are equivalent.
Condition 2 implies condition 3 since $\dim\{0\} = 0$.
Condition 3 implies condition 2 since $\{0\}$ is the *only* vector space over $\mathbb{F}$ with dimension 0.

Remark 15.5 says that conditions 4 and 5 are equivalent.
Condition 5 implies condition 6 since $\dim W = n$.
Condition 6 implies condition 5 by theorem 13.6 applied to $\mathrm{im}\, T \subseteq W$.

Rank-Nullity says $rank(T) + null(T) = n$ and so conditions 3 and 6 are equivalent. □

17 February 12: Classifying linear transformations with domain a finite-dimensional vector space

Theorem 17.1. *Let V and W be vector spaces over a field $\mathbb{F}$. Suppose $(v_1, \ldots, v_n)$ is a basis for V and that $(w_1, \ldots, w_n)$ is a tuple of vectors in W. Then there is exactly one linear transformation $T : V \longrightarrow W$ with the property that for all $j \in \{1, \ldots, n\}$,*

$$T(v_j) = w_j.$$

Sketch proof. Let V and W be vector spaces over a field $\mathbb{F}$. Suppose $(v_1, \ldots, v_n)$ is a basis for V and that $(w_1, \ldots, w_n)$ is a tuple of vectors in W. Given $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$, define $T : V \longrightarrow W$ by

$$T(\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n) = \lambda_1 w_1 + \lambda_2 w_2 + \ldots + \lambda_n w_n.$$

One has to check this is well-defined. This relies on $(v_1, \ldots, v_n)$ is being a basis for V. Given $v \in V$, we can express it as a linear combination of $v_1, v_2, \ldots, v_n$, and doing this allows us to calculate a value for $T(v)$ using the formula above. Writing v as a linear combination in a different way could result in a different answer. Thankfully, since $(v_1, \ldots, v_n)$ is a basis, there is only *one* way to express v as a linear combination of the vectors $v_1, v_2, \ldots, v_n$ and so this problem does not arise.

You should then check that T is linear, and that for all $j \in \{1, \ldots, n\}$, $T(v_j) = w_j$. Also, you should check that if $T' : V \to W$ is linear with the property that for all $j \in \{1, \ldots, n\}$, $T'(v_j) = w_j$, then $T' = T$. □

Corollary 17.2. *Let V, W be vector spaces over a field $\mathbb{F}$, and let $T_1 : V \to W$ and $T_2 : V \to W$ be linear transformations. Suppose that $(v_1, \ldots, v_n)$ is a basis for V, and that for all $j \in \{1, \ldots, n\}$, $T_1(v_j) = T_2(v_j)$. Then $T_1 = T_2$.*

18 February 23: A consequence of homework 4, question 8

Notation 18.1. Suppose V is a vector space over a field $\mathbb{F}$, and that $\alpha = (v_1, \ldots, v_n)$ is a tuple of vectors in V. We use the notation $\psi_\alpha : \mathbb{F}^n \longrightarrow V$ for the unique linear transformation with the property that for all $j \in \{1, \ldots, n\}$,

$$\psi_\alpha(e_j) = v_j.$$

In fact, it has the formula

$$\psi_\alpha\left(\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix}\right) = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

Theorem 18.2. *Suppose V is a vector space over a field $\mathbb{F}$, and that $\alpha = (v_1, \ldots, v_n)$ is a tuple of vectors in V.*

1. *ψ_α is injective if and only if $(v_1, \ldots, v_n)$ is linearly independent.*
2. *ψ_α is surjective if and only if $\{v_1, \ldots, v_n\}$ spans V.*
3. *ψ_α is an isomorphism if and only if $(v_1, \ldots, v_n)$ is a basis for V.*

Proof. Suppose V is a vector space over a field $\mathbb{F}$, and that $\alpha = (v_1, \ldots, v_n)$ is a tuple of vectors in V. Recall $(e_1, \ldots, e_n)$ is a basis for $\mathbb{F}^n$.

Suppose ψ_α is injective. Homework 4, question 8(d) says $(\psi_\alpha(e_1), \ldots, \psi_\alpha(e_n)) = (v_1, \ldots, v_n)$ is linearly independent. Suppose $(v_1, \ldots, v_n) = (\psi_\alpha(e_1), \ldots, \psi_\alpha(e_n))$ is linearly independent. Homework 4, question 8(g) says ψ_α is injective.

Suppose ψ_α is surjective. Homework 4, question 8(e) says $\{\psi_\alpha(e_1), \ldots, \psi_\alpha(e_n)\} = \{v_1, \ldots, v_n\}$ spans W. Suppose $\{v_1, \ldots, v_n\} = \{\psi_\alpha(e_1), \ldots, \psi_\alpha(e_n)\}$ spans W. Homework 4, question 8(h) says ψ_α is surjective.

3. follows from 1. and 2.

This theorem can also be proved directly. Maybe that proof is easier, so it's a good exercise. □

19 February 12: Classifying finite dimensional vector spaces

Theorem 19.1 (Classification theorem)**.** *Let V be a finite-dimensional vector space over a field $\mathbb{F}$. Let $n = \dim V$. Then there exists an isomorphism $\psi : \mathbb{F}^n \longrightarrow V$.*

The orginal proof I gave. Let V be a finite-dimensional vector space over a field $\mathbb{F}$. Let $n = \dim V$. Choose a basis $(v_1, \ldots, v_n)$ for V. Let $(e_1, \ldots, e_n)$ be the standard basis for $\mathbb{F}^n$. By theorem 17.1 there is a unique linear transformation ψ with the property that for all $j \in \{1, \ldots, n\}$,

$$\psi(e_j) = v_j.$$

In fact, it is given by

$$\psi\left(\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix}\right) = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

We wish to show ψ is bijective. Theorem 16.2 tells us that it is enough to show that $\ker \psi = \{0\}$. So suppose

$$\psi\left(\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix}\right) = 0.$$

Then $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda v_n = 0$. Since $(v_1, \ldots, v_n)$ is linearly independent, this tells us that

$$\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix} = 0.$$

□

A condensed proof. Let V be a finite-dimensional vector space over a field $\mathbb{F}$. Choose a basis $(v_1, \ldots, v_n)$ for V. We have the linear transformation $\psi_\beta : \mathbb{F}^n \longrightarrow V$. The previous theorem says ψ_β is an isomorphism. □

Notation 19.2. Suppose V is a vector space over a field $\mathbb{F}$, and that $\beta = (v_1, \ldots, v_n)$ is a basis for V. We have the isomorphism $\psi_\beta : \mathbb{F}^n \longrightarrow V$ defined by

$$\psi_\beta\left(\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix}\right) = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

We use $[v]_\beta$ for $\psi_\beta^{-1}(v)$, so that

$$[\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n]_\beta = \begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix}.$$

20 February 14: LOVE-ly matrices

Recall the following example.

Example. Suppose $m, n \in \mathbb{N}$ and $A \in M_{m\times n}(\mathbb{R})$. Then $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ defined by $T(x) = Ax$ is a linear transformation. This follows from facts about matrix-vector multiplication.

The next theorem tells us that all linear transformations $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ arise in this way and that there is only one matrix defining a given linear transformation. In fact, matrix-vector multiplication is defined the way it is so that this theorem is true.

Remark 20.1. Recall that $(e_1, \ldots, e_n)$ is a basis for $\mathbb{R}^n$ and $(e_1, \ldots, e_m)$ is a basis for $\mathbb{R}^m$. Notice, here, that there is some potential for confusion since e_1 could be referring to a vector in $\mathbb{R}^n$ or $\mathbb{R}^m$. The equations should make it clear where an element lives. Also, recall that for any $(m\times n)$-matrix $C = (c_{ij})$, and $j \in \{1, \ldots, n\}$, we have $Ce_j = \sum_{i=1}^m c_{ij} e_i$. This equation expresses the fact that the j-th column of a matrix tells us where e_j goes.

Theorem 20.2. *Suppose $m, n \in \mathbb{N}$ and $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation. Then there is a unique matrix $A \in M_{m\times n}(\mathbb{R})$ such that for all $x \in \mathbb{R}^n$,*

$$T(x) = Ax.$$

Proof. Suppose $m, n \in \mathbb{N}$ and $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is linear.

First, we show there is an $A \in M_{m\times n}(\mathbb{R})$ such that for all $x \in \mathbb{R}^n$, $T(x) = Ax$.

Let $j \in \{1, \ldots, n\}$. Since $(e_1, \ldots, e_m)$ is a basis for $\mathbb{R}^m$, $T(e_j)$ can be expressed uniquely as a linear combination of these vectors: that is, there are $a_{1j}, a_{2j}, \ldots, a_{mj} \in \mathbb{R}$ such that

$$T(e_j) = \sum_{i=1}^{m} a_{ij} e_i.$$

Do this for each j, and let A be the $(m\times n)$-matrix (a_{ij}). The previous example shows that A defines a linear transformation by $T'(x) = Ax$ and for $j \in \{1, \ldots, n\}$, $T(e_j) = \sum_{i=1}^m a_{ij} e_i = Ae_j = T'(e_j)$. This shows T and T' agree on a basis, and so $T = T'$. Thus, for all $x \in \mathbb{R}^n$, $T(x) = T'(x) = Ax$, and A is the desired matrix.

Now we turn to uniqueness. This is implied by the fact that the a_{ij}'s found above were unique. Writing the argument out in as much detail as possible... Suppose $A = (a_{ij})$ and $B = (b_{ij})$ are two matrices defining T, and let $j \in \{1, \dots, n\}$. Then

$$\sum_{i=1}^{m} a_{ij} e_i = A e_j = T(e_j) = B e_j = \sum_{i=1}^{m} b_{ij} e_i.$$

Since $(e_1, \dots, e_m)$ is a basis for $\mathbb{R}^m$, we deduce from this equation and theorem 10.2.3 that $a_{1j} = b_{1j}$, $a_{2j} = b_{2j}$, ..., $a_{mj} = b_{mj}$. Since j was arbitrary, this shows $(a_{ij}) = (b_{ij})$, i.e. $A = B$. $\square$

Remark 20.3. The previous theorem is true with $\mathbb{R}$ replaced by $\mathbb{F}$. I just wrote it for $\mathbb{R}$ because you liek $\mathbb{R}$ better!

Remark 20.4. Let V and W be vector spaces over a field $\mathbb{F}$, and suppose $T : V \longrightarrow W$ is a linear transformation.

Suppose $\beta_V = (v_1, \dots, v_n)$ is a basis for V and $\beta_W = (w_1, \dots, w_m)$ is a basis for W.

Let $j \in \{1, \dots, n\}$. Since $(w_1, \dots, w_m)$ is a basis for W, $T(v_j)$ can be expressed uniquely as a linear combination of these vectors: that is, there are unique $a_{1j}, a_{2j}, \dots, a_{mj} \in \mathbb{F}$ such that

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Doing this for each j, we obtain an $(m \times n)$-matrix $A = (a_{ij})$.

Definition 20.5. Let V and W be vector spaces over a field $\mathbb{F}$, and suppose $T : V \longrightarrow W$ is a linear transformation.

Suppose $\beta_V = (v_1, \dots, v_n)$ is a basis for V and $\beta_W = (w_1, \dots, w_m)$ is a basis for W.

A matrix

$$A = (a_{ij})_{\substack{1 \le i \le m, \\ 1 \le j \le n}}$$

is said to be *the matrix of the linear transformation T with respect to the bases β_V and β_W* **iff** for all $j \in \{1, \dots, n\}$,

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i. \tag{20.6}$$

In this case, we write $[T]_{\beta_V}^{\beta_W}$ for the matrix A.

Example 20.7. Suppose $m, n \in \mathbb{N}$, $A = (a_{ij}) \in M_{m \times n}(\mathbb{R})$ and $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is defined by

$$T(x) = Ax.$$

Since $Ae_j = \sum_{i=1}^{m} a_{ij} e_i$, the matrix of T with respect to the bases $(e_1, \dots, e_n)$ and $(e_1, \dots, e_m)$ is the original matrix A.

