

Lectures 8-10

# Linear algebra: Dot products and Projections

DSC 40A, Fall 2025

## Question 🤔

Answer at [q.dsc40a.com](https://q.dsc40a.com)

**Remember, you can always ask questions at [q.dsc40a.com!](https://q.dsc40a.com)**

If the direct link doesn't work, click the " Lecture Questions" link in the top right corner of [dsc40a.com](https://dsc40a.com).





## Exercise

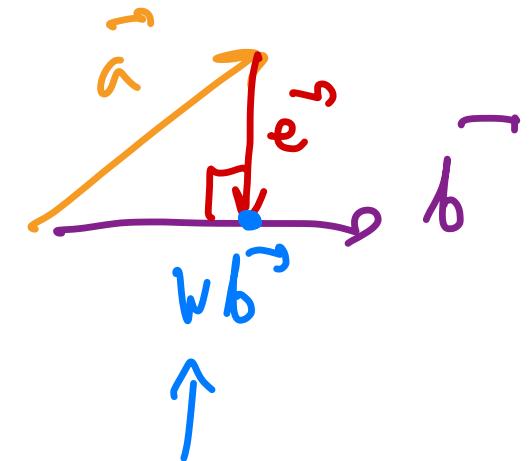
Let  $\vec{a} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$  and  $\vec{b} = \begin{bmatrix} -1 \\ 9 \end{bmatrix}$ .

What is the orthogonal projection of  $\vec{a}$  onto  $\text{span}(\vec{b})$ ?

Your answer should be of the form  $w^* \vec{b}$ , where  $w^*$  is a scalar.

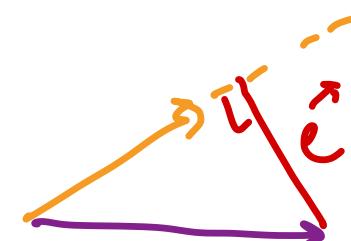
Projecting  $\vec{a}$  onto  $\text{span}\{\vec{b}\}$

$$w^* = \frac{\vec{b} \cdot \vec{a}}{\vec{b} \cdot \vec{b}} = \frac{-1 \cdot 5 + 9 \cdot 2}{(-1)^2 + 9^2} = \frac{-5 + 18}{1 + 81} = \frac{13}{82}$$



Projecting  $\vec{b}$  onto  $\text{span}\{\vec{a}\}$

$$w^* = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}} = \frac{5 \cdot (-1) + 2 \cdot 9}{5^2 + 2^2} = \frac{13}{29}$$



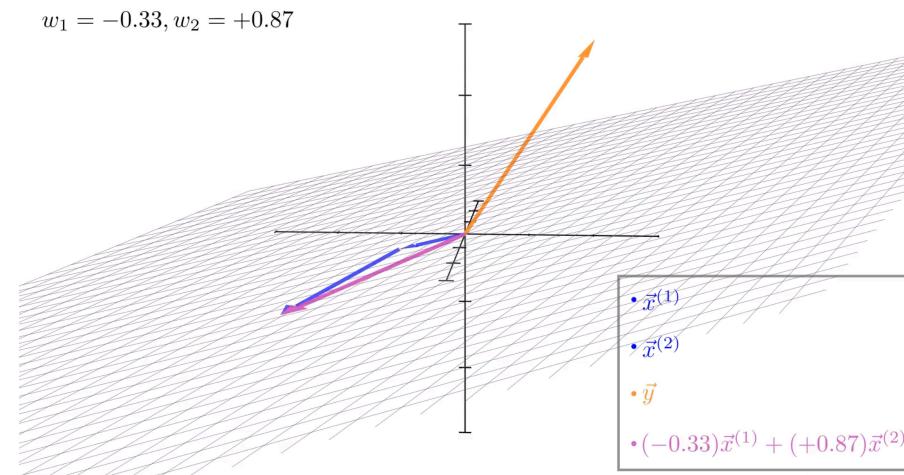
## Moving to multiple dimensions

feature index superscript



- Let's now consider three vectors,  $\vec{y}$ ,  $\vec{x}^{(1)}$ , and  $\vec{x}^{(2)}$ , all in  $\mathbb{R}^n$ .
- Question:** What vector in  $\text{span}(\vec{x}^{(1)}, \vec{x}^{(2)})$  is closest to  $\vec{y}$ ?
  - Vectors in  $\text{span}(\vec{x}^{(1)}, \vec{x}^{(2)})$  are of the form  $w_1\vec{x}^{(1)} + w_2\vec{x}^{(2)}$ , where  $w_1, w_2 \in \mathbb{R}$  are scalars.
- Before trying to answer, let's watch [this animation that Jack, one of our tutors, made.](#)

