

# CSCSE 689: Special Topics in Modern Algorithms for Data Science

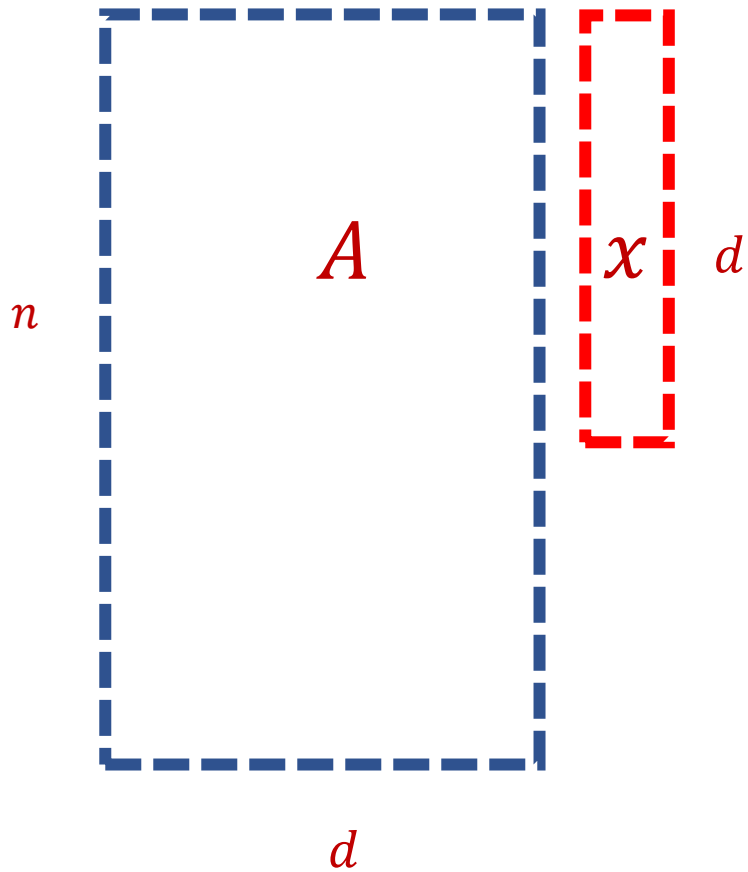
## Lecture 30

Samson Zhou

# Presentation Schedule

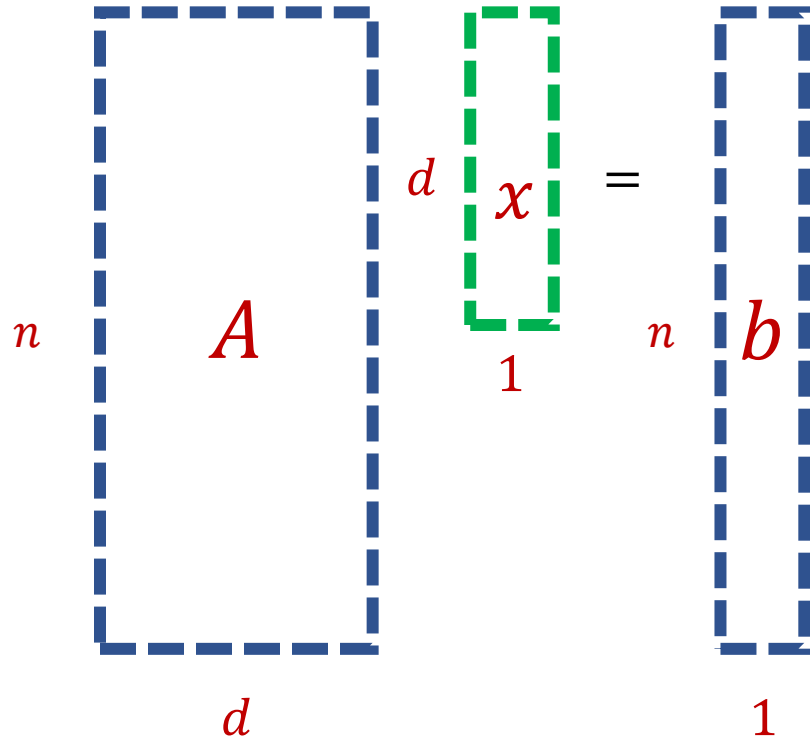
- **November 27:** Chunkai, Jung, Galaxy AI
- **November 29:** STMI, Anmol, Jason
- **December 1:** Bokun, Ayesha, Dawei, Lipai

