OPTIONAL 01

Linear Algebra 01

Vectors, Span, Norms, Dot Products, Matrix Multiplication

Data 100/Data 200, Spring 2023 @ UC Berkeley

Narges Norouzi and Lisa Yan

Content credit: The wealth of instructional knowledge from Math 54, Math 91, and EECS16A

Today's Roadmap

Optional 01, Data 100 Spring 2023

Assumptions, disclaimers, and credits

Matrix/Vector Dimensions, Ax=b

Matrix multiplication

Transpose and dot product

Dot product properties, Norms

Span and Orthogonality

Assumptions and Disclaimers

- This Linear Algebra optional lecture is primarily geared towards co-enrolled Math 54 and Data 100 students in Spring 2023.
    - It may also be useful as a refresher to other students.
    - At the time of this review session, I am assuming you have covered all of "Week 1 curriculum" on [Prof. Lott's page]() (linear combinations, spans, Gaussian elimination). 1.3
- This is part one in a multipart series.
    - Today: Vector properties and Matrix multiplication
    - Next time: Matrix properties
- These optional lectures are **not** a substitute for a linear algebra course. We only cover basics.
- The goal of these optional lectures is to prepare you for the linear algebra requirements for the Ordinary Least Squares unit.
    - Linear algebra will be needed in proofs on Homework 06: Regression.
    - On the Spring 2023 midterm, we will try to de-emphasize linear algebra—as in, we will expect you to be familiar with equations (and how to use them), but we will not ask you to prove anything.

Credits for this lecture, and other Linear Algebra Resources

- Material credit:
  - Math 54 and its textbook (Lay, Lay, McDonald)
  - The new Math 91 course (bCourses)
  - EECS 16A (website)

- https://ds100.org/sp23/resources/#calculus-and-linear-algebra
  - Includes non-Berkeley sources like Khan Academy, 3Brown1Blue, etc.

Lay 1.4, 2.1

EECS 16A Lec1B

Matrix/Vector Dimensions, Ax=b

Optional 01, Data 100 Spring 2023

Three Common Notations for a System of Linear Equations

So far:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 + 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned}$$

1. "augmented matrix"

$$\left[\begin{array}{ccc:c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{array}\right]$$

$\rightarrow$ row reduction to find a soln for system

Three Common Notations for a System of Linear Equations

So far:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 + 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned}$$

1. "augmented matrix"

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{array}\right]$$

This lecture's goal: Get comfortable with these two notations. Also learn some vector properties.

Data science more commonly uses these two notations:

2. **Matrix (and Vector) notation**:

$$\boxed{A\vec{x} = \vec{b}}$$

where $A = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 5 & 0 & -5 \end{bmatrix}, \vec{b} = \begin{bmatrix} 0 \\ 8 \\ 10 \end{bmatrix}, \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

3. **Dot product notation**:

$$\begin{aligned} \vec{a_1}^T \vec{x} &= 0 \\ \vec{a_2}^T \vec{x} &= 8 \\ \vec{a_3}^T \vec{x} &= 10 \end{aligned}$$

where $\vec{a_1} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \vec{a_2} = \begin{bmatrix} 0 \\ 2 \\ -8 \end{bmatrix}, \vec{a_3} = \begin{bmatrix} 5 \\ 0 \\ -5 \end{bmatrix},$

$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Definition of a Vector, Matrix, and $\mathbb{R}^n$

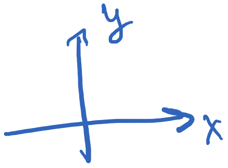
$\begin{bmatrix} 2 & 4 & 6 \\ 1 & 3 & 4 \end{bmatrix}$

- A matrix with only one column is called a column vector, or simply a **vector**.

$$\vec{u} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

$\mathbb{R}^2$

- $\vec{u}$ is part of the Euclidean space $\mathbb{R}^2$.



$$\vec{u} \in \mathbb{R}^2$$

Say: "R two"

Or "the space of two-dimensional vectors"

Definition of a Vector, Matrix, and $\mathbb{R}^n$

- A matrix with only one column is called a column vector, or simply a **vector**.

$$\vec{u} = \begin{array}{cc} \begin{bmatrix} 3 \\ -1 \end{bmatrix} & \begin{array}{c} 1 \\ 2 \end{array} \end{array}$$

Another notation for this space(why?)

$$\boxed{\vec{u} \in \mathbb{R}^{2\times 1}}$$

- $\vec{u}$ is part of the Euclidean space $\mathbb{R}^2$.

$$\boxed{\vec{u} \in \mathbb{R}^2}$$

Say: "R two"

Or "the space of two-dimensional vectors"

- A **matrix** can also be defined as part of a Euclidean space.

$$A = \begin{array}{cc} \begin{array}{cc} 1 & 2 \end{array} & \\ \begin{bmatrix} 1 & -2 \\ 0 & 2 \\ 5 & 0 \end{bmatrix} & \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \end{array}$$

$$\boxed{A \in \mathbb{R}^{3\times 2}}$$

"the space of 3-by-2 matrices"

rows # columns

This order is very important!!!

What are the dimensions of the following vectors/matrices?

1. $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ 2-dimensional vector

2. $\begin{bmatrix} 1 & -2 \\ 0 & 2 \\ 5 & 0 \end{bmatrix}$ 3x2 matrix

3. $\begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}$

4. $\begin{bmatrix} 4 & -2 \\ 5.6 & 2.3 \end{bmatrix}$

5. $\begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 2 \end{bmatrix}$

6. $\begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix}$

7. $\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$

8. $\begin{bmatrix} a_{11} & \cdots & a_{1p} \\ a_{21} & & a_{2p} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mp} \end{bmatrix}$

🤔

What are the dimensions of the following vectors/matrices?

1. $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ 2-dimensional vector

2. $\begin{bmatrix} 1 & -2 \\ 0 & 2 \\ 5 & 0 \end{bmatrix}$ 3x2 matrix

3. $\begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix} \in \mathbb{R}^3$ 3-Dimensional column vector

3x1 matrix

4. $\begin{bmatrix} 4 & -2 \\ 5.6 & 2.3 \end{bmatrix} \in \mathbb{R}^{2\times 2}$ 2-by-2 matrix

"square"

$n_{rows} = n_{cols}$

5. $\begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 2 \end{bmatrix} \in \mathbb{R}^{2\times 3}$

6. $\begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix} \in \mathbb{R}^3$ $\theta_i \in \mathbb{R}$

7. $m \left\{ \overbrace{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}}^{n} \right. \in \mathbb{R}^{m\times n}$ $a_{ij} \in \mathbb{R}$

8. $\begin{bmatrix} a_{11} & \cdots & a_{1p} \\ a_{21} & & a_{2p} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mp} \end{bmatrix} \in \mathbb{R}^{m\times p}$

Linear combination of vectors means $A\vec{x} = \vec{b}$

A fundamental idea in linear algebra is to view a linear combination of vectors as the product of a matrix and a vector. The following definition permits us to rephrase some of the concepts of Section 1.3 in new ways.

> If A is an $m \times n$ matrix, with columns $\mathbf{a}_1, \dots, \mathbf{a}_n$, and if $\mathbf{x}$ is in $\mathbb{R}^n$, then the **product of A and $\mathbf{x}$**, denoted by $A\mathbf{x}$, is **the linear combination of the columns of A using the corresponding entries in $\mathbf{x}$ as weights**; that is,
>
> $$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

Note that $A\mathbf{x}$ is defined only if the number of columns of A equals the number of entries in $\mathbf{x}$.

