Welcome to 6.S084 | 18.S096

"Linear Algebra and Optimization"

Instructors : Ankur Moitra
moitra@mit.edu
Pablo Parrilo
parrilo@mit.edu

Course Goals : essentials of linear algebra and optimization in an integrated way

Geometry plays a key role, e.g.

(1) the problem formulation

Q How do you visualize a system of linear equations?

② the abstractions

Q How do you represent the set of all solutions?

③ understanding the behavior of operators and algorims

e.g many iterative algorithms, like PageRank, can be understood thru linear algebra

We hope this course will help you learn how to use linear algebra and optimization in w/e you're interested in, e.g.

Q: often, the problem you want to solve comes from data, when is the answer is stable?

Q: How can you change the problem to make it stable?

You'll explore these issues through experimentation in Julia

Some Logistics

Prereqs: 18.02

Antiprereqs: 18.06, 18.065 or 18.700

Grading: see syllabus for more detailed information

Homework: 50%

Miniprojects + Recitations: 10%

Take home exams: 15% + 15%

Final exam: 20%

Attendance not required for lectures, but strongly encouraged

Videos will be made available after

Recitations: Attendance required, three unexcused absences allowed, need approval from S^3 otherwise

Today: some examples of things we'll learn about via linear algebra

def: A vector is just a tuple of #s

e.g. $\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ or $\begin{bmatrix} i \\ 2i+3 \\ 4 \end{bmatrix}$

Could be over reals / complexes
(won't consider other fields in this class)

the dimenscon is just the size of the tuple

The important thing is we can manipulate vectors

1. multiply by a scalar

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

2. add them

$$\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 8 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 7 \\ 5 \end{bmatrix}$$

How do we typically use linear algebra?

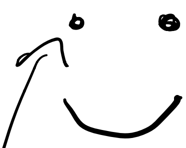
1. represent data as a vector

2. do something to the vector to change into something new

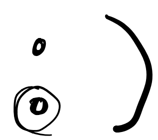
3. do something to a collecton of vector, to extract some structure

Ex: We have a black and white image where each black point we associate with a vector

e.g.



Q: How do we describe what to do to each vector so that it <u>rotates</u>?



Alternatively this point has an x/y coordinate; what should its new x'/y' coordinate be?

We will mostly be interested in **linear functions** (rotation is a special case)

def: A matrix is a grid of numbers

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

Now we can take a matrix-vector product:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$$

which gives us another vector

Note Above, the matrix is 2×2 and the vector is 2×1, the inner-dimensions need to match

Another way to interpret it is as forming a linear combination of the columns of the matrix

e.g. $x \begin{bmatrix} a \\ c \end{bmatrix} + y \begin{bmatrix} b \\ d \end{bmatrix}$

Q: How do you multiply a 3x3 matrix with a 3-dimensional vector?

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax + by + cz \\ dx + ey + fz \\ gz + hy + iz \end{bmatrix}$$

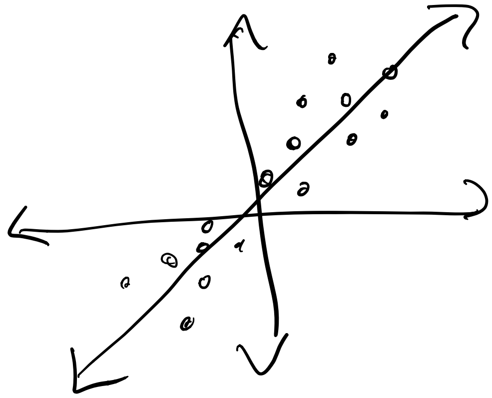
$$= x \begin{bmatrix} a \\ d \\ g \end{bmatrix} + y \begin{bmatrix} b \\ e \\ h \end{bmatrix} + z \begin{bmatrix} c \\ f \\ i \end{bmatrix}$$

Take away: matrices define ways to map a vector into another vector

But you can do so much more with matrices!

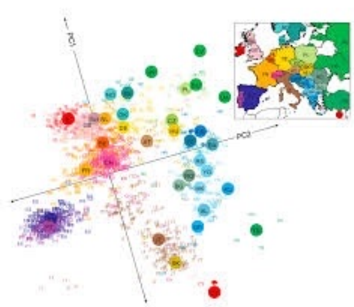
Instead of designing one by hand, how about finding one automatically that does something nice for your data?

Ex: we have a bunch of data points (each data point is a vector)



Q: what one line captures the data the best?

Especially important in high dimension



Sometimes matrices represent complex systems that you'd like to analyze

e.g. Tacoma bridge

Here you have a system that is updating over time, and your state is a <u>vector</u> and a <u>matrix</u> describes how you get to the next state

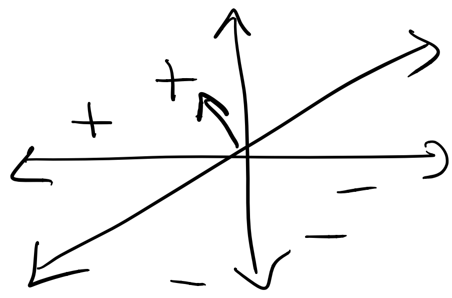
Increasingly linear algebra plays a key role in decision making

e.g. we want to classify emails as spam or not

Doesn't sound like linear algebra

$\boxed{\text{email}} \rightarrow \begin{bmatrix}\text{vector of} \\ \text{word counts}\end{bmatrix}$

sometimes the classification rule is a <u>linear</u> <u>separator</u>



All the positive examples (spam) are above the line, all negative examples (ham) are below

Q: How do we represent the decision rule?

① take the vector normal to the line

e.g. $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

② take the inner-product between it and the example

inner-product of $\begin{bmatrix} a \\ b \end{bmatrix}$ and $\begin{bmatrix} c \\ d \end{bmatrix}$

is $\underset{1 \times 1}{ac + bd}$ or $\underset{1 \times 2}{[a \ b]} \underset{2 \times 1}{\begin{bmatrix} c \\ d \end{bmatrix}}$

If the inner product is positive
output "+"

Else output "−"

Even for state-of-the-art classifiers (e.g. deep nets) that have <u>very</u> complex decision rules, you use vectors, matrices and carefully chosen <u>nonlinearities</u>