# Linear Algebra Review



- For  $y = Ax$ , we have  $y_i = \langle a_i, x \rangle$

# Regression



- Let  $A$  be a set of  $n$  observations, each with  $d$  features
- Let  $b$  be the vector of outcomes/labels for each observation
- Find the vector  $x$  such that  $Ax = b$



**Karl Rohe** @karlrohe · 3h



i'm teaching linear regression in the spring

First thing we'll do: highlight the two \*very\* different motivations for linear models

- 1) for causal inference
- 2) for prediction

Breiman's "two cultures" paper is too advanced for mid-undergrads... are there any easier readings?



# Regression

- **Economics and Finance:**
  - **Stock price prediction:** Predict stock prices based on historical data and other relevant factors.
  - **Econometric analysis:** Study the relationships between economic variables like GDP, inflation, and interest rates.



# Regression

- **Medicine and Healthcare:**
  - **Medical research:** Understand the relationship between factors such as age, genetics, and lifestyle on health outcomes.
  - **Disease prediction:** Predicting the probability of disease occurrence based on risk factors like smoking, diet, and exercise.



# Regression

- **Sports Analytics:**

- **Player performance forecast:** Predicting player performance in sports like baseball, basketball, or soccer based on historical data.
- **Game outcome prediction:** Predicting the outcome of games based on team statistics and other factors.





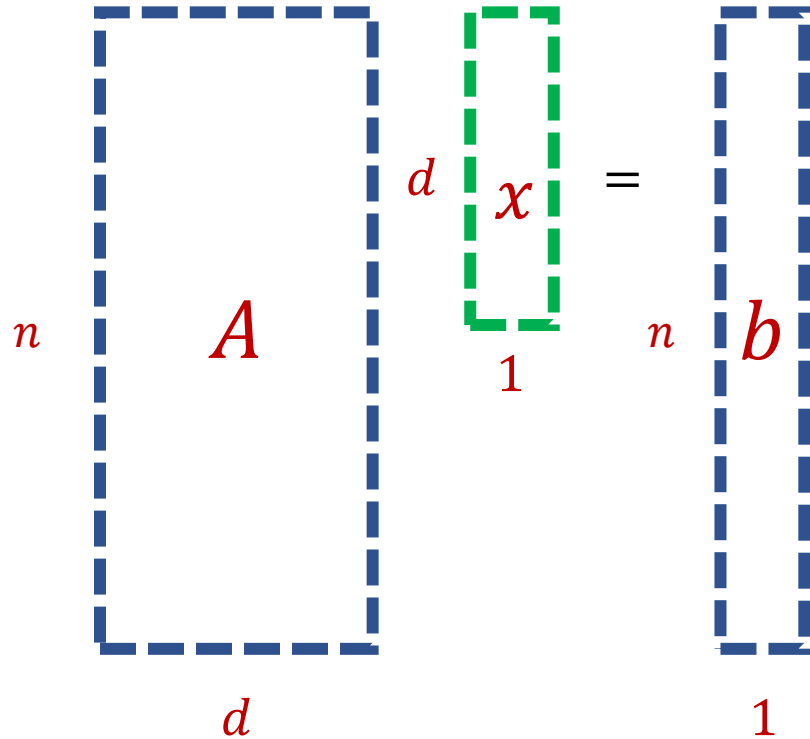
# Regression

- **Environmental Science:**

- **Climate modeling:** Model climate change variables like temperature, precipitation, and sea levels.
- **Pollution analysis:** Analyze the relationship between pollution levels and various factors like industrial activity and population.



# Regression



- Let  $A$  be a set of  $n$  observations, each with  $d$  features
- Let  $b$  be the vector of outcomes/labels for each observation
- Find the vector  $x$  such that  $Ax = b$

# Linear Algebra Review

The diagram illustrates the matrix equation  $Ax = b$ . It shows three components: a matrix  $A$  on the left, a vector  $x$  in the middle, and a vector  $b$  on the right, separated by an equals sign. Matrix  $A$  is represented by a blue dashed rectangle with height  $n$  and width  $d$ . Vector  $x$  is represented by a green dashed rectangle with height  $d$  and width  $1$ . Vector  $b$  is represented by a blue dashed rectangle with height  $n$  and width  $1$ .

