APMA 4007 | Sep 12, 2024 | Lecture 4

Outline

Matrix - matrix mult

- columns
- rows
- elements (dot)
- rows & columns (outer product)
- block

Last time introduced matrix-matrix mult. w/ the promise that it could be used in elimination, which itself consists of strategic row operations (L.C. of rows) add/sub.

match

$$\underset{m \times n}{A} \underset{n \times K}{B} = \begin{bmatrix} A b_1 & \cdots & A b_K \end{bmatrix}$$

shape of result

By inspection each column of AB is a L.C. of the columns of A. What about rows of AB?

- Each row of AB has K entries
- Each row of B has K entries

A (correct) guess is that the rows of AB are L.C. of the rows of B.

To see this (warm-up): Recall that the dot product is m.-m. mult

$$\underset{\substack{1\times 3\\ A}}{\begin{bmatrix} 1 & 0 & -1 \end{bmatrix}} \underset{\substack{3\times 1\\ B}}{\begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}} = 1\cdot 1 + 0 \cdot 4 + (-1)7$$

$$= 1 - 7 = 6$$

= row 1 of B − row 3 of B.

what if B has more than one column? = (also L.C. of the "columns" of A)

same A

$$\begin{bmatrix} 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 8 \\ 4 & 5 \\ 7 & 2 \end{bmatrix} = \begin{bmatrix} -6 & 6 \end{bmatrix} = \text{row 1 of B} - \text{row 3 of B}$$

(same ans. dotting rows of A into columns of B) 1 always "hits" the 1st row, −1 the 3rd

In general

$$\underset{1 \times 3}{\begin{bmatrix} 1 & 0 & -1 \end{bmatrix}} \underset{3 \times K}{B} = \underset{1 \times K}{\text{row 1 of B} - \text{row 3 of B}}$$

- More generally, $\underset{1\times n}{y^T}\ \underset{n\times K}{B} = \underset{1\times K}{[\qquad\qquad]}$

$$\begin{bmatrix} \text{---} & y^T & \text{---} \end{bmatrix} \begin{bmatrix} | & & | \\ b_1 & \cdots & b_K \\ | & & | \end{bmatrix} = \begin{bmatrix} y^T \vec{b}_1 & y^T b_2 & \cdots & y^T b_K \end{bmatrix}$$

dot products

let's examine the dot products in more detail

$$\vec{y} = \begin{bmatrix} y_1 \\ | \\ y_n \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & \cdots & b_{1K} \\ b_{21} & \cdots & b_{2K} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nK} \end{bmatrix},$$

in $y^T B$
y is dotted into the columns of B

y_1 "hits" elements in 1st row of B
y_2 " " " " 2nd " " "
and so on

$$\begin{aligned} y^T B = \ & y_1 [\, b_{11} \quad b_{1K} \,] \\ + \ & y_2 [\, b_{21} \quad b_{2K} \,] \\ & \vdots \\ & y_n [\, b_{n1} \ \cdots \ b_{nK} \,] \end{aligned}$$

$y^T B$ = L. C. of rows of B

what if A has more than one row?   using rows

$$\underset{2\times3}{\begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}} \underset{3\times2}{\begin{bmatrix} 1 & 8 \\ 4 & 5 \\ 7 & 2 \end{bmatrix}} = \underset{2\times2}{\begin{bmatrix} -6 & 6 \\ 5 & 13 \end{bmatrix}} = \begin{bmatrix} \text{row 1} - \text{row 3} \\ \text{row 1} + \text{row 2} \end{bmatrix}$$

Using columns   dot products

$$\begin{bmatrix} 1\begin{bmatrix}1\\1\end{bmatrix} + 4\begin{bmatrix}0\\1\end{bmatrix} + 7\begin{bmatrix}-1\\0\end{bmatrix} & 8\begin{bmatrix}1\\1\end{bmatrix} + 5\begin{bmatrix}0\\1\end{bmatrix} + 2\begin{bmatrix}-1\\0\end{bmatrix} \end{bmatrix} = \begin{bmatrix} \text{row 1 A} \cdot \text{col 1 B} & \text{row 1 A} \cdot \text{col 2 B} \\ \text{row 2 A} \cdot \text{col 1 B} & \text{row 2 A} \cdot \text{col 2 B} \end{bmatrix}$$

$$\begin{bmatrix} -6 & 6 \\ 5 & 13 \end{bmatrix} \checkmark$$

And so on for larger matrices.

columns of AB = L.C. of columns A ← existence

rows of AB = L.C. of rows of B ← elimination

$[AB]_{ij}$ = row i of A dotted into column j of B.

