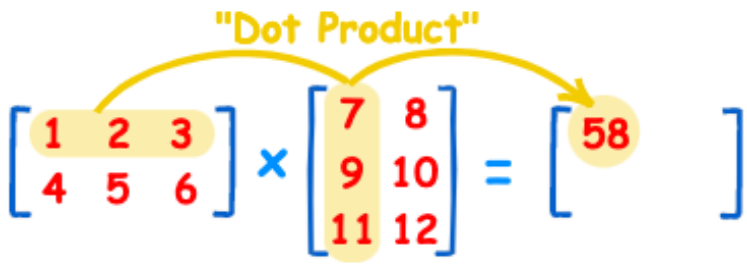
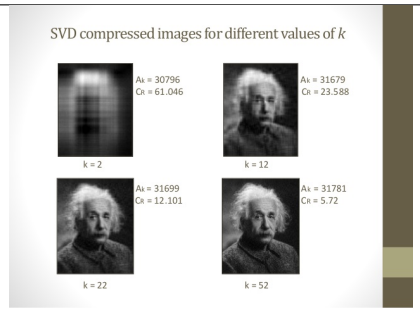
Math 210 Mathematics for data science - Fall 2021

Introduction and Course Logistics

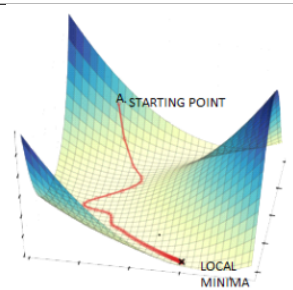
Topics



BASIC LINEAR ALGEBRA



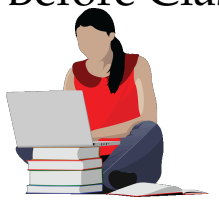
MORE LINEAR ALGEBRA



DERIVATIVES AND OPTIMIZATION

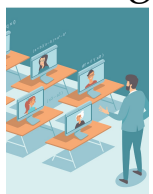
Class structure

Before Class



- Check Sakai for instructions
- Watch short video(s)
- Do before-class assignment on Edfinity

During Class



- Take notes using skeleton notes posted on Sakai
- Work problems in groups
- Ask questions and answer questions

After Class

- Check Sakai for filled in notes
- Read the book
- Work homework problems on Edfinity
- Attend office hours and ask questions on Piazza

Help

Here are some ways to get help for this class:

- Office hours:
- Piazza
- Math Help Center: free online tutoring Monday - Saturday, in person and on zoom
- Peer Tutoring from the Learning Center: usually Tu, W, Th evenings or by appointment
- Private tutor list - see math department list of grad students who want to tutor for pay
- Peers

Honor Code

For homework, in-class problems, and before-class assignments

- Books and calculators are allowed
- Use of apps and software including Python is encouraged for checking answers, and where specified, to perform parts of computations
- Please resist the temptation to overuse technology: take care to do enough computations by hand to thoroughly master all algorithms
- Collaboration with other students is encouraged
- Collaboration should benefit the learning of all involved
- Copying or trading answers is a violation of the honor code

For tests

- Collaboration is not allowed
- Closed book and notes unless otherwise specified
- Calculator and computer use is not allowed unless otherwise specified

If in doubt about what is permitted on an assignment, please ask!

More on Class Logistics

Masks

Tour of Sakai

Questions about class logistics

Poll about interest and math background

Linear Equations - Introduction

After completing this section, students should be able to:

- Distinguish linear vs. non-linear equations
- Identify solutions of systems of linear equations by plugging them in
- Convert between a system of linear equations and an augmented matrix

Example. The equation

is a linear equation.

Definition. A *linear equation* is a an equation that can be written as ...

Example. Which of the following are linear equations?

A. $4x^2 + 5x + 6 = 0$

B. $\frac{1}{3}uv + \frac{2}{3}u = \frac{3}{5}v$

C. $5a + 2 = 6b - \sqrt{2}c$

Definition. A *system of linear equations* is ...

For example,

Definition. A solution to a system of linear equations is ...

For example,

is a solution to the system of equations.

$$x + y = 4$$
$$3x + 2z = 6$$
$$x - y = 3z$$

But

is not.

END OF VIDEO

Question. Which of these are linear equations?

A. $x + y = 1 - t$

B. $xy + yz + xz = 1$

C. $(x - 1)(x + 1) = 0$

D. $2^x + 2^y = 16$

E. $\cos(15)y + \dfrac{x}{4} = -1$

F. $\sqrt{5}x_2 = \pi x_1 + 1$

Question. Which of the following are solutions to the system of linear equations?

$$\begin{aligned} x + y + z &= 1 \\ 2x + y &= 2 \\ y + 2z &= 0 \end{aligned}$$

A. $(x, y, z) = (0, 0, 0)$

B. $(x, y, z) = (0, 2, -1)$

C. $(x, y, z) = (3, -4, 2)$

D. $(x, y, z) = (7, 8, 14)$

E. None

Definition. A *matrix* is ...

Example. For example, this is a matrix:

Example. Encode the following system of linear equations as an "augmented" matrix:

$$\begin{aligned} 3x - 5y - 8z &= 4 \\ 5x - 2y &= 3z \\ x &= 2y + 7 \end{aligned}$$

Example. Rewrite the following matrix as a system of linear equations:

$$\begin{bmatrix} 1 & -2 & 3 & 3 \\ 0 & 1 & 4 & 2 \\ -1 & 3 & 1 & 4 \end{bmatrix}$$

Extra Example. Solve the system of linear equations

$$\begin{aligned} x_1 - 3x_2 + 2x_3 &= -1 \\ 2x_1 - 5x_2 - x_3 &= 2 \\ -4x_1 + 13x_2 - 12x_3 &= 11 \end{aligned}$$

Extra Example. Solve the system of equations

$$\begin{aligned} x - y + z &= 1 \\ 2x + 6y - z &= -4 \\ 4x - 5y + 2z &= 0 \end{aligned}$$

Solving Systems Linear Equations

After completing this section, students should be able to:

- Convert a system of linear equations to an augmented matrix and vice versa.
- Solve a system of linear equations using elimination and identify the operations on the augmented matrix that correspond to each step.

Example. Solve the system of linear equations using substitution:

$$-2a + 3b + 4c = 1$$
$$a + b + 5c = 2$$
$$b = 2a + c$$

Example. Solve the system of linear equations using elimination:

$$-2a + 3b + 4c = 1$$
$$a + b + 5c = 2$$
$$b = 2a + c$$

Example. Solve the system of linear equations using elimination and identify the operations on the augmented matrix that correspond to each step :

$$-2a + 3b + 4c = 1$$
$$a + b + 5c = 2$$
$$b = 2a + c$$

END OF VIDEOS

Example. Write this system of linear equations as an augmented matrix.

$$\begin{aligned} -2x_1 + 5x_2 - 10x_3 &= 4 \\ x_1 - 2x_2 + 3x_3 &= -1 \\ 7x_1 - 17x_2 + 34x_3 &= -16 \end{aligned}$$

Example. Solve this system of linear equations by simultaneously manipulating the equations and the matrix.

$$\begin{aligned} -2x_1 + 5x_2 - 10x_3 &= 4 \\ x_1 - 2x_2 + 3x_3 &= -1 \\ 7x_1 - 17x_2 + 34x_3 &= -16 \end{aligned}$$

Example. Perform the given row operation on the matrix $A = \begin{bmatrix} 3 & 2 & -1 \\ 5 & 0 & -1 \\ 4 & 6 & 0 \end{bmatrix}$

A. $R_1 \leftrightarrow R_3$

B. $\frac{1}{2}R_3 \to R_3$

C. $2R_1 - R_2 \to R_2$

Example. The matrix on the right results after performing a single row operation on the matrix on the left. Identify the row operation.

A. $\begin{bmatrix} -2 & 1 & 0 \\ 13 & -3 & 6 \\ -11 & 7 & -5 \end{bmatrix} \to \begin{bmatrix} 4 & -2 & 0 \\ 13 & -3 & 6 \\ -11 & 7 & -5 \end{bmatrix}$

B. $\begin{bmatrix} 4 & 2 & -1 & 2 \\ -1 & 0 & 5 & 7 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 14 & 23 \\ -1 & 0 & 5 & 7 \end{bmatrix}$

Example. Rewrite the system of equations as an augmented matrix.

$$\begin{aligned} 2x_1 - 2x_2 + x_3 &= 3 \\ -x_1 + x_2 - x_3 &= -3 \\ x_1 - 2x_2 + x_3 &= 2 \end{aligned}$$

Example. Solve this system of linear equations by simultaneously manipulating the equations and the matrix.

$$\begin{aligned} 2x_1 - 2x_2 + x_3 &= 3 \\ -x_1 + x_2 - x_3 &= -3 \\ x_1 - 2x_2 + x_3 &= 2 \end{aligned}$$

Reduced Row Echelon Form

After completing this section, students should be able to:

- Recognize when a matrix is in reduced row echelon form.
- Use Gaussian elimination to put a matrix in reduced row echelon form.
- Explain the Gaussian elimination algorithm.

Consider the matrix

$$\begin{bmatrix} 1 & 0 & 0 & -\frac{8}{3} \\ 0 & 1 & 0 & -\frac{11}{3} \\ 0 & 0 & 1 & \frac{5}{3} \end{bmatrix}$$

If this matrix represents a system of linear equations in the variables x, y, z, what are the solutions to the system?

Consider the matrix

$$\begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

If this matrix represents a system of linear equations in the variables x, y, z, what are the solutions to the system?

What properties of these matrices makes it easy to find the solutions?

Definition. A matrix is said to be in *reduced row-echelon form (RREF)* if:

1.

2.

3.

4.

Example. Which of these matrices are in RREF?

(a) $\begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 5 & 2 \\ 0 & 0 & 1 & 0 & 3 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$ (f) $\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 3 \end{bmatrix}$

END OF VIDEOS

Question. When we solve a system of equations using the method of elimination, what are the types of manipulations that we use on the equations?

-
-
-
-

Question. What are the corresponding operations that we use on rows of the augmented matrix?

-
-
-
-

Definition. The following operations on a matrix are called ...

1. Switch two rows.
2. Mulitply a row by a non-zero constant.
3. Multiply a row by a constant and add it to another row.

Example. Sometimes the easiest way to solve a system of equations is to use more complex operations:

$$\begin{aligned} -7x + 5y &= 2 \\ 3x - 2y &= 4 \end{aligned}$$

But it can also be solved using only elementary row operations. How could we do this, in this example?

Example. Each row operation below is not an elementary row operation. Can it be written as a sequence of two elementary row operations, one performed after the other? If so, give the two row operations. If not, explain why not.

A. $2R_1 + 5R_3 \to R_1$

B. $6R_2 - R_4 \to R_3$

Example. Which of these matrices are in reduced row echelon form (RREF)?

A. $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

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

C. $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

D. $\begin{bmatrix} 1 & 4 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

E. $\begin{bmatrix} 1 & 0 & 0 & 7 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 \end{bmatrix}$

A matrix is in RREF if:

Question. How is putting an augmented matrix in RREF, using row operations, useful for solving equations?

Example. Use elementary row operations to put the matrix in reduced row echelon form. Check your answer with technology.

$$\begin{bmatrix} 0 & 0 & 2 & -4 & -6 \\ 1 & -7 & 0 & 6 & 5 \\ -1 & 7 & -4 & 3 & 5 \end{bmatrix}$$

Example. Convert the system to an augmented matrix and then find all solutions by reducing the system to RREF. Check your answer with technology.

$$\begin{aligned} 2x_1 + x_2 &= 1 \\ -4x_1 - x_2 &= 3 \end{aligned}$$

Example. Covert the system to an augmented matrix and then find all solutions by reducing the system to RREF. Check your answer with technology.

$$\begin{aligned} 2x_1 + 6x_2 - 9x_3 &= 1 \\ -3x_1 - 11x_2 + 9x_3 &= 2 \\ x_1 + 4x_2 - 2x_3 &= 3 \end{aligned}$$

Extra Example. Put the augmented matrix for the system of linear equations in reduced row echelon form. Check your answer with technology.

$$\begin{aligned} -7x_1 - 6x_2 - 12x_3 &= -33 \\ 5x_1 + 5x_2 + 7x_3 &= 24 \\ x_1 + 4x_3 &= 5 \end{aligned}$$

Existence and Uniqueness of Solutions

After completing this section, students should be able to:

- Give examples of systems of linear equations that have (a) one unique solution, (b) no solutions, and (c) infinitely many solutions
- After putting an augmented matrix in reduced row echelon form, identify whether the associated system of linear equations has (a) one unique solution, (b) no solutions, and (c) infinitely many solutions
- Write down the solutions to a system of linear equations based on its augmented matrix in RREF.
- For a system of two variables, explain the relationship between the positions of the lines represented by the equations and whether the system has (a) one unique solution, (b) no solutions, and (c) infinitely many solutions.

Example. How many solutions does each of these three systems of linear equations have?

$$(a)\quad \begin{aligned} x + y &= 1 \\ x - y &= 3 \end{aligned} \qquad (b)\quad \begin{aligned} x + y &= 1 \\ x + y &= 4 \end{aligned} \qquad (c)\quad \begin{aligned} x + y &= 1 \\ 2x + 2y &= 2 \end{aligned}$$

Note. A system of linear equations can have:

-
-
-

Example. Find the solutions to this system of linear equations.

$$\begin{bmatrix} 1 & 3 & 5 & 7 \\ 3 & 5 & 7 & 9 \\ 5 & 7 & 9 & 1 \end{bmatrix}$$

Example. Find the solutions to this system of linear equations.

$$\begin{bmatrix} 0 & 0 & 2 & -4 & -6 \\ 1 & -7 & 0 & 6 & 5 \\ -1 & 7 & -4 & 3 & 5 \end{bmatrix}$$

Example. Find the solutions to this system of linear equations.

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

END OF VIDEO

In each of the following examples, a system of linear equations has an augmented matrix that has been converted to RREF. How many solutions does the system have?

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

B. $\begin{bmatrix} 1 & 0 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

C. $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

Summary: To determine the number of solutions for a system of equations from the augmented matrix in RREF,

1.

2.

3.

Example. For each example,

(i) write the system of equations to an augmented matrix

(ii) use technology to convert the matrix to RREF

(iii) write out the solutions to the system of equations, or write "NONE" if the system is inconsistent

(iv) if the system has infinitely many solutions, list all solutions and also give 2 particular solutions

A)

$$\begin{aligned} x_1 + x_2 + 6x_3 + 9x_4 &= 0 \\ -x_1 - x_3 - 2x_4 &= -3 \end{aligned}$$

B)

$$\begin{aligned} 2x_1 + x_2 + 2x_3 &= 0 \\ x_1 + x_2 + 3x_3 &= 1 \\ 3x_1 + 2x_2 + 5x_3 &= 3 \end{aligned}$$

C)

$$\begin{aligned} x_1 + 3x_2 + 3x_3 &= 1 \\ 2x_1 - x_2 + 2x_3 &= -1 \\ 4x_1 + 5x_2 + 6x_3 &= 2 \\ x_1 + 3x_2 + x_3 &= 2 \end{aligned}$$

Example. For which values of k, if any, will the following system have (a) exactly 1 solution, (b) infinitely many solutions, (c) no solution ?

$$\begin{aligned} x_1 + 3x_2 &= 4 \\ x_1 + kx_2 &= 2 \end{aligned}$$

Extra Example. For which values of k, if any, will the following system have (a) exactly 1 solution, (b) infinitely many solutions, (c) no solution ?

$$\begin{aligned} x_1 + 3x_2 &= 4 \\ 2x_1 + kx_2 &= 8 \end{aligned}$$

Extra Example. For which values of k, if any, will the system represented by the following augmented matrix have (a) exactly 1 solution, (b) infinitely many solutions, (c) no solution ?

$$\begin{bmatrix} 1 & 3 & 4 & 1 \\ 2 & 5 & 1 & -1 \\ 3 & -2 & k & 7 \end{bmatrix}$$

Setting Up Linear Equations

After completing this section, students should be able to:

- Set up a system of linear equations to model a real world problem
- Solve the system of linear equations using substitution, elimination, or other methods
- Critically evaluate the solutions of a system of linear equations to see if they are reasonable in the context of the problem

Example. Pete is going on a backpacking trip, and for simplicity, he decides to only bring three kinds of food: peanuts, corn nuts, and freeze dried mango. He wants to use exactly 12 liters of space and exactly 5 kg of weight for these food items. Since freeze dried mango is even more delicious than corn nuts and peanuts, he would like to take as much freeze dried mango (in liters) as peanuts and corn nuts combined. Is this possible?

Example. Find a polynomial of degree 3 that goes through the points $(0, -3)$, $(1, 2)$, $(3, 5)$, and $(4, 0)$.

Example. When propane burns in oxygen, it produces carbon dioxide and water:

$C_3H_8 + O_2 \rightarrow CO_2 + H_2O$

Balance the chemical equation by figuring out how many molecules of each compound you need on each side. Use the smallest possible numbers that work.

Extra Example. Two solutions of salt water contain 0.03% and 0.18% salt respectively. A lab technician wants to make 1 liter of solution which contains 0.09% salt. How much of each solution should she use?

Extra Example. A man flies a small airplane from Fargo to Bismarck, North Dakota — a distance of 180 miles. Because he is flying into a head wind, the trip takes him 2 hours. On the way back, the wind is still blowing at the same speed, so the return trip takes only 1 hour 12 minutes. What is his speed in still air, and how fast is the wind blowing?

Extra Example. One serving of Lucky Charms contains 10% of the percent daily values (PDV) for calcium, 25% of the PDV for iron, and 25% of the PDV for zinc.

One serving of Raisin Bran contains 2% of the PDV for calcium, 25% of the PDV for iron, and 10% of the PDV for zinc.

Determine the number of servings of each cereal required to get 40% of the PDV for calcium, 200% of the PEV for iron, and 125% of the PDV for zinc.

Matrix Addition and Scalar Multiplication

After completing this section, students should be able to:

- State the dimensions of a matrix
- Add two matrices of the same dimensions
- Multiply a matrix by a scalar
- Identify what properties matrix addition and scalar multiplication satisfy (e.g. associativity, commutativity)

Definition. A *matrix* is ...

$$\begin{bmatrix} 1 & 0 & -4 \\ 5 & 9 & 1 \end{bmatrix}$$

The *diimensions* of a matrix are ...

The *elements* of a matrix are ...

To add two matrices of the same dimensions ...

$$\begin{bmatrix} 1 & 0 & -4 \\ 5 & 9 & 1 \end{bmatrix} + \begin{bmatrix} 3 & -5 & 7 \\ 1 & 2 & 1 \end{bmatrix} =$$

To subtract two matrices ...

$$\begin{bmatrix} 1 & 0 & -4 \\ 5 & 9 & 1 \end{bmatrix} - \begin{bmatrix} 3 & -5 & 7 \\ 1 & 2 & 1 \end{bmatrix} =$$

To multiply a matrix by a *scalar* ...

$$3 \cdot \begin{bmatrix} 1 & 0 & -4 \\ 5 & 9 & 1 \end{bmatrix} =$$

Properties of addition and multiplication of numbers:

1. Addition is associative	
2. Addition is commutative	
3. Additive identity (the number zero)	
4. Multiplication distributes over addition	
5. Multiplication by 0	
6. Multiplication is associative	
7. Multiplication is commutative	

Properties of matrix addition and scalar multiplication:

1. Addition is associative	
2. Addition is communtative	
3. Additive identity (the zero matrix)	
4. Scalar multiplication is distributive	
5. Multiplication by 0	
6. Scalar multiplication is associative	
7. Scalar multiplication is commutative	

Example. Find $5A - B$, where

$$A = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 4 & -2 \\ 3 & 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 8 & 10 & 9 \\ 4 & -5 & 3 \\ 6 & 0 & 2 \end{bmatrix}$$

END OF VIDEOS

Example. For matrices $A = \begin{bmatrix} 4 & -5 \\ 1 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ -4 & 6 \end{bmatrix}$, simplify each expression

1. $A + B$

2. $3A - 4B$

3. $2(A - B) - (A - 3B)$

Example. For matrices $C = \begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}$ and $D = \begin{bmatrix} 1 & 7 \\ 3 & -4 \end{bmatrix}$, find the matrix X that satisfies the equation $3C + 2X = -D$

Example. Find the scalars a and b that satisfy the equation

$$a\begin{bmatrix}1\\2\\3\end{bmatrix}+b\begin{bmatrix}1\\1\\2\end{bmatrix}=\begin{bmatrix}0\\-1\\-1\end{bmatrix}$$

Matrix Multiplication

After completing this section, students should be able to:

- Decide, based on the dimensions of two matrices A and B, if it is possible to take the product AB
- Compute the product of two matrices of appropriate dimensions
- Identify what properties matrix multiplication satisfies (e.g. associativity, distributive property)
- Give an example to show that multiplication of matrices is not always communtative
- Write down a matrix that is a multiplicative identify

Definition. A matrix that has only one column is called a ...

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

Definition. A matrix has only one row is called a ....

$$W = \begin{bmatrix} 3 & 6 & 1 & -1 \end{bmatrix}$$

Definition. We multiply a $1 \times n$ row vector with an $n \times 1$ column vector as follows:

Definition. Suppose A is an $m \times r$ matrix and B is a dimension $r \times n$ matrix. The product $A \cdot B$ (also written AB) is the $m \times n$ matrix whose entry in row i and column j is ...

Example. Compute AB, where $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \\ 0 & 7 \\ -1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$

Example. For $Y = \begin{bmatrix} 1 & 5 \\ 7 & 2 \end{bmatrix}$ and $Z = \begin{bmatrix} -1 & 3 \\ 9 & 5 \end{bmatrix}$

find (a) YZ and (b) ZY

Properties of Matrix Multiplication

Properties of multiplication:	Numbers vs.	Matrices
1. Multiplication is associative		
2. Multiplication is commutative		
3. Multiplication distributes over addition		
4. Multiplicative identity		

END OF VIDEOS

Example. Multiply $\vec{u} \cdot \vec{v}$, where $\vec{u} = \begin{bmatrix} 5 & 6 & 0 & -2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 3 \\ -1 \\ 4 \\ 2 \end{bmatrix}$

Example. Multiply $\begin{bmatrix} -5 & 6 & -4 \\ 1 & 2 & -3 \end{bmatrix} \cdot \begin{bmatrix} 2 & -1 & 4 \\ 9 & -5 & -2 \\ 3 & -1 & 0 \end{bmatrix}$

Example. For matrices $A = \begin{bmatrix} 5 & 4 & -1 \\ 3 & -2 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 0 \\ 2 & 1 \\ 9 & 8 \\ 2 & -3 \end{bmatrix}$

(a) Is it possible to multiply AB ?

(b) Is it possible to multiply BA?

Example. Multiply AB and BA

$$A = \begin{bmatrix} -4 & 3 & 3 \\ -5 & -1 & -5 \\ -5 & 0 & -1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 5 & 0 \\ -5 & -4 & 3 \\ 5 & -4 & 3 \end{bmatrix}$$

Example. Multiply AB and BA

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad B = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Example. Multiply AB and BA

$$A = \begin{bmatrix} -3 & 6 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix}$$

For $n \times n$ matrices A, B, and C are these statements true or false? Note that True means always true and False means sometimes or always false.

1. True or False: $AB = BA$

2. True or False: $A(B + C) = AB + AC$

3. True or False: $A(BC) = (AB)C$

4. True of False: If $C \cdot D = 0$, then either $C = 0$ or $D = 0$.

5. True or False: If C is not the zero matrix, and $CA = CB$, then $A = B$.

Interpretations of Matrix Multiplication

$$\begin{bmatrix} 2 & 3 & 1 \\ -4 & 2 & 5 \end{bmatrix} \cdot \begin{bmatrix} 1 & -2 \\ 2 & 5 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 12 \\ 15 & 23 \end{bmatrix}$$

Interpretation 1: the $i\ j$ entry of $A \cdot B$ tells us ...

Interpretation 2: the jth column of $A \cdot B$ tells us ...

Example. Given that $\begin{bmatrix} 4 & 5 \\ 6 & 7 \\ 8 & 9 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 13 \\ 19 \\ 25 \end{bmatrix}$, and $\begin{bmatrix} 4 & 5 \\ 6 & 7 \\ 8 & 9 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 9 \\ 13 \\ 17 \end{bmatrix}$, and $\begin{bmatrix} 4 & 5 \\ 6 & 7 \\ 8 & 9 \end{bmatrix}\begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 15 \\ 21 \\ 27 \end{bmatrix}$,

find $\begin{bmatrix} 4 & 5 \\ 6 & 7 \\ 8 & 9 \end{bmatrix}\begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 3 \end{bmatrix}$.

Definition. A *linear combination* of vectors is ...

Example. $3\begin{bmatrix}1\\2\\3\end{bmatrix} + 4\begin{bmatrix}-2\\5\\0\end{bmatrix} - 2\begin{bmatrix}4\\1\\-1\end{bmatrix}$

Interpretation 3: the jth column of $A \cdot B$ tells us ...

$$\begin{bmatrix}2 & 3 & 1\\-4 & 2 & 5\end{bmatrix} \cdot \begin{bmatrix}1 & -2\\2 & 5\\3 & 1\end{bmatrix} = \begin{bmatrix}11 & 12\\15 & 23\end{bmatrix}$$

Example. For $A = \begin{bmatrix} 1 & 2 \\ 6 & -5 \\ 3 & 4 \end{bmatrix}$, write down a matrix B so that

- the first column of $A \cdot B$ wlll be twice the first column of A plus 3 times the second column of A,
- the second column of $A \cdot B$ wlll be -3 times the first column of A plus 4 times the second column of A,
- the third column of $A \cdot B$ will be the first column of A plus the second column of A,
- the forth column of $A \cdot B$ will be 5 times the second column of A.

