CS 260: Foundations of Data Science

Prof. Sara Mathieson

Fall 2023

HAVERFORD
COLLEGE

Admin

- **Lab 1** grades returned hopefully Wednesday (on Moodle)

- **Lab 2** was due last night unless you're taking a late day

- **Lab 3** posted, start in lab today

Outline for September 19

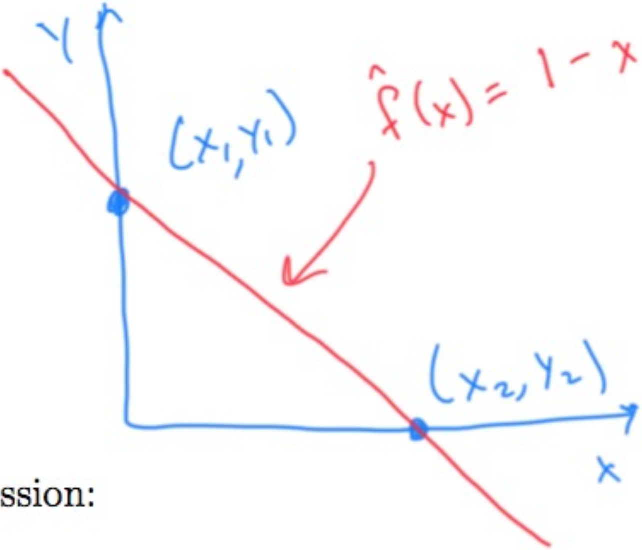
- Recap *simple* (i.e. p=1) linear regression
- Introduction to applied linear algebra
- *Multiple* linear regression
- Analytic solution to multiple linear regression

Outline for September 19

- 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):

$$\boldsymbol{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \boldsymbol{X} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$



(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:

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

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

$$\hat{w}_1 = \frac{\mathrm{Cov}(\boldsymbol{x}, \boldsymbol{y})}{\mathrm{Var}(\boldsymbol{x})} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad \hat{w}_0 = \bar{y} - \hat{w}_1\bar{x}$$

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).

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

$$= \frac{-\frac{1}{4} - \frac{1}{4}}{\frac{1}{4} + \frac{1}{4}} \Rightarrow \boxed{\hat{w}_1 = -1}$$

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

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

Outline for September 19

- 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

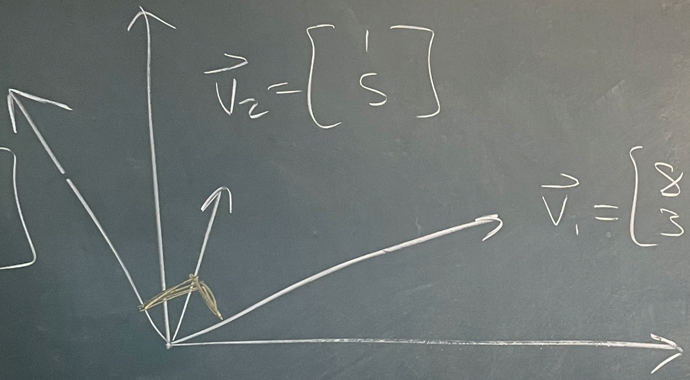
$$\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix}^{\mathrm{T}}$$

- Vector dot product

Vector dot product

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

$$\vec{v}_3 = \begin{bmatrix} -3 \\ 8 \end{bmatrix}$$



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

$$\vec{v}_1 \cdot \vec{v}_3 = -3\cdot 8 + \underbrace{8\cdot 3 = 0}_{\text{perpendicular}}$$

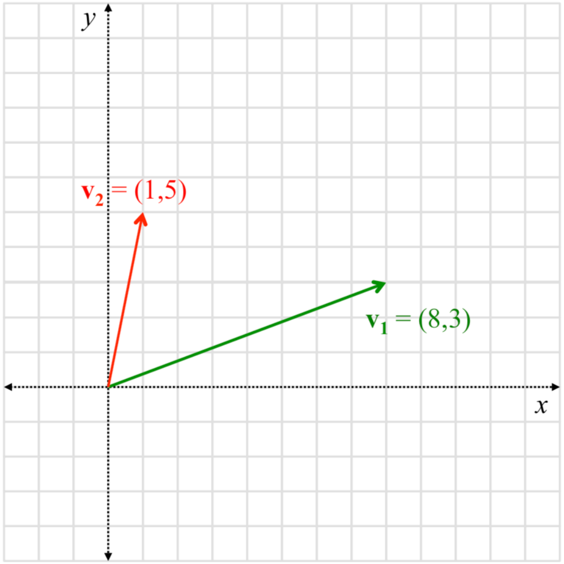
$$\vec{v}_1 \cdot \vec{v}_2 = \begin{bmatrix} 8 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 5 \end{bmatrix} = 8\cdot 1 + 3\cdot 5$$

