

CS 260: Foundations of Data Science

Prof. Sara Mathieson

Spring 2025



Haverford
COLLEGE

Admin

- **Lab 2** due tomorrow, ideally finish in lab today
- **Lab 3** posted, can start in lab if Lab 2 finished
- Updated **Office Hours**
 - Mon 12-1pm (L302)
 - Wed 4-5pm (zoom)
- **Summer research** applications due Friday!

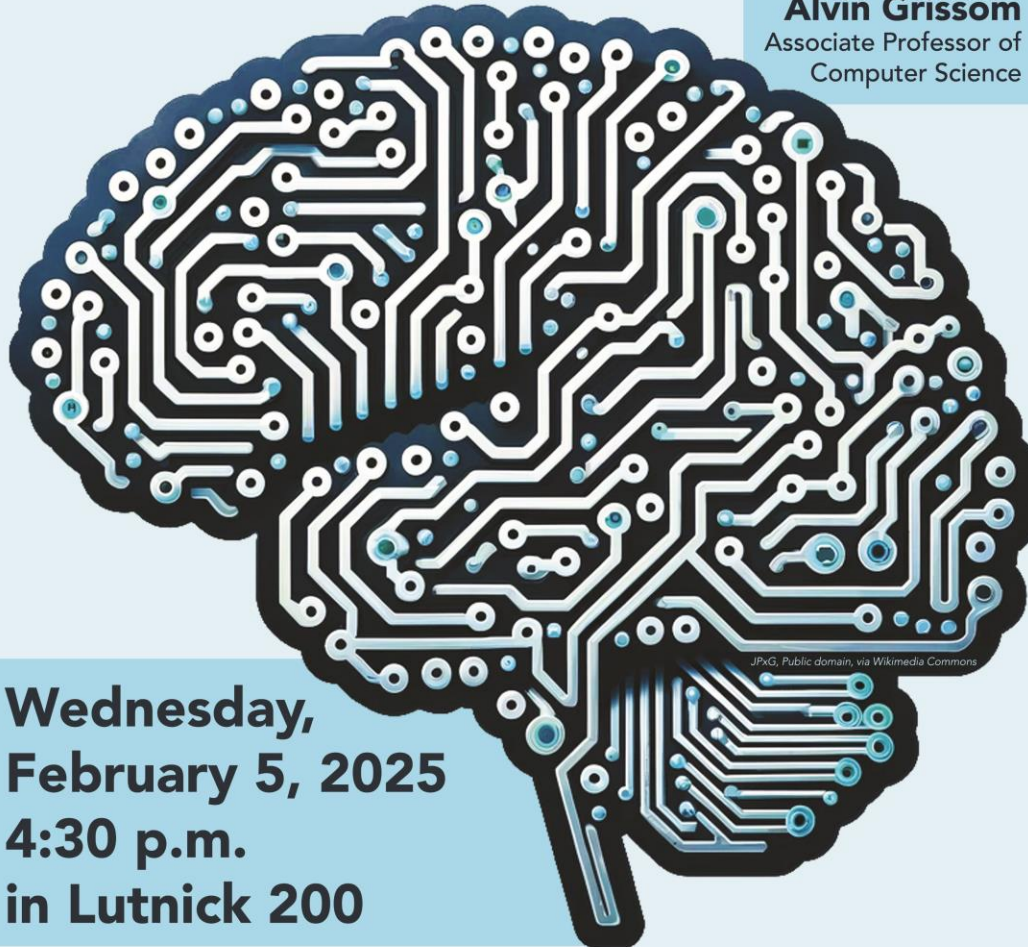
Tomorrow!

