MATH 20F LINEAR ALGEBRA

Lecture 7: 2.1 Matrix algebra. An $m \times n$ matrix A is a collection of numbers (called entries) $[\, a_{ij} \,]_{1\le i\le m,\, 1\le j\le n}$

$$\begin{array}{cc} \begin{array}{ccccc} & & j\text{ th }\mathbf{column} & & \end{array} & \\ A = \left[\begin{array}{ccccc} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ & & \vdots & & \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ & & \vdots & & \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{array}\right] & i\text{ th }\mathbf{row} \end{array}$$

we often write it in terms of column vectors: $A = [\, \mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n \,]$, $\qquad \mathbf{a}_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$

or just $A = [\, a_{ij} \,]$. We will use the notation $(A)_{ij}$ to denote the (i,j)**th entry** a_{ij}.

Addition: Two $m \times n$ matrices A and B can be summed together by summing their entries.

$$(A+B)_{ij} = (A)_{ij} + (B)_{ij}$$

Example: $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} 1+0 & 2+1 \\ 3-3 & 4+2 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 6 \end{bmatrix}$

Subtraction can be defined in a similar way.

Scalar multiplication: An $m \times n$ matrix A can be multiplied by a scalar λ by multiplying all the entries by λ.

$$(\lambda A)_{ij} = \lambda (A)_{ij}$$

Example: $3 \cdot \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 3\cdot 1 & 3\cdot 2 & 3\cdot 3 \\ 3\cdot 4 & 3\cdot 5 & 3\cdot 6 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 9 \\ 12 & 15 & 18 \end{bmatrix}$

The zero matrix: There is a zero matrix of each shape. For example the 4×4 zero matrix is

$$0 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Laws of arithmetic: If A, B and C are $m \times n$ matrices, 0 is the zero $m \times n$ matrix and r, s as scalars, then we can the usual laws of arithmetic hold for matrices, for example

$$A + B = B + A, \qquad A + 0 = A, \qquad A - A = 0, \qquad (A+B)+C = A+(B+C)$$
$$(rs)A = r(sA), \qquad r(A+B) = rA + rB \qquad \text{etc.}$$

Multiplication of Matrices. How do we define the product of two matrices? We already described how to multiply a matrix and a column vector. If $A =$

$\begin{bmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \vdots & \vdots \\ A_{m1} & \dots & A_{mn} \end{bmatrix}$ is an $m \times n$ matrix and if $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ is an $n \times 1$ column vector then $A\mathbf{v}$ is an $m \times 1$ column vector. The formula is

$$(A\mathbf{v})_i \ = \ A_{i1}v_1 \ + \ A_{i2}v_2 \ + \ \dots \ + \ A_{in}v_n.$$

Rule: "*Dot the ith row of A with* $\mathbf{v}$ *to get the ith element of* $A\mathbf{v}$."
Geometric meaning: Multiplication by A gives a **linear map** from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Question: In general when can you multiply two matrices A and B? **Answer**: When the number of columms of A equals the number of rows of B.
Rule: If A is an $m \times n$ matrix and B is an $n \times p$ matrix then AB is an $m \times p$ matrix. The ijth entry of AB is the ith row of A dot the jth column of B:

$$(AB)_{ij} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj}$$

$$i \text{ th } \mathbf{row} \quad \begin{bmatrix} \\ a_{i1} \cdots a_{in} \\ \\ \end{bmatrix} \underset{j \text{ th } \mathbf{column}}{\begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix}} = \underset{j \text{ th } \mathbf{column}}{\begin{bmatrix} \vdots \\ \cdots \ (AB)_{ij} \ \cdots \\ \vdots \end{bmatrix}} \quad i \text{ th } \mathbf{row}$$

Alternative rule: If we write B as columns $B\mathbf{x} = [\,\mathbf{b}_1 \ \cdots \ \mathbf{b}_p\,]$, then
$AB = [\,A\mathbf{b}_1 \quad \cdots \quad A\mathbf{b}_p\,]$

Example: Let $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}$. Find AB
Solution:

$$AB = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1(-1) + 2 \cdot 1 & 1 \cdot 1 + 2 \cdot 2 \\ 0(-1) + (-1)1 & 0 \cdot 1 + (-1)2 \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ -1 & -2 \end{bmatrix}$$

If it makes things easier:

$$\begin{array}{rc} & \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix} \\ AB = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} & \begin{bmatrix} 1 & 5 \\ -1 & -2 \end{bmatrix} \end{array}$$