$$= 23$$

positive → "same" direction

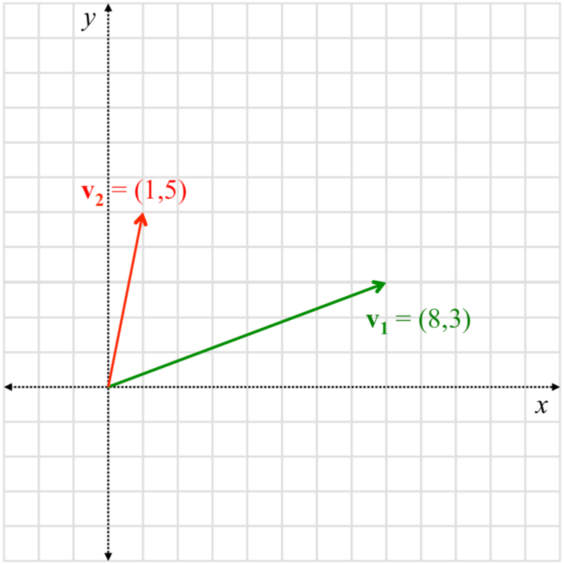
Vector Addition

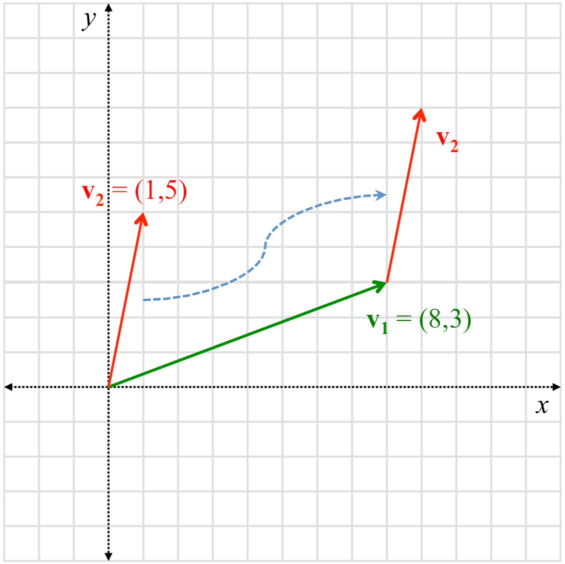
$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{bmatrix} 8 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 8+1 \\ 3+5 \end{bmatrix} = \begin{bmatrix} 9 \\ 8 \end{bmatrix}$$



Vector Addition

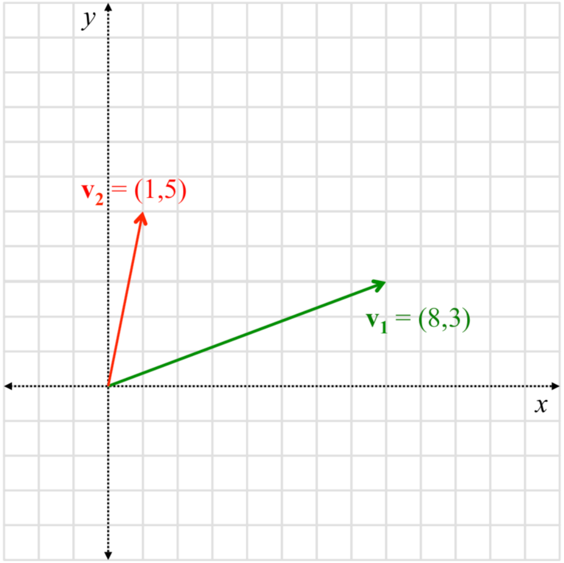
$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{bmatrix} 8 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 8+1 \\ 3+5 \end{bmatrix} = \begin{bmatrix} 9 \\ 8 \end{bmatrix}$$