previous

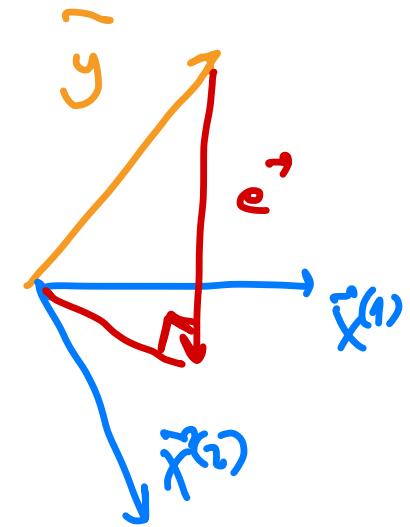


## Minimizing projection error in multiple dimensions

- Question: What vector in  $\text{span}(\vec{x}^{(1)}, \vec{x}^{(2)})$  is closest to  $\vec{y}$ ?

- That is, what vector minimizes  $\|\vec{e}\|$ , where:

$$\vec{e} = \vec{y} - w_1 \vec{x}^{(1)} - w_2 \vec{x}^{(2)}$$



- Answer: It's the vector such that  $w_1 \vec{x}^{(1)} + w_2 \vec{x}^{(2)}$  is **orthogonal** to  $\vec{e}$ .
- Issue: Solving for  $w_1$  and  $w_2$  in the following equation is difficult:

$$(w_1 \vec{x}^{(1)} + w_2 \vec{x}^{(2)}) \cdot \underbrace{(\vec{y} - w_1 \vec{x}^{(1)} - w_2 \vec{x}^{(2)})}_{\vec{e}} = 0$$

Vector is in the span  $\{\vec{x}^{(1)}, \vec{x}^{(2)}\}$   
is a linear comb. of  $\vec{x}^{(1)}$  and  $\vec{x}^{(2)}$

## Minimizing projection error in multiple dimensions

- It's hard for us to solve for  $w_1$  and  $w_2$  in:

$$\left( w_1 \vec{x}^{(1)} + w_2 \vec{x}^{(2)} \right) \cdot \underbrace{\left( \vec{y} - w_1 \vec{x}^{(1)} - w_2 \vec{x}^{(2)} \right)}_{\vec{e}} = 0$$

- Observation:** All we really need is for  $\vec{x}^{(1)}$  and  $\vec{x}^{(2)}$  to individually be orthogonal to  $\vec{e}$ .

- That is, it's sufficient for  $\vec{e}$  to be orthogonal to the spanning vectors themselves.
- If  $\vec{x}^{(1)} \cdot \vec{e} = 0$  and  $\vec{x}^{(2)} \cdot \vec{e} = 0$ , then:

$$\begin{aligned} (w_1 \vec{x}^{(1)} + w_2 \vec{x}^{(2)}) \cdot \vec{e} &= (w_1 \vec{x}^{(1)}) \cdot \vec{e} + (w_2 \vec{x}^{(2)}) \cdot \vec{e} \\ &= w_1 (\vec{x}_1 \cdot \vec{e}) + w_2 (\vec{x}_2 \cdot \vec{e}) = 0 \end{aligned}$$

## Minimizing projection error in multiple dimensions

- Question: What vector in  $\text{span}(\vec{x}^{(1)}, \vec{x}^{(2)})$  is closest to  $\vec{y}$ ?
- Answer: It's the vector such that  $w_1\vec{x}^{(1)} + w_2\vec{x}^{(2)}$  is orthogonal to  $\vec{e} = \vec{y} - w_1\vec{x}^{(1)} - w_2\vec{x}^{(2)}$ .
- Equivalently, it's the vector such that  $\vec{x}^{(1)}$  and  $\vec{x}^{(2)}$  are both orthogonal to  $\vec{e}$ :

$$\begin{aligned}\vec{x}^{(1)} \cdot (\underbrace{\vec{y} - w_1\vec{x}^{(1)} - w_2\vec{x}^{(2)}}_{\vec{e}}) &= 0 \\ \vec{x}^{(2)} \cdot (\underbrace{\vec{y} - w_1\vec{x}^{(1)} - w_2\vec{x}^{(2)}}_{\vec{e}}) &= 0\end{aligned}$$

Solving these  
2 equations  
will yield  
 $w_1^*$ ,  $w_2^*$

- This is a system of two equations, two unknowns ( $w_1$  and  $w_2$ ), but it still looks difficult to solve.

