CS 260: Foundations of Data Science

Prof. Thao Nguyen

Fall 2024

HAVERFORD
COLLEGE

Materials by Sara Mathieson

Admin

- **Lab 3** grades & feedback will hopefully be posted tomorrow (Thursday)
- **Mid-semester feedback form (link on Piazza)**
- **Midterm 1 handed out today in class**
  - Do not open until you are ready to take it!
  - 3 hour time limit
  - Due Wednesday at the beginning of class (Oct 9)

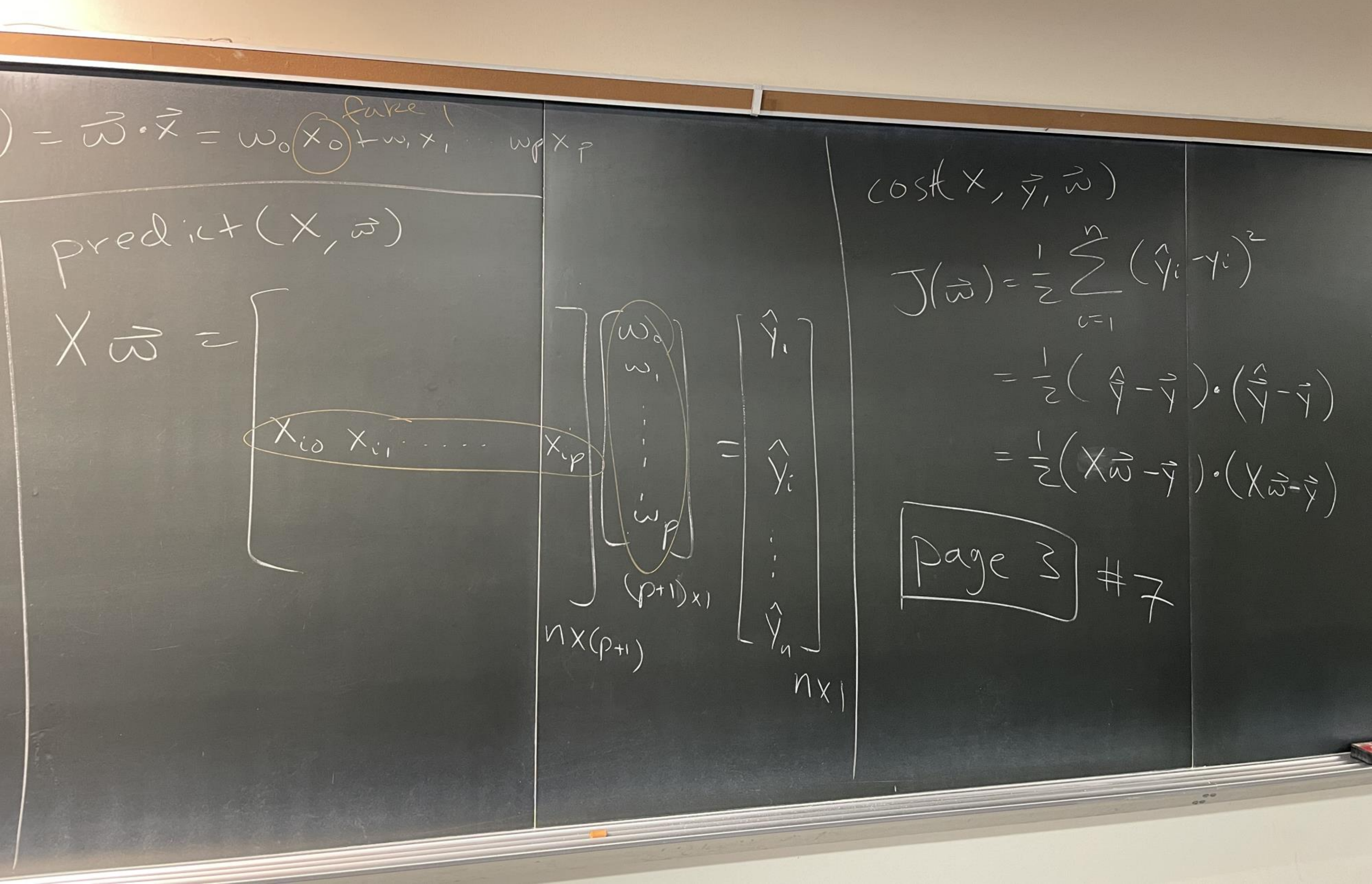
Why do we have a exam?