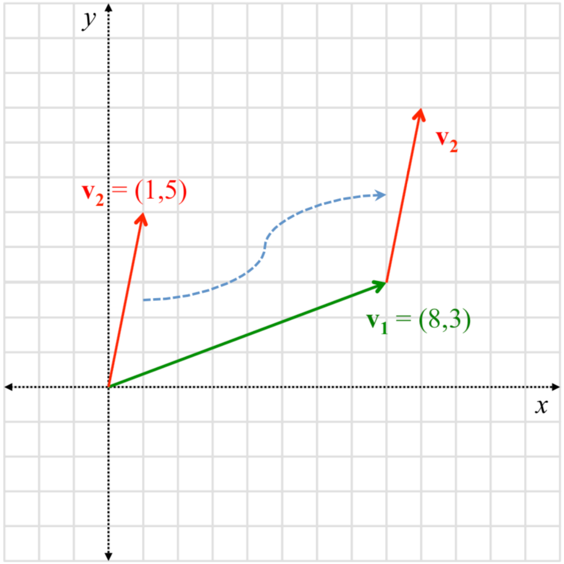


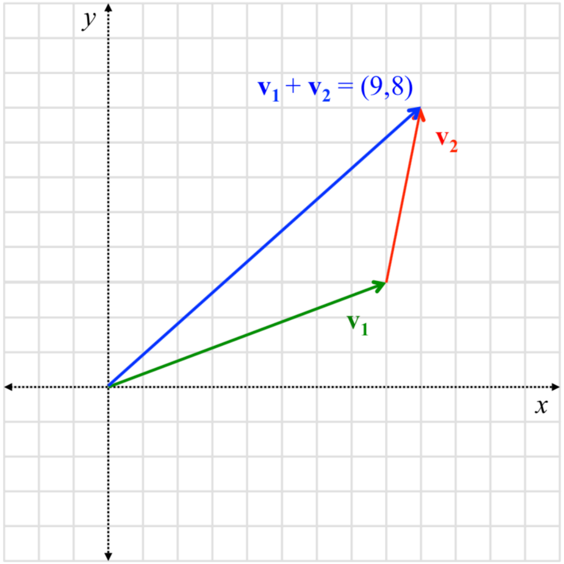


Vector Addition

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







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.shape = (m, n) and B.shape = (n, p), then AB.shape = (m,p)

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

dot product of row 1 and col 1

Matrix Multiplication

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

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

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

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)

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

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

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

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

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

Matrix Transpose

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

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

Matrix Inverse

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

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

$$AA^{-1} = \begin{bmatrix} ad-bc & \cancel{ab-ba} \\ \cancel{cd-dc} & -cb+da \end{bmatrix}$$

$$\nearrow$$

$$\frac{1}{ad-bc}$$

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

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

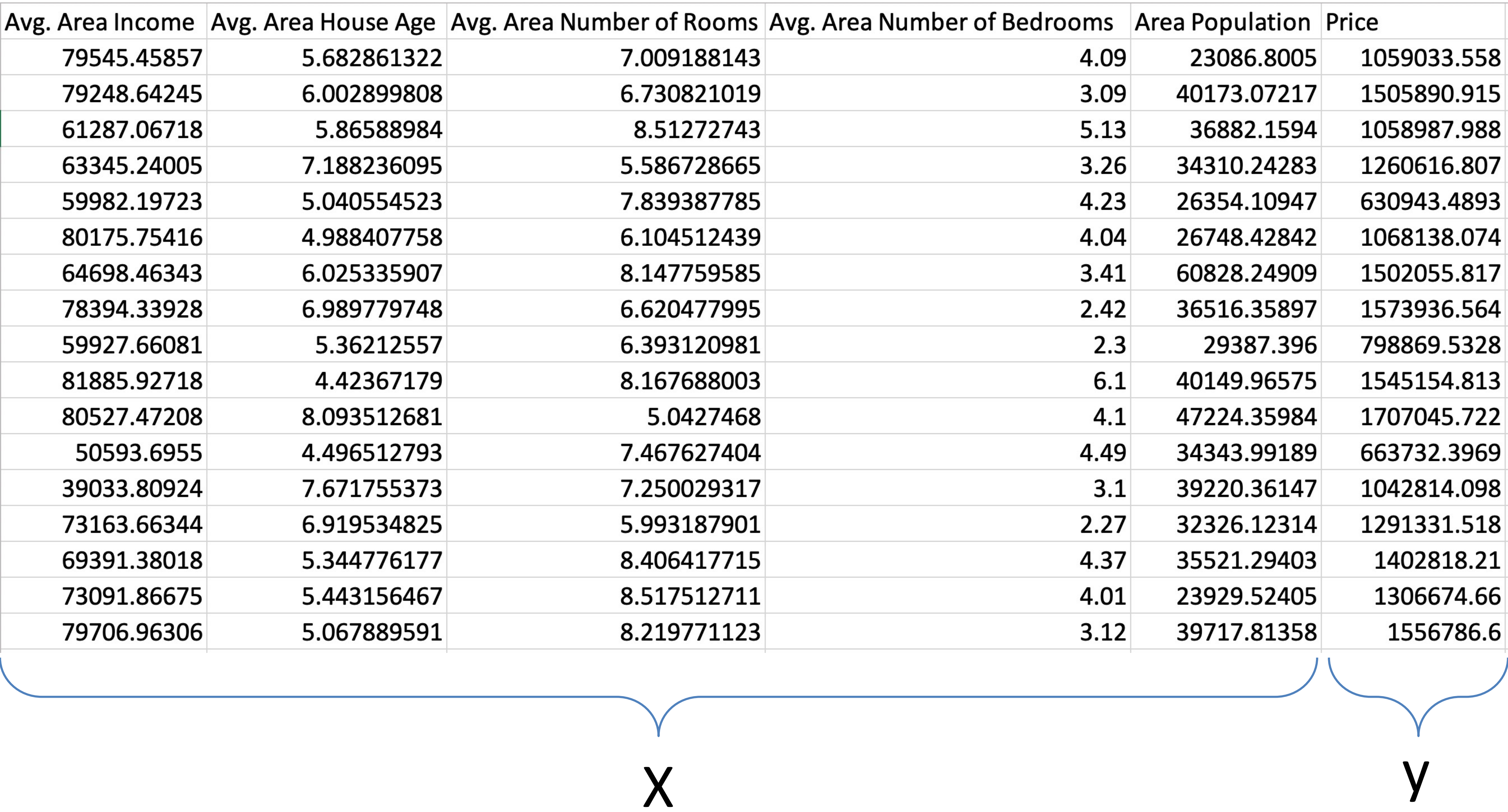
$$\frac{\cancel{2}x}{\cancel{2}} = \frac{3}{2}$$

