MATRICES

2

A key aspect of differential calculus is the approximation of nonlinear maps from $\mathbb{R}^n \to \mathbb{R}^m$ by linear maps (plus a constant). Today we review some basics on linear maps and matrices.

§1. Matrix transformations

There are two approaches to multiplying a matrix by a column vector:

① Row-by-column (more computationally efficient)

② Linear combinations (more conceptual)

① $$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{pmatrix}$$

take dot products

notice that this $$= \begin{pmatrix} a_{11}x_1 \\ a_{21}x_1 \end{pmatrix} + \begin{pmatrix} a_{12}x_2 \\ a_{22}x_2 \end{pmatrix} + \begin{pmatrix} a_{13}x_3 \\ a_{23}x_3 \end{pmatrix}$$

$$= x_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}$$

which leads to...

② Writing $A = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$, $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$,

$$\boxed{A\vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + \dots + x_n\vec{a}_n = \sum_{j=1}^{n} x_j \vec{a}_j .}$$

① $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{pmatrix}$

take dot products

notice that this $= \begin{pmatrix} a_{11}x_1 \\ a_{21}x_1 \end{pmatrix} + \begin{pmatrix} a_{12}x_2 \\ a_{22}x_2 \end{pmatrix} + \begin{pmatrix} a_{13}x_3 \\ a_{23}x_3 \end{pmatrix}$

$$= x_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}$$

which leads to...

② Writing $A = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$, $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$,

$$\boxed{A\vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + \dots + x_n\vec{a}_n = \sum_{j=1}^{n} x_j \vec{a}_j .}$$

The general form of ①, writing

$A = \begin{pmatrix} \leftarrow & \vec{A}_1 & \rightarrow \\ & \vdots & \\ \leftarrow & \vec{A}_m & \rightarrow \end{pmatrix}$, is

$$A\vec{x} = \begin{pmatrix} \vec{A}_1 \cdot \vec{x} \\ \vdots \\ \vec{A}_m \cdot \vec{x} \end{pmatrix}.$$

As an example of the 2nd approach's utility, we can use it to check the

LINEARITY PROPERTY:

$$\boxed{A(\underbrace{c\vec{u}+d\vec{w}}_{\begin{pmatrix} cu_1+dw_1 \\ \vdots \\ cu_n+dw_n \end{pmatrix}}) = cA\underbrace{\vec{u}}_{\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}} + dA\underbrace{\vec{w}}_{\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}}}$$

By ②,

$$\begin{aligned} \text{LHS} &= (cu_1+dw_1)\vec{a}_1 + \cdots + (cu_n+dw_n)\vec{a}_n \\ &= c(u_1\vec{a}_1+\cdots+u_n\vec{a}_n) + d(w_1\vec{a}_1+\cdots+w_n\vec{a}_n) \\ &= \text{RHS}. \end{aligned}$$

① $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{pmatrix}$

take dot products

notice that this $= \begin{pmatrix} a_{11}x_1 \\ a_{21}x_1 \end{pmatrix} + \begin{pmatrix} a_{12}x_2 \\ a_{22}x_2 \end{pmatrix} + \begin{pmatrix} a_{13}x_3 \\ a_{23}x_3 \end{pmatrix}$

$= x_1\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} + x_3\begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}$

which leads to...

② Writing $A = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$, $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$,

$$\boxed{A\vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \sum_{j=1}^{n} x_j\vec{a}_j.}$$

The general form of ①, writing

$A = \begin{pmatrix} \leftarrow \vec{A}_1 \rightarrow \\ \vdots \\ \leftarrow \vec{A}_m \rightarrow \end{pmatrix}$, is

$$A\vec{x} = \begin{pmatrix} \vec{A}_1 \cdot \vec{x} \\ \vdots \\ \vec{A}_m \cdot \vec{x} \end{pmatrix}.$$

As an example of the 2^{nd} approach's utility, we can use it to check the

$\boxed{\begin{array}{l}\textbf{LINEARITY PROPERTY:} \\ A(\underbrace{c\vec{u}+d\vec{w}}_{\begin{pmatrix} cu_1+dw_1 \\ \vdots \\ cu_n+dw_n \end{pmatrix}}) = cA\underbrace{\vec{u}}_{\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}} + dA\underbrace{\vec{w}}_{\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}}\end{array}}$

By ②,

$$\begin{aligned} LHS &= (cu_1+dw_1)\vec{a}_1 + \cdots + (cu_n+dw_n)\vec{a}_n \\ &= c(u_1\vec{a}_1 + \cdots + u_n\vec{a}_n) + d(w_1\vec{a}_1 + \cdots + w_n\vec{a}_n) \\ &= RHS. \end{aligned}$$

In other words, $T_A(\vec{x}) := A\vec{x}$ yields a linear transformation

$$T_A : \mathbb{R}^n \to \mathbb{R}^m$$

for any $A \in M_{m\times n}$:= $m\times n$ matrices with real entries.

As an example of the 2^{nd} approach's utility, we can use it to check the

$$\boxed{\begin{array}{l}\text{LINEARITY PROPERTY:} \\ A(\underbrace{c\vec{u}+d\vec{w}}_{\begin{pmatrix} cu_1+dw_1 \\ \vdots \\ cu_n+dw_n \end{pmatrix}}) = cA\underbrace{\vec{u}}_{\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}} + dA\underbrace{\vec{w}}_{\begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}}\end{array}}$$

By ②,

$$\begin{aligned} \text{LHS} &= (cu_1+dw_1)\vec{a}_1 + \cdots + (cu_n+dw_n)\vec{a}_n \\ &= c(u_1\vec{a}_1 + \cdots + u_n\vec{a}_n) + d(w_1\vec{a}_1 + \cdots + w_n\vec{a}_n) \\ &= \text{RHS}. \end{aligned}$$

In other words, $T_A(\vec{x}) := A\vec{x}$ yields a linear transformation

$$T_A : \mathbb{R}^n \to \mathbb{R}^m$$

for any $A \in \mathcal{M}_{m\times n} :=$ $m \times n$ matrices with real entries.

I claim that, in fact, <u>every</u> linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ arises in this way (is a matrix transformation). If we write

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \sum_{j=1}^n x_j \hat{e}_j$$

then (by linearity of T)

$$T(\vec{x}) = T\left(\sum_{j=1}^n x_j \hat{e}_j\right) = \sum_{j=1}^n x_j T(\hat{e}_j)$$

$$= \underbrace{\begin{pmatrix} \uparrow & & \uparrow \\ T(\hat{e}_1) & \cdots & T(\hat{e}_n) \\ \downarrow & & \downarrow \end{pmatrix}}_{=:A \text{ (standard matrix of } T)} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

I claim that, in fact, <u>every</u> linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ arises in this way (is a matrix transformation). If we write

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \sum_{j=1}^{n} x_j \hat{e}_j$$

then (by linearity of T)

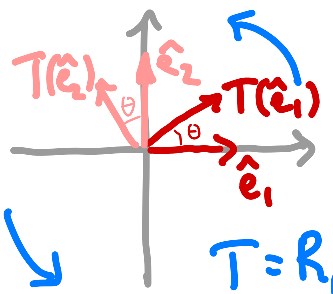
$$T(\vec{x}) = T\left(\sum_{j=1}^{n} x_j \hat{e}_j\right) = \sum_{j=1}^{n} x_j T(\hat{e}_j)$$

$$= \underbrace{\begin{pmatrix} \uparrow & & \uparrow \\ T(\hat{e}_1) & \cdots & T(\hat{e}_n) \\ \downarrow & & \downarrow \end{pmatrix}}_{=: A \text{ (standard matrix of } T)} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= A \vec{x},$$

with A an $m \times n$ matrix.