- Process of synthesizing the material on your own is essential
- Preparing the “study sheet” is designed to facilitate that process

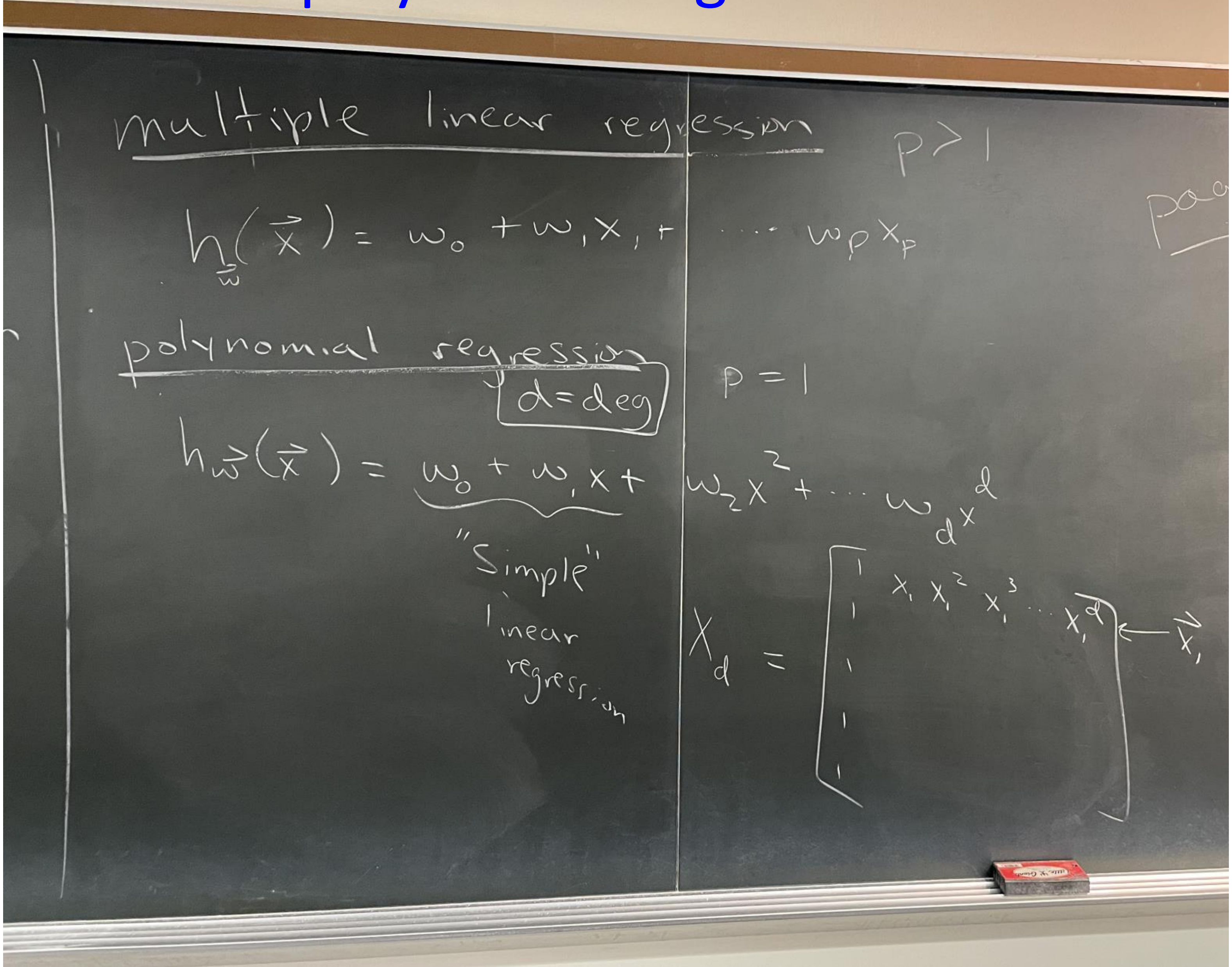
Outline for today

- Review
  - Linear regression
  - Gradient descent
  - Classification
  - Single feature models / decision trees
  - Evaluation metrics

Multiple linear regression



Multiple linear regression vs. polynomial regression



Time varying α $\boxed{\text{SGD}}$

(learning rate) step size

$t = 1$

while not converged:

$\alpha = \frac{1}{t}$

for $i = 1 \ldots n$

$$\underbrace{\vec{w} \leftarrow \vec{w} - \alpha \overbrace{(h_{\vec{w}}(x_i) - y_i)\,\vec{x}_i}^{\text{derivative of } J \text{ wrt } \vec{x}_i}}_{\text{sgd solution to linear regression}}$$

$t += 1$

Matrix / Vector form of SGD

while not converged:

  # shuffle the data

  sort indices 4, 7, 2, 1 ... instead

  for $i = 1, 2, \dots n$:

