Part 1: Data Models

Felix Muzny

↳ NLP → modeling human lang. w/ math

sometimes I'll add notes after lecture - normally in pink

Playing a lottery

10% chance - given a puppy

100 people

$P_{puppy} = \frac{1}{10} = .1$

$P_{no\ puppy} = 1 - P_{puppy} = .9$

$.9 * .9 * .9 \ldots\ldots$

$P_{\substack{no\ puppies \\ in\ room}} = (P_{no\ puppy})^{100} = .000026$

$= .0026\%$

Assumptions made?

How many people do we need in class

$P_{\text{puppy in the class}} \geq 50\%$?

- puppy distr. is independent
- puppy distr. is uniform

Discovering a secret grading scheme

grade = homework + exams

student 1: $80x + 72y = 78$

2: $60x + 100y = 70$

$x + y = 1$

what are the values of x and y?

Assumptions made?

- Prof. Felix will the same grading scheme
- weights should sum to one

What else could we model?

ICA Question 1: based on these examples, what are two things that you'd like to model/learn how to model in this class?

1) your answer here

Math of Data Models

- Given that the real world is so complex... what are we aiming to give you in this course?
    - A breadth of mathematical models to choose from
    - The ability to be creative & rigorous in making and evaluating assumptions
    - A sense of which aspects of the application we're trying to model most accurately ... and how to do so

Part 2: Linearity & functions

Functions

- In programming-land: A **function** is a conveniently packaged set of instructions (lines of code) that together accomplish some common operation.

- In math-land: A **function** is a formal~~ly~~ mapping of *inputs* to *outputs*

numbers

Linearity

- A function is **linear** if (informal definition):
  - scaling, applied before or after the function, has an equivalent effect
  - addition, applied before or after the function, has an equivalent effect

$\alpha f(x)$

$f(x) + f(y)$

scalar – α

↳ $\# \in \mathbb{R}$

↳ a real number

(there is no slide 10)

Linearity

- A function is **linear** if:
  - scaling, applied before or after the function, has an equivalent effect

choose α

$$f(\alpha x) = \alpha f(x)$$

$f(x) = 2x$

$\alpha = 3$

$f(3*2) = f(6) = 12$

$3*f(2) = 3*4 = 12$

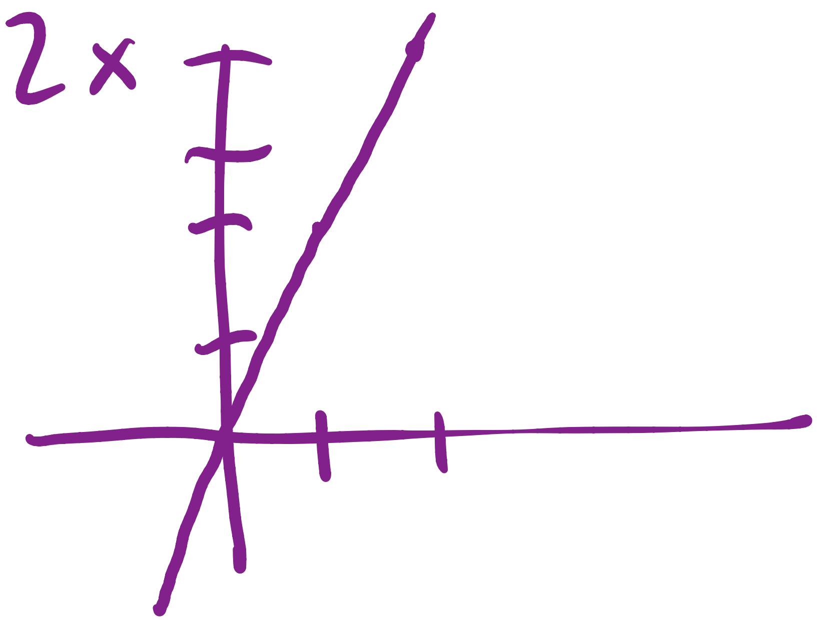
Linearity

- A function is **linear** if:
  - addition, applied before or after the function, has an equivalent effect

$x, y \in$ domain of f

$$f(x+y) = f(x) + f(y)$$

$f(x) = 2x$



addition:

$f(1+4) = f(5) = 10$

$f(1) + f(4) = 2 + 8 = 10$

Linearity

- A function is **linear** if (formal definition, & what we use in practice):
  - $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$
  - for any $\alpha, \beta \in \mathbb{R}$ and $x, y \in$ domain of f

real numbers

$f(x) = 2x$

domain is $\mathbb{R}$

Linearity

- How to show a function is linear:
- $f(x) = 10x$
  - choose $\alpha, \beta \in \mathbb{R}$
  - $(x, y \in \mathbb{R}$ for this equation)

$$f(\alpha x + \beta y) = 10(\alpha x + \beta y)$$

$$= \alpha 10x + \beta 10y$$

$$= \alpha f(x) + \beta f(y)$$

yes!