Remark 20.8. Let V and W be vector spaces over $\mathbb{F}$, and suppose that $T : V \longrightarrow W$ is a linear transformation. Suppose that $\beta_V = (v_1, \dots, v_n)$ is a basis for V and $\beta_W = (w_1, \dots, w_m)$ is a basis for W. Then

$$[T]_{\beta_V}^{\beta_W} [v]_{\beta_V} = [T(v)]_{\beta_W}.$$

Remark 20.9. Let V and W be vector spaces over $\mathbb{R}$, and suppose that $T : V \longrightarrow W$ is a linear transformation. Suppose that $\beta_V = (v_1, \ldots, v_n)$ is a basis for V and $\beta_W = (w_1, \ldots, w_m)$ is a basis for W. By the proof of theorem 19.1, these bases determine isomorphisms $\psi_{\beta_V} : \mathbb{R}^n \longrightarrow V$ and $\psi_{\beta_W} : \mathbb{R}^m \longrightarrow W$. The point of the matrix A of the linear transformation T with respect to the bases β_V and β_W is that it makes the following square commute: that is, it does not matter which way we compose the linear transformations in the square, they give the same result.

$$
\begin{tikzcd}
e_j \arrow[rrr, mapsto] \arrow[ddd, mapsto] & & & \sum_{i=1}^m a_{ij} e_i \arrow[ddd, mapsto] \\
& \mathbb{R}^n \arrow[r, "A\cdot(-)"] \arrow[d, "\psi_{\beta_V}"', "\cong"] & \mathbb{R}^m \arrow[d, "\psi_{\beta_W}"', "\cong"] & \\
& V \arrow[r, "T"] & W & \\
v_j \arrow[rrr, mapsto] & & & \sum_{i=1}^m a_{ij} w_i
\end{tikzcd}
$$

$$
\begin{tikzcd}
{[v]_{\beta_V}} \arrow[rrr, mapsto] & & & \bullet \\
& \mathbb{R}^n \arrow[r, "{[T]_{\beta_V}^{\beta_W}\cdot(-)}"] & \mathbb{R}^m & \\
& V \arrow[r, "T"] \arrow[u, "{[-]_{\beta_V}}", "\cong"'] & W \arrow[u, "{[-]_{\beta_W}}", "\cong"'] & \\
v \arrow[rrr, mapsto] \arrow[uuu, mapsto] & & & T(v) \arrow[uuu, mapsto]
\end{tikzcd}
$$

By using the inverses of the isomorphisms instead and considering $\bullet$ in the diagram above, we see that the definition of the matrix of a linear transformation makes the following formula true:

$$[T]_{\beta_V}^{\beta_W} [v]_{\beta_V} = [T(v)]_{\beta_W}.$$

21 February 16: Composition and matrix multiplication

Let's recall the definition of matrix multiplication.

Definition 21.1. Suppose A is an $m \times n$, and B is an $n \times p$ matrix B:

$$A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}, \; B = (b_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le p}}.$$

For $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, p\}$, let

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + a_{i3} b_{3j} + \ldots + a_{in} b_{nj}.$$

The matrix AB is defined to be

$$(c_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le p}}.$$

Matrix-matrix multiplication is defined the way it is so that the following theorem is true.

Theorem 21.2. *Let U, V, and W be vector spaces over a field $\mathbb{F}$, and suppose $S : V \longrightarrow W$ and $T : U \longrightarrow V$ are linear transformations.*

Suppose $\beta_U = (u_1, \ldots, u_p)$, $\beta_V = (v_1, \ldots, v_n)$, and $\beta_W = (w_1, \ldots, w_m)$ are bases for U, V, W, respectively.

Suppose that $A = (a_{ij})$ is the matrix of the linear transformation S with respect to the bases β_V and β_W, and that $B = (b_{ij})$ is the matrix of the linear transformation T with respect to the bases β_U and β_V.

Then the matrix of ST with respect to the bases β_U and β_W is AB.

Proof. Let U, V, and W be vector spaces over a field $\mathbb{F}$, and suppose $S : V \longrightarrow W$ and $T : U \longrightarrow V$ are linear transformations. Suppose $\beta_U = (u_1, \ldots, u_p)$, $\beta_V = (v_1, \ldots, v_n)$, and $\beta_W = (w_1, \ldots, w_m)$ are bases for U, V, W, respectively.

By definition, the matrix $A = (a_{ik})$ of the linear transformation S with respect to the bases β_V and β_W is determined by the equations

$$S(v_k) = \sum_{i=1}^{m} a_{ik} w_i, \; k \in \{1, \ldots, n\},$$

and the matrix $B = (b_{kj})$ of the linear transformation T with respect to the bases β_U and β_V is determined by the equations

$$T(u_j) = \sum_{k=1}^{n} b_{kj} v_k, \; j \in \{1, \ldots, p\}.$$

So, for $j \in \{1, \ldots, p\}$,

$$\begin{aligned}(ST)(u_j) = S(T(u_j)) = S\left(\sum_{k=1}^{n} b_{kj} v_k\right) &= \sum_{k=1}^{n} b_{kj} S(v_k) \\ &= \sum_{k=1}^{n} b_{kj} \left(\sum_{i=1}^{m} a_{ik} w_i\right) = \sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{ik} b_{kj}\right) w_i.\end{aligned}$$

Thus, the matrix of ST with respect to the bases β_U and β_W is $C = (c_{ij})$ where

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

This is the definition of AB. □

Remark 21.3. The last theorem says that $[ST]_{\beta_U}^{\beta_W} = [S]_{\beta_V}^{\beta_W}[T]_{\beta_U}^{\beta_V}$.

Remark 21.4. Let U, V, and W be vector spaces over $\mathbb{R}$, and suppose that $S : V \longrightarrow W$ and $T : U \longrightarrow V$ are linear transformations.

Suppose $\beta_U = (u_1, \ldots, u_p)$, $\beta_V = (v_1, \ldots, v_n)$, and $\beta_W = (w_1, \ldots, w_m)$ are bases for U, V, W, respectively. By the proof of theorem 19.1, these bases determine isomorphisms $\psi_{\beta_U} : \mathbb{R}^p \longrightarrow U$, $\psi_{\beta_V} : \mathbb{R}^n \longrightarrow V$, and $\psi_{\beta_W} : \mathbb{R}^m \longrightarrow W$.

Suppose that $A = (a_{ij})$ is the matrix of the linear transformation S with respect to the bases β_V and β_W, and that $B = (b_{ij})$ is the matrix of the linear transformation T with respect to the bases β_U and β_V.

The reason the previous theorem is true is that the following diagram commutes.

$$\begin{tikzcd}
\mathbb{R}^p \arrow[r, "B\cdot(-)"] \arrow[d, "\psi_{\beta_U}"', "\cong"] \arrow[rr, bend left, "(AB)\cdot(-)"] & \mathbb{R}^n \arrow[r, "A\cdot(-)"] \arrow[d, "\psi_{\beta_V}"', "\cong"] & \mathbb{R}^m \arrow[d, "\psi_{\beta_W}"', "\cong"] \\
U \arrow[r, "T"] \arrow[rr, bend right, "ST"'] & V \arrow[r, "S"] & W
\end{tikzcd}$$

The squares commute because of the previous diagramed remark. The bottom triangle commutes by definition of ST. The top square commutes because, for all $x \in \mathbb{R}^p$, $(AB)x = A(Bx)$.

Wanting this last formula to be true forces matrix-matrix multiplication to be what it is. This is because, once we understand matrix-vector multiplication, the equation $(AB)e_j = A(Be_j)$ says that the j-th column of AB is A times the j-th column of B.

22 February 23: Isomorphisms and invertible matrices

Theorem 22.1. *Let V and W be vector spaces over a field $\mathbb{F}$, and suppose that $T : V \longrightarrow W$ is a linear transformation. If V is finite-dimensional and T is surjective, then W is finite-dimensional. If W is finite-dimensional and T is injective, then V is finite-dimensional.*

Theorem 22.2. *Let V and W be finite-dimensional vector spaces over a field $\mathbb{F}$, and suppose that $T : V \longrightarrow W$ is a linear transformation.*

1. *If T is injective,* $\dim V \leq \dim W$.
2. *If T is surjective,* $\dim V \geq \dim W$.

3. If T is an isomorphism, then $\dim V = \dim W$.

Proof. The first two statements are the contrapositive of Homework 4, question 9. The last follows from the first two. $\square$

Theorem 22.3. *Let V and W be finite-dimensional vector spaces over a field $\mathbb{F}$. If* $\dim V = \dim W$*, then there exists an isomorphism $T : V \longrightarrow W$.*

Proof. Let V and W be finite-dimensional vector spaces over a field $\mathbb{F}$. Suppose that $\dim V = \dim W$ and call this number n. Let $\beta_V = (v_1, \ldots, v_n)$ be a basis for V, and let $\beta_W = (w_1, \ldots, w_n)$ be a basis for W. We have the isomorphisms $\psi_{\beta_V} : \mathbb{F}^n \to V$, $\psi_{\beta_W} : \mathbb{F}^n \to W$. Let $T = \psi_{\beta_W}\psi_{\beta_V}^{-1}$. Then $T : V \longrightarrow W$ is linear and T is an isomorphism.

Notice that for $j \in \{1, \ldots, n\}$, we have $T(v_j) = \psi_{\beta_W}(\psi_{\beta_V}^{-1}(v_j)) = \psi_{\beta_W}(e_j) = w_j$. $\square$

Theorem 22.4. *Let V and W be vector spaces over a field $\mathbb{F}$.*

Suppose $T : V \longrightarrow W$ is a linear transformation, and that $(v_1, \ldots, v_n)$ is a basis for V. Then T is an isomorphism if and only if $(T(v_1), \ldots, T(v_n))$ is a basis for W.

Proof. Homework 4, question 8(f) and (i). $\square$

Remark 22.5. Suppose $m, n \in \mathbb{N}$ and $A \in M_{m\times n}(\mathbb{R})$. Recall that $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ defined by $T_A(x) = Ax$ is a linear transformation. Theorem 22.2 tells us that a necessary condition for T_A to be an isomorphism is $m = n$, i.e. A needs to be a square matrix.

Remark 22.6. Suppose $n \in \mathbb{N}$ and $A \in M_{n\times n}(\mathbb{R})$ and consider $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. Corollary 14.13 says T_A is an isomorphism if and only if T_A has an inverse $S : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. Theorem 20.2 says that such an inverse must be of the form T_B for a unique matrix $B \in M_{n\times n}(\mathbb{R})$. We have $T_A T_B = T_{AB}$ and $T_B T_A = T_{BA}$. Moreover, theorem 20.2 says that $T_C = 1_{\mathbb{R}^n}$ if and only if $C = I_n$. Thus, T_A is an isomorphism if and only if there is a $B \in M_{n\times n}(\mathbb{R})$ such that $AB = BA = I_n$.

Definition 22.7. We say a matrix $A \in M_{n\times n}(\mathbb{F})$ is *invertible* **iff** there is a matrix $B \in M_{n\times n}(\mathbb{F})$ such that

$$AB = BA = I_n.$$

The matrix B is often written as A^{-1} and is called the *inverse* of A.

Non-square matrices are never called invertible.

The previous remarks prove the following theorem.

Theorem 22.8. *A matrix defines an isomorphism if and only if it is invertible.*

Theorem 22.9. *Suppose V is a vector space over a field $\mathbb{F}$, that $\beta = (v_1, \ldots, v_n)$ is a basis for V, and $T : V \longrightarrow V$ is a linear transformation. Then $T = 1_V$ if and only if $[T]_\beta^\beta = I_n$.*

Proof. Suppose V is a vector space over a field $\mathbb{F}$, that $\beta = (v_1, \ldots, v_n)$ is a basis for V, and that $T : V \longrightarrow V$ is a linear transformation.

$I_n = (\delta_{i,j})$ (see example 13.3 for notation) and so $[T]_\beta^\beta = I_n$ is equivalent to

$$T(v_j) = \sum_{i=1}^{n} \delta_{i,j} v_i \text{ for all } j \in \{1, \ldots, n\},$$

i.e. $T(v_j) = v_j$ for all $j \in \{1, \ldots, n\}$.

Since $(v_1, \ldots, v_n)$ is a basis for V, this is equivalent to $T = 1_V$. $\square$

Theorem 22.10. *Let $n \in \mathbb{N}$, let V and W be finite-dimensional vector spaces over a field $\mathbb{F}$ each with dimension n, and let $T : V \longrightarrow W$ be a linear transformation. Suppose $\beta_V = (v_1, \ldots, v_n)$ is a basis for V and $\beta_W = (w_1, \ldots, w_n)$ is a basis for W.*

T is an isomorphism if and only if $[T]^{\beta_W}_{\beta_V}$ is invertible.

Proof. Let $n \in \mathbb{N}$, let V and W be finite-dimensional vector spaces over a field $\mathbb{F}$ each with dimension n, and let $T : V \longrightarrow W$ be a linear transformation. Suppose $\beta_V = (v_1, \ldots, v_n)$ is a basis for V and $\beta_W = (w_1, \ldots, w_n)$ is a basis for W.

Suppose T is an isomorphism. Then T has an inverse $S : W \longrightarrow V$. We have

$$[S]^{\beta_V}_{\beta_W}[T]^{\beta_W}_{\beta_V} = [ST]^{\beta_V}_{\beta_V} = [1_V]^{\beta_V}_{\beta_V} = I_n$$

and

$$[T]^{\beta_W}_{\beta_V}[S]^{\beta_V}_{\beta_W} = [TS]^{\beta_W}_{\beta_W} = [1_W]^{\beta_W}_{\beta_W} = I_n,$$

so $[T]^{\beta_W}_{\beta_V}$ is invertible.