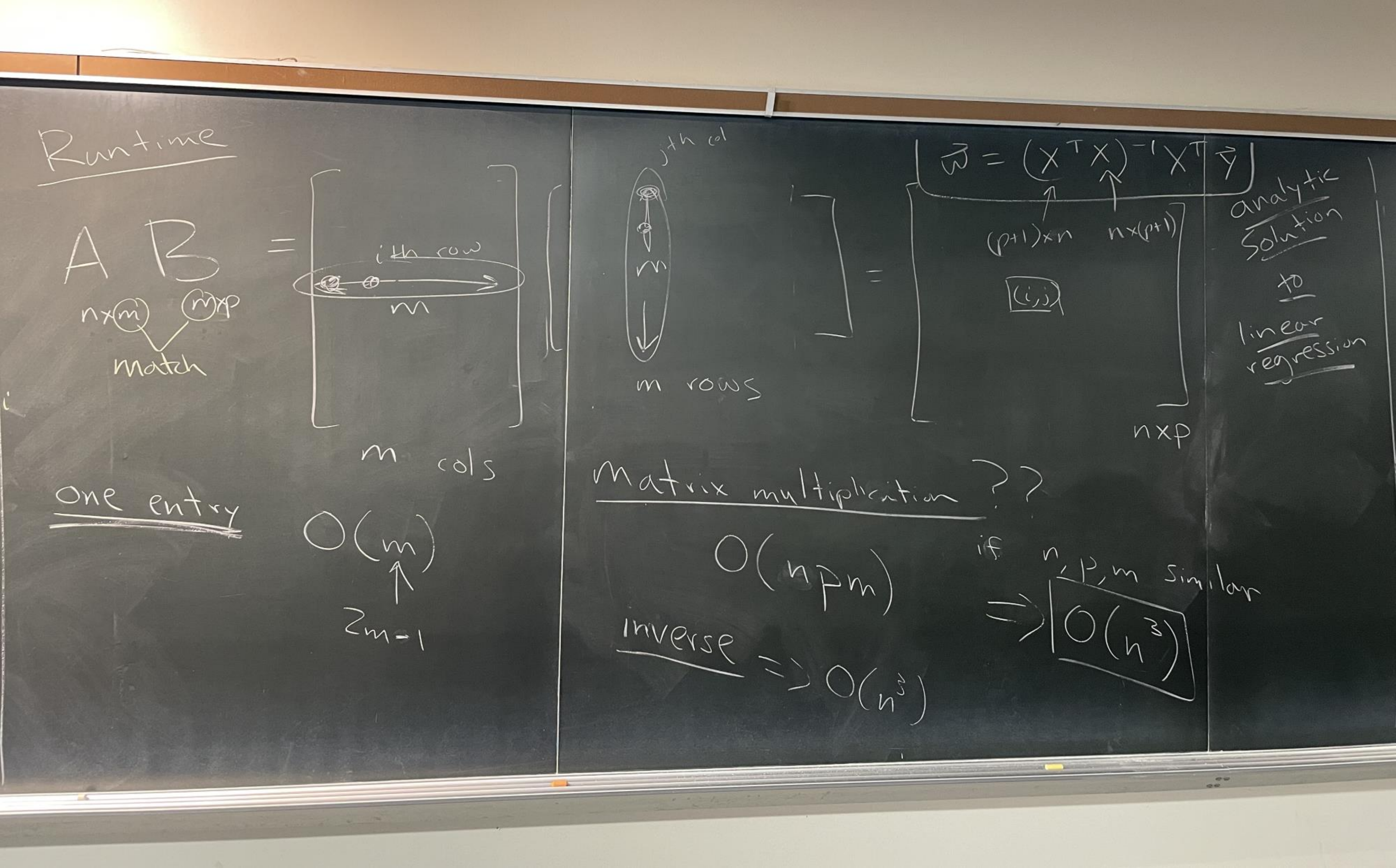
check cost not changing

$$\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_p \end{bmatrix} \leftarrow \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_p \end{bmatrix} - \alpha \left(h_{\vec{w}}(\vec{x}_i) - y_i \right) \begin{bmatrix} x_{i0} \\ x_{i1} \\ \vdots \\ x_{ip} \end{bmatrix}$$

