# CS 260: Foundations of Data Science

Prof. Thao Nguyen

Fall 2025

HAVERFORD COLLEGE

Materials by Sara Mathieson

# Admin

- **Sit somewhere new**
- **Lab 2** is due tonight at midnight
- **Lab 3** posted (due next Tuesday)

# Outline for today

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

# Outline for today

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

# Simple linear regression

- Model: $h_{\vec{w}}(x) = w_0 + w_1 x = \hat{y}$
- Cost function:

$$J(w_0, w_1) = \frac{1}{2}\sum_{i=1}^{n}(y_i - w_0 - w_1 x_i)^2$$

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

# Outline for today

- 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

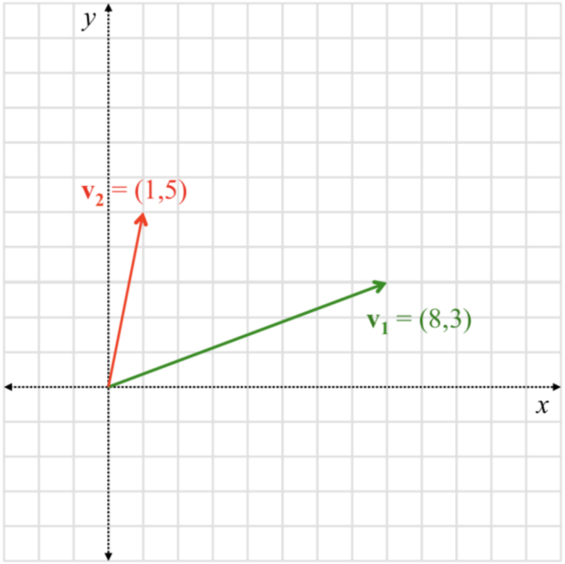
$$\mathbf{v} = \left[ \begin{array}{c} x \\ y \end{array} \right], \quad \text{then} \quad |\mathbf{v}| = \sqrt{x^2 + y^2}.$$

- Different ways to write a vector

$$\mathbf{v} = \left[ \begin{array}{c} x \\ y \end{array} \right] = \left[ \begin{array}{cc} x & y \end{array} \right]^{\mathrm{T}}$$

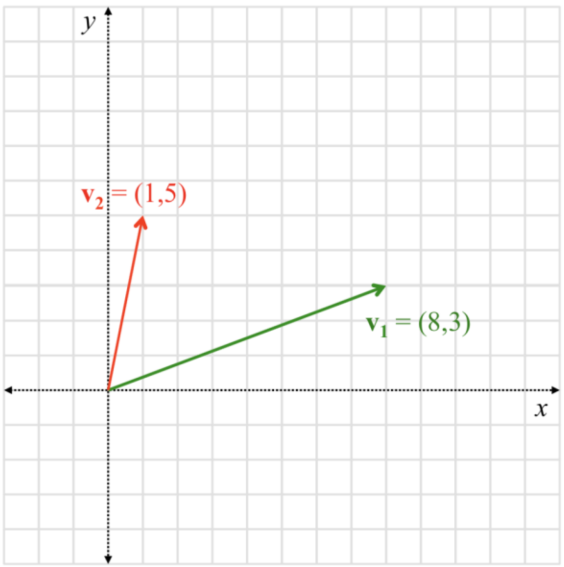
# Vector Addition

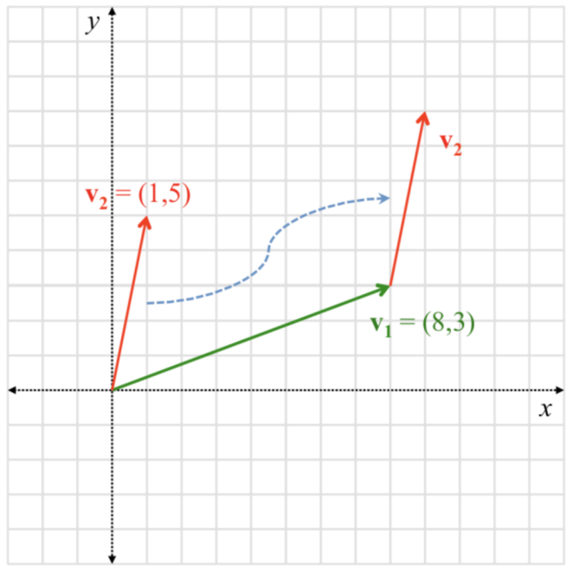
$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{bmatrix} 8 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 8+1 \\ 3+5 \end{bmatrix} = \begin{bmatrix} 9 \\ 8 \end{bmatrix}$$