- What are conditions for finding  $x$  such that  $Ax = b$ ?

# Linear Algebra Review

The diagram illustrates the matrix equation  $Ax = b$ . Matrix  $A$  is represented by a blue dashed rectangle with height  $n$  and width  $d$ . Vector  $x$  is represented by a green dashed rectangle with height  $d$  and width  $1$ . Vector  $b$  is represented by a blue dashed rectangle with height  $n$  and width  $1$ . The equation is shown as  $Ax = b$ .

- What are conditions for finding  $x$  such that  $Ax = b$ ?
- If the system is inconsistent, there are no solutions

# Linear Algebra Review

The diagram illustrates the linear system  $Ax = b$ . It shows three matrices represented by dashed boxes:

- A coefficient matrix  $A$  with dimensions  $n$  (rows) by  $d$  (columns).
- A variable vector  $x$  with dimensions  $d$  (rows) by  $1$  (column).
- A constant vector  $b$  with dimensions  $n$  (rows) by  $1$  (column).

The equation is shown as  $Ax = b$ , with the dimensions of each matrix indicated by red labels.

- What are conditions for finding  $x$  such that  $Ax = b$ ?
- If the system is inconsistent, there are no solutions
- Can check for consistency by looking at the rank of the coefficient matrix and the augmented matrix

# Linear Algebra Review

- For a square matrix  $A \in \mathbb{R}^{n \times n}$ , if  $A$  is full rank, i.e.,  $\text{rank}(A) = n$ , then  $A$  has an inverse  $A^{-1} \in \mathbb{R}^{n \times n}$  such that  $AA^{-1} = A^{-1}A = \mathbb{I}_n$
- For a general matrix  $A \in \mathbb{R}^{n \times d}$  with linearly independent columns, the Moore-Penrose inverse/pseudoinverse  $A^\dagger$  of  $A$  satisfies  $A^\dagger = (A^\top A)^{-1}A^\top$  so that  $A^\dagger A = \mathbb{I}_d$

# Linear Algebra Review

The diagram illustrates the matrix equation  $Ax = b$ . On the left, a large blue dashed rectangle represents matrix  $A$ , with height  $n$  and width  $d$ . To its right is a smaller green dashed rectangle representing vector  $x$ , with height  $d$  and width  $1$ . An equals sign follows, and to the right is a blue dashed rectangle representing vector  $b$ , with height  $n$  and width  $1$ .

- What are conditions for finding  $x$  such that  $Ax = b$ ?
- If the system is inconsistent, there are no solutions
- If the system is consistent and  $n = d$ , then there is a single solution,  $x = A^{-1}b$

# Linear Algebra Review

The diagram illustrates the matrix equation  $Ax = b$ . On the left, a large blue dashed rectangle represents matrix  $A$ , with height  $n$  and width  $d$ . To its right is a smaller green dashed rectangle representing vector  $x$ , with height  $d$  and width  $1$ . An equals sign follows, and to the right is a blue dashed rectangle representing vector  $b$ , with height  $n$  and width  $1$ .

- What are conditions for finding  $x$  such that  $Ax = b$ ?
- If the system is inconsistent, there are no solutions
- If the system is consistent and  $n < d$ , then there are infinite solutions and can find a solution by looking at  $n$  columns of  $A$



# Linear Algebra Review

The diagram shows the equation  $Ax = b$  with dimensions indicated. Matrix  $A$  is represented by a blue dashed rectangle with height  $n$  and width  $d$ . Vector  $x$  is a green dashed vertical rectangle with height  $d$  and width  $1$ . Vector  $b$  is a blue dashed vertical rectangle with height  $n$  and width  $1$ . An equals sign is placed between  $x$  and  $b$ .

- What are conditions for finding  $x$  such that  $Ax = b$ ?
- If the system is inconsistent, there are no solutions
- If the system is consistent and  $n > d$ , then at most one solution ( $x = A^\dagger b$ ) and can find by looking at  $d$  rows of  $A$

# Regression

The diagram illustrates the linear regression equation  $Ax = b$ . It shows three matrices represented by dashed boxes:

- A matrix  $A$  on the left, with height  $n$  and width  $d$ .
- A vector  $x$  in the middle, with height  $d$  and width  $1$ .
- A vector  $b$  on the right, with height  $n$  and width  $1$ .

