Chapter 5: Linear transformations

Linear transformations are what make vector spaces useful.

Operations like differentiation, integration, as well as geometrical transformations like rotation, reflection, shear, etc. are all examples of linear transformations.

⟶ Applications in calculus (& therefore in physics, engineering, etc.) & computer graphics, etc. etc.

The archetypal example of a linear transformation is

- multiplication by a (fixed) matrix.

Section 5.1 : Definition & examples

<u>Terminology 5.1</u> A function $L : V \to W$, where V and W are two vector spaces, is called a <u>linear transformation</u> if

- $L(u+v) = L(u) + L(v)$ for all $u, v \in V$

& • $L(\alpha v) = \alpha L(v)$ for all scalars α, and $v \in V$.

<u>Example</u> If A is an $m \times n$ matrix, then the map $x \to Ax$ is a function from $\mathbb{R}^n$ to $\mathbb{R}^m$. $\begin{pmatrix} x \in \mathbb{R}^{n\times 1} \\ Ax \in \mathbb{R}^{m\times 1} \end{pmatrix}$

It is a linear transformation because:

- $A(x+y) = Ax + Ay$
- $A(\alpha x) = \alpha(Ax)$

Example 5.3 $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad Ax = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$$

So the map is $L: \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longrightarrow \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$

Check linearity:

- $$L\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = L\left(\begin{pmatrix} x_1+y_1 \\ x_2+y_2 \end{pmatrix}\right) = \begin{pmatrix} x_1+y_1 \\ 0 \end{pmatrix}$$
$$= \begin{pmatrix} x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} y_1 \\ 0 \end{pmatrix} = L\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) + L\left(\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right).$$

- $$L\left(\alpha \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = L\left(\begin{pmatrix} \alpha x_1 \\ \alpha x_2 \end{pmatrix}\right) = \begin{pmatrix} \alpha x_1 \\ 0 \end{pmatrix} = \alpha \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$$
$$= \alpha L\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right).$$

There L is a <u>linear transformation</u>.

Example 5.5 Let $M : \mathbb{R}^2 \to \mathbb{R}^1$ be the map $M\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \sqrt{x_1^2 + x_2^2}$

Is this linear? No, because

$$M\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = 1. \qquad M\left(\begin{pmatrix} -1 \\ 0 \end{pmatrix}\right) = \sqrt{(-1)^2} = \sqrt{1} = 1.$$

i.e. $M\left(-\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = 1 \neq -M\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = -1$

Alternatively: $M\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = 1$, $M\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = 1$

but $M\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = M\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = \sqrt{2} \neq 2 = M\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) + M\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$.

We shall show (eventually) that <u>every</u> linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ <u>is</u> just <u>multiplication</u> by a <u>suitable matrix</u>.

Section 5.2 : More examples and basic properties

Observation 5.6 If $L : V \to W$ is a linear transformation between two vector space V, W, then

- $L(\underset{\sim}{0}_V) = \underset{\sim}{0}_W$ $\quad (\underset{\sim}{0}_V$ = zero vector in $V)$ $\quad (\underset{\sim}{0}_W$ = zero vector in $W)$
- $L(-\underset{\sim}{v}) = -L(\underset{\sim}{v})$
- $L\left(\sum_{i=1}^{n} \alpha_i \underset{\sim}{v}_i\right) = \sum_{i=1}^{n} \alpha_i L(\underset{\sim}{v}_i)$

Why?
- $\underset{\sim}{0} = 0\underset{\sim}{v}$ for any $\underset{\sim}{v} \in V$.

  $\Rightarrow L(\underset{\sim}{0}) = L(0\underset{\sim}{v}) = 0.L(\underset{\sim}{v}) = \underset{\sim}{0}$.

- $(-1)\underset{\sim}{v} = -\underset{\sim}{v}$ (see earlier)

  $\Rightarrow L(-v) = L((-1)v) = (-1) L(v)$
  $= -L(v)$.

- First $L(\alpha_i \underset{\sim}{v}_i) = \alpha_i L(\underset{\sim}{v}_i)$ for each i

  & then $L(\sum \alpha_i v_i) = \sum \alpha_i L(v_i)$

  by repeated use of

  $L(u+w) = L(u) + L(w)$.

Example 5.7 (The definite integral)

Define $L : C[a,b] \longrightarrow \mathbb{R}$ by

$$L(f) = \int_a^b f(x)\,dx$$

Is L a linear transformation?

