CS 360: Machine Learning

Sara Mathieson, Sorelle Friedler

Spring 2024

HAVERFORD
COLLEGE

Admin

- **Lab 4** due TONIGHT!
- No office hours today
- **Lab 5** posted tonight
- Review session next week in class (study guide out on Tuesday)

Outline for Feb 22

- Extending logistic regression to multi-class classification (softmax)
- Regularization introduction
- Fairness regularization

Outline for Feb 22

- Extending logistic regression to multi-class classification (softmax)
- Regularization introduction
- Fairness regularization

Sorry, this is s_k!

Multi-class Logistic Regression

$$\hat{y} = \begin{bmatrix} e^{\vec{w}^{(1)} \cdot \vec{x}} \\ e^{\vec{w}^{(2)} \cdot \vec{x}} \\ \vdots \\ e^{\vec{w}^{(K)} \cdot \vec{x}} \end{bmatrix} \Big/ \sum_{k=1}^{K} e^{\vec{w}^{(k)} \cdot \vec{x}}$$

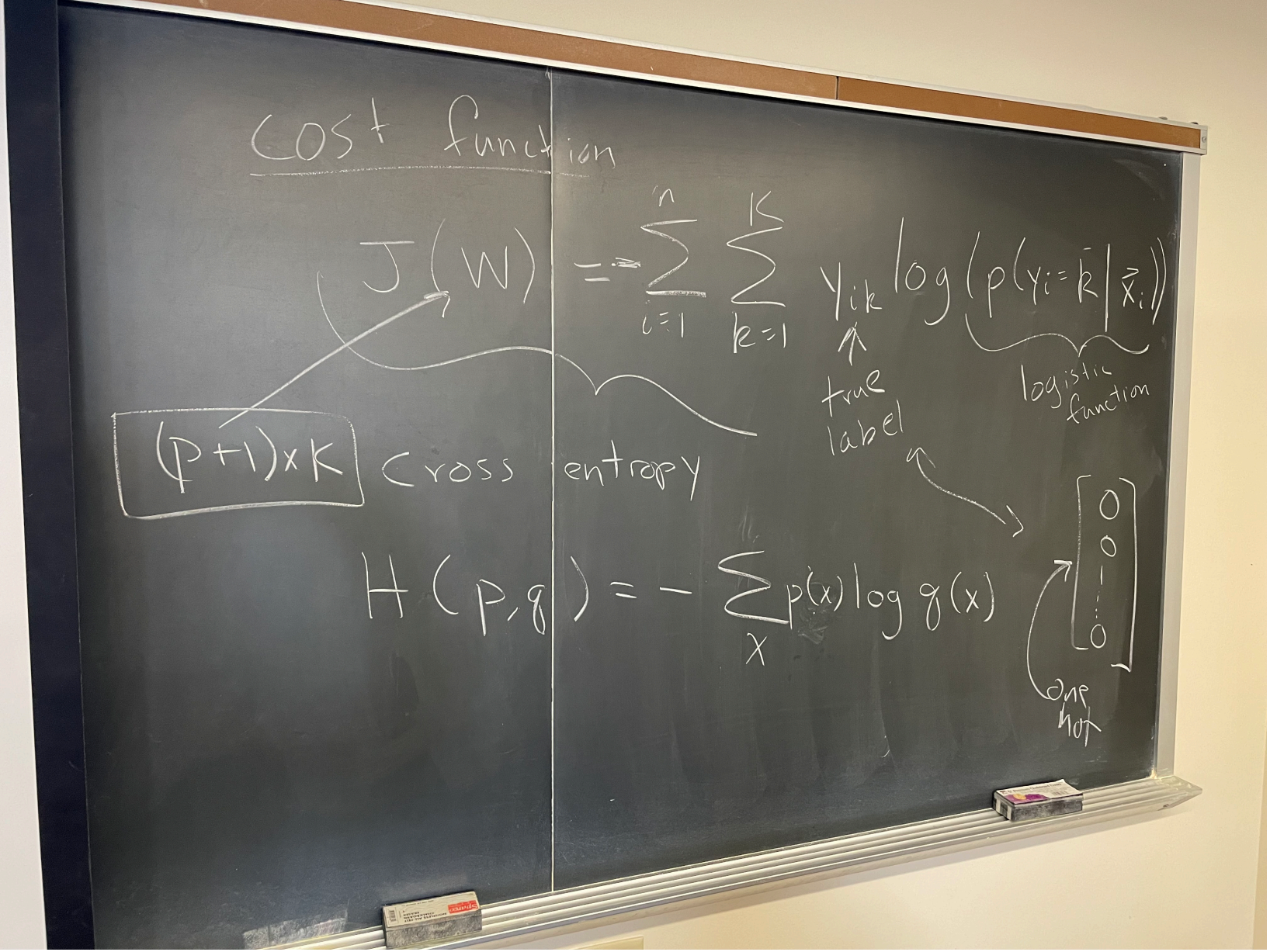
K = # classes

$$s_k = \frac{e^{\vec{w}^{(k)} \cdot \vec{x}}}{\sum_{i=1}^{K} e^{\vec{w}^{(i)} \cdot \vec{x}}} = p(y_i = k \mid \vec{x}_i)$$