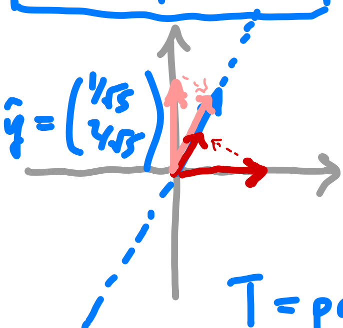
Example 1



$T = R_\theta$

$$A = \begin{pmatrix} \uparrow & \uparrow \\ T(\hat{e}_1) & T(\hat{e}_2) \\ \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Example 2



$\hat{y} = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$

$$A = \begin{pmatrix} \uparrow & \uparrow \\ T(\hat{e}_1) & T(\hat{e}_2) \\ \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} \uparrow & \uparrow \\ \frac{1}{\sqrt{5}}\hat{y} & \frac{2}{\sqrt{5}}\hat{y} \\ \downarrow & \downarrow \end{pmatrix}$$

$$= \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} = \hat{y}\,\hat{y}^T = \begin{pmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{pmatrix}$$

$T = \mathrm{proj}_{\hat{y}}$ $\left(\mathrm{proj}_{\hat{y}} \vec{x} = (\hat{y}\cdot\vec{x})\,\hat{y} \text{ from Lecture 1}\right)$

Example 3

Given $\vec{a} \in \mathbb{R}^n$, dot product with $\vec{a}$ defines a L.T. $T : \mathbb{R}^n \to \mathbb{R}$.

I claim that, in fact, every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ arises in this way (is a matrix transformation). If we write

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \sum_{j=1}^{n} x_j \hat{e}_j$$

then (by linearity of T)

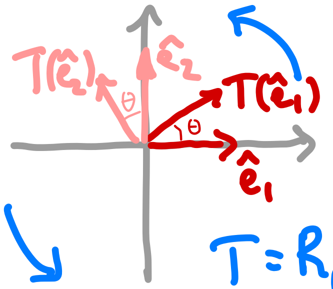
$$T(\vec{x}) = T\left(\sum_{j=1}^{n} x_j \hat{e}_j\right) = \sum_{j=1}^{n} x_j T(\hat{e}_j)$$

$$= \underbrace{\begin{pmatrix} \uparrow & & \uparrow \\ T(\hat{e}_1) & \cdots & T(\hat{e}_n) \\ \downarrow & & \downarrow \end{pmatrix}}_{=:\, A \text{ (standard matrix of } T)} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= A\vec{x},$$

with A an $m \times n$ matrix.

Example 1

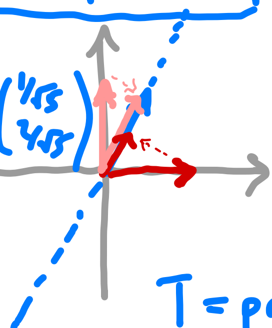


$T = R_\theta$

$$A = \begin{pmatrix} \uparrow & \uparrow \\ T(\hat{e}_1) & T(\hat{e}_2) \\ \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Example 2

$\hat{y} = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$



$$A = \begin{pmatrix} \uparrow & \uparrow \\ T(\hat{e}_1) & T(\hat{e}_2) \\ \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} \uparrow & \uparrow \\ \frac{1}{\sqrt{5}}\hat{y} & \frac{2}{\sqrt{5}}\hat{y} \\ \downarrow & \downarrow \end{pmatrix}$$

$$= \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} = \hat{y}\,{}^t\hat{y} = \begin{pmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{pmatrix}$$

$T = \mathrm{proj}_{\hat{y}}$ $\left(\mathrm{proj}_{\hat{y}} \vec{x} = (\hat{y} \cdot \vec{x})\,\hat{y} \right.$ from Lecture 1 $\left.\right)$

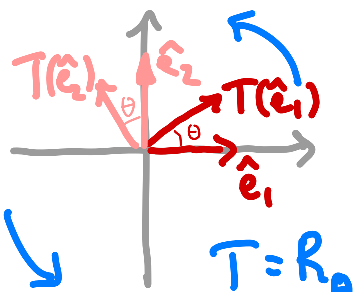
Example 3

Given $\vec{a} \in \mathbb{R}^n$, dot product with $\vec{a}$ defines a L.T. $T : \mathbb{R}^n \to \mathbb{R}$.

$$A = \underbrace{\begin{pmatrix} T(\hat{e}_1) \cdots T(\hat{e}_n) \end{pmatrix}}_{1 \times n} = \begin{pmatrix} \vec{a} \cdot \hat{e}_1 & \cdots & \vec{a} \cdot \hat{e}_n \end{pmatrix}$$

$$= \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} = {}^t\vec{a}.$$

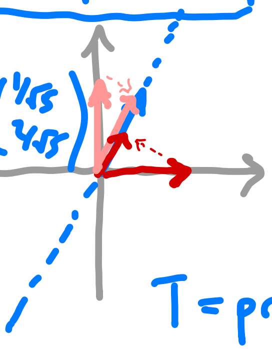
Example 1



$T = R_\theta$

$$A = \begin{pmatrix} \uparrow & \uparrow \\ T(\hat{e}_1) & T(\hat{e}_2) \\ \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Example 2



$\hat{y} = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$

$$A = \begin{pmatrix} \uparrow & \uparrow \\ T(\hat{e}_1) & T(\hat{e}_2) \\ \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} \uparrow & \uparrow \\ \frac{1}{\sqrt{5}}\hat{y} & \frac{2}{\sqrt{5}}\hat{y} \\ \downarrow & \downarrow \end{pmatrix}$$

$$= \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} = \hat{y}\, {}^t\hat{y} = \begin{pmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{pmatrix}$$

$T = \mathrm{proj}_{\hat{y}}$ $\left(\mathrm{proj}_{\hat{y}} \vec{x} = (\hat{y}\cdot\vec{x})\,\hat{y}\right.$ from Lecture 1$\left.\right)$

Example 3

Given $\vec{a} \in \mathbb{R}^n$, dot product with $\vec{a}$ defines a L.T. $T : \mathbb{R}^n \to \mathbb{R}$.

$$A = \underbrace{\begin{pmatrix} T(\hat{e}_1) \cdots T(\hat{e}_n) \end{pmatrix}}_{1\times n} = \begin{pmatrix} \vec{a}\cdot\hat{e}_1 & \cdots & \vec{a}\cdot\hat{e}_n \end{pmatrix}$$

$$= \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} = {}^t\vec{a}.$$

§2. Algebra with matrices

Adding matrices:

- must be of same "dimensions" (both $m\times n$)
- add entry by entry, e.g.

$$\begin{pmatrix} 0 & 1 & -1 \\ 3 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & 5 \end{pmatrix}$$

- commutative: $A + B = B + A$
- associative: $(A+B)+C = A + (B+C)$
- additive identity: $A + \underbrace{O}_{\text{zero matrix}} = A$
- additive inverse: $A + \underbrace{(-A)}_{\text{negate all entries}} = O$
- cancellation: $A+B = C+B \Rightarrow A = C$

§2. Algebra with matrices

Adding matrices:

- must be of same "dimensions" (both $m \times n$)
- add entry by entry, e.g.

$$\begin{pmatrix} 0 & 1 & -1 \\ 3 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & 5 \end{pmatrix}$$

- commutative: $A+B = B+A$
- associative: $(A+B)+C = A+(B+C)$
- additive identity: $A + \underset{\uparrow \text{ zero matrix}}{O} = A$
- additive inverse: $A + (-A) = O$ ← negate all entries
- cancellation: $A+B = C+B \Rightarrow A = C$

Scalar multiplication:

- multiply each entry by the scalar, e.g.

$$3\begin{pmatrix} 0 & 1 & -1 \\ 3 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 3 & -3 \\ 9 & 0 & 6 \end{pmatrix}$$

- distributivity: $c(A+B) = cA + cB$
  $(c+d)A = cA + dA$
- $(cd)A = c(dA)$

(Upshot: $M_{m \times n}$ is a vector space)

Multiplying matrices: A $m\times n$, B $l\times p$

AB is defined when $n=l$
BA is defined when $m=p$

Assume $n=l$, $B=\begin{pmatrix} \uparrow & & \uparrow \\ \vec{b}_1 & \cdots & \vec{b}_p \\ \downarrow & & \downarrow \end{pmatrix}$:

then $AB := \begin{pmatrix} \uparrow & & \uparrow \\ A\vec{b}_1 & \cdots & A\vec{b}_p \\ \downarrow & & \downarrow \end{pmatrix}$.

