Math 115A: Linear Algebra

Michael Andrews
UCLA Mathematics Department

September 13, 2018

Contents

1 Boring things

1.1 Quizzes

There will be quizzes on **Monday 20th August** and **Thursday 30th August**.
They will take place in the first 50 minutes of class.

1.2 Final

The final exam will take place on **Thursday 13th September**.
The exam will be 2 hours long and will take place in the usual classroom.
To give adequate preparation time, the final Tuesday and Wednesday will consist of office hours in the usual classroom.

1.3 Make-up exams?

There will **not** be any make-up exams. Note that university policy requires that a student who has an undocumented absence from the final exam be given a failing grade in the class.

1.4 Questions for turning in

Each lecture I will assign problems for you to turn in during the next lecture. In total, you'll turn in work 15 times. Not everything will be graded because the grader has a limited number of hours, but we'll try to choose good questions for them to grade in order to keep you updated on how well you are doing. Sometimes I might let you know if a question is going to be graded in advance. I will sum up your scores in five groups of 3. Your *lowest* scoring group of questions will not count towards your grade.

I'll be reasonable about how many questions I assign: on Wednesdays, I will not give you much to turn in on Thursday. There will be more questions assigned on a Thursday. Thursday's questions are likely to ask about the whole week's material and to develop ideas that we do not have time for in class. I'll post them as soon as we have covered the relevant material in class.

If you have to miss a lecture, please have a friend turn in your questions.

You should expect to receive your work back within two lectures.

Please **staple** your work. I have given the grader permission to take off points for not stapling.

Kevin is not the grader of the homeworks.

1.5 Grading

Using your turned in questions, quizzes, and final, two scores will be calculated for you using the following schemes:

	Scheme 1	Scheme 2
Turned in questions (lowest group of 3 dropped)	15%	15%
Better Quiz	25%	25%
Other Quiz	20%	0%
Final	40%	60%

Your final score will be the *higher* of those two scores.
You will be assigned a letter grade using your class rank.

Any issues about grading must be addressed within *two weeks*. After that time no score changes will be allowed. Scores will be available online through `my.ucla.edu`.

1.6 Discussion

Ask Kevin to solve a Rubrix cube blindfolded.

1.7 Office hours

Office hours are posted on the website: `math.ucla.edu/~mjandr/Math115A`.

2 Please talk to me, Kevin, and your peers

What do I like least in a (math) classroom? When my students don't talk!

When my students don't talk to me, I don't know whether they are confused or bored. If I can tell that they are confused, I still don't know *when* they became confused or *what* confused them. Was the previous two weeks worth of material overwhelming, or was it the last five minutes? Have I made a mistake on the board? Is it that which is causing confusion? Oh wait, finally someone pointed it out - silly old me! (I officially become old this birthday.) I wish someone had spoken up 20 minutes ago when they were first confused!

There are many reasons that I can think of for students to be hesitant to speak out. In my Math 95 class, I'm going to make a note of the reasons that they think of, so take a look at my notes for that course if you're interested: `http://math.ucla.edu/~mjandr/Math95`. I hope that this list of *the most important things that I learned in grad school* might help you to be less quiet.

- I am in control of my learning. How much I participate usually corresponds with how much I enjoy a class and how much I learn in the class.
- Asking a question earlier rather than later usually results in less confusion on my part.
- If I am confused, it is likely that half the room is confused too, and my question will probably help others. It can help in two ways.
  - Directly, by clearing up a mistake or something that the lecturer said unclearly.
  - Indirectly, by calling the lecturer's attention to the fact that people are confused, forcing a flexible lecturer to try to change something.
- My "stupid" questions are usually my most important questions and it makes me feel better if I do not think of them as "stupid."
- Math is really difficult, and on some days it feels worse than others. It's better to be kind to myself on such days, and to ask the people who are explaining ideas to me to slow down or explain things more thoroughly if necessary.
- The lecturer probably likes questions.

3 Lecture on August 6th: Sets and functions

3.1 Sets

Definition 3.1.1. A *set* is a collection of mathematical things. The members of a set are called its *elements*. We write $x \in X$ to mean that "x is an element of a set X."

Definition 3.1.2. We often use curly brackets (braces) for sets whose elements can be written down easily. For example, the set whose elements are $a_1, a_2, a_3, \ldots, a_n$ is written as $\{a_1, a_2, a_3, \ldots, a_n\}$, and the set whose elements are those of an infinite sequence $a_1, a_2, a_3, \ldots$ is written as $\{a_1, a_2, a_3, \ldots\}$.

Remark 3.1.3. Some infinite sets can not be written out in an infinite sequence, so not all sets can be described using the notation just defined. An example is the set of real numbers. If this interests you, come to the first week or so of my math 95 class: `http://math.ucla.edu/~mjandr/Math95`.

Definition 3.1.4. We often describe a set as "elements with some property." For example, the set of elements with a property P is written as $\{x : x \text{ has property } P\}$, and the elements of a previously defined set X with a property P is written as $\{x \in X : x \text{ has property } P\}$.

Remark 3.1.5.

- In the previous definition the colons should be read as "such that."
- The second way of describing a set $\{x \in X : \ x \text{ has property } P\}$ is much better than the first because the first allows for Russell's paradox: the "set" $\{A : \ A \text{ is a set and } A \notin A\}$ makes no sense! `http://en.wikipedia.org/wiki/Barber_paradox#Paradox`

Notation 3.1.6. Here is some notation for familiar sets. ":=" means "is defined to be equal to."

1. the *empty set* $\emptyset := \{ \ \}$ is the set with no elements,
2. the *natural numbers* $\mathbb{N} := \{1, 2, 3, \ldots\}$,
3. the *integers* $\mathbb{Z} := \{0, -1, 1, -2, 2, -3, 3, \ldots\}$,
4. the *rationals* $\mathbb{Q} := \{\frac{m}{n} : \ m \in \mathbb{Z}, \ n \in \mathbb{N}\}$,
5. the *reals* $\mathbb{R}$,
6. the *complex numbers* $\mathbb{C} := \{x + iy : \ x, y \in \mathbb{R}\}$.

Definition 3.1.7. Suppose X and Y are sets. We say that X is a *subset* of Y **iff** whenever $z \in X$, we have $z \in Y$. We write $X \subseteq Y$ to mean "X is a subset of Y."

Remark 3.1.8. 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.

Remark 3.1.9. Suppose X and Y are sets, that X is a subset of Y, and that you want to write a proof for this. What do you write?

Well, another way of expressing the "whenever" sentence in the previous definition is "if $z \in X$, then $z \in Y$." To verify a sentence like "if P, then Q" directly you assume that P is true, and check that Q is true. Thus, the definition forces our proof to look as follows.

- We wish to show $X \subseteq Y$.
- By definition, we must show that if $z \in X$, then $z \in Y$.
- So suppose that $z \in X$.
- We want to show that $z \in Y$.
- [Insert mathematical arguments to show that $z \in Y$.]
- We conclude that $z \in Y$, and so we have shown that $X \subseteq Y$.

Theorem 3.1.10. *Let $A = \{n \in \mathbb{N} : \ n \geq 8$ and n is prime$\}$ and $B = \{n \in \mathbb{Z} : \ n$ is odd$\}$.*
We have $A \subseteq B$.

Proof. Let A and B be as in the theorem statement.

We wish to show $A \subseteq B$. By definition, we must show that if $n \in A$, then $n \in B$. So suppose that $n \in A$. We want to show that $n \in B$.

Because $n \in A$, the definition of A tells us that n is a natural number which is greater than or equal 8 and prime. In order to show $n \in B$, we must demonstrate that n is an odd integer. Natural numbers are integers and n is a natural number, and so n is an integer. We are just left to show n is odd.

Suppose for contradiction that n is not odd. Then n is even, and $\frac{n}{2} \in \mathbb{N}$. Since $n \geq 8$, $\frac{n}{2} \geq 4$. Thus, writing $n = 2 \cdot \frac{n}{2}$ shows that n is not prime. This contradicts the fact that n *is* prime, and so n must be odd.

We conclude that $n \in B$, and so we have shown that $A \subseteq B$. □

Here is a complete list of the things we do during the previous proof. You might want to rewrite this list in your notes and add to it whenever we learn new things to do in a proof.

- We introduce the mathematical objects that are we are going to be using during the proof.
- We state what we wish to demonstrate before doing so.
- We unpack and use definitions where necessary (which is a lot of the time). We try to highlight when we are doing this unless we think that the reader can figure it out without such help. We should assume that the reader is another student in the class.
- We state assumptions.
- We prove an if-then statement directly.
- We use modus ponens.

  The first paragraph of `http://en.wikipedia.org/wiki/Modus_ponens` and the first paragraph of the section "explanation" are useful. You might find the rest overly confusing.

  An example of using modus ponens we might see later is: "we know that bases are linearly independent, and $(v_1, \ldots, v_n)$ is a basis, so $(v_1, \ldots, v_n)$ is linearly independent."

  You could abbreviate this using the words "in particular" by saying: "$(v_1, \ldots, v_n)$ is a basis. In particular, $(v_1, \ldots, v_n)$ is linearly independent."

  We could also abbreviate to "since $(v_1, \ldots, v_n)$ is a basis, $(v_1, \ldots, v_n)$ is linearly independent."

- We make a deduction from an assumption declared in the previous sentence.

  For example, "Suppose that $(v_1, \ldots, v_n)$ is a basis. Then $(v_1, \ldots, v_n)$ is linearly independent."

- We use a proof by contradiction:
  - We make an assumption that is going to cause a contradiction, the negation of what we want to verify. We use the words "suppose for contradiction…"
  - We highlight what causes the contradiction, and conclude what we wanted to verify.

- We use a mathematical equation.

- We summarize what we have done.

For each sentence of the previous proof, find the appropriate bullet point(s) describing what we are doing.

There is one case when the previous proof format for showing that $X \subseteq Y$ breaks down.

Theorem 3.1.11. *Suppose X is a set. Then $\emptyset \subseteq X$.*

Proof. Suppose X is a set. We wish to show that $\emptyset \subseteq X$. By definition, we must show that if $z \in \emptyset$, then $z \in X$. Since the empty set has no elements, $z \in \emptyset$ is never true, and so there is nothing to check. We conclude that $\emptyset \subseteq X$. □

Remark 3.1.12. Maybe the "whenever" wording makes this proof seems less strange.

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. We conclude that $\emptyset \subseteq X$.

Definition 3.1.13. Suppose X and Y are sets. We say that X is *equal to* Y **iff** X is a subset of Y and Y is a subset of X. We write $X = Y$ to mean "X is equal Y."

Remark 3.1.14. Suppose X and Y are sets, that X is equal to Y, and that you want to write a proof for this. What do you write?

- We wish to show $X = Y$.
- By definition, we must show that $X \subseteq Y$, and $Y \subseteq X$.
- [Insert proof that $X \subseteq Y$.]
- [Insert proof that $Y \subseteq X$.]
- We have shown that $X \subseteq Y$ and $Y \subseteq X$, so we have shown that $X = Y$.

Example 3.1.15.

1. $\{0, 1\} = \{0, 0, 0, 1, 1\} = \{1, 0\}$.
2. $\emptyset \neq \{\emptyset\}$. This is because $\emptyset \notin \emptyset$ and so $\{\emptyset\} \not\subseteq \emptyset$.
3. $\{n \in \mathbb{N} : n \text{ divides } 12\} = \{1, 2, 3, 4, 6, 12\}$.

3.2 Some more on sets (covered in discussion)

Definition 3.2.1. 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
$$\{z : z \in X \text{ and } z \in Y\}.$$
$X \cap Y$ 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 3.2.2. I hope Kevin will give examples in discussion.

Theorem 3.2.3 (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. Due on Wednesday.

2. Suppose X, A, and B are sets. We wish to show that
$$X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B).$$
By definition of set equality, we must show that
$$X \setminus (A \cap B) \subseteq (X \setminus A) \cup (X \setminus B) \quad \text{and} \quad X \setminus (A \cap B) \supseteq (X \setminus A) \cup (X \setminus B).$$
   (a) First, we demonstrate that $X \setminus (A \cap B) \subseteq (X \setminus A) \cup (X \setminus B)$, that is, 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 2: $z \notin B$. 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 2: $z \in X \setminus B$. 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 $z \in X \setminus (A \cap B)$.

□

Remark 3.2.4. 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!

3.3 Functions

Definition 3.3.1. 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 way of assigning to each $x \in X$, exactly one $y \in Y$.

  We write $f(x)$ for the unique $y \in Y$ assigned to x.

Notation 3.3.2. We often use the notation $f : X \longrightarrow Y$, $x \longmapsto f(x)$.
"$\longmapsto$" is read as "maps to."

Definition 3.3.3. 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 3.3.4. 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.

Remark 3.3.5. Suppose X and Y are sets, $f : X \longrightarrow Y$ is a function, and that f is injective.
How do you write a proof that f is injective?
Well, another way of expressing the "whenever" sentence in part 1 of the previous definition is "if $x_1, x_2 \in X$ and $f(x_1) = f(x_1)$, then $x_1 = x_2$."
Thus, the definition forces our proof to look as follows.

- We wish to show that f is injective.
- By definition, we must show that if $x_1, x_2 \in X$ and $f(x_1) = f(x_1)$, then $x_1 = x_2$.
- So suppose that $x_1, x_2 \in X$ and $f(x_1) = f(x_2)$.
- (We want to show that $x_1 = x_2$.)
- [Insert mathematical arguments to show that $x_1 = x_2$.]
- We conclude that $x_1 = x_2$, and so we have shown that f is injective.

Theorem 3.3.6. *The function* $f : \{x \in \mathbb{R} : \ x \geq 0\} \longrightarrow \mathbb{R}$, $x \longmapsto x^2$ *is injective.*

Proof. Let $X = \{x \in \mathbb{R} : \ x \geq 0\}$ and $f : X \longrightarrow \mathbb{R}$ be as in the theorem statement. We wish to show that f is injective. By definition of injectivity, we must show that if $x_1, x_2 \in X$ and $f(x_1) = f(x_1)$, then $x_1 = x_2$. So suppose $x_1, x_2 \in X$ and $f(x_1) = f(x_2)$. By definition of f, $f(x_1) = f(x_2)$ tells us that $x_1^2 = x_2^2$. Taking the square roots of both sides gives $|x_1| = |x_2|$. Since $x_1, x_2 \in X$, $x_1, x_2 \geq 0$, and the previous equations says that $x_1 = x_2$. We have now shown that f is injective. $\square$

Remark 3.3.7. Suppose X and Y are sets, $f : X \longrightarrow Y$ is a function, and that f is surjective.

How do you write a proof that f is surjective?

Well, another way of expressing the "whenever" sentence in part 2 of the previous definition is "if $y \in Y$, then we can find an $x \in X$ such that $f(x) = y$."

Thus, the definition forces our proof to look as follows.

- We wish to show that f is surjective.
- By definition, we must show that if $y \in Y$, then we can find an $x \in X$ such that $f(x) = y$.
- So suppose that $y \in Y$.
- (We want to show that we can find an $x \in X$ such that $f(x) = y$.)
- [Insert mathematical arguments.]

  These arguments must do two things:

  - Specify an element $x \in X$.
  - Check that the specified x satisfies the equation $f(x) = y$.
- We have completed the demonstration that f is surjective.

Theorem 3.3.8. *The function* $f : \mathbb{R} \longrightarrow \{y \in \mathbb{R} : \ y \geq 0\}$, $x \longmapsto x^2$ *is surjective.*

Proof. Let $Y = \{y \in \mathbb{R} : \ y \geq 0\}$ and $f : \mathbb{R} \longrightarrow Y$ be defined as in the theorem statement. We wish to show that f is surjective. By definition of surjectivity, we must show that if $y \in Y$, then we can find an $x \in \mathbb{R}$ such that $f(x) = y$. So suppose that $y \in Y$. Since $y \geq 0$, we can let $x = \sqrt{y}$. Then $f(x) = x^2 = (\sqrt{y})^2 = y$, and we have completed the demonstration that f is surjective. $\square$

Remark 3.3.9. Suppose X and Y are sets, $f : X \longrightarrow Y$ is a function, and that f is a bijection.

How do you write a proof that f is a bijection?

The definition forces our proof to look as follows.

- We wish to show that f is bijection.
- By definition, we must show that f is injective and surjective.
- [Insert proof that f is injective.]
- [Insert proof that f is surjective.]
- We have demonstrated that f is bijection.

Theorem 3.3.10. *The function* $f : \{x \in \mathbb{R} : \ x \geq 0\} \longrightarrow \{y \in \mathbb{R} : \ y \geq 0\}$*,* $x \longmapsto x^2$ *is a bijection.*

Proof. Omitted. $\square$

Definition 3.3.11. Suppose X, Y, and Z are sets, and that $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ are functions. The *composition of* f *and* g is the function $g \circ f : X \to Z$ defined by $(g \circ f)(x) = g(f(x))$.

Theorem 3.3.12. *Suppose* X*,* Y*, and* Z *are sets, and* $f : X \longrightarrow Y$ *and* $g : Y \longrightarrow Z$ *are functions.*

1. *If* $g \circ f$ *is injective, then* f *is injective.*
2. *If* $g \circ f$ *is surjective, then* g *is surjective.*
3. *If* f *and* g *are injective, then* $g \circ f$ *is injective.*
4. *If* f *and* g *are surjective, then* $g \circ f$ *is surjective.*

Proof. Left to you. $\square$

4 Questions due on August 8th

1. (a) How many functions are there with domain $\{1, 2, 3\}$ and codomain $\{1, 2, 3, 4, 5\}$?
   (b) How many injective functions are there with domain $\{1, 2, 3\}$ and codomain $\{1, 2, 3, 4, 5\}$?
   (c) How many surjective functions are there with domain $\{1, 2, 3\}$ and codomain $\{1, 2, 3, 4, 5\}$?
   (d) How many functions are there with domain $\{1, 2, 3, 4, 5\}$ and codomain $\{1, 2, 3\}$?
   (e) How many injective functions are there with domain $\{1, 2, 3, 4, 5\}$ and codomain $\{1, 2, 3\}$?
   (f) How many surjective functions are there with domain $\{1, 2, 3, 4, 5\}$ and codomain $\{1, 2, 3\}$? Hint: how many functions are there with domain $\{1, 2, 3, 4, 5\}$ and codomain $\{1, 2\}$?
   (g) Let $n \in \mathbb{N}$. How many bijections are there with domain and codomain $\{1, 2, \dots, n\}$?

   Solution:

   (a) 5^3. (1 pt.)
   (b) $5 \cdot 4 \cdot 3$. (1 pt.)
   (c) 0. (1 pt.)
   (d) 3^5.
   (e) 0.
   (f) Let try to count how many functions $f : \{1, 2, 3, 4, 5\} \longrightarrow \{1, 2, 3\}$ are NOT surjective. There are three cases: Let $X = \{1, 2, 3, 4, 5\}$
      i. $1 \notin f(A)$: then f is basically a function $X \longrightarrow \{2, 3\}$. There are 2^5 such functions.
      ii. $2 \notin f(A)$: then f is basically a function $X \longrightarrow \{1, 3\}$. There are 2^5 such functions.
      iii. $3 \notin f(A)$: then f is basically a function $X \longrightarrow \{1, 2\}$. There are 2^5 such functions.

      It looks like we have counted that there are $3 \cdot 2^5$ functions which are NOT surjective. But we have double counted: there are three constant functions which have been counted twice. So there are $3 \cdot 2^5 - 3$ functions which are NOT sujective.
      The answer is $3^5 - 3 \cdot 2^5 + 3$. (2 pts.)
   (g) $n!$ (1 pt.)

2. Prove part 1 of theorem 3.3.12. Here's a general outline for how your proof should look.
   (a) Introduce relevant mathematical objects.
   (b) State what you want to prove.
   (c) It is an if-then sentence, so suppose the premise and say that you would like to verify the conclusion.
   (d) Unpack the definition of the what you have just said you want to verify.
   (e) It is an if-then sentence, so suppose the premise and say that you would like to verify the conclusion.
   (f) Figure out how to do this verification. You'll need to use your assumptions. This is the crux of proof, but without the context for this argument (the previous 5 steps) it does not really make any sense.
   (g) Conclude your proof.

Solution: (2 pts. for format, 2 pts. for (f).)

(a) Suppose X, Y, and Z are sets, and that $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ are functions.

(b) We want to prove that if $g \circ f$ is injective, then f is injective.

(c) So suppose $g \circ f$ is injective. We would like to verify that f is injective.

(d) By definition, we must show that if $x_1, x_2 \in X$ and $f(x_1) = f(x_1)$, then $x_1 = x_2$.

(e) So let $x_1, x_2 \in X$ and suppose that $f(x_1) = f(x_2)$. We need to show that $x_1 = x_2$.

(f) Since $f(x_1) = f(x_2)$, we can apply g to see $g(f(x_1)) = g(f(x_2))$. By definition of $g \circ f$, this says that $(g \circ f)(x_1) = (g \circ f)(x_2)$. Since $g \circ f$ is injective, this gives $x_1 = x_2$.

(g) Thus, we have shown f is injective, and this finishes the proof that if $g \circ f$ is injective, then f is injective.

3. Prove part 1 of theorem 3.2.3.

[You should be trying to write the proof similarly to how I proved part 2 in the notes, or how Kevin proved it in discussion. My proof of the current result does not need case work; in this sense, the current proof is easier than the one in the notes.]

Solution: (1 pt. for completion.)

Suppose X, A, and B are sets. We wish to show that

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

By definition of set equality, we must show that

$$X \setminus (A \cup B) \subseteq (X \setminus A) \cap (X \setminus B) \quad \text{and} \quad X \setminus (A \cup B) \supseteq (X \setminus A) \cap (X \setminus B).$$

(a) First, we demonstrate that $X \setminus (A \cup B) \subseteq (X \setminus A) \cap (X \setminus B)$.
By definition of $\subseteq$, we have to show that whenever $z \in X \setminus (A \cup B)$, it is the case that $z \in (X \setminus A) \cap (X \setminus B)$.
So suppose $z \in X \setminus (A \cup B)$. By definition of $\setminus$, this means $z \in X$ and $z \notin A \cup B$. We cannot have $z \in A$ or $z \in B$ since, by the definition of $\cup$, both these conditions imply $z \in A \cup B$.
Thus, $z \in X$ and $z \notin A$, and the definition of $\setminus$ gives $z \in X \setminus A$.
Similarly, $z \in X$ and $z \notin B$, and the definition of $\setminus$ gives $z \in X \setminus B$.
By the definition of $\cap$, the last two facts say $z \in (X \setminus A) \cap (X \setminus B)$.

(b) Next, we show that $X \setminus (A \cup B) \supseteq (X \setminus A) \cap (X \setminus B)$, i.e if $z \in (X \setminus A) \cap (X \setminus B)$, then $z \in X \setminus (A \cup B)$.
Suppose that $z \in (X \setminus A) \cap (X \setminus B)$. By definition of $\cap$, this means that $z \in X \setminus A$ and $z \in X \setminus B$.
The first statement together with the definition of $\setminus$ says $z \in X$ and $z \notin A$.
The second statement together with the definition of $\setminus$ says $z \in X$ and $z \notin B$.
If we had $z \in A \cup B$, by definition of $\cup$, we'd have either $z \in A$ or $z \in B$, and this is not the case. Thus, $z \notin A \cup B$.
In conclusion, $z \in X$ and $z \notin A \cup B$, so the definition of $\setminus$ gives $z \in X \setminus (A \cup B)$.

5 Lecture on August 8th: Vector spaces over $\mathbb{R}$

Definition 5.1. 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 5.2.

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 $\mathbb{R}^2$.
3. $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$ is the home of 3D calculus $\mathbb{R}^3$.

Definition 5.3. A *vector space over* $\mathbb{R}$ is a set V together with operations

- $+ \ : V \times V \longrightarrow V, \ (v, w) \longmapsto v + w$ (addition)
- $\cdot \ : \mathbb{R} \times V \longrightarrow V, \ (\lambda, v) \longmapsto \lambda v$ (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 $\mathbb{R}$ acts sensibly under scalar multiplication).
6. $\forall \lambda \in \mathbb{R}, \ \forall \mu \in \mathbb{R}, \ \forall v \in V, \ (\lambda\mu)v = \lambda(\mu v)$
   (the interaction of $\mathbb{R}$'s multiplication and scalar multiplication is sensible).
7. $\forall \lambda \in \mathbb{R}, \ \forall u \in V, \ \forall v \in V, \ \lambda(u + v) = \lambda u + \lambda v$
   (the interaction of scalar multiplication and vector space addition is sensible).
8. $\forall \lambda \in \mathbb{R}, \ \forall \mu \in \mathbb{R}, \ \forall v \in V, \ (\lambda + \mu)v = \lambda v + \mu v$
   (the interaction of $\mathbb{R}$'s addition and scalar multiplication is sensible).

Example 5.4.

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.

   In class, for the case of $n = 2$, I'll check axioms 1, 3, and 4.

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.

   In class, I'll check axioms 2, 3, and 6.

   I'd certainly copy down what I write, and check with me and friends that you copied correctly.

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 5.5.

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

   Axiom 8 fails because

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

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 5.6. Maybe a reminder of how a proof by contradiction works is still useful. 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 $1(0, 1) = (0, 0) \neq (0, 1)$.

There are facts which we would like to take for granted whenever dealing with vector spaces. Proving these facts gives us permission to take them for granted forever afterwards (unless I ask you to prove them on an exam). I'm not a big fan of axiomatic proofs, but in the following proofs there are some things which we do which you can add to your "good proofs" list.

- We explain our ideas to the reader.
- We reference and use axioms clearly.
- We see what it means to be the unique element with a property, and how to use uniqueness.
- We reduce a proof to a more easily obtained goal using a previously proven fact.
- We reference previously obtained facts.
- We leave some things to the reader. Maybe that is frustrating for you? So you can see why we deduct points when not everything is properly explained.

Theorem 5.7. *Suppose V is a vector space over $\mathbb{R}$.*

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 $v \in V$ with the property that $u + v = 0$.*

   We give it a more useful name: $(-u)$.
4. *Let $\lambda \in \mathbb{R}$ 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 $\lambda \in \mathbb{R}$ and $v \in V$. Then $-(\lambda v) = (-\lambda)v = \lambda(-v)$.*

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

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

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

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 $\mathbb{R}$, 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 $\lambda \in \mathbb{R}$ 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 properties of $\mathbb{R}$ (I'll leave it to you to figure out which is which).

$\square$

6 Questions due on August 9th

1. In class, I said that $\mathbb{R}^2$ is a vector space over $\mathbb{R}$ when equipped with coordinatewise addition and scalar multiplication. I checked axioms 1, 3, and 4, but I omitted checking axiom 2, and axioms 5-8.

   Carefully verify axioms 2, 5, 6, 7. Each of your equalities should be justified. Are you using the definition of addition for $\mathbb{R}^2$, the definition of scalar multiplication for $\mathbb{R}^2$, properties of real number addition and/or multiplication?

   I suppose you can do 8 as well if you want more practice.

   Solution: (1 pt. for completion.)

   I'll post a little bit of axiom checking in the solutions to the weekend problems.

2. Let V be a vector space over $\mathbb{R}$, $\lambda, \mu \in \mathbb{R}$, and $u, v \in V$. (Since axiom 2 is true, it does not matter how you parenthesize expressions involving the addition of three or more vectors.) Prove that

$$(\lambda + \mu)(u + v) = \lambda u + \lambda v + \mu u + \mu v,$$

   carefully referencing which axioms of a vector space you use for every equality.

   (I can think of two proofs and one is shorter than the other.)

   Solution: (1 pt. for completion.)

   Let V be a vector space over $\mathbb{R}$, $\lambda, \mu \in \mathbb{R}$, and $u, v \in V$. Then

$$(\lambda + \mu)(u + v) = \lambda(u + v) + \mu(u + v) = \lambda u + \lambda v + \mu u + \mu v.$$

   The first equality is axiom 8; the second is axiom 7 twice.

7 Lecture on August 9th: Subspaces and linear transformations

7.1 Subspaces

Definition 7.1.1. Suppose V is a vector space over $\mathbb{R}$ with operations

$$+ : V \times V \longrightarrow V \text{ and } \cdot\ : \mathbb{R} \times V \longrightarrow V,$$

and that W is a set.

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

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

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

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

Example 7.1.2. $\{(x, 0) :\ x \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^2$.

Example 7.1.3. Suppose that V is a vector space over $\mathbb{R}$ 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 7.1.4. *Suppose V is a vector space over $\mathbb{R}$ with operations $+ : V \times V \to V$, $\cdot : \mathbb{R} \times V \to V$, 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 $\lambda \in \mathbb{R}$ and $w \in W$, then $\lambda w \in W$.*

Proof. Suppose V is a vector space over $\mathbb{R}$ with operations $+ : V \times V \longrightarrow V$ and $\cdot\ : \mathbb{R} \times V \longrightarrow V$, 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

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

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 $\mathbb{R}$, 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 5.7 part 6, we know that $-w = (-1)w$ and by assumption 3, this shows $-w \in W$. $\square$

Remark 7.1.5. 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.

Example 7.1.6.

$$\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : \ 8x_1 + 18x_2 + 88x_3 = 0 \text{ and } 7x_2 + 14x_3 + 94x_4 = 0\}$$

is a subspace of $\mathbb{R}^4$ (with the usual addition and scalar multiplication).

I'll check this carefully in class using the subspace test.

Example 7.1.7.

1. 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 $\{f : \mathbb{R} \longrightarrow \mathbb{R}\}$.

2. 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 $\{f : \mathbb{R} \longrightarrow \mathbb{R}\}$.

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

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

5. You might want to try and say that $\mathcal{P}(\mathbb{R})$ 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 since it can happen if you consider polynomials over a finite field, this magic fact requires a proof.

   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? Your best strategy is to show that i is linear and that its kernel is $\{0\}$, but you'll have to wait a little to be told the definitions of these terms.

Theorem 7.1.8. *Suppose W is a subspace of $\mathbb{R}^3$. Then W is either:*

1. *a line through the origin.*

   This means that there is a vector $(a_1, a_2, a_3) \in \mathbb{R}^3 \setminus \{0\}$ such that

$$W = \{\lambda(a_1, a_2, a_3) : \ \lambda \in \mathbb{R}\}.$$

2. *a plane through the origin.*

   This means that there is a vector $(n_1, n_2, n_3) \in \mathbb{R}^3 \setminus \{0\}$ such that

$$W = \{\lambda(x_1, x_2, x_3) \in \mathbb{R}^3 : \ x_1 n_1 + x_2 n_2 + x_3 n_3 = 0.\}$$

3. *a trivial subspace, either $\{0\}$ or $\mathbb{R}^3$.*

Remark 7.1.9. We will not prove this theorem until later. Until we do prove it, you should not use it in your proofs. On the other hand, you should use it to help you think sensibly about whether certain statements concerning subspaces are likely to be true or not.

Theorem 7.1.10. *Suppose V is a vector space over a $\mathbb{R}$. Suppose that $\{W_i : i \in I\}$ is some collection of subspaces of V. 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.

Remark 7.1.11. 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 is allowed to be infinite in this notation.

Proof. Suppose V is a vector space over a $\mathbb{R}$. Suppose that $\{W_i : i \in I\}$ is some collection of subspaces of V 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 $\lambda \in \mathbb{R}$. We have to show that $w + w' \in W$ and $\lambda w \in W$. By definition of intersection, we have $w \in W_i$ and $w' \in W_i$ for all $i \in I$. 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 $\lambda w \in W$. $\square$

7.2 Linear transformations

Definition 7.2.1. Let V and W be vector spaces over $\mathbb{R}$. 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. $\forall \lambda \in \mathbb{R}$, $\forall v \in V$, $T(\lambda v) = \lambda T(v)$.

Often, we say just "T is linear."

Lemma 7.2.2. *Let V and W be vector spaces over $\mathbb{R}$ 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*

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

Concise proof. Let V and W be vector spaces over $\mathbb{R}$ and $T : V \longrightarrow W$ be a function. We'll prove part 1 as we prove part 2. Suppose T is linear, $\lambda \in \mathbb{R}$, $v_1 \in V$, $v_2 \in V$. Then

$$T(\lambda v_1 + v_2) = T(\lambda v_1) + T(v_2) = \lambda T(v_1) + T(v_2).$$

Conversely, suppose that

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

Taking $\lambda = 1$ proves that the first part of linearity, and we now have $T(0)+T(0) = T(0+0) = T(0)$, so that $T(0) = 0$. Taking $v_2 = 0$ proves the second part of linearity. $\square$

Example 7.2.3. Suppose $m, n \in \mathbb{N}$ and $A \in M_{m\times n}(\mathbb{R})$. Then $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T_A(x) = Ax$ is a linear transformation. This follows from facts about matrix-vector multiplication. We'll use the notation T_A throughout the rest of the class, so don't ignore this.

Remark 7.2.4. Notice that for the previous example to make sense, we have to think of elements of $\mathbb{R}^n$ and $\mathbb{R}^m$ as column vectors, rather than row vectors / n-tuples. We'll often blur the distinction between row and column vectors. If you are ever confused about this, please ask.

Example 7.2.5. Suppose that X is a nonempty set. Recall that $\mathcal{F} = \{f : X \longrightarrow \mathbb{R}\}$ is a vector space over $\mathbb{R}$. Given $x_0 \in X$, $\mathrm{ev}_{x_0} : \mathcal{F} \longrightarrow \mathbb{R},\ f \longmapsto f(x_0)$ is linear.

Example 7.2.6. Suppose V and W are vector spaces over $\mathbb{R}$.

- The *identity function* $1_V : V \longrightarrow V,\ v \longmapsto v$ is a linear transformation.
- The *zero function* $0_{V,W} : V \longrightarrow W,\ v \longmapsto 0$ is a linear transformation.

Notation 7.2.7. Suppose U, V, and W are vector spaces over $\mathbb{R}$, that $T : U \to V$ and $S : V \to W$ are linear transformations. Then we write $ST : U \longrightarrow W$ for the composite $S \circ T$.

Lemma 7.2.8. *Suppose U, V, and W are vector spaces over $\mathbb{R}$, that $T : U \to V$ and $S : V \to W$ are linear transformations. Then $ST : U \longrightarrow W$ is a linear transformation.*

Example 7.2.9. 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 will prove that $\mathcal{F}_{\text{cont}}([a, b]) \supseteq \mathcal{F}_{\text{diff}}([a, b])$ 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

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

is the natural inclusion.

8 Questions due on August 13th

1. Prove part 2 of theorem 3.3.12.

   The format of your proof should be similar to that of part 1. You should note that part (f) requires specifying a $y \in Y$ with a given property. To write down such a y, you probably need to have found another element first. I'll leave for you to think about which set this element is contained in. You will get this auxilary element using an assumed surjectivity.

   Solution: [1 pt. for completion]

   (a) Suppose X, Y, and Z are sets, and that $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ are functions.

