MAT240 Lecture Notes

Arkyter

'23 Fall Semester

§1 Day 1: Introduction to Class (Sep. 7, 2023)

General course guidelines can be found in the page in Quercus; weekly homework will be issued with the lowest two grades being dropped. This class is "an introduction to linear algebra over an arbitrary field aimed at students with a very serious interest in mathematics," so yeah. Have fun!

§1.1 Review on Sets

A *set* is a collection of objects (referred to as *elements* of said set). For example, we may have sets as such,

$$A = \{1, 2, 3, 4\}$$

$$\mathbb{Z} := \text{"the integers," aka } \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$$

Note that a set need not only include numbers; it is to be thought of strictly as a collection of objects, which means that order does not matter. As for notation, we have "$\in$", denoting "is an element of." This is used in contexts such as $1 \in A$, or $5 \notin A$, which would respectively read as "1 is an element of A" and "5 is not an element of A." Set-builder notation is used to write out sets which would be more conveniently expressed with a predicate, i.e.

$$B = \{x \in \mathbb{Z} \mid 1 \leq x \leq 4\},$$

which would read as the set of all integers from 1 to 4, or

$$\mathbb{Q} = \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z} \wedge q \neq 0 \right\},$$

which expresses the set of rationals. Some other examples include

$$\mathbb{P} = \{p \in \mathbb{Z} \mid p \text{ prime}\}^{1},$$

or $\emptyset$, which represents the null set (aka the empty set), a set with no elements.

When comparing sets, we use $X = Y$ to denote an equality between sets X and Y, which occurs if and only if they both share the same elements. It may be thought of as $\forall x, [x \in X \iff x \in Y]$. Moreover, a set X is said to be a subset of another set Y, $X \subset Y$, if all elements of X are also elements of Y. Note the differentiation between $\subseteq$ and $\subset$, the former including the equality case, the latter being strict. This may be intuitively described as the difference between $\geq$ and $>$. As a basic example, the empty set is a subset of all sets S, written as

$$\emptyset \subseteq S.$$

[1] i'm not sure if this is acceptable notation in the course, i'm just used to seeing it in books

Unions of sets can be expressed with the symbol $\cup$, as in

$$X \cup Y = \{x \mid x \in X \text{ or } x \in Y\},$$

$$\bigcup_{i=1}^{n} X_i = X_1 \cup X_2 \cup X_3 \cup \cdots \cup X_n.$$

Intersections are expressed with the symbol $\cap$, as in

$$X \cap Y = \{x \mid x \in X \text{ and } x \in Y\},$$

$$\bigcap_{i=1}^{n} X_i = X_1 \cap X_2 \cap X_3 \cap \cdots \cap X_n.$$

And finally, set differences are expressed with the symbol $\setminus$, as in

$$X \setminus Y = \{x \mid x \in X \text{ but } x \notin Y\}.$$

The cartesian product of sets can be expressed as

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}.$$

For example, the real plane is expressed as $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$. All sorts of stuff can be written out with sets. Though, if you want an extreme example, here's Fermat's Last Theorem:

Theorem 1.1 (Fermat's Last Theorem)
There does not exist $x, y, z \in \mathbb{N}$ and $n \in \mathbb{N} \geq 3$ such that $x^n + y^n = z^n$.

$$\{(x, y, z, n) \in \mathbb{N}^4 \mid n \geq 3, x^n + y^n = z^n\} = \emptyset.$$

§1.2 Functions

[2] A function f is written as

$$f : X \to Y,$$

with sets X and Y its domain and codomain respectively. Note that the codomain of a function is different from the range; the range would be the set of all outputs, as in $\{f(x) \mid x \in X\}$, while the codomain is a general set simply expressing "what may possibly come out," such as "real numbers" or "rational numbers" and so forth. Two functions f_1, f_2 are considered equal ($f_1 = f_2$) if and only if

$$f_1(x) = f_2(x) \; \forall x \in X.$$

[2] not much to say here

§2 Day 2: Functions and Fields (Sep. 12, 2023)

Today serves as a reminder and an introduction to formal definition of functions and properties. First off, recall that a function is written as

$$f : \underbrace{X}_{\text{domain}} \to \underbrace{Y}_{\text{codomain}},$$

in which f is a function that assigns, for every $x \in X$, a unique $f(x) \in Y$. This can alternatively be written as a map using the symbol $\mapsto$, as in,

$$x \underbrace{\mapsto}_{\text{maps to}} f(x)$$

for all $x \in X$ under an implicit assumption. Note: range is discussed in my Day 1 notes.

§2.1 Example Functions and Properties

Recall that two functions, $f_1, f_2 : X \to Y$, are considered to be the *same function* if and only if they share the same domain and codomain (as aforementioned) and, for all $x \in X$, we have $f_1(x) = f_2(x)$. Take away the constraint on sharing domains and codomains, and we may end up in an awkward situation where $f_1 : X_1 \to Y$ and $f_2 : X_2 \to Y$, where $\exists x \in X_1$, but $x \notin X_2$, thus nullifying the equality. Even if functions appear similar, changing codomains may remove the invertibility, surjectivity, et al. which is why we must strictly enforce all conditions above.

Now, let's consider a few examples in which we will use to provide intuition.

Example 2.1 (Identity Function)
Define an identity function $\mathrm{id}_X : X \to X$ where $x \mapsto x$ for any set X.

Example 2.2 (Inclusion Function)
Given two sets $X \subset Y$, the inclusion map is $\iota : X \to Y$, where $x \mapsto x$. This function differs from the identity in that we are recognizing elements $x \in X$ as to be regarded as included in its superset Y.

Example 2.3 (Example 1: f_1)
Let f_1 be an alphabet enumeration function where $f_1 : \{A, B, C, \ldots, Z\} \to [26]$, where $A \mapsto 1$, $B \mapsto 2$, $\ldots$, $Z \mapsto 26$.

Example 2.4 (Example 2: f_2)
Let f_2 be an exponential function where $f_2 : \mathbb{R} \to \mathbb{R}$, where $x \mapsto e^x$.

Example 2.5 (Example 3: f_3)
Let f_3 be a modular counting function where $f_3 : \mathbb{Z} \to \{0, 1, \ldots, 9\}$, where $m \mapsto m \pmod{10}$.

§2.2 Injectivity, Surjectivity, and Bijectivity

We start off with a few more definitions to keep things organized.

Definition 2.6 (Injectivity)**.** An **injective** ("one-to-one") function is a function f where for all pairs $x_1, x_2 \in X$ where $f(x_1) = f(x_2)$, we have $x_1 = x_2$. In the same contrapositive, if $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$.

From our above example functions, we see that identity, inclusion, f_1, and f_2 are injective.

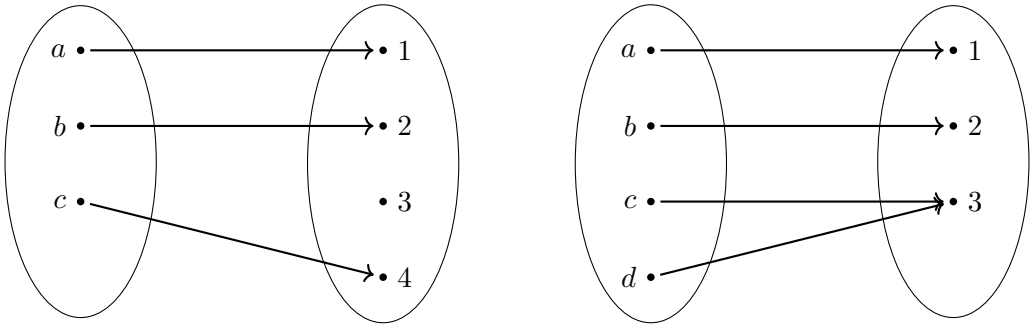


Figure 1: Injective but not Surjective, and vice versa

Definition 2.7 (Surjectivity)**.** A function $f : X \to Y$ is called **surjective** ("onto") if the range is equal to the codomain, i.e. $f(X) = \{f(x) \mid x \in X\} = Y$. In other words, $\forall y \in Y, \exists x \in X$ such that $f(x) = y$.

From our above examples, we see that id_X, f_1, and f_3 are surjective.

Definition 2.8 (Bijectivity)**.** A function is **bijective** if and only if it is *both* injective and surjective. That is, $\forall y \in Y, y = f(x)$ for a unique $x \in X$.[3]

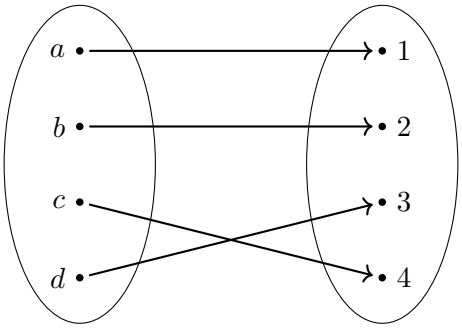


Figure 2: Bijection

By definition, this makes id_X and f_1 the only bijection functions in our examples from above.

[3] a digression from class: bijections are really useful in all sorts of situations; it's nice to notice them when they occur. an example would be that if there exists a bijection between two infinite sets, they are said to be equal in cardinality (read: size).

§2.3 Images and Pre-Images

Although not covered in class, the word *image* refers to three possible contexts, of which are listed below: given a function $f : X \to Y$, we have:

1. The image of an element $x \in X$ is $f(x)$. It is the value that we obtain when f is applied to x.
2. The image of a subset of $A \subset X$[4], denoted $f(A) = \{f(a) \mid a \in A\}$, is the set of all values we get when applying f to every element $a \in A$. It follow naturally that $f(A) \subseteq Y$ (with equality if A happens to produce a bijection $f : A \to Y$).
3. The image of a function is its range. It can also be written as $f(X)$.

Now, when we talk about pre-image, we usually first define a set $B \subseteq Y$ for a function $f : X \to Y$. The pre-image of B under f is

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\},$$

and $f^{-1}(B)$ is naturally a subset of the domain X of f. As an example, examine function f_3 from earlier. The preimage

$$f^{-1}(\{0\}) = \{\text{all multiples of } 10\},$$

as all, and only the multiples of ten, that would have a remainder of 0 when divided by 10. Other examples include

$$f^{-1}(\{1, 3\}) = \{10x + 1 \vee 10x + 3 \mid x \in \mathbb{Z}\}.$$

§2.4 Restrictions and Compositions

The restriction of a function $f : X \to Y$ with $A \subseteq X$ is, informally defined, taking a function, not altering its outputs, but changing it to only be defined on A, a subset of X. This new function would be expressed as the **restriction of f to A**, as in, $f|_A : A \to Y$ where $f|_A(x) = f(x)$ for all $x \in A$. More generally, we can also have $A \subseteq X$ and $B \subseteq Y$ with $f(A) \subseteq B$, yielding $f|_{A,B} : A \to B$. Now, a few properties on restrictions.

1. If the subset A of X happens to be equal to X, as in $A = X$, then we have $f|_A = f$.
2. Restricting a function more than one time is the same as restricting it only once. An example would be $A \subseteq B \subseteq X$, then $(f|_B)|_A = f|_A$.
3. ($\star$) The restriction of the identity function $\mathrm{id}_X : X \to X$ to a subset $A \subseteq X$ is the same as the inclusion map $\iota : A \to X$.
4. The restriction of a continuous function is continuous.

[4] i disagree with using X' for subsets, it's just way too confusing and unappealing... know that this *is* what was used in lecture today, though, as in $X' \subset X$.

For composition, we say if $f : X \to Y$ and $g : Y \to Z$, then $g \circ f : X \to Z$, where $x \mapsto g(f(x))$.

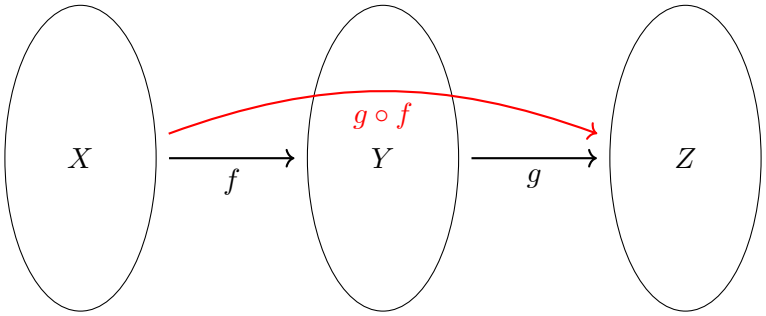


Figure 3: Function Composition

And here are the exercises from class:

Problem 2.9 (Associativity of Composition)
Suppose we have $f : X \to Y, g : Y \to Z$, and $h : Z \to W$. Show that $h \circ (g \circ f) = (h \circ g) \circ f : X \to W$.

Problem 2.10 (Invertibility and Bijectivity)
Show that f is invertible if and only if f is bijective.

Problem 2.11 (Unique Inverse)
Show that the inverse of a function f is unique if it exists. We call it $f^{-1} : Y \to X$, and that $f^{-1} \circ f = \mathrm{id}_X$ and $f \circ f^{-1} = \mathrm{id}_Y$.

Problem 2.12 (Compositional Inverses)
If $f : X \to Y$, $g : Y \to Z$ are invertible functions, show that

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

§2.5 Fields

Onto numero cinco! Fields are generalizations of $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, and are sets equipped with addition and multiplication, where they satisfy $+, \cdot : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ (closure). Fields must satisfy the field axioms, which have already been talked about (ref: see MAT157 day 1 and day 2 for **P1** to **P9**, as well as **P11** and **P12**), so I'm just going to list the food for thought below.

Problem 2.13
Show that the additive identity 0 and multiplicative identity 1 are unique and do not equal each other.

Problem 2.14

Prove the uniqueness of the additive and multiplicative inverses relative to each element $x \in \mathbb{F}$.

Problem 2.15

Show that $x \cdot 0 = 0$.

Problem 2.16

Show that $x(-y) = -(xy)$.

In reality, these have already been talked about in MAT157, so it's really just redundant. I'm going to go hug my blåhaj now, bye!

§3 Day 3: More Fields and Complex Numbers (Sep. 14, 2023)

Quick recap: fields are generalizations of $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, etc...: a set $\mathbb{F}$ equipped with operations $+, \cdot : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ along with the field axioms. Now, we introduce some of the useful notation to make our lives a little simpler:

$$\frac{x}{y} = x \cdot y^{-1}, \text{ given } x, y \in F, y \neq 0.$$
$$xy = x \cdot y.$$
$$\underbrace{x^n}_{n \geq 1} = \underbrace{x \dots x}_{n \text{ times}}$$

Problem 3.1 (Checking Cayley Tables)

Let us have a finite field with two elements $\mathbb{F}_2 = \{0, 1\}$. Verify the following tables with the field axioms,

$\mathbb{F}_2, +$	0	1
0	0	1
1	1	0

$\mathbb{F}_2, \cdot$	0	1
0	0	0
1	0	1

§3.1 Characteristic

The characteristic of field $\mathbb{F}$ is the smallest integer $n \geq 1$ such that

$$\underbrace{1 + \dots + 1}_{n \text{ summands}} = 0$$

in $\mathbb{F}$ (as in, the smallest number of times we need the additive identity to itself to obtain the multiplicative identity). If it's impossible to sum to 0, then we say the characteristic is 0 instead, such as is the case in $\mathbb{R}, \mathbb{Q}, \mathbb{C}$, and many (but not all) infinite fields. Conversely, if the characteristic is 0, we know the field has infinite elements.[5]

Problem 3.2 (Prime Characteristic)

Prove that the characteristic of a field must be either a prime number or 0.

Suppose characteristic k of field F is composite with $k = pq$; p being prime, but not necessarily for q. Then

$$\underbrace{(1 + 1 + \dots + 1)}_{p}\underbrace{(1 + 1 + \dots + 1)}_{q} = \underbrace{1 + 1 + \dots + 1}_{k} = 0$$

however, since $pq = 0$, this implies we have two different zeroes.

[5] $1, 1 + 1, 1 + 1 + 1, \dots$ cannot all be distinct if the field is finite.

§3.2 Complex Numbers

The complex numbers are a field extension of the reals, defined by $\mathbb{C} = \mathbb{R}(i) = \{a + bi \mid a, b \in \mathbb{R}\}$, where i is the root of $i^2 = -1$. $\mathbb{C}$ comes equipped with addition and multiplication $+, \cdot : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ defined as follows,

$$\begin{aligned}(a+bi)+(c+di) &= a+bi+c+di \\ &= \underbrace{(a+c)}_{\mathbb{R}} + \underbrace{(b+d)}_{\mathbb{R}} i, \\ (a+bi)\cdot(c+di) &= ac+bci+adi+bdii \\ &= \underbrace{(ac-bd)}_{\mathbb{R}} + \underbrace{(bc+ad)}_{\mathbb{R}} i.\end{aligned}$$

Remark 3.3. The reals are a subfield of the complex numbers; if $b = 0$ in $z = a + bi$, then we simply say $z \in \mathbb{R}$ for $z = a + 0i$.

Problem 3.4 ($\mathbb{C}$ is a Field)

Verify that $\mathbb{C}$ satisfies all the field axioms.

I'm not typing this out, but this should be pretty trivial to do as an exercise.

Problem 3.5

Prove that the multiplicative inverse of $a + bi$ is

$$(a+bi)^{-1} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2} i,$$

where $a + bi \neq 0$. Consider $a^2 + b^2 = (a + bi)(a - bi)$.

Now, we move onto the idea of complex conjugates. Suppose we have $z = a + bi \in \mathbb{C}$. Then, define the conjugate of z as $\overline{z} = a - bi$. The magnitude of a complex number is defined as $|z| = \sqrt{a^2 + b^2}$. These two satisfy a handful of properties: for all complex numbers z, w,

- $\overline{z + w} = \overline{z} + \overline{w}$,
- $\overline{zw} = \overline{z}\,\overline{w}$,
- $\overline{(z^{-1})} = (\overline{z})^{-1}$,
- $\overline{\left(\frac{z}{w}\right)} = \frac{\overline{z}}{\overline{w}}$,
- $|z| = \sqrt{z\overline{z}}$,
- $|zw| = |z|\,|w|$,
- $\left|z^{-1}\right| = |z|^{-1}$ if $z \neq 0$,
- $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}$, and finally,

- The triangle inequality, $|z + w| \leq |z| + |w|$.

Oh, and a joke from class: The greek character Ξ could be used in the context of complex numbers to get this absolutely disgusting fraction, $\frac{\Xi}{\Xi}$.

§4 Day 4: Fundamental Theorem of Algebra; Modular Arithmetic; Intro to Vector Spaces (Sep. 19, 2023)

Last time, we defined the complex numbers as

$$\mathbb{C} = \{a + bi \mid a, b, \in \mathbb{R}\},$$

a set endowed with the usual addition and multiplication $(+, \cdot)$. In general, when talking about fields, we don't expressly mention the usual addition and multiplication if the notation is conventional, such as in $\mathbb{R}, \mathbb{Q}, \mathbb{C}$.

§4.1 Fundamental Theorem of Algebra

Unlike in the real numbers, polynomials in the complex numbers always have roots.

Theorem 4.1

If $p(z)$ is a polynomial $a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$ for all $(a_n) \in \mathbb{C}$, there exists $z_0 \in \mathbb{C}$ such that $p(z_0) = 0$.