$$\boxed{x = \frac{3}{2}}$$

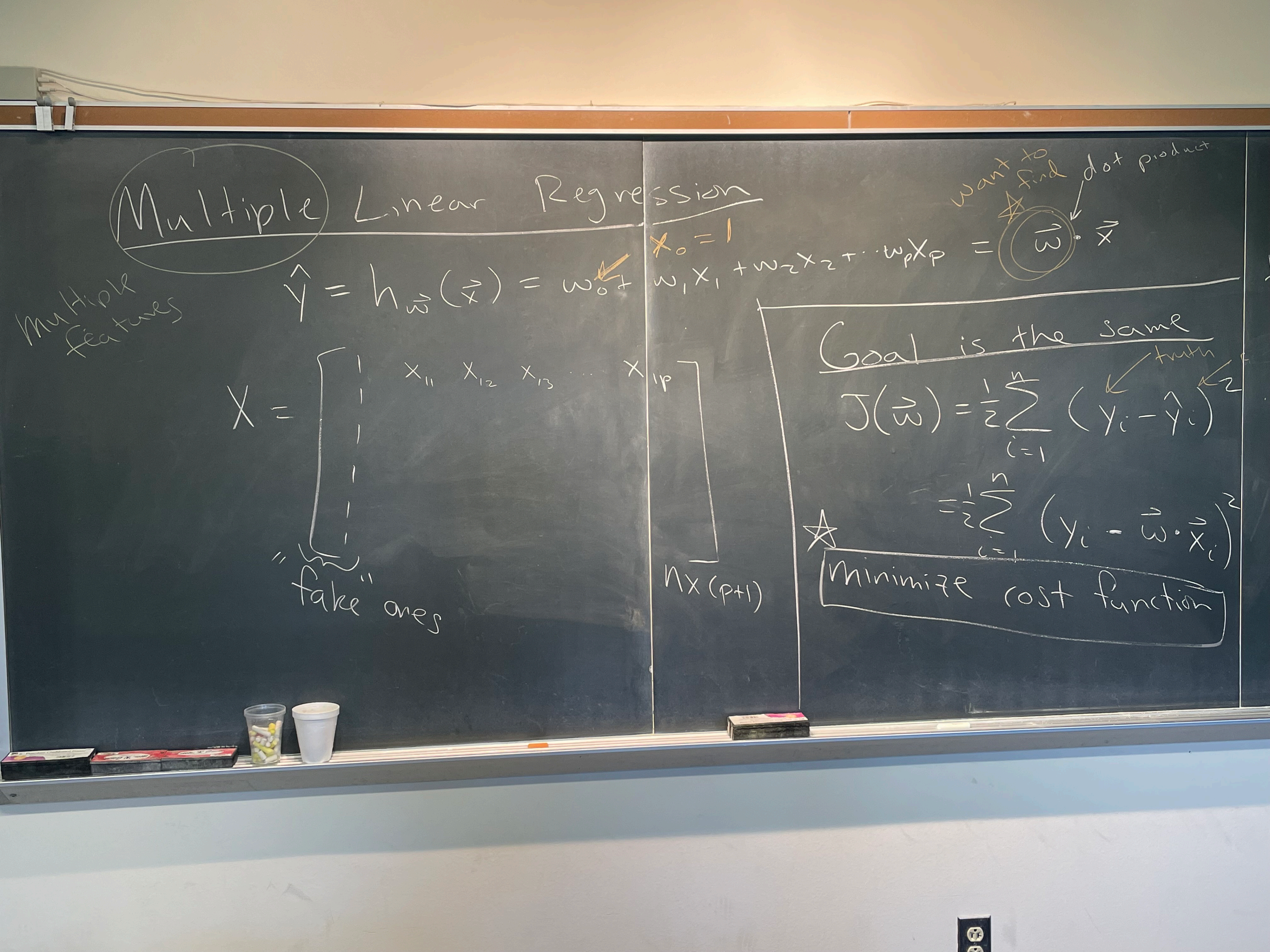
Outline for September 19

- 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



Multiple Linear Regression

multiple features

want to find

dot product

$x_0 = 1$

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

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} & \cdots & x_{1p} \\ 1 & & & & & \\ \vdots & & & & & \\ 1 & & & & & \end{bmatrix}$$