## Now what?

- We're looking for the scalars  $w_1$  and  $w_2$  that satisfy the following equations:

$$\vec{x}^{(1)} \cdot \left( \vec{y} - w_1 \vec{x}^{(1)} - w_2 \vec{x}^{(2)} \right) = 0$$
$$\vec{x}^{(2)} \cdot \underbrace{\left( \vec{y} - w_1 \vec{x}^{(1)} - w_2 \vec{x}^{(2)} \right)}_{\vec{e}} = 0$$

- In this example, we just have two spanning vectors,  $\vec{x}^{(1)}$  and  $\vec{x}^{(2)}$ .
- If we had any more, this system of equations would get extremely messy, extremely quickly.
- Idea: Rewrite the above system of equations as a single equation, involving matrix-vector products.

# Matrices



## Matrix addition and scalar multiplication

- We can add two matrices only if they have the same dimensions.
- Addition occurs elementwise:

$$\begin{bmatrix} 2 & 5 & 8 \\ -1 & 5 & -3 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 7 & 11 \\ -1 & 6 & -1 \end{bmatrix}$$

$$C[i,j] = A[i,j] + B[i,j]$$

- Scalar multiplication occurs elementwise, too:

$$2 \begin{bmatrix} 2 & 5 & 8 \\ -1 & 5 & -3 \end{bmatrix} = \begin{bmatrix} 4 & 10 & 16 \\ -2 & 10 & -6 \end{bmatrix}$$

$$D[i,j] = cA[i,j]$$

↓

$c \in \mathbb{R}$



## Question 🤔

Answer at [q.dsc40a.com](http://q.dsc40a.com)

Assume  $A$ ,  $B$ , and  $C$  are all matrices. Select the incorrect statement below.

- A.  $A(B + C) = AB + AC$ .
- B.  $A(BC) = (AB)C$ .
- C.  $AB = BA$ .
- D.  $(A + B)^T = A^T + B^T$ .
- E.  $(AB)^T = B^T A^T$ .

example

$$\begin{matrix} A & \cdot & B & = & C \\ 5 \times 7 & & 7 \times 3 & & 5 \times 3 \end{matrix}$$

$B \cdot A \rightarrow$  product undefined!

$$\begin{matrix} 7 \times 3 & & 5 \times 7 \\ \uparrow & & \uparrow \\ \text{different} \end{matrix}$$

# Matrix-vector multiplication

- A vector  $\vec{v} \in \mathbb{R}^n$  is a matrix with  $n$  rows and 1 column.

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^{n \times 1}$$

- Suppose  $A \in \mathbb{R}^{n \times d}$ .

- What must the dimensions of  $\vec{v}$  be in order for the product  $A\vec{v}$  to be valid?

$$A \quad \vec{v}$$

$n \times d$        $d \times 1$

$\vec{v} \in \mathbb{R}^d$ , column vec with  $d$  elements

- What must the dimensions of  $\vec{v}$  be in order for the product  $\vec{v}^T A$  to be valid?

$$[ - ] [ ]$$

$$\vec{v}^T \cdot A$$

$1 \times n$        $n \times d$

$\vec{v} \in \mathbb{R}^n$ ,  $\vec{v}^T$  row vec with  $n$  elements



## Another view of matrix-vector multiplication

- Another way of thinking about the product  $A\vec{v}$  is that it is a linear combination of the columns of  $A$ , using the weights in  $\vec{v}$ .
- Example: What is  $A\vec{v}$ ?

$$A = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \begin{bmatrix} 5 \\ 5 \end{bmatrix} \begin{bmatrix} 8 \\ -3 \end{bmatrix}$$
$$\vec{v} = \begin{bmatrix} 2 \\ -1 \\ -5 \end{bmatrix}$$

$$A\vec{v} = 2 \underbrace{\begin{bmatrix} 2 \\ -1 \end{bmatrix}}_{\text{linear combination of the cols}} - 1 \underbrace{\begin{bmatrix} 5 \\ 5 \end{bmatrix}}_{\text{of } A} - 5 \underbrace{\begin{bmatrix} 8 \\ -3 \end{bmatrix}}_{\text{of } A} = \begin{bmatrix} 4 \\ -2 \end{bmatrix} + \begin{bmatrix} -5 \\ -5 \end{bmatrix} + \begin{bmatrix} -40 \\ 15 \end{bmatrix} = \begin{bmatrix} 41 \\ 8 \end{bmatrix}$$

linear combination of the cols  
of  $A$

## Matrix-vector products create linear combinations of columns!

- Key idea: It'll be very useful to think of the matrix-vector product  $A\vec{v}$  as a linear combination of the columns of  $A$ , using the weights in  $\vec{v}$ .