In other words, every non-constant, single variable polynomial with complex coefficients (which can include the reals) has at least one complex root. As a corollary, we may write $p(z)$ as

$$(z - c_1) \dots (z - c_n)$$

for $c_1, c_2, \dots, c_n \in \mathbb{C}$. Alternatively, we may say

$$\sum_{i=0}^{k} a_i z^k = c \prod_{i=1}^{k} (z - b_i)$$

for a constant $c \in \mathbb{C}$ as the unique factorization of $p(z)$.

§4.2 Modular Arithmetic

Notes by Conrad and Vogan: here, and here. The core idea of modular arithmetic is introducing a new system/architecture for arithmetic, where numbers "wrap around" after reaching a certain $m \in \mathbb{Z}$; kind of like fields! To start, consider $m = 10$ (evaluating for the unit digit).

Example 4.2

Calculate $729 \cdot 328 \pmod{10}$.

We can do this by simply multiplying the units digits of both numbers. Thus, $9 \cdot 8 \equiv 2 \pmod{10}$. □

Now, we generalize to arbitrary bases b. Let's first establish,

Lemma 4.3

For any $a, b \in \mathbb{Z}$, we say $a \equiv b \pmod{n}$ if and only if a, b has the same remainder when divided by n.

For example, $142 \equiv 2 \pmod 7$. Now, some properties (note that $\equiv$ denotes congruence)[6]:

$$\begin{aligned} a &\equiv a \pmod m. && \text{(Reflexivity)} \\ a &\equiv b \pmod m \implies b \equiv a \pmod m. && \text{(Symmetry)} \\ a &\equiv b, b \equiv c \implies a \equiv c. && \text{(Transitivity)} \end{aligned}$$

We may extend our above example with these properties as such,

$$142 \equiv 2 \equiv 9 \equiv \cdots \equiv 7k + 2, \qquad k \in \mathbb{Z}.^{7}$$

Now, for some more properties:

Theorem 4.4 (Modular Arithmetic "Closure")
If $a \equiv b \pmod m$ and $c \equiv d \pmod m$, we have the following properties,

$$\begin{aligned} a + c &\equiv b + d \pmod m. \\ ac &\equiv bd \pmod m. \end{aligned}$$

We may prove it as follows: for addition, we may rewrite:

$$a \equiv b \pmod m \implies a - b = km \text{ for some } k \in \mathbb{Z}$$

$$c \equiv d \pmod m \implies c - d = lm \text{ for some } l \in \mathbb{Z}$$

Then,

$$\begin{aligned} (a + c) - (b + d) &= (k + l)m \\ \implies a + c &\equiv b - d \pmod m. \quad \square \end{aligned}$$

For multiplication, a, similar method of proof follows; it shall be left to the reader of thy notes.

Theorem 4.5 (I don't know what to call this either)
If $a \equiv b \pmod m$ and $k \geq 1$ is an integer, then $a^k \equiv b^k \pmod m$.

This can be proved by induction. First, consider the base case where $k = 1$: we see that $a^1 = a$, same for b, and we are done. For the inductive step, given $a^k = b^k$, we may simply write

$$a^{k+1} = \underbrace{a^k \cdot a \equiv b^k \cdot b}_{\text{above property; idk if it's closure}} = b^{k+1} \pmod m \quad \square.$$

Moreover, modular arithmetic doesn't have to operate strictly in the non-negative integers: consider this example;

Problem 4.6
What is $97 \cdot 99 \pmod{105}$?

[6] technically iff but it would read differently otherwise; should be seen as left implies right, not forward/backward bs

[7] yes, negative numbers included, like -5, -12, and so on

We can evaluate this by writing $97 \equiv -8 \pmod{105}$ and $99 \equiv -6 \pmod{105}$, which evaluates out to $(-8)(-6) \equiv 48 \pmod{105}$. □
Either way, we continue onto a demonstration of uniqueness:

Theorem 4.7 (Uniqueness of Modulo)
For all $a \in \mathbb{Z}$, there exists a unique $r \in \{0, 1, 2, \ldots, m-1\}$ such that $a \equiv r \pmod{m}$.

§4.3 $\mathbb{Z}/m\mathbb{Z}$ and $\mathbb{F}_p$

The term $\mathbb{Z}/m\mathbb{Z}$ are the integers modulo m, as in, the same modular arithmetic from the previous subsection. It is defined as $\mathbb{Z}/m\mathbb{Z} = \{0, 1, \ldots, m-1\}$, representing all the remainders of the integers upon division by m exactly once. We define the following binary operators on $\mathbb{Z}/m\mathbb{Z}$,

$$a + b = c \text{ if } a + b \equiv c \pmod{m}.$$
$$ab = c \text{ if } ab \equiv c \pmod{m}.$$

Personally, I don't want to write bars above my numbers because I'm aware of what is and isn't being considered modular arithmetic. In this way, we see that our operations on $\mathbb{Z}/m\mathbb{Z}$ are well defined, according to the uniqueness of modulo. For example, in $\mathbb{Z}/2\mathbb{Z}$, we have the following Cayley tables,

$+$	0	1
0	0	1
1	1	0

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

Remark 4.8. We have a ring isomorphism

$$\mathbb{Z}/p\mathbb{Z} \cong \mathbb{F}_p$$

if p is prime. In general, $\mathbb{Z}/p\mathbb{Z}$ is a ring that counts as a field if and only if p is prime. Moreover, any field $\mathbb{F}$ with $\text{char}(\mathbb{F}) = p$ containing set $\{0, 1, 1+1, \ldots, \}$ which is a subring isomorphic to $\mathbb{Z}/p\mathbb{Z}$.

One thing important to consider, though, is that $\mathbb{Z}/m\mathbb{Z}$ is not a field unless m is prime (as mentioned in the remark). This is because $\mathbb{Z}/m\mathbb{Z}$ satisfies all field axioms aside from multiplicative inverse, unless m is a prime, then all nonzero elements have multiplicative inverses. Prototypical example:

$$\mathbb{Z}/4\mathbb{Z} = \{0, 1, 2, 3\}; \qquad 2 \cdot 2 = 0 \implies 2 \text{ has no multiplicative inverse.}$$

A more elegant way to define $\mathbb{Z}/m\mathbb{Z}$ though, uses equivalence classes: $[a]$ is an equivalence class where

$$[a]_m := \{x \in \mathbb{Z} \mid x \equiv a \pmod{m}\} = \{x \in \mathbb{Z} \mid x = mn + a, \text{ for some } n \in \mathbb{Z}\}.$$

Now, we may define $\mathbb{Z}/m\mathbb{Z}$:

$$\mathbb{Z}/m\mathbb{Z} = \{[a]_m : a \in \mathbb{Z}\}$$

To see an example, observe

$$\mathbb{Z}/3\mathbb{Z} = \{[0], [1], [2]\}$$

which encompasses all of $\mathbb{Z}$ but written in terms of 3 equivalence classes.

§4.4 Vector Spaces

A vector space (as defined in sections 1.1 to 1.2) over a field F is a non-empty set V endowed with binary operators $+ : V \times V \to V$ and $\times : F \times V \to V$ (scalar multiplication). Here are the axioms,

VS1. (Commutativity of Vector Addition) $u + v = v + u$ for all $u, v \in V$.

VS2. (Associativity of Vector Addition) $u + (v + w) = (u + v) + w$ for all $u, v, w \in V$.

VS3. (Identity of Vector Addition) There exists an element, $0 \in V$, called the zero vector, such that $v + 0 = v$ for all $v \in V$.

VS4. (Inverse Elements of Vector Addition) For every $v \in V$, there exists an element $-v \in V$, called the additive inverse of v, such that $v + (-v) = 0$.

VS5. (Identity of Scalar Multiplication) The scalar multiplication identity is 1, which is the multiplicative identity in F.

VS6. (Compatibility of Vector and Scalar Multiplication) $a(bv) = (ab)v$ for $a, b \in F$ and $v \in V$.

VS7. (Distributivity of Scalar Multiplication w.r.t. Vector Addition) $a(u + v) = au + av$ for $a \in F$ and $u, v \in V$.

VS8. (Distributivity of Scalar Multiplication w.r.t. Field Addition) $(a + b)v = av + bv$ for $a, b \in F$ and $v \in V$.

§5 Day 5: Examples of Vector Spaces (Sep. 21, 2023)

Quick reminder that the integers modulo p is a (finite) field if and only if p is a prime,

$$\mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p.$$

Recall that a vector space is motivated by the concept of $\mathbb{R}^n$ ($n \geq 1$). We say that a *vector space* over $\mathbb{F}$ (any field) is a set V endowed with the binary operations

$$\begin{aligned} + : \underbrace{V \times V}_{\{x,y|x,y\in V\}} &\to V, \qquad (x,y) \mapsto x+y \\ \cdot : \mathbb{F} \times V &\to V, \qquad (a,x) \mapsto a \cdot x. \end{aligned}$$

such that **VS1** to **VS8** hold.

Remark 5.1. The empty set, $\emptyset$, is not a vector space becuase it does not have an additive identity.

Remark 5.2. It would be more precise to write

$$+_F, \cdot_F, 0_F, 1_F \text{ for } F,$$

$$+_V, \cdot_V, 0_V \text{ for } V,$$

but this is basically overly pednatic and would make everything annoyingly long. For the sake of demonstration, we shall proceed with these protracted symbols. Using this notation, we can rewrite some of the vector space axioms,

$$(a \cdot_F b) \cdot_V X = (a \cdot_V b) \cdot_V X, \tag{**VS6**}$$

$$1_F \cdot_V x = x. \tag{**VS5**}$$

We now proceed to giving a handful of examples of vector spaces.

1. V = {0} forms the trivial vector space, which only contains the zero vector.
2. A coordinate space can be expressed as[8]

$$F^n = \underbrace{F \times \ldots F}_{n \text{ times}} = \{(a_1, \ldots, a_n) \mid a_i \in F \forall i\}.$$

It is equipped with binary operators $+, \cdot$; for example, we may write

$$\begin{aligned} (a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) &= \underbrace{(a_1 + b_1, \ldots, a_n + b_n)}_{\text{Component by component}} \\ 1_V \cdot (a_1, \ldots, a_n) &= (1_F a_1, \ldots, 1_F a_n) && (\textbf{VS5}) \\ &= (a_1, \ldots, a_n). && (\textbf{F3}) \end{aligned}$$

Visually, we may write $(a_1, \ldots, a_n) \in F^n$ as a column vector $\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$.

[8] F is commonly the reals, but can also be complex, or whatever. any arbitrary field will do

3. Matrices are also vector spaces; though, we should say, the set of all matrices of a fixed size, with entries in F, forms a vector space. Let $m, n \geq 1$ with

$$\begin{aligned} M_{m\times n}(F) &= \{m \times n \text{ matrices over } F\} \\ &= \left\{ \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \right\} \\ &= F^{mn} \end{aligned}$$

   Note that a_ij is to be read as the entry on the ith row, and jth column. Matrices are also endowed with the usual addition and multiplication, i.e.

$$\underbrace{(A+B)_{ij}}_{\text{matrices}} = \underbrace{a_{ij} +_F b_{ij}}_{\text{entries}}.$$

   someone confirm for me, not sure about this one.

4. Finally, we have function spaces. Let $\mathcal{F}(S, F) = \{\text{functions } f : S \to F\}$. Given that S is a non-empty arbitrary set and V a vector space over the field F, the functions $S \to F$ can be given the structure of a vector space over S, with the usual binary operators equipped. For example,

$$\begin{aligned} (f+g)(s) &= f(s) +_F g(s), \\ (c \cdot f)(s) &= c \cdot_F f(s), \end{aligned}$$

   for all functions f, g, and $s \in S$, $c \in F$.

   Intuition: the set of functions from any set S into a vector space has a natural vector space structure, given by addition and scalar multiplication.

§6 Tutorial 1: Modular Arithmetic and Fermat's Little Theorem (Sep. 21, 2023)

Since I'm assuming tutorials are different for everyone, I'll just go over the problems here. My proofs tend to be more laconic because there are 11 questions in total. Let me know if there are any questions!

1. Find the last digit of
$$(((7^6)^5)^4.$$

   Notice that we are working mod 10, and that 7 cycles as follows,
$$\begin{aligned} 7 &= 7 \\ 7^2 &\equiv 9 \pmod{10} \\ 7^3 &\equiv 3 \pmod{10} \\ 7^4 &\equiv 1 \pmod{10} \end{aligned}$$
   As a result, we may rewrite the exponentiation to obtain
$$7^{6\cdot 5\cdot 4}$$
   which indeed has an exponent of a multiple of 4. Thus, the last digit is 1.

---

2. Find the last digit of
$$7^{6^{5^4}}.$$

   Notice that $6^{5^4} \equiv 0 \pmod 4$. This means the last digit is simply 1 again.

---

3. Prove that $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is a prime.

   Remember that a consequence of the field axioms is that $ab = 0$ if and only if one of a, b is equal to 0. Moreover, by Bezout's Lemma, given $a \in \mathbb{Z}/n\mathbb{Z}$, if n is prime, then (a, n) coprime and we may find integers b, c such that
$$ba + cn = 1.$$
   Thus, $ba = 1$ and we see the forward implication that if n is prime, we have a field. Otherwise, suppose n is not prime. We then may factor it as $n = ab$, where $ab = 0$. This contradicts the field axioms, and so we are done. □

---

4. Let p be a prime. Prove that
$$a, b \in \mathbb{F}_p \implies (a+b)^p \implies a^p + b^p.$$

   Proceed with the binomial theorem,
$$(a+b)^p = \sum_{i=0}^{p} \binom{p}{i} a^i b^{p-i}.$$
   Note that $\binom{p}{i} \equiv 0 \pmod p$ for all $0 < i < p$, which means the only terms we have left are $a^p + b^p$. □

5. Find the multiplicative inverse of every non-zero element in $\mathbb{F}_7$.

$$\begin{aligned} 1^{-1} &= 1 \\ 2^{-1} &= 4 \\ 3^{-1} &= 5 \\ 4^{-1} &= 2 \\ 5^{-1} &= 3 \\ 6^{-1} &= 6. \end{aligned}$$

---

6. Solve $6x = 3$ in $\mathbb{F}_{11}$, and then in $\mathbb{Z}/8\mathbb{Z}$.

   For $\mathbb{F}_1 1$, $x = 6$. For $\mathbb{Z}/8\mathbb{Z}$, I don't think a solution exists.

---

7. Use the Euclidean Algorithm to find the greatest common divisor of 24 and 136.

$$\begin{aligned} 136 &= 24 \cdot 5 + 16 \\ 24 &= 1 \cdot 16 + 8 \\ 16 &= 2 \cdot 8. \end{aligned}$$

   Therefore, $\gcd(24, 136) = 8$.

---

8. Prove that the set of complex numbers is a field.

   See **MAT157** week 1 homework question 3(a)...

---

9. Let $0 \neq a \in \mathbb{F}_p$. Prove that

$$a^{p-1} = 1$$

   in $\mathbb{F}_p$ (Note that this is Fermat's Little Theorem).

   We may do this by induction. First, rewrite the expression to $a^p = a \implies a^p - a = 0$. It is quick to prove the base case of $a = 1$. For the inductive step, suppose the expression holds for some $a \in \mathbb{F}_p$. Then,

$$(a+1)^p = a^p + \binom{p}{1}a^{p-1} + \cdots + \binom{p}{p-1}a + 1 \quad \text{(Binomial Theorem)}$$

$$(a+1)^p - a^p - 1 = \underbrace{\binom{p}{1}a^{p-1} + \cdots + \binom{p}{p-1}a}_{\text{vanishes}}$$

   which then we may insert into our inductive step, yielding

$$((a+1)^p - a^p - 1) + (a^p - a) = (a+1)^p - (a+1),$$

   and so we are done.

---

10. Let F be a field, and let the set $V = \{(a, b) \mid a, b \in F\}$ with component-wise addition and scalar multiplication $c(a, b) = (ca, cb)$. Show that V is a vector space over F.

    This is just re-verifying **VS1** to **VS8**. Will not be written out here for brevity reasons, but should be treated as an exercise in checking the axioms.

11. Prove that the set of all functions from $\mathbb{R} \to \mathbb{R}$ with point-wise addition and scalar multiplication (as in exercise 3 in section 1.2 from the textbook) is a vector space over $\mathbb{R}$. Which of the following subsets of the above set form a subspace? The set of even functions, the set of odd functions, the set of continuous functions, the set of differentiable functions, the set of functions such that $f(1) = 0$, the set of functions such that $f(0) = 1$?

    This is basically the function space defined in notes from **Day 5**.

    - The set of even functions form a subspace. Let f, g be even and c a real.

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

    and

    $$\begin{aligned}(cf)(-x) &= c(f(-x)) \\ &= c(f(x)) \\ &= (cf)(x).\end{aligned}$$

    which shows it indeed does form a subspace. □

    - The same goes for the set of odd functions; the proof is similar to the above for even functions.
    - The set of continuous functions form a subspace. Let f, g be continuous functions. Then, $f + g$ is also continuous, since f nor g have points of discontinuity. Multiplication by scalar is a vertical stretch or compression that does not create discontinuities either.
    - The set of all functions such that $f(1) = 0$ form a subspace. Let f, g be functions such that $f(1) = g(1) = 0$. Then,

    $$(f+g)(1) = 0 + 0 = 0$$

    Moreover, for any real c, we also have $(cf)(1) = c(f(0)) = c \cdot 0 = 0$.

    - The set of all functions such that $f(0) = 1$ does not form a subspace. It is not closed under addition nor multiplication, which can be demonstrated by picking any two arbitrary functions f, g such that $f(0) = g(0) = 1$. Then,

    $$(f+g)(0) = 1 + 1 = 2,$$

    which shows function $f + g$ is not in our set.

§7 Day 6: Review of Vector Space Examples; Polynomials; Subspaces (Sep. 26, 2023)

Some review today! A vector space is a set V endowed with operations $+ : V \times V \to V$ and $\cdot : F \times V \to V$, where F may be any field; this set also further satisfies the 8 vector space axioms **VS1** to **VS8**.

§7.1 Additional Examples of Vector Spaces (and a revision to Function Space)

1. A function space $\mathcal{F}(S, F)$ is the set of all functions $f : S \to F$ where S is some set, and F is a field. Define the usual addition and scalar multiplication as follows, for $f, g \in \mathcal{F}(S, F)$ and $a \in F$,

$$\begin{aligned} f + g : S &\to F \\ s &\mapsto f(s) +_F g(s) \in F, \\ a \cdot f : S &\to F \\ s &\mapsto a \cdot_F f(s) \in F \end{aligned}$$

   where s is an element of set S. Quick digression,

   > **Remark 7.1.** If we have $S = \{1\}$, then $\mathcal{F}(S, F)$ is said to be "F", since this simply induces a simple map to F.
   > If $S = \{1, 2, \dots, n\}$, then $\mathcal{F}(S, F)$ is said to be "F^n",
   >
   > $$\begin{aligned} 1 &\mapsto a_1 \\ &\vdots \\ n &\mapsto a_n \end{aligned}$$
   >
   > where (a_n) should be treated as a sequence of outputs; for example, we may view it as a variant of *coordinate space*, i.e. an n-vector can be thought of a function with domain $\{1, 2, \dots, n\}$; this is equivalent to seeing it as an ordered list of n numbers, hence giving rise to the idea of sequence (a_n),
   >
   > $$\mathbb{R}^n := \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \middle| \, a_1, \dots, a_n \in \mathbb{R} \right\} = \{a : \{1, \dots, n\} \to \mathbb{R}\}.$$
   >
   > If the set S has a natural ordering, we may write the components in the sequence in order for convenience. If S is an infinite set $\{1, 2, \dots\}$, we may see ourselves at a situation with the *sequence space*; example 1.5 in §1.2 of the textbook.