The equation is represented as  $Ax = b$ , with an equals sign between the vector  $x$  and the vector  $b$ .

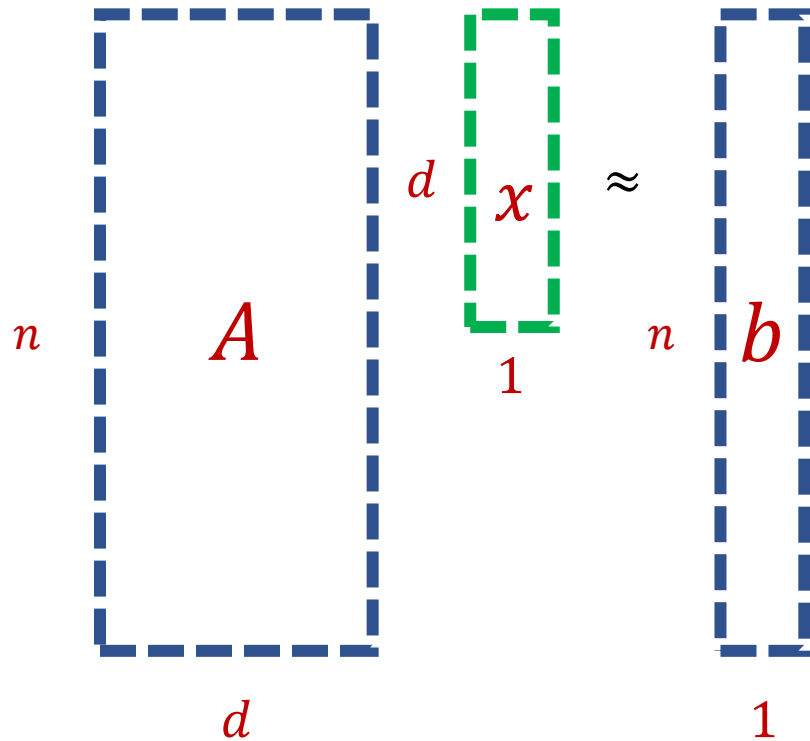
- What are conditions for finding  $x$  such that  $Ax = b$ ?
- What to do when there is no  $x$  such that  $Ax = b$ ?

# Regression

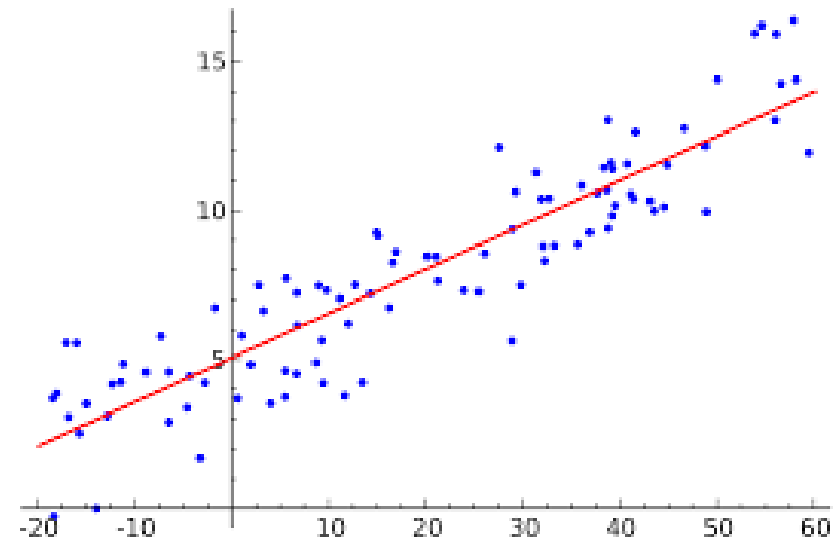
The diagram illustrates the regression equation  $Ax = b$ . On the left, a blue dashed rectangle represents matrix  $A$ , with height  $n$  and width  $d$ . To its right is a green dashed rectangle representing vector  $x$ , with height  $d$  and width  $1$ . An equals sign follows, and to the right is a blue dashed rectangle representing vector  $b$ , with height  $n$  and width  $1$ .