- $$L(f+g) = \int_a^b (f+g)(x)\,dx$$
  $$= \int_a^b \big(f(x) + g(x)\big)\,dx \quad \Leftarrow \text{calculus}$$
  $$= \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$$
  $$= L(f) + L(g).$$

- $$L(\alpha f) = \int_a^b (\alpha f)(x)\,dx$$
  $$= \int_a^b \alpha . f(x)\,dx \quad \Leftarrow \text{calculus}$$
  $$= \alpha \int_a^b f(x)\,dx$$
  $$= \alpha\, L(f).$$

$\Rightarrow$ L is a linear transformation.

Example 5.8 (Differentiation)

Let $C'[a,b]$ denote the space of functions $[a,b] \longrightarrow \mathbb{R}$ which are differentiable with continuous derivate.

Then $D : C'[a,b] \longrightarrow C[a,b]$

defined by $D(f) = f'$

$\left(\text{where } f'(x) = \frac{df}{dx}\right)$

is a linear transformation.

<u>Why?</u> • $D(f+g) = (f+g)'$

$= f' + g' = D(f) + D(g).$

• $D(\alpha f) = (\alpha f)' = \alpha f'$

real scaler (constant) $= \alpha D(f).$

Section 5.3 : Image and Kernel (KERNEL)

Many of the concepts we studied for <u>matrices</u> can be easily <u>translated</u> into corresponding concepts for linear transformations.

- Nullspace $\rightsquigarrow$ kernel. $N(A) = \{x \mid Ax = 0\}$

~~• Image $\rightsquigarrow$~~

- Column space $\rightsquigarrow$ range (also called image)

<u>Terminology 5.10</u> If $L : V \to W$ is a linear transformation between vector spaces V, W, the <u>kernel</u> of L is the subset of V given by

$$\ker(L) = \{v \in V \mid L(v) = 0\}$$

Example 5.11 Pick $A \in \mathbb{R}^{m \times n}$ and define $L: \mathbb{R}^n \to \mathbb{R}^m$ by

$$L(x) = Ax.$$

Then $\ker(L) = N(A)$.

Why? $\ker L = \{x \in \mathbb{R}^n \mid L(x) = 0\}$

$= \{x \in \mathbb{R}^n \mid Ax = 0\}$

$= N(A)$. **Done**.

We know that $N(A)$ is a **subspace** of $\mathbb{R}^n$.

Is the same thing true for **kernels** in general?

Theorem 5.14 (a) $\ker L$ is a subspace of V (where $L: V \to W$ is a linear transformation).

Why?
- $L(0) = 0$ so $0 \in \ker L$
- if $u, v \in \ker L$
  then $L(u) = 0$, $L(v) = 0$
  $\Rightarrow L(u+v) = L(u) + L(v) = 0 + 0 = 0$
  $\Rightarrow u + v \in \ker L$
- if α is a scalar & $u \in \ker L$
  then $L(\alpha u) = \alpha L(u) = \alpha 0 = 0$
  $\Rightarrow \alpha u \in \ker L$.

Done.

Example Compute $\ker L$ when

$L : P_3 \to P_2$ is differentiation

$(L(p) = p')$.

Solution: $\ker L = \{ p \mid p' = 0 \}$

$= \{$ polynomials of degree $0 \}$

$= \{$ constant polynomials $\}$.

Terminology 5.12

If $L : V \to W$ is a linear transformation, then the image of V is the range of L, that is $L(V) = \{w \in W \mid w = L(v) \text{ for some } v \in V\}$

More generally, if H is a subspace of V, then $L(H) = \{w \in W \mid w = L(v) \text{ for some } v \in H\}$ is the image of H.

Example 5.13

If $A \in \mathbb{R}^{m \times n}$ and $L : \mathbb{R}^n \to \mathbb{R}^m$ is the map

$$L(x) = Ax$$

then the range of L is

$$L(\mathbb{R}^n) = \text{col}(A).$$

Why? If $A = \begin{pmatrix} \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_n \end{pmatrix}$ as n columns

& $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

then $L(x) = Ax = (\underline{a}_1 \cdots \underline{a}_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

$$= x_1\underline{a}_1 + x_2\underline{a}_2 + \cdots + x_n\underline{a}_n$$

So $\{L(x) \mid x \in \mathbb{R}^n\} = \text{Span}(\underline{a}_1, \ldots, \underline{a}_n)$

$$= \text{col}(A).$$

Theorem 5.14 (b) If $L: V \to W$ is a linear transformation, and H is a subspace of V, then $L(H)$ is a <u>subspace</u> of W.