Linearity

- How to show a function is **not** linear:
- $f(x) = x^2$

guess, get a counter example

  - choose $\alpha, \beta \in \mathbb{R}$

$\alpha = \beta = 1 \quad x = y = 1$

  - choose $x, y \in \mathbb{R}$

- if proving **non**-linearity, we expect $f(\alpha x + \beta y) \neq \alpha f(x) + \beta f(y)$

$$f(\alpha x + \beta y) = f(1 \cdot 1 + 1 * 1) = f(2) \to 2^2 = 4$$

$$= 1 * f(1) + 1 * f(1) = 1 * 1^2 + 1 * 1^2 = 2$$

No!

Linearity

- Wait, why do we care?
    - Remember: we're trying to build mathematical models of the world
        - many real-world things are linear
        - many are not linear, but can be re-cast as linear!

Linearity

- Wait, why do we care?
  - we can find all solutions to these equalities (next)
  - we can find values that are "closest" (lines of best fit)
  - **<u>all</u>** outputs are defined by the systems behavior on any set of basis inputs

$f(x) = 2x$ $f(3) = 6$

$f(1) = 2$ $f(17) = 34$

$f(2) = 4$

  - matrix multiplications!

↳ GPUs are <u>really</u> good + <u>really</u> fast

Solving systems of linear equations

Hefferon, 1.1

- A **linear system** is a set of **linear equalities**
  - Solutions to a linear system must satisfy all equalities
  - $x + y = 0$
  - $2x - y + 3z = 3$
  - $x - 2y - z = 3$

$\underbrace{x = 2, \; y = -2, \; z = 30}$

$2 + -2 = 0$ ✓

$2(2) - (-2) + 3(30) = 3$

$\underbrace{4 + 2 + 90 = 3}$ ✗

Solving systems of linear equations

Hefferon, 1.1

- A **linear system** is a set of **linear equalities**
  - Solutions to a linear system must satisfy all equalities
  - $x + y = 0$
  - $2x - y + 3z = 3$
  - $x - 2y - z = 3$

ICA Question 2: *pause* this video, then spend no more than 5 minutes attempting to solve this system of equations. Write down your work as you go!

Solving systems of linear equations

- To think about: how might you teach a computer to solve every possible linear system?

Gauss's Method

- Transform the system to a system with the same set of solutions (but whose solution is more obvious)
  - $x + y = 0$
  - $2x - y + 3z = 3$
  - $x - 2y - z = 3$
- swap two equations
- multiply both sides of an equation by a non-zero constant
- replace an equation with the sum of itself and the multiple of another

$x = ?$
$y = ?$

Gauss's Method

- Transform the system to a system with the same set of solutions (but whose solution is more obvious)

r_0
- $x + y = 0$

r_1
- $2x - y + 3z = 3$

r_2
- $x - 2y - z = 3$

- swap two rows
- scale a row
- sum two rows

swap

$2x - y + 3z = 3$
$x + y = 0$
$x - 2y - z = 3$

$r_0' = r_1$
$r_1' = r_0$

scale

$x + y = 0$
$4x - 2y + 6z = 6$
$\shortparallel$
$r_1' = 2r_1$

sum

$2x - y - z = 3$

$r_0' = r_0 + r_2$

Gauss's Method

- $x + y = 0$

✓ $1 - 1 = 0$

- $2x - y + 3z = 3$

✓ $2 + 1 + 0 = 3$

- $x - 2y - z = 3$

✓ $1 + 2 - 0 = 3$

$\xrightarrow[r_2' = r_2 - r_0]{r_1' = r_1 - 2r_0}$

$x + y = 0$

$-3y + 3z = 3$

$-3y - z = 3$

$\xrightarrow{r_1' = -\frac{1}{3} r_1}$

$y - z = -1$

$-3y - z = 3$

$\xrightarrow{r_2' = r_2 + 3r_1}$

$x + y = 0$

$y - z = -1$

$-4z = 0$

$\xrightarrow[r_2' = -\frac{1}{4} r_2]{r_0' = r_0 - r_1}$

$x + z = 1$

$y - z = -1$

$z = 0$

$\xrightarrow[r_1' = r_1 + r_2]{r_0' = r_0 - r_2}$

$x = 1$

$y = -1$

$z = 0$

Gauss's Method - generalized

- Always keep the order of variables the same in equations
- $x + y = 0$
- $2x - y + 3z = 3$
- $x - 2y - z = 3$

