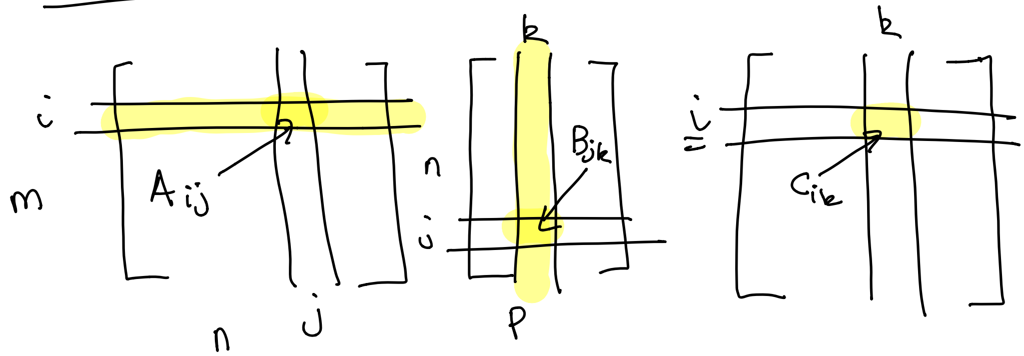
Lecture #4: Matrix Multiplication & Inverses

Today we will talk about how to multiply matrices and what it means

There are a few ways to think about it

View #1: As a formula



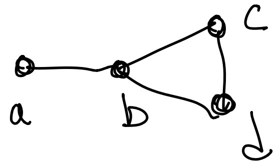
$m \times n$ matrix A $\quad$ $n \times p$ matrix B $\quad$ $m \times p$ matrix C

$$C_{ik} = \sum_{j=1}^{n} A_{ij} B_{jk}$$

It's hard to see why this definition is natural, but an example helps

Application: Counting Walks

We'll be working with graphs — has vertices and edges, e.g.



This is a natural abstraction for doing things like describing who is friends with who on facebook (vertices = people, edges = friendships)

def: A walk is a sequence of vertices, connected by edges, where repetitions are o.k.

e.g. a – b – a – b – d is a walk of length 4
\# edges

Goal: Count the number of walks of a given length

It turns out you can do it through matrix multiplication

Q: But how do we represent a graph as a matrix?

def: An adjacency matrix is a matrix where each row/column represents a vertex, and there is a 1 in row i, column j iff there is an edge between the corresponding vertices e.g.

$$A = \begin{array}{c} \\ a \\ b \\ c \\ d \\ \\ \end{array}\!\!\begin{array}{c} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \\ \begin{array}{cccc} a & b & c & d \end{array} \end{array}$$

what happens if we multiply A by itself?

$$\begin{array}{c} \\ a \\ \\ \\ \\ \end{array}\!\!\begin{array}{c} \begin{array}{cccc} a & b & & \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \end{array} \begin{array}{c} \begin{array}{cccc} & & c & \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \end{array} = \begin{array}{c} \\ a \\ \\ \\ \\ \end{array}\!\!\begin{array}{c} \begin{array}{cccc} & & c & \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 3 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix} \\ A^2 \end{array}$$

$$0\cdot 0 + 1\cdot 1 + 0\cdot 0 + 0\cdot 1 = 1$$

$$a\cancel{-}a-c \quad a-b-c \quad a-c\cancel{-}c \quad a\cancel{-}d-c$$

Each term is a walk of length two

And the adjacency matrix tells us if edges are there

<u>Q2</u>: What does $(A^2)_{a,b}=0$ tell us?

There are no walks of length two that start at a and end at b

More generally $\underbrace{A \times A \times \dots A}_{\ell \text{ times}}$ counts length ℓ walks

row = starting vertex

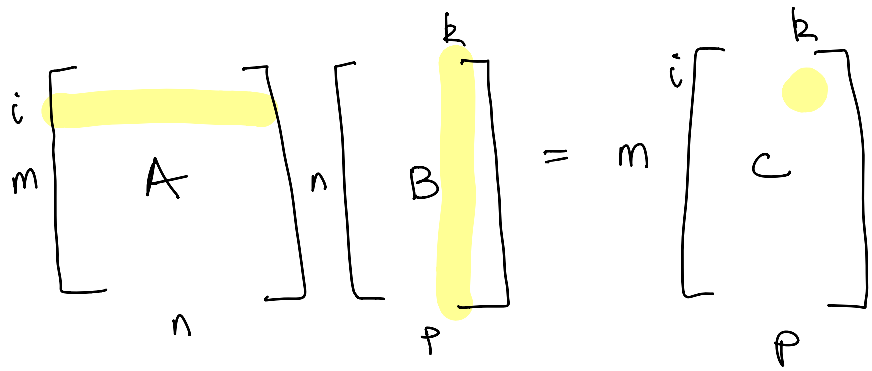
column = ending vertex

Another way to think about matrix multiplication:

<u>def:</u> The inner-product of two vectors x and y with the same dimension n

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

<u>View #2:</u> In terms of inner products



C_{ik} = the inner product btwn i^{th} row in A and k^{th} column in B

Let's do another application

<u>Application:</u> Linear Dynamical Systems

In science and engineering we often want to understand how a (linear) system evolves

e.g. $x^{t+1} = A x^{t}$

nxn matrix representing how the system updates

length n vector representing the state of the system at time t e.g. the Tacoma bridge example from earlier