<u>Why?</u>

- $0 \in H$ and $L(0) = 0$
  $\Rightarrow 0 \in L(H)$. $\Rightarrow L(H)$ is nonempty

- if $w_1, w_2 \in L(H)$
  then $w_1 = L(v_1)$ and $w_2 = L(v_2)$
  for some $v_1, v_2 \in H$
  $\Rightarrow w_1 + w_2 = L(v_1) + L(v_2)$
  $= L(v_1 + v_2)$ (since L is linear)
  and $v_1 + v_2 \in H$
  So $w_1 + w_2 \in L(H)$.

- if $w \in L(H)$ and α is a scalar,
  then $w = L(v)$ for some $v \in H$
  $\Rightarrow \alpha v \in H$ (since H is a subspace)
  $\Rightarrow \alpha w = \alpha L(v) = L(\alpha v)$ (since L linear)
  $\Rightarrow \alpha w \in L(H)$.

Example 5.16 (Differential equations)

check
$L(f+g) = L(f)$
$\overset{\shortparallel}{f+g+f'+g'} = + L(g)$

$L(\alpha f) = \alpha f + \alpha f'$
$= \alpha L(f).$

Let $L : C^1[-1,1] \longrightarrow C[-1,1]$
be the map $L(f) = f + f'$.
Find the <u>kernel</u> and <u>range</u> of L.

- The <u>kernel</u> of L is (by definition)

  $$\{ f \in C^1[-1,1] \mid f + f' = 0 \}$$

  that is, the set of <u>solutions</u> to this <u>homogeneous</u> <u>linear</u> <u>differential</u> <u>equation</u>.

  Use your favourite method to get the general solution

  $$f(t) = Ae^{-t}$$

  Thus $\ker L = \{ f \mid f(t) = Ae^{-t}, A \in \mathbb{R} \}$
  $= \mathrm{Span}(h)$

  where h is the function

  $$h(t) = e^{-t}.$$

  In particular $\dim(\ker L) = 1$.

- The crucial question about the <u>range</u> of L is, is it the whole of the codomain $C[-1,1]$?

In other words, given $g \in C[-1,1]$, is there a function $f \in C^1[-1,1]$ such that
$$L(f) = g \;?$$
Or: is there a solution to the differential equation $f + f' = g$?

<u>Answer</u>: $\quad e^t f(t) + e^t f'(t) = e^t g(t)$
$$\Rightarrow \frac{d}{dt}\left(e^t f(t)\right) = e^t g(t) \in C[-1,1]$$
Now R.H.S. has an anti-derivative $H(t)$ (indefinite integral)

so $\frac{d}{dt}(H(t)) = e^t g(t)$
$$\Rightarrow \quad e^t f(t) = H(t) + C \qquad \left(\text{Some constant } C\right)$$
$$\Rightarrow \quad f(t) = e^{-t}\left(H(t) + C\right)$$
<u>Conversely</u>, this function f does satisfy the equation $f + f' = g$.
$$\left[\, f(t) + f'(t) = e^{-t}\frac{d}{dt}H(t) - e^{-t}(H(t)+C) + e^{-t}(H(t)+C) = e^{-t}\frac{d}{dt}(H(t)) = g(t). \right.$$

Therefore the range of L is the whole of $C[-1,1]$.

Section 5.4: Matrix representations of linear transformations

We show that every linear transformation between finite-dimensional vector spaces can be represented by a matrix.

(This doesn't work for spaces of functions, which are usually infinite-dimensional, so does not directly apply to differential equations. However "finite element methods" approximate linear differential operators by matrices, and are widely used in practice.)

Basic Example / Theorem 5.17

If $L: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation,
then $L(x) = Ax$ (for all $x \in \mathbb{R}^n$)
for some fixed matrix $A \in \mathbb{R}^{m \times n}$.

Why? How do we find A?

- Putting $\underline{x} = \underline{e}_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \end{pmatrix}$ ← j'th coordinate

gives $L(e_j) =$ j'th column of A $(= \underline{a}_j$, say$)$.