2. Polynomials also form an example vector space. But, before we may define the polynomial vector space, we should formally define what a polynomial is. A polynomial is a formal expression

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

   where $n \geq 0$, $a_i \in F \forall i$. In the polynomial space $P(F)$ of polynomials over F, x is not necessarily in F (in the context of "$P(x)$"); only the elements $a_0, \dots, a_n$ are. Moreover, we often omit terms of the form $0a^k$ out of convenience; these are not counted when determining the degree of a polynomial. The leading coefficient is the largest n such that a_n is nonzero (of the term $a_n x^n$); we say the degree of P is n.

> **Remark 7.2.** In the case of $P(x) = 0$, we conventionally define it as having negative infinite degree.

Two polynomials are equal if and only if they share the same coefficients, which brings us to this example:

Example 7.3

Suppose $F = \mathbb{F}_2$; then let $P(x) = x^2 + x$. Indeed, this is *not* the zero polynomial, despite

$$\begin{aligned} P : 0 &\mapsto 0 \\ 1 &\mapsto 0. \end{aligned}$$

It is important to note that while polynomials may be thought of as functions intuitively, they are not the same thing; a polynomial is a formal linear combination over F of monomials, it is not to be formally interpreted as a map $F \to F$.

In a vector space, we define the polynomials to have the following operations; given polynomials $P, Q \in P(F)$ with coefficients $(a_n), (b_n) \in F$ and a scalar $c \in F$, then

$$P(x) + Q(x) = (a_n + b_n)x^n + \cdots + (a_0 + b_0), c \cdot P(x) \quad = (c \cdot a_n)x^n + \cdots + (c \cdot a_0).$$

where we see things go component by component as usual.

§7.2 Subspaces

Let's start with the formal definition, since this one is fairly intuitive; a subset W of a vector space V is called a *subspace* if W together with the restriction of vector space operations of V is a vector space. Let's start with the operations: start with the operations on space V,

$$\begin{aligned} +_V &: V \times V \to V \\ \cdot_V &: F \times V \to V \end{aligned}$$

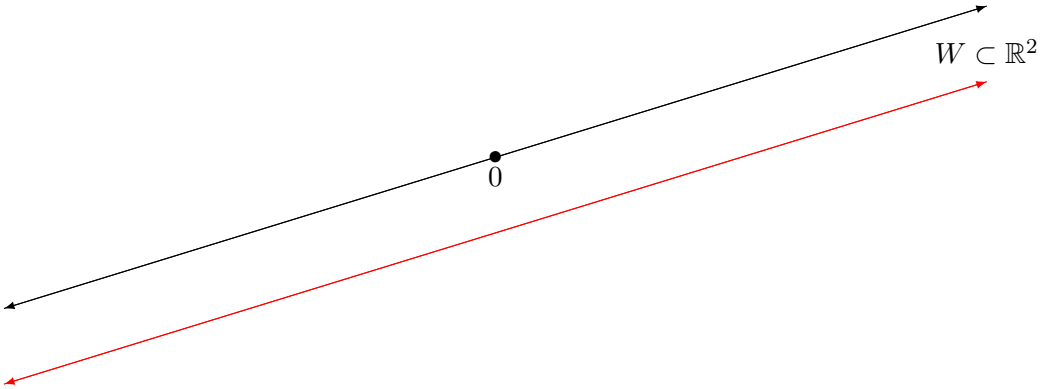
we define restrictions on the above to obtain the operations for W (this is done to ensure closure),

$$\begin{aligned} +_V\big|_{W \times W} &= +_W \\ \cdot_V\big|_{F \times W} &= \cdot_W \end{aligned}$$

From these definitions, we see $x, y \in W \implies x +_V y \in W$ which, by restriction of $+_V$, also means $x +_V y = x +_W y \in W$. The same goes with scalar multiplication:

$$x \in V, a \in F \implies a \cdot_V x = a \cdot_W x \in W.$$

Now, we move onto some examples of subspaces: for example, given vector space $\mathbb{R}^2$, (Terrible diagram, I'm sorry); either way, the point stands: suppose W is the line in $\mathbb{R}^2$;



it is a set containing all points on said line. This is indeed a subspace! However, the red line is not. Let's demonstrate a helpful theorem:

Theorem 7.4 (Subspace Criteria)

W is a subspace of vector space V if and only if

- It is additively closed; $x, y \in W \implies x + y \in W$.
- It is multiplicatively closed; $x \in W, a \in F \implies a \cdot x \in W$.
- It contains the additive identity 0.

Now, we prove this theorem: for its forward implication, note that additive and multiplicative closure comes directly from the definition of these operators; for the inclusion of 0, suppose we take any $x \in W$. Then

$$0_V \cdot x = 0_V \in W$$

by definition. For the converse, assuming the above, we may construct the restrictions easily. I'm leaving it as an exercise to construct these proofs in further detail; see question 11 on the first tutorial homework set for practice.

§8 Day 7: Matrix Transpose, Trace; Intersection and Direct Sum of Subspaces (Sep. 28, 2023)

Today we went over some review of vector spaces, as well as the definition of some matrix operations and subspace properties. Remember that a subset W of a vector space V is a subspace if it contains 0_V and is closed under the restriction of the operations of V. Here are a few examples,

Example 8.1

The set of all functions that vanish at a given s,

$$\{f \in \mathcal{F}(S, F) \mid f(s) = 0\},$$

is a subspace of $\mathcal{F}(S, F)$.

Example 8.2

In $\mathcal{F}(\mathbb{R}, \mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R}\}$,

- $\mathcal{C}(\mathbb{R})$ (the set of all continuous f)
- $\{f(x) = -f(x) \mid f \in \mathcal{F}(\mathbb{R}, \mathbb{R})\}$ (odd functions)
- $\{f(x) = f(-x) \mid f \in \mathcal{F}(\mathbb{R}, \mathbb{R})\}$ (even functions)

are all examples of subspaces.

Remark 8.3. If U is a subspace of W and W is a subspace of V, then U is a subspace of V. The proof is left as an exercise, but note that (wlog the field is F)

$$\left(+_V|_{W\times W}\right)\Big|_{U\times U} = +_V|_{U\times U}, \qquad \left(\cdot_V|_{F\times W}\right)\Big|_{F\times U} = \cdot_V|_{F\times U}$$

§8.1 Matrix Transpose and Trace

Given a matrix $A \in M_{m\times n}(F)$, its transpose A^T is defined by $A^T_{ij} = A_{ji}$, for example,

$$A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}, \qquad A^T = \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix}.$$

Moreover, we say A is symmetric if $A = A^T$ (and so $m = n$; i.e. A is a square matrix), and antisymmetric if $-A = A^T$.

Problem 8.4

Prove that the set of symmetric $n \times n$ matrices is a subspace of $M_{n\times n}(F)$.

More notation now! If A is an $n \times n$ square matrix, then we write $\mathrm{Tr}(A)$ as the "sum of all entries on the diagonal", i.e.

$$\mathrm{Tr}(A) = A_{11} + A_{22} + \cdots + A_{nn} = \sum_{i=1}^{n} A_{ii}.$$

Problem 8.5

Show that the set of all matrices with trace 0 form a subspace in $M_{n\times n}(F)$.

This is fast to prove; try it! A quick digression:

Remark 8.6. The trace of a matrix is generally a pretty important notion since:

- It is the sum of the eigenvalues.
- It is invariant by transposition, and basis change, and cyclic permutation (invariants are really nice... see: determinant!).

(I think we're gonna cover these later on, but it's just a quick motivation for people going "why should I care"

§8.2 Intersection and Direct Sum of Subspaces

Here are two important theorems about subspaces:

Theorem 8.7 (Intersections of Subspaces is Subspace)

The intersection of any collection of subspaces of a vector space V is a subspace of V. More specifically, let us have subspaces $W_1, W_2, \cdots : W_i \subset V$ with $i \in I$, then

$$\bigcap_{i\in I} W_i = \{x \mid x \in W_i \ \forall i \in I\}$$

is a subspace.

We may quickly show this by satisfying the subspace test: let $\bigcap_{i\in I} W_i = \mathcal{W}$,

- Subspaces must contain 0. Therefore, $0 \in \mathcal{W}$.
- For any $u, v \in \mathcal{W}$, since all of the individual subspaces are closed under addition, we see that $u + v \in W_1, u + v \in W_2, \cdots \implies u + v \in \mathcal{W}$.
- The same argument holds for scalar multiplication.

Now, consider a different variant of the above:

Theorem 8.8 (Sum of Subspaces is Subspace)

The *sum* of two subspaces $U, W \subset V$ is defined as

$$U + W = \{u + w \mid u \in U, w \in W\}.$$

Moreover, $U + W$ is a subspace.

This may be generalized to a sum of an arbitrary set of subspaces, but we may see that this is indeed a subspace by considering,

- 0 is once again in the sum ($0_U + 0_W = 0_V$).

- Let $u \in U$, $w \in W$, and $x \in U \cap W$. Then

$$u + w = \underbrace{(u + x)}_{\in U} + \underbrace{(u - x)}_{\in W},$$

and a similar argument can be made for scalar multiplication.

Remark 8.9 (Direct Sum)**.** We may denote $U+W$ as $U \oplus W$ (direct sum) if the representation of all $v \in V$ as $u + w$ is unique; i.e., $U \cap W = \{0\}$. More specifically, there is one and only one way to write each element of V as the sum of some elements of U and W.

§9 Tutorial 2: Tut. 2 HW Solutions (i guess) (Sep. 28, 2023)

Laconic solutions to Week 2 tutorial homework! (because I don't really wanna type everything out step by step, if you follow lectures you should get it)

1. Is the set
$$\{(3x, 5x, 2x) \mid x \in \mathbb{R}\}$$
a vector space over the reals?

   Notice that $A = \{3x \mid x \in \mathbb{R}\}, B = \{5x \mid x \in \mathbb{R}\}, C = \{2x \mid x \in \mathbb{R}\}$ are individually subspaces of $\mathbb{R}$. Therefore, $A \times B \times C$ is also a vector space.

---

2. Let c be a real parameter. Find all values of c so that the set
$$X = \{(3x + c, 5x, 2x) \mid x \in \mathbb{R}\}$$
with coordinate-wise vector addition and scalar multiplication is a vector space over the reals.

   We must have $c = 0$ otherwise $(0, 0, 0) \notin X$.

---

3. Let $n \in \mathbb{N}$ and let $a_i \in F^n$, $1 \leq i \leq n$, where F is a field. Prove that
$$\left\{\sum_{i=1}^{n} a_i v_i \mid v \in F^n\right\}$$
with coordinate-wise vector addition and scalar multiplication is a subspace of F^n. Here, v_i is the ith entry of the column vector v.

   We see that this is true; each component of the sum $a_1v_1 + a_2v_2 + \cdots + a_nv_n$ forms a subspace on its own (since $v_i \in F$) and the scalar multiple of any subspace is still a subspace.

---

4. Let $n \in \mathbb{N}$ and let $a_i \in F^n$, $1 \leq i \leq n$, where F is a field. Prove that
$$\left\{v \in F^n \mid \sum_{i=1}^{n} a_i v_i = 0\right\}$$
with coordinate-wise vector addition and scalar multiplication is a subspace of F^n. Here, v_i is the ith entry of the column vector v.

   Passes the subspace test; contains $(0, 0, \ldots, 0)$, and addition and scalar multiplication are trivially closed.

---

5. Let $n \in \mathbb{N}$ and let $a_i \in F^n$, $1 \leq i \leq n$, where F is a field. Find all c, so that
$$\left\{v \in F^n \mid \sum_{i=1}^{n} a_i v_i = c\right\}$$
with coordinate-wise vector addition and scalar multiplication is a subspace of F^n. Here, v_i is the ith entry of the column vector v.

   This is only a subspace if $c = 0$. Otherwise we don't have the zero vector.

6. Let F be a field. Explain why it is also a vector space over itself. That is, describe the underlying vector space with its operation, and the field over which it is a vector space.

   Addition is the usual addition, and scalar multiplication over F is the same as ordinary multiplication, both of which are closed. Also $0 \in F$ by default.

---

7. A matrix (with entries in a field) is symmetric if and only if $A^T = A$, and is anti-symmetric if $A^T = -A$. Prove that any matrix can be represented as a sum of a symmetric and an anti-symmetric matrix. Is this representation unique? Here the transpose of a matrix is defined by $(A^T)_{ij} = A_{ji}$.

   Let P be such a symmetric matrix and Q an anti-symmetric matrix. Then we have, for all $A \in \mathcal{M}_n(F)$,

$$\begin{aligned} A &= P + Q \\ &= \underbrace{\frac{(A + A^t)}{2}}_{P} + \underbrace{\frac{(A - A^t)}{2}}_{Q}. \end{aligned}$$

   This representation is necessarily unique because the main diagonal of Q is 0.

---

8. Let V_n be a subspace of a vector space V (over a field) for all $n \geq 1$. Is

$$\bigcup_{n \geq 1} \bigcap_{k \geq n} V_k$$

   a subspace of V? Prove or disprove it. Hint: Define $W_n = \bigcap_{k \geq n} V_k$ and show that W_n is an increasing sequence of subspaces.

   A union of two vector spaces is a vector space only if one contains the other. Forward implication is obvious, now for converse; let us have subspaces P, Q of V where $P \not\subset Q$ and $Q \not\subset P$. Take $a \in P$ and $b \in Q$; $a + b \in P \cup Q$, so $a + b \in P$ or Q. If $a + b \in P$, then $a + b + (-a) \in P$, and if $a + b \in Q$, then $a + b + (-b) \in Q$. Both are contradictions, and so we are done. Clearly $W_1 \subset W_2 \subset \ldots$, and so we are done.

---

9. Find all possible real solutions to the system

$$\begin{aligned} x + 2y + 2z &= 1 \\ -2x - y + 2z &= 1 \\ -3x - 2y + 3x &= 1. \end{aligned}$$

   $(x, y, z) = \left(\frac{1}{2}, -\frac{1}{2}, \frac{3}{4}\right)$.

---

10. Find all possible real solutions to the system

$$\begin{aligned} 2x + 2y &= 0 \\ -x - y &= 0. \end{aligned}$$

    $\{(x, y) \mid x = -y\}$.

11. Find all possible real solutions to the system

$$\begin{aligned} 2x + 2y &= 1 \\ -x - y &= 1. \end{aligned}$$

No solutions exist.

---

12. Let

$$(*) \qquad \begin{cases} 2x + 2y = a \\ -x - y = b \end{cases}$$

be a real system. Show that

$$\{(a, b) \in \mathbb{R}^2 \mid (*) \text{ has a solution}\}$$

is a subspace of $\mathbb{R}^2$.

Let (x_1, y_1) be a solution for (a, b) and (x_2, y_2) be a solution for (c, d). Then $(a + c, b + d)$ has solution $(x_1 + x_2, y_1 + y_2)$. Zero and scalar multiplication is quick to verify.

§10 Day 8: Linear Combinations and Linear Independence (Oct. 3, 2023)

Recap from last class:

- The sum (or direct sum, if unique) of subspaces $W_1, W_2, \ldots, W_n \subset V$ can be written as
$$W_1 + W_2 + \cdots + W_n = \{x_1 + x_2 + \cdots + x_n \mid x_i \in W_i,$$

> **Remark 10.1.** In the case of an infinite sum of subspaces, we can construct each element using finite sums of subspaces; this isn't covered in the class, just a note from the board.

- The intersection of a set of subspaces may be written as
$$\bigcap_{\forall i} W_i = \{x \mid x \in W_i \ \forall i\};$$
This is always a subspace.
- The union of a two subspaces is a subspace is a subspace if and only if one contains the other; this logic may be extended to the union of more than two at a time.

§10.1 Linear Combinations (Section 1.4)

We start with the motivation behind linear combinations:

- In general, to parametricize a plane in $\mathbb{R}^n$, we may write
$$P + ax + by$$
with $a, b \in \mathbb{R}$ and point $P \in \mathbb{R}^n$ (note that x, y are to be non-parallel vectors)
- We can solve systems of equations with them.[9]

Example 10.2

Find the solutions of the system,
$$\begin{cases} 2x + 3y + z = 2 \\ 3x + 4y + 7z = 4. \end{cases}$$

Note that we can convert this system into a linear combination, as in,
$$\begin{cases} 2x + 3y + z = 2 \\ 3x + 4y + 7z = 4 \end{cases} \iff \begin{pmatrix} 2x + 3y + z \\ 3x + 4y + 7z \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \qquad \text{(Expression in } \mathbb{R}^2\text{)}$$
$$\iff x\begin{pmatrix} 2 \\ 3 \end{pmatrix} + y\begin{pmatrix} 3 \\ 4 \end{pmatrix} + z\begin{pmatrix} 1 \\ 7 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix},$$
which is a linear combination in terms of the above column vectors. As an exercise, check that $(x, y, z) = (4 - 17s, -2 + 11s, s)$ is a valid solution for all $s \in F$.

[9] inb4 prof. herzig starts writing the most convoluted elementary procedure to solve this

In general, given $x_1, \dots, x_n$ in a vector space V, we define a linear combination of $x_1, \dots, x_n$ to be $a_1x_1 + \cdots + a_nx_n$ (any vector) where all such coefficients $a_i \in F$. With the linear combination defined, we can further define the notion of span: let $S \subset V$; then $\text{span}(S)$ is the set of all linear combinations formed from elements of S; i.e.,

$$\{a_1x_1 + \cdots + a_nx_n \mid a_i \in F, x_i \in S \ \forall i; n \geq 1\}$$

Moreover, the span of the empty set is $\{0\}$; as another example, if $V = \mathbb{R}^2$ and $S = \{\binom{1}{0}, \binom{0}{1}\}$, then $\text{span}(S) = \mathbb{R}^2$. Though, if we let $S = \{\binom{1}{2}, \binom{2}{4}\}$, then $\text{span}\, S = \{\binom{a}{2a} \mid a \in \mathbb{R}\}$. For a different idea, let V be the polynomial space $P(F)$ and S the subset $\{1, x^2, x^4, \dots, \}$. Then

$$\text{span}(S) = \{a_0 + a_1x^2 + a_2x^4 + \cdots + a_nx^{2n} \mid a_i \in F, n \geq 0\}$$

which can be thought of as a "polynomial in x^2." With some examples in place, we may move on to define some properties regarding the span:

Lemma 10.3

Given that S is a subset of vector space V,

(a) $\text{span}(S)$ is always a subspace of V.

(b) $S \subset \text{span}(S)$.

(c) If $S_1 \subset S_2$ are subsets, then $\text{span}(S_1) \subset \text{span}(S_2)$ are subspaces.