$$\begin{matrix} 1 & 2 & 3 \\ [-10, & 2, & 5] \end{matrix} \quad K = 3$$

$$s_2 = \frac{2}{-10+2+5} \qquad s_2 = \frac{e^2}{e^{-10} + e^2 + e^5}$$

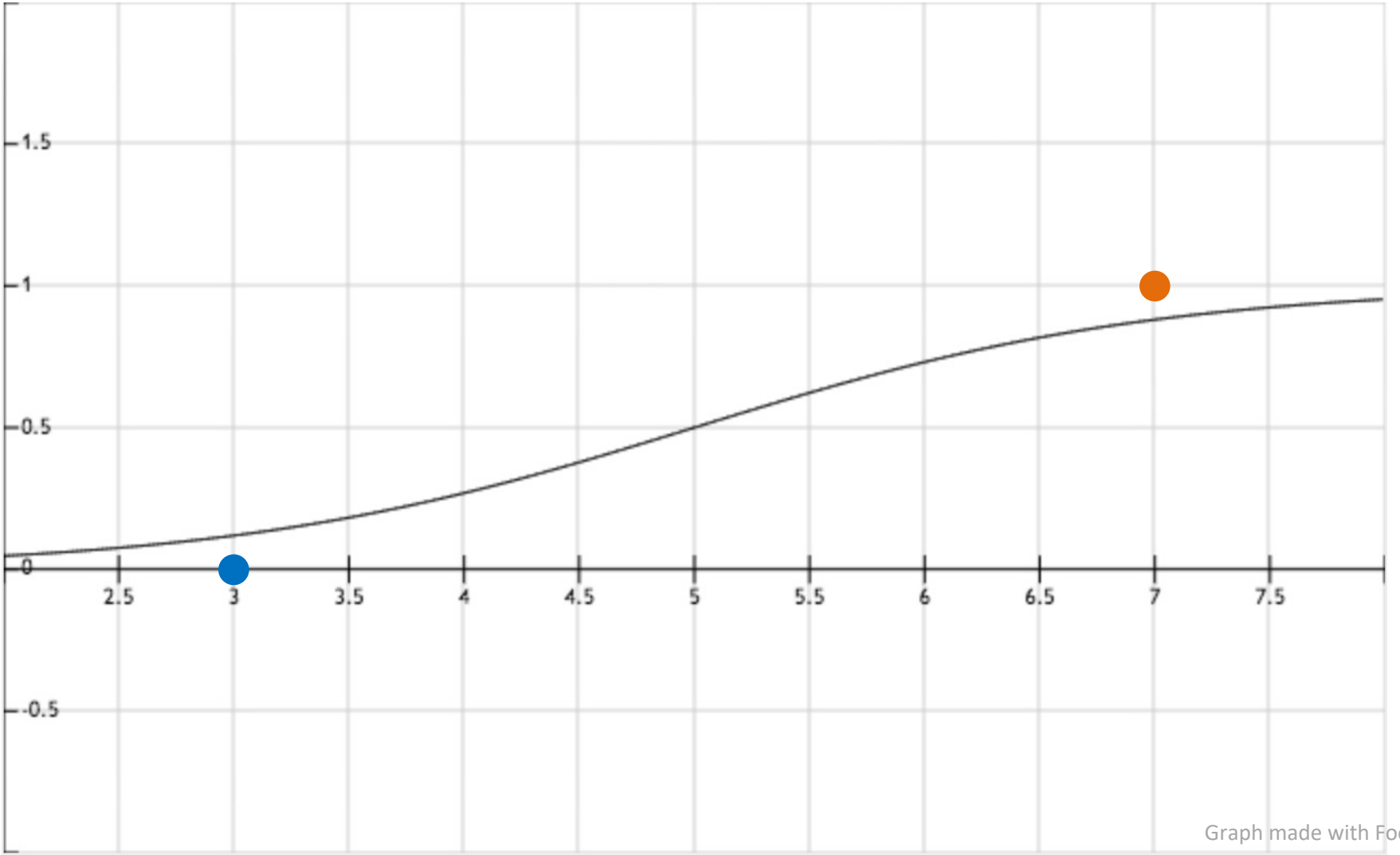
$$s_1 = \frac{-10}{-10+2+5}$$


cost function
$J(W) = -\sum_{i=1}^{n}\sum_{k=1}^{K} y_{ik} \log\left(p(y_i = k \mid \vec{x}_i)\right)$
true label
logistic function
$(p+1)\times K$
cross entropy
$H(p,q) = -\sum_{x} p(x) \log q(x)$
one hot