Introduction to Vectors

After completing this section, students should be able to:

- Represent vectors as columns of numbers and as arrows
- Add and subtract vectors and multiply vectors by a scalar, both in the column representation and in the arrow representation
- Find the length of a vector

Definition. A *vector* can be defined as ...

Definition. The *components* of the vector are ...

Definition. The *dimension* of the vector is ...

Definition. We add vectors by ...

$$\begin{bmatrix} 3 \\ 7 \end{bmatrix} + \begin{bmatrix} 1 \\ -5 \end{bmatrix}$$

and we subtract them by ...

$$\begin{bmatrix} 3 \\ 7 \end{bmatrix} - \begin{bmatrix} 1 \\ -5 \end{bmatrix}$$

Definition. A *scalar* is another word for ...

Definition. We mutliply a scalar by a vector by ...

$$5 \cdot \begin{bmatrix} 3 \\ 7 \end{bmatrix}$$

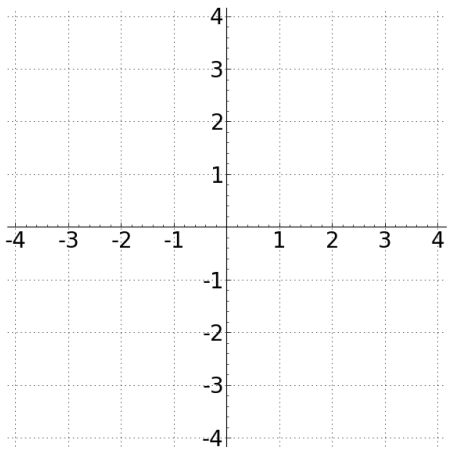
Definition. The negative of a vector is formed by ...

$$-\begin{bmatrix} 1 \\ -5 \end{bmatrix}$$

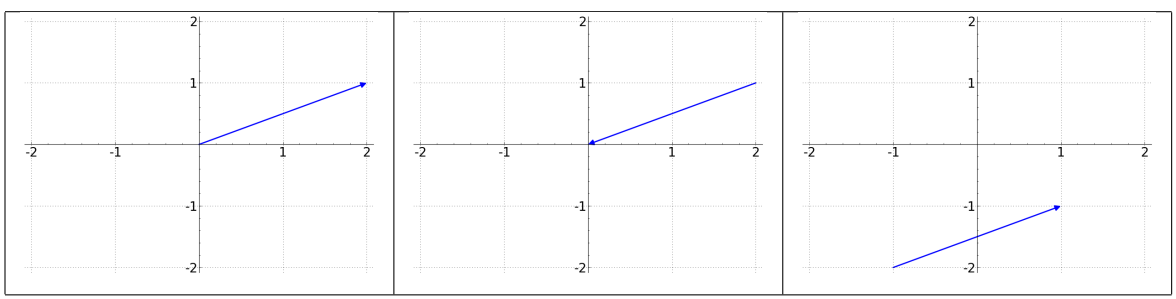
Example. For $\vec{v} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$, what is $-\vec{v} + \vec{v}$?

We can visualize a vector with two components $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ by ...

Example. Draw $\vec{v} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$

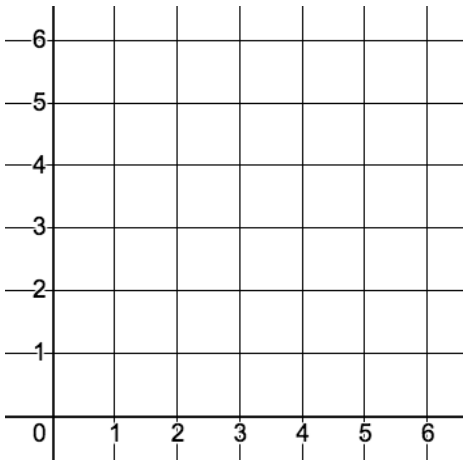


Example. Which vectors are equivalent?



Vector addition

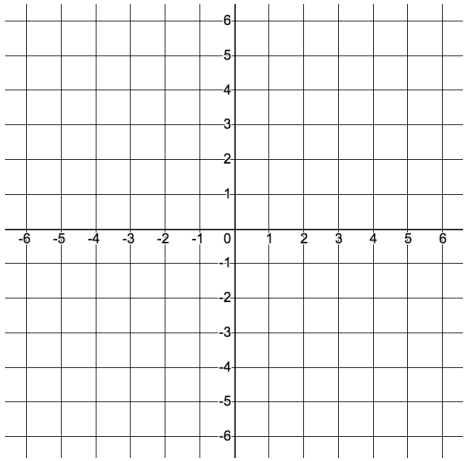
Example. For $\vec{a} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, represent $\vec{a}$, $\vec{b}$, and $\vec{a} + \vec{b}$ with arrows.



In general, to draw the sum of two vectors $\vec{v} + \vec{w}$

Scalar multiplication

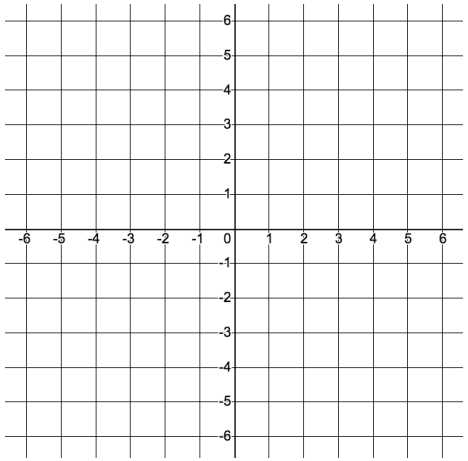
Example. For $\vec{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, represent $\vec{b}$, $2\vec{b}$, and $-\vec{b}$ with arrows.



In general, to draw a scalar multiple times a vector ...

Vector subtraction

Example. For $\vec{a} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, represent $\vec{a} - \vec{b}$ with arrows.



In general, to draw the difference of two vectors $\vec{v} - \vec{w}$

Length of vectors

Example. What is the length of the vector $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$?

In general, the length of a vector $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ is ...

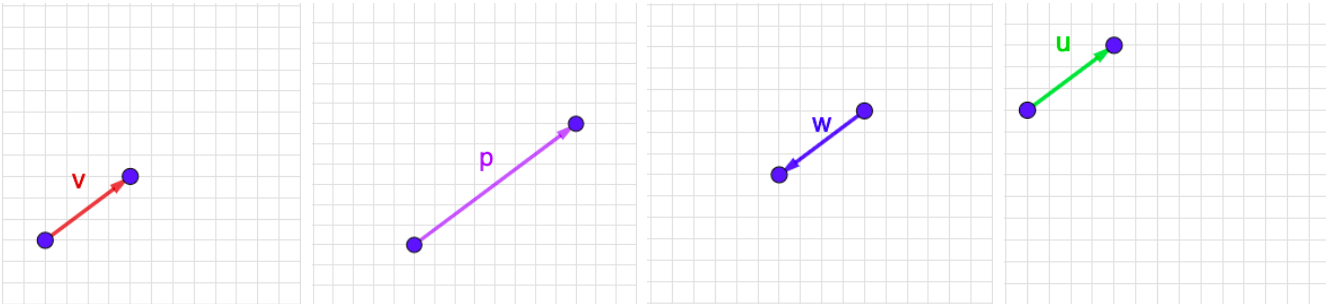
END OF VIDEOS

Question. What are two different ways of representing vector?

- .

- .

Example. Group the following vectors into categories based on which ones represent the same vector.



$$\vec{r} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \qquad \vec{z} = \begin{bmatrix} -4 \\ -3 \end{bmatrix} \qquad \vec{y} = \begin{bmatrix} 8 \\ 6 \end{bmatrix} \qquad \vec{x} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

The vector $\vec{AB}$ that extends from the point $A = (2, 5)$ to the point $B = (-2, 2)$.

Example. For the vectors $\vec{a}$, $\vec{b}$, and $\vec{c}$ shown below, compute the following vectors.

(a) $\vec{a} + \vec{b}$

(b) $\vec{a} - \vec{b}$

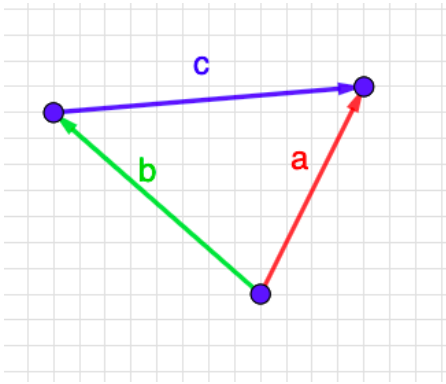
(c) $2\vec{c} - \vec{a} + \vec{b}$

(d) How does $||\vec{a} + \vec{b}||$ compare to $||\vec{a}|| + ||\vec{b}||$?

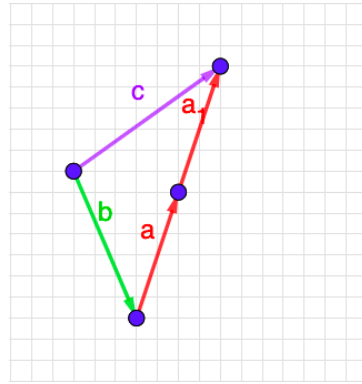
Example. Find a vector in the opposite direction of $\vec{v} =< 5, -4 >$ that is 3 times a long.

Example. Find a vector in the same direction as $\vec{u} =< 5, 12 >$ that has length 10.

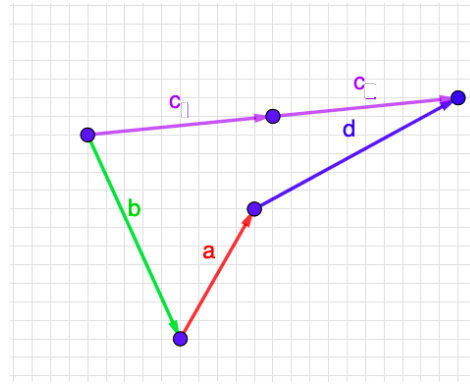
Example. Write the vector $\vec{a}$ in terms of the other vectors.



(a)



(b)



(c)

In 3-d

All the examples up to now have been for vectors with only two entries (2-dimensional vectors), but is it also possible to work with vectors in 3-d or higher dimensions in a similar way.

Example. Consider the vectors $\vec{a} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$.

(a) Sketch $\vec{a}$ and $\vec{b}$.

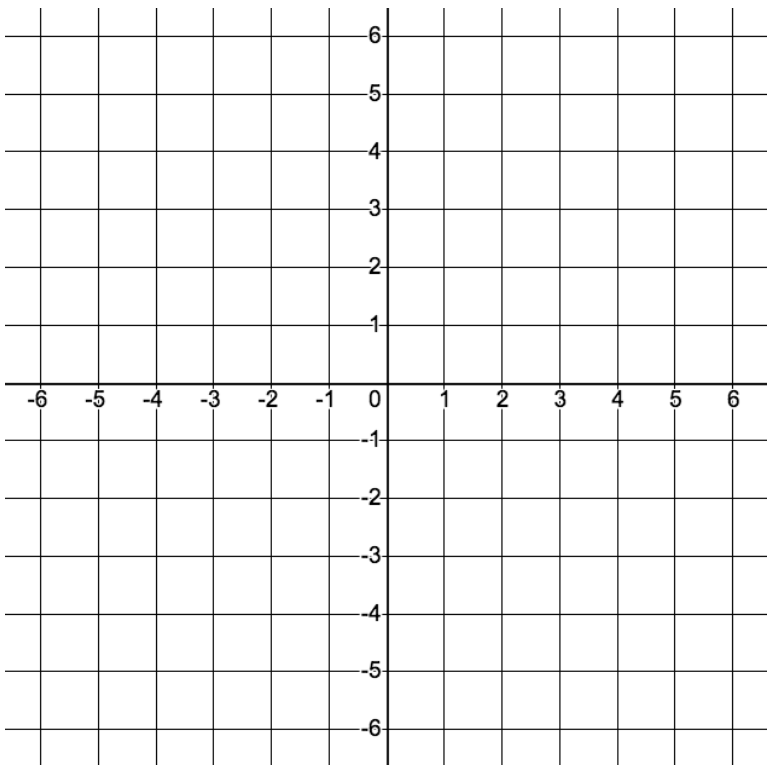
(b) Compute $\vec{a} + \vec{b}$ and draw it on the sketch.

(c) What is $||\vec{a} + \vec{b}||$?

Multiplying Matrices by Vectors

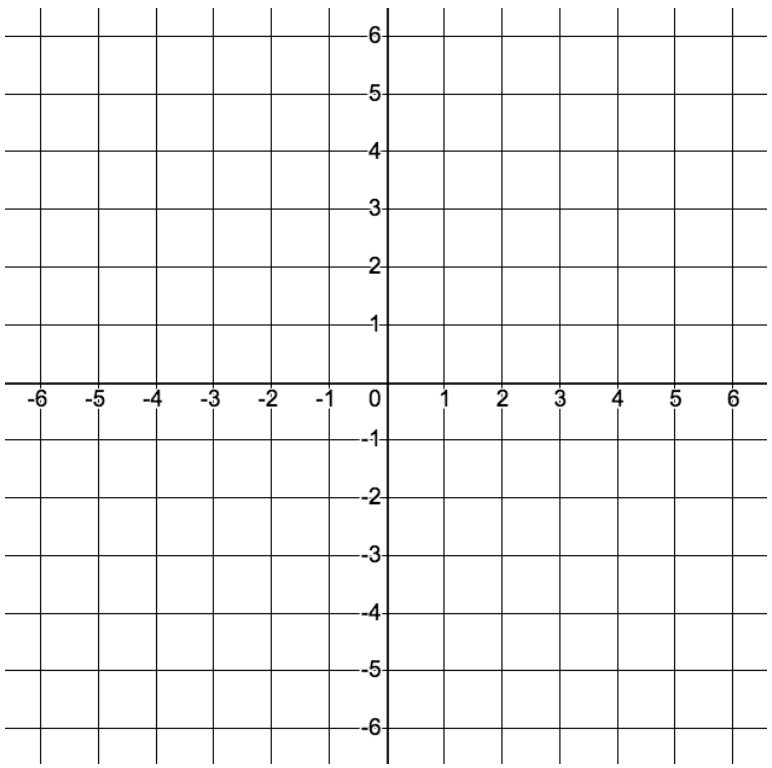
Extra Example. Consider the matrix $A = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ and the vectors $\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\vec{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.

Graph $\vec{x}$, $\vec{y}$, $\vec{v}$, $\vec{w}$, and $A\vec{x}$, $A\vec{y}$, $A\vec{v}$, and $A\vec{w}$ all on the same coordinate axes, using the same color for $\vec{x}$ and $A\vec{x}$, etc.



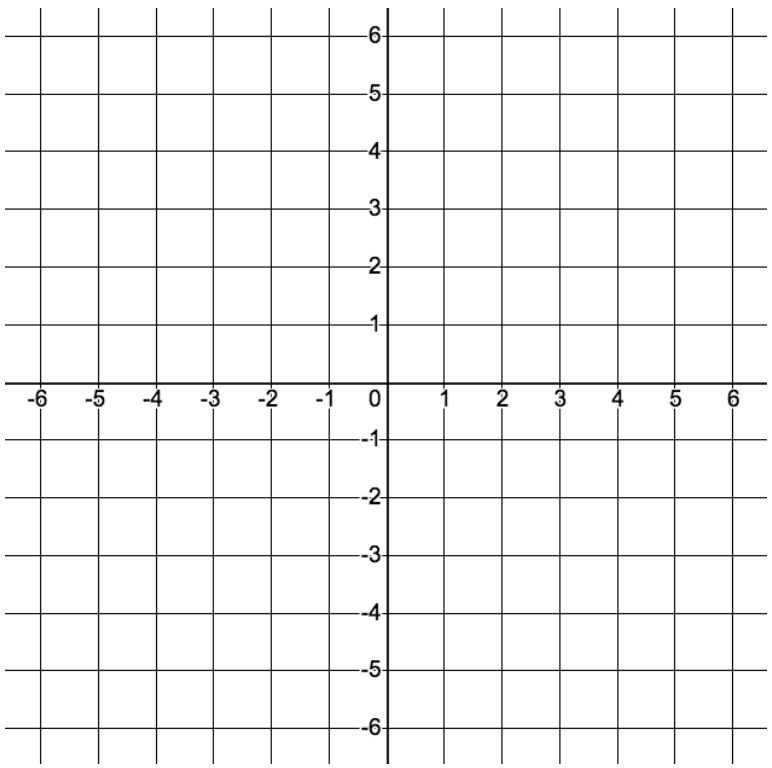
Extra Example. Consider the matrix $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ and the vectors $\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\vec{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.

Graph $\vec{x}$, $\vec{y}$, $\vec{v}$, $\vec{w}$, and $A\vec{x}$, $A\vec{y}$, $A\vec{v}$, and $A\vec{w}$ all on the same coordinate axes, using the same color for $\vec{x}$ and $A\vec{x}$, etc.



Extra Example. Consider the matrix $A = \begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix}$ and the vectors $\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\vec{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.

Graph $\vec{x}$, $\vec{y}$, $\vec{v}$, $\vec{w}$, and $A\vec{x}$, $A\vec{y}$, $A\vec{v}$, and $A\vec{w}$ all on the same coordinate axes, using the same color for $\vec{x}$ and $A\vec{x}$, etc.



Solving Vector Equations

After completing the section, students should be able to:

- solve equations like $A\vec{x} = \vec{b}$ where A is a matrix and $\vec{x}$ and $\vec{b}$ are vectors
- when there are infinitely many solutions to the equation $A\vec{x} = \vec{b}$, use them to find solutions to the equation $A\vec{x} = \vec{0}$
- when there is one unique solution to the equation $A\vec{x} = \vec{b}$, determine the solutions to the equation $A\vec{x} = \vec{0}$ with no other information

Example. Solve the equation $A\vec{x} = \vec{b}$ for $\vec{x}$, where
$A = \begin{bmatrix} 1 & 2 & -2 \\ 3 & 4 & -5 \\ 0 & 1 & 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 1 \\ 7 \\ -2 \end{bmatrix}$.

Example. Solve the equation $A\vec{x} = \vec{b}$ for $\vec{x}$.

$$A = \begin{bmatrix} 3 & 4 & 1 \\ 1 & 1 & 0 \\ 2 & 3 & 1 \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} -2 \\ -5 \\ 3 \end{bmatrix}.$$

Recap:

1. $A = \begin{bmatrix} 1 & 2 & -2 \\ 3 & 4 & -5 \\ 0 & 1 & 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 1 \\ 7 \\ -2 \end{bmatrix}$

   $A\vec{x} = \vec{b}$ $\qquad$ $A\vec{x} = \vec{0}$

2. $A = \begin{bmatrix} 3 & 4 & 1 \\ 1 & 1 & 0 \\ 2 & 3 & 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} -2 \\ -5 \\ 3 \end{bmatrix}$.

   $A\vec{x} = \vec{b}$ $\qquad$ $A\vec{x} = \vec{0}$

END OF VIDEO

Example. Solve the equation $A\vec{x} = \vec{b}$ and the equation $A\vec{x} = \vec{0}$ where $A = \begin{bmatrix} 0 & 2 \\ -1 & 3 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

Example. Solve the equation $A\vec{x} = \vec{b}$ and the equation $A\vec{x} = \vec{0}$ where $A = \begin{bmatrix} -4 & 8 & 3 & 2 \\ -4 & 8 & 5 & 2 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$.

Example. Solve the equation $A\vec{x} = \vec{b}$ and the equation $A\vec{x} = \vec{0}$ where

$$A = \begin{bmatrix} 1 & 5 & -2 \\ 1 & 4 & 5 \\ 1 & 3 & 12 \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Example. Solve the equation $A\vec{x} = \vec{b}$ and the equation $A\vec{x} = \vec{0}$ where

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 5 \end{bmatrix}$$

Question. If the equation $A\vec{x} = \vec{b}$ has infinitely many solutions, then the equation $A\vec{x} = \vec{0}$ has / could have: (select all that apply)

A. no solutions

B. one unique solution

C. infinitely many solutions

Question. If the equation $A\vec{x} = \vec{b}$ has one unique solution, then the equation $A\vec{x} = \vec{0}$ has / could have: (select all that apply)

A. no solutions

B. one unique solution

C. infinitely many solutions

Question. If the equation $A\vec{x} = \vec{b}$ has no solutions, then the equation $A\vec{x} = \vec{0}$ has / could have: (select all that apply)

A. no solutions

B. one unique solution

C. infinitely many solutions

Question. If the equation $A\vec{x} = \vec{0}$ has infinitely many solutions, then the equation $A\vec{x} = \vec{b}$ has / could have: (select all that apply)

A. no solutions

B. one unique solution

C. infinitely many solutions

Question. If the equation $A\vec{x} = \vec{0}$ has one unique solution, then the equation $A\vec{x} = \vec{b}$ has / could have: (select all that apply)

A. no solutions

B. one unique solution

C. infinitely many solutions

Question. If the equation $A\vec{x} = \vec{0}$ has no solutions, then the equation $A\vec{x} = \vec{b}$ has / could have: (select all that apply)

A. no solutions

B. one unique solution

C. infinitely many solutions

D. this is impossible

Example. Suppose that the complete set of solutions to the equation $A\vec{x} = \vec{b}$ is $x_3 \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} + \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix}$. Find all solutions to $A\vec{x} = \vec{0}$.

Example. Suppose that the complete set of solutions to the equation $A\vec{x} = \vec{0}$ is $x_2 \begin{bmatrix} -7 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}$. Suppose also that $\vec{w} = \begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix}$ is one solution to the equation $A\vec{x} = \vec{b}$. Find all solutions to $A\vec{x} = \vec{b}$.

Extra Example. Suppose that $\vec{v} = \begin{bmatrix} 6 \\ 1 \\ -1 \end{bmatrix}$ is a solution to the equation $A\vec{x} = \vec{0}$ and $\vec{w} = \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix}$ is a solution to the equation $A\vec{x} = \vec{b}$.

1. Find as many solutions to the equation $A\vec{x} = \vec{0}$ as you can.

2. Find as many solutions to the equation $A\vec{x} = \vec{b}$ as you can.

Extra Example. Suppose that $\vec{v} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ are both solutions to the equation $A\vec{x} = \vec{b}$.

1. Find one solution to the equation $A\vec{x} = \vec{0}$.

2. Find as many solutions to the equation $A\vec{x} = \vec{0}$ as you can.

3. Find as many solutions to the equation $A\vec{x} = \vec{b}$ as you can.

Solving Matrix Equations

After completing this section, students should be able to

- Solve matrix equations of the form $AX = B$ for X, where A and B are matrices, or explain why there are no solutions.

Example. Solve the matrix equation $AX = B$ for matrix X, where $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 4 & 1 \\ -1 & 2 & 1 \end{bmatrix}$.

Example. Solve the matrix equation $AX = B$ for matrix X, where

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 2 & 5 & 0 \\ 4 & -10 & 3 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 3 & 0 \end{bmatrix}.$$

END OF VIDEO

Question. For matrix $A = \begin{bmatrix} -2 & 0 & 4 \\ -5 & -4 & 5 \\ -3 & 5 & -3 \end{bmatrix}$, how do we go about solving the equation

$A\vec{x} = \begin{bmatrix} -18 \\ -38 \\ 10 \end{bmatrix}$?

How do we go about solving

$A\vec{x} = \begin{bmatrix} 2 \\ 18 \\ 2 \end{bmatrix}$?

$A\vec{x} = \begin{bmatrix} -14 \\ -13 \\ -18 \end{bmatrix}$?

Question. For matrix $A = \begin{bmatrix} -2 & 0 & 4 \\ -5 & -4 & 5 \\ -3 & 5 & -3 \end{bmatrix}$, how do we go about solving the equation

$$AX = \begin{bmatrix} -18 & 2 & -14 \\ -38 & 18 & -13 \\ 10 & 2 & -18 \end{bmatrix}?$$

Example. Solve the matrix equation $AX = B$, where

$$A = \begin{bmatrix} -2 & 0 & 4 \\ -5 & -4 & 5 \\ -3 & 5 & -3 \end{bmatrix} \text{ and } B = \begin{bmatrix} -18 & 2 & -14 \\ -38 & 18 & -13 \\ 10 & 2 & -18 \end{bmatrix}$$

Example. The augmented matrix $\begin{bmatrix} 1 & 3 & 4 & 1 & 3 \\ 2 & 5 & 8 & -1 & 0 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 0 & 4 & -8 & -15 \\ 0 & 1 & 0 & 3 & 6 \end{bmatrix}$.

What is the solution to the following matrix equation ?

$$\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix} X = \begin{bmatrix} 4 & 1 & 3 \\ 8 & -1 & 0 \end{bmatrix}$$

Example. The augmented matrix $\begin{bmatrix} 1 & 3 & 4 & 1 & 3 \\ 2 & 6 & 8 & -1 & 0 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 3 & 4 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}$.