We now prove these statements: suppose $x, y \in \text{span}(S)$. Then let $x = a_1x_1 + \cdots + a_nx_n$ for $a_i \in F, x_i \in S$, and $n \geq 1$, and $y = b_1y_1 + \cdots + b_mny_m$ for $b_i \in F$, $y_i \in S$, and $m \geq 1$. We may write

$$x + y = (a_1x_1 + \cdots + a_nx_n) + (b_1y_1 + \cdots + b_my_m),$$

which is indeed a linear combination, and so is in the span of S. We can also check that given the same $x \in \text{span}(S)$ (as defined above) and $a \in F$, then

$$a \cdot x = a(a_1x_1 + \cdots + a_nx_n) = (aa_1)x_1 + \cdots + (aa_n)x_n \in \text{span}(S).$$

It remains to check that $0 \in \text{span}(S)$; if $S = \emptyset$, then it is done by default; otherwise, take $0_F \cdot x = 0 \in \text{span}(S)$. We also see $x \in S \implies x = 1 \cdot x \in \text{span}(S)$, and thus $S \subset \text{span}(S)$. The same argument applies to the third part of this lemma. □

Moreover, we also say S *spans* (or *generates*) the subspace $\text{span}(S)$; if $S \subset V$, then $\text{span}(S)$ is the smallest subspace of V containing S. That is, for all subspaces $W \subset V$ where $S \subseteq W$, we have $\text{span}(S) \subseteq W$. Since all $x_i \in S$ are also elements of W and W is a subspace, all linear combinations

$$a_1x_1 + \cdots + a_nx_n \in W.$$

Now, since $\text{span}(S)$ contains all such linear combinations, we see that it is necessarily a subset of W.

Remark 10.4. From this, we see that all subspaces are closed under linear combinations, and that if any $W \subset V$ if a subspace, then $\text{span}(W) = W$. Moreover, we can express the span of any subset $S \subset V$ as the intersection of all subspaces W containing S.

§10.2 Linear Independence (Section 1.5)

Let use have $x_1, \ldots, x_n \in V$ (distinct choices of x_i; $x_i \neq x_j$ for $i \neq j$). If we can express some $x \in V$ is a linear combination of $x_1, \ldots, x_n$ in two different ways,

$$\begin{cases} x = a_1x_1 + \cdots + a_nx_n \\ x = b_1x_1 + \cdots + b_nx_n \end{cases}$$

such that $(a_1, \ldots, a_n) \neq (b_n, \ldots, b_n) \implies 0 = (a_1 - b_1)x_1 + \cdots + (a_n - b_n)x_n$, we say $x_1, \ldots, x_n$ is linearly dependent in V. In short, if $(a_1 - b_1, \ldots, a_n - b_n) \neq (0, \ldots, 0)$ then our choice of $x_1, \ldots, x_n$ is linearly dependent.

Example 10.5 (Linearly Dependent Example)

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\}$$

is linearly dependent since $2\binom{1}{0} + \binom{0}{1} + \binom{2}{-1} = 0$.

On the other hand, $\{\binom{0}{1}, \binom{1}{0}\}$ is linearly independent.

§11 Day 9: Linear Independence and Dependence (Oct. 5, 2023)

Last time, we went over spanning set and linear independence: given a subset $S \subset V$, we write

$$\begin{aligned}\text{span}(S) &= \{\text{all linear combinations of } S\} \\ &= \{a_1x_1 + \cdots + a_nx_n \mid a_i \in F, x_i \in S \ \forall i\}.\end{aligned}$$

Note that $\text{span}(S)$ has the property that it is a subspace containing S, and the smallest one to do so. If there was another subspace $W \supset S$, then $\text{span}(S) \subset W$. Moreover,

$$S_1 \subset S_2 \implies \text{span}(S_1) \subset \text{span}(S_2),$$

and if $\text{span}(S_1)$ happens to be the whole vector space V, then $\text{span}(S_2) = V$ as well.

§11.1 Conditions for Linear Independence / Dependence

Remember that a subset S is linearly dependent in V if we may write

$$0 = a_ix_i + \ldots a_nx_n$$

for some $x_i \in S$ (distinct), such that $a_i \in F$ are not all zero; if we cannot write 0 in such a manner, we say S is linearly independent in V.[10] Take a few examples:

- Let us have set $S = \{1, x, x^2, \ldots, x^n\} \subset P(F)$. This set is linearly independent, since $a_01 + a_1x + \cdots + a_nx^n = 0$ in $P(F)$ if and only if $a_0, \ldots, a_n = 0$, by polynomial equivalence (two polys are equal if and only if they have the same coefficients).
- If the definition for linear independence did not require $x_i \in S$ to be distinct, we may pick $x_1 = x_2 \in S$ such that $1x_1 + (-1)x_2 = 0$, which makes all sets linearly dependent, which is absurd.
- For all $x \in V$, the set $\{x\}$ is linearly independent as long as $x \neq 0$.
- If $0 \in S$, then S is linearly dependent since $1 \cdot 0 = 0$.
- $\{x, y\}$ is linearly dependent if and only if one is a multiple of another (i.e. there exists some $a \in F$ where $ax = y$).

With these down, we now continue onto the main lemmas proved in lecture:

Lemma 11.1 (Implication of Linear Dependence)

Suppose we have subsets $S_1 \subset S_2 \subset V$. If S_1 is linearly dependent, then S_2 is linearly dependent. Furthermore, if S_2 is linearly independent, then S_1 is linearly independent.[a]

[a] leave it as an exercise to check, ig. it's very quick

Theorem 11.2 (Element in Span of Others)

Suppose we have a linearly independent subset $S \subset V$. Pick an $x \in V \setminus S$. Then $S \cup \{x\}$ is linearly dependent if and only if $x \in \text{span}(S)$.

[10] note that we don't actually need all of S in the sum; we could do that and set all the extraneous stuff to a coefficient of 0 though.

Before we prove this theorem, remember that it's not generally true that for some linearly independent set, each vector is in the span of others; we only need *one* such vector. Example: $\{\binom{1}{0}, \binom{2}{0}, \binom{0}{1}\}$. Anyways, onto the proof: Starting with the forward implication: if $S \cup \{x\}$ is linearly dependent (while S itself is not), then

$$a_1x_1 + \cdots + a_nx_n + ax = 0$$

for some distinct choices $x_i \in S$, and $a_i, a \in F$. If $a = 0$, then $a_1x_1 + \cdots + a_nx_n = 0$ which would be a contradiction as S itself is linearly independent; thus, $a \neq 0$, and we solve for x:

$$x = -a^{-1}a_nx_1 - \cdots - a^{-1}a_nx_n$$

which is indeed a linear combination of S, and thus $x \in \text{span}(S)$. Now for the other way around! If $x \in \text{span}(S)$, then we may write $x = b_1x_1 + \cdots + b_nx_n$ for some $b_i \in F$, $x_i \in S$, and $n \geq 1$. This means

$$0 = -1x + b_1x_1 + \cdots + b_nx_n = 0,$$

which means we've found a linear combination equaling zero, and thus $S \cup \{x\}$ is linearly dependent.

Remark 11.3. $S \cup \{x\}$ is linearly dependent implies that $\text{span}(S \cup \{x\}) = \text{span}(S)$.

§11.2 Bases

We were low on time when this came around (like 5 minutes from end of class) but a *basis* of V is a linearly independent subset that spans V. Also, no it's not necessarily unique. Example: let $V = \mathbb{R}^2$. Then $\{(0,1),(1,0)\}$ is a basis, but so is $\{(1,1),(0,1)\}$, and so is something ridiculous like $\{(100,0),(0,\pi)\}$.

§12 Tutorial 3: Tut. 3 HW Solutions (Oct. 5, 2023)

Laconic solution part trio

1. Is $\{(1,2,3),(3,4,5),(6,7,8)\}$ a linearly dependent subset of $\mathbb{R}^3$?

   Yes. All vectors in the set are of the form $(1,2,3)+a(1,1,1)$ where $(1,1,1) = \frac{1}{3}(6,7,8) - \frac{1}{3}(3,4,5)$.

---

2. Let V be a vector space and let $S \subset V$ be a linearly independent subset. Show that $S_1 \subset S$ implies that S_1 is a linearly independent subset.

   None of the elements in S are in the span of the others. Therefore any subset $S_1 \subset S$ is necessary linearly independent too, since we are picking vectors from S.

---

3. Let V be a vector space and let $S \subset V$. Show that $v \in V$ and $v \notin \operatorname{span} S$ implies that $\{v, s\}$ is linearly independent for all $s \in \operatorname{span} S$.

   By theorem 1.7, $S \cup \{v\}$ is LI; by question 2, $\{s, v\} \subset S \cup \{v\}$ and thus is LI as well.

---

4. Let S_1, S_2 be subsets of the vector space V. Recall that we define

$$S_1 + S_2 := \{x_1 + x_2 \mid x_1 \in S_1, x_2 \in S_2\}.$$

   Prove that $\operatorname{span}(S_1 \cup S_2) = \operatorname{span} S_1 + \operatorname{span} S_2$.

   Any element of the RHS is naturally the sum of a linear combination in S_1 and linear combination in S_2. By definition of linear combinations, by adding these together, it's simply a linear combination of $S_1 \cup S_2$; same argument for the opposite direction.

---

5. Let S_1, S_2, V be as in question 4. Prove that

$$\operatorname{span}(S_1 \cap S_2) \subset \operatorname{span} S_1 \cap \operatorname{span} S_2,$$

   and give an example when the opposite inclusion is false.

   Observe that $\operatorname{span}\{e_1, e_2\} = \operatorname{span}\{e_1 + e_2, e_2\}$ but

$$\{e_1, e_2\} \cap \{e_1, e_3\} \neq \{e_1 + e_2, e_2 \cap \{e_1, e_3\} = \emptyset.$$

---

6. Let $v \in V$, where V is a vector space over the field F. Prove that $\operatorname{span}\{v\} = \operatorname{span}\{cv\}$ for all $0 \neq c \in F$.

   If $x \in \operatorname{span}\{v\}$ then

$$x = av = \frac{a}{c}cv \in \operatorname{span}\{cv\}.$$

   Same argument holds in the opposite direction.

7. Let $M_{n\times n}(F) = \operatorname{span}\{A_i \mid 1 \le i \le k\}$ for some integers n, k. Prove that

$$\operatorname{span}\{A_i \mid 1 \le i \le k\} = \operatorname{span}\{A_i^T \mid 1 \le i \le k\}.$$

If all $M \in M_{n\times n}(F)$ can be written as $\sum \lambda_i A_i$ then it can also be written as $(\sum \lambda_i A_i)^T = \sum \lambda_i A_i^T$.

---

8. Let V be a vector space and let W be a subset of V. Prove that $\operatorname{span} W = W$ if and only if W is a subspace.

In general, the span of a set is always a subspace, so W is also a subspace. All linear combinations of W are in W because of additive and multiplicative closure; therefore $\operatorname{span} W \subset W$. Moreover, for all $w \in W$, we have $w = 1 \cdot w \implies w \in \operatorname{span} W$, so $W \subset \operatorname{span} W \implies \operatorname{span} W = W$.

---

9. Let $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ be the vector space of all functions from reals to reals. Let $S \subset V$ be the subset of all even and all odd functions. Show that $V = \operatorname{span} S$.

Any function $f \in \mathcal{F}(\mathbb{R}, \mathbb{R})$ can be written as the unique sum of an even and odd function in this manner,

$$f(x) = \underbrace{\frac{f(x) + f(-x)}{2}}_{\text{even}} + \underbrace{\frac{f(x) - f(-x)}{2}}_{\text{odd}},$$

and so we are done. (Covered in **MAT157** Tut. 3)

---

10. Let V be the space of all polynomials in x with coefficients in $\mathbb{F}_5$. Determine if each of the subsets are linearly dependent or independent,
    (a) $\{x^2 + 1, x + 1, 1\}$. Rewrite as $\{x^2, x, 1\}$, which is clearly LI.
    (b) $\{3x + 1, x + 1, 4\}$. $3(x + 1) + 2(4) = 3x + 11 \equiv 3x + 1$, which means the set is LD.
    (c) $\{p(x) \in V \mid p(0) = 0\}$. Take x and $2x$, which is in our set. Then $2(x) = 2x$ and so the set is LD.

---

11. Let $S = \{v_1, \ldots, v_n\} \subset V$, a vector space. Show that there exists a linearly independent subset $B \subset S$ such that $\operatorname{span} B = \operatorname{span} S$. Is this set B unique?

Either S is linearly independent or dependent; if it is linearly independent, then take $B = S$, and B is necessarily unique (otherwise $\dim B < \dim S$ and is not a basis of V). Otherwise, we may take $x \in S$ such that $x \in \operatorname{span}(S \setminus \{x\})$ (by theorem 1.7), and let B be S minus all x with such a property; Then $\operatorname{span} B = \operatorname{span} S$, and B is unique.

12. Let $S = \{(1,0,1), (-1,2,0), (0,1,0), (1,1,1)\} \subset \mathbb{Q}^3$. Find a linearly independent subset $B \subset S$ such that $\operatorname{span} B = \operatorname{span} S$.

    $(1,1,1) = (1,0,1) + (0,1,0)$ so $B = \{(1,0,1), (-1,2,0), (0,1,0)\}$ satisfies $\operatorname{span} B = \operatorname{span} S$.

§13 Day 10: Basis, Steinitz Exchange Lemma, and Dimension (Oct. 10, 2023)

Reminders from last class: let V be a vector space with subset $S \subset V$; then $\operatorname{span}(S) = \{a_1x_1 + \cdots + a_nx_n \mid x_i \in S, a_i \in F, n \geq 1\}$. We say S is linearly independent if and only if $a_1x_1 + \cdots + a_nx_n = 0$ (with $x_i \in S$ distinct and $a_i \in F$) implies $a_1 = \cdots = a_n = 0$. Moreover, for any two subsets $S_1 \subset S_2$, we have

- If S_1 spans V, then so does S_2.
- If S_2 is linearly independent, so is S_1.
- If S_1 is linearly dependent, so is S_2.[11]

§13.1 Basis

A subset β of V is a basis if it is linearly independent and spanning (that is, $\operatorname{span}(\beta) = V$). For example, $V = \mathbb{R}^2$ has basis $\binom{1}{0}, \binom{0}{1}$ since

$$a\begin{pmatrix}1\\0\end{pmatrix} + b\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}a\\b\end{pmatrix},$$

which takes on all elements in $\mathbb{R}^2$. Note that the basis of a set is not necessarily unique (we can find infinitely many such bases, actually). $\binom{1}{c}$ and $\binom{0}{1}$ for some $c \in \mathbb{R}$ is such an example.

Example 13.1 (Basis of $\mathbb{R}$)
$\{x\}$ is a basis for $\mathbb{R}$, provided $x \neq 0$.

Example 13.2 (Basis of F^n)
F^n has basis

$$\underbrace{\begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}}_{e_1}, \underbrace{\begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}}_{e_2}, \underbrace{\begin{pmatrix}0\\0\\1\\\vdots\\0\end{pmatrix}}_{e_3}, \ldots, \underbrace{\begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix}}_{e_n}.$$

Note that $e_1, \ldots, e_n$ is conventional labelling.

Example 13.3 (Basis of $P_n(F)$)
$\{1 = x^0, x^1, x^2, \ldots, x^n\}$ is a basis for $P_n(F)$ (note that $P_n(F) = \{p(x) \in P(F) \mid \deg p(x) \leq n\}$, which is a subspace of $P(F)$.)

Anyways, onto theorems from class.

[11] In general, it's easy to make big sapnning sets, such as V itself, and it's easy to make small LI sets, like $\emptyset$.

Theorem 13.4 (Theorem 1.8 (from the book))

Let $\beta = \{x_1, \ldots, x_n\} \subset V$. Then β is a basis of V if and only if, for all $x \in V$, there are unique $a_1, \ldots, a_n \in F$ such that $x = a_1x_1 + \cdots + a_nx_n$.

The forward implication is already covered in the book, so we simply prove that β spans, and is linearly independent. The former is demonstrated by the fact that all x are linear combinations of β, while we may consider $x = 0 = 0x_1 + 0x_2 + \cdots + 0x_n$ for the latter, which is the unique linear combination from our assumption.

Theorem 13.5 (Theorem 1.9)

If a finite subset $S \subset V$ spans V, then some subset of S spans V. In particular, V admits a finite basis.

Remark 13.6. By the well ordering principle, any nonempty subset of $\mathbb{N}$ has a minimum. In the same manner, any nonempty subset of $\mathbb{Z}$ that is bounded below has as minimum, and a maximum if bounded above.

Let β be a linearly independent subset of S of largest cardinality (by remark). We claim that β is a basis of V. β is automatically linearly independent by definition; so now we show it spans V. As β is the largest subset possible, we see that $\beta \cup \{x\}$ is linearly dependent for all $x \in S$ (where $x \notin \beta$) implies all $x \in S \subset \text{span}(\beta)$. Then

$$S \subset \text{span}(\beta) \implies \text{span}(S) \subset \text{span}(\beta) \underbrace{\implies}_{S \text{ is spanning}} \text{span}(\beta) = V. \quad \square$$

Theorem 13.7 (Replacement Theorem / Steinitz Exchange Lemma)

Suppose L, S are finite subsets of V. If L is linearly independent and S is spanning with $|L| = m$, $|S| = n$, $m \leq n$, then there exists a subset $T \subset S$ such that

(a) $|T| = n - m$,

(b) $L \cup T$ spans V.

We may prove this by induction on m. In the case $m = 0$, we have $0 \leq n$ and so may pick $T = S$, and we are done. Now, suppose that the statement is true for some m; we now prove it for $m + 1$. Let $L = \{x_1, x_2, \ldots, x_{m+1}\}$. First, we apply the induction hypothesis to $L_m = \{x_1, \ldots, x_m\}$. This means there exists $T_m = \{y_{m+1}, \ldots, y_n\}$ (with $n - m$ elements) such that $L_m \cup T_m$ is spanning. Since $x_{m+1} \in V$, there exists a linear combination

$$x_{m+1} = \sum_{i=1}^{m} a_i x_i + \sum_{i=m+1}^{n} b_i y_i$$

for some $a_i, b_i \in F$. In particular, if $b_i = 0$ for all i, this would contradict L being linearly independent, so we say that $b_i \neq 0$ for some i (and thus, we reorder the y_i's such that $b_{m+1} \neq 0$). Then we may write

$$y_{m+1} = -b_{m+1}^{-1} \left(\sum_{i=1}^{m} a_i x_i + \sum_{i=m+2}^{n} b_i y_i \right)$$

which is in the span of $\{x_1, \ldots, x_{m+1}, y_{m+2}, \ldots, y_n\}$. Now, it only remains to show $\text{span}\{x_1, \ldots, x_m, y_{m+1}, \ldots, y_n\} = \text{span}\{x_1, \ldots, x_{m+1}, y_{m+2}, \ldots, y_n\}$. In particular, the

left hand side is V (by above), and take $T = \{y_{m+2}, \dots, y_n\}$. Then $y_{m+1} \in \text{RHS} \implies x_1, \dots, x_m, y_{m+2}, \dots, y_n \in \text{RHS}$, and thus LHS $\subset$ RHS. For the other direction of the implication, we see LHS $= V$ which immediately finishes it. $\square$

With the main theorem out of the way, we have a few corollaries to state:

Corollary 13.8 (Cor. 1 of Steinitz Exchange)

If V admits two finite bases β, γ, then $|\beta| = |\gamma|$. We see this by applying Steinitz exchange in both directions (since β, γ are both linearly independent *and* spanning), where β LI and γ spanning implies $|\beta| \leq |\gamma|$, and vice versa, yielding that they're equal.

