Introduction

Tuesday, February 8, 2022 11:26 PM

Hello everyone! My name is Gregory Elias. I was a student in MATH 1554 for the fall 2021 semester.

I realized early on that MATH 1554 was really fast paced and it was hard to keep track of class notes. That's why I painstakingly wrote all of my professor's lecture notes and all studio worksheet questions & answers in a OneNote/LaTeX format and compiled them into a PDF so that you - the reader - never miss material from any lecture or studio.

Enjoy!

If you have any questions, email me at: elias.gregory.w@gmail.com or gelias7@gatech.edu

P.S. Great power comes great responsibility. *Don't use this as an alternative of attending lecture, it's really important that you do so, especially in this class!*

Unit 1

Saturday, November 13, 2021 8:49 PM

<u>Material Covered:</u>

Chapter 1: Linear Equations in Linear Algebra

- Section 1.1 : Systems of Linear Equations
- Section 1.2 : Row Reduction and Echelon Forms
- Section 1.3 : Vector Equations
- Section 1.4 : The Matrix Equation
- Section 1.5 : Solution Sets of Linear Systems
- Section 1.7 : Linear Independence
- Section 1.8 : An Introduction to Linear Transforms
- Section 1.9 : Linear Transforms

Chapter 2: Matrix Algebra

- Section 2.1 : Matrix Operations

Lecture 1

Monday, August 23, 2021 3:49 PM

General Information:

- Professor: Victor Vilaça Da Rocha
  - School Email: vrocha3@gatech.edu
  - Personal Email: v.vilaca.da.rocha@gmail.com
- Use MyLab for assignments and textbook

Notes:

Section 1.1: Systems of Linear Equations

Linear Equations:

- A linear equation has the form:
  - $A_1X_1 + A_2X_2 + \dots + A_nX_n = b$
    - $A_1 \dots A_n$ are coefficients
    - $X_1 \dots X_n$ are variables
    - "n" is the dimension
- Examples:
  - In 2D: $6x_1 + 4x_2 = 5$
    - A line in 2D
  - In 3D $9x_1 + 7x_2 + 2x_3 = 8$
    - A plane in 3D
  - *Non linear equations:
    - $X_1^2 + X_2^7 = 4$
    - $\ln(x) + 1/x + x^{143} = e^x$
    - $4x_1x_2 = 3$

Systems of Linear Equations:

- A system of linear equations have more than one equation. For example:
  - $X_1 + 1.5x_2 + \pi x_3 = 4$
  - $5X_1 + \quad 7x_3 = 5$
- A system can have a unique solution, no solution, or an infinite number of solutions.
  - Two lines that intersect have one solution
  - Two parallel lines that have different heights have no solution
  - Two parallel lines that have the same height have infinitely many solutions.
- A equation $A_1X_1 + A_2X_2 + A_3X_3 = b$ defines a plane in $\mathbb{R}^3$.
  - The solution to a system of 3 equations is the set of intersections of the planes.

Row Reduction by Elementary Row Operations:

1. (Replacement/Addition) Add a multiple of one row to another
2. (Interchange) Interchange two rows
3. (Scaling) Multiply a row by a non-zero scalar

Example:

$X_1 - 2X_2 + X_3 = 0$	R_1
$2X_2 - 8X_3 = 8$	R_2
$5X_1 - 5X_3 = 10$	R_3

$X_1 - 7X_3 = 8$	$R_3 \leftarrow R_2 + R_1$
$2X_2 - 8X_3 = 8$	~
$X_1 - X_3 = 2$	$R_3 \leftarrow {}^1/_5 R_3$

$X_1 - 7X_3 = 8$	$R_3 \leftarrow R_3 - R_1$
$X_2 - 4X_3 = 4$	~
$6X_3 = -6$	$R_2 \leftarrow {}^1/_2 R_2$

→ $X_3 = -1$

→ by substitution: $X_2 = 0$, $X_1 = 1$

Studio 1

Tuesday, August 24, 2021 12:34 PM

TA: Jad Salem

- Email: Jsalem7@gatech.edu
- Office Hours: ?

Def. A system of equations is consistent if there is at least 1 solution

No solutions:

X + Y = 1	Inconsistent
X + Y = 2	Sol: UND

One Solution:

2X + 4Y = 2	Consistent
3X + 4Y = 2	Sol: x = 0, y = 0.5

Infinite Solutions:

X + Y = 1	Consistent
2X + 2Y = 2	Sol: y = 1 - x

Lecture 2

Wednesday, August 25, 2021 3:33 PM

Notes:

Augmented Matrices

It is redundant to write x1, x2, x3 again and again, so we rewrite systems using matrices. For example,

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

can be written as the augmented matrix,

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

The vertical line reminds us that the first three columns are the coefficients to our variables x_1, x_2, and x_3.

Consistent Systems and Row Equivalence

Definition (Consistent)

- A linear system is consistent if it has at least one **solution**.

Definition (Row Equivalence)

- Two matrices are row equivalent if a sequence of **elementary row operations** transforms one matrix into the other.

Note: if the augmented matrices of two linear systems are row equivalent, then they have the same solution set.

Fundamental Questions

Two questions that we will revisit many times throughout our course.

1. Does a given linear system have a solution? (Is consistent)
2. If it is consistent, is the solution unique?
3. How do you find the solutions?

Section 1.2: Row Reduction and Echelon Forms

A rectangular matrix is in echelon form if

1. All zero rows (if any are present) are at the bottom.
2. The first non-zero entry (or leading entry) of a row is to the right of any leading entries in the row above it (if any).
3. All elements below a leading entry (if any) are zero.

A matrix in echelon form is in row reduced echelon form (RREF) if

1. All leading entries, if any, are equal to 1.
2. Leading entries are the only nonzero entry in their respective column.

A = $\begin{pmatrix} 2 & 0 \\ 3 & 0 \end{pmatrix}$ is not in Echelon Form.

B = $\begin{pmatrix} 0 & \pi \\ 0 & 0 \end{pmatrix}$ is in Echelon Form.

C = $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ is not in Echelon Form.

D = $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is in (Row Reduced) Echelon Form

Example of a Matrix in Echelon Form

$$\begin{pmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}$$

To be in RREF:

1. $\blacksquare = 1$
2. Everything above $\blacksquare$ must be 0.

Example of RREF:

$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$	EF
$B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	RREF
$C = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$	(NO)
$D = \begin{bmatrix} 0 & 6 & 3 & 0 \end{bmatrix}$	EF
$E = \begin{bmatrix} 1 & 17 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	RREF

Definition: Pivot Position, Pivot Column

- A **pivot position** in a matrix A is a location in A that corresponds to a leading 1 in the reduced echelon form of A.
- A **pivot column** is a column of A that contains a pivot position.

Example 2: Express the matrix in row reduced echelon form and identify the pivot columns.

$\left[\begin{array}{ccc|c} 0 & -3 & -6 & 4 \\ -1 & -2 & -1 & 3 \\ -2 & -3 & 0 & 3 \end{array}\right]$	R_1, R_2, R_3
$\left[\begin{array}{ccc|c} -2 & -3 & 0 & 3 \\ -1 & -2 & -1 & 3 \\ 0 & -3 & -6 & 4 \end{array}\right]$	$R_1 \leftrightarrow R_3$
$\left[\begin{array}{ccc|c} -2 & -3 & 0 & 3 \\ 0 & -1 & -2 & 3 \\ 0 & -3 & -6 & 4 \end{array}\right]$	$R_2 \leftarrow 2R_2 - R_1$
$\left[\begin{array}{ccc|c} -2 & -3 & 0 & 3 \\ 0 & -1 & -2 & 3 \\ 0 & 0 & 0 & -5 \end{array}\right]$	$R_3 \leftarrow R_3 - 3R_1$

Row Reduction Algorithm

- The algorithm we used in the previous example produces a matrix in RREF.

Step 1a:	Swap the 1st row with a lower one so the leftmost nonzero entry is in the 1st row
Step 1b:	Scale the 1st row so that its leading entry is equal to 1
Step 1c:	Use row replacement so all entries below this 1 are 0
Step 2a:	Swap the 2nd row with a lower one so that the leftmost nonzero entry below 1st row is in the 2nd row
etc. . . .	Now the matrix is in echelon form, with leading entries equal to 1.
Last step:	Use row replacement so all entries above each leading entry are 0, starting from the right.

Studio 2

Thursday, August 26, 2021 12:30 PM

<u>General Information:</u>
Office hours: 4-5 Tuesday (MathLab)

<u>Notes:</u>

x + 3y = 2
3x + 2y = 1

$\begin{pmatrix} 1 & 3 & 2 \\ 3 & 2 & 1 \end{pmatrix}$ = Augmented Matrix

$\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$ = Coefficient Matrix

Pivot: The first non-zero entry in a row.

Worksheet 1.2

1a.) What are some of the differences between echelon form and row reduced echelon form ((RREF)? List at least three.
1. All pivots = 1 in RREF
2. Pivots are the only nonzero entry in a column in RREF
3. RREF is unique

1b.) How can we use row reduced to determine whether an augmented matrix corresponds to a consistent system?
1. Reduce to RREF & if there is a pivot in the rightmost column (in an augment matrix), it is inconsistent.

2.) Which matrices are in RREF? In echelon form?

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

B = $\begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ {EF}

C = $\begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 5 & 5 \end{pmatrix}$ {EF}

3.) List all 3×2 matricies in RREF. Use * for entries that can be arbitrary.

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & * \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$

4.) Indicate whether the statements are true or false

a.) A linear system, whose 3×5 coefficient matrix has three pivotal columns, must be consistent

True, every row has a pivot

b.) The echelon form of a coefficient matrix is unique

False, only row reduced echelon form is unique; there are infinitely many echelon forms of a singular matrix.