- What are conditions for finding  $x$  such that  $Ax = b$ ?
- What to do when there is no  $x$  such that  $Ax = b$ ?
- Minimize  $\mathcal{L}(Ax - b)$  for some loss function  $\mathcal{L}$

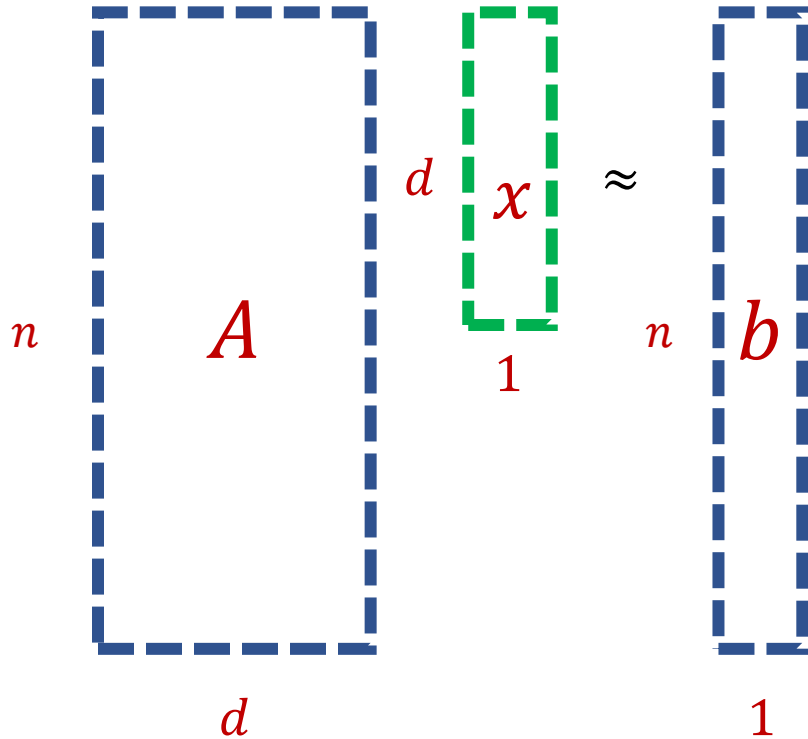
# Linear Regression



- Find the vector  $x$  that minimizes  $\|Ax - b\|_2$
- “Least squares” optimization




# Linear Regression



- Find the vector  $x$  that minimizes  $\|Ax - b\|_2$
- What is the solution to  $\min_{x \in \mathbb{R}^d} \|Ax - b\|_2$ ?

