Data Mining 2026 Spring!

This class:

- linear algebra + probability
- machine learning
- foundations of gen AI

challenging course!

Prereqs: linear algebra, probability, algorithms, calc 3

Time Expectation: 9.5 hr/week

2.5	1	1	2	1	1	1
class	Problem Set			OH	discussion	"rock"

- Earlier = better
- More drawing = better
- More talking = better

www.rtealwitter.com/datamining2026spring

- discord for communication
- reading for each lecture
- these slides are online

Plan:

- Warm up
  - ↳ math review
  - ↳ Page Rank
- Supervised Learning
- Unsupervised Learning

Math Review

machine learning?

calculus and linear algebra!

Derivatives

$f: \mathbb{R} \to \mathbb{R}$

"maps real number to real number"

$$f'(x) = \frac{\partial}{\partial x} f(x)$$

$$= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

Function: $f(x)$	Derivative: $\frac{\partial}{\partial x} f(x)$
x^2	$2x$
x^a	
$ax+b$	
$\ln(x)$	
e^x	

Q: If my investment earns 10% per year, how much more will I have in 60 years?

In ML, we'll combine many functions (e.g., neural net)

$f: \mathbb{R} \to \mathbb{R}$ $\quad g: \mathbb{R} \to \mathbb{R}$

$f(x) = \ln(x)$ $\quad g(x) = x^2$

Chain Rule

$$\frac{\partial [g(f(x))]}{\partial x} = g'(f(x))\, f'(x)$$

$$=$$

Product Rule

$$\frac{\partial}{\partial x}[g(x) \cdot f(x)]$$

$$= g(x)\, f'(x) + f(x)\, g'(x)$$

$$=$$

Gradients

In ML, process high-dim input

$f: \mathbb{R}^d \to \mathbb{R}$

"maps d real numbers to one real"

$f(x_1, x_2, \ldots, x_d)$ ← visualize as mountain with x_1 north/south, x_2 east/west

$$= f\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \right) = f(\underline{x})$$

$\frac{\partial}{\partial x_i} f(\underline{x})$ = how changing x_i changes f

$$= \lim_{h \to 0} \frac{f(x_1, \ldots, x_i + h, \ldots, x_d) - f(x_1, \ldots, x_i, \ldots, x_d)}{h}$$

$$\nabla_{\underline{x}} f(\underline{x}) = \begin{bmatrix} \frac{\partial f(\underline{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\underline{x})}{\partial x_2} \end{bmatrix}$$

Vector and Matrix Multiplication

↑ crucial to ML, neural nets only succeeded with GPUs

$\underline{u} \in \mathbb{R}^d$ $\qquad$ $\underline{v} \in \mathbb{R}^d$

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_d \end{bmatrix} \qquad \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{bmatrix}$$

$$\underline{u} \cdot \underline{v} = \sum_{i=1}^{d} u_i v_i = u_1 v_1 + u_2 v_2 + \dots + u_d v_d$$

$$= \underline{u}^T \underline{v} = [u_1 \ u_2 \dots u_d] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{bmatrix}$$

$$= \underline{v}^T \underline{u} = [v_1 \ v_2 \ \dots \ v_d] \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_d \end{bmatrix}$$

$$= \langle \underline{u}, \underline{v} \rangle$$

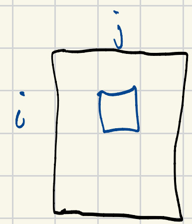
ℓ_2-norm

$$\|\underline{u}\|_2^2 = \underline{u} \cdot \underline{u} = \sum_{i=1}^{d} u_i^2$$

$A \in \mathbb{R}^{d \times m}$ $\qquad$ $B \in \mathbb{R}^{m \times n}$

$$\begin{array}{cc} & \\ i & \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & & & & \\ \vdots & & & & \\ & & & & a_{dm} \end{bmatrix} \end{array} \qquad \begin{array}{c} j \\ \begin{bmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \\ b_{m1} & & b_{mn} \end{bmatrix} \end{array}$$

$$\underbrace{\underset{d \times m}{A}\ \underset{m \times n}{B}}_{d \times n}$$



$$[AB]_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}$$

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$AB = ?$ $\qquad$ $BA = ?$

Week 1 Thursday

- No OH Monday 1/26
  - ⇒ Problem 2 due Monday 2/2
- Join discord!
- Discussions!
  - ↳ host can write quiz question
  - ↳ TA hosts 6-8pm Sunday in RN 104