5.) For any three distinct points in the plane, no two on a vertical line, there is a second degree polynomial $p(t) = a_0 + a_1t + a_2t^2$ that passes through (1,12), (2,15), and (3,17). That is, solve

$p(1) = 12 = a_0 + a_1 + a_2$
$p(2) = 15 = a_0 + 2a_1 + 4a_2$
$p(3) = 16 = a_0 + 3a_1 + 9a_2$

$\left[\begin{array}{ccc\|c} 1 & 1 & 1 & 12 \\ 1 & 2 & 4 & 15 \\ 1 & 3 & 9 & 16 \end{array}\right]$	R_1, R_2, R_3
$\left[\begin{array}{ccc\|c} 1 & 1 & 1 & 12 \\ 0 & 1 & 3 & 3 \\ 0 & 2 & 8 & 4 \end{array}\right]$	$R_2 - R_1$ $R_3 - R_1$
$\left[\begin{array}{ccc\|c} 1 & 0 & -2 & 9 \\ 0 & 1 & 3 & 3 \\ 0 & 0 & 2 & -2 \end{array}\right]$	$-(1/2R_3)$ $-R_3 - R_2$ $2R_3 + R_1$
$\left[\begin{array}{ccc\|c} 1 & 0 & 0 & 7 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & -1 \end{array}\right]$	$R_1 - R_2$ $R_3 - 2R_2$
$\left[\begin{array}{ccc\|c} 1 & 0 & 0 & 7 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & -1 \end{array}\right]$	$A_0 = 7$ $A_1 = 6$ $A_2 = -1$

Lecture 3

Friday, August 27, 2021 3:26 PM

General Information:

- Label pages in homework/exploration

Notes:

Basic and Free Variables

Consider the augmented matrix

$$[A \mid \vec{b}] = \left[\begin{array}{ccccc|c} 1 & 3 & 0 & 7 & 0 & 4 \\ 0 & 0 & 1 & 4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 6 \end{array}\right]$$

The leading one's are in first, third, and fifth columns. So:

- The pivot variables of the system $A\vec{x} = \vec{b}$ are x_1, x_3, and x_5.
- The free variables are x_2 and x_4. **Any choice** of the free variables leads to a solution of the system.

Note that A does not have basic or free variables. Systems have variables.

Existence and Uniqueness

A linear system is consistent if and only if (exactly when) the last column of the augmented matrix does not have a pivot. In other words, the RREF of the augmented matrix does not have the form:

$$(0 \quad 0 \quad 0 \; \dots \; 0 \mid 1)$$

Moreover, if a linear system is consistent, then it has

1. a unique solution if and only if there are no free variables.
2. infinitely many solutions that are parameterized by free variables.

Section 1.3: Vector Equations

Think about the algebra in linear algebra.

- To do this, we need to introduce n-dimensional space $\mathbb{R}^n$, and vectors inside it.

$\mathbb{R}^1$ is a number line.

$\mathbb{R}^2$ is a plane.

Example:

$p = (3, 2)$

$\vec{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$

Vectors

- Also think about $\mathbb{R}^n$ as vectors, with given length and direction

Vector Algebra

- Ex.

$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$

1. Scalar Multiple:
   $\overrightarrow{cu} = \begin{pmatrix} cu_1 \\ cu_2 \end{pmatrix}$
2. Vector Multiple:
   $\vec{u} + \vec{v} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}$
3. Pentagon Rule
   Two vectors added together will be the length of the line that connects the beginning of the first vector and the end of the second vector.

Linear Combinations and Span

1. Given vectors $\overrightarrow{v_1}, \overrightarrow{v_2}, \dots \overrightarrow{v_n} \in \mathbb{R}^n$, and scalars $c_1, c_2, \dots c_n$ the vector below

   $$\vec{y} = c_1\overrightarrow{v_1} + c_2\overrightarrow{v_2} + c_n\overrightarrow{v_n}$$

   is called a linear combination of $\overrightarrow{v_1}, \overrightarrow{v_2}, \dots \overrightarrow{v_n}$ with weights $c_1, c_2, \dots c_n$.

2. The set of all linear combinations of $\overrightarrow{v_1}, \overrightarrow{v_2}, \dots \overrightarrow{v_n}$ is called the **Span** of $c_1, c_2, \dots c_n$.

Geometric Interpretation of Linear Combinations

- Ex.

Is $\vec{y}$ the span of vectors $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$?

$\overrightarrow{v_1} = \begin{pmatrix} 1 \\ -2 \\ -3 \end{pmatrix}, \quad \overrightarrow{v_2} = \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix}, \quad$ and $\vec{y} = \begin{pmatrix} 7 \\ 4 \\ 15 \end{pmatrix}$.

Solution:
$\vec{y}$ is in the span of vectors $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$ if there exists two constants c_1, c_2 such that

$\vec{y} = c_1\overrightarrow{v_1} + c_2\overrightarrow{v_2}$

$$\begin{cases} c_1 + 2c_2 = 7 \\ -2c_1 + 5c_2 = 4 \\ -3c_1 + 6c_2 = 15 \end{cases}$$

*Make sure to write "R_1, R_2, R_3" because the grader will know where you might have went wrong.

$\left(\begin{array}{cc\|c} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -3 & 6 & 15 \end{array}\right)$	$R_2 \leftarrow R_2 + 2R_1$ ~ $R_3 \leftarrow R_3 + 3R_1$	$\left(\begin{array}{cc\|c} 1 & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 12 & 36 \end{array}\right)$	Since c = 2 for R_2 and c = 3 for R_3, the system is inconsistent.

$\therefore$ $\vec{y}$ is not in the span of vectors $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$
$\vec{y} \notin \langle\overrightarrow{v_1}, \overrightarrow{v_2}\rangle$

The Span of two Vectors in $\mathbb{R}^n$
In general: Any two non-parallel vectors in $\mathbb{R}^3$ span a plane that passesthrough the origin. Any vector in that plane is also in the span of the twovectors.

Vectors
- In general: Any two non-parallel vectors in $\mathbb{R}^3$ span a plane that passesthrough the origin. Any vector in that plane is also in the span of the two vectors.

Lecture 4

Monday, August 30, 2021 4:25 PM

General Information:

- Quiz on Thursday (mainly on 1.1, 1.2, 1.3)
  - 11am - 7pm (15min)

Notes:

Section 1.4: The Matrix Equation

Notation

Symbol	Meaning
$\in$	Belongs to
$\mathbb{R}^n$	The set of vector with n real-valued elements
$\mathbb{R}^{m\times n}$	The set of real-valued matricies with m rows and n columns

Ex.1 $\begin{pmatrix}\mathbf{1}\\\mathbf{2}\\\mathbf{3}\end{pmatrix} \in \mathbb{R}^3$

Ex.2 $\begin{pmatrix}\mathbf{1} & \mathbf{3} & \mathbf{1} & \mathbf{2}\\\mathbf{2} & \mathbf{4} & \mathbf{5} & \mathbf{6}\end{pmatrix} \in \mathbb{R}^{2\times 4}$

Ex.3 $\mathbb{R}^{3\times 1} = \mathbb{R}^3$

Linear Combinations

A is a m×n matrix with columns $\vec{a}_1, \ldots \vec{a}_n$ and $x \in \mathbb{R}^n$, then the matrix vector product $A\vec{x}$ is a linear equation of the columbs of A.

$$\begin{bmatrix} | & | & \cdots & | \\ \vec{a}_1 & \vec{a}_n & \cdots & \vec{a}_n \\ | & | & \cdots & | \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\overrightarrow{a_1} + x_2\overrightarrow{a_2} + \cdots + x_n\overrightarrow{a_n}$$

- Note that $A\vec{x}$ is in the span of the columns of A.

Ex.

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & -3 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix} = 4\begin{pmatrix}1\\0\end{pmatrix} + 3\begin{pmatrix}0\\-3\end{pmatrix} + 7\begin{pmatrix}-1\\3\end{pmatrix} = \begin{pmatrix}-3\\12\end{pmatrix}$$

Solution Sets

If A is a m×n , matrix with columns $\vec{a}_1, \ldots \vec{a}_n$ and $x \in \mathbb{R}^n$, then the solutions to

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

has the same set of solutions as the vector equation

$x_1\overrightarrow{a_1} + \cdots + x_n\overrightarrow{a_n} = \vec{b}$

which has the same set of solutions as the set of linear equations with the augmented matrix

$[\overrightarrow{a_1} \quad \overrightarrow{a_2} \quad \cdots \quad \overrightarrow{a_n} \quad \vec{b}]$

Existence of Solutions

The equation $A\vec{x} = \vec{b}$ has a solution if and only if $\vec{b}$ is a linear combination of the colums of A.

$A\vec{x} = x_1\overrightarrow{a_1} + \cdots + x_n\overrightarrow{a_n}$

Ex.

For what vectors $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ does the equation have a solution?

$\left(\begin{array}{ccc\|c} 1 & 3 & 4 & b_1 \\ 2 & 8 & 4 & b_2 \\ 0 & -1 & 2 & b_3 \end{array}\right)$	$R_2 \leftarrow R_2 - 2R_1$ ~	$\left(\begin{array}{ccc\|c} 1 & 3 & 4 & b_1 \\ 0 & 2 & -4 & b_2 - 2b_1 \\ 0 & -1 & 2 & b_3 \end{array}\right)$	$R_3 \leftarrow R_3 - R_2$	$\left(\begin{array}{ccc\|c} 1 & 3 & 4 & b_1 \\ 0 & 2 & -4 & b_2 - 2b_1 \\ 0 & 0 & 0 & 2b_3 - b_2 + 2b_1 \end{array}\right)$

The system is consistent if and only if

$2b_3 - b_2 + 2b_1 = 0$	$(2 \quad -1 \quad 2 \mid 0)$
or	$b_1 = {}^1\!/_2\, b_2 - b_3$

The system is consistent iff $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} {}^1\!/_2\, b_2 - b_3 \\ b_2 \\ b_3 \end{pmatrix} = b_2\begin{pmatrix} {}^1\!/_2 \\ 1 \\ 0 \end{pmatrix} + b_3\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$