   (b) We want to prove that if $g \circ f$ is surjective, then g is surjective.

   (c) So suppose $g \circ f$ is surjective. We would like to verify that g is surjective.

   (d) By definition, we must show that if $z \in Z$, then we can find an $y \in Y$ such that $g(y) = z$.

   (e) So let $z \in Z$. We need to specify a $y \in Y$ with $g(y) = z$.

   (f) Since $g \circ f$ is surjective, we can find an $x \in X$ with $(g \circ f)(x) = z$.
   Let $y = f(x)$. Then $g(y) = g(f(x)) = (g \circ f)(x) = z$.

   (g) Thus, we have shown g is surjective, and this finishes the proof that if $g \circ f$ is surjective, then g is surjective.

2. [Optional]

   Suppose X, Y, A, and B are sets and that $A \subseteq X$ and $B \subseteq Y$. Prove that

   $$(X \times Y) \setminus (A \times B) = ((X \setminus A) \times Y) \cup (X \times (Y \setminus B)).$$

   Cases made my proof clearer.

   If you're attempting to be careful, then you should make explicit reference to the definition of the Cartesian product. For example, you might say the following. "Let $z \in (X \times Y) \setminus (A \times B)$. Then $z \in X \times Y$. By the definition of the Cartesian product, $z = (x, y)$ for some $x \in X$ and some $y \in Y$."

   Solution:

   Suppose X, Y, A, and B are sets and that $A \subseteq X$ and $B \subseteq Y$. We wish to show that

   $$(X \times Y) \setminus (A \times B) = ((X \setminus A) \times Y) \cup (X \times (Y \setminus B)).$$

   By definition of set equality, we must show "$\subseteq$" and "$\supseteq$."

   (a) First, we show '$\subseteq$."
   Suppose $z \in (X \times Y) \setminus (A \times B)$.
   By definition of $\setminus$, this says $z \in X \times Y$ and $z \notin A \times B$. By definition of the Cartesian product $X \times Y$, we can write $z = (x, y)$ where $x \in X$ and $y \in Y$.
   We cannot have $x \in A$ and $y \in B$, for otherwise the definition of the Cartesian product $A \times B$ would give $z = (x, y) \in A \times B$. Thus, either $x \notin A$, or $y \notin B$.

i. Case 1: $x \notin A$.
Since we have $x \in X$, the definition of $\setminus$ gives $x \in X \setminus A$.
Since $y \in Y$, the definition of Cartesian product gives $z = (x, y) \in (X \setminus A) \times Y$.
By definition of $\cup$, this gives $z \in ((X \setminus A) \times Y) \cup (X \times (Y \setminus B))$.

ii. Case 2: $y \notin B$.
Since we have $y \in Y$, the definition of $\setminus$ gives $y \in Y \setminus B$.
Since $x \in X$, the definition of Cartesian product gives $z = (x, y) \in X \times (Y \setminus B)$.
By definition of $\cup$, this gives $z \in ((X \setminus A) \times Y) \cup (X \times (Y \setminus B))$.

In either case, we have $z \in ((X \setminus A) \times Y) \cup (X \times (Y \setminus B))$.

(b) First, we show '$\supseteq$."
Suppose $z \in ((X \setminus A) \times Y) \cup (X \times (Y \setminus B))$. By definition of $\cup$, there are two cases.

i. Case 1: $z \in (X \setminus A) \times Y$.
By definition of Cartesian product, $z = (x, y)$ where $x \in X \setminus A$ and $y \in Y$.
By definition of $\setminus$, we have $x \in X$ and $x \notin A$.
Because $x \in X$ and $y \in Y$, the definition of the Cartesian product $X \times Y$ tells us that we have $z = (x, y) \in X \times Y$.
Because $x \notin A$, the definition of the Cartesian product $A \times B$ tells us that we do not have $(x, y) \in A \times B$.
So, by definition of $\setminus$, we have $z \in (X \times Y) \setminus (A \times B)$.

ii. Case 2: $z \in X \times (Y \setminus B)$.
By definition of Cartesian product, $z = (x, y)$ where $x \in X$ and $y \in Y \setminus B$.
By definition of $\setminus$, we have $y \in Y$ and $y \notin B$.
Because $x \in X$ and $y \in Y$, the definition of the Cartesian product $X \times Y$ tells us that we have $z = (x, y) \in X \times Y$.
Because $y \notin B$, the definition of the Cartesian product $A \times B$ tells us that we do not have $(x, y) \in A \times B$.
So, by definition of $\setminus$, we have $z \in (X \times Y) \setminus (A \times B)$.

In either case, we have $z \in (X \times Y) \setminus (A \times B)$.

3. The first time I taught this class, someone suggested that

$$\{(x_1, x_2) \in \mathbb{R}^2 : \; x_1 x_2 \geq 0\}$$

with addition defined by

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$$

and scalar multiplication defined by $\lambda(x_1, x_2) = (\lambda x_1, \lambda x_2)$ is a vector space over $\mathbb{R}$.

Although this is incorrect, I was very grateful for their suggestion. Please do not stop making such suggestions! This is a great example for improving your intuition of what a vector space is, and making sure that you understand *all* the ingredients to be one.

(a) Which axioms are true for this wannabe-vector-space? You don't need to write down all of your proofs. If you think for a while, you'll see that most of them are the same proofs that you or I did for $\mathbb{R}^2$. Perhaps you could expand on axiom 4 a little.

(b) Does addition always make sense?
If yes, give a quick explanation as to why it is always okay.
If not, give an explicit example to show it screws up.

(c) Does scalar multiplication always make sense?
If yes, give a quick explanation as to why it is always okay.
If not, give an explicit example to show it screws up.

(d) Why is this wannabe-vector-space just a wannabe?

(e) This wannabe-vector-space is actually a subset of the vector space $\mathbb{R}^2$.
Show that it fails the subspace test.

Solution:

(a) In (b), you find out that addition it not well-defined. Because of this axioms 1,2,3,4,7,8 are just weird from the get-go. However, if you don't see the problem with addition, and naïvely check the axioms, you'll succeed! Even axiom 4 works since $(-x_1)(-x_2) = x_1x_2$.

(b) [2 pts.] Addition is not well-defined since it would say

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

$(1, 0), (0, -1)$ are in the wannabe-vector-space, but $(1, -1)$ is not in there.

(c) Scalar multiplication is okay: because $(\lambda x_1)(\lambda x_2) = \lambda^2(x_1x_2)$ and $\lambda^2 \geq 0$, we find that $(\lambda x_1, \lambda x_2)$ is an element of the wannabe-vector-space whenever (x_1, x_2) is.

(d) Addition is not well-defined.

(e) Our example in (b) shows that it fails the subspace test.

4. (a) [Optional (just part (a))]
Let $V = \{0\}$ consist of a single vector 0. Define addition and scalar multiplication by

$$0 + 0 = 0, \ \lambda 0 = 0 \ (\lambda \in \mathbb{R}),$$

respectively. Prove that, with these operations, V is a vector space over $\mathbb{R}$.
Your proof should not be long! I won't have this one graded, so you should be trying for the most concise proof that says everything that is necessary.

(b) Suppose V is a vector space over $\mathbb{R}$. Is the following statement true or false?

$$\text{If } \lambda, \mu \in \mathbb{R}, \ v \in V, \text{ and } \lambda v = \mu v, \text{ then } \lambda = \mu.$$

If not, you should be giving the most concise (but explicit) counter-example.

(c) Suppose V is a vector space over $\mathbb{R}$. Is the following statement true or false?

$$\text{If } \lambda \in \mathbb{R}, \ u, v \in V, \text{ and } \lambda u = \lambda v, \text{ then } u = v.$$

Help: your answer will depend on what V is.
This is reasonable since V was given to you before the statement. A similar thing happens in the following English excerpts because "it" means something different.
Have you tried Poke? It's delicious and healthy.
Have you tried Fat Sal's? It's delicious and healthy.

Solution:

(a) Here's the long proof...

Define addition and scalar multiplication on $\{0\}$ by $0 + 0 = 0$, $\lambda 0 = 0$ ($\lambda \in \mathbb{R}$), respectively. We verify the axioms of a vector space as follows:

i. Let $u, v \in \{0\}$. Then $u = v = 0$, so $u + v = 0 + 0 = v + u$.

ii. Let $u, v, w \in \{0\}$. Then $u = v = w = 0$, so that

$$(u + v) + w = (0 + 0) + 0 = 0 + 0 = 0 + (0 + 0) = u + (v + w),$$

where the second and third equalities are by definition of addition.

iii. We already have an element called 0. We must check that it has the correct property. Let $v \in \{0\}$. Then $v = 0$, so $v + 0 = 0 + 0 = 0 = v$, where the second equality is by definition of addition.

iv. Let $u \in \{0\}$ and $v = 0$. Then $v \in \{0\}$. Also, $u = 0$, so $u + v = 0 + 0 = 0$, where the second equality is by definition of addition.

v. Let $v \in \{0\}$. Then $v = 0$, so $1v = 1 \cdot 0 = 0 = v$, where the second equality is by definition of scalar multiplication.

vi. Let $\lambda, \mu \in \mathbb{R}$, $v \in \{0\}$. Then $v = 0$, so

$$(\lambda\mu)v = (\lambda\mu)0 = 0 = \lambda 0 = \lambda(\mu 0) = \lambda(\mu v),$$

where the second, third, and fourth equalities are by definition of scalar multiplication.

vii. Let $\lambda \in \mathbb{R}$, $u, v \in \{0\}$. Then $u = v = 0$, so

$$\lambda(u + v) = \lambda(0 + 0) = \lambda 0 = 0 = 0 + 0 = \lambda 0 + \lambda 0 = \lambda u + \lambda v,$$

where the second and fourth equalities are by definition of addition, and the third and fifth equalities are by definition of scalar multiplication.

viii. Let $\lambda, \mu \in \mathbb{R}$, $v \in \{0\}$. Then $v = 0$, so

$$(\lambda + \mu)v = (\lambda + \mu)0 = 0 = 0 + 0 = \lambda 0 + \mu 0 = \lambda v + \mu v,$$

where the second and fourth equalities are by definition of scalar multiplication, and the third is by definition of addition.

Here's the short proof...

Every element we think of, and every expression we can think of writing down is equal to 0. So the axioms all hold trivially.

(b) Suppose V is a vector space over $\mathbb{R}$. The statement is false because here is a counter-example: let $\lambda = 0$, $\mu = 1$, and $v = 0$; then $\lambda v = 0 = \mu v$, but $\lambda \neq \mu$.

(c) [2pts.] Suppose V is a vector space over $\mathbb{R}$, and $V \neq \{0\}$. Then the statement is false because here is a counter-example: let $\lambda = 0$, $u = 0$, $v \in V \setminus \{0\}$; then $\lambda u = 0 = \lambda v$, but $u \neq v$.

If $V = \{0\}$, then the statement is true since whenever $u, v \in V$, we have $u = 0 = v$.

5. Suppose X is a nonempty set. In class, I said that the set of real-valued functions from X, $\{f : X \longrightarrow \mathbb{R}\}$ is a vector space over $\mathbb{R}$ when equipped with pointwise addition and scalar multiplication. I checked axioms 2, 3, and 6.

   (a) Carefully check axioms 5, 7, and 8.
   I suppose you can check the others as well if you want more practice.

   (b) Taking $X = \{0, 1\}$ we see that $\{f : \{0, 1\} \longrightarrow \mathbb{R}\}$ is a vector space over $\mathbb{R}$ under pointwise addition and scalar multiplication. Let

   $$f, g, h : \{0, 1\} \longrightarrow \mathbb{R}$$

   be defined by $f(t) = 2t + 1$, $g(t) = 1 + 4t - 2t^2$, and $h(t) = 5^t + 1$.
   Prove that $f = g$ and $f + g = h$.
   For the first, you need to remember (or sensibly decide) what it means for two functions to be equal. For the second, you need to use this definition again, as well as the definition of $f + g$ in terms of f and g.

Solution:

Suppose X is a nonempty set.

(a) [3 pts.] For axiom 5, suppose that $f : X \longrightarrow \mathbb{R}$ and $x \in X$. Then

$$(1f)(x) = 1(f(x)) = f(x)$$

where the first equality is by definition of scalar multiplication, and the second is because of how 1 multiplies with real numbers. Since $x \in X$ is arbitrary, we conclude that $1f = f$.
For axiom 7, suppose $\lambda \in \mathbb{R}$, $f : X \longrightarrow \mathbb{R}$, $g : X \longrightarrow \mathbb{R}$, and $x \in X$. Then

$$\begin{aligned}(\lambda(f + g))(x) &= \lambda((f + g)(x)) = \lambda(f(x) + g(x)) \\ &= \lambda(f(x)) + \lambda(g(x)) = (\lambda f)(x) + (\lambda g)(x) = (\lambda f + \lambda g)(x),\end{aligned}$$

where the first and fourth equalities are by definition of scalar multiplication, the second and fifth equalities are by definition of addition, and the third is by the distributivity of the real numbers' addition and multiplication. Since $x \in X$ was arbitrary, we conclude that $\lambda(f + g) = \lambda f + \lambda g$.
For axiom 8, suppose $\lambda, \mu \in \mathbb{R}$, $f : X \longrightarrow \mathbb{R}$, and $x \in X$. Then

$$((\lambda + \mu)f)(x) = (\lambda + \mu)(f(x)) = \lambda(f(x)) + \mu(f(x)) = (\lambda f)(x) + (\mu f)(x) = (\lambda f + \mu f)(x),$$

where the first and third equalities are by definition of scalar multiplication, and the second is by the distributivity of the real numbers' addition and multiplication, and the fourth is by definition of addition. Because $x \in X$ was arbitrary, we can conclude that $(\lambda + \mu)f = \lambda f + \mu f$.

(b) Let X, f, g, and h, be as in the question.
Then $f(0) = 1 = g(0)$ and $f(1) = 3 = g(1)$, so $f = g$.
Moreover, $(f + g)(0) = f(0) + g(0) = 1 + 1 = 2 = h(0)$ and $(f + g)(1) = f(1) + g(1) = 3 + 3 = 6 = h(1)$, so that $f + g = h$.

6. (a) Recall theorem 7.1.4. It says that in order to check a subset W of a vector space V over $\mathbb{R}$ is a subspace, we just have to check that the following three conditions hold.

   i. $0 \in W$.
   ii. If $w \in W$ and $w' \in W$, then $w + w' \in W$.
   iii. If $\lambda \in \mathbb{R}$ and $w \in W$, then $\lambda w \in W$.

   Why is the first condition necessary; doesn't it follow iii. and the fact that $0w = 0$?

   (b) Suppose V is a vector space over $\mathbb{R}$ with operations $+ : V \times V \to V$, $\cdot : \mathbb{R} \times V \to V$, zero element 0, and that W is a subset of V. Prove that W is a subspace of V if and only if the following two conditions hold:

   i. $0 \in W$.
   ii. If $\lambda \in \mathbb{R}$, $w \in W$, and $w' \in W$, then $\lambda w + w' \in W$.

   I would call the conditions in (a), ai, aii, aiii, and the conditions in (b), bi, bii. Then I would go about proving that (ai, aii, and aiii hold) if and only if (bi and bii hold). So your proof should have two parts. One should begin with "suppose ai, aii, and aiii hold" and the other with "suppose bi and bii hold."

Solution:

(a) [1pt.] The empty set satisfies the second and third condition, but it is not a subspace because the empty set is not a vector space. This is because the empty set does not have a zero element (because it has no elements).

(b) [4 pts.] Suppose V is a vector space over $\mathbb{R}$ with operations $+ : V \times V \to V$, $\cdot : \mathbb{R} \times V \to V$, zero element 0, and that W is a subset of V. Consider the following conditions.

   i. $0 \in W$.
   ii. If $w \in W$ and $w' \in W$, then $w + w' \in W$.
   iii. If $\lambda \in \mathbb{R}$ and $w \in W$, then $\lambda w \in W$.
   iv. If $\lambda \in \mathbb{R}$, $w \in W$, and $w' \in W$, then $\lambda w + w' \in W$.

   Using the subspace test (theorem 7.3), it is enough to show that

$$i, ii, iii \iff i, iv.$$

   Suppose i, ii, iii. We need to show i, iv. i is trivial because we assumed it! To show iv, let $\lambda \in \mathbb{F}$, $w, w' \in W$. Because condition iii holds $\lambda w \in W$. Because condition ii holds, $\lambda w + w' \in W$.

   Suppose i, iv. We need to show i, ii, iii. i is trivial because we assumed it!

   To show ii, let $w, w' \in W$. Since $1 \in \mathbb{R}$, and condition iv holds, $w + w' = 1w + w' \in W$.

   To show iii, let $\lambda \in \mathbb{R}$ and $w \in W$. Because i holds, we have $0 \in W$. Because iv holds, we have $\lambda w = \lambda w + 0 \in W$.

7. (a) Let X be a nonempty subset of $\mathbb{R}$. Recall that the set of real-valued functions from X, $\mathcal{F} = \{f : X \longrightarrow \mathbb{R}\}$ is a vector space over $\mathbb{R}$ when equipped with pointwise addition and scalar multiplication.

Fix an $x_0 \in X$.

i. Prove that $\mathrm{ev}_{x_0} : \mathcal{F} \longrightarrow \mathbb{R}$, $f \longmapsto f(x_0)$ is linear.

ii. Prove that $\mathrm{Kev}_{x_0} = \{f \in \mathcal{F} : \ f(x_0) = 0\}$ is a subspace of $\mathcal{F}$.
(“Kev” does not stand for Kevin. Can you figure out what it stands for?)

(b) Taking $X = \mathbb{R}$ we see that the real-valued functions on the real line $V = \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$ form a vector space over $\mathbb{R}$ under pointwise addition and scalar multiplication. Consider the subset of *even functions*

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

Prove that E is a subspace of V.

Solution:

(a) [1pt. for completion]

i. Let $f, g \in \mathcal{F}$. Then

$$\mathrm{ev}_{x_0}(f + g) = (f + g)(x_0) = f(x_0) + g(x_0) = \mathrm{ev}_{x_0}(f) + \mathrm{ev}_{x_0}(g),$$

where the first and last equality are by definition of ev_{x_0}, and the middle equality is by definition of addition in $\mathcal{F}$.
Let $\lambda \in \mathbb{R}$, $f \in \mathcal{F}$. Then

$$\mathrm{ev}_{x_0}(\lambda f) = (\lambda f)(x_0) = \lambda(f(x_0)) = \lambda \cdot \mathrm{ev}_{x_0}(f),$$

where the first and last equality are by definition of ev_{x_0}, and the middle equality is by definition of scalar multiplication in $\mathcal{F}$.
We have shown that ev_{x_0} is linear.

ii. Here is the slow way...
Let X be a nonempty subset of $\mathbb{R}$, $\mathcal{F} = \{f : X \longrightarrow \mathbb{R}\}$.
Fix an $x_0 \in X$.
We wish to show that $\mathrm{Kev}_{x_0} = \{f : X \longrightarrow \mathbb{R} : \ f(x_0) = 0\}$ is a subspace of $\mathcal{F}$.
We use the subspace test.
Notice $0 \in \mathrm{Kev}_{x_0}$ since $0(x_0) = 0$.
Let $f, g \in \mathrm{Kev}_{x_0}$. Then $(f + g)(x_0) = f(x_0) + g(x_0) = 0 + 0 = 0$, where the first equality is by definition of addition in $\mathcal{F}$, and the second is because $f, g \in \mathrm{Kev}_{x_0}$.
We conclude that $f + g \in \mathrm{Kev}_{x_0}$.
Let $\lambda \in \mathbb{R}$, $f \in \mathrm{Kev}_{x_0}$. Then $(\lambda f)(x_0) = \lambda(f(x_0)) = \lambda 0 = 0$, where the first equality is by definition of scalar multiplication in $\mathcal{F}$, and the second is because $f \in \mathrm{Kev}_{x_0}$.
We conclude that $\lambda f \in \mathrm{Kev}_{x_0}$.

Here is the fast way: $\mathrm{Kev}_{x_0} = \ker(\mathrm{ev}_{x_0})$, and every kernel is a subspace.

(b) [3pts.] We wish to show that the subset of even functions

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

form a subspace of $V = \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$. We use the subspace test.
Let $t \in \mathbb{R}$. Then $0(-t) = 0 = 0(t)$, so $0 \in E$.
Let $f, g \in E$, and $t \in \mathbb{R}$. Then $(f+g)(-t) = f(-t) + g(-t) = f(t) + g(t) = (f+g)(t)$. Here, the first and last equality are by definition of addition, and the middle equality is because $f, g \in E$. We conclude that $f + g \in E$.
Let $\lambda \in \mathbb{R}$, $f \in E$, and $t \in \mathbb{R}$. Then $(\lambda f)(-t) = \lambda(f(-t)) = \lambda(f(t)) = (\lambda f)(t)$. Here, the first and last equality are by definition of scalar multiplication, and the middle equality is because $f \in E$. We conclude that $\lambda f \in E$.

8. [Optional]

(a) Suppose $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is linear, $T(1,0) = (1,4)$, and $T(1,1) = (2,4)$. What is $T(2,3)$?

(b) Prove that there exists a linear transformation $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ such that $T(1,1) = (1,0,3)$ and $T(2,3) = (1,-2,4)$. What is $T(8,10)$?

Solution:

(a) $T(2,3) = T(3(1,1) - (1,0)) = 3 \cdot T(1,1) - T(1,0) = 3(2,4) - (1,4) = (5,8)$.

(b) Notice that

$$T(1,0) = T(3(1,1) - (2,3)) = 3 \cdot T(1,1) - T(2,3) = 3(1,0,3) - (1,-2,4) = (2,2,5) \quad \text{and}$$
$$T(0,1) = T((2,3) - 2(1,1)) = T(2,3) - 2 \cdot T(1,1) = (1,-2,4) - 2(1,0,3) = (-1,-2,-2).$$

Also,

$$\begin{pmatrix} 2 & -1 \\ 2 & -2 \\ 5 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 \\ 2x_2 - 2x_2 \\ 5x_1 - 2x_2 \end{pmatrix}.$$

So let $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ be defined by $T(x_1, x_2) = (2x_1 - x_2, 2x_1 - 2x_2, 5x_1 - 2x_2)$.
Then $T(1,1) = (1,0,3)$, $T(2,3) = (1,-2,4)$, and $T(8,10) = (6,-4,20)$.

9. Prove that each of the following functions $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is *not* linear as efficiently as possible.

(a) $T(x_1, x_2) = (1, x_1)$,

(b) $T(x_1, x_2) = (x_1, x_1^3)$,

(c) $T(x_1, x_2) = (\sin x_1, 0)$,

(d) $T(x_1, x_2) = (|x_1|, 2x_2)$.

Solution: [1pt. for completion]

(a) $T(0,0) = (1,0) \neq (0,0)$.

(b) $T(2(1,0)) = T(2,0) = (2,8) \neq 2(1,1) = 2 \cdot T(1,0)$.

(c) $T(2(\pi/2, 0)) = T(\pi, 0) = (0,0) \neq 2(1,0) = 2 \cdot T(\pi/2, 0)$.

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

9 Lecture on August 13th: Kernels, images, linear transformations from $\mathbb{R}^n$, matrices

9.1 Kernels and images

Definition 9.1.1. Let V and W be vector spaces over $\mathbb{R}$ and let $T : V \longrightarrow W$ be linear.

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

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

Remark 9.1.2. Let V and W be vector spaces over $\mathbb{R}$ and let $T : V \longrightarrow W$ be a linear transformation. Immediately from the definition, we see that $\ker T \subseteq V$ and $\operatorname{im} T \subseteq W$.

Theorem 9.1.3. *Let V and W be vector spaces over $\mathbb{R}$ 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 9.1.4. *Let V and W be vector spaces over $\mathbb{R}$ and let $T : V \longrightarrow W$ be a linear transformation. Then T is injective if and only if $\ker T = \{0\}$.*

Proof. Let V and W be vector spaces over $\mathbb{R}$ and let $T : V \longrightarrow W$ be a linear transformation.

First, suppose that T is injective. We must show that $\ker T = \{0\}$. Since $\ker T$ is a subspace, we have $\{0\} \subseteq \ker T$. To show $\ker T \subseteq \{0\}$ suppose $v \in \ker T$. Then $T(v) = 0 = T(0)$. Since T is injective, this gives $v = 0$, so $v \in \{0\}$.

Conversely, suppose that $\ker T = \{0\}$. We must show T is injective. So suppose $v_1, v_2 \in V$ and $T(v_1) = T(v_2)$. Using linearity of T, we get $T(v_1 - v_2) = T(v_1) - T(v_2) = 0$. Thus, $v_1 - v_2 \in \ker T$. Since we assumed $\ker T = \{0\}$, this gives $v_1 - v_2 = 0$, i.e. $v_1 = v_2$. $\square$

Remark 9.1.5. This simple fact is one of the most important in the class.

Remark 9.1.6. Let V and W be vector spaces over $\mathbb{R}$ and let $T : V \longrightarrow W$ be a linear transformation. Then T is surjective if and only if $\operatorname{im} T = W$.

Example 9.1.7. Define $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ by $T(x_1, x_2, x_3) = (x_1 + x_2, -x_1 - x_2 + x_3, -x_3)$.

Suppose $(x_1, x_2, x_3) \in \ker T$. Then $T(x_1, x_2, x_3) = (0, 0, 0)$,

$$\text{i.e. } (x_1 + x_2, -x_1 - x_2 + x_3, -x_3) = (0, 0, 0).$$

This gives $x_1 = -x_2$ and $x_3 = 0$, which leads us to conjecture that $\ker T = \{(x, -x, 0) : \ x \in \mathbb{R}\}$. The argument just given shows "$\subseteq$." For any $x \in \mathbb{R}$, $T(x, -x, 0) = (0, 0, 0)$, which shows "$\supseteq$."

We conjecture that $\operatorname{im} T$ is the plane

$$\{(y_1, y_2, y_3) \in \mathbb{R}^3 : \ y_1 + y_2 + y_3 = 0\}.$$

Given $(y_1, y_2, y_3) \in \operatorname{im} T$, we can find an $(x_1, x_2, x_3) \in \mathbb{R}^3$ such that $(y_1, y_2, y_3) = T(x_1, x_2, x_3)$,

$$\text{i.e. } (y_1, y_2, y_3) = (x_1 + x_2, -x_1 - x_2 + x_3, -x_3).$$

One can check $y_1 + y_2 + y_3 = 0$, which shows "$\subseteq$."

Conversely, given $(y_1, y_2, y_3) \in \mathbb{R}^3$, with $y_1 + y_2 + y_3 = 0$, let $(x_1, x_2, x_3) = (y_1, 0, -y_3)$. Then $T(x_1, x_2, x_3) = (x_1 + x_2, -x_1 - x_2 + x_3, -x_3) = (y_1 + 0, -y_1 - 0 - y_3, y_3) = (y_1, -y_1 - y_3, y_3)$. Since $y_1 + y_2 + y_3 = 0$, we have $y_2 = -y_1 - y_3$, and $(y_1, y_2, y_3) = (y_1, -y_1 - y_3, y_3) = T(x_1, x_2, x_3) \in \operatorname{im} T$. Thus, we have "$\supseteq$."

Remark 9.1.8. Such calculations will be far quicker once we have the rank-nullity theorem at our disposal.

9.2 Linear transformations from $\mathbb{R}^n$

Definition 9.2.1. Suppose V is a vector space over $\mathbb{R}$ and $n \in \mathbb{N}$. By an *n-tuple of vectors in V*, we mean an element of the Cartesian product V^n, $(v_1, v_2, \ldots, v_n)$.

We take as convention that there is one 0-tuple, the *empty tuple* ( ).

We'll use the following notation throughout the rest of the class, so don't ignore this.

Notation 9.2.2. Suppose V is a vector space over $\mathbb{R}$ and $n \in \mathbb{N}$, and $\alpha = (v_1, \ldots, v_n)$ is an n-tuple of vectors in V. We write $\Gamma_\alpha : \mathbb{R}^n \longrightarrow V$ for the function

$$(\lambda_1, \lambda_2, \ldots, \lambda_n) \longmapsto \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

Remark 9.2.3. In case you ever worry about whether $n = 0$ is allowed, it is a sensible convention to take $\mathbb{R}^0 = \{0\}$. Moreover, if α is the empty tuple ( ), $\Gamma_\alpha : \mathbb{R}^0 \longrightarrow V$ is the inclusion of 0 into V.

Theorem 9.2.4. *Suppose V is a vector space over $\mathbb{R}$, and $\alpha = (v_1, \ldots, v_n)$ is an n-tuple of vectors in V. The function $\Gamma_\alpha : \mathbb{R}^n \longrightarrow V$ is linear.*

Remark 9.2.5. Linear transformations $\Gamma_\alpha : \mathbb{R}^n \longrightarrow V$ where α is an n-tuple of vectors in V will be *very* important for us. This is because these type of transformations clarify almost every concept that we study in the remainder of the class. Moreover, every linear transformation out of $\mathbb{R}^n$ is of this form.

Theorem 9.2.6. *Suppose V is a vector space over $\mathbb{R}$ and $T : \mathbb{R}^n \to V$ is a linear transformation. For $j \in \{1, \ldots, n\}$, let $v_j = T(e_j)$ where $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ with the 1 in the j-th position. Let $\alpha = (v_1, \ldots, v_n)$. Then $T = \Gamma_\alpha$.*

Proof. Let everything be as in the theorem statement. Then for all $(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{R}^n$, we have

$$\begin{aligned} T(\lambda_1, \lambda_2, \ldots, \lambda_n) &= T(\lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n) \\ &= \lambda_1 T(e_1) + \lambda_2 T(e_2) + \ldots + \lambda_n T(e_n) \\ &= \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n \\ &= \Gamma_\alpha(\lambda_1, \lambda_2, \ldots, \lambda_n). \end{aligned}$$

□

Remark 9.2.7. Recall remark 7.2.4. Suppose $v_1, v_2, \ldots, v_n \in \mathbb{R}^m$ and A is the $m \times n$ matrix with j-th column given by v_j. Recall example 7.2.3 which tells us that $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $x \longmapsto Ax$ is a linear transformation. We also have $\Gamma_{(v_1, v_2, \ldots, v_n)} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $(\lambda_1, \lambda_2, \ldots, \lambda_n) \longmapsto \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n$. The definition of matrix-vector multiplication ensures that $T_A = \Gamma_{(v_1, v_2, \ldots, v_n)}$.

Corollary 9.2.8. *Suppose $m, n \in \mathbb{N}$ and $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Then there is a unique matrix $A \in M_{m\times n}(\mathbb{R})$ such that $T = T_A$.*

Proof. We use the previous theorem and remark.

For $j \in \{1, \dots, n\}$, let $v_j = T(e_j)$ where $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ with the 1 in the j-th position, let $\alpha = (v_1, \dots, v_n)$, and let A be the $m \times n$ matrix with j-th column given by v_j. The theorem says that $T = \Gamma_\alpha$. The remark gives $T_A = \Gamma_\alpha$. So $T = T_A$. This completes the existence proof.

For uniqueness, suppose $A, B \in M_{m\times n}(\mathbb{R})$ and $T_A = T_B$. For each $j \in \{1, \dots, n\}$, we have

$$j\text{-th column of } A = Ae_j = T_A(e_j) = T_B(e_j) = Be_j = j\text{-th column of } B.$$

Thus, $A = B$. $\square$

9.3 Basic matrix operations (in case you have forgotten)

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

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1n} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2n} \\ A_{31} & A_{32} & A_{33} & \cdots & A_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & A_{m3} & \cdots & A_{mn} \end{pmatrix}$$

We can add two $m \times n$ matrices entrywise, so that $[A+B]_{ij} = A_{ij} + B_{ij}$. We can also multiply by a scalar, so that $[\lambda A]_{ij} = \lambda A_{ij}$.

Given an $m \times n$ matrix A, and an $n \times p$ matrix B, we can multiply A and B to obtain an $m \times p$ matrix AB. There is a concise formula for the (i, j)-entry of AB, in terms of the entries of A and B:

$$[AB]_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj} = A_{i1}B_{1j} + A_{i2}B_{2j} + \ldots + A_{in}B_{nj}.$$

One way to think of this is as 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. The language we have developed can be used to fix this.

9.4 How to think about matrices the "right" way

The previous corollary says that linear transformations $\mathbb{R}^n \to \mathbb{R}^m$ are the "same" as $m \times n$ matrices.

In 33A, you spend all quarter thinking about and using matrices. Matrices are arrays of numbers which can often be overwhelming. Although they are computationally useful, for conceptual purposes it is often better to think about the linear transformations that they define: instead of thinking about the matrix A, think about the linear transformation T_A. For example, do you prefer thinking about a special orthogonal matrix or a rotation? A rotation seems far more reasonable to me! The difference between thinking about a matrix versus a linear transformation is somewhat philosophical since one determines the other, but while matrices just sit there being all ugly, linear transformations do something: they move vectors to new vectors.

As before, let $e_j = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^n$ with the 1 in the j-th position. While proving the previous corollary, we used the following important fact.

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

Remark 9.4.1. We can now summarize *the two most important properties of $m \times n$ matrices.*

- An $m \times n$ matrix A does something: via the linear transformation T_A, 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 under T_A.

Remark 9.4.2. The importance of these two bullet points cannot be overstated. If you ever had a moment in 33A where your TA could see something about a matrix was true and it appeared to be black magic on their part, it was probably just that your TA knew these two facts.

Because of the last bullet point, 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:

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

Then matrix-vector multiplication can be described as follows:

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

Notice that this is precisely $\Gamma_{(v_1, v_2, \ldots, v_n)}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. 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.

As long as we're thinking about $m \times n$ matrices in terms of their linear transformations, 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

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

Then

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

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

Example 9.4.3. I gave examples of "reflect across the x-axis" and "project onto the xy-plane" in class.

10 Questions due on August 15th

1. Suppose V is a vector space over $\mathbb{R}$, and that W_1 and W_2 are subspaces of V.

   Prove that $W_1 \cup W_2$ is a subspace of V if and only if $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

   Hint: For the forwards implication $\implies$, *I would prove the contrapositive which is*

   $$\text{If } W_1 \not\subseteq W_2 \text{ and } W_2 \not\subseteq W_1, \text{ then } W_1 \cup W_2 \text{ is not a subspace of } V.$$

   Solution:

   Suppose V is a vector space over $\mathbb{R}$, and that W_1 and W_2 are subspaces of V.

   First, suppose $W_1 \subseteq W_2$. Then $W_1 \cup W_2 = W_2$, and so $W_1 \cup W_2$ is a subspace.

   Second, suppose $W_2 \subseteq W_1$. Then $W_1 \cup W_2 = W_1$, and so $W_1 \cup W_2$ is a subspace.