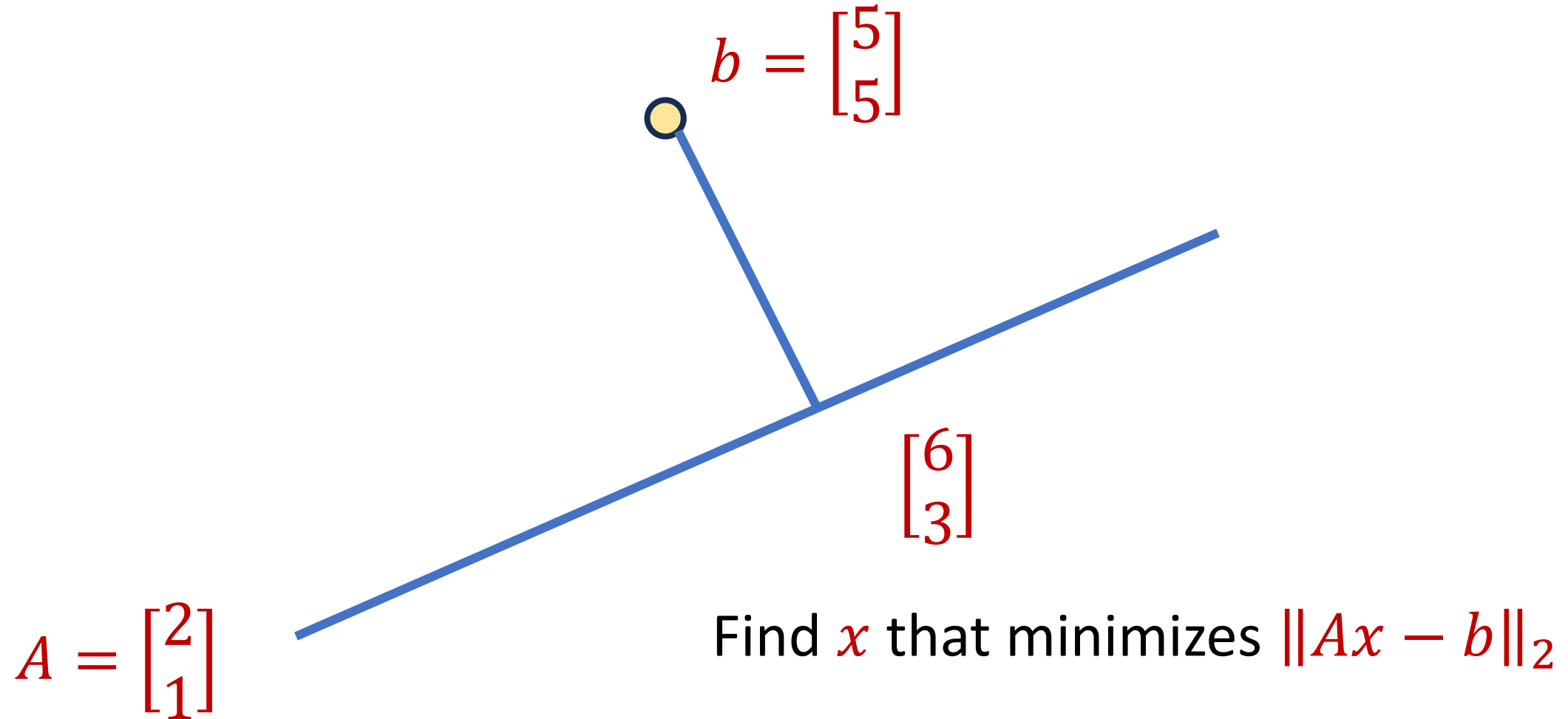
# Linear Regression


$$b = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

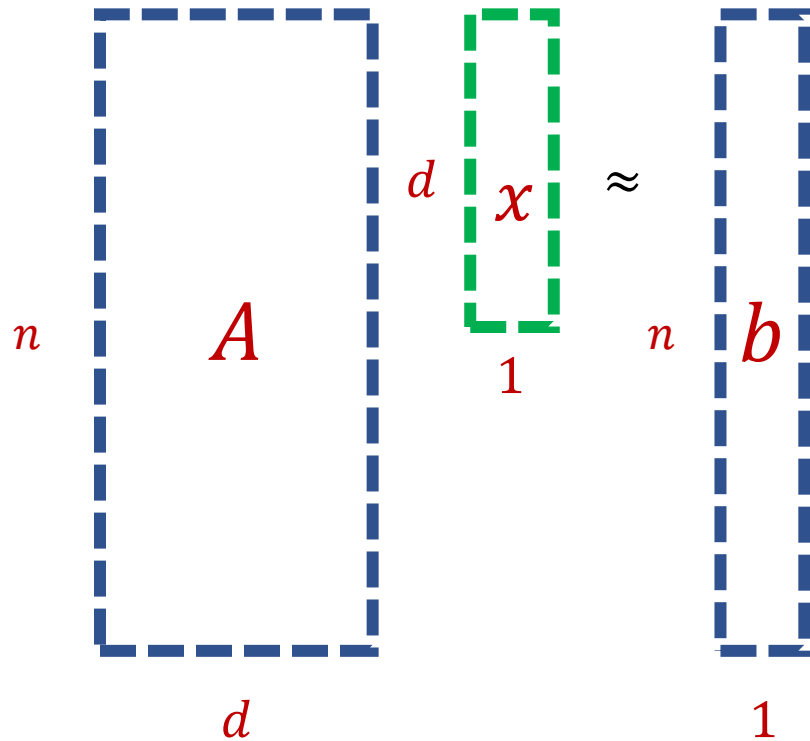
$$A = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Find  $x$  that minimizes  $\|Ax - b\|_2$

# Linear Regression



# Linear Regression



- Find the vector  $x$  that minimizes  $\|Ax - b\|_2$
- What is the solution to  $\min_{x \in \mathbb{R}^d} \|Ax - b\|_2$ ?
- Let  $b^\perp$  be the projection of  $b$  onto the column space of  $A$
- Decompose  $b$  into  $b^\perp$  and  $b^\parallel$