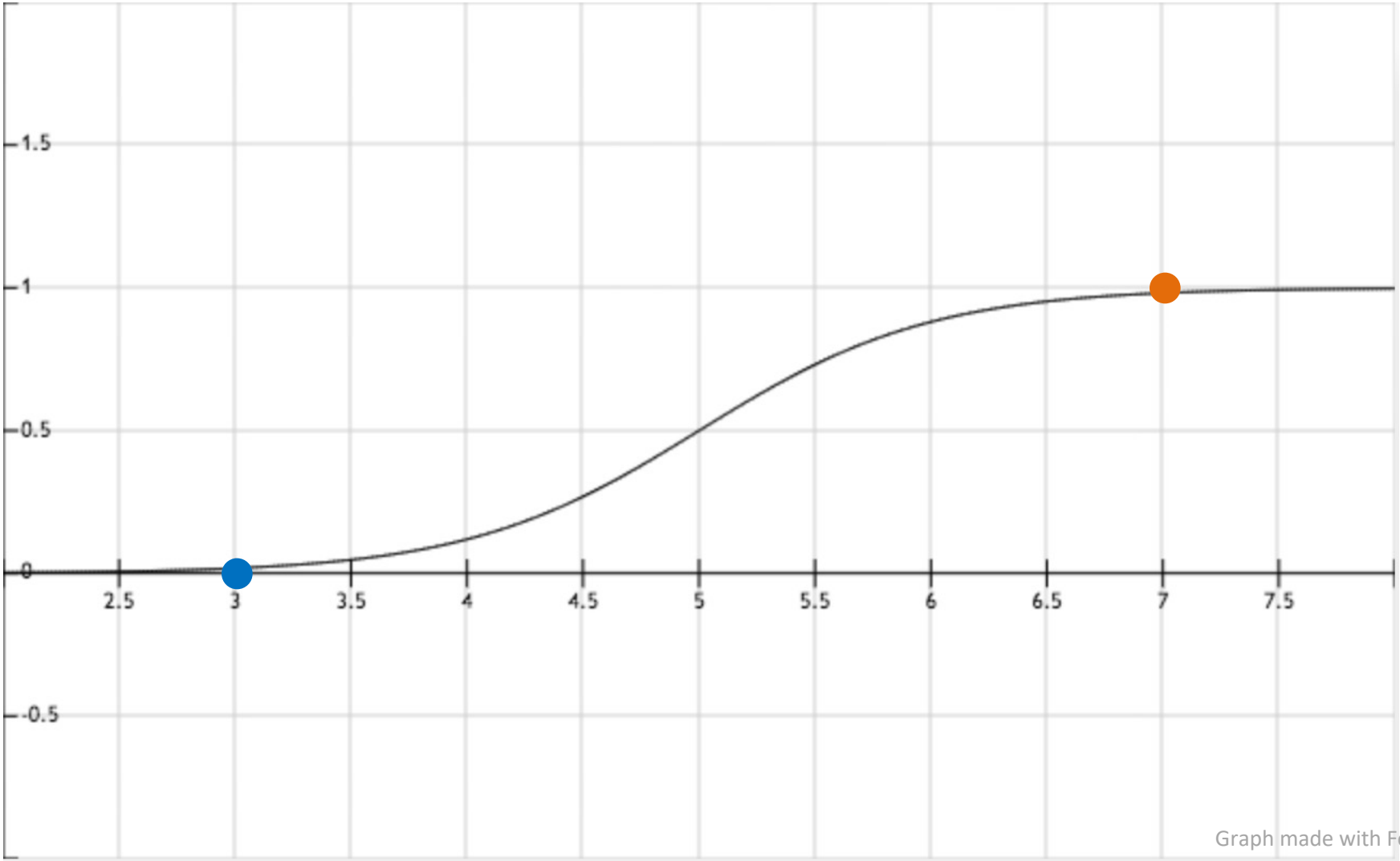
Outline for Feb 22

- Extending logistic regression to multi-class classification (softmax)
- Regularization introduction
- Fairness regularization