Today: the predator-prey model

$g(t) = \#$ frogs at time t
$y(t) = \#$ flies at time t

$$g(t+1) = 0.4\, g(t) + 0.2\, y(t)$$
$$y(t+1) = -0.6\, g(t) + 1.8\, y(t)$$

Q3: Can we model this using matrices?

$$\begin{bmatrix} g(t+1) \\ y(t+1) \end{bmatrix} = \begin{bmatrix} 0.4 & 0.2 \\ -0.6 & 1.8 \end{bmatrix} \begin{bmatrix} g(t) \\ y(t) \end{bmatrix}$$

Q4: If we start at $g(0), y(0)$ what does the population look like at time n?

Fact: matrix multiplication is associative

$$\underset{m \times n}{A}\,(\underset{n \times p}{B}\,\underset{p \times l}{C}) = (AB)C$$

Note: dimensions need to match, but it's the same constraint on both sides

what this means for us is we can rewrite

$$\underbrace{A \times (A \times (A \times \ldots \times (A \times}_{n \text{ times}} \begin{bmatrix} g(0) \\ y(0) \end{bmatrix}) \ldots))$$

instead as $\underbrace{(..(((A)\times A)\times A)\dots \times A)}_{n \text{ times}} \begin{bmatrix} g(0) \\ y(0) \end{bmatrix}$

we'll call this A^n

<u>Takeaway</u>: So if we want to simulate the model for many different initial conditions we can just compute A^n once

some other key facts

<u>Fact</u>: matrix multiplication is distributive

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

<u>Q5</u>: why did I write this rule out two different ways?

<u>Anti-Fact</u>: Matrix multiplication is generally not commutative $\quad AB \neq BA$

In fact, the dimensions do not necessarily make sense

$$(m\times n)(n\times p) \text{ vs. } (n\times p)(m\times n)$$

Also there is a "unit" element
acts like 1

<u>def:</u> The identity matrix is unit

$$
\begin{array}{cc} & n \\ n & \begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & \ddots 1 \end{bmatrix} \end{array} = I_n
$$

<u>Fact:</u> $\underset{n\times n}{I}\,\underset{n\times m}{A} = \underset{n\times m}{A}$; $\underset{m\times m}{AI} = A$

Now we are ready for a GREAT idea:

> when we want to solve the linear system $Ax=b$, wouldn't it be easier if there was a matrix A^{-1} so that $A^{-1}A=I$?

First, why would this help?

$$Ax=b \Rightarrow A^{-1}(Ax) = A^{-1}b$$

$$\Rightarrow \underbrace{(A^{-1}A)}_{I}x = A^{-1}b \Rightarrow x = A^{-1}b$$

<u>Q6:</u> what rule did I use here?

Even better: I claim we already know how to find A^{-1}

Recall in Gauss-Jordan elimination the first step when solving

$$\begin{array}{c} r_1 \\ r_2 \\ r_3 \end{array} \begin{bmatrix} 1 & -1 & 2 \\ -2 & 2 & -3 \\ -3 & -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}$$

was to add $2r_1$ to r_2

Q7: How can I describe this operation using matrix multiplication?

1x3 vector

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} r_1 \\ 2r_1 + r_2 \\ r_3 \end{bmatrix}$$

In Gauss-Jordan elimination, sometimes we need to swap rows, which can be done by

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_3 \\ r_2 \end{bmatrix}$$

Each step in Gauss-Jordan elimination can be implemented as a matrix mult.

$$A \to B_1 A \to B_2 B_1 A \rightsquigarrow \underbrace{(B_p \cdots B_1)}_{A^{-1}} A = I$$

But we don't always get so lucky

Sometimes we get a row of zeros, e.g.

$$\overset{A}{\begin{bmatrix} 2 & -3 & 0 \\ 1 & 1 & 2 \\ 3 & -2 & 2 \end{bmatrix}} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \overset{b}{\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}}$$

$$\rightsquigarrow \begin{bmatrix} 2 & -3 & 0 \\ 0 & \frac{5}{2} & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{1}{2} \\ 1 \end{bmatrix}$$

Poll

Q8: What does this tell us about A^{-1}?

(a) It doesn't exist

(b) For this choice of b it doesn't exist, but for others it might

(c) You can't trust matrices

Note: we introduced A^{-1} as the matrix s.t.

$$A^{-1}A = I \qquad \text{(left inverse)}$$

If A is square, there could also be a matrix B s.t.

$$AB = I \qquad \text{(right inverse)}$$

Q9: Could these be different?

__Fact:__ If A is square then its right and left inverse are the same.

__Q10:__ Can you come up with a non-square A that has a left but no right inverse?

__One last view:__ As matrix vector products

Can also think about

$$m\underset{n}{\begin{bmatrix} & A & \end{bmatrix}} \; n\underset{p}{\begin{bmatrix} & B & \end{bmatrix}}$$

as p matrix vector products

Let $B_1, B_2 \ldots B_p$ be the columns of B

Then $AB = [AB_1, AB_2, \ldots AB_p]$