So the k^{th} column of AB is

$$A\vec{b}_k = \begin{pmatrix} \uparrow & & \uparrow \\ \vec{a}_1 & \cdots & \vec{a}_n \\ \downarrow & & \downarrow \end{pmatrix}\begin{pmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{pmatrix} = \sum_{j=1}^{n} b_{jk}\vec{a}_j$$

and the $(i,k)^{th}$ entry of AB is

$$(AB)_{ik} = \Big(\sum_j b_{jk}\vec{a}_j\Big)_i = \sum_j b_{jk}(\vec{a}_j)_i$$

i^{th} entry

$$= \sum_{j=1}^{n} a_{ij} b_{jk}.$$

§2. Algebra with matrices

Adding matrices:

- must be of same "dimensions" (both $m\times n$)
- add entry by entry, e.g.
$$\begin{pmatrix} 0 & 1 & -1 \\ 3 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & 5 \end{pmatrix}$$
- commutative: $A+B=B+A$
- associative: $(A+B)+C = A+(B+C)$
- additive identity: $A+O=A$ ← zero matrix
- additive inverse: $A+(-A)=O$ ← negate all entries
- cancellation: $A+B=C+B \Rightarrow A=C$

Scalar multiplication:

- multiply each entry by the scalar, e.g.
$$3\begin{pmatrix} 0 & 1 & -1 \\ 3 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 3 & -3 \\ 9 & 0 & 6 \end{pmatrix}$$
- distributivity: $c(A+B) = cA+cB$
  $(c+d)A = cA+dA$
- $(cd)A = c(dA)$

(Upshot: $M_{m\times n}$ is a vector space)

Multiplying matrices: A $m \times n$, B $l \times p$

AB is defined when $n = l$

BA is defined when $m = p$

Assume $n = l$, $B = \begin{pmatrix} \uparrow & & \uparrow \\ \vec{b}_1 & \cdots & \vec{b}_p \\ \downarrow & & \downarrow \end{pmatrix}$:

then $AB := \begin{pmatrix} \uparrow & & \uparrow \\ A\vec{b}_1 & \cdots & A\vec{b}_p \\ \downarrow & & \downarrow \end{pmatrix}$.

So the k^{th} column of AB is

$$A\vec{b}_k = \begin{pmatrix} \uparrow & & \uparrow \\ \vec{a}_1 & \cdots & \vec{a}_n \\ \downarrow & & \downarrow \end{pmatrix} \begin{pmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{pmatrix} = \sum_{j=1}^{n} b_{jk} \vec{a}_j$$

and the $(i,k)^{\text{th}}$ entry of AB is

$$(AB)_{ik} = \Big(\sum_j b_{jk} \vec{a}_j \Big)_i = \sum_j b_{jk} (\vec{a}_j)_i$$

i^{th} entry

$$= \sum_{j=1}^{n} a_{ij} b_{jk}.$$

recognize as the familiar "i^{th} row of A dotted with the k^{th} column of B".

Multiplying matrices: A $m\times n$, B $l\times p$

AB is defined when $n=l$ \
BA is defined when $m=p$

Assume $n=l$, $B=\begin{pmatrix} \uparrow & & \uparrow \\ \vec{b}_1 & \cdots & \vec{b}_p \\ \downarrow & & \downarrow \end{pmatrix}$:

then $AB := \begin{pmatrix} \uparrow & & \uparrow \\ A\vec{b}_1 & \cdots & A\vec{b}_p \\ \downarrow & & \downarrow \end{pmatrix}$.

So the k^{th} column of AB is

$$A\vec{b}_k = \begin{pmatrix} \uparrow & & \uparrow \\ \vec{a}_1 & \cdots & \vec{a}_n \\ \downarrow & & \downarrow \end{pmatrix}\begin{pmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{pmatrix} = \sum_{j=1}^{n} b_{jk}\vec{a}_j$$

and the $(i,k)^{\text{th}}$ entry of AB is

$$(AB)_{ik} = \Big(\sum_j b_{jk}\vec{a}_j\Big)_i = \sum_j b_{jk}(\vec{a}_j)_i$$

i^{th} entry

$$= \sum_{j=1}^{n} a_{ij}b_{jk}.$$

recognize as the familiar "i^{th} row of A dotted with the k^{th} column of B".

- powers of a matrix: A^m means $A\cdot A\cdot \ldots \cdot A$ (m times)

- identity matrix:

$$\underset{n\times n}{I_n} = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots & \\ & & & 1 \end{pmatrix} = \begin{pmatrix} \uparrow & & \uparrow \\ \hat{e}_1 & \cdots & \hat{e}_n \\ \downarrow & & \downarrow \end{pmatrix}$$

$$AI_n = \begin{pmatrix} \uparrow & & \uparrow \\ A\hat{e}_1 & \cdots & A\hat{e}_n \\ \downarrow & & \downarrow \end{pmatrix} = \begin{pmatrix} \uparrow & & \uparrow \\ \vec{a}_1 & \cdots & \vec{a}_n \\ \downarrow & & \downarrow \end{pmatrix} = A = I_n A$$

- distributivity: $A(\beta B+\gamma C) = \beta(AB)+\gamma(AC)$ \
$(\beta B+\gamma C)A = \beta(BA)+\gamma(CA)$

- NOT commutative: $AB \neq BA$ in general

e.g. $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

- powers of a matrix: A^m means
$$A \cdot A \cdot \ldots \cdot A \quad (m \text{ times})$$

- identity matrix:

$$\underset{n\times n}{I_n} = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots \\ & & & 1 \end{pmatrix} = \begin{pmatrix} \uparrow & & \uparrow \\ \hat{e}_1 & \cdots & \hat{e}_n \\ \downarrow & & \downarrow \end{pmatrix}$$

$$AI_n = \begin{pmatrix} \uparrow & & \uparrow \\ A\hat{e}_1 & \cdots & A\hat{e}_n \\ \downarrow & & \downarrow \end{pmatrix} = \begin{pmatrix} \uparrow & & \uparrow \\ \vec{a}_1 & \cdots & \vec{a}_n \\ \downarrow & & \downarrow \end{pmatrix} = A = I_n A$$

- distributivity: $A(\beta B + \gamma C) = \beta(AB) + \gamma(AC)$
$$(\beta B + \gamma C)A = \beta(BA) + \gamma(CA)$$

- NOT commutative: $AB \neq BA$ in general

e.g. $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

- associativity holds: $(AB)C = A(BC)$.
Can check by hand, but it follows automatically from thinking of matrices as linear maps and products as composition (see below).

- multiplicative inverse? In general, NO; and only square matrices have the possibility of a "both-sided inverse":

Definition: Let A be an $n\times n$ matrix. A is *invertible* if there exists an $n\times n$ matrix M (called its *inverse*) satisfying
$$AM = I_n = MA.$$

Example:
$$\underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad-bc & \cancel{ab-ab}^{\,0} \\ \cancel{cd-cd}^{\,0} & ad-bc \end{pmatrix}$$
$$= (ad-bc)\, I_2.$$

So if $\det(A) := ad-bc \neq 0$, then $M := \frac{1}{\det(A)}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ satisfies $AM = I_2$ (and you can check $MA = I_2$).