Scalar

EXAMPLE 1

a. $$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ -15 \end{bmatrix} + \begin{bmatrix} -7 \\ 21 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = \begin{bmatrix} 4+6-7 \\ 0+-15+21 \end{bmatrix}$$

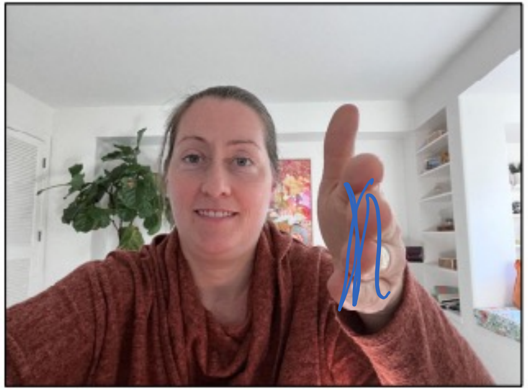
b. $$\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ -20 \end{bmatrix} + \begin{bmatrix} -21 \\ 0 \\ 14 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$$ ■

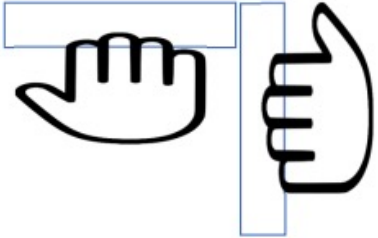
$A\vec{x}$, another view

- Multiplication is valid only for specific matching dimensions!
- Multiply row elements of first by column elements of second, then add

Like this....

and like that!





"**dot product**" (more later)

$A\vec{x} = \vec{b}$, more commonly

Matrix-Vector Multiplication

$A \in R^{M\times N}, \vec{x} \in \mathbb{R}^{N\times 1}$

$$A\vec{x} = \begin{array}{cc} & N \\ M & \begin{bmatrix} a_{11}a_{12}\cdots a_{1N} \\ a_{21}a_{22}\cdots a_{2N} \\ \vdots \\ a_{M1}a_{M2}\cdots a_{MN} \end{bmatrix} \\ & \mathbb{R}^{M\times N} \end{array} \begin{array}{c} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \\ N \times 1 \\ \mathbb{R}^{N\times 1} \end{array} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1N}x_N \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2N}x_N \\ \vdots \\ a_{M1}x_1 + a_{M2}x_2 + \cdots + a_{MN}x_N \end{bmatrix} = \begin{array}{cc} & 1 \\ M & \begin{bmatrix} \\ \\ \\ \end{bmatrix} \end{array}$$

Like this....

and like that!

Note that EE16A notation is x instead of $\vec{x}$, but regardless dimensions are clearly specified.

Thank you, Professor Waller! (EE16A slides)

$A\vec{x} = \vec{b}$, exercise

Before:

EXAMPLE 1

a. $\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -1 \\ 3 \end{bmatrix}$

$= \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ -15 \end{bmatrix} + \begin{bmatrix} -7 \\ 21 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$

b. $\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ -20 \end{bmatrix} + \begin{bmatrix} -21 \\ 0 \\ 14 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$ ■

⇐ original

Like this....

and like that!

Now:

EXAMPLE 5

a. $\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix}\overset{b}{\begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix}} = \begin{bmatrix} 1(4) + 2(3) + {}^{-}1(7) \\ 0(4) + {}^{-}5(3) + 3(7) \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$

b. $\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 7 \end{bmatrix} =$

We just got through one notation!

So far:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 + 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned}$$

1. "augmented matrix"

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{array}\right]$$

This lecture's goal: Get comfortable with these two notations. Also learn some vector properties.

Data science more commonly uses these two notations:

2. **Matrix (and Vector) notation**: ✅

$$A\vec{x} = \vec{b}$$

where $A = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 5 & 0 & -5 \end{bmatrix}, \vec{b} = \begin{bmatrix} 0 \\ 8 \\ 10 \end{bmatrix}, \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

3. **Dot product notation**:

$$\begin{aligned} \vec{a_1}^T \vec{x} &= 0 \\ \vec{a_2}^T \vec{x} &= 8 \\ \vec{a_3}^T \vec{x} &= 10 \end{aligned}$$

where $\vec{a_1} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \vec{a_2} = \begin{bmatrix} 0 \\ 2 \\ -8 \end{bmatrix}, \vec{a_3} = \begin{bmatrix} 5 \\ 0 \\ -5 \end{bmatrix},$

$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Note for now: Two interpretations of the $A\vec{x} = \vec{b}$ matrix equation

$$\underset{m\times n}{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}} \underset{n\times 1}{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}} = \underset{m\times 1}{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}}$$

1. Columns represent how much a particular variable affects all measurements.

This is the "linear combination of vectors" view.

$$\underbrace{x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}}_{x_1\vec{a}_1 + x_2\vec{a}_2 + \ldots + x_n\vec{a}_n} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \ldots + x_n\vec{a}_n = \vec{b}$$

Linear Combination of vectors weighted by the unknowns!

2. Rows represent how much the variables affect a particular measurement.

This is the "system of equations" view.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots a_{2n}x_n &= b_2 \\ \vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \quad & \quad \vdots \\ a_{m1}x_1 + a_{m2}x_1 + \ldots a_{mn}x_n &= b_m \end{aligned}$$

Important!
We will revisit this in Data 100.

(EE16A slides)

Lay 1.4, 2.1

EECS 16A Lec1B

Matrix Multiplication

Optional 01, Data 100 Spring 2023

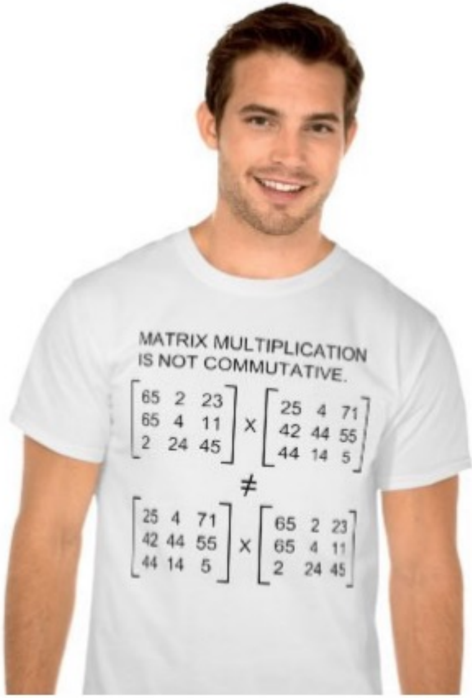
Matrix Multiplication

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1p} \\ b_{21} & & b_{2p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & c_{1p} \\ c_{21} & & c_{2p} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mp} \end{bmatrix}$$

$m \times n$ $\quad$ $n \times p$ $\quad$ $m \times p$

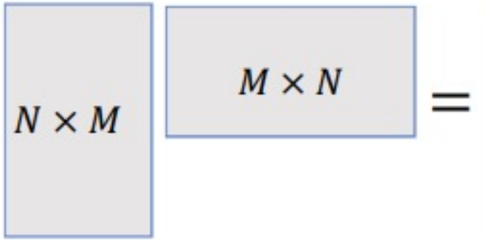
Must be same!

