Apr. 14 — Matrix Computations

Apr. 14 — Matrix Computations

Why matrix algebra?
Makes many other formulas and computations clearer.

Ex: The β_0 and β_1 in the line-of-best-fit $y=\beta_0+\beta_1 x$ are:

(Matrix) $\boxed{\begin{pmatrix}\beta_0\\ \beta_1\end{pmatrix} = (A^T A)^{-1} A^T \vec{b}}$, $A=\begin{pmatrix}1 & x_1\\ \vdots & \vdots\\ 1 & x_n\end{pmatrix}$, $\vec{b}=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$.

(No matrices): $$\beta_1 = \frac{n\sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right)}{n\left(\sum_{i=1}^n x_i^2\right) - \left(\sum_{i=1}^n x_i\right)^2}$$

$\beta_0 =$ something equally messy.

I: Matrix basics:

* Matrices are rectangles of numbers, functions, ...
* An $m \times n$ matrix has m rows & n columns.
* The entry i rows from the top and j columns from the left in a matrix A is written a_{ij}, A_{ij} ...
* <u>Example:</u>

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}$$

is a matrix. It's 3×4. $A_{31} = 9$.

$A_{23} = 7$, A_{42} = doesn't exist.

Two matrices are the same if they have the same dimensions, and the corresponding entries match.

II: Adding, Subtracting, Transposing, and Scalar Multiplying Matrices

Adding: You can add or subtract two matrices only if the have the same size.
Then you add/subtract corresponding entries.

Example: $\begin{pmatrix} 1 & 0 & -1 \\ 2 & 3 & -4 \end{pmatrix} + \begin{pmatrix} 5 & 7 & -8 \\ -2 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 6 & 7 & -9 \\ 0 & 4 & 0 \end{pmatrix}.$

$\begin{pmatrix} 2 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} =$ does not exist.

$\begin{pmatrix} 1 & 4 \end{pmatrix} - \begin{pmatrix} 5 & 3 \end{pmatrix} = \begin{pmatrix} -4 & 1 \end{pmatrix}$

Scalar Multiplying: You can multiply any matrix by any scalar (number). Multiply every entry by that number: $3\begin{pmatrix} 1 & 0 \\ -1 & 7 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ -3 & 21 \end{pmatrix}.$

Transpose: Flip the matrix over this diagonal:

$$\begin{pmatrix} \ddots \end{pmatrix}$$

Ex:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}^T = \begin{pmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \\ 4 & 8 & 12 \end{pmatrix}$$

Ex: The transpose of an $m \times n$ matrix is $\underline{n \times m}$

Formulas: $(A+B)_{ij} = A_{ij} + B_{ij}$ $\quad (A-B)_{ij} = A_{ij} - B_{ij}$

$(cA)_{ij} = c(A_{ij})$ $\quad (A^T)_{ij} = A_{ji}$.

Example: Show that $A+B=B+A$ for any two matrices A & B, and that one side exists only when the other does.

Solution: * $A+B$ and $B+A$ each exist only when A & B have the same size.

* If A and B are both $m \times n$, then so are $A+B$ and $B+A$.

* $(A+B)_{ij} = A_{ij} + B_{ij} = B_{ij} + A_{ij} = (B+A)_{ij}$.

$A+B$ and $B+A$ have the same size, and their corresponding entries match, so $A+B=B+A$. ✓

III: Matrix multiplication

If A is $m \times n_1$, and B is $n_2 \times p$,
AB exists only when $n_1 = n_2$, and
in that case AB is $m \times p$.

$$\underset{m \times n_1}{A} \qquad \underset{n_2 \times p}{B}$$

have to be equal

dimensions of $A \cdot B$

To find the i,j entry of AB, multiply the entries in the ith row of A and the jth column of B, then add those products together.

<u>Example</u>:
$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} p & q \\ r & s \\ t & u \end{pmatrix} = \begin{pmatrix} ap+br+ct & aq+bs+cu \\ dp+er+ft & dq+es+fu \end{pmatrix}$$

<u>Example</u>: If $A = \begin{pmatrix} 1 & -1 & 4 \\ 2 & 3 & 5 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ what are AB and BA?

<u>Answer</u>: AB $\boxed{\text{does not exist!}}$

$$BA = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 4 \\ 2 & 3 & 5 \end{pmatrix}$$
$$= \begin{pmatrix} 2-2 & -2-3 & 8-5 \\ -1+2 & 1+3 & -4+5 \end{pmatrix}$$
$$= \boxed{\begin{pmatrix} 0 & -5 & 3 \\ 1 & 4 & 1 \end{pmatrix}}$$

Example: If A is 5×3 and AB is 5×7, what size is B?

Solution:

$$\underbrace{5\times \overbrace{3 \quad m}^{\text{have to be equal}} \times n}_{\text{Needs to be } 5 \,\&\, 7}$$

So B is $\boxed{3\times 7}$.