Today:

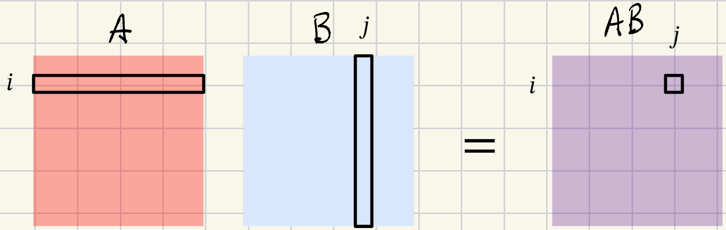
- Matrices
  - ↳ multiplication
  - ↳ eigendecomposition
  - ↳ inversion

- Next up: Page Rank!
  i.e., how to return
  most relevant pages

Matrix Multiplication

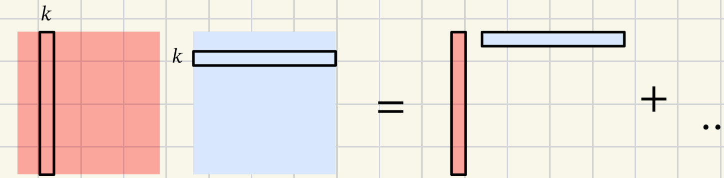
$A \in \mathbb{R}^{d\times m}$ $\quad B \in \mathbb{R}^{m\times n}$

Perspective #1



$$[AB]_{i,j} = \sum_{k=1}^{m} A_{ik} B_{kj} = A_{i,:} \cdot B_{:,j}$$

Perspective #2



$$AB = \sum_{k=1}^{m} A_{:,k} B_{k,:}$$

Check: $[AB]_{i,j} = \sum_{k=1}^{m} A_{i,k} B_{k,j}$

Example: $A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & -2 \\ 0 & 1 & 0 \end{bmatrix}$ $\quad B = \begin{bmatrix} 2 & 0 \\ -3 & 1 \\ 1 & 2 \end{bmatrix}$

Eigenvalues & Eigendecomposition

Consider square, symmetric $A \in \mathbb{R}^{d \times d}$

when does A act like a scalar?

$A\underline{v} = \lambda \underline{v}$ $\quad$ $\underline{v} \in \mathbb{R}^d$ eigenvector

$\lambda \in \mathbb{R}$ eigenvalue

$\underline{v}_1, \dots, \underline{v}_r$ $\quad$ eigenvectors

$\lambda_1, \dots, \lambda_r$ $\quad$ eigenvalues

Eigenvectors are orthnormal

$\langle \underline{v}_i, \underline{v}_j \rangle = 0$ if $i \neq j$

$\langle \underline{v}_i, \underline{v}_i \rangle = 1$

$$A = \begin{bmatrix} | & | & & | \\ \boldsymbol{v}_1 & \boldsymbol{v}_2 & \dots & \boldsymbol{v}_r \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_r \end{bmatrix} \begin{bmatrix} \text{---} & \boldsymbol{v}_1^T & \text{---} \\ \text{---} & \boldsymbol{v}_2^T & \text{---} \\ & \vdots & \\ \text{---} & \boldsymbol{v}_r^T & \text{---} \end{bmatrix}$$

$V \in \mathbb{R}^{d \times r}$ $\quad$ $\Lambda \in \mathbb{R}^{r \times r}$ $\quad$ $V^T \in \mathbb{R}^{r \times d}$

$$= \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \dots & v_r \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 v_1^T \\ \vdots \\ \lambda_r v_r^T \end{bmatrix}$$

$$= \sum_{k=1}^{r} \underline{v}_k \lambda_k \underline{v}_k^T$$

What is $A\underline{v}_i$?