Demystifying Large Language Models

What They Are and
Why I Rarely Use Them



Alvin Grissom
Associate Professor of
Computer Science



**Wednesday,
February 5, 2025
4:30 p.m.
in Lutnick 200**

Sponsored by the Computer Science department and

HAVERFORD
LIBRARIES

Outline

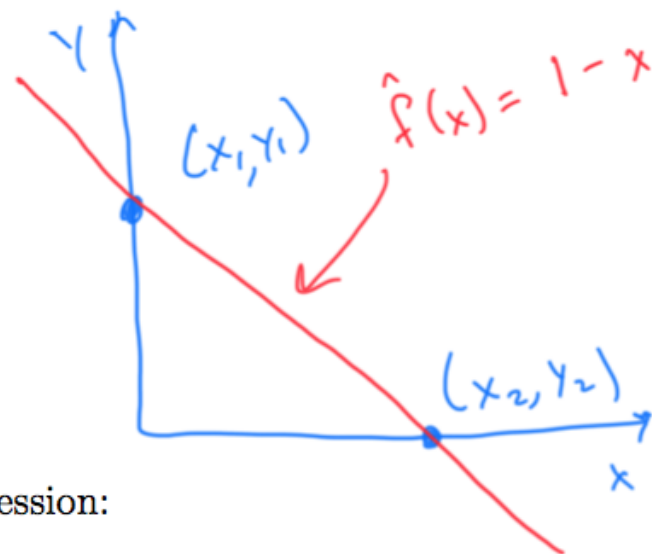
- Recap *simple* (i.e. $p=1$) linear regression
- Introduction to applied linear algebra
- *Multiple* linear regression
- Analytic solution to multiple linear regression

Outline

- Recap *simple* (i.e. $p=1$) linear regression
- Introduction to applied linear algebra
- *Multiple* linear regression
- Analytic solution to multiple linear regression

Handout 4 Let $n = 2$ and $p = 1$, with the following data (we will omit the first column of 1's in simple linear regression):

$$\mathbf{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$



(a) Plot these two points – what should \hat{w}_0 and \hat{w}_1 be?

$$\hat{w}_0 = 1$$

$$\hat{w}_1 = -1$$

(b) This week we derived the solution for simple linear regression:

note:

$$\bar{x} = \frac{1}{2}$$

$$\bar{y} = \frac{1}{2}$$

$$\hat{w}_1 = \frac{\text{Cov}(\mathbf{x}, \mathbf{y})}{\text{Var}(\mathbf{x})} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\hat{w}_0 = \bar{y} - \hat{w}_1 \bar{x}$$

where $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ and $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$. Use these equations to compute \hat{w}_0 and \hat{w}_1 and verify your answer to (a).

$$\hat{w}_1 = \frac{\frac{1}{2} [(1 - \frac{1}{2})(0 - \frac{1}{2}) + (0 - \frac{1}{2})(1 - \frac{1}{2})]}{\frac{1}{2} [(1 - \frac{1}{2})^2 + (0 - \frac{1}{2})^2]}$$

$$= \frac{-\frac{1}{4} - \frac{1}{4}}{\frac{1}{4} + \frac{1}{4}}$$

\Rightarrow

$$\boxed{\hat{w}_1 = -1}$$

$$\hat{w}_0 = \frac{1}{2} - (-1) \frac{1}{2}$$

$$\Rightarrow \boxed{\hat{w}_0 = 1}$$

Outline

- Recap *simple* (i.e. $p=1$) linear regression
- Introduction to applied linear algebra
- *Multiple* linear regression
- Analytic solution to multiple linear regression

Vectors

- Vector magnitude

$$\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \text{then} \quad |\mathbf{v}| = \sqrt{x^2 + y^2}.$$

- Different ways to write a vector

$$\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix}^T$$

- Vector dot product (mini-quiz)

$$\vec{v}_1 = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

$$\vec{v}_2 = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

$$\vec{u}_1 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$$

$$\vec{u}_2 = \begin{bmatrix} 0 \\ 10 \end{bmatrix}$$

$$\vec{v}_1 \cdot \vec{v}_2 = ?$$

$$23 = 8 \cdot 1 + 3 \cdot 5$$



$$\vec{u}_1 \cdot \vec{u}_2 = ?$$

$$0$$

$$\begin{bmatrix} 0 \\ 10 \end{bmatrix}$$

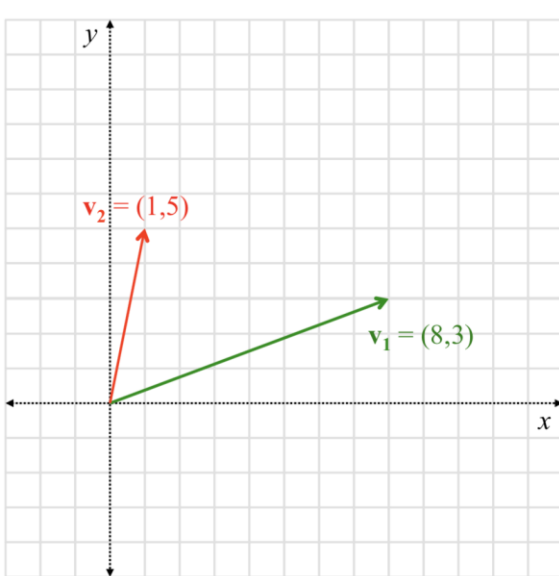
$$\begin{bmatrix} 5 \\ 0 \end{bmatrix}$$