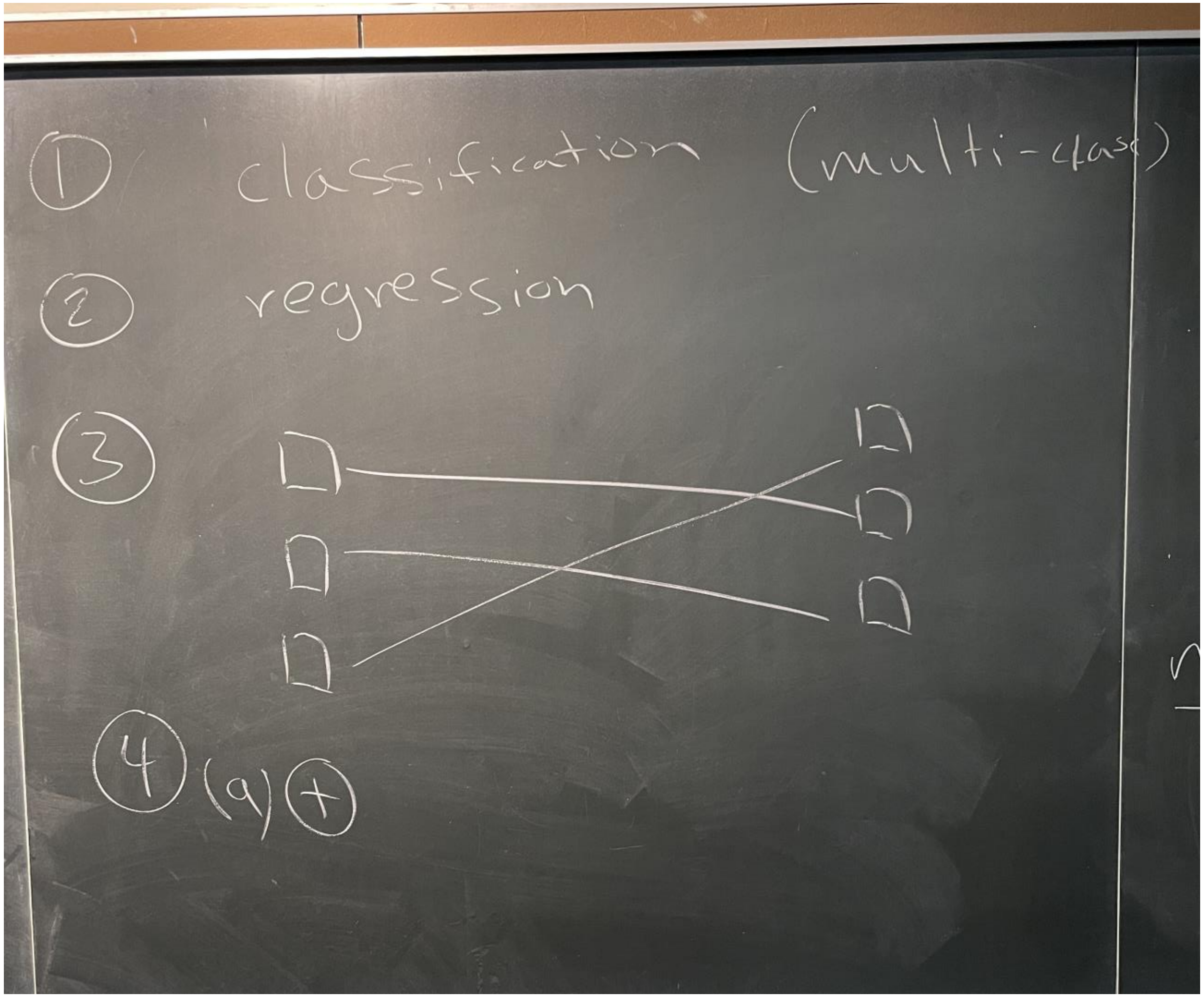
Scalar

$\to h_{\vec{w}}(\vec{x})$

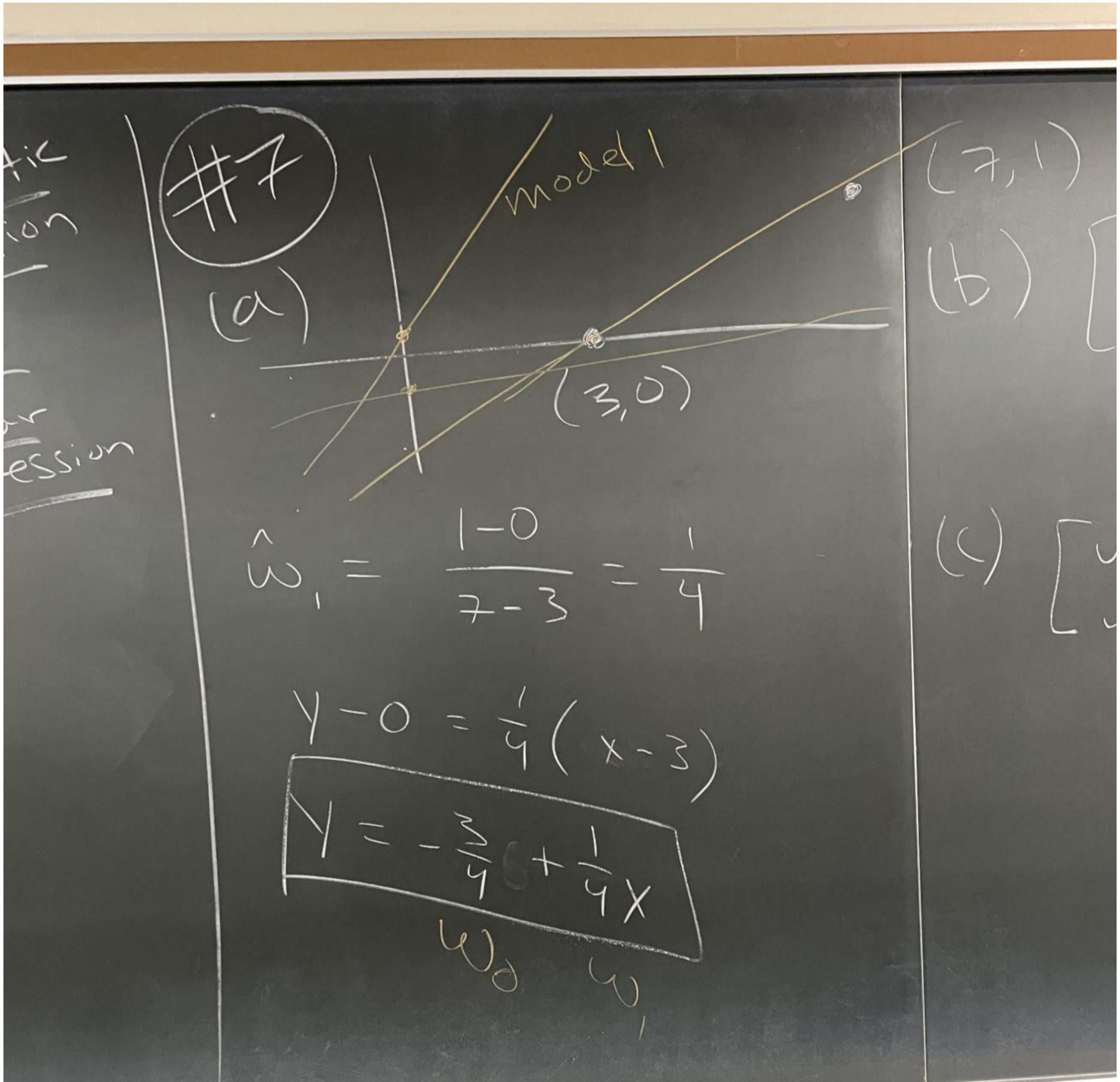
Runtime of matrix operations



Handout 10, page 1



Handout 10, #7


#7
(a)
model 1
(7,1)
(3,0)
(b)
(c)
$\hat{w}_1 = \frac{1-0}{7-3} = \frac{1}{4}$
$y - 0 = \frac{1}{4}(x-3)$
$\boxed{y = -\frac{3}{4} + \frac{1}{4}x}$
w_0
w_1

Handout 10, #7

$(7, 1)$