# Vector Addition

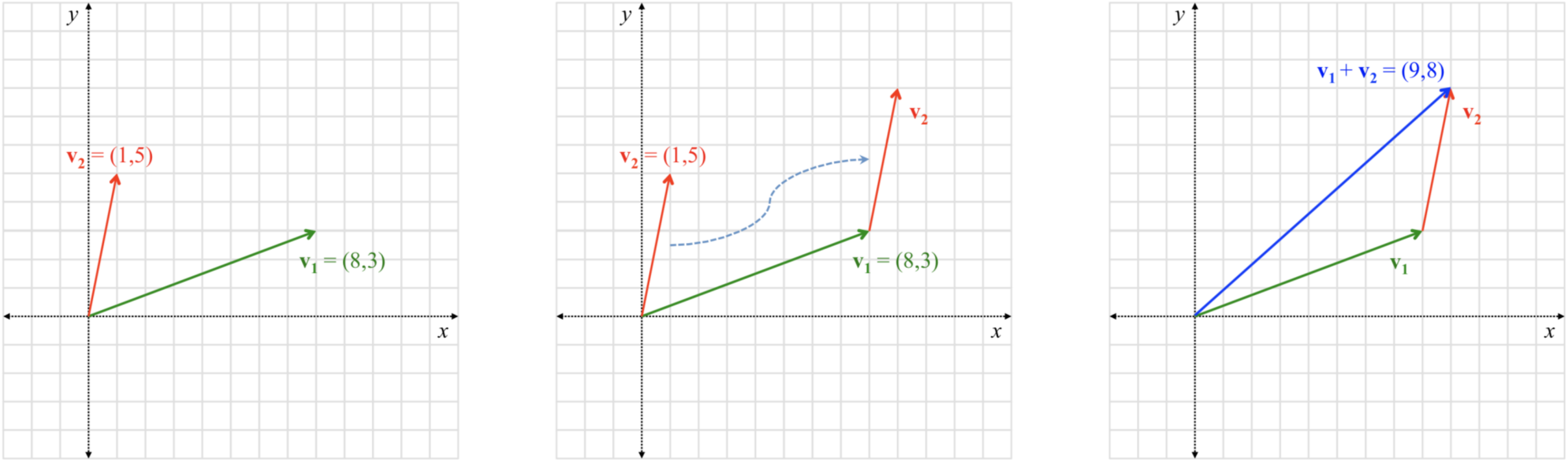
$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{bmatrix} 8 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 8+1 \\ 3+5 \end{bmatrix} = \begin{bmatrix} 9 \\ 8 \end{bmatrix}$$



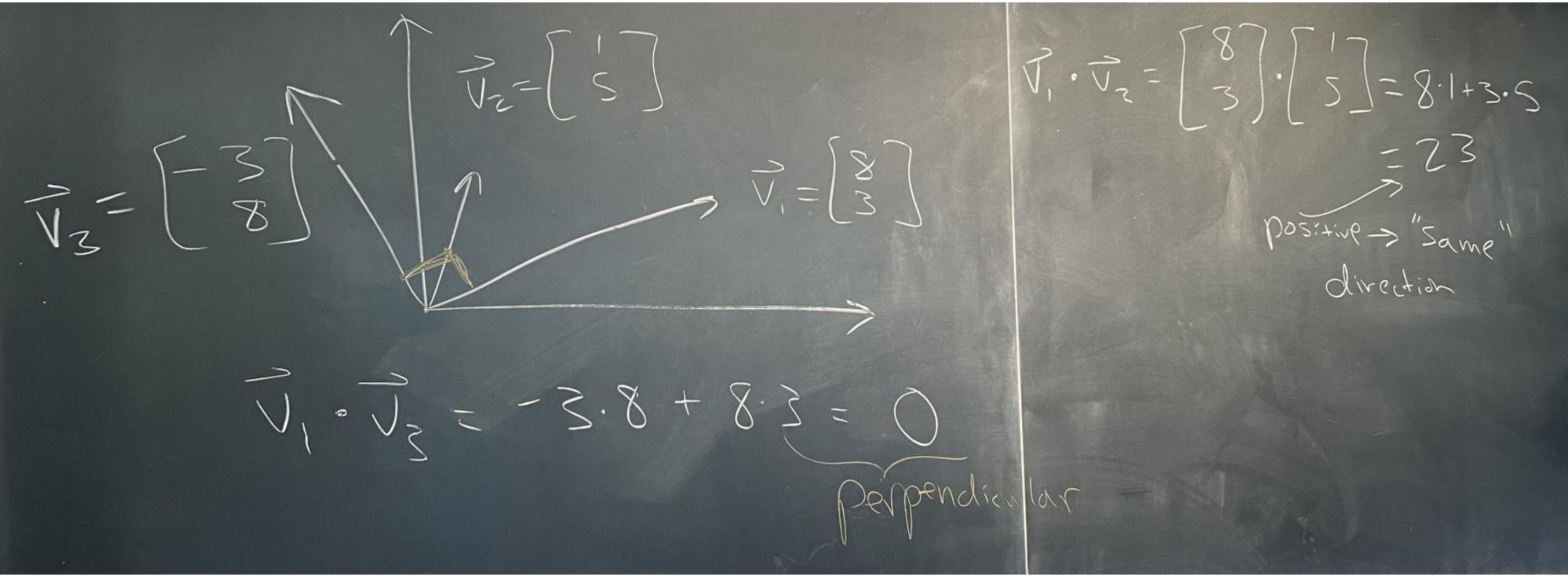



# Vector Addition

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



# Vector Dot Product



# Matrices

- Matrix addition (must be exactly the same dimension!)

$$\mathbf{A}+\mathbf{B}=\left[\begin{array}{cc} a & b \\ c & d \end{array}\right]+\left[\begin{array}{cc} e & f \\ g & h \end{array}\right]=\left[\begin{array}{cc} a+e & b+f \\ c+g & d+h \end{array}\right]$$

# 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: $(\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T\boldsymbol{A}^T$

# Matrix Inverse

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$$

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

# Outline for today

- 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 |

X

y

# Multiple Linear Regression

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

"fake" ones

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

$n \times (p+1)$

- Goal: minimize cost function

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

# Outline for today

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

# Analytic solution to multiple linear regression

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

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

- Take derivative and set to $\vec{0}$

$$\frac{\partial J}{\partial \vec{w}} = -X^T\vec{y} + (X^TX)\vec{w} = \vec{0}$$

$$\leftrightarrow (X^TX)\vec{w} = X^T\vec{y}$$

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

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

(Keep this formula and its interpretation in mind!)

# Handout 5

# Handout 5, #1

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

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

# Handout 5

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

"fake" ones

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

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

$$= \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{2-1} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Same as before!