8/29

1.0 General Notation

- $\mathbb{R}$ - real numbers
- $\mathbb{C}$ - complex numbers
- $\mathbb{N}$ - natural numbers $(1,2,3,\dots)$
- $\mathbb{Z}$ - integers $(\dots,-2,-1,0,1,2,\dots)$
- $\mathbb{Q}$ - rational numbers $(\frac{3}{4}, -\frac{2}{7}, \text{etc.})$
- $x \in \mathbb{R}$ "x is an element of $\mathbb{R}$" or "x is in the set $\mathbb{R}$"
- Set Notation:
    - $\{x \in \mathbb{Z} \mid x \geq 1\} = \mathbb{N}$
      (type of elements: $x \in \mathbb{Z}$; "such that": $\mid$; conditions: $x \geq 1$)
    - $\{x \in \mathbb{R} \mid x = \frac{n}{m} \text{ for some } n \in \mathbb{Z} \text{ and } m \in \mathbb{N}\} = \mathbb{Q}$
- $A \subset B$ "the set A is contained in the set B" or "the set A is a subset of the set B"
    - $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$
    - $\mathbb{R} \subset \mathbb{R}$, $\mathbb{Q} \subsetneq \mathbb{R}$ "$\mathbb{Q}$ is contained in but <u>not</u> equal to $\mathbb{R}$" (strict subset).

1.1 Vector Spaces

Def A vector space V is a collection of objects (called vectors) equipped with operations of addition and scalar multiplication such that the following vector space axioms hold:

Addition:

1. Commutativity: $\vec{v} + \vec{w} = \vec{w} + \vec{v}$ for all $\vec{v}, \vec{w} \in V$
2. Associativity: $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ for all $\vec{u}, \vec{v}, \vec{w} \in V$;
3. Zero Vector: there exists a special vector, denoted by $\vec{0}$ such that $\vec{v} + \vec{0} = \vec{v}$ for all $\vec{v} \in V$. It is called the zero vector;
4. Additive Inverse: for every vector $\vec{v} \in V$ there exists a vector $\vec{w} \in V$ such that $\vec{v} + \vec{w} = \vec{0}$. This vector is called the additive inverse of $\vec{v}$ and is usually denoted $-\vec{v}$;

Scalar multiplication:

5. Multiplicative Identity: $1\vec{v} = \vec{v}$ for all $\vec{v} \in V$;
6. Multiplicative associativity: $(\alpha\beta)\vec{v} = \alpha(\beta\vec{v})$ for all $\vec{v} \in V$ and all scalars α, β;

distributive laws:

7. $\alpha(\vec{u} + \vec{v}) = \alpha\vec{u} + \alpha\vec{v}$ for all $\vec{u}, \vec{v} \in V$ and scalars α;
8. $(\alpha + \beta)\vec{v} = \alpha\vec{v} + \beta\vec{v}$ for all $\vec{v} \in V$ and all scalars α, β.

- Vectors can potentially be very abstract objects (we'll see some examples soon), whereas "scalars" is just a fancy term for numbers. Sometimes we will mean the real numbers $\mathbb{R}$ and sometimes we'll mean the complex numbers $\mathbb{C}$. We call V a real vector space in the former case, and a complex vector space in the latter case. If we do not specify $\mathbb{R}$ or $\mathbb{C}$ or we write $\mathbb{F}$ then the statement holds for both. (It may even hold for any "field" $\mathbb{F}$, but we'll focus on $\mathbb{R}$ and $\mathbb{C}$ in this course).

- It is important to distinguish between vectors and scalars. So vectors will always be decorated with an arrow: $\vec{v}$ (or will be bold when typed). Scalars will usually be Greek letters ($\alpha, \beta, \gamma, \ldots$) while vectors will be roman letters ($u, v, w, \ldots$).

Remark The above axioms should be (at least vaguely) familiar from algebra/arithmetic, where they apply to just numbers rather than vectors and scalars. Consequently you should not need memorize the axioms (and in particular I will not ask you to do so), but you do need to remember what operations apply to what objects. For example, you can add vectors, and multiply a vector by a scalar, but you cannot multiply two vectors. ~~$\vec{u}\,\vec{v}$~~

Examples

(0) $V = \mathbb{R}$ is a real vector space where the vectors are just real numbers and so are the scalars, so all the axioms trivially hold. Similarly $V = \mathbb{C}$ is a complex vector space. We can also make it a real vector space, since a real number times a complex number is still a complex number.

(1) For $n \in \mathbb{N}$, $n \geq 2$ let $\mathbb{R}^n$ denote the columns with n entries from $\mathbb{R}$:

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

Then $\mathbb{R}^n$ is a vector space entry-wise operations:

$$\underbrace{\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1+w_1 \\ v_2+w_2 \\ \vdots \\ v_n+w_n \end{pmatrix} \in \mathbb{R}^n}_{\text{addition}} \qquad \underbrace{\alpha \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \alpha v_1 \\ \alpha v_2 \\ \vdots \\ \alpha v_n \end{pmatrix} \in \mathbb{R}^n}_{\text{scalar multiplication}}$$

Let's check a few axioms:

Commutativity:

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1+w_1 \\ v_2+w_2 \\ \vdots \\ v_n+w_n \end{pmatrix} = \begin{pmatrix} w_1+v_1 \\ w_2+v_1 \\ \vdots \\ w_n+v_n \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

(since addition in $\mathbb{R}$ is commutative)

Zero vector: we claim $\vec{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. Indeed

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} v_1+0 \\ v_2+0 \\ \vdots \\ v_n+0 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Multiplicative Associativity:

$$(\alpha\beta)\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} (\alpha\beta)v_1 \\ (\alpha\beta)v_2 \\ \vdots \\ (\alpha\beta)v_n \end{pmatrix} = \begin{pmatrix} \alpha(\beta v_1) \\ \alpha(\beta v_2) \\ \vdots \\ \alpha(\beta v_n) \end{pmatrix} = \alpha\begin{pmatrix} \beta v_1 \\ \beta v_2 \\ \vdots \\ \beta v_n \end{pmatrix} = \alpha\left(\beta\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}\right).$$

$\mathbb{C}^n$ is defined similarly but with complex entries, and it is a complex vector space.

(2) For $n \in \mathbb{N}$, let $\mathbb{P}_n$ denote the collection of polynomials of degree at most n:

$$p(t) = a_n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0 \in \mathbb{P}_n$$

Note that the coefficients a_i are allowed to be zero.

Addition and scalar multiplication are given by:

$$(a_n t^n + \cdots + a_1 t + a_0) + (b_n t^n + \cdots + b_1 t + b_0) = (a_n + b_n)t^n + \cdots + (a_1+b_1)t + (a_0+b_0)$$

$$\alpha(a_n t^n + \cdots + a_1 t + a_0) = (\alpha a_n)t^n + \cdots + (\alpha a_1)t + (\alpha a_0)$$

If we consider only real coefficients, then $\mathbb{P}_n$ is a real vector space. If we allow complex coefficients, then it is a complex vector space.

What is $\vec{0}$ here? $p(t) = 0t^n + \cdots + 0t + 0 = 0.$

What is the additive inverse of $p(t) = 3t^3 - t^2 + 4it + 1.2$? $-p(t) = -3t^3 + t^2 - 4it - 1.2$?

(3) (Netflix)

List genres of movies: $g_1, g_2, \ldots, g_N$.

Consider the vector space

$$V = \mathbb{R}^N$$

with the usual operations of scalar mult. and addition.

For a netflix user, define $x = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \in V$ for each $i = 1, \ldots, N$ by:

x_i = (# of movies of genre g_i they have finished watching) − (# of movies in genre g_i they started but didn't finish)

Heuristically: x_i is positive if they like movies in genre g_i

Netflix will use this data to recommend movies to you

Addition $x + y$ corresponds to user x and user y sharing an account. □

Rem It is implicit in the definition of a vector space V that it is "closed" under the addition and scalar multiplication operations. That is:

1. For any $\vec{v}, \vec{w} \in V$ we must have that $\vec{v}+\vec{w}$ is <u>also</u> in V.
2. For any $\vec{v} \in V$ and any scalar α, we must have that $\alpha\vec{v}$ is <u>also</u> in V.

In particular, this means that if a set V fails to be closed under either of these operations, then it is <u>not</u> a vector space.

Ex Let

$$V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x^2+y^2=z^2 \right\}$$

with addition and scalar multiplication defined as for $\mathbb{R}^3$. Then V is <u>not</u> a vector space because it is not closed under addition. Indeed,

$$\vec{v} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \vec{w} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

are both in V, but $\vec{v}+\vec{w} = (1,1,2)^T$ is <u>not</u> since $1^2+1^2 \neq 2^2$. □

Theorem For every vector space V, the zero vector $\vec{0}$ is unique.