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \vdots \\ \vdots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

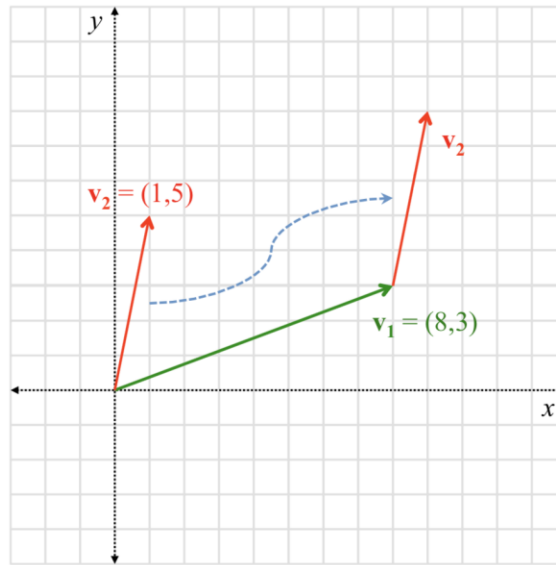
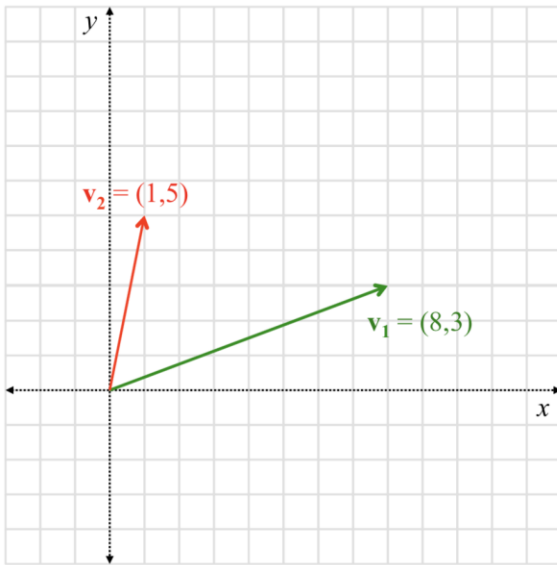
Vector Addition

$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{bmatrix} 8 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 8 + 1 \\ 3 + 5 \end{bmatrix} = \begin{bmatrix} 9 \\ 8 \end{bmatrix}$$



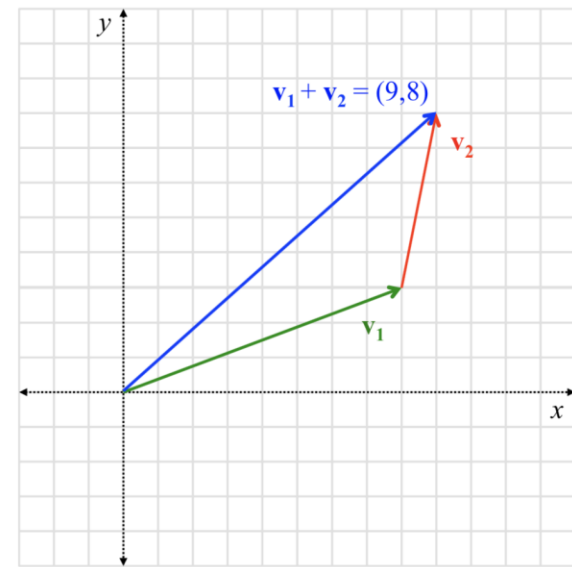
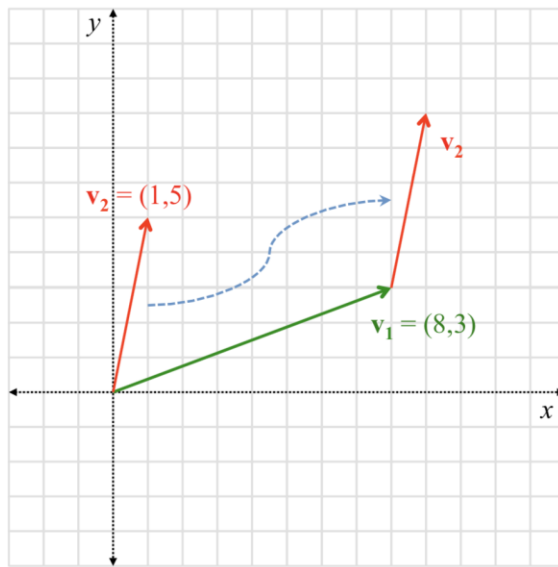
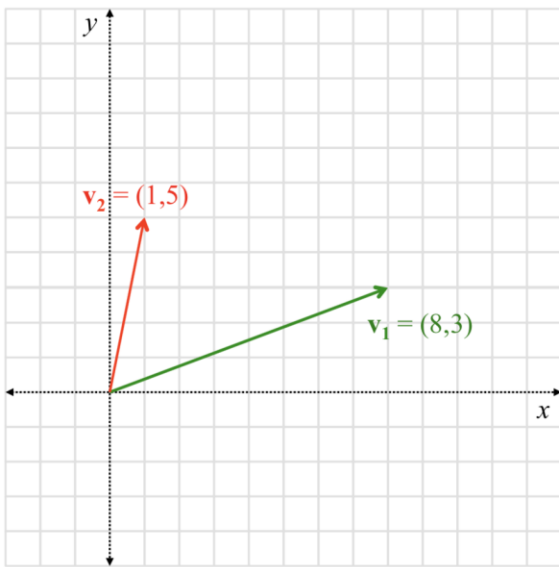
Vector Addition

