Math Foundations for ML

10-606

Geoff Gordon

Course page

- https://www.cs.cmu.edu/~ggordon/10606s22/syllabus-and-lecture-outline.html
- Anyone with an Andrew ID can access all course materials starting from there
  - First step: click on signup link for Piazza

2nd step: sign up for replit.com

Returning to in-person

- As a reminder, CMU is returning to in-person lectures
  - first in-person lecture 1/31 (Mon), GHC 4101 ***(note: updated location!)***
  - distancing: occupy every 2nd seat, informal auditors and waitlist welcome if there is space (please give precedence to those who are registered)
  - we'll try to record and broadcast over Zoom, at least for now; technical difficulties are possible
- Bring a device!
  - we'll still be doing synchronous in-class exercises
  - physical keyboard and big-enough screen recommended

$\mathbb{R}$

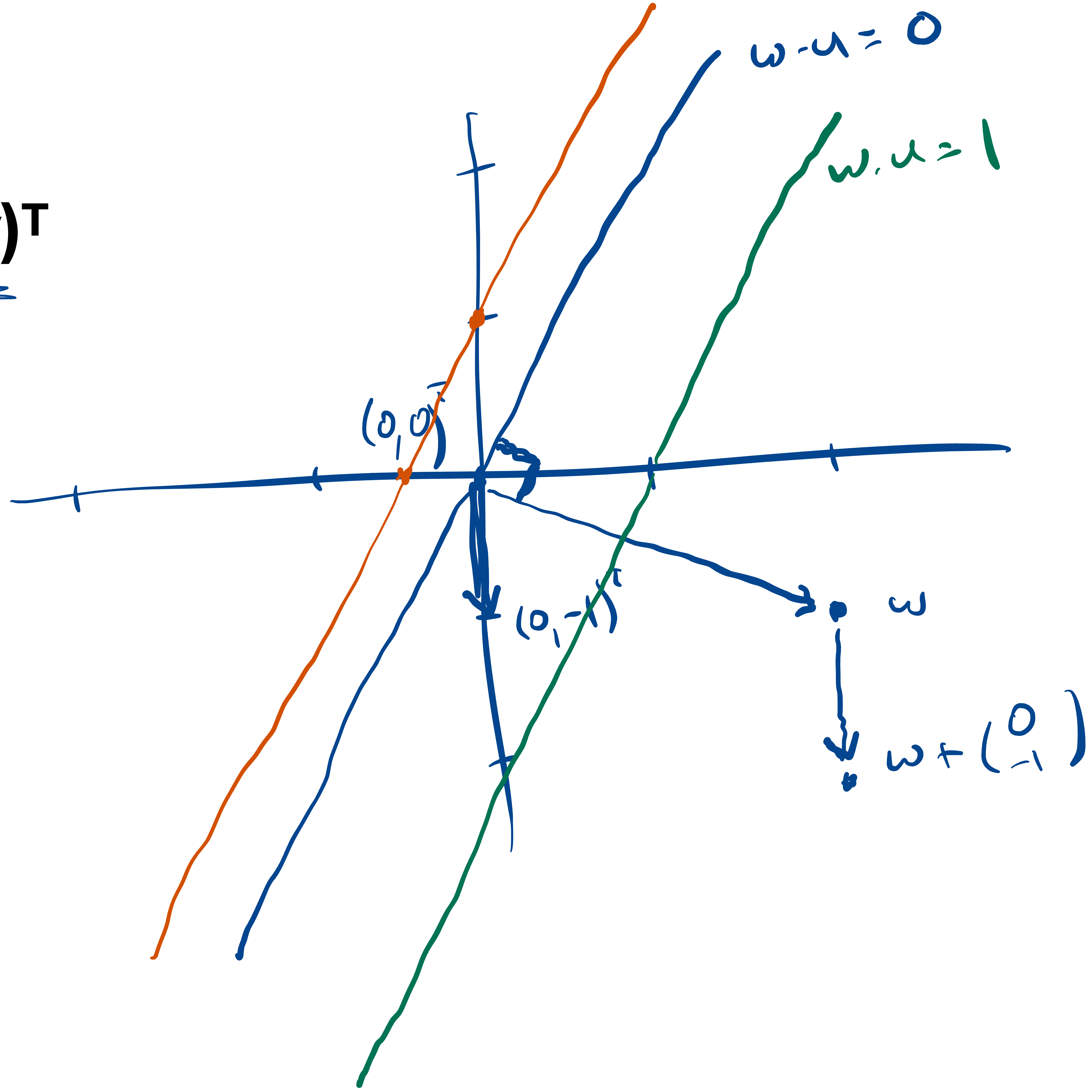
$ax + y \in \mathbb{R}^1$

$\in \mathbb{R}$

R^2 example

Let w = $(2, -1)^T$ and u = $(x, y)^T$

- Draw w
- Draw the line w · u = 0
- Draw the line w · u = 2
- Draw the line w · u = –1