"fake" ones

$n \times (p+1)$

Goal is the same

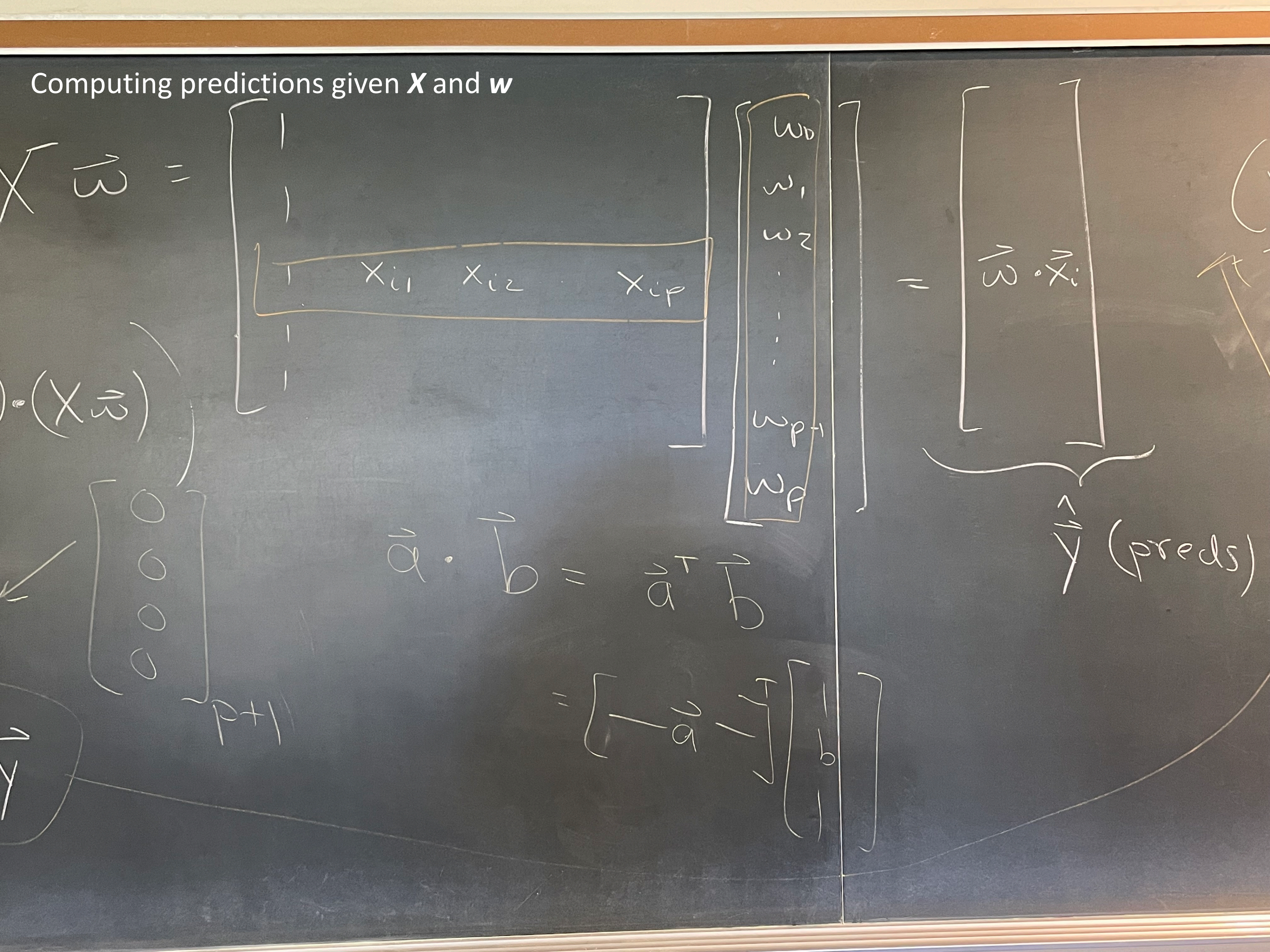
truth

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

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

☆ minimize cost function

Computing predictions given ***X*** and ***w***



Outline for September 19

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

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

$$X\vec{w} =$$

take derivative

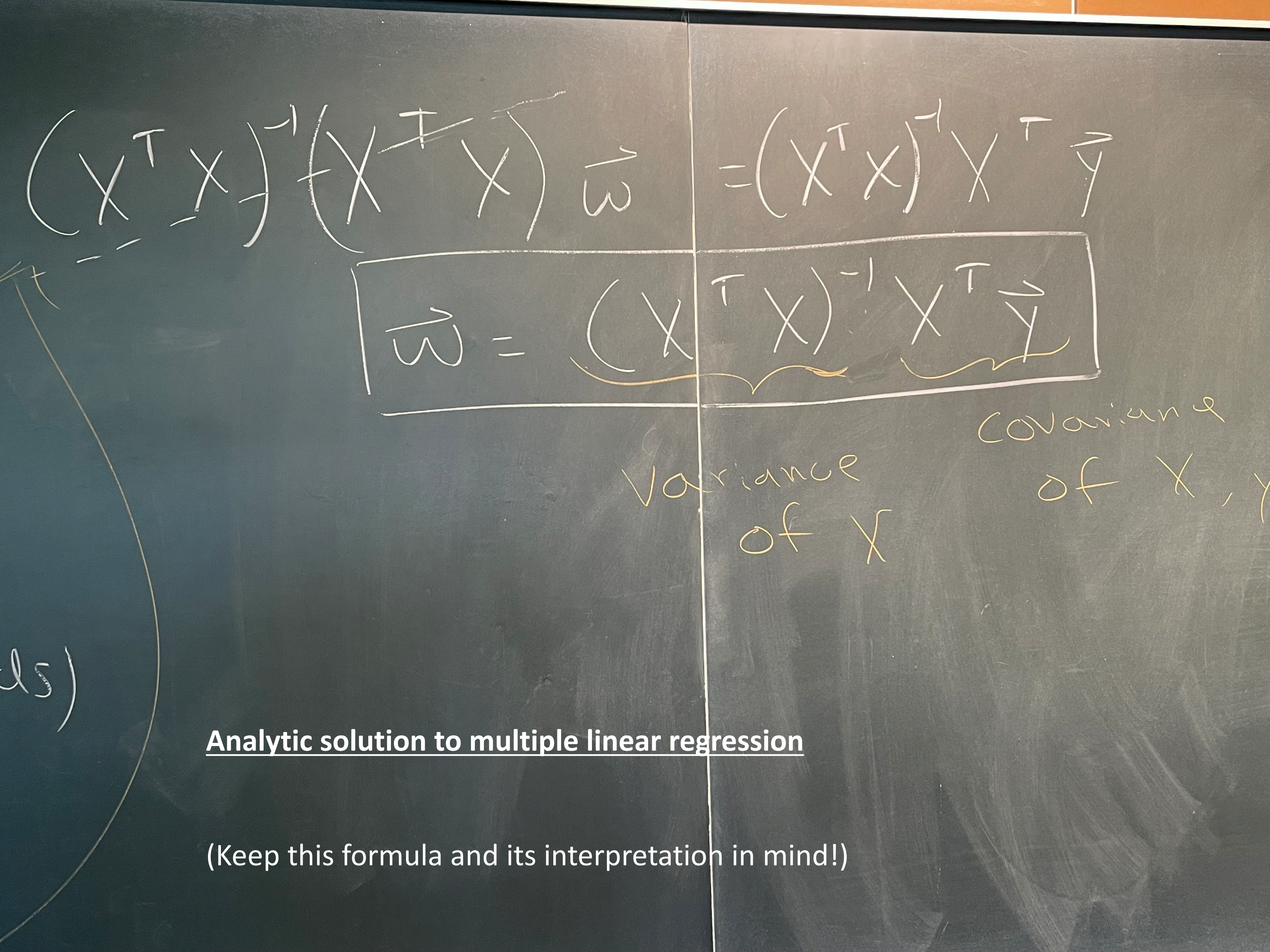
number (scalar)

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

$$\frac{\partial J}{\partial \vec{w}} = -X^T\vec{y} + (X^TX)\vec{w} = \vec{0} \quad \underbrace{\begin{bmatrix}0\\0\\0\\0\end{bmatrix}}_{p+1}$$

Solve for $\vec{w}$

$$\boxed{(\underset{(p+1)\times n}{X^T}\,\underset{n\times(p+1)}{X})\vec{w} = X^T\vec{y}}$$



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

$$\boxed{\vec{w} = \underbrace{(X^TX)^{-1}}_{\text{Variance of } X}\underbrace{X^T\vec{y}}_{\text{Covariance of } X, Y}}$$