The Raw Vector Rule for Computing $A\vec{x}$

$$\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = [\overrightarrow{R_1} \cdot \vec{x} \quad \overrightarrow{R_2} \cdot \vec{x}\,]$$

$$= x_1\begin{pmatrix}1\\0\end{pmatrix} + x_2\begin{pmatrix}0\\1\end{pmatrix} + x_3\begin{pmatrix}2\\0\end{pmatrix} + x_4\begin{pmatrix}0\\2\end{pmatrix}$$

$$= \begin{pmatrix} x_1 & & 2x_3 & \\ & x_2 & & 2x_4 \end{pmatrix}$$

Summary

We now have four equivalent ways of expressing linear systems

1.1 A system of equations:

$$2x_1 + 3x_2 = 7$$
$$x_1 - x_2 = 5$$

1.2 An augmented matrix:

$$\left[\begin{array}{cc|c} 2 & 3 & 7 \\ 1 & -1 & 5 \end{array}\right]$$

1.3 A vector equation:

$$x_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$$

1.4 As a matrix equation:

$$\begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$$

Each representation gives us a different way to think about linear systems

Studio 3

Tuesday, August 31, 2021 12:30 PM

<u>General Information:</u>
TA Office Hours: 4-5pm today
- Link on canvas

Definition:
$\vec{u}$ is a linear combination of
$\vec{a}_1, \ldots \vec{a}_n$ if
$\vec{u} = c_1\overrightarrow{n_1} + \cdots + c_n\overrightarrow{v_n}$
For some $c_1, \cdots, c_n \in \mathbb{R}$

Worksheet 1.3 and 1.4, Vector Equation and The Matrix Equation

1. Written Explanation Exercise
   a. What does the span of a set of vectors represent?
      i. All linear combinations of a set of vectors.
   b. How do we determine whetehr a vector is in the span of a set of vectors?
      i. Determine whether the augmented matrix of $\left[\overrightarrow{v_1} \quad \overrightarrow{v_2} \quad \overrightarrow{v_3} \;\middle|\; \vec{b}\right]$ is consistent.
2. Indicate whether the statements are true or false.
   a. If the equation $\mathrm{A}\vec{x} = \vec{b}$ is consistent, then $\vec{b}$ is not in the set spanned by the columns of A.
      i. True
   b. If the augmented matrix $[A \quad \vec{b}]$ has a pivot position in every row, then the equation $\mathrm{A}\vec{x} = \vec{b}$ must be consistent.
      i. True
   c. There are exactly three vectors in Span$\{\overrightarrow{a_1} \quad \overrightarrow{a_2} \quad \overrightarrow{a_3}\}$
      i. False. There can only be 0 or ∞ vectors in a span.
3. Span$\{\overrightarrow{v_1} \quad \overrightarrow{v_2}\}$ is equal to which of the expressions below?

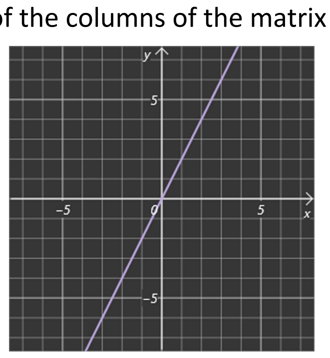
a.	Span$\{\overrightarrow{v_1} \quad \overrightarrow{v_2} \quad 3\overrightarrow{v_1}\}$	Equal
b.	Span$\{\overrightarrow{v_1} \quad 3\overrightarrow{v_1}\}$	Not Equal
c.	Span$\{\overrightarrow{v_1} \quad \overrightarrow{v_2} \quad 3\overrightarrow{v_1} + 2\overrightarrow{v_2}\}$	Equal

4. For what values of h is $\vec{b}$ in the plane spanned by $\overrightarrow{a_1}$ and $\overrightarrow{a_1}$?

a.	$\overrightarrow{a_1} = \begin{pmatrix} 1 \\ 4 \\ -1 \end{pmatrix}$	$\overrightarrow{a_2} = \begin{pmatrix} -6 \\ -17 \\ 2 \end{pmatrix}$	$\vec{b} = \begin{pmatrix} 4 \\ 2 \\ h \end{pmatrix}$

i.	$\left(\begin{array}{cc\|c} 1 & -6 & 4 \\ 4 & -17 & 2 \\ -1 & 2 & h \end{array}\right)$	$R_2 \leftarrow R_2 - 4R_1$ ~ $R_3 \leftarrow R_3 + R_1$	$\left(\begin{array}{cc\|c} 1 & -6 & 4 \\ 0 & 7 & -14 \\ 0 & -4 & h+4 \end{array}\right)$

   ii. Now we know that $x_2 = -2$. Hence, $h = 4$
5. Sketch the span of the columns of the matrix $A = \begin{pmatrix} 2 & -1 & 3 \\ 4 & -2 & 6 \end{pmatrix}$



Lecture 5

Wednesday, September 1, 2021 3:24 PM

General Information:

- Quiz on Thursday

Notes:

Section 1.5: Solution Sets of Linear Systems

Homogenous Systems

Linear systems of the form $A\vec{x} = \vec{0}$ are **homogeneous**

Linear systems of the form $A\vec{x} \neq \vec{0}$ are inhomogeneous

Because homogeneous systems always have the **trivial solution**, $\vec{x} = \vec{0}$, the interesting question is whether they have non-trivial solutions.

Observation

$A\vec{x} = \vec{0}$ has a nontrivial solution

$\Leftrightarrow$ there is a free variable

$\Leftrightarrow$ A has a column with no pivot

Ex.

$x_1 + 3x_2 + x_3 = 0$

$2x_1 - x_2 - 5x_3 = 0$

$x_1 - 2x_3 = 0$

$\left(\begin{array}{ccc\|c} 1 & 3 & 1 & 0 \\ 2 & -1 & -5 & 0 \\ 1 & 0 & -2 & 0 \end{array}\right)$	$R_2 \leftarrow R_2 - 2R_1$ ~	$\left(\begin{array}{ccc\|c} 1 & 3 & 1 & 0 \\ 2 & -7 & -7 & 0 \\ 0 & -3 & -3 & 0 \end{array}\right)$	~	$\left(\begin{array}{ccc\|c} 1 & 3 & 1 & 0 \\ 2 & -7 & -7 & 0 \\ 0 & -3 & -3 & 0 \end{array}\right)$	$R_1 \leftarrow R_1 - 3R_2$ ~ RREF →	$\left(\begin{array}{ccc\|c} 1 & 3 & 1 & 0 \\ 2 & -7 & -7 & 0 \\ 0 & -3 & -3 & 0 \end{array}\right)$

Solution x_3: free, $x_1 = 2x_3$, $x_2 = -x_3$

$\therefore \vec{x}$ is a solution if and only if:

$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_3 \\ -x_3 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \gg \textit{the solution set is a line}$

Parametric Forms, Homogeneous Case

In the example on the previous slide we expressed the solution to a system using a vector equation. This is a **parametric form** of the solution.

In general, suppose the free variables for $A\vec{x} = \vec{0}$ are x_k, ... x_n. Then all solutions to $A\vec{x} = \vec{0}$ can be written as

$\vec{x} = x_k\overrightarrow{v_k} + x_k\overrightarrow{v_{k+1}} + \cdots + x_n\overrightarrow{v_n}$

For some $\vec{v}_1, \ldots \vec{v}_n$. This is a **parametric form** of the solution.

Ex.1 $\vec{x} \in \mathbb{R}^5$

x_1, x_2: pivot variables

X_3, x_4, x_5: free variables

$$\vec{x} = \begin{pmatrix} ax_3 & bx_4 & cx_5 \\ dx_3 & ex_4 & fx_5 \\ x_3 & & \\ & x_4 & \\ & & x_5 \end{pmatrix} = x_3\begin{pmatrix} a \\ d \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4\begin{pmatrix} b \\ e \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_5\begin{pmatrix} c \\ f \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Ex.2 (non-homogeneous system)

$x_1 + 3x_2 + x_3 = 9$

$2x_1 - x_2 - 5x_3 = 11$

$x_1 - 2x_3 = 6$

$\left(\begin{array}{ccc\|c} 1 & 3 & 1 & 5 \\ 2 & -1 & -5 & 11 \\ 1 & 0 & -2 & 6 \end{array}\right)$	$R_2 \leftarrow R_2 - 2R_1$ ~ $R_3 \leftarrow R_3 - R_1$	$\left(\begin{array}{ccc\|c} 1 & 3 & 1 & 5 \\ 0 & -7 & -7 & -7 \\ 0 & -3 & -3 & -3 \end{array}\right)$	~	$\left(\begin{array}{ccc\|c} 1 & 3 & 1 & 5 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$	$R_1 \leftarrow R_1 - 3R_2$ ~ RREF →	$\left(\begin{array}{ccc\|c} 1 & 0 & -2 & 6 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$

Solutions:

$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 + 2x_3 \\ 1 - x_3 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \\ 0 \end{pmatrix} + x_3\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$

We found that:

$\vec{x} = \overrightarrow{x_p} + \overrightarrow{x_0} \gg \textit{solution of } A\vec{x} = \vec{0}$

Take $\vec{x}$ another solution $A\vec{x} = \vec{b}$

$A(\vec{x} - \overrightarrow{x_p}) = A\vec{x} + A\overrightarrow{x_p}$

$= \vec{b} - \vec{b}$

$\Rightarrow (\vec{x} - \overrightarrow{x_p})$ is a solution to $A\vec{x} = \vec{0}$

Studio 4

Thursday, September 2, 2021 12:33 PM

General Information:

Def. The solution set of $A\vec{x} = \vec{b}$ is the set of all $\vec{x}$ such that $A\vec{x} = \vec{b}$

If A is $m \times n$, then solution set $\{\vec{x} \in \mathbb{R}^n | A\vec{x} = \vec{b}\}$

Def. A homogeneous system is a system of equation of form $A\vec{x} = \vec{D}$

A = $\begin{pmatrix} 1 & 0 & 4 \\ 0 & 2 & -10 \end{pmatrix}$

What is the solution set of $A\vec{x} = \vec{0}$?