- powers of a matrix: A^m means
$$A \cdot A \cdot \ldots \cdot A \quad (m \text{ times})$$

- identity matrix:
$$\underset{n\times n}{I_n} = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots \; 1 \end{pmatrix} = \begin{pmatrix} \uparrow & & \uparrow \\ \hat{e}_1 & \cdots & \hat{e}_n \\ \downarrow & & \downarrow \end{pmatrix}$$
$$AI_n = \begin{pmatrix} \uparrow & & \uparrow \\ A\hat{e}_1 & \cdots & A\hat{e}_n \\ \downarrow & & \downarrow \end{pmatrix} = \begin{pmatrix} \uparrow & & \uparrow \\ \vec{a}_1 & \cdots & \vec{a}_n \\ \downarrow & & \downarrow \end{pmatrix} = A = I_n A$$

- distributivity: $A(\beta B + \gamma C) = \beta(AB) + \gamma(AC)$
$$(\beta B + \gamma C)A = \beta(BA) + \gamma(CA)$$

- NOT commutative: $AB \neq BA$ in general
e.g. $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

- associativity holds: $(AB)C = A(BC)$.
Can check by hand, but it follows automatically from thinking of matrices as linear maps and products as composition (see below).

- multiplicative inverse? In general, NO; and only square matrices have the possibility of a "both-sided inverse":

Definition: Let A be an $n \times n$ matrix. A is invertible if there exists an $n \times n$ matrix M (called its inverse) satisfying

$$AM = I_n = MA.$$

Example:

$$\underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad-bc & \cancel{ab-ab}^{\,0} \\ \cancel{cd-cd}^{\,0} & ad-bc \end{pmatrix}$$

$$= (ad-bc)\, I_2 .$$

So if $\det(A) := ad-bc \neq 0$, then

$M := \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ satisfies $AM = I_2$

(and you can check $MA = I_2$).

We usually write A^{-1} for the inverse.

Since $AB\,B^{-1}A^{-1} = AA^{-1} = I_n$, we get

$$(AB)^{-1} = B^{-1}A^{-1}.$$

• multiplicative inverse? In general, NO; and only square matrices have the possibility of a "both-sided inverse":

Definition: Let A be an $n \times n$ matrix. A is *invertible* if there exists an $n \times n$ matrix M (called its *inverse*) satisfying
$$AM = I_n = MA.$$

Example:
$$\underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad-bc & \cancel{ab-ab}^{\,0} \\ \cancel{cd-cd}^{\,0} & ad-bc \end{pmatrix}$$
$$= (ad-bc)\, I_2 .$$

So if $\det(A) := ad-bc \neq 0$, then
$M := \frac{1}{\det(A)}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ satisfies $AM = I_2$
(and you can check $MA = I_2$).

We usually write A^{-1} for the inverse.
Since $AB\,B^{-1}A^{-1} = AA^{-1} = I_n$, we get
$$(AB)^{-1} = B^{-1}A^{-1}.$$

One more "matrix algebra" operation...

Transpose of a matrix:

• A $m \times n$ $\rightsquigarrow$ ${}^tA$ $n \times m$, given by
$$({}^tA)_{ij} = A_{ji}$$

e.g. ${}^t\begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 3 \\ 1 & 1 & 2 \end{pmatrix}$

• ${}^t(AB) = {}^tB\,{}^tA$

Proof: $({}^t(AB))_{ij} = (AB)_{ji} = \sum_k a_{jk} b_{ki}$
$= \sum_k b_{ki} a_{jk} = \sum ({}^tB)_{ik}({}^tA)_{kj} = ({}^tB\,{}^tA)_{ij}$. $\square$

• $\vec{x}\cdot\vec{y} = (x_1 \cdots x_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = ({}^t\vec{x})\,\vec{y}$.

So $(A\vec{x})\cdot\vec{y} = {}^t(A\vec{x})\,\vec{y} = {}^t\vec{x}\,{}^tA\,\vec{y}$
$= \vec{x}\cdot({}^tA\,\vec{y})$

$\boxed{\text{Exercise}}$ A, B $n \times n$; B invertible. Calculate $(BAB^{-1})^k$.

$\boxed{\text{Exercise}}$ A $m \times n$, $\vec{x} \in \mathbb{R}^n$, with ${}^tAA\vec{x} = \vec{0}$. Show that $A\vec{x} = \vec{0}$.

[Hint: what is $\|A\vec{x}\|$?]

$\boxed{\text{Exercise}}$ An $n \times n$ matrix A is orthogonal if ${}^tAA = I_n$. Check that the column vectors $\vec{a}_j$ are orthogonal unit vectors.

One more "matrix algebra" operation..

Transpose of a matrix:

- A $m \times n$ $\rightsquigarrow$ ${}^tA$ $n \times m$, given by
$$({}^tA)_{ij} = A_{ji}$$

e.g. ${}^t\begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 3 \\ 1 & 1 & 2 \end{pmatrix}$

- ${}^t(AB) = {}^tB\,{}^tA$

Proof: $({}^t(AB))_{ij} = (AB)_{ji} = \sum_k a_{jk} b_{ki}$

$= \sum_k b_{ki} a_{jk} = \sum ({}^tB)_{ik} ({}^tA)_{kj} = ({}^tB\,{}^tA)_{ij}$. $\square$

- $\vec{x} \cdot \vec{y} = (x_1 \cdots x_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = ({}^t\vec{x})\vec{y}$.

So $(A\vec{x}) \cdot \vec{y} = {}^t(A\vec{x})\,\vec{y} = {}^t\vec{x}\,{}^tA\,\vec{y}$
$= \vec{x} \cdot ({}^tA\,\vec{y})$

- If ${}^tA = A$ (only poss. if A $n \times n$), then A is symmetric.

$\boxed{\text{Exercise}}$ A, B $n\times n$; B invertible. Calculate $(BAB^{-1})^k$.

$$B A \cancel{B^{-1}} \cancel{B} A \cancel{B^{-1}} \cancel{B} \cdots \cancel{B^{-1}} \cancel{B} A B^{-1} = B A^k B^{-1}$$

$\boxed{\text{Exercise}}$ A $m\times n$, $\vec{x} \in \mathbb{R}^n$, with ${}^tAA\vec{x} = \vec{0}$. Show that $A\vec{x} = \vec{0}$.
[Hint: what is $\|A\vec{x}\|$?]

$$\|A\vec{x}\|^2 = A\vec{x}\cdot A\vec{x} = {}^t(A\vec{x})A\vec{x} = {}^t\vec{x}\,\underbrace{{}^tAA\vec{x}}_{\vec{0}}$$
$$= 0 \Rightarrow A\vec{x} = \vec{0}.$$

$\boxed{\text{Exercise}}$ An $n\times n$ matrix A is orthogonal if ${}^tAA = I_n$. Check that the column vectors $\vec{a}_j$ are orthogonal unit vectors.

$$\begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & \ddots & 1 \end{pmatrix} = \begin{pmatrix} \leftarrow {}^t\vec{a}_1 \rightarrow \\ \vdots \\ \leftarrow {}^t\vec{a}_n \rightarrow \end{pmatrix} \begin{pmatrix} \uparrow & & \uparrow \\ \vec{a}_1 & \cdots & \vec{a}_n \\ \downarrow & & \downarrow \end{pmatrix} = (\vec{a}_i\cdot\vec{a}_j)$$

$\Rightarrow$ $\vec{a}_i\cdot\vec{a}_i = 1$ and $\vec{a}_i\cdot\vec{a}_j = 0$ $(i\neq j)$.

§3. Composing linear transformations

$\boxed{\text{Exercise}}$ A, B $n \times n$; B invertible. Calculate $(BAB^{-1})^k$.

$$BA\cancel{B^{-1}}\cancel{B}AB^{-1}\cancel{B}\cdots\cancel{B^{-1}}BAB^{-1} = BA^kB^{-1}$$

$\boxed{\text{Exercise}}$ A $m \times n$, $\vec{x} \in \mathbb{R}^n$, with ${}^tAA\vec{x} = \vec{0}$. Show that $A\vec{x} = \vec{0}$.