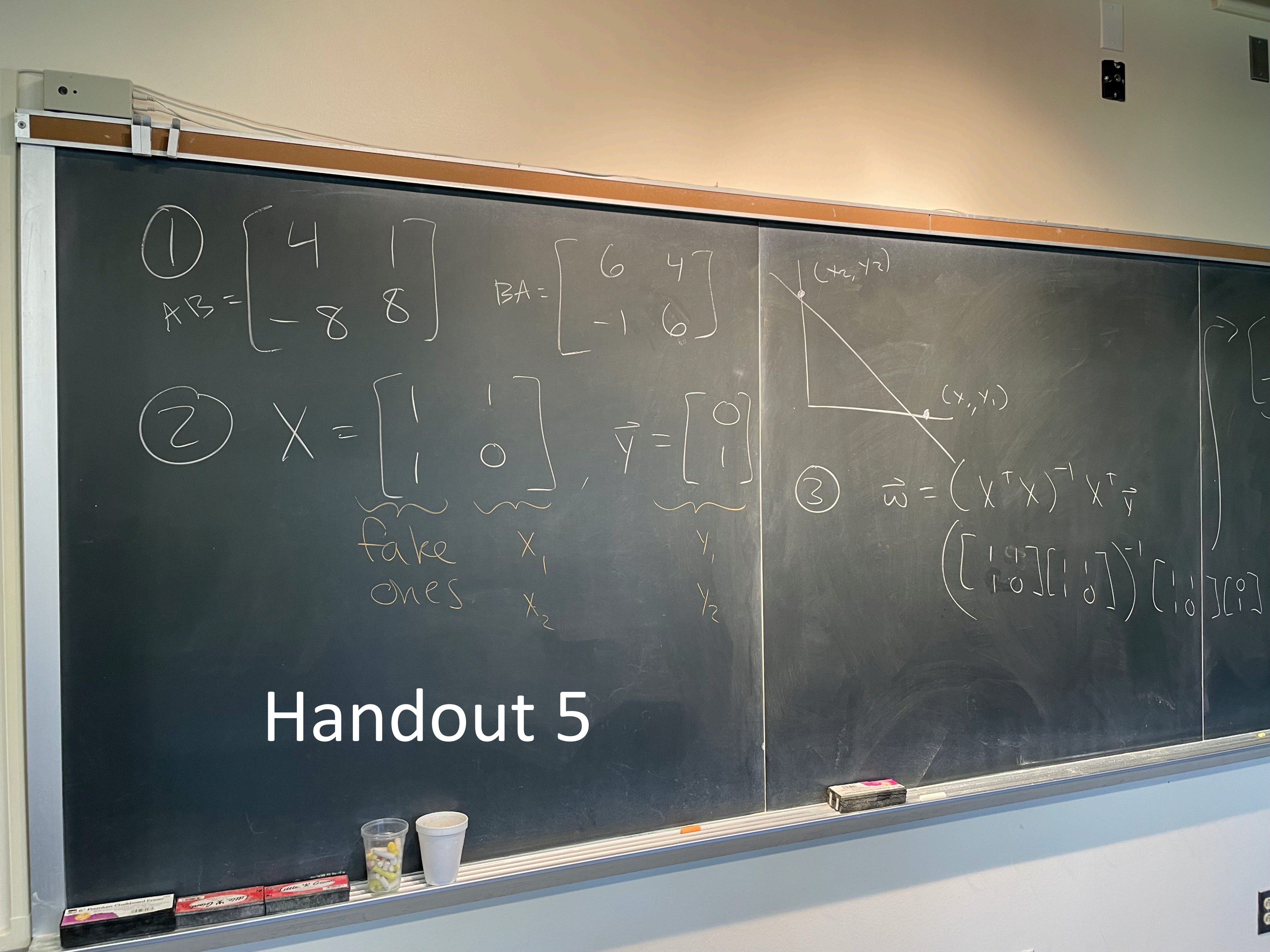
Analytic solution to multiple linear regression

(Keep this formula and its interpretation in mind!)

Handout 5, #1

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

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



① $AB = \begin{bmatrix} 4 & 1 \\ -8 & 8 \end{bmatrix}$ $BA = \begin{bmatrix} 6 & 4 \\ -1 & 6 \end{bmatrix}$

② $X = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, $\vec{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