A = $\begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & -5 \end{pmatrix}$

$X_1 + 4x_3 = 0$

$X_2 - 5x_3 = 0$

$$\begin{cases} x_1 = -4x_3 \\ x_2 = 5x_3 \\ x_3 = x_3 \end{cases}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} -4 \\ 5 \\ 1 \end{pmatrix}$$

$$\Rightarrow Solution\ set\ is\ \left\{ x_3 \begin{pmatrix} -4 \\ 5 \\ 1 \end{pmatrix} \middle| x_3 \in \mathbb{R} \right\}$$

Worksheet 1.5, Solution Sets of Linear Systems

1. Written Explanation Exercise
    a. When a homogeneous system has a nontrivial solution, what properties does that system have? List at least two.
        i. Column with no pivots (at least 1 free variable)
        ii. Includes the zero vector
2. Indicate whether the statements are true or false
    a. A non-trivial solution $\vec{x}$ to $A\vec{x} = \vec{0}$ has all non-zero entries.
        i. False
    b. If $A\vec{x} = \vec{b}$ and $A\vec{y} = \vec{b}$, then $A(\vec{x} - \vec{y}) = \vec{0}$
        i. True
    c. Any 3×2 matrix A with two pivotal positions has a non-trivial solution to $A\vec{x} = \vec{0}$
        i. False
3. Example Construction
    a. Give an example of a non-zero 2×3 matrix A such that $\vec{x} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ is a solution to $A\vec{x} = \vec{0}$
        i. $\begin{pmatrix} 1 & -1 & 1 \\ 2 & -2 & 2 \end{pmatrix}$
    b. Give an example of a non-trivial solution to $A\vec{x} = \vec{0}$, where A = $\begin{pmatrix} 2 & 5 \\ 0 & 0 \\ 4 & 10 \end{pmatrix}$
        i. $\begin{pmatrix} 2.5 \\ -1 \end{pmatrix}$
4. Express the solution to $A\vec{x} = \vec{0}$ in the parametric vector form, where A = $\begin{pmatrix} 1 & 3 & 4 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$
    i.

A = $\left(\begin{array}{cccc\|c} 1 & 3 & 4 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right)$	$R_1 \leftarrow R_1 - 4R_2$	$\left(\begin{array}{cccc\|c} 1 & 3 & 0 & -3 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right)$

    ii. $x_1 + 3x_2 - 3x_4 = 0$
    iii. $x_1 = 3x_2 + 3x_4$
    iv. $x_3 + x_4 = 0$
    v. $x_3 - x_4 = 0$
    vi. $\Rightarrow \vec{x} = \begin{pmatrix} -3x_2 + 3x_4 \\ x_2 \\ -x_4 \\ x_4 \end{pmatrix} = \overrightarrow{x_2} \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \overrightarrow{x_4} \begin{pmatrix} 3 \\ 0 \\ -1 \\ 1 \end{pmatrix}$

Lecture 6

Friday, September 3, 2021 4:59 PM

Notes:

Section 1.7: Linear Independence

Linear Independence

A set of vectors $\{\overrightarrow{v_1}, \dots, \overrightarrow{v_k}\}$ in $\mathbb{R}^n$ are linearly independent if

$$\sum_{i=1}^{k} c_i\overrightarrow{v_i} = c_1\overrightarrow{v_1} + c_2\overrightarrow{v_2} + \dots + c_k\overrightarrow{v_k} = \vec{0} \Rightarrow \left(\overrightarrow{v_1}, \dots, \overrightarrow{v_k} \;\middle|\; \vec{0}\right)$$

has only the trivial solution. It is heavily dependent otherwise.

In other words, $\{\overrightarrow{v_1}, \dots, \overrightarrow{v_k}\}$ are linearly dependent if there are real numbers $c_1, c_2, \dots, c_k$ not all zero so that

$$c_1\overrightarrow{v_1} + c_2\overrightarrow{v_2} + \dots + c_k\overrightarrow{v_k} = \vec{0}$$

Consider the vectors $\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_k}$

To determine whether the vectors are linearly independent, we can se the linear combination the zero vector:

$$c_1\overrightarrow{v_1} + c_2\overrightarrow{v_2} + \dots + c_k\overrightarrow{v_k} = [\overrightarrow{v_1} \quad \overrightarrow{v_2} \quad \cdots \quad \overrightarrow{v_k}]\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix} = V\vec{c} = \vec{0}$$

Linear independence: There is NO non-zero solution $\vec{c}$.
Linear dependence: There is a non-zero solution $\vec{c}$.

Ex.1

For what values of h are the vectors linearly independent?

$$\begin{bmatrix} 1 \\ 1 \\ h \end{bmatrix}, \begin{bmatrix} 1 \\ h \\ 1 \end{bmatrix}, \begin{bmatrix} h \\ 1 \\ 1 \end{bmatrix}$$

$\left(\begin{array}{ccc\|c} 1 & 1 & h & 0 \\ 1 & h & 1 & 0 \\ h & 1 & 1 & 0 \end{array}\right)$	$R_2 \leftarrow R_2 - R_1$ $\sim$ $R_3 \leftarrow R_3 - hR_1$	$\left(\begin{array}{ccc\|c} 1 & 1 & h & 0 \\ 0 & h-1 & 1-h & 0 \\ 0 & 1-h & 1-h^2 & 0 \end{array}\right)$	$R_3 \leftarrow R_3 + R_2$ $\sim$	$\left(\begin{array}{ccc\|c} 1 & 1 & h & 0 \\ 0 & h-1 & 1-h & 0 \\ 0 & 0 & 2-h-h^2 & 0 \end{array}\right)$

The 3 vectors are linearly independent iff $\begin{cases} h-1 \neq 0 & : h \neq 1 \\ h^2+h-2 \neq 0 & : h \neq 1, -2 \end{cases}$

Hence, the 3 vectors are linearly independent iff $h \neq 1, -2$

Check:

$$h = 1 \qquad 2\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$h = -2 \qquad \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Suppose $\vec{v} \in \mathbb{R}^n$. When is the set $\{\vec{v}\}$ linearly dependent for some $c_1 \neq 0$.

$c_1\vec{v} = \vec{0} \Longrightarrow \vec{v}$ must be $\vec{0}$

Suppose $\overrightarrow{v_1}, \overrightarrow{v_2} \in \mathbb{R}^n$. When is the set $\{\overrightarrow{v_1}, \overrightarrow{v_2}\}$ linearly dependent? Provide a geometric interpretation.

$c_1\overrightarrow{v_1} + c_2\overrightarrow{v_2} = \vec{0}$ with $(c_1, c_2) \neq (0, 0)$

1) If $c_1 = 0$ or $c_2 = 0$. Say $c_1 = 0$
   $c_2\overrightarrow{v_2} = \vec{0}$ with $c_2 \neq 0$
   $\overrightarrow{v_2} = 0$
2) If $c_1 \neq 0$, $c_2 \neq 0$
   $\overrightarrow{v_2} = -\frac{c_1}{c_2}\overrightarrow{v_1}$: $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$ are parallel.

Two Theorems

Fact 1. Suppose $\overrightarrow{v_1}, \dots, \overrightarrow{v_k}$ are vectors $\mathbb{R}^n$. If $k > n$ then $\{\overrightarrow{v_1}, \dots, \overrightarrow{v_k}\}$ is linearly dependent.

Ex. $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ \pi \end{pmatrix} \begin{pmatrix} 3 \\ 47 \end{pmatrix}\right\}$ are linearlyly dependent.

$$\left.\left(\underbrace{\overrightarrow{v_1}, \dots, \overrightarrow{v_k}}_{k\ columns} \quad | \; 0\right)\right\} n \text{ rows}$$

Fact 2. If any one or more of $\overrightarrow{v_1}, \dots, \overrightarrow{v_2}$ is $\vec{0}$, then $\{\overrightarrow{v_1}, \dots, \overrightarrow{v_k}\}$ is linearly dependent.

Ex. $\left\{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \\ 42 \end{pmatrix}\right\}$ are linearly dependent. Indeed: $58\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 2 \\ 1 \\ 3 \\ 42 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

1.5 Parametric Vector form

x_1 = pivot; x_2, x_3: free

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} ax_2 & bx_3 \\ x_2 & \\ & x_3 \end{pmatrix} = x_2\begin{pmatrix} a \\ 1 \\ 0 \end{pmatrix} + x_3\begin{pmatrix} b \\ 0 \\ 1 \end{pmatrix}$$

Studio 5

Tuesday, September 7, 2021 12:31 PM

Def. A set $\{\overrightarrow{v_1}, \dots, \overrightarrow{v_k}\}$ is linearly independent if

$$\begin{pmatrix} c_1\overrightarrow{v_1} + \dots + c_k\overrightarrow{v_k} = \vec{0} \\ \Downarrow \\ c_1 = c_2 = \dots = c_k = 0 \end{pmatrix}$$

(i.e. $\{\overrightarrow{v_1}, \dots, \overrightarrow{v_k}\}$ indepdent if $\left(\overrightarrow{v_1}, \dots, \overrightarrow{v_k} \;\middle|\; \vec{0}\right)$ only has trivial solution.)