Let $B = [T]^{\beta_W}_{\beta_V}$, and suppose B is invertible. This means that there exists $A \in M_{n\times n}(\mathbb{F})$ such that $AB = BA = I_n$. Write $A = (a_{ij})$ and use theorem 17.1 to define $S : W \longrightarrow V$ by

$$S(w_j) = \sum_{i=1}^{n} a_{ij} v_i, \; j \in \{1, \ldots, n\}.$$

We have $[S]^{\beta_V}_{\beta_W} = A$. Thus,

$$[ST]^{\beta_V}_{\beta_V} = [S]^{\beta_V}_{\beta_W}[T]^{\beta_W}_{\beta_V} = AB = I_n$$

and

$$[TS]^{\beta_W}_{\beta_W} = [T]^{\beta_W}_{\beta_V}[S]^{\beta_V}_{\beta_W} = BA = I_n.$$

The previous theorem implies that $ST = 1_V$ and $TS = 1_W$, so T is an isomorphism. □

Recall theorem 16.2 which says that a linear transformation between finite-dimensional vector spaces of the same dimension is injective if and only if it is surjective.

Theorem 22.11. *Let $n \in \mathbb{N}$ and suppose $A, B \in M_{n\times n}(\mathbb{R})$. Then $AB = I_n$ if and only if $BA = I_n$.*

Proof. Let $n \in \mathbb{N}$ and suppose that $A, B \in M_{n\times n}(\mathbb{R})$.

Assume $AB = I_n$. Then you can check that $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is surjective. Thus, T_A is injective, and so T_A is an isomorphism. This means A is invertible and so there exists a $C \in M_{n\times n}(\mathbb{R})$ such that $AC = CA = I_n$. We have

$$B = I_n B = (CA)B = C(AB) = CI_n = C.$$

Thus $BA = CA = I_n$.

Assume $BA = I_n$. Then you can check that $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is injective. Thus, T_A is surjective, and so T_A is an isomorphism. This means A is invertible. So there exists a $C \in M_{n\times n}(\mathbb{R})$ such that $AC = CA = I_n$. We have

$$B = BI_n = B(AC) = (BA)C = I_nC = C.$$

Thus $AB = AC = I_n$. □

23 February 21: The change of coordinate matrix

Definition 23.1. Suppose V is a vector space over a field $\mathbb{F}$, that $\beta = (v_1, \ldots, v_n)$ and $\beta' = (v'_1, \ldots, v'_n)$ are bases for V. Then $[1_V]^{\beta}_{\beta'}$ is called the *change of coordinate matrix from* β' *to* β.

Remark 23.2. Theorem 22.9 only concerns one basis β. $[1_V]^{\beta}_{\beta'}$ does not have to be the identity. It is not the identity if $\beta \neq \beta'$.

Remark 23.3. Suppose V is a vector space over a field $\mathbb{F}$, and that $\beta = (v_1, \ldots, v_n)$ and $\beta' = (v'_1, \ldots, v'_n)$ are bases for V. Write $[1_V]^{\beta}_{\beta'}$ as $Q = (q_{ij})$. Then, for $j \in \{1, \ldots, n\}$, we have

$$v'_j = 1_V(v'_j) = \sum_{i=1}^{n} q_{ij} v_i.$$

Also, for $v \in V$, we have $[v]_\beta = [1_V(v)]_\beta = [1_V]^{\beta}_{\beta'}[v]_{\beta'} = Q[v]_{\beta'}$.

Finally, both of these equations allow one to see that the j-th column of Q is $[v'_j]_\beta$.

Theorem 23.4. *Let V and W be vector spaces over $\mathbb{R}$, and suppose that $T : V \longrightarrow W$ is a linear transformation. Suppose that $\beta_V = (v_1, \ldots, v_n)$ and $\beta'_V = (v'_1, \ldots, v'_n)$ are bases for V, and $\beta_W = (w_1, \ldots, w_m)$ and $\beta'_W = (w'_1, \ldots, w'_m)$ are bases for W.*

Suppose $A = (a_{ij})$ is the matrix of T with respect to the bases β_V and β_W, and that $A' = (a'_{ij})$ is the matrix of T with respect to the bases β'_V and β'_W.

Suppose that Q is the change of coordinate matrix from β'_V to β_V, and that P is the change of coordinate matrix from β'_W to β_W. Then

$$A' = P^{-1}AQ.$$

Proof. $A' = [T]^{\beta'_W}_{\beta'_V} = [1_W T 1_V]^{\beta'_W}_{\beta'_V} = [1_W]^{\beta'_W}_{\beta_W}[T]^{\beta_W}_{\beta_V}[1_V]^{\beta_V}_{\beta'_V} = ([1_W]^{\beta_W}_{\beta'_W})^{-1}[T]^{\beta_W}_{\beta_V}[1_V]^{\beta_V}_{\beta'_V} = P^{-1}AQ.$ $\square$

Remark 23.5. Here's the diagram associated to the previous theorem.

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "Q\cdot(-)", "\cong"'] \arrow[d, "\psi_{\beta'_V}"', "\cong"] \arrow[rrr, bend left, "(A')\cdot(-)"] & \mathbb{R}^n \arrow[r, "A\cdot(-)"] \arrow[d, "\psi_{\beta_V}"', "\cong"] & \mathbb{R}^m \arrow[d, "\psi_{\beta_W}"', "\cong"] & \mathbb{R}^m \arrow[l, "P\cdot(-)"', "\cong"] \arrow[d, "\psi_{\beta'_W}"', "\cong"] \\
V \arrow[r, "1_V"] \arrow[rrr, bend right, "T"'] & V \arrow[r, "T"] & W & W \arrow[l, "1_W"']
\end{tikzcd}$$

Corollary 23.6. *Suppose $m, n \in \mathbb{N}$ and $A \in M_{m\times n}(\mathbb{R})$. Let $\beta_{\mathbb{R}^n} = (v_1, \dots, v_n)$ be a basis of $\mathbb{R}^n$ and $\beta_{\mathbb{R}^m} = (w_1, \dots, w_m)$ be a basis of $\mathbb{R}^m$. Then the matrix of the linear transformation $x \longmapsto Ax$ with respect to the bases $\beta_{\mathbb{R}^n}$ and $\beta_{\mathbb{R}^m}$ is given by*

$$\left(w_1 \middle| \cdots \middle| w_m\right)^{-1} A \left(v_1 \middle| \cdots \middle| v_n\right).$$

Example 23.7. See the homework solutions for homework 4 question 10, after part (h), for some good examples.

24 February 26: Determinants

Notation 24.1. Suppose $A \in M_{n\times n}(\mathbb{F})$. Denote the $(n-1)\times(n-1)$ matrix obtained from A by deleting the i-th row and j-th column by $\widetilde{A_{ij}}$.

Theorem 24.2. *For each $n \in \mathbb{N}$, there is a function called the* determinant*, $\det : M_{n\times n}(\mathbb{F}) \longrightarrow \mathbb{F}$. The functions have the following properties.*

1. (a) *If* $A \in M_{1\times 1}(\mathbb{F})$ *so that* $A = (a_{11})$, *then* $\det A = a_{11}$.

   (b) *If* $A = (a_{ij}) \in M_{n\times n}(\mathbb{F})$ *where* $n \geq 2$, *and* $j \in \{1, \dots, n\}$, *then*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \cdot \det(\widetilde{A_{ij}}).$$

2. (a) $\det(I_n) = 1$.

   (b) *Suppose* $j \in \{1, \dots, n\}$, $v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n \in \mathbb{F}^n$, $u, w \in \mathbb{F}^n$, *and* $\lambda \in \mathbb{F}$. *Then*

$$\begin{aligned} &\det\left(v_1 \middle| \cdots \middle| v_{j-1} \middle| u + \lambda w \middle| v_{j+1} \middle| \cdots \middle| v_n\right) \\ =\ &\det\left(v_1 \middle| \cdots \middle| v_{j-1} \middle| u \middle| v_{j+1} \middle| \cdots \middle| v_n\right) + \lambda \det\left(v_1 \middle| \cdots \middle| v_{j-1} \middle| w \middle| v_{j+1} \middle| \cdots \middle| v_n\right). \end{aligned}$$

   This says det *is linear in each column.*

   (c) *Suppose* $j, k \in \{1, \dots, n\}$, $j < k$, $v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_{k-1}, v_{k+1}, \dots, v_n \in \mathbb{F}^n$, $v \in \mathbb{F}^n$. *Then*

$$\det\left(v_1 \middle| \cdots \middle| v_{j-1} \middle| v \middle| v_{j+1} \middle| \cdots \middle| v_{k-1} \middle| v \middle| v_{k+1} \middle| \cdots \middle| v_n\right) = 0.$$

   This says that if two columns are the same, then det *is zero.*

3. (a) *Suppose* $j \in \{1, \dots, n\}$ *and* $v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n \in \mathbb{F}^n$. *Then*

$$\det\left(v_1 \middle| \cdots \middle| v_{j-1} \middle| 0 \middle| v_{j+1} \middle| \cdots \middle| v_n\right) = 0.$$

   This says that if a column is zero, then det *is zero.*

   (b) *Suppose* $j, k \in \{1, \dots, n\}$, $j < k$, *and* $v_1, \dots, v_n \in \mathbb{F}^n$. *Then*

$$\det\left(v_1 \middle| \cdots \middle| v_{j-1} \middle| v_k \middle| v_{j+1} \middle| \cdots \middle| v_{k-1} \middle| v_j \middle| v_{k+1} \middle| \cdots \middle| v_n\right) = -\det\left(v_1 \middle| \cdots \middle| v_n\right).$$

   This says that swapping two columns introduces a minus sign to the determinant.

   (c) *Suppose* $j, k \in \{1, \dots, n\}$, $j \neq k$, $v_1, \dots, v_n \in \mathbb{F}^n$, *and* $\lambda \in \mathbb{F}$. *Then*

$$\det\left(v_1 \middle| \cdots \middle| v_{j-1} \middle| v_j + \lambda v_k \middle| v_{j+1} \middle| \cdots \middle| v_n\right) = \det\left(v_1 \middle| \cdots \middle| v_n\right).$$

   This says that adding a scalar multiple of one column to another column does not change the determinant.

4. *If* $A \in M_{n\times n}(\mathbb{F})$, *then* $\det(A^T) = \det(A)$.

5. *If* $A, B \in M_{n\times n}(\mathbb{F})$, *then* $\det(AB) = \det(A)\det(B)$.

6. *If* $A \in M_{n\times n}(\mathbb{F})$, A *is invertible if and only if* $\det(A) \neq 0$.

Remark 24.3. If we always take $j = 1$, then 1 provides a definition of the determinant.

The determinant is the only function with the properties of 2.

Property 3a follows from 2b. Property 3b follows from 2b and 2c, and the same is true for 3c.

Property 4 implies that for every property involving columns, there is a corresponding property for rows.

Properties 5 and 6 are useful.

It is good practice to prove the following theorem.

Theorem 24.4. *Suppose properties 3a and 3c. Then we can prove half of property* 6 *quickly, that if* $A \in M_{n\times n}(\mathbb{F})$ *is not invertible then* $\det A = 0$.

Proof. Suppose properties 3a and 3c and that $A \in M_{n\times n}(\mathbb{F})$ is not invertible. Write

$$A = \left(v_1 \middle| \cdots \middle| v_n\right)$$

where $v_1, \ldots, v_n \in \mathbb{F}^n$ and let $\alpha = (v_1, \ldots, v_n)$.

The linear transformation that A defines is ψ_α (recall 18.1). Since A is not invertible, ψ_α is not invertible. Since ψ_α is a linear transformation between vector spaces of this same dimension, this tells us ψ_α is not injective, so α is linearly dependent (18.2). By homework 3, 6.(a), this means we can rewrite some v_j as a linear combination of the other vectors

$$v_j = \lambda_1 v_1 + \ldots + \lambda_{j-1}v_{j-1} + \lambda_{j+1}v_{j+1} + \ldots + \lambda_n v_n.$$

So

$$\det A = \det\left(v_1 \middle| \cdots \middle| v_{j-1} \middle| \lambda_1 v_1 + \ldots + \lambda_{j-1}v_{j-1} + \lambda_{j+1}v_{j+1} + \ldots + \lambda_n v_n \middle| v_{j+1} \middle| \cdots \middle| v_n\right).$$

Property 3c allows us to see that

$$\det A = \det\left(v_1 \middle| \cdots \middle| v_{j-1} \middle| 0 \middle| v_{j+1} \middle| \cdots \middle| v_n\right)$$

and property 3a gives $\det A = 0$. $\square$

Lemma 24.5. *Suppose property 2.(a) and 5. Then we can prove half of property* 6 *quickly, that if* $A \in M_{n\times n}(\mathbb{F})$ *is invertible then* $\det A \neq 0$.