   Now suppose that $W_1 \not\subseteq W_2$ and $W_2 \not\subseteq W_1$. This means we can choose elements w_1 and w_2 such that $w_1 \in W_1 \setminus W_2$ and $w_2 \in W_2 \setminus W_1$. We have $w_1, w_2 \in W_1 \cup W_2$, and we will show that $W_1 \cup W_2$ is not a subspace by showing that $w_1 + w_2 \notin W_1 \cup W_2$. Suppose for contradiction that $w_1 + w_2 \in W_1 \cup W_2$.

   (a) Case 1: $w_1 + w_2 \in W_1$. Since $w_1 \in W_1$ and W_1 is a subspace, $-w_1 \in W_1$, and so

   $$w_2 = (w_1 + w_2) + (-w_1) \in W_1.$$

   This is a contradiction to how we chose w_2.

   (b) Case 2: $w_1 + w_2 \in W_2$. Since $w_2 \in W_2$ and W_2 is a subspace, $-w_2 \in W_2$, and so

   $$w_1 = (w_1 + w_2) + (-w_2) \in W_2.$$

   This is a contradiction to how we chose w_1.

2. Prove lemma 7.2.8.

   On the one hand, I think this is very easy, and should feel very boring if you understand what is going on.

   On the other hand, if you don't say the phrases "suppose $T : U \to V$ and $S : V \to W$ are linear", "by definition of ST" a few times, "since T is linear," and "since S is linear," then you don't know what you're doing, and you need to talk to me, Kevin, or a friend who does know what is happening.

   Solution:

   Suppose U, V, and W are vector spaces over $\mathbb{R}$, that $T : U \to V$ and $S : V \to W$ are linear transformations. We wish to show $ST : U \longrightarrow W$ is a linear transformation.

   Let $u_1, u_2 \in U$. Then

   $$\begin{aligned}(ST)(u_1 + u_2) = S(T(u_1 + u_2)) &= S(T(u_1) + T(u_2)) \\ &= S(T(u_1)) + S(T(u_2)) = (ST)(u_1) + (ST)(u_2).\end{aligned}$$

   Here the first and last equalities are by definition of ST, the second is by linearity of T, and the third is by linearity of S.

   Similarly, you can check that for $\lambda \in \mathbb{R}$ and $u \in U$, $(ST)(\lambda u) = \lambda(ST)(u)$. Thus, ST is linear.

3. [More difficult, but not optional]

Let V be the vector space $\{f : \mathbb{R} \longrightarrow \mathbb{R}\}$ of real-valued functions on the real line.

Let $f_1, f_2, f_3 \in V$ be defined by $f_1(t) = 1,\ f_2(t) = \cos^2 t,\ f_3(t) = \sin^2(t)$.

Let α be the 2-tuple (f_1, f_2) and β be the 3-tuple (f_1, f_2, f_3). Look back on 9.2.2.

We have the linear transformations $\Gamma_\alpha : \mathbb{R}^2 \longrightarrow V$, $\Gamma_\beta : \mathbb{R}^3 \longrightarrow V$.

(a) Prove that $\ker \Gamma_\alpha = \{0\}$. (So Γ_α is injective!)
Showing "$\subseteq$" is the harder direction.
"Suppose $(\lambda_1, \lambda_2) \in \ker \Gamma_\alpha$. We want to show $(\lambda_1, \lambda_2) = (0, 0) \ldots$"
You might find the linear transformations $\mathrm{ev}_0 : V \longrightarrow \mathbb{R}$, $\mathrm{ev}_{\frac{\pi}{2}} : V \longrightarrow \mathbb{R}$ useful.

(b) Find an element of $(\lambda_1, \lambda_2, \lambda_3) \in (\ker \Gamma_\beta) \setminus \{0\}$.

(c) Prove that $\ker \Gamma_\beta = \{\mu(\lambda_1, \lambda_2, \lambda_3) :\ \mu \in \mathbb{R}\}$.

Solution:

(a) $\ker \Gamma_\alpha$ is a subspace of $\mathbb{R}^2$ so $\{0\} \subseteq \ker \Gamma_\alpha$.
We must show $\ker \Gamma_\alpha \subseteq \{0\}$. So suppose $(\lambda_1, \lambda_2) \in \ker \Gamma_\alpha$. Then

$$\lambda_1 f_1 + \lambda_2 f_2 = \Gamma_\alpha(\lambda_1, \lambda_2) = 0.$$

This equation expresses an equality of functions. In particular, 0 means the 0-function.
We apply ev_0 and $\mathrm{ev}_{\frac{\pi}{2}}$ to obtain useful equations:

$$\lambda_1 + \lambda_2 = 0 \text{ and } \lambda_1 = 0.$$

These imply $(\lambda_1, \lambda_2) = (0, 0)$, as required.

(b) $\Gamma_\beta(-1, 1, 1) = -f_1 + f_2 + f_3$. Evaluating this function at $t \in \mathbb{R}$ gives

$$-1 + \cos^2 t + \sin^2 t = 0.$$

Thus, $\Gamma_\beta(-1, 1, 1)$ is the zero function and $(-1, 1, 1) \in \ker \Gamma_\beta \setminus \{0\}$.

(c) We wish to show $\ker \Gamma_\beta = \{\mu(-1, 1, 1) :\ \mu \in \mathbb{R}\}$.
"$\supseteq$" is clear since $(-1, 1, 1) \in \ker \Gamma_\beta$ and $\ker \Gamma_\beta$ is a subspace of $\mathbb{R}^3$.
For "$\subseteq$," suppose $(\lambda_1, \lambda_2, \lambda_3) \in \ker \Gamma_\beta$.
Then

$$\lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3 = \Gamma_\beta(\lambda_1, \lambda_2, \lambda_3) = 0.$$

This equation expresses an equality of functions. In particular, 0 means the 0-function.
We apply ev_0 and $\mathrm{ev}_{\frac{\pi}{2}}$ to obtain useful equations:

$$\lambda_1 + \lambda_2 = 0 \text{ and } \lambda_1 + \lambda_3 = 0.$$

Thus, $(\lambda_1, \lambda_2, \lambda_3) = (\lambda_1, -\lambda_1, -\lambda_1) = -\lambda_1(-1, 1, 1) \in \{\mu(-1, 1, 1) :\ \mu \in \mathbb{R}\}$.

4. [Optional]

   Corollary 9.2.8 together with remark 9.4.1 makes question 8.(b) in the last lot of questions much easier. Revisit that question.

   Solution:

   A solution to last week's 8.(b) has been posted.

Remark. Kevin was concerned that he may have confused some people in discussion.
There are a few things he wanted me to emphasize.

1. A linear transformation $T : W \longrightarrow V$ can only be of the form Γ_α if $W = \mathbb{R}^n$ for some n.

2. A linear transformation $T : W \longrightarrow V$ can only be of the form T_A (for some matrix A) when $W = \mathbb{R}^n$ for some n *and* $V = \mathbb{R}^m$ for some m.

3. Eventually, we will address how matrices show up from linear transformations $T : W \longrightarrow V$ between abstract vector spaces. This will require picking *bases* for W and V. So far, we have not spoken about bases and none of the definitions I have made have used this concept.

   Be patient! Make sure that you understand the definitions given so far.

11 Lecture on August 15th: Spans and linear (in)dependence

11.1 Linear combinations and spans of tuples

Recall that whenever we have a vector space V over $\mathbb{R}$ and an n-tuple $\alpha = (v_1, \ldots, v_n)$ of vectors in V, we have a linear transformation $\Gamma_\alpha : \mathbb{R}^n \longrightarrow V$ defined by

$$(\lambda_1, \lambda_2, \ldots, \lambda_n) \longmapsto \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

In this section, we consider the image of such transformations and the condition for such a transformation to be surjective.

Definition 11.1.1. Suppose V is a vector space over $\mathbb{R}$ and $(v_1, v_2, \ldots, v_n)$ is an n-tuple of vectors in V. An element $v \in V$ is said to be *a linear combination of* $(v_1, v_2, \ldots, v_n)$ **iff** there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ 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.

Remark 11.1.2. It is a sensible convention to regard 0 as the only linear combination of the empty tuple.

Example 11.1.3.

1. In $\mathbb{R}^3$, $(1, 8, 1)$ is a linear combination of $((1, 2, 1),\ (0, \pi, 0))$.
2. In $\mathbb{R}^3$, $(0, 0, 1)$ is not a linear combination of $((1, 1, 0),\ (0, 1, 1))$.

Definition 11.1.4. Suppose V is a vector space over $\mathbb{R}$ and $(v_1, v_2, \ldots, v_n)$ is an n-tuple of vectors in V. The *span* of $(v_1, v_2, \ldots, v_n)$ is the set

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

Remark 11.1.5. It is a sensible convention to take the span of the empty tuple to be $\{0\}$.

Remark 11.1.6. Suppose V is a vector space over $\mathbb{R}$ and $(v_1, v_2, \ldots, v_n)$ is an n-tuple of vectors in V. It is immediate from the previous two definitions that an element $v \in V$ is a linear combination of $(v_1, v_2, \ldots, v_n)$ if and only if $v \in \operatorname{span}(v_1, v_2, \ldots, v_n)$.

Definition 11.1.7. Suppose V is a vector space over $\mathbb{R}$ and $(v_1, v_2, \ldots, v_n)$ is an n-tuple of vectors in V. We say that $(v_1, v_2, \ldots, v_n)$ *spans* V **iff** every $v \in V$ is a linear combination of $(v_1, v_2, \ldots, v_n)$.

Remark 11.1.8. Suppose V is a vector space over $\mathbb{R}$ and $(v_1, v_2, \ldots, v_n)$ is an n-tuple of vectors in V. It is immediate from the previous two definitions that $(v_1, v_2, \ldots, v_n)$ spans V if and only if $\operatorname{span}(v_1, v_2, \ldots, v_n) = V$.

Example 11.1.9.

1. Let $n \in \mathbb{N}$. The tuple $(e_1, e_2, \ldots, e_n)$ spans $\mathbb{R}^n$.
2. The tuple $(e_1, e_1 + e_2, e_2)$ spans $\mathbb{R}^2$.
3. Let $n \in \mathbb{N}$. The tuple $(1, x, x^2, \ldots, x^n)$ spans $\mathcal{P}_n(\mathbb{R})$.

Remark 11.1.10. Suppose V is a vector space over $\mathbb{R}$ and $(v_1, v_2, \ldots, v_n)$ is an n-tuple of vectors in V. First, notice that $\operatorname{span}(v_1, v_2, \ldots, v_n)$ is precisely the image of $\Gamma_{(v_1, v_2, \ldots, v_n)}$. This tells us that $\operatorname{span}(v_1, v_2, \ldots, v_n)$ is a subspace of V. Moreover, notice that $(v_1, v_2, \ldots, v_n)$ spans V if and only if $\Gamma_{(v_1, v_2, \ldots, v_n)}$ is surjective.

11.2 Linear combinations and spans of arbitrary sets (omitted)

This section is not necessary for the main results of the class. This is because we focus on *finite*-dimensional vector spaces.

Definition 11.2.1. Suppose V is a vector space over $\mathbb{R}$ and that $\emptyset \neq S \subseteq V$.
An element $v \in V$ is said to be *a linear combination of vectors in S* **iff** there exists an n-tuple of vectors $(v_1, v_2, \ldots, v_n)$ in S such that v is a linear combination of $(v_1, v_2, \ldots, v_n)$.

Definition 11.2.2. Suppose V is a vector space over $\mathbb{R}$ and that $\emptyset \neq S \subseteq V$.
The *span of S* is the set

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

Our conventions regarding the empty tuple mean that $\mathrm{span}(\emptyset) = \{0\}$.

Definition 11.2.3. Suppose V is a vector space over $\mathbb{R}$, and that $S \subseteq V$.
We say that S *spans* V **iff** $\mathrm{span}(S) = V$.

Theorem 11.2.4. *Suppose V is a vector space over $\mathbb{R}$ and $S \subseteq V$.*
Then span(S) is the smallest subspace of V which contains S.

11.3 Linear dependence and linear independence

Recall that whenever we have a vector space V over $\mathbb{R}$ and an n-tuple $\alpha = (v_1, \ldots, v_n)$ of vectors in V, we have a linear transformation $\Gamma_\alpha : \mathbb{R}^n \longrightarrow V$ defined by

$$(\lambda_1, \lambda_2, \ldots, \lambda_n) \longmapsto \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

In this section, we consider the kernel of such transformations and the condition for such a transformation to be injective.

Definition 11.3.1. Suppose V is a vector space over $\mathbb{R}$ and $(v_1, v_2, \ldots, v_n)$ is an n-tuple of vectors in V. An equation

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0$$

with $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ is called *a dependency relation for* $(v_1, v_2, \ldots, v_n)$.
The equation $0v_1 + 0v_2 + \ldots + 0v_n = 0$ is called the *trivial* dependency relation for $(v_1, v_2, \ldots, v_n)$. Other dependency relations are called *non-trivial*.

Definition 11.3.2. Suppose V is a vector space over $\mathbb{R}$ and $(v_1, v_2, \ldots, v_n)$ is an n-tuple of vectors in V. $(v_1, v_2, \ldots, v_n)$ is said to be *linearly dependent* **iff** there exists a non-trivial dependency relation for $(v_1, v_2, \ldots, v_n)$. $(v_1, v_2, \ldots, v_n)$ is said to be *linearly independent* **iff** the only dependency relation for $(v_1, v_2, \ldots, v_n)$ is the trivial one.

Remark 11.3.3. It is a sensible convention to regard the empty tuple as linearly independent.

Example 11.3.4.

1. Let $f_1, f_2, f_3 : \mathbb{R} \longrightarrow \mathbb{R}$ be defined by $f_1(t) = 1,\ f_2(t) = \cos^2 t,\ f_3(t) = \sin^2 t$.

   These are elements of the vector space $\mathcal{F} = \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$.

   (f_1, f_2, f_3) is linearly dependent. (f_1, f_2) is linearly independent.

2. The tuple $((1, -1, 2), (2, 0, 1), (-1, 2, -1))$ of vectors in $\mathbb{R}^3$ is linearly independent.

Remark 11.3.5. Suppose V is a vector space over $\mathbb{R}$ and $(v_1, v_2, \ldots, v_n)$ is an n-tuple of vectors in V. Notice that we have a dependency relation

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0$$

if and only if $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ in the kernel of $\Gamma_{(v_1, v_2, \ldots, v_n)}$.

The fact that a tuple $(v_1, v_2, \ldots, v_n)$ always has the trivial dependency relation is related to the fact that 0 is always an element of the kernel of a linear transformation.

Moreover, $(v_1, v_2, \ldots, v_n)$ is linearly independent if and only if $\ker \Gamma_{(v_1, v_2, \ldots, v_n)} = \{0\}$, and this latter condition is equivalent to the statement that $\Gamma_{(v_1, v_2, \ldots, v_n)}$ is injective.

Lemma 11.3.6. *Suppose V is a vector space over $\mathbb{R}$ 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 $\mathbb{R}$ 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 $\lambda \in \mathbb{R}$ such that $\lambda v = 0$. Thus, $v = \lambda^{-1}(\lambda v) = \lambda^{-1} 0 = 0$. $\square$

Corollary 11.3.7. *Suppose V is a vector space over $\mathbb{R}$ and $v \in V$.*

The 1-tuple (v) is linearly independent if and only if $v \neq 0$.

Example 11.3.8. Suppose V is a vector space over $\mathbb{R}$ and that $v \in V$. The 2-tuple (v, v) is linearly dependent. This is because $1v + (-1)v = 0$ is a non-trivial dependency relation.

Remark 11.3.9. My definition differs 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*

$$\Big\{(1,0),\ (1,0)\Big\} = \Big\{(1,0)\Big\}$$

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

$$\Big((1,0),\ (1,0)\Big)$$

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 13.1.1 and theorem 13.2.5 are easier than the corresponding proofs in the textbook, those of theorems 1.7 and 1.9, largely because of these choices of definitions.

As a final piece of propaganda: "there are at least two exercises in the textbook that are wrong because the authors chose their definition incorrectly."

Theorem 11.3.10. *Suppose V is a vector space over $\mathbb{R}$, 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.

12 Questions due on August 16th

1. (a) [Optional]

   Question 5 of page 34 from the textbook (but replace the word "set" by "tuple" and braces {} with parentheses ()).

   (b) Give an example a linearly dependent 3-tuple in $\mathbb{R}^3$ such that none of the 3 vectors is a multiple of another.

   Solution: $((1,0,0),\ (0,1,0),\ (1,1,0))$.

   (c) [Optional]

   Question 2(a)-(f) of page 40 of the textbook (but replace the word "set" by "tuple" and braces {} with parentheses ()).

   (d) Prove that

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

   is linearly dependent in $M_{3\times 2}(\mathbb{R})$.

   Solution: The following is a nontrivial dependecy relation:

$$1 \cdot \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} = 0$$

   (e) Prove that $\Big((1,1,0),\ (0,1,1),\ (1,0,1)\Big)$ is linearly independent in $\mathbb{R}^3$.

   Solution: Suppose we have a dependency relation

$$\lambda_1(1,1,0) + \lambda_2(0,1,1) + \lambda_3(1,0,1) = 0.$$

   We wish to show that this must be the trivial dependency relation, i.e. that $\lambda_1 = \lambda_2 = \lambda_3 = 0$. The above equation says $(\lambda_1 + \lambda_3,\ \lambda_1 + \lambda_2,\ \lambda_2 + \lambda_3) = 0$.

   So $\lambda_1 + \lambda_3 = 0$, $\lambda_1 + \lambda_2 = 0$, and $\lambda_2 + \lambda_3 = 0$. Thus,

$$\begin{aligned}
\lambda_1 &= \frac{+\ (\lambda_1 + \lambda_3) + (\lambda_1 + \lambda_2) - (\lambda_2 + \lambda_3)}{2} = \frac{0+0+0}{2} = 0, \\
\lambda_2 &= \frac{-\ (\lambda_1 + \lambda_3) + (\lambda_1 + \lambda_2) + (\lambda_2 + \lambda_3)}{2} = \frac{0+0+0}{2} = 0, \\
\lambda_3 &= \frac{+\ (\lambda_1 + \lambda_3) - (\lambda_1 + \lambda_2) + (\lambda_2 + \lambda_3)}{2} = \frac{0+0+0}{2} = 0.
\end{aligned}$$

13 Lecture on August 16th: Towards bases and dimension

13.1 Spans and linear (in)dependence

Theorem 13.1.1. *Suppose V is a vector space over $\mathbb{R}$, 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 $\mathbb{R}$, 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 $v_{n+1} \in \text{span}(v_1, \ldots, v_n)$. By definition of "span," there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ 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 of linear dependence, 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

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

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 interpretation it is just lemma 11.3.6 all over again, but you might prefer to divide the proof into two cases. (The textbook fails to point this out.) $\square$

13.2 Towards bases and dimension

Recall that whenever we have a vector space V over $\mathbb{R}$ and an n-tuple $\alpha = (v_1, \ldots, v_n)$ of vectors in V, we have a linear transformation $\Gamma_\alpha : \mathbb{R}^n \longrightarrow V$ defined by

$$(\lambda_1, \lambda_2, \ldots, \lambda_n) \longmapsto \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n.$$

In this section, we consider the condition under which such a transformation is a bjiection.

One can think of Γ_α as an attempt to coordinate-ify V using the tuple of vectors $(v_1, v_2, \ldots, v_n)$: a collection of coordinates $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ determines a vector $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n$.

First, we can ask that every vector have a coordinate. This is the same as demanding that Γ_α be surjective. We have seen that this, in turn, is the same as demanding that $(v_1, v_2, \ldots, v_n)$ spans V. Second, we can ask that every vector have at most one coordinate. This is the same as demanding that Γ_α be injective. We have seen, in turn, that this is the same as demanding that $(v_1, v_2, \ldots, v_n)$ be linearly independent. We find that every vector has exactly one coordinate when Γ_α is a bijection, and that this is the case when $(v_1, v_2, \ldots, v_n)$ spans V and is linearly independent.

Definition 13.2.1. Suppose V is a vector space over $\mathbb{R}$. 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 and
2. $(v_1, v_2, \ldots, v_n)$ is linearly independent.

Example 13.2.2. The empty tuple is a basis for $\{0\}$.

By our conventions, we have span( ) = $\{0\}$ and ( ) is linearly independent.

Theorem 13.2.3. *Suppose V is a vector space over $\mathbb{R}$. 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 $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that*

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

Proof. Suppose V is a vector space over $\mathbb{R}$. We have seen that an n-tuple $\alpha = (v_1, v_2, \ldots, v_n)$ of vectors in V is a basis for V if and only if Γ_α is a bijection. This latter statement is the same as saying that for all $v \in V$, there are unique $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that

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

You should try writing the proof out in more elementary terms using the definition of span and linear independence. $\square$

We would like to make the following definition.

Definition 13.2.4. Suppose V is a vector space over $\mathbb{R}$. We say V is *finite-dimensional* **iff** V is the span of a finite tuple; 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 13.2.5. *Suppose V is a vector space over $\mathbb{R}$ and $(v_1, v_2, \ldots, v_n)$ is an n-tuple of vectors in V which 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 and $(v_1, v_2, \ldots, v_n)$ are as in the theorem statment.

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 empty 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 11.3.7, 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 11.3.8). 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. By construction, we have that $(v_{m_1}, \ldots, v_{m_r})$ is linearly independent. So we are just left to show that $(v_{m_1}, \ldots, v_{m_r})$ spans V.

As a first step towards this, we show $\{v_1, v_2, \ldots, v_n\} \subseteq \operatorname{span}(v_{m_1}, \ldots, v_{m_r})$. With this goal, let $i \in \{1, \ldots, n\}$. We need $v_i \in \operatorname{span}(v_{m_1}, \ldots, v_{m_r})$. By theorem 13.1.1, 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 $(v_{m_1}, \ldots, v_{m_r})$.

Since $\operatorname{span}(v_{m_1}, \ldots, v_{m_r})$ is a subspace of V, from $\{v_1, \ldots, v_n\} \subseteq \operatorname{span}(v_{m_1}, \ldots, v_{m_r})$, we obtain

$$\operatorname{span}(v_1, v_2, \ldots, v_n) = \{\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n : \ \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}\} \subseteq \operatorname{span}(v_{m_1}, \ldots, v_{m_r}).$$

We supposed that $\operatorname{span}(v_1, v_2, \ldots, v_n) = V$, so this shows $V \subseteq \operatorname{span}(v_{m_1}, \ldots, v_{m_r})$.

Thus, $V = \operatorname{span}(v_{m_1}, \ldots, v_{m_r})$, i.e. $(v_{m_1}, \ldots, v_{m_r})$ spans V, which completes the proof. $\square$

Theorem 13.2.6 (Replacement theorem)**.** *Let V be a vector space over $\mathbb{R}$, 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 tuples of vectors 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 $\mathbb{R}$.

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 11.3.10), 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

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

spans V. Therefore, there exists scalars $\lambda_1, \lambda_2, \ldots, \lambda_m, \mu_1, \mu_2, \ldots, \mu_{n-m} \in \mathbb{R}$ 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 13.1.1 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

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

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 α to α' which is why the theorem has its name.

First, we'll show that every element of α is in $\operatorname{span}(\alpha')$. We immediately have

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

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

Since $\operatorname{span}(\alpha')$ is a subspace of V, any linear combination of α is also in $\operatorname{span}(\alpha')$. Thus, we obtain $\operatorname{span}(\alpha) \subseteq \operatorname{span}(\alpha')$. Because $V = \operatorname{span}(\alpha)$, this tells us that $\operatorname{span}(\alpha') = V$, 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 m. $\square$

14 Questions due on August 20th

The first 4 questions are designed to encourage you to start thinking about matrices, matrix-vector multiplication, and matrix multiplication more in line with how I think about them, and how I discussed them in lecture. The key things to remember and internalize are:

- An $m \times n$ matrix A does something: multiplying on the left takes vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$. This is the linear transformation T_A.
- The j-th column of A tells you where e_j (the vector with 1 in the j-th place and 0's elsewhere) goes.
- Matrix multiplication is just the same as lots of matrix-vector multiplications:

$$A\left(\begin{array}{c|c|c|c} w_1 & w_2 & \cdots & w_p \end{array}\right) = \left(\begin{array}{c|c|c|c} Aw_1 & Aw_2 & \cdots & Aw_p \end{array}\right).$$

If you do these questions by trial-and-error, or by doing a bunch of matrix multiplications in the row-dot-column way of thinking, you'll be missing the point. I can promise you that the perspective I am trying to provide you with here is one of the most fundamental in linear algebra and, in terms of your success in this class, it will be completely invaluable once we come to the matrix of a linear transformation.

(These questions (1-4) should not take a long time.

On the other hand, question 7 and 8 might take longer, but they're important.)

1. This question will be concerned with 3×3 matrices and linear functions $\mathbb{R}^3 \longrightarrow \mathbb{R}^3$. Let

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \text{ and } e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

   (a) Find the 3×3 matrix which defines the function

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} x \\ y \\ 0 \end{pmatrix},$$

   by thinking about what this function does to the vectors e_1, e_2, and e_3.

   (b) Similarly, find the 3×3 matrix which defines the function

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} y \\ x \\ z \end{pmatrix}.$$

   (c) Find the 3×3 matrix which defines the function

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} x \\ x+y \\ x+y+z \end{pmatrix}.$$

2. Let e_1, e_2 and e_3 be as in the previous problem and let

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \left(e_2 \middle| e_3 \middle| e_1 \right).$$

(Matrix A is a special type of matrix called a permutation matrix.)

(a) What does the matrix A do to the vectors e_1, e_2, and e_3?

(b) Without multiplying matrices, what does A^2 do to e_1, e_2, and e_3?

(c) Without multiplying matrices, what is A^2?

(d) Check your answer to (c) by doing the matrix multiplication.
Be sure to think of the matrix multiplication as three matrix-vector multiplications. Wait... hopefully this forced you to redo some of the thinking in (b) again - hmmm.

(e) Without multiplying matrices, what does $A^3 = A \cdot A^2$ do to e_1, e_2, and e_3?

(f) Without multiplying matrices, what is A^3?

(g) Check your answer to (f) by doing the matrix multiplication $A \cdot A^2$.
Be sure to think of the matrix multiplication as three matrix-vector multiplications. Wait... hopefully this forced you to redo some of the thinking in (e) again - hmmm.

3. Let $e_1, e_2, e_3, e_4, e_5 \in \mathbb{R}^5$ be as in lecture, and let

$$B = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} = \left(e_2 \middle| e_1 \middle| e_4 \middle| e_5 \middle| e_3 \right).$$

(Matrix B is also a permutation matrix.)

(a) What does the matrix B do to the vectors e_1, e_2, e_3, e_4, and e_5?

(b) What is the smallest $n \in \mathbb{N} = \{1, 2, 3, \ldots\}$ such that

$$B^n = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = I_5?$$

Explain your answer fully without ever multiplying matrices.

(c) Do you want to check your answer with matrix multiplication?

(d) (Only if you answered “yes” to (c).)
Are you okay? Do you want me to set a bigger example to convince you otherwise?

4. Let $C = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $D = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

   (Matrices C and D are also permutation matrices.)

   (a) What does C do to e_1? What does C do to e_2?

   (b) What does D do to e_1? What does D do to e_2?

   (c) Without multiplying matrices, what does CD do to e_1?

   (d) Without multiplying matrices, what does DC do to e_1?

   (e) Without multiplying matrices, is $CD = DC$? Why?

   (f) Check your answer to (e) by doing matrix multiplication.
   Be sure to think of the matrix multiplication as three matrix-vector multiplications. You should be getting quick by now. You'll certainly repeat some of the thinking in (c) and (d).

5. (a) Give an example of a linear map $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ with kernel $\{(x, 0, 0) : \ x \in \mathbb{R}\}$.

   (b) Give an example of a linear map $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ with image $\{(x, 0, 0, 0) : \ x \in \mathbb{R}\}$.

6. Suppose U, V, and W are vector spaces over $\mathbb{R}$, and that

$$T : U \longrightarrow V, \ S : V \longrightarrow W$$

   are linear.

   (a) Prove $\ker(ST) \supseteq \ker(T)$.

   (b) Prove $\operatorname{im}(ST) \subseteq \operatorname{im} S$.

   (c) Give examples where the inequalities in (a) and (b) are actually equality.

   (d) Give examples where the inequalities in (a) and (b) are strict.

Question 7 is on the next page.

7. Suppose that V is a vector space over $\mathbb{R}$ and that $(v_1, v_2, \ldots, v_n)$ is a tuple of vectors in V.

   (a) Suppose that $n \geq 2$. Prove that $(v_1, v_2, \ldots, v_n)$ is linearly dependent if and only if one of the vectors can be written as a linear combination of the others.

   (b) Prove that $(v_1, v_2, \ldots, v_n)$ is linearly dependent if and only if there is a $k \in \{1, 2, \ldots, n\}$ such that
   $$v_k \in \operatorname{span}(v_1, v_2, \ldots, v_{k-1}).$$
   (When $k = 1$, the formula above should be read as saying $v_1 = 0$.)

8. This is probably the best / most important question.

   Suppose V and W are vector spaces over $\mathbb{R}$, that $T : V \to W$ is a linear transformation, and that $(v_1, v_2, \ldots, v_n)$ is a tuple of vectors in V.

   (a) Prove that if $(v_1, v_2, \ldots, v_n)$ is linearly dependent,
   then $(T(v_1), T(v_2), \ldots, T(v_n))$ is linearly dependent.

   (b) Prove that if $(T(v_1), T(v_2), \ldots, T(v_n))$ is linearly independent,
   then $(v_1, v_2, \ldots, v_n)$ is linearly independent.

   (c) Suppose $\ker T = \{0\}$.
   Prove that if $(T(v_1), T(v_2), \ldots, T(v_n))$ is linearly dependent,
   then $(v_1, v_2, \ldots, v_n)$ is linearly dependent.

   (d) Suppose $\ker T = \{0\}$.
   Prove that if $(v_1, v_2, \ldots, v_n)$ is linearly independent,
   then $(T(v_1), T(v_2), \ldots, T(v_n))$ is linearly independent.

   (e) Prove that if $(v_1, \ldots, v_n)$ spans V,
   then $(T(v_1), \ldots, T(v_n))$ spans $\operatorname{im} T$.

   (f) Prove that if $(v_1, v_2, \ldots, v_n)$ spans V
   and $(T(v_1), T(v_2), \ldots, T(v_n))$ is linearly independent,
   then $\ker T = \{0\}$.

9. Let V be a vector space over $\mathbb{R}$, and let $u, v, w \in V$.

   (a) Prove that (u, v) is linearly independent if and only if $(u+v, u-v)$ is linearly independent.

   If you thought I was weird for giving the textbook such a hard time about its definition of linearly (in)dependent, take a look at their version of this question: page 42, 13(a).
   As long as $V \neq \{0\}$, their statement is false because you can take $u \neq 0$ and $v = 0$. Then $\{u, v\}$ is linearly dependent, and $\{u + v, u - v\} = \{u\}$ is linearly independent (according to THEIR definition which involves sets instead of tuples).
   I rest my case.

   (b) Prove that (u, v, w) is linearly independent if and only if $(u + v, v + w, w + u)$ is linearly independent.

15 Solutions to the previous questions

1. (a) $e_1 \longmapsto e_1$, $e_2 \longmapsto e_2$, $e_3 \longmapsto 0$, so $(e_1|e_2|0)$.
   (b) $e_1 \longmapsto e_2$, $e_2 \longmapsto e_1$, $e_3 \longmapsto e_3$, so $(e_2|e_1|e_3)$.
   (c) $e_1 \longmapsto e_1 + e_2 + e_3$, $e_2 \longmapsto e_2 + e_3$, $e_3 \longmapsto e_3$, so $(e_1 + e_2 + e_3|e_2 + e_3|e_3)$.
2. (a) $A : e_1 \longmapsto e_2 \longmapsto e_3 \longmapsto e_1$.
   (b) $A^2 : e_1 \longmapsto e_3 \longmapsto e_2 \longmapsto e_1$.
   (c) $A^2 = (e_3|e_1|e_2)$.
   (d) $A^2 = (Ae_2|Ae_3|Ae_1) = (e_3|e_1|e_2)$.
   (e) $A^3 : e_1 \longmapsto e_1,\ e_2 \longmapsto e_2,\ e_3 \longmapsto e_3$.
   (f) $A^3 = (e_1|e_2|e_3) = I_3$.
   (g) $A^3 = A \cdot A^2 = (Ae_3|Ae_1|Ae_2) = (e_1|e_2|e_3) = I_3$.
3. (a) $B : e_1 \longmapsto e_2 \longmapsto e_1,\ e_3 \longmapsto e_4 \longmapsto e_5 \longmapsto e_3$.
   (b) We find that $B^n e_1 = e_1$ and $B^n e_2 = e_2$ if and only if n is even.
   We find that $B^n e_3 = e_3$, $B^n e_4 = e_4$, and $B^n e_5 = e_5$ if and only if n is divisible by 3.
   Thus, $B^n = I_5$ if and only if n is even and divisible by 3. The smallest such $n \in \mathbb{N}$ is 6.
   (c) No.
   (d) If I did, I wouldn't be okay. I'd be extremely upset and bored.
4. (a) $C : e_1 \longmapsto e_2 \longmapsto e_1$.
   (b) $D : e_1 \longmapsto e_2 \longmapsto e_3$.
   (c) $CD : e_1 \longmapsto e_1$.
   (d) $DC : e_1 \longmapsto e_3$.
   (e) The previous two observations tell us the first column of CD is different from the first column of DC.
   (f) $CD = C(e_2|e_3|e_1) = (Ce_2|Ce_3|Ce_1) = (e_1|e_3|e_2)$.
   $DC = D(e_2|e_1|e_3) = (De_2|De_1|De_3) = (e_3|e_2|e_1)$.

5. (a) $T(x_1, x_2, x_3) = (0, x_2, x_3, 0)$.
   (b) $T(x_1, x_2, x_3) = (x_1, 0, 0, 0)$.

6. Suppose U, V, and W are vector spaces over $\mathbb{R}$, and that

$$T : U \longrightarrow V,\ S : V \longrightarrow W$$

are linear.

   (a) We'll prove that $\ker(ST) \supseteq \ker(T)$. Let $u \in \ker(T)$. Then

$$(ST)(u) = S(T(u)) = S(0) = 0.$$

   The first equality holds by the definition of ST. The second hold because $u \in \ker T$. The last equality is a consequence of S being linear. This calculation shows that $u \in \ker(ST)$.

   (b) We'll prove that $\mathrm{im}(ST) \subseteq \mathrm{im}\, S$. Let $w \in \mathrm{im}(ST)$. By the definition of $\mathrm{im}(ST)$, we can find a $u \in U$ such that $w = (ST)(u)$. Letting $v = T(u)$, we have

$$S(v) = S(T(u)) = (ST)(u) = w.$$

   The first equality is because $v = T(u)$. The second is by definition of ST. The third is because of how we chose u. This calculation shows that $w \in \mathrm{im}\, S$.

   (c) Take $U = V = W = \{0\}$.