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1d} \\ a_{21} & a_{22} & \dots & a_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nd} \end{bmatrix} \quad \vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{bmatrix}$$

↓

$$A\vec{v} = v_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + v_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \dots + v_d \begin{bmatrix} a_{1d} \\ a_{2d} \\ \vdots \\ a_{nd} \end{bmatrix}$$

# Spans and projections, revisited

## Moving to multiple dimensions

- Let's now consider three vectors,  $\vec{y}$ ,  $\vec{x}^{(1)}$ , and  $\vec{x}^{(2)}$ , all in  $\mathbb{R}^n$ .
- Question: What vector in  $\text{span}(\vec{x}^{(1)}, \vec{x}^{(2)})$  is closest to  $\vec{y}$ ?
  - That is, what values of  $w_1$  and  $w_2$  minimize  $\|\vec{e}\| = \|\vec{y} - w_1\vec{x}^{(1)} - w_2\vec{x}^{(2)}\|$ ?

Find  $w_1^+$ ,  $w_2^+$  such that

$$\left\{ \begin{array}{l} \vec{x}^{(1)} \cdot \vec{e} = 0 \\ \vec{x}^{(2)} \cdot \vec{e} = 0 \end{array} \right.$$

## Matrix-vector products create linear combinations of columns!

$$\vec{x}^{(1)} = \begin{bmatrix} 2 \\ 5 \\ 3 \end{bmatrix} \quad \vec{x}^{(2)} = \begin{bmatrix} -1 \\ 0 \\ 4 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix} \quad w_1 \vec{x}^{(1)} + w_2 \vec{x}^{(2)}$$

- Combining  $\vec{x}^{(1)}$  and  $\vec{x}^{(2)}$  into a single matrix gives:

$$X = \begin{bmatrix} | & | \\ \vec{x}^{(1)} & \vec{x}^{(2)} \\ | & | \end{bmatrix} = \begin{bmatrix} \underline{2} & \underline{-1} \\ \underline{5} & \underline{0} \\ \underline{3} & \underline{4} \end{bmatrix} \quad X\vec{w} = w_1 \vec{x}^{(1)} + w_2 \vec{x}^{(2)}$$
$$\vec{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

- Then, if  $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ , linear combinations of  $\vec{x}^{(1)}$  and  $\vec{x}^{(2)}$  can be written as  $X\vec{w}$ .
- The **span of the columns** of  $X$ , or  $\text{span}(X)$ , consists of all vectors that can be written in the form  $X\vec{w}$ .

## Minimizing projection error in multiple dimensions

$$X = \begin{bmatrix} | & | \\ \vec{x}^{(1)} & \vec{x}^{(2)} \\ | & | \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 5 & 0 \\ 3 & 4 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix}$$

- Goal: Find the vector  $\vec{w} = [w_1 \ w_2]^T$  such that  $\|\vec{e}\| = \|\vec{y} - X\vec{w}\|$  is minimized.
- As we've seen,  $\vec{w}$  must be such that:

2 equations  
with 2 variables  
 $w_1, w_2$

$$\left\{ \begin{array}{l} \vec{x}^{(1)} \cdot (\vec{y} - w_1 \vec{x}^{(1)} - w_2 \vec{x}^{(2)}) = 0 \\ \vec{x}^{(2)} \cdot (\vec{y} - w_1 \vec{x}^{(1)} - w_2 \vec{x}^{(2)}) = 0 \end{array} \right. \underbrace{\quad}_{\vec{e}}$$

known      unknown

- How can we use our knowledge of matrices to rewrite this system of equations as a single equation?