<u>Dependent Set:</u>

$$\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \quad \begin{bmatrix}2\\2\\2\end{bmatrix}\right\}$$

$$\Rightarrow -2\begin{bmatrix}1\\1\\1\end{bmatrix} + 1\begin{bmatrix}2\\2\\2\end{bmatrix} = \vec{0}$$

$\{\vec{0}\}$

$$\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \quad \begin{bmatrix}1\\1\\0\end{bmatrix}, \quad \begin{bmatrix}2\\1\\1\end{bmatrix}\right\}$$

$$\Rightarrow -\begin{bmatrix}1\\0\\1\end{bmatrix} - \begin{bmatrix}1\\1\\0\end{bmatrix} + \begin{bmatrix}2\\1\\1\end{bmatrix} = \vec{0}$$

<u>Independent Set:</u>

$\left\{\begin{bmatrix}1\\0\end{bmatrix}\right\}$

$\left\{\begin{bmatrix}1\\0\end{bmatrix}, \quad \begin{bmatrix}0\\1\end{bmatrix}\right\}$

Worksheet 1.7, Linear Independence

1. Written Explanation Exercise
   a. How are span and linear dependence related to each other?
      If $\overrightarrow{v_1} \in$ Span $\{\overrightarrow{v_2}, \dots, \overrightarrow{v_k}\}$, then $\{\overrightarrow{v_1}, \dots, \overrightarrow{v_k}\}$ is dependent
   b. Suppose T is a linear map
      i. If $\overrightarrow{v_1}, \dots, \overrightarrow{v_k}$ are dependent, why are $T(v_1), \dots, T(v_k)$ depdendent?
         Exist $c_1, \dots, c_k$ such that
         $c_1\overrightarrow{v_1} + \dots + c_k\overrightarrow{v_k} = \vec{0}$
         Then $\vec{0} = T(\vec{0})$
         $= T(c_1\overrightarrow{v_1} + \dots + c_k\overrightarrow{v_k})$
         $= c_1T(\overrightarrow{v_1}) + \dots + c_kT(\overrightarrow{v_k})$
         (Think of $T(\vec{x})$ as $A\vec{x}$)

      ii. If $\overrightarrow{v_1}, \dots, \overrightarrow{v_k}$ are independent, need $T(v_1), \dots, T(v_k)$ be independent?
         Take $A = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}, \quad v_1 = \begin{bmatrix}1\\0\end{bmatrix}, \quad v_2 = \begin{bmatrix}0\\1\end{bmatrix}$
         $\{A\overrightarrow{v_1}, \quad A\overrightarrow{v_2}\}$ dependent

2. In the problems below, $\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}$ are three linearly independent vectors in $\mathbb{R}^3$. Which of the collections of vectors below are linearly independent?
   a. $(\overrightarrow{v_1}, \overrightarrow{v_2}, \vec{0})$
      Dependent
   b. $(\overrightarrow{v_1}, \overrightarrow{v_1} + \overrightarrow{v_2}, \overrightarrow{v_2})$
      Dependent
   c. $(\overrightarrow{v_1}, \overrightarrow{v_1} + \overrightarrow{v_2})$
      Independent
3. For what values of h are the colums of A linearly dependent?
   $$A = \begin{bmatrix}2 & 4 & -2\\-2 & -6 & 2\\4 & 7 & h\end{bmatrix}$$

$\left[\begin{array}{ccc\|c}2 & 4 & -2 & 0\\-2 & -6 & 2 & 0\\4 & 7 & h & 0\end{array}\right]$	$R_1 \leftarrow \frac{1}{2}R_1$ ~ $R_2 \leftarrow \frac{1}{2}R_2$	$\left[\begin{array}{ccc\|c}1 & 2 & -1 & 0\\-1 & -3 & 1 & 0\\4 & 7 & h & 0\end{array}\right]$	$R_2 \leftarrow R_2 - R_1$ $R_3 \leftarrow R_3 - 4R_1$ $R_3 \leftarrow R_3 + R_2$	$\left[\begin{array}{ccc\|c}1 & 2 & -1 & 0\\0 & 1 & 0 & 0\\0 & 0 & h-4 & 0\end{array}\right]$

   Hence, $h = 4$.

4. A 5×3 matrix $A = [\overrightarrow{a_1} \quad \overrightarrow{a_2} \quad \overrightarrow{a_3}]$ has all non-zero columns, and $\overrightarrow{a_3} = 5\overrightarrow{a_1} + 7\overrightarrow{a_2}$. Identify a non-trivial solution to $A\vec{x} = \vec{0}$.
   $A = [\overrightarrow{a_1} \quad \overrightarrow{a_2} \quad \overrightarrow{a_3}]$
   $\overrightarrow{a_3} = 5\overrightarrow{a_1} + 7\overrightarrow{a_2}$
   $[\overrightarrow{a_1}], \; [\overrightarrow{a_2}], \; [\overrightarrow{a_3}]$
   $-5[\overrightarrow{a_1}] - 7[\overrightarrow{a_2}] + 1(5[\overrightarrow{a_1}] + 7[\overrightarrow{a_2}])$
   $\vec{x} = \begin{bmatrix}-5\\-7\\1\end{bmatrix}$
5. Fill in the blanks.
   a. The columns of a 7×3 matrix are linearly independent. How many pivots does the matrix have?
      5 pivots
   b. If the columns of a 3×7 matrix span $\mathbb{R}^3$, how many pivots does the matrix have?
      3 pivots

Lecture 7

Wednesday, September 8, 2021 3:27 PM

Notes:

Section 1.8: Introduction to Linear Transformations

From Matrices to Functions
Let A be a $m \times n$ matrix. We define a function

T_A: $\mathbb{R}^n \to \mathbb{R}^m$

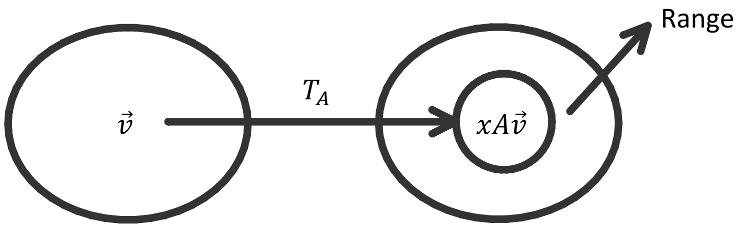
$T(\vec{v}) = A\vec{v}$	$A \in \mathbb{R}^{m \times n}$

This is called a **matrix transformation**.

- The **domain** of T is $\mathbb{R}^n$.
- The **co-domain** of T or target T is $\mathbb{R}^m$.
- The vector $T(\vec{x})$ is the **image** $\vec{x}$ under T
- The set of all possible images $T(\vec{x})$ is the **range**.

This gives us **another** interpretation of $A\vec{x} = \vec{b}$.

- Set of equations
- Augmented matrix
- Matrix equation
- Vector equation
- Linear transformation equation



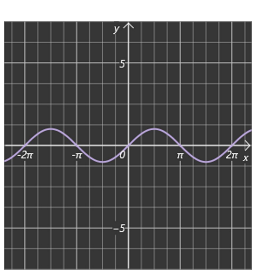
Domain: $\mathbb{R}^n$ Target: $\mathbb{R}^m$
Co-Domain

"Bad example"
F(x) = x^2
Domain: $\mathbb{R}$
Target: $\mathbb{R}$
Range: [0, +∞)

Functions from Calculus
Many of the functions we know have domain and codomain $\mathbb{R}$. We can express the rule that defines the functions in this way:

f: $\mathbb{R} \to \mathbb{R}$ $f(x) = \sin(x)$
Domain: $\mathbb{R}$
Co-domain: $\mathbb{R}$
Range: $[-1,\ 1]$

In calculus we often think of a function in terms of its graph, whosehorizontal axis is the domain, and the vertical axis is the codomain.



This is ok when the domain and codomain are $\mathbb{R}$. It's hard to do when the domain is $\mathbb{R}^2$ and the codomain is $\mathbb{R}^3$. We would need five dimensions to draw that graph.

Ex.1

Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$, $\vec{u} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 7 \\ 5 \\ 7 \end{bmatrix}$ $\quad Domain: \mathbb{R}^2 \quad Codomain: \mathbb{R}^3$

Compute $T(\vec{u}) = A\vec{u} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \\ 7 \end{pmatrix}$

Calculate $\vec{v} \in \mathbb{R}^2$ so that $T(\vec{v}) = \vec{b}$

$T(\vec{v}) = \begin{pmatrix} v_1 + v_2 \\ v_2 \\ v_1 + v_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \\ 7 \end{pmatrix}$

By substitution: $\vec{v} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$

Give a $\vec{c} \in \mathbb{R}^2$ so there is no $\vec{c}$ with $T(\vec{v}) = \vec{c}$.
or: Give a $\vec{c}$ that is not in the range of T.
or: Give a $\vec{c}$ that is not in the span of the columns of A.

$\vec{c} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$, $\vec{c} = \begin{pmatrix} 5 \\ 4 \\ 8 \end{pmatrix}$, $or \cdots$ (any vector $\vec{c}$ with $c_1 \neq c_3$)

Range (a.k.a. span of columns): plane in $\mathbb{R}^3$ of equation $x_1 - x_3$ or x = z or $x_1 = x_3$

Linear transformations

A function T: $\mathbb{R}^n \to \mathbb{R}^m$ is linear if

- $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$ for all $\vec{u}, \vec{v}$ in $\mathbb{R}^n$

- $T(c\vec{v}) = cT(\vec{v})$ for all $\vec{v} \in \mathbb{R}^n$ and c in $\mathbb{R}$.

So if T is linear then

$T(c_1\overrightarrow{v_1} + \cdots + c_k\overrightarrow{v_k}) = c_1T(\overrightarrow{v_1}) + \cdots + c_kT(\overrightarrow{v_k})$

This is called the **principle of superposition**. The idea is that if we know $T(c_1), \dots, T(c_k)$ then we know every $T(\overrightarrow{v_1})$

Fact: Every matrix transformation T_A is linear

Fact: $T(\vec{0}) = T(0\vec{v}) = 0T(\vec{v}) = \vec{0}$

$$\overrightarrow{e_1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \overrightarrow{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \cdots$$

In $\mathbb{R}^2 = \overrightarrow{e_1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \overrightarrow{e_2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

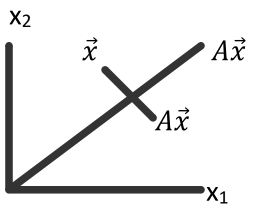
$$T\begin{pmatrix} x \\ y \end{pmatrix} = T\left(x\begin{pmatrix} 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$$
$$= xT\begin{pmatrix} 1 \\ 0 \end{pmatrix} + yT\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Indeed: $\begin{cases} A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v} \\ A(c\vec{u}) = cA\vec{u} \end{cases}$

Ex.2

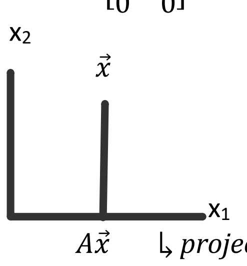
Suppose T is the linear transformation $T(\vec{x}) = A\vec{x}$. Give a short geometric interpreation of what $T(\vec{x})$ does to vectors in $\mathbb{R}^2$.