# Linear Regression

- We have  $\arg \min_{x \in \mathbb{R}^d} \|Ax - b\|_2 = \arg \min_{x \in \mathbb{R}^d} \|Ax - b\|_2^2$
- $\|Ax - b\|_2^2 = \|Ax - b^\perp - b^\parallel\|_2^2$

# Linear Regression

- We have  $\arg \min_{x \in \mathbb{R}^d} \|Ax - b\|_2 = \arg \min_{x \in \mathbb{R}^d} \|Ax - b\|_2^2$
- $$\begin{aligned} \|Ax - b\|_2^2 &= \|Ax - b^\perp - b^\parallel\|_2^2 \\ &= \|Ax - b^\parallel\|_2^2 - 2\langle Ax - b^\parallel, b^\perp \rangle + \|b^\perp\|_2^2 \end{aligned}$$

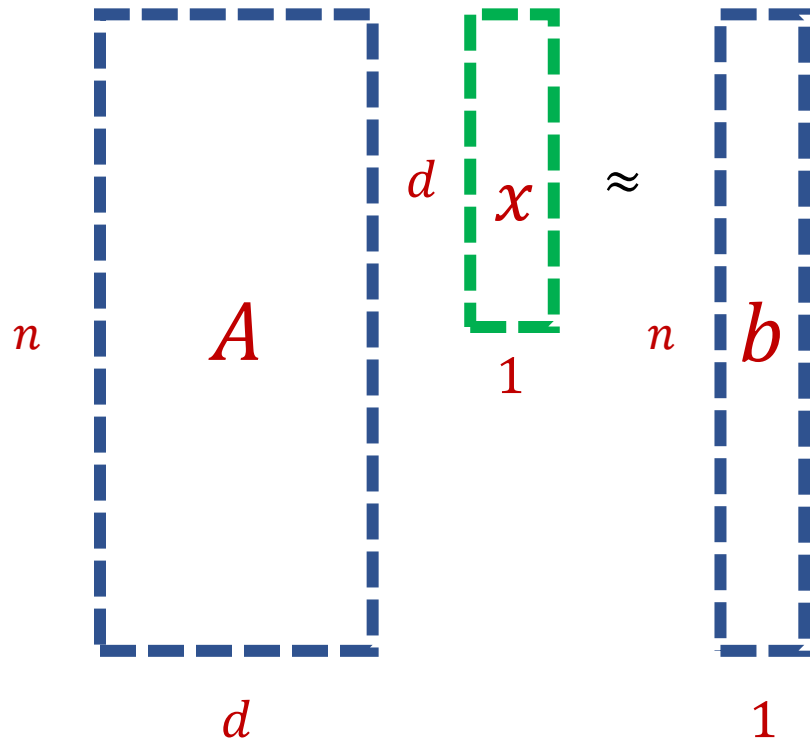
# Linear Regression

- We have  $\arg \min_{x \in \mathbb{R}^d} \|Ax - b\|_2 = \arg \min_{x \in \mathbb{R}^d} \|Ax - b\|_2^2$
- $$\begin{aligned} \|Ax - b\|_2^2 &= \|Ax - b^\perp - b^\parallel\|_2^2 \\ &= \|Ax - b^\parallel\|_2^2 - 2\langle Ax - b^\parallel, b^\perp \rangle + \|b^\perp\|_2^2 \\ &= \|Ax - b^\parallel\|_2^2 + \|b^\perp\|_2^2 \end{aligned}$$

# Linear Regression

- We have  $\arg \min_{x \in \mathbb{R}^d} \|Ax - b\|_2 = \arg \min_{x \in \mathbb{R}^d} \|Ax - b\|_2^2$
- $$\begin{aligned} \|Ax - b\|_2^2 &= \|Ax - b^\perp - b^\parallel\|_2^2 \\ &= \|Ax - b^\parallel\|_2^2 - 2\langle Ax - b^\parallel, b^\perp \rangle + \|b^\perp\|_2^2 \\ &= \|Ax - b^\parallel\|_2^2 + \|b^\perp\|_2^2 \end{aligned}$$
- Minimized for  $\|Ax - b^\parallel\|_2^2 = 0$  when  $x = A^\dagger b^\parallel = A^\dagger b$

# Linear Regression



- Find the vector  $x$  that minimizes  $\|Ax - b\|_2$
- “Least squares” optimization
- MLE under Gaussian noise
- Closed form solution:  $x = A^\dagger b$