x, y, z

$\cancel{y + x = 0}$

$\cancel{3z + 2x - y = 3}$

Gauss's Method - generalized

- For n = 0 to n = number of equations - 1:
  - scale the leading coefficient of eq'n N to 1
  - add (the correct multiple) of eq'n N to others

(0-indexing)

- $x + y = 0$
- $2x - y + 3z = 3$
- $x - 2y - z = 3$

a 1 in one eq'n for each variable, a 0 for that variable in the other eq'ns

Matrices

- $x + y = 0$
- $2x - y + 3z = 3$
- $x - 2y - z = 3$
- As a matrix:

$$
\begin{array}{cccc}
x & y & z & \text{answer} \\
\end{array}
$$
$$
\left[\begin{array}{ccc|c}
1 & 1 & 0 & 0 \\
2 & -1 & 3 & 3 \\
1 & -2 & -1 & 3
\end{array}\right]
$$

"augmented" matrix

Reduced Row Echelon Form

Hefferon, 1.3

- (this is what we are aiming for as we row-reduce our matrices)

- the **leading coefficient** is the first non-zero value in a row

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 7 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -10 \end{array}\right] \qquad \left[\begin{array}{cc|c} 1 & 0 & \ast \\ 0 & 1 & \ast \\ 0 & 0 & \ast \end{array}\right] \qquad \left[\begin{array}{ccc|c} 1 & 0 & \ast & \ast \\ 0 & 1 & \ast & \ast \\ 0 & 0 & 0 & \ast \end{array}\right]$$

Reduced Row Echelon Form

- (this is what we are aiming for as we row-reduce our matrices)
- the **leading coefficient** is the first non-zero value in a row
- For a Reduced Row Echelon Form (**RREF**) matrix**:**
  - Leading coefficient = 1 in row N in position N (or does not exist)
  - Zeroes above & below leading coefficient -> each leading coefficient is the only non-zero entry in its column

Part 4: Admin

→ course website

where is the syllabus, how are you graded, and how will this course work, what is your schedule

- Homework (42%)
- Quiz-tests (44%)
- In-Class Activities (6%)
  - Graded on completion/effort (hard deadline of 11:59 pm the evening of lecture*) → T/F
  - * except on asynchronous lecture days, when they'll be due before the next lecture (11:45am Thursday)
- Mini-projects (8%)

lecture

- We're remote until Feb 5th
- Lectures are not recorded*
- We'll be meeting on Zoom at 11:45am on M/R -> see you all synchronously on Thursday! ↳ EST

Remote lectures: expectations

- Remote learning can be weird! We'll be doing our best to reduce weirdness.
- Here are my expectations of you all:
  - Be in a **location** conducive to learning
  - Set your zoom profile picture to a picture **of yourself**
  - When we are in breakout rooms, **turn on your cameras**
  - When we are in breakout rooms, each group will pick one person to **screenshare**
  - Use the chat or "raise hand" features to ask me questions!
  - Wear a fun hat
  - Pets are absolutely welcome
  - Tell me about your music preferences every week

ICA questions: the fun bits

ICA Question 3: what is your current preferred *genre* of music?

ICA Question 4: do you have a current favorite artist?

ICA questions: wrap-up

ICA Question 5: after watching all the videos for lecture today, are there any questions/clarifications that you want me to cover during our next lecture?

Schedule

OH for TAs : start week of Jan 24th

Complete ICA 1 before class on Thursday -> find this on Gradescope!

We are remote until Feb 5th

Mon	Tue	Wed	Thu	Fri	Sat	Sun
January 17th **MLK Day**	**Felix OH Calendly**	**Felix OH Calendly/drop-in**	Lecture 2 - Vector Algebra			
January 24th Lecture 3 - Matrices & vector geometry **HW 1 released**	**Felix OH Calendly**	**Felix OH Calendly/drop-in**	Lecture 4 - ML, linear perceptron			