Proof Suppose $\vec{0}$ and $\vec{0}'$ both satisfy the zero vector axiom:

$$\vec{v}+\vec{0}=\vec{v} \quad \text{and} \quad \vec{v}+\vec{0}'=\vec{v} \quad \text{for all } \vec{v}\in V.$$

Then we have

$$\begin{aligned} \vec{0}' &= \vec{0}'+\vec{0} && \text{(first equation)} \\ &= \vec{0}+\vec{0}' && \text{(commutativity)} \\ &= \vec{0} && \text{(second equation)} \end{aligned}$$

□

9/3

Thm Let V be a vector space. For any $\vec{v}\in V$, $0\vec{v}=\vec{0}$.

Proof Fix $\vec{v}\in V$. First observe

$$\vec{v} = 1\vec{v} = (1+0)\vec{v} = 1\vec{v}+0\vec{v} = \vec{v}+0\vec{v}$$

So $\vec{v}=\vec{v}+0\vec{v}$. But then adding $-\vec{v}$ to each side we get:

$$\begin{aligned} \vec{v}+-\vec{v} &= \vec{v}+0\vec{v}+-\vec{v} \\ \vec{0} &= \vec{v}+-\vec{v}+0\vec{v} \\ \vec{0} &= \vec{0}+0\vec{v} \\ \vec{0} &= 0\cdot\vec{v} \end{aligned}$$

as claimed. □

Matrix Notation

An $m\times n$ <u>matrix</u> is a rectangular array with m rows and n columns. Elements in the array are called entries and can be <u>real</u> or <u>complex</u> numbers.

Ex

$\begin{pmatrix} 2 & -1 \\ 0 & 4.3 \\ 16 & -0.\bar{3} \end{pmatrix}$ is 3×2 $\quad$ $\begin{pmatrix} 2+i & e^{\pi i} & 1 \end{pmatrix}$ is 1×3 $\square$

- $M_{m\times n}(\mathbb{R})$: $m\times n$ matrices with real entries

$M_{m\times n}(\mathbb{C})$: $m\times n$ matrices with complex entries

$M_{m\times n}$ refers to both. Using entrywise addition and scalar mult. (like $\mathbb{R}^n$ and $\mathbb{C}^n$) $M_{m\times n}(\mathbb{R})$ is a real vector space and $M_{m\times n}(\mathbb{C})$ is a complex vector space.

- For $A\in M_{m\times n}$, we write $(A)_{jk}$ for the entry of A in row j and column k, with $j=1,\dots,m$ and $k=1,\dots,n$.

$$A=\begin{pmatrix} -1 & 3 \\ 2 & 0 \\ 5 & -2 \end{pmatrix} \qquad (A)_{21}=2 \qquad (A)_{32}=-2 \qquad (A)_{23}=\text{DNE}$$

We may also use lower-case letters to denote entries: $a_{jk}=(A)_{jk}$. In this case we could define A by writing

$$A=(a_{jk})_{j=1,k=1}^{m\ \ n}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Def For a matrix $A\in M_{m\times n}$, its transpose, denoted A^T, is the $n\times m$ matrix satisfying $(A^T)_{jk}=(A)_{kj}$ for $j=1,\dots,n$ and $k=1,\dots,m$.

Taking the transpose of a matrix turns its rows into columns (and vice-versa)

$$A=\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \qquad A^T=\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

$$B=\begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix} \qquad B^T=\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

- Note that $\mathbb{F}^n$ is the same as $M_{n\times 1}(\mathbb{F})$. So we can take the transpose of a vector:

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}^T=(v_1, v_2, \dots, v_n) \leftarrow \text{row vector}$$

column vector

Also $(x_1, x_2, \dots, x_n)^T=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\in\mathbb{F}^n$. Since this saves space, we will frequently write elements of $\mathbb{F}^n$ as the transpose of a row vector.

1.2 Linear Combinations, bases

Let's make use of our two operations: addition and scalar multiplication.

Def Let $\vec{v}_1, \ldots, \vec{v}_p \in V$ be a collection of vectors. A <u>linear combination</u> of $\vec{v}_1, \ldots, \vec{v}_p \in V$ is a sum of the form

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_p \vec{v}_p \quad \left(= \sum_{k=1}^{p} \alpha_k \vec{v}_k \right)$$

where $\alpha_1, \ldots, \alpha_p$ are scalars.

Note that linear combination of vectors in V is ultimately another vector in V. In this section we will be interested in answering the following questions for a fixed collection of vectors $\vec{v}_1, \ldots, \vec{v}_p \in V$:

1. Can <u>every</u> vector in V be written as a lin. comb. of $\vec{v}_1, \ldots, \vec{v}_p$?
2. If $\vec{v} = \sum_{k=1}^{p} \alpha_k \vec{v}_k$ is a linear comb. of $\vec{v}_1, \ldots, \vec{v}_p$, is this the <u>only</u> way to express it as a lin. comb.? (i.e. can we pick scalars different from $\alpha_1, \ldots, \alpha_p$ and still get $\vec{v}$?)

When the answer to both these questions is "yes", we get a basis:

Def A system of vectors $\vec{v}_1, \ldots, \vec{v}_p \in V$ is called a <u>basis</u> for V if every $\vec{v} \in V$ <u>has</u> a <u>unique</u> representation as a linear combination

$$\vec{v} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_p \vec{v}_p$$

The scalars $\alpha_1, \ldots, \alpha_p$ are called the <u>coordinates</u> of the vector $\vec{v}$ with respect to the basis $\vec{v}_1, \ldots, \vec{v}_p$.

Ex ① For $V = \mathbb{F}^n$ (recall $\mathbb{F}$ can be either $\mathbb{R}$ or $\mathbb{C}$), consider the vectors:

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Then $\vec{e}_1, \ldots, \vec{e}_n$ is a basis for $\mathbb{F}^n$. Indeed, for any $\vec{v} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{F}^n$ we have:

$$\vec{v} = \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix} = x_1 \vec{e}_1 + x_2 \vec{e}_2 + \cdots + x_n \vec{e}_n$$

Thus $\vec{v}$ admits a rep. as a lin. comb. This is also unique: suppose $\vec{v} = y_1 \vec{e}_1 + \cdots + y_n \vec{e}_n$ for some scalars $y_1, \ldots, y_n$. Then

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = v = y_1 \vec{e}_1 + y_2 \vec{e}_2 + \cdots + y_n \vec{e}_n = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

So $x_i = y_i$ for $i = 1, \ldots, n$, which means the lin. comb. is unique.

We call $\vec{e}_1, \dots, \vec{e}_n$ the standard basis for $\mathbb{F}^n$.

Note that $\vec{e}_2, \dots, \vec{e}_n$ is not a basis for $\mathbb{F}^n$. This is because for any lin. comb. we have

$$\alpha_2 \vec{e}_2 + \dots + \alpha_n \vec{e}_n = \begin{pmatrix} 0 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \leftarrow \text{always zero}$$

So whenever $\vec{v} = (x_1, \dots, x_n)^T$ has $x_1 \neq 0$ it will not admit a rep. as a lin. comb.

Furthermore, if $\vec{v} = \begin{pmatrix} 2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ then $\vec{v}, \vec{e}_1, \dots, \vec{e}_n$ is not a basis for $\mathbb{F}^n$.

Indeed, notice that

$$\begin{pmatrix} 2 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \vec{v} + \vec{e}_2 = 2\vec{e}_1 + \vec{e}_2$$

So $(2, 1, 0, \dots, 0)^T$ does not have unique coordinates with respect to $\vec{v}, \vec{e}_1, \dots, \vec{e}_n$, which means it is not a basis.

(2) Recall $\mathbb{P}_n$ is the space of polynomials of degree at most n.

Define $\vec{e}_0, \vec{e}_1, \dots, \vec{e}_n \in \mathbb{P}_n$ by

$$\vec{e}_0 = 1, \quad \vec{e}_1 = t, \quad \dots, \quad \vec{e}_n = t^n$$

Then for any $p(t) = a_n t^n + \dots + a_1 t + a_0$, we have

$$p(t) = a_n \vec{e}_n + \dots + a_1 \vec{e}_1 + a_0 \vec{e}_0$$

and by the same argument as in the previous example, it is unique. Hence $\vec{e}_0, \vec{e}_1, \dots, \vec{e}_n$ is a basis for $\mathbb{P}_n$, and we call it the standard basis for $\mathbb{P}_3$.