   (d) Take $T : \mathbb{R} \longrightarrow \mathbb{R}^2$, $T(x) = (x, 0)$, and $S : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $S(x_1, x_2) = x_2$. Then, for all $x \in \mathbb{R}$,

$$(ST)(x) = S(T(x)) = S(x, 0) = 0.$$

   Thus, $\ker(ST) = \mathbb{R}$ and $\mathrm{im}(ST) = \{0\}$.
   On the other hand, we have $\ker(T) = \{0\}$ and $\mathrm{im}\, S = \mathbb{R}$.

7. Suppose that V is a vector space over $\mathbb{R}$ and that $(v_1, v_2, \ldots, v_n)$ is a tuple of vectors in V.

   (a) First, suppose that $(v_1, v_2, \ldots, v_n)$ is linearly dependent. This means that we can find a non-trivial dependency relation

   $$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0.$$

   Since the dependency relation is non-trivial, we can choose a $k \in \{1, 2, \ldots, n\}$ such that $\lambda_k \neq 0$. We have

   $$v_k = \left(-\frac{\lambda_1}{\lambda_k}\right)v_1 + \ldots + \left(-\frac{\lambda_{k-1}}{\lambda_k}\right)v_{k-1} + \left(-\frac{\lambda_{k+1}}{\lambda_k}\right)v_{k+1} + \ldots + \left(-\frac{\lambda_n}{\lambda_k}\right)v_n,$$

   so v_k is a linear combination of $(v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n)$.
   Conversely, suppose that v_k is a linear combination of $(v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n)$. Then there are $\lambda_1, \ldots, \lambda_{k-1}, \lambda_{k+1}, \ldots \lambda_n \in \mathbb{R}$ such that

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

   This gives

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

   Since $1 \neq 0$, this is a non-trivial dependency relation, and so this shows $(v_1, v_2, \ldots, v_n)$ is linearly dependent.
   We don't need $n \geq 2$, as long as you have the correct convention regarding "others."

   (b) Suppose there is a $k \in \{1, 2, \ldots, n\}$ such that

   $$v_k \in \operatorname{span}(v_1, v_2, \ldots, v_{k-1}).$$

   If $k = 1$, this means that $v_1 = 0$, so (v_1) is linearly dependent. In any case, the argument given in part (a) shows that $(v_1, \ldots, v_k)$ is linearly dependent, and theorem 11.3.10 says that $(v_1, \ldots, v_n)$ is linearly dependent.
   Conversely, suppose $(v_1, v_2, \ldots, v_n)$ is linearly dependent. Let $k \in \{1, 2, \ldots, n\}$ be the minimum value such that $(v_1, v_2, \ldots, v_k)$ is linearly dependent. If $k = 1$, this means (v_1) is linearly dependent, so $v_1 = 0$, which is what we mean by

   $$v_k \in \operatorname{span}(v_1, v_2, \ldots, v_{k-1}).$$

   In any case, we have that $(v_1, \ldots, v_{k-1})$ is linearly independent, and $(v_1, \ldots, v_{k-1}, v_k)$ is linearly dependent, so theorem 13.1.1 tells us that $v_k \in \operatorname{span}(v_1, v_2, \ldots, v_{k-1})$.

   (c) Another proof of the harder direction of (b) which refines the argument of (a) says…
   Suppose that $(v_1, v_2, \ldots, v_n)$ is linearly dependent. Then there exist $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$, not all zero, such that $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0$. Let λ_k be the largest nonzero λ. Then $v_k = \frac{-1}{\lambda_k}(\lambda_1 v_1 + \ldots + \lambda_{k-1} v_{k-1}) \in \operatorname{span}(v_1, \ldots, v_{k-1})$.

8. Suppose V and W are vector spaces over $\mathbb{R}$, that $T : V \to W$ is a linear transformation, and that $(v_1, v_2, \ldots, v_n)$ is a tuple of vectors in V.

   (a) Suppose $(v_1, v_2, \ldots, v_n)$ is linearly dependent. This means that we can find a non-trivial dependency relation

   $$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0. \tag{15.1}$$

   We obtain

   $$\lambda_1 T(v_1) + \lambda_2 T(v_2) + \ldots + \lambda_n T(v_n) = T(\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n) = T(0) = 0.$$

   The first and last equalities use the linearity of T. The middle equality is obtained by applying T to (15.1).

   $$\lambda_1 T(v_1) + \lambda_2 T(v_2) + \ldots + \lambda_n T(v_n) = 0$$

   is a non-trivial dependency relation for $(T(v_1), T(v_2), \ldots, T(v_n))$, so $(T(v_1), \ldots, T(v_n))$ is linearly dependent.

   (b) This is the contrapositive of (a).

   (c) Suppose $\ker T = \{0\}$. Also, suppose that $(T(v_1), T(v_2), \ldots, T(v_n))$ is linearly dependent. This means that we can find a non-trivial dependency relation

   $$\lambda_1 T(v_1) + \lambda_2 T(v_2) + \ldots + \lambda_n T(v_n) = 0.$$

   Using linearity of T, we obtain

   $$T(\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n) = \lambda_1 T(v_1) + \lambda_2 T(v_2) + \ldots + \lambda_n T(v_n) = 0.$$

   This shows $\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n \in \ker T$. Since $\ker T = \{0\}$,

   $$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0$$

   and this is a non-trivial dependency relation for $(v_1, v_2, \ldots, v_n)$, so $(v_1, \ldots, v_n)$ is linearly dependent.

   (d) This is the contrapositive of (c).

   (e) Suppose $(v_1, \ldots, v_n)$ spans V. It is clear that $\operatorname{im} T \supseteq \operatorname{span}(T(v_1), T(v_2), \ldots, T(v_n))$. We wish to show that $\operatorname{im} T \subseteq \operatorname{span}(T(v_1), T(v_2), \ldots, T(v_n))$.
   Let $w \in \operatorname{im} T$. By definition of $\operatorname{im} T$, we can find a $v \in V$ such that $T(v) = w$.
   Since $(v_1, v_2, \ldots, v_n)$ spans V, we can find $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ such that

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

   Applying T and using the fact that T is linear gives

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

   Since this quantity is equal to w, we see that $w \in \operatorname{span}(T(v_1), T(v_2), \ldots, T(v_n))$.
   Thus, $\operatorname{im} T = \operatorname{span}(T(v_1), T(v_2), \ldots, T(v_n))$, i.e. $(T(v_1), \ldots, T(v_n))$ spans $\operatorname{im} T$.

(f) Suppose that $(v_1, v_2, \ldots, v_n)$ spans V and that $(T(v_1), T(v_2), \ldots, T(v_n))$ is linearly independent. We wish show $\ker T = \{0\}$.

Let $v \in \ker T$. Since $(v_1, v_2, \ldots, v_n)$ spans V, we can find $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ such that

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = v. \tag{15.2}$$

We obtain

$$\lambda_1 T(v_1) + \lambda_2 T(v_2) + \ldots + \lambda_n T(v_n) = T(\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n) = T(v) = 0.$$

The first equality uses the linearity of T. The middle equality is obtained by applying T to (15.2). The last equality is because $v \in \ker T$.

$$\lambda_1 T(v_1) + \lambda_2 T(v_2) + \ldots + \lambda_n T(v_n) = 0$$

is a dependency relation for $(T(v_1), T(v_2), \ldots, T(v_n))$. Since $(T(v_1), T(v_2), \ldots, T(v_n))$ is linearly independent, it must be the trivial dependency relation, i.e. $\lambda_1 = \ldots = \lambda_n = 0$. Thus, $v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0$, and $v \in \{0\}$.

We have shown that $\ker T \subseteq \{0\}$. $\ker T \supseteq \{0\}$ is always true. Thus, $\ker T = \{0\}$.

9. This question turned out to be longer than I thought it would be. I'm sorry.

(a) Let V be a vector space over $\mathbb{R}$, and let $u, v \in V$.
Suppose (u, v) is linearly independent, $\lambda_1, \lambda_2 \in \mathbb{R}$, and that

$$\lambda_1(u+v) + \lambda_2(u-v) = 0.$$

Rewriting this gives

$$(\lambda_1 + \lambda_2)u + (\lambda_1 - \lambda_2)v = 0.$$

Since (u, v) is linearly independent, we obtain

$$\lambda_1 + \lambda_2 = \lambda_1 - \lambda_2 = 0.$$

Thus,

$$\begin{aligned}
\lambda_1 &= \frac{(\lambda_1 + \lambda_2) + (\lambda_1 - \lambda_2)}{2} = \frac{0+0}{2} = 0, \\
\lambda_2 &= \frac{(\lambda_1 + \lambda_2) - (\lambda_1 - \lambda_2)}{2} = \frac{0+0}{2} = 0.
\end{aligned}$$

We conclude $(u+v, u-v)$ is linearly independent.
Suppose $(u+v, u-v)$ is linearly independent, $\lambda_1, \lambda_2 \in \mathbb{R}$, and that

$$\lambda_1 u + \lambda_2 v = 0.$$

Rewriting this gives

$$\left(\frac{\lambda_1 + \lambda_2}{2}\right)(u+v) + \left(\frac{\lambda_1 - \lambda_2}{2}\right)(u-v) = 0.$$

Since $(u+v, u-v)$ is linearly independent, we obtain

$$\frac{\lambda_1 + \lambda_2}{2} = \frac{\lambda_1 - \lambda_2}{2} = 0.$$

Thus,

$$\begin{aligned}
\lambda_1 &= \frac{\lambda_1 + \lambda_2}{2} + \frac{\lambda_1 - \lambda_2}{2} = 0 + 0 = 0, \\
\lambda_2 &= \frac{\lambda_1 + \lambda_2}{2} - \frac{\lambda_1 - \lambda_2}{2} = 0 - 0 = 0.
\end{aligned}$$

(b) Let V be a vector space over $\mathbb{R}$, and let $u, v, w \in V$.

Suppose (u, v, w) is linearly independent, $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, and that

$$\lambda_1(u+v) + \lambda_2(v+w) + \lambda_3(w+u) = 0.$$

Rewriting this gives

$$(\lambda_3 + \lambda_1)u + (\lambda_1 + \lambda_2)v + (\lambda_2 + \lambda_3)w = 0.$$

Since (u, v, w) is linearly independent, we obtain

$$\lambda_3 + \lambda_1 = \lambda_1 + \lambda_2 = \lambda_2 + \lambda_3 = 0.$$

Thus,

$$\begin{aligned}
\lambda_1 &= \frac{+\ (\lambda_3 + \lambda_1) + (\lambda_1 + \lambda_2) - (\lambda_2 + \lambda_3)}{2} = \frac{0+0+0}{2} = 0, \\
\lambda_2 &= \frac{-\ (\lambda_3 + \lambda_1) + (\lambda_1 + \lambda_2) + (\lambda_2 + \lambda_3)}{2} = \frac{0+0+0}{2} = 0, \\
\lambda_3 &= \frac{+\ (\lambda_3 + \lambda_1) - (\lambda_1 + \lambda_2) + (\lambda_2 + \lambda_3)}{2} = \frac{0+0+0}{2} = 0.
\end{aligned}$$

We conclude that $(u+v, v+w, w+u)$ is linearly independent.

Suppose $(u+v, v+w, w+u)$ is linearly independent, $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, and that

$$\lambda_1 u + \lambda_2 v + \lambda_3 w = 0.$$

Rewriting this gives

$$\begin{aligned}
\left(\frac{+\ \lambda_1 + \lambda_2 - \lambda_3}{2}\right)(u+v) + \left(\frac{-\ \lambda_1 + \lambda_2 + \lambda_3}{2}\right)(v+w) \\
+ \left(\frac{+\ \lambda_1 - \lambda_2 + \lambda_3}{2}\right)(w+u) = 0.
\end{aligned}$$

Since $(u+v, v+w, w+u)$ is linearly independent, we obtain

$$\frac{+\ \lambda_1 + \lambda_2 - \lambda_3}{2} = \frac{-\ \lambda_1 + \lambda_2 + \lambda_3}{2} = \frac{+\ \lambda_1 - \lambda_2 + \lambda_3}{2} = 0.$$

Then

$$\begin{aligned}
\lambda_1 &= \frac{+\ \lambda_1 + \lambda_2 - \lambda_3}{2} + \frac{+\ \lambda_1 - \lambda_2 + \lambda_3}{2} = 0 + 0 = 0, \\
\lambda_2 &= \frac{+\ \lambda_1 + \lambda_2 - \lambda_3}{2} + \frac{-\ \lambda_1 + \lambda_2 + \lambda_3}{2} = 0 + 0 = 0, \\
\lambda_3 &= \frac{-\ \lambda_1 + \lambda_2 + \lambda_3}{2} + \frac{+\ \lambda_1 - \lambda_2 + \lambda_3}{2} = 0 + 0 = 0.
\end{aligned}$$

We conclude that (u, v, w) is linearly independent.

16 Quiz 1

1. Recall that $\mathbb{R}^1 = \{\ (x)\ :\ x \in \mathbb{R}\}$ is a vector space over $\mathbb{R}$ with coordinatewise addition and scalar multiplication.

   (a) List every subspace of $\mathbb{R}^1$. You don't need to prove your claim in this part.
   Solution: $\{0\}$ and $\mathbb{R}^1$.

   (b) Suppose V is a subspace of $\mathbb{R}^1$.
   Prove that V has to be equal to one of the subspaces you wrote down in part (a).
   Solution: There are two cases.

      i. $V = \{0\}$. Then, trivially, V is one of the subspaces we wrote down in part (a).

      ii. $V \neq \{0\}$. We wish to show that $V = \mathbb{R}^1$.
      Since V is a subspace, we have $\{0\} \subseteq V$. Thus, $V \neq \{0\}$ tells us that $V \not\subseteq \{0\}$. This means we can find an element $(x) \in V$ with $x \neq 0$. Since $x \neq 0$, x^{-1} is a well-defined real number. Because $(x) \in V$ and V is closed under scalar multiplication, we find that $x^{-1}(x) \in V$, i.e. $(1) \in V$.
      We trivially have $V \subseteq \mathbb{R}^1$ because V is a subspace. Conversely, given $(y) \in \mathbb{R}^1$, we have $(y) = y(1) \in V$ because V is closed under scalar multiplication. Thus, $V = \mathbb{R}^1$.

2. Suppose X is a nonempty set.

   In class and on the homework, we proved that the set of real-valued functions from X,

   $$\mathcal{F} = \{f : X \longrightarrow \mathbb{R}\}$$

   is a vector space over $\mathbb{R}$ when equipped with pointwise addition and scalar multiplication.

   (a) Give the definition of addition and scalar multiplication in $\mathcal{F}$.
   Solution: For $f, g \in \mathcal{F}$, $\lambda \in \mathbb{R}$, $x \in X$, we have

   $$(f+g)(x) := f(x) + g(x) \text{ and } (\lambda f)(x) := \lambda(f(x)).$$

   (b) Verify the eighth axiom of a vector space for $\mathcal{F}$, that is, that for all $\lambda, \mu \in \mathbb{R}$, and for all $f \in \mathcal{F}$, $(\lambda + \mu)f = \lambda f + \mu f$.
   Solution: see homework.

   (c) Fix $x_0, x_1 \in X$. Prove that $T : \mathcal{F} \longrightarrow \mathbb{R}$, $f \longmapsto f(x_0) + f(x_1)$ is linear.
   Solution: Let $f, g \in \mathcal{F}$, $\lambda \in \mathbb{R}$. Then

   $$\begin{aligned} T(f+g) = (f+g)(x_0) + (f+g)(x_1) &= f(x_0) + g(x_0) + f(x_1) + g(x_1) \\ &= f(x_0) + f(x_1) + g(x_0) + g(x_1) = T(f) + T(g). \end{aligned}$$

   The first and last equality follow from the definition of T. The second follows from the definition of addition given in part (a). The third uses commutativity of addition in $\mathbb{R}$. Also,

   $$T(\lambda f) = (\lambda f)(x_0) + (\lambda f)(x_1) = \lambda(f(x_0)) + \lambda(f(x_1)) = \lambda(f(x_0) + f(x_1)) = \lambda T(f).$$

   The first and last equality follow from the definition of T. The second follows from the definition of scalar multiplication given in part (a). The third follows from distributivity of addition and multiplication in $\mathbb{R}$.

Now take $X = \mathbb{R}$ so that $\mathcal{F} = \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$.

(d) Is the subset $\{f \in \mathcal{F} : \ f(0) = 1\}$ a subspace of $\mathcal{F}$?
Solution: No, because the zero function is not contained in this subset.

(e) Is the subset $\{f \in \mathcal{F} : \ f(-1) \cdot f(1) = 0\}$ a subspace of $\mathcal{F}$?
Solution: No. Here's why...
Call the set in question $\mathcal{G}$. Let $f_+, f_- : \mathbb{R} \longrightarrow \mathbb{R}$ be defined by $f_+(x) = 1+x$ and $f_-(x) = 1 - x$. Then $f_+(-1) = 0$ and $f_-(1) = 0$, so $f_+, f_- \in \mathcal{G}$. However, $(f_+ + f_-)(x) = 2$, so

$$(f_+ + f_-)(-1) \cdot (f_+ + f_-)(1) = 2 \cdot 2 = 4 \neq 0,$$

which means $f_+ + f_- \notin \mathcal{G}$.

3. Let $(e_1, e_2, \ldots, e_6)$ be the standard basis of $\mathbb{R}^6$. Consider the 6×6 matrix

$$A = \left(2e_5 \middle| e_3 \middle| 3e_6 \middle| 5e_4 \middle| e_1 \middle| e_2 \right).$$

Calculate A^{30} and A^{31} efficiently.

Solution:

- $A : e_4 \longmapsto 5e_4$,
- $A : e_5 \longmapsto e_1 \longmapsto 2e_5$,
- $A : e_6 \longmapsto e_2 \longmapsto e_3 \longmapsto 3e_6$.

Thus,

- $A^n : e_4 \longmapsto 5^n e_4$,
- $A^{2n} : e_1 \longmapsto 2^n e_1, \ e_5 \longmapsto 2^n e_5$,
- $A^{3n} : e_2 \longmapsto 3^n e_2, \ e_3 \longmapsto 3^n e_3, \ e_6 \longmapsto 3^n e_6$.

In particular,

- $A^{30} : e_4 \longmapsto 5^{30} e_4$,
- $A^{30} : e_1 \longmapsto 2^{15} e_1, \ e_5 \longmapsto 2^{15} e_5$,
- $A^{30} : e_2 \longmapsto 3^{10} e_2, \ e_3 \longmapsto 3^{10} e_3, \ e_6 \longmapsto 3^{10} e_6$.

So $A^{30} = \left(2^{15} e_1 \middle| 3^{10} e_2 \middle| 3^{10} e_3 \middle| 5^{30} e_4 \middle| 2^{15} e_5 \middle| 3^{10} e_6 \right)$ and

$$\begin{aligned} A^{31} &= \left(2^{15} Ae_1 \middle| 3^{10} Ae_2 \middle| 3^{10} Ae_3 \middle| 5^{30} Ae_4 \middle| 2^{15} Ae_5 \middle| 3^{10} Ae_6 \right) \\ &= \left(2^{16} e_5 \middle| 3^{10} e_3 \middle| 3^{11} e_6 \middle| 5^{31} e_4 \middle| 2^{15} e_1 \middle| 3^{10} e_2 \right). \end{aligned}$$

4. (a) Prove that the following tuple spans $\mathbb{R}^3$.

$$\Big((1,0,0),\ (1,1,0),\ (1,1,1)\Big)$$

Solution: For any $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$, we have

$$(\lambda_1, \lambda_2, \lambda_3) = (\lambda_1 - \lambda_2)(1,0,0) + (\lambda_2 - \lambda_3)(1,1,0) + \lambda_3(1,1,1).$$

This shows $\mathbb{R}^3 \subseteq \operatorname{span}\Big((1,0,0),\ (1,1,0),\ (1,1,1)\Big)$. The other inclusion is trivial.

(b) Define $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ by

$$T\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} 1 & 0 & -1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

For this particular example, describe $\ker T$, and prove your claim.
Solution: We claim that

$$\ker T = \left\{ \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \ :\ \lambda \in \mathbb{R} \right\}.$$

"$\supseteq$" is easy since you can just check that for any $\lambda \in \mathbb{R}$, $T\left(\lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right) = 0.$

For "$\subseteq$", suppose that $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \ker T$, i.e. $T\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = 0$. The definition of T gives

$$\begin{pmatrix} x_1 - x_3 \\ x_1 - x_2 \\ x_2 - x_3 \end{pmatrix} = 0.$$

Looking at the second coordinate gives $x_1 = x_2$, while looking at the third coordinate gives $x_2 = x_3$. Letting $\lambda = x_1$, we obtain

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Thus, $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \left\{ \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \ :\ \lambda \in \mathbb{R} \right\}.$

5. Suppose V and W are vector spaces over $\mathbb{R}$, that $S : V \to W$ is a linear transformation, and that $(v_1, v_2, \ldots, v_n)$ is a tuple of vectors in V.

   Always true or sometimes false (i.e. depends on V, W, S, etc.)?

   (a) If $(v_1, v_2, \ldots, v_n)$ spans V,
   then $(S(v_1), S(v_2), \ldots, S(v_n))$ is linearly independent and $\ker S = \{0\}$.

   (b) If $\ker S = \{0\}$ and $(S(v_1), S(v_2), \ldots, S(v_n))$ is linearly independent,
   then $(v_1, v_2, \ldots, v_n)$ is linearly independent.

   (c) If im $S = W$, then $(S(v_1), S(v_2), \ldots, S(v_n))$ spans W.

   (d) If $(v_1, v_2, \ldots, v_n)$ is linearly independent,
   then $(S(v_1), S(v_2), \ldots, S(v_n))$ is linearly independent.

   Solution:

   (a) Sometimes false. Consider the case when $S : \mathbb{R} \longrightarrow \{0\}$, and $v_1 = 1$.
   (v_1) spans $\mathbb{R}$, but $(S(v_1)) = (0)$ is linearly dependent and $\ker S = \mathbb{R}$.

   (b) Always true. You proved this on the homework.
   In fact, the $\ker S = \{0\}$ hypothesis is not even needed.

   (c) Sometimes false. Consider the case when $S : \mathbb{R} \longrightarrow \mathbb{R}$ is the identity, and $v_1 = 0$. Then $\operatorname{im} S = \mathbb{R}$, but $(S(v_1)) = (0)$ does not span $\mathbb{R}$.

   (d) Sometimes false. Consider the case when $S : \mathbb{R} \longrightarrow \{0\}$, and $v_1 = 1$.
   (v_1) is linearly independent, but $(S(v_1)) = (0)$ is linearly dependent.

17 Comments on quiz 1

1. (a) $\{1, 2, \ldots, n\}$?!
   $\{(x) : \ x \in \mathbb{N} \setminus \{0\}\}$?!
   $\{0, 1, 2, \ldots, n\}$?!
   $\mathbb{N}$, $\mathbb{Z}$?!
   You really should know from 33A that the only subspaces of $\mathbb{R}^1$ are $\{0\}$ and $\mathbb{R}^1$. 33A is a prerequisite for this class.

   (b) Most common error...
   You proved that $\{0\}$ is a subspace and that $\mathbb{R}^1$ is a subspace. Yes, this is true. It's the simplest fact I can think of about subspaces. However, it is not at all what the question asks you to do, and is completely irrelevant for answering the question. In proving these results you're addressing the claim:

   $$\text{if } V = \{0\} \text{ or } V = \mathbb{R}^1, \text{ then } V \text{ is a subspace of } \mathbb{R}^1 .$$

   The question asks you to prove:

   $$\text{if } V \text{ is a subspace of } \mathbb{R}^1, \text{ then } V = \{0\} \text{ or } V = \mathbb{R}^1.$$

2. Main complaints...

   - Why do you not know definitions which I have given in class and have used excessively in my homework solutions?
   - Every time I have checked an axiom, I have done one of two things:
     - labelled each equality with a justification;
     - addressed, in a paragraph after the relevant equations, why each equality is true.

     What are you not doing the same?
   - To show that $\{f \in \mathcal{F} : \ f(-1) \cdot f(1) = 0\}$ is not closed under addition, you must provide an explicit counter-example. This means defining functions.
     I do something similar in 3.(b) of the 8/13 questions where I use explicit vectors.
     Also, in class I showed

     $$A = \Big\{(x, 0) : \ x \in \mathbb{R}\Big\} \cup \Big\{(0, y) : \ y \in \mathbb{R}\Big\}$$

     is not a subspace of $\mathbb{R}^2$ by saying that $(1, 0), (0, 1) \in A$, but $(1, 0) + (0, 1) \notin A$.

3. Not much to say about this.

4. Main complaints...

   - I never use the symbol $\implies$ in my notes or in lecture. I would encourage you to avoid it too. For one thing, most mathematicians, even experienced ones, misuse this symbol terribly. For another, in lots of your arguments, you needed $\impliedby$ instead. Using words might cause you to realize this.
   - In (b), realizing that

     $$\ker T = \left\{ \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \;:\; \lambda \in \mathbb{R} \right\}$$

     is 33A-level stuff. But this is 115A, so we require more...

     At the time of this quiz, we did not have the rank-nullity theorem or the consequences of the replacement theorem. There was exactly one way to address this equality of sets: prove "$\subseteq$" and "$\supseteq$" directly.

     I have done this literally every time in lecture, pointed out that I am doing this, and I have done it in my homework solutions: see example 9.1.7, 8/8: 3, 8/13: 2, 8/15: 3.

5. If you said something is true when it is false, then you are missing the point of proving results. The purpose of proving results is to have a complete understanding of why they are true. If you do not have a solid reason for knowing why something is true, then don't claim it is true.

   If you said something is false when it is true, then you are missing the point of counter-examples. The purpose of a counter-example is to have an indisputable and explicit example of something being false. If you cannot give one, then maybe the result is true. Don't dismiss it necessarily.

18 Lecture on August 20th: Dimension

Recall theorem 13.2.5.

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

In particular, V has a basis.

Recall the replacement theorem.

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

Suppose that $(v_1, v_2, \dots, v_m)$ and $(w_1, w_2, \dots, w_n)$ are tuples of vectors in V.

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

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

Corollary 18.1. *Let V be a vector space over $\mathbb{R}$, and $m, n \in \mathbb{N} \cup \{0\}$.*

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

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

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

Recall definition 13.2.4.

Definition. Suppose V is a vector space over $\mathbb{R}$. We say V is *finite-dimensional* **iff** V is the span of a finite tuple; 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.

Theorem 18.2. *Definition 13.2.4 makes sense.*

Proof. There was no doubt that the definition of "finite-dimensional" and "infinite-dimensional" made sense. Suppose that V is a finite-dimensional vector space over $\mathbb{R}$. By definition, we can find a tuple $(v_1, v_2, \dots, v_n)$ spanning V. Theorem 13.2.5 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. $\square$

Example 18.3.

1. The vector space $\{0\}$ has dimension 0, since its basis is the empty tuple.

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

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

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

   is a basis; here $[E_{pq}]_{ij} = \delta_{p,i}\delta_{q,j}$ where $\delta_{k,l} = 1$ if $k = l$, and 0 otherwise.

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

Theorem 18.4. *Let V be a finite-dimensional vector space over $\mathbb{R}$ and suppose that $\dim V = n$.*

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 $\mathbb{R}$ and suppose that $\dim V = n$.

1. Suppose $(v_1, \ldots, v_m)$ spans V. By theorem 13.2.5, 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 $(v_1, \ldots, v_n)$ span V. By theorem 13.2.5, 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 $(v_1, \ldots, v_m, v_{m+1}, \ldots v_n)$ spans 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.

$\square$

19 Questions due on August 22th

1. [Optional]

   page 53, question 1, but replace "subset" by "tuple" where appropriate, and "generates" with "spans." page 54, question 2, but replace "sets" by "tuple."

2. *This question should be quick! Use theorem 18.4.*

   (a) Is $\Big((1,4,-6),\ (1,5,8),\ (2,1,1),\ (0,1,0)\Big)$ linearly independent in $\mathbb{R}^3$?

   Solution: No. There are 4 vectors and $4 > 3 = \dim \mathbb{R}^3$.

   (b) Does $(x^3 - 2x^2 + 1, 4x^2 - x + 3, 3x - 2)$ span $\mathcal{P}_3(\mathbb{R})$?

   Solution: No. There are 3 polynomials and $3 < 4 = \dim \mathcal{P}_3(\mathbb{R})$.

3. You should know the following definition. We'll need it later, for sure!

   Definition. Suppose V is a vector space over $\mathbb{R}$, and that S_1 and S_2 are subsets of V. Then the *sum* of S_1 and S_2 is the set

   $$S_1 + S_2 := \{s_1 + s_2 :\ s_1 \in S_1,\ s_2 \in S_2\}.$$

   Remark. This means that in proofs you write things like...

   Let $x \in S_1 + S_2$. By definition, we can find $s_1 \in S_1$ and $s_2 \in S_2$ such that $x = s_1 + s_2$.

   Or... We have $s_1 \in S_1$ and $s_2 \in S_2$, and so, by definition, $s_1 + s_2 \in S_1 + S_2$.

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

   (a) Suppose that W_1 and W_2 are subspaces of V.
   Prove that $W_1 + W_2$ is a subspace of V that contains both W_1 and W_2.

   (b) Suppose that W_1 and W_2 are subspaces of V.
   Prove that any subspace of V that contains both W_1 and W_2 must also contain $W_1 + W_2$.

   (c) Suppose that $v_1, v_2, \ldots, v_n, v_{n+1}, \ldots, v_{n+m} \in V$.
   Prove that $\operatorname{span}(v_1, \ldots, v_{n+m}) = \operatorname{span}(v_1, \ldots, v_n) + \operatorname{span}(v_{n+1}, \ldots, v_{n+m})$.

20 Lecture on August 22th: The Rank-Nullity Theorem

20.1 Leftover from last time

Theorem 20.1.1. *Suppose V is a finite-dimensional vector space over $\mathbb{R}$ and W is 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 $\mathbb{R}$, 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 ( ), 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 18.4, 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 13.1.1 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 18.4 part 4, shows $(w_1, w_2, \ldots, w_m)$ is a basis for V, so $W = V$. $\square$

Corollary 20.1.2. *Suppose V is a finite-dimensional vector space over $\mathbb{R}$ and W is 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 18.4 part 3 says we can extend it to a basis of V. $\square$

20.2 Rank-Nullity

Theorem 20.2.1. *Let V and W be vector spaces over $\mathbb{R}$ and let $T : V \longrightarrow W$ be a linear transformation. Suppose that $(v_1, v_2, \ldots, v_n)$ spans V. Then*

$$\operatorname{im} T = \mathit{span}(T(v_1), T(v_2), \ldots, T(v_n)).$$

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

Proof. This is 8(e) of from your last set of weekend questions. $\square$

Definition 20.2.2. Let V and W be vector spaces over $\mathbb{R}$ 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 $\dim(\operatorname{im} T)$.

Theorem 20.2.3 (Rank-Nullity Theorem)**.** *Suppose V and W are vector spaces over $\mathbb{R}$ and that $T : V \longrightarrow W$ is a linear transformation. Suppose, also, that V is finite-dimensional. Then*

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

Proof. Let V and W be vector spaces over $\mathbb{R}$ 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 $(T(v_{n+1}), \ldots, T(v_m))$ spans $\mathrm{im}\, T$. Since $(v_1, \ldots, v_n, v_{n+1}, \ldots, v_m)$ spans V, theorem 20.2.1 tells us that $(T(v_1), \ldots, T(v_n), T(v_{n+1}), \ldots, T(v_m))$ spans $\mathrm{im}\, T$. Since $v_1, \ldots, v_n \in \ker T$, we have $T(v_1) = \ldots = T(v_n) = 0$. So

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

Now we show that $(T(v_{n+1}), \ldots, T(v_m))$ is linearly independent. Suppose that $\lambda_{n+1}, \ldots, \lambda_m \in \mathbb{R}$ 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 $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ 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 20.2.4. *Let $n \in \mathbb{N} \cup \{0\}$, let V and W be finite-dimensional vector spaces over $\mathbb{R}$ 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 $\mathbb{R}$ each with dimension n, and let $T : V \longrightarrow W$ be a linear transformation.

Theorem 9.1.4 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 $\mathbb{R}$ with dimension 0.

Remark 9.1.6 says that conditions 4 and 5 are equivalent.
Condition 5 implies condition 6 since $\dim W = n$.
Condition 6 implies condition 5 by theorem 20.1.1 applied to $\mathrm{im}\, T \subseteq W$.

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

The rank-nullity theorem allows one to calculate bases for kernels and images of linear transformations very quickly. The next example illustrates this. I'll be disappointed if you don't use this type of argument on quiz 2 or the final.

Example 20.2.5. Let

$$A = \begin{pmatrix} 1 & 6 & 1 & 15 & 1 \\ 0 & 3 & 0 & 6 & 1 \\ 0 & 5 & 1 & 11 & 1 \\ 0 & 3 & 0 & 6 & 1 \end{pmatrix}$$

Then $T_A : \mathbb{R}^5 \longrightarrow \mathbb{R}^4$. Here is *the best way* to find a basis for $\ker T_A$ and $\operatorname{im} T_A$.

- $\beta_K = \left(\begin{pmatrix} 1 \\ -1 \\ 2 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ 5 \\ -1 \\ 6 \end{pmatrix} \right)$ is tuple of vectors in $\ker T_A$.

- $\beta_I = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right)$ is tuple of vectors in $\operatorname{im} T_A$.

- The tuples that we just wrote down are both linearly independent and this is easy to check. Theorem 18.4 (part 3) tells us that $null(T_A) \geq 2$ and $rank(T_A) \geq 3$. The rank-nullity theorem tells us that $rank(T_A) + null(T_A) = 5$. These two pieces of information imply $null(T_A) = 2$ and $rank(T_A) = 3$.

- Theorem 18.4 (part 4) tells us that β_K is a basis for $\ker T_A$ and that β_I is a basis for $\operatorname{im} T_A$.

- We see that from this calculation that T_A is neither injective or surjective.

20.3 Isomorphisms

Lemma 20.3.1. *Suppose that V and W are vector spaces over $\mathbb{R}$, that $T : V \longrightarrow W$ is an injective linear transformation, and that $S : W \longrightarrow V$ is a function with $TS = 1_W$. Then S is linear.*

Proof. Suppose that V and W are vector spaces over $\mathbb{R}$, that $T : V \longrightarrow W$ is an injective linear transformation, and that $S : W \longrightarrow V$ is a function with $TS = 1_W$. We will show that S is linear.

Let $w_1, w_2 \in W$. We wish to show

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

Since T is injective, it is enough to show $T(S(w_1 + w_2)) = T(S(w_1) + S(w_2))$. Since $TS = 1_W$, the LHS is $w_1 + w_2$. Since T is linear, the RHS is $T(S(w_1)) + T(S(w_2))$. Moreover, since $TS = 1_W$, this is $w_1 + w_2$. Thus, $S(w_1 + w_2) = S(w_1) + S(w_2)$.

