New Topic: Matrix Multiplication.

Fairly modern idea.

Surprisingly fruitful (you get more out than you put in).

Consider matrices

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{\ell 1} & \cdots & a_{\ell m} \end{pmatrix} \quad \& \quad B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$$

"$\ell \times m$" = ℓ rows, m cols   "$m \times n$" = m rows, n cols.

Since # cols $A = m =$ # rows B, we can define the matrix product AB, which has ℓ rows & n cols.

Definition: Let $AB = (c_{ij})$

i,j entry of AB

$$= c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}$$

Special Case: $l=1$, $n=1$.

$$A = (a_{11} \ a_{12} \ \cdots \ a_{1m})$$

$$B = \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \end{pmatrix}$$

AB = dot product

$$= \sum_{k=1}^{m} a_{1k} b_{k1}$$

Notation: From now on we think of vectors in $\mathbb{R}^m$ as column vectors:

$$\vec{v} = (v_1, v_2, \dots, v_m) = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix}$$

$m \times 1$ matrix.

To talk about rows we use transposition:

If $A = (a_{ij})$

then $A^T = (a_{ji})$

Using this language:

$$\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v}$$

dot product

row times column.

There are many ways to think about matrix product:

dot product

ij entry of $AB = (i\text{th row } A)(j\text{th col } B)$

$(j\text{th col } AB) = A(j\text{th col } B)$

$(i\text{th row } AB) = (i\text{th row } A)B$

$$AB = \sum_{k=1}^{m} (k\text{th col } A)(k\text{th row } B)$$

e.g.

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \quad \& \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$AB = \begin{pmatrix} (1\,1\,1)\begin{pmatrix}1\\1\\0\end{pmatrix} & (1\,1\,1)\begin{pmatrix}0\\1\\1\end{pmatrix} \\ (1\,2\,3)\begin{pmatrix}1\\1\\0\end{pmatrix} & (1\,2\,3)\begin{pmatrix}0\\1\\1\end{pmatrix} \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 3 & 5 \end{pmatrix}$$

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

$$AB = \begin{pmatrix} (1\,1\,1)\begin{pmatrix}1&0\\1&1\\0&1\end{pmatrix} \\ (1\,2\,3)\begin{pmatrix}1&0\\1&1\\0&1\end{pmatrix} \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 3 & 5 \end{pmatrix}$$

$$AB = \begin{pmatrix}1\\1\end{pmatrix}(1\;0) + \begin{pmatrix}1\\2\end{pmatrix}(1\;1) + \begin{pmatrix}1\\3\end{pmatrix}(0\;1)$$

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

$$= \begin{pmatrix}2&2\\3&5\end{pmatrix} \quad \checkmark$$

Each of these is a special case of a general rule:

BLOCK MULTIPLICATION:

$$A = \left(\begin{array}{c|c|c} A_{11} & \cdots & A_{1m} \\ \hline \vdots & \ddots & \vdots \\ \hline A_{\ell 1} & \cdots & A_{\ell m} \end{array}\right)$$

$$B = \left(\begin{array}{c|c|c} B_{11} & \cdots & B_{1n} \\ \hline \vdots & \ddots & \vdots \\ \hline B_{m1} & \cdots & B_{mn} \end{array}\right)$$

If $\forall k$, # cols A_{ik} = # rows B_{kj}

Then the product matrix AB is partitioned &

$$ij\text{th chunk of } AB = \sum_{k=1}^{m} A_{ik} B_{kj}$$

products of matrices.

e.g. $\left(\begin{array}{ccc} 1 & 1 & 1 \\ \hline 1 & 2 & 3 \end{array}\right)\left(\begin{array}{c|c} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{array}\right)$

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}\left(\begin{array}{c|c} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc} 1 & 1 & 1 \\ \hline 1 & 2 & 3 \end{array}\right)\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\left(\begin{array}{c|c|c} 1 & 1 & 1 \\ 1 & 2 & 3 \end{array}\right)\left(\begin{array}{cc} 1 & 0 \\ \hline 1 & 1 \\ \hline 0 & 1 \end{array}\right)$$

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

KEY PROPERTY: Associativity.

$$\left.\begin{array}{l} A = l \times m \\ B = m \times n \end{array}\right\} \text{ AB exists \& is } l \times n.$$

Also $C = n \times p$, then

BC exists & is $m \times p$

$(AB)C$ exists & is $l \times p$

$A(BC)$ exists & is $l \times p$.

Theorem: $A(BC) = (AB)C$.

Extremely useful.

Brute force proof is unenlightening.

There is a conceptual explanation.