Matrix multiplication distributes, and is associative:

$$(A+B)\cdot C = A\cdot C + B\cdot C$$
$$C\cdot(A+B) = C\cdot A + C\cdot B$$
$$A(BC) = (AB)\cdot C$$

Matrix multiplication doesn't always commute:

AB isn't always the same as BA even if both exist and are the same size.

Example: Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$

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

$$BA = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0+0 & 0+0 \\ 0+0 & 1+0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Because of this, algebra with matrices is basically the same as algebra w/ numbers, but you need to make sure you don't do $AB = BA$.

Example: True/False: For any $n \geq 1$, and any two $n \times n$ matrices A and B,

$$A^2 - B^2 = (A+B)(A-B).$$

Solution 1: Figure out what both sides are.

But I don't want to do that. @_@

Solution 2:

$$(A+B)(A-B) = A(A-B) + B(A-B)$$
$$= A^2 - AB + BA - B^2.$$

This isn't the same as $A^2 - B^2$, unless $AB = BA$.

So it's $\boxed{\text{False}}$.

Example: Find a 2×2 matrix B with $B^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

that is different from $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

[Different from numbers, since if x is a number, $x^2 = 0$ means $x = 0$.]

Answer: $\boxed{B = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}}$?

$$\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 1-1 & 1-1 \\ -1+1 & -1+1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \checkmark$$

There are infinitely many B's that work:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$c(a+d) = b(a+d) = 0$$

Either $b=0$ and $c=0$, but that forces $a=d=0$. X

Or $d=-a$. Then $a^2+bc = bc+d^2 = 0$

means $c = -\dfrac{a^2}{b}$, if $b \neq 0$

~~$a = 0$~~ if $b=0$.

$$\boxed{\begin{pmatrix} a & b \\ -\frac{a^2}{b} & -a \end{pmatrix} \text{ for } b\neq 0, \text{ any } a}$$

$$\boxed{\begin{pmatrix} 0 & 0 \\ c & 0 \end{pmatrix} \text{ for any } c.}$$

Lesson: Be careful when trying to divide matrices:

$$AB = \text{matrix of } 0\text{'s}$$

doesn't mean A or B is a matrix of 0's.

What is matrix multiplication good for?

Look at this system of equations:

$$\begin{aligned} x + 2y + z &= 4 \\ -2x + 5y \quad\; &= -1 \\ 3x - 4y + 2z &= 3 \end{aligned}$$

Rewrite like this:

$$\underbrace{\begin{pmatrix} 1 & 2 & 1 \\ -2 & 5 & 0 \\ 3 & -4 & 2 \end{pmatrix}}_{\text{Call this } A} \underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{\vec{x}} = \underbrace{\begin{pmatrix} 4 \\ -1 \\ 3 \end{pmatrix}}_{\text{Call this } \vec{b}}$$

If I could "divide" by A, I could divide

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

through by A and get

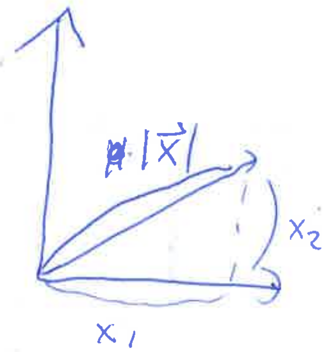
$$\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1}\vec{b}.$$

The matrices you can divide by are called "invertible." See Thursday's lecture.

IV: Lengths and dot products.

* If $\vec{x} = (x_1, \dots, x_n)$, then its <u>length</u> is $\sqrt{(x_1)^2 + \dots + (x_n)^2}$.

  Eg. if $n=2$,



  length is written as "$|\vec{x}|$".

* If $\vec{x} = (x_1, \dots, x_n)$, $\vec{y} = (y_1, \dots, y_n)$ then their dot product is:

$$\vec{x} \cdot \vec{y} = x_1 y_1 + \dots + x_n y_n.$$

<u>Example:</u> If $\vec{x} = (1, -2, 1)$ $\vec{y} = (0, 3, -4)$ then:

$$|\vec{x}| = \sqrt{1^2 + (-2)^2 + 1^2} = \boxed{\sqrt{6}}$$

$$|\vec{y}| = \sqrt{0^2 + 3^2 + (-4)^2} = \sqrt{25} = \boxed{5}$$

$$\vec{x} \cdot \vec{y} = 1 \cdot 0 + (-2)3 + 1(-4) = \boxed{-10}$$

<u>Notice</u>: $|\vec{x} \cdot \vec{y}| = 10 \leq |\vec{x}| \cdot |\vec{y}| = 5\sqrt{6}$

<u>That's always true</u>: Cauchy–Schwarz says this always works,

$$\underbrace{|\vec{x} \cdot \vec{y}|}_{\text{absolute value}} \leq \underbrace{|\vec{x}| \cdot |\vec{y}|}_{\text{lengths}}.$$

Notes use this to get an upper bound for errors, by writing the error as $\vec{x} \cdot \vec{y}$.