So this tells us how to write down A:

$$A = \left(L(e_1), L(e_2), \dots L(e_n) \right)$$

$$\Rightarrow A\underline{x} = \left(L(e_1) \ \dots \ L(e_n) \right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$= x_1 L(\underline{e}_1) + \dots + x_n L(\underline{e}_n)$

$= L(x_1\underline{e}_1 + \dots + x_n\underline{e}_n)$ since L is linear

$= L(\underline{x})$ as required.

where $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1\underline{e}_1 + \dots + x_n\underline{e}_n$.

Example 5.18

If $L : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by

$$L : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} x_1 - x_2 \\ x_2 + 2x_3 \end{pmatrix}$$

then (L is linear and)

$$L(e_1) = L\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$L(e_2) = L\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$L(e_3) = L\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

So if we let $A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$

then we find

$$Ax = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_2 + 2x_3 \end{pmatrix} = L(x).$$

The general case 5.19

(Matrix Representation Theorem)

If $\mathcal{B} = \{v_1, \dots, v_n\}$ is a basis for V,
and $\mathcal{C} = \{w_1, \dots, w_m\}$ is a basis for W,
then for each linear transformation $L: V \to W$
there is a corresponding $m \times n$ matrix A such that

$$[L(v)]_{\mathcal{C}} = A.[v]_{\mathcal{B}}$$

where $[v]_{\mathcal{B}}$ give the $\mathcal{B}$-coordinates of v
and $[w]_{\mathcal{C}}$ " " $\mathcal{C}$ " " w.

Indeed, the j'th column of A is

$$[L(v_j)]_{\mathcal{C}} . \left(= A.[v_j]_{\mathcal{B}} = A \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j \right)$$

Recall Defn. 4.47

If $\mathcal{B} = \{\underline{b}_1, \dots, \underline{b}_n\}$ is a basis for V, & $\underline{v} = c_1 \underline{b}_1 + \dots + c_n \underline{b}_n$ then

$$[\underline{v}]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

& $c_1, \dots, c_n$ are the <u>coordinates</u> of $\underline{v}$ relative to $\mathcal{B}$.

Proof of 5.19

- Let $\underline{a}_j = [L(\underline{v}_j)]_{\mathcal{C}} = (a_{1j}, a_{2j}, \dots, a_{mj})^T$

  so that $L(\underline{v}_j) = a_{1j} \underline{w}_1 + \dots + a_{mj} \underline{w}_m = \sum_{i=1}^{m} a_{ij} \underline{w}_i$

- Pick $\underline{v} \in V$ and let $\underline{x} = (x_1, \dots, x_n)^T = [\underline{v}]_{\mathcal{B}}$

  so that $\underline{v} = x_1 \underline{v}_1 + \dots + x_n \underline{v}_n$

- Now calculate $L(\underline{v}) =$

$$L(x_1 \underline{v}_1 + \dots + x_n \underline{v}_n) =$$

$$x_1 L(\underline{v}_1) + \dots + x_n L(\underline{v}_n) = \sum_{j=1}^{n} x_j L(\underline{v}_j)$$

$$= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} \underline{w}_i$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right) \underline{w}_i$$

- Hence the $\mathcal{C}$-coordinates of $L(\underline{v})$ are

$$\sum_{j=1}^{n} a_{ij} x_j \qquad (= y_i \text{, say})$$

that is $[L(\underline{v})]_{\mathcal{C}} = (y_1, \dots, y_m)^T$. $(= \underline{y}\text{, say})$

- Moreover, we have calculated that

$$\underline{y} = A \underline{x}$$

that is $[L(\underline{v})]_{\mathcal{C}} = A . [\underline{v}]_{\mathcal{B}}$

<u>as required.</u>

<u>Notation 5.20</u> We write

$$[L]_{\mathcal{C}}^{\mathcal{B}} \quad \text{for the matrix } A \text{ of Theorem 5.19.}$$

It is called the <u>matrix representation</u> of L <u>with respect to</u> the bases $\mathcal{B}$ and $\mathcal{C}$.

Example 5.21

Let $\mathcal{B} = \{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ be the standard basis for $\mathbb{R}^3$,

& let $\mathcal{C} = \{\underline{w}_1, \underline{w}_2\}$ where $\underline{w}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\underline{w}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$,

so that $\mathcal{C}$ is a basis for $\mathbb{R}^2$.