We also define that if V admits a finite basis, we say its *dimension*, $\dim(V)$, is the cardinality of any of its basis (from the above corollary, all bases have the same cardinality). Moreover, if V admits a finite basis β, then every basis γ is also finite. Suppose $|\beta| = n$. If γ is infinite, then it is linearly independent and every subset of γ is LI too, which is a contradiction as it is skew with Steinitz, which we see by taking a finite subset $\gamma' \subset \gamma$ with $|\gamma'| > n$, then letting γ' be LI while β spans. By Steinitz, we see $|\gamma'| \leq |\beta| = n$ which is contradictory.

Moreover, we may also say if V admits finite basis, then it is finite-dimensional. Examples:

$$\begin{aligned} \dim(F^n) &= n \\ \dim(M_{m\times n}(F)) &= mn \\ \dim(P_n(F)) &= n+1 && \text{(Set } \{x^0, x^1, \dots, x^n\}\text{)} \\ \dim(\{0\}) &= 0 && \text{(Empty set } \emptyset\text{.)} \end{aligned}$$

§14 Day 11: More Theorems (Oct. 12, 2023)

Today we continue off of the theorems introduced from last class. We introduce a second corollary to the Steinitz exchange theorem, where

Corollary 14.1 (Cor. 2 of Steinitz Exchange)

Let V be of finite dimension n.

(i) If subset $L \subset V$ is linearly independent, then $|L| \leq n$ (where $\dim V = n$). We have $|L| = n$ if and only if L is a basis of V.

(ii) If $S \subset V$ spans V, then $|S| \geq n$. We have $|S| = n$ if and only if S is a basis of V.

(iii) Any linearly independent subset can be extended to a basis. That is, for any linearly independent subset $L \subset V$, there is a basis β where $L \subset \beta$.

We start by proving (i). If L is a basis of V, we automatically have $|L| = \dim V = n$ (by Cor. 1), so it remains to prove the other direction of implication. If $|L| = n$, then we may apply replacement to some spanning S where there exists $T \subset S$ such that $L \cup T$ spans (though $|T| = 0$, so $L \cup T = T \implies L$ is a basis). Moreover, we see L cannot be infinite, since otherwise any finite subset is linearly independent, which means we may pick $L' \subset L$ such that $|L'| = n + 1$ is contradictory by $|L'| \leq n$.

A similar proof follows for (ii), where, for any spanning S, we may find a spanning $H \subset S$ (by Theorem 1.9) where $|H| = n$ (by Cor. 1). Therefore $G \supset H \implies |G| \geq |H| = n$, and we may also conclude $|S| = n$ implies basis by Cor. 1 as well.

For (iii), if L is a linearly independent subset of V (let $m = |L|$), then by replacement we may find $T \subset S$ (where S spans, $n = |S|$) with $|T| = n - m$, so that $L \cup T$ spans V. By (i), we see that $L \cup T$ is necessarily a basis for V. □

Theorem 14.2 (Theorem 1.11)

Let V be finite-dimensional such that $W \subset V$ is a subspace. Then

(i) W is finite-dimensional.

(ii) $\dim W \leq \dim V$, with equality if and only if $W = V$.

(iii) Any basis W can be extended to be a basis of V.

Let $n = \dim V$. Then we see $\dim W \leq \dim V = n$ because either $W = \{0\} \implies \dim W = 0 \leq n$, or we have linearly independent vectors $\{x_1, \ldots, x_k\} \in W$, where k is maximized (as in, we pick the largest LI subset by well ordering) and $k \leq n$ because no LI subset of V may contain more than $\dim V = n$ vectors (by definition); if not, adjoining any other vector from W creates a linearly dependent set, and so $\{x_1, \ldots, x_k\}$ spans W by Theorem 1.7, and $\dim W = k \leq n$, which concludes (i) and (ii). Note that if $\dim W = \dim V$, then the basis $\{x_1, \ldots, x_n\}$ for W also spans V by Cor. 2, forcing $W = V$.

A similar argument holds for (iii); any basis β of W is linearly independent in V, which means it can be extended to be a basis of V by Cor. 2.

Next lecture we will probably properly introduce linear transformations (so I'll write up linear transformation stuff when we properly get into it next time).

empty until i'm done with mat157 midterm 1 c: im bad at keeping promises sorry—soon soon soon

§15 Day 12: Linear Transformations (Oct. 17, 2023)

Prof. Herzig is out for the week, so Prof. Meinrenken will be subbing in (**MAT247** instructor). We're now on section 2 of the book!

§15.1 Linear Transformations (Linear Maps)

Let V, W be vector spaces over the same field F; we say the function $T : V \to W$ is a linear transformation (or linear map) if, for all $x, y \in V$, and any scalar $c \in F$, we have

(a) $T(x + y) = T(x) + T(y)$,

(b) $T(cx) = cT(x)$

Equivalently, T is also a linear map if $T(cx + y) = T(cx) + T(y)$ (combination of the above), or if T, in words, sends linear combinations to linear combinations; i.e., $T(a_1x_1 + \cdots + a_nx_n) = a_1T(x_1) + \cdots + a_nT(x_n)$. Moreover, note that $T(0_V) = 0_W$ for all linear maps T.

Example 15.1

Here are a few examples of linear maps:

$$T : \mathbb{R}^3 \to \mathbb{R}^3 \text{ where } (t_1, t_2, t_3) \mapsto (t_2, t_1, t_3)$$
$$T : \mathbb{R}^3 \to \mathbb{R}^2 \text{ where } (t_1, t_2, t_3) \mapsto (t_1, t_2)$$
$$T : \mathbb{R}^2 \to \mathbb{R}^3 \text{ where } (t_1, t_2) \mapsto (t_1, t_1, t_2)$$
$$T : \mathbb{R}^2 \to \mathbb{R}^2 \text{ where } (t_1, t_2) \mapsto (2t_1 + 3t_2, 3t_1 - 4t_2)$$

However,

$$T : \mathbb{R} \to \mathbb{R}^3 \text{ where } t \mapsto (t, t^2, t^3)$$

is *not* linear, and therefore does not satisfy our criterion.

Anyways, we now provide more examples of linear transformations from lecture:[12]

(a) The map $T : M_{m\times n}(F) \to M_{n\times m}(F)$ given by $A \mapsto A^T$.

(b) The map $T : \mathcal{F}(S, F) \to F$ given by $f \mapsto f(p)$ for a fixed $p \in S$.

(c) The map $T : P(\mathbb{R}) \to \mathbb{R}$ given by $f \mapsto f(t)$ for a fixed $t \in \mathbb{R}$.

(d) The identity map $I_V : V \to V$ given by $v \mapsto v$.

(e) The zero map $0_V : V \to V$ given by $v \mapsto 0$.

(f) Straight lines! $T : \mathbb{R} \to \mathbb{R}$ given by $x \mapsto mx + b$ is a linear map if and only if $b = 0$. We may check this by observing

$$\begin{aligned} T(x_1 + x_2) &= T(x_1) + T(x_2) \\ &= mx_1 + b + mx_2 + b \\ &= m(x_1 + x_2) + 2b, \end{aligned}$$

which implies $b = 0$ necessarily; otherwise T is not a linear map by our definitions.

[12] boy do i have an obsession with colons and semicolons... colon three uwu

§15.2 Kernel, Nullity, Image, and Rank

We start with two quick definitions; let $T : V \to W$ be a linear map, then

- We say $N(T) = \{v \in V \mid T(v) = 0\} = T^{-1}(0)$ is called the **kernel** or **nullspace** of T.[13]
- We say $R(T) = \{w \in W \mid \exists v \in V;\ T(v) = w\}$ is called the **range** or **image** of T.

Remark 15.2. For our linear map $T : V \to W$, note that $N(T) \subset V$ and $R(T) \subset W$; both of which are subspaces. We see this from

$$T(ax_1 + bx_2) = aT(x_1) + bT(x_2) = a0 + b0 = 0$$

for whatever $a, b \in F$, $x_1, x_2 \in V$; and thus $N(T)$ (are we supposed to say $N(T)$ or $\ker T$?) is a subspace of V. Moreover, let us arbitrarily pick $T(x_1), T(x_2) \in R(T)$. Let a, b be scalars in F once again. Then

$$aT(x_1) + bT(x_2) = T(ax_1 + bx_2) \in R(T)$$

and so we are done.

On top of kernel and image, we also define the following,

- $\text{null}(T) = \dim N(T)$
- $\text{rank}(T) = \dim R(T)$.

Now let's look at a few examples for intuition before establishing the Rank-Nullity Theorem:

Example 15.3

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ with $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \mapsto \begin{pmatrix} a_2 + a_3 \\ a_1 - a_3 \\ -a_1 - a_2 \end{pmatrix}$. Then $N(T) = \left\{ \begin{pmatrix} c \\ -c \\ c \end{pmatrix} \mid c \in \mathbb{R} \right\}$ and $R(T) = \left\{ \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \mid b_1 + b_2 + b_3 = 0 \right\}$; and thus $\text{null}(T) = 1$ and $\text{rank}(T) = 2$. Notice $\dim R(T) + \dim N(T) = 2 + 1 = 3 = \dim \mathbb{R}^3$.

Example 15.4

Let V, W be matrix spaces $M_{n\times n}(F)$. Then let $T : M_{n\times n}(F) \to M_{n\times n}(F)$ where $A \mapsto A - A^T$. Then $N(T)$ is the set of all symmetric $n \times n$ matrices; and so $\text{null}(T) = \frac{n(n+1)}{2}$. We also see $R(T)$ is the set of all antisymmetric $n \times n$ matrices, and so $\text{rank}(T) = \frac{n(n-1)}{2}$. Notice $\dim R(T) + \dim N(T) = \frac{n(n-1)}{2} + \frac{n(n+1)}{2} = n^2 = \dim M_{n\times n}(F)$.

With these two examples in place we now prove the Rank-Nullity Theorem.

[13] as far as i know i think kernel is more used for an abstract linear transformation while nullspace is more used for matrices. probably just convention though!

Theorem 15.5 (Rank-Nullity Theorem)

Let V, W be vector spaces over a field F. Then for any linear map $T : V \to W$, we have

$$\operatorname{null} T + \operatorname{rank} T = \dim V.$$

Since $N(T)$ is a subspace of V (as demonstrated earlier), we may find a basis; let it be $\mathcal{K} := \{v_1, \ldots, v_k\} \subset N(T)$. By the Steinitz Exchange lemma, we may extend the set $\mathcal{K}$ with $n-k$ linearly independent vectors; let these vectors belong to $\mathcal{S} := \{v_{k+1}, \ldots, v_n\} \subset V \setminus N(T)$. We now have a basis $\mathcal{B} := \mathcal{K} \cup \mathcal{S} = \{v_1, \ldots, v_n\} \subset V$. By definition,

$$R(T) = \operatorname{span}(T(\mathcal{B})) = \operatorname{span}\{T(v_1), \ldots, T(v_k), \ldots, T(v_n)\}.$$

Note that $a_1T(v_1) + \cdots + a_kT(v_k) = T(a_1v_1 + \cdots + a_kv_k) = T(0) = 0$ because $a_1v_1 + \cdots + a_kv_k$ is a linear combination of $\{v_1, \ldots, v_k\}$, and so is an element of $N(T)$, and thus vanishes. Thus,

$$\operatorname{span}\{T(v_1), \ldots, T(v_k), \ldots, T(v_n)\} = \operatorname{span}\{T(v_{k+1}), \ldots, T(v_n)\} = \operatorname{span}(T(\mathcal{S})).$$

With this in place, we claim $T(\mathcal{S}) = \{T(v_{k+1}), \ldots, T(v_n)\}$ forms a basis of the image $R(T)$. We've already shown $T(\mathcal{S})$ spans $R(T)$, so it remains to show that the set is linearly independent. Suppose we have

$$\sum_{i=k+1}^{n} a_iT(v_i) = 0_W$$

for some $a_{k+1}, \ldots, a_n \in F$. By the linearity of T, we may write

$$T\left(\sum_{i=k+1}^{n} a_iT(v_i)\right) = 0_W \implies \sum_{i=k+1}^{n} a_iT(v_i) \in N(T) = \operatorname{span}(\mathcal{K}),$$

and so our linear combination is in the kernel of T; thus, we may write

$$0_W = \sum_{i=k+1}^{n} a_iT(v_i) = \sum_{i=1}^{k} a_iT(v_i) \implies \sum_{i=1}^{n} a_iT(v_i) = 0_W$$

for some scalars $a_1, \ldots, a_k \in F$. However, since $\mathcal{B}$ is linearly independent, we see all scalars a_i must be equal to zero, showing that $T(\mathcal{S})$ is linearly independent as well; thus, we may conclude that it is indeed a basis. To finish, we have

$$\operatorname{null} T + \operatorname{rank} T = |\mathcal{K}| + |T(\mathcal{S})| = k + (n-k) = n,$$

which is indeed equal to $\dim V$. □

Here are a few examples using the above theorem;

Example 15.6

Let $T : P_3(\mathbb{R}) \to P_3(\mathbb{R})$, where $f \mapsto f'$; i.e., $T(a_0 + a_1x + a_2x^2 + a_3x^3) = a_1 + 2a_2x + 3a_3x^2$. Then we can quickly check $\operatorname{null}(T) = 1$ and $\operatorname{rank}(T) = \dim(P_3(\mathbb{R})) - 1 = 4 - 1 = 3$.[a]

[a] i have no idea what the "if $2 = 0$" case from class meant...? $\operatorname{char}(F) \neq 2$

§15.3 Isomorphisms

A linear map $T : V \to W$ is called an isomorphism if it is a bijection; in another way, two vector spaces V and W are called isomorphic if there exists an invertible linear transformation (aka an isomorphism) between them.

> **Remark 15.7.** In general, isomorphisms are neat because it lets you "replace" a vector space with another that is more familiar; $P_2(\mathbb{R}) \cong \mathbb{R}^3$ (i.e., there exists an isomorphism $T : P_2(\mathbb{R}) \to \mathbb{R}^3$.

Moreover, we also have the following properties,

- A map $T : V \to W$ is injective if and only if $N(T) = \{0\}$.
- A map is surjective if and only if $R(T) = W$.
- A map is bijective if and only if it satisfies both of the above.

Proofs of the above are left to the reader; I go hug Blåhaj now.

§16 Day 13: Linear Transformations; Isomorphisms (Oct. 19, 2023)

For vector spaces V, W over F, define

$$\mathcal{L}(V,W) = \{T : T : V \to W\}$$

i.e., "the space of linear transformations from V to W." As a vector space $\mathcal{L}$[14] must satisfy:

$$(T_1 + T_2)(v) = T_1(v) + T_2(v),$$
$$(a \cdot T)(v) = aT(v).$$

We may check that $\mathcal{L}$ is indeed a vsp by going through the axioms, but I don't think that's really needed here.

§16.1 Isomorphisms II

As defined yesterday, we further call any map $T \in \mathcal{S}(V,W)$ an isomorphism if it is invertible (remember that invertibility comes with bijectivity). We call $V \cong W$ (isomorphic) if such a T exists between them. Now for examples.

(a) There exists an isomorphism between $P_n(F)$, F^{n+1}, and let said isomorphism be T. Then

$$T(a_0 + a_1x + \cdots + a_nx^n) = (a_0, a_1, \ldots, a_n),$$
$$T^{-1}(a_0, a_1, \ldots, a_n) = a_0 + a_1x + \cdots + a_nx_n.$$

Since such an inverse is defined we see $P_n(F) \cong F^{n+1}$.

(b) Alternatively, when viewing from the other direction, let us pick $(x_1, \ldots, x_{n+1}) \in F^{n+1}$ with all x_i distinct. Let $T : P_n(F) \to F^{n+1}$, where

$$T(p) = (p(x_1), \ldots, p(x_n)).$$

Notice that T has trivial kernel, since if $T(p) = 0$, we have $p(x_1) = p(x_2) = \cdots = p(x_{n+1}) = 0$; if p has $n+1$ distinct roots by assumption but $\deg(p) \leq n$, it must be the zero polynomial. Thus, we may claim T is bijective and is thus an isomorphism between $P_n(F)$ and F^{n+1}, and we may find a polynomial from Lagrange Interpolation to get T^{-1}.

(c) $P(F)$ is not isomorphic to F^∞, the set of all infinite sequences, since there does not actually exist a bijection between them.[15] Instead, we have $P(F) \cong F^\infty_{\text{fin}}$, the set of sequences with finitely many nonzero elements.

(d) Every n-dimensional vector space V over F is isomorphic to F^n. We may see this by taking the basis $(v_1, \ldots, v_n)$ of V; then we have an isomorphism $T : F^n \to V$ given by

$$(c_1, \ldots, c_n) \mapsto \sum_{i=1}^{n} c_i v_i,$$

which is necessarily injective because of linear independence (and thus $N(T) = \ker(T) = \{0\}$) and surjective because $(v_1, \ldots, v_n)$ is spanning.

[14] the amount of times i've fucked up this symbol *man* like £, cursive 2, etc... ouch.
[15] if you're curious, $\dim F^\infty = 2^{\aleph_0}$, while $\dim P(F) = \aleph_0$... actually, someone check if i'm bullshitting

Onto theorems for today!

Theorem 16.1 (Same finite dimension implies Isomorphism)
Let V, W be finite-dimensional vector spaces. Then $V \cong W$ if and only if $\dim V = \dim W$.

Suppose there exists an isomorphism T between V and W. Then we know the kernel of T, $N(T) = \{0\}$, which implies, by the rank-nullity theorem, that

$$\underbrace{\dim N(T)}_{0} + \dim R(T) = \dim R(T) = \dim V.$$

Since T is also surjective, we know $R(T) = W$, and so $\dim V = \dim W$. For the converse, let $\dim V = \dim W$ be our pre-assumption. Then let $v_1, \ldots, v_n$ be the basis of V, and $w_1, \ldots, w_n$ be the basis of W. Take

$$T(a_1v_1 + \cdots + a_nv_n) = a_1w_1 + \cdots + a_nw_n,$$

$$T^{-1}(a_1w_1 + \cdots + a_nw_n) = a_1v_1 + \cdots + a_nv_n,$$

and we are done). □

Remark 16.2. The above theorem only works for vector spaces of finite dimension. This is false if either are infinite; ex: F^∞ and F^∞_{fin}.

Theorem 16.3
Suppose $\dim V = \dim W < \infty$. Let $T \in \mathcal{L}(V, W)$. Then the fact T is an isomorphism can be concluded from either T being injective or surjective.

If T is injective, then $N(T) = \{0\}$, implying $\dim R(T) = \dim V = \dim W$, which yields an isomorphism by the above theorem. This also implies surjectivity; if we start by assuming T is surjective, then $\dim R(T) = \dim W$, and $\dim N(T) = 0$ necessarily from rank-nullity (which yields injectivity). Thus, one implies the other, and we are done. □

Remark 16.4. Once again this isn't true if $\dim V, \dim W = \infty$.

§17 Tutorial 5: Tut. 5 HW Solutions (Oct. 19, 2023)

(a) Let $T : M_{2\times 3}(F) \to M_{2\times 2}(F)$, given by

$$T = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 2a_{11} - a_{12} \end{pmatrix}$$

§18 Day 14: Matrix Representations (Oct. 24, 2023)

Our midterm will be on Oct. 31, 11:10am to 1pm. Practice tests have been announced and posted on here (hmm.. doesn't seem to exist yet).