How many solutions are there to the following matrix equation ?

$$\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} X = \begin{bmatrix} 4 & 1 & 3 \\ 8 & -1 & 0 \end{bmatrix}$$

Example. The augmented matrix $\begin{bmatrix} 1 & 3 & 4 & 1 & 3 \\ 2 & 6 & 8 & 2 & 6 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 3 & 4 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

How many solutions are there to the following matrix equation ?

$$\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} X = \begin{bmatrix} 4 & 1 & 3 \\ 8 & 2 & 6 \end{bmatrix}$$

Example. Solve the matrix equation $AX = B$, where

$$A = \begin{bmatrix} -4 & 2 & -2 \\ 1 & 0 & 1 \\ 3 & -1 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 10 & 2 \\ -3 & 5 \\ -8 & -1 \end{bmatrix}$$

Example. Solve the matrix equation $AX = B$, where

$$A = \begin{bmatrix} 1 & -2 & -3 \\ 1 & -1 & -2 \\ 2 & -3 & -5 \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 & 2 & 3 \\ 1 & 0 & 5 \\ 2 & 0 & 4 \end{bmatrix}$$

Example. Solve the matrix equation $AX = B$, where

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

The Inverse of a Matrix

After completing this section, students should be able to:

- compute the inverse of a matrix by converting an augmented matrix to RREF
- compute the inverse of a 2x2 matrix using a shortcut formula
- use the inverse of a matrix to solve a matrix equation of the form $AX = B$ for X

Inverses of Matrices Video

Definition. Suppose that A is an $n \times n$ matrix and B is another $n \times n$ matrix such that $AB = I_n$ and $BA = I_n$. Then

1. A is called ...
2. B is called ... , and denoted A^{-1}.

Question. Is it possible to have $AB = I_n$ but $BA \neq I_n$?

Question. Is it possible to have two different matrices B and C that are both inverses for A?

Question. Does every square matrix have an inverse?

Example. Find the inverse matrix for $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$ if it exists.

Example. Solve the equation $AX = B$, where $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 3 & 3 \\ 0 & -1 \end{bmatrix}$

The Inverse of a 2×2 Matrix Video

Example. Find the inverse matrix for $A = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$, if it exists.

A shortcut formula for inverse of a $2x2$ matrix.

END OF VIDEOS

Review of definitions

Review. The identity matrix I is the matrix ...

Review. If B is a square matrix, and I is the identity matrix with the same dimensions, then $IB =$ and $BI =$.

Review. A matrix A is called invertible if ...

Computing Inverses

Example. Decide if the matrix is invertible. If it is, find the inverse.

$$\begin{bmatrix} 25 & -10 & -4 \\ -18 & 7 & 3 \\ -6 & 2 & 1 \end{bmatrix}$$

Example. Solve the equation $A\vec{x} = \vec{b}$, where $A = \begin{bmatrix} 25 & -10 & -4 \\ -18 & 7 & 3 \\ -6 & 2 & 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$.

(Note: A is the same matrix as in the previous problem.)

Example. Decide if the matrix is invertible. If it is, find the inverse.

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

Example. Decide if the matrix is invertible. If it is, find the inverse.

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

Example. Decide if the matrix is invertible. If it is, find the inverse.

$$\begin{bmatrix} 2 & 3 & 4 \\ -3 & 6 & 9 \\ -1 & 9 & 13 \end{bmatrix}$$

Example. Decide if the matrix is invertible. If it is, find the inverse.

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Example. Decide if the matrix is invertible. If it is, find the inverse.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -19 & -9 & 0 & 4 \\ 33 & 4 & 1 & -7 \\ 4 & 2 & 0 & -1 \end{bmatrix}$$

Solving equations with matrix inverses

Example. Use A^{-1} to solve the equation $A\vec{x} = \vec{b}$.

$A = \begin{bmatrix} 9 & 70 \\ -4 & -31 \end{bmatrix}, \vec{b} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

Example. Use A^{-1} to solve the equation $AX = B$.

$$A = \begin{bmatrix} 1 & -6 & 0 \\ 0 & 1 & 0 \\ 2 & -8 & 1 \end{bmatrix}, B = \begin{bmatrix} -69 & 0 & 4 & 1 \\ 10 & 4 & 5 & 0 \\ -102 & 9 & 9 & 5 \end{bmatrix}$$

Properties of the Matrix Inverse

After completing this section, students should be able to:

- For an $n \times n$ matrix, list several conditions that are equivalent to the matrix being invertible
- For an $n \times n$ matrix A, describe the relationship between A being invertible and the number of solutions to the equation $A\vec{x} = \vec{b}$
- Compute $(AB)^{-1}$ from A^{-1} and B^{-1}
- Simplify $(A^{-1})^{-1}$

Equivalent Conditions for Matrices to be Invertible Video

Theorem. Let A be an $n \times n$ matrix. The following statements are equivalent:

1. A is invertible.
2. There exists a matrix B such that $BA = I$.
3. There exists a matrix C such that $AC = I$.
4. The reduced row echelon form of A is I.
5. The equation $A\vec{x} = \vec{b}$ has exactly one solution for every $n \times 1$ vector $\vec{b}$.
6. The equation $A\vec{x} = \vec{0}$ has exactly one solution (namely $\vec{x} = \vec{0}$).

Note. If A is invertible, then $A\vec{x} = \vec{b}$ ...

Question. If A is NOT invertible, what can we say about the number of solutions to the equation $A\vec{x} = \vec{b}$?

Properties of Inverse Matrices Video

Suppose that A and B are invertible matrices. Which of the following statements are necessarily true?

1. True or False: $(A^{-1})^{-1} = A$

2. True or False: $(AB)^{-1} = A^{-1}B^{-1}$

3. True or False: $(A + B)^{-1} = A^{-1} + B^{-1}$

4. True or False: $(3A)^{-1} = \frac{1}{3}A^{-1}$.

Definition. A *diagonal matrix* is a square matrix with ...

Example. Which of these matrices are diagonal matrices?

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}. \qquad . \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 7 \end{bmatrix}$$

Note. If A is a diagonal matrix with diagonal entries $d_1, d_2, d_3, \cdots d_n$, and none of the diagonal entries are 0, then A^{-1} ...

Example. Find the inverse of $\begin{bmatrix} 5 & 0 \\ 0 & 6 \end{bmatrix}$.

END OF VIDEOS

Review

Review. For the equation $A\vec{x} = \vec{b}$ there are three options as far as number of solutions and three corresponding options for the RREF of the augmented matrix $\left[A \mid \vec{b}\right]$,

-
-
-

Review. For the square matrix A, the option that $A\vec{x} = \vec{b}$ has a unique solution corresponds to the RREF of A ...

Review. For the square matrix A, the option that $A\vec{x} = \vec{b}$ has either no solutions or infinitely many solutions corresponds to the RREF of A ...

True or False

Review. True or False: For a square matrix A if $A\vec{x} = \vec{0}$ has solution(s), then $A\vec{x} = \vec{b}$ has solution(s) for any vector $\vec{b}$.

Review. True or False: For a square matrix A if $A\vec{x} = \vec{0}$ has one unique solution, then $A\vec{x} = \vec{b}$ has one unique solution for any vector $\vec{b}$.

Review. Which of the following statements are true?

1. True or False: $(AB)^{-1} = A^{-1}B^{-1}$

2. True or False: $(A + B)^{-1} = A^{-1} + B^{-1}$

3. True or False: $(A^{-1})^{-1} = A$

4. True or False: $(5A)^{-1} = 5A^{-1}$

Example. In the following statements, A and B are $n \times n$ invertible matrices. For each statement, decide if the statement is true or false. Recall that true means always true, and false means sometimes or always false. Justify your answer.

1. True or False: $(ABA^{-1})^2 = AB^2A^{-1}$

2. True or False: A^3 is invertible

3. True or False: A^2B^4 is invertible

4. True or False: $ABA^{-1} = B$

5. True or False: $(A + A^{-1})^2 = A^2 + (A^{-1})^2$

6. True or False: $A + I_n$ is invertible

7. True or False: $A + A^{-1}$ is invertible

The Transpose of a Matrix

After completing this section, students should be able to:

- Compute the transpose of a matrix
- Find the transpose of the sum $A + B$ from the transpose of A and the transpose of B
- Find the transpose of the product AB from the transpose of A and the transpose of B
- Find the inverse of A^T from the inverse of A
- Find the transpose of the scalar product kA from the transpose of A.

Transpose Intro Video

Definition. The *transpose* of a matrix A, denoted A^T, is the matrix you get by ...

Example. Find the transpose of $A = \begin{bmatrix} -3 & 4 & 7 \\ 1 & 0 & 5 \end{bmatrix}$

Note. If A has size $m \times n$ matrix, then A^T has size:

Properties of Matrix Transpose

Let A and B be two matrices for which the following operations are defined. Which of these properties necessarily hold?

1. True or False: $(A + B)^T = A^T + B^T$

2. True or False: $(AB)^T = A^T B^T$

3. True or False: $(kA)^T = kA^T$

4. True or False: $(A^{-1})^T = (A^T)^{-1}$

5. True or False: $(A^T)^T = A$

END OF VIDEO

More Transpose Examples

Example. Find the transpose of the matrix $\begin{bmatrix} 4 & 2 \\ -3 & 17 \\ 9 & 6 \end{bmatrix}$

Question. If the matrix A is a 3×5 matrix, what is the dimension of A^T?

Question. If the matrix A has the number 7 in its second row and fifth column (i.e. as element a_{25}) then where can that number 7 be found in the matrix A^T?

Question. Which one of these statements is false?

A. $(A^T)^T = A$

B. $(kA)^T = kA^T$

C. $(A + B)^T = A^T + B^T$

D. $(AB)^T = A^T B^T$

E. $(A^{-1})^T = (A^T)^{-1}$

Symmetric and Skew-Symmetric Matrices

Example. Find the transpose of these two matrices. How does the transpose compare to the original?

$$A = \begin{bmatrix} 4 & -2 & 3 \\ -2 & 6 & 1 \\ 3 & 1 & 5 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 7 & -1 \\ -7 & 0 & 4 \\ 1 & -4 & 0 \end{bmatrix}$$

Definition. A square matrix A is called *symmetric* if ...

Definition. A square matrix B is called *skew-symmetric* if ...

Example. Which of these matrices are symmetric and which are skew-symmetric?

$$\begin{bmatrix} 4 & 5 & -1 \\ 5 & -11 & 3 \\ -1 & 3 & 2 \end{bmatrix} \qquad \begin{bmatrix} 0 & -8 & 9 & 6 \\ 8 & 0 & 3 & -4 \\ -9 & -3 & 0 & 5 \\ -6 & 4 & -5 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

Question. For a square matrix A, is $A + A^T$ always symmetric? or always skew-symmetric? or neither of these?

Question. For a square matrix A, is $A - A^T$ always symmetric? or always skew-symmetric? or neither of these?

Question. For any matrix A, is AA^T always symmetric? or always skew-symmetric? or neither of these?

Example. For the square matrix $A = \begin{bmatrix} 5 & -7 & 3 \\ -1 & 2 & 1 \\ 10 & -4 & 6 \end{bmatrix}$

1. Calculate the matrix $\frac{1}{2}(A + A^T)$. Is it symmetric, skew-symmetric, or neither?

2. Calculate the matrix $\frac{1}{2}(A - A^T)$ Is it symmetric, skew-symmetric, or neither?

3. Add together the two matrices from the previous two steps. What do you notice?

4. True or False: for any square matrix A, it is possible to write A as the sum of a symmetric matrix and a skew-symmetric matrix.

The Trace of a Matrix

After completing this section, students should be able to:

- Find the trace of a square matrix
- Find the trace of the sum or difference of two matrices from the trace of each of the matrices
- Find the trace of a scalar multiple of a matrix from the trace of the matrix
- Find the trace of the transpose of a matrix from the trace of the matrix.
- Use the fact that $tr(AB) = tr(BA)$ to simplify expressions.

Definition. The *trace* of an $n \times n$ matrix A, written $\operatorname{tr}(A)$ is ...

Example. Find $\operatorname{tr}(B)$ where $B = \begin{bmatrix} 4 & -3 & 2 & 9 \\ -5 & 1 & 7 & 11 \\ 2 & 1 & 3 & -8 \\ 3 & 2 & -4 & 5 \end{bmatrix}$

Let A and B be $n \times n$ matrices, and let k be a number. Then

1. $\operatorname{tr}(A + B) =$

2. $\operatorname{tr}(A - B) =$

3. $\operatorname{tr}(kA) =$

4. $\operatorname{tr}(AB) =$

5. $\text{tr}(A^T) =$

6. $\text{tr}(A^{-1}) =$

END OF VIDEO

Example. Find the trace of these two matrices:

$$A = \begin{bmatrix} 5 & -2 & 3 & 4 \\ 1 & -6 & 7 & 10 \\ 5 & 8 & -4 & 0 \\ -3 & -7 & 11 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 9 & 2 \\ -9 & 3 & 4 & 7 \\ 2 & 3 & 0 & 6 \\ -5 & 10 & 2 & 5 \end{bmatrix}$$

Question. What is:

1. $tr(A + B)$

2. $tr(4A - B)$

3. $tr(A^T)$

Example. For the matrices

$$A = \begin{bmatrix} 5 & -2 & 3 & 4 \\ 1 & -6 & 7 & 10 \\ 5 & 8 & -4 & 0 \\ -3 & -7 & 11 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 9 & 2 \\ -9 & 3 & 4 & 7 \\ 2 & 3 & 0 & 6 \\ -5 & 10 & 2 & 5 \end{bmatrix}$$

use python to find

- $tr(AB)$
- $tr(BA)$
- $tr(A^{-1})$
- $tr(B^{-1})$

What relationships, if any, do you find between these quantities and $tr(A)$ and $tr(B)$?

- True or False: $tr(AB) = tr(B)tr(A)$

- True or False: $tr(A^{-1}) = tr(A)^{-1}$

- True or False: $tr(AB) = tr(BA)$

The Determinant of a Matrix

After completing this section, students should be able to

- Compute the determinant of a square matrix.

Definition. The determinant of a 1×1 matrix $\begin{bmatrix} a \end{bmatrix}$ is

Definition. The determinant of a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is given by

Example. Find the determinant of the matrix $\begin{bmatrix} 3 & -2 \\ 4 & 5 \end{bmatrix}$

Definition. The determinant of a 3×3 matrix $\begin{bmatrix} a & b & c \\ d & e & f \\ i & j & k \end{bmatrix}$ can be computed

Example. Find the determinant of the matrix $\begin{bmatrix} 2 & 5 & 3 \\ 1 & -4 & 2 \\ 2 & 1 & -1 \end{bmatrix}$

Definition. If A is a 3×3 matrix, the determinant of the 2×2 matrix that you get by crossing out row i and column j is called the

and denoted

Example. For $A = \begin{bmatrix} 2 & 5 & 3 \\ 1 & -4 & 2 \\ 2 & 1 & -1 \end{bmatrix}$

$A_{1,1} =$ $A_{1,2} =$ $A_{1,3} =$

$\det(A) =$

Definition. If A is an $n \times n$ matrix, then the i, j minor of A is

and the determinant of A is given by $\det(A) =$

Example. Find the determinant of

$$B = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 1 & 3 & 1 \\ 4 & 2 & 0 & 1 \\ 2 & 1 & 1 & 2 \end{bmatrix}$$

END OF VIDEO

Idea: For a square matrix A, the determinant $\det(A)$ is a number that tells if ..

Definition. The determinant of a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is given by

Example. Find the determinant of the matrix $\begin{bmatrix} 5 & -7 \\ 2 & 4 \end{bmatrix}$

Notation: The determinant of the matrix $M = \begin{bmatrix} 5 & -7 \\ 2 & 4 \end{bmatrix}$ can be written using the following notation ...

Definition. The determinant of a 3×3 matrix $\begin{bmatrix} a & b & c \\ d & e & f \\ i & j & k \end{bmatrix}$ can be computed by ...

Example. Find the determinant of the matrix $A = \begin{bmatrix} 6 & 2 & 1 \\ -3 & 4 & 2 \\ 5 & -1 & 0 \end{bmatrix}$

What's so special about the first row? What happens if we expand along the second row instead?

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

What happens if we expand along the third row?

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

What happens if we expand along a column?

Example. Find the determinant of the matrix $B = \begin{bmatrix} 1 & -4 & 1 \\ 0 & 3 & 0 \\ 1 & 2 & 2 \end{bmatrix}$

Example. Find the determinant of the matrix $C = \begin{bmatrix} -3 & -5 & 2 & 5 \\ -2 & 4 & -3 & 4 \\ -5 & 1 & 0 & 0 \\ 5 & 4 & -3 & 3 \end{bmatrix}$

Extra Example. Find the determinant of the matrix $M = \begin{bmatrix} 2 & -1 & 4 & 4 \\ 3 & -3 & 3 & 2 \\ 0 & 4 & -5 & 1 \\ -2 & -5 & -2 & -5 \end{bmatrix}$

Properties of Determinants

After completing this section, students should be able to:

- Explain how elementary row operations on a matrix affect the determinant.
- Compute the determinant of a triangular matrix efficiently.
- Explain why the determinant of a matrix with two identical rows is 0
- Given values for $det(A)$ and $det(B)$, find values for $det(A^T)$, $det(A^{-1})$, $det(AB)$, and $det(kA)$ for a scalar k,
- Determine if a matrix is invertible or not based on its determinant.

The Determinant and Elementary Row Operations Video

Example. Find the determinant of the matrix $A = \begin{bmatrix} 2 & 1 & 5 \\ 2 & 1 & 5 \\ 0 & 3 & 4 \end{bmatrix}$.

For any square matrix A, if two rows (or two columns) of A are identical, then $\det(A) =$

Recall: There are three elementary row operations that we use to convert matrices to reduced row echelon form:

1.

2.

3.

Example. How do these elementary row operations affect the determinant of the matrix $A = \begin{bmatrix} 1 & 2 & -1 \\ 4 & 3 & 0 \\ 1 & 5 & -2 \end{bmatrix}$?

1. Swap the first two rows.

2. Multiply the first row by 5.

3. Add twice the first row to the second row.

Properties of Determinants Video

Example. Find the determinant of this matrix:

$$B = \begin{bmatrix} 3 & 2 & 4 & -5 \\ 0 & 5 & 7 & 10 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 10 \end{bmatrix}$$

Definition. An *upper triangular matrix* is a matrix with ...

Definition. A *lower triangular matrix* is a matrix with ...

Definition. A *triangular matrix* is a matrix that is

Note. The determinant of a triangular matrix is

Question. How does the determinant interact with other matrix operations?

1. $\det(kA) =$

2. $\det(A^T) =$

3. $\det(AB) =$

4. If A is invertible, then $\det(A^{-1}) =$

END OF VIDEOS

Determinant tricks

Example. What is the determinant of the matrix $B = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 & 0 \\ 1 & 7 & 3 & 0 & 0 \\ -9 & 13 & -4 & 10 & 0 \\ 8 & -5 & -5 & 6 & 1 \end{bmatrix}$?

Example. What is the determinant of the matrix $C = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 10 & 9 & 8 & 7 & 6 & 5 \\ 2 & 3 & 5 & 7 & 11 & 13 \\ 3 & 1 & 4 & 1 & 5 & 9 \\ 3 & 1 & 4 & 1 & 5 & 9 \\ 6 & 2 & 8 & 2 & 10 & 18 \end{bmatrix}$?

Determinants and elementary row operations

Example. For the matrix $M = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 0 & 4 \\ 5 & 2 & 0 \end{bmatrix}$, $\det(M) = -6$.

What is the determinant of each of the following matrices? Hint: how are they related to M?

(a) $A = \begin{bmatrix} 3 & 1 & 1 \\ 5 & 2 & 0 \\ -1 & 0 & 4 \end{bmatrix}$

(b) $B = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 0 & 4 \\ -15 & -6 & 0 \end{bmatrix}$

(c) $C = \begin{bmatrix} 33 & 13 & 1 \\ -1 & 0 & 4 \\ 5 & 2 & 0 \end{bmatrix}$.

Hint: the first row of C is equal to the first row of M plus a multiple of the third row of M.

Example. The matrix 3×3 matrix A row reduces to the identity matrix using the following steps:

1. $R_2 \leftrightarrow R_1$.
2. $-4R_1 + R_3 \to R_3$
3. $3R_2 + R_3 \to R_3$
4. $\frac{1}{2}R_3 \to R_3$
5. $-2R_3 + R_2 \to R_2$
6. $-3R_3 + R_1 \to R_1$

What is $\det(A)$?

The 4×4 matrix B row reduces to the matrix $\tilde{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ using the following steps:

1. $\frac{1}{3}R_1 \to R_1$
2. $4R_1 + R_2 \to R_2$
3. $-5R_1 + R_4 \to R_4$
4. $R_2 \leftrightarrow R_3$
5. $\frac{1}{2}R_3 \to R_3$
6. $-R_3 + R_4 \to R_4$

What is $\det(B)$?

Invertible Matrices

Question. Suppose a square matrix A row-reduces to a matrix $\tilde{A}$ with a row of 0's. What can you say about $\det(A)$?

Question. Suppose a square matrix A row-reduces to the identity matrix I. What can you say about $\det(A)$?

Question. Suppose a square matrix A is invertible. What can you say about $\det(A)$?

Question. Suppose a square matrix A is non-invertible. What can you say about $\det(A)$?

Example. Which of these matrices are invertible?

(a) $B = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 & 0 \\ 1 & 7 & 3 & 0 & 0 \\ -9 & 13 & -4 & 10 & 0 \\ 8 & -5 & -5 & 6 & 1 \end{bmatrix}$

(b) $C = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 10 & 9 & 8 & 7 & 6 & 5 \\ 2 & 3 & 5 & 7 & 11 & 13 \\ 3 & 1 & 4 & 1 & 5 & 9 \\ 3 & 1 & 4 & 1 & 5 & 9 \\ 6 & 2 & 8 & 2 & 10 & 18 \end{bmatrix}$

Determinants and other matrix properties

Example. Consider $A = \begin{bmatrix} 1 & 0 & 1 \\ -4 & 2 & 5 \\ -3 & 1 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 4 & 1 & -2 \\ 3 & 0 & 7 & 1 \\ 1 & 2 & -3 & 6 \\ 2 & 2 & 0 & 5 \end{bmatrix}$ and $C = \begin{bmatrix} 2 & 0 & 5 \\ -4 & 10 & 4 \\ 3 & 5 & 2 \end{bmatrix}$

Use Python to calculate the following quantities.

(a) $\det(A)$ and $\det(B)$.

(b) $\det(7A)$ and $\det(3B)$

(c) $\det(A^T)$ and $\det(B^T)$

(d) $\det(AC)$ and $\det(CA)$

(e) $\det(A^{-1})$ and $\det(B^{-1})$

Based on your experiments in Python, make a conjecture about how the following determinants can be written terms of $\det(A)$ and $\det(B)$, if A and B are square matrices of the same dimensions.

(a) $\det(kA)$

(b) $\det(A^T)$

(c) $\det(AB)$

(d) $\det(A^{-1})$

Example. If A and B are 5×5 matrices with $det(A) = -5$ and $det(B) = 4$, find each of the following

(a) $det(A^2B)$

(b) $det(A^3B^{-2})$

(c) $det(2A)$

(d) $det(AA^T)$

Extra Example. Use properties of determinants to solve for a. You do not need to find the determinants.

(a) $\begin{vmatrix} 4 & 3 & 1 \\ 2 & 6 & 7 \\ -5 & 10 & 8 \end{vmatrix} = a \begin{vmatrix} 1 & 3 & 4 \\ 7 & 6 & 2 \\ 8 & 10 & -5 \end{vmatrix}$

(b) $\begin{vmatrix} 4 & 3 & 1 \\ 2 & 6 & 7 \\ -5 & 10 & 8 \end{vmatrix} = a \begin{vmatrix} 4 & 3 & -5 \\ 2 & 6 & -35 \\ -5 & 10 & -40 \end{vmatrix}$

(c) $\begin{vmatrix} 4 & 3 & 1 \\ 2 & 6 & 7 \\ -5 & 10 & 8 \end{vmatrix} = a \begin{vmatrix} 9 & 3 & 1 \\ 37 & 6 & 7 \\ 35 & 10 & 8 \end{vmatrix}$

Hint: add the first column and a multiple of the third column of the left matrix

Linear Transformations

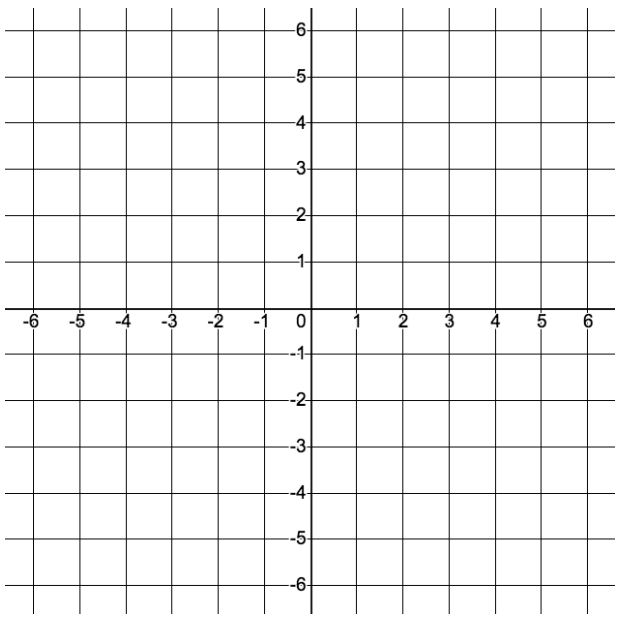
After completing this section, students should be able to:

- Match matrices with the transformation that they produce on vectors or images.
- Locate eigenvectors of a matrix by looking for vectors whose direction remains unchanged in direction after a transformation by that matrix.
- Write down a matrix based on information about where it takes the vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Multiplying Vectors by Matrices

Example. Consider the matrix $A = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ and the vectors $\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\vec{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.

Graph $\vec{x}$, $\vec{y}$, $\vec{v}$, $\vec{w}$, and $A\vec{x}$, $A\vec{y}$, $A\vec{v}$, and $A\vec{w}$ all on the same coordinate axes, using the same color for $\vec{x}$ and $A\vec{x}$, etc.



In the previous example, find a vector that is fixed (unchanged) under multiplication by A.

