Hello and welcome to class!

Last time

We saw the definition of vector space, and subspace. We saw many examples: $\mathbb{R}^n$, its subspaces, spaces of functions.

As you learned, it somewhat harder to figure out whether a given subset of a function space is a subspace than with the same question for subsets of $\mathbb{R}^n$.

That's not surprising: thinking of function spaces as vector spaces is a new concept; it takes some time to get used to it.

Hello and welcome to class!

This time

We study the notions of linear independence, spanning sets, and bases in the context of general vector spaces.

In particular, we emphasize how these notions serve to give coordinates to abstract, otherwise unfamiliar vector spaces — e.g. subspaces of function spaces.

This will allow us to pretend that these vector spaces are just $\mathbb{R}^n$.

Review: the definition of a vector space.

Definition

The data of an $\mathbb{R}$ vector space is a set V, equipped with a distinguished element $\mathbf{0} \in V$ and two maps

$$+ : V \times V \to V \qquad \cdot : \mathbb{R} \times V \to V$$

This data determines a vector space if it obeys the following rules.

$$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w} \qquad \mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v} \qquad \mathbf{v} + \mathbf{0} = \mathbf{v}$$

$$c \cdot (d \cdot \mathbf{v}) = (cd) \cdot \mathbf{v} \qquad 1 \cdot \mathbf{v} = \mathbf{v} \qquad 0 \cdot \mathbf{v} = \mathbf{0}$$

$$(c + d)\mathbf{v} = c\mathbf{v} + d\mathbf{v} \qquad c(\mathbf{v} + \mathbf{w}) = c\mathbf{v} + c\mathbf{w}$$

Review: the definition of a subspace.

Definition

If V is a vector space, we say a subset W of V is a subspace if:

- Zero is in W.
- The sum of any two elements in W is in W.
- Any scalar multiple of an element in W is in W.

Fact

A subspace is itself a vector space.

Polynomials

Last time we learned that functions $\mathbb{R} \to \mathbb{R}$ formed a vector space and observed that polynomial functions formed a subspace.

In particular, polynomial functions form a vector space.

Following the notation of the book, we write:

$\mathbb{P}$	all polynomials
$\mathbb{P}_n$	polynomials of degree at most n

Polynomials

You add them like this:

$$
\begin{array}{rrrrrrrrrr}
 & 5x^4 & + & 4x^3 & + & 3x^2 & + & 2x & + & 1 \\
+ & & & 2x^3 & + & 4x^2 & + & 8x & + & 16 \\
\hline
 & 5x^4 & + & 6x^3 & + & 7x^2 & + & 10x & + & 17
\end{array}
$$

You scalar multiply them like this:

$$7 \cdot (x^2 + 2x + 1) = 7x^2 + 14x + 2$$

And there's a zero polynomial.

Review: linear combinations and linear span

Given a vector space V and $\mathbf{v}_1, \dots \mathbf{v}_n \in V$ and $c_1, \dots, c_n \in R$, we can name an element

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \in V$$

This is said to be a linear combination of the $\mathbf{v}_i$.

The set of linear combinations of the $\mathbf{v}_i$ is the linear span of the $\mathbf{v}_i$.

Linear spans are subspaces.

Spanning sets

We can turn this around and ask:

Given a vector space V,
find a collection of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots$
such that every element of V
is a linear combination of the $\mathbf{v}_i$.

In other words, find a collection of vectors which span V.

Such a collection is called a spanning set.

Review: Does it span $\mathbb{R}^2$?

$(0, 1)$ no

$(1, 0), (0, 1)$ yes

$(2, 3), (4, 5)$ yes

$(2, 3), (4, 6)$ no

$(2, 3), (4, 6), (1, 0)$ yes

Does it span $\mathbb{P}_2$?

$x^3 + 3x^2 + 3x + 1,\ x^2 + 2x + 1,\ x + 1,\ 1$
the first one isn't even in $\mathbb{P}_2$

1 no

$1, x$ no

$1, x, x^2$ yes
every polynomial of degree ≤ 2 can be written as $ax^2 + bx + c$

$1, x + 1, x^2 + 2x + 1$ yes

$1 + 2x + 3x^2, 4 + 5x + 6x^2, 7 + 8x + 9x^2$ no
If you don't see why now, we'll discuss it in a little bit.