w_0=-5, w_1=1 $h_w(x) = 1/(1+e^{(5-x)})$

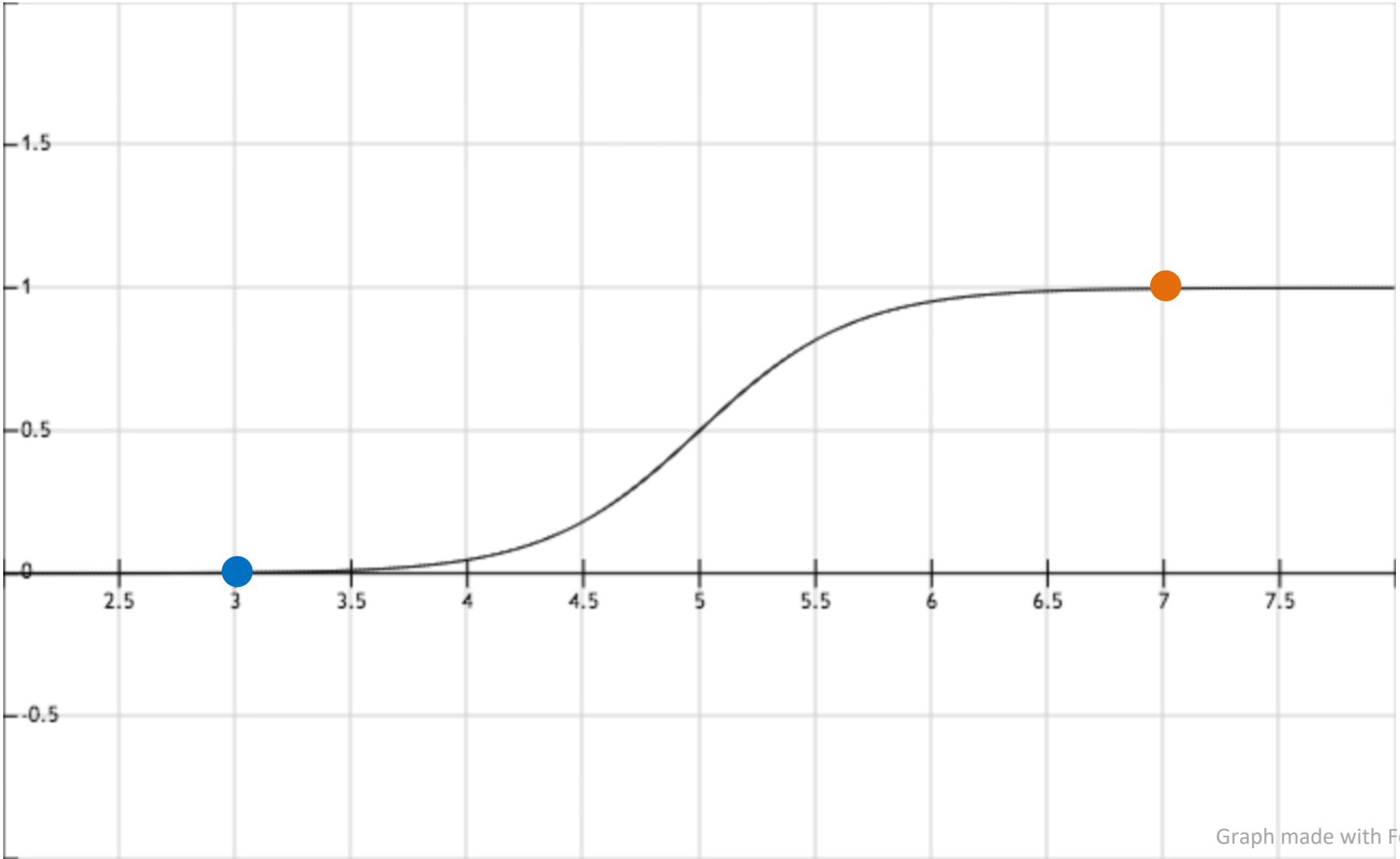


w_0=-10, w_1=2 $h_w(x) = 1/(1+e^{(10-2x)})$

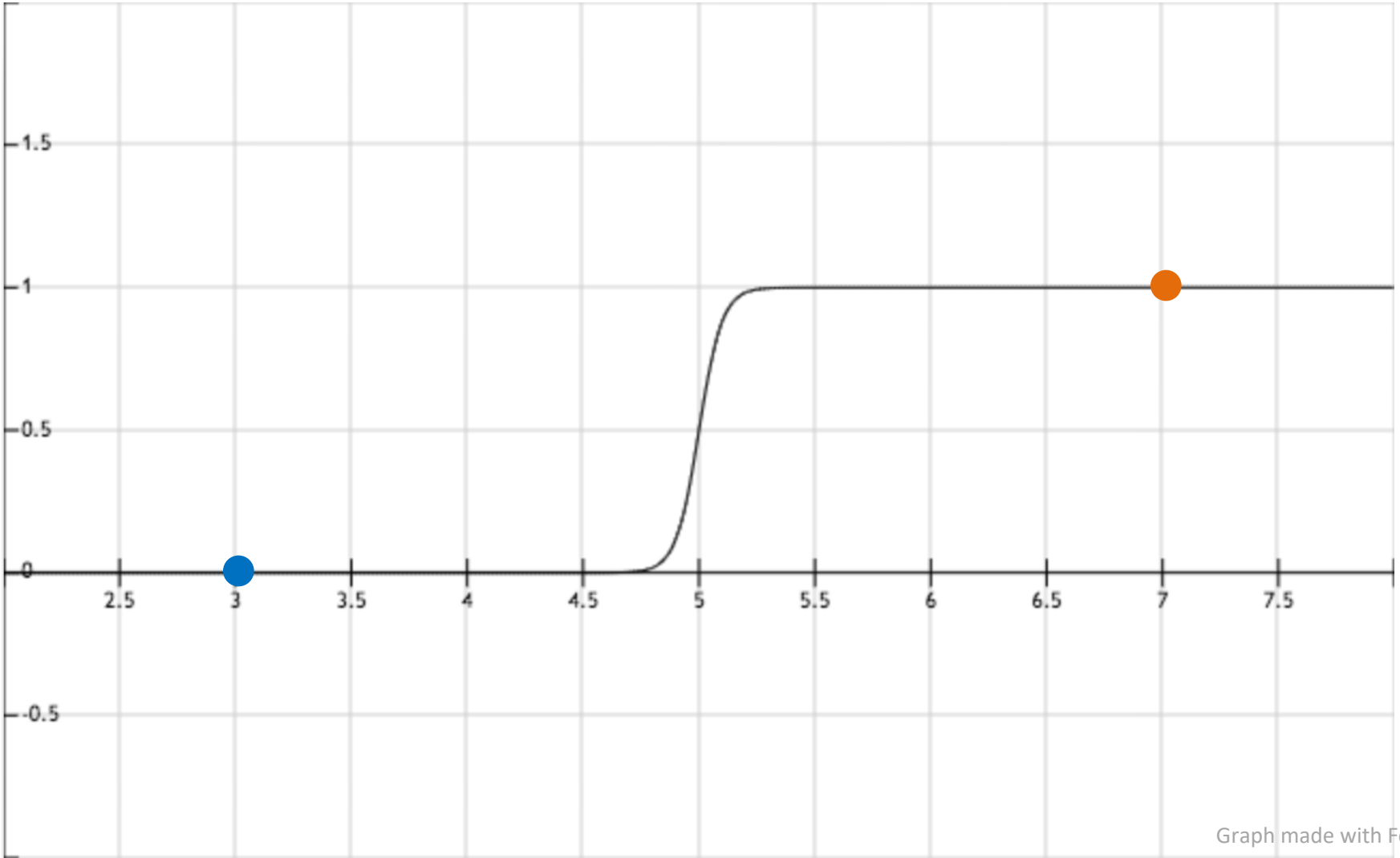


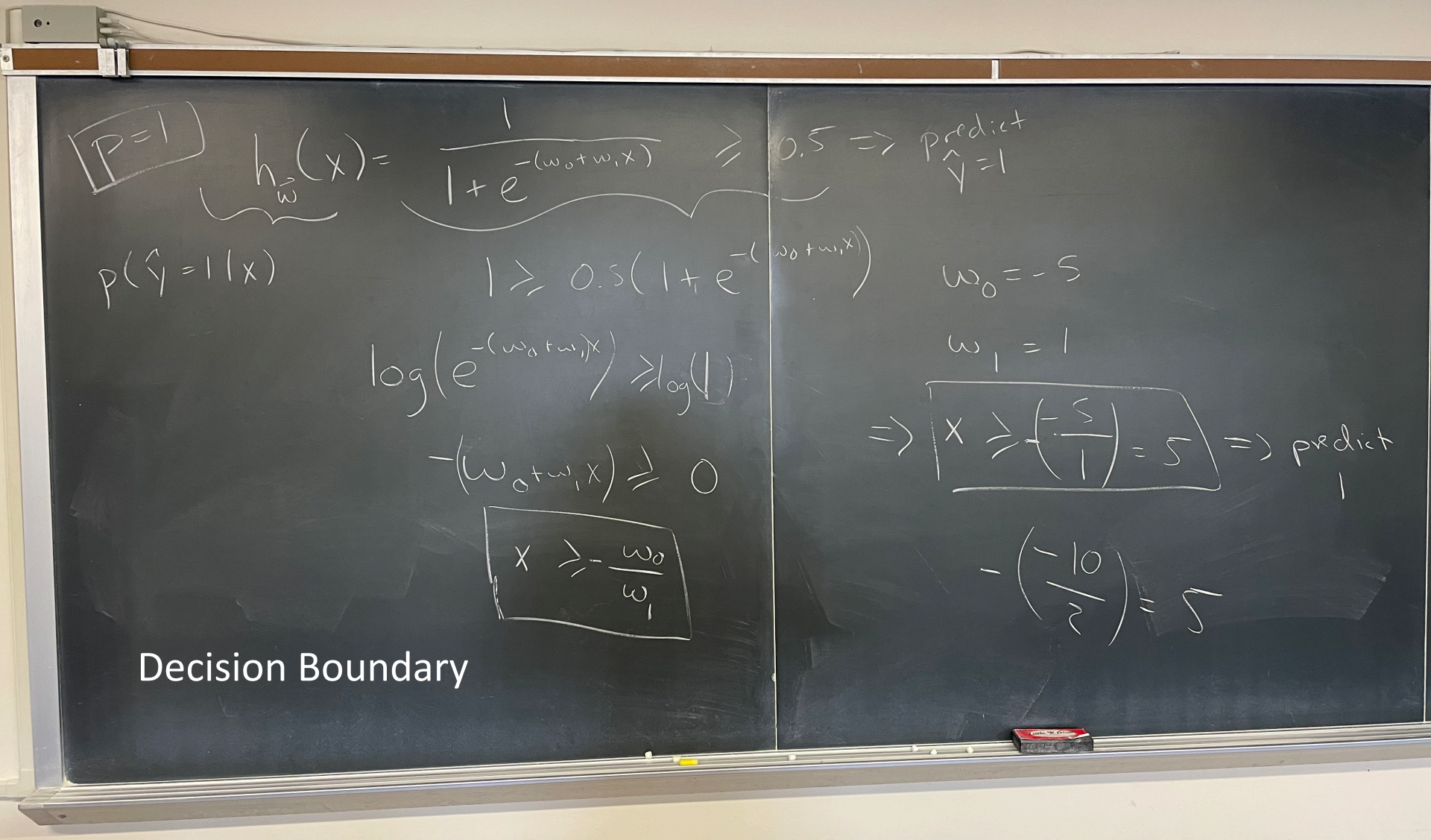
Graph made with FooPlot

w_0=-15, w_1=3 $h_w(x) = 1/(1+e^{(15-3x)})$



w_0=-100, w_1=20 $h_w(x)$ = 1/(1+e^(100-20x))




$p=1$
$h_{\vec{\omega}}(x) = \frac{1}{1+e^{-(\omega_0+\omega_1 x)}} \geq 0.5 \Rightarrow$ predict $\hat{y}=1$
