CS 260: Foundations of Data Science

Prof. Thao Nguyen

Fall 2025

HAVERFORD COLLEGE

Materials by Sara Mathieson

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

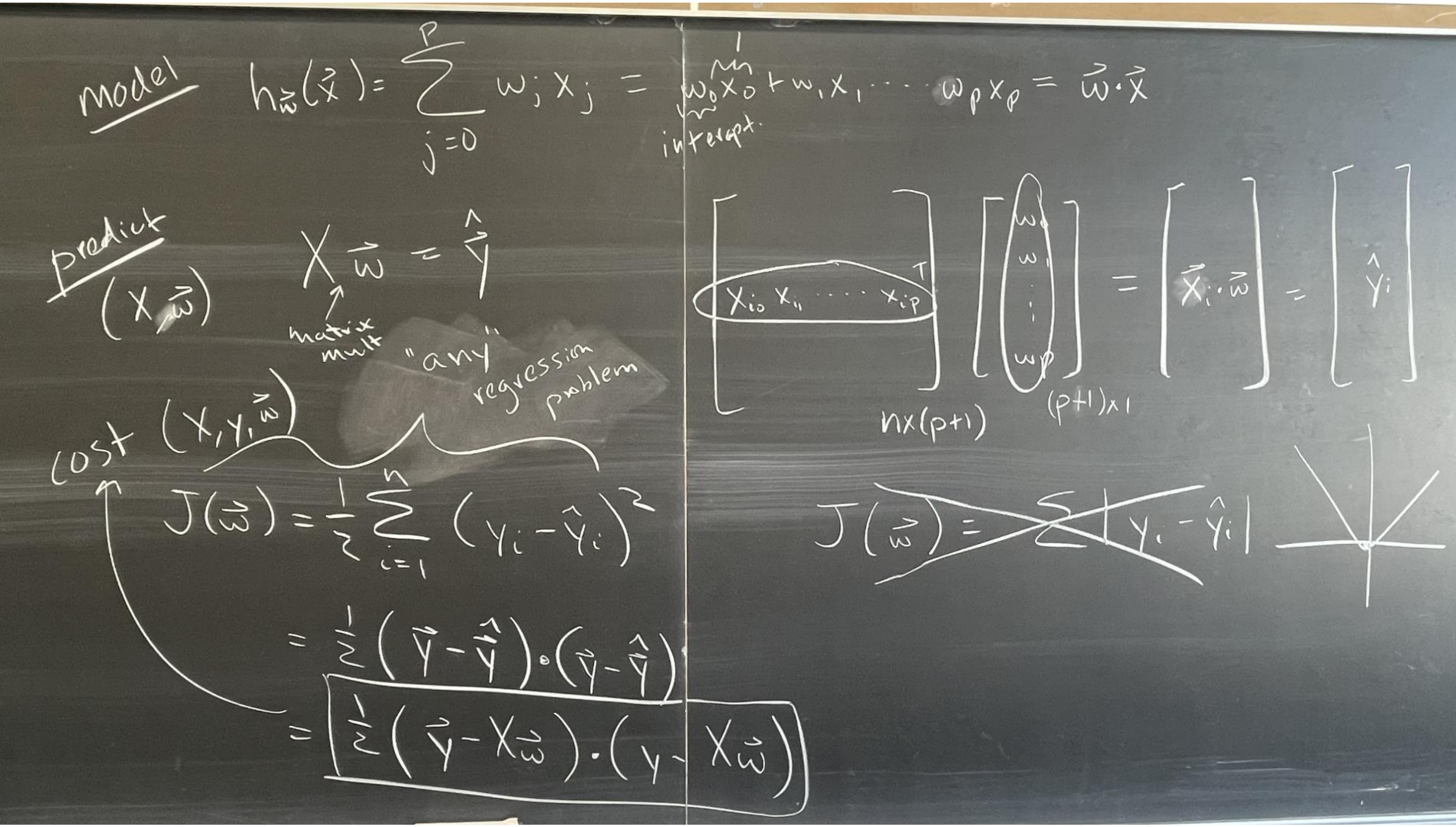
Lab 2: not posted online

Outline for today

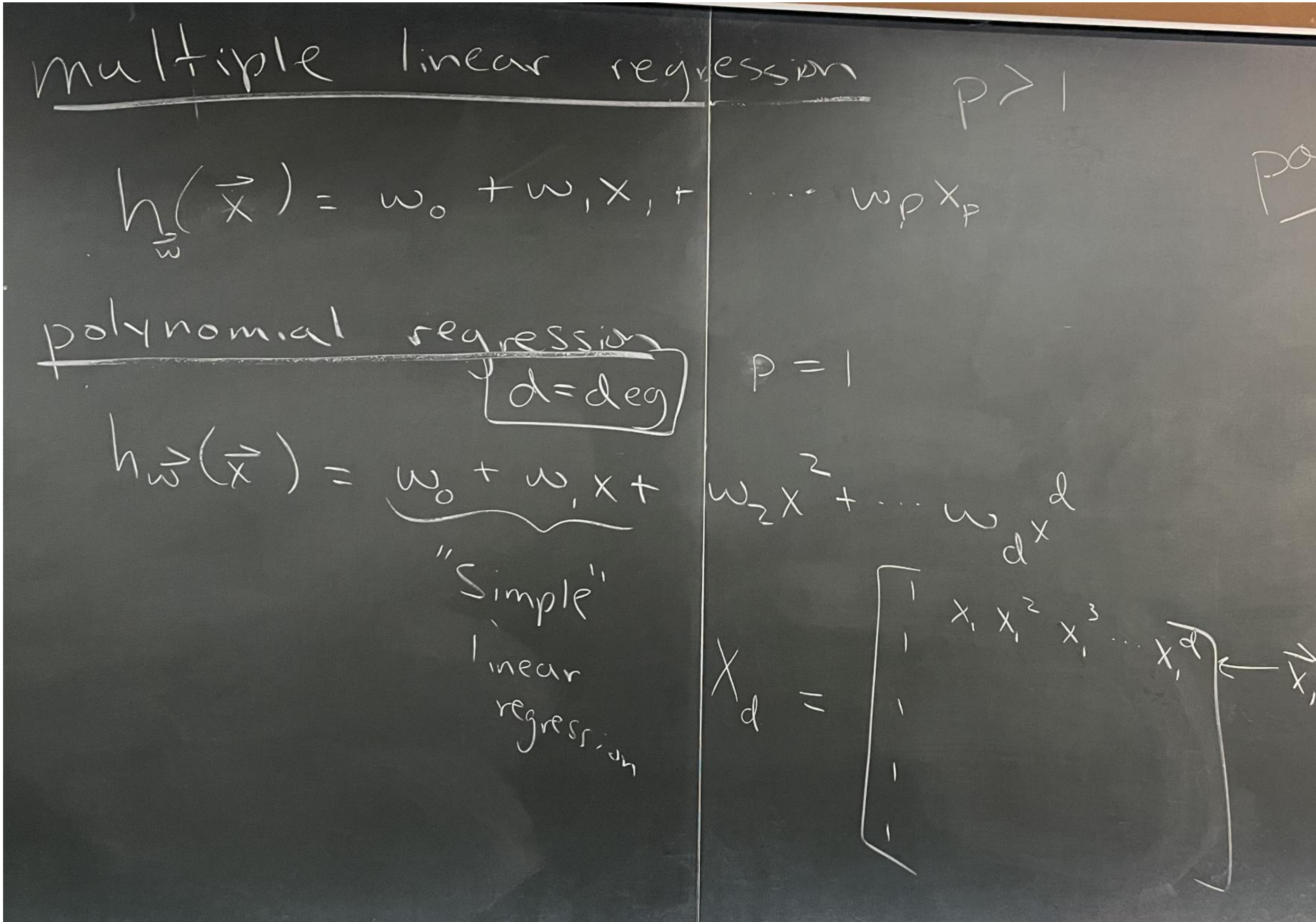
- Go over Lab 2

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