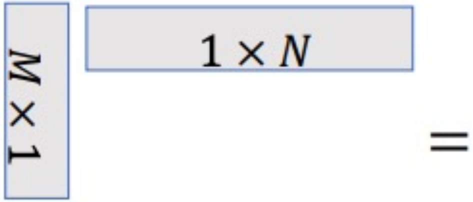
$\mathbb{R}^{m \times n}$ $\quad$ $\mathbb{R}^{n \times p}$ $\quad$ $\mathbb{R}^{m \times p}$

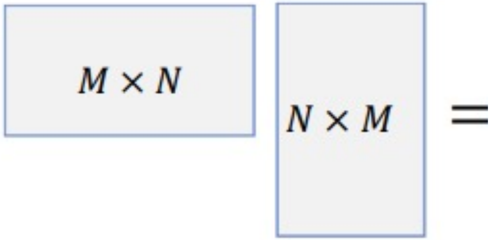


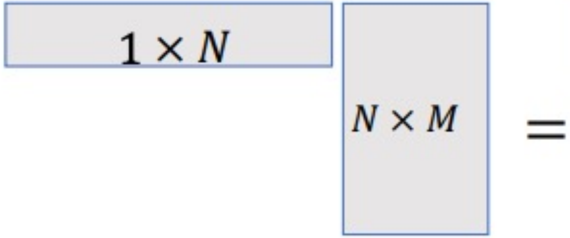
(EE16A slides)

Matrix-Matrix Multiplication









What are the dimensions of the resulting matrices?

(EE16A slides)

Matrix-Matrix Multiplication

row1, col2

$$N \times M \quad M \times N = \begin{bmatrix} \circ & \circ & \\ \circ & & \\ & & \end{bmatrix}$$

$$\mathbb{R}^{N \times M} \quad \mathbb{R}^{M \times N} \qquad \mathbb{R}^{N \times N}$$

$$M \times 1 \quad 1 \times N = \mathbb{R}^{M \times N}$$

$$\mathbb{R}^{M \times 1} \quad \mathbb{R}^{1 \times N}$$

$$M \times N \quad N \times M = \mathbb{R}^{M \times M}$$

$$\mathbb{R}^{M \times N} \quad \mathbb{R}^{N \times M}$$

$$1 \times N \quad N \times M = \mathbb{R}^{1 \times M}$$

$$\mathbb{R}^{1 \times N} \quad \mathbb{R}^{N \times M}$$

Lay 2.1 Example 3

EXAMPLE 3 Compute AB, where $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$.

Like this....

and like that!

① $\mathbb{R}^{2\times 2} \quad \mathbb{R}^{2\times 3} \rightarrow \mathbb{R}^{2\times 3}$ ✓

$2\cdot 4 + 3\cdot 1 \qquad 2\cdot 3 + 3(-2)$

② $\begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}$

Matrix Multiplication is NOT commutative!!!!

$AB \neq BA$

$\mathbb{R}^{2\times3}$

EXAMPLE 3b Compute BA where $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$.

① $\mathbb{R}^{2\times3} \quad \mathbb{R}^{2\times2}$ ✗

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

$2\times2 \qquad 2\times3$

EXAMPLE 7 Let $A = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$. Show that these matrices do not commute. That is, verify that $AB \neq BA$.

SOLUTION

$$AB = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix}$$

$$BA = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 29 & -2 \end{bmatrix}$$

■

Lay 2.1, 6.1

Transpose and Dot Product

Optional 01, Data 100 Spring 2023

Where do we go from here?

So far:

- We have worked out a compact representation of a system of linear equations:

$$A\vec{x} = \vec{b} \leftarrow$$

- We have learned how to perform (some) algebra on matrices and vectors
  - (Really, just matrix multiplication, which is the least intuitive.)
  - (You will learn other properties in Math 54.)

What next?

- There is one more ~~notation~~ operation that is incredibly important for data science: the **transpose**.
- With the transpose, we can define a lot of new measurements for vectors.

Transpose of a Matrix

Matrix transpose

→ swap the rows with the columns

$$A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \longrightarrow A^T = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \longrightarrow B^T = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

What are the dimensions? $\underline{X \in \mathbb{R}^{N \times M}}$ $\quad X^T \in \mathbb{R}^{M \times N}$

If the elements of the matrix $A \in \mathbb{R}^{N \times M}$ are a_{ij}
The elements of $A^T \in \mathbb{R}^{M \times N}$ are a_{ji}
Matrix transpose is not (generally) an inverse! (more next time on inverses)

(EE16A slides)

Practice: What are the transposes? What Euclidean space is the resulting transposed matrix in?

(Lay 2.1, Example 8)

Let

$$\underset{\mathbb{R}^{3\times 2}}{B} = \begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \end{bmatrix}$$

$$\underset{\mathbb{R}^{2\times 3}}{B^T} = \begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix}$$

$$C^T = \begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix}$$

(Lay 6.1, Example 1)

Let

$$\mathbb{R}^{3\times 1}\ \underset{\vec{u}}{\mathbf{u}} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} \text{ and } \mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}.$$

$$\underset{\mathbb{R}^{1\times 3}}{\vec{u}^T} = \begin{bmatrix} \underline{2} & \underline{-5} & \underline{-1} \end{bmatrix} \qquad \vec{v}^T = \begin{bmatrix} 3 & 2 & -3 \end{bmatrix}$$

Yes, vectors have transposes too!!!

Poll

Given $A \in \mathbb{R}^{m \times p}$ and $B \in \mathbb{R}^{n \times p}$, which of the following is a valid multiplication? Assume $m \neq n \neq p \neq m$.

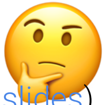
A. AB — $m \times p \quad n \times p$ ✗

B. BA — $n \times p \quad m \times p$ ✗

C. $A^T B$ — $p \times m \quad n \times p$ ✗

D. AB^T — $m \times p \quad p \times n$ ✓

Left mat — Right mat



(EE16A slides)

Dot products: From $\mathbb{R}^n$ to $\mathbb{R}$

If $\vec{u}, \vec{v} \in \mathbb{R}^n$ ("are vectors in $\mathbb{R}^n$"), then their **dot product** is defined as:

$$u \cdot v = \vec{u}^T \vec{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n \in \mathbb{R}$$

EXAMPLE 1 Compute $\mathbf{u} \cdot \mathbf{v}$ and $\mathbf{v} \cdot \mathbf{u}$ for $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$.

$u^T v$

SOLUTION

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \begin{bmatrix} 2 & -5 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} = (2)(3) + (-5)(2) + (-1)(-3) = -1$$

$$\mathbf{v} \cdot \mathbf{u} = \mathbf{v}^T \mathbf{u} = \begin{bmatrix} 3 & 2 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} = 3(2) + 2(-5) + (-3)(-1) = -1$$

Dot products are commutative (unlike matrix products), i.e.,

$$\vec{u}^T \vec{v} = \vec{v}^T \vec{u}$$

Dot products: From $\mathbb{R}^n$ to $\mathbb{R}$

If $\vec{u}, \vec{v} \in \mathbb{R}^n$ ("are vectors in $\mathbb{R}^n$"), then their **dot product** is defined as:

$$u \cdot v = \vec{u}^T \vec{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

We will most commonly use this notation in Data 100.

EXAMPLE 1 ... $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$.

Dot products are commutative (unlike matrix products), i.e.,
$\vec{u}^T \vec{v} = \vec{v}^T \vec{u}$

SOLUTION

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \begin{bmatrix} 2 & -5 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} = (2)(3) + (-5)(2) + (-1)(-3) = -1$$

$$\mathbf{v} \cdot \mathbf{u} = \mathbf{v}^T \mathbf{u} = \begin{bmatrix} 3 & 2 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} = (3)(2) + (2)(-5) + (-3)(-1) = -1$$

Let's practice dot products!

1. What is the dot product of $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$?

2. Is it possible to take the dot product of $\begin{bmatrix} 7 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ 0 \\ 5 \end{bmatrix}$?

Define $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, where $\vec{x} \in \mathbb{R}^3$.

3. How do we represent $x_1 - 2x_2 + x_3$ as a dot product involving $\vec{x}$?

4. How do we represent $2x_2 + 8x_3 = 8$ as an equation involving $\vec{x}$? (Be careful!!)

Let's practice dot products!