$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{bmatrix} 8 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 8 + 1 \\ 3 + 5 \end{bmatrix} = \begin{bmatrix} 9 \\ 8 \end{bmatrix}$$



Vector Addition

$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{bmatrix} 8 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 8 + 1 \\ 3 + 5 \end{bmatrix} = \begin{bmatrix} 9 \\ 8 \end{bmatrix}$$



Matrices

- Matrix addition (must be exactly the same dimension!)

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a + e & b + f \\ c + g & d + h \end{bmatrix}$$

Matrix Multiplication

- inner dimensions must match
- If $A.\text{shape} = (m, n)$ and $B.\text{shape} = (n, p)$, then $AB.\text{shape} = (m, p)$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & \\ & \end{bmatrix}$$

dot product of row 1 and col 1



Matrix Multiplication

- inner dimensions must match
- If $A.\text{shape} = (m, n)$ and $B.\text{shape} = (n, p)$, then $AB.\text{shape} = (m, p)$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & \\ & \end{bmatrix} \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ & \end{bmatrix}$$

dot product of row 1 and col 2



Matrix Multiplication

- inner dimensions must match
- If $A.shape = (m, n)$ and $B.shape = (n, p)$, then $AB.shape = (m, p)$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & \\ & \end{bmatrix} \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ & \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & \end{bmatrix} \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$$

Matrix Transpose

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \quad \mathbf{A}^T = \begin{bmatrix} a & d \\ b & e \\ c & f \end{bmatrix}$$

Useful note: $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$

Matrix Inverse

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \mathbf{A}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Useful note: $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$

Outline

- Recap *simple* (i.e. $p=1$) linear regression
- Introduction to applied linear algebra
- ***Multiple*** linear regression
- Analytic solution to multiple linear regression

Lab 3: USA Housing data

Avg. Area Income	Avg. Area House Age	Avg. Area Number of Rooms	Avg. Area Number of Bedrooms	Area Population	Price
79545.45857	5.682861322	7.009188143	4.09	23086.8005	1059033.558
79248.64245	6.002899808	6.730821019	3.09	40173.07217	1505890.915
61287.06718	5.86588984	8.51272743	5.13	36882.1594	1058987.988
63345.24005	7.188236095	5.586728665	3.26	34310.24283	1260616.807
59982.19723	5.040554523	7.839387785	4.23	26354.10947	630943.4893
80175.75416	4.988407758	6.104512439	4.04	26748.42842	1068138.074
64698.46343	6.025335907	8.147759585	3.41	60828.24909	1502055.817
78394.33928	6.989779748	6.620477995	2.42	36516.35897	1573936.564
59927.66081	5.36212557	6.393120981	2.3	29387.396	798869.5328
81885.92718	4.42367179	8.167688003	6.1	40149.96575	1545154.813
80527.47208	8.093512681	5.0427468	4.1	47224.35984	1707045.722
50593.6955	4.496512793	7.467627404	4.49	34343.99189	663732.3969
39033.80924	7.671755373	7.250029317	3.1	39220.36147	1042814.098
73163.66344	6.919534825	5.993187901	2.27	32326.12314	1291331.518
69391.38018	5.344776177	8.406417715	4.37	35521.29403	1402818.21
73091.86675	5.443156467	8.517512711	4.01	23929.52405	1306674.66
79706.96306	5.067889591	8.219771123	3.12	39717.81358	1556786.6