1.) $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $\quad T(\vec{x}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$



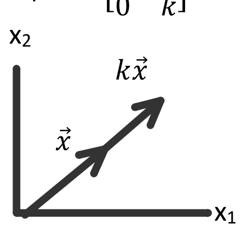
↳ Reflected through the line $x_1 = x_2$

2.) $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ $\quad T(\vec{x}) = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$



↳ *projected onto the y-axis.*

3.) $A = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$ for all $k \in \mathbb{R}$ $\quad T(\vec{x}) = k\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = k\vec{x}$

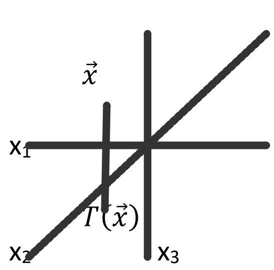


↳ *scaling by k.*

Ex. 3

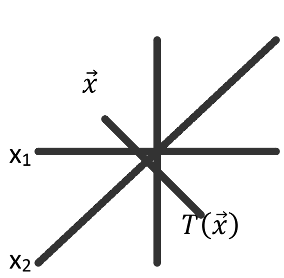
What does T_A do to the vector in $\mathbb{R}^3$?

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad T(\vec{x}) = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}$$



↳ *projecting by the* $x_1 - x_2$ *plane.*

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad T(\vec{x}) = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$



↳ *projecting by the* $x_1 - x_3$ *plane.*

Ex.4

A linear transformation $T\colon \mathbb{R}^2 \longmapsto \mathbb{R}^3$ satisfies

$$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}5\\-7\\2\end{bmatrix},\quad T\left(\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}-3\\8\\0\end{bmatrix}$$

What is the matrix that represents T?

$$A=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{23}\\ a_{31} & a_{32}\end{pmatrix}$$

$$T\begin{pmatrix}1\\0\end{pmatrix}=A\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}a_{11}\\a_{21}\\a_{31}\end{pmatrix}=\begin{pmatrix}5\\-7\\2\end{pmatrix}$$

$$T\begin{pmatrix}0\\1\end{pmatrix}=A\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}a_{12}\\a_{22}\\a_{32}\end{pmatrix}=\begin{pmatrix}-3\\8\\0\end{pmatrix}$$

$$\Longrightarrow A=\begin{pmatrix}5 & -3\\ -7 & 8\\ 2 & 0\end{pmatrix}$$

Studio 6

Thursday, September 9, 2021 12:31 PM

Def. A function $T: A \longmapsto B$ is assignment rule that assings one value $T(a) \in B$ to each $a \in A$.

A: Domain

B: Codomain

↱ This is one-to-one

Ex. $f(x) = x^2$, $f: [3,\ 7] \to \mathbb{R}$

↳ Not onto because (e.g. -1) is not hit.

Def. A function $T: A \longmapsto B$ is onto if for all $b \in B$, there is some element of $a \in A$ such that $T(a) = B$.

Def. A function $T: A \longmapsto B$ is 1-1 if it passes the horizontal line test.

$T(a_1) = T(a_2) \Rightarrow a_1 = a_2$

Worksheet 1.8, An Introduction to Linear Transforms

1. Suppose $T(x) = Ax$ for all x where A is a matrix amd T is onto.
   a. What can we say about pivotal rows of A?
      i. There is a pivot in every row
   b. What can we say about the existence of solutions to $Ax = b$?
      i. $Ax = b$ is consistent

2. Let A be an 3×4 matrix. What must c and d be if we define the linear transformation $T: \mathbb{R}^c \longmapsto \mathbb{R}^d$ by $T(\vec{x}) = A\vec{x}$?

   $c = 4$

   $\mathrm{d} = 3$

3. Let $T: \mathbb{R}^2 \longmapsto \mathbb{R}^2$ be a linear transformation such that

   $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 \begin{bmatrix} -1 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ -1 \end{bmatrix}$

   Construct a matrix A so that $T(\vec{x}) = A\vec{x}$ for all vectors $\vec{x}$.

   $\begin{bmatrix} -1 & 4 \\ 3 & -1 \end{bmatrix}$ because $[\vec{c_1} \quad \cdots \quad \vec{c_n}] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

   $= x_1\vec{c_1} + \cdots x_n\vec{c_n}$

4. Let $T: \mathbb{R}^4 \longmapsto \mathbb{R}^3$ be a linear transformation such that

   $T \begin{bmatrix} 4 \\ 0 \\ 1 \\ 0 \end{bmatrix} = T \begin{bmatrix} 0 \\ 0 \\ 1 \\ 4 \end{bmatrix} \neq \vec{0}$

   Identify a non-trivial solution $\vec{x}$ to $T\vec{x} = \vec{0}$

   $x_1 = \begin{bmatrix} 4 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \qquad x_2 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 4 \end{bmatrix}$

   Know: $Ax_1 = Ax_2$

   Know: $T(x) = Ax$

   So, $Ax_1 - Ax_2 = 0$

   So $T\left(\begin{bmatrix} 4 \\ 0 \\ 0 \\ -4 \end{bmatrix}\right) = 0$

5. Let T_A be the lienar transformation with the matrix below. Match each choice of A on the left with the geometric description of the action of T_A on the right.

   $\begin{bmatrix} .5 & 0 \\ 0 & .5 \end{bmatrix} = dilation\ by\ 1/2$

   $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = projection\ onto\ y - axis$

   $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = rotation\ by\ 90°$

   $\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} = sheer$

Lecture 8

Friday, September 10, 2021 3:25 PM

<u>General Information:</u>

- Midterm 1: Kenaeda Building (8:00-8:50pm)
- MATLAB due tonight
    - Professor is not an expert in MATLAB

<u>Notes:</u>

Section 1.9: Linear Transforms

Definition: The Standard Vectors

The standard vecotrs in $\mathbb{R}^n$ are the vectors $\vec{e_1}, \vec{e_2}, \dots, \vec{e_n}$ where:

$$\vec{e_1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e_n} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

For example in $\mathbb{R}^3$:

$$\vec{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{e_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$
$$= x_1 e_1 + x_2 e_2 + x_3 e_3$$

$$T(\vec{x}) = x_1 T(\vec{e_1}) + x_2 T(\vec{e_2}) + x_3 T(\vec{e_3})$$

A Property of the Standard Vectors

Note: If A is a $m \times n$ matrix with columns $\vec{v_1}, \vec{v_2}, \dots, \vec{v_n}$ then

A$\vec{v_i} = \vec{v_i}$, for $i = 1,2,\dots,n$

So multiplying matrix by $\vec{e_i}$ gives column i of A.

Ex.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \vec{e_2} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix}$$

The Standard Matrix

Theorem:

Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. There there is a unique matrix A such that:

$$T(\vec{x}) = A\vec{x}, \quad \vec{x} \in \mathbb{R}^n.$$

In fact, A is a $m \times n$ matrix, and its j$^{\text{th}}$ column is the vector $T(\vec{e_j})$

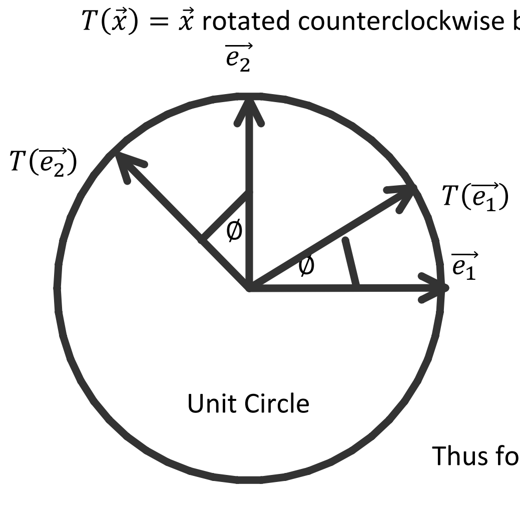
$$A = [\,T(\vec{e_1}) \quad T(\vec{e_2}) \quad \cdots \quad T(\vec{e_n})\,]$$

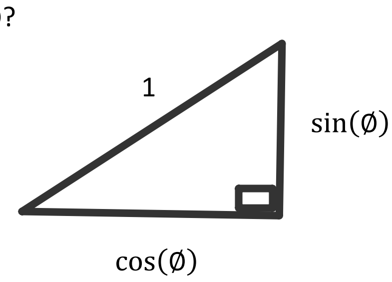
The matrix A is the **standard matrix** for a linear transformation.

Rotations

Ex.1 What is the linear transform $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$T(\vec{x}) = \vec{x}$ rotated counterclockwise by angle Ø?





$$T(\vec{e_1}) = \begin{pmatrix} \cos(Ø) \\ \sin(Ø) \end{pmatrix}, \quad T(\vec{e_2}) = \begin{pmatrix} -\sin(Ø) \\ \cos(Ø) \end{pmatrix}$$

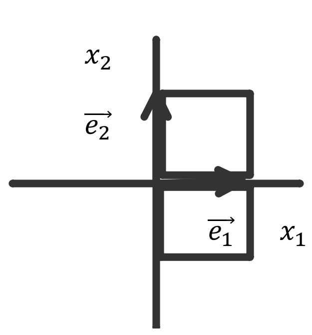
Thus for $A = \begin{pmatrix} \cos(Ø) & -\sin(Ø) \\ \sin(Ø) & \cos(Ø) \end{pmatrix}$: $T(\vec{x}) = A\vec{x}, \ \forall \vec{x} \in \mathbb{R}^2$

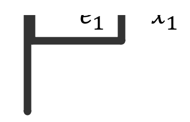
Standard Matrices in $\mathbb{R}^2$

- There is a long list of geometric transformations ofR2in ourtextbook, as well as on the next few slides (reflections, rotations, contractions and expansions, shears, projections, . . .)
- Please familiarize yourself with them: you are expected to memorize them (or be able to derive them)

Two Dimensional Examples: Reflections

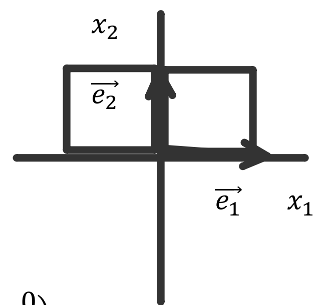
Reflection through x_1 axis





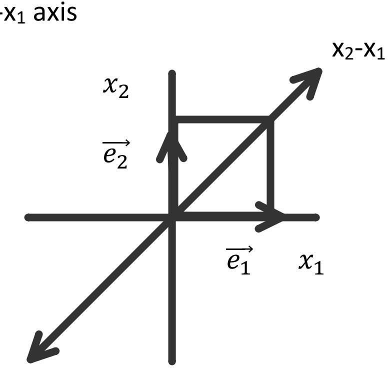
Standard Matrix: $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Reflection through x_2 axis