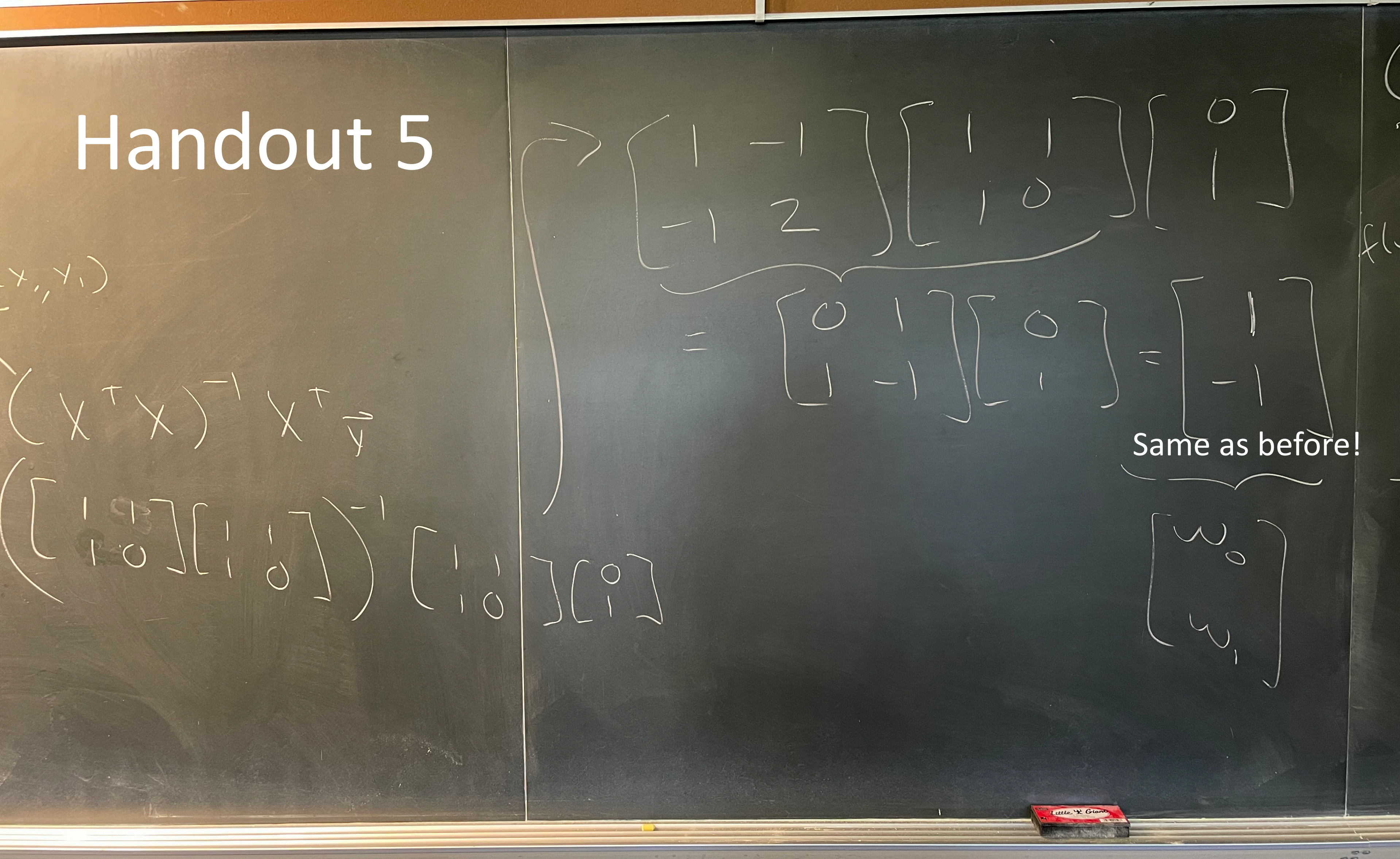
$$X = [\underbrace{\begin{matrix}1\\1\end{matrix}}_{\text{fake ones}} \quad \underbrace{\begin{matrix}1\\0\end{matrix}}_{\begin{matrix}x_1\\x_2\end{matrix}}], \quad \vec{y} = [\underbrace{\begin{matrix}0\\1\end{matrix}}_{\begin{matrix}y_1\\y_2\end{matrix}}]$$

③ $\vec{w} = (X^T X)^{-1} X^T \vec{y}$

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

Handout 5

Handout 5



$x_1, y_1)$

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

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

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

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

$$\underbrace{\text{Same as before!}}_{\begin{bmatrix}w_0\\w_1\end{bmatrix}}$$

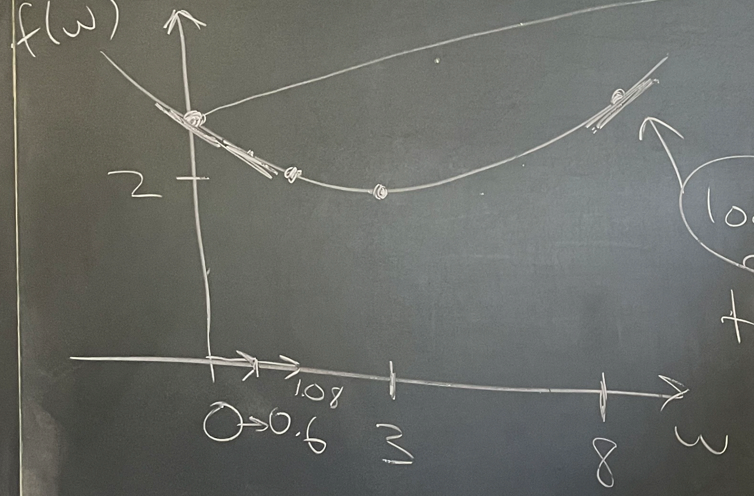
$f($

Begin gradient descent

Gradient Descent

$$f(w) = (w-3)^2 + 2 = w^2 - 6w + 9 + 2$$

$$\boxed{f'(w) = 2w - 6}$$



local derivative

travel in opposite direction of the derivative

$$\underbrace{w = 0}_{\text{startin } w}$$

① $w \leftarrow 0$

$\boxed{w \leftarrow 0.6}$

② $w \leftarrow 0.6$

$\boxed{w \leftarrow 1.08}$

$+ 9 + 2$

$\underbrace{w = 0}_{\text{starting } w}$, $\underbrace{\alpha = 0.1}_{\text{step size}}$

$$\boxed{w \leftarrow w - \alpha f'(w)}$$

gradient descent

① $w \leftarrow 0 \underset{\substack{\uparrow \\ \text{opposite} \\ \text{direction}}}{-} 0.1\underbrace{(2\cdot 0 - 6)}_{\text{derivative}}$

$$\boxed{w \leftarrow 0.6}$$

② $w \leftarrow 0.6 - 0.1(2(0.6) - 6)$

$$\boxed{w \leftarrow 1.08}$$