$\mathbb{R}^{1\times 1}$ $\mathbb{R}^1$ $\mathbb{R}$

1. What is the dot product of $\begin{bmatrix}1\\2\end{bmatrix}$ and $\begin{bmatrix}2\\3\end{bmatrix}$?

① $\begin{bmatrix}1 & 2\end{bmatrix}\begin{bmatrix}2\\3\end{bmatrix} = 1(2)+2(3)$

$= 8$.

$[8]$?

2. Is it possible to take the dot product of $\begin{bmatrix}7\\3\end{bmatrix}$ and $\begin{bmatrix}4\\0\\5\end{bmatrix}$?

② $\begin{bmatrix}7 & 3\end{bmatrix}\begin{bmatrix}4\\0\\5\end{bmatrix}$

NO

Define $\vec{x} = \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$, where $\vec{x} \in \mathbb{R}^3$.

3. How do we represent $\underbrace{x_1 - 2x_2 + x_3}$ as a dot product involving $\vec{x}$?

$\begin{bmatrix}1 & -2 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = x_1 - 2x_2 + x_3$

4. How do we represent $2x_2 + 8x_3 = 8$ as an equation involving $\vec{x}$? (Be careful!!)

$\begin{bmatrix}0 & 2 & 8\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$

Looks familiar…

So far:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 + 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned}$$

1. "augmented matrix"

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{array}\right]$$

This lecture's goal: Get comfortable with these two notations. Also learn some vector properties.

Data science more commonly uses these two notations:

2. **Matrix (and Vector) notation**: ✅

$$A\vec{x} = \vec{b}$$

where $A = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 5 & 0 & -5 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 0 \\ 8 \\ 10 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

3. **Dot product notation**: ?

$$\begin{aligned} \vec{a_1}^T \vec{x} &= 0 \\ \vec{a_2}^T \vec{x} &= 8 \\ \vec{a_3}^T \vec{x} &= 10 \end{aligned}$$

where $\vec{a_1} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, $\vec{a_2} = \begin{bmatrix} 0 \\ 2 \\ -8 \end{bmatrix}$, $\vec{a_3} = \begin{bmatrix} 5 \\ 0 \\ -5 \end{bmatrix}$,

$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Looks familiar…

So far:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 + 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned}$$

$\vec{a}_1^T \vec{x} = 0$

3. **Dot product notation**:

?

$$\begin{aligned} \overrightarrow{a_1}^T \vec{x} &= 0 \\ \overrightarrow{a_2}^T \vec{x} &= 8 \\ \overrightarrow{a_3}^T \vec{x} &= 10 \end{aligned}$$

where $\overrightarrow{a_1} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \overrightarrow{a_2} = \begin{bmatrix} 0 \\ 2 \\ -8 \end{bmatrix}, \overrightarrow{a_3} = \begin{bmatrix} 5 \\ 0 \\ -5 \end{bmatrix},$

$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

We just got through another notation!!!!!

So far:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 + 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned}$$

1. "augmented matrix"

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{array}\right]$$

This lecture's goal: Get comfortable with these two notations. Also learn some vector properties.

Data science more commonly uses these two notations:

2. **Matrix (and Vector) notation**: ✅

$$A\vec{x} = \vec{b}$$

where $A = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 5 & 0 & -5 \end{bmatrix}, \vec{b} = \begin{bmatrix} 0 \\ 8 \\ 10 \end{bmatrix}, \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

3. **Dot product notation**: ✅

$$\begin{aligned} \vec{a_1}^T \vec{x} &= 0 \\ \vec{a_2}^T \vec{x} &= 8 \\ \vec{a_3}^T \vec{x} &= 10 \end{aligned}$$

where $\vec{a_1} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \vec{a_2} = \begin{bmatrix} 0 \\ 2 \\ -8 \end{bmatrix}, \vec{a_3} = \begin{bmatrix} 5 \\ 0 \\ -5 \end{bmatrix},$

$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

We just got through another notation!!!!!

So far:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 + 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned}$$

1. "augmented matrix"

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{array}\right]$$

This lecture's goal: Get comfortable with these two notations. Also learn some vector properties.

Data science more ... only uses the

2. **Matrix (and V... ...tion**:

... e A ... $, \vec{b} = \begin{bmatrix} 0 \\ 8 \\ 10 \end{bmatrix}, \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

3. **Dot product notation**:

$$\vec{a_1}^T \vec{x} = 0$$
$$\vec{a_2}^T \vec{x} = 8$$
$$\vec{a_3}^T \vec{x} = 10$$

where $\vec{a_1} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \vec{a_2} = \begin{bmatrix} 0 \\ 2 \\ -8 \end{bmatrix}, \vec{a_3} = \begin{bmatrix} 5 \\ 0 \\ -5 \end{bmatrix},$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Next up: Vector (and dot product) properties

So far:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 + 8x_3 &= 8 \\ 5x_1 - 5x_3 &= 10 \end{aligned}$$

1. "augmented matrix"

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{array}\right]$$

This lecture's goal: Get comfortable with these two notations. **Also learn some vector properties.**

Data science more commonly uses these two notations:

2. **Matrix (and Vector) notation**:

$$A\vec{x} = \vec{b}$$

where $A = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ 5 & 0 & -5 \end{bmatrix}, \vec{b} = \begin{bmatrix} 0 \\ 8 \\ 10 \end{bmatrix}, \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

3. **Dot product notation**:

$$\begin{aligned} \overrightarrow{a_1}^T \vec{x} &= 0 \\ \overrightarrow{a_2}^T \vec{x} &= 8 \\ \overrightarrow{a_3}^T \vec{x} &= 10 \end{aligned}$$

where $\overrightarrow{a_1} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \overrightarrow{a_2} = \begin{bmatrix} 0 \\ 2 \\ -8 \end{bmatrix}, \overrightarrow{a_3} = \begin{bmatrix} 5 \\ 0 \\ -5 \end{bmatrix},$

$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Lay 2.9, 6.1

Dot product properties, Norms

Optional 01, Data 100 Spring 2023

Properties of Dot Product (Math 91, 6.1) Lay

Properties of the dot product

$u^T v = v^T u$

For any real n-vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, and scalar $c \in \mathbb{R}$, the dot product satisfies the following:

(1) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$.
(2) $\mathbf{v} \cdot (c\mathbf{w}) = c(\mathbf{v} \cdot \mathbf{w})$.
(3) $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$.
(4) $\mathbf{v} \cdot \mathbf{v} \geq 0$, and $\mathbf{v} \cdot \mathbf{v} = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

These properties look scary at first glance!

But really what they tell you that the dot product works like “normal” algebra.

This means PEMDAS properties hold!

Properties and intuition

Let $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \vec{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} -5 \\ 1 \end{bmatrix}$

1. What is $(2\vec{u})^T(-3\vec{v})$?

2. What is $(2\vec{u})^T(-3\vec{v} + 4\vec{w})$?

(It is okay to compute from first principles)

Properties and intuition

Let $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \vec{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} -5 \\ 1 \end{bmatrix}$

1. What is $(2\vec{u})^T(-3\vec{v})$?

① $2\vec{u} = \begin{bmatrix} 2 \\ 8 \end{bmatrix}, -3\vec{v} = \begin{bmatrix} 0 \\ -3 \end{bmatrix}$

$\begin{bmatrix} 2 & 8 \end{bmatrix} \begin{bmatrix} 0 \\ -3 \end{bmatrix}$

$= -24$

② $\overbrace{2 \cdot -3}^{\text{multiply scalars}} \underbrace{(\vec{u}^T \vec{v})}_{}$ dot vectors

$-6 \left(\begin{bmatrix} 1 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)$

$= -6(4) = -24$

2. What is $(2\vec{u})^T(-3\vec{v} + 4\vec{w})$?

① $\underline{-3\vec{v} + 4\vec{w}} = \begin{bmatrix} -20 \\ 1 \end{bmatrix}$

$\begin{bmatrix} 2 & 8 \end{bmatrix} \begin{bmatrix} -20 \\ 1 \end{bmatrix} = -40 + 8$

$= -32$

② $(2 \cdot -3)\vec{u}^T\vec{v} + (2 \cdot 4)\vec{u}^T\vec{w}$

$-6\vec{u}^T\vec{v} + 8\vec{u}^T\vec{w}$

$-24 + 8 \begin{bmatrix} 1 & 4 \end{bmatrix} \begin{bmatrix} -5 \\ 1 \end{bmatrix}$

$-24 + 8(-1) = \boxed{-32}$

Properties of the dot product

For any real n-vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, and scalar $c \in \mathbb{R}$, the dot product satisfies the following:

(1) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$.