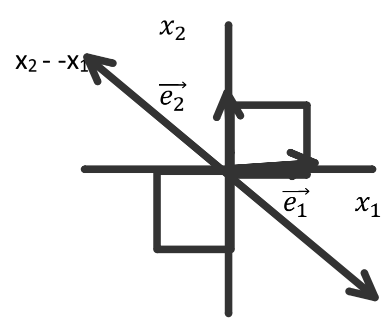
Standard Matrix: $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$

Two Dimensional Examples: Reflection
Reflection through x_2-x_1 axis



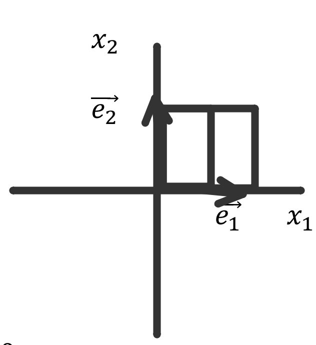
Standard Matrix: $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

Reflection through x_2 - -$x_1$axis



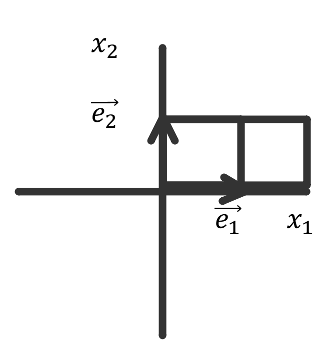
Standard Matrix: $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$

Two Dimensional Examples: Contractions and Expansions
Horizontal Contraction



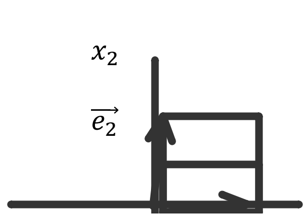
Standard Matrix: $\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix} |k| < 1$

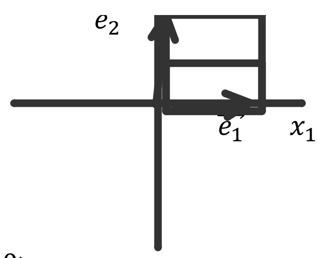
Horizontal Expansion



Standard Matrix: $\begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix} k > 1$

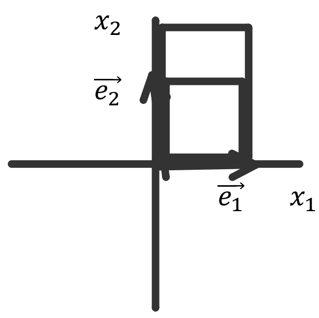
Two Dimensional Examples: Contractions and Expansions
Vertical Contraction





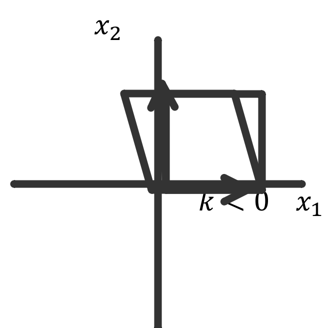
Standard Matrix: $\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix} |k| < 1$

Vertical Expansion



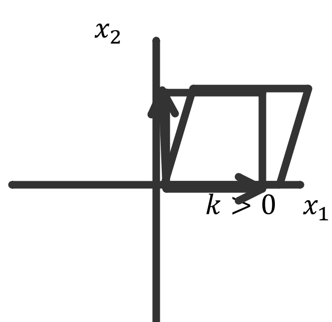
Standard Matrix: $\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix} k > 1$

Two Dimensional Examples: Shears
Horizontal Shear(left)



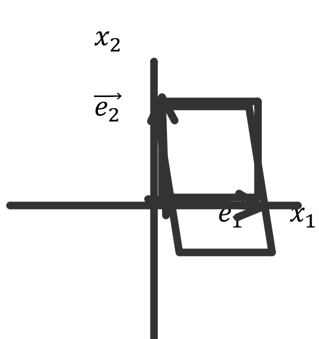
Standard Matrix: $\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} k > 0$

Horizontal Shear(right)



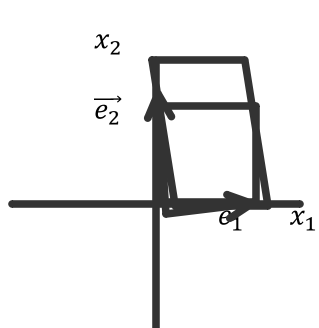
Standard Matrix: $\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} k < 0$

Two Dimensional Examples: Shears
Vertical Shear(down)



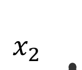
Standard Matrix: $\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix} k > 0$

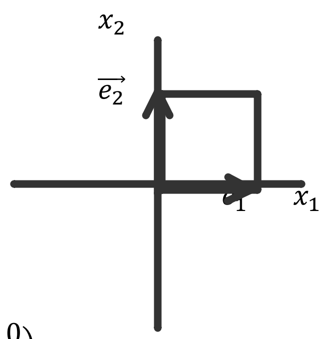
Vertical Shear(up)



Standard Matrix: $\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix} k < 0$

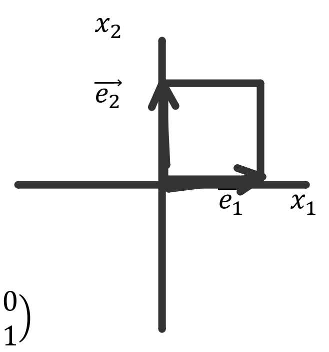
Two Dimensional Examples: Projections
Projection onto the x_1 axis





Standard Matrix: $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

Projection onto the x_2 axis



Standard Matrix: $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

Onto

Definition:

A linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is onto if for all $\vec{b} \in \mathbb{R}^m$ there is a $\vec{x} \in \mathbb{R}^n$ so that $T(\vec{x}) = \vec{b}$.

Onto is an existence property: for any $\vec{b} \in \mathbb{R}^m$, $A\vec{x} = \vec{b}$ has a solution.

Examples:

- A rotation on the plane is an onto linear transformation.
- A projection in the plane is not onto.

Useful Fact:

T is onto if and only if its standard matrix has a pivot in every row.

Lecture 9

Monday, September 13, 2021 3:41 PM

Notes:

One-to-One

Definition:

A linear transformation $T\colon\ \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one if for all $b \in \mathbb{R}^m$ there is at most one (possibly no) $\vec{x} \in \mathbb{R}^n$ so that $T(\vec{x}) = \vec{b}$

One-to-one is a uniqueness property, it does not assert existence for all $\vec{b}$.

Ex.

- A rotation on the plane is a one-to-one linear transformation.
- A projection in the plane is not one-to-one.

Useful Facts

- T is one-to-one if and only if the only solution to $T(\vec{x}) = \vec{0}$ is the zero vector, $\vec{x} = \vec{0}$.
- T is one-to-one if and only if the standard matrix A of T has no free variables.

Ex.

Complete the matrices below by entering numbers into the missing entries so that the properties are satisfied. If it isn't possible to do so, state why.

a) A is a 2×3 standard matrix for a one-to-one linear transform.

$A = \begin{pmatrix} 1 & 0 & \\ 0 & & 1 \end{pmatrix}$ impossible: #col > # rows

b) B is a 3×2 standard matrix for an onto linear transform.

$B = \begin{pmatrix} 1 & \\ & \\ & \end{pmatrix}$ impossible: #rows > #cols

c) C is a 3×3 standard matrix of a linear transform that is one-to-one and onto.

$C = \begin{pmatrix} 1 & 1 & 1 \\ & & \\ & & \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & \pi \\ 0 & 0 & 15 \end{pmatrix}$ if $T\colon\ \mathbb{R}^n \to \mathbb{R}^n$, then "one-to-one" $\Leftrightarrow$ "onto"

Theorem

For a linear transformation $T\colon\ \mathbb{R}^n \to \mathbb{R}^m$ with standard matrix A these are equivalent statements.

a) T is onto.
b) The matrix A has columns which span $\mathbb{R}^m$. n ≥ m
c) The matrix A has m pivotal columns.

Theorem

For a linear transformation $T\colon\ \mathbb{R}^n \to \mathbb{R}^m$ with standard matrix A these are equivalent statements.

1. T is one-to-one.
2. The unique solution to $T(\vec{x}) = \vec{0}$ is the trivial one
3. The matrix A linearly independent columns. m ≥ n
4. Each column of A is pivotal.

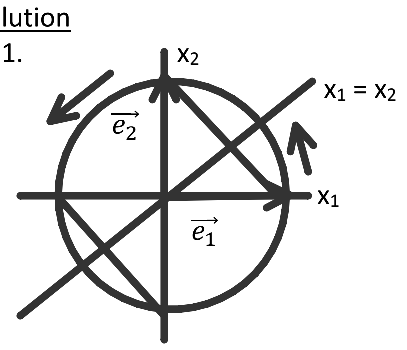
Additional Examples

1. Construct a matrix $A \in \mathbb{R}^{2\times 2}$, such that $T(\vec{x}) = A\vec{x}$, where T is a linear transformation that rotates vectors in $\mathbb{R}^2$ counterclockwise by $\frac{\pi}{2}$ radians about the origin, then reflects them through the line $x_1 = x_2$.
2. Define a linear transformation by $T(x_1,\ \ x_2) = (3x_1 + x_2;\ \ 5x_1 + 7x_2;\ \ x_1 + 3x_2)$

   Is T one-to-one? Is T onto?

Solution

1.



$\vec{e_1} \xrightarrow{Ref} \vec{e_2} \xrightarrow{Ref} \vec{e_1}$

$\vec{e_2} \to -\vec{e_1} \to -\vec{e_2}$

$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

$T(\vec{e_1})$ $T(\vec{e_2})$

2. $T\colon\ \mathbb{R}^2 \to \mathbb{R}^3$

   ↳ Not onto because columns > rows

   $A = \begin{pmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{pmatrix}$ $T(\vec{e_1}) = T(1,\ \ 0) = \begin{pmatrix} 3 \\ 5 \\ 1 \end{pmatrix},\ \ T(\vec{e_2}) = T(0,\ \ 1) = \begin{pmatrix} 1 \\ 7 \\ 3 \end{pmatrix}$

   ↳ Two linearly independent columns: one-to-one.