Note a vector space has more than one basis:

$$t-1, \ t+1, \ t^2, \dots, t^n$$

is also a basis for $\mathbb{P}_n$. This is because $\vec{e}_0 = \frac{1}{2}(t+1) - \frac{1}{2}(t-1)$ and $\vec{e}_1 = \frac{1}{2}(t+1) + \frac{1}{2}(t-1)$ □

Rem If a vector space V has a basis $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$, then by definition every vector $\vec{v} \in V$ is uniquely determined by its coordinates with respect the basis:

$$\vec{v} = \sum_{k=1}^{n} \alpha_k \vec{v}_k$$

Therefore, so long as we remember $\vec{v}_1, \dots, \vec{v}_n$ and $\alpha_1, \dots, \alpha_n$, we will remember $\vec{v}$. Hence we can treat $\vec{v}$ as the column vector $(\alpha_1, \alpha_2, \dots, \alpha_n)^T \in \mathbb{F}^n$. This isn't just a convenient way to remember $\vec{v}$, it also respects addition and scalar multiplication:

$$\vec{v} + \vec{w} = \left(\sum_{k=1}^{n} \alpha_k \vec{v}_k\right) + \left(\sum_{k=1}^{n} \beta_k \vec{v}_k\right) = \sum_{k=1}^{n} (\alpha_k + \beta_k) \vec{v}_k \longleftrightarrow \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 + \beta_1 \\ \alpha_2 + \beta_2 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix}$$

$$\beta \vec{v} = \beta \sum_{k=1}^{n} \alpha_k \vec{v}_k = \sum_{k=1}^{n} \beta \alpha_k \vec{v}_k \longleftrightarrow \beta \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \beta \alpha_1 \\ \beta \alpha_2 \\ \vdots \\ \beta \alpha_n \end{pmatrix}$$

This is the true power of a basis: it allows us to take a potentially very abstract vector space and replace it and its vectors with something much more familiar ($\mathbb{F}^n$). ☐

- So, how does one find a basis or check that a given system is a basis? By answering our two questions from earlier, but one at a time.

9/5

Generating and linearly independent systems

Def A system of vectors $\vec{v}_1, \dots, \vec{v}_p \in V$ is called a generating system (or a spanning system, or a complete system) in V if any vector $\vec{v} \in V$ admits a representation as a linear combination:

$$\vec{v} = \alpha_1 \vec{v}_1 + \cdots + \alpha_p \vec{v}_p$$

for some scalars $\alpha_1, \dots, \alpha_p$.

- Note that we do not assume this linear comb. is unique.

Ex ⓪ Every basis

① $\begin{pmatrix} 2 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ is a spanning system for $\mathbb{F}^n$, $\vec{e}_2, \vec{e}_3, \dots, \vec{e}_n$ is not

② Let $\vec{v}_1, \dots, \vec{v}_p \in V$ be a basis in V. Let $\vec{v}_{p+1}, \dots, \vec{v}_n \in V$ be any vectors. Then $\vec{v}_1, \dots, \vec{v}_p, \vec{v}_{p+1}, \dots, \vec{v}_n$ is a spanning system in V, since you can always just set the coefficients $\alpha_{p+1} = \cdots = \alpha_n = 0$. ☐

The other aspect of the def. of a basis considered uniqueness of the linear combination:

$$\sum_{k=1}^{p} \alpha_k \vec{v}_k = \vec{v} = \sum_{k=1}^{p} \beta_k \vec{v}_k \implies \alpha_k = \beta_k \quad k=1,\dots,p$$

Equivalently,

$$\sum_{k=1}^{p} (\alpha_k - \beta_k) \vec{v}_k = \vec{0} \implies \alpha_k - \beta_k = 0 \quad k=1,\dots,p$$

So uniqueness is really talking about how one can obtain the zero vector as a linear combination.

Def A linear combination $\alpha_1 \vec{v}_1 + \cdots + \alpha_p \vec{v}_p$ is called trivial if $\alpha_k = 0$ for each $k=1,\dots,p$. It is called non-trivial if $\alpha_k \neq 0$ for at least one $k=1,\dots,p$. Equivalently, $\sum_{k=1}^{p} |\alpha_k| \neq 0$.

Notice that, regardless of $\vec{v}_1, \dots, \vec{v}_p$, a trivial linear combination equals $\vec{0}$.

Def A system of vectors $\vec{v}_1, \dots, \vec{v}_p \in V$ is called <u>linearly independent</u> if a linear combination $\sum_{k=1}^{p} \alpha_k \vec{v}_k$ equaling the zero vector $\vec{0}$ implies it is a trivial linear combination (i.e. $\sum_{k=1}^{p} \alpha_k \vec{v}_k = \vec{0} \Rightarrow \alpha_1 = \cdots = \alpha_k = 0$). Equivalently, every non-trivial linear combination does not equal $\vec{0}$.

- This definition gives you a guide for how to prove a system is linearly independent: suppose $\sum_{k=1}^{p} \alpha_k \vec{v}_k = \vec{0}$, then use what you know about $\vec{v}_1, \dots, \vec{v}_p$ to show that one must have $\alpha_1 = \cdots = \alpha_k = 0$.

Ex Let V be the set of functions $f: \mathbb{R} \to \mathbb{R}$, with operations:

$$(f+g)(x) = f(x) + g(x) \qquad (\alpha f)(x) = \alpha f(x).$$

Then V is a real vector space.

We claim

$$f_1(x) = 2 \qquad f_2(x) = x^2(x-1) \qquad f_3(x) = 1 - e^x$$

are linearly independent. Indeed, suppose $\alpha_1 f_1 + \alpha_2 f_2 + \alpha_3 f_3 = \vec{0}$. Note that the zero vector is the function $\vec{0}(x) = 0$ for all $x \in \mathbb{R}$. So $\uparrow$ implies

$$(\alpha_1 f_1 + \alpha_2 f_2 + \alpha_3 f_3)(x) = 0 \qquad \forall x \in \mathbb{R}$$

$$\alpha_1 2 + \alpha_2 x^2(x-1) + \alpha_3 (1 - e^x) = 0 \qquad \forall x \in \mathbb{R}.$$

Plugging in $x=0$ gives:

$$\alpha_1 2 + \alpha_2 (0) + \alpha_3 (0) = 0 \Rightarrow \alpha_1 2 = 0 \Rightarrow \alpha_1 = 0.$$

Next, plugging in $x=1$ gives:

$$0 \cdot 2 + \alpha_2 (0) + \alpha_3 (1-e) = 0 \Rightarrow \alpha_3 (1-e) = 0 \Rightarrow \alpha_3 = 0$$

Finally, plugging in $x=2$ gives:

$$0 \cdot 2 + \alpha_2 2^2 (2-1) + 0 \cdot (1 - e^2) = 0 \Rightarrow \alpha_2 4 = 0 \Rightarrow \alpha_2 = 0.$$

So $\alpha_1 = \alpha_2 = \alpha_3 = 0$ and the lin. comb. is trivial. So f_1, f_2, f_3 are lin. indep. in V. □

- If a system of vectors is <u>not</u> linearly independent, we say it is "linearly dependent". We obtain the formal definition by negating the one above.

Def A system of vectors $\vec{v}_1, \dots, \vec{v}_p \in V$ is called <u>linearly dependent</u> if there exists a non-trivial linear combination $\sum_{k=1}^{p} \alpha_k \vec{v}_k$ that equals the zero vector $\vec{0}$.

- An equivalent way to say this is: $\vec{v}_1, \dots, \vec{v}_p$ are linearly dependent if and only if the equation

$$x_1 \vec{v}_1 + x_2 \vec{v}_2 + \cdots + x_p \vec{v}_p = \vec{0}$$

(with unknowns $x_1, \dots, x_p$) has a non-trivial solution.

- The next proposition gives yet another characterization:

Prop A system of vectors $\vec{v}_1, \ldots, \vec{v}_p \in V$ is linearly dependent if and only if one of the vectors $\vec{v}_k$ can be represented as a linear combination of the other vectors:

$$\vec{v}_k = \sum_{\substack{j=1 \\ j \neq k}}^{p} \beta_j \vec{v}_j$$

for some scalars β_j, $j = 1, \ldots, k-1, k+1, \ldots, p$.

Proof ($\Longrightarrow$) Suppose $\vec{v}_1, \ldots, \vec{v}_p$ are linearly dependent. Then there exists scalars $\alpha_1, \ldots, \alpha_p$ such that $\sum_{k=1}^{p} |\alpha_k| \neq 0$ and

$$\alpha_1 \vec{v}_1 + \cdots + \alpha_p \vec{v}_p = \vec{0}$$

We know there is at least one non-zero scalar, say α_k. For $j \neq k$, add $-\alpha_j \vec{v}_j$ to each side of the above equation and divide by α_k to get:

$$\vec{v}_k = \sum_{j \neq k} -\frac{\alpha_j}{\alpha_k} \vec{v}_j$$

Thus we take $\beta_j := -\frac{\alpha_j}{\alpha_k}$.