Remark 24.6. Unless we allow property 5, the proof of that if $A \in M_{n\times n}(\mathbb{F})$ is invertible, then $\det A \neq 0$ is more annoying. When $\mathbb{F} = \mathbb{R}$, a good reason to believe this is true is that the det is connected with volume. The argument in the previous proof says a matrix

$$A = \left(v_1 \middle| \cdots \middle| v_n\right)$$

is invertible if and only if $(v_1, \ldots, v_n)$ is a basis for $\mathbb{R}^n$. Drawing a higher dimensional parallelogram using these vectors, it has nonzero volume exactly when $(v_1, \ldots, v_n)$ is a basis for $\mathbb{R}^n$.

25 February 28: Eigenvalues and eigenvectors

Definition 25.1. Suppose V is a vector space over a field $\mathbb{F}$ and $T : V \longrightarrow V$ is a linear transformation. A vector $v \in V$ is said to be *an eigenvector of* T **iff** the following conditions hold:

1. $v \neq 0$;
2. $Tv = \lambda v$ for some $\lambda \in \mathbb{F}$.

If $v \in V$ is an eigenvector of T and $Tv = \lambda v$, then λ is called *the eigenvalue corresponding to* v.
A scalar $\lambda \in \mathbb{F}$ is said to be an *eigenvalue of* T **iff** there is an eigenvector v of T such that λ is the eigenvalue corresponding to v.

Definition 25.2. Suppose V and W are vector spaces over a field $\mathbb{F}$, that $S : V \longrightarrow W$ and $T : V \longrightarrow W$ are linear transformations and $\lambda \in \mathbb{F}$. Then $S + T : V \longrightarrow W$ and $\lambda T : V \longrightarrow W$ are defined by

$$(S + T)(v) := S(v) + T(v) \quad \text{and} \quad (\lambda T)(v) := \lambda(T(v)).$$

Theorem 25.3. *Suppose V and W are vector spaces over a field $\mathbb{F}$, that $S : V \longrightarrow W$ and $T : V \longrightarrow W$ are linear transformations and $\lambda \in \mathbb{F}$. Then $S + T : V \longrightarrow W$ and $\lambda T : V \longrightarrow W$ are linear transformations.*
If β_V and β_W are bases for V and W, respectively, then

$$[S + T]_{\beta_V}^{\beta_W} = [S]_{\beta_V}^{\beta_W} + [T]_{\beta_V}^{\beta_W} \quad \text{and} \quad [\lambda T]_{\beta_V}^{\beta_W} = \lambda [T]_{\beta_V}^{\beta_W}.$$

Theorem 25.4. *Suppose V is a vector space over a field $\mathbb{F}$, $T : V \longrightarrow V$ is a linear transformation, $v \in V$, and $\lambda \in \mathbb{F}$. v is an eigenvector of T with corresponding eigenvalue λ if and only if*

$$v \in \ker(T - \lambda 1_V) \setminus \{0\}.$$

Proof. Suppose V is a vector space over a field $\mathbb{F}$, $T : V \longrightarrow V$ is a linear transformation, $v \in V$, and $\lambda \in \mathbb{F}$.
First, suppose v is an eigenvector of T with corresponding eigenvalue λ. This means $v \neq 0$, and $Tv = \lambda v$. The second equation gives

$$(T - \lambda 1_V)(v) = T(v) - \lambda 1_V(v) = T(v) - \lambda v = 0.$$

So $v \in \ker(T - \lambda 1_V) \setminus \{0\}$.
Suppose $v \in \ker(T - \lambda 1_V) \setminus \{0\}$. Then $v \neq 0$ and $(T - \lambda 1_V)(v) = 0$. The second equation gives $Tv = \lambda v$, so v is an eigenvector of T with corresponding eigenvalue λ. □

Corollary 25.5. *Suppose V is a vector space over a field $\mathbb{F}$, $T : V \longrightarrow V$ is a linear transformation, and $\lambda \in \mathbb{F}$. λ is an eigenvalue of T if and only if* $\ker(T - \lambda 1_V) \neq \{0\}$.

Definition 25.6. Suppose V is a vector space over a field $\mathbb{F}$, $T : V \longrightarrow V$ is a linear transformation, and $\lambda \in \mathbb{F}$. $E_\lambda := \ker(T - \lambda 1_V)$ is called the *eigenspace of T corresponding to λ*.

Remark 25.7. We normally only talk about eigenspaces E_λ when λ is an eigenvalue of T.

Theorem 25.8. *Suppose V is a vector space over a field $\mathbb{F}$, $T : V \longrightarrow V$ is a linear transformation, $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ are distinct eigenvalues of T, and $v_1, \ldots, v_n \in V$ are eigenvectors of T such that for each $j \in \{1, \ldots, n\}$, λ_j is the eigenvalue corresponding to v_j.*

Then $(v_1, \ldots, v_n)$ is linearly independent.

Proof. We prove the result by mathematical induction on n.

When $n = 1$, the result is true since, eigenvectors are non-zero, and (v) is linearly independent as long as $v \neq 0$.

Suppose $n \in \mathbb{N}$, and that the result is true for this n. Now suppose V is a vector space over a field $\mathbb{F}$, $T : V \longrightarrow V$ is a linear transformation, $\lambda_1, \ldots, \lambda_n, \lambda_{n+1} \in \mathbb{F}$ are distinct eigenvalues of T, and $v_1, \ldots, v_n, v_{n+1} \in V$ are eigenvectors of T such that for each $j \in \{1, \ldots, n, n+1\}$, λ_j is the eigenvalue corresponding to v_j. We wish to show that $(v_1, \ldots, v_n, v_{n+1})$ is linearly independent. Suppose $\mu_1, \ldots, \mu_n, \mu_{n+1} \in \mathbb{F}$ and that

$$\mu_1 v_1 + \ldots + \mu_n v_n + \mu_{n+1} v_{n+1} = 0.$$

Applying $(T - \lambda_{n+1} 1_V)$ to this equation gives

$$\mu_1(\lambda_1 - \lambda_{n+1})v_1 + \ldots + \mu_n(\lambda_n - \lambda_{n+1})v_n = 0.$$

By the inductive hypothesis, $(v_1, \ldots, v_n)$ is linearly independent, so we obtain

$$\mu_1(\lambda_1 - \lambda_{n+1}) = \ldots = \mu_n(\lambda_n - \lambda_{n+1}) = 0.$$

Since, the λ_j's are distinct, this gives

$$\mu_1 = \ldots = \mu_n = 0.$$

Going back to the original equation, we have $\mu_{n+1} v_{n+1} = 0$. Since v_{n+1} is an eigenvector, it is non-zero, so $\mu_{n+1} = 0$. We have shown all the μ_j's are zero. Thus, $(v_1, \ldots, v_n, v_{n+1})$ is linearly independent and we have completed the proof of the inductive step. $\square$

Lemma 25.9. *Suppose V is a vector space over a field $\mathbb{F}$, $T : V \longrightarrow V$ is a linear transformation, $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ are distinct scalars, and $v_1, \ldots, v_n \in V$ have the property that for all $j \in \{1, \ldots, n\}$, $v_j \in E_{\lambda_j}$. If*

$$v_1 + \ldots + v_n = 0,$$

then for all $j \in \{1, \ldots, n\}$, $v_j = 0$.

Proof. Suppose V is a vector space over a field $\mathbb{F}$, $T : V \to V$ is a linear transformation, $\lambda_1, \ldots, \lambda_n$ are distinct scalars, $v_1, \ldots, v_n \in V$ have the property that for all $j \in \{1, \ldots, n\}$, $v_j \in E_{\lambda_j}$, and

$$v_1 + \ldots + v_n = 0.$$

Say that there are m non-zero v_j's. Suppose for contradiction that $m \neq 0$. Then $m \in \{1, \ldots, n\}$, and by reordering the v_j's, we can assume that for $j \in \{1, \ldots, m\}$, $v_j \neq 0$, and for $j \in \{m+1, \ldots, n\}$, $v_j = 0$. Now, $v_1, \ldots, v_m$ are eigenvectors for T corresponding to distict eigenvalues, so the previous theorem tells us $(v_1, \ldots, v_m)$ is linearly independent. However, $1 \cdot v_1 + \ldots + 1 \cdot v_m = 0$. This is the required contradiction. $\square$

26 March 2: Diagonalizability

Theorem 26.1. *Suppose V is a vector space over a field $\mathbb{F}$, $T : V \longrightarrow V$ is a linear transformation, and that $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ are distinct eigenvalues of T. Suppose, also, that for each $j \in \{1, \ldots, n\}$, $(v_{j,1}, \ldots, v_{j,r_j})$ is a linearly independent tuple in the eigenspace E_{λ_j}. Then*

$$(v_{1,1}, \ldots, v_{1,r_1}, v_{2,1}, \ldots, v_{2,r_2}, \ldots\ldots, v_{n,1}, \ldots, v_{n,r_n})$$

is linearly independent.

Proof. Suppose we have scalars $\mu_{1,1}, \ldots, \mu_{1,r_1}, \mu_{2,1}, \ldots, \mu_{2,r_2}, \ldots\ldots, \mu_{n,1}, \ldots, \mu_{n,r_n} \in \mathbb{F}$ such that

$$\sum_{j=1}^{n} \sum_{k=1}^{r_j} \mu_{j,k} v_{j,k} = 0.$$

The previous lemma tells us that for each $j \in \{1, \ldots, n\}$,

$$\sum_{k=1}^{r_j} \mu_{j,k} v_{j,k} = 0.$$

Since $(v_{j,1}, \ldots, v_{j,r_j})$ is a linearly independent for each $j \in \{1, \ldots, n\}$, this tells us all the μ's are zero. $\square$

Definition 26.2. Suppose V is a finite-dimensional vector space over a field $\mathbb{F}$ and $T : V \longrightarrow V$ is a linear transformation. T is said to be *diagonalizable* **iff** V has a basis consisting of eigenvectors of T, i.e. there exist eigenvectors $v_1, \ldots, v_n \in V$ such that $(v_1, \ldots, v_n)$ is a basis.

This is because, if $\beta = (v_1, \ldots, v_n)$ is such a basis, then $[T]_\beta^\beta$ is a diagonal matrix.

Theorem 26.3. *Suppose V is a finite-dimensional vector space over a field $\mathbb{F}$ and $T : V \longrightarrow V$ is a linear transformation. If T has $\dim V$ distinct eigenvalues, then T is diagonalizable.*

Proof. Suppose V is a finite-dimensional vector space over a field $\mathbb{F}$ and $T : V \longrightarrow V$ is a linear transformation. Let $n = \dim V$, and suppose that $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ are distinct eigenvalues for T. We can choose $v_1, \ldots, v_n \in V$ such that for $j \in \{1, \ldots, n\}$, λ_j is the eigenvalue corresponding to v_j.

Because the λ_j's are distinct, theorem 25.8 tells us that $(v_1, \ldots, v_n)$ is linearly independent. Thus, $(v_1, \ldots, v_n)$ is a basis for V consisting of eigenvectors. $\square$

The previous theorem is great, but how do we find eigenvalues? The key is corollary 25.5.

Definition 26.4. Suppose V is a finite-dimensional vector space over a field $\mathbb{F}$ and $T : V \longrightarrow V$ is a linear transformation. The *determinant of* T is defined by the following equation

$$\det(T) = \det([T]_{\beta_V}^{\beta_V}).$$

Here, β_V is any basis of V.

Remark 26.5. Suppose V is a finite-dimensional vector space over a field $\mathbb{F}$, $T : V \longrightarrow V$ is a linear transformation, β_V and β'_V are bases for V, and P is the change of coordinate matrix from β'_V to β_V. Then

$$\det([T]_{\beta'_V}^{\beta'_V}) = \det(P^{-1}[T]_{\beta_V}^{\beta_V} P) = \det(P^{-1}P[T]_{\beta_V}^{\beta_V}) = \det([T]_{\beta_V}^{\beta_V}).$$

Thus, the previous definition makes sense.