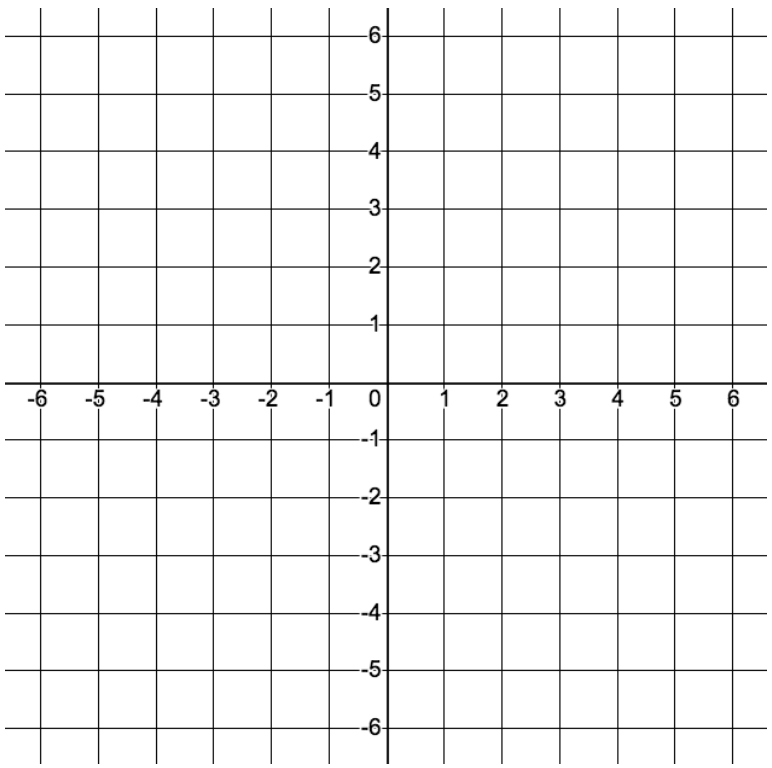
Definition. A non-zero vector $\vec{v}$ that is fixed under multiplication by A, or that is just multiplied by a scalar under multiplication by A, is called ...

That is, if $A\vec{v} = k\vec{v}$ for some scalar k, then $\vec{v}$ is called ...

The scalar amount k that $\vec{v}$ gets multiplied by is called ..

Example. Consider the matrix $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ and the vectors $\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\vec{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.

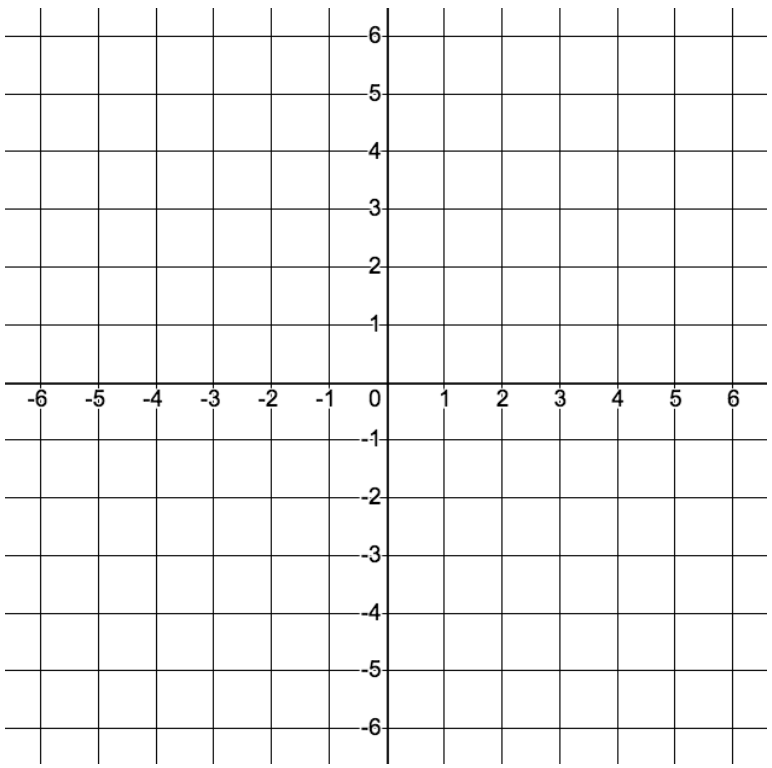
Graph $\vec{x}$, $\vec{y}$, $\vec{v}$, $\vec{w}$, and $A\vec{x}$, $A\vec{y}$, $A\vec{v}$, and $A\vec{w}$ all on the same coordinate axes, using the same color for $\vec{x}$ and $A\vec{x}$, etc.



Are there any eigenvectors and eigenvalues for this matrix A? If so, what are they?

Example. Consider the matrix $A = \begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix}$ and the vectors $\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\vec{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.

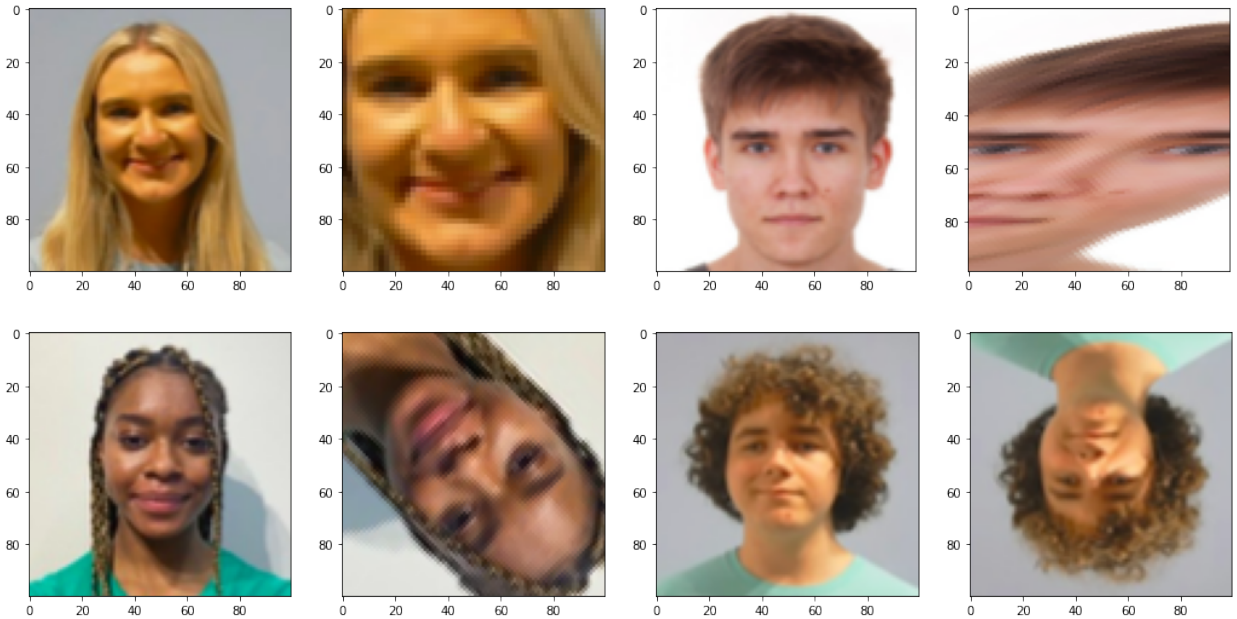
Graph $\vec{x}$, $\vec{y}$, $\vec{v}$, $\vec{w}$, and $A\vec{x}$, $A\vec{y}$, $A\vec{v}$, and $A\vec{w}$ all on the same coordinate axes, using the same color for $\vec{x}$ and $A\vec{x}$, etc.



Are there any eigenvectors and eigenvalues for this matrix A? If so, what are they?

Match the matrices with the before and after photos

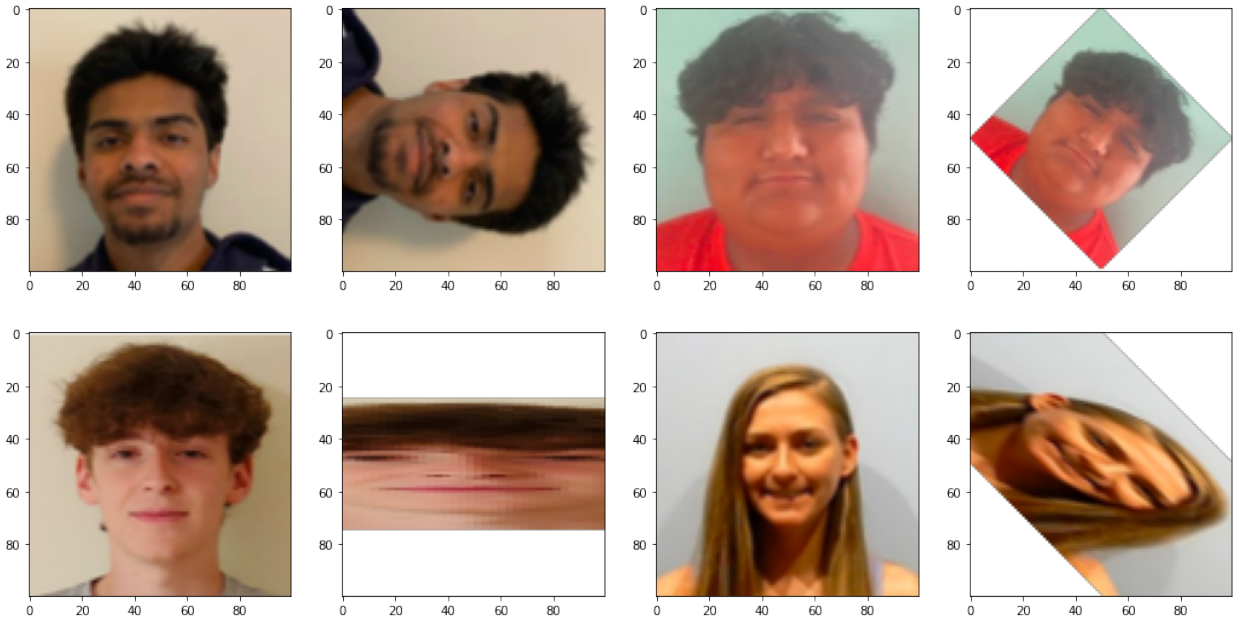
$$A = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad C = \begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$



Can you find any eigenvectors?

Match the matrices with the before and after photos

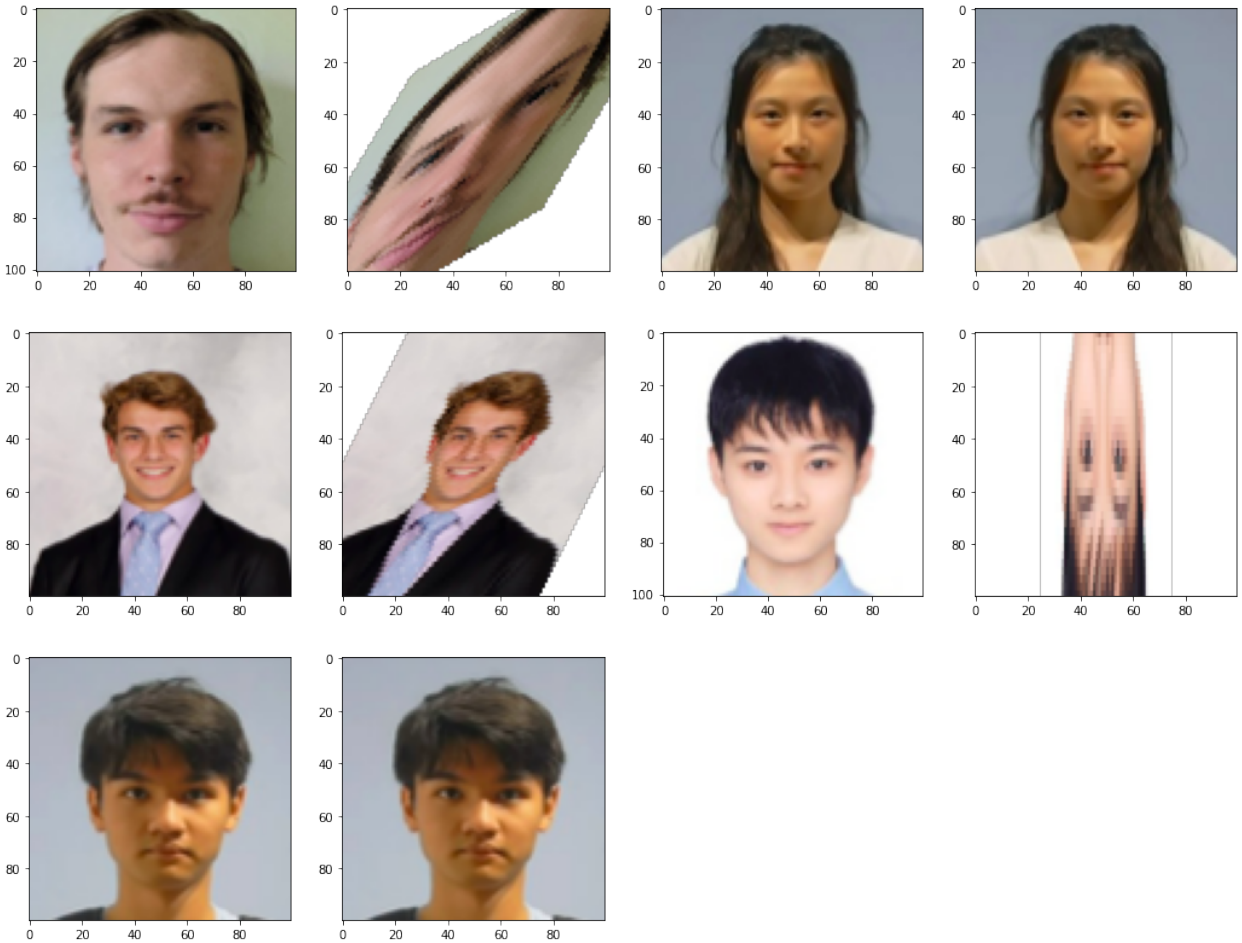
$$E = \begin{bmatrix} 0.5 & 0.5 \\ -0.5 & 0.5 \end{bmatrix} \qquad F = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \qquad G = \begin{bmatrix} 3 & 0 \\ 0 & 0.5 \end{bmatrix} \qquad H = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$



Can you find any eigenvectors?

Match the matrices with the before and after photos

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad J = \begin{bmatrix} 1.25 & 0.75 \\ 0.75 & 1.25 \end{bmatrix} \qquad K = \begin{bmatrix} 0.5 & 0 \\ 0 & -3 \end{bmatrix} \qquad L = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \qquad M = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}$$



Can you find any eigenvectors?

Eigenvalues and Eigenvectors

After completing this section, students should be able to:

- Define and eigenvalue and an eigenvector.
- Find the eigenvalues and eigenvectors of a matrix.

Example. Consider the matrix $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$ and the vectors

$\vec{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\qquad$ $\vec{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ $\qquad$ $\vec{w} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

Compute $A\vec{u}$, $A\vec{v}$, and $A\vec{w}$.

Definition. For an $n \times n$ square matrix A, suppose there is a scalar λ and a non-zero vector $\vec{x}$ such that $A\vec{x} = \lambda\vec{x}$.

Then λ is called

and $\vec{x}$ is called

Question. Does every matrix have eigenvalues and eigenvectors?

Question. How do we find the eigenvalues and eigenvectors for a matrix A?

Example. Find the eigenvalues for the matrix $A = \begin{bmatrix} 8 & 3 \\ 2 & 7 \end{bmatrix}$.

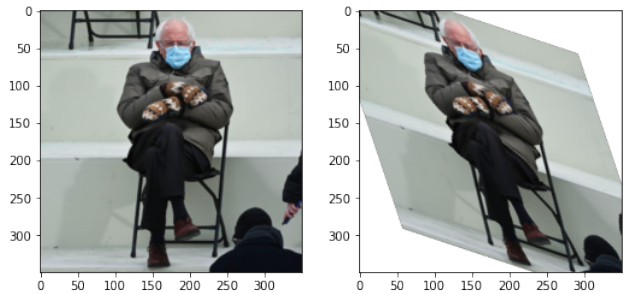
Example. Find the eigenvectors for the matrix $A = \begin{bmatrix} 8 & 3 \\ 2 & 7 \end{bmatrix}$.

END OF VIDEO

Recall: An eigenvector for a matrix A is a non-zero vector $\vec{x}$ such that ...

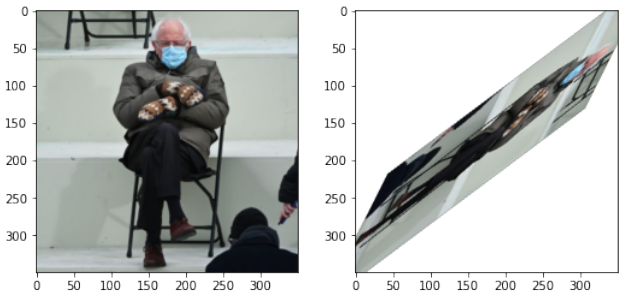
An eigenvalue for that eigenvector is ...

Example. For the matrix $A. = \begin{bmatrix} 1 & -\frac{1}{3} \\ -\frac{1}{3} & 1 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector. Find the corresponding eigenvalue. ′



Can you guess another eigenvector in another direction? Verify that is is an eigenvalue and find its eigenvalue.

Example. The matrix $B = \begin{bmatrix} \frac{1}{4} & 1 \\ \frac{1}{2} & \frac{3}{4} \end{bmatrix}$ has eigenvalues of $\frac{5}{4}$ and $-\frac{1}{4}$. Find an eigenvectors for these eigenvalues.



Theory

To find the eigenvalues of A,

To find the eigenvectors of A, for each eigenvalue λ_0,

Definition. Let A be an $n \times n$ matrix. The *characteristic polynomial* of A is the polynomial $p(\lambda) =$

Example. Find the eigenvalues and eigenvectors of the matrix

$$C = \begin{bmatrix} 2 & -12 \\ 2 & -8 \end{bmatrix}$$

Example. Find the eigenvalues and eigenvectors of the matrix

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

Extra Example. Find the eigenvalues and eigenvectors of the matrix

$$E = \begin{bmatrix} 3 & 12 \\ 1 & -1 \end{bmatrix}$$

Extra Example. Find the eigenvalues and eigenvectors of the matrix

$$F = \begin{bmatrix} -1 & 18 & 0 \\ 1 & 2 & 0 \\ 5 & -3 & -1 \end{bmatrix}$$

Properties of Eigenvalues

After completing this section, students should be able to:

- From the eigenvalues of A, find eigenvalues for A^{-1}, A^T, A^2, $5A$, and other related matrices.
- Predict if a matrix will have an eigenvalue of 0 or not from the determinant of A.
- Describe the relationship between the eigenvalues of a matrix and the trace and determinant of the matrix.

Example. Find the eigenvalues for this triangular matrix $B = \begin{bmatrix} 5 & 8 & 11 \\ 0 & 2 & 3 \\ 0 & 0 & 9 \end{bmatrix}$

Note. The eigenvalues for a triangular matrix are ...

Example. Find the trace, determinant, and eigenvalues for each of these two matrices

$$C = \begin{bmatrix} 2 & 2 \\ 5 & -1 \end{bmatrix} \qquad D = \begin{bmatrix} -3 & 0 & 0 \\ 9 & 7 & 0 \\ -4 & 10 & 5 \end{bmatrix}.$$

Note. The sum of the eigenvalues of a matrix is ...

the product of the eigenvalues of a matrix is ...

Question. How are the eigenvalues of A related to the eigenvalues of A^T and A^{-1}?

Question. Is it possible to have an eigenvalue of 0?

Note. A matrix A has an eigenvalue of 0 if and only if ...

END OF VIDEO

Question. What is the relationship between the trace of matrix and its eigenvalues?

Question. What is the relationship between the determinant of a matrix and its eigenvalues?

Example. Suppose the 2×2 matrix M has trace of -5 and determinant of -14. What are the eigenvalues of M?

Question. Which of the following statements are true?

(a) If A is invertible, then 0 is an eigenvalue for A.

(b) If A is not invertible, then 0 is an eigenvalue of A.

(c) $\lambda = 0$ is never an eigenvalue for any matrix A

(d) $\lambda = 0$ is always an eigenvalue for any matrix A

Example. Without doing any math, find an eigenvalue for the matrix $A = \begin{bmatrix} 6 & 6 & 7 \\ 3 & 3 & -1 \\ -2 & -2 & 9 \end{bmatrix}$

Example. If 7 is an eigenvalue of A with eigenvector $\vec{v}$, then find an eigenvalue of

(a) A^2

(b) A^{-1}

(c) A^T

(d) $A - I$

(e) $6A$

Question. True or False: If α and β are eigenvalues for A and B, respectively, then $\alpha \cdot \beta$ is an eigenvalue for AB.

Diagonalizing a Matrix

After completing this section, students should be able to:

- Explain what it means to diagonalize a square matrix.
- For an $n \times n$ matrix with distinct eigenvalues, use the eigenvalues and associated eigenvectors to diagonalize the matrix

Recall: A *diagonal matrix* is a square matrix ...

Definition. A square matrix A is said to be *diagonalizable* if ...

Example. Show that the matrix $A = \begin{bmatrix} -2 & 2 \\ -6 & 5 \end{bmatrix}$ is diagonalizable. Hint: try using $D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ and $P = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$.

Example. Find matrices P and D to diagonalize the matrix $A = \begin{bmatrix} 3 & 1 \\ -2 & 0 \end{bmatrix}$.

Summary:

Suppose A is an $n \times n$ matrix with n distinct real eigenvalues. Then A can be diagonalized as follows:

END OF VIDEO

True or False: True or False: If we multiply an $n \times n$ matrix A on the left by an $n \times n$ diagonal matrix D, to get DA, then each row of A gets multiplied by a corresponding diagonal entry of D.

True or False: If we multiply an $n \times n$ matrix A on the right by an $n \times n$ diagonal matrix D, to get AD, then each column of A gets multiplied by a corresponding diagonal entry of D.

True or False: If B is a matrix whose columns are the vectors $\vec{b_1}, \vec{b_2}, \vec{b_3}, \cdots \vec{b_n}$, then AB is a matrix whose columns are $A\vec{b_1}, A\vec{b_2}, A\vec{b_3}, \cdots A\vec{b_n}$.

True or False: If A is an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \cdots \lambda_n$, and corresponding eigenvectors $\vec{b_1}, \vec{b_2}, \cdots \vec{b_n}$, and B is the matrix with columns $\vec{b_1}, \vec{b_2}, \vec{b_3}, \cdots \vec{b_n}$ then $AB = BD$.

Hint: think about the previous two statements.

Fact: If the $n \times n$ matrix A has n distinct eigenvalues, then the matrix whose columns are the eigenvectors is invertible.

Question. How does the facts above show that an $n \times n$ matrix A with n distinct eigenvalues is diagonalizable?

Example. Consider the matrix $B = \begin{bmatrix} 6 & -3 & 7 \\ 4 & 1 & 5 \\ 4 & -3 & 9 \end{bmatrix}$

(a) Use technology to find the eigenvalues and eigenvectors for B

(b) Diagonalize B as PDP^{-1} for a diagonal matrix D and an invertible matrix P.

(c) Use technology to verify that when you multiply PDP^{-1} you do in fact get back B.

Example. Diagonalize the matrix $C = \begin{bmatrix} 7 & -8 \\ 4 & -5 \end{bmatrix}$ (by hand). Check your answer using technology.

Extra Example. Diagonalize the matrix $M = \begin{bmatrix} 2 & -2 \\ 0 & 0 \end{bmatrix}$

Extra Example. Diagonalize the matrix $N = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -3 & -2 \\ 2 & 4 & 2 \end{bmatrix}$

Example. Compute A^5 if $A = PDP^{-1}$, where $P = \begin{bmatrix} 2 & 1 \\ 7 & 3 \end{bmatrix}$ and $D = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$

Example. Find the determinant of A, if $A = PDP^{-1}$ for $P = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 3 & -1 \\ 4 & 3 & 0 \end{bmatrix}$ and $D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

Dot product and orthogonal vectors

After completing this section, students should be able to:

- Compute the dot product of two vectors
- Find the length of a vector
- Decide if two vectors are parallel, perpendicular, or neither
- Find two unit vectors that are parallel to a given (non-zero) vector
- Define orthogonal vectors and orthonormal vectors
- Determine if a set of vectors is orthogonal and / or orthonormal
- Convert a set of orthogonal vectors to a set of orthonormal vectors by rescaling the vectors

Dot Product Video

Example. The *dot product* of the two vectors $\vec{a} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix}$ is given by

Definition. The *dot product* of $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$ is given by

Properties of dot product:

Suppose $\vec{u}$, $\vec{v}$ and $\vec{w}$ are vectors, all of the same dimension, and c is a scalar.

1. $\vec{u} \cdot \vec{v} =$	
2. $\vec{u} \cdot (\vec{v} + \vec{w}) =$	
3. $(c\vec{u}) \cdot \vec{v} =$	
4. $\vec{u} \cdot \vec{u} \geq$	

Unit Vectors Video

Example. The *length* of the vector $\vec{a} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ is

Example. The *length* of the vector $\vec{b} = \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}$ is

Definition. The *length* of the vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ is given by

Definition. A *unit vector* is

Example. Is $\vec{a} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ a unit vector?

Example. For the vector $\vec{a} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ find a constant c such that $c\vec{b}$ is a unit vector.

Note. For any vector $\vec{v}$ (that is not the zero vector), the vector is a unit vector.

Example. Rescale the vector $\vec{b} = \begin{bmatrix} 7 \\ 3 \end{bmatrix}$ to be a unit vector.