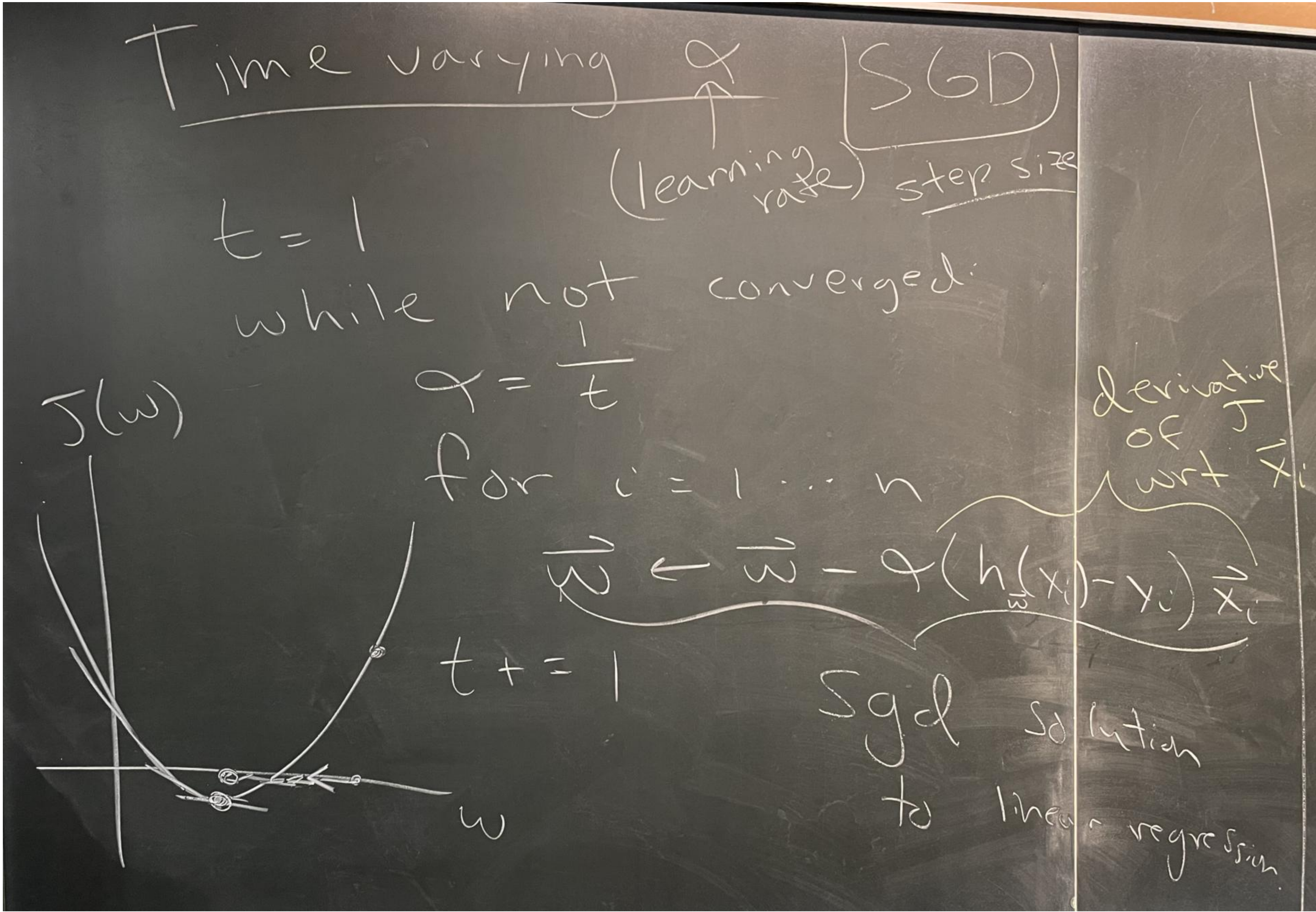
Multiple linear regression



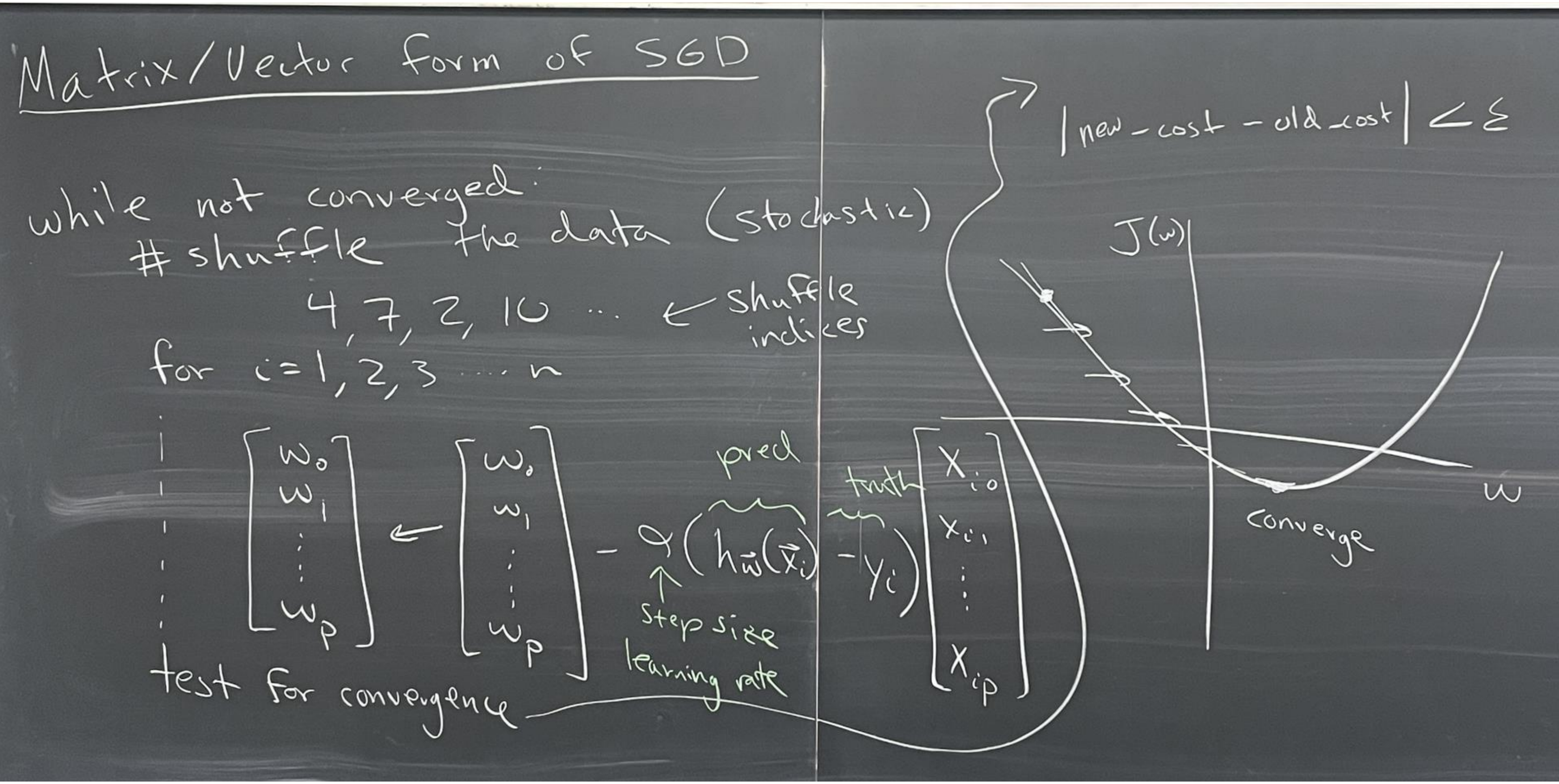
Multiple linear regression vs. polynomial regression



SGD with time varying α