So if $[AB]_{ij} = 0$ row i of A $\perp$ column j of B

and if $A\vec{z} = \vec{0}$, z is $\perp$ to the rows of A ← uniqueness

---

This can be seen more algebraically w/ equation

$$[AB]_{ij} = \sum_{l=1}^{n} \underbrace{A_{il}}_{\text{row } i} \underbrace{B_{lj}}_{\text{column } j}$$

in physics repeated indices summed $A_{il} B_{lj}$

Column view

$$\underbrace{[AB]}_{j\text{-th column of AB}}\begin{matrix} m \\ | \\ i \\ | \\ 1 \end{matrix}\ j = \sum_{l=1}^{n} \underbrace{A \begin{matrix} m \\ | \\ i \\ | \\ 1 \end{matrix}\ l}_{l\text{-th column of A}} \ B_{lj} = \text{L.C. of columns of A}$$

Row view

$$\underbrace{[AB]_{i}}_{i\text{th row of AB}}\ 1 — j — n = \sum_{l=1}^{n} A_{il}\ \underbrace{B_{l}\ 1 — j — n}_{l\text{th row of B}} = \text{L.C. of the rows of B}$$

(Since all elements of AB are to be computed 3 loops (i, j, l) could be used. The row & column views might correspond to different ~~or~~ ~~or~~ ordering of the loops + vectorization)

But wait! There is more

The outer product view (aka rows & columns at the same time)

dot product is also called inner product

match

3 x 1   1 x 2   ans. will be 3x2

$$\begin{bmatrix}1\\2\\3\end{bmatrix}\begin{bmatrix}-4 & 5\end{bmatrix} = ?$$

column view

$$\begin{bmatrix} -4\begin{bmatrix}1\\2\\3\end{bmatrix} & 5\begin{bmatrix}1\\2\\3\end{bmatrix}\end{bmatrix} = \begin{bmatrix}-4 & 5\\-8 & 10\\-12 & 15\end{bmatrix}$$

(all) both columns are multiples of $\begin{bmatrix}1\\2\\3\end{bmatrix}$

row view

$$\begin{bmatrix} 1\begin{bmatrix}-4 & 5\end{bmatrix}\\ 2\begin{bmatrix}-4 & 5\end{bmatrix}\\ 3\begin{bmatrix}-4 & 5\end{bmatrix}\end{bmatrix} = \begin{bmatrix}-4 & 5\\-8 & 10\\-12 & 15\end{bmatrix}$$

all rows are multiples of $\begin{bmatrix}-4 & 5\end{bmatrix}$

In general $\underset{m\times 1}{y}\ \underset{1\times n}{x^T}$ is an outer product and makes ~~sense~~ sense (is defined) for any 2 vetors regardless of their size.

- the result is a $m \times n$ matrix
- each row is a multiple of ~~y~~ x^T
- each column is a multiple of y

This is a rank-1 matrix but we don't know what rank means yet.

the outer product view of AB

sum of outer products

$$[AB]_{i\,(1 \ldots m),\ j\,(1 \ldots n)} = \sum_{\ell=1}^{n} A_{i\ell} \, B_{\ell j}$$

column ℓ of A

row ℓ of B

$$AB = \sum_{\ell=1}^{n} \big[\text{column } \ell \text{ of } A\big] \big[\text{ row } \ell \text{ of } B\big]$$

Back to our example

$$\begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 8 \\ 4 & 5 \\ 7 & 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 8 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 4 & 5 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} \begin{bmatrix} 7 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 8 \\ 1 & 8 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 4 & 5 \end{bmatrix} + \begin{bmatrix} -7 & -2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 5 & 13 \end{bmatrix}$$

Go to block matrix mult. maybe?

Sometimes matrix-matrix mult is not that hard.
For instance if one of the matrices is diagonal
only non-zero on diagonal

$$\begin{array}{cc} A & B \\ \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix} & \begin{bmatrix} 3 & 1 & 2 \\ 4 & -2 & 1 \\ 0 & 2 & 5 \end{bmatrix} \end{array} = \begin{bmatrix} 6 & 2 & 4 \\ -4 & 2 & -1 \\ 0 & 6 & 15 \end{bmatrix}$$

row views says diagonal ~~elements of~~ A scales the rows of B.

$$\begin{bmatrix} 3 & 1 & 2 \\ 4 & -2 & 1 \\ 0 & 2 & 5 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 6 & -1 & 6 \\ 8 & 2 & 3 \\ 0 & -2 & 15 \end{bmatrix}$$

column views says diagonal B scales the columns of A

NOTE $AB \neq BA$

Q: how much storage needed for a diagonal matrix?