Similarly, we can show that if $\lambda \in \mathbb{R}$ and $w \in W$, then $S(\lambda w) = \lambda S(w)$. $\square$

Theorem 20.3.2. *Suppose that V and W are vector spaces over $\mathbb{R}$, and $T : V \longrightarrow W$ is a linear transformation. The following statements are equivalent.*

1. $\ker T = \{0\}$ *and* $\operatorname{im} T = W$.

2. *T is injective and surjective, that is, T is a bijection.*

3. *There exists a linear transformation $S : W \longrightarrow V$ such that $ST = 1_V$ and $TS = 1_W$.*

Proof. Suppose that V and W are vector spaces over $\mathbb{R}$, and $T : V \longrightarrow W$ is a linear transformation.

- 1. $\implies$ 2.

  We said this back in section 9.1.

- 2. $\implies$ 3.

  Suppose T is a bijection. It is a theorem in set theory that there is a function $S : W \longrightarrow V$ such that $ST = 1_V$ and $TS = 1_W$. You can see `http://math.ucla.edu/~mjandr/Math95` (theorem 11.2.4) for a proof. The lemma tells us that S is linear.

- 3. $\implies$ 1.

  Suppose we have $S : W \longrightarrow V$ with the property that $ST = 1_V$ and $TS = 1_W$.

  Given $v \in \ker T$, we have $v = 1_V(v) = (ST)(v) = S(T(v)) = S(0) = 0$, so $\ker T = \{0\}$.

  Given $w \in W$, we have $w = 1_W(w) = (TS)(w) = T(S(w)) \in \operatorname{im} T$, so $\operatorname{im} T = W$.

$\square$

Definition 20.3.3. Suppose that V and W are vector spaces over $\mathbb{R}$, and $T : V \longrightarrow W$ is a linear transformation. We say that T is an *isomorphism* **iff** any of the three statements in the theorem hold.

Theorem 20.3.4. *Let V and W be vector spaces over $\mathbb{R}$.*

Suppose $T : V \longrightarrow W$ is a linear transformation, and that $(v_1, v_2, \ldots, v_n)$ is a basis for V. Then T is an isomorphism if and only if $(T(v_1), T(v_2), \ldots, T(v_n))$ is a basis for W.

Proof. Let V and W be vector spaces over $\mathbb{R}$. Also, suppose $T : V \longrightarrow W$ is a linear transformation, and that $(v_1, v_2, \ldots, v_n)$ is a basis for V.

First, suppose T is an isomorphism.

We have $\ker T = \{0\}$ and $(v_1, v_2, \ldots, v_n)$ is linearly independent, so 8.(d) of the last weekend questions shows that $(T(v_1), T(v_2), \ldots, T(v_n))$ is linearly independent.

We have $\operatorname{im} T = W$ and $(v_1, v_2, \ldots, v_n)$ spans V, so 8.(e) of the last weekend questions shows that $(T(v_1), T(v_2), \ldots, T(v_n))$ spans W.

Thus, $(T(v_1), T(v_2), \ldots, T(v_n))$ is a basis for W.

Conversely, suppose that $(T(v_1), T(v_2), \ldots, T(v_n))$ is a basis for W. This immediately gives $\operatorname{im} T = W$. Moreover, since $(v_1, v_2, \ldots, v_n)$ spans V and $(T(v_1), T(v_2), \ldots, T(v_n))$ is linearly independent, 8.(f) of the last weekend questions shows that $\ker T = \{0\}$. Thus, T is an isomorphism. $\square$

21 Questions due on August 23th

0. Have the night off, or start the weekend's questions, or read over the lecture notes.

22 Lecture on August 23th: The classification theorem and the matrix of a linear transformation

22.1 The classification of finite-dimensional vector spaces over $\mathbb{R}$

Theorem 22.1.1. *Suppose that V and W are vector spaces over $\mathbb{R}$, and $T : V \longrightarrow W$ 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 question 2 of your weekend questions.

The last statement follows from the first two statements, or the previous theorem. □

Theorem 22.1.2 (Classification theorem)**.** *Let V be a finite-dimensional vector space over $\mathbb{R}$. Let $n = \dim V$. Then there exists an isomorphism $\mathbb{R}^n \longrightarrow V$.*

Proof. Let V be a finite-dimensional vector space over $\mathbb{R}$. Let $n = \dim V$. Pick a basis

$$\beta = (v_1, v_2, \dots, v_n)$$

for V. We have a linear map $\Gamma_\beta : \mathbb{R}^n \longrightarrow V$, $(\lambda_1, \lambda_2, \dots, \lambda_n) \longmapsto \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n$.

We have already seen that β spanning V gives surjectivity and β being linearly independent gives injectivity. Thus, Γ_β is an isomorphism. (Perhaps you'd prefer to argue surjectivity directly and use theorem 20.2.4 to get injectivity, or maybe you'd prefer to say that Γ_β takes the standard basis $(e_1, e_2, \dots, e_n)$ to the basis $(v_1, v_2, \dots, v_n)$, and use theorem 20.3.4 to conclude it's an isomorphism.) □

Theorem 22.1.3. *Suppose that V and W are finite-dimensional vector spaces over $\mathbb{R}$ and that $\dim V = \dim W$. Then there exists an isomorphism $V \longrightarrow W$.*

Proof. Suppose that V and W are finite-dimensional vector spaces over $\mathbb{R}$ and that $\dim V = \dim W$. Let n be the dimension of V and W. We have isomorphisms $\varphi_V : \mathbb{R}^n \longrightarrow V$ and $\varphi_W : \mathbb{R}^n \longrightarrow W$. $\varphi_W \varphi_V^{-1} : V \to W$ and $\varphi_V \varphi_W^{-1} : W \to V$ are inverse linear transformations. Thus, $\varphi_W \varphi_V^{-1} : V \to W$ is an isomorphism. □

Theorem 22.1.4. *Let V and W be vector spaces over $\mathbb{R}$. Suppose $(v_1, v_2, \dots, v_n)$ is a basis for V and that $(w_1, w_2, \dots, w_n)$ is a tuple of vectors in W. Then there is exactly one linear transformation $T : V \longrightarrow W$ with the property that for all $j \in \{1, 2, \dots, n\}$,*

$$T(v_j) = w_j.$$

Proof. Let V and W be vector spaces over $\mathbb{R}$. Suppose $\beta = (v_1, v_2, \dots, v_n)$ is a basis for V and that $\gamma = (w_1, w_2, \dots, w_n)$ is a tuple of vectors in W. These choice of tuples determine the isomorphism $\Gamma_\beta : \mathbb{R}^n \longrightarrow V$, and the linear map $\Gamma_\gamma : \mathbb{R}^n \longrightarrow W$. We claim that $T = \Gamma_\gamma \Gamma_\beta^{-1} : V \longrightarrow W$ has the desired property. Let $j \in \{1, 2, \dots, n\}$. Since $\Gamma_\beta(e_j) = v_j$, we have $\Gamma_\beta^{-1}(v_j) = e_j$. Thus,

$$T(v_j) = \Gamma_\gamma \Gamma_\beta^{-1}(v_j) = \Gamma_\gamma(e_j) = w_j.$$

We now address the uniqueness of T. Suppose that S is linear and satisfies $S(v_j) = w_j$ for all $j \in \{1, 2, \ldots, n\}$ too. Let $v \in V$; we'll show that $S(v) = T(v)$. Because $(v_1, v_2, \ldots, v_n)$ spans V, we can find $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ such that $v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n$. Then

$$\begin{aligned} S(v) &= S(\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n) \\ &= \lambda_1 S(v_1) + \lambda_2 S(v_2) + \ldots + \lambda_n S(v_n) \\ &= \lambda_1 w_1 + \lambda_2 w_2 + \ldots + \lambda_n w_n. \\ &= \lambda_1 T(v_1) + \lambda_2 T(v_2) + \ldots + \lambda_n T(v_n) \\ &= T(\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n) \\ &= T(v). \end{aligned}$$

(I'll leave it to you to justify the equalities.) Since $v \in V$ was arbitrary, this shows $S = T$. $\square$

22.2 The matrix of a linear transformation

Remark 22.2.1. Let V and W be vector spaces over $\mathbb{R}$, and $T : V \longrightarrow W$ be a linear transformation. Suppose $\beta_V = (v_1, \ldots, v_n)$ is a basis for V and $\beta_W = (w_1, \ldots, w_m)$ is a basis for W.

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

Definition 22.2.2. Let V and W be vector spaces over $\mathbb{R}$, and $T : V \longrightarrow W$ be a linear transformation. Suppose $\beta_V = (v_1, \ldots, v_n)$ is a basis for V and $\beta_W = (w_1, \ldots, w_m)$ is a basis for W.

An matrix $(m \times n)$-matrix A is said to be *the matrix of the linear transformation T with respect to the bases β_V and β_W* **iff** for all $j \in \{1, \ldots, n\}$,

$$T(v_j) = \sum_{i=1}^{m} A_{ij} w_i. \tag{22.2.3}$$

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

Example 22.2.4. Define $T : \mathcal{P}_3(\mathbb{R}) \longrightarrow \mathcal{P}_6(\mathbb{R})$ by $T(p(x)) = (x^2 + 5x + 6) \cdot p(x)$.

Let $\beta_{\mathcal{P}_3(\mathbb{R})} = (1, x, x^2, 1 + x^3)$.

Let $\beta_{\mathcal{P}_6(\mathbb{R})} = (1, x, x^2, x^3, x^4, x^5 + 4x^4 + 5x^3, x^6 + x^5 + x^4 + x^3 + x^2 + x + 1)$.

What is $[T]_{\beta_{\mathcal{P}_3(\mathbb{R})}}^{\beta_{\mathcal{P}_6(\mathbb{R})}}$?

Let $f(x) = x^5 + 4x^4 + 5x^3$ and $g(x) = x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$ (to save space).

Here are the relevant equations:

$$\begin{aligned}
T(1) &= 6 \cdot 1 + 5 \cdot x + 1 \cdot x^2 + 0 \cdot x^3 + 0 \cdot x^4 + 0 \cdot f(x) + 0 \cdot g(x) \\
T(x) &= 0 \cdot 1 + 6 \cdot x + 5 \cdot x^2 + 1 \cdot x^3 + 0 \cdot x^4 + 0 \cdot f(x) + 0 \cdot g(x) \\
T(x^2) &= 0 \cdot 1 + 0 \cdot x + 6 \cdot x^2 + 5 \cdot x^3 + 1 \cdot x^4 + 0 \cdot f(x) + 0 \cdot g(x) \\
T(1 + x^3) &= 6 \cdot 1 + 5 \cdot x + 1 \cdot x^2 + 1 \cdot x^3 + 1 \cdot x^4 + 1 \cdot f(x) + 0 \cdot g(x)
\end{aligned}$$

So

$$[T]_{\beta_{\mathcal{P}_3(\mathbb{R})}}^{\beta_{\mathcal{P}_6(\mathbb{R})}} = \begin{pmatrix} 6 & 0 & 0 & 6 \\ 5 & 6 & 0 & 5 \\ 1 & 5 & 6 & 1 \\ 0 & 1 & 5 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

23 Questions due on August 27th

1. Find bases for the following subspaces of $\mathbb{R}^5$ and tell me their dimension:

$$W_1 = \left\{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 : \; x_1 - x_3 - x_4 = 0\right\},$$

$$W_2 = \left\{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 : \; x_2 = x_3 = x_4 \text{ and } x_1 + x_5 = 0\right\}.$$

Solution: The tuple

$$\beta_1 = ((0,1,0,0,0),\; (1,0,1,0,0),\; (1,0,0,1,0),\; (0,0,0,0,1))$$

is a linearly independent tuple of vectors in W_1. Part 3 of theorem 18.4 tells us that $\dim W_1 \geq 4$. Since $(1,0,0,0,0) \notin W_1$, theorem 20.1.1 gives $\dim W_1 < 5$. Thus, $\dim W_1 = 4$, and part 4 of theorem 18.4 tells us β_1 is a basis of W_1.

Let $T : \mathbb{R}^5 \to \mathbb{R}^3$ defined by $T(x_1, x_2, x_3, x_4, x_5) = (x_2 - x_3, x_4 - x_3, x_1 + x_5)$. Then T is linear and one can show that $W_2 = \ker T$.

$T(0,1,0,0,0) = (1,0,0)$, $T(0,0,0,1,0) = (0,1,0)$, $T(0,0,0,0,1) = (0,0,1)$, and so $\operatorname{im} T = \mathbb{R}^3$. Thus, $rank(T) = 3$. The rank-nullity theorem tells us $null(T) = 2$. So $\dim W_2 = 2$. Moreover,

$$\beta_2 = ((1,0,0,0,-1),\; (0,1,1,1,0))$$

is a linearly independent tuple of vectors in W_1. Part 4 of theorem 18.4 tells us β_2 is a basis of W_2.

2. Suppose V and W are finite-dimensional vector spaces over $\mathbb{R}$, and $T : V \longrightarrow W$ is a linear transformation. Use the rank-nullity to theorem to prove:

   (a) if T is injective, then $\dim(V) \leq \dim(W)$.

   (b) if T is surjective, then $\dim(V) \geq \dim(W)$;

Solution: Suppose V and W are finite-dimensional vector spaces over $\mathbb{R}$, and $T : V \longrightarrow W$ is a linear transformation.

(a) Suppose T is injective. Then $null(T) = 0$, so we have

$$\dim V = rank(T) + null(T) = rank(T) \leq \dim W.$$

The first equality is the rank-nullity theorem.
The second equality uses $null(T) = 0$. The fourth equality is because $\operatorname{im} T \subseteq W$.

(b) Suppose T is surjective. Then $\operatorname{im} T = W$, so we have

$$\dim V = rank(T) + null(T) \geq rank(T) = \dim W.$$

The first equality is the rank-nullity theorem.
The second equality uses $null(T) \geq 0$. The fourth equality is because $\operatorname{im} T = W$.

3. For each of the following, check T is linear, give a basis for $\ker T$ and $\operatorname{im} T$, and say whether T is injective or surjective.

   The following is almost always the most efficient way to proceed with such questions:

   - prove that T is a linear transformation;
   - write down a linearly independent tuple in $\ker T$ which seems as big as possible;
   - write down a linearly independent tuple in $\operatorname{im} T$ which seems as big as possible;
   - use theorem 18.4 (part 3) and rank-nullity to deduce what null(T) and rank(T) are;
   - use theorem 18.4 (part 4) to conclude that you have a basis for $\ker T$ and $\operatorname{im} T$;
   - say whether T is injective or surjective.

   (a) $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ defined by $T(x_1, x_2, x_3) = (x_1 - x_2, 2x_3)$.

   (b) $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ defined by $T(x_1, x_2) = (x_1 + x_2, 0, 2x_1 - x_2)$.

   (c) $T : M_{2\times 3}(\mathbb{R}) \longrightarrow M_{2\times 2}(\mathbb{R})$ defined by

$$T\left(\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{pmatrix}\right) = \begin{pmatrix} 2x_{11} - x_{12} & x_{13} + 2x_{12} \\ 0 & 0 \end{pmatrix}.$$

   Solution:

   (a)
   - T is a linear. Y'eard me.
   - $\beta_K = ((1, 1, 0))$ is a linearly independent tuple in $\ker T$.
   - $\beta_I = ((1, 0),\ (0, 2)) = (T(1, 0, 0),\ T(0, 0, 1))$ is a linearly independent tuple in $\operatorname{im} T$.
   - Theorem 18.4 (part 3) gives $null(T) \geq 1$ and $rank(T) \geq 2$. The rank-nullity theorem gives $rank(T) + null(T) = 3$. Thus, $null(T) = 1$ and $rank(T) = 2$.
   - Theorem 18.4 (part 4) tells us that β_K and β_I are bases for $\ker T$ and $\operatorname{im} T$, respectively.
   - T is surjective, but not injective.

   (b)
   - T is a linear. Have it.
   - $\beta_I = ((3, 0, 0),\ (0, 0, 3)) = (T(1, 2),\ T(1, -1))$ is linearly independent in $\operatorname{im} T$.
   - $null(T) \geq 0$. Theorem 18.4 (part 3) gives $rank(T) \geq 2$. The rank-nullity theorem gives $rank(T) + null(T) = 2$. Thus, $null(T) = 0$ and $rank(T) = 2$.
   - $(\ )$ is a basis for $\ker T$. Theorem 18.4 (part 4) tells us that β_I is a basis for $\operatorname{im} T$.
   - T is injective, but not surjective.

(c)
- T is linear. Big mood.
- $\beta_K = \left(\begin{pmatrix} 1 & 2 & -4 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right)$ is a linearly independent tuple in $\ker T$.
- $\beta_I = \left(\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right) = \left(T\left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right), T\left(\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \right) \right)$ is a linearly independent tuple in $\operatorname{im} T$.
- Theorem 18.4 (part 3) gives $null(T) \geq 4$ and $rank(T) \geq 2$. The rank-nullity theorem gives $rank(T) + null(T) = 2 \cdot 3 = 6$. Thus, $null(T) = 4$ and $rank(T) = 2$.
- Theorem 18.4 (part 4) tells us that β_K and β_I are bases for $\ker T$ and $\operatorname{im} T$, respectively.
- T is neither injective nor surjective.

4. Suppose U, V, and W are finite-dimensional vector spaces over $\mathbb{R}$, and that

$$T : U \longrightarrow V, \; S : V \longrightarrow W$$

are linear.

(a) Prove $null(ST) \geq null(T)$.

(b) Prove $rank(ST) \leq \min\{rank(S), rank(T)\}$.

Solution:

(a) We know from a previous homework that $\ker(ST) \supseteq \ker(T)$.
Applying $\dim(-)$ gives $null(ST) \geq null(T)$.

(b) We know from a previous homework that $\operatorname{im}(ST) \subseteq \operatorname{im}(S)$.
Applying $\dim(-)$ gives $rank(ST) \leq rank(S)$.
Part (a) together with the rank-nullity theorem gives

$$\dim(U) - rank(ST) \geq \dim(U) - rank(T).$$

So $rank(ST) \leq rank(T)$.
Together, the two inequalities concerning $rank(ST)$ show that

$$rank(ST) \leq \min\{rank(S), rank(T)\}.$$

Definition. Suppose V and W are vector spaces over $\mathbb{R}$, that $S : V \longrightarrow W$ and $T : V \longrightarrow W$ are linear transformations. Then $S + T : V \longrightarrow W$ is defined by $(S + T)(v) := S(v) + T(v)$.

5. Suppose V and W are finite-dimensional vector spaces over $\mathbb{R}$, and that

$$S : V \longrightarrow W,\ T : V \longrightarrow W$$

are linear.

(a) Prove $\operatorname{im}(S + T) \subseteq \operatorname{im}(S) + \operatorname{im}(T)$.

(b) Suppose U_1 and U_2 are subspaces of W. Prove that

$$\dim(U_1 + U_2) \leq \dim U_1 + \dim U_2.$$

(c) Prove $rank(S + T) \leq rank(S) + rank(T)$.

(d) Give examples where the inequality in (c) is equality, and when it is strict.

Solution: Suppose V and W are finite-dimensional vector spaces over $\mathbb{R}$, and that

$$S : V \longrightarrow W,\ T : V \longrightarrow W$$

are linear.

(a) Let $w \in \operatorname{im}(S+T)$. Then there is a $v \in V$ such that $w = (S+T)(v)$, i.e. $w = S(v)+T(v)$. Since $S(v) \in \operatorname{im} S$ and $T(v) \in \operatorname{im} T$, this shows that $w \in \operatorname{im} S + \operatorname{im} T$.

(b) Suppose U_1 and U_2 are subspaces of W. Let $m_1 = \dim U_1$ and $m_2 = \dim U_2$, and choose bases for U_1 and U_2, respectively: $(u_{1,1}, u_{1,2}, \ldots, u_{1,m_1})$ and $(u_{2,1}, u_{2,2}, \ldots, u_{2,m_2})$.
By a previous homework, $\operatorname{span}(u_{1,1}, u_{1,2}, \ldots, u_{1,m_1}, u_{2,1}, u_{2,2}, \ldots, u_{2,m_2})$ is equal to

$$\operatorname{span}(u_{1,1}, u_{1,2}, \ldots, u_{1,m_1}) + \operatorname{span}(u_{2,1}, u_{2,2}, \ldots, u_{2,m_2}) = U_1 + U_2.$$

Thus, $U_1 + U_2$ has a spanning tuple of size $m_1 + m_2$, and so the dimension of $U_1 + U_2$ is less than or equal to $m_1 + m_2$ (theorem 18.4 part 1), i.e.

$$\dim(U_1 + U_2) \leq \dim U_1 + \dim U_2.$$

(c) We have

$$\begin{aligned} rank(S + T) = \dim(\operatorname{im}(S + T)) &\leq \dim(\operatorname{im}(S) + \operatorname{im}(T)) \\ &\leq \dim(\operatorname{im}(S)) + \dim(\operatorname{im}(T)) = rank(S) + rank(T). \end{aligned}$$

The equalities are definitional.
The first inequality follows form part (a). The second follows from part (b).

(d) Let $V = W = \mathbb{R}$ and $S = 1_{\mathbb{R}}$.
When $T = 0$, we have

$$rank(S + T) = rank(1_{\mathbb{R}}) = 1 = 1 + 0 = rank(S) + rank(T).$$

When $T = -1_{\mathbb{R}}$, we have

$$rank(S + T) = rank(0) = 0 < 2 = rank(S) + rank(T).$$

24 Lecture on August 27th

24.1 The matrix of a linear transformation

Remark 24.1.1. 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.

Example 24.1.2. Suppose $m, n \in \mathbb{N}$ and A is an $(m\times n)$-matrix. We have the linear transformation $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Because

$$T_A(e_j) = Ae_j = j\text{-th column of } A = \sum_{i=1}^{m} A_{ij}e_i,$$

the matrix of T_A with respect to the bases $(e_1, \ldots, e_n)$ and $(e_1, \ldots, e_m)$ is the original matrix A.

Notation 24.1.3. Suppose V is a vector space over $\mathbb{R}$, and that $\beta = (v_1, v_2, \ldots, v_n)$ is a basis for V. We have the isomorphism $\Gamma_\beta : \mathbb{R}^n \longrightarrow V$.

For the rest of the class, for $v \in V$, we will use the notation $[v]_\beta$ for $\Gamma_\beta^{-1}(v)$, so that

$$[\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n]_\beta = (\lambda_1, \lambda_2, \ldots, \lambda_n) \quad \text{and} \quad \Gamma_\beta([v]_\beta) = v.$$

The first equation shows that $[v]_\beta$ returns the *coordinates* of an abstract vector $v \in V$ with respect to the basis β. The second equation will be useful in the proof of the next theorem.

Theorem 24.1.4. *Let V and W be vector spaces over $\mathbb{R}$, and $T : V \longrightarrow W$ be a linear transformation. Suppose $\beta_V = (v_1, \ldots, v_n)$ is a basis for V and $\beta_W = (w_1, \ldots, w_m)$ is a basis for W. Then, for all $v \in V$,*

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

Proof. Let V, W, T, β_V, β_W be as in the theorem statement, and let $A = [T]_{\beta_V}^{\beta_W}$ be the matrix of the linear transformation T with respect to the bases β_V and β_W.

Consider the following diagram of linear maps.

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "T_A"] \arrow[d, "\Gamma_{\beta_V}"'] & \mathbb{R}^m \arrow[d, "\Gamma_{\beta_W}"] \\
V \arrow[r, "T"] & W
\end{tikzcd}$$

First, we will show that $\Gamma_{\beta_W} \circ T_A = T \circ \Gamma_{\beta_V}$. By theorem 22.1.4, it is enough to check this on the basis elements $e_1, e_2, \ldots, e_n \in \mathbb{R}^n$.

$$
\begin{tikzcd}
e_j \arrow[rrr, mapsto] \arrow[ddd, mapsto] & & & \sum_{i=1}^m A_{ij} e_i \arrow[ddd, mapsto] \\
& \mathbb{R}^n \arrow[r, "T_A"] \arrow[d, "\Gamma_{\beta_V}"'] & \mathbb{R}^m \arrow[d, "\Gamma_{\beta_W}"] & \\
& V \arrow[r, "T"] & W & \\
v_j \arrow[rrr, mapsto] & & & \sum_{i=1}^m A_{ij} w_i
\end{tikzcd}
$$

So let $j \in \{1, 2, \ldots, n\}$. Then

$$
\begin{aligned}
(\Gamma_{\beta_W} \circ T_A)(e_j) = \Gamma_{\beta_W}(T_A(e_j)) = \Gamma_{\beta_W}(A(e_j)) &= \Gamma_{\beta_W}\left(\sum_{i=1}^m A_{ij} e_i\right) \\
&= \sum_{i=1}^m A_{ij} \cdot \Gamma_{\beta_W}(e_i) \\
&= \sum_{i=1}^m A_{ij} w_i = T(v_j) = T(\Gamma_{\beta_V}(e_j)) = (T \circ \Gamma_{\beta_V})(e_j).
\end{aligned}
$$

Remark 24.1.5. The previous diagram is actually the entire point of the matrix A. Given T, β_V, β_W, the function $\Gamma_{\beta_W}^{-1} \circ T \circ \Gamma_{\beta_V}$ is a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$. Thus, by corollary 9.2.8, there is a unique matrix A such that $T_A = \Gamma_{\beta_W}^{-1} \circ T \circ \Gamma_{\beta_V}$. It is then immediate that $\Gamma_{\beta_W} \circ T_A = T \circ \Gamma_{\beta_V}$. We could have used these observations to define $[T]_{\beta_V}^{\beta_W}$, but I thought the more calculational definition would be more comfortable you.

We are now ready to prove the result. Let $v \in V$. We wish to show that $[T]_{\beta_V}^{\beta_W}[v]_{\beta_V} = [T(v)]_{\beta_W}$,

$$
\text{i.e. } T_A([v]_{\beta_V}) = [T(v)]_{\beta_W}
$$

Since Γ_{β_W} is injective, it is enough to show that

$$
\Gamma_{\beta_W}(T_A([v]_{\beta_V})) = \Gamma_{\beta_W}([T(v)]_{\beta_W}). \tag{24.1.6}
$$

By the previous calculation, the LHS is equal to $T(\Gamma_{\beta_V}([v]_{\beta_V}))$.

Now recall that for all $v \in V$, $\Gamma_{\beta_V}([v]_{\beta_V}) = v$, and for all $w \in W$, $\Gamma_{\beta_W}([w]_{\beta_W}) = w$. Thus, both sides of (24.1.6) are equal to $T(v)$. This finishes the proof. $\square$

Theorem 24.1.7. *Let U, V, and W be vector spaces over $\mathbb{R}$, $T : U \longrightarrow V$, $S : V \longrightarrow W$ be 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, and W, respectively. Then*

$$[S]^{\beta_W}_{\beta_V} [T]^{\beta_V}_{\beta_U} = [ST]^{\beta_W}_{\beta_U}.$$

Proof. Let U, V, W, S, T, β_U, β_V, and β_W be as in the theorem statement. For all $u \in U$, we have

$$\begin{aligned}([S]^{\beta_W}_{\beta_V} [T]^{\beta_V}_{\beta_U})[u]_{\beta_U} &= [S]^{\beta_W}_{\beta_V} ([T]^{\beta_V}_{\beta_U} [u]_{\beta_U}) \\ &= [S]^{\beta_W}_{\beta_V} [T(u)]_{\beta_V} = [S(T(u))]_{\beta_W} = [(ST)(u)]_{\beta_W} = [ST]^{\beta_W}_{\beta_U} [u]_{\beta_U}.\end{aligned}$$

The first equality uses the fact that for an $(m \times n)$-matrix A, an $(n \times p)$-matrix B, and an element $x \in \mathbb{R}^p$, $(AB)x = A(Bx)$. The second, third, and fifth equalities all use the previous theorem. The fourth equality uses the definition of ST.

Let $j \in \{1, 2, \ldots, p\}$. We have $[u_j]_{\beta_U} = e_j$. So setting $u = u_j$ in the equation above tells us that the j-th column of $[S]^{\beta_W}_{\beta_V} [T]^{\beta_V}_{\beta_U}$ is equal to the j-th column of $[ST]^{\beta_W}_{\beta_U}$. □

24.2 The purpose of matrices and vectors

Remark 24.2.1. In math 33A, you learn about matrices, matrix-vector multiplication, and matrix-matrix multiplication. For this reason, I have never defined these for you. You know what they are. The definitions of matrix-vector and matrix-matrix multiplication are chosen so that the following properties hold:

- for $A \in M_{m \times n}(\mathbb{R})$, the function $T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $x \longmapsto Ax$ is linear;
- for $A \in M_{m \times n}(\mathbb{R})$, $j \in \{1, 2, \ldots, n\}$, we have $Ae_j = j$-th column of A;
- for $A \in M_{m \times n}(\mathbb{R})$, $B \in M_{n \times p}(\mathbb{R})$, $x \in \mathbb{R}^p$, we have $(AB)x = A(Bx)$.

Another way of saying the third property is:

- for $A \in M_{m \times n}(\mathbb{R})$, $B \in M_{n \times p}(\mathbb{R})$, we have $T_{AB} = T_A T_B$.

In fact, one can check that the first two properties force matrix-vector multiplication to be what it is. The second and third property then force matrix-matrix multiplication to be what it is... Let $j \in \{1, \ldots, p\}$. The third property gives $(AB)e_j = A(Be_j)$. Together with the second property, this tells us that the j-th column of AB is given by multiplying A with the j-th column of B.

We can say more... I can't imagine caring about matrices if the first property did not hold. The second property is needed in example 24.1.2 and to prove theorem 24.1.4. Together these two properties ensure that linear transformations and matrices are the abstract and coordinate-ified versions of each other. The third property is then needed to prove theorem 24.1.7.

24.3 Isomorphisms and invertible matrices

Notation 24.3.1. We write I_n for the $n \times n$ *identity matrix.*

Definition 24.3.2. We say a matrix $A \in M_{n\times n}(\mathbb{R})$ is *invertible* **iff** there is a matrix $B \in M_{n\times n}(\mathbb{R})$ 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.

Theorem 24.3.3. *Let V and W be vector spaces over $\mathbb{R}$, and $T : V \longrightarrow W$ be 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. If T is an isomorphism, then $[T]_{\beta_V}^{\beta_W}$ is invertible.*

Proof. Let V, W, T, β_V, β_W be as in the theorem statement. Suppose that T is an isomorphism.

From theorem 22.1.1, we know $\dim V = \dim W$, so $[T]_{\beta_V}^{\beta_W}$ is a square matrix.

Also, T has an inverse linear transformation $S : W \longrightarrow V$ with $ST = 1_V$ and $TS = 1_W$. Using theorem 24.1.7, we obtain

$$[S]_{\beta_W}^{\beta_V}[T]_{\beta_V}^{\beta_W} = [1_V]_{\beta_V}^{\beta_V} = I_n \quad\text{and}\quad [T]_{\beta_V}^{\beta_W}[S]_{\beta_W}^{\beta_V} = [1_W]_{\beta_W}^{\beta_W} = I_n,$$

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

Theorem 24.3.4. *Suppose A is an $(m \times n)$-matrix. $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is an isomorphism if and only if A is invertible.*

Proof. Suppose A is an $(m \times n)$-matrix.

First, suppose T_A is an isomorphism. From theorem 22.1.1, we know $n = m$. Let $\beta = (e_1, \ldots, e_n)$ be the standard basis of $\mathbb{R}^n$. The previous theorem tells us that $[T_A]_\beta^\beta$ is invertible. Example 24.1.2 shows $[T_A]_\beta^\beta = A$, and so we conclude A is invertible.

Conversely, suppose that A is invertible. Then $n = m$ and there exists another matrix B such that $AB = BA = I_n$. We have $T_AT_B = T_{AB} = T_{I_n} = 1_{\mathbb{R}^n}$ and $T_BT_A = T_{BA} = T_{I_n} = 1_{\mathbb{R}^n}$. Thus, T_A is an isomorphism. □

Theorem 24.3.5. *Suppose $A, B \in M_{n\times n}(\mathbb{R})$. If $AB = I_n$, then $BA = I_n$.*

Proof. Suppose $A, B \in M_{n\times n}(\mathbb{R})$ and that $AB = I_n$. Then $T_AT_B = T_{AB} = T_{I_n} = 1_{\mathbb{R}^n}$. Thus, given $v \in \mathbb{R}^n$, we have $v = 1_{\mathbb{R}^n}(v) = T_A(T_B(v))$, and so T_A is surjective. By theorem 20.2.4, T_A is also injective. So T_A is an isomorphism, and the previous theorem tells us that A is invertible.

Let C be an inverse for A, so that $CA = I_n$. Then

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

Here, the second equality uses $CA = I_n$ and the fourth uses $AB = I_n$. The equations $B = C$ and $CA = I_n$ show that $BA = I_n$. □

Theorem 24.3.6. *Let V and W be vector spaces over $\mathbb{R}$, and $T : V \longrightarrow W$ be 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. If $[T]_{\beta_V}^{\beta_W}$ is invertible, then T is an isomorphism*

Proof. Let V, W, T, β_V, β_W be as in the theorem statement and suppose that $[T]_{\beta_V}^{\beta_W}$ is invertible. Let $A = [T]_{\beta_V}^{\beta_W}$, so that A is invertible. Theorem 24.3.4 says that T_A is an isomorphism.

We saw in a previous proof that $T = \Gamma_{\beta_W} \circ T_A \circ \Gamma_{\beta_V}^{-1}$.

T is an isomorphism because it has an inverse $\Gamma_{\beta_V} \circ (T_A)^{-1} \circ \Gamma_{\beta_W}^{-1}$. □

24.4 Sean's proof of theorem 24.1.4

The proof of theorem 24.1.4 may seem weird to you...

Category theory is an area of mathematics that has infiltrated my life since I was 21, and it leads me to think a certain way. This way of thinking means that if V is a vector space over $\mathbb{R}$, I prefer linear transformations $\mathbb{R}^n \longrightarrow V$ over linear transformations $V \longrightarrow \mathbb{R}^n$. That's one reason for the proof I gave.

I also took the proof of theorem 24.1.4 as an excuse to introduce a diagram I find particularly useful. The extent to which you find it useful will vary, but if you do lots more math, I expect it'll come back to get you! I used this diagram in a later proof.

Having said all that (in some defence of the proof I gave) a student presented to me a delightful proof of theorem 24.1.4 in office hours, and you will probably prefer his proof to mine. Here it is...

Sean's proof of theorem 24.1.4. Let V, W, T, β_V, β_W be as in the theorem statement, and let $A = [T]_{\beta_V}^{\beta_W}$ be the matrix of the linear transformation T with respect to the bases β_V and β_W.