(b) $\begin{bmatrix} w_0 \\ w_1 \end{bmatrix} \leftarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \alpha \Big(\underbrace{\overbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}^{\vec{w}} \cdot \overbrace{\begin{bmatrix} 1 \\ 7 \end{bmatrix}}^{\vec{x}_2}}_{0\cdot 1 + 0\cdot 7 = 0} - \underset{y_2}{1} \Big) \begin{bmatrix} 1 \\ 7 \end{bmatrix}$

$\leftarrow \begin{bmatrix} 0.1 \\ 0.7 \end{bmatrix}$

$X = \begin{bmatrix} 1 & 3 \\ 1 & 7 \end{bmatrix} \begin{matrix} \to \vec{x}_1 \\ \to \vec{x}_2 \end{matrix}$

(c) $\begin{bmatrix} w_0 \\ w_1 \end{bmatrix} \leftarrow \begin{bmatrix} 0.1 \\ 0.7 \end{bmatrix} - 0.1 \Big(\underbrace{\begin{bmatrix} 0.1 \\ 0.7 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix}}_{\substack{0.1\cdot 1 + 0.7\cdot 3 \\ z_2}} - 0 \Big) \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

$\leftarrow \begin{bmatrix} -0.12 \\ 0.04 \end{bmatrix}$