Recall that an isomorphism is a bijective linear map $T : V \to W$ (where we say $V \cong W$ if there exists such an isomorphism T); if $\dim V = \dim W < \infty$, then it suffices to check for injectivity or surjectivity to conclude that T is an isomorphism (since one implies the other).

Remark 18.1. Given $T : V \to W$, we see T is injective if and only if $\dim V \leq \dim W$, and T is surjective if and only if $\dim V \geq \dim W$.

A few more preliminary comments on linear maps! Recall that the space of all linear maps $\mathcal{L}(V, W) = \{T : V \to W \mid T \text{ is a linear map}\}$ is a subspace of the function space $\mathcal{F}(V, W)$. We also have

Theorem 18.2 (Theorem 2.9; Composition of Linear Maps)

Given linear maps $S : U \to V$, $T : V \to W$, we have $T \circ S : U \to W$ is also a linear map. To check this, observe

$$\begin{aligned}(T \circ S)(ax + y) &= T(S(ax + y)) \\ &= T(aS(x) + S(y)) \\ &= aT(S(x)) + T(S(y)) \\ &= a(T \circ S)(x) + (T \circ S)(y)\end{aligned}$$

by the linearity of S and T, and so we are done.

Moreover, if we have linear maps $S : U \to V$ and $T, T_1, T_2 : V \to W$, then

$$(T_1 + T_2) \circ S = T_1 \circ S + T_2 \circ S$$
$$(aT) \circ S = a(T \circ S).$$

Though, caution that

$$S \circ T \neq T \circ S,$$

since function composition is not commutative. Moreover, we may define function composition recursively; i.e., if $T : V \to V$, we say

$$T^n = \underbrace{T \circ T \circ \cdots \circ T}_{n \text{ times}} : V \to V,$$

and $T^0 = \mathrm{id}_V$. On the topic of composition, we also have the following theorem,

Theorem 18.3 (Invertibility of Isomorphism)

Let $T \in \mathcal{L}(V, W)$; T is an isomorphism if and only if there exists an $S \in \mathcal{L}(W, V)$ such that

$$S \circ T = \mathrm{id}_V, \quad T \circ S = \mathrm{id}_W.$$

We already proved the forward direction last week by demonstrating the existence of the inverse of T, and for the converse, bijectivity implies invertibility. Alternatively, if this isn't good enough, use Theorem 2.6 to map basis to basis and the reverse:

Theorem 18.4 (Theorem 2.6: Map of Basis)

Suppose $\{x_1, \dots, x_n\}$ be a basis of V, and let us have $y_1, \dots, y_n \in W$. Then there exists a unique linear transformation $T : V \to W$ such that $T(x_i) = y_i$ for $1 \leq i \leq n$. (Note that our choice of $y_1, \dots, y_n$ doesn't have to be a basis.)

To prove this, we start with uniqueness:

- Suppose T_1, T_2 satisfy the above. Let $x \in V$; Since $\{x_i\}$ spans, we may write $x = a_1x_1 + \cdots + a_nx_n$, and so we have

$$T_1(x) = \sum_{i=1}^{n} a_iT_1(x_i) = \sum_{i=1}^{n} a_iy_1 = \sum_{i=1}^{n} a_iT_2(x_i) = T_2(x),$$

  and so $T_1 = T_2$. Thus T is unique.

- Let T satisfy $T(x_i) = y_i$; then for all $x \in V, x = \sum a_ix_i$ has a unique combination of unique scalars $a_1, \dots, a_n$, and we may write

$$T(x) = a_1T(x_1) + \cdots + a_nT(x_n);$$

  it is easy to check see this is linear (check problem 5 on the homework of week 6).

As a corollary, we also see if $S = \{x_1, \dots, x_n\}$ spans V, and we have any two linear maps $T_1, T_2 : V \to W$ where $T_1(x_i) = T_2(x_i)$ for all $1 \leq i \leq n$, we have $T_1 = T_2$ (by additivity and linearity).

§18.1 Matrix Representations

We start off by defining ordered bases: given any finite dimensional vector space V, we say $\{x_1, \dots, x_n\}$ is an ordered basis of V where $\{x_1, \dots, x_n\}$ is LI and spanning; however, in this case, we care about the order of $x_1 \dots x_n$. As an example, observe how $\mathbb{R}^2$ has $\{e_1, e_2\}, \{e_2, e_1\}$ as bases; while as sets they're the same thing, as ordered bases, they're different because the order is switched up. Now, we may proceed to build matrix representations. First, let us have an element $x \in V$ with basis β; we may write the **coordinate vector of x relative to β**,

$$[x]_\beta = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in F^n \text{ where } x = a_1x_1 + \cdots + a_nx_n$$

with $a_1, \dots, a_n$ unique (as $\beta = \{x_1, \dots, x_n\}$ forms a basis). Notice that $[x_i]_\beta = e_i$ as above; in fact, $x \mapsto [x]_\beta$ is a linear transformation from $V \to F^n$. First, let us represent the entire vector space V with respect to β:

Theorem 18.5 (Theorem 2.21: φ_β is an isomorphism)

If $\beta = \{x_1, \dots, x_n\}$ is an ordered basis of V, let us have

$$\begin{aligned} \varphi_\beta : V &\to F^n, \\ x &\mapsto [x]_\beta. \end{aligned}$$

Then φ_β is an isomorphism.

To prove this, we first verify φ_β is indeed a linear map. Taking any scalar $c \in F$ and $a, b \in V$, we start by writing a, b in terms of the basis β, where we have

$$\begin{aligned} a &= a_1 x_1 + \cdots + a_n x_n, \\ b &= b_1 x_1 + \cdots + b_n x_n, \end{aligned}$$

implying $\varphi_\beta(a) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ and $\varphi_\beta(b) = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$. Thus, we may write

$$\varphi_\beta(ca + b) = \begin{pmatrix} ca_1 + b_1 \\ \vdots \\ ca_n + b_n \end{pmatrix} = c \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = c\varphi_\beta(a) + \varphi_\beta(b),$$

which concludes φ_β is indeed linear. Now, we may show that it is indeed an isomorphism; for completeness, we will show injectivity and surjectivity:

- First, given any u, v such that $[u]_\beta = [v]_\beta$, let us have

  $$[u]_\beta = [v]_\beta = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix},$$

  which implies $u = c_1 x_1 + \cdots + c_n x_n = v$, concluding injectivity.

- Now, observe that for any $\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in F^n$, we simply pick $u = c_1 x_1 + \cdots + c_n x_n$; then $\varphi_\beta(u)$ is as desired. Thus we also have surjectivity, and so φ_β is an isomorphism. $\square$

From this, we may also conclude that any n-dimensional vector space is isomorphic to F^n. Either way, before we proceed, let's check a quick example. Given $V = P_2(F)$, let us write $x^2 - 1$ in terms of two different ordered bases: let us have $\beta_1 = \{1, x, x^2\}$ and $\beta_2 = \{x, x^2, 1\}$. Then

$$[x^2 - 1]_{\beta_1} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \qquad [x^2 - 1]_{\beta_2} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

With this in place, we may now define the matrix representation of a linear map T. Let $T : V \to W$ be a linear map where V has an ordered basis $\beta = \{x_1, \ldots, x_n\}$, and W has an ordered basis $\gamma = \{y_1, \ldots, y_m\}$. By Theorem 2.6, if we want T to be unique, we would want $T(x_i) = y_i$. We now construct the matrix representation. For all $1 \leq i \leq n$, let

$$\begin{aligned} T(x_i) &= a_{1i} y_1 + \cdots + a_{ni} y_n \\ &= \sum_{j=1}^{m} a_{ji} y_i \end{aligned}$$

with unique $a_{ji} \in F$.[16]

[16] yes, by convention j comes first oh what a world we live in

With this in place, we may now write

$$[T]_\beta^\gamma = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in M_{m\times n}(F).$$

Note that the ith column of $[T]_\beta^\gamma$ is equal to $[T(x_i)]_\gamma$. Alternatively, for all $v \in V$, we may decompose it into $v = a_1x_1 + \cdots + a_nx_n$ (with $a_1, \dots, a_n \in F$ scalars) and write the following,

$$T(v) = T(a_1x_1 + \cdots + a_nx_n) = \sum_{j=1}^{n} a_jT(x_j) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}x_j \right) y_i,$$

and thus

$$[T(v)]_\gamma = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Let's check an example instead for intuition. Let $T : P_3(\mathbb{R}) \to P_2(\mathbb{R})$ defined by $p(x) \mapsto p'(x) + p''(x)$, referring to the derivative and double derivative. Then we may say $\beta = \{1, x, x^2, x^3\}$ and $\gamma = \{1, x, x^2\}$. First, observe that

$$\begin{aligned} T(x_1) &= 0 \\ T(x_2) &= 1 \\ T(x_3) &= 2x + 2 \\ T(x_4) &= 3x^2 + 6x \end{aligned}$$

where $\beta = \{x_1, x_2, x_3, x_4\}$ as per our convention established above. Then

$$[0]_\gamma = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, [1]_\gamma = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, [2x+2]_\gamma = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, [3x^2+6x]_\gamma = \begin{pmatrix} 0 \\ 6 \\ 3 \end{pmatrix},$$

and we have

$$[T]_\beta^\gamma = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Now for some other examples! If $T : V \to V$ is the zero map[17], then

$$[T_0]_\beta^\beta = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix} = 0 \in M_{n\times n}(F)$$

where $\dim V = n$. If $T : V \to V$ is the identity map, then we may say $T = \mathrm{id}_V$ and write

$$[\mathrm{id}_V]_\beta^\beta = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix} = I_n \in M_{n\times n}(F),$$

where once again $\dim V = n$.

[17] question: do we say $T_0 = 0_V$?

Theorem 18.6 (Theorem 2.20: Isomorphism Theorem)

Let V, W be vector spaces of dimension n, m respectively (both finite-dimensional). Then, define Φ (book says Φ_β^γ) where

$$\begin{aligned}\Phi : \mathcal{L}(V,W) &\to M_{m\times n}(F)\\ \Phi(T) &\mapsto [T]_\beta^\gamma,\end{aligned}$$

with β, γ as ordered bases for V and W respectively. Then we have that Φ is an isomorphism.

As a quick corollary, note that $\dim \mathcal{L}(V,W) = \dim V \cdot \dim W$. This is true because they are both equal to $\dim M_{m\times n}(F) = mn$. We now prove Φ is an isomorphism.

- (Linearity is actually Theorem 2.8) Let $\beta = \{x_1, \ldots, x_n\}$ and $\gamma = \{y_1, \ldots, y_n\}$. Let $T, S : V \to W$ be linear maps; necessarily we have unique scalars a_{ij} and b_{ij} for $1 \le i \le m$and $1 \le j \le n$ such that

  $$T(x_j) = \sum_{i=1}^{m} a_{ij} y_i \text{ and } S(y_j) = \sum_{i=1}^{m} b_{ij} y_i \text{ for } 1 \le j \le n.$$

  Thus, we have, for any scalar $c \in F$

  $$(cT + S)(x_j) = \sum_{i=1}^{m} (ca_{ij} + b_{ij}) y_i \implies ([cT+S]_\beta^\gamma)_{ij}$$

  since $ca_{ij} + b_{ij} = (c[T]_\beta^\gamma + [U]_\beta^\gamma)_{ij}$.

- To show the surjectivity of Φ, we may simply invoke Theorem 2.6 to find a unique linear map $T : V \to W$ where

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

  where $1 \le j \le n$ for any matrix $A \in M_{m\times n}(F)$. Since this means $[T]_\beta^\gamma = A$. For injectivity, we want to show that the nullity of $T = 0$; in a similar argument, suppose $T \in \ker \Phi$. Then we have

  $$[T]_\beta^\gamma = 0 = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

  where $T(x_i) = 0$ for all basis vectors in β; by Theorem 2.6, we see such a T is unique, and so $\dim \ker T = 0$ implying T is injective.

Thus, we may conclude that Φ is indeed an isomorphism. □

§19 Day 15: Matrix Representation; Composition of Linear Transformations (Oct. 26, 2023)

Recall from last class; let V have ordered basis $\beta = \{x_1, \dots, x_n\}$. Then φ_β is an isomorphism as follows,

$$\begin{aligned} \varphi_\beta : V &\xrightarrow{\sim} F^n \\ x &\mapsto [x]_\beta \end{aligned}$$

where the matrix representation $[x]_\beta$ is of entries a_1 to a_n (since $\dim F^n = n$) such that $a_1x_1 + \cdots + a_nx_n = x$. Moreover, if we have the linear map $T : V \to W$ equipped with ordered bases $\beta = \{x_1, \dots, x_n\}$ of V and $\gamma = \{y_1, \dots, y_m\}$ of W. Then we may also construct the isomorphism Φ where

$$\begin{aligned} \mathcal{L}(V, W) &\xrightarrow{\sim} M_{m\times n}(F) \\ T &\mapsto [T]^\gamma_\beta. \end{aligned}$$

Remember that the matrix representation $[T]^\gamma_\beta$ is a matrix of n columns and m rows, where each column is equal to $[T(x_i)]_\gamma$ for every $1 \le i \le n$.[18]

§19.1 Matrix Representation Operations

Let $S : V \to W$, $T : W \to Z$ be linear maps where

$$\begin{aligned} \alpha &= \{x_1, \dots, x_n\}, \\ \beta &= \{y_1, \dots, y_m\}, \\ \gamma &= \{z_1, \dots, z_p\} \end{aligned}$$

are ordered bases for V, W, Z respectively. In this subsection, we'll evaluate for $[T \circ S]^\gamma_\alpha$ in terms of $[T]^\gamma_\beta$ and $[S]^\beta_\alpha$. Let these two matrices be A and B respectively, and also let $C = [T]^\gamma_\beta$; we will be writing in terms of their entries. First, start by having

$$\begin{aligned} [T \circ S](x_i) &= T(S(x_i)) \\ &= T\left(\sum_{j=1}^m B_{ji}y_j\right) && \text{(Linear Combination)} \\ &= \sum_{j=1}^m B_{ji}T(y_j) && \text{(Linearity of } T\text{)} \\ &= \sum_{j=1}^m B_{ji}\left(\sum_{k=1}^p A_{kj}z_k\right) && \text{(Linear Combination)} \\ &= \sum_{j=1}^m \underbrace{\left(\sum_{k=1}^p A_{kj}B_{ji}\right)}_{C_{ki}} z_k. \end{aligned}$$

Since this is commensurate to the definition of matrix multiplication, we see $C = AB \in M_{p\times n}(F)$ with $(AB)_{ki} = \sum_{j=1}^m A_{kj}B_{ji}$. Thus, $[T \circ S]^\gamma_\alpha = [T]^\gamma_\beta \cdot [S]^\beta_\alpha$.

[18] sorry, i don't know how to draw this...

Theorem 19.1 (Theorem 2.11: Composition of Matrix Representations)
Given linear maps $S : V \to W$ and $T : W \to Z$, we have

$$[T \circ S]_\alpha^\gamma = [T]_\beta^\gamma \cdot [S]_\alpha^\beta.$$

Proof is above. Do note that the ith column of AB is equal to A multiplied by the ith column of B, and the ith row of AB is the ith row of A multiplied by B. Moreover, Ae_i gives the jth column, and e_iA gives the ith row of A (tl;dr, think of multiplying by zero and one for each respective row and column).

Theorem 19.2 (Theorem 2.12: Distributivity and Scalar Associativity)
Let A be a matrix of size $m \times n$ and B, C be matrices of size $n \times p$. Then

$$A(B + C) = AB + AC,$$

and

$$A(cB) = c(AB)$$

for any choice of scalar $c \in F$.

For the first, for any indices i, j we have

$$A(B + C)_{ij} = \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}) = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj} = AB_{ij} + AC_{ij},$$

and for the second, we may also write

$$A(cB)_{ij} = \sum_{k=1}^{n} a_{ik}(cb_{kj}) = c\sum_{k=1}^{n} a_{ik}b_{kj} = c(AB)_{ij}.$$

Theorem 19.3 (Theorem 2.14: "Extracting the Input")
Let V, W be finite dimensional vector spaces with ordered bases β, γ respectively; moreover, let us have a linear map $T : V \to W$. Then for all $x \in V$, we may write[a]

$$[T(x)]_\gamma = [T]_\beta^\gamma \cdot [x]_\beta.$$

[a] apparently, that one kid who thought $f(x) = f \cdot x$ was right...

Let $S : F \to V$ be a map taking $a \mapsto ax$, with basis $\alpha = \{1\}$ (we want to use S to "pick" out x). To start, we may write

$$[g]_\alpha^\beta = [g(1)]_\beta = [x]_\beta,$$

and so, by theorem 2.11, we may write

$$[T(u)]_\gamma = [T(g(1))]_\gamma = [T \circ g]_\alpha^\gamma = [T]_\beta^\gamma \cdot [g]_\alpha^\beta = [T]_\beta^\gamma \cdot [x]_\beta.$$

§20 Day 16: Online Class; More Matrix Representations (Nov. 2, 2023)

Reminder that matrix multiplication is a binary operation on two matrices A and B where the resultant matrix AB has entries that may be calculated as

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj}$$

where n represents the number of elements in a row in matrix A (or similarly the number of elements in a column in matrix B). More importantly, also recall that given linear maps $S : V \to W$ and $T : W \to Z$ with respective ordered bases of α, β, and γ, we may write

$$[T \circ S]_{\alpha}^{\gamma} = [T]_{\beta}^{\gamma}[S]_{\alpha}^{\beta},$$

which may be seen as $[S]$ transforming α basis vectors into β basis vectors, then $[T]$ doing the same into γ.

§20.1 Matrices as a Linear Transformation

Given any $A \in M_{m\times n}(F)$, we have

$$\begin{aligned} L_A : F^n &\to F^m \\ x &\mapsto A \cdot x \end{aligned} \qquad \text{(Matrix Product)}$$

where we may view the right hand side of the map as multiplying A onto an $n \times 1$ coordinate matrix x; we may check that this is indeed linear by Theorem 2.12 (see previous page).

Theorem 20.1 (Theorem 2.15: Representation of Matrix as Linear Transformation)
Using the definition of L_A as above, we have

$$[L_A]_{\beta}^{\gamma} = A,$$

where β is the standard basis (read: $\{e_1, \ldots, e_n\}$) of F^n and γ is the standard basis of F^m.

To prove this, directly compute $[L_A]_{std}^{std}$[19]; each column is equal to $L_A(e_i)$ by definition, which is equal to Ae_i, which is the ith column of A for all i. As a quick remark, recall the isomorphism theorem (Theorem 2.20; Day 14), and apply it here:

$$\begin{aligned} \Phi : \mathcal{L}(F^n, F^m) &\xrightarrow{\sim} M_{m\times n}(F) \\ T &\mapsto [T]_{std}^{std} \end{aligned}$$

where we may note Φ is isomorphic. By Theorem 2.15, its inverse

$$A \mapsto L_A,$$

is isomorphic (and linear). To check this, simply observe that

$$L_{A+cB} = L_A + cL_B,$$

[19] std represents the standard basis.

and that we also have $L_A \circ L_B = L_{AB}$ by Theorem 2.11, which can be checked by quickly writing

$$[L_A \circ L_B]_{std}^{std} = [L_A]_{std}^{std}[L_B]_{std}^{std} = AB = [L_{AB}]_{std}^{std}.$$

Theorem 20.2 (Theorem 2.16: Associativity of Matrix Multiplication)
Let A, B, C be matrices that can be multiplied, i.e. of dimensions $m \times n$, $n \times p$, and $p \times q$ respectively. Then $(AB)C = A(BC)$.