[Hint: what is $\|A\vec{x}\|$?]

$$\|A\vec{x}\|^2 = A\vec{x} \cdot A\vec{x} = {}^t(A\vec{x})A\vec{x} = {}^t\vec{x}\underbrace{{}^tAA\vec{x}}_{\vec{0}}$$

$$= 0 \Rightarrow A\vec{x} = \vec{0}.$$

$\boxed{\text{Exercise}}$ An $n \times n$ matrix A is orthogonal if ${}^tAA = I_n$. Check that the column vectors $\vec{a}_j$ are orthogonal unit vectors.

$$\begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots\ 1 \end{pmatrix} = \begin{pmatrix} \leftarrow {}^t\vec{a}_1 \rightarrow \\ \vdots \\ \leftarrow {}^t\vec{a}_n \rightarrow \end{pmatrix} \begin{pmatrix} \uparrow & & \uparrow \\ \vec{a}_1 & \cdots & \vec{a}_n \\ \downarrow & & \downarrow \end{pmatrix} = (\vec{a}_i \cdot \vec{a}_j)$$

$\Rightarrow \vec{a}_i \cdot \vec{a}_i = 1$ and $\underset{(i \neq j)}{\vec{a}_i \cdot \vec{a}_j} = 0$.

§3. Composing linear transformations

Suppose that

$S : \mathbb{R}^p \to \mathbb{R}^n$ has matrix B ($n \times p$)

and

$T : \mathbb{R}^n \to \mathbb{R}^m$ has matrix A ($m \times n$).

The composition $T \circ S$

$$\begin{tikzcd} \mathbb{R}^p \arrow[r, "S"] \arrow[rr, bend left, "T \circ S"] & \mathbb{R}^n \arrow[r, "T"] & \mathbb{R}^m \end{tikzcd}$$

is defined, & so is the product AB.

Theorem: AB is the matrix of $T \circ S$.

Proof: $(T \circ S)(\hat{e}_k) = T(S(\hat{e}_k))$

$= T(\vec{b}_k) = A\vec{b}_k = k^{\text{th}}$ column of AB $\square$

(exactly how we defined AB)

§3. Composing linear transformations

Suppose that

$S : \mathbb{R}^p \to \mathbb{R}^n$ has $\overset{n \times p}{\text{matrix}}$ B

and

$T : \mathbb{R}^n \to \mathbb{R}^m$ has $\overset{m \times n}{\text{matrix}}$ A.

The composition $T \circ S$

$$\begin{tikzcd} \mathbb{R}^p \arrow[r, "S"] \arrow[rr, bend left, "T \circ S"] & \mathbb{R}^n \arrow[r, "T"] & \mathbb{R}^m \end{tikzcd}$$

is defined, & so is the product AB.

Theorem: AB is the matrix of $T \circ S$.

Proof: $(T \circ S)(\hat{e}_k) = T(S(\hat{e}_k))$

$= T(\vec{b}_k) = A\vec{b}_k = k^{\underline{th}}$ column of AB $\square$

(exactly how we defined AB)

$\boxed{\text{Example}}$ Any kind of projection T satisfies $T^2 = T$ and is not invertible (why?). So the same is true for its matrix A.

As a plausibility check, let's see if this works for our projection matrix $A = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ from before. Certainly $\det(A) = 0$, so it's non-invertible; and

$$A^2 = \frac{1}{25}\begin{pmatrix} 1^2+2^2 & 1\cdot 2+2\cdot 4 \\ 2\cdot 1+4\cdot 2 & 2^2+4^2 \end{pmatrix} = \frac{1}{25}\begin{pmatrix} 5 & 10 \\ 10 & 20 \end{pmatrix} = A.$$

Exercise Let $T = R_\theta$ and $S = R_\phi$ be rotation through angles θ & ϕ respectively. Use the Theorem to rederive the formulas for $\cos(\theta+\phi)$ and $\sin(\theta+\phi)$.

Exercise Show that any 2×2 orthogonal matrix is the matrix of a rotation or a rotation composed with a reflection.

§3. Composing linear transformations

Suppose that

$S : \mathbb{R}^p \to \mathbb{R}^n$ has $n\times p$ matrix B

and

$T : \mathbb{R}^n \to \mathbb{R}^m$ has $m \times n$ matrix A.

The composition $T\circ S$

$$\begin{tikzcd} \mathbb{R}^p \arrow[r, "S"] \arrow[rr, bend left, "T\circ S"] & \mathbb{R}^n \arrow[r, "T"] & \mathbb{R}^m \end{tikzcd}$$

is defined, & so is the product AB.

Theorem: AB is the matrix of $T\circ S$.

Proof: $(T\circ S)(\hat{e}_k) = T(S(\hat{e}_k))$

$= T(\vec{b}_k) = A\vec{b}_k = k^{\text{th}}$ column of AB $\square$

(exactly how we defined AB)

Example Any kind of projection T satisfies $T\circ T = T$ and is not invertible (why?). So the same is true for its matrix A: $A^2 = A$, no A^{-1}.

As a plausibility check, let's see if this works for our projection matrix $A = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ from before. Certainly $\det(A) = 0$, so it's non-invertible; and

$$A^2 = \frac{1}{25}\begin{pmatrix} 1^2+2^2 & 1\cdot 2+2\cdot 4 \\ 2\cdot 1+4\cdot 2 & 2^2+4^2 \end{pmatrix} = \frac{1}{25}\begin{pmatrix} 5 & 10 \\ 10 & 20 \end{pmatrix} = A.$$

$\boxed{\text{Exercise}}$ Let $T = R_\theta$ and $S = R_\phi$ be rotation through angles θ & ϕ respectively. Use the Theorem to rederive the formulas for $\cos(\theta+\phi)$ and $\sin(\theta+\phi)$.

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} = \begin{pmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix}$$

because $R_\theta \circ R_\phi = R_{\theta+\phi}$.

$\boxed{\text{Exercise}}$ Show that any 2×2 orthogonal matrix is the matrix of a rotation or a rotation composed with a reflection.

$\begin{pmatrix} \uparrow & \uparrow \\ \vec{a}_1 & \vec{a}_2 \\ \downarrow & \downarrow \end{pmatrix}$ satisfy $\|\vec{a}_i\|^2 = a_i \cdot a_i = 1 \Rightarrow \vec{a}_i$ on unit circle

and $\vec{a}_1 \cdot \vec{a}_2 = 0 \Rightarrow$ perpendicular.

so get either $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ or $\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

§4. Determinants (3×3 case)

As a plausibility check, let's see if this works for our projection matrix $A = \frac{1}{5}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ from before. Certainly $\det(A) = 0$, so it's non-invertible; and

$$A^2 = \frac{1}{25}\begin{pmatrix} 1^2+2^2 & 1\cdot 2+2\cdot 4 \\ 2\cdot 1+4\cdot 2 & 2^2+4^2 \end{pmatrix} = \frac{1}{25}\begin{pmatrix} 5 & 10 \\ 10 & 20 \end{pmatrix} = A.$$

$\boxed{\text{Exercise}}$ Let $T = R_\theta$ and $S = R_\phi$ be rotation through angles θ & ϕ respectively. Use the Theorem to rederive the formulas for $\cos(\theta+\phi)$ and $\sin(\theta+\phi)$.

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} = \begin{pmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix}$$

because $R_\theta \circ R_\phi = R_{\theta+\phi}$.

$\boxed{\text{Exercise}}$ Show that any 2×2 orthogonal matrix is the matrix of a rotation or a rotation composed with a reflection.

$\begin{pmatrix} \uparrow & \uparrow \\ \vec{a}_1 & \vec{a}_2 \\ \downarrow & \downarrow \end{pmatrix}$ satisfy $\|\vec{a}_i\|^2 = a_i \cdot a_i = 1 \Rightarrow \vec{a}_i$ on unit circle

and $\vec{a}_1 \cdot \vec{a}_2 = 0 \Rightarrow$ perpendicular.

so get either $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ or $\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

§4. Determinants (3×3 case)

For 2×2 determinants, write

$$\mathcal{D}(\vec{x},\vec{y}) := \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1y_2 - x_2y_1 .$$

Clearly $\underline{\mathcal{D}(\vec{y},\vec{x}) = -\mathcal{D}(\vec{x},\vec{y})}$ [alternating property]

and also $\underline{\mathcal{D}(a\vec{y}+b\vec{z},\vec{x}) = a\mathcal{D}(\vec{y},\vec{x}) + b\mathcal{D}(\vec{z},\vec{x})}$

$\Leftrightarrow$ $\underline{\mathcal{D}(\vec{x},a\vec{y}+b\vec{z}) = a\mathcal{D}(\vec{x},\vec{y}) + b\mathcal{D}(\vec{x},\vec{z})}$ [multi-linearity].