$p(\hat{y}=1|x)$
$1 \geq 0.5(1+e^{-(\omega_0+\omega_1 x)})$
$\log(e^{-(\omega_0+\omega_1 x)}) \geq \log(1)$
$-(\omega_0+\omega_1 x) \geq 0$
$x \geq -\frac{\omega_0}{\omega_1}$
$\omega_0 = -5$
$\omega_1 = 1$
$\Rightarrow x \geq -\left(\frac{-5}{1}\right) = 5 \Rightarrow$ predict 1
$-\left(\frac{-10}{2}\right) = 5$
Decision Boundary

Regularization w's small

$$J^R(\vec{w}) = \left[-\sum_{i=1}^{n} y_i \log h_{\vec{w}}(\vec{x}_i) + (1-y_i) \log \left(1 - h_{\vec{w}}(\vec{x}_i)\right) \right] + \frac{\lambda}{2n} \sum_{j=1}^{p} w_j^2$$

regularization parameter

incentivizes small weights

$0 < \lambda < 1$

ex: $\lambda = 0.02$

SGD

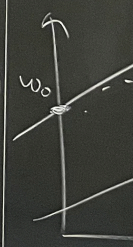
$$\nabla J^R_{\vec{x}_i}(\vec{w}) = \left(h_w(\vec{x}_i) - y_i\right) \vec{x}_i + \lambda \vec{w}^*$$