Definition 26.6. Suppose V is a finite-dimensional vector space over a field $\mathbb{F}$ and $T : V \longrightarrow V$ is a linear transformation. The *characteristic polynomial of* T is defined by the following equation

$$c_T(x) = \det(T - x1_V).$$

Theorem 26.7. *Suppose V is a finite-dimensional vector space over a field $\mathbb{F}$, that $T : V \longrightarrow V$ is a linear transformation, and that $\lambda \in \mathbb{F}$. λ is an eigenvalue of T if and only if λ is a root of the characteristic polynomial of T.*

Proof. Suppose V is a finite-dimensional vector space over a field $\mathbb{F}$, that $T : V \longrightarrow V$ is a linear transformation, and that $\lambda \in \mathbb{F}$. Suppose β_V is a basis for V. Then

$$\begin{aligned}
& \lambda \text{ is an eigenvalue of } T \\
\iff & \ker(T - \lambda 1_V) \neq \{0\} \\
\iff & T - \lambda 1_V \text{ is not an isomorphism (since the domain and codomain have the same dimension)} \\
\iff & [T - \lambda 1_V]_{\beta_V}^{\beta_V} \text{ is not invertible} \\
\iff & \det([T - \lambda 1_V]_{\beta_V}^{\beta_V}) = 0 \\
\iff & \det(T - \lambda 1_V) = 0 \iff c_T(\lambda) = 0,
\end{aligned}$$

i.e. λ is a root of $c_T(x)$. □

Example 26.8. Consider $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ defined by

$$T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Then

$$c_T(x) = \det(T - x1_{\mathbb{R}^2}) = \det([T - x1_{\mathbb{R}^2}]_{(e_1,e_2)}^{(e_1,e_2)}) = \det\left(\begin{pmatrix} 1-x & 1 \\ 0 & 1-x \end{pmatrix}\right) = (x-1)^2.$$

So the only eigenvalue of T is 1.

Suppose for contradiction that T is diagonalizable. Then there is a basis (v_1, v_2) of $\mathbb{R}^2$ consisting of eigenvectors for T. Since 1 is the only eigenvalue of T, we have $v_1, v_2 \in E_1 = \ker(T - 1_V)$. Thus, $\ker(T - 1_V) = \mathbb{R}^2$, which means $T - 1_V = 0_V$, and so $T = 1_V$. This is the required contradiction.

We could also have obtained a contradiction by calculating

$$E_1 = \ker(T - 1_V) = \left\{ \begin{pmatrix} x_1 \\ 0 \end{pmatrix} : \ x_1 \in \mathbb{R} \right\}.$$

The reason this result occurs is that $1 \neq 2$: here, 1 is the dimension of E_1, and 2 is the power of $(x-1)$ in $c_T(x)$. This motivates the next definition.

Definition 26.9. Suppose V is a finite-dimensional vector space over a field $\mathbb{F}$, $T : V \longrightarrow V$ is a linear transformation, and λ is an eigenvalue of T.

The *geometric multiplicity of* λ is

$$\dim E_\lambda = \operatorname{null}(T - \lambda 1_V).$$

The *algebraic multiplicity of* λ is

$$\max\{k \in \mathbb{N} : \ (x-\lambda)^k \text{ is a factor of } c_T(x)\}.$$

27 March 5: Diagonalizability

Definition 27.1. A polynomial $p(x) \in \mathcal{P}(\mathbb{F})$ *splits over* $\mathbb{F}$ **iff** there are scalars $c, \sigma_1, \sigma_2, \ldots, \sigma_n$ such that

$$p(x) = c(x - \sigma_1)(x - \sigma_2) \cdots (x - \sigma_n).$$

Theorem 27.2. *Suppose V is a finite-dimensional vector space over a field $\mathbb{F}$, $T : V \longrightarrow V$ is a linear transformation, and T is diagonalizable. Then $c_T(x)$ splits over $\mathbb{F}$.*

Proof. Suppose V is a vector space over a field $\mathbb{F}$, $T : V \longrightarrow V$ is a linear transformation, and T is diagonalizable.

Let $\beta_V = (v_1, \ldots, v_n)$ be a basis of eigenvectors of T. Let $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ be the eigenvalues corresponding to $v_1, \ldots, v_n$. Then

$$\begin{aligned} c_T(x) = \det(T - x1_V) &= \det([T - x1_V]_{\beta_V}^{\beta_V}) \\ &= \det([T]_{\beta_V}^{\beta_V} - x[1_V]_{\beta_V}^{\beta_V}) \\ &= \det(\operatorname{diag}(\lambda_1, \ldots, \lambda_n) - xI_n) \\ &= \det(\operatorname{diag}(\lambda_1 - x, \ldots, \lambda_n - x) = (-1)^n(x - \lambda_1) \cdots (x - \lambda_n). \end{aligned}$$

□

Theorem 27.3. *Suppose V is a finite-dimensional vector space over a field $\mathbb{F}$, $T : V \longrightarrow V$ is a linear transformation, and λ is an eigenvalue of T. Let G be the geometric multiplicity of λ and A be the algebraic multiplicity of λ. Then $1 \leq G \leq A$.*

Proof. Let $(v_1, \ldots, v_G)$ be a basis for E_λ. Extend it to a basis $\beta_V = (v_1, \ldots, v_n)$ for V. We have $[T]_{\beta_V}^{\beta_V}$ equal to

$$\begin{pmatrix} \lambda I_G & B \\ 0 & C \end{pmatrix}$$

for some $B \in M_{G \times (n-G)}$ and $C \in M(n - G) \times (n - G)$. Thus,

$$\begin{aligned} c_T(x) = \det([T]_{\beta_V}^{\beta_V} - xI_n) &= \det \begin{pmatrix} (\lambda - x)I_G & B \\ 0 & C - xI_{n-G} \end{pmatrix} \\ &= \det((\lambda - x)I_G) \det(C - xI_{n-G}) = (x - \lambda)^G p(x) \end{aligned}$$

where $p(x) = (-1)^G \det(C - xI_{n-G})$. Thus, $A \geq G$.

Since λ is an eigenvalue of T, $E_\lambda \neq \{0\}$, so $G = \dim E_\lambda \geq 1$. □

Theorem 27.4. *Suppose V is a finite-dimensional vector space over a field $\mathbb{F}$ and $T : V \longrightarrow V$ is a linear transformation. T is diagonalizable if and only if*

- *the characteristic polynomial of T splits; and*
- *for each eigenvalue λ of T, its geometric multiplicity is equal to its algebraic multiplicity.*

Proof. Suppose V is a finite-dimensional vector space over a field $\mathbb{F}$ and $T : V \longrightarrow V$ is a linear transformation. Let $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ be the distinct eigenvalues of T.

Suppose T is diagonalizable. Let

$$(v_{1,1}, \dots, v_{1,r_1}, v_{2,1}, \dots, v_{2,r_2}, \dots\dots, v_{n,1}, \dots, v_{n,r_n})$$

be a basis of T consisting of eigenvectors such that λ_j is the eigenvalue corresponding to $v_{j,1}, \dots, v_{j,r_j}$.

We know that $c_T(x)$ splits as $\pm(x-\lambda_1)^{r_1}(x-\lambda_2)^{r_2}\cdots(x-\lambda_n)^{r_n}$ and so the algebraic multiplicity of λ_j is r_j. We know from the previous theorem that $\dim E_{\lambda_j} \leq r_j$.

Since $(v_{j,1}, \dots, v_{j,r_j})$ is linearly independent and consists of vectors in E_{λ_j}, we also know

$$\dim E_{\lambda_j} \geq r_j.$$

Thus, the algebraic multiplicity and geometric multiplicity of λ_j coincides for all $j \in \{1, \dots, n\}$.

Conversely, suppose that $c_T(x)$ splits and that the algebraic multiplicity and geometric multiplicity of λ_j coincides for all $j \in \{1, \dots, n\}$. Let $j \in \{1, \dots, n\}$ and

$$(v_{j,1}, \dots, v_{j,r_j})$$

be a basis of E_{λ_j}. Since the algebraic multiplicity and geometric multiplicity of λ_j coincide, r_j is the algebraic multiplicity of λ_j. Because the characteristic polynomial of T splits, we have

$$\sum_{j=1}^{n} r_j = \dim V.$$

By theorem 26.1 $(v_{1,1}, \dots, v_{1,r_1}, v_{2,1}, \dots, v_{2,r_2}, \dots\dots, v_{n,1}, \dots, v_{n,r_n})$ is linearly independent. Since it is a $(\dim V)$-tuple, it is a basis for V, and it consists of eigenvectors for T. $\square$

Example 27.5. $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ defined by

$$T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

is not diagonalizable over $\mathbb{R}$ since $c_T(x) = \det \begin{pmatrix} x & -1 \\ 1 & x \end{pmatrix} = x^2 + 1$ does not split over $\mathbb{R}$.

Example 27.6. $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ defined by

$$T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

is not diagonalizable since the geometric multiplicity of 1 is 1, but the algebraic multiplicity of 1 is 2.

28 March 7: Inner Products

Definition 28.1. Suppose V is a vector space over $\mathbb{R}$. An *inner product* on V is a function

$$\langle\ ,\ \rangle : V \times V \longrightarrow \mathbb{R}$$

with the following properties:

1. $\forall \lambda \in \mathbb{R}$, $\forall u_1 \in V$, $\forall u_2 \in V$, $\forall v \in V$,
$$\langle \lambda u_1 + u_2, v\rangle = \lambda\langle u_1, v\rangle + \langle u_2, v\rangle.$$
2. $\forall u \in V$, $\forall v \in V$, $\langle u, v\rangle = \langle v, u\rangle$.
3. $\forall v \in V$, $v \neq 0 \implies \langle v, v\rangle > 0$.

Definition 28.2. A *real inner product space* is a vector space V over $\mathbb{R}$, together with an inner product $\langle\ ,\ \rangle$ on V.

Example 28.3. Let $n \in \mathbb{N}$. $\mathbb{R}^n$ has an inner product, *the standard inner product* defined by

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{i=1}^{n} x_i y_i.$$

Example 28.4. Let $C([0,1])$ be $\{f : [0,1] \longrightarrow \mathbb{R} :\ f \text{ is continuous}\}$. In 131A, you proof that $C([0,1])$ is a vector space over $\mathbb{R}$, and we have an inner product defined by

$$\langle f, g\rangle = \int_0^1 f(x)g(x)\ dx.$$

Definition 28.5. Suppose V is a vector space over $\mathbb{C}$. An *inner product* on V is a function

$$\langle\ ,\ \rangle : V \times V \longrightarrow \mathbb{C}$$

with the following properties:

1. $\forall \lambda \in \mathbb{R}$, $\forall u_1 \in V$, $\forall u_2 \in V$, $\forall v \in V$,
$$\langle \lambda u_1 + u_2, v\rangle = \lambda\langle u_1, v\rangle + \langle u_2, v\rangle.$$
2. $\forall u \in V$, $\forall v \in V$, $\langle u, v\rangle = \overline{\langle v, u\rangle}$.
3. $\forall v \in V$, $v \neq 0 \implies \langle v, v\rangle > 0$.

Definition 28.6. A *complex inner product space* is a vector space V over $\mathbb{C}$, together with an inner product $\langle\ ,\ \rangle$ on V.

Example 28.7. Let $n \in \mathbb{N}$. $\mathbb{C}^n$ has an inner product, *the standard inner product* defined by

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{i=1}^{n} x_i \overline{y_i}.$$

Theorem 28.8. *Suppose V is an inner product space over $\mathbb{R}$ (or $\mathbb{C}$). Then*

1. $\langle \ , \ \rangle$ *is* (conjugate) linear *in the second variable.*

2. *For all $v \in V$, $\langle 0, v\rangle = \langle v, 0\rangle = 0$.*

3. *For all $v \in V$, $v = 0$ if and only if $\langle v, v\rangle = 0$.*

4. *Suppose $v_1, v_2 \in V$. If $\langle u, v_1\rangle = \langle u, v_2\rangle$ for all $u \in V$, then $v_1 = v_2$.*

Definition 28.9. Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$. For $v \in V$, we define the *norm* of v by $\|v\| = \sqrt{\langle v, v\rangle}$.

Theorem 28.10. *Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$. Let $\lambda \in \mathbb{F}$, $u, v \in V$. Then*

1. $\|v\| \geq 0$*, with equality if and only if $v = 0$.*

2. $\|\lambda v\| = |\lambda|\|v\|$.

3. $\|u + v\| \leq \|u\| + \|v\|$ *(triangle inequality).*

4. $|\langle u, v\rangle| \leq \|u\| \cdot \|v\|$ *(Cauchy-Schwartz).*

Proof. We'll just do 3. and 4 in the case that $v \neq 0$.

Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$ and $\lambda \in \mathbb{F}$, $u, v \in V$ with $v \neq 0$. Note that

$$\begin{aligned} 0 \leq \|u + \lambda v\|^2 = \langle u + \lambda v, u + \lambda v\rangle &= \langle u, u\rangle + \overline{\lambda}\langle u, v\rangle + \lambda\langle v, u\rangle + \lambda\overline{\lambda}\langle v, v\rangle \\ &= \|u\|^2 + \overline{\lambda}\langle u, v\rangle + \lambda\overline{\langle u, v\rangle} + |\lambda|^2\|v\|^2. \end{aligned}$$

Since λ was arbitrary, we can take $\lambda = -\frac{\langle u,v\rangle}{\|v\|^2}$ to see that $0 \leq \|u\|^2 - \frac{|\langle u,v\rangle|^2}{\|v\|^2}$, from which we obtain Cauchy-Schwarz.