($\Longleftarrow$) Suppose

$$v_k = \sum_{\substack{j=1 \\ j \neq k}}^{p} \beta_j \vec{v}_j$$

For some k and scalars β_j. Adding $-\beta_j \vec{v}_j$ to each side for $j = 1, \ldots, p$, $j \neq k$ yields:

$$\vec{v}_k - \sum_{\substack{j=1 \\ j \neq k}}^{p} \beta_j \vec{v}_j = \vec{0}$$

$$(-\beta_1)\vec{v}_1 + \cdots + (-\beta_{k-1})\vec{v}_{k-1} + (1)\vec{v}_k + (-\beta_{k+1})\vec{v}_{k+1} + \cdots + \vec{v}_p = \vec{0}$$

Since the coefficient of $\vec{v}_k$ is non-zero, this is a non-trivial linear combination. Thus $\vec{v}_1, \ldots, \vec{v}_p$ are linearly dependent. □

- Observe that any basis is linearly independent: $\vec{0} \in V$ so if $\vec{v}_1, \ldots, \vec{v}_p$ is a basis in V then $\vec{0}$ admits a <u>unique</u> representation as a linear combination:

$$\vec{0} = \alpha_1 \vec{v}_1 + \cdots + \alpha_p \vec{v}_p$$

We can of course always choose $\alpha_1, \ldots, \alpha_p = 0$, so this must be the only choice. Consequently $\vec{v}_1, \ldots, \vec{v}_p$ are linearly independent.

- We've seen bases give us examples of both generating <u>and</u> lin. indep. systems. It turns out the converse is also true:

Prop A system of vectors $\vec{v}_1, \ldots, \vec{v}_n \in V$ is a basis if and only if it is generating and linearly independent.

Proof: ($\Longrightarrow$) We have already showed that a basis is both generating and linearly indep.

($\Longrightarrow$) Suppose $\vec{v}_1, \ldots, \vec{v}_n$ is generating and linearly indep. We need to show every vector in V has a unique rep. as a lin. comb. of $\vec{v}_1, \dots, \vec{v}_n$. Fix an arbitrary $\vec{v} \in V$. Since $\vec{v}_1, \dots, \vec{v}_n$ is generating, we $\vec{v}$ has a rep. as a lin comb.

$$\vec{v} = \sum_{k=1}^{n} \alpha_k \vec{v}_k$$

Towards showing this rep. is unique, suppose $\vec{v} = \sum_{k=1}^{n} \tilde{\alpha}_k \vec{v}_k$ is another rep. Since both rep's equal $\vec{v}$, subtracting one from the other gives:

$$\vec{0} = \sum_{k=1}^{n} (\alpha_k - \tilde{\alpha}_k) \vec{v}_k$$

Since the system is lin. indep., the lin. comb. on the right must be trivial: $\alpha_k - \tilde{\alpha}_k = 0$ for $k = 1, \dots, n$. Thus $\tilde{\alpha}_k = \alpha_k$ for each $k = 1, \dots, n$, and so the rep. $\vec{v} = \sum \alpha_k \vec{v}_k$ is unique. □

Prop Any finite generating system contains a basis.

Proof: suppose $\vec{v}_1, \dots, \vec{v}_n \in V$ is a generating system. If it is linearly indep., then by the previous Proposition it is a basis and so we're done.

Otherwise it is linearly dep., and so by the Proposition before the previous one there is a k s.t.

$$v_k = \sum_{\substack{j=1 \\ j \neq k}}^{n} \beta_j \vec{v}_j$$

Without loss of generality (WLOG), we may assume $k = n$. Observe then that any lin. comb. of $\vec{v}_1, \dots, \vec{v}_n$ can be re-written as a lin comb. of $\vec{v}_1, \dots, \vec{v}_{n-1}$:

$$\alpha_1 \vec{v}_1 + \dots + \alpha_n \vec{v}_n = \alpha_1 \vec{v}_1 + \dots + \alpha_{n-1} \vec{v}_{n-1} + \alpha_n \left(\sum_{j=1}^{n-1} \beta_j \vec{v}_j \right)$$
$$= (\alpha_1 + \alpha_n \beta_1) \vec{v}_1 + \dots + (\alpha_{n-1} + \alpha_n \beta_{n-1}) \vec{v}_{n-1}$$

It follows that $\vec{v}_1, \dots, \vec{v}_{n-1}$ is a (smaller) generating system. If it is linearly indep. then we are done. Otherwise, we repeat the above argument until we obtain a generating and lin. indep. system. Note that this process will stop at or before we reduce the system down to a single vector, since a single vector is always linearly independent (Exercise)

Thus, after finitely many steps, we will obtain a basis that was contained in the original system □

4/10

1.3 Linear Transformations and Matrix-vector multiplication

- A "transformation" T from a set X to a set Y is a rule that for each input $x \in X$ assigns an output $y \in Y$, which we denote $T(x) := y$. We write

$$T: X \to Y$$

(X: domain; Y: target space / codomain)

synonyms: transform, mapping, map, operation, or function.

- If the sets X, Y have more structure, then so can T.

Def Let V, W be vector spaces (over the same field $\mathbb{F}$). A transformation $T: V \to W$ is called <u>linear</u> if

1. $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v}) \quad \forall \vec{u}, \vec{v} \in V$ ("for all")
2. $T(\alpha\vec{v}) = \alpha T(\vec{v})$ for all $\vec{v} \in V$ and all scalars $\alpha \in \mathbb{F}$.

Note that we can combine properties 1. and 2. into a single, equivalent statement:

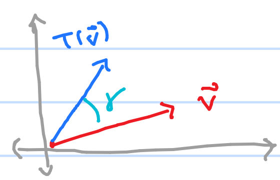
$$T(\alpha\vec{u}+\beta\vec{v}) = \alpha T(\vec{u}) + \beta T(\vec{v}) \qquad \forall \vec{u}, \vec{v} \in V \text{ and } \forall \alpha, \beta \in \mathbb{F}.$$

Ex ① Let $V = \mathbb{P}_n$, $W = \mathbb{P}_{n-1}$ and define $T: \mathbb{P}_n \to \mathbb{P}_{n-1}$ by $T(p) = p'$. That is,

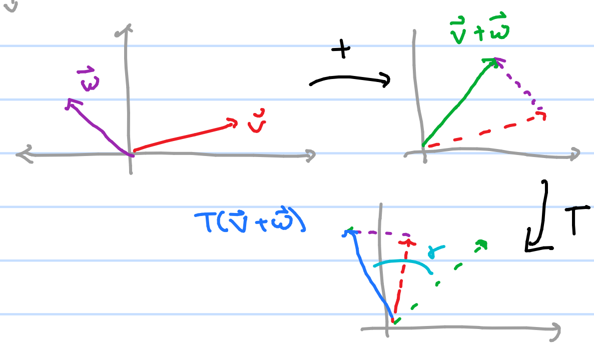
$$T(a_n t^n + \cdots + a_1 t + a_0) = n a_n t^{n-1} + \cdots + a_1$$

Since $(p+q)' = p' + q'$ and $(\alpha p)' = \alpha p'$, T is linear.

② Let $V = W = \mathbb{R}^2$ and fix $\gamma \in [0, 2\pi)$. Define $T: \mathbb{R}^2 \to \mathbb{R}^2$ by setting $T(\vec{v})$ to be the vector one obtains after rotating the plane counterclockwise by γ radians.



Recall that addition in $\mathbb{R}^2$ is visually equivalent to concatenating vectors and forming a triangle. Since rotating the plane preserves the internal angles of the triangle, we see that T satisfies Property 1. Property 2 is also easily checked, so T is a linear transformation.



③ Let $V = \{f: \mathbb{R} \to \mathbb{R}\}$ and $W = \mathbb{R}$. Define $T: V \to \mathbb{R}$ by $T(f) = f(2)$. Then for $f, g \in V$ and $\alpha, \beta \in \mathbb{R}$

$$T(\alpha f + \beta g) = (\alpha f + \beta g)(2) = \alpha(f(2)) + \beta(g(2)) = \alpha T(f) + \beta T(g)$$

So T is a linear transformation.

④ Let $V = W = \mathbb{R}$.