X

y

Multiple Linear Regression

$$X = \begin{bmatrix} 1 & \vec{x}_1^T \\ \vdots & \vdots \\ 1 & \vec{x}_n^T \\ \vdots & \vdots \end{bmatrix}$$

"fake" ones

$n \times (p+1)$

$$\hat{y} = h_{\vec{w}}(\vec{x}) = w_0 + w_1 x_1 + w_2 x_2 + \dots + w_p x_p$$

$x_0 = 1$

$= (\vec{w}) \cdot \vec{x}$ goal find the best \vec{w}

Goal: minimize the cost function

$$J(\vec{w}) = \frac{1}{2} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$= \frac{1}{2} \sum_{i=1}^n (y_i - \vec{w} \cdot \vec{x}_i)^2$$

take gradient (derivative) & set
equal to $\vec{0}$ (zero vector)

Outline

- Recap *simple* (i.e. $p=1$) linear regression
- Introduction to applied linear algebra
- *Multiple* linear regression
- Analytic solution to multiple linear regression

$p \times p$

\vec{w}

$\{$

$$J(\vec{w}) = \frac{1}{2} (\vec{y} - X\vec{w}) \cdot (\vec{y} - X\vec{w})$$

$$= \begin{bmatrix} y_1 - \hat{y}_1 \\ y_2 - \hat{y}_2 \\ \vdots \\ y_n - \hat{y}_n \end{bmatrix} \cdot \begin{bmatrix} y_1 - \hat{y}_1 \\ y_2 - \hat{y}_2 \\ \vdots \\ y_n - \hat{y}_n \end{bmatrix}$$

$X\vec{w}$

$$J(\vec{w}) = \frac{1}{2} (\vec{y} \cdot \vec{y} - 2\vec{y} \cdot (X\vec{w}) + (X\vec{w}) \cdot (X\vec{w}))$$

$$\nabla J(\vec{w}) = X^T X \vec{w} - X^T \vec{y} = \vec{0}$$

"derivative"

$$X^T X \vec{w} = X^T \vec{y}$$

$$X \vec{w} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_p \end{bmatrix} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix}$$

Computing predictions given X and w

$n \times (p+1)$ $(p+1) \times 1$ $n \times 1$

Predictions!

hc-cs@

github.com

pull / push

clone / pull / push



$$(X^T X)^{-1} (X^T X) \vec{w} = (X^T X)^{-1} X^T \vec{y}$$

$$\hat{\vec{w}} = (X^T X)^{-1} X^T \vec{y} \quad \star$$

$(p+1) \times n$

variance
 X

covariance
 $X \text{ \& } \vec{y}$

$n \times (p+1)$

$(p+1) \times n$

$(p+1) \times 1$

Analytic solution to multiple linear regression

(Keep this formula and its interpretation in mind!)

Handout 5

Handout 5, #1

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 0 & 2 \\ -4 & 1 \end{bmatrix}.$$

$$\mathbf{AB} = \begin{bmatrix} 4 & 1 \\ -8 & 8 \end{bmatrix} \quad \text{and} \quad \mathbf{BA} = \begin{bmatrix} 6 & 4 \\ -1 & 6 \end{bmatrix}.$$

$$(ab)c = a(bc) \checkmark$$

$$AB \neq BA$$

Handout

$$AB = \begin{bmatrix} 4 & 1 \\ -8 & 8 \end{bmatrix}$$

$$BA = \begin{bmatrix} 6 & 4 \\ -1 & 6 \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

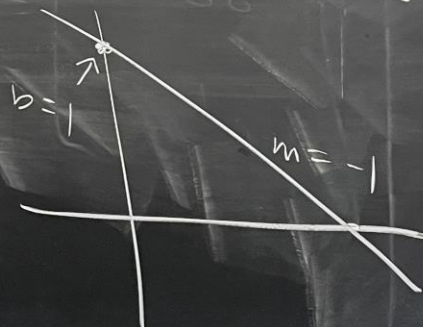
$$\vec{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\vec{w} = (X^T X)^{-1} X^T \vec{y} \quad \text{analytic solution}$$

$$= \left(\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$= \frac{1}{2-1} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$



Same as before!