We can prove this by direct computation, but it is more elegant to apply the following: since the composition of functions is associative, namely

$$L_A \circ (L_B \circ L_C) = (L_A \circ L_B) \circ L_C : F^q \to F^m,$$

we find the matrix representations in terms of the standard bases of F^m, F^n, F^p, F^q, and conclude that $(AB)C = A(BC)$. □

We call a matrix $A \in M_{n\times m}(F)$ invertible if there exists $B \in M_{n\times n}(F)$ such that $AB = BA = I_n$. As usual, the inverse B is unique (if it does exist) and we denote it as A^{-1}.

Theorem 20.3 (Theorem 2.18: Matrix Representation Inverses)
Let V, W be finite dimensional vector spaces with ordered bases β, γ respectively, and let $T : V \to W$ be a linear map. Then T is invertible (as a linear transformation) if and only if $[T]_\beta^\gamma$ is invertible (as a matrix). In particular,

$$[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}.$$

To prove this, start by assuming T is invertible; then we may write

$$T^{-1} \circ T = \mathrm{id}_V \implies [T^{-1} \circ T]_\beta^\beta = [\mathrm{id}_V]_\beta^\beta = I_n,$$

where the left hand side reduces to $[T^{-1}]_\gamma^\beta[T]_\beta^\gamma$ (where $n = |\beta| = \dim V$). Similarly, we may interchange the two matrices to see

$$[T]_\beta^\gamma[T^{-1}]_\gamma^\beta = I_n$$

using the fact that $V \cong W \implies \dim W = n$. For the other direction, if $[T]_\beta^\gamma$ is invertible, let it be A, and recall Theorem 2.20: Let Φ be the isomorphism

$$\begin{aligned} \Phi : \mathcal{L}(W, V) &\xrightarrow{\sim} M_{n\times n}(F) \\ S &\mapsto [S]_\gamma^\beta \end{aligned}$$

which means there necessarily exists a unique $S \in \mathcal{L}(W, V)$ such that $[S]_\gamma^\beta = A^{-1}$. In particular, we may check $S \circ T = \mathrm{id}_V$, and $T \circ S = \mathrm{id}_W$ by observing

$$[S \circ T]_\beta^\beta = [S]_\gamma^\beta[T]_\beta^\gamma = A^{-1}A = I_n = [\mathrm{id}_V]_\beta^\beta,$$

with a similar argument holding for $T \circ S$. With this, we are done. □

§21 Day 17: Change of Basis and Dual Spaces (Nov. 14, 2023)

Recall that $T : V \to W$ is an invertible linear map if there exists a linear map $S : W \to V$ such that $T \circ S = \mathrm{id}_W$ and $S \circ T = \mathrm{id}_V$. Also, note that Theorem 2.18 states the matrix representation itself is invertible as well, i.e. $([T]_\beta^\gamma)^{-1} = [T^{-1}]_\gamma^\beta$ (where V, W have ordered bases β, γ respectively). Moreover, a matrix $A \in M_{n\times n}(F)$ is invertible if and only if there exists a B such that $AB = I_n = BA$ (i.e., the identity matrix of size $n \times n$, denoted I_n).

§21.1 Change of Basis

In general, any vector space has a lot of bases[20], so it's useful to be able to swap matrix representations around whenever necessary. And so with this, we pose two questions:

1. For a choice $x \in V$ with β, β' as given ordered bases for V, what is the relationship between $[x]_\beta$ and $[x]_{\beta'}$?
2. For a given linear map $T : V \to W$ with β, β' and γ, γ' as ordered bases for V and W respectively, what is the relationship between $[T]_\beta^\gamma$ and $[T]_{\beta'}^{\gamma'}$?

To start, recall Theorem 2.14, where we have $[T(x)]_\gamma = [T]_\beta^\gamma [x]_\beta$. Let $\gamma = \beta'$ and we have

$$[x]_{\beta'} = [\mathrm{id}_V]_\beta^{\beta'} \cdot [x]_\beta.$$

This is a change of basis from β to β', and we may note $[\mathrm{id}_V]_\beta^{\beta'}$ is invertible (from Theorem 2.18, since id_V itself is invertible). To check an example, let $\mathbb{R}^2$ have bases $\beta = \{e_1, e_2\}$ and $\beta' = \{(2,1),(1,0)\}$. We see

$$[\mathrm{id}_V]_\beta^{\beta'} = \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix}, \qquad [\mathrm{id}_V]_{\beta'}^{\beta} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix},$$

along with

$$\begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2,$$

which also holds the other way around (check it!). This resolves question 1. As for question 2, recall that given $S : V \to W$ and $T : W \to Z$ with ordered bases α, β, γ for V, W, Z respectively, we have $[T \circ S]_\alpha^\gamma = [T]_\beta^\gamma [S]_\alpha^\beta$; with this, we may describe the following,

$$T = \mathrm{id}_W \circ T \circ \mathrm{id}_V \implies [T]_{\beta'}^{\gamma'} = [\mathrm{id}_W]_\gamma^{\gamma'} [T]_\beta^\gamma [\mathrm{id}_V]_{\beta'}^\beta, \tag{1}$$

since the RHS of the first equation is simply two identities. Moreover, note that both of $[\mathrm{id}_W]_\gamma^{\gamma'}[T]_\beta^\gamma$ and $[T]_\beta^\gamma[\mathrm{id}_V]_{\beta'}^\beta$ act as a change of basis actions. We also have the special case (trivial case) where $T : V \to V$ with ordered basis β of V; we may simply write

$$[T]_\beta = [T]_\beta^\beta.$$

[20] i mean, ofc ignoring the trivial case n stuff

Theorem 21.1 (Theorem 2.23: Change of Basis)

Using $T : V \to W$ (with finite dimensional V, W) as a linear map with ordered bases β, β' of V, we may write

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q,$$

where $Q = [\mathrm{id}_V]^\beta_{\beta'}$. We call Q the change of basis matrix (that sends coordinates in β' to coordinates in β).

To prove this, we can either approach it from the pov of using our earlier equation (1) and letting $\gamma = \beta$, $\gamma' = \beta'$, and so we have

$$[T]^{\beta'}_{\beta'} = [\mathrm{id}_V]^{\beta'}_\beta [T]^\beta_\beta [\mathrm{id}_V]^\beta_{\beta'};$$

where the above may be simplified to the form $[T]_{\beta'} = Q^{-1}[T]_\beta Q$. For intuition, observe that

$$Q[T]_{\beta'} = [\mathrm{id}_V]^\beta_{\beta'}[T]'_\beta = [\mathrm{id}_V \circ T]^\beta_{\beta'} = [T \circ \mathrm{id}_V]^\beta_{\beta'} = [T]_\beta [\mathrm{id}_V]^\beta_{\beta'} = [T]_\beta Q.$$

Also, you may check (with $\beta = \{x_1, \ldots, x_n\}$ and $Q \in M_{n\times n}(F)$), that

$$x'j = \sum_{i=1}^{n} Q_{ij} x_i \text{ for } 1 \le j \le n$$

gives the changed basis $\beta' = \{x'_1, \ldots, x'_n\}$. We say $A, B \in M_{n\times n}(F)$ are similar if there exists an invertible Q such that $B = Q^{-1}AQ$. For example, $[T]_\beta$ and $[T]_{\beta'}$ are similar.

§21.2 Dual Spaces

The dual vector space $V^* = \mathcal{L}(V, F)$ represents the space of linear functionals (linear maps from a vector space to its ground field / "field of scalars"), i.e. all linear maps of the form $f : V \to F$. For example, taking $V = P_n(F)$, we may define a linear functional

$$\begin{aligned} f : P_n(F) &\to F \\ p(x) &\mapsto p(a) \end{aligned}$$

for whatever choice of fixed $a \in F$. Note that if V is finite dimensional, then we may evaluate for the dimension of the dual space,

$$\dim V^* = \dim \mathcal{L}(V, F) = \dim V \cdot \dim F = \dim V,$$

since the scalar field is one dimensional. To find a basis for V^*, suppose we have a given basis $\beta = \{x_1, \ldots, x_n\}$ of V. Define the functions $f_i : V \to F$ where

$$f_i(x_j) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & \text{otherwise} \end{cases},$$

for all $1 \le i \le n$ (note that this is the definition of the Kronecker delta, and I will be using it from now on); moreover, we may ensure f_i are linear maps from Theorem 2.6 (and that they are unique), with $f_i \in V^*$ from definition.

Theorem 21.2 (Theorem 2.24: Basis of V^*)

We will make two claims:

(a) $\beta^* = \{f_1, \dots, f_n\}$ as defined above is a basis of V^*.

(b) We may write, for any $f \in V^*$ in terms of β^* as

$$f = \sum_{i=1}^{n} f(x_i) f_i.$$

Since $\dim V^* = n$ and β' has n linearly independent vectors, we only need to verify the formula for f above (check Steinitz Exchange Cor. 2). Let g be the function

$$g = \sum_{i=1}^{n} f(x_i) f_i.$$

We will now check, according to Thm. 2.6, that this is the unique function satisfying $g(x_j) = f(x_j)$ (by linearity, we only need to match the basis to confirm that they're the same function). We have

$$g(x_j) = \left(\sum_{i=1}^{n} f(x_i) f_i \right)(x_j) = \sum_{i=1}^{n} f(x_i) f_i(x_j) = \sum_{i=1}^{n} f(x_i) \delta_{ij} = f(x_j),$$

since all other terms vanish under the delta. This confirms that β^* cannot generate non-unique functions (and is thus LI) and is thus a basis, and (b) follows automatically.[21]

§21.3 Transposition of Linear Maps

Let $V^* = \mathcal{L}(V, F)$, $W^* = \mathcal{L}(W, F)$, and suppose we have a map $T : V \to W$. Then, we may define the transpose of T to be

$$\begin{aligned} T^t : W^* &\to V^* \\ g &\mapsto g \circ T, \end{aligned}$$

which is illustrated by the commutative diagram, [22]

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ & \searrow^{g \circ T} & \downarrow g \\ & & F \end{array}$$

(and so we may say $f^*(\varphi) = \varphi \circ f$ for all $\varphi \in W^*$, as in $g \circ T = T^t(g)$). We can check that T^t is indeed linear by writing

$$T^t(cf + g) = (cf + g) \circ T = c(f \circ T) + g \circ T = cT^t(f) + T^t(g).$$

This brings us to Theorem 2.25,

[21] herzig proof was the other direction but both work i think

[22] holy shit i figured out how to draw this thing LET'S GOooo

Theorem 21.3 (Theorem 2.25: Transpose of Linear Map)

Let $T : V \to W$ (with V, W finite dimensional) be equipped with β, γ as ordered bases for V and W, and β^*, γ^* as ordered bases for V^* and W^* respectively. Then

$$[T^t]_{\gamma^*}^{\beta^*} = ([T]_\beta^\gamma)^t.$$

To prove this, start by noting $(A^t)_{ij} = A_{ji}$ (recall definition of transpose). For any choice of $g \in W^*$, observe $T^t(g) = g \circ T \in V^*$, and so $T^t : W^* \to V^*$; write $\beta = \{x_1, \ldots, x_n\}$ and $\gamma = \{y_1, \ldots, y_m\}$ along with $\beta^* = \{f_1, \ldots, f_n\}$ and $\gamma^* = \{g_1, \ldots, g_m\}$. For convenience, let us also have $A = [T]_\beta^\gamma$ and $B = [T^t]_{\gamma^*}^{\beta^*}$. To find the jth column of B, we simply have to express $T^t(g_j)$ as a linear combination of β^*, and so by Theorem 2.24, write

$$T^t(g_j) = g_j \circ T = \sum_{\ell=1}^{n} B_{\ell k} f_\ell, \tag{$*$}$$

and so by evaluating $(*)$[23] on x_i, we get

$$\begin{aligned} B_{ik} = g_k(T(x_i)) &= \sum_{\ell=1}^{n} B_{\ell k} f_\ell(x_i) \\ &= \sum_{j=1}^{m} g_k(A_{ji} y_j) \\ &= \sum_{j=1}^{m} A_{ji} g_k(y_j) \\ &= \sum_{j=1}^{m} A_{ji} \delta_{ki}, \end{aligned}$$

and so the above evaluates out to be A_{ki}. Since $A_{ki} = B_{ik}$, we see $B = A^t$, and the proof is complete. As an exercise, verify from the proof that T^t is indeed unique. □

[23] i will fix this garbage later my god the asterisk in the display above WILL NOT WORK THIS IS RIDICULOUS *scream*

§22 Day 18: Elementary Row Operations (Nov. 16, 2023)

We now begin chapter 3. In order to solve systems of linear equations, say, suppose,

$$\begin{cases} y - z & = 2 \\ 3x + y + z & = 1 \\ 2x + 3y - z & = 0 \end{cases}$$

with x, y, z unknown, observe that the solutions x, y, z don't change if we

1. Interchange two equations.
2. Multiply one equation by a nonzero scalar.
3. Add a multiple of one equation to another equation.

Note that these three operations are called the elementary row operations, and each of them are reversible (we may label these **R1**, **R2**, **R3** respectively). It is useful to write the above system as

$$\begin{pmatrix} 0 & 1 & -1 \\ 3 & 1 & 1 \\ 2 & 3 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}.$$

Theorem 22.1 (Theorem 3.1/3.2: Elementary Row Operations)
If a matrix $A' \in M_{m\times n}(F)$ is obtained by a series of elementary row operations, then there exists a matrix E such that $A' = EA$. In fact, E is obtained from elementary row operations on I_m. This matrix E is invertible.

To prove this, verify that it works for each of **R1**, **R2**, and **R3** (leaving it as an exercise, omitting for brevity). In a similar fashion, we may define the elementary column operations **C1**, **C2**, and **C3** of which Thm. 3.1 and 3.2 hold as well.

§22.1 Rank of a Matrix (Section 3.2)

If $A \in M_{m\times n}(F)$ is a matrix, then we define $\operatorname{rank} A := \operatorname{rank} L_A$ with

$$\begin{aligned} L_A : F^n &\to F^m \\ x &\mapsto Ax. \end{aligned}$$

By the rank-nullity theorem, we also have $\operatorname{null} A = \operatorname{null} L_A$ from $\operatorname{rank} A + \operatorname{null} A = n = \dim F^n$ (as in, the number of columns). In particular, we may write

$$\operatorname{Im}(L_A) = \{L_A(x) \mid x \in F^n\} = \operatorname{span}\{L_A(e_1), \ldots, L_A(e_n)\},$$

which translates to the span of the columns of A. Moreover, we may derive

Theorem 22.2 (Theorem 3.5: Rank of a Matrix)
The rank of a matrix $A \in M_{m\times n}(F)$ is equal to the dimension of its span of columns, which is equal to the maximum number of linearly independent columns in A.

For example, if we take $A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 0 \end{pmatrix}$, we see its rank is equal to 2 (since the first and third columns are linearly independent).

Theorem 22.3 (Theorem 3.7: Rank Inequality)

For any two compatible matrices A, B, we have

$$\operatorname{rank}(AB) \leq \operatorname{rank} A \text{ and } \operatorname{rank} B.$$

To prove this, directly compute

$$\operatorname{rank}(AB) = \operatorname{rank}(L_{AB}) = \operatorname{rank}(L_A \circ L_B),$$

which is enough to show that if we take linear transformations $S : V \to W$ and $T : W \to Z$, then

$$\operatorname{rank}(T \circ S) \leq \operatorname{rank} T \text{ and } \operatorname{rank} S,$$

which immediately follows from the property that $\operatorname{Im}(T \circ S) \subset \operatorname{Im}(T)$ and $\ker(T \circ S) \supset \ker(S)$ implies $\operatorname{rank}(T \circ S) \leq \operatorname{rank} T$ along with $\operatorname{null}(T \circ S) \geq \operatorname{null} T$, with the same conclusion on $\operatorname{rank} S$ and $\operatorname{null} S$.

We may also check that $\ker L_A$ is unchanged under elementary row operations, and that $\operatorname{Im} L_A$ is unchanged under elementary column operations.

§23 Day 19: tbd (Nov. 21, 2023)

§24 Day 20: tbd (Nov. 23, 2023)

§25 Day 21: Determinant (Nov. 28, 2023)

Decided to let my friend try writing notes (so I'll be kinda, yk, fixing here and there). When a matrix is taken in reduced row-echelon form, we can easily measure the number of pivots it has (i.e., a pivot is the first nonzero entry in each row, of which the columns they are in are called pivot columns). It is quick to check that

$$\operatorname{rank}(A) = \#\text{ of Pivot columns} = \#\text{ of nonzero rows},$$

which is given by an example below

$$\begin{pmatrix} 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

If the given matrix is finite, we may apply the rank-nullity theorem to count $\operatorname{null} A = n - \operatorname{rank} A$, which is represented by the number of non-pivot columns. We can use this technique to find a basis for $\ker L_A$; start by solving the system $Ax = 0$, which, given by the above example, has pivots

$$\begin{aligned} x_2 &= -2x_3 \\ x_4 &= 0 \\ x_5 &= 0 \end{aligned}$$

and non-pivots

$$\begin{aligned} x_1 &= t_1 \\ x_3 &= t_2, \end{aligned}$$

which we assign arbitrary scalars to. Using this, we have that the solutions to the kernel is given by

$$x = t_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

It is clear that this linear combination spans $\ker L_A$, since they are necessarily linearly independent and there are two of them (which is equal to $\operatorname{null} A$). In a similar manner, we may find the basis of $\operatorname{Im}(L_A)$ (i.e., the column span). To do this, we simply have to find a maximal linearly independent subset of A; let A' be the RREF of A. For example,

$$A = \begin{pmatrix} 0 & 0 & 2 & 2 \\ -1 & 3 & 2 & 4 \\ 2 & -6 & 4 & 0 \end{pmatrix} \implies A' = \begin{pmatrix} 1 & -3 & 0 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Observe that the pivot columns here are the first and third (since there are two of them, they form a basis as well); however, let us proceed with a more general result instead. Given $v_1, \ldots, v_n$ as columns of A with $v_1', \ldots, v_n'$ being columns for A', notice that if there exists a linear combination with scalars $c_1, \ldots, c_n$ such that $c_1 v_1 + \cdots + c_n v_n = 0$, the same scalars yield $c_1 v_1' + \cdots + c_n v_n' = 0$. This comes from the fact that

$$A \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = 0 \iff \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \ker L_A = \ker L_{A'}$$

since A and A' share the same kernel (unchanged under row operations). Equivalently, we have

$$\iff A' \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \iff c_1 v_1' + \cdots + c_n v_n' = 0.$$

Thus, there must exist some sequence of elementary column operations, and this concludes the claim.

Lemma 25.1 (RREF is Unique)

Using the above, we can prove that the RREF of any given matrix is necessarily unique.[a]

[a] i'm not actually sure how to go about this... i'll try later.

In a similar manner, given a subset $S \subset F^n$, let A be the matrix with columns $(x_1 \mid \cdots \mid x_n)$. Then $\operatorname{span} S$ is equal to the column span of A, which lets us find a maximally linearly independent subset for S.