<u>Claim</u>: Any linear transformation $T: \mathbb{R} \to \mathbb{R}$ is given by

$$T(x) = ax \quad \text{where} \quad a = T(1).$$

Indeed, $x \in \mathbb{R}$ is both a vector in V and a scalar in $\mathbb{R}$: $x = x\,1$ (vectors; scalar)

So because T is linear and satisfies Property 2, we have:

$$T(x) = T(x1) = x\,T(1) = xa = ax.$$

as claimed.

Similarly, any linear transformation $T: \mathbb{C} \to \mathbb{C}$ is determined by multiplication by a scalar $a \in \mathbb{C}$. □

Linear Transformations $\mathbb{F}^n \to \mathbb{F}^m$

Let $n, m \in \mathbb{N}$. We will generalize the claim in the last example to higher dimensions and show any linear transformation $\mathbb{F}^n \to \mathbb{F}^m$ is given by multiplication by a matrix (rather than a scalar).

Let $T: \mathbb{F}^n \to \mathbb{F}^m$ be a linear transformation.

Claim 1 To compute $T(\vec{x})$ for any $\vec{x} \in \mathbb{F}^n$, it suffices to know $T(\vec{e}_1), \dots, T(\vec{e}_n)$ for the standard basis $\vec{e}_1, \dots, \vec{e}_n$ of $\mathbb{F}^n$.

Indeed, suppose $\vec{x} = (x_1, x_2, \dots, x_n)^T$. Then

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n$$

So using the linearity of T we have:

$$\begin{aligned} T(\vec{x}) &= T(x_1\vec{e}_1 + \cdots + x_n\vec{e}_n) \\ &= T(x_1\vec{e}_1) + \cdots + T(x_n\vec{e}_n) \\ &= x_1 T(\vec{e}_1) + \cdots + x_n T(\vec{e}_n) \end{aligned}$$

Thus if

$$\vec{a}_1 := T(\vec{e}_1) \quad \vec{a}_2 := T(\vec{e}_2) \quad \cdots \quad \vec{a}_n := T(\vec{e}_n)$$

Then $T(\vec{x})$ is the linear comb. $\sum_{j=1}^{n} x_j \vec{a}_j$. □

Let's examine this further. Define a matrix $A \in M_{m\times n}$ by

$$A = \begin{pmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{pmatrix}$$

Label the entries $A = (a_{ij})_{i=1, j=1}^{m \;\; n}$, so that

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{and} \quad \vec{a}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

Recall that multiplication by matrices is defined by a "row by column" rule:
If $A\vec{x}=\vec{y}$ for $A\in M_{m\times n}$, $\vec{x}\in\mathbb{F}^n$, and $\vec{y}\in\mathbb{F}^m$, then the kth entry of $\vec{y}$ is the dot product of the kth row of A and the column vector $\vec{x}$:

$$y_k = (A)_{k1}x_1 + (A)_{k2}x_2 + \cdots + (A)_{kn}x_n = \sum_{j=1}^{n} (A)_{kj}x_j \qquad k\left(\overline{\quad A \quad} \right)\left(\vec{x} \right) = \left(\vec{y} \right)k$$

Claim 2 With $T:\mathbb{F}^n\to\mathbb{F}^m$ and A as before, $T(\vec{x}) = A\vec{x}$.

Indeed, we have already seen that

$$T(\vec{x}) = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n$$

Expanding the vectors:

$$= x_1\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{m2} \end{pmatrix}$$

$$= \begin{pmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \vdots \\ x_1a_{m1} + x_2a_{m2} + \cdots + x_na_{mn} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{pmatrix} = A\vec{x}.$$

□

- Thus any linear transformation $T:\mathbb{F}^n\to\mathbb{F}^m$ can be represented as multiplication by by the matrix $A\in M_{m\times n}$ whose columns are $T(\vec{e}_j)$. Pay careful attention to to 'n' vs 'm' in this representation.
- We will denote the matrix A by $[T]$ or even just T if there is no room for confusion, so we can write $T\vec{v}$ for $T(\vec{v})$.

Rem In claim 1, we did not need to use the standard basis $\vec{e}_1,\ldots,\vec{e}_n$. All we needed was to be able to write $\vec{x}$ as a linear comb. of $\vec{e}_1,\ldots,\vec{e}_n$ (which is easy). Thus we could have used any basis for $\mathbb{F}^n$ or even any generating set. This is even true for arbitrary vector spaces.

Prop A linear transformation $T:V\to W$ is completely determined by its output on a generating set (in particular by its output on a basis).

Proof Exercise. □

1.4 Linear transformations as a vector space

Let us explore what operations we can perform on linear transformations themselves.

Suppose $S, T: V \to W$ are linear transformations. Then for any $\vec{v} \in V$ we can add $S(\vec{v}) + T(\vec{v})$. Denote $(S+T)(\vec{v}) = S(\vec{v}) + T(\vec{v})$, then this defines a new trans. $S+T: V \to W$. Is it linear? For $\vec{v}, \vec{w} \in V$ and scalars α, β we have

$$\begin{aligned}(S+T)(\alpha\vec{v}+\beta\vec{w}) &= S(\alpha\vec{v}+\beta\vec{w}) + T(\alpha\vec{v}+\beta\vec{w}) \\ &= \alpha S(\vec{v}) + \beta S(\vec{w}) + \alpha T(\vec{v}) + \beta T(\vec{w}) \\ &= \alpha[S(\vec{v})+T(\vec{v})] + \beta[S(\vec{w})+T(\vec{w})] \\ &= \alpha(S+T)(\vec{v}) + \beta(S+T)(\vec{w})\end{aligned}$$

So $S+T$ is also linear.

For $T: V \to W$ a lin. trans. and α a scalar define a trans. $\alpha T: V \to W$ by

$$(\alpha T)(\vec{v}) = \alpha T(\vec{v}) \qquad \vec{v} \in V.$$

One can show αT is also linear (Exercise).

Let $L(V,W)$ denote the collection of linear transformations from a vector space V to a vector space W. We have shown above that it admits operations of addition and scalar multiplication. Moreover, one can check that these axioms satisfy the vector space axioms. For example:

- Zero vector: Let $\vec{0}_W$ be the zero vector in W. Then $O: V \to W$ defined by

$$O(\vec{v}) = \vec{0}_W \qquad \vec{v} \in V$$

is the zero "vector" in $L(V,W)$. Indeed, it is linear:

HW1

$$O(\alpha\vec{v}+\beta\vec{w}) = \vec{0}_W = \vec{0}_W + \vec{0}_W = \alpha\vec{0}_W + \beta\vec{0}_W = \alpha O(\vec{v}) + \beta O(\vec{w}).$$

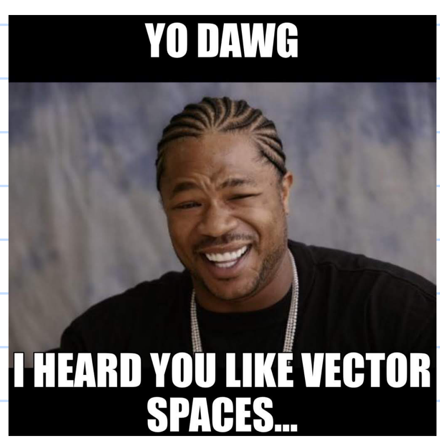
So $O \in L(V,W)$. And for any $T \in L(V,W)$ we have $T + O = T$ since

$$(T+O)(\vec{v}) = T(\vec{v}) + O(\vec{v}) = T(\vec{v}) + \vec{0}_W = T(\vec{v})$$

for all $\vec{v} \in V$.

(Exercise: check the remaining axioms).

Thus $L(V,W)$ is itself a vector space.



Ex For $V = \mathbb{F}^n$ and $W = \mathbb{F}^m$, $L(\mathbb{F}^n, \mathbb{F}^m)$ is a vector space. On the other hand, we know every $T \in L(\mathbb{F}^n, \mathbb{F}^m)$ can be represented as matrix multiplication by $[T]$. Moreover, the operations on $L(\mathbb{F}^n, \mathbb{F}^m)$ match those on $M_{m \times n}$:

$$[S+T] = [S] + [T]$$
$$[\alpha T] = \alpha [T]$$

So $L(\mathbb{F}^n, \mathbb{F}^m) = M_{m\times n}$. □

Rem What about the other operation we have for $M_{m\times n}$: multiplication? It turns out this corresponds to composition of linear trans., as we will see in the next section.