So if $\vec{x}$ & $\vec{y}$ are parallel, we get 0 (why?)

(and $\mathcal{D}(\hat{e}_1,\hat{e}_2) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$)

§ 4. Determinants (3×3 case)

For 2×2 determinants, write

$$\mathcal{D}(\vec{x},\vec{y}) := \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1 .$$

Clearly $\underline{\mathcal{D}(\vec{y},\vec{x}) = -\mathcal{D}(\vec{x},\vec{y})}$ [alternating property]

and also $\underline{\mathcal{D}(a\vec{y}+b\vec{z},\vec{x}) = a\,\mathcal{D}(\vec{y},\vec{x}) + b\,\mathcal{D}(\vec{z},\vec{x})}$

$\Rightarrow \underline{\mathcal{D}(\vec{x},a\vec{y}+b\vec{z}) = a\,\mathcal{D}(\vec{x},\vec{y}) + b\,\mathcal{D}(\vec{x},\vec{z})}$ [multi-linearity].

So if $\vec{x}$ & $\vec{y}$ are parallel, we get 0 (why?)

(and $\mathcal{D}(\hat{e}_1,\hat{e}_2) = 1$).

Let's define 3×3 determinants

by insisting on

- [alternating property] $\mathcal{D}(\vec{x},\vec{y},\vec{z}) = -\mathcal{D}(\vec{y},\vec{x},\vec{z})$
$= -\mathcal{D}(\vec{x},\vec{z},\vec{y})$
$= -\mathcal{D}(\vec{z},\vec{y},\vec{x})$
- [multilinearity] $\mathcal{D}(\vec{x},\vec{y},a\vec{z}+b\vec{w}) =$
$a\,\mathcal{D}(\vec{x},\vec{y},\vec{z}) + b\,\mathcal{D}(\vec{x},\vec{y},\vec{w})$ etc.

and

- $\mathcal{D}(\hat{e}_1,\hat{e}_2,\hat{e}_3) = 1$

The alternating property implies that if any vector is repeated, we get 0:

e.g., $\mathcal{D}(\vec{x},\vec{x},\vec{y}) \overset{\text{swap}}{=} -\mathcal{D}(\vec{x},\vec{x},\vec{y})$

$\Rightarrow 2\mathcal{D}(\vec{x},\vec{x},\vec{y}) = 0 \Rightarrow \mathcal{D}(\vec{x},\vec{x},\vec{y}) = 0.$

So $\mathcal{D}(\vec{x},\vec{y},\vec{z}) = \mathcal{D}\left(\sum_i x_i\hat{e}_i, \sum_j y_j\hat{e}_j, \sum_k z_k\hat{e}_k\right)$

$\overset{\text{multi-linearity}}{=} \sum_{i,j,k} x_i y_j z_k \underbrace{\mathcal{D}(\hat{e}_i,\hat{e}_j,\hat{e}_k)}$ is 0 if any 2 equal

$= \sum_{i\neq j\neq k} x_i y_j z_k \cdot \underbrace{(\pm 1)}$ depending on i,j,k

$= x_1y_2z_3 + x_2y_3z_1 + x_3y_1z_2 - x_1y_3z_2 - x_2y_1z_3 - x_3y_2z_1.$

§ 4. Determinants (3×3 case)

For 2×2 determinants, write

$$\mathcal{D}(\vec{x},\vec{y}) := \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1y_2 - x_2y_1.$$

Clearly $\mathcal{D}(\vec{y},\vec{x}) = -\mathcal{D}(\vec{x},\vec{y})$ [alternating property]

and also $\mathcal{D}(a\vec{y}+b\vec{z},\vec{x}) = a\mathcal{D}(\vec{y},\vec{x}) + b\mathcal{D}(\vec{z},\vec{x})$

$\Leftrightarrow \mathcal{D}(\vec{x},a\vec{y}+b\vec{z}) = a\mathcal{D}(\vec{x},\vec{y}) + b\mathcal{D}(\vec{x},\vec{z})$ [multi-linearity].

So if $\vec{x}$ & $\vec{y}$ are parallel, we get 0 (why?)

Let's define 3×3 determinants by insisting on

- [alternating property] $\mathcal{D}(\vec{x},\vec{y},\vec{z}) = -\mathcal{D}(\vec{y},\vec{x},\vec{z})$
  $= -\mathcal{D}(\vec{x},\vec{z},\vec{y})$
  $= -\mathcal{D}(\vec{z},\vec{y},\vec{x})$
- [multilinearity] $\mathcal{D}(\vec{x},\vec{y},a\vec{z}+b\vec{w}) = a\mathcal{D}(\vec{x},\vec{y},\vec{z}) + b\mathcal{D}(\vec{x},\vec{y},\vec{w})$ etc.

and

- $\mathcal{D}(\hat{e}_1,\hat{e}_2,\hat{e}_3) = 1.$

The alternating property implies that if any vector is repeated, we get 0:

e.g., $$\mathcal{D}(\overset{\text{swap}}{\vec{x},\vec{x}},\vec{y}) = -\mathcal{D}(\vec{x},\vec{x},\vec{y})$$

$$\Rightarrow 2\mathcal{D}(\vec{x},\vec{x},\vec{y}) = 0 \Rightarrow \mathcal{D}(\vec{x},\vec{x},\vec{y}) = 0.$$

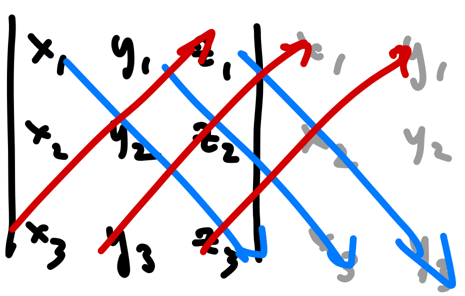
So $$\mathcal{D}(\vec{x},\vec{y},\vec{z}) = \mathcal{D}\Big(\sum_i x_i \hat{e}_i, \sum_j y_j \hat{e}_j, \sum_k z_k \hat{e}_k\Big)$$

multi-linearity

$$= \sum_{i,j,k} x_i y_j z_k \underbrace{\mathcal{D}(\hat{e}_i,\hat{e}_j,\hat{e}_k)}_{\text{is 0 if any 2 equal}}$$

$$= \sum_{i\neq j\neq k} x_i y_j z_k \cdot \underbrace{(\pm 1)}_{\text{depending on } i,j,k}$$

$$= x_1 y_2 z_3 + x_2 y_3 z_1 + x_3 y_1 z_2 - x_1 y_3 z_2 - x_2 y_1 z_3 - x_3 y_2 z_1 .$$



visual mnemonic

or

$$x_1 \begin{vmatrix} y_2 & z_2 \\ y_3 & z_3 \end{vmatrix} - x_2 \begin{vmatrix} y_1 & z_1 \\ y_3 & z_3 \end{vmatrix} + x_3 \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}.$$

etc.