What spans $\mathbb{P}$?

No finite set spans $\mathbb{P}$.

Indeed, given any finite collection of polynomials, they have some maximal degree, say n. Then no linear combination of them has degree greater than n: so they do not span $\mathbb{P}$.

What spans $\mathbb{P}$?

There's an infinite set that spans $\mathbb{P}$: all elements of $\mathbb{P}$.

One can do better: the set $\{1, x, x^2, \ldots\}$ spans $\mathbb{P}$.

Linear indepedendence

Definition

Vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots\}$ in a vector space V are linearly indepedendent if none is a linear combination of the others.

Equivalently, if, whenever $\sum_i c_i \mathbf{v}_i = 0$ for some constants $c_i \in \mathbb{R}$, all the c_i must be zero.

We already met this notion for vectors in $\mathbb{R}^n$.

Linear indepedendence of a single vector

The set $\{\mathbf{v}\}$ is always linearly independent unless $\mathbf{v} = 0$.

Indeed, suppose $\{\mathbf{v}\}$ is linearly dependent. This means that there's some $a \neq 0$ such tha $\mathbf{0} = a \cdot \mathbf{v}$. Multiplying by $\frac{1}{a}$ and expanding according to the axioms, we have on the one hand, that anything times the zero vector is zero:

$$\frac{1}{a} \cdot \mathbf{0} = \frac{1}{a} \cdot (0 \cdot \mathbf{0}) = (\frac{1}{a} \cdot 0) \cdot \mathbf{0} = 0 \cdot \mathbf{0} = \mathbf{0}$$

and on the other hand,

$$\frac{1}{a} \cdot (a \cdot \mathbf{v}) = (\frac{1}{a} \cdot a) \cdot \mathbf{v} = 1 \cdot \mathbf{v} = \mathbf{v}$$

So $\mathbf{v} = 0$. (We already saw this argument in $\mathbb{R}^n$.)

Linear indepedendence of two vectors

The set $\{\mathbf{v}, \mathbf{w}\}$ is linearly independent unless one vector is a multiple of the other.

Indeed, suppose $\{\mathbf{v}, \mathbf{w}\}$ is linearly dependent. This means that there's some a, b not both zero such that $\mathbf{0} = a \cdot \mathbf{v} + b \cdot \mathbf{w}$

If a is zero, then $b \cdot \mathbf{w} = \mathbf{0}$ and b is not zero; by the previous slide $\mathbf{w} = \mathbf{0} = 0 \cdot \mathbf{v}$.

If on the other hand a is not zero then we can — using the vector space axioms —rearrange the above equation to

$$\mathbf{v} = -\frac{b}{a} \cdot \mathbf{w}$$

(We already saw this argument in $\mathbb{R}^n$.)

Is it linearly independent?

The subset $1, x^2, x^5$ of $\mathbb{P}_5$? yes:

Consider a linear combination $a \cdot 1 + b \cdot x^2 + c \cdot x^5 = a + bx^2 + cx^5$.
This is only the zero polynomial if $a = b = c = 0$.

The subset $(x+1)^2, (x-1)^2, x$ of $\mathbb{P}$? no:

$$(x+1)^2 - (x-1)^2 + 4 \cdot x = \mathbf{0}$$

Is it linearly independent?

The subset $\{e^x, e^{2x}\}$ of the space of functions? yes:

Suppose one was a multiple of the other, say $e^x = ce^{2x}$. Then we would have $c = e^{-x}$ for some constant c, which is not true.

The subset $\{\sin(x)^2, \cos(x)^2, 1\}$ of the space of functions? no:

$$\sin(x)^2 + \cos(x)^2 = 1$$

Is it linearly independent?

The subset $\{e^x, e^{2x}, e^{3x}\}$ of the space of functions? yes

You won't have to do this kind of thing until we start talking about differential equations, but here's a hint of things to come:

Suppose $ae^x + be^{2x} + ce^{3x} = 0$. If any coefficient is zero, we can use the method of last slide. So let's assume this isn't the case.

We take a derivative and get $ae^x + 2be^{2x} + 3ce^{3x} = 0$. Subtracting the first equation from this, we see $be^{2x} + 2ce^{3x} = 0$. Rearranging, we learn that $e^x = -b/2c$, which is a contradiction.