1.5 Composition of linear transformations (and matrix multiplication)

For two matrices A, B recall how their product is defined:
Then entry $(AB)_{jk}$ is given by the dot product of the jth row of A and the kth column of B:

$$(AB)_{jk} = \sum_{\ell} (A)_{j\ell} (B)_{\ell k} \qquad j\begin{pmatrix} \text{—} A \text{—} \end{pmatrix} \begin{pmatrix} \Big| B \end{pmatrix}^{k} = \begin{pmatrix} \overset{jk}{AB} \end{pmatrix}$$

Warning AB only makes sense if # columns of A = # rows of B. That is

$$\underset{A}{M_{m \times n}} \qquad \underset{B}{M_{n \times r}}$$

Composition of Linear Transformations

Suppose $T_1 : \mathbb{F}^n \to \mathbb{F}^m$ and $T_2 : \mathbb{F}^r \to \mathbb{F}^n$. Then we define $T_1 \circ T_2 : \mathbb{F}^r \to \mathbb{F}^m$ by

$$(T_1 \circ T_2)(\vec{v}) = T_1(\underbrace{T_2(\vec{v})}_{\in \mathbb{F}^n}) \qquad \vec{v} \in \mathbb{F}^r$$

note order

Claim $[T_1 \circ T_2] = [T_1][T_2]$ ← matrix product

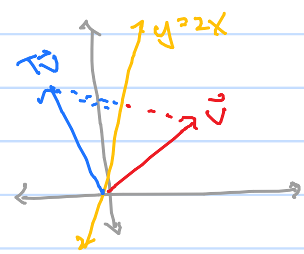
Proof: Denote $T = T_1 \circ T_2$ and let $A = [T_1]$, $B = [T_2]$. Recall that $[T]$ is simply the matrix whose columns are given by $T(\vec{e}_1), \ldots, T(\vec{e}_r)$.
For $j = 1, \ldots, r$ we have

$$T(\vec{e}_j) = T_1 \circ T_2(\vec{e}_j) = T_1(T_2(\vec{e}_j)) = T_1(B\vec{e}_j) = A(B\vec{e}_j)$$
$$= AB\vec{e}_j$$

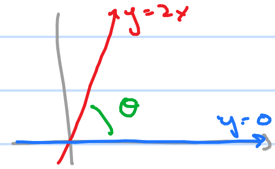
So the jth column of $[T]$ is $AB\vec{e}_j$, but this is precisely the jth column of AB. So $[T] = AB$ since they have the same columns. □

Ex Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation that sends $\vec{v} \in \mathbb{R}^2$ to its reflection over the line $y = 2x$.
Lets compute $[T]$. It suffices to compute $T(\vec{e}_1), T(\vec{e}_2)$ for

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$



but this is hard. Instead, we note that if the line were $y = 0$, reflecting would be much easier to compute: $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \mapsto \begin{pmatrix} v_1 \\ -v_2 \end{pmatrix}$.



But remember that rotation is a linear transformation, so we can first rotate so $y=2x$ becomes $y=0$, reflect, and then rotate back.

Let θ be the angle between $y=2x$ and $y=0$.

Let $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be rotation by θ radians counterclockwise.

Let $S : \mathbb{R}^2 \to \mathbb{R}^2$ be reflection over $y=0$.

Then

$$T = R_\theta \circ S \circ R_{-\theta}$$

So $[T] = [R_\theta][S][R_{-\theta}]$.

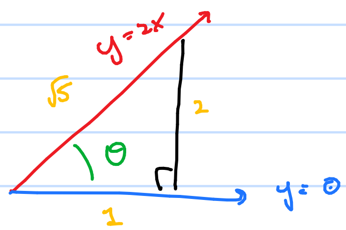
First we compute $[S]$:

$$S\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad S\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \quad \Rightarrow \quad [S] = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Fact:

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Note that



so $\cos\theta = \frac{1}{\sqrt{5}} = \frac{\sqrt{5}}{5}$

$\sin\theta = \frac{2}{\sqrt{5}} = \frac{2\sqrt{5}}{5}$

So $R_\theta = \frac{\sqrt{5}}{5}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$ and $R_{-\theta} = \begin{pmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \frac{\sqrt{5}}{5}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$

Hence

$$[T] = \frac{1}{5}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}$$

□

Properties of matrix multiplication

Matrix multiplication satisfies the following properties:

1. Associativity: $A(BC) = (AB)C$, provided that either the left or right side is well-defined; we therefore just write ABC.
2. Distributivity: $A(B+C) = AB + AC$
   $(A+B)C = AC + BC$
3. Commutativity with scalar multiplication: $A(\alpha B) = \alpha(AB) = (\alpha A)B$

Exercise: prove these properties.

These are just the usual multiplication properties that numbers satisfy, except multiplication is not commutative: generally $AB \neq BA$.

Ex

- $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$ while $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix}$.
- $\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ while $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ DNE □

Recall that the transpose A^T of a matrix A is found by turning the rows of A into the columns of A^T (or vice-versa): $(A^T)_{jk} = (A)_{kj}$.

Prop Let $A \in M_{m\times n}$ and $B \in M_{n \times r}$ be matrices. Then

$$(AB)^T = B^T A^T$$

Proof We simply appeal to the definition of matrix multiplication and the transpose:

$$((AB)^T)_{jk} = (AB)_{kj} = \sum_{\ell=1}^{n} (A)_{k\ell}(B)_{\ell j}$$

while

$$(B^T A^T)_{jk} = \sum_{\ell=1}^{n} (B^T)_{j\ell}(A^T)_{\ell k} = \sum_{\ell=1}^{n} (B)_{\ell j}(A)_{k\ell} = \sum_{\ell=1}^{n} (A)_{k\ell}(B)_{\ell j}$$

So the entries of $(AB)^T$ agree with the entries of $B^T A^T$, which means $(AB)^T = B^T A^T$. □

The Trace

Def For $A \in M_{n\times n}$ (a square matrix) its trace is the scalar

$$\operatorname{tr}(A) := (A)_{11} + (A)_{22} + \cdots + (A)_{nn} = \sum_{i=1}^{n} (A)_{ii}$$

- Observe that the trace defines a transformation $\operatorname{tr}: M_{n\times n} \longrightarrow \mathbb{F}$. It is in fact linear: for $A, B \in M_{n\times n}$ and scalars $\alpha, \beta \in \mathbb{F}$

$$\operatorname{tr}(\alpha A + \beta B) = \sum_{i=1}^{n} (\alpha A + \beta B)_{ii} = \sum_{i=1}^{n} (\alpha A)_{ii} + (\beta B)_{ii}$$

$$= \sum_{i=1}^{n} \alpha (A)_{ii} + \beta (B)_{ii} = \alpha \sum_{i=1}^{n} (A)_{ii} + \beta \sum_{i=1}^{n} (B)_{ii} = \alpha \operatorname{tr}(A) + \beta \operatorname{tr}(B).$$

Thm Let $A \in M_{m\times n}$ and $B \in M_{n\times r}$ be matrices. Then

$$\operatorname{tr}(AB) = \operatorname{tr}(BA)$$

Proof This is Exercise 6 on Homework 3. □

1.6 Invertible Transformations: Isomorphisms

- Recall that if $T: \mathbb{R} \to \mathbb{R}$ is linear, then for all $x \in \mathbb{R}$ $T(x) = ax$ for some $a \in \mathbb{R}$. Also recall that if $a \neq 0$ the "inverse" (reciprocal) of a is $\frac{1}{a}$. Define $S: \mathbb{R} \to \mathbb{R}$ by $S(x) = (\frac{1}{a})x$. Let's compute $T \circ S$ and $S \circ T$:

$$T \circ S(x) = T(\tfrac{1}{a}x) = a(\tfrac{1}{a}x) = x$$

$$S \circ T(x) = S(ax) = \tfrac{1}{a}(ax) = x$$

So $T \circ S$ and $S \circ T$ both equal the transformation $I: \mathbb{R} \to \mathbb{R}$ defined by $I(x) = x$. We call I the identity transformation, and can easily see it is linear.

- More generally we have:

Def For a vector space V, the <u>identity</u> (linear) transformation $I_V: V \to V$ is defined by $I_V(\vec{x}) = \vec{x}$. We will often just write I for I_V

- Note that the domain and target space of I are the same.

Ex Let $I: \mathbb{F}^n \to \mathbb{F}^n$ be the identity transformation. Let's compute $[I]$:

$$I(\vec{e}_1) = \vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad I(\vec{e}_2) = \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad I(\vec{e}_n) = \vec{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Thus

$$[I] = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

We will denote this matrix by I_n. □

- Let $T: V \to V$ be a linear transformation. Observe that

$$T \circ I(\vec{v}) = T(I(\vec{v})) = T(\vec{v})$$

$$I \circ T(\vec{v}) = I(T(\vec{v})) = T(\vec{v})$$

So $T \circ I = T = I \circ T$. Compare this to $t \cdot 1 = t = 1 \cdot t$ for $t \in \mathbb{R}$.