$p+1$ $p+1$

$\vec{w}^* = [0, w_1, w_2 \ldots w_p]$

$\vec{w} = [w_0, w_1, w_2 \ldots w_p]$

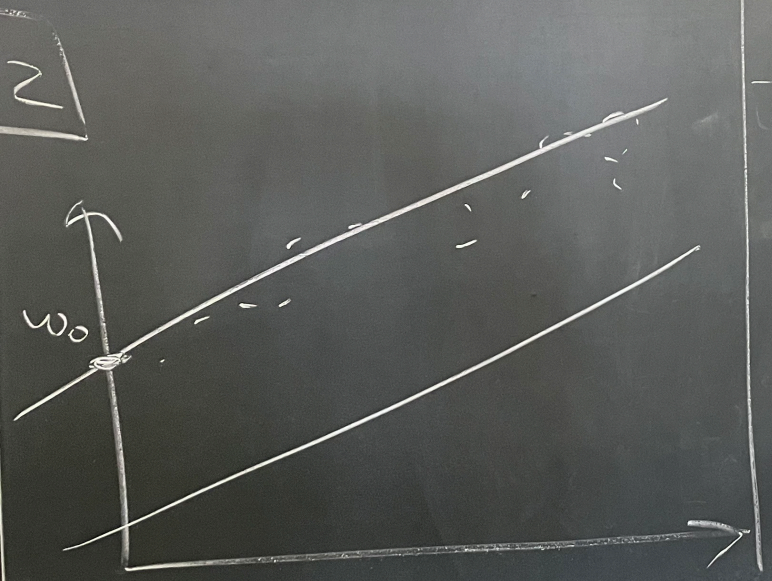
not regularizing



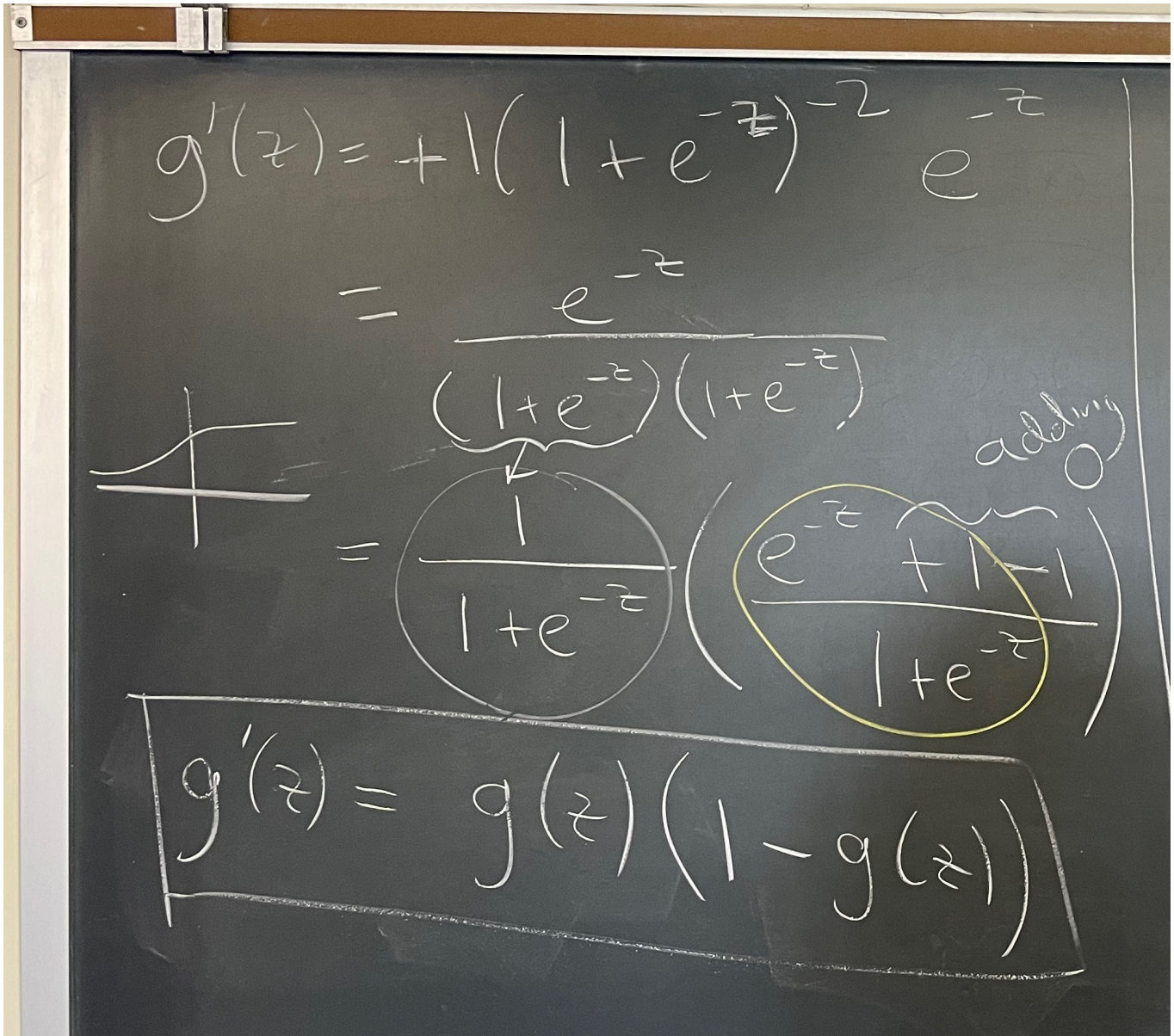
$$\vec{w} \leftarrow \vec{w} - \alpha\left[\left(h_{\vec{w}}(\vec{x}_i) - y_i\right)\vec{x}_i + \lambda\vec{w}^*\right]$$

$$\leftarrow \left(1 - \underbrace{\alpha\lambda}_{\substack{\text{small} \\ \& \text{ positive}}}\right)\vec{w} + \cdots\cdots$$