Define $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ by

$$L\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = x_1 \underline{w}_1 + (x_2 + x_3)\underline{w}_2$$

Find $[L]^{\mathcal{B}}_{\mathcal{C}}$

<u>Solution</u>

$$L(\underline{e}_1) = 1.\underline{w}_1 + 0.w_2 \Rightarrow [L(\underline{e}_1)]_{\mathcal{C}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$L(\underline{e}_2) = \underline{w}_2 \qquad \Rightarrow [L(\underline{e}_2)]_{\mathcal{C}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$L(\underline{e}_3) = \underline{w}_2 \qquad \Rightarrow [L(\underline{e}_3)]_{\mathcal{C}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\Rightarrow [L]^{\mathcal{B}}_{\mathcal{C}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Example 5.22 (Using the Matrix Representation Theorem).

The differentiation map on polynomial of degree ≤ 2

$$D : P_2 \longrightarrow P_1$$

is defined by $D(p)(t) = p'(t)$.

Basis $\mathcal{B}_2$ for P_2 and $\mathcal{B}_1$ for P_1:

$\mathcal{B}_2 = \{ \underline{p}_1, \underline{p}_2, \underline{p}_3 \}$ where $\underline{p}_1(t) = 1$, $\underline{p}_2(t) = t$, $\underline{p}_3(t) = t^2$

$$\mathcal{B}_1 = \{ \underline{p}_1, \underline{p}_2 \}.$$

$$D(\underline{p}_1) = \underline{0} = 0\underline{p}_1 + 0\underline{p}_2$$

$$D(\underline{p}_2) = \underline{p}_1 = 1\underline{p}_1 + 0\underline{p}_2$$

$$D(\underline{p}_3) = 2\underline{p}_2 = 0\underline{p}_1 + 2\underline{p}_2$$

$$[D]^{\mathcal{B}_2}_{\mathcal{B}_1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Given any $\underline{p}(t) = a + bt + ct^2$

$$\Rightarrow \underline{p} = a\underline{p}_1 + b\underline{p}_2 + c\underline{p}_3$$

i.e. $[\underline{p}]_{\mathcal{B}_2} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$

Matrix Representation Theorem tells us

$$[D(p)]_{\mathcal{B}_1} = [D]_{\mathcal{B}_1}^{\mathcal{B}_2} [p]_{\mathcal{B}_2}$$

$$= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

$$= \begin{pmatrix} b \\ 2c \end{pmatrix}$$

i.e. $D(\underline{p}) = b\,\underline{p}_1 + 2c\,\underline{p}_2$.

$\left[\ p'(t) = \frac{d}{dt}(a + bt + ct^2) = b + 2ct\ \right]$.

Section 5.5: Composition of linear transformations

If $T : U \to V$

and $S : V \to W$

are linear transformations

then $S \circ T : U \to W$

defined by $(S \circ T)(u) = S(T(u))$

is a linear transformation. (the composite of S and T.)

Proof • if $u_1, u_2 \in U$

then $T(u_1 + u_2) = T(u_1) + T(u_2)$

$$\Rightarrow S(T(u_1+u_2)) = S(T(u_1) + T(u_2))$$

$$= S(T(u_1)) + S(T(u_2))$$

i.e. $(S \circ T)(u_1+u_2) = (S \circ T)(u_1) + (S \circ T)(u_2)$.

• if $u \in U$, $\alpha \in \mathbb{R}$

Then $T(\alpha u) = \alpha T(u)$ $\because$ T linear

$$\Rightarrow (S \circ T)(\alpha u) = S(T(\alpha u))$$

$$= S(\alpha . T(u))$$

$$= \alpha . S(T(u))$$

$$= \alpha . (S \circ T)(u)$$

Done.

Composition of linear transformations corresponds to **multiplication** of matrices.

(This is the **real** reason why multiplication of matrices is defined in the way it is.)