Or using the alternative rule:

$$A\mathbf{b}_1 = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot (-1) + 2 \cdot 1 \\ 0 \cdot (-1) + (-1) \cdot 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

$$A\mathbf{b}_2 = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + 2 \cdot 2 \\ 0 \cdot 1 + (-1) \cdot 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

Hence

$$AB = A\Big[\mathbf{b}_1\mathbf{b}_2\Big] = \Big[A\mathbf{b}_1 A\mathbf{b}_2\Big] = \begin{bmatrix} 1 & 5 \\ -1 & -2 \end{bmatrix}$$

Theorem. *If A is an $m \times n$ matrix and B is an $n \times p$ matrix and $\mathbf{v} \in \mathbb{R}^p$, then*

$$(AB)\mathbf{v} \;=\; A(B\mathbf{v}).$$

Geometric Interpretation: The map $\mathbf{v} \to B\mathbf{v}$ is a linear map sending $\mathbb{R}^p \to \mathbb{R}^n$. The map $\mathbf{w} \to A\mathbf{w}$ is a linear map sending $\mathbb{R}^n \to \mathbb{R}^m$. But the composition of these maps is

$$\begin{array}{ccccc} \mathbf{v} & \to & (B\mathbf{v}) & \to & A(B\mathbf{v}) \\ \mathbb{R}^p & \to & \mathbb{R}^n & \to & \mathbb{R}^m \end{array}$$

is given by multiplying $\mathbf{v}$ by the matrix AB:

$$\begin{array}{ccc} \mathbf{v} & \longrightarrow & (AB)\mathbf{v} \\ \mathbb{R}^p & \longrightarrow & \mathbb{R}^m. \end{array}$$

Proof of the Theorem. We calculate $A(B\mathbf{x})$:

$$B\mathbf{x} = [\,\mathbf{b}_1 \; \cdots \; \mathbf{b}_p\,] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = B\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_p\mathbf{b}_p.$$

and hence by linearity

$$A(B\mathbf{x}) = A\big(x_1\mathbf{b}_1 + \cdots + x_p\mathbf{b}_p\big) = x_1 A\mathbf{b}_1 + \cdots + x_p A\mathbf{b}_p = [\,A\mathbf{b}_1 \; \cdots \; A\mathbf{b}_p\,] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = (AB) \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Warning: Rules that hold for scalar mulplication don't always hold for matrix multiplication!

Definition: The identity matrix is $I = [\,\delta_{ij}\,]$, where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$:

$$I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \text{in case } 4 \times 4.$$

Definition: The transpose A^T is the matrix with rows and columns interchanged, $(A^T)_{ij} = (A)_{ji}$

Example: If $A = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 0 & -1 \\ 4 & 5 & 2 \end{bmatrix}$ then $A^T = \begin{bmatrix} 1 & -2 & 4 \\ 2 & 0 & 5 \\ 3 & -1 & 2 \end{bmatrix}$.

Quiz: Which are true? Which of these statements are true when they make sense? Hint: think in terms of linear maps!

(a) $$A(B+C) = AB + AC.$$

(b) $$AB = BA.$$

(c) $$AB = 0 \quad \Rightarrow \quad A = 0 \text{ or } B = 0.$$

(d) $$AI = A.$$

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

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

Solution (a). $B = [\mathbf{b}_1 \dots \mathbf{b}_p]$, $C = [\mathbf{c}_1 \dots \mathbf{c}_p]$ then $B + C = [\mathbf{b}_1 + \mathbf{c}_1 \dots \mathbf{b}_p + \mathbf{c}_p]$. Hence

$$A(B+C) = [A(\mathbf{b}_1 + \mathbf{c}_1) \dots A(\mathbf{b}_p + \mathbf{c}_p)] = [A\mathbf{b}_1 \dots A\mathbf{b}_p] + [A\mathbf{c}_1 \dots A\mathbf{c}_p].$$

(b). There are linear maps which do not commute. For example rotation by $\frac{\pi}{2}$ counterclockwise and reflection in the x_1 axis do not commute. In terms of matrices,

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

But

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

Why do people expect things to be commutative in math when they are not commutative in real life? It is not the same thing to first put on the shoes and then the socks as it is to first put on the socks and then the shoes?

(c). If you project onto the x-axis and then the y-axis you just get the zero map.

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

(e) $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. $(AB)^T = \left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right)^T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

$$\begin{aligned} A^T B^T &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^T \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}^T = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ B^T A^T &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

$((AB)^T)_{ij} = (AB)_{ji} = A_{j1}B_{1i} + \cdots + A_{jn}B_{ni} = (B^T)_{i1}(A^T)_{1j} + \cdots + (B^T)_{in}(A^T)_{nj} = (B^T A^T)_{ij}$.