Moreover, we have

$$\begin{aligned} \|u + v\|^2 = \langle u + v, u + v\rangle &= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\ &= \|u\|^2 + \langle u, v\rangle + \overline{\langle u, v\rangle} + \|v\|^2 \\ &\leq \|u\|^2 + 2|\langle u, v\rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2, \end{aligned}$$

so $\|u + v\| \leq \|u\| + \|v\|$. □

Remark 28.11. If one draws the correct picture, then the triangle inequality expresses the fact that the length of one side of a triangle is less than or equal to the sum of the lengths of the other two sides of that triangle.

If one draws the correct picture, then Cauchy-Schwartz expresses the fact that the hypothenuse of a right-angled triangle is the longest side.

Definition 28.12. Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$, and let $u, v \in V$. We say that u, v are *orthogonal* **iff** $\langle u, v\rangle = 0$. v is said to be a *unit vector* **iff** $\|v\| = 1$.

Definition 28.13. Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$ and let $v_1, \dots, v_n \in V$. The tuple $(v_1, \dots, v_n)$ is said to be *orthogonal* **iff** for every $i, j \in \{1, \dots, n\}$ with $i \neq j$, v_i and v_j are orthogonal. We say that $(v_1, \dots, v_n)$ is *othonormal* **iff** each it is orthogonal, and each v_i is a unit vector.

Remark 28.14. Suppose V is an inner product space over $\mathbb{R}$ or $\mathbb{C}$. Then $(v_1, \dots, v_n)$ is *othonormal* if and only if

$$\langle v_i, v_j\rangle = \delta_{i,j}.$$

Example 28.15. Let $\mathbb{R}^n$ have the standard inner product. Then the standard basis is an orthonormal tuple.

Example 28.16. Let $\mathbb{C}^n$ have the standard inner product.

Let $z = e^{\frac{2\pi i}{n}}$, and for $k \in \mathbb{Z}$, let

$$v_k = \begin{pmatrix} z^k \\ z^{2k} \\ \vdots \\ z^{(n-1)k} \\ z^{nk} \end{pmatrix} \in \mathbb{C}^n.$$

Let $k, l \in \mathbb{Z}$. Then

$$\langle v_k, v_l\rangle = \sum_{i=1}^{n} z^{ik}\overline{z^{il}} = \sum_{i=1}^{n} z^{ik} z^{-il} = \sum_{i=1}^{n} z^{i(k-l)}.$$

If $k - l$ is divisible by n then this number is n. It's an exercise to show, that otherwise, the answer is 0. Your proof will break into cases, depending on what highest common factor of $k - l$ and n is.

Thus, $(v_1, v_2, \dots, v_n)$ is orthogonal.

29 March 9: Gram-Schmidt

Theorem 29.1. *Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$, and $w_1, \ldots, w_n \in V$. Suppose that $(w_1, \ldots, w_n)$ is orthogonal and consists of non-zero vectors, $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$, and*

$$v = \sum_{j=1}^{n} \lambda_j w_j.$$

Then for each $i \in \{1, \ldots, n\}$, $\lambda_i = \frac{\langle v, w_i \rangle}{\|w_i\|^2}$.

Proof. Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$ and $w_1, \ldots, w_n \in V$. Suppose that $(w_1, \ldots, w_n)$ is orthogonal and consists of non-zero vectors, $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$, and

$$v = \sum_{j=1}^{n} \lambda_j w_j.$$

Then

$$\langle v, w_i \rangle = \left\langle \sum_{j=1}^{n} \lambda_j w_j, w_i \right\rangle = \sum_{j=1}^{n} \lambda_j \langle w_j, w_i \rangle = \lambda_i \langle w_i, w_i \rangle = \lambda_i \|w_i\|^2.$$

Since $w_i \neq 0$, this gives $\lambda_i = \frac{\langle v, w_i \rangle}{\|w_i\|^2}$. □

Corollary 29.2. *Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$, and $w_1, \ldots, w_n \in V$. If $(w_1, \ldots, w_n)$ is orthogonal and consists of non-zero vectors, then $(w_1, \ldots, w_n)$ is linearly independent.*

Corollary 29.3. *Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$, and $w_1, \ldots, w_n \in V$. Suppose that $(w_1, \ldots, w_n)$ is orthonormal, $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$, and*

$$v = \sum_{j=1}^{n} \lambda_j w_j.$$

Then for each $i \in \{1, \ldots, n\}$, $\lambda_i = \langle v, w_i \rangle$.

Corollary 29.4. *Suppose V is an inner product space, $\beta = (w_1, \ldots, w_n)$ is an orthonormal basis of V, and $v \in V$. Then*

$$[v]_\beta = \Big(\langle v, w_1 \rangle, \ldots, \langle v, w_n \rangle \Big).$$

Theorem 29.5 (Gram-Schmidt)**.** *Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$, let $v_1, \ldots, v_n \in V$, and suppose $(v_1, \ldots, v_n)$ is linearly independent. Define new vectors $w_1, \ldots, w_n \in V$ by:*

- *$w_1 = v_1$;*
- *for $k \in \{2, \ldots, n\}$, $w_k = v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, w_j \rangle}{\|w_j\|^2} w_j$.*

These formulae make sense, and the vectors $w_1, \ldots, w_n$ have the following two properties:

- *$(w_1, \ldots, w_n)$ is orthogonal and consists of non-zero vectors;*
- *for all $k \in \{1, \ldots, n\}$, $span(\{w_1, \ldots, w_k\}) = span(\{v_1, \ldots, v_k\})$.*

Proof. Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$, let $v_1, \ldots, v_n \in V$, and suppose $(v_1, \ldots, v_n)$ is linearly independent. We prove, by induction on k, that:

- the formula defining w_k makes sense;
- $(w_1, \ldots, w_k)$ is orthogonal and consists of non-zero vectors;
- $\operatorname{span}(\{w_1, \ldots, w_k\}) = \operatorname{span}(\{v_1, \ldots, v_k\})$.

First, we address the base case, that's when $k = 1$. The formula defining w_1 makes sense and (w_1) is orthogonal. Since $(v_1, \ldots, v_n)$ is linearly independent, v_1 is non-zero. Thus, since $w_1 = v_1$, (w_1) consists of non-zero vectors. Moreover, $\operatorname{span}(\{w_1\}) = \operatorname{span}(\{v_1\})$.

Suppose the properties above hold for $k \in \{1, \ldots, n-1\}$. We show the properties for $k+1$. Since $(w_1, \ldots, w_k)$ consists of non-zero vectors the formula

$$w_{k+1} = v_{k+1} - \sum_{j=1}^{k} \frac{\langle v_{k+1}, w_j \rangle}{\|w_j\|^2} w_j$$

makes sense.

We know $(w_1, \ldots, w_k)$ are orthogonal. To show $(w_1, \ldots, w_k, w_{k+1})$ is orthogonal, we just have to show that for all $i \in \{1, \ldots, k\}$, $\langle w_{k+1}, w_i \rangle = 0$. Let $i \in \{1, \ldots, k\}$. Then

$$\begin{aligned}
\langle w_{k+1}, w_i \rangle &= \left\langle v_{k+1} - \sum_{j=1}^{k} \frac{\langle v_{k+1}, w_j \rangle}{\|w_j\|^2} w_j, w_i \right\rangle \\
&= \langle v_{k+1}, w_i \rangle - \sum_{j=1}^{k} \frac{\langle v_{k+1}, w_j \rangle}{\|w_j\|^2} \langle w_j, w_i \rangle \\
&= \langle v_{k+1}, w_i \rangle - \frac{\langle v_{k+1}, w_i \rangle}{\|w_i\|^2} \langle w_i, w_i \rangle \\
&= \langle v_{k+1}, w_i \rangle - \langle v_{k+1}, w_i \rangle \\
&= 0.
\end{aligned}$$

We know $(w_1, \ldots, w_k)$ consists of non-zero vectors. To show $(w_1, \ldots, w_k, w_{k+1})$ consists of non-zero vectors, we just have to show that $w_{k+1} \neq 0$. Suppose for contradiction that $w_{k+1} = 0$. This is the same as saying

$$v_{k+1} = \sum_{j=1}^{k} \frac{\langle v_{k+1}, w_j \rangle}{\|w_j\|^2} w_j.$$

Thus, $v_{k+1} \in \operatorname{span}(\{w_1, \ldots, w_k\}) = \operatorname{span}(\{v_1, \ldots, v_k\})$ and this contradicts the fact $(v_1, \ldots, v_n)$ is linearly independent.

The definition of w_{k+1} gives

$$w_{k+1} \in \operatorname{span}(\{w_1, \ldots, w_k, v_{k+1}\}) = \operatorname{span}(\{v_1, \ldots, v_k, v_{k+1}\}).$$

So $\operatorname{span}(\{w_1, \ldots, w_k, w_{k+1}\}) \subseteq \operatorname{span}(\{v_1, \ldots, v_k, v_{k+1}\})$. Since $(w_1, \ldots, w_k, w_{k+1})$ is is orthogonal and consists of non-zero vectors, it is linearly independent. $(v_1, \ldots, v_k, v_{k+1})$ is linearly independent too, and so

$$\dim(\operatorname{span}(\{w_1, \ldots, w_k, w_{k+1}\}) = k+1 = \dim(\operatorname{span}(\{v_1, \ldots, v_k, v_{k+1}\}).$$

We conclude $\operatorname{span}(\{w_1, \ldots, w_k, w_{k+1}\}) = \operatorname{span}(\{v_1, \ldots, v_k, v_{k+1}\})$. $\square$

30 March 12: Orthogonal Complements, Riesz

Theorem 30.1. *Suppose V is a finite-dimensional inner product space and that $(w_1, \ldots, w_m)$ is orthnormal. Then we can extend $(w_1, \ldots, w_m)$ to an orthonormal basis $(w_1, \ldots, w_n)$ of V.*

Proof. Suppose V is a finite-dimensional inner product space and that $(w_1, \ldots, w_m)$ is orthonormal. Extend $(w_1, \ldots, w_m)$ to a basis of V:

$$(w_1, \ldots, w_m, v_{m+1}, \ldots, v_n).$$

The Gram-Schmidt process leaves $w_1, \ldots, w_m$ alone, so run Gram-Schmidt to obtain orthogonal $(w_1, \ldots, w_m, w_{m+1}, \ldots, w_n)$.

$$\left(w_1, \ldots, w_m, \frac{w_{m+1}}{\|w_{m+1}\|}, \ldots, \frac{w_n}{\|w_n\|}\right)$$

is an orthonormal basis for V: it is orthonormal, thus linearly independent, and it has the correct size to be a basis. □

Corollary 30.2. *Suppose V is a finite-dimensional inner product space. Then V has an orthonormal basis.*

Definition 30.3. Suppose V is an inner product space, and that S is a nonempty subset of V. Then the *orthogonal complement* of S is the set

$$S^\perp = \{v \in V : \forall s \in S,\ \langle v, s \rangle = 0\}.$$

$S^\perp$ is read as "S perp."

Remark 30.4. You can check $S^\perp$ is a subspace.

Theorem 30.5. *Suppose U is a finite-dimensional vector space, that V, W are subspaces of U, and that $U = V \oplus W$. Then $\dim(U) = \dim(V) + \dim(W)$.*

Proof. Let $(v_1, \ldots, v_m)$ be a basis for V, and $(w_1, \ldots, w_n)$ be a basis for W. We claim that

$$(v_1, \ldots, v_m, w_1, \ldots, w_n)$$

is a basis for U.

Suppose $\lambda_1 v_1 + \ldots + \lambda_m v_m + \mu_1 w_1 + \ldots + \mu_n w_n = 0$. Then

$$\lambda_1 v_1 + \ldots + \lambda_m v_m = -(\mu_1 w_1 + \ldots + \mu_n w_n) \in V \cap W = \{0\}.$$

Linear independence of $(v_1, \ldots, v_m)$ and $(w_1, \ldots, w_n)$ show that all the coefficients are 0. Thus, $(v_1, \ldots, v_m, w_1, \ldots, w_n)$ is linearly independent. So $\dim(U) \geq m + n = \dim(V) + \dim(W)$.

Showing the opposite inequality is the Shakespeare question in homework 5. □

Theorem 30.6. *Suppose V is a finite dimensional inner product space, and that W is a subspace of V. Then $V = W \oplus W^{\perp}$. Thus,* $\dim(V) = \dim(W) + \dim(W^{\perp})$.

Proof. First, note that if $w \in W \cap W^{\perp}$, then $\|w\|^2 = \langle w, w\rangle = 0$, so $w = 0$.