Section 2.1: Matrix Operations

Definition: Zero and Identity Matrices

1. A zero matrix is any matrix whose every entry is zero.

   $0_{2\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad 0_{2\times 1} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

2. The $n \times n$ identity matrix has ones on the main diagonal, otherwise all zeros.

   $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Note: any matrix with dimensions $n \times n$ is **square**. Zero matrices need not be square, identity matrices must be square.

Sums and Scalar Multiples
Suppose $A \in \mathbb{R}^{m\times n}$, and $a_{i,j}$ is the element of A in row i and column j.
1. If A and B are $m \times n$ matrices, then the elements of $A + B$ are $a_{i,j} + b_{i,j}$.
2. If $c \in \mathbb{R}$, then the elements of cA are $ca_{i,j}$.

For example, if
$$\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix} + c\begin{bmatrix}7 & 4 & 7\\0 & 0 & k\end{bmatrix} = \begin{bmatrix}15 & 10 & 17\\4 & 5 & 16\end{bmatrix}$$

What are the values of c and k?
$c = 2; \quad k = 5$

Properties of Sums and Scalar Multiples
Scalar multiples and matrix addition have the expected properties. If $r, s \in \mathbb{R}$ are scalars, and A, B, C are $m \times n$ matrices, then
1. $A + 0_{m\times n} = A$
2. $(A + B) + C = A + (B + C)$
3. $r(A + B) = rA + rB$
4. $(r + s)A = rA + sA$
5. $r(sA) = (rs)A$

Matrix Multiplication
Definition
Let A be a $m \times n$ matrix, and B be **a** $n \times p$ matrix. The product is AB a $m \times p$ matrix, equal to
$$\mathrm{AB} = \mathrm{A}\left[\vec{b_1} \quad \cdots \quad \vec{b_p}\right] = \left[A\vec{b_1} \quad \cdots \quad A\vec{b_p}\right]$$

Note: the dimensions of A and B determine whether AB is defined, and what its dimensions will be.

$$\underbrace{\underset{m\times n}{\underbrace{A}} \qquad \underset{n\times p}{\underbrace{B}}}_{dimensions\ of\ product}$$

$A \in \mathbb{R}^{m\times n} \Rightarrow \vec{x} \in \mathbb{R}^n; \quad b_i \in \mathbb{R}^n$

Row Column Rule for Matrix Multiplication
The Row Column Rule is a convenient way to calculate the product AB that many students have encountered in pre-requisite courses.

Row Column Method
If $A \in \mathbb{R}^{m\times n}$ has rows $\vec{a_i}$, and $B \in \mathbb{R}^{n\times p}$ has columns $\vec{b_j}$, each element of the product $C = AB$ is $c_{i,j}$=$\vec{a_i}.\vec{b_j}$

Ex.
Compute the following using the row-column method.
$$C = AB = \begin{pmatrix}2 & 0\\1 & -1\end{pmatrix}\begin{pmatrix}3 & 0 & 1\\4 & 5 & 6\end{pmatrix}$$
$$\therefore C = \begin{pmatrix}6 & 0 & 2\\-1 & -5 & -5\end{pmatrix}$$
BA: not possible
$AB \neq BA$ in general

Properties of Matrix Multiplication
Let A, B, C be matrices of the sizes needed for the matrix multiplication to be defined, and A is a $m \times n$ matrix.
1. (Associative) $(AB)C = A(BC)$
2. (Left Distributive) $A(B + C) = AB + AC$
3. (Right Distributive) $(A + B)C = AC + BC$
4. (Identity for matrix multiplication) $I_m A = AI_n$

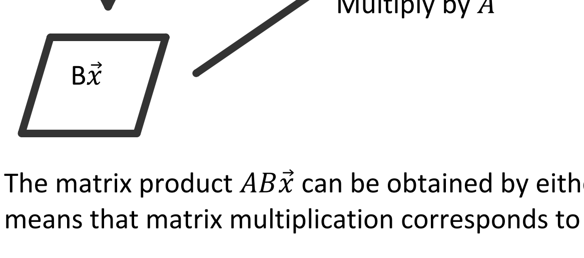
Warnings:
1. (non-commutative) In general, $AB \neq BA$.
2. (non-cancellation) $AB = AC$ does not mean $B = C$.
3. (Zero divisors) $AB = 0$ does not mean that either $A = 0$ or $B = 0$.

Ex.
$$A = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} \quad B = \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix} \quad C = \begin{pmatrix}0 & 0\\0 & 2\end{pmatrix}$$

$AB = 0, \qquad AB = AC,$
but: $A \neq B,\ B \neq C,\ B \neq 0$

The Associative Property
The associative property is $(AB)C = A(BC)$. If $C = \vec{x}$, then
$(AB)\vec{x} = A(B\vec{x})$

Schematically:



The matrix product $AB\vec{x}$ can be obtained by either: multiplying by matrix AB, or by multiplying by B then by A. This means that matrix multiplication corresponds to composition of the linear transformations.

$$AB\vec{x}\;\left|\;\begin{array}{l} first\colon B\vec{x} \\ then\colon\; A(B\vec{x}) \end{array}\right.$$

Studio 7

Tuesday, September 14, 2021 12:34 PM

Worksheet 1.9, Linear Transforms

1. Indicate whether the statements are true or false.
    a. If A is a 3×2 matrix then the map $x \to Ax$ cannot be one-to-one.
        i. False
    b. If A is a 2×3 matrix then the map $x \to Ax$ cannot be onto.
        i. False
    c. T_A: $\mathbb{R}^n \to \mathbb{R}^m$ is one-to-one if and only if $A\vec{x} = \vec{0}$ only has the trivial solution.
        i. True
2. Construct the standard matrix of the linear transformation T.
    a. T: $\mathbb{R}^2 \to \mathbb{R}^4$, where $T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}3\\1\\4\\1\end{bmatrix}$ and $T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}1\\6\\1\\8\end{bmatrix}$

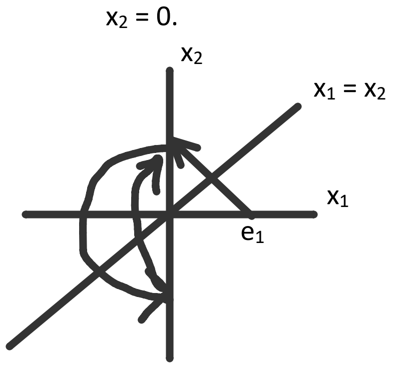
        $\begin{bmatrix}3 & 1\\1 & 6\\4 & 1\\1 & 8\end{bmatrix}$

    b. T is a vertical shear given by $T(\vec{e_2}) = 2\vec{e_2}$ and $T(\vec{e_1}) = \vec{e_1} - 2\vec{e_2}$

        $[\vec{e_1} \quad -2\vec{e_2} \quad 2\vec{e_2}]$

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

    c. A matrix $A \in \mathbb{R}^{2\times 2}$ such that $T(\vec{x}) = A\vec{x}$. T is a linear transformation that first reflects vectors across the line $x_1 = x_2$ then rotates them counterclockwise by π radians about the origin, then reflects them across the line $x_2 = 0$.



$$\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix} = \underbrace{\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}}_{x_2 = 0}\underbrace{\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}}_{\pi}\underbrace{\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}}_{x_1 = x_2}$$

<u>Notes:</u>

T: $\mathbb{R}^n \to \mathbb{R}^m$, $\quad T(x) = Ax$

The following are equivalent

- T is one-to-one.
- $Ax = 0$ has only the trivial solution
- No free variables in A
- Columns of A are independent.
- T is onto
- Pivot in every row of A
- $Ax = b$ consistent for every $b \in \mathbb{R}^m$
- $\mathrm{Ran}(T) = \mathbb{R}^m$
- If T is 1-1 and onto, then $m = n$ and RREF of A is I.

Worksheet 2.1, Matrix Operations

1. Written Explanation Exercise
    For square matrices A, B, is it always true that $(A + B)^2 = A^2 + 2AB + B^2$? Explain why/why not.
    a. No, because AB does not always equal BA
2. Consider:
    $$A = \begin{pmatrix}1 & 1\\0 & 1\end{pmatrix} B = \begin{pmatrix}1 & h\\k & 1\end{pmatrix}$$
    For what values (if any) of $k \in \mathbb{R}$ and $h \in \mathbb{R}$:
    a. Do matrices A and B commute?
        False
    b. Is the product AB equal to I_2?
        True
    c. Is the product AB equal to the 2×2 zero matrix $0_{2\times 2}$?
        True
3. A is an $n \times n$ matrix that has elements $a_{i,j}$ where
    $$a_{i,j} = \begin{cases}0, when\ i + j\ is\ even\\1, when\ i + j\ is\ odd\end{cases}$$
    For $n \geq 2$, how many pivot columns does A have?
    2 pivots

Lecture 10

Wednesday, September 15, 2021 3:28 PM

<u>Notes:</u>

Proof of the Associative Law

Let A be $m \times n$, $B = [\vec{b}_1 \quad \cdots \quad \vec{b}_p]$ a $n \times p$ and $C = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$ $p \times 1$ matrix. Then,

$$BC = \underbrace{c_1\vec{b}_1 + \cdots + c_p\vec{b}_p}_{lin\ combin\ of\ cols\ of\ B}$$

So,

$$\begin{aligned} A(BC) &= A(c_1\vec{b}_1 + \cdots + c_p\vec{b}_p) \\ &= c_1A\vec{b}_1 + \cdots + Ac_p\vec{b}_p \qquad \text{(multiply by } A \text{ is linear)} \\ &= [A\vec{b}_1 \quad \cdots \quad A\vec{b}_p]\begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix} \qquad \text{(lin combin of cols of } AB\text{)} \\ &= (AB)C \end{aligned}$$

Ex.

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