Orthogonal Vectors Video

Definition. Two vectors $\vec{v}$ and $\vec{w}$ are said to be *orthogonal* if

Example. Which of the following pairs of vectors are orthogonal?

a) $\vec{a} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$

b) $\vec{c} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\vec{d} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

Definition. An *orthogonal set* of vectors is a collection of vectors such

Definition. An *orthonormal set* of vectors is a collection of vectors

Example. Show that the set of vectors $\vec{a} = \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}, \vec{b} = \begin{bmatrix} 11 \\ -1 \\ 7 \end{bmatrix}, \vec{c} = \begin{bmatrix} 3 \\ -2 \\ -5 \end{bmatrix}$ is an orthogonal set.

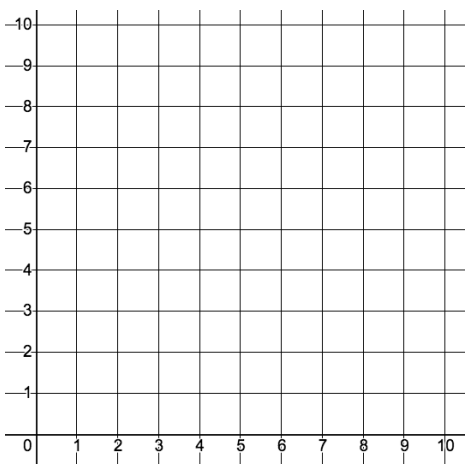
Is it an ortho*normal* set?

END OF VIDEOS

Perpendicular and Parallel Vectors

Example. Find the dot product of the two vectors $\vec{a} = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 3 \\ 5 \\ 6 \end{bmatrix}$.

Example. Graph the two vectors $\vec{u} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ and estimate the angle between them.



What is their dot product?

Note. For any two non-zero vectors $\vec{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, show that $\vec{a}$ and $\vec{b}$ are perpendicular if and only if their dot product is 0. Hint: what do you know about the rise over run for perpendicular vectors?

.

Fact: For vectors $\vec{a}$ and $\vec{b}$ of any dimension, $\vec{a}$ and $\vec{b}$ are perpendicular if and only if. ...

Definition. Two vectors $\vec{u}$ and $\vec{v}$ are called *parallel vectors* if ...

Extra Example. Find two unit vectors parallel to $\vec{m} = \begin{bmatrix} 2 \\ -4 \\ 4 \end{bmatrix}$?

Example. Determine whether the following pairs of vectors are parallel, perpendicular, or neither.

A. $\vec{u} = \begin{bmatrix} 6 \\ 3 \\ -9 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} -4 \\ -2 \\ 6 \end{bmatrix}$.

B. $\vec{w} = \begin{bmatrix} 5 \\ -4 \\ 3 \end{bmatrix}$ and $\vec{z} = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$.

C. $\vec{q} = \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix}$ and $\vec{r} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$.

Orthogonal and orthonormal vectors

Definition. Two vectors are called orthogonal if ...

Definition. Two vectors are called orthonormal if ...

Definition. A set of vectors is called orthogonal if ...

Definition. A set of vectors is called orthonormal if ...

Example. Is there a value of k that the following vectors form an orthogonal set? If so, find k. If not, explain why not.

$$\vec{u} = \begin{bmatrix} 1 \\ -3 \\ 2 \\ -1 \end{bmatrix}, \vec{v} = \begin{bmatrix} 4 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \vec{w} = \begin{bmatrix} -1 \\ 0 \\ k \\ 7 \end{bmatrix}$$

Example. Is there a value of k that the following vectors form an orthogonal set? If so, find k. If not, explain why not.

$$\vec{u} = \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}, \vec{v} = \begin{bmatrix} 7 \\ -1 \\ 3 \end{bmatrix}, \vec{w} = \begin{bmatrix} 4 \\ 1 \\ k \end{bmatrix}$$

Example. Verify that the following set of vectors forms an orthogonal set. Then rescale them to form an orthonormal set.

$$\vec{u_1} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \vec{u_2} = \begin{bmatrix} 5 \\ -4 \\ 1 \end{bmatrix}, \vec{u_3} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Orthogonal Matrices and Symmetric Matrices

After completing this section, students should be able to:

- Determine if a matrix is orthogonal (orthonormal) or not
- Explain why the inverse of an orthogonal matrix is its transpose.
- Describe properties of the eigenvalues and eigenvectors of symmetric matrices
- Describe a property of the eigenvalues of a symmetric matrix of the form $A^T A$
- Orthogonally diagonalize a symmetric matrix

Orthogonal Matrices Video

Definition. An *orthogonal matrix* is a square matrix with ...

Example. Verify that $B = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ is an orthogonal matrix.

Example. Compute $B^T B$.

Properties of orthogonal matrices:

For an orthogonal matrix Q,

1. $Q^T Q =$

2. $Q^{-1} =$

3. $QQ^T =$

Symmetric Matrices and Eigenvalues and Eigenvectors Video

Example. Find the eigenvalues and eigenvectors of the symmetric matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & 3 & 0 \end{bmatrix}$

Theorem. If A is a symmetric matrix, then the eigenvectors associated with distinct eigenvalues are ...

Theorem. If A is a symmetric matrix, then all of its eigenvalues are ...

Diagonalizing Symmetric Matrices Video

Recall: If A is an $n \times n$ matrix with n distinct real eigenvalues, then A can be written as ...

Theorem. If A is an $n \times n$ symmetric matrix with n distinct real eigenvalues, then A can be written as ...

Theorem. If A is an $n \times n$ symmetric matrix with n distinct real eigenvalues, then A can be written as PDP^{-1}, where D is a diagonal matrix and P is an orthogonal matrix.

Example. Orthogonally diagonalize the matrix $A = \begin{bmatrix} 5 & 2 \\ 2 & 8 \end{bmatrix}$

END OF VIDEOS

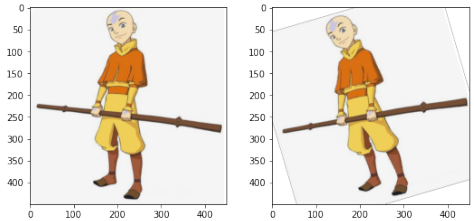
Orthogonal Matrices

Definition. An *orthogonal matrix* (or *orthonormal matrix* is a matrix ...

Example. Verify that the following matrix is an orthonormal matrix.

$$A = \begin{bmatrix} 0.96 & -0.28 \\ 0.28 & 0.96 \end{bmatrix}$$

What does the matrix A do to an image on the plane?


0
50
100
150
200
250
300
350
400
0
100
200
300
400
0
50
100
150
200
250
300
350
400
0
100
200
300
400

Example. Verify that the following matrix is an orthonormal matrix.

$$B = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Example. For each matrix A,

(i) compute the product $A^T A$

(ii) compute A^{-1}

(iii) compute AA^T

$$A = \begin{bmatrix} 0.96 & -0.28 \\ 0.28 & 0.96 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Example. For an orthonormal matrix A,

1. True or False: The dot product of any two distinct columns is 0, and the dot product of any column with itself is 1.

2. True or False: $A^T A = I$

3. True or False: $A^{-1} = A^T$

4. True or False: $AA^T = I$

5. True or False: The dot product of any two distinct rows is 0, and the dot product of any row with itself is 1.

Symmetric matrices and their eigenvalues and eigenvectors

Example. True or False:

1. True or False: The eigenvalues of a symmetric matrix are always real numbers.

2. True or False: The eigenvalues of a symmetric matrix are always non-negative real numbers.

3. True or False: The eigenvalues of the symmetric matrix $A^T A$ are always non-negative real numbers.

4. True or False: The eigenvectors of a symmetric matrix are always orthogonal.

5. True or False: The eigenvectors of a symmetric matrix are always orthoNORMAL.

6. True or False: A symmetric matrix an always be diagonalized as $A = PDP^{-1}$ where P is an orthogonal, invertible matrix and D is a diagonal matrix.

Example. The eigenvalues and corresponding eigenvectors for a symmetric matrix A are given. Find matrices D and P that orthogonally diagonalize A.

$\lambda_1 = 2, \vec{u_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \lambda_2 = -3, \vec{u_2} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

Example. The eigenvalues and corresponding eigenvectors for a symmetric matrix A are given. Find matrices D and P that orthogonally diagonalize A.

$$\lambda_1 = 0, \vec{u_1} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \lambda_2 = 2, \vec{u_2} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \lambda_3 = -1, \vec{u_2} = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

Extra Example. Use Python to find the eigenvalues and eigenvectors of the symmetric matrix A. Then find an orthogonal matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & 3 & 0 \end{bmatrix}$$

Linearly Independent Vectors

After completing this section, students should be able to:

1. Define a linear combination of vectors.
2. Determine if a set of vectors is linearly dependent or linearly independent.

Example. Consider the three vectors $\vec{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, and $\vec{c} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

A *linear combination* of these vectors is any sum of scalar multiples of the vectors, such as:

- $3\vec{a} - 4\vec{b}$

- $\vec{a} + \frac{1}{2}\vec{b} - 2\vec{c}$

Definition. A *linear combination* of vectors $\vec{v_1}, \vec{v_2}, \vec{v_3}, \cdots \vec{v_n}$ is any sum of scalar multiples of the vectors $c_1\vec{v_1} + c_2\vec{v_2} + c_3\vec{v_3} + \cdots c_n\vec{v_n}$.

Example. Is the vector $\vec{d} = \begin{bmatrix} 10 \\ 9 \\ 8 \\ 7 \end{bmatrix}$ a linear combination of $\vec{a} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, and $\vec{c} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$?

Definition. The vectors $\vec{v_1}, \vec{v_2}, \vec{v_3}, \cdots \vec{v_n}$ are called *linearly dependent* if ...

Equivalently, $\vec{v_1}, \vec{v_2}, \vec{v_3}, \cdots \vec{v_n}$ are *linearly dependent* if there are scalars $x_1, x_2, \cdots x_n$, not all zero, such that ...

Definition. The vectors $\vec{v_1}, \vec{v_2}, \vec{v_3}, \cdots \vec{v_n}$ are called *linearly independent* if ...

Equivalently, $\vec{v_1}, \vec{v_2}, \vec{v_3}, \cdots \vec{v_n}$ are *linearly independent* if the only possible scalars $x_1, x_2, \cdots x_n$ that make $x_1\vec{v_1} + x_2\vec{v_2} + x_3\vec{v_3} + \cdots x_n\vec{v_n} = 0$ are ...

Example. Are these three vectors linearly dependent or linearly independent?

$\vec{u} = \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 7 \\ 2 \\ -6 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} -2 \\ -1 \\ 8 \end{bmatrix}$

END OF VIDEOS

Example. Show that the set of vectors $\vec{v_1} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$, $\vec{v_2} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\vec{v_3} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is linearly dependent.

Example. Show that the set of vectors $\vec{v_1} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \vec{v_2} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \vec{v_3} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is linearly independent.

Example. Determine by inspection (that is, with only minimal computation) if the given vectors form a linearly dependent or linearly independent set.

A. $\vec{u} = \begin{bmatrix} 6 \\ -4 \\ 2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 3 \\ -2 \\ -1 \end{bmatrix}$

B. $\vec{u} = \begin{bmatrix} 1 \\ -8 \\ 3 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, $\vec{w} = \begin{bmatrix} -7 \\ 1 \\ 12 \end{bmatrix}$

C. $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$

Example. Determine if the vectors are linearly dependent or independent. If they are linearly dependent, find a linear combination of them that equals the zero vector.

$\vec{u} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \vec{v} = \begin{bmatrix} -4 \\ -2 \end{bmatrix}, \vec{w} = \begin{bmatrix} 9 \\ 2 \end{bmatrix}$

Example. Determine if the vectors are linearly dependent or independent. If they are linearly dependent, find a linear combination of them that equals the zero vector.

$$\vec{u} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \vec{v} = \begin{bmatrix} 0 \\ 4 \\ 1 \end{bmatrix}, \vec{w} = \begin{bmatrix} 2 \\ 4 \\ 7 \end{bmatrix}$$

Example. Determine if the vectors are linearly dependent or independent. If they are linearly dependent, find a linear combination of them that equals the zero vector.

$$\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}, \vec{v} = \begin{bmatrix} 1 \\ 6 \\ 6 \\ 5 \end{bmatrix}, \vec{w} = \begin{bmatrix} -1 \\ 2 \\ 0 \\ -1 \end{bmatrix}$$

Gram-Schmidt Orthonormalization

After completing this section, students should be able to:

- Explain the Gram-Schmidt algorithm.
- Convert a set of linearly independent vectors into the same size set of orthogonal vectors using the Gram-Schmidt algorithm.
- Explain why a set of orthogonal vectors is linearly independent.

Question. Suppose we have n vectors $\vec{v_1}, \vec{v_2}, \vec{v_3}, \ldots \vec{v_n}$, that are linearly independent. How can we use them to construct n vectors $\vec{w_1}, \vec{w_2}, \vec{w_3}, \ldots \vec{w_n}$ that are orthogonal?

Step 1: Fix $\vec{v_1}$

Step 2: Fix $\vec{v_2}$

Step 3: Fix $\vec{v_3}$.

$\vdots$

Step n: Fix $\vec{v_n}$

Question. Are the new vectors $\vec{w_1}, \vec{w_2}, \vec{w_3}, \ldots \vec{w_n}$ still linearly independent?

Question. Once we have $\vec{w_1}, \vec{w_2}, \vec{w_3}, \ldots \vec{w_n}$, a set of n non-zero orthogonal vectors, how can we build a set of n orthonormal vectors?

Example. Apply the Gram-Schmidt algorithm to the following set of vectors to get an orthogonal set of vectors.

$\vec{v_1} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$, $\vec{v_2} = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 3 \end{bmatrix}$, and $\vec{v_3} = \begin{bmatrix} -3 \\ -1 \\ 1 \\ 5 \end{bmatrix}$

END OF VIDEO

Example. Apply the Gram-Schmidt algorithm to the following set of vectors to get an orthogonal set of vectors.

$\vec{v_1} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \vec{v_2} = \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}.$

Convert them to an orthonormal set of vectors.

Example. Apply the Gram-Schmidt algorithm to the following set of vectors to get an orthogonal set of vectors.

$$\vec{v_1} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \vec{v_2} = \begin{bmatrix} 1 \\ 3 \\ 0 \\ 1 \end{bmatrix}, \vec{v_3} = \begin{bmatrix} 4 \\ 2 \\ 2 \\ 0 \end{bmatrix}$$

Example. Apply the Gram-Schmidt algorithm to the following set of vectors to get an orthogonal set of vectors.

$$\vec{v_1} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \vec{v_2} = \begin{bmatrix} 1 \\ 3 \\ 0 \\ 1 \end{bmatrix}, \vec{v_3} = \begin{bmatrix} 4 \\ 2 \\ 2 \\ 0 \end{bmatrix}, \vec{v_4} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix},$$

Example. Apply the Gram-Schmidt algorithm to the following set of vectors to get an orthogonal set of vectors.

$$\vec{v_1} = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}, \vec{v_2} = \begin{bmatrix} -5 \\ 1 \\ 5 \\ -7 \end{bmatrix}, \vec{v_3} = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 8 \end{bmatrix}$$

The Eigenvalues of A^TA

The Eigenvalues of $A^T A$

UNC-CH
Math 210

Theorem. The eigenvalues of a matrix $B = A^TA$ are all non-negative real numbers.

Singular Value Decomposition, Part 1

After completing this section, students should be able to:

- Explain what a singular value decomposition is
- Find a singular value decomposition for a matrix

Introduction Video

Recall: An $n \times n$ square matrix A is diagonalizable if there is a diagonal matrix D and an invertible matrix P such that ...

Recall: For an $n \times n$ symmetric matrix S, we can diagonalize S in such a way that ...

Recall: An $n \times n$ matrix is called a *diagonal matrix* if ...

Definition. An $m \times n$ matrix is called a *diagonal matrix* if

Definition. For an $m \times n$ matrix A, the singular value decomposition (SVD) of A is way of writing A as ...

Example. Consider the matrix $A = \begin{bmatrix} 0 & 2 \\ 1 & 3 \\ -2 & 0 \end{bmatrix}$. The following matrices can be used to construct a SVD of A:

$$U = \begin{bmatrix} \frac{3}{\sqrt{35}} & \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{14}} \\ \frac{\sqrt{5}}{\sqrt{7}} & 0 & \frac{\sqrt{2}}{\sqrt{7}} \\ -\frac{1}{\sqrt{35}} & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{14}} \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \sqrt{14} & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \qquad V = \begin{bmatrix} \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{bmatrix}$$

How Do We Find It Video

Question. For an $m \times n$ matrix A, how do we find orthogonal matrices U, and V, and a diagonal matrix Σ, such that $A = U\Sigma V^T$?

First consider the case when A is a "tall" matrix:

Question. For an $m \times n$ matrix A, how do we find orthogonal matrices U, and V, and a diagonal matrix Σ, such that $A = U\Sigma V^T$?

Next consider the case when A is a "wide" matrix:

Example. Find the singular value decomposition for the matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{bmatrix}$.

END OF VIDEOS

SVD Examples

Review. A singular value decomposition for a matrix A is ...

Example. Choose the matrices that correspond to U, Σ, and V^T in the singular value decomposition $M = U\Sigma V^T$ for the matrix

$$M = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$$

1. $\begin{bmatrix} \frac{3\sqrt{10}}{10} & -\frac{\sqrt{10}}{10} \\ \frac{\sqrt{10}}{10} & \frac{3\sqrt{10}}{10} \end{bmatrix}$
2. $\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{4}{9} & \frac{4}{9} & -\frac{2}{9} \end{bmatrix}$
3. $\begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}$

Order the steps for finding the SVD of a "tall" matrix.

- Compute $\dfrac{A\vec{v}_i}{\sigma_i}$, where the vectors $\vec{v}_i$ are _____ and the numbers σ_i are _____. These vectors will be _____
- Compute A^TA, which is a _____ matrix.
- Build the diagonal matrix Σ out of _____.
- Apply the Gram Schmidt orthonormalization process to make these vectors orthogonal, and rescale as needed.
- Rescale the vectors to have length 1.
- Build the matrix V from these vectors.
- Build the matrix U from these vectors.
- Find additional vectors that are linearly independent to these ones, if needed, until you have enough vectors to fill out the columns of the matrix you are building.
- Find the eigenvectors of A^TA, which will be _____ vectors.
- Find the eigenvalues of A^TA, which are _____ numbers.

If A is a "wide" matrix instead of a "tall" matrix, then ...

Example. Find the singular value decomposition for the matrix $A = \begin{bmatrix} 1 & 1 \\ 3 & 2 \\ 3 & -1 \\ 2 & 1 \end{bmatrix}$.

Example. Find the singular value decomposition for the matrix $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$.

Extra Example. Find the singular value decomposition for the matrix $A = \begin{bmatrix} 7 & 5 & 0 \\ 1 & 5 & 0 \end{bmatrix}$.

Extra Example. Find the singular value decomposition for the matrix $A = \begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix}$.

Singular Value Decomposition, Part 2

After completing this section, students should be able to:

- Explain why in the the singular value decomposition algorithm works, including
  - why the eigenvalues of A^TA are non-negative real numbers
  - why V is orthogonal
  - why the columns of U of the form $\frac{A\vec{v_i}}{\sigma_i}$ are orthogonal
  - why $U\Sigma V^T$ is equal to A
  - why it doesn't matter much what the "extra" columns of U are, past the ones of the form $\frac{A\vec{v_i}}{\sigma_i}$
- Show that the non-zero eigenvalues for A^TA are also the non-zero eigenvalues for AA^T.
- Show that the vectors we use for the first few columns of U, $\frac{Av_i}{\sigma_i}$, are eigenvectors for AA^T.
- Explain how a singular value decomposition can be used to approximate a matrix using smaller amounts of data

- Explain how a singular value decomposition is used in image compression

SVD - Why Does It Work - Video

Recall: We used the following steps to find the singular value decomposition for a "tall" matrix A.

1. Find the eigenvalues λ_i for A^TA and take their square roots $\sigma_i = \sqrt{\lambda_i}$ to build Σ.
2. Find the eigenvectors v_i for each eigenvalue λ_i to build V.
3. Compute $\dfrac{Av_i}{\sigma_i}$ (for non-zero σ_i) for the first few columns of U.

   Complete the matrix U using the Graham-Schmidt orthogonalization process to find other columns for U.

To show this makes a legit SVD for A, we need to show that:

- $A = U\Sigma V^T$
- V is an orthogonal matrix.
- U is an orthogonal matrix.

In addition we will see that:

- The non-zero eigenvalues for A^TA are also the non-zero eigenvalues for AA^T.
- The vectors we use for the first few columns of U, $\dfrac{Av_i}{\sigma_i}$, are eigenvectors for AA^T.

Show: V is an orthogonal matrix.

Show: U is an orthogonal matrix.

Show: $A = U\Sigma V^T$

Show: The non-zero eigenvalues for A^TA are also the non-zero eigenvalues for AA^T.

Show: The vectors we use for the first few columns of U, $\dfrac{Av_i}{\sigma_i}$, are eigenvectors for AA^T.

END of VIDEO

Why it works - Review

Review. How do we know that we can take the square roots of the eigenvalues of $A^T A$ (to get the singular values of A) without running into trouble taking the square root of a negative number, or a complex number?

Review. How do we know that V is an orthonormal matrix?

Review. How do we know that the columns of U of the form $\dfrac{A\vec{v}_i}{\sigma_i}$ are orthonormal?

Review. How do we know that $U\Sigma V^T$ really works out to equal A?

Question. Why does it not really matter what we do to "fill out" U with remaining columns?

Writing A with less bytes of data

Suppose A is a tall $m \times n$ matrix.

- The singular value matrix Σ contains a lot of zeros.

- Each row of zeros in Σ means a column of U that we can ignore.

- Each column of zeros in Σ means a row of V^T (i.e. a column of V) that we can ignore.

- If Σ has a lot of zero rows and columns, then the SVD of A allows us to use fewer numbers than A itself.

Example. Suppose the original A is an $N \times N$ matrix for some large number N, but it has only 20 non-zero singular values.

To write out A directly, how many numbers do we need to write down?

If we write down the SVD instead: $U\Sigma V^T$, how many numbers do we need to write down?

Example. Suppose the original A is an $M \times N$ tall matrix for some large numbers M and N, but it has only r non-zero singular values.

To write out A directly, how many numbers do we need to write down?

If we write down the SVD instead: $U\Sigma V^T$, how many numbers do we need to write down?

Approximating A

- Often some of the non-zero singular values of A are much smaller than others.

- By replacing the smaller singular values in Σ with 0's we can approximate A with smaller size matrices.

- We can make successive approximations to A by using just the first singular value (σ_1), then using the first two (σ_1 and σ_2), the first three (σ_1, σ_2, σ_3), etc.

- We can make successive approximations to A by using just the first singular value (σ_1), then using the first two (σ_1 and σ_2), the first three (σ_1, σ_2, σ_3), etc.
  - Using just σ_1 in $U\Sigma V^T$ corresponds to calculating $\sigma_1 \vec{u_1}\vec{v_1}^T$
  - Using just σ_1 and σ_2 in $U\Sigma V^T$ corresponds to calculating $\sigma_1 \vec{u_1}\vec{v_1}^T + \sigma_2 \vec{u_2}\vec{v_2}^T$
  - Using just σ_1, σ_2, σ_3 in $U\Sigma V^T$ corresponds to calculating $\sigma_1 \vec{u_1}\vec{v_1}^T + \sigma_2 \vec{u_2}\vec{v_2}^T + \sigma_3 \vec{u_3}\vec{v_3}^T$

Image Compression

An image is a large matrix, one element for each pixel.

If nearby pixels are related, then the picture is a good candidate for image compression using SVD.

See https://timbaumann.info/svd-image-compression-demo/ to see what it looks like to do successive approximations using more and more singular values.

Principal Component Analysis

After completing this section, students should be able to:

- Give a rough qualitative description of the goals of principal component analysis
- Outline the main steps than need to be done for principal component analysis

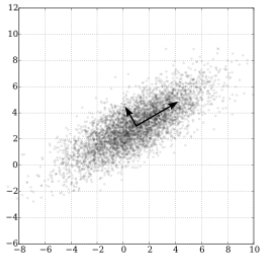
Suppose you have a data set with N observations, each with p variables. This can be written as an $p \times N$ matrix. Each column $\vec{x_i}$ is one observation.

Example. If you have 150 students, each with 15 homework scores, what dimension is the data matrix?

Example. If you have 45 factories, each with 8 values of various types of pollution, what dimension is the data matrix X?

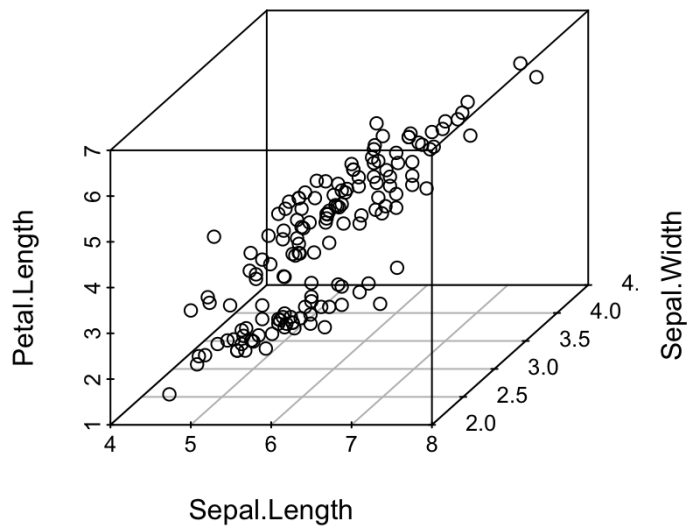
If we imagine plotting the data from N observations of p variables, how many points will we plot? How many axes will we need?