Example 5.24 If $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times k}$

we have linear maps

$$L_A : \mathbb{R}^n \to \mathbb{R}^m, \quad L_A(\underline{x}) = A\underline{x}$$

and $L_B : \mathbb{R}^k \to \mathbb{R}^n, \quad L_B(\underline{y}) = B\underline{y}$

So
$$\begin{aligned}
(L_A \circ L_B)(\underline{y}) &= L_A(L_B(\underline{y})) \\
&= L_A(B\underline{y}) \\
&= A(B\underline{y}) \\
&= (AB)\underline{y} \qquad \text{(associativity)} \\
&= L_{AB}(\underline{y}) \qquad \text{(defn. of } L_{AB})
\end{aligned}$$

that is $L_A \circ L_B = L_{AB}$ as linear maps.

composite   matrix multiplication

The same applies in the general case:

Theorem 5.27 (Composition formula for linear transformations)

If $T: U \to V$
and $S: V \to W$
are linear transformations,
and $\mathcal{B}$ is a basis for U
$\mathcal{C}$ is a basis for V
& ~~$\mathcal{B}$~~ $\mathcal{D}$ is a basis for W,

then $[S \circ T]^{\mathcal{B}}_{\mathcal{D}} = [S]^{\mathcal{C}}_{\mathcal{D}} [T]^{\mathcal{B}}_{\mathcal{C}}$.

Proof ~~Compute $[(S \circ T)(\underline{u})]_{\mathcal{D}}$~~
Write $A = [S]^{\mathcal{C}}_{\mathcal{D}}$, $B = [T]^{\mathcal{B}}_{\mathcal{C}}$,

& pick $\underline{u} \in U$, write $\underline{v} = T(\underline{u}) \in V$
& $\underline{w} = S(\underline{v}) \in W$.

$$[(S \circ T)(\underline{u})]_{\mathcal{D}} = [S(T(\underline{u}))]_{\mathcal{D}} = [S(\underline{v})]_{\mathcal{D}}$$

$$= [S]^{\mathcal{C}}_{\mathcal{D}} [\underline{v}]_{\mathcal{C}} \quad (\text{defn. of } [S]^{\mathcal{C}}_{\mathcal{D}})$$

Also $[\underline{v}]_{\mathcal{C}} = [T(\underline{u})]_{\mathcal{C}} = [T]^{\mathcal{B}}_{\mathcal{C}} [\underline{u}]_{\mathcal{B}}$

$$\Rightarrow \left[(S\circ T)(\underline{u})\right]_{\mathcal{D}} = [S]_{\mathcal{D}}^{\mathcal{C}}\,[T]_{\mathcal{C}}^{\mathcal{B}}\,[\underline{u}]_{\mathcal{B}}\,.$$

But $[S\circ T]_{\mathcal{D}}^{\mathcal{B}}$ is defined by

$$\left[(S\circ T)(\underline{u})\right]_{\mathcal{D}} = [S\circ T]_{\mathcal{D}}^{\mathcal{B}}\,[\underline{u}]_{\mathcal{B}}$$

$$\Rightarrow \underbrace{[S\circ T]_{\mathcal{D}}^{\mathcal{B}}}\,[\underline{u}]_{\mathcal{B}} = \underbrace{[S]_{\mathcal{D}}^{\mathcal{C}}\,[T]_{\mathcal{C}}^{\mathcal{B}}}\,[\underline{u}]_{\mathcal{B}}$$

So these two matrices represent the <u>same</u> linear transformation.

$\Rightarrow$ they are equal as matrices.

i.e. $[S\circ T]_{\mathcal{D}}^{\mathcal{B}} = [S]_{\mathcal{D}}^{\mathcal{C}}\,[T]_{\mathcal{C}}^{\mathcal{B}}$ <u>As required</u>.

Section 5.6 : Change of basis

In Section 4.6 we saw what happens to vectors when we change basis.

Now we look at what happens to matrices representing linear transformations.

First look at a single vector space V, with two bases

$$\mathcal{B} = \{u_1, \ldots, u_n\}$$

$$\& \ \mathcal{C} = \{v_1, \ldots, v_n\}$$

and the identity map $Id : V \to V$,

$$Id(\underline{v}) = \underline{v}.$$

Then

$$[\underline{v}]_{\mathcal{C}} = [Id(\underline{v})]_{\mathcal{C}}$$

$$= [Id]_{\mathcal{C}}^{\mathcal{B}} [\underline{v}]_{\mathcal{B}}$$

by Matrix Representation Theorem

The matrix $[Id]_{\mathcal{C}}^{\mathcal{B}}$ is called the transition matrix from $\mathcal{B}$ to $\mathcal{C}$.

(We ~~do~~ already did the case $V = \mathbb{R}^n$ in Section 4.6)

Observation 5.30 $[Id]^{\mathcal{B}}_{\mathcal{C}}$ is invertible,

& its inverse is $[Id]^{\mathcal{C}}_{\mathcal{B}}$.