This sort of interleaving of differential calculus and linear algebra will characterize the third part of this class.

Bases

Definition

A subset $\{\mathbf{v}_1, \mathbf{v}_2, \ldots\}$ of a vector space V is a basis for V if it is linearly independent and spans V

Example

As we know, the vectors $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$, which are a basis for $\mathbb{R}^3$.

Example

$\{1, x, x^2\}$ is a basis for $\mathbb{P}_2$. They span: every polynomial of degree at most 2 is some $ax^2 + bx + c$. And they're linearly independent: $ax^2 + bx + c$ is the zero polynomial only if a, b, c are all zero.

Unique representation

Another way to write the definition: a subset $\{\mathbf{v}_1, \mathbf{v}_2, \ldots\}$ of a vector space V is a basis for V if there's one (spanning) and only one (linear independence) way to write any element $\mathbf{v} \in V$ as a linear combination of the $\mathbf{v}_i$.

Bases from spanning sets

Given a finite spanning set $\{\mathbf{v}_i\}$ of V, one can find a basis by iteratively throwing out vectors in the linear span of the others.

Indeed, this procedure does not change the linear span of all the vectors. On the other hand, it must terminate, since the number of vectors decreases each time, and can only terminate when the remaining vectors are linearly independent.

(We already saw this argument when finding a basis for the column space of a matrix)

Example

Consider the polynomials

$$x+1, x^2+2x+1, 4, 4x+3, x^2$$

Let's find, from among them, a basis for whatever space they span.

- $x^2+2x+1 = (x^2)+2\cdot(x+1)+\frac{1}{4}\cdot 4$. So, let's throw it out. We're left with

$$x+1, 4, 4x+3, x^2$$

- $4x+3 = 4\cdot(x+1)-\frac{1}{4}\cdot 4$. So throw it out. We're left with

$$x+1, 4, x^2$$

- These are linearly independent.

Let’s see how these notions interact with linear transformations.

Review: Linear transformations

Definition

If V and W are vector spaces, a function $T : V \to W$ is said to be a linear transformation if

$$T(c\mathbf{v} + c'\mathbf{v}') = cT(\mathbf{v}) + c'T(\mathbf{v}')$$

for all c, c' in $\mathbb{R}$ and all $\mathbf{v}, \mathbf{v}'$ in V.

Review: one-to-one and onto

A function $f : X \to Y$ is said to be:

- onto if every element of Y is f of ≥ 1 element of X.
- one-to-one if every element of Y is f of ≤ 1 element of X.

For a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$, we know that it is one-to-one if and only if the columns of the associated matrix are linearly independent, and onto if and only if the columns of the associated matrix span the codomain.

The identity function

For a set X, there's a function from X to itself which does nothing.

$$id_X : X \mapsto X$$
$$x \mapsto x$$

When X is a vector space, id_X is a linear transformation.
(Do you see why?)

When $X = \mathbb{R}^n$, the matrix of id_X is the identity matrix.

Invertible functions

A function $f : X \to Y$ is said to be invertible if there's another function which undoes it, i.e., some $g : Y \to X$ with $g \circ f = id_X$ and $f \circ g = id_Y$.

Note the similarity to the definition of an invertible matrix. Indeed, when X, Y are vector spaces and f, g are linear, then we're asking their matrices to multiply to the identity matrix.

As with matrices, if f has an inverse, it's unique. We write it f^{-1}.

Invertible functions

A function is invertible if and only if it's one-to-one and onto.

If $f : X \to Y$ is invertible, then it's onto: for any $y \in Y$, we have $f(f^{-1}(y)) = y$. It's one-to-one: for any x with $f(x) = y$, we have $x = f^{-1}f(x) = f^{-1}(y)$, so there's only one such x.

Conversely If $f : X \to Y$ is one-to-one and onto, then for any $y \in Y$, there is a unique $x \in X$ with $f(x) = y$. Consider

$$\begin{aligned} f^{-1} : Y &\to X \\ y &\mapsto \text{the unique } x \text{ with } f(x) = y \end{aligned}$$

You can check this is the inverse.

Invertible functions

The inverse of an invertible linear transformation is again linear.