(2) $\mathbf{v} \cdot (c\mathbf{w}) = c(\mathbf{v} \cdot \mathbf{w})$.

(3) $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$.

(4) $\underline{\mathbf{v} \cdot \mathbf{v} \geq 0}$, and $[\mathbf{v} \cdot \mathbf{v} = 0$ if and only if $\mathbf{v} = \mathbf{0}.]$

zero vector

$\vec{0} \in \mathbb{R}^n$

$$\underbrace{\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}}\Big\} n$$

Let's prove Property 4. Suppose we have vector

$v^T v$

$$\mathbf{v} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

$[a_1 \quad \cdots \quad a_n] \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$

we calculate from the definition

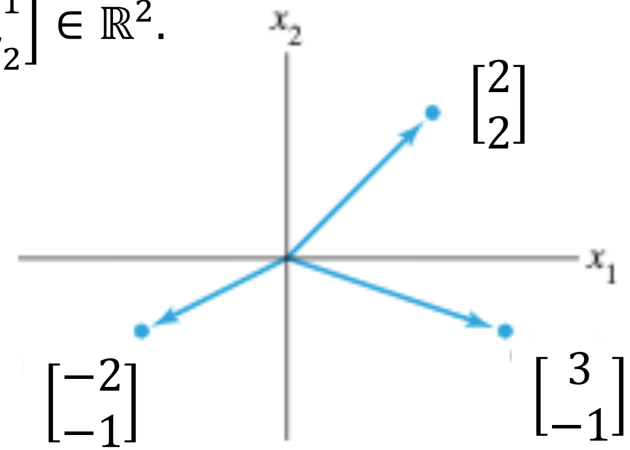
$$\mathbf{v} \cdot \mathbf{v} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = a_1^2 + \cdots + a_n^2$$

$a_i \in \mathbb{R}$

For any real number a, its square satisfies $a^2 \geq 0$, with $a^2 = 0$ if and only if $a = 0$. So we conclude $\mathbf{v} \cdot \mathbf{v} = a_1^2 + \cdots a_n^2 \geq 0$, with $\mathbf{v} \cdot \mathbf{v} = a_1^2 + \cdots a_n^2 = 0$ if and only if $a_1 = \cdots = a_n = 0$, so if and only if $\mathbf{v} = \mathbf{0}$. $\square$

Reminder: Geometry

Let $\vec{\mathrm{x}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$.



Reminder: Linear "Algebra" is actually also Linear "Geometry."

L2 Norm Properties (Math 91, 6.1)

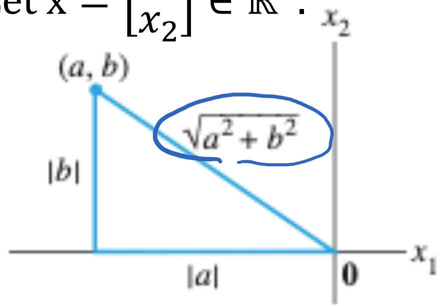
The **length** (or **norm**) of $\mathbf{v}$ is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \boxed{\sqrt{\mathbf{v}\cdot\mathbf{v}}} = \underline{\sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}}, \quad \text{and} \quad \boxed{\|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}}$$

This is the more important definition for us!!

$$\|v\| = \sqrt{v^T v}$$

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$.

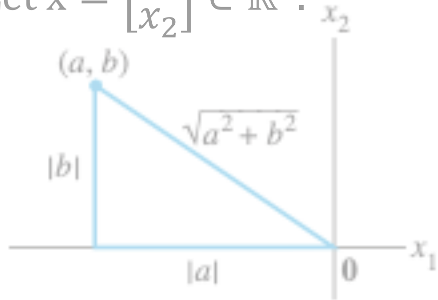


Defining the L2 Norm

The **length** (or **norm**) of $\mathbf{v}$ is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}, \quad \text{and} \quad \|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}$$

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$.



$\vec{w} + (\vec{v} - \vec{w}) = \vec{v}$

Example 1.6.13

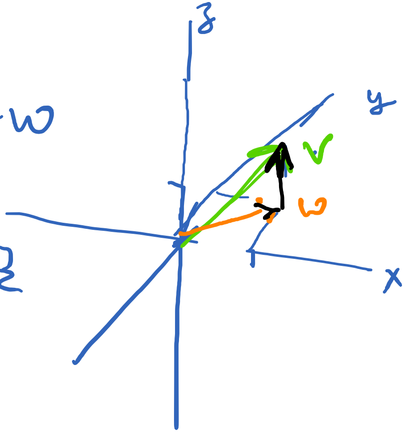
For the vectors

$$\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \in \mathbb{R}^3$$

we have

$$\|\mathbf{v}\| = \sqrt{3}, \|\mathbf{w}\| = \sqrt{2}, \|\mathbf{v} - \mathbf{w}\| = 1$$

$V - W$



$\sqrt{1^2+1^2+1^2}$

$= \sqrt{3}$

$\sqrt{1^2+1^2+0}$

$\sqrt{2}$

$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \rightarrow \sqrt{0^2+0^2+1^2}$

$= 1$

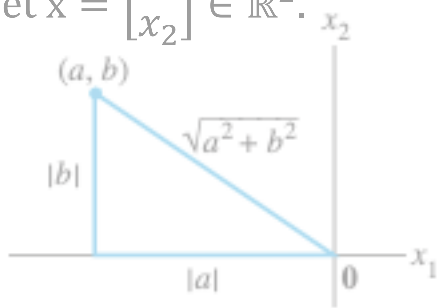
L2 Norm Properties (Math 91, 6.1)

The **length** (or **norm**) of $\mathbf{v}$ is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}, \quad \text{and} \quad \boxed{\|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}}$$

This is the more important definition for us!!

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$.



1.6.15: Properties of length

For any vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, and scalar $c \in \mathbb{R}$, we have

(1) $\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$.

(2) $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

$\vec{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ $\quad \|\vec{v}\|_2 = \sqrt{3}$

↑ more in a bit

norm of $-5\vec{v} \to 5\sqrt{3}$?

$\begin{bmatrix} -5 \\ -5 \\ -5 \end{bmatrix}$ $\quad \|5\vec{v}\| = \sqrt{(-5)^2+(-5)^2+(-5)^2}$

$= \sqrt{3(-5)^2}$

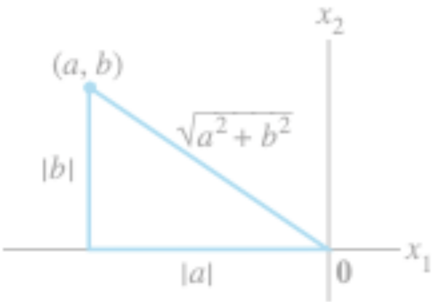
$= |-5|\sqrt{3} = 5\sqrt{3}$

Defining the L2 Norm

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$.