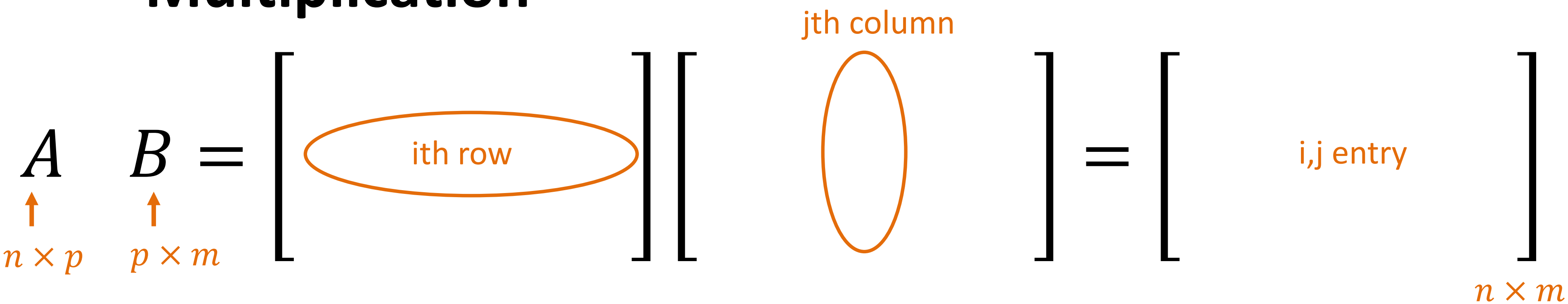


Matrix/Vector form of SGD



Runtime of matrix operations

- **Multiplication**



\# 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