Consider vector spaces V, W over $\mathbb{R}$ (or $\mathbb{C}$). A function $T: V \longrightarrow W$ is called a linear transformation when:

- $T(\vec{0}) = \vec{0}$
- $T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2)$
- $T(a\vec{v}) = aT(\vec{v})$.

Or we can say it in one step:

$$\underbrace{T\left(\sum a_i \vec{v}_i\right) = \sum a_i T(\vec{v}_i)}_{\text{T preserves linear combinations.}}$$

Why? Many natural operations are linear:

- diff : Funcs $\to$ funcs
- int : funcs $\to$ scalars
- Inner products are "bilinear":

$\langle \vec{v}, - \rangle$: vecs $\to$ scalars

$\langle -, \vec{v} \rangle$: vecs $\to$ scalars.

[Over $\mathbb{C}$, $\langle -, \vec{v} \rangle$ is "conjugate-linear" so $\langle -, - \rangle$ is called "sesquilinear" (one & a half times linear).]

If V & W are finite dim.
then choosing bases turns

linear transformations $\rightsquigarrow$ matrices.

To keep things simple, use Euclidean space & standard bases.

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

Each basis vector $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n \in \mathbb{R}^n$
gets sent to a vector in $\mathbb{R}^m$:

$$T(\vec{e}_1), T(\vec{e}_2), \ldots, T(\vec{e}_n) \in \mathbb{R}^m.$$

Encode this by a matrix

$$[T] = \begin{bmatrix} | & | & & | \\ T(\vec{e}_1) & T(\vec{e}_2) & \cdots & T(\vec{e}_n) \\ | & | & & | \end{bmatrix}$$

m rows & n columns.

This matrix determines the function T. Given $\vec{v} \in \mathbb{R}^n$

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

$$T(\vec{v}) = T(v_1\vec{e}_1 + v_2\vec{e}_2 + \cdots + v_n\vec{e}_n)$$

$$= v_1 T(\vec{e}_1) + v_2 T(\vec{e}_2) + \cdots + v_n T(\vec{e}_n)$$

$$= \sum v_i (i\text{th col of } [T])$$

$$= [T]\vec{v}$$

special case of matrix mult.

Summary: $T: \mathbb{R}^n \to \mathbb{R}^m$ linear.

Then for any col vector $\vec{v} \in \mathbb{R}^n$,

apply function

matrix mult.

$$\star \quad \underbrace{T(\vec{v})}_{\substack{m\times 1 \\ \text{column}}} = \underbrace{[T]}_{\substack{m\times n \\ \text{matrix}}} \underbrace{\vec{v}}_{\substack{n\times 1 \\ \text{column}}}$$

Suppose we have another linear trans:

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "T"] \arrow[rr, bend right, "S \circ T"'] & \mathbb{R}^m \arrow[r, "S"] & \mathbb{R}^\ell \end{tikzcd}$$

S, T linear $\Longrightarrow$ $S \circ T$ linear.

Proof:

$$S \circ T \left(\sum a_i \vec{v}_i \right)$$

$$= S\left(T\left(\sum a_i \vec{v}_i\right)\right)$$

$$= S\left(\sum a_i T(\vec{v}_i)\right)$$

$$= \sum a_i S(T(\vec{v}_i))$$

$$= \sum a_i (S \circ T)(\vec{v}_i) \quad \checkmark$$

Thus we have a matrix $[S \circ T]$.

Note: $[S]$ is $\ell \times m$

$[T]$ is $m \times n$

$[S \circ T]$ is $\ell \times m$.

Theorem:

function composition matrix multiplication

$$[S \circ T] = [S][T]$$

Proof: We need to show

$$j\text{th col } [S \circ T] = j\text{th col } [S][T].$$

L.H.S.

DEF of $[S \circ T]$

$$\begin{aligned} j\text{th col } [S \circ T] &= S \circ T(\vec{e}_j) \\ &= S(T(\vec{e}_j)) \end{aligned}$$

R.H.S

DEF

$$j\text{th col } [S][T] = [S](j\text{th col } [T])$$

$$= [S]\, T(\vec{e}_j) \qquad \text{DEF}$$

$$= S(T(\vec{e}_j)). \quad \text{from } \star .$$

Corollary: Matrix multiplication is associative.

$$A(BC) = (AB)C.$$

<u>Proof</u>: let $A = [R]$
$B = [S]$
$C = [T]$.

Then

$$A(BC) = [R]\big([S][T]\big)$$

$$= [R][S \circ T]$$

$$= [R \circ (S \circ T)]$$

$$= [(R \circ S) \circ T]$$

$$= [R \circ S][T]$$

$$= ([R][S])[T]$$

$$= (AB)C \qquad \checkmark$$

Automatic once we see matrix mult as composition of linear functions.