§25.1 Determinants (Section 4)

To start, let $A \in M_{2\times 2}(\mathbb{R})$. We define the determinant of A to be

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Theorem 25.2 (Theorem 4.2: Invertibility from Determinant)

If $\det A \neq 0$, then A is invertible.

In particular, for the 2×2 case, we may write

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

To prove the above theorem, start by some direct computation.

$$A \begin{pmatrix} d & b \\ -c & a \end{pmatrix} = \begin{pmatrix} \det A & 0 \\ 0 & \det A \end{pmatrix} = (\det A) I_2.$$

Now, suppose A is invertible (we may let A be any $n \times n$ matrix, actually). If $\det A = 0$, then let B be such that $AB = BA = I_n$. Then $\det(AB) = (\det A)(\det B) = 0$, but $\det I_n = 1$, which is a contradiction. A similar proof goes for the other direction.[24] For a version without cauchy-binet (as in, done in class), notice if $\det(A) \neq 0$, then

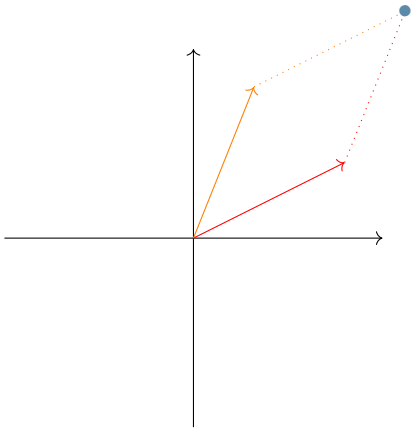
$$\frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

is the intended formula, which works as checked earlier; for the other direction, set $AB = 0$, then write $B = AA^{-1}B = 0$. This gives $a = b = c = d = 0 \implies A = 0$, which is not invertible; therefore $\det A \neq 0$.

[24] i'd assume we'll go over cauchy-binet at some point, which is $\det AB = \det A \det B$.

§25.2 Interpretation of Determinant

Another thing of note is that $\det A$ represents the signed area of a parallelogram (back to the $A \in M_{2\times 2}(\mathbb{R})$ case),



where the above parallelogram is spanned by the columns of A, as in $\begin{pmatrix} a \\ b \end{pmatrix}$ and $\begin{pmatrix} c \\ d \end{pmatrix}$ respectively. If we let α be the angle between the two vectors, then the area of the parallelogram is said to be positive if $\alpha \in (0, \pi)$ and negative if $\alpha \in (\pi, 2\pi)$. Also, we may write det as a function $\det : F^2 \times F^2 \to F$, with $\det(v_1, v_2) = \det(v_1 \mid v_2)$ (as in a 2×2 matrix).

Theorem 25.3 (Theorem 4.1 (ish): Determinant is Bilinear)
Let $\det : F^2 \times F^2 \to F$; we have that det is bilinear (as in, linear in each argument).

To check this, observe

$$\det(e_1, e_2) = \det \left(\begin{array}{c|c} 1 & 0 \\ 0 & 1 \end{array} \right) = 1$$
$$\det \left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} b \\ a \end{pmatrix} \right) = 0.$$

To see it is linear in the first argument, fix $v_2 = \begin{pmatrix} b \\ d \end{pmatrix} \in F^2$. Then

$$\det(v_1 + \lambda v_1', v_2) = \det(v_1, v_2) + \lambda \det(v_1', v_2)$$

comes from the following computation,

$$\begin{aligned} \det(v_1 + \lambda v_1', v_2) &= \det \left(\begin{array}{c|c} a + \lambda a_1 & b \\ c + \lambda c' & d \end{array} \right) \\ &= (a + \lambda a')d - b(c + \lambda c') \\ &= \det(v_1, v_2) + \lambda \det(v_1', v_2). \end{aligned}$$

The proof is similar for the second argument. □

Now, to generalize this theorem to the $n \times n$ case, let us define the function $\delta : F^n \times F^n \times \cdots \times F^n \to F$, and check that it is multilinear (or n-linear) in each argument, i.e.

$$F^n \to F \text{ is linear for all } i$$
$$v_i \mapsto \delta(v_1, \ldots, v_n).$$

First, claim that δ is alternating[25] if $\delta(v_1, \dots, v_n) = 0$ whenever there exists $v_i = v_j$ for some $i \neq j$. If δ is alternating, then for any $i < j$, claim that

$$\delta(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = \delta(v_1, \dots, v_j, \dots, v_i, \dots, v_n)$$

(i.e., swapping two inputs). To prove this, start in the case $n = 2$. Then

$$\begin{aligned} 0 &= \delta(v_1 + v_2, v_1 + v_2) \\ &= \delta(v_1, v_1 + v_2) + (v_2, v_1 + v_2) \\ &= \delta(v_1, v_1) + \delta(v_1, v_2) + \delta(v_2, v_1) + \delta(v_2, v_2), \end{aligned}$$

where the first and fourth terms vanish. This yields

$$\delta(v_1, v_2) = -\delta(v_2, v_1),$$

and we see δ is indeed alternating. With this in place, we now claim there is a unique function $\delta : F^n \times \cdots \times F^n \to F$ that is multilinear, alternating, and satisfies $\delta(e_1, \dots, e_n) = 1$. Suppose such a δ exists; observe

$$\delta(e_{i_1}, \dots, e_{i_n}),$$

for $i_1, \dots, i_n \in \{1, \dots, n\}$. If there exists $i_j = i_k$ with $j \neq k$, then we know the above must be equal to zero as it is alternating; otherwise, all the i_j's are pairwise distinct, and we see we have a permutation $\{1, \dots, n\} \to \{1, \dots, n\}$ such that $j \mapsto i_j$ is bijective. *proof was left off here, will be continued next class I think.*

Actually let me give this a shot

[25] which means if we switch two of the variables, we flip the sign. alternatively, the following definition works too

§26 Day 22: Deriving the Determinant (Nov. 30, 2023)

Recall that the determinant is a functional, and in the 2×2 case it operates as such,

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc,$$

where in the 2×2 case, we have $\det M_{2\times 2}(F) \to F$ (i.e. $F^2 \times F^2 \to F$). Recall that the determinant is linear in each argument, i.e. "bilinear" or "multilinear" in the higher dimensional case.

Theorem 26.1 (Determinant Definition)
There is a unique function $\delta : \underbrace{F^n \times \cdots \times F^n}_{n \text{ times}} \to F$ that is multilinear, alternating, and satisfies $\delta(e_1, \ldots, e_n) = 1$. This function δ is called the determinant, det.

Let v_i be columns vectors of size $n \times 1$ defined as such,

$$v_i = \begin{pmatrix} A_{1i} \\ \vdots \\ A_{ni} \end{pmatrix} = \sum_{j=1}^{n} A_{ji} e_j,$$

for $1 \leq i \leq n$. Start by noticing that

$$\delta_{e_{i_1}, \ldots, e_{i_n}} = 0$$

if $i_j = i_k$ for some $j \neq k$; i.e., in order for the above to be equal to 1 and not vanish under the alternating condition, we must have $(e_{\sigma(1)}, \ldots, e_{\sigma(n)})$ be ordered with $\sigma \in S_n$ a permutation (read: S_n as in the symmetric group of size n). Then

$$\delta(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = \operatorname{sgn}(\sigma) = (-1)^s,$$

where s is the number of swaps needed to permute $\{1, \ldots, n\}$ into $\{\sigma(1), \ldots, \sigma(n)\}$. Back to the more general case; observe that

$$\begin{aligned} &\delta(v_1, \ldots, v_n) \\ &= \delta\left(\sum_{j_1=1}^{n} A_{j_1 1} e_{j_1}, \ldots, \sum_{j_n=1}^{n} A_{j_n 1} e_{j_n} \right) \\ &= \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} A_{j_1 1} \ldots A_{j_n 1} \delta(e_{j_1}, \ldots, e_{j_n}) \qquad \text{(Multilinearity)} \end{aligned}$$

Notice that $\delta(e_{j_1}, \ldots, e_{j_n})$ is nonzero if and only if $j_1, \ldots, j_n$ forms a permutation of $\{1, \ldots, n\}$. Thus, we may rewrite the above to get

$$\sum_{\sigma \in S_n} A_{\sigma(1),1} \ldots A_{\sigma(n),n} \underbrace{\delta(e_{\sigma(1)}, \ldots, e_{\sigma(n)})}_{\operatorname{sgn}(\sigma)}.$$

Now, it only remains to check that this formula works.

- It is multilinear (i'll prove this when it's not like 2am or something)

- For alternating, check that if $v_k = v_\ell$ for some $k < \ell$, then $\delta(v_1, \dots, v_n) = 0$. This comes from the fact that for any even permutation, we may swap v_k and v_ℓ to get an odd permutation that yields the same sum (but swapped sign), i.e.

$$A_{\sigma(1),1} \dots A_{\sigma(k),k} \dots A_{\sigma(\ell),\ell} \dots A_{\sigma(n),n} = A_{\sigma(1),1} \dots A_{\sigma(\ell),k} \dots A_{\sigma(k),\ell} \dots A_{\sigma(n),n},$$

  since $v_k = v_\ell$, so the permutation does not affect it.

- Finally,

$$\delta(e_1, \dots, e_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{\sigma(1),1} \dots A_{\sigma(n),n} = \operatorname{sgn}(\mathrm{id}) \cdot 1 \cdot \dots \cdot 1 = 1.$$

As a quick lemma, if A is upper-triangular, i.e. $A_{ij} = 0$ for all $i > j$, then $\det A = A_{11} \cdot A_{22} \cdot \dots \cdot A_{nn}$. This is quick to see from the fact that $\sigma(1) \leq 1, \dots, \sigma(n) \leq n$, implying the only valid permutation being $\sigma = \mathrm{id}$.

§27 Day 23: More on Determinants (Dec. 5, 2023)

Recall from last time that we proved the existence and uniqueness of a function δ satisfying

$$\delta : \underbrace{F^n \times \cdots \times F^n}_{n \text{ times}} \to F,$$

that is multilinear, alternating, and takes the standard bases to 1 (i.e. $\delta(e_1, \ldots, e_n) = 1$). δ takes on the general formula

$$\delta(v_1, \ldots, v_n) = \det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{\sigma(1),1} \ldots A_{\sigma(n),n},$$

with A defined as the $n \times n$ matrix with columns $v_1, \ldots, v_n$. As a corollary, the summation above also satisfies

$$\det \begin{pmatrix} A_{11} & \ldots & * \\ \vdots & \ddots & \vdots \\ 0 & \ldots & A_{nn} \end{pmatrix} = A_{11} \cdot \cdots \cdot A_{nn}.$$

Theorem 27.1 (Determinant Properties)

Using properties of the determinant, we have a few relations:

(a) $\det A' = \begin{cases} -\det A & \text{if } A' \text{ is obtained by swapping two rows or columns of } A, \\ c \cdot \det A & \text{if a column or a row is scaled by } c \in F, \\ \det A & \text{if we add a multiple of a row or column to another.} \end{cases}$

(b) $\det A^t = \det A$,

(c) $\det A \neq 0$ if and only if A is invertible.

To check this, for (a) on column operations, each result is immediately given from the alternating and multlinear properties (respectively); for **(C3)**, if we add c times the jth column to the ith column, then apply linearity in the ith column to get,

$$\begin{aligned} \det(v_1, \ldots, v_i + cv_j, \ldots, v_j, \ldots, v_n) &= \det A + c \underbrace{\det(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)}_{=0 \text{ by alternating}} \\ &= \det A. \end{aligned}$$

For (b), we have

$$\begin{aligned} \det A^t &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \underbrace{(A^t)_{\sigma(1),1}}_{A_{1,\sigma(1)}} \cdots \underbrace{(A^t)_{\sigma(n),n}}_{A_{n,\sigma(n)}} \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) A_{\sigma^{-1}(1),1} \ldots A_{\sigma^{-1}(n),n}. \qquad (\text{"Un-permute" the indices with } \sigma^{-1}) \end{aligned}$$

Since $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$ (number of swaps to get from $\{1, \ldots, n\}$ to $\{\sigma(1), \ldots, \sigma(n)\}$ is the same in either direction), we see the summation above simply gets us $\det A$ again. To prove (a), for rows, simply let $v_1, \ldots, v_n$ be the columns of A, and $r_1, \ldots, r_n$ be the rows of A^t. Then clearly $v_i = r_i$ for $1 \leq i \leq n$ (check: transpose), and so we may proceed from there.

For part (c), let A' be the RREF of A. If A is invertible, then $A' = I_n$ and thus A cannot have zero determinant by (a); the same argument holds for the opposite direction.

To evaluate $\det A$ quickly for any given matrix, use **(R1)** and **(R3)** to put it into upper triangular form so we can simply multiply down the main diagonal. For example,

$$\begin{aligned}\det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{pmatrix} &= \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 3 \end{pmatrix} \\ &= \det \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & -1 \end{pmatrix} = -1.\end{aligned}$$

Theorem 27.2 (Cauchy-Binet Formula; Product of Determinant)

Given two square matrices $A, B \in M_{n\times n}(\mathbb{R})$[a], we have

$$\det(AB) = (\det A)(\det B).$$

[a] pretty sure this is valid for any matrix with values in a commutative ring

To prove this, first assume $\det A \neq 0$, and consider the function

$$\begin{aligned}\delta' : F^n \times \cdots \times F^n &\to F, \\ (x_1, \ldots, x_n) &\mapsto \frac{\det(Ax_1, \ldots, Ax_n)}{\det A}.\end{aligned}$$

Clearly, δ' is multilinear and alternating, and we may see

$$\delta'(e_1, \ldots, e_n) = \frac{\det(Ae_1, \ldots, Ae_n)}{\det A} = 1,$$

since both numerator and denominator are taking the determinant of the same thing. Thus, $\delta' = \det$, and let $\delta'(v_1, \ldots, v_n) = \det B$ if $v_1, \ldots, v_n$ are columns of B. Then by definition,

$$\delta'(v_1, \ldots, v_n) = \frac{\det(Av_1, \ldots, Av_n)}{\det A} = \frac{\det(\text{Col. of } AB)}{\det A} = \frac{\det AB}{\det A},$$

which gives $\det AB = \det A \det B$. For the other case, if $\det A = 0$, then simply invoke $\operatorname{rank} A < n \implies \operatorname{rank} AB \leq \operatorname{rank} A < n$, so $\det AB = 0$ (by linear dependence). This concludes the proof. □

Theorem 27.3 (Determinant of Block Upper-Triangular Matrices)

If A is a square matrix of the form

$$A = \begin{pmatrix} A' & X \\ 0 & A'' \end{pmatrix}$$

with A', A'' being square matrices and X any matrix, then $\det A = \det A' \det A''$.

Quick remark that this doesn't just hold for A' and A'' blocks; the number of blocks can be as many as wanted by induction; for example, with blocks B, C, D, we have

$$\det \begin{pmatrix} B & X & Y \\ 0 & C & Z \\ 0 & 0 & D \end{pmatrix} = (\det B) \left(\det \begin{pmatrix} C & Z \\ 0 & D \end{pmatrix} \right) = (\det B)(\det C)(\det D),$$

so it suffices to prove it for the 2×2 case. Using A from the theorem statement, observe

$$\det A = \left(\det \begin{pmatrix} I_m & 0 \\ 0 & A'' \end{pmatrix}\right)\left(\det \begin{pmatrix} A' & X \\ 0 & I_{n-m} \end{pmatrix}\right) = \left(\det \begin{pmatrix} I_m & 0 \\ 0 & A'' \end{pmatrix}\right)\left(\det \begin{pmatrix} A' & 0 \\ 0 & I_{n-m} \end{pmatrix}\right),$$

with the last step obtained by using **(R3)** to eliminate X. Let B be the latter matrix; then $B_{ij} = \delta_{ij}$, hence

$$\begin{aligned} \det B &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) B_{\sigma(1),1} \dots B_{\sigma(m),m} \underbrace{B_{\sigma(m+1),m+1}, \dots, B_{\sigma(n),n}}_{=0 \text{ unless } \sigma(m+1)=m+1,\dots,\sigma(n)=n}, \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) B_{\sigma(1),1} \dots B_{\sigma(m),m} = \det A'. \end{aligned}$$

In a similar fashion, we have $\det \begin{pmatrix} I_m & 0 \\ 0 & A'' \end{pmatrix} = \det A''$.

§27.1 Cofactor Expansion Along a Column (Laplace Expansion)

For a square matrix A, we may inductively compute its determinant as follows (by going along its ith column),

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \cdot \underbrace{\det(\tilde{A}^{(i,j)})}_{\text{"minor"}}.$$

Note that $\tilde{A}^{(i,j)}$ represents the matrix A after removing its ith row and jth column, (so it goes from a $n \times n$ matrix to a $n-1$ by $n-1$ matrix). For example, let us expand along the 2nd column in the following matrix,

$$\begin{aligned} \det \begin{pmatrix} a & 0 & b \\ c & 1 & d \\ e & 0 & f \end{pmatrix} &= (-1)^{1+2} A_{12} \det \begin{pmatrix} c & d \\ e & f \end{pmatrix} + (-1)^{2+2} A_{22} \det \begin{pmatrix} a & b \\ e & f \end{pmatrix} \\ &\quad + (-1)^{3+2} A_{32} \det \begin{pmatrix} a & c \\ b & d \end{pmatrix} \\ &= \det \begin{pmatrix} a & b \\ e & f \end{pmatrix} = af - be, \end{aligned}$$

since $A_{12} = A_{32} = 0$. Do note that there is another different way to compute the determinant here as well; using elementary operations, we see

$$\det \begin{pmatrix} a & 0 & b \\ c & 1 & d \\ e & 0 & f \end{pmatrix} = \begin{pmatrix} 1 & c & d \\ 0 & a & b \\ 0 & e & f \end{pmatrix} = \det(1) \det \begin{pmatrix} a & b \\ e & f \end{pmatrix} = af - be$$

from the block matrix technique earlier. Proof: tbd later (or Meinrenken notes)

§27.2 Cramer's Rule

Theorem 27.4 (Cramer's Rule)

Let A be an invertible matrix. Then the unique solution $Ax = b$ (with x, b being column matrices of size $n \times 1$) is given by

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad x_i = \frac{\det(v_1, \dots, b, \dots, v_n)}{\det A},$$

i.e. the ith column of A is replaced with b in the numerator.

To prove this, let

$$Ax = (v_1 \mid \cdots \mid v_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = b \iff b = x_1 v_1 + \cdots + x_n v_n.$$

This means

$$\begin{aligned} \det(v_1, \dots, b, \dots, v_n) &= \det(v_1, \dots, x_1 v_1 + \cdots + x_n v_n, \dots, v_n) \\ &= \sum_{k=1}^{n} x_k \underbrace{\det(v_1, \dots, v_k, \dots, v_n)}_{=0 \text{ by alternating if } i \neq k} \\ &= x_i \det(x_1, \dots, x_i, \dots, x_k) \\ &= x_i \det A. \quad \square \end{aligned}$$

As a quick remark, this allows us to get a formula for A^{-1}. Assuming that A is invertible, we would have $AA^{-1} = I_n$. Then by Cramer's rule,

$$(A^{-1})_{ij} = \frac{\det(v_1, \dots, e_j, \dots, v_n)}{\det A} = \frac{(-1)^{i+j} \det(\tilde{A}^{(ji)})}{\det A}, \qquad (e_j \text{ replaces } v_i)$$

by expanding along the jth column.