Define $S_1 : V \longrightarrow \mathbb{R}^m$ and $S_2 : V \longrightarrow \mathbb{R}^m$ by

$$S_1(v) = A[v]_{\beta_V} \quad \text{and} \quad S_2(v) = [T(v)]_{\beta_W}.$$

We just have to show that $S_1 = S_2$. Since S_1 and S_2 are linear, by theorem 22.1.4, it is enough to check this on the basis elements $v_1, v_2, \ldots, v_n \in V$.

Let $j \in \{1, 2, \ldots, n\}$. Then

$$S_1(v_j) = A[v_j]_{\beta_V} = Ae_j = \sum_{i=1}^{m} A_{ij} e_i.$$

The first equality uses the definition of S_1. The second uses $[v_j]_{\beta_V} = e_j$. The third expresses the fact that $Ae_j =$ the j-th column of A.

Also,

$$S_2(v_j) = [T(v_j)]_{\beta_W} = \left[\sum_{i=1}^{m} A_{ij} w_i\right]_{\beta_W} = \sum_{i=1}^{m} A_{ij} [w_i]_{\beta_W} = \sum_{i=1}^{m} A_{ij} e_i.$$

The first equality is the definition of S_2. The second uses the definition of $A = [T]_{\beta_V}^{\beta_W}$. The third uses linearity of $[-]_{\beta_W}$. The fourth uses $[w_i]_{\beta_W} = e_i$. $\square$

Well done, Sean!

25 Questions due on August 29th

1. On the next page, in each part, I define a linear map $T : V \longrightarrow W$, give bases β_V, β_W, and an element $v \in V$. You should:

   - calculate $[v]_{\beta_V}$,
   - calculate $T(v)$,
   - calculate $[T(v)]_{\beta_W}$,
   - calculate $[T]_{\beta_V}^{\beta_W}$,
   - confirm that $[T]_{\beta_V}^{\beta_W}[v]_{\beta_V} = [T(v)]_{\beta_W}$.

(a) $T : \mathcal{P}_5(\mathbb{R}) \longrightarrow \mathcal{P}_4(\mathbb{R})$ is defined by $p(x) \longmapsto p'(x)$.
$\beta_{\mathcal{P}_5(\mathbb{R})} = (1, x, x^2, x^3, x^4, x^5)$, $\beta_{\mathcal{P}_4(\mathbb{R})} = (1, x, x^2, x^3, x^4)$.
$v = 4 + 3x - 2x^2 + x^3 + 5x^4 - 8x^5$.

Remark. A polynomial is *not* a function by definition, but you can do something which looks a lot like differentiation to polynomials by enforcing the "power rule." This is a *formal* operation which coincides with differentiation once we allow polynomials to define a function.

(b) $T : \mathcal{P}_5(\mathbb{R}) \longrightarrow \mathcal{P}_4(\mathbb{R})$ is defined by $p(x) \longmapsto p'(x)$.
$\beta_{\mathcal{P}_5(\mathbb{R})} = (1, x, x^2, x^3, x^4, x^5)$, $\beta_{\mathcal{P}_4(\mathbb{R})} = (x^4, x^3, x^2, x, 1)$,
$v = 4 + 3x - 2x^2 + x^3 + 5x^4 - 8x^5$.

(c) $T : \mathcal{P}_5(\mathbb{R}) \longrightarrow \mathcal{P}_4(\mathbb{R})$ is defined by $p(x) \longmapsto p'(x)$.
$\beta_{\mathcal{P}_5(\mathbb{R})} = (x^5, x^4, x^3, x^2, x, 1)$, $\beta_{\mathcal{P}_4(\mathbb{R})} = (x^4, x^3, x^2, x, 1)$,
$v = 4 + 3x - 2x^2 + x^3 + 5x^4 - 8x^5$.

(d) $T : \mathcal{P}_5(\mathbb{R}) \longrightarrow \mathcal{P}_4(\mathbb{R})$ is defined by $p(x) \longmapsto p'(x)$.
$\beta_{\mathcal{P}_5(\mathbb{R})} = (1, 1+x, 1+x+x^2, 1+x+x^2+x^3, 1+x+x^2+x^3+x^4, 1+x+x^2+x^3+x^4+x^5)$,
$\beta_{\mathcal{P}_4(\mathbb{R})} = (1, x, x^2, x^3, x^4)$,
$v = 4 + 3x - 2x^2 + x^3 + 5x^4 - 8x^5$.

(e) $T : \mathcal{P}_5(\mathbb{R}) \longrightarrow \mathcal{P}_4(\mathbb{R})$ is defined by $p(x) \longmapsto p'(x)$.
$\beta_{\mathcal{P}_5(\mathbb{R})} = (1, x, x^2, x^3, x^4, x^5)$,
$\beta_{\mathcal{P}_4(\mathbb{R})} = (1, 1+x, 1+x+x^2, 1+x+x^2+x^3, 1+x+x^2+x^3+x^4)$,
$v = 4 + 3x - 2x^2 + x^3 + 5x^4 - 8x^5$.

(f) $T : \mathcal{P}_4(\mathbb{R}) \longrightarrow \mathcal{P}_5(\mathbb{R})$ is defined by $p(x) \longmapsto \int_0^x p(t)\ dt$.
$\beta_{\mathcal{P}_5(\mathbb{R})} = (1, x, x^2, x^3, x^4, x^5)$, $\beta_{\mathcal{P}_4(\mathbb{R})} = (1, x, x^2, x^3, x^4)$,
$v = 8 + 4x - 3x^2 - 4x^3 + 120x^4$.

Remark. A polynomial is *not* a function by definition. However, you can do something which looks a lot like integration to polynomials by enforcing the "power rule." This is a *formal* operation which coincides with integration once we allow polynomials to define a function.

(g) $T : \mathcal{P}_4(\mathbb{R}) \longrightarrow \mathcal{P}_5(\mathbb{R})$ is defined by $p(x) \longmapsto \int_1^x p(t)\ dt$.
$\beta_{\mathcal{P}_5(\mathbb{R})} = (1, x, x^2, x^3, x^4, x^5)$, $\beta_{\mathcal{P}_4(\mathbb{R})} = (1, x, x^2, x^3, x^4)$,
$v = 8 + 4x - 3x^2 - 4x^3 + 120x^4$.

(h) In this part, $V = W$ and $\beta_V = \beta_W$.

Let $A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ and $T = T_A : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$.

$\beta_{\mathbb{R}^3} = \left(\begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right)$, $v = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$.

2. Below, in each part, I define linear maps $T : U \longrightarrow V$, $S : V \longrightarrow W$, and give bases β_U, β_V, β_W. You should:

   - calculate $[T]_{\beta_U}^{\beta_V}$,
   - calculate $[S]_{\beta_V}^{\beta_W}$,
   - calculate ST,
   - calculate $[ST]_{\beta_U}^{\beta_W}$,
   - confirm that $[S]_{\beta_V}^{\beta_W}[T]_{\beta_U}^{\beta_V} = [ST]_{\beta_U}^{\beta_W}$.

   (a) In this part, $U = W$ and $\beta_U = \beta_W$.
   $T : \mathcal{P}_4(\mathbb{R}) \longrightarrow \mathcal{P}_5(\mathbb{R})$ is defined by $p(x) \longmapsto \int_0^x p(t)\ dt$.
   $S : \mathcal{P}_5(\mathbb{R}) \longrightarrow \mathcal{P}_4(\mathbb{R})$ is defined by $p(x) \longmapsto p'(x)$.
   $\beta_{\mathcal{P}_4(\mathbb{R})} = (1, x, x^2, x^3, x^4)$, $\beta_{\mathcal{P}_5(\mathbb{R})} = (1, x, x^2, x^3, x^4, x^5)$.

   (b) In this part, $U = W$ and $\beta_U = \beta_W$.
   $T : \mathcal{P}_5(\mathbb{R}) \longrightarrow \mathcal{P}_4(\mathbb{R})$ is defined by $p(x) \longmapsto p'(x)$.
   $S : \mathcal{P}_4(\mathbb{R}) \longrightarrow \mathcal{P}_5(\mathbb{R})$ is defined by $p(x) \longmapsto \int_0^x p(t)\ dt$.
   $\beta_{\mathcal{P}_5(\mathbb{R})} = (1, x, x^2, x^3, x^4, x^5)$, $\beta_{\mathcal{P}_4(\mathbb{R})} = (1, x, x^2, x^3, x^4)$.

   (c) In this part, $U = W$ and $\beta_U = \beta_W$.
   $T : \mathcal{P}_4(\mathbb{R}) \longrightarrow \mathcal{P}_5(\mathbb{R})$ is defined by $p(x) \longmapsto \int_1^x p(t)\ dt$.
   $S : \mathcal{P}_5(\mathbb{R}) \longrightarrow \mathcal{P}_4(\mathbb{R})$ is defined by $p(x) \longmapsto p'(x)$.
   $\beta_{\mathcal{P}_4(\mathbb{R})} = (1, x, x^2, x^3, x^4)$, $\beta_{\mathcal{P}_5(\mathbb{R})} = (1, x, x^2, x^3, x^4, x^5)$.

   (d) In this part, $U = W$ and $\beta_U = \beta_W$.
   $T : \mathcal{P}_5(\mathbb{R}) \longrightarrow \mathcal{P}_4(\mathbb{R})$ is defined by $p(x) \longmapsto p'(x)$.
   $S : \mathcal{P}_4(\mathbb{R}) \longrightarrow \mathcal{P}_5(\mathbb{R})$ is defined by $p(x) \longmapsto \int_1^x p(t)\ dt$.
   $\beta_{\mathcal{P}_5(\mathbb{R})} = (1, x, x^2, x^3, x^4, x^5)$, $\beta_{\mathcal{P}_4(\mathbb{R})} = (1, x, x^2, x^3, x^4)$.

   (e) [Optional]
   I suppose I never thought up one where $U \neq W \ldots$
   $T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathcal{P}_3(\mathbb{R})$ is defined by $T(p(x)) = (3 + x)p'(x) + 2p(x)$.
   $S : \mathcal{P}_3(\mathbb{R}) \longrightarrow \mathbb{R}^3$ is defined by $S(a + bx + cx^2 + dx^3) = (a + b, c, a - b)$.
   $\beta_{\mathcal{P}_2(\mathbb{R})} = (1, x, x^2)$, $\beta_{\mathcal{P}_3(\mathbb{R})} = (1, x, x^2, x^3)$, $\beta_{\mathbb{R}^3} = (e_1, e_2, e_3)$.

26 Solutions to the previous questions

1. (a)
   - $[4+3x-2x^2+x^3+5x^4-8x^5]_{\beta_{\mathcal{P}_5(\mathbb{R})}} = (4,3,-2,1,5,-8)$,
   - $T(4+3x-2x^2+x^3+5x^4-8x^5) = 3-4x+3x^2+20x^3-40x^4$,
   - $[T(4+3x-2x^2+x^3+5x^4-8x^5)]_{\beta_{\mathcal{P}_4(\mathbb{R})}} = (3,-4,3,20,-40)$,
   - $[T]_{\beta_{\mathcal{P}_5(\mathbb{R})}}^{\beta_{\mathcal{P}_4(\mathbb{R})}} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}$.

   (b)
   - $[4+3x-2x^2+x^3+5x^4-8x^5]_{\beta_{\mathcal{P}_5(\mathbb{R})}} = (4,3,-2,1,5,-8)$,
   - $T(4+3x-2x^2+x^3+5x^4-8x^5) = 3-4x+3x^2+20x^3-40x^4$,
   - $[T(4+3x-2x^2+x^3+5x^4-8x^5)]_{\beta_{\mathcal{P}_4(\mathbb{R})}} = (-40,20,3,-4,3)$,
   - $[T]_{\beta_{\mathcal{P}_5(\mathbb{R})}}^{\beta_{\mathcal{P}_4(\mathbb{R})}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$.

   (c)
   - $[4+3x-2x^2+x^3+5x^4-8x^5]_{\beta_{\mathcal{P}_5(\mathbb{R})}} = (-8,5,1,-2,3,4)$,
   - $T(4+3x-2x^2+x^3+5x^4-8x^5) = 3-4x+3x^2+20x^3-40x^4$,
   - $[T(4+3x-2x^2+x^3+5x^4-8x^5)]_{\beta_{\mathcal{P}_4(\mathbb{R})}} = (-40,20,3,-4,3)$,
   - $[T]_{\beta_{\mathcal{P}_5(\mathbb{R})}}^{\beta_{\mathcal{P}_4(\mathbb{R})}} = \begin{pmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$.

   (d)
   - $[4+3x-2x^2+x^3+5x^4-8x^5]_{\beta_{\mathcal{P}_5(\mathbb{R})}} = (1,5,-3,-4,13,-8)$,
   - $T(4+3x-2x^2+x^3+5x^4-8x^5) = 3-4x+3x^2+20x^3-40x^4$,
   - $[T(4+3x-2x^2+x^3+5x^4-8x^5)]_{\beta_{\mathcal{P}_4(\mathbb{R})}} = (3,-4,3,20,-40)$,
   - $[T]_{\beta_{\mathcal{P}_5(\mathbb{R})}}^{\beta_{\mathcal{P}_4(\mathbb{R})}} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 & 3 \\ 0 & 0 & 0 & 0 & 4 & 4 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}$.

   (e)
   - $[4+3x-2x^2+x^3+5x^4-8x^5]_{\beta_{\mathcal{P}_5(\mathbb{R})}} = (4,3,-2,1,5,-8)$,
   - $T(4+3x-2x^2+x^3+5x^4-8x^5) = 3-4x+3x^2+20x^3-40x^4$,
   - $[T(4+3x-2x^2+x^3+5x^4-8x^5)]_{\beta_{\mathcal{P}_4(\mathbb{R})}} = (7,-7,-17,60,-40)$,
   - $[T]_{\beta_{\mathcal{P}_5(\mathbb{R})}}^{\beta_{\mathcal{P}_4(\mathbb{R})}} = \begin{pmatrix} 0 & 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 2 & -3 & 0 & 0 \\ 0 & 0 & 0 & 3 & -4 & 0 \\ 0 & 0 & 0 & 0 & 4 & -5 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}$.

(f)
- $[8+4x-3x^2-4x^3+120x^4]_{\beta_{\mathcal{P}_4(\mathbb{R})}} = (8,4,-3,-4,120)$,
- $T(8+4x-3x^2-4x^3+120x^4) = 8x+2x^2-x^3-x^4+24x^5$,
- $[T(v)]_{\beta_{\mathcal{P}_5(\mathbb{R})}} = (0,8,2,-1,-1,24)$,
- $[T]_{\beta_{\mathcal{P}_4(\mathbb{R})}}^{\beta_{\mathcal{P}_5(\mathbb{R})}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{5} \end{pmatrix}$.

(g)
- $[8+4x-3x^2-4x^3+120x^4]_{\beta_{\mathcal{P}_4(\mathbb{R})}} = (8,4,-3,-4,120)$,
- $T(8+4x-3x^2-4x^3+120x^4) = -32+8x+2x^2-x^3-x^4+24x^5$,
- $[T(v)]_{\beta_{\mathcal{P}_5(\mathbb{R})}} = (-32,8,2,-1,-1,24)$,
- $[T]_{\beta_{\mathcal{P}_4(\mathbb{R})}}^{\beta_{\mathcal{P}_5(\mathbb{R})}} = \begin{pmatrix} -1 & -\frac{1}{2} & -\frac{1}{3} & -\frac{1}{4} & -\frac{1}{5} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{5} \end{pmatrix}$.

(h)
- $[(0,1,2)]_{\beta_{\mathbb{R}^3}} = (0,1,1)$,
- $T(0,1,2) = (-1,2,3)$,
- $[T(0,1,2)]_{\beta_{\mathbb{R}^3}} = (0,1,2)$,
- $[T]_{\beta_{\mathbb{R}^3}}^{\beta_{\mathbb{R}^3}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

(i) This was not a question but it provides more examples of coordinate change matrices. In (a), we calculated the matrix of $T : \mathcal{P}_5(\mathbb{R}) \to \mathcal{P}_4(\mathbb{R})$ defined by $p(x) \longmapsto p'(x)$ with respect to the standard basis for $\mathcal{P}_4(\mathbb{R})$ and $\mathcal{P}_5(\mathbb{R})$. It's

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}.$$

Then in questions (b-e), we calculated the matrix with respect to different bases. Here are the relevant versions of the formula $A' = P^{-1}AQ$:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}^{-1} A \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}^{-1} A \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 & 3 \\ 0 & 0 & 0 & 0 & 4 & 4 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}^{-1} A \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & -2 & 0 & 0 & 0 \\ 0 & 0 & 2 & -3 & 0 & 0 \\ 0 & 0 & 0 & 3 & -4 & 0 \\ 0 & 0 & 0 & 0 & 4 & -5 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}^{-1} A \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

In (h), we learned that

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 1 & -1 \\ -2 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 1 & -1 \\ -2 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

2. (a) $ST = 1_{\mathcal{P}_4(\mathbb{R})}$.

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{5} \end{pmatrix} = I_5.$$

(b) $(ST)(p(x)) = p(x) - p(0)$.

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{5} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

(c) $ST = 1_{\mathcal{P}_4(\mathbb{R})}$.

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} -1 & -\frac{1}{2} & -\frac{1}{3} & -\frac{1}{4} & -\frac{1}{5} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{5} \end{pmatrix} = I_5.$$

(d) $(ST)(p(x)) = p(x) - p(1)$.

$$\begin{pmatrix} -1 & -\frac{1}{2} & -\frac{1}{3} & -\frac{1}{4} & -\frac{1}{5} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{5} \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -1 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

27 Lecture on August 29th

27.1 The change of coordinate matrix

Definition 27.1.1. Suppose V is a vector space over $\mathbb{R}$, 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 27.1.2. Suppose V is a vector space over $\mathbb{R}$, that $\beta = (v_1, \ldots, v_n)$ and $\beta' = (v'_1, \ldots, v'_n)$ are bases for V. Let $Q = [1_V]^{\beta}_{\beta'}$. Then, for $j \in \{1, 2, \ldots, n\}$, 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 27.1.3. *Let V and W be vector spaces over $\mathbb{R}$, and $T : V \longrightarrow W$ be a linear transformation. Suppose $\beta_V = (v_1, \ldots, v_n)$ and $\beta' = (v'_1, \ldots, v'_n)$ are bases for V, and $\beta_W = (w_1, \ldots, w_m)$ and $\beta' = (w'_1, \ldots, w'_m)$ are bases for W.*

Suppose A is the matrix of T with respect to the bases β_V and β_W, and that A' 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. $A' = [T]^{\beta'_W}_{\beta'_V} = [1_W T 1_V]^{\beta'_W}_{\beta'_V} = [1_W]^{\beta'_W}_{\beta_W}[T]^{\beta_W}_{\beta_V}[1_V]^{\beta_V}_{\beta'_V} = ([1_W]^{\beta_W}_{\beta'_W})^{-1}[T]^{\beta_W}_{\beta_V}[1_V]^{\beta_V}_{\beta'_V} = P^{-1}AQ$. $\square$

Corollary 27.1.4. *Suppose $m, n \in \mathbb{N}$, $A \in M_{m\times n}(\mathbb{R})$, that $\beta_{\mathbb{R}^n} = (v_1, \ldots, v_n)$ is a basis of $\mathbb{R}^n$, and $\beta_{\mathbb{R}^m} = (w_1, \ldots, w_m)$ is a basis of $\mathbb{R}^m$. Then the matrix of the linear transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ 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 27.1.5. See the previous questions' solutions for examples: 1i.

27.2 Determinants of matrices (not lectured)

Notation 27.2.1. Suppose $A \in M_{n\times n}(\mathbb{R})$. 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 27.2.2. *For each $n \in \mathbb{N}$, there is a function called the* determinant, $\det : M_{n\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$. *The functions have the following properties.*

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

   (b) *If* $A \in M_{n\times n}(\mathbb{R})$ *where* $n \geq 2$, *and* $j \in \{1, \ldots, n\}$, *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* $j \in \{1, \ldots, n\}$, $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n \in \mathbb{R}^n$, $u, w \in \mathbb{R}^n$, *and* $\lambda \in \mathbb{R}$. *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* $j, k \in \{1, \ldots, n\}$, $j < k$, $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n \in \mathbb{R}^n$, $v \in \mathbb{R}^n$. *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* $j \in \{1, \ldots, n\}$ *and* $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n \in \mathbb{R}^n$. *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* $j, k \in \{1, \ldots, n\}$, $j < k$, *and* $v_1, \ldots, v_n \in \mathbb{R}^n$. *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* $j, k \in \{1, \ldots, n\}$, $j \neq k$, $v_1, \ldots, v_n \in \mathbb{R}^n$, *and* $\lambda \in \mathbb{R}$. *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* $A \in M_{n\times n}(\mathbb{R})$, *then* $\det(A^T) = \det(A)$.

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

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

Remark 27.2.3. 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 27.2.4. *Suppose properties 3a and 3c. Then we can prove half of property* 6 *quickly: if* $A \in M_{n\times n}(\mathbb{R})$ *is not invertible, then* $\det A = 0$*.*

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

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

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

Since A is not invertible, T_A is not an isomorphism (theorem 24.3.4). Thus, T_A is not injective (theorem 20.2.4). $T_A = \Gamma_\alpha$, so this shows that α is linearly dependent (remark 11.3.5).

By a homework problem, 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 27.2.5. *Suppose properties 2a and 5. Then we can prove half of property* 6 *quickly: if* $A \in M_{n\times n}(\mathbb{R})$ *is invertible, then* $\det A \neq 0$*.*

Remark 27.2.6. Without property 5, the proof that "if $A \in M_{n\times n}(\mathbb{R})$ is invertible, then $\det A \neq 0$" is more annoying. A good reason to believe that this result is true is that the det is connected with volume. The argument in the previous proof (24.3.4, 20.2.4, 11.3.5) can be enhanced to say that 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, we see that it has non-zero volume exactly when $(v_1, \ldots, v_n)$ is a basis for $\mathbb{R}^n$.

27.3 Eigenvectors

Definition 27.3.1. Suppose V is a vector space over $\mathbb{R}$ 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:

- $v \neq 0$;
- $Tv = \lambda v$ for some $\lambda \in \mathbb{R}$ (often we write Tv instead of $T(v)$).

If $v \in V$ is an eigenvector of T and $Tv = \lambda v$, then λ is called *the eigenvalue corresponding to v*.

A scalar $\lambda \in \mathbb{R}$ 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 27.3.2. Suppose V and W are vector spaces over $\mathbb{R}$, that $S : V \longrightarrow W$ and $T : V \longrightarrow W$ are linear transformations and $\lambda \in \mathbb{R}$. 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 27.3.3. *Suppose V and W are vector spaces over $\mathbb{R}$, that $S : V \longrightarrow W$ and $T : V \longrightarrow W$ are linear transformations and that $\lambda \in \mathbb{R}$. 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

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

Theorem 27.3.4. *Suppose that V is a vector space over $\mathbb{R}$, $T : V \longrightarrow V$ is a linear transformation, $v \in V$, and $\lambda \in \mathbb{R}$. 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 that V is a vector space over $\mathbb{R}$, $T : V \longrightarrow V$ is a linear transformation, $v \in V$, and $\lambda \in \mathbb{R}$. Notice that

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

Thus, the condition $(T - \lambda 1_V)(v) = 0$ is equivalent to $Tv = \lambda v$.

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

Conversely, 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. $\square$

Corollary 27.3.5. *Suppose V is a vector space over $\mathbb{R}$, $T : V \longrightarrow V$ is a linear transformation, and $\lambda \in \mathbb{R}$. λ is an eigenvalue of T if and only if $\ker (T - \lambda 1_V) \neq \{0\}$.*

Definition 27.3.6. Suppose V is a vector space over $\mathbb{R}$, $T : V \longrightarrow V$ is a linear transformation, and $\lambda \in \mathbb{R}$. $E_\lambda := \ker (T - \lambda 1_V)$ is called the *eigenspace of T corresponding to λ*.

Remark 27.3.7. We normally only talk about eigenspaces E_λ when λ is an eigenvalue of T.

28 Questions due on August 30th

0. Get ready for the quiz.

29 Quiz 2

1. Suppose X is a nonempty set.

   In class and on the homework, we proved that the set of real-valued functions from X,

   $$\mathcal{F} = \{f : X \longrightarrow \mathbb{R}\}$$

   is a vector space over $\mathbb{R}$ when equipped with pointwise addition and scalar multiplication.

   (a) Give the definition of addition and scalar multiplication in $\mathcal{F}$.
   Solution: see quiz 1.

   (b) Verify the *seventh* axiom of a vector space for $\mathcal{F}$, that is, that for all $\lambda \in \mathbb{R}$, and for all $f, g \in \mathcal{F}$, $\lambda(f + g) = \lambda f + \lambda g$.
   Solution: see homework.

   (c) Fix $x_0 \in X$. Prove that $T : \mathcal{F} \longrightarrow \mathbb{R}$, $f \longmapsto f(x_0)$ is linear.
   Solution: see homework.

   Now take $X = \mathbb{R}$ so that $\mathcal{F} = \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$.

   (d) Is the subset $\{f \in \mathcal{F} : \ f(-8) \cdot f(18) = 0\}$ a subspace of $\mathcal{F}$?
   Prove your claim.

   Solution: No, it's not. Let $f(x) = x + 8$ and $g(x) = x - 18$. Then $f(-8) \cdot f(18) = 0$ and $g(-8) \cdot g(18) = 0$, but $(f + g)(-8) \cdot (f + g)(18) = -26^2 \neq 0$.

2. Suppose U, V, and W are finite-dimensional vector spaces over $\mathbb{R}$, and that

   $$T : U \longrightarrow V, \ S : V \longrightarrow W$$

   are linear transformations.

   For each of the following two statements, either prove that it is always true, or give an example of U, V, W, S, T demonstrating that it is sometimes false.

   (a) $\text{null}(ST) \leq \text{null}(S)$.

   Solution: Let $U = \mathbb{R}$, $V = \{0\}$, $W = \{0\}$, $S = 0$, and $T = 0$.
   Then $\text{null}(ST) = 1 > 0 = \text{null}(S)$. Thus, the statement is sometimes false.

   (b) $\text{null}(ST) \geq \text{null}(S)$.

   Solution: Let $U = \{0\}$, $V = \mathbb{R}$, $W = \{0\}$, $S = 0$, and $T = 0$.
   Then $\text{null}(ST) = 0 < 1 = \text{null}(S)$. Thus, the statement is sometimes false.

3. Define $T : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ by

$$T\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}\right) = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

Find bases for $\ker T$ and $\operatorname{im} T$.

Justify your claim that they are bases as efficiently as possible while stating any results that you use clearly, e.g. consequences of the replacement theorem.

Solution: Consider the following two tuples.

$$\beta_K = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}\right), \quad \beta_I = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}\right) = (T(e_1),\ T(e_2))$$

The vectors in β_K are easily checked to be in $\ker T$. They are linearly independent.

The vectors in β_I are seen to be in $\operatorname{im} T$. They are linearly independent.

It is a consequence of the replacement theorem that linearly independent tuples are smaller than or equal to bases in size. Thus, we obtain $null(T) \geq 2$ and $rank(T) \geq 2$.

The rank-nullity theorem tells us $rank(T) + null(T) = 4$.

Thus, $rank(T) = null(T) = 2$.

It is a consequence of the replacement theorem that linearly independent tuples of the correct size are bases. Thus, β_K and β_I are bases for $\ker T$ and $\operatorname{im} T$, respectively.

4. Suppose V is a finite-dimensional vector space over $\mathbb{R}$, and that U and W are subspaces of V. Recall that the *sum* of U and W is defined by

$$U + W := \{u + w : \ u \in U, \ w \in W\},$$

and that it is a subspace of V that contains both U and W.

Prove that $\dim(U + W) \leq \dim U + \dim W$ stating any results that you use clearly.

This combines two homework problems. You do not get 8 *points for noticing this!*

You can assume that bases exist, the concept of dimension makes sense, subspaces of finite-dimensional vector spaces are finite-dimensional, any direct consequences of the replacement theorem, and what I told you above. Everything else should be proved.

Solution:

Let $m = \dim U$ and $n = \dim W$, and choose bases for U and W, respectively:

$$(u_1, u_2, \ldots, u_m) \text{ and } (w_1, w_2, \ldots, w_n).$$

We claim $(u_1, u_2, \ldots, u_m, w_1, w_2, \ldots, w_n)$ spans $U + W$.

Provided that we prove the claim, $U + W$ has a spanning tuple of size $m + n$, and so the dimension of $U + W$ is less than or equal to $m + n$ (this is a consequence of the replacement theorem), i.e.

$$\dim(U + W) \leq \dim U + \dim W.$$

To prove the claim let $v \in U + W$.

By definition of $U + W$, we can find $u \in U$ and $w \in W$ such that $v = u + w$.

Since $(u_1, u_2, \ldots, u_m)$ is a basis for U, $u = \lambda_1 u_1 + \ldots + \lambda_m u_m$ for some $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$. Similarly, $w = \mu_1 w_1 + \ldots + \mu_n w_n$ for some $\mu_1, \ldots, \mu_n \in \mathbb{R}$. Thus,

$$v = u + w = (\lambda_1 u_1 + + \ldots \lambda_m u_m) + (\mu_1 w_1 + \ldots + \mu_n w_n).$$

This shows v is in the span of $(u_1, u_2, \ldots, u_m, w_1, w_2, \ldots, w_n)$, so $(u_1, \ldots, u_m, w_1, \ldots, w_n)$ spans $U + W$.

5. (a) Let $A = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix}$.

Consider the linear transformation $T_A : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$.
Let $\beta_{\mathcal{D}} = (e_2, e_3, e_1)$ and $\beta_{\mathcal{C}} = (e_1, e_3, e_2)$.
Here $\mathcal{D}$ stands for domain and $\mathcal{C}$ stands for codomain.
Write down $[T_A]_{\beta_{\mathcal{D}}}^{\beta_{\mathcal{C}}}$.

Solution: $\begin{pmatrix} 1 & 2 & 0 \\ 7 & 8 & 6 \\ 4 & 5 & 3 \end{pmatrix}$.

(b) Suppose $T : V \longrightarrow \mathcal{P}_2(\mathbb{R})$ is linear, β_V is a basis for V, and

$$\beta_{\mathcal{P}_2(\mathbb{R})} = (1, 1 + x, x + x^2).$$

Suppose, also, that $v \in V$, $[v]_{\beta_V} = (1, -1)$, and

$$[T]_{\beta_V}^{\beta_{\mathcal{P}_2(\mathbb{R})}} = \begin{pmatrix} 3 & 2 \\ 8 & 6 \\ 4 & 1 \end{pmatrix}.$$

What's $T(v)$? **Solution**: $1 \cdot 1 + 2 \cdot (1 + x) + 3 \cdot (x + x^2) = 3 + 5x + 3x^2$.

(c) Is there a linear transformation $T : \mathcal{P}_3(\mathbb{R}) \longrightarrow \mathbb{R}^2$ with the following properties?

- $T(x^2 + x + 1) = (3, 0)$
- $T(x + 1) = (1, 0)$
- $T(x^2 + 1) = (1, 0)$
- $T(x^2 + x) = (1, 0)$

Solution: No. If there was, we'd have

$$\begin{aligned}(6, 0) &= 2 \cdot T(x^2 + x + 1) \\ &= T((x^2 + x) + (x^2 + 1) + (x + 1)) = (1, 0) + (1, 0) + (1, 0) = (3, 0).\end{aligned}$$

(d) Is there a linear transformation $T : \mathcal{P}_3(\mathbb{R}) \longrightarrow \mathbb{R}^2$ with the following properties?

- $T(x^3 + x^2 + x + 1) = (1, -1)$
- $T(x + 3) = (0, 1)$
- $T(x - 3) = (1, 0)$
- $T(6) = (-1, 1)$

Solution: Yes. $(6, 2x, x^2, x^3 + x^2 + x + 1)$ is a basis for $\mathcal{P}_3(\mathbb{R})$ and we can define T by

$$T(6) = (-1, 1),\ T(2x) = (1, 1),\ T(x^2) = (0, 0),\ T(x^3 + x^2 + x + 1) = (1, -1).$$

30 Comments on quiz 2

1. This was a clone of a quiz 1 question.
2. This problem was more difficult than intended.

   What did you know going in?

   - $\ker(ST) \supseteq \ker(T)$, so $null(ST) \geq null(T)$.
   - $\mathrm{im}(ST) \subseteq \mathrm{im}(S)$, so $rank(ST) \leq rank(S)$.
   - $\mathrm{im}(T) \subseteq V$ and $\mathrm{im}(ST) \subseteq W$, so these subspaces cannot be subsets of one another unless V and W are related. However, $rank(ST)$ and $rank(T)$ can always be compared:
     - $null(ST) \geq null(T)$ and the rank-nullity theorem allow us to see that $rank(ST) \leq rank(T)$. When applying the rank-nullity theorem, it is important that the domain of ST and T are the same. In both cases, their domain is U.
     - Alternatively, you can note that we have a surjective linear transformation

       $$\overline{S} : \mathrm{im}\, T \longrightarrow \mathrm{im}(ST),\ v \longmapsto S(v).$$

       Thus, $\dim(\mathrm{im}\, T) \geq \dim(\mathrm{im}(ST))$.

   This prepares one to become pessimistic about comparing $null(ST)$ and $null(S)\ldots$

   - $\ker(ST) \subseteq U$ and $\ker(S) \subseteq V$, so these subspaces cannot be subsets of each other unless U and V are related.
   - Although we know that $rank(ST) \leq rank(S)$, applying the rank-nullity gives

     $$\dim(U) - null(ST) \leq dim(V) - null(S).$$

     $\dim U$ and $\dim V$ could be very different, so this does not help us.
   - Although we have a linear transformation

     $$\overline{T} : \ker(ST) \longrightarrow \ker(S),\ u \longmapsto T(u),$$

     it is impossible to say anything about its injectivity or surjectivity.

   All of this should lead one to doubt that either inequality holds in general and to look for counter-examples. What if $S = 0$? Then $null(ST) = \dim U$ and $null(S) = \dim V$. You can surely arrange for either $\dim U > \dim V$ or $\dim U < \dim V\ldots$
3. This was exactly the same in spirit as example 20.2.5 and 8/27 question 3.
4. This was the concatenation of two homework problems: 8/22 3c and 8/27 5b.
5. (a) A simple calculation.
   (b) A simple calculation.
   (c) You spot a dependency relation and show it creates a problem with defining a T.
   (d) You spot a dependency relation but show it does not prevent defining a T.

31 Lecture on August 30th: Diagonalizable Transformations

Definition 31.1. Suppose V is a finite-dimensional vector space over $\mathbb{R}$ and $T : V \to V$ 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.

I want to state the conditions for a linear transformation $T : V \longrightarrow V$ from a finite-dimensional vector space to itself to be diagonalizable as soon as possible. We'll worry about proving it after.