Handout 5

Begin gradient descent

Gradient Descent

example
want to minimize

$$f(w) = (w-3)^2 + 2$$

$$f'(w) = 2(w-3) = 0$$

$$2w = 6$$

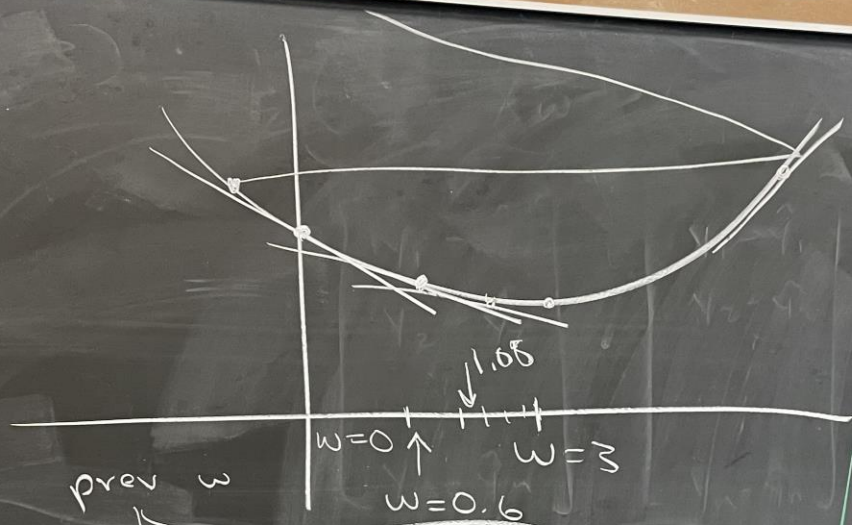
$$w = 3$$

can take the derivative,
but we can't set equal to 0
+ solve

$$\begin{aligned} f'(0) &= 2(0-3) \\ &= -6 \end{aligned}$$

prev w
↓
① $w \leftarrow 0$

op
dr
o
0
↓
 $w \leftarrow 0.6$



$$\textcircled{1} \quad w \leftarrow 0 - \underbrace{0.1}_{\substack{\text{Step size} \\ \alpha}} \underbrace{(2(0-3))}_{f'(w)}$$

opposite direction of gradient

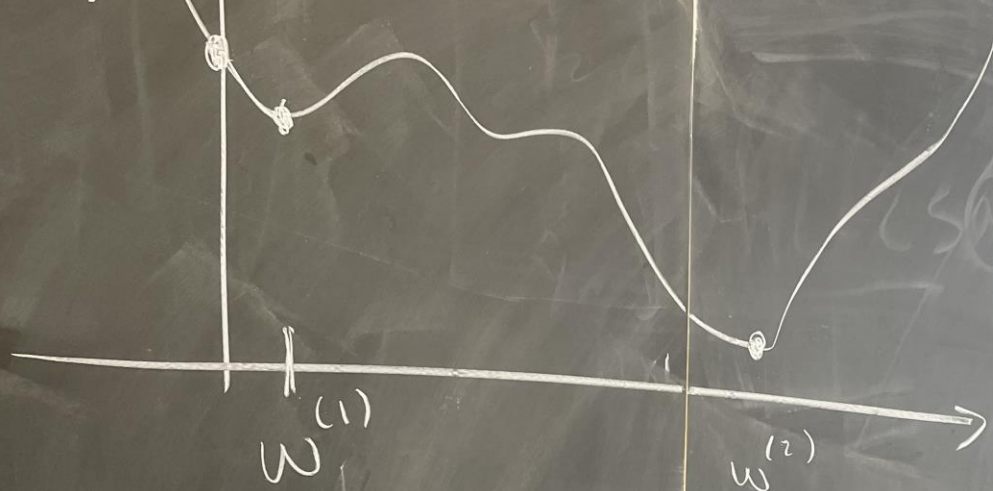
$$w \leftarrow 0 - 0.1(-6)$$

$$w \leftarrow 0.6$$

$$\textcircled{2} \quad w \leftarrow 0.6 - 0.1(2(0.6-3))$$

$$w \leftarrow 1.08$$

(0.6 - 3)



$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$A^{-1} = \frac{1}{da-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$AA^{-1} = \frac{1}{da-bc} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$= \frac{1}{da-bc} \begin{bmatrix} ad-bc & 0 \\ 0 & da-bc \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Quote of the week

“They tried to bury us, but they didn’t know we were seeds.” –
Dinos Christianopoulos