Often data is clustered around a line, or a plane, or a k=dimesional hyperplane where k is less than the dimension that we started with.
Example. This data set is clustered around a line.



PCA identifies the line that it is clustered around, that explains most of the variance (variability) of the data. The perpendicular direction to this line explains less of the variability.

Example. Estimate the directions that explain the most variance, second most, and third most variance for this 3-d data set, in such a way that these directions are perpendicular to each other.



What's the point?

The goals of PCA are as follows:

1. Find the perpendicular directions that explain from the most to the least variance of the data. These will correspond to the _____of a related matrix (called the covariance matrix). These perpendicular directions are called the principal components.
2. Describe the proportion of the variance in the data that comes from each of these directions. This will come from the _____of this related matrix (the covariance matrix).
3. Rewrite the data in new coordinates. Intuitively, this corresponds to "turning our heads" so that the principal components now just look like the x-axis, y-axis, etc. This new frame of reference gives us "uncorrelated" variables.
4. Simplify the dimension of the data set by ignoring all but the first few components of the data corresponding to the largest proportion of the variance. This corresponds to projecting the data onto ints first few principal components.

How do we do it?

Example. Suppose our observations are the ages, heights, and weights of 1000 NBA basketball players.

The first few individuals can be given in this table (matrix):

	player1	player2	player3	player4
age	22.0	27.0	30.0	29.0
height	213.36	210.82	208.28	210.82 $\cdots$
weight	106.59	106.59	106.59	111.13

Or, in this table (matrix):

	age	height	weight
player1	22.0	213.36	106.59
player2	27.0	210.82	106.59
player3	30.0	208.28	106.59
player4	29.0	210.82	111.13
$\vdots$	$\vdots$	$\vdots$	$\vdots$

So we have 1000 observations of 3 variables.

1. First, shift this data set so that each variable has a mean of 0. We can do this by subtracting the mean of each variable from each observation.

   Compute the mean of each variable.

   age mean = 28.11
   height mean = 200.76
   weight mean = 100.33

   Subtract these means from each data point. Let B be the matrix with these new columns.

	age	height	weight
player1	-6.11	12.60	6.26
player2	-1.11	10.06	6.26
player3	1.89	7.52	6.26
player4	0.89	10.06	10.80

2. Now compute (essentially) the SVD of B: which means we do the matrix multiplication _____ and compute its eigenvalues and eigenvectors.
   - The eigenvectors of $B^T B$ give you the *principal components* ...

     Recall: where do these eigenvectors appear in the SVD $U\Sigma V^T$?

   - The eigenvalues of _____ give you the *proportion of the variance of the data* in the direction of each of the principal components...

     Recall: what do these eigenvalues have to do with the singular values in Σ?

3. The eigenvectors with the largest eigenvalues are most important, since they explain most of the variance of the data.

4. We can simplify the data set by "projecting" onto the hyperplane of the first, say, k principal components, so that we only need k variables to represent the data, without much loss of accuracy.
5. We can do a change of variables so that the unit eigenvector directions correspond

to the x-direction, y-direction, etc.

The standard principal component analysis actually deals with the matrix $\frac{1}{N-1}B^TB$ instead of B^TB where N is the number of observations (individuals).

If we use $\frac{1}{N-1}B^TB$, we still get the same eigenvalues and are eigenvalues are just multiplied by $\frac{1}{N-1}$.

We use $\frac{1}{N-1}B^TB$ instead of B^TB because $\frac{1}{N-1}B^TB$ is the *covariance matrix*, its diagonal entries give the variance (the squares of the standard deviations) of each variable (age, height, weight) and the off-diagonal entries give the covariance, which is related to correlation.

The Idea of a Derivative

After completing this section, students should be able to:

- Estimate a derivative of a function at a point from a graph by drawing a tangent line and finding its slope
- Estimate the derivative of a function at a point from a table of values by finding an average rate of change
- Sketch the derivative of a function from the graph of the function

Consider a function $y = f(t)$.

For example, t could represent time in seconds past noon and y could represent the height of a mosquito over your head. Assume the mosquito is just flying straight up and down.

Here are some values of y at times t.

t	0	1	2	3	4	5	6	7	8	9	10
y	-1	-0.66	0.62	1.8	1.45	-0.77	-3.19	-3.06	0.72	5.51	6.2

The *derivative* of $y = f(t)$ at a given value of t is its rate of change at that time.

For example, the derivative of $y = f(t)$ at $t = 1$, represented with the notation $f'(1)$ (or $\frac{df}{dt}$ or $\frac{dy}{dt}$) is the rate at which the mosquito's height is changing 1 second after noon.

We can estimate $f'(1)$ by looking at the difference in y-values per difference in t-values for t-values near 1, for example from $t = 0$ to $t = 2$

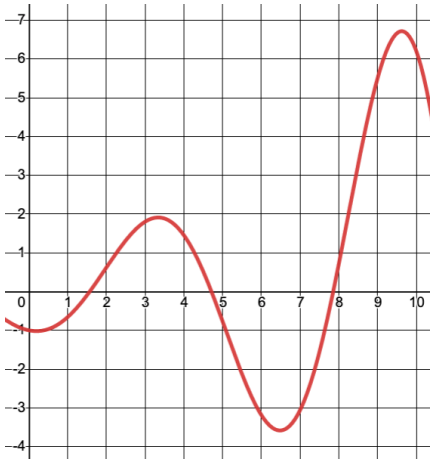
For a more precise estimate of $\frac{df}{dt}$, we need more data about values of y for values of t closer to $t = 1$.

t	0.5	0.6	0.7	0.8	0.9	1	1.1	1.2	1.3	1.4	1.5
y	-0.97	-0.93	-0.88	-0.82	-0.74	-0.66	-0.57	-0.46	-0.35	-0.22	-0.1

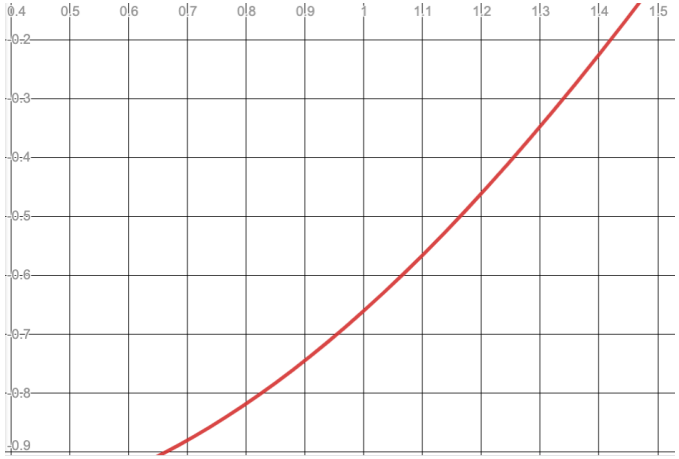
We could get even closer with more refined data.

t	0.95	0.96	0.97	0.98	0.99	1	1.01	1.02	1.03	1.04	1.05
y	-0.7	-0.69	-0.69	-0.68	-0.67	-0.66	-0.65	-0.64	-0.63	-0.62	-0.61

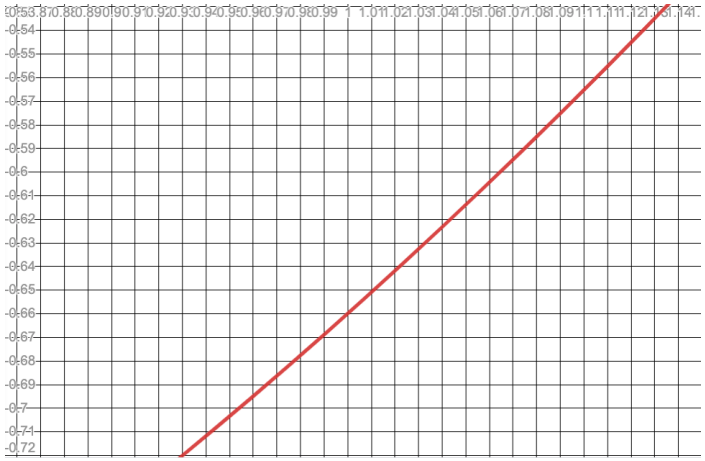
Geometrically, these calculations amount to approximating the slope of the function.



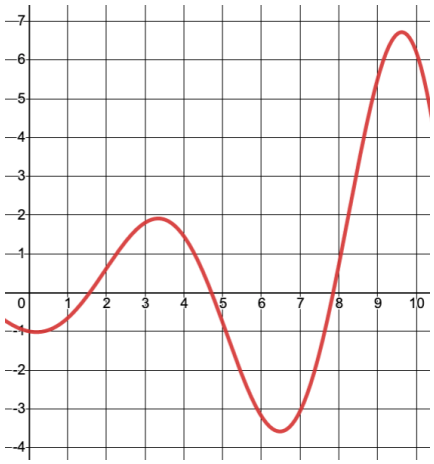
We can get a better approximation by zoming in.



By the time we zoom in this far, the function is looking like a line, so the slope estimate will give an accurate answer for the rate of change at time $t = 1$.



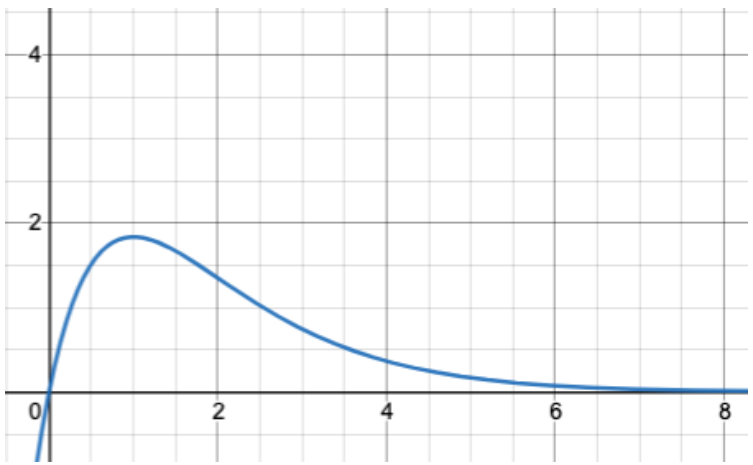
Another way to estimate the slope of the function at $t = 1$ from the graph is to draw its tangent line at $t = 1$. The tangent line at $t = 1$ is a line that goes though the point $(t, f(t)) = (1, f(1))$ and goes in the same direction as the graph.



Example. Estimate the derivative of $y = f(x)$ at $x = 3$ from this table of values.

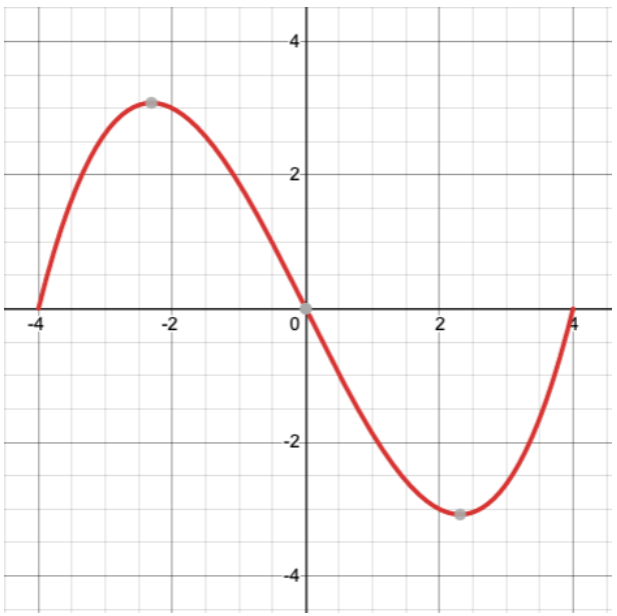
x	0	3	5	6	9	13
$f(x)$	72	95	112	77	54	32

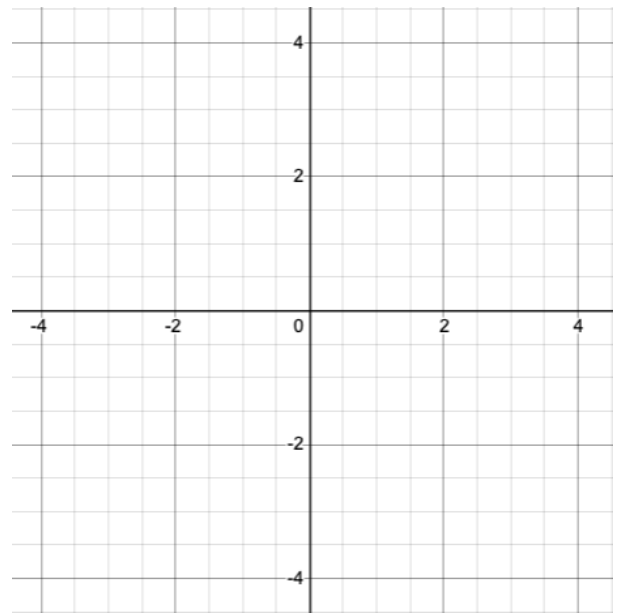
Example. Estimate $g'(4)$ from this graph.



For a function $f(x)$, the derivative $f'(x)$ is itself a function, which can be evaluated at any value of x by finding the rate of change of $f(x)$ at that x-value. For example:

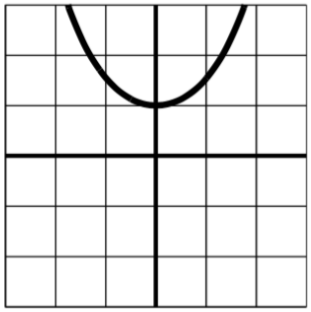
- $f'(2)$ is the rate of change, or slope, of $y = f(x)$ at $x = 2$
- $f'(3)$ is the rate of change, or slope, of $y = f(x)$ at $x = 3$
- $f'(3.2)$ is the rate of change, or slope, of $y = f(x)$ at $x = 3.2$



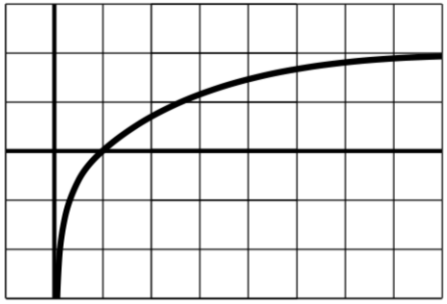


Draw a graph of $y = f'(x)$ on the empty axes at the right.

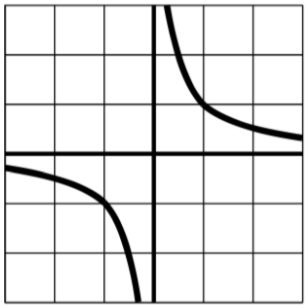
The graphs of several functions f are shown below. For each function, estimate the slope of the graph of f at various points. From your estimates, sketch graphs of f'.



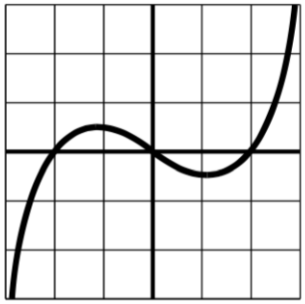
Graph 1



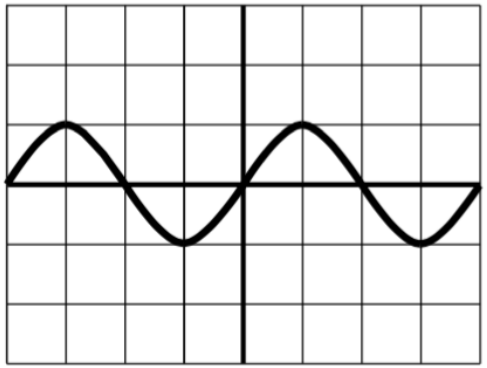
Graph 2



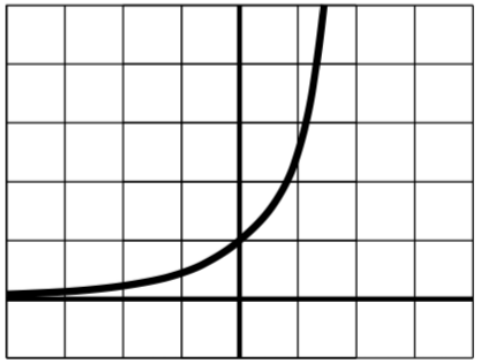
Graph 3



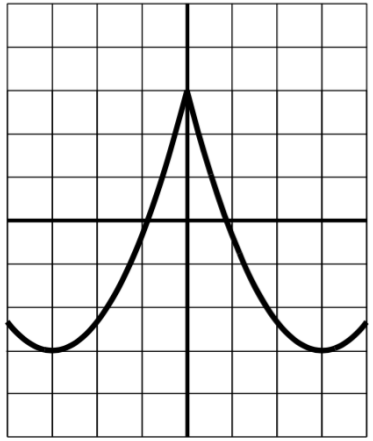
Graph 4



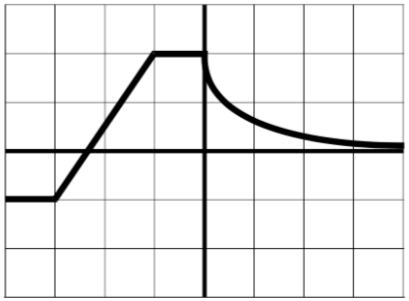
Graph 5



Graph 6



Graph 7



Graph 8

Calculating Derivatives

After completing this section, students should be able to:

- Calculate the derivative of a polynomial
- Calculate the derivative of an exponential function

1. Derivative of a constant

2. Derivative of $f(x) = x$

3. Power Rule

Example. Find the derivatives of these functions:

1. $y = x^{15}$

2. $f(x) = \sqrt[3]{x}$

3. $g(x) = \frac{1}{x^{3.7}}$

4. Derivative of a constant multiple

Example. Find the derivative of $f(x) = 5x^3$.

5. Derivative of a sum

6. Derivative of a difference

Example. Find the derivative of $y = 7x^3 - 5x^2 + 4x - 2$.

END OF VIDEOS

Using Derivative Rules

Question. Which of the following are correct? (c represents a constant)

A. $\frac{d}{dx}x^4 = 4x^3$

B. $\frac{d}{dx}e^{0.6x} = e^{0.6x}$

C. $\frac{d}{dx}\pi = 0$

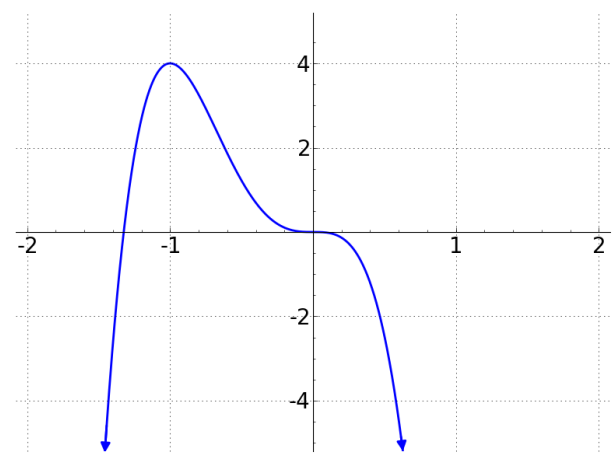
D. $\frac{d}{dx}5f(x) = 5\frac{d}{dx}f(x)$

E. $\frac{d}{dx}\left(f(x) + g(x)\right) = \frac{d}{dx}f(x) + \frac{d}{dx}g(x)$

F. $\frac{d}{dx}(5x^3 + 2x) = 5x^2 + 2$

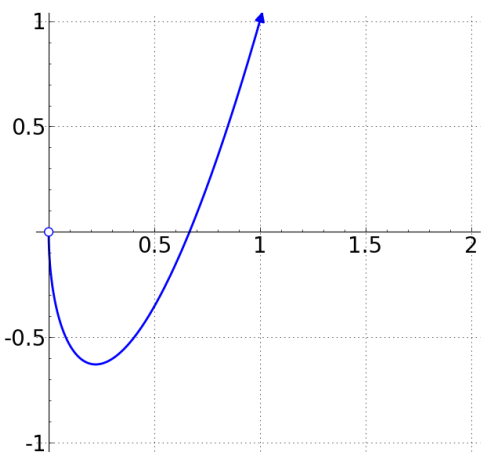
Example. At what x-values is the tangent line of this graph horizontal?

$$y = x^5 - 10x^4 - 15x^3$$



Example. Calculate the derivative: $g(t) = 4t^2 + \dfrac{1}{4t^2}$

Example. Find the derivative $y = 3x\sqrt{x} - 2\sqrt{x}$ at $x = \frac{1}{4}$.



Example. Find the derivative of $g(z) = ez^2 + 2e^z + ze^2 + z^{e^2}$.

Derivatives in Python

```
A, B, x = symbols('A B x')

diff(x**2 + 3*x, x)

diff( (A*x + B)/(x**3 + x - 2), x)
```

Example. Calculate $\frac{d}{dx}\left(4x^4 - \frac{1}{x}\right)$ by hand and then check your answer with Python. Use a double star ** for exponentiation, and a single star to indicate multiplication.

Use Python to calculate derivatives that we have not yet learned how to do:

(a) $\frac{d}{dx}\sin(x)$

(b) $\frac{d}{dy}(ye^y)$ Hint: you can write e^y as $\exp(y)$. Don't forget the star to indicate multiplication.

(c) $\frac{d}{dz}\sqrt{z^2 + 5z}$

Functions of Several Variables and Level Curves

After completing this section, students should be able to:

- Match equations of the form $z = f(x, y)$ to graphs of surfaces and graphs of level curves.
- Describe the graphs of functions of three variables $w = f(x, y, z)$ in terms of the level curves $f(x, y, z) = k$

Example. Consider the function of two variables $f(x, y) = \sqrt{xy}$

1. What is its domain?

For the function $f(x, y) = \sqrt{xy}$ ...

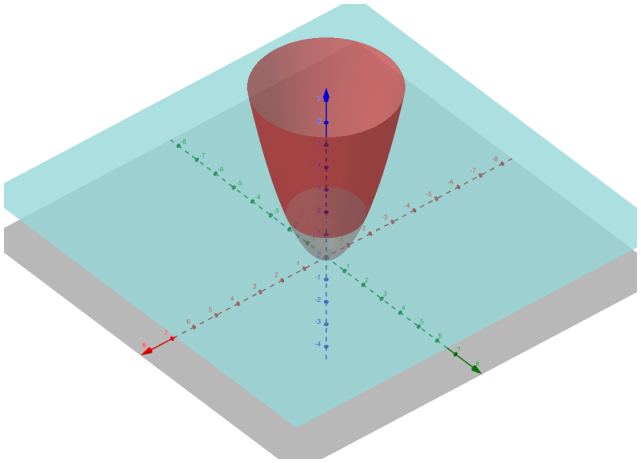
2. What are its level curves?

For the function $f(x, y) = \sqrt{xy}$ ...

3. What does its graph look like?

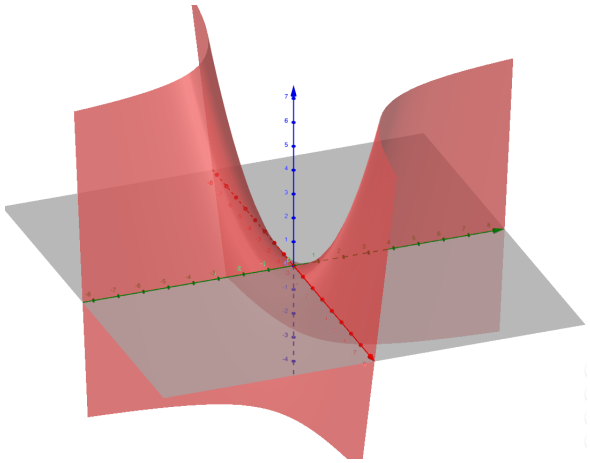
END OF VIDEO

Definition. A **level curve** of a surface $z = f(x, y)$ is ...



Definition. A **contour map** of a surface $z = f(x, y)$ is ...

Example. Consider the function of two variables $f(x, y) = xy$ graphed below.



1. What is the equation for the level curve at height 1?

2. What is the equation for the level curve at height -3?

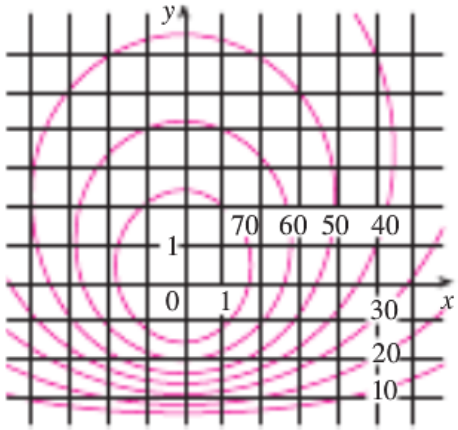
3. What is the equation for the level curve at height 0?

4. Graph some the level curves.

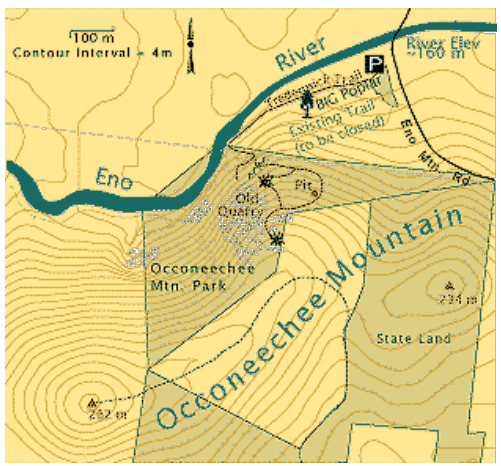
5. What dimension is needed to draw the graph the function $z = f(xy)$?