§5. The cross-product

The alternating property implies that if any vector is repeated, we get 0:

e.g., $\mathcal{D}(\vec{x},\vec{x},\vec{y}) \overset{\text{swap}}{=} -\mathcal{D}(\vec{x},\vec{x},\vec{y})$

$$\Rightarrow 2\mathcal{D}(\vec{x},\vec{x},\vec{y}) = 0 \Rightarrow \mathcal{D}(\vec{x},\vec{x},\vec{y}) = 0.$$

So $\mathcal{D}(\vec{x},\vec{y},\vec{z}) = \mathcal{D}\left(\sum_i x_i \hat{e}_i, \sum_j y_j \hat{e}_j, \sum_k z_k \hat{e}_k\right)$

$$\underset{\substack{\text{multi}\\ \text{-linearity}}}{=} \sum_{i,j,k} x_i y_j z_k \underbrace{\mathcal{D}(\hat{e}_i, \hat{e}_j, \hat{e}_k)}_{\text{is 0 if any 2 equal}}$$

$$= \sum_{i\neq j\neq k} x_i y_j z_k \cdot \underbrace{(\pm 1)}_{\text{depending on } i,j,k}$$

$$= x_1y_2z_3 + x_2y_3z_1 + x_3y_1z_2 - x_1y_3z_2 - x_2y_1z_3 - x_3y_2z_1 .$$

$$\left|\begin{matrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{matrix}\right| \begin{matrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{matrix}$$

visual mnemonic

or

$$x_1 \begin{vmatrix} y_2 & z_2 \\ y_3 & z_3 \end{vmatrix} - x_2 \begin{vmatrix} y_1 & z_1 \\ y_3 & z_3 \end{vmatrix} + x_3 \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix} .$$

etc.

§5. The cross-product

- Define, for vectors in $\mathbb{R}^3$,

$$\vec{x}\times\vec{y} = \text{“}\begin{vmatrix} \hat{e}_1 & x_1 & y_1 \\ \hat{e}_2 & x_2 & y_2 \\ \hat{e}_3 & x_3 & y_3 \end{vmatrix}\text{”} = \hat{e}_1 \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} - \hat{e}_2 \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} + \hat{e}_3 \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$$

$$= \begin{pmatrix} x_2y_3 - x_3y_2 \\ -x_1y_3 + x_3y_1 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

- Since determinants are alternating in columns (i.e. swapping 2 of them changes the sign), we have

$$\vec{y}\times\vec{x} = -\vec{x}\times\vec{y}.$$

§5. The cross-product

- Define, for vectors in $\mathbb{R}^3$,

$$\vec{x}\times\vec{y} = \text{``}\begin{vmatrix} \hat{e}_1 & x_1 & y_1 \\ \hat{e}_2 & x_2 & y_2 \\ \hat{e}_3 & x_3 & y_3 \end{vmatrix}\text{''} = \hat{e}_1\begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} - \hat{e}_2\begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} + \hat{e}_3\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$$

$$= \begin{pmatrix} x_2y_3 - x_3y_2 \\ -x_1y_3 + x_3y_1 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

- Since determinants are alternating in columns (i.e. swapping 2 of them changes the sign), we have

$$\vec{y}\times\vec{x} = -\vec{x}\times\vec{y}.$$

- $$\mathcal{D}(\vec{x},\vec{y},\vec{z}) = x_1\begin{vmatrix} y_2 & z_2 \\ y_3 & z_3 \end{vmatrix} + x_2\left(-\begin{vmatrix} y_1 & z_1 \\ y_3 & z_3 \end{vmatrix}\right) + x_3\begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}$$

$$= \vec{x}\cdot(\vec{y}\times\vec{z})$$ scalar triple-product

- $\vec{x}\cdot(\vec{x}\times\vec{y}) = D(\vec{x},\vec{x},\vec{y}) = 0$

  $\vec{y}\cdot(\vec{x}\times\vec{y}) = D(\vec{y},\vec{x},\vec{y}) = 0$

  $\Rightarrow \vec{x}\times\vec{y} \perp \vec{x},\vec{y}$.

- $\|\vec{x}\times\vec{y}\|^2 + (\vec{x}\cdot\vec{y})^2 =$

$$(\vec{x}\times\vec{y})\cdot(\vec{x}\times\vec{y}) + \Big(\sum_i x_i y_i\Big)^2 =$$

$$\sum_{i<j}(x_i y_j - x_j y_i)^2 + \sum_i x_i^2 y_i^2 + \sum_{i\neq j} x_i y_i x_j y_j =$$

$$\sum_{i<j} x_i^2 y_j^2 + \sum_{i<j} x_j^2 y_i^2 \cancel{- 2\sum_{i<j} x_i y_i x_j y_j} + \sum_i x_i^2 y_i^2 \cancel{+ 2\sum_{i<j} x_i y_i x_j y_j} =$$

$$\sum_{i\neq j} x_i^2 y_j^2 + \sum_i x_i^2 y_i^2 = \Big(\sum_i x_i^2\Big)\Big(\sum_j y_j^2\Big) = \|\vec{x}\|^2\|\vec{y}\|^2.$$

§5. The cross-product

- Define, for vectors in $\mathbb{R}^3$,

$$\underline{\vec{x}\times\vec{y}} = \text{``}\begin{vmatrix} \hat{e}_1 & x_1 & y_1 \\ \hat{e}_2 & x_2 & y_2 \\ \hat{e}_3 & x_3 & y_3 \end{vmatrix}\text{''} = \hat{e}_1\begin{vmatrix} x_2 & y_2 \\ x_3 & y_3\end{vmatrix} - \hat{e}_2\begin{vmatrix} x_1 & y_1 \\ x_3 & y_3\end{vmatrix} + \hat{e}_3\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2\end{vmatrix}$$

$$= \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ -x_1 y_3 + x_3 y_1 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}.$$

- Since determinants are alternating in columns (i.e. swapping 2 of them changes the sign), we have

$$\vec{y}\times\vec{x} = -\vec{x}\times\vec{y}.$$

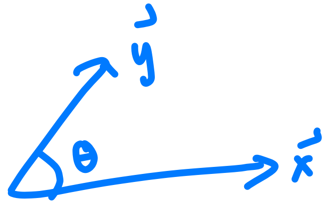
- $D(\vec{x},\vec{y},\vec{z}) = x_1\begin{vmatrix} y_2 & z_2 \\ y_3 & z_3\end{vmatrix} + x_2\left(-\begin{vmatrix} y_1 & z_1 \\ y_3 & z_3\end{vmatrix}\right) + x_3\begin{vmatrix} y_1 & z_1 \\ y_2 & z_2\end{vmatrix}$

  $= \vec{x}\cdot(\vec{y}\times\vec{z})$ scalar triple-product

- $\vec{x}\cdot(\vec{x}\times\vec{y}) = D(\vec{x},\vec{x},\vec{y}) = 0$

  $\vec{y}\cdot(\vec{x}\times\vec{y}) = D(\vec{y},\vec{x},\vec{y}) = 0$

  $\Rightarrow \ \vec{x}\times\vec{y} \perp \vec{x},\vec{y}.$



- $\|\vec{x}\times\vec{y}\|^2 + (\vec{x}\cdot\vec{y})^2 =$

$$(\vec{x}\times\vec{y})\cdot(\vec{x}\times\vec{y}) + \Big(\sum_i x_i y_i\Big)^2 =$$

$$\sum_{i<j}(x_i y_j - x_j y_i)^2 + \sum_i x_i^2 y_i^2 + \sum_{i\neq j} x_i y_i x_j y_j =$$

$$\sum_{i<j} x_i^2 y_j^2 + \sum_{i<j} x_j^2 y_i^2 \cancel{-2\sum_{i<j} x_i y_i x_j y_j} + \sum_i x_i^2 y_i^2 \cancel{+2\sum_{i<j} x_i y_i x_j y_j} =$$

$$\sum_{i\neq j} x_i^2 y_j^2 + \sum_i x_i^2 y_i^2 = \Big(\sum_i x_i^2\Big)\Big(\sum_j y_j^2\Big)$$

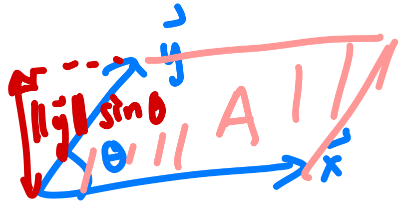
$$= \|\vec{x}\|^2\|\vec{y}\|^2.$$

Since $\vec{x}\cdot\vec{y} = \|\vec{x}\|\|\vec{y}\|\cos\theta$ we get

$$\|\vec{x}\times\vec{y}\|^2 = \|\vec{x}\|^2\|\vec{y}\|^2 - \|\vec{x}\|^2\|\vec{y}\|^2\cos^2\theta$$

$$= \|\vec{x}\|^2\|\vec{y}\|^2\sin^2\theta.$$

- $\vec{x}\cdot(\vec{x}\times\vec{y}) = D(\vec{x},\vec{x},\vec{y}) = 0$
  $\vec{y}\cdot(\vec{x}\times\vec{y}) = D(\vec{y},\vec{x},\vec{y}) = 0$
  $\Rightarrow \vec{x}\times\vec{y} \perp \vec{x},\vec{y}.$