Consider some $T : V \to W$ and its inverse $T^{-1} : W \to V$:

$$\begin{aligned} T^{-1}(c_1\mathbf{w}_1 + c_2\mathbf{w}_2) &= T^{-1}(c_1 TT^{-1}(\mathbf{w}_1) + c_2 TT^{-1}(\mathbf{w}_2)) \\ &= T^{-1}(T(cT^{-1}(\mathbf{w}_1) + cT^{-1}(\mathbf{w}_2))) \\ &= cT^{-1}(\mathbf{w}_1) + cT^{-1}(\mathbf{w}_2) \end{aligned}$$

So T^{-1} is linear.

Isomorphism

Invertible linear transformations are also called isomorphisms.

If there's an isomorphism $f : V \to W$, we say V and W are isomorphic vector spaces.

Isomorphic vector spaces look the same to linear algebra.

More precisely, any question which can be asked just in terms of operations which make sense in any vector space must have the same answer in both.

You use the isomorphism f to translate back and forth.

Isomorphism

For example, the following map is an isomorphism

$$\begin{array}{rcl} \mathbb{P}_2 & \to & \mathbb{R}^3 \\ ax^2 + bx + c & \mapsto & (a, b, c) \end{array}$$

For the purposes of linear algebra, $\mathbb{P}_2$ and $\mathbb{R}^3$ are the same.

For instance, we learn immediately that no collection of more than three vectors in $\mathbb{P}_2$ can ever be linearly independent.

For other purposes, $\mathbb{P}_2$ and $\mathbb{R}^3$ are quite different: it doesn't make sense to solve an element of $\mathbb{R}^3$ for x, or to take its derivative.

Independence, span, bases, and linear transformations

A linear transformation $T : V \to W$ is onto if and only if the image of a spanning set is a spanning set.

Suppose a collection $\{\mathbf{v}_i\}$ of vectors spans V. To say $\{T(\mathbf{v}_i)\}$ spans W means any $\mathbf{w}$ in W is a linear combination of the $T(\mathbf{v}_i)$.

Since T is onto, we can at least write $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v}$ in V.

Since the $\mathbf{v}_i$ span, we can write $\mathbf{v} = \sum_i a_i \mathbf{v}_i$, hence

$$\mathbf{w} = T(\mathbf{v}) = T(\sum_i a_i \mathbf{v}_i) = \sum_i a_i T(\mathbf{v}_i)$$

For the converse: if the image of any collection of vectors in V span W then by linearity V must map onto W. (I'll let you think through the details.)

Independence, span, bases, and linear transformations

A linear transformation $T : V \to W$ is one-to-one if and only if the only element sent to zero is zero.

Indeed, by definition, T is one-to-one if every element of w has at most one thing in V mapping to it.

If two things map to zero, then T is certainly not one-to-one.

Conversely, suppose $\mathbf{0}$ maps to $\mathbf{0}$. Then $T(\mathbf{v}_1) = T(\mathbf{v}_2)$, or in other words $\mathbf{0} = T(\mathbf{v}_1) - T(\mathbf{v}_2) = T(\mathbf{v}_1 - \mathbf{v}_2)$, implies $\mathbf{v}_1 - \mathbf{v}_2 = 0$, and thus $\mathbf{v}_1 = \mathbf{v}_2$.

(We already saw this argument for $T : \mathbb{R}^n \to \mathbb{R}^m$.)

Independence, span, bases, and linear transformations

A linear transformation $T : V \to W$ is one-to-one if, and only if, the image of linearly independent vectors are linearly independent.

Say $\{\mathbf{v}_i\}$ in V are linearly independent, but their images in W are not. Then we must have $\sum a_i T(\mathbf{v}_i) = \mathbf{0}$ for some a_i not all zero. By linearity, $T(\sum a_i \mathbf{v}_i) = \mathbf{0}$; by independence, $\sum a_i \mathbf{v}_i \neq \mathbf{0}$. By the previous slide, T is not one-to-one.

Conversely, if T is not one-to-one, there is some nonzero vector $\mathbf{v}$ with $T(\mathbf{v}) = \mathbf{0}$. Thus the linearly independent set $\{\mathbf{v}\}$ is sent to the non-linearly-independent set $\{\mathbf{0}\}$.