Let $m = \dim V$, and $n = \dim W$. Let $(w_1, \ldots, w_n)$ be an orthonormal basis for W. Extend it to an orthonormal basis of V:

$$(w_1, \ldots, w_n, w_{n+1}, \ldots, w_m).$$

$(w_{n+1}, \ldots, w_m)$ consists of vectors in $W^{\perp}$. Thus, given $v \in V$, we have

$$v = \sum_{i=1}^{m} \langle v, w_i\rangle w_i = \sum_{i=1}^{n} \langle v, w_i\rangle w_i \ + \sum_{i=n+1}^{m} \langle v, w_i\rangle w_i \ \in W + W^{\perp},$$

which completes the proof that $V = W \oplus W^{\perp}$.

Aside: notice that $(w_{n+1}, \ldots, w_m)$ is a basis of $W^{\perp}$. $\square$

Theorem 30.7. *Suppose V is a finite-dimensional inner product space over $\mathbb{F}$ and that $f : V \longrightarrow \mathbb{F}$ is a linear transformation. Then there exists a unique $r \in V$ such that $f(v) = \langle v, r\rangle$ for all $v \in V$.*

Proof. Suppose V is a finite-dimensional inner product space over $\mathbb{F}$ and that $f : V \to \mathbb{F}$ is a linear transformation.

When $f = 0$, we can take $r = 0$, so suppose $f \neq 0$. The rank-nullity theorem gives $\text{null}(f) = \dim(V) - 1$. Thus, $\dim(\ker(f)^{\perp}) = 1$. Pick any $s \in \ker(f)^{\perp}$ with $\|s\| = 1$, and let $r = \overline{f(s)}s$. Since $s \notin \ker(f)$, $f(s) \neq 0$, and so $\text{span}(\{r\}) = \text{span}(\{s\}) = \ker(f)^{\perp}$. We note:

- for $v \in \ker(f)$, $f(v) = 0 = \langle v, r\rangle$;
- when $v = r$, we have $v = \overline{f(s)}s$, so

$$f(v) = f(\overline{f(s)}s) = \overline{f(s)}f(s) = |f(s)|^2 = |f(s)|^2\|s\|^2 = \|\overline{f(s)}s\|^2 = \|v\|^2 = \langle v, v\rangle = \langle v, r\rangle.$$

From these two observations we conclude that $f(v) = \langle v, r\rangle$ for all $v \in \ker(f) + \text{span}(\{r\}) = V$.

Suppose $r' \in V$ and $f(v) = \langle v, r'\rangle$ for all $v \in V$. Then $\langle v, r\rangle = \langle v, r'\rangle$ for all $v \in V$, which gives $r = r'$ by theorem 28.8, 4. $\square$

31 March 14: Adjoints, self-adjoint operators

Theorem 31.1. *Suppose V is a finite-dimensional inner product space over $\mathbb{F}$ and that $T : V \longrightarrow V$ is a linear transformation. Then there exists a unique function $T^* : V \longrightarrow V$ such that $\langle T(v), v'\rangle = \langle v, T^*(v')\rangle$ for all $v, v' \in V$. Moreover, T^* is linear.*

Proof. For $v' \in V$, define $f_{v'} : V \longrightarrow \mathbb{F}$ by $f_{v'}(v) = \langle T(v), v'\rangle$. $f_{v'}$ is linear and so we obtain an $r_{v'}$ such that for all $v \in V$,

$$f_{v'}(v) = \langle v, r_{v'}\rangle, \text{ i.e. } \langle T(v), v'\rangle = \langle v, r_{v'}\rangle.$$

Define $T^* : V \longrightarrow V$ by $T^*(v') = r_{v'}$.

Say T^* is unique.

Show T^* is linear. $\square$

Definition 31.2. Suppose V is a finite-dimensional inner product space over $\mathbb{F}$ and that $T : V \longrightarrow V$ is a linear transformation. The unique function $T^* : V \longrightarrow V$ such that $\langle T(v), v' \rangle = \langle v, T^*(v') \rangle$ for all $v, v' \in V$ is called the *adjoint* of T.

Theorem 31.3. *Suppose V is a finite-dimensional inner product space over $\mathbb{F}$ and that $T : V \longrightarrow V$ is a linear transformation. For all $v, v' \in V$, $\langle v, T^*(v') \rangle = \langle T(v), v' \rangle$.*

Theorem 31.4. *Suppose V is a finite-dimensional inner product space over $\mathbb{F}$, that $S, T : V \longrightarrow V$ are linear transformations, and $\lambda \in \mathbb{F}$. Then*

1. $(S + T)^* = S^* + T^*$;
2. $(\lambda T)^* = \overline{\lambda} T^*$;
3. $(ST)^* = T^* S^*$;
4. $(T^*)^* = T$;
5. $1_V^* = 1_V$.

Theorem 31.5. *Suppose V is an inner product space, $\beta = (w_1, \ldots, w_n)$ is an orthonormal basis of V, $T : V \longrightarrow V$ is a linear transformation, and $[T]_\beta^\beta = (a_{ij})$. Then for $i, j \in \{1, \ldots, n\}$,*

$$a_{ij} = \langle T(w_j), w_i \rangle.$$

Proof. Suppose V is an inner product space, $\beta = (w_1, \ldots, w_n)$ is an orthonormal basis of V, that $T : V \longrightarrow V$ is a linear transformation, and $[T]_\beta^\beta = (a_{ij})$. Then for $j \in \{1, \ldots, n\}$,

$$T(w_j) = \sum_{k=1}^{n} a_{kj} w_k.$$

Thus, for $i, j \in \{1, \ldots, n\}$,

$$\langle T(w_j), w_i \rangle = \left\langle \sum_{k=1}^{n} a_{kj} w_k, w_i \right\rangle = \sum_{k=1}^{n} a_{kj} \langle w_k, w_i \rangle = a_{ij} \langle w_i, w_i \rangle = a_{ij}.$$

$\square$

Theorem 31.6. *Suppose V is an inner product space, $\beta = (w_1, \ldots, w_n)$ is an orthonormal basis of V, and $T : V \longrightarrow V$ is a linear transformation. Then if $[T]_\beta^\beta = (a_{ij})$ and $[T^*]_\beta^\beta = (b_{ij})$, we have $a_{ij} = \overline{b_{ji}}$.*

Proof. The previous theorem says that $a_{ij} = \langle T(w_j), w_i \rangle$. and $b_{ij} = \langle T^*(w_j), w_i \rangle$. Thus,

$$a_{ij} = \langle T(w_j), w_i \rangle = \langle w_j, T^*(w_i) \rangle = \overline{\langle T^*(w_i), w_j \rangle} = \overline{b_{ji}}.$$

$\square$

Definition 31.7. Suppose V is a finite-dimensional inner product space over $\mathbb{F}$ and that $T : V \longrightarrow V$ is a linear transformation. We say T is *self-adjoint* **iff** $T^* = T$.

Theorem 31.8. *Suppose V is a finite-dimensional (complex) inner product space, that $T : V \longrightarrow V$ is self-adjoint and that λ is an eigenvalue of T. Then $\lambda \in \mathbb{R}$.*

Proof. Let $v \in V$ be a unit eigenvector of T with eigenvalue λ. Then

$$\lambda = \lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T^*(v)\rangle = \langle v, T(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle = \overline{\lambda}.$$

□

Remark 31.9. The word complex is not necessary in the previous theorem, but when V is a real inner product space, the result is not very interesting!

32 March 16: The spectral theorem

Theorem 32.1. *Suppose V is a finite-dimensional inner product space, and that $T : V \longrightarrow V$ is self-adjoint. Then T has an eigenvector.*

Proof. Suppose V is a finite-dimensional inner product space, and that $T : V \longrightarrow V$ is self-adjoint.

1. If V is a vector space over $\mathbb{C}$, then $c_T(x)$ has a root (the fundamental theorem of algebra); this means that T has an eigenvalue, and there's a corresponding eigenvector.

   We do not need to use self-adjointness anywhere.

2. We are left with the case when V is a vector space over $\mathbb{R}$.

   Choose an orthonormal basis of V, $\beta = (w_1, \ldots, w_n)$, and let $A = [T]_\beta^\beta$.

   A is an $n \times n$ symmetric matrix with real entries because T is self-adjoint. Define

   $$S : \mathbb{C}^n \longrightarrow \mathbb{C}^n,\ x \longmapsto Ax.$$

   Notice that we have

   $$c_T(x) = \det(T - x1_V) = \det(A - xI_n) = \det(S - x1_{\mathbb{C}^n}) = c_S(x).$$

   To show that T has an eigenvector, it is enough to show that $c_T(x)$ has a real root. By the observation just made, it is enough to show that $c_S(x)$ has a real root, and to do this, it is enough to show that S has a real eigenvalue.

   Since $\mathbb{C}^n$ is vector space over $\mathbb{C}$, S has an eigenvalue $\lambda \in \mathbb{C}$. We claim that $\lambda \in \mathbb{R}$. By theorem 31.8 it is enough to show that S is self adjoint. This is true: for $x_1, x_2 \in \mathbb{C}^n$, we have

   $$\langle Sx_1, x_2\rangle = \langle Ax_1, x_2\rangle = (Ax_1)^T\overline{x_2} = x_1^T A^T \overline{x_2} = x_1^T \overline{Ax_2} = \langle x_1, Ax_2\rangle = \langle x_1, Sx_2\rangle.$$

□

Theorem 32.2. *Suppose V is a finite-dimensional inner product space, and that $T : V \longrightarrow V$ is self-adjoint. Then there exists an orthonormal basis $(v_1, \ldots, v_n)$ of V consisting of eigenvectors of T.*

Proof. Suppose V is a finite-dimensional inner product space, and that $T : V \longrightarrow V$ is self-adjoint. Let $n = \dim V$ and let v_n be an eigenvector of T. By normalizing, we can assume that $\|v_n\| = 1$.

Let $W = \operatorname{span}(\{v_n\})^\perp$ and attempt to define $S : W \longrightarrow W$, by $S(w) = T(w)$. In order to see this makes sense, we must show that whenever $w \in W$, we have $T(w) \in W$. So let $w \in W$. Writing λ_n for the eigenvalue corresponding to v_n, we have

$$\langle T(w), v_n\rangle = \langle w, T^*(v_n)\rangle = \langle w, T(v_n)\rangle = \langle w, \lambda_n v_n\rangle = \overline{\lambda_n}\langle w, v_n\rangle = 0.$$

This shows $T(w) \in W$. Thus, $S : W \longrightarrow W$ is well-defined.

S is self-adjoint because T is: given $w_1, w_2 \in W$, we have

$$\langle w_1, S^*(w_2)\rangle = \langle S(w_1), w_2\rangle = \langle T(w_1), w_2\rangle = \langle w_1, T^*(w_2)\rangle = \langle w_1, T(w_2)\rangle = \langle w_1, S(w_2)\rangle.$$

Moreover, $\dim W = \dim V - 1 = n - 1$. Thus, an inductive hypothesis tells us that there exists an orthonormal basis $(v_1, \ldots, v_{n-1})$ of W consisting of eigenvectors of S. $(v_1, \ldots, v_n)$ is an orthonormal basis of V consisting of eigenvectors of T. □

33 Exam expectations

Axiom checking for a vector space is fair game.

The results of theorem 6.1 should be familiar. I won't ask you to prove such silly things.

Similarly, I will not ask you to prove theorem 7.3. You can assume that the subsapce test works and use it to show things are or are not subspaces.

Results like theorem 8.2.8 should hopefully feel easy by now. I also hope that doing the exercises about spans on the homeworks directly feels easier by now. If not, get to work! It is reasonable if you're not completely happy with my proofs (which often make use of theorem 9.1.1), but if they also feel fine to you though, then your confidence should be growing as you read this.

Checking tuples are linearly (in)dependent, and assuming tuples are linearly (in)dependent and making arguments using such assumptions has made up a bulk of the class. I cannot imagine an exam without such a question.

You should definitely know the following results (except 31 and 32), but as for their proofs…

Examinable proofs: 10.1.1, 10.1.2, 10.2.3, 14.12, 15.3, 15.4, 15.7, 16.1, 16.2, 18.2 (without reference to the homework), 19.1, 22.1, 22.2, 22.3, 22.4, 25.4, 25.8, 26.3, 26.7, 27.2, 28.8, 29.1, 29.2, 29.3, 30.1, 30.5, 30.6,

Non-examinable proofs: 11.2, 12.1, 13.4, 13.6, 17.1, 20.2, 21.2, 22.8, 22.9, 22.10, 22.11, 23.4, 24.2, 25.3, 25.9, 26.1, 27.3, 27.4, 28.10, 29.5, 30.7, all of 31, all of 32.