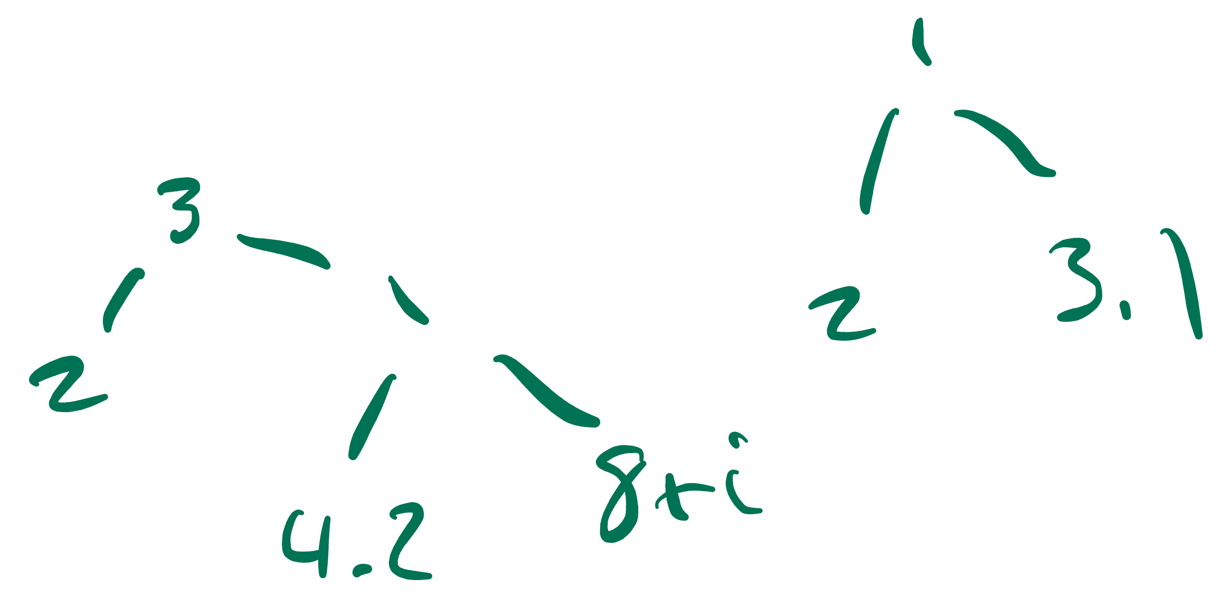
Exercise: vector spaces

Which of the following are vector spaces?

- ✓ A: polynomials of degree 7 in real variables x, y
- ✗ B: arrays of four real numbers in sorted order
- ✗ C: strings of alphanumeric characters: [A-Za-z0-9]*
- ✓ D: complex numbers
- ✗ ~~E~~: binary trees with a complex number at each node → H
  ↳ w/ a little work, ✓
- ✓ F: square 7x7 matrices
- ✓ G: square-summable sequences of real numbers

$\begin{pmatrix}1\\3\\4\\7\end{pmatrix}$ not $\begin{pmatrix}1\\4\\7\\2\end{pmatrix}$