The **length** (or **norm**) of $\mathbf{v}$ is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}, \quad \text{and} \quad \|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}$$



Example 1.6.13

For the vectors

$$\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \in \mathbb{R}^3$$

we have

$$\|\mathbf{v}\| = \sqrt{3}, \|\mathbf{w}\| = \sqrt{2}, \|\mathbf{v} - \mathbf{w}\| = 1$$

The norm is a good way to represent vector "length". It's also a good way to represent an N-D vector by a single number.

Math 54 only defines this "Pythagorean Theorem" norm, which Data 100 calls the **L2 norm** (notation $\|\vec{v}\|_2$). Other types of norms exist (e.g., we will cover the L1 norm).

Lay 1.3, 6.1, 6.2, 6.3

Orthogonality and Span

Optional 01, Data 100 Spring 2023

Reminder: Span

If $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are in $\mathbb{R}^n$, then the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ is denoted by Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ and is called the **subset of $\mathbb{R}^n$ spanned** (or **generated**) **by** $\mathbf{v}_1, \ldots, \mathbf{v}_p$. That is, Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is the collection of all vectors that can be written in the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

with $c_1, \ldots, c_p$ scalars.

A Geometric Description of Span$\{\mathbf{v}\}$ and Span$\{\mathbf{u}, \mathbf{v}\}$

Let $\mathbf{v}$ be a nonzero vector in $\mathbb{R}^3$. Then Span $\{\mathbf{v}\}$ is the set of all scalar multiples of $\mathbf{v}$, which is the set of points on the line in $\mathbb{R}^3$ through $\mathbf{v}$ and $\mathbf{0}$. See Figure 10.

If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in $\mathbb{R}^3$, with $\mathbf{v}$ not a multiple of $\mathbf{u}$, then Span $\{\mathbf{u}, \mathbf{v}\}$ is the plane in $\mathbb{R}^3$ that contains $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{0}$. In particular, Span $\{\mathbf{u}, \mathbf{v}\}$ contains the line in $\mathbb{R}^3$ through $\mathbf{u}$ and $\mathbf{0}$ and the line through $\mathbf{v}$ and $\mathbf{0}$. See Figure 11.

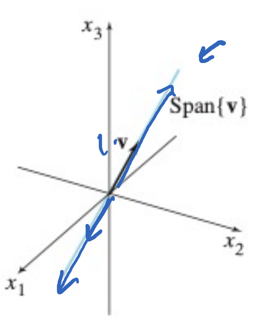


FIGURE 10 Span $\{\mathbf{v}\}$ as a line through the origin.

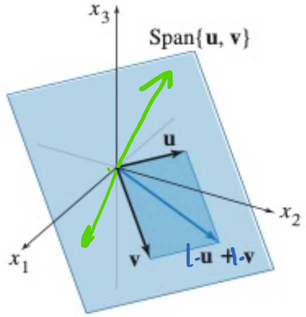


FIGURE 11 Span $\{\mathbf{u}, \mathbf{v}\}$ as a plane through the origin.

Span Example (Lay 1.3, extended)

EXAMPLE 6 Let $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix}$.

Span $\{\mathbf{a}_1, \mathbf{a}_2\}$ is a plane through the origin in $\mathbb{R}^3$.

$$\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} x_1 + \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix} x_2 = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix}$$

1. Let $\vec{b} = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix}$. Is $\vec{b} \in \text{Span}\{\vec{a_1}, \vec{a_2}\}$?

SOLUTION Does the equation $\boxed{x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}}$ have a solution? To answer this, row reduce the augmented matrix $[\,\mathbf{a}_1 \;\; \mathbf{a}_2 \;\; \mathbf{b}\,]$:

$$\left[\begin{array}{rr:r} 1 & 5 & -3 \\ -2 & -13 & 8 \\ 3 & -3 & 1 \end{array}\right] \sim \begin{bmatrix} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & -18 & 10 \end{bmatrix} \sim \left[\begin{array}{rr:r} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & 0 & -2 \end{array}\right]$$

$x_1 + 5x_2 = -3$

$-3x_2 = 2$

$0 \neq -2$

$0 = -2$

The third equation is $0 = -2$, which shows that the system has no solution. The vector equation $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}$ has no solution, and so $\mathbf{b}$ is *not* in Span $\{\mathbf{a}_1, \mathbf{a}_2\}$. ■

2. Let $\vec{c} = \begin{bmatrix} -4 \\ 11 \\ 6 \end{bmatrix}$. Is $\vec{c} \in \text{Span}\{\vec{a_1}, \vec{a_2}\}$?

$$\vec{c} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} - \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix} = \begin{bmatrix} -4 \\ 11 \\ 6 \end{bmatrix}$$

$1 \cdot a_1 + (-1) a_2$

yes

Spans of Columns of a Matrix

If $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are in $\mathbb{R}^n$, then the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ is denoted by Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ and is called the **subset of $\mathbb{R}^n$ spanned** (or **generated**) **by** $\mathbf{v}_1, \ldots, \mathbf{v}_p$. That is, Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is the collection of all vectors that can be written in the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

with $c_1, \ldots, c_p$ scalars.

A Geometric Description of Span $\{\mathbf{v}\}$ and Span $\{\mathbf{u}, \mathbf{v}\}$

Let $\mathbf{v}$ be a nonzero vector in $\mathbb{R}^3$. Then Span $\{\mathbf{v}\}$ is the set of all scalar multiples of $\mathbf{v}$, which is the set of points on the line in $\mathbb{R}^3$ through $\mathbf{v}$ and $\mathbf{0}$. See Figure 10.

If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in $\mathbb{R}^3$, with $\mathbf{v}$ not a multiple of $\mathbf{u}$, then Span $\{\mathbf{u}, \mathbf{v}\}$ is the plane in $\mathbb{R}^3$ that contains $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{0}$. In particular, Span $\{\mathbf{u}, \mathbf{v}\}$ contains the line in $\mathbb{R}^3$ through $\mathbf{u}$ and $\mathbf{0}$ and the line through $\mathbf{v}$ and $\mathbf{0}$. See Figure 11.

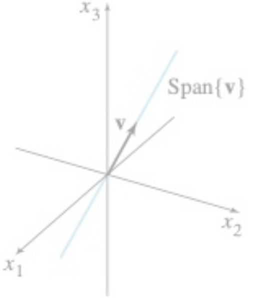


FIGURE 10 Span $\{\mathbf{v}\}$ as a line through the origin.

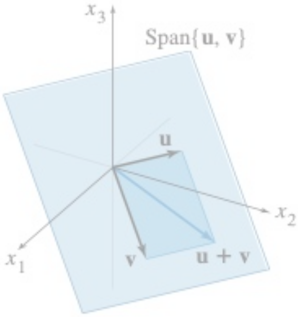
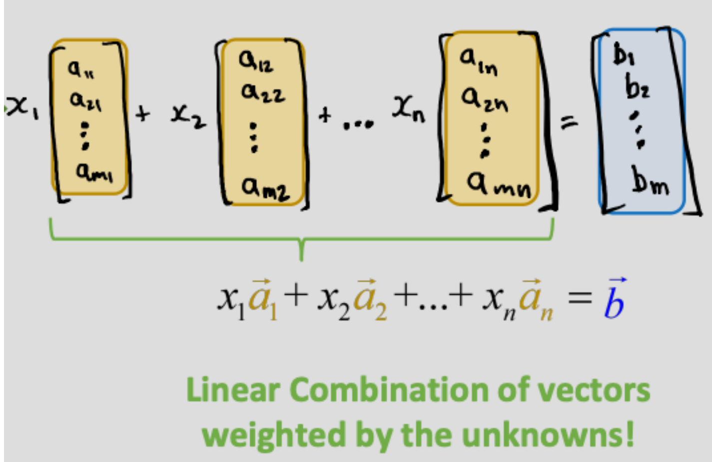


FIGURE 11 Span $\{\mathbf{u}, \mathbf{v}\}$ as a plane through the origin.

