

CS 260: Foundations of Data Science

Prof. Thao Nguyen

Fall 2025



HAVERFORD
COLLEGE

Admin

- Lab 4 due tonight



Admin

- **Mid-semester feedback form** (link on Piazza)
- **Midterm 1:** in class 10/07

Midterm 1 Notes

- You may use one letter page (front and back) “study sheet”, handwritten, created by you
- You may also use a regular calculator
- Outside of your “study sheet” and calculator, **no other notes or resources**
- **As per the Honor Code, all work must be your own**

Why do we have exams?

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

Outline for today

- Go over Lab 2
- Review
 - Linear regression
 - Gradient descent
 - Classification
 - Single feature models / decision trees
 - Evaluation metrics

Outline for today

- Go over Lab 2
- Review
 - Linear regression
 - Gradient descent
 - Classification
 - Single feature models / decision trees
 - Evaluation metrics

Lab 2: not posted online

Outline for today

- Go over Lab 2
- Review
 - Linear regression
 - Gradient descent
 - Classification
 - Single feature models / decision trees
 - Evaluation metrics

Multiple linear regression

model $h_{\vec{w}}(\vec{x}) = \sum_{j=0}^p w_j x_j = \underbrace{w_0 x_0}_{\text{intercept}} + w_1 x_1 + \dots + w_p x_p = \vec{w} \cdot \vec{x}$

predict
 $(X, \vec{w}) \rightarrow X \vec{w} = \hat{\vec{y}}$
 matrix mult
 "any" regression problem

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

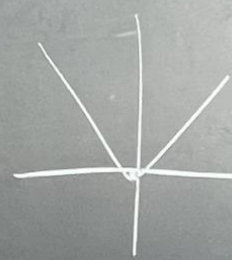
$$= \frac{1}{2} (\vec{y} - \hat{\vec{y}}) \cdot (\vec{y} - \hat{\vec{y}})$$

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

$$\begin{bmatrix} x_{i0} & x_{i1} & \dots & x_{ip} \end{bmatrix}^T \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_p \end{bmatrix} = \begin{bmatrix} \vec{x}_i \cdot \vec{w} \end{bmatrix} = \begin{bmatrix} \hat{y}_i \end{bmatrix}$$

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

~~$J(\vec{w}) = \sum |y_i - \hat{y}_i|$~~



Multiple linear regression vs. polynomial regression

multiple linear regression

$p > 1$

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

polynomial regression
 $d = \text{deg}$

$p = 1$

$$h_{\vec{w}}(\vec{x}) = \underbrace{w_0 + w_1 x_1}_{\text{"Simple" linear regression}} + w_2 x^2 + \dots + w_d x^d$$

$$X_d = \begin{bmatrix} 1 \\ x_1 \\ x_1^2 \\ x_1^3 \\ \vdots \\ x_1^d \end{bmatrix} \leftarrow \vec{x}_1$$

SGD with time varying α

Time varying α (learning rate) step size SGD

$t = 1$

while not converged:

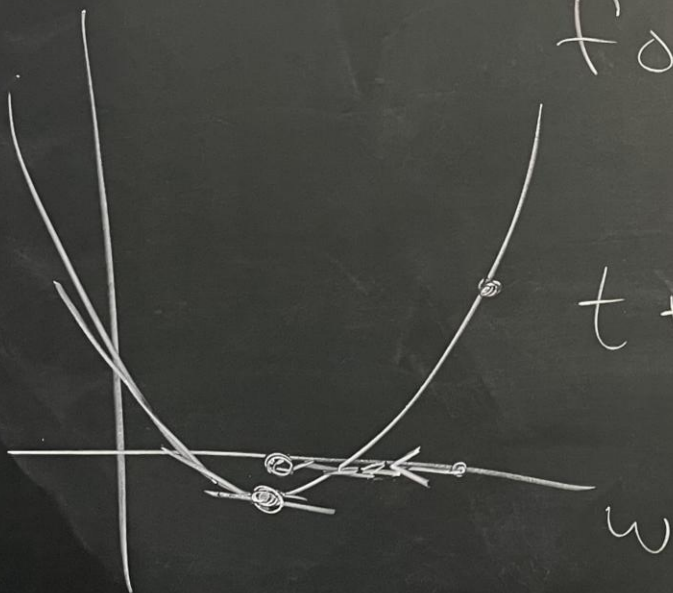
$$\alpha = \frac{1}{t}$$

for $i = 1 \dots n$

$$\vec{w} \leftarrow \vec{w} - \alpha (h_{\vec{w}}(x_i) - y_i) \vec{x}_i$$

$t += 1$

$J(w)$



derivative
of J
wrt \vec{x}_i

SGD solution
to linear regression

Matrix/Vector form of SGD

Matrix/Vector form of SGD

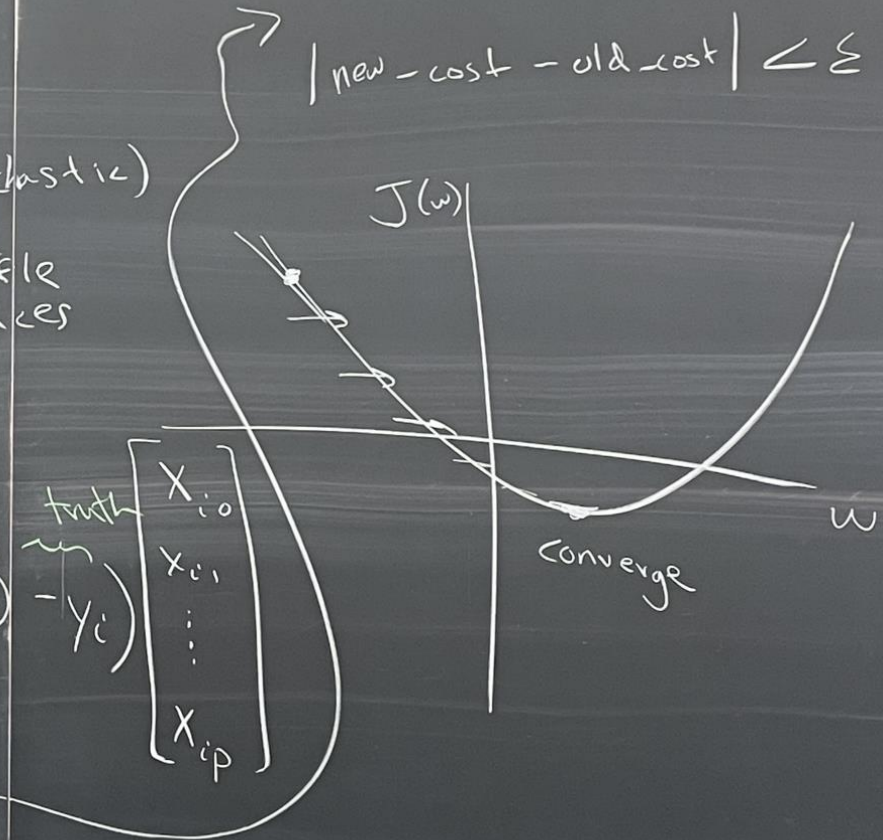
while not converged:
shuffle the data (stochastic)

4, 7, 2, 10 ... ← shuffle indices
for $i=1, 2, 3 \dots n$

$$\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_p \end{bmatrix} \leftarrow \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_p \end{bmatrix} - \alpha \left(\overbrace{h_{\vec{w}}(\vec{x}_i)}^{\text{pred}} - \underbrace{y_i}_{\text{truth}} \right)$$

↑
step size
learning rate

test for convergence



Runtime of matrix operations

- **Multiplication**

$$\begin{matrix} A \\ \uparrow \\ n \times p \end{matrix} \begin{matrix} B \\ \uparrow \\ p \times m \end{matrix} = \begin{bmatrix} \text{ith row} \end{bmatrix} \begin{bmatrix} \text{jth column} \end{bmatrix} = \begin{bmatrix} \text{i,j entry} \end{bmatrix} \quad n \times m$$

entries: $O(nm)$
one entry: $O(p)$ } $O(nmp)$

- **Inverse**

Matrix must be square, $(n \times n) \rightarrow O(n^3)$

Analytic solution to linear regression

$$\textcircled{11} \quad \vec{w} = \underbrace{\left(\underbrace{X^T X}_{(d)} \right)^{-1}}_{(b)} \underbrace{X^T \vec{y}}_{(c)}$$

(p+1) x (p+1) + (p+1) x n

(c) $X^T \Rightarrow O(np)$

(a) multiply (p+1) x n + n x (p+1) $\Rightarrow O(p^2 n)$

(b) invert (p+1) x (p+1) matrix $\Rightarrow O(p^3)$
inverse is cubic

(c) multiply (p+1) x n + (n x 1) $\Rightarrow O(pn)$

(d) multiply (p+1) x (p+1) + (p+1) x 1 $\Rightarrow O(p^2)$

$\Rightarrow O(p^2 n + p^3 + \cancel{pn + p^2}) \Rightarrow \boxed{O(p^2 n + p^3)} \star$