h3LLØ → not w/o lots of work



$s_1, s_2, s_3, \ldots \in \mathbb{R}$

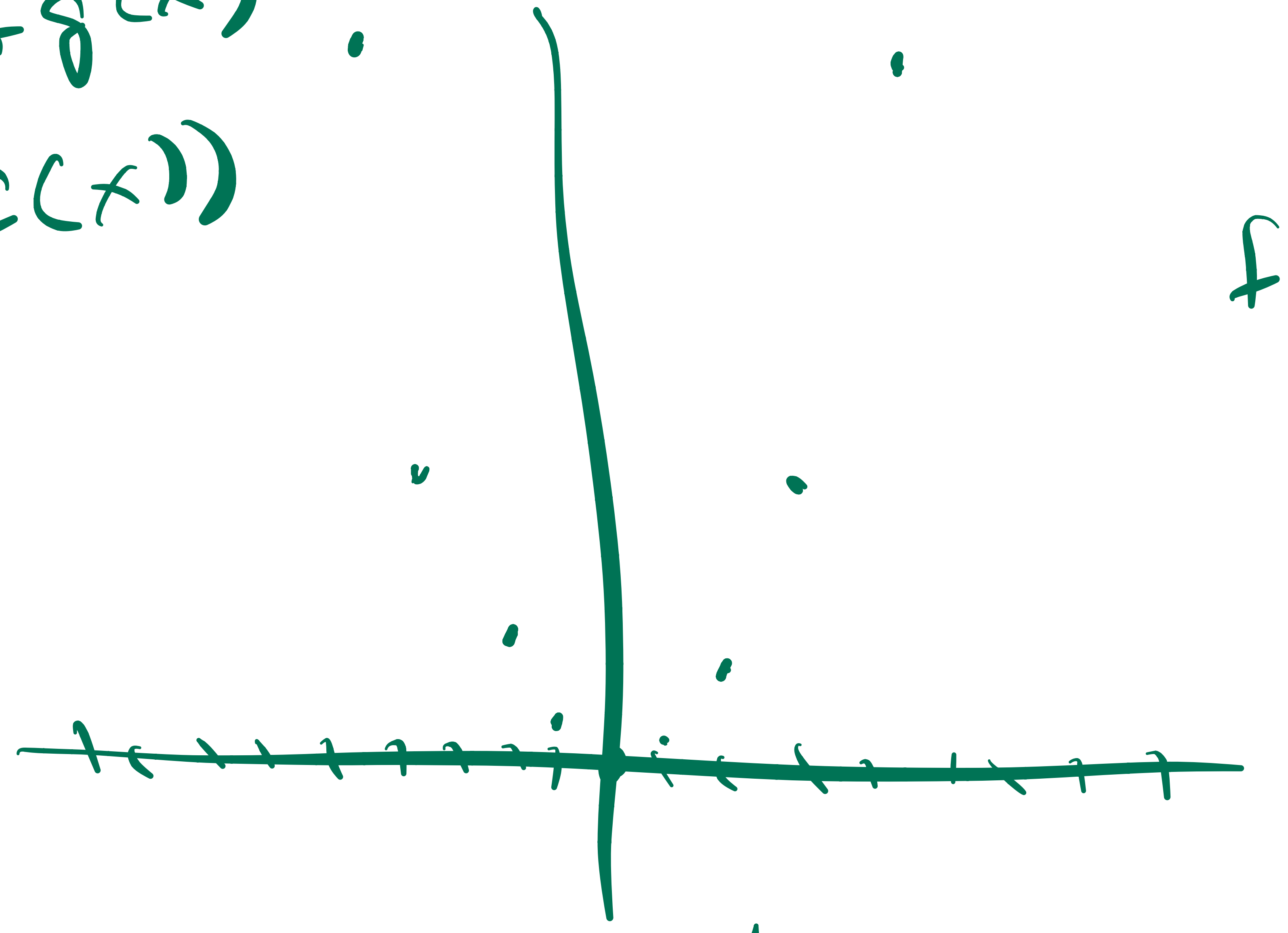
$\sum_{t=1}^{\infty} s_t^2 < \infty$

$V = \mathbb{R} \to \mathbb{R}$

$\forall x \quad (f+g)(x) = f(x) + g(x)$

$\forall x \quad (3f)(x) = 3(f(x))$

$f(x) = x^2$



$-3, -2, -1 \ldots 3$

7 samples

$$\begin{pmatrix} 9 \\ 4 \\ 1 \\ 0 \\ 1 \\ 4 \\ 9 \end{pmatrix} \in \mathbb{R}^7$$

$$x \cdot y = \sum_{i=1}^{n} x_i y_i \qquad x, y \in \mathbb{R}^n$$

$$\|x\| = \sqrt{\langle x, x \rangle}$$

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

$$\mathbb{R}^{3\times 3} \longleftrightarrow \mathbb{R}^9$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{pmatrix}$$

$$X, Y \in \mathbb{R}^{3\times 3}$$

$$\langle X, Y \rangle = \sum_{i=1}^{3} \sum_{j=1}^{3} X_{ij} Y_{ij}$$

$$\langle x, y \rangle = (Ax) \cdot (Ay)$$

fixed $A \in \mathbb{R}^{n \times n}$

V = polynomials $\mathbb{R}$ coeff x

$\mathbb{R}[x]$

$$\langle x^d, x^{d'} \rangle = \begin{cases} 1 & \text{if } d = d' \\ 0 & \text{o/w} \end{cases}$$

$$\langle x^3+1, x^3 \rangle = \underbrace{\langle x^3, x^3 \rangle}_{1} + \underbrace{\langle 1, x^3 \rangle}_{0} = 1$$

$$1, \quad 1 + \frac{x}{2}, \quad 1 + \frac{x}{2} + \frac{x^2}{4}, \quad \dots$$

$V \to \mathbb{R}$ functional

$V \to V$ operator

↳ e.g. $\frac{d}{dx} \in \underbrace{(\mathbb{R} \to \mathbb{R})}_{V} \to (\mathbb{R} \to \mathbb{R})$

$$\frac{d}{dx}(f+g) = \frac{d}{dx}f + \frac{d}{dx}g$$

$$\frac{d}{dx}(3f) = 3\underbrace{\frac{d}{dx}f}_{\in \mathbb{R} \to \mathbb{R}}$$

↳ $\frac{d}{dx}$ is a **linear** operator

$$\frac{d}{dx} \underbrace{3x^2}_{\in V = \mathbb{R} \to \mathbb{R}} = \underbrace{6x}_{\in V}$$