Recall the "linear combination of vectors" view of $A\vec{x} = \vec{b}$:

$A = [\vec{a}_1\ \vec{a}_2\ \ \vec{a}_n]$



The **span** of the **columns of A** is the set of all vectors that can be reached by all possible linear combinations of the columns of A.

Continuing our Span example

EXAMPLE 6 Let $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix}$.

Span $\{\mathbf{a}_1, \mathbf{a}_2\}$ is a plane through the origin in $\mathbb{R}^3$.

Define $A = \begin{bmatrix} 1 & 5 \\ -2 & -13 \\ 3 & -3 \end{bmatrix}$.

$\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ $\begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix}$

1. Let $\vec{b} = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix}$. Is $\vec{b} \in \mathrm{Span}\{\overrightarrow{a_1}, \overrightarrow{a_2}\}$?

No.

$\vec{b}$ **is not** in the span of the columns of A.

2. Let $\vec{c} = \begin{bmatrix} -4 \\ 11 \\ 6 \end{bmatrix}$. Is $\vec{c} \in \mathrm{Span}\{\overrightarrow{a_1}, \overrightarrow{a_2}\}$?

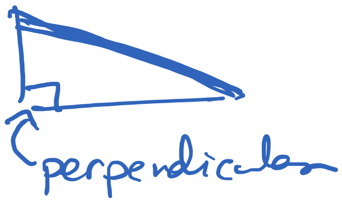
Yes.

$\vec{c}$ **is** in the span of the columns of A.

Orthogonality links Geometry to Algebra

To "compare" two vectors:

- Draw them in a geometric space.
- Do some algebraic manipulation with them.



The dot product links algebra and geometry together via a concept called **orthogonality**.

Orthogonality links Geometry to Algebra

To "compare" two vectors:

- Draw them in a geometric space.
- Do some algebraic manipulation with them.

The dot product links algebra and geometry together via a concept called **orthogonality**.

Two vectors **u** and **v** in $\mathbb{R}^n$ are **orthogonal** (to each other) if $\mathbf{u} \cdot \mathbf{v} = 0$.

The Pythagorean Theorem

Two vectors **u** and **v** are orthogonal if and only if $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

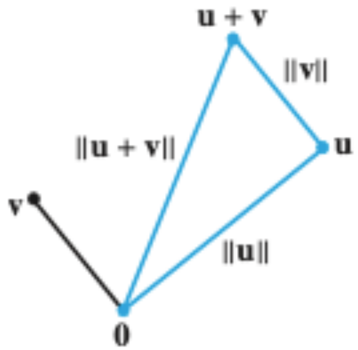


FIGURE 6

(Math 91)

Orthogonality Exercises

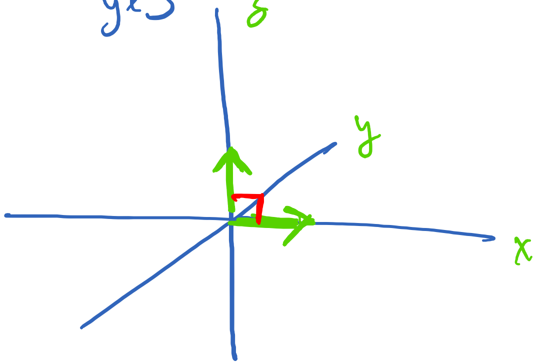
Two vectors **u** and **v** in $\mathbb{R}^n$ are **orthogonal** (to each other) if $\mathbf{u} \cdot \mathbf{v} = 0$.

Are the following vectors orthogonal to each other?

1. $\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \vec{v} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

$[1 \ \ 0 \ \ 0] \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$

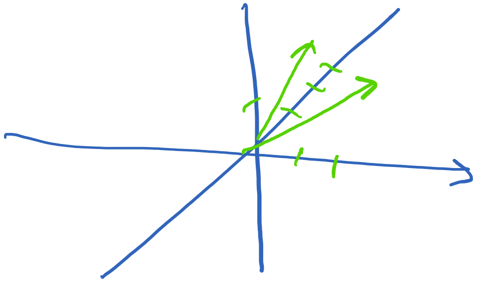
yes



2. $\vec{u} = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}, \vec{v} = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$

$[2 \ \ 3 \ \ 0] \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix} = 9$

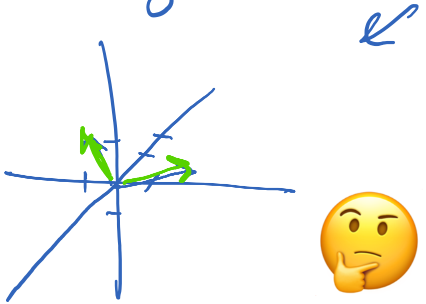
no



3. $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \vec{v} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

$[1 \ \ 2 \ \ -1] \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = 0$

yes



Orthogonality Exercises

Two vectors **u** and **v** in $\mathbb{R}^n$ are **orthogonal** (to each other) if $\mathbf{u} \cdot \mathbf{v} = 0$.

Are the following vectors orthogonal to each other?

1. $\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \vec{v} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

2. $\vec{u} = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}, \vec{v} = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$

3. $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \vec{v} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

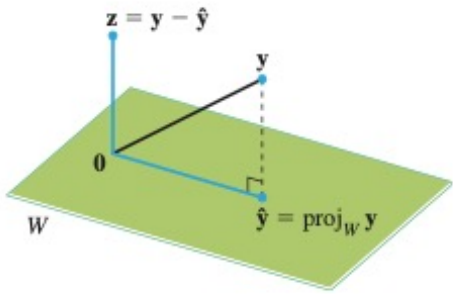


FIGURE 2
Finding α to make $\mathbf{y} - \hat{\mathbf{y}}$ orthogonal to $\mathbf{u}$.

An Orthogonal Projection

Given a nonzero vector $\mathbf{u}$ in $\mathbb{R}^n$, consider the problem of decomposing a vector $\mathbf{y}$ in $\mathbb{R}^n$ into the sum of two vectors, one a multiple of $\mathbf{u}$ and the other orthogonal to $\mathbf{u}$. We wish to write

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

where $\hat{\mathbf{y}} = \alpha\mathbf{u}$ for some scalar α and $\mathbf{z}$ is some vector orthogonal to $\mathbf{u}$. See Figure 2.

The vector $\hat{\mathbf{y}}$ is called the **orthogonal projection of y onto u**, and the vector $\mathbf{z}$ is called the **component of y orthogonal to u**.

$$\boxed{\hat{\mathbf{y}} = \text{proj}_L\, \mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}} \tag{2}$$

Let $\vec{u} = \begin{bmatrix}1\\0\\0\end{bmatrix}$. Let $\vec{y} = \begin{bmatrix}2\\3\\4\end{bmatrix}$.