## Simplifying the system of equations, using matrices

$$X = \begin{bmatrix} | & | \\ \vec{x}^{(1)} & \vec{x}^{(2)} \\ | & | \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 5 & 0 \\ 3 & 4 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix}$$

$$\vec{x}^{(1)} \cdot (\vec{y} - w_1 \vec{x}^{(1)} - w_2 \vec{x}^{(2)}) = 0$$

$$\vec{x}^{(2)} \cdot (\vec{y} - w_1 \vec{x}^{(1)} - w_2 \vec{x}^{(2)}) = 0$$

$$w_1 \vec{x}^{(1)} + w_2 \vec{x}^{(2)} = \vec{X} \vec{w}$$

$\vec{e}$

$$\vec{x}^{(1)} (\vec{y} - \vec{X} \vec{w}) = 0$$

$$\vec{e} = \vec{y} - \vec{X} \vec{w}$$

plug in

$$\vec{x}^{(2)} (\vec{y} - \vec{X} \vec{w}) = 0$$

## Simplifying the system of equations, using matrices

$$X = \begin{bmatrix} | & | \\ \vec{x}^{(1)} & \vec{x}^{(2)} \\ | & | \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 5 & 0 \\ 3 & 4 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix}$$

1.  $w_1 \vec{x}^{(1)} + w_2 \vec{x}^{(2)}$  can be written as  $X \vec{w}$ , so  $\vec{e} = \vec{y} - X \vec{w}$ .
2. The condition that  $\vec{e}$  must be orthogonal to each column of  $X$  is equivalent to condition that  $X^T \vec{e} = 0$ .



$$\vec{y} - X \vec{w}$$



## The normal equations

$$X = \begin{bmatrix} | & | \\ \vec{x}^{(1)} & \vec{x}^{(2)} \\ | & | \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 5 & 0 \\ 3 & 4 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix}$$

- **Goal:** Find the vector  $\vec{w} = [w_1 \quad w_2]^T$  such that  $\|\vec{e}\| = \|\vec{y} - X\vec{w}\|$  is minimized.
- We now know that it is the vector  $\vec{w}^*$  such that:

$$\begin{aligned} X^T \vec{e} &= 0 \\ X^T (\vec{y} - X\vec{w}^*) &= 0 && \text{if invertible} \\ X^T \vec{y} - X^T X \vec{w}^* &= 0 \\ \implies X^T X \vec{w}^* &= X^T \vec{y} \end{aligned}$$
$$\vec{w}^* = \underbrace{(X^T X)^{-1}}_{\sim} X^T \vec{y}$$

- The last statement is referred to as the **normal equations**.

## The general solution to the normal equations

$$X \in \mathbb{R}^{n \times d} \quad \vec{y} \in \mathbb{R}^n$$

- **Goal, in general:** Find the vector  $\vec{w} \in \mathbb{R}^d$  such that  $\|\vec{e}\| = \|\vec{y} - X\vec{w}\|$  is minimized.
- We now know that it is the vector  $\vec{w}^*$  such that:

$$\begin{aligned} X^T \vec{e} &= 0 \\ \implies X^T X \vec{w}^* &= X^T \vec{y} \end{aligned}$$

- Assuming  $X^T X$  is invertible, this is the vector:

$$\boxed{\vec{w}^* = (X^T X)^{-1} X^T \vec{y}}$$

- This is a big assumption, because it requires  $X^T X$  to be **full rank**.
- If  $X^T X$  is not full rank, then there are infinitely many solutions to the normal equations,  $X^T X \vec{w}^* = X^T \vec{y}$ .

## What does it mean?

- Original question: What vector in  $\text{span}(\vec{x}^{(1)}, \vec{x}^{(2)})$  is closest to  $\vec{y}$ ?
- Final answer: It is the vector  $\mathbf{X}\vec{w}^*$ , where:

$$\vec{w}^* = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \vec{y}$$

- Revisiting our example:

$$\mathbf{X} = \begin{bmatrix} | & | \\ \vec{x}^{(1)} & \vec{x}^{(2)} \\ | & | \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 5 & 0 \\ 3 & 4 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} 1 \\ 3 \\ 9 \end{bmatrix}$$

- Using a computer gives us  $\vec{w}^* = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \vec{y} \approx \begin{bmatrix} 0.7289 \\ 1.6300 \end{bmatrix}$ .
- So, the vector in  $\text{span}(\vec{x}^{(1)}, \vec{x}^{(2)})$  closest to  $\vec{y}$  is  $0.7289\vec{x}^{(1)} + 1.6300\vec{x}^{(2)}$ .

## An optimization problem, solved

- We just used linear algebra to solve an **optimization problem**.
- Specifically, the function we minimized is:

$$\text{error}(\vec{w}) = \|\vec{y} - \mathbf{X}\vec{w}\|$$

- This is a function whose input is a vector,  $\vec{w}$ , and whose output is a scalar!
- The input,  $\vec{w}^*$ , to **error**( $\vec{w}$ ) that minimizes it is:

$$\vec{w}^* = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \vec{y}$$

- We're going to use this frequently!