Another easy matrix-matrix mult.

$I =$ square diagonal matrix w/ ones on diagonal

$m=3 \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$AI = A$ b/c each column scaled by 1

$IA = A$ b/c each row scaled by 1

jump to (*)

The outer product view is a special case of block matrix-matrix mult.

all we know is that the inner dim. mater

$$\underset{?\times 3}{\begin{bmatrix} | & | & | \\ a_1 & a_2 & a_3 \\ | & | & | \end{bmatrix}} \underset{3\times ?}{\begin{bmatrix} - & b_1^T & - \\ - & b_2^T & - \\ - & b_3^T & - \end{bmatrix}} = a_1 b_1^T + a_2 b_2^T + a_3 b_3^T$$

which is exactly like a dot product, except not.

More ~~generally~~ make matching cuts of the columns of A & rows of B

$$\underset{m\times n}{A} = \left[\underset{m\times n_1}{A_1} \;\middle|\; \underset{m\times n_2}{A_2} \right] \qquad B = \begin{bmatrix} \underset{n_1\times k}{B_1} \\ \underset{n_2\times k}{B_2} \end{bmatrix}$$

$n_1 + n_2 = n$

$$AB = [A_1 \; A_2] \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \underset{m\times n_1}{A_1} \underset{n_1\times k}{B_1} + \underset{m\times n_2}{A_2} \underset{n_2\times k}{B_2}$$

like a dot product

CAREFUL w/ order

Example

$$\underset{2\times 3}{\left[\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 5 \end{array}\right]} \underset{3\times 1}{\left[\begin{array}{c} -2 \\ 1 \\ \hline 1 \end{array}\right]} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 5 \end{bmatrix} 1$$

$$= \begin{bmatrix} -2 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$$

Likewise (if time)

$$\left[\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 5 \end{array}\right]\left[\begin{array}{cc} -2 & 7 \\ 1 & 3 \\ \hline 1 & 1 \end{array}\right] = \begin{bmatrix} -2 & 7 \\ 1 & 3 \end{bmatrix} + \begin{bmatrix} 3 & 3 \\ 5 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 10 \\ 6 & 8 \end{bmatrix}$$

Why does this work? Magic? No.

$$[AB]_{ij} = \sum_{l=1}^{n} A_{il} B_{lj}$$

$$= \underbrace{\sum_{l=1}^{n_1} A_{il} B_{lj}}_{A_1 \; B_1} + \underbrace{\sum_{l=n_1+1}^{n} A_{il} B_{lj}}_{A_2 \; B_2}$$

---

in AB # of col. A = # rows in B

in block m-m mult cuts in columns of A must match cuts in rows of B otherwise <u>error</u>

But additional cuts in the rows of A & columns of B do not have to match

match

$$\left[\begin{array}{c|c} \overset{m_1 \times n_1}{A_1} & \overset{m_1 \times m_2}{A_2} \\ A_3 & A_4 \end{array}\right]\left[\begin{array}{cc} \overset{n_1 \times K_1}{B_1} & B_2 \\ \hline \overset{n_2 \times K_1}{B_3} & B_4 \end{array}\right] \text{match} = \begin{bmatrix} \overset{m_1 \times n_1 \;\; n_1 \times K_1 \quad m_1 \times n_2 \;\; n_2 \times K_1}{A_1 B_1 + A_2 B_3} \\ \\ \end{bmatrix}$$

and so on. A little surprising.

A slight variation on the identity matrix is a permution matrix P which has the same rows & columns as I but no in the same order.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 4 & 2 & 8 \\ 6 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 \\ 6 & 4 & 0 \\ 4 & 2 & 8 \end{bmatrix}$$

rows 2 & 3 switched

$$\begin{bmatrix} 2 & 1 & 3 \\ 4 & 2 & 8 \\ 6 & 4 & 6 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 8 & 2 \\ 6 & 0 & 4 \end{bmatrix}$$

columns 2 & 3 switched