(Lay 6.2)

Vector = (orthogonal projection) + (orthogonal vector)

We'll revisit this next time

The Orthogonal Decomposition Theorem

Let W be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely in the form

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \qquad (1)$$

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^{\perp}$. In fact, if $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is any orthogonal basis of W, then

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\mathbf{u}_p \qquad (2)$$

and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

$\hat{y}$: **orthogonal projection** of $\boldsymbol{y}$ onto W

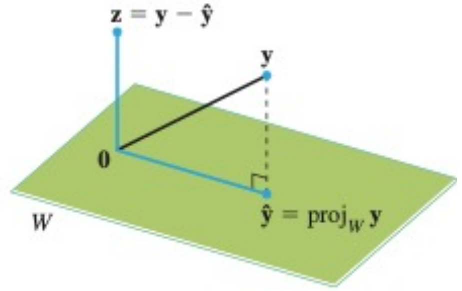
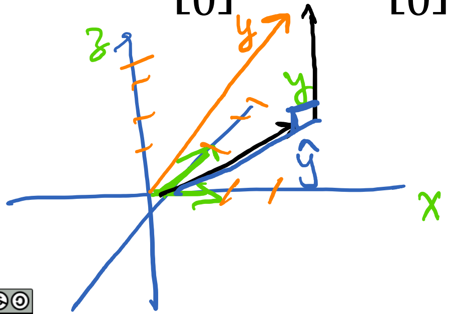


FIGURE 2 Finding α to make $\mathbf{y} - \hat{\mathbf{y}}$ orthogonal to $\mathbf{u}$.

There's a lot of terminology here, but here's the general idea:

Let $\overrightarrow{u_1} = \begin{bmatrix}1\\0\\0\end{bmatrix}, \overrightarrow{u_2} = \begin{bmatrix}0\\1\\0\end{bmatrix}$. Let $\vec{y} = \begin{bmatrix}2\\3\\4\end{bmatrix}$.



$z \perp u_2$

$z \perp u_1$

$$y = \begin{bmatrix}2\\3\\4\end{bmatrix} = \begin{bmatrix}2\\3\\0\end{bmatrix} + \overbrace{\begin{bmatrix}0\\0\\4\end{bmatrix}}$$

$\hat{y} \in \text{span}\{u_1, u_2\}$

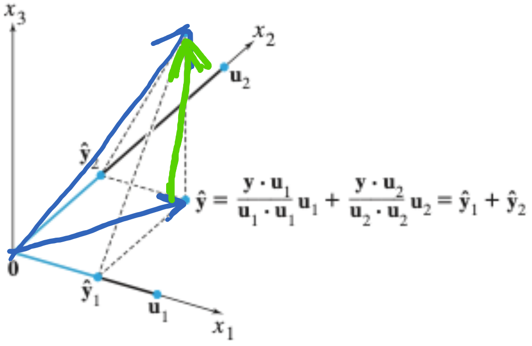


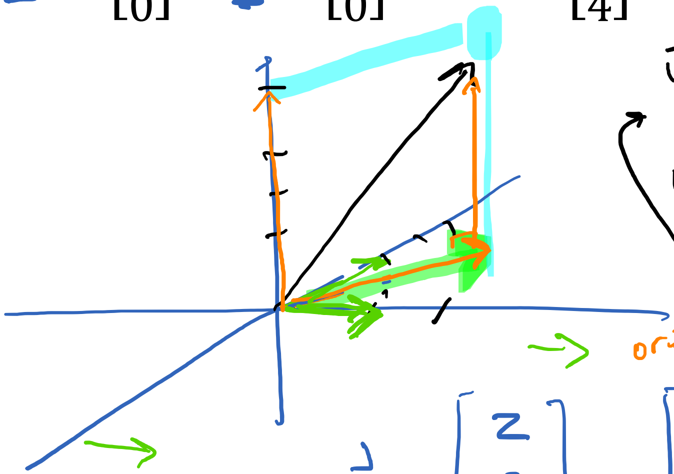
FIGURE 3 The orthogonal projection of $\mathbf{y}$ is the sum of its projections onto one-dimensional subspaces that are mutually orthogonal.

Vector = (orthogonal projection) + (orthogonal vector)

$\hat{y}$: **orthogonal projection** of y onto W

$W = \text{span}\{u_1, u_2\}$

Let $\vec{u_1} = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $\vec{u_2} = \begin{bmatrix}0\\1\\0\end{bmatrix}$. Let $\vec{y} = \begin{bmatrix}2\\3\\4\end{bmatrix}$.



$\vec{u_1}^T z = [1\ 0\ 0]\begin{bmatrix}0\\0\\4\end{bmatrix} = 0$

$\vec{u_2}^T z = [0\ 1\ 0]\begin{bmatrix}0\\0\\4\end{bmatrix} = 0$

orthogonal vector to span$\{u_1, u_2\}$

$$\vec{y} = \begin{bmatrix}2\\3\\0\end{bmatrix} + \underbrace{\begin{bmatrix}0\\0\\4\end{bmatrix}}_{z} = \begin{bmatrix}2\\3\\4\end{bmatrix}$$

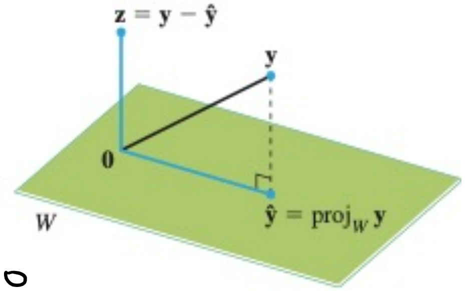


FIGURE 2 Finding α to make $\mathbf{y} - \hat{\mathbf{y}}$ orthogonal to $\mathbf{u}$.

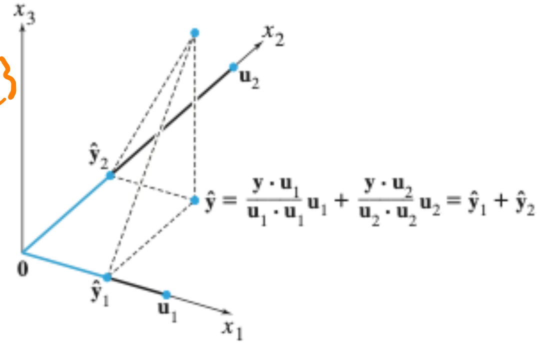


FIGURE 3 The orthogonal projection of $\mathbf{y}$ is the sum of its projections onto one-dimensional subspaces that are mutually orthogonal.

Orthogonal Projection Examples

Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

What is the orthogonal projection of $\vec{y}$ onto the spans of columns of A if:

1. $\vec{y} = \begin{bmatrix} -2 \\ 1 \\ -4 \end{bmatrix}$

2. $\vec{y} = \begin{bmatrix} -2 \\ -3 \\ 0 \end{bmatrix}$

3. $\vec{y} = \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}$

Orthogonal Projection Examples

Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

What is the orthogonal projection of $\vec{y}$ onto the spans of columns of A if:

1. $\vec{y} = \begin{bmatrix} -2 \\ 1 \\ -4 \end{bmatrix}$

2. $\vec{y} = \begin{bmatrix} -2 \\ -3 \\ 0 \end{bmatrix}$

3. $\vec{y} = \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}$

Orthogonal Projection Takeaway

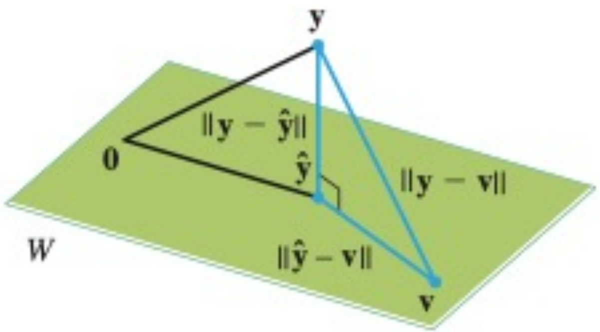


FIGURE 4 The orthogonal projection of **y** onto W is the closest point in W to **y**.

Big geometric takeaway:

Vector = **orthogonal projection** + **orthogonal vector**

OPTIONAL 01

Linear Algebra 01

Content credit: The wealth of instructional knowledge from Math 54, Math 91, and EECS16A