- $\|\vec{x}\times\vec{y}\|^2 + (\vec{x}\cdot\vec{y})^2 =$

$$(\vec{x}\times\vec{y})\cdot(\vec{x}\times\vec{y}) + \Big(\sum_i x_i y_i\Big)^2 =$$

$$\sum_{i<j}(x_i y_j - x_j y_i)^2 + \sum_i x_i^2 y_i^2 + \sum_{i\neq j} x_i y_i x_j y_j =$$

$$\sum_{i<j} x_i^2 y_j^2 + \sum_{i<j} x_j^2 y_i^2 \cancel{- 2\sum_{i<j} x_i y_i x_j y_j} + \sum_i x_i^2 y_i^2 \cancel{+ 2\sum_{i<j} x_i y_i x_j y_j} =$$

$$\sum_{i\neq j} x_i^2 y_j^2 + \sum_i x_i^2 y_i^2 = \Big(\sum_i x_i^2\Big)\Big(\sum_j y_j^2\Big)$$

$$= \|\vec{x}\|^2\|\vec{y}\|^2.$$

Since $\vec{x}\cdot\vec{y} = \|\vec{x}\|\|\vec{y}\|\cos\theta$ we get

$$\|\vec{x}\times\vec{y}\|^2 = \|\vec{x}\|^2\|\vec{y}\|^2 - \|\vec{x}\|^2\|\vec{y}\|^2\cos^2\theta$$

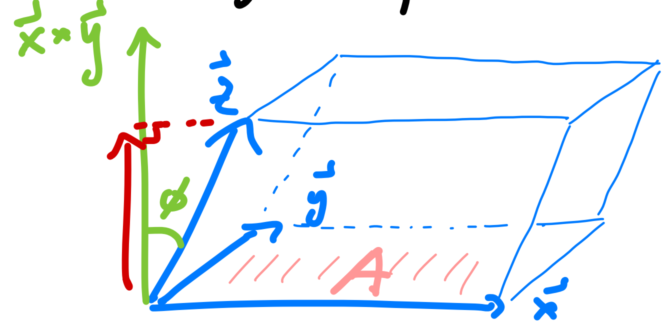
$$= \|\vec{x}\|^2\|\vec{y}\|^2\sin^2\theta.$$

Taking square roots gives

$$\|\vec{x}\times\vec{y}\| = \|\vec{x}\|\|\vec{y}\|\,|\sin\theta| = A$$

$\Rightarrow$ $\|\vec{x}\times\vec{y}\|$ is the area of the parallelogram spanned by $\vec{x}$ & $\vec{y}$.

- Now consider the parallelepiped spanned by $\vec{x}, \vec{y}$, & $\vec{z}$:



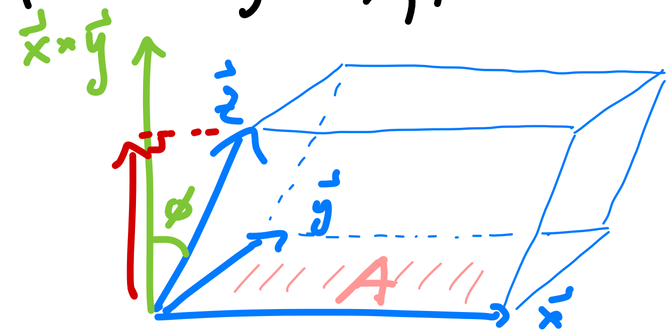
Its volume $V = A\|\vec{z}\|\cos\phi$

Taking square roots gives

$$\|\vec{x} \times \vec{y}\| = \|\vec{x}\| \, \|\vec{y}\| \, |\sin\theta| = A$$

$\Longrightarrow$ $\|\vec{x} \times \vec{y}\|$ is the area of the parallelogram spanned by $\vec{x}$ & $\vec{y}$.

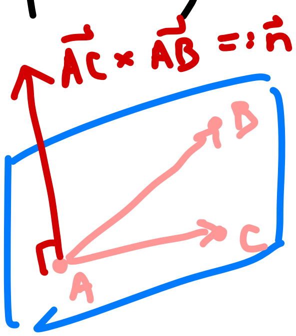
- Now consider the parallelepiped spanned by $\vec{x}, \vec{y},$ & $\vec{z}$:



Its volume
$$\begin{aligned} V &= A \, \|\vec{z}\| \, |\cos\phi| \\ &= \|\vec{x} \times \vec{y}\| \, \|\vec{z}\| \, |\cos\phi| \\ &= |\vec{z} \cdot (\vec{x} \times \vec{y})| \\ &= |\mathcal{D}(\vec{z}, \vec{x}, \vec{y})| \\ &= |\mathcal{D}(\vec{x}, \vec{y}, \vec{z})| . \end{aligned}$$

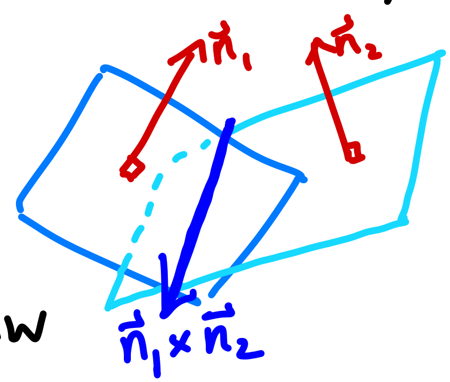
The scalar triple-product (i.e. determinant) itself is something called the signed volume, which is positive if $\vec{x}\times\vec{y}$ has the same direction as the projection of $\vec{z}$ to it.

We can use cross-product to compute areas of triangles ($\frac{1}{2}$ that of parallelogram), equations of planes



(by computing a normal vector as shown and writing $(\vec{x}-A)\cdot\vec{n} = 0$),

and parametrizing the intersection of two planes, as you'll see in the HW

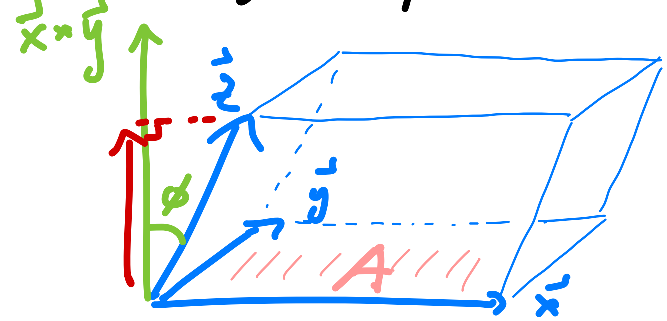


Taking square roots gives

$$\|\vec{x}\times\vec{y}\| = \|\vec{x}\|\,\|\vec{y}\|\,|\sin\theta| = A$$

$\Longrightarrow$ $\|\vec{x}\times\vec{y}\|$ is the area of the parallelogram spanned by $\vec{x}$ & $\vec{y}$.

- Now consider the parallelepiped spanned by $\vec{x}, \vec{y},$ & $\vec{z}$:



Its volume
$$\begin{aligned} V &= A\,\|\vec{z}\|\,|\cos\phi| \\ &= \|\vec{x}\times\vec{y}\|\,\|\vec{z}\|\,|\cos\phi| \\ &= |\vec{z}\cdot(\vec{x}\times\vec{y})| \\ &= |\mathcal{D}(\vec{z},\vec{x},\vec{y})| \\ &= |\mathcal{D}(\vec{x},\vec{y},\vec{z})|. \end{aligned}$$