Definition 31.2. Suppose V is a finite-dimensional vector space over $\mathbb{R}$ 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 31.3. Suppose V is a finite-dimensional vector space over $\mathbb{R}$, $T : V \longrightarrow V$ 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 does not depend on the choice of basis.

Definition 31.4. Suppose V is a finite-dimensional vector space over $\mathbb{R}$ and $T : V \to V$ is a linear transformation. The *characteristic polynomial of* T is defined by the following equation

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

Theorem 31.5. *Suppose V is a finite-dimensional vector space over $\mathbb{R}$, $T : V \to V$ is a linear transformation, and that $\lambda \in \mathbb{R}$. 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 $\mathbb{R}$, that $T : V \longrightarrow V$ is a linear transformation, and that $\lambda \in \mathbb{R}$. Suppose β_V is a basis for V. Then

$$\begin{aligned}
&\lambda \text{ is an eigenvalue of } T \\
\iff\ & \ker(T - \lambda 1_V) \neq \{0\} \text{ (corollary 27.3.5)} \\
\iff\ & T - \lambda 1_V \text{ is not an isomorphism (since the domain and codomain have the same dimension)} \\
\iff\ & [T - \lambda 1_V]_{\beta_V}^{\beta_V} \text{ is not invertible (section 24.3)} \\
\iff\ & \det([T - \lambda 1_V]_{\beta_V}^{\beta_V}) = 0 \text{ (section 27.2)} \\
\iff\ & \det(T - \lambda 1_V) = 0 \text{ (definition 31.2)} \\
\iff\ & c_T(\lambda) = 0 \text{ (definition 31.4).}
\end{aligned}$$

$c_T(\lambda) = 0$ is exactly what we mean when we say λ is a root of $c_T(x)$, so this finishes the proof. $\square$

Example 31.6. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, so that $T_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$. We have

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

So the only eigenvalue of T_A is 1. Moreover, a direct calculation gives

$$E_1 = \ker(T_A - 1_{\mathbb{R}^2}) = \left\{ \begin{pmatrix} x_1 \\ 0 \end{pmatrix} : \ x_1 \in \mathbb{R} \right\}.$$

We'll show that T_A is *not* diagonalizable.

Suppose for contradiction that T_A is diagonalizable. Then there is a basis (v_1, v_2) of $\mathbb{R}^2$ consisting of eigenvectors of T_A. Since 1 is the *only* eigenvalue of T_A, we have $v_1, v_2 \in E_1 = \ker(T_A - 1_{\mathbb{R}^2})$. This gives $E_1 = \mathbb{R}^2$ which contradicts the calculation above.

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 is an attempt to motivate the next definition.

Definition 31.7. Suppose V is a finite-dimensional vector space over $\mathbb{R}$, $T : V \longrightarrow V$ is a linear transformation, and λ is an eigenvalue of T.

The *geometric multiplicity of* λ is

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

The *algebraic multiplicity of* λ is

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

Theorem 31.8. *Suppose V is a finite-dimensional vector space over $\mathbb{R}$, $T : V \longrightarrow V$ 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 $1 \leq G \leq A$.*

Definition 31.9. A polynomial $p(x) \in \mathcal{P}(\mathbb{R})$ *splits over* $\mathbb{R}$ **iff** there are scalars $c, \sigma_1, \sigma_2, \ldots, \sigma_n \in \mathbb{R}$ such that

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

Theorem 31.10. *Suppose V is a finite-dimensional vector space over $\mathbb{R}$ and $T : V \longrightarrow V$ is a linear transformation. T is diagonalizable if and only if*

- *the characteristic polynomial of T splits over $\mathbb{R}$; and*
- *for each eigenvalue λ of T, its geometric multiplicity is equal to its algebraic multiplicity.*

Example 31.11. Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, so that $T_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$.

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

Example 31.12. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, so that $T_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$.

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

32 Questions due on September 5th

You should know the following definition. We'll need it next week, for sure!

Definition. Suppose V is a vector space over $\mathbb{R}$, and that W_1 and W_2 are subspaces of V. We say that V is the *direct sum* of W_1 and W_2 **iff** $V = W_1 + W_2$ and that $W_1 \cap W_2 = \{0\}$. In this case, we write $V = W_1 \oplus W_2$.

1. (a) Consider the following subspaces of $\mathbb{R}^2$.

   $$W_1 = \{(x, 0) : \ x \in \mathbb{R}\}, \ W_2 = \{(t, t) : t \in \mathbb{R}\}.$$

   Prove that $W_1 \oplus W_2 = \mathbb{R}^2$.

   (b) Recall that the set of functions $V = \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$ form a vector space over $\mathbb{R}$ under pointwise addition and scalar multiplication. The subsets of *odd functions*

   $$W_1 = \{f : \mathbb{R} \longrightarrow \mathbb{R} : \ \text{for all } t \in \mathbb{R}, \ f(-t) = -f(t)\}$$

   and *even functions*

   $$W_2 = \{f : \mathbb{R} \longrightarrow \mathbb{R} : \ \text{for all } t \in \mathbb{R}, \ f(-t) = f(t)\}$$

   are subspaces of V. Prove that $V = W_1 \bigoplus W_2$.

2. (a) Suppose V is a vector space over $\mathbb{R}$, and that W_1 and W_2 are subspaces of V. Prove that V is the direct sum of W_1 and W_2 if and only if each $v \in V$ can be written uniquely as $w_1 + w_2$ for $w_1 \in W_1$ and $w_2 \in W_2$.

   (b) Suppose W_1 and W_2 are vector spaces over $\mathbb{R}$. Let

   $$V := W_1 \times W_2 = \{(w_1, w_2) : w_1 \in W_1, \ w_2 \in W_2\}.$$

   V is a vector space over $\mathbb{R}$ with the operations

   $$(w_1, w_2) + (w_1', w_2') := (w_1 + w_1', w_2 + w_2') \ \text{ and } \ \lambda(w_1, w_2) := (\lambda w_1, \lambda w_2).$$

   (If you want more practice, you could try and prove this.)
   Write 0_{W_1} and 0_{W_2} for the zeros of W_1 and W_2, respectively. The zero of V is given by $0_V = (0_{W_1}, 0_{W_2})$. Let $\overline{W}_1 = \{(w_1, 0_{W_2}) : w_1 \in W_1\}$ and $\overline{W}_2 = \{(0_{W_1}, w_2) : w_2 \in W_2\}$. $\overline{W}_1$ and $\overline{W}_2$ are subspaces of V. Prove that $V = \overline{W}_1 \oplus \overline{W}_2$.

3. Suppose V and W are finite-dimensional vector spaces over $\mathbb{R}$, and that $T : V \to W$ is a linear transformation. Show that there is a subspace $U \subseteq V$ with the following two properties:

   (a) $V = \ker T \oplus U$;

   (b) The linear transformation $S : U \to \operatorname{im} T$ defined by $S(u) := T(u)$ is an isomorphism.

   Read the proof of rank-nullity and define U as the span of some tuple; use theorem 20.3.4.

4. Suppose V is a vector space over $\mathbb{R}$ and that $T : V \longrightarrow V$ is a linear transformation with the property that $T^2 = T$. Prove that $V = \ker T \oplus \operatorname{im} T$.

   T^2 means TT, i.e. T composed with itself.

5. Suppose V is a vector space over $\mathbb{R}$ and that $T : V \longrightarrow V$ is a linear transformation with two distinct eigenvectors $\lambda, \mu \in \mathbb{R}$.

   Recall the eigenspaces $E_\lambda = \ker(T - \lambda 1_V)$ and $E_\mu = \ker(T - \mu 1_V)$.

   (a) Prove that $E_\lambda \cap E_\mu = \{0\}$.

   (b) Suppose, in addition, that λ and μ are the *only* eigenvalues of T, and that T is diagonalizable. Prove that $V = E_\lambda \oplus E_\mu$.

6. (a) Consider $\mathcal{P}(\mathbb{R})$, the vector space of real polynomials.
   Define $T : \mathcal{P}(\mathbb{R}) \longrightarrow \mathcal{P}(\mathbb{R})$ by $T(p(x)) = x \cdot p'(x)$.
   Tell me infinitely many eigenvalues of T, and give a corresponding eigenvector for each.

   (b) Let $n \in \mathbb{N}$.
   Consider $\mathcal{P}_n(\mathbb{R})$, the vector space of real polynomials of degree less than or equal to n.
   Define $T : \mathcal{P}_n(\mathbb{R}) \longrightarrow \mathcal{P}_n(\mathbb{R})$ by $T(p(x)) = x \cdot p'(x)$.
   Prove that T is diagonalizable.

   (c) [Optional but instructive]
   Recall that the set of real-valued functions on $\mathbb{R}$

   $$\mathcal{F} = \{f : \mathbb{R} \longrightarrow \mathbb{R}\}$$

   is a vector space over $\mathbb{R}$.
   Given $a \in \mathbb{R}$, define $f_a : \mathbb{R} \longrightarrow \mathbb{R}$ by $f_a(x) = e^{ax}$.
   Let $V = \operatorname{span}\{f_a : a \in \mathbb{R}\}$. You'll need section 11.2 to make sense of this.
   Define $T : V \longrightarrow V$ by $f \longmapsto f'$ (this is well-defined).
   Tell me infinitely many eigenvalues of T, and give a corresponding eigenvector for each.

7. Read the rest of section 31.

   For each of the linear transformations $T : V \longrightarrow V$ on the next page,

   - calculate $c_T(x)$;
   - list the (real) eigenvalues $\lambda_1, \ldots, \lambda_n$ of T;
   - give their algebraic multiplicity;
   - write down a basis for their eigenspaces;
   - give their geometric multiplicity;
   - say whether T is diagonalizable (over $\mathbb{R}$);
   - if T is diagonalizable (over $\mathbb{R}$), give a basis of eigenvectors;

(a) $T : \mathbb{R}^6 \longrightarrow \mathbb{R}^6$ defined by

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} \longmapsto \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix}.$$

(b) $T : \mathbb{R}^6 \longrightarrow \mathbb{R}^6$ defined by

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} \longmapsto \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix}.$$

(c) $T : \mathbb{R}^6 \longrightarrow \mathbb{R}^6$ defined by

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} \longmapsto \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix}.$$

(d) $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ defined by $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

(e) $T : \mathbb{R}^8 \longrightarrow \mathbb{R}^8$ defined by

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{pmatrix} \longmapsto \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{pmatrix}.$$

(f) $T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathcal{P}_2(\mathbb{R})$ defined by $p(x) \longmapsto p'(x)$.

(g) $T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathcal{P}_2(\mathbb{R})$ defined by $p(x) \longmapsto p'(x) + p(x)$.

(h) $T : \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathcal{P}_2(\mathbb{R})$ defined by $p(x) \longmapsto p'(x) + p(x) + p(0)$.

33 Sketch solutions for the previous questions

The following are definitely sketch solutions. Your solutions should contain more details.

1. (a) Suppose $(x_1, x_2) \in W_1 \cap W_2$. Since $(x_1, x_2) \in W_1$, $x_2 = 0$. Since $(x_1, x_2) \in W_2$, $x_1 = x_2$. Thus, $(x_1, x_2) = (0, 0) = 0 \in \{0\}$.
   To see that $\mathbb{R}^2 = W_1 + W_2$, we note that for any $(x_1, x_2) \in \mathbb{R}^2$,
   $$(x_1, x_2) = (x_1 - x_2, 0) + (x_2, x_2) \in W_1 + W_2.$$
   (b) Suppose $f \in W_1 \cap W_2$ and $t \in \mathbb{R}$. Since $f \in W_1$, we have $f(t) = -f(-t)$. Since $f \in W_2$, we have $f(-t) = f(t)$. These two equations show $f(t) = -f(t)$, so $f(t) = 0$. Since $t \in \mathbb{R}$ was arbitrary, this shows f is the zero function, i.e. $f \in \{0\}$.
   To see that $V = W_1 + W_2$, we note that for any $f \in V$, and $t \in \mathbb{R}$,
   $$f(t) = \left(\frac{f(t) - f(-t)}{2}\right) + \left(\frac{f(t) + f(-t)}{2}\right),$$
   that $\left(t \longmapsto \frac{f(t)-f(-t)}{2}\right) \in W_1$, and $\left(t \longmapsto \frac{f(t)+f(-t)}{2}\right) \in W_2$.

2. (a) Suppose $V = W_1 \oplus W_2$ and $v \in V$. We can definitely write $v = w_1 + w_2$ for some $w_1 \in W_1$ and $w_2 \in W_2$. That much is definitional. Suppose we can also express v as $w_1' + w_2'$ for $w_1' \in W_1$ and $w_2' \in W_2$. We find that $w_1 - w_1' = w_2' - w_2 \in W_1 \cap W_2 = \{0\}$, so $w_1 = w_1'$ and $w_2 = w_2'$.
   Conversely, suppose that each $v \in V$ can be written uniquely as $w_1 + w_2$ for $w_1 \in W_1$ and $w_2 \in W_2$. The existence assumption gives $V = W_1 + W_2$. Suppose $v \in W_1 \cap W_2$. Then $v = v + 0 = 0 + v$ and the uniqueness assumption forces $v = 0$. Thus, $W_1 \cap W_2 = \{0\}$.
   (b) Easy.

3. Suppose V and W are finite-dimensional vector spaces over $\mathbb{R}$, and $T : V \longrightarrow W$ is a linear transformation. Choose a basis $(v_1, \ldots, v_n)$ for $\ker T$, and extend it to a basis
   $$(v_1, \ldots, v_n, v_{n+1}, \ldots, v_m)$$
   of V. Let $U = \mathrm{span}(v_{n+1}, \ldots, v_m)$ and define
   $$S : U \longrightarrow \operatorname{im} T, \;\; u \longmapsto T(u).$$
   First, we show that S is an isomorphism... Because $(v_1, \ldots, v_n, v_{n+1}, \ldots, v_m)$ is linearly independent, theorem 11.3.10 tells us that $(v_{n+1}, \ldots, v_m)$ is linearly independent. By definition of U, $(v_{n+1}, \ldots, v_m)$ spans U. So $(v_{n+1}, \ldots, v_m)$ is a basis for U. We know from the proof of the rank-nullity theorem that $(S(v_{n+1}), \ldots, S(v_m)) = (T(v_{n+1}), \ldots, T(v_m))$ is a basis for $\mathrm{im}(T)$. Theorem 20.3.4 says that S is an isomorphism.
   Now we show that $V = \ker T \oplus U$. By a previous homework, we have
   $$\begin{aligned} \ker T + U &= \mathrm{span}(v_1, \ldots, v_n) + \mathrm{span}(v_{n+1}, \ldots, v_m) \\ &= \mathrm{span}(v_1, \ldots, v_n, v_{n+1}, \ldots, v_m) = V. \end{aligned}$$
   Now let $u \in \ker T \cap U$. Then $S(u) = T(u) = 0$. Since S is injective, this gives $u = 0$. Thus, $\ker T \cap U = \{0\}$.

4. Suppose V is a vector space over $\mathbb{R}$ and that $T : V \longrightarrow V$ is a linear transformation with the property that $T^2 = T$.

   To show $\ker T \cap \operatorname{im} T = \{0\}$, let $v \in \ker T \cap \operatorname{im} T$.

   Since $v \in \operatorname{im} T$, we can find a $v' \in V$ such that $v = T(v')$. Since $v \in \ker T$, $T(v) = 0$. Thus,

   $$v = T(v') = T^2(v') = T(T(v')) = T(v) = 0$$

   and we conclude that $\ker T \cap \operatorname{im} T = \{0\}$.

   To show $V = \ker T + \operatorname{im} T$, let $v \in V$. We have $v - T(v) \in \ker T$ because

   $$T(v - T(v)) = T(v) - T(T(v)) = T(v) - T^2(v) = T(v) - T(v) = 0.$$

   We also have $T(v) \in \operatorname{im} T$.

   Thus, $v = (v - T(v)) + T(v) \in \ker T + \operatorname{im} T$ and so $V = \ker T + \operatorname{im} T$.

   We conclude that $V = \ker T \oplus \operatorname{im} T$.

5. Suppose V is a vector space over $\mathbb{R}$ and that $T : V \longrightarrow V$ is a linear transformation with two distinct eigenvectors $\lambda, \mu \in \mathbb{R}$.

   (a) Let $v \in E_\lambda \cap E_\mu$. Then $\lambda v = Tv = \mu v$. So $(\lambda - \mu)v = 0$.
   Since $\lambda - \mu \neq 0$, this gives $v = 0$. So $E_\lambda \cap E_\mu = \{0\}$.

   (b) Suppose that λ and μ are the *only* eigenvalues of T, and that T is diagonalizable.
   Let $(v_1, \ldots, v_m, v_{m+1}, \ldots, v_n)$ be a basis of V consisting of eigenvectors of T where λ is the eigenvalue corresponding to $v_1, \ldots, v_m$ and μ is the eigenvalue corresponding to $v_{m+1}, \ldots, v_n$. Given $v \in V$, we can find $\sigma_1, \ldots, \sigma_n \in \mathbb{R}$ such that $v = \sigma_1 v_1 + \ldots + \sigma_n v_n$; then

   $$v = (\sigma_1 v_1 + \ldots + \sigma_m v_m) + (\sigma_{m+1} v_{m+1} + \ldots + \sigma_n v_n) \in E_\lambda + E_\mu.$$

   Thus, $V = E_\lambda + E_\mu$. We have shown in part (a) that $E_\lambda \cap E_\mu = \{0\}$.

6. (a) For each $n \in \mathbb{N}$, x^n is an eigenvector of T with corresponding eigenvalue n.

   (b) $(1, x, x^2, \ldots, x^n)$ is a basis consisting of eigenvectors.

   (c) For each $a \in \mathbb{R}$, f_a is an eigenvector with eigenvalue a.

7. (a)
   - $(x-2)^4(x-3)^2$
   - $2, 3$
   - $4, 2$
   - (e_1, e_2, e_3, e_4), (e_5, e_6)
   - $4, 2$
   - Yes
   - $(e_1, e_2, e_3, e_4, e_5, e_6)$

   (b)
   - $(x-2)^4(x-3)^2$
   - $2, 3$
   - $4, 2$
   - (e_1, e_2), (e_5)
   - $2, 1$
   - No
   - N/A

   (c)
   - $(x-2)^4(x-3)^2$
   - $2, 3$
   - $4, 2$
   - (e_1, e_3), (e_5)
   - $2, 1$
   - No
   - N/A

   (d)
   - $(x-1)^2+1$
   - $\emptyset$
   - N/A
   - N/A
   - N/A
   - No
   - N/A

(e)
- $((x-1)^2+1)(x-2)^4(x-3)^2$
- $2, 3$
- $4, 2$
- (e_3, e_5), (e_7)
- $2, 1$
- No
- N/A

(f)
- $-x^3$
- 0
- 3
- (1)
- 1
- No
- N/A

(g)
- $-(x-1)^3$
- 1
- 3
- (1)
- 1
- No
- N/A

(h)
- $-(x-1)^2(x-2)$
- $1, 2$
- $2, 1$
- $(x-1)$, (1)
- $1, 1$
- No
- N/A

34 Proving theorem 31.10 (I'll lecture a strict subset of this)

The results of this section are ordered from the easiest to the most difficult rather than what I think is the most natural order. The first theorem highlights a connection between a linear transformation being diagonalizable and its characteristic polynomial splitting over $\mathbb{R}$.

Theorem 34.1. *Suppose V is a finite-dimensional vector space over $\mathbb{R}$, $T : V \longrightarrow V$ is a linear transformation, and T is diagonalizable. Then $c_T(x)$ splits over $\mathbb{R}$.*

Proof. Suppose V is a vector space over $\mathbb{R}$, $T : V \longrightarrow V$ is a linear transformation, and T is diagonalizable. Let $\beta_V = (v_1, \ldots, v_n)$ be a basis of eigenvectors of T. Let $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ be the eigenvalues corresponding to $v_1, \ldots, v_n$, respectively. Then

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

□

The next theorem relates the concept eigenvectors and linear independence.

Theorem 34.2. *Suppose V is a vector space over $\mathbb{R}$, $T : V \longrightarrow V$ is a linear transformation, $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ 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 that $n \in \mathbb{N}$, and that the result is true for this n. Now suppose V is a vector space over $\mathbb{R}$, $T : V \longrightarrow V$ is a linear transformation, $\lambda_1, \ldots, \lambda_n, \lambda_{n+1} \in \mathbb{R}$ 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. So suppose $\mu_1, \ldots, \mu_n, \mu_{n+1} \in \mathbb{R}$ 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, $(v_1, v_2, \ldots, v_n)$ 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 first equation, we find that $\mu_{n+1} v_{n+1} = 0$. Since v_{n+1} is an eigenvector, it is non-zero, so $\mu_{n+1} = 0$. We have shown that 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. □

Here is the simplest condition which ensures a linear transformation is diagonalizable.

Theorem 34.3. *Suppose V is a finite-dimensional vector space over $\mathbb{R}$ 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 $\mathbb{R}$ and that $T : V \longrightarrow V$ is a linear transformation. Let $n = \dim V$ and suppose that $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ are the distinct eigenvalues of 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, the previous theorem tells us that $(v_1, \ldots, v_n)$ is linearly independent. Thus, by theorem 18.4 (part 4), $(v_1, \ldots, v_n)$ is a basis of V consisting of eigenvectors of T. $\square$

The next lemma and theorem build on theorem 34.2.

Lemma 34.4. *Suppose that V is a vector space over $\mathbb{R}$, $T : V \longrightarrow V$ is a linear transformation, $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ 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 $\mathbb{R}$, $T : V \to V$ is a linear transformation, $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ 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 of T corresponding to distinct 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. $\square$

Theorem 34.5. *Suppose V is a vector space over $\mathbb{R}$, $T : V \longrightarrow V$ is a linear transformation, and that $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ 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 $\mu_{1,1}, \ldots, \mu_{1,r_1}, \mu_{2,1}, \ldots, \mu_{2,r_2}, \ldots\ldots, \mu_{n,1}, \ldots, \mu_{n,r_n} \in \mathbb{R}$ 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 tuple for each $j \in \{1, \ldots, n\}$, this tells us all the μ's are zero. $\square$

We now turn to the proofs of theorem 31.8 and theorem 31.10.

Theorem. *Suppose V is a finite-dimensional vector space over $\mathbb{R}$, $T : V \longrightarrow V$ 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. Suppose V, T, λ, G, and A are as in the theorem statement. 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. Then $[T]_{\beta_V}^{\beta_V}$ is equal to

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

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

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

where $p(x) = (-1)^G \det(C - xI_{(n-G)})$. Here, we have used a fact about calculating determinants in block form; the fact that the bottom left block is full of zeros is essential. The last equation shows $(x - \lambda)^G$ is a factor of $c_T(x)$ and so, by the definition of algebraic multiplicity, we have $A \geq G$.

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

The next theorem is over the page.

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

- *the characteristic polynomial of T splits over $\mathbb{R}$; 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 $\mathbb{R}$ and $T : V \longrightarrow V$ is linear.

Before proving the theorem, we make some observations. Let $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ be the distinct eigenvalues of T. For $j \in \{1, \ldots, n\}$, let G_j be the geometric multiplicity of λ_j, and let A_j be the algebraic multiplicity of λ_j. The previous theorem tells us that for each $j \in \{1, \ldots, n\}$, $1 \leq G_j \leq A_j$.

By theorem 31.5 and definition 31.7, we have

$$c_T(x) = (x - \lambda_1)^{A_1}(x - \lambda_2)^{A_2} \cdots (x - \lambda_n)^{A_n} \cdot p(x)$$

for some polynomial $p(x) \in \mathcal{P}(\mathbb{R})$ which has no real roots. So

$$\sum_{j=1}^{n} A_j \leq \deg c_T(x) = \dim V.$$

We have equality in this equation if and only if $p(x)$ is a degree zero polynomial, that is, if and only if $c_T(x)$ splits over $\mathbb{R}$.

In summary,

$$\sum_{j=1}^{n} G_j \leq \sum_{j=1}^{n} A_j \leq \dim V.$$

The second inequality is an equality if and only if $c_T(x)$ splits over $\mathbb{R}$. The first inequality is an equality if and only for each eigenvalue λ of T, its geometric multiplicity is equal to its algebraic multiplicity. We are now ready to prove the theorem.

First, suppose that T is diagonalizable. Let

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

be a basis of T consisting of eigenvectors such that λ_j is the eigenvalue corresponding to $v_{j,1}, \ldots, v_{j,r_j}$. We know (theorem 34.1) 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 $A_j = r_j$ for all $j \in \{1, \ldots, n\}$. Let $j \in \{1, \ldots, n\}$. We know from the previous theorem that $G_j \leq A_j$. Since $(v_{j,1}, \ldots, v_{j,r_j})$ is linearly independent and consists of vectors in E_{λ_j}, we also know

$$G_j = \dim E_{\lambda_j} \geq r_j = A_j.$$

Thus, $G_j = A_j$ for all $j \in \{1, \ldots, n\}$.

Conversely, suppose that $c_T(x)$ splits over $\mathbb{R}$ and that $G_j = A_j$ for all $j \in \{1, \ldots, n\}$. By the observations made before starting the proof, we obtain

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

For each $j \in \{1, \ldots, n\}$, choose a basis of E_{λ_j}: $(v_{j,1}, \ldots, v_{j,G_j})$. By theorem 34.5,

$$(v_{1,1}, \ldots, v_{1,G_1}, v_{2,1}, \ldots, v_{2,G_2}, \ldots\ldots, v_{n,1}, \ldots, v_{n,G_n})$$

is linearly independent. Because it is a $(\dim V)$-tuple, theorem 18.4 tells us that it is a basis for V. Since it consists of eigenvectors of T, this shows that T is diagonalizable. $\square$

35 Lecture on September 5th

35.1 Vector spaces over a field

Definition 35.1.1. 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 35.1.2.

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 35.1.3. 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 axioms. The axioms are identical to those of definition 5.3 except all occurrences of $\mathbb{R}$ are replaced with $\mathbb{F}$.

Theorem 35.1.4. *Every major definition that we have given and every major theorem that we have proved works for vector spaces over any field.*

Proof. Read the definitions and the proofs of the theorems. We never used anything special about $\mathbb{R}$. Everything we used about $\mathbb{R}$ is true for any field. $\square$

In the next section, we will use vector spaces over $\mathbb{C}$ as well as vector spaces over $\mathbb{R}$.

Example 35.1.5. $\mathbb{C}$ is a vector space over $\mathbb{R}$ and a vector space over $\mathbb{C}$.

- As a vector space over $\mathbb{R}$, $(1, i)$ is a basis for $\mathbb{C}$, and so $\dim_{\mathbb{R}} \mathbb{C} = 2$.
- As a vector space over $\mathbb{C}$, (1) is a basis for $\mathbb{C}$, and so $\dim_{\mathbb{C}} \mathbb{C} = 1$.

Example 35.1.6. Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, and consider $T : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$, $x \longmapsto Ax$.

$c_T(x) = \det \begin{pmatrix} -x & -1 \\ 1 & -x \end{pmatrix} = x^2 + 1 = (x - i)(x + i)$ splits over $\mathbb{C}$, and T is diagonalizable.

35.2 Inner products

Definition 35.2.1. 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 35.2.2. A *real inner product space* is a vector space V over $\mathbb{R}$, together with an inner product $\langle\ ,\ \rangle$ on V.

Example 35.2.3. Let $n \in \mathbb{N}$. $\mathbb{R}^n$ has an inner product, *the standard inner product* defined by

$$\Big\langle (x_1, x_2, \ldots, x_n),\ (y_1, y_2, \ldots, y_n) \Big\rangle = \sum_{i=1}^{n} x_i y_i.$$

Example 35.2.4. Let $\mathcal{C}$ be

$$\Big\{ f : \quad f \text{ is a function } \{x \in \mathbb{R} :\ 0 \le x \le 1\} \longrightarrow \mathbb{R} \ \text{ and } f \text{ is continuous} \Big\}.$$

In 131A, you prove that $\mathcal{C}$ 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 35.2.5. 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. $\forall \lambda \in \mathbb{C}$, $\forall u_1 \in V$, $\forall u_2 \in V$, $\forall v \in V$,
$$\langle \lambda u_1 + u_2, v\rangle = \lambda\langle u_1, v\rangle + \langle u_2, v\rangle.$$
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 35.2.6. A *complex inner product space* is a vector space V over $\mathbb{C}$, together with an inner product $\langle\ ,\ \rangle$ on V.

Example 35.2.7. Let $n \in \mathbb{N}$. $\mathbb{C}^n$ has an inner product, *the standard inner product* defined by

$$\Big\langle (x_1, x_2, \ldots, x_n),\ (y_1, y_2, \ldots, y_n) \Big\rangle = \sum_{i=1}^{n} x_i \overline{y_i}.$$

Theorem 35.2.8. *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 35.2.9. Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$. For $v \in V$, we define the *norm* of v by $\|v\| = \sqrt{\langle v, v\rangle}$. You should think of this as the length of v.

Theorem 35.2.10. *Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$. Let λ be a scalar and $u, v \in V$. 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 $\mathbb{C}$, λ be a scalar, and $u, v \in V$ 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 $\|u + v\| \leq \|u\| + \|v\|$. $\square$

Remark 35.2.11. 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.

Cauchy-Schwartz also allows us to define the angle between two non-zero vectors u and v as

$$\arccos\left(\frac{|\langle u, v\rangle|}{\|u\|\|v\|}\right).$$

Definition 35.2.12. 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 35.2.13. Let V be an inner product space over $\mathbb{R}$ or $\mathbb{C}$ and let $v_1, \ldots, v_n \in V$. The tuple $(v_1, \ldots, v_n)$ is said to be *orthogonal* **iff** for every $i, j \in \{1, \ldots, n\}$ with $i \neq j$, v_i and v_j are orthogonal. We say that $(v_1, \ldots, v_n)$ is *othonormal* **iff** it is orthogonal, and each v_i is a unit vector.

Remark 35.2.14. Suppose V is an inner product space over $\mathbb{R}$ or $\mathbb{C}$. Notice that $(v_1, \ldots, v_n)$ is *orthonormal* if and only if for all $i, j \in \{1, \ldots, n\}$,

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

Example 35.2.15. Let $n \in \mathbb{N}$, and let $\mathbb{R}^n$ have the standard inner product. The standard basis is an orthonormal tuple.

Theorem 35.2.16 (The Fourier basis)**.** *Let $n \in \mathbb{N}$, and let $\mathbb{C}^n$ have the standard inner product.*

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

$$f_k = \frac{(\zeta^k,\ \zeta^{2k},\ \ldots,\ \zeta^{(n-1)k},\ \zeta^{nk})}{\sqrt{n}} \in \mathbb{C}^n.$$

Then $(f_1, f_2, \ldots, f_n)$ is an orthonormal tuple.

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

$$\langle f_k, f_l\rangle = \frac{1}{n}\sum_{j=1}^{n} \zeta^{jk}\overline{\zeta^{jl}} = \frac{1}{n}\sum_{j=1}^{n}(\zeta^{k-l})^j.$$

If $k - l$ is divisible by n, $\zeta^{k-l} = 1$, and so $\langle f_k, f_l\rangle = 1$.

If $k - l = 1$, then $\langle f_k, f_l\rangle$ is $\frac{1}{n}$ times the sum of the roots of $z^n - 1 = 0$, which is 0.

It's an exercise to show that the other cases give 0 too.

Your proof will break into cases depending on what highest common factor of $k-l$ and n is. $\square$

36 Questions due on September 6th

Definition. Let V and W be inner product spaces over $\mathbb{R}$ and suppose that $T : V \longrightarrow W$ is a linear transformation. T is said to be an *isometry* **iff** for all $v_1, v_2 \in V$, $\langle Tv_1, Tv_2 \rangle = \langle v_1, v_2 \rangle$.

Note that on the left side of the last equation $\langle\ ,\ \rangle$ is the inner product on W; on the right side of the equation, it is the inner product on V.

1. Let V and W be inner product spaces over $\mathbb{R}$. Suppose $T : V \longrightarrow W$ is a linear transformation.
   (a) Show that for all $v_1, v_2 \in V$,
   $$\langle v_1, v_2 \rangle = \frac{\|v_1 + v_2\|^2 - \|v_1 - v_2\|^2}{4}.$$
   (b) Show that T is an isometry if and only if for all $v \in V$, $\|Tv\| = \|v\|$.
   (c) Suppose that T is an isometry. Prove that T is injective.
2. Read and understand the section that goes through the proof of theorem 31.10.

Solution...

1. (a) Let $v_1, v_2 \in V$. Then
   $$\begin{aligned}\|v_1 + v_2\|^2 - \|v_1 - v_2\|^2 &= \langle v_1 + v_2, v_1 + v_2 \rangle - \langle v_1 - v_2, v_1 - v_2 \rangle \\ &= \ + \langle v_1, v_1 \rangle + \langle v_1, v_2 \rangle + \langle v_2, v_1 \rangle + \langle v_2, v_2 \rangle \\ &\quad\ - \langle v_1, v_1 \rangle + \langle v_1, v_2 \rangle + \langle v_2, v_1 \rangle - \langle v_2, v_2 \rangle \\ &= 4\langle v_1, v_2 \rangle.\end{aligned}$$
   Thus,
   $$\langle v_1, v_2 \rangle = \frac{\|v_1 + v_2\|^2 - \|v_1 - v_2\|^2}{4}.$$
   (b) Suppose T is an isometry and $v \in V$. Then
   $$\|Tv\|^2 = \langle Tv, Tv \rangle = \langle v, v \rangle = \|v\|^2$$
   and so $\|Tv\| = \|v\|$.
   Conversely, suppose that for all $v \in V$, $\|Tv\| = \|v\|$, and let $v_1, v_2 \in V$. Then
   $$\begin{aligned}\langle Tv_1, Tv_2 \rangle &= \frac{\|Tv_1 + Tv_2\|^2 - \|Tv_1 - Tv_2\|^2}{4} \\ &= \frac{\|T(v_1 + v_2)\|^2 - \|T(v_1 - v_2)\|^2}{4} = \frac{\|v_1 + v_2\|^2 - \|v_1 - v_2\|^2}{4} = \langle v_1, v_2 \rangle,\end{aligned}$$
   so T is an isometry.
   (c) Suppose that T is an isometry and that $v \in \ker T$.
   Then $\|v\| = \|Tv\| = \|0\| = 0$. Thus, $v = 0$, so $\ker T = \{0\}$ and T is injective.