6. What dimension is needed to draw the level curves of (x, y)?

Example. A contour map for a function f is shown. Use it to estimate the values of $f(-3,3)$ and $f(3,-2)$. What can you say about the shape of the graph?

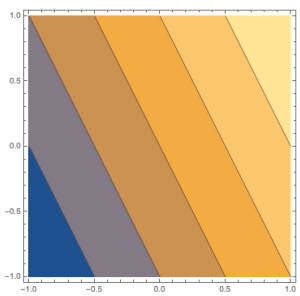
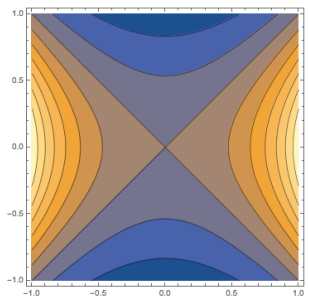
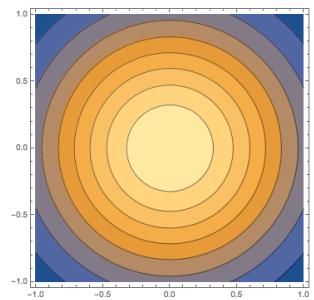
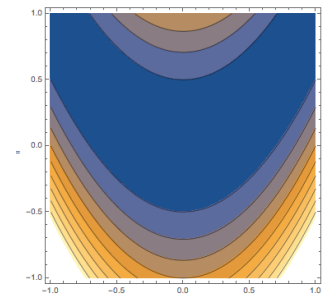


Question. What do the lines in this contour map represent? Where should you go if you like steep hills? Mountain tops?

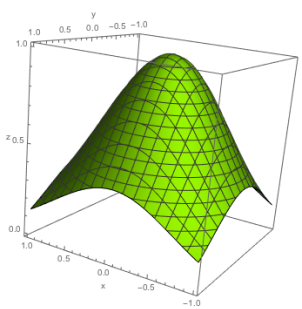
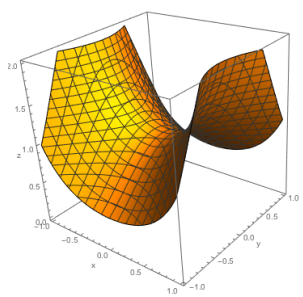
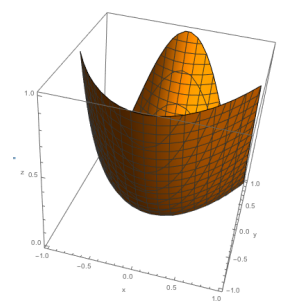
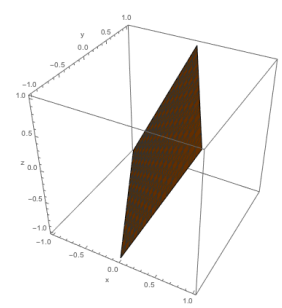


Example. What type of curve do you get when you intersect the graph of $z = (x^2 - y)^2$ with a horizontal plane?

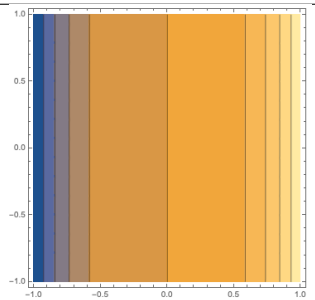
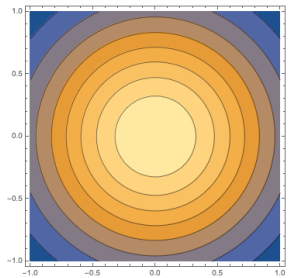
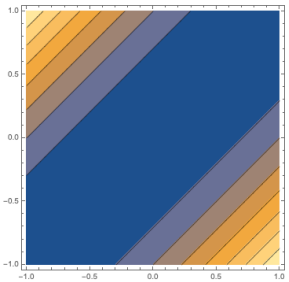
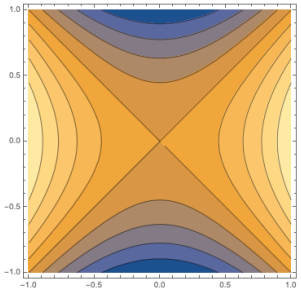
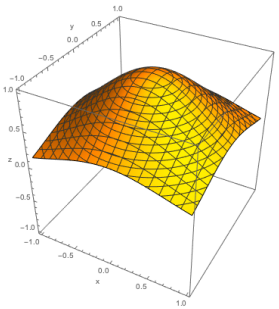
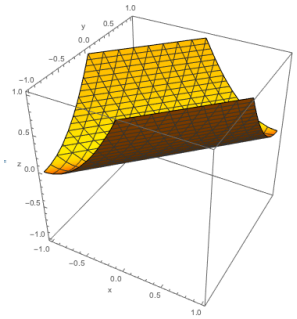
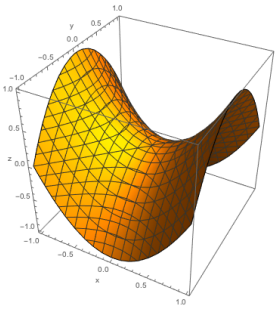
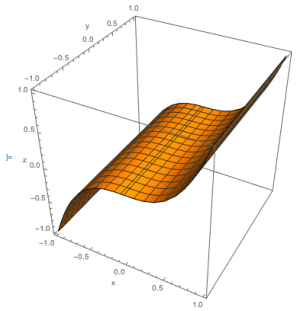
Example. Find the contour map for the surface $z = (x^2 - y)^2$.

A.  B.  C.  D. 

Example. Find the 3-d graph of the surface $z = (x^2 - y)^2$.

1.  2.  3.  4. 

Match the equations with the contour maps and the surface graphs.

1. $z = (x - y)^2$	2. $z = x^3$	3. $z = x^2 - y^2$	4. $z = e^{-(x^2+y^2)}$
I. 	II. 	III. 	IV. 
A. 	B. 	C. 	E. 

Functions of 3 or more variables

To visualize functions $f(x, y, z)$ of three variables, it is handy to look at **level surfaces**.
Example. $f(x, y, z) = x^2 + y^2 + z^2$

(a) Guess what the level surfaces should look like.

(b) Graph a few level surfaces (e.g. $x^2 + y^2 + z^2 = 10$, $x^2 + y^2 + z^2 = 20$, $x^2 + y^2 + z^2 = 30$) on a 3-d plot.

Example. $f(x, y, z) = x^2 - y^2 + z^2$

1. Guess what the level surfaces should look like.
2. Graph a few level surfaces (e.g. $x^2 - y^2 + z^2 = 0$, $x^2 - y^2 + z^2 = 10$, $x^2 - y^2 + z^2 = 20$) on a 3-d plot.

Partial Derivatives

After completing this section, students should be able to:

- Compute partial derivatives.
- Use average rates of change to approximate partial derivatives.
- For functions of two variables, explain the geometric meaning of a partial derivative as the slope of a tangent line to a curve in the intersection of a surface and a plane.

Example. The wave heights h in the open sea depend on the speed v of the wind and the length of time t that the wind has been blowing at that speed. So we write $h = f(v, t)$.

Duration (hours)

Wind speed (knots) v \ t	5	10	15	20	30	40	50
10	2	2	2	2	2	2	2
15	4	4	5	5	5	5	5
20	5	7	8	8	9	9	9
30	9	13	16	17	18	19	19
40	14	21	25	28	31	33	33
50	19	29	36	40	45	48	50
60	24	37	47	54	62	67	69

1. What is $f(40, 20)$?

2. If we fix duration at $t = 20$ hours and think of $g(v) = f(v, 20)$ as a function of v, what is the approximate value of the derivative $\left.\frac{dg}{dv}\right|_{v=40}$?

3. If we fix wind speed at 40 knots, and think of $k(t) = f(40, t)$ as a function of duration t, what is the approximate value of the derivative $\left.\frac{dk}{dt}\right|_{t=20}$?

Definition. For a function $f(x, y)$ defined near (a, b), the **partial derivatives** of f at (a, b) are:

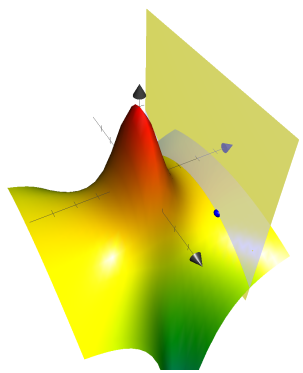
$f_x(a, b)$ = the derivative of $f(x, b)$ with respect to x when $x = a$,

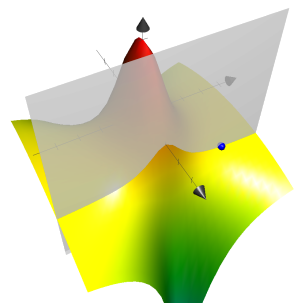
i.e. $f_x(a, b) = \frac{d}{dx} f(x, b)|_{x=a}$. We fix y and take the derivative with resect to x.

$f_y(a, b)$ = the derivative of $f(a, y)$ with respect to y when $y = b$

i.e. $f_x(a, b) = \frac{d}{dy} f(a, y)|_{y=b}$. We fix x and take the derivative with resect to y.

Geometrically, $f(x, b)$ can be thought of as





So $f_x(a, b) = \frac{d}{dx} f(x, b)|_{x=a}$ can be thought of as

Note. To compute f_x, we just take the derivative with x as our variable, holding all other variables constant. Similarly for the partial derivative with respect to any other variable.

Example. $f(x, y) = \frac{x}{y}$. Find $f_x(1, 2)$ and $f_y(1, 2)$.

Notation. There are many notations for partial derivatives, including the following:

f_x	$\frac{\partial f}{\partial x}$	$\frac{\partial z}{\partial x}$	f_1	$D_1 f$	$D_x f$

Note. Partial derivatives can also be taken for functions of three or more variables. For example, if $f(x, y, z, w)$ is a function of 4 variables, then $f_z(3, 4, 2, 7)$ means:

END OF VIDEO

Review. For a function $f(x, y)$ of two variables,

- the partial derivative $\dfrac{\partial f}{\partial x}$ means

- the partial derivative $\dfrac{\partial f}{\partial y}$ means

Example. This chart gives the heat index, or perceived temperature I as a function of actual temperature T and relative humidity h.

Relative humidity (%)

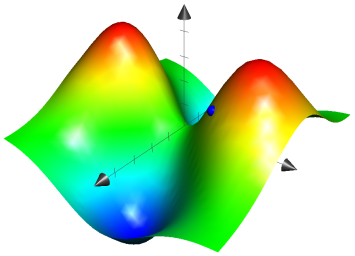
Actual temperature (°F) T \ H	50	55	60	65	70	75	80	85	90
90	96	98	100	103	106	109	112	115	119
92	100	103	105	108	112	115	119	123	128
94	104	107	111	114	118	122	127	132	137
96	109	113	116	121	125	130	135	141	146
98	114	118	123	127	133	138	144	150	157
100	119	124	129	135	141	147	154	161	168

(a) What is $I(96, 70)$?

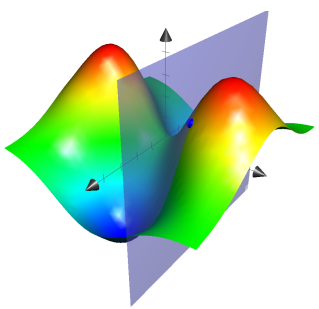
(b) Estimate $\dfrac{\partial I}{\partial T}(96, 70)$.

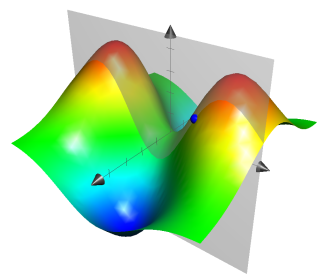
(c) Estimate $\dfrac{\partial I}{\partial h}(96, 70)$.

Based on the graph of $z = f(x, y)$ shown

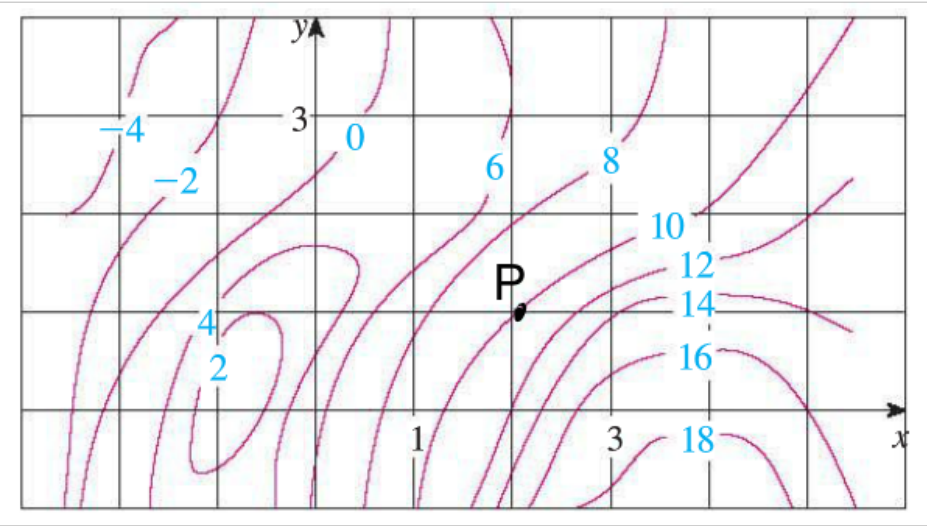


- is $\dfrac{\partial f}{\partial x}(1, 2)$ positive, negative, or 0?
- is $\dfrac{\partial f}{\partial y}(1, 2)$ positive, negative, or 0?





Example. Level curves are shown for a function f. Estimate the partial derivatives at the point P.



(a) f_x

(b) f_y

Example. For $g(x, y) = 3x^2y + 5y^2$, find $\dfrac{\partial g}{\partial x}$ at $(x, y) = (1, 2)$.

Example. For $f(x, y, z) = e^{3x}y^2 - 2e^{xy}z^3$, find f_x, f_y, and f_z.

Example. For $h(r, s, t) = \dfrac{s - t}{r^2}$, find h_r, h_s and h_t.

The Gradient

After completing this section, students should be able to:

- Find the gradient of a function of two or more variables.
- Find the direction of greatest increase for a function and the magnitude of that increase.
- Describe the relationship between the gradient vector and the level curves of a function.

Gradient

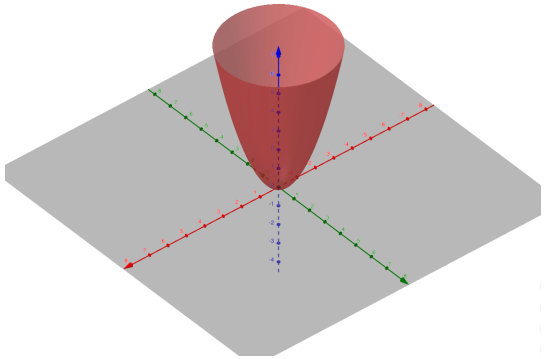
Definition. The **gradient** of the function $f(x, y)$ at (a, b) is defined as:

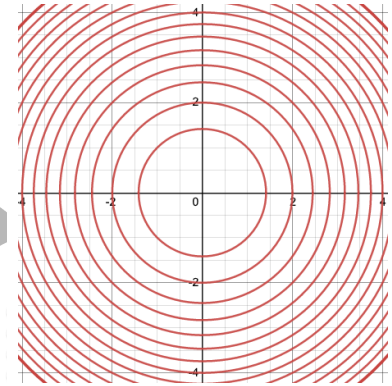
$$\nabla f(a, b) =$$

Example. What is $\nabla f(3, 4)$ for $f(x, y) = 2xy + 3y^2$?

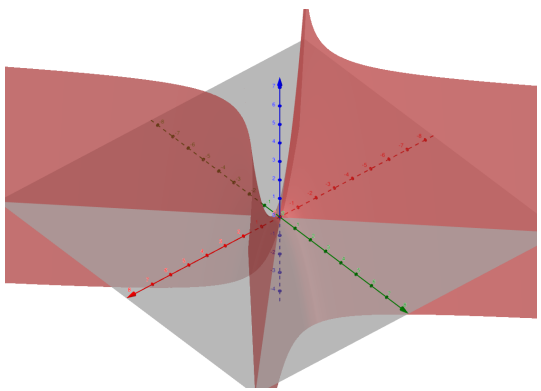
Example. For each function, calculate $\nabla f(2,0)$, $\nabla f(2,2)$, $\nabla f(0,3)$, and $\nabla f(-2,4)$ and draw these gradient vectors on the contour map.

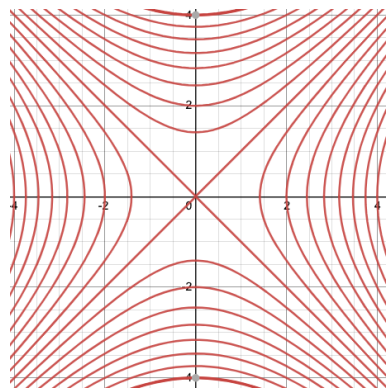
A. $f(x,y) = x^2 + y^2$





B. $g(x,y) = x^2 - y^2$





What is the relationship between the gradient vectors and the level curves?

Facts about the gradient

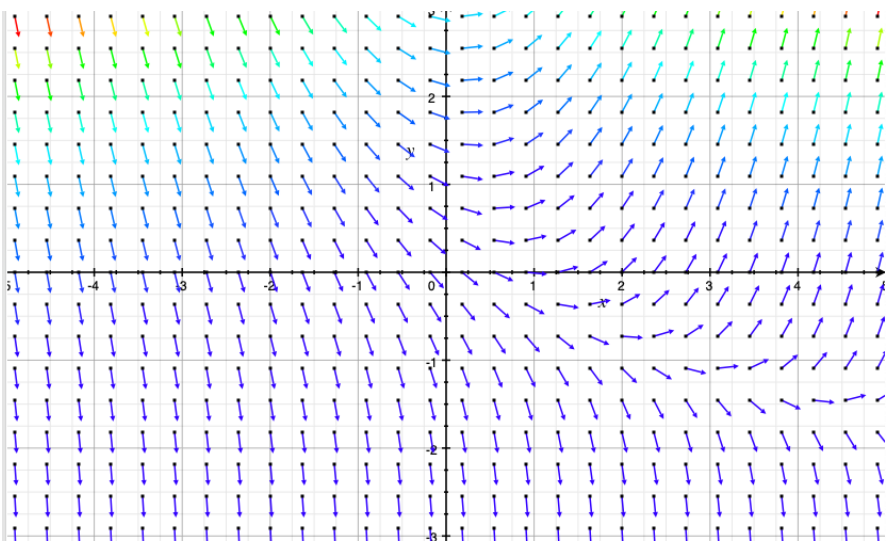
1. The gradient ∇f points ...

2. The negative of the gradient, $-\nabla f$, points ...

3. The gradient ∇f is perpendicular to ...

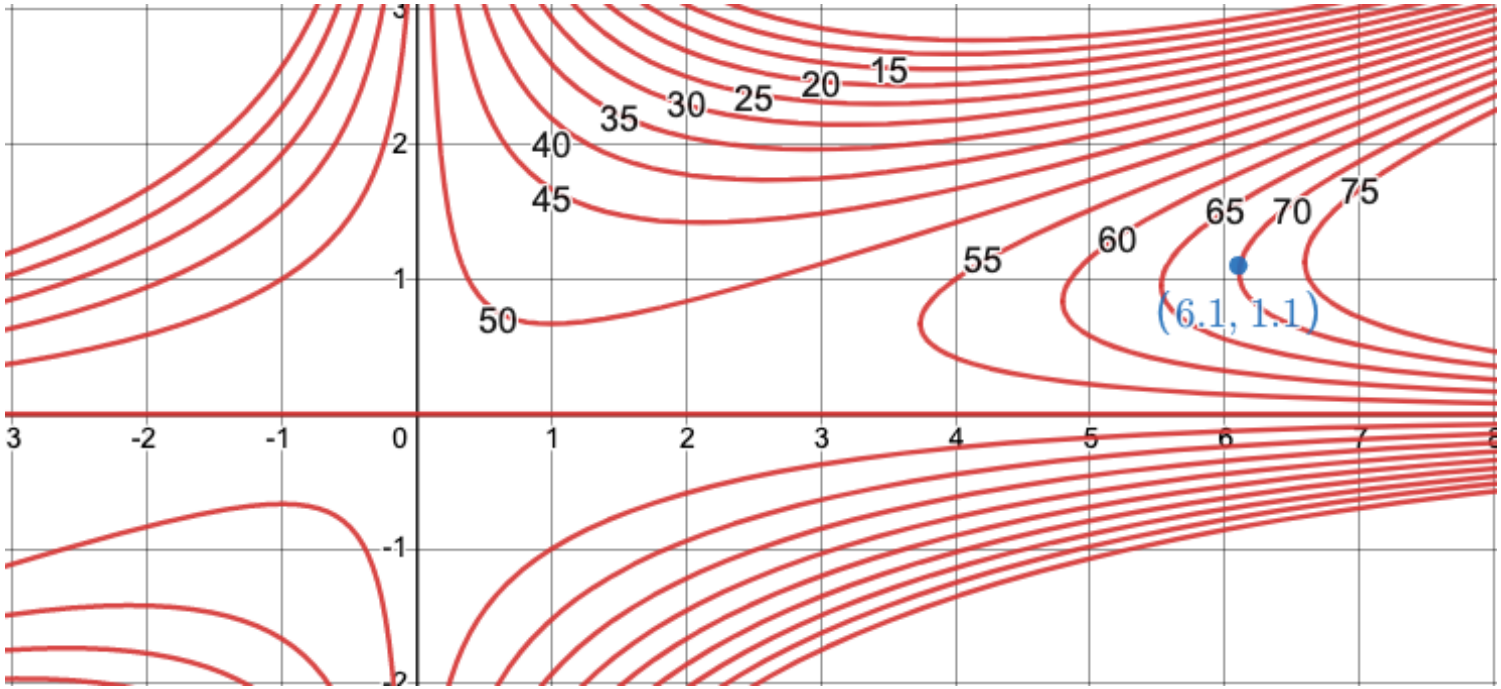
Example. Find ∇f for $f(x, y) = xe^y - y$, find ∇f for an arbitrary point (x, y).

Example. The graph of ∇f on the x-y plane drawn below. Use this graph to draw some of the level curves for f.



Draw the curve along which water would flow.

Example. Below is the graph of some level curves of a function.



(a) Draw the gradient at the point $(6.1, 1.1)$.

(b) Draw the path that water would flow, starting at this point.

Everything we have done so far can also be done for functions of three or more variables! For $f(x, y, z)$ and a point (x_0, y_0, z_0)

- $\nabla f(x_0, y_0, z_0) =$

- The direction of greatest increase at the point (x_0, y_0, z_0) is _____ .

- The direction of greatest decrease at the point (x_0, y_0, z_0) is _____ .

- $\nabla f(x, y, z)$ is perpendicular to _____ .

Is there such thing as a gradient for a function of one variable $f(x)$?

Extra Example. A bug at the point $(1, 2)$ observes that if it moves in the direction of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, the temperature increases at the rate of 2° per centimeter. If it moves in the direction $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, the temperature decreases at the rate of 3° per centimeter.

(a) In what direction should the bug move if it wants to warm up most rapidly?

(b) In what direction should the bug move if it wants to cool down most rapidly?

(c) In what direction should the bug move if it wants to change the temperature as little as possible?

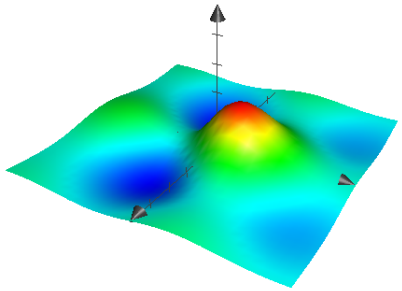
Gradient in Python.

Maximum and Minimum Values and the Method of Gradient Descent

After completing this section, students should be able to:

- Identify local and absolute maximum and minimum values from a graph.
- Explain the relationship between maximum and minimum points and the gradient.
- Explain the method of gradient descent.

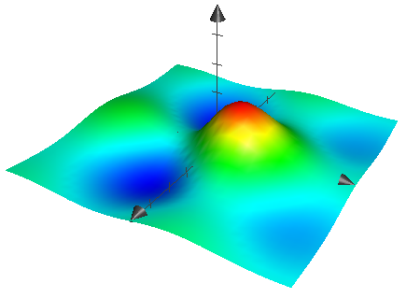
Definition. A function $f(x, y)$ has an absolute maximum at a point (a, b) if ...



and $f(x, y)$ has an absolute minimum at (a, b) if ...

The z-value $c = f(a, b)$ is called the maximum (or minimum) _____ and the point (a, b, c) is called the maximum (or minimum) _____ .

Definition. A function $f(x, y)$ has a local maximum at a point (a, b) if ...

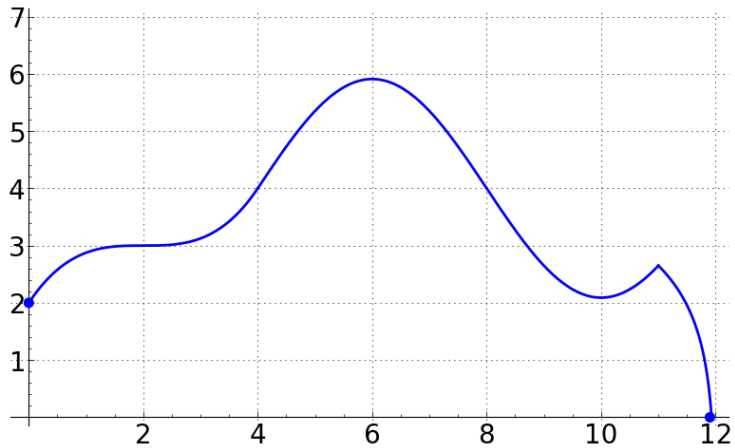


and $f(x, y)$ has a local minimum at (a, b) if ...