A special case of this fact is that for $A \in M_{n \times n}$

$$A I_n = A = I_n A.$$

<u>Exercise</u>: Use matrix multiplication to show the above equalities.

- Recall that in the first example above, we had $S \circ T = T \circ S = I$. More generally, we have:

Def Let $A: V \to W$ be a linear transformation. We say that A is

left invertible if there exists a linear transformation $B: W \to V$ such that

$$B \circ A = I,$$

where here $I = I_V$.

We say A is **right invertible** if there exists a linear transformation $C: W \to V$ such that

$$A \circ C = I,$$

where here $I = I_W$

We call B and C the **left** and **right inverses** of A, respectively.

Def A linear transformation $A: V \to W$ is **invertible** if it is both left and right invertible

Ex Consider the linear transformation $A: \mathbb{P}_3 \to \mathbb{P}_2$ defined by $A(p(x)) = p'(x)$.

Claim 1 A is right invertible

Proof We must find a lin. trans. $B: \mathbb{P}_2 \to \mathbb{P}_3$ so that $A \circ B = I_{\mathbb{P}_2}$. That is, for $p(x) \in \mathbb{P}_2$, we have

$$A \circ B(p(x)) = I(p(x))$$

$$A(B(p(x))) = p(x)$$

So $q(x) := B(p(x)) \in \mathbb{P}_3$ must be a polynomial s.t. $A(q(x)) = q'(x) = p(x)$. Thus $q(x)$ should be an anti-derivative of $p(x)$. If $p(x) = a_2x^2 + a_1x + a_0$, then

$$B(p(x)) = \tfrac{1}{3}a_2x^3 + \tfrac{1}{2}a_1x^2 + a_0x + C$$

for some scalar C. However, since we want B to be linear, we must choose $C = 0$. Indeed,

$$B(2x) = x^2 + C$$

$$\parallel$$

$$B(x+x) = B(x) + B(x) = \tfrac{1}{2}x^2 + C + \tfrac{1}{2}x^2 + C = x^2 + 2C$$

So $2C = C \Rightarrow C = 0$. So define B by

$$B(a_2x^2 + a_1x + a_0) = \tfrac{1}{3}a_2x^3 + \tfrac{1}{2}a_1x^2 + a_0x$$

(Exercise: Check that B satisfies the full def. of linear.

It is easy to see that $A(B(p(x))) = p(x)$ for all $p(x) \in \mathbb{P}_2$, so A is right invertible with right inverse B. □

Claim 2 A is not left invertible.

Proof We'll do a proof by contradiction. Suppose, towards a a contradiction, that $C: \mathbb{P}_2 \to \mathbb{P}_3$ was a left inverse for A:

$$C \circ A = I_{\mathbb{P}_3}$$

Then $C(A(p(x))) = p(x)$ for every $p(x) \in \mathbb{P}_3$. Consider $p(x) = 1$. Then $A(p(x)) = (1)' = 0$. Since C is linear, we must have

$$C(A(p(x)) = C(0) = 0 \neq p(x)$$

a contradiction. Thus the left inverse of A must not exist, and so A is not left invertible. This further implies A is not invertible. □ □

- The above example implies there are transformations that are right invertible, but not left invertible. Similarly, there are trans. that are left but not right invertible transformations (e.g. B)

Thm If a linear transformation $A: V \to W$ is invertible, then its left and right inverses B and C are unique and satisfy $B = C$

Proof By definition of the left and right inverses, we have

$$B \circ A = I_V \quad \text{and} \quad A \circ C = I_W$$

So

$$(B \circ A) \circ C = I_V \circ C$$
$$B \circ (A \circ C) = C$$
$$B \circ (I_W) = C$$
$$B = C$$

Now, if $B_1 : W \to V$ is another left inverse of A, then repeating the argument with B_1 instead of B gives $B_1 = C$. But since $C = B \Rightarrow B_1 = B$. Thus the left inverse is unique. A similar proof shows the right inverse is unique. □

- This theorem can be used to simplify our proof of Claim 2 in the previous example. Rather than showing a left inverse cannot exist, we just have to show it cannot be B from Claim 1.

Corollary A linear transformation $A: V \to W$ is invertible if and only if there exists a unique linear transformation, denoted A^{-1}, such that $A^{-1}: W \to V$ and

$$A^{-1}A = I_V \quad \text{and} \quad AA^{-1} = I_W.$$

A^{-1} is called the inverse of A.

Matrix Inverses

Def A matrix $\in M_{m \times n}$ is invertible (resp. left invertible, right invertible) if the linear transformation

$$F^n \ni \vec{x} \mapsto A\vec{x} \in F^m$$

is invertible (resp. left invertible, right invertible).

The above theorem says that if $A \in M_{m \times n}$ is invertible, then there is a unique matrix A^{-1} satisfying

$AA^{-1} = I$ and $A^{-1}A = I$. We, of course, call A^{-1} the inverse of A.

EX ① The identity matrix $I_n \in M_{n\times n}$ is invertible with $(I_n)^{-1} = I_n$. (Compare this to the reciprocal of 1 being 1).

② For $\theta \in [0, 2\pi)$ the rotation matrix

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

has inverse $(R_\theta)^{-1} = R_{-\theta}$. This clear from how the rotation is defined.
Exercise: verify using matrix multiplication: $R_\theta R_{-\theta} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $R_{-\theta}R_\theta = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

③ For $A = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $B = \begin{pmatrix} 1/2 & 1/2 \end{pmatrix}$ is a left inverse:

$$BA = \begin{pmatrix} \tfrac12 & \tfrac12 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \tfrac12 + \tfrac12 \end{pmatrix} = (1) = I,$$

But it isn't a right inverse since AB is not defined. So A is not right invertible

Thm If $A \in M_{m\times n}$ is invertible, then A is square (i.e. $n=m$).

Proof Since $A^{-1} \in M_{n\times m}$, we have

$$A^{-1}A = I_n \quad \text{and} \quad AA^{-1} = I_m$$

So using Exercise #6 on HW 3 we have:

$$n = Tr(I_n) = Tr(A^{-1}A) = Tr(AA^{-1}) = Tr(I_m) = m.$$ □

Properties of the inverse transformation

Thm If $A: V \to W$ and $B: U \to V$ are invertible linear transformations, then $A \circ B: V \to U$ is invertible with

$$(A\circ B)^{-1} = B^{-1} \circ A^{-1}$$

Proof We compute

$$(A\circ B)\circ(B^{-1}\circ A^{-1}) = A\circ(B\circ B^{-1})\circ A^{-1} = A\circ I \circ A^{-1} = A\circ A^{-1} = I$$

and

$$(B^{-1}\circ A^{-1})\circ(A\circ B) = B^{-1}\circ(A^{-1}\circ A)\circ B = B^{-1}\circ I\circ B = B^{-1}\circ B = I$$ □

Rem Be careful: if $A\circ B$ is invertible this does not imply A and B are invertible. In fact, in the above proof, we really only used that B is right inv. and A is left inv.

Thm If A is an invertible matrix, then A^T is invertible with $(A^T)^{-1} = (A^{-1})^T$!

Proof Recall $(AB)^T = B^T A^T$. So we have:

$$(A^{-1})^T A^T = (A A^{-1})^T = (I)^T = I$$

and

$$A^T (A^{-1})^T = (A^{-1} A)^T = (I)^T = I.$$ □

Thm If $A: V \to W$ is invertible, then so is A^{-1} with $(A^{-1})^{-1} = A$.

Proof The same equations that show A^{-1} is the inverse of A also show $A = (A^{-1})^{-1}$. □

9/19

Isomorphism and Isomorphic Spaces

Def Given two vector spaces V and W (over the same field $\mathbb{F}$), we say V and W are isomorphic, and write $V \cong W$, if there exists an invertible linear transformation $T: V \to W$. We call T an isomorphism.

- When V and W are isomorphic, this means they are effectively the same vector space, just with another name. Let's see some evidence of this:

Thm Let $T: V \to W$ be an isomorphism, and let $\vec{v}_1, \dots, \vec{v}_n \in V$ be a generating (resp. linearly independent) system. Then $T(\vec{v}_1), \dots, T(\vec{v}_n) \in W$ is a generating (resp. linearly independent) system. In particular, if $\vec{v}_1, \dots, \vec{v}_n$ is a basis for V, then $T(\vec{v}_1), \dots, T(\vec{v}_n)$ is a basis for W.

Proof Homework 4 □

Thm Let $T: V \to W$ be a linear transformation, and let $\vec{v}_1, \dots, \vec{v}_n \in V$ be a basis. If $T(\vec{v}_1), \dots, T(\vec{v}_n)$ is a basis for W, then T is an isomorphism.

