Lecture 16: More on matrix multiplication

Story so far:

$$\begin{tikzcd} X \arrow[r, "S"] \arrow[rr, bend left, "T\circ S"] & Y \arrow[r, "T"] & Z \end{tikzcd} \quad \text{linear trans}$$

$$\alpha \qquad \beta \qquad \gamma \quad \text{bases}$$

Thm: $[T\circ S]_\alpha^\gamma = [T]_\beta^\gamma [S]_\alpha^\beta$

Matrix multiplication

$$\underset{p\times m}{A}, \underset{m\times n}{B} \rightsquigarrow \underset{p\times n}{AB} \quad \text{defined via } (AB)_{ij} = \sum_{k=1}^{m} A_{ik}B_{kj}$$

In general, matrix mult is not commutative.

Thm: Matrix multiplication is

Associative: $A(BC) = (AB)C$

Distributative: $A(B+C) = AB + AC$

$(A+B)C = AC + BC.$

Pf: I'll check the first dist. prop, see text for the others. Suppose A is $p\times m$ with B and C $m\times n$. So at least both sides are $p \times n$ matrices. Now focusing on the (i,j) entry we have:

$$\big(A(B+C)\big)_{ij} = \sum_{k=1}^{m} A_{ik}(B+C)_{kj}$$

$$= \sum_{k=1}^{m} A_{ik}(B_{kj}+C_{kj})$$

$$= \left(\sum_{k=1}^{m} A_{ik}B_{kj}\right) + \left(\sum_{k=1}^{m} A_{ik}B_{kj}\right)$$

$$= (AB)_{ij} + (AC)_{ij} = (AB+AC)_{ij}$$

So $A(B+C) = AB + AC$ as desired. $\blacksquare$

[What about 0 and 1?]

If $0_{a \times b}$ is the all-zero matrix in $M_{a\times b}(\mathbb{R})$

and $A \in M_{m\times n}(\mathbb{R})$ then

$$A\,0_{n\times p} = 0_{m \times p} \qquad \underline{Ex}: \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$0_{p\times m} A = 0_{p\times n}$$

<u>Identity matrix</u>: $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

In general $I_n \in M_{n\times n}(\mathbb{R})$ with

$I_{ij} = \delta_{ij}$ where $\delta_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i\neq j \end{cases}$ (Kronecker delta)

Again, if $A \in M_{m\times n}(\mathbb{R})$ we have

$$AI_n = A \quad \text{and} \quad I_m A = A$$

<u>Ex</u>: $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$

$$A \cdot I_2 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$$

$$I_3 \cdot A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$$

<u>Reason for nomenclature</u>:

Suppose V has dim n. The <u>identity</u> <u>transformation</u> $I_V : V \longrightarrow V$ is defined by $I_V(v) = v$ for all $v \in V$.

If β is any basis for V, then $[I_V]_\beta = I_n$.

[Think about for $\mathbb{R}^n$ or even just $\mathbb{R}^2$...]

If $T: V \longrightarrow W$ is linear then $T \circ I_V = T$

and so $[T]_\beta^\gamma = [T]_\beta^\gamma [I_V]_\beta^\beta = [T]_\beta^\gamma I_n$,

which matches the fact that $AI_n = A$ we saw above.

[Know $T \rightsquigarrow$ matrix, now let's reverse this process...]

Suppose $A \in M_{m \times n}(\mathbb{R})$. Define the <u>left-multiplication transformation</u> $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ by $L_A(x) = Ax$ where $x \in \mathbb{R}^n$ is viewed as a column vector.

<u>Ex:</u> $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 4 \end{pmatrix}$ $\quad L_A : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$

$$x = \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix} \qquad L_A(x) = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 4 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 17 \\ 43 \end{pmatrix}$$

<u>Thm.</u> $A \in M_{m \times n}(\mathbb{R})$. Then

1) $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is linear

2) $[L_A]_{\text{std}}^{\text{std}} = A$

3) If $B \in M_{n \times p}(\mathbb{R})$ then $L_{AB} = L_A \circ L_B$ which is a linear trans $\mathbb{R}^p \longrightarrow \mathbb{R}^m$.

Pf: ① Suppose $x_1, x_2 \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

Then

$$\begin{aligned} L_A(cx_1 + x_2) &= A(cx_1 + x_2) \\ &= A(cx_1) + Ax_2 \quad \left(\text{by dist of mat mult.}\right) \\ &= c(Ax_1) + Ax_2 \\ &= c\,L_A(x_1) + L_A(x_2) \end{aligned}$$

as required.

② $L_A(e_i) = Ae_i = i^{\text{th}}$ column of A

So $[L_A(e_i)]_{\text{std}} = i^{\text{th}}$ column of A,

and so $[L_A]_{\text{std}}^{\text{std}} = A$.

③ Throughout, use the std bases for $\mathbb{R}^p$, $\mathbb{R}^n$, and $\mathbb{R}^m$. Now

$$[L_A \circ L_B] = [L_A][L_B] = AB.$$

and so $[L_{AB}] = [L_A \circ L_B]$.

As any linear tranformation is

determined by what it does to a basis, we must have $L_{AB} = L_A \circ L_B$ as claimed. $\blacksquare$

Alternate proof that $(AB)C = A(BC)$ when A, B, C are matrices where these products make sense:

$$\begin{aligned} L_{(AB)C} &= L_{AB} \circ L_C = (L_A \circ L_B) \circ L_C \\ &= L_A \circ (L_B \circ L_C) \\ &= L_A \circ (L_{BC}) = L_{A(BC)} \end{aligned}$$

Taking the matrices with respect to the standard bases gives $(AB)C = A(BC)$ as claimed. $\blacksquare$