37 Lecture on September 6th

37.1 Gram-Schmidt and orthonormal bases

Theorem 37.1.1. *Let V be an inner product space and $w_1, \ldots, w_n \in V$. Suppose that $(w_1, \ldots, w_n)$ is orthogonal and consists of non-zero vectors, $\lambda_1, \ldots, \lambda_n$ are scalars, and that*

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

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 and $w_1, \ldots, w_n \in V$. Suppose that $(w_1, \ldots, w_n)$ is orthogonal and consists of non-zero vectors, $\lambda_1, \ldots, \lambda_n$ are scalars, and that

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

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

Corollary 37.1.2. *Suppose V is an inner product space and that $w_1, \ldots, w_n \in V$. If $(w_1, \ldots, w_n)$ is orthogonal and consists of non-zero vectors, then $(w_1, \ldots, w_n)$ is linearly independent.*

Corollary 37.1.3. *Let V be an inner product space and $w_1, \ldots, w_n \in V$. Suppose that $(w_1, \ldots, w_n)$ is orthonormal, $\lambda_1, \ldots, \lambda_n$ are scalars, and that*

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

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

Corollary 37.1.4. *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 37.1.5 (Gram-Schmidt)**.** *Let V be an inner product space, $v_1, \ldots, v_n \in V$, 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, $v_1, \ldots, v_n \in V$, 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 some $k \in \{1, \ldots, n-1\}$. We show the properties hold 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, $v_{k+1} \in \operatorname{span}(w_1, \ldots, w_k) = \operatorname{span}(v_1, \ldots, v_k)$ which contradicts the fact $(v_1, \ldots, v_n)$ is linearly independent.

The definition of w_{k+1} gives

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

and so $\operatorname{span}(w_1, \ldots, w_k, w_{k+1}) \subseteq \operatorname{span}(v_1, \ldots, v_k, v_{k+1})$. Because $(w_1, \ldots, w_k, w_{k+1})$ 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 $\operatorname{span}(w_1, \ldots, w_k, w_{k+1}) = \operatorname{span}(v_1, \ldots, v_k, v_{k+1})$. $\square$

Theorem 37.1.6. *Suppose V is a finite-dimensional inner product space and that $(w_1, \dots, w_m)$ is orthonormal. Then we can extend $(w_1, \dots, w_m)$ to an orthonormal basis $(w_1, \dots, w_n)$ of V.*

Proof. Suppose V is a finite-dimensional inner product space and that $(w_1, \dots, w_m)$ is orthonormal. Extend $(w_1, \dots, w_m)$ to a basis of V: $(w_1, \dots, w_m, v_{m+1}, \dots, v_n)$. The Gram-Schmidt process leaves $w_1, \dots, w_m$ alone, so run Gram-Schmidt to obtain orthogonal $(w_1, \dots, w_m, w_{m+1}, \dots, w_n)$. Consider

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

It is orthonormal and so it's linearly independent. Theorem 18.4 tells us it is a basis. Thus, it's an orthonormal basis for V. □

Corollary 37.1.7. *Suppose V is a finite-dimensional inner product space.*
Then V has an orthonormal basis.

37.2 Orthogonal complements and the Riesz representation lemma

Definition 37.2.1. 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 37.2.2. You can check that $S^\perp$ is a subspace.

Theorem 37.2.3. *Suppose V is a finite-dimensional vector space, that U, W are subspaces of V, and that $V = U \oplus W$. Then $\dim(V) = \dim(U) + \dim(W)$.*

Proof. Let $(u_1, \dots, u_m)$ be a basis for U, and $(w_1, \dots, w_n)$ be a basis for W. We claim that

$$(u_1, \dots, u_m, w_1, \dots, w_n)$$

is a basis for V.

Suppose $\lambda_1 u_1 + \ldots + \lambda_m u_m + \mu_1 w_1 + \ldots + \mu_n w_n = 0$. Then

$$\lambda_1 u_1 + \ldots + \lambda_m u_m = (-\mu_1)w_1 + \ldots + (-\mu_n)w_n \in U \cap W = \{0\}.$$

Linear independence of $(u_1, \dots, u_m)$ and $(w_1, \dots, w_n)$ show that all the coefficients are 0. Thus, $(u_1, \dots, u_m, w_1, \dots, w_n)$ is linearly independent. So $\dim(V) \geq m + n = \dim(U) + \dim(W)$.

We proved that $(u_1, \dots, u_m, w_1, \dots, w_n)$ spans V and that the opposite inequality holds on the homework. □

Theorem 37.2.4. *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 $n = \dim V$, and $m = \dim W$. Let $(w_1, \ldots, w_m)$ be an orthonormal basis for W. Extend it to an orthonormal basis of V: $(w_1, \ldots, w_m, w_{m+1}, \ldots, w_n)$. $(w_{m+1}, \ldots, w_n)$ consists of vectors in $W^\perp$. Thus, given $v \in V$, we have

$$v = \sum_{i=1}^{n} \langle v, w_i\rangle w_i = \sum_{i=1}^{m} \langle v, w_i\rangle w_i \ + \sum_{i=m+1}^{n} \langle v, w_i\rangle w_i \in W + W^\perp,$$

which completes the proof that $V = W \oplus W^\perp$.

Aside: notice that $(w_{m+1}, \ldots, w_n)$ is a basis of $W^\perp$. □

Once one progresses to the infinite-dimensional context, the next theorem is often called the *Riesz representation lemma*. In its statement, $\mathbb{F}$ can be taken to be $\mathbb{R}$ or $\mathbb{C}$.

Theorem 37.2.5. *Suppose V is a finite-dimensional inner product space over $\mathbb{F}$ and that $f : V \to \mathbb{F}$ 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 $\mathit{null}(f) = \dim(V) - 1$. Thus, $\dim(\ker(f)^\perp) = 1$. Pick any $s \in \ker(f)^\perp$ with $\|s\| = 1$, and let $r = \overline{f(s)}s$. We note:

- for $v \in \ker(f)$, $f(v) = 0 = \langle v, r\rangle$ (the second equality comes from $r \in \ker(f)^\perp$);
- when $v = r$, we have $v = \overline{f(s)}s$, so that

$$f(v) = f(\overline{f(s)}s) = \overline{f(s)}f(s) = |f(s)|^2 = |f(s)|^2\|s\|^2 = \|\overline{f(s)}s\|^2 = \|v\|^2 = \langle v, v\rangle = \langle v, r\rangle.$$

From these two observations, we conclude that $f(v) = \langle v, r\rangle$ for all $v \in \ker(f) + \mathrm{span}(r)$.

Since $s \in \ker(f)^\perp$ and $\|s\| = 1$, we have $s \notin \ker(f)$. Thus, $f(s) \neq 0$ and $\mathrm{span}(r) = \mathrm{span}(s) = \ker(f)^\perp$. This means $\ker(f) + \mathrm{span}(r) = V$, and so we have shown $f(v) = \langle v, r\rangle$ for all $v \in V$.

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 35.2.8 (part 4). □

38 Questions due on September 10th

1. Consider $\mathbb{R}^4$ with its standard inner product. Apply Gram-Schmidt to the 4-tuple:

$$\Big((1,1,1,1),\ (0,4,0,0),\ (0,0,12,0),\ (0,0,0,24)\Big).$$

Solution:

$$\begin{pmatrix}1\\1\\1\\1\end{pmatrix} = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$$

$$\begin{pmatrix}0\\4\\0\\0\end{pmatrix} - \begin{pmatrix}1\\1\\1\\1\end{pmatrix} = \begin{pmatrix}-1\\3\\-1\\-1\end{pmatrix}$$

$$\begin{pmatrix}0\\0\\12\\0\end{pmatrix} - 3\begin{pmatrix}1\\1\\1\\1\end{pmatrix} + \begin{pmatrix}-1\\3\\-1\\-1\end{pmatrix} = \begin{pmatrix}-4\\0\\8\\-4\end{pmatrix}$$

$$\begin{pmatrix}0\\0\\0\\24\end{pmatrix} - 6\begin{pmatrix}1\\1\\1\\1\end{pmatrix} + 2\begin{pmatrix}-1\\3\\-1\\-1\end{pmatrix} + \begin{pmatrix}-4\\0\\8\\-4\end{pmatrix} = \begin{pmatrix}-12\\0\\0\\12\end{pmatrix}$$

2. Define $\langle\ ,\ \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ by

$$\Big\langle (x_1,x_2),\ (y_1,y_2)\Big\rangle = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2.$$

   (a) Check that $\langle\ ,\ \rangle$ is an inner product on $\mathbb{R}^2$.

   (b) Find an orthonormal basis for $\mathbb{R}^2$ with respect to this inner product.

Solution:

   (a) I checked (privately!) that $\langle\ ,\ \rangle$ is linear in the first variable, and that it has the requisite symmetry. For the last property, note that

$$\Big\langle (x_1,x_2),\ (x_1,x_2)\Big\rangle = 2x_1^2 + 2x_1x_2 + x_2^2 = x_1^2 + (x_1+x_2)^2.$$

   If $(x_1,x_2) \neq 0$, then either $x_1 \neq 0$, or $x_1 = 0$ and we must have $x_1 + x_2 \neq 0$. In both cases, we see that

$$\Big\langle (x_1,x_2),\ (x_1,x_2)\Big\rangle > 0.$$

   (b) $\big((1,-1),\ (0,1)\big)$.

3. Suppose that V is an inner product space.

   (a) Suppose that S is a nonempty subset of V. Prove that $S^\perp$ is a subspace of V.
   (b) Suppose that $\emptyset \neq S' \subseteq S$. Prove that $(S')^\perp \supseteq S^\perp$.
   (c) Suppose that $v_1, \ldots, v_n \in V$. Prove that $\{v_1, \ldots, v_n\}^\perp = (\text{span}(v_1, \ldots, v_n))^\perp$.
   (d) Suppose that W is a subspace of V. Prove that $W \subseteq (W^\perp)^\perp$.
   (e) Suppose, in addition, that V is finite-dimensional.
   Use theorem 37.2.4 to prove that $W = (W^\perp)^\perp$.
   (f) [Optional and only accessible if you've done 131A]
   Let
   $$l_2 = \left\{ (a_n)_{n=1}^\infty : \quad \sum_{n=1}^\infty a_n^2 \text{ converges} \right\}.$$
   It is a theorem that l_2 is a vector space over $\mathbb{R}$ when given the following operations.
   $$(a_n)_{n=1}^\infty + (b_n)_{n=1}^\infty := (a_n + b_n)_{n=1}^\infty, \ \lambda (a_n)_{n=1}^\infty := (\lambda a_n)_{n=1}^\infty.$$
   Moreover, it is also a theorem that l_2 has a well-defined inner product:
   $$\left\langle (a_n)_{n=1}^\infty, (b_n)_{n=1}^\infty \right\rangle = \sum_{n=1}^\infty a_n b_n.$$
   Let $W = \{(a_n)_{n=1}^\infty : \quad \text{there exists an } N \in \mathbb{N} \text{ such that } a_n = 0 \text{ whenever } n \geq N\}$.
   It is easy to show that W is a subspace of l_2. Prove that $W \neq (W^\perp)^\perp$.

Solution:

   (a) Routine.
   (b) Suppose $v \in S^\perp$ and $s \in S'$. Since $S' \subseteq S$, $s \in S$, so $\langle v, s \rangle = 0$.
   Since $v \in S^\perp$ and $s \in S'$ were arbitrary, this shows $S^\perp \subseteq (S')^\perp$.
   (c) $\{v_1, \ldots, v_n\} \subseteq \text{span}(v_1, \ldots, v_n)$, so the previous part gives
   $$\{v_1, \ldots, v_n\}^\perp \supseteq \text{span}(v_1, \ldots, v_n)^\perp.$$
   For the other inclusion, suppose $w \in \{v_1, \ldots, v_n\}^\perp$ and $v \in \text{span}(v_1, \ldots, v_n)$.
   We can find scalars $\lambda_1, \ldots, \lambda_n$ such that $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$. Then
   $$\langle w, v \rangle = \left\langle w, \lambda_1 v_1 + \ldots + \lambda_n v_n \right\rangle = \overline{\lambda_1} \langle w, v_1 \rangle + \ldots + \overline{\lambda_n} \langle w, v_n \rangle = \overline{\lambda_1} 0 + \ldots + \overline{\lambda_n} 0 = 0.$$
   Since $w \in \{v_1, \ldots, v_n\}^\perp$ and $v \in \text{span}(v_1, \ldots, v_n)$ were arbitrary, this shows
   $$\{v_1, \ldots, v_n\}^\perp \subseteq \text{span}(v_1, \ldots, v_n)^\perp.$$
   (d) "Obvious."

(e) We have $W \subseteq (W^\perp)^\perp$.
Also, $\dim((W^\perp)^\perp) = \dim V - \dim W^\perp = \dim V - (\dim V - \dim W) = \dim W$.
Thus, $W = (W^\perp)^\perp$.

(f) Let e_j be the sequence $(\delta_{j,n})_{n=1}^\infty$, the sequence whose j-th term is 1, and all other terms are 0. Notice that for all $j \in \mathbb{N}$, $e_j \in W$.
Now suppose $(a_n)_{n=1}^\infty \in W^\perp$. Then for all $j \in \mathbb{N}$, we have

$$a_j = \left\langle (a_n)_{n=1}^\infty, e_j \right\rangle = 0.$$

Thus, $(a_n)_{n=1}^\infty = 0$. We have shown $W^\perp = \{0\}$, so $(W^\perp)^\perp = l_2$. $W \neq (W^\perp)^\perp$ because

$$\left(\frac{1}{2^n}\right)_{n=1}^\infty \in (W^\perp)^\perp \setminus W.$$

4. [Optional] Suppose that V is a finite-dimensional inner product space and that $T : V \longrightarrow V$ is a linear transformation. Let $T^* : V \longrightarrow V$ be the adjoint of T (this is defined in definition 39.1.2).

(a) Prove $\ker T \subseteq (\operatorname{im} T^*)^\perp$.

(b) Prove $\ker T \supseteq (\operatorname{im} T^*)^\perp$. *Hint: Let $v \in (\operatorname{im} T^*)^\perp$.*
You want to show $Tv = 0$. For this, it is enough to show $\|Tv\|^2 = 0$.

(c) You've shown $\ker T = (\operatorname{im} T^*)^\perp$.

(d) Use (c), a result from question 3, and theorem 39.1.3 to prove $\operatorname{im} T = (\ker T^*)^\perp$.

Solution:

(a) Let $v \in \ker T$ and $w \in \operatorname{im} T^*$. We can find a $v' \in V$ such that $w = T^*(v')$. Thus,

$$\langle v, w\rangle = \langle v, T^*(v')\rangle = \langle T(v), v'\rangle = \langle 0, v'\rangle = 0.$$

Since v and w were arbitrary, this shows $\ker T \subseteq (\operatorname{im} T^*)^\perp$.

(b) Let $v \in (\operatorname{im} T^*)^\perp$. Then $T^*(Tv) \in \operatorname{im} T^*$, so

$$\|Tv\|^2 = \langle Tv, Tv\rangle = \langle v, T^*(Tv)\rangle = 0.$$

Thus, $Tv = 0$ and so $v \in \ker T$.

(c) We have $\ker T = (\operatorname{im} T^*)^\perp$.

(d) Using (c) and 3(e), we get $\operatorname{im} T^* = ((\operatorname{im} T^*)^\perp)^\perp = (\ker T)^\perp$.
Replacing T by T^* and using theorem 39.1.3 shows $\operatorname{im} T = \operatorname{im}((T^*)^*) = (\ker T^*)^\perp$.

5. Prepare for the final. See the final section of these notes for what to expect.

39 Lecture (final one) on September 10th: The spectral theorem

Throughout all of this lecture, $\mathbb{F}$ will refer to $\mathbb{R}$ or $\mathbb{C}$.

39.1 Adjoints and self-adjoint operators

Theorem 39.1.1. *Suppose V is a finite-dimensional inner product space 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. First, fix $v' \in V$. Define $f_{v'} : V \longrightarrow \mathbb{F}$ by $f_{v'}(v) = \langle T(v), v' \rangle$. $f_{v'}$ is linear and so the Riesz representation lemma provides an $r_{v'}$ such that for all $v \in V$,

$$f_{v'}(v) = \langle v, r_{v'} \rangle, \text{ i.e. } \langle T(v), v' \rangle = \langle v, r_{v'} \rangle.$$

We can now vary v', and define $T^* : V \longrightarrow V$ by $T^*(v') = r_{v'}$.

The uniqueness of $r_{v'}$ implies T^* is unique.

I'll prove in lecture that T^* is linear. □

Definition 39.1.2. Suppose V is a finite-dimensional inner product space and that $T : V \longrightarrow V$ is a linear transformation. The unique linear transformation $T^* : V \longrightarrow V$ 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 39.1.3. *Suppose V is a finite-dimensional inner product space, 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 39.1.4. *Suppose V is an inner product space, $\beta = (w_1, \dots, w_n)$ is an orthonormal basis of V, $T : V \longrightarrow V$ is a linear transformation, and $A = [T]_\beta^\beta$. Then for $i, j \in \{1, \dots, n\}$,*

$$A_{ij} = \langle T(w_j), w_i \rangle.$$

Proof. Suppose V is an inner product space, $\beta = (w_1, \dots, w_n)$ is an orthonormal basis of V, that $T : V \longrightarrow V$ is a linear transformation, and $A = [T]_\beta^\beta$. Then for $j \in \{1, \dots, n\}$,

$$T(w_j) = \sum_{k=1}^{n} A_{kj} w_k.$$

Thus, for $i, j \in \{1, \dots, n\}$,

$$\langle T(w_j), w_i \rangle = \left\langle \sum_{k=1}^{n} A_{kj} w_k, w_i \right\rangle = \sum_{k=1}^{n} A_{kj} \langle w_k, w_i \rangle = A_{ij} \langle w_i, w_i \rangle = A_{ij}.$$

□

Theorem 39.1.5. *Suppose V is an inner product space, $\beta = (w_1, \dots, w_n)$ is an orthonormal basis of V, $T : V \longrightarrow V$ is a linear transformation, $A = [T]_\beta^\beta$, and $B = [T^*]_\beta^\beta$. Then for $i, j \in \{1, \dots, n\}$, $A_{ij} = \overline{B_{ji}}$.*

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,

$$A_{ij} = \langle T(w_j), w_i \rangle = \langle w_j, T^*(w_i) \rangle = \overline{\langle T^*(w_i), w_j \rangle} = \overline{B_{ji}}.$$

□

Definition 39.1.6. Suppose V is a finite-dimensional inner product space and that $T : V \longrightarrow V$ is a linear transformation. We say T is *self-adjoint* **iff** $T^* = T$.

Theorem 39.1.7. *Suppose V is a finite-dimensional complex inner product space, that $T : V \to V$ 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 39.1.8. The word “complex” is not necessary in the previous theorem. However, when V is a real inner product space, the result is not very interesting!

39.2 The spectral theorem

Theorem 39.2.1. *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 a matrix with *real* entries. Because T is self-adjoint, theorem 39.1.5 tells us that A is an $n \times n$ *symmetric* matrix. Now define a linear transformation between *complex* vector spaces by

   $$S : \mathbb{C}^n \longrightarrow \mathbb{C}^n,\ x \longmapsto Ax.$$

   Because $[T]_\beta^\beta = A = [S]_{(e_1,\ldots,e_n)}^{(e_1,\ldots,e_n)}$, 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 S is a linear transformation between vector spaces over $\mathbb{C}$, S has an eigenvalue $\lambda \in \mathbb{C}$. We claim that $\lambda \in \mathbb{R}$. By theorem 39.1.7, 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 39.2.2 (Spectral Theorem)**.** *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. □

40 Exam Expectations

The exam will contain at most 7 questions. Most likely, it will actually contain 6 questions, but I don't want to go back on my word. So let's stick with the first thing I said. This probably means the exam will be less of a time crunch than the quizzes. Usually I like my finals to take 2 out of 3 hours. Since this is a 2 hour exam, I'll aim for it to take 90 mins out of the 2 hours. That way, if you waste time on a question, or panic, it should still be salvageable.

The main difference between the final and the quizzes, is that I can ask for proofs from lecture notes. We covered a lot so I don't expect you to be able to prove everything. In the rest of this document, I'll say how important each of the previous 39 sections are to the final (even though I haven't written it yet).

1. The exam will be 2 hours long and will take place in the usual classroom.

2. This class was very fast paced. If you found it very difficult, that is quite reasonable; it does *not* mean that you are bad at math. I took upper division math classes for an entire year in my undergraduate before taking an exam on any of the material. I genuinely sympathize with your predicament. I hope this section helps your preparation for the final.

3. You should be able to verify subset inclusions and set equality. You should be comfortable with functions. I will not ask you to prove injectivity or surjectivity of functions which are not linear; injectivity or surjectivity of linear functions is fair game. Of course, injectivity is easier for linear functions because you can prove the the kernel is $\{0\}$.

4. The most important thing learned in this homework was how to process many if-then sentences to write a proof. Although you don't need to know question 2, it might be useful to be familiar with the results, and it might be helpful to go over the solution again with fresh eyes.

5. Real business begins. I expect you to be able to check axioms. I do not expect you to be able to remember axioms by number. I will always state the axiom I want you to prove.

   I will not ask you to prove anything from theorem 5.7. Those results are safely assumed.

6. I really, really hope both of these are easy by now.

7. I will not ask you to prove theorem 7.1.4 and you can use it and refer to it as the "subspace test" freely.

   Notice that you can now prove theorem 7.1.8... Suppose W is a subspace of $\mathbb{R}^3$. Then $0 \leq \dim W \leq \dim \mathbb{R}^3 = 3$. If $\dim W = 1$, you can pick $(a_1, a_2, a_3) \in W \setminus \{0\}$. If $\dim W = 2$, then $\dim W^\perp = 1$, and you can pick $(n_1, n_2, n_3) \in W^\perp \setminus \{0\}$. If $\dim W = 0$ or 3, then $W = \{0\}$ or $\mathbb{R}^3$, respectively.

   I won't ask about infinite intersections; $U \cap W$ will do.

   Except for the final example, you should know all of 7.2. I won't ask for the proof of the first lemma, but it is useful to know the result.

8. Most important questions: 3.(b)(e), 5.(a), some smart students have been using 6.(b) to save time, 7., 8., 9.

9. Everything in 9.1 is fair game, proofs included.

   From 9.2, you should definitely know what Γ_α means. As for the rest of the 9.2, you should know by now whether you find it useful to think about or not. Certainly, I have encountered some students who are finding this way of thinking useful. But you may also find it confusing. I think you should at least know that the final corollary is true. That corollary is the entire reason matrices are even a thing. I won't ask for the proof.

   9.4 will not be examined explicitly. However, if you still think about matrix-vector multiplication as some terrible thing you do with a load of numbers for no reason, then you have missed the point of the matrix of a linear transformation, and the linear transformation associated with a matrix.

10. I hope question 1 feels easy compared to the other things we have been doing.

11. Probably the most essential section of the entire class.

    Again, if you like thinking about Γ_α, that is great - me too! If not, I wrote things in such a way that you can ignore my remarks concerning it. However, eventually it becomes essential for defining various isomorphisms.

    11.2 will not be examined.

12. This was all routine or tedious.

13. The proof of theorem 13.1.1 is examinable.

    Notice that wherever there is an "if and only if" statement, I could ask a shorter question by asking for the proof of only one direction.

    I think it is a good exercise to prove theorem 13.2.3 using the definition of span and linear independence (with no reference to Γ_α).

    You should know theorem 13.2.5, but its proof is too long for me to examine.

    You know by now that the consequences of the replacement theorem are more important than the theorem itself.

    You should know the first conclusion of the replacement theorem: that linearly independent tuples are smaller than or equal to spanning tuples. You do not need to know the other part, or the proof.

14. Question 6, 7, and 8 are the most important.

15. ... as are the solutions!

16. Question 1 would now be slightly silly because you can make a dimension argument.

    I already said above that I wouldn't bother with a question 3 type question again.

    Question 3 on quiz 2 shows that asking something like question 4(b) again would also be a bit silly. This is because you can calculate a basis for the kernel and the image of a linear transformation using the rank-nullity theorem and the replacement theorem in the same time as calculating a kernel directly.

17. The comments about question 2 might still be relevant.

18. This section contains some of the most important results of the class.

    The corollary is a very quick proof which you should know.

    I will not ask for the proof of theorem 18.4, but I think that understanding it should be easier than when we first covered it, and could be useful for you. These type of size arguments have shown up a few times now.

19. Question 1 was probably tedious. Question 2 was easy, but important.

    The main point of question 3 was to have you write proofs correctly by addressing the remark I made before the question.

20. You should know the results of 20.1. I won't ask for proofs.

    The proof of the rank-nullity theorem is reasonably long. But I could ask for you to prove $rank(T) + null(T) \leq \dim V$ or $rank(T) + null(T) \geq \dim V$. Then that cuts the proof in half. It is too nice of a proof, and too important of a result for me to say this one is non-examinable. Learn it!

    Theorem 20.2.4 is incredibly important. It might be a weird one for me to ask for the proof, but I still think you should know the proof.

    You should understand example 20.2.5 flawlessly.

    You can ignore the first lemma of 20.3. You should know the other results from 20.3. Only the proof of the last theorem is examinable. Since this builds on a homework question, I would arrange a question on it so that only a subset of the homework is required, and all of it is provable in one go (similar to quiz 2, question 4).

21. `http://soundcloud.com/mjandr/sets/best-of`

22. You should be able to prove the first two theorems.

    You should know the third theorem. Proving it is unnecessary.

    You should definitely know theorem 22.1.4, and be able to use it. You do not need to be able to prove it.

    You should know all of 22.2.

23. All of this homework is important.

24. You should know the results of 24.1. The proofs can be ignored.

    24.2 can be ignored.

    Theorem 24.3.5 is one of my favorite results. I won't ask for the proofs of this section, but you should understand what the important ideas are in these proofs. Mostly, there's applications of theorem 20.2.4 and theorem 22.1.1.

    Why can't a non-square matrix A define an isomorphism T_A?

    24.4 stunned me. I won't ask for his proof, but note how he makes use of theorem 22.1.4.

25. All of this homework was important to check that you can do the calculations.

    Revisit 1.(h) to help put diagonalization into persepctive.

26. So many matrices. Notice the examples that I gave of coordinate change matrices.

27. I won't ask for proofs of the results in 27.1.

    All I'll expect from 27.2 is that you can calculate very easy determinants.

    In 27.3, you can believe the unproved theorem. You should know everything else.

28. `http://vimeo.com/mjandr`

29. All of these types of questions are reasonable for the final.

30. I said some things, I guess, hopefully useful.

31. You should know all of this section as written (i.e. only one proof).

    I'll address the big theorem's proofs shortly.

32. Direct sums are fair game on the final.

    I think questions 1, 4, 5, 6, and 7 are the most important questions

33. ... as are the solutions!

34. You should be able to prove theorem 34.1.

    Other than that, you should read the proofs to help your understanding. Ideas in them could be useful, but I won't examine them explicitly.

35. In the exam, I will not use complex numbers. The reason I introduced them is so that I can prove the spectral theorem, and so that I could tell you the full story about inner product spaces. (Both of these topics are part of UCLA's 115A syllabus.)

    So you can ignore 35.1.

    In 35.2, you can ignore anything to do with the word "conjugate." You can ignore the 131A example I gave. You should know the poof of theorem 35.2.8 (it's in the book).

    You do not need to know the proofs of the triangle inequality and Cauchy-Schwartz.

    You do not need to know about "angles" other than the concept of orthogonal.

    The Fourier basis is off-syllabus even though it is my favorite bit of math ever.

36. This question was a piece of cake.

37. The proof of Gram-Schmidt is too long. You do not need to know this proof.

    The Riesz representation lemma is probably a little confusing. You do not need to know this result.

    Everything else is fair game.

38. Questions 2 and 3 are important.

    I changed my mind from earlier, and decided that question 4 is non-examinable.

39. Non-examinable, but it really is one of the coolest things we have done.

    Please try and appreciate this after the exam has passed.

40. By commenting on this section within this section, does it create a glitch in the MATRIX?

41 The final

1. Suppose V and W are vector spaces over $\mathbb{R}$, that $S : V \longrightarrow W$ is a linear transformation, and that $v_1, v_2, \ldots, v_n \in V$. Always true or sometimes false (i.e. depends on V, W, S, etc.)?

   (a) If $\ker S = \{0\}$ and $(v_1, v_2, \ldots, v_n)$ is linearly independent, then $(S(v_1), S(v_2), \ldots, S(v_n))$ is linearly independent.

   (b) If $(v_1, v_2, \ldots, v_n)$ spans V and $(S(v_1), S(v_2), \ldots, S(v_n))$ is a basis for W, then S is an isomorphism.

   (c) If S is an isomorphism and $(v_1, v_2, \ldots, v_n)$ spans V, then $(S(v_1), S(v_2), \ldots, S(v_n))$ is a basis for W.

   (d) If $(S(v_1), S(v_2), \ldots, S(v_n))$ is a basis for W and S is an isomorphism, then $(v_1, v_2, \ldots, v_n)$ spans V.

   Solution. T,T,F,T.

2. Let V and W be vector spaces over $\mathbb{R}$ and let $T : V \longrightarrow W$ be a linear transformation.

   (a) Suppose that $(v_1, v_2, \ldots, v_n)$ is a tuple of vectors in V which spans V, and that $(T(v_1), T(v_2), \ldots, T(v_n))$ is linearly independent. Prove that $\ker T = \{0\}$.
   Solution. 8/20 homework, question 8.(f).

   (b) Suppose that $\ker T = \{0\}$. Prove that T is injective, i.e. that

   $$\text{if } v_1, v_2 \in V \text{ and } T(v_1) = T(v_2), \text{ then } v_1 = v_2.$$

   Solution. Second half of theorem 9.1.4.

3. (a) Suppose V and W are finite-dimensional vector spaces over $\mathbb{R}$ and that $T : V \longrightarrow W$ is a linear transformation. Prove that $rank(T) + null(T) \leq \dim V$.
   Solution. Write out the first half of the proof of the rank-nullity theorem, and then use consequences of the replacement theorem to conclude the inequality.

   (b) Suppose U, V, and W are finite-dimensional vector spaces over $\mathbb{R}$, and that

   $$T : U \longrightarrow V,\ S : V \longrightarrow W$$

   are linear transformations. Prove that $rank(ST) \leq rank(T)$.
   Solution. 8/20 homework, question 6.(a) and 8/27 homework, question 4.

   (c) Suppose $m, n \in \mathbb{N}$, $m < n$, A is an $n \times m$ matrix, and B is an $m \times n$ matrix. Prove that $AB \neq I_n$, but $BA = I_m$ is possible.
   Solution. If $AB = I_n$, we'd have $n = rank(T_{AB}) = rank(T_A T_B) \leq rank(T_B) \leq m < n$. On the other hand, if $A = \begin{pmatrix} I_m \\ 0_{(n-m)\times m} \end{pmatrix}$ and $B = \begin{pmatrix} I_m & 0_{m\times(n-m)} \end{pmatrix}$, then $BA = I_m$.

4. Let $T : M_{2\times 2}(\mathbb{R}) \longrightarrow M_{2\times 2}(\mathbb{R})$ be the linear transformation defined by

$$T\left(\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\right) = \begin{pmatrix} 2 & 128 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

Calculate $\det(T)$.

Solution.

Let $\beta = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)$.

Then $[T]_\beta^\beta = \begin{pmatrix} 2 & 128 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 128 \\ 0 & 0 & 0 & 4 \end{pmatrix}$. So $\det(T) = 2 \cdot 4 \cdot 2 \cdot 4 = 64$.

5. Suppose V is a finite-dimensional vector space over $\mathbb{R}$, $T : V \longrightarrow V$ is linear, and $T^2 = 1_V$. Recall that the eigenspace of λ is defined by $E_\lambda = \ker(T - \lambda 1_V)$.

(a) Prove that $E_1 \cap E_{-1} = \{0\}$.

(b) Let $w_1 \in E_1$ and $w_{-1} \in E_{-1}$, and let $v = w_1 + w_{-1}$. Prove that $w_1 = \frac{v+Tv}{2}$.

(c) Prove that $V = E_1 + E_{-1}$.

(d) Is T diagonalizable? Prove your claim.

Solution.

(a) Let $v \in E_1 \cap E_{-1} = \{0\}$. Then $v = Tv = -v$, so $v = 0$.

(b) Let $w_1 \in E_1$ and $w_{-1} \in E_{-1}$, and let $v = w_1 + w_{-1}$.
We have $Tv = w_1 - w_{-1}$, so $v + Tv = (w_1 + w_{-1}) + (w_1 - w_{-1}) = 2w_1$.
Thus, $w_1 = \frac{v+Tv}{2}$.

(c) Let $v \in V$.
Notice $T(v + Tv) = Tv + T^2 v = Tv + v = 1(v + Tv)$, so $v + Tv \in E_1$.
Similarly, $T(v - Tv) = Tv - T^2 v = Tv - v = (-1)(v - Tv)$, so $v - Tv \in E_{-1}$.
Thus, $v = \frac{v+Tv}{2} + \frac{v-Tv}{2} \in E_1 + E_{-1}$.

(d) Let $(a_1, \ldots, a_m)$ be a basis for E_1 and $(b_1, \ldots, b_n)$ be a basis for E_{-1}.
Then $\alpha = (a_1, \ldots, a_m, b_1, \ldots, b_n)$ spans $E_1 + E_{-1} = V$. In fact, α is a basis for V; even without proving this, we know some that sub-tuple of α is a basis of V.
Since α consists of eigenvectors of T, this shows that T is diagonalizable.

6. (a) Let V be an inner product space over $\mathbb{R}$.

Suppose that $(v_1, \ldots, v_n)$ is orthonormal tuple of vectors in V.

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

Solution. Suppose $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, $\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$, and let $j \in \{1, 2 \ldots, n\}$. Then

$$\lambda_j = \lambda_j \langle v_j, v_j \rangle = \sum_{i=1}^{n} \lambda_i \langle v_i, v_j \rangle = \left\langle \sum_{i=1}^{n} \lambda_i v_i, v_j \right\rangle = \langle 0, v_j \rangle = 0.$$

(b) Define an unusual inner product on $\mathbb{R}^2$ by $\langle \ , \ \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$

$$\Big\langle (x_1, x_2),\ (y_1, y_2) \Big\rangle = x_1 y_1 + 2x_1 y_2 + 2x_2 y_1 + 5x_2 y_2.$$

Find an orthonormal basis for $\mathbb{R}^2$.

Solution. $(1, 0)$, $(2, -1)$.

(c) Let V be an inner product space over $\mathbb{R}$.

Suppose that U and W are subspaces of V.

Prove that $U^\perp + W^\perp \subseteq (U \cap W)^\perp$.

Solution. Let $v' \in U^\perp + W^\perp$. By definition of addition of sets, we can write $v' = u' + w'$ for some $u' \in U^\perp$ and some $w' \in W^\perp$.

Now let $v \in U \cap W$.

Because $u' \in U^\perp$ and $v \in U$, we have $\langle u', v \rangle = 0$.

Because $w' \in W^\perp$ and $v \in W$, we have $\langle w', v \rangle = 0$.

Thus, $\langle v', v \rangle = \langle u', v \rangle + \langle w', v \rangle = 0$.

Since v was an arbitrary element in $U \cap W$, this shows that $v' \in (U \cap W)^\perp$.