Proof Denote

$$\vec{w}_1 := T(\vec{v}_1), \dots, \vec{w}_n := T(\vec{v}_n)$$

Recall that any linear transformation is defined by its outputs on a basis. Thus we can define a lin. trans. $S: W \to V$ by $S(\vec{w}_j) := \vec{v}_j$. It follows that

$$T \circ S(\vec{w}_j) = T(\vec{v}_j) = \vec{w}_j = I_W(\vec{w}_j)$$

$$S \circ T(\vec{v}_j) = S(\vec{w}_j) = \vec{v}_j = I_V(\vec{v}_j)$$

Since $T \circ S$ and $S \circ T$ are determined by their outputs on the bases $\vec{w}_1, \dots, \vec{w}_n$ and $\vec{v}_1, \dots, \vec{v}_n$, respectively, it follows that $T \circ S = I_W$ and $S \circ T = I_V$. That is, $S = T^{-1}$ and so T is an isomorphism. □

Cor $A \in M_{n \times n}$ is invertible if and only if its columns form a basis for $\mathbb{F}^n$.

Proof The columns of A are the vectors $A(\vec{e}_1), \ldots, A(\vec{e}_n)$. So the previous two theorems yield ($\Rightarrow$) and ($\Leftarrow$), respectively. □

Ex ① $\mathbb{P}_n \cong \mathbb{F}^{n+1}$. Define $T: \mathbb{F}^{n+1} \to \mathbb{P}_n$ by

$$T(\vec{e}_1) = 1, \quad T(\vec{e}_2) = x, \quad \ldots, \quad T(\vec{e}_{n+1}) = x^n$$

Since $1, x, \ldots, x^n$ is a basis for $\mathbb{P}_n$, the previous theorem implies T is an iso.

② Let V be any real vector space with basis $\vec{v}_1, \ldots, \vec{v}_n \in V$. Then $V \cong \mathbb{R}^n$ by the iso

$$T(\vec{e}_1) = \vec{v}_1, \ldots, T(\vec{e}_n) = \vec{v}_n.$$

Similarly, if V is a complex vector space with basis $\vec{v}_1, \ldots, \vec{v}_n$, then $V \cong \mathbb{C}^n$.

③ $M_{2\times 3} \cong \mathbb{F}^6$ $(= 2\cdot 3)$ since $\begin{pmatrix}1&0&0\\0&0&0\end{pmatrix}, \begin{pmatrix}0&1&0\\0&0&0\end{pmatrix}, \begin{pmatrix}0&0&1\\0&0&0\end{pmatrix}, \begin{pmatrix}0&0&0\\1&0&0\end{pmatrix}, \begin{pmatrix}0&0&0\\0&1&0\end{pmatrix}, \begin{pmatrix}0&0&0\\0&0&1\end{pmatrix}$ is a basis.

Rem In general, $M_{m\times n} \cong \mathbb{F}^{m\cdot n}$ and an isomorphism is defined by:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \longmapsto \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \\ a_{12} \\ \vdots \\ a_{m2} \\ \vdots \\ a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

i.e. stack the columns of A on top of each other

However, this forgets some important information about $M_{m\times n}$: mult. and transpose.

Thm Let V be a vector space, and suppose $\vec{v}_1, \ldots, \vec{v}_n$ and $\vec{w}_1, \ldots, \vec{w}_m$ are both bases for V. Then $n = m$.

Proof By the above example, we have $V \cong \mathbb{F}^n$, say with iso. $T: V \to \mathbb{F}^n$, and $V \cong \mathbb{F}^m$ say with iso. $S: V \to \mathbb{F}^m$. Then $S \circ T^{-1}: \mathbb{F}^n \to \mathbb{F}^m$ is an isomorphism, and therefore $[S \circ T^{-1}] \in M_{m\times n}$ is invertible. Since only square matrices are invertible, we have $n = m$. □

- So any two finite bases of a vector space have the same size.

Def Let V be a vector space and let $\vec{v}_1, \ldots, \vec{v}_n$ be any basis for V. The <u>dimension</u> of V, denoted $\dim(V)$, is the number n.

Invertibility and Solving Equations

Thm Let $A: X \to Y$ be a linear transformation. Then A is invertible if and only if for every $\vec{b} \in Y$ the equation
$$A(\vec{x}) = \vec{b}$$
has a unique solution.

Proof ($\Longrightarrow$) Suppose A is invertible. For $\vec{b} \in Y$, $\vec{x} := A^{-1}(\vec{b})$ solves the equation. Moreover, if $\vec{x}_1$ is another solution then we have:
$$A(\vec{x}_1) = \vec{b}$$
$$A^{-1}(A(\vec{x}_1)) = A^{-1}(\vec{b})$$
$$\vec{x}_1 = \vec{x}.$$
So the solution is unique.

($\Longleftarrow$) Suppose for any $\vec{b} \in Y$, $A(\vec{x}) = \vec{b}$ has a unique solution. Define $B: Y \to X$ by letting $B(\vec{y}) \in X$ be the unique solution of $A(\vec{x}) = \vec{y}$.

We claim B is the inverse of A. First, let us verify that it is linear. For $\vec{y}_1, \vec{y}_2 \in Y$ let $\vec{x}_1 = B(\vec{y}_1)$ and $\vec{x}_2 = B(\vec{y}_2)$. Then by definition of B we have:
$$A(\vec{x}_1) = \vec{y}_1$$
$$A(\vec{x}_2) = \vec{y}_2.$$
Now, for any scalars α, β we have by linearity of A that
$$A(\alpha\vec{x}_1 + \beta\vec{x}_2) = \alpha A(\vec{x}_1) + \beta A(\vec{x}_2) = \alpha\vec{y}_1 + \beta\vec{y}_2.$$
So $\alpha\vec{x}_1 + \beta\vec{x}_2$ must be the unique solution to
$$A(\vec{x}) = \alpha\vec{y}_1 + \beta\vec{y}_2.$$
Hence
$$B(\alpha\vec{y}_1 + \beta\vec{y}_2) = \alpha\vec{x}_1 + \beta\vec{x}_2 = \alpha B(\vec{y}_1) + \beta B(\vec{y}_2).$$
And so B is linear.

Finally, we check $A \circ B = I$ and $B \circ A = I$. Let $\vec{x} \in A$ and set $\vec{y} := A(\vec{x})$. Then by def. of B we have
$$\vec{x} = B(\vec{y}) = B(A(\vec{x})) = B \circ A(\vec{x}).$$
Similarly, if $\vec{y} \in Y$, then $\vec{x} := B(\vec{y})$ solves $A(\vec{x}) = \vec{y}$. Hence
$$\vec{y} = A(\vec{x}) = A(B(\vec{y})) = A \circ B(\vec{y}).$$
Thus $B = A^{-1}$ as claimed. $\square$

1.7 Subspaces

Def A <u>subspace</u> of a vector space V is a subset $V_0 \subset V$ satisfying:

1. $\vec{0} \in V_0$
2. For every $\vec{v}, \vec{w} \in V_0$, $\vec{v} + \vec{w} \in V_0$ (V_0 is closed under addition)
3. For every $\vec{v} \in V_0$ and every scalar α, $\alpha \vec{v} \in V_0$ (V_0 is closed under scalar multiplication)

- A subspace $V_0 \subset V$ is in particular a vector space in its own right. Indeed, all the axioms are satisfied because they are satisfied for V. Thus subspaces can very easily give us lots of new examples of vector spaces, since we only need to check 1-3 above, rather than all of the vector space axioms.

EX ① For a vector space V, $V_0 = \{\vec{0}\}$ is a subspace called the trivial subspace. It is the smallest subspace in any vector space. ($\phi \subset V$ is <u>not</u> a subspace because it fails 1.)
$V \subset V$ is also a subspace (the biggest subspace).

② Let $T: V \to W$ be a linear transformation.

- $\text{Null}(T) = \{\vec{v} \in V : T(\vec{v}) = \vec{0}\}$ is a subspace of V called the <u>null space</u> of T. (also called the <u>kernel</u> of T and denoted $\text{Ker}(T)$).

- $\text{Ran}(T) = \{\vec{w} \in W : \exists \vec{v} \in V \text{ such that } T(\vec{v}) = \vec{w}\}$ is a subspace of W called the <u>range</u> of T.

③ Let $\vec{v}_1, \ldots, \vec{v}_r \in V$. The <u>span</u> of $\vec{v}_1, \ldots, \vec{v}_n$ is the set

$$\text{span}\{\vec{v}_1, \ldots, \vec{v}_n\} = \{\alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n : \alpha_1, \ldots, \alpha_n \text{ are scalars}\},$$

and it is a subspace of V.