Of course, you should know the how to find the matrix of a linear transformation with respect to two bases and how this plays with the coordinate isomorphisms $[\ ]_\beta$ (remark 20.8).

Homework 4 and 5 were my favoUrite. Homework 3 was alright, but questions 8 and 9 were probably a little over-the-top: don't worry about these - sorry! Homework 2 was reasonable other than the matrix calculations - I don't want to grade them, so they won't be on the final - they should help you in the future. Homework 1 seems like a previous life - I hope you see the connection between the column stuff and the matrix of a linear transformation now.

Recall $D_8 = \langle r, s : \ r^4 = s^2 = 1, \ srs = r^{-1}\rangle$. We have a diagram of SESs.

$$\begin{tikzcd}
1 \arrow[r] & \mathbb{Z}/2\langle r^2\rangle \arrow[r] \arrow[d, equal] & \mathbb{Z}/4\langle r\rangle \arrow[r] \arrow[d] & \mathbb{Z}/2\langle r\rangle \arrow[r] \arrow[d] & 1 \\
1 \arrow[r] & \mathbb{Z}/2\langle r^2\rangle \arrow[r] & D_8\langle r, s\rangle \arrow[r] & (\mathbb{Z}/2\times\mathbb{Z}/2)\langle r, s\rangle \arrow[r] & 1 \\
1 \arrow[r] & \mathbb{Z}/2\langle r^2\rangle \arrow[r] \arrow[u, equal] & (\mathbb{Z}/2\times\mathbb{Z}/2)\langle r^2, s\rangle \arrow[r] \arrow[u] & \mathbb{Z}/2\langle s\rangle \arrow[r] \arrow[u] & 1
\end{tikzcd}$$

Consider the $\mathbb{Z}$-SSS and $\mathbb{Z}/2$-SSS for the fibration $B\,\mathbb{Z}/2 \longrightarrow BD_8 \longrightarrow B(\mathbb{Z}/2\times\mathbb{Z}/2)$:

$$\begin{tikzcd}
H^*(B(\mathbb{Z}/2\times\mathbb{Z}/2); H^*(B\,\mathbb{Z}/2;\mathbb{Z})) \arrow[r, Rightarrow] \arrow[d] & H^*(BD_8;\mathbb{Z}) \arrow[d] \\
H^*(B(\mathbb{Z}/2\times\mathbb{Z}/2); H^*(B\,\mathbb{Z}/2;\mathbb{Z}/2)) \arrow[r, Rightarrow] & H^*(BD_8;\mathbb{Z}/2).
\end{tikzcd}$$

We describe the E_2-pages of these SSs as subquotients of $\mathbb{Z}[x_1, x_2, y]$.

Here, x_1, x_2, y correspond to r, s, and r^2, respectively.

$H^*(B(\mathbb{Z}/2\times\mathbb{Z}/2); H^*(B\,\mathbb{Z}/2;\mathbb{Z}/2)) = \mathbb{Z}[x_1, x_2, y]/2$ and $H^*(B(\mathbb{Z}/2\times\mathbb{Z}/2); H^*(B\,\mathbb{Z}/2;\mathbb{Z}))$ is

$$\mathbb{Z}[x_1^2, x_2^2, x_1^2x_2 + x_1x_2^2, y^2, x_1y^2, x_2y^2, x_1x_2y^2]/(2x_1, 2x_2, 2y).$$

The map of SSs on the E_2-page is indicated by the names we have given elements.

The $\mathbb{Z}/2$-SS is determined by the following differential (which is obtained using the SESs above):

$$d_2(y) = x_1^2 + x_1x_2 = x_1(x_1+x_2).$$

Let $X_1 = x_1$, $X_2 = x_1 + x_2$ and $Y = y$, so $d_2(Y) = X_1X_2$.

Then $d_2(X_1Y) = X_1^2X_2$, $d_2(X_2Y) = X_1X_2^2$, and $d_3(Y^2) = X_1^2X_2 + X_1X_2^2 = 0$.

Thus, $E_\infty = \mathbb{Z}/2[X_1, X_2, Y^2]/(X_1X_2)$,

and $H^*(BD_8;\mathbb{Z}/2) = \mathbb{Z}/2[A_1, B_1, D_2]/(A_1B_1)$, where $[A_1] = X_1$, $[B_1] = X_2$ and $[D_2] = Y^2$.

We have a BSS, $H^*(BD_8;\mathbb{Z}/2)[h_0] \implies H^*(BD_8;\mathbb{Z})$.

$d_1(A_1) = h_0A_1^2$, $d_1(B_1) = h_0B_1^2$,

$d_1(D_2) = h_0(A_1+B_1)D_2$, $d_2(A_1D_2) = d_2(B_1D_2) = h_0^2D_2^2$.

Thus, $E_\infty = \mathbb{Z}/2[A_1^2, B_1^2, (A_1+B_1)D_2, D_2^2, h_0]/(h_0A_1^2, h_0B_1^2, h_0(A_1+B_1)D_2, h_0^2D_2^2, A_1B_1)$.

and $H^*(BD_8;\mathbb{Z}) = \mathbb{Z}[A_2, B_2, C_3, D_4]/(2A_2, 2B_2, 2C_3, 4D_4, A_2B_2, C_3^2 + (A_2+B_2)D_4)$,

where $[A_2] = A_1^2$, $[B_2] = B_1^2$, $[C_3] = (A_1+B_1)D_2$, $[D_4] = D_2^2$.

Let $a_2 = A_2 + B_2$, $b_2 = B_2$, $c_3 = C_3$ and $D_4 = d_4$. Then

$H^*(BD_8;\mathbb{Z}) = \mathbb{Z}[a_2, b_2, c_3, d_4]/(2a_2, 2b_2, 2c_3, 4d_4, (a_2+b_2)b_2, c_3^2 + a_2d_4)$.

In summary. . .

$a_2 = A_2 + B_2 = A_1^2 + B_1^2 = X_1^2 + X_2^2 = x_2^2$ is the image of the generator of $H^2(B(\mathbb{Z}/2\langle s\rangle))$.

$b_2 = B_2 = B_1^2 = X_2^2 = x_1^2 + x_2^2$ is the image of the generator of $H^2(B(\mathbb{Z}/2\langle rs\rangle))$.

$c_3 = C_3 = (A_1+B_1)D_2 = (X_1+X_2)Y^2 = x_2y^2$ is strange!

$d_4 = D_4 = D_2^2 = Y^4 = y^4$ is a lift of the generator of $H^4(B(\mathbb{Z}/2\langle r^2\rangle))$.

In the $\mathbb{Z}$-SS, we have the following differentials:

$$\begin{aligned} d_3(y^2) &= x_1^2x_2 + x_1x_2^2, \\ d_3(x_1y^2) &= x_1^4 + x_1^2x_2^2, \\ d_3(x_2y^2) &= 0, \\ d_3(x_1x_2y^2) &= d_3(x_1^2y^2), \\ d_5(y^4) &= 0. \end{aligned}$$

E_∞ is a quotient of $\mathbb{Z}[x_1^2, x_2^2, y^4, x_2y^2, (x_1^2 + x_1x_2)y^2]$ and we have:

$[a_2] = x_2^2$,
$[b_2] = x_1^2 + x_2^2$,
$[c_3] = x_2y^2$,
$[d_4] = y^4$,
$[2d_4] = (x_1^2 + x_1x_2)y^2$.

What about D_{2m} in general?

For $n \geq 1$, we have $H^*(B\,\mathbb{Z}/2; H^*(B\,\mathbb{Z}/2^n; \mathbb{Z}/2)) \implies H^*(BD_{2^{n+1}}; \mathbb{Z}/2)$.

When $n = 1$, $H^*(B\,\mathbb{Z}/2; H^*(B\,\mathbb{Z}/2; \mathbb{Z}/2)) = \mathbb{Z}/2[b, f]$, and the SS is degenerate.

$$\begin{tikzcd} 1 \arrow[r] & \mathbb{Z}/2^{n-1} \arrow[r] \arrow[d] & D_{2^n} \arrow[r] \arrow[d] & \mathbb{Z}/2 \arrow[r] \arrow[d, "="] & 1 \\ 1 \arrow[r] & \mathbb{Z}/2^n \arrow[r] & D_{2^{n+1}} \arrow[r] & \mathbb{Z}/2 \arrow[r] & 1 \end{tikzcd}$$

For $n \geq 2$, $H^*(B\,\mathbb{Z}/2; H^*(B\,\mathbb{Z}/2^n; \mathbb{Z}/2)) = \mathbb{Z}/2[b, f_1, f_2]/(f_1^2)$, and the SS is degenerate.

$$\begin{tikzcd} 1 \arrow[r] & \mathbb{Z}/2\langle r^2\rangle \arrow[r] \arrow[d, equal] & \mathbb{Z}/4\langle r\rangle \arrow[r] \arrow[d] & \mathbb{Z}/2\langle r\rangle \arrow[r] \arrow[d] & 1 \\ 1 \arrow[r] & \mathbb{Z}/2\langle r^2\rangle \arrow[r] \arrow[d] & D_8\langle r, s\rangle \arrow[r] \arrow[d, equal] & (\mathbb{Z}/2 \times \mathbb{Z}/2)\langle r, s\rangle \arrow[r] \arrow[d] & 1 \\ 1 \arrow[r] & \mathbb{Z}/4\langle r\rangle \arrow[r] & D_8\langle r, s\rangle \arrow[r] & \mathbb{Z}/2\langle s\rangle \arrow[r] & 1 \end{tikzcd}$$

When $n = 2$, $[A_1 + B_1] = b$, $[A_1] = f_1$, $[D_2] = f_2$.

$f_1^2 = 0$ makes sense since $[A_1^2] = [(A_1 + B_1)A_1] = bf_1$; there's a multiplicative extension.

In the BSS for the E_2-pages, we have $d_n(f_1) = h_0^n f_2$.

This leads to the Bockstein differential $d_n(A_1D_2) = h_0^n D_2^2$.

So, when $n \geq 2$, $H^*(BD_{2^{n+1}}; \mathbb{Z}) = \mathbb{Z}[a_2, b_2, c_3, d_4]/(2a_2, 2b_2, 2c_3, 2^n d_4, (a_2 + b_2)b_2, c_3^2 + a_2d_4)$.

When $n = 1$, the relation $(a_2 + b_2)b_2$ is replaced by $(a_2 + b_2)b_2 + d_4$, because $f_1^2 = f_2$.

In $D_{2^n} \longrightarrow D_{2^{n+1}} \longrightarrow D_{2^n}$, the first map preserves a_2, c_3, d_4 and takes b_2 to a_2; the second map preserves a_2, b_2 and doubles c_3, d_4.

Calculating $H^*(D_{2^{n+1}m}; \mathbb{Z})$ for $n \geq 1$ and odd m changes the relation on d_4 to $2^n m d_4$ since (accounting for local coefficients) $H^*(B\,\mathbb{Z}/2; H^*(B\,\mathbb{Z}/2^n m; \mathbb{Z}[\frac{1}{2}])) = \mathbb{Z}[\frac{1}{2}][d_4]/(md_4)$.

When m is odd, $H^*(D_{2m}; \mathbb{Z}) = \mathbb{Z}[a_2, d_4]/(2a_2, nd_4)$ since the $\mathbb{Z}$-SSS is quickly computable.

We have a vector field N on $\mathbb{R}^n$ given by

$$(x_1, x_2, \dots, x_n) \longmapsto \sum_{i=1}^{n} x_i \frac{\partial}{\partial x_i}$$

which extends the unit normal vector field on S^{n-1}.

The volume form on $\mathbb{R}^n$ is $dx_1 \wedge dx_2 \wedge \dots \wedge dx_n$.

Let $i : S^{n-1} \longrightarrow \mathbb{R}^n$ be the inclusion. Then

$$i^* \iota_N (dx_1 \wedge dx_2 \wedge \dots \wedge dx_n) = \sum_{i=1}^{n} (-1)^{i-1} x_i \ dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_n$$

is the volume form on S^{n-1}.

Since $\sum_{i=1}^{n} x_i^2 = 1$, we have $\sum_{i=1}^{n} x_i dx_i = 0$ on S^{n-1}.

So in $(x_1, \dots, \widehat{x_j}, \dots, x_n)$ coordinates, by using the fomula

$$dx_j = -\frac{1}{x_j}(x_1 dx_1 + \dots + \widehat{x_j dx_j} + \dots x_n dx_n),$$

we see that

$$\begin{aligned} \sum_{i=1}^{n} (-1)^{i-1} x_i \ dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_n &= \sum_{i=1}^{n} (-1)^{i-1} (-1)^{j-i} \frac{x_i^2}{x_j} \ dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n \\ &= (-1)^{j-1} \frac{1}{x_j} \ dx_1 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n. \end{aligned}$$