The z-value $c = f(a, b)$ is called the maximum (or minimum) _____ and the coordinates (a, b, c) is called the maximum (or minimum) _____ .

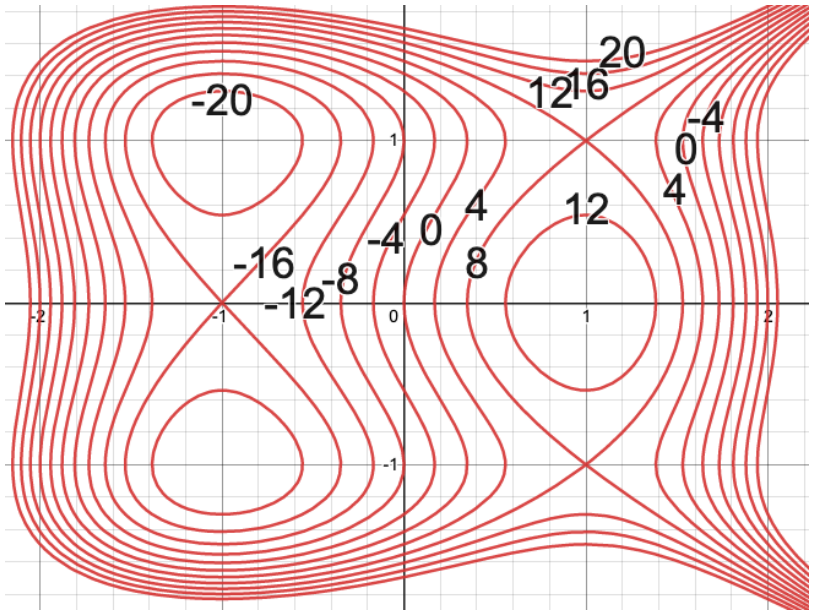
This same terminology can be used for functions of more than 2 variables, and for functions of less than two variables.

Example. Use the graph to find the absolute and local max and min points for the function.



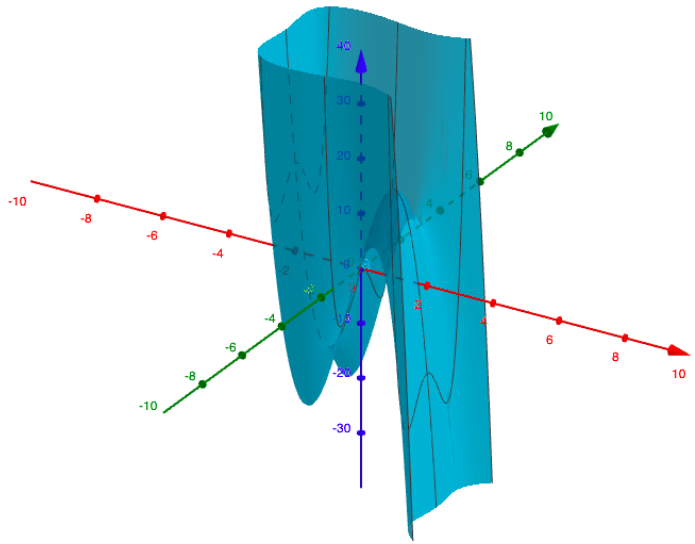
What do you notice about the derivative of f at maximum and minimum points?

Example. Use the contour map of the function $f(x, y) = 8(3x - x^3 - 2y^2 + y^4)$ to locate local and absolute maximum and minimum points for the function.

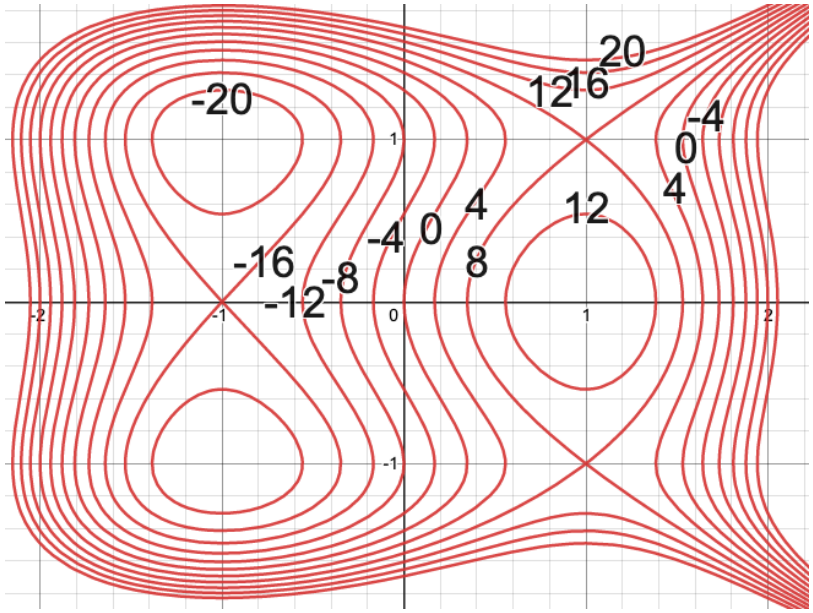


What do you notice about f_x and f_y at maximum and minimum points?

Here are graphs of the same function $f(x, y) = 8(3x - x^3 - 2y^2 + y^4)$.



If we start at the point $(0.5, 0.5)$ and we want to travel to a local max, how could we use the gradient to guide your path?



If we start at the point $(0.5, 0.5)$ and we want to travel to a local min, how could we use the gradient to guide your path?

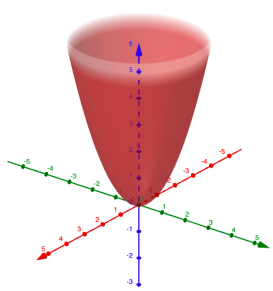
How can we avoid the danger of overshooting?

Method of Gradient Descent:

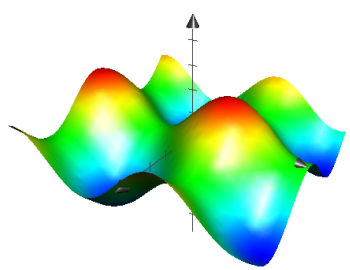
First, establish a step size, a number of steps, and an initial point.

- .
- .
- .

Example. Use the method of gradient descent to find the minimum value of $f(x, y) = x^2 + y^2$, using a step size of 0.1, and starting at the point $(4, 6)$.



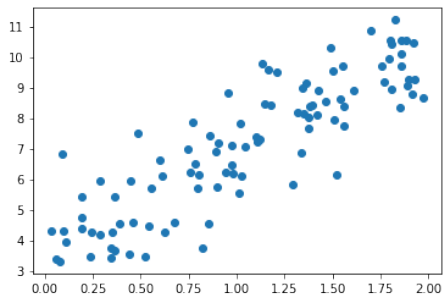
Example. Use the method of gradient descent on the function $f(x, y) = \cos(x) + \sin(y)$, starting at the point $(1, 1)$, using a step size of 0.2.



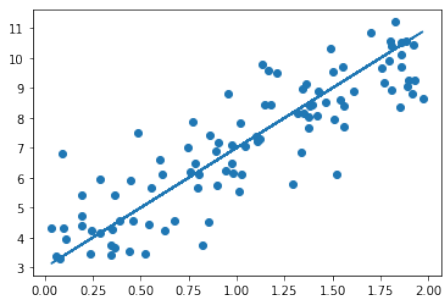
What are some ways that the method of gradient descent can go wrong?

Application: fitting a line to data

Suppose you want to fit a line $y = mx + b$ to this data set:



You could guess a value for m and b, for example, ...



But is this the best fit line?

A standard way to quantify how far off the line is from the points is by taking each point's vertical distance from the line and square it.

We can call this the error, or the cost.

For the 100 points in this data set, if we write each point as (x_i, y_i), then

$$\text{cost} = \sum_{i=1}^{100} (y_i - (mx_i + b))^2$$

We want to find the value of m and b that minimize this cost.

How can we use gradient descent to do this? How many variables does the function we are minimizing have?

For this problem, we can find an explicit formula for the gradient of the cost function, in terms of the x and y values of our data set. Find it!

Neural Networks

After completing this section, students should be able to:

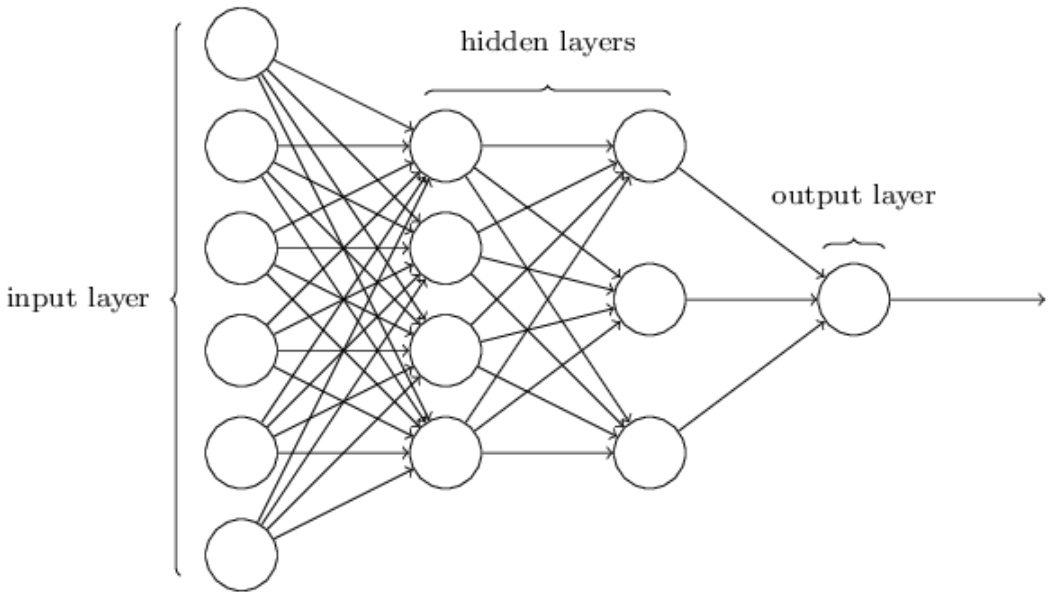
- Describe the structure of a neural network, including the nodes, weights, and biases.
- Explain how training a neural network is related to optimizing a function of many variables.

Machine Learning

Machine learning is often used to detect patterns:

- identify the ocean photos with sharks in them,
- recognize handwritten digits and letters,
- understand speech

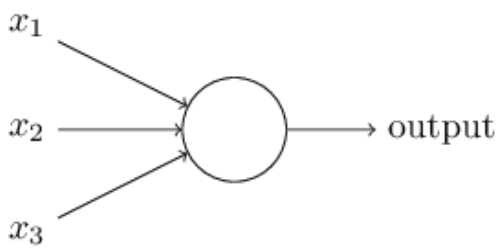
Sometimes a neural network is used.



- The neural network takes some input, e.g. the amount of white vs. black on each pixel of an image, and produces some output, e.g. a number between 0 and 1 at the end node, which gives us a level of certainty as to whether the picture contains a shark or not.
- Each node, or neuron, is connected to each node, or neuron in the next layer. How many total connections are there in this neural network?

Perceptrons

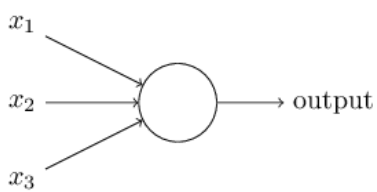
The simplest kind of neuron is a *perceptron*.



- A perceptron takes several binary inputs (numbers that are 0's or 1's) and creates a binary output (0 or 1). We will think of the output of 1 as firing, and the output of 0 as not firing.
- The output is determined by the inputs, together with *weights* and a *bias*.
- Each connection in has a weight $w_1, w_2, w_3, w_4, \cdots w_n$ (for n incoming connections), and the neuron itself has a bias b.
- The perceptron will fire if $w_1x_1 + w_2x_2 + w_3x_3 + \cdots w_nx_n$ compares favorably to b, i.e. $w_1x_1 + w_2x_2 + w_3x_3 + \cdots w_nx_n > -b$, i.e. $w_1x_1 + w_2x_2 + w_3x_3 + \cdots w_nx_n + b > 0$.

Example. Rewrite the equation $w_1x_1 + w_2x_2 + w_3x_3 + \cdots w_nx_n + b > 0$ in vector notation.

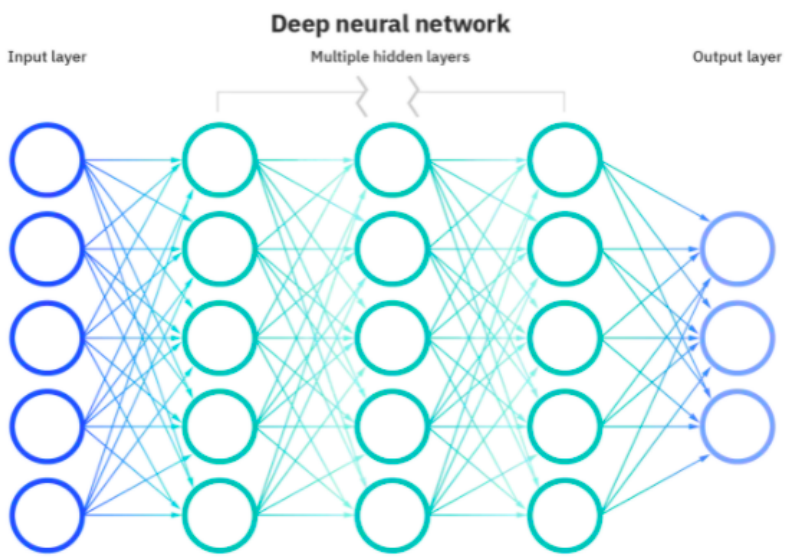
Example. For example, for this perceptron with three inputs,



if the weights are $w_1 = 5$, $w_2 = 3$, $w_3 = -2$, and the bias is $b = -4$, then determine the output for various inputs. The first two lines of the table are filled in already.

x_1	x_2	x_3	$w_1x_1 + w_2x_2 + w_3x_3$	$w_1x_1 + w_2x_2 + w_3x_3 + b$	output
0	1	1	1	-3	0
1	0	0	5	1	1
0	0	1			
1	1	1			
1	0	1			

- When we build a neural network out of perceptrons, we have to assign a weight to each connection and a bias to each perceptron (except possibly to the input layer). **Example.** How many weights and biases are there in this network?



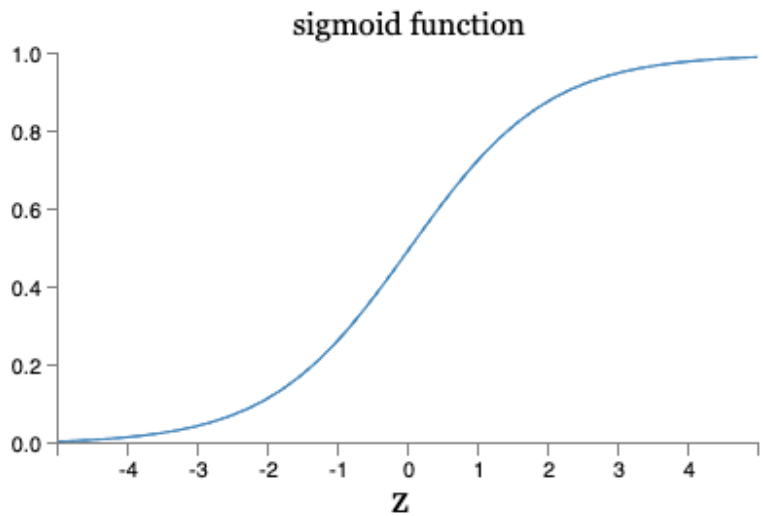
Different weights and biases will produce different behavior in the neural network.

A useful neural network is one that has weights and biases so that the final output successfully performs identification tasks, like picking out the pictures that contain sharks in them, at least most of the time.

Sigmoidal neurons

- One problem with perceptrons is that small changes in weights or biases can flip a perceptron's output from a 0 to a 1 or vice versa, possibly causing drastic changes downstream. This can make it difficult for the network to "learn".
- Sigmoidal neurons give an alternative architecture for which small changes in weights or biases cause only small changes in outputs.
- Sigmoidal neurons are based on the sigmoid function

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$



- Sigmoidal neurons take as input any number between $-\infty$ and ∞ and output a number between 0 and 1, according to the rule

$$\text{output} = \sigma(\vec{w} \circ \vec{x} + b)$$

where $\vec{w} = \begin{bmatrix} w_1 & w_2 & \cdots w_n \end{bmatrix}$ is the vector of weights, $\vec{x} = \begin{bmatrix} x_1 & x_2 & \cdots x_n \end{bmatrix}$ is the vector of inputs, and b is the bias of the neuron.

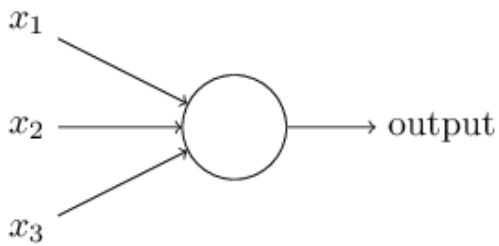
This can also be written as

Question. Suppose $\vec{w} \circ \vec{x} + b$ is a large positive number. What will the output of a perceptron be? What will be output of a sigmoidal neuron be?

Question. Suppose $\vec{w} \circ \vec{x} + b$ is a very negative number. What will the output of a perceptron be? What will be output of a sigmoidal neuron be?

Question. Suppose $\vec{w} \circ \vec{x} + b$ is close to 0. What will the output of a perceptron be? What will be output of a sigmoidal neuron be?

Example. Suppose this neuron is a sigmoidal neuron.



For weights of $w_1 = 5$, $w_2 = 3$, $w_3 = -2$, and a bias of $b = -4$, determine the output for the following inputs.

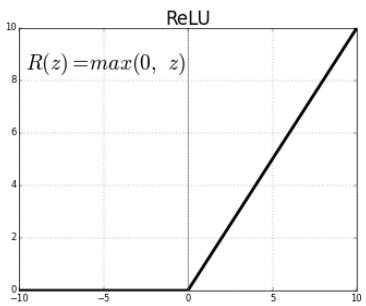
x_1	x_2	x_3	$w_1x_1 + w_2x_2 + w_3x_3$	$w_1x_1 + w_2x_2 + w_3x_3 + b$	output
0	1	1	1	-3	0.047
1	0	0	5	1	0.731
0.5	0.6	0.9			
-5	2	4			
0.6	0.3	0.1			

ReLU Neurons

Some neural networks use Rectified Linear Unit (ReLU) functions instead of sigmoidal functions.

- A typical ReLU function is

$$f(z) = \max(0, z)$$



- ReLU functions have the advantage of being easier to compute than sigmoidal functions, but the disadvantage that they are not differentiable everywhere. .

Building a Neural Network

To build a neural network to perform a task we need to:

- Decide on the network "architecture", that is ...

- Find weights and biases that work.
  - This is done using a training data set, where you have the inputs labeled with the desired outputs, e.g. lots of pictures with and without sharks and the answers for whether they contain sharks or not.
  - We have to define a cost function (error function) that is small when the weights and biases give us good output, and large when the weights and biases give us lousy output
  - We look for weights and biases that minimize cost using ...

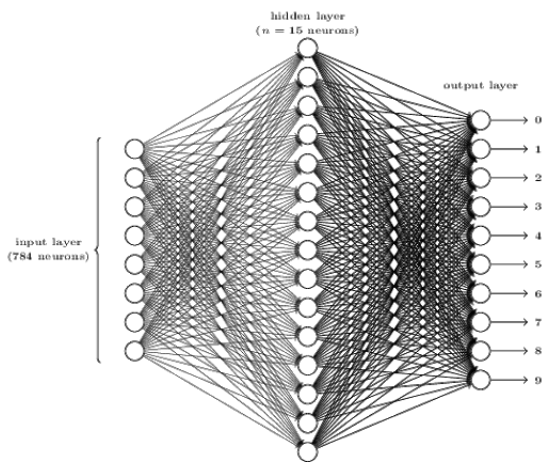
Handwritten Digit Recognition

There is a database called MNIST that contains 60,000 images of handwritten digits, all on 28 x 28 pixel grids, plus a separate 10,000 digits for testing purposes. Suppose we want to build a neural network to recognize handwritten digits, and train it on the 60,000 examples.
Question. How many input neurons should our network have?

Question. How many output neurons would it be reasonable for our network to have? (Hint: would one output neuron be enough to distinguish the digits 0, 1, 2, ... 9?)

Question. Note that our final output will be numbers between 0 and 1 (assuming we are using sigmoidal neurons), and not necessarily numbers equal to 0 or 1. There will be one number per output neuron. How do we turn that into an answer, like "This is the digit 4!".

- In addition to deciding on the number of input neurons and output neurons, we need to decide how many "hidden layers" of neurons to use, and how many neurons to put in each layer.
- We will go with one hidden layer of 15 neurons, following the example in the 3Blue1Brown vidso. This is a fairly arbitrary choice.



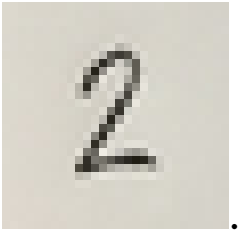
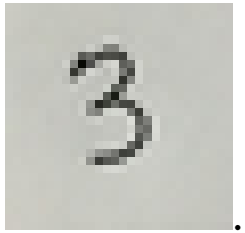
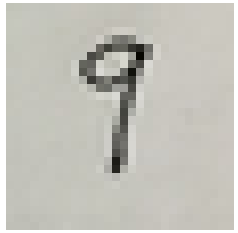
Example. How many weights and biases will there be in our neural network?

Evaluating the Error, or Cost

Question. Suppose I set all my weights and biases at random or according to some simple rule, and you set the all a different way. How can we decide whose neural network does a better job of recognizing handwritten digits?

Hint: For a fixed set of weights and biases, we need a ”cost function” (or ”error function” or ”objective function”) that will give a score for how bad our neural network is doing on our training data set. A high cost will mean the neural net sucks, and a low cost will mean it is pretty good.

Example. Suppose my training set contains only the following three images, and my neural net gives the following output:

and gives the following outputs:

0	0.37
1	0.52
2	0.71
3	0.21
4	0.17
5	0.55
6	0.13
7	0.68
8	0.22
9	0.11

0	0.44
1	0.17
2	0.53
3	0.89
4	0.17
5	0.46
6	0.09
7	0.24
8	0.82
9	0.31

0	0.25
1	0.47
2	0.16
3	0.27
4	0.22
5	0.13
6	0.08
7	0.56
8	0.33
9	0.75

Evaluate the cost function on this output.

Example. If your neural network gives the following outputs, instead, whose network is doing a better job of recognizing digits?

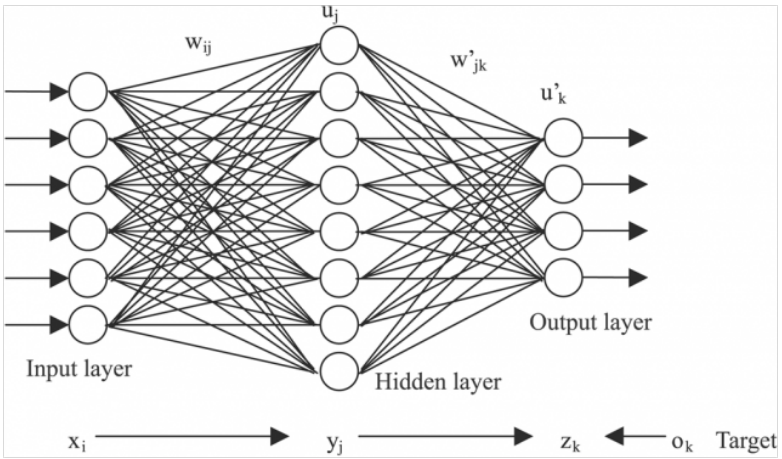
0	0.27
1	0.62
2	0.81
3	0.31
4	0.27
5	0.65
6	0.23
7	0.58
8	0.12
9	0.31

0	0.54
1	0.47
2	0.63
3	0.79
4	0.27
5	0.56
6	0.19
7	0.34
8	0.62
9	0.41

0	0.35
1	0.37
2	0.26
3	0.17
4	0.12
5	0.23
6	0.18
7	0.46
8	0.23
9	0.65

Writing down the cost function

The cost function is supposed to be a function of the weights and biases. Even writing it down can be cumbersome!



Suppose we call the activations in the neurons in the input level $a_0^{(0)}, a_1^{(0)}, \ldots a_{n_0}^{(0)}$ and the activations in the first hidden layer $a_0^{(1)}, a_1^{(1)}, a_2^{(1)}, \ldots a_{n_1}^{(1)}$.

Suppose we refer to the weight for the connection between $a_i^{(0)}$ and $a_j^{(1)}$ as $w_{ji}^{(0)}$, and the bias for $a_k^{(1)}$ as $b_k^{(1)}$.

Then we can relate, say, $a_0^{(1)}$ to the input neuron activations, weights, and biases by the dot product:

and we can relate all activations on hidden layer 1 to the activations for the input layer by the matrix product and sum:

This yeilds the ultra compact notation:

$\vec{a}^{(1)} = W^{(0)}\vec{a}^{(0)} + \vec{b}^{(1)}$