Handout 10
Q1



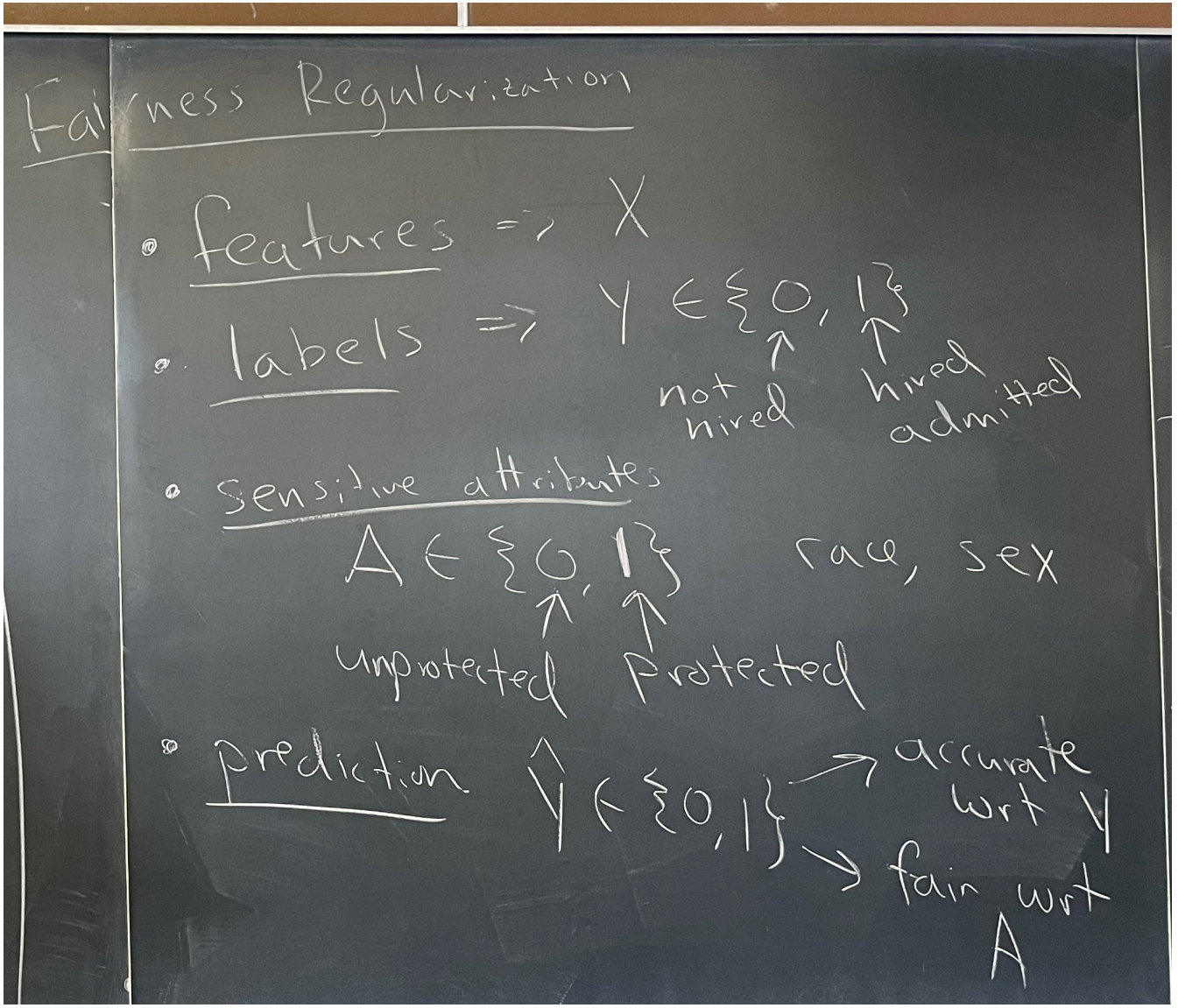
Handout 10, Q1



Outline for Feb 22

- Extending logistic regression to multi-class classification (softmax)
- Regularization introduction
- Fairness regularization

Fairness Regularization setup



demographic parity

$$\frac{P(\hat{Y}=1 \mid A=1)}{P(\hat{Y}=1 \mid A=0)} \approx 1$$

goal (fairness)

equalized odds

$$\frac{P(\hat{Y}=1 \mid A=1, Y=y)}{P(\hat{Y}=1 \mid A=0, Y=y)} \approx 1 \quad \text{for } y \in \{0,1\}$$

goal

$Y=0 \Rightarrow$ FPR per demographic

$Y=1 \Rightarrow$ TPR " "

regularization term (demographic parity)

$$R(h_w, D) = 1 - P(\hat{Y}=1 \mid A=1)$$

h_w: log reg; D: data (train)

want small

$$J^R(w) = \text{cost} + R(h_w, D)$$

cost: accuracy; R: fairness

$$1 = \frac{1}{|D_0|} \sum_{\vec{x} \in D_0} p(\hat{y}=1 \mid \vec{x})$$

given $A=1$

$p(\hat{y}=1)$

D_0 = all examples in the protected group ($A=1$)

$$1 - \frac{1}{|D_0|} \sum_{\vec{x} \in D_0} h_{\vec{w}}(\vec{x})$$

$$h_{\vec{w}}(\vec{x}) = \frac{1}{1+e^{-\vec{w}\cdot\vec{x}}}$$

$$w \leftarrow \vec{w} - \eta \left[(h-y)\vec{x} - \frac{1}{|D_0|} h(1-h)\vec{x} \right]$$

if $A=1$ for $\vec{x}_i$

Handout 10 Q2&3

gradient

$$\nabla h_{\vec{w}}(\vec{x}) = h_{\vec{w}}(\vec{x})\left(1 - h_{\vec{w}}(\vec{x})\right)\vec{x}$$

SGD → add this on! to the gradient

Handout 10, Q2 & Q3