Matrix Inversion

Motivation: With $a \in \mathbb{R}$, $(a)^{-1} a = 1$

"1" in higher dimensions: $I_{4\times 4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

"identity"

$A^{-1} A = I$ Claim: $A^{-1} = V \Lambda^{-1} V^T$ when $r = d$

$$A^{-1} A = V \Lambda^{-1} V^T V \Lambda V^T = V \Lambda^{-1} I_{d\times d} \Lambda V^T = V \Lambda^{-1} \Lambda V^T = V I_{d\times d} V^T = V V^T$$

$$\underset{d\times d}{V^T} \underset{d\times d}{V} = I_{d\times d} \qquad \underset{d\times d}{\Lambda^{-1}} \underset{d\times d}{\Lambda} = \underset{d\times d}{I}$$

$$\begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_d^T & - \end{bmatrix} \begin{bmatrix} | & & | \\ v_1 & \cdots & v_d \\ | & & | \end{bmatrix} \qquad \begin{bmatrix} 1/\lambda_1 & & \\ & \ddots & \\ & & 1/\lambda_d \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_d \end{bmatrix}$$

Another View on Matrix Inversion

$$A^{-1} = \sum_{i=1}^{d} \frac{1}{\lambda_i} v_i v_i^T$$

$$A^{-1}A = \sum_{i=1}^{d} \frac{1}{\lambda_i} v_i v_i^T \sum_{j=1}^{d} \lambda_j v_j v_j^T$$

$$= \sum_{i,j=1}^{d} \frac{\lambda_j}{\lambda_i} v_i \underbrace{v_i^T v_j}_{= \begin{cases} 1 & \text{if } i=j \\ 0 & \text{else} \end{cases}} v_j^T$$

$$= \sum_{i,j=1 \,:\, i=j}^{d} \frac{\lambda_j}{\lambda_i} v_i v_j^T$$

$$= \sum_{i=1}^{d} v_i v_i^T$$

$$= V V^T$$

Claim: $\sum_{i=1}^{d} v_i v_i^T = I_{d \times d}$

Proof: Since $v_1, \dots, v_d$ span $\mathbb{R}^d$ (as orthonormal basis), we can write

projection of x onto v_i

$$x = c_1 v_1 + c_2 v_2 + \dots + c_d v_d$$

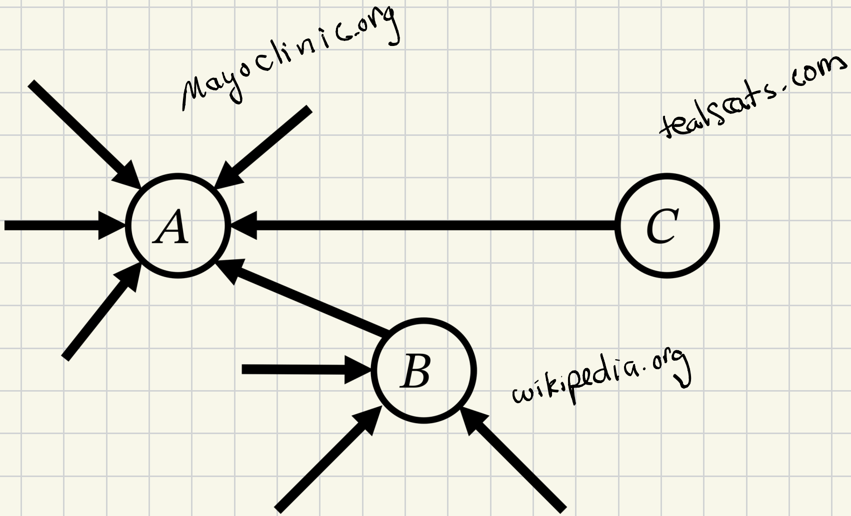
$$= \sum_{i=1}^{d} v_i (v_i \cdot x)$$

$$= \sum_{i=1}^{d} v_i v_i^T x$$

Hence $\sum_{i=1}^{d} v_i v_i^T$ acts as identity for all x $\iff$ must be identity!

Page Rank

How do we find <u>important</u> webpages?



Important if:

- lots of incoming links
- incoming links from <u>important</u> pages

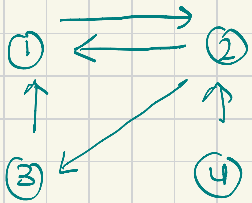
$\boxed{\text{PageRank}}$

iterative process of computing importance

$p^{(0)} \in \mathbb{R}^n$ where $p_i^{(0)} = 1/n$

$$p_i^{(t+1)} = \sum_{j=1}^{n} \mathbb{1}[j \text{ links to } i] \frac{p_j^{(t)}}{d_j}$$

d_j ← # outgoing edges from j



$$p^{(0)} = \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$$

$p^{(1)} =$