Proof By the composition formula

$$[Id]^{\mathcal{B}}_{\mathcal{C}}[Id]^{\mathcal{C}}_{\mathcal{B}} = [Id \circ Id]^{\mathcal{C}}_{\mathcal{C}} = [Id]^{\mathcal{C}}_{\mathcal{C}}$$

$$= I,$$

& similarly the other way round.

Done.

Example 5.31 Consider the following bases of P_2:

$\mathcal{B} = \{\underline{p}_1, \underline{p}_2, \underline{p}_3\}$ where $\underline{p}_1(t) = 1$, $\underline{p}_2(t) = t$, $\underline{p}_3(t) = t^2$

$\mathcal{Q} = \{\underline{q}_1, \underline{q}_2, \underline{q}_3\}$ where $\underline{q}_1(t) = 1 + t^2$, $\underline{q}_2(t) = t$, $\underline{q}_3(t) = 2 + 3t^2$

Then
$$q_1 = p_1 + p_3 = 1.p_1 + 0.p_2 + 1.p_3$$
$$q_2 = p_2$$
$$q_3 = 2p_1 + 3p_3$$

So $[q_1]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$

$[q_2]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

$[q_3]_{\mathcal{B}} = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}$

$$\Rightarrow [Id]^{Q}_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{pmatrix}$$

$$\Rightarrow [Id]^{\mathcal{B}}_{Q} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & 0 & -2 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

If $p(t) = 2 - 3t + t^2$

then $[p]_{\mathcal{B}} = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$

$$\Rightarrow [p]_Q = [Id]^{\mathcal{B}}_{Q} [p]_{\mathcal{B}} = \begin{pmatrix} 3 & 0 & -2 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ -3 \\ -1 \end{pmatrix}$$

So $p = 4q_1 - 3q_2 - q_3$.

Main Theorem

If $L : V \to W$ is a linear transformation,

& $\mathcal{B}_1$, $\mathcal{B}_2$ are two bases of V,

& $\mathcal{C}_1$, $\mathcal{C}_2$ are two bases of W.

Then by the composition formula

$$[Id]^{\mathcal{C}_1}_{\mathcal{C}_2}[L]^{\mathcal{B}_1}_{\mathcal{C}_1} = \underset{\substack{\| \\ Id \circ L = L \circ Id}}{[L]^{\mathcal{B}_1}_{\mathcal{C}_2}} = [L]^{\mathcal{B}_2}_{\mathcal{C}_2}[Id]^{\mathcal{B}_1}_{\mathcal{B}_2}$$

Consequently multiplying by $[Id]^{\mathcal{B}_2}_{\mathcal{B}_1}$ on right:

$$[L]^{\mathcal{B}_2}_{\mathcal{C}_2} = [Id]^{\mathcal{C}_1}_{\mathcal{C}_2}\,[L]^{\mathcal{B}_1}_{\mathcal{C}_1}\,[Id]^{\mathcal{B}_2}_{\mathcal{B}_1}$$

Special case when $V = W$, & $\mathcal{B}_1 = \mathcal{C}_1$, & $\mathcal{B}_2 = \mathcal{C}_2$,

we have

$$[L]^{\mathcal{C}_2}_{\mathcal{C}_2} = S\,[L]^{\mathcal{C}_1}_{\mathcal{C}_1}\,S^{-1}$$

where $S = [Id]^{\mathcal{C}_1}_{\mathcal{C}_2}$ (Theorem 5.32)

Example 5.33

Let $L : \mathbb{R}^3 \to \mathbb{R}^3$ be $L(x) = Ax$

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

(a) Compute the matrix representation of L w.r.t. $\mathcal{B} = \{\underline{b}_1, \underline{b}_2, \underline{b}_3\}$ where

$$b_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

~~(b) Compute $L^n(\underline{x})$ where $\underline{x} = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}$.~~

Solution (a) Let $\mathcal{E}$ be the standard basis, so that $[L]_{\mathcal{E}}^{\mathcal{E}} = A$.

Then let $S = [\mathrm{Id}]_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 1 & -2 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

& compute $S^{-1} = \frac{1}{3}\begin{pmatrix} 0 & -3 & 3 \\ -1 & -1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$

so that $[L]_{\mathcal{B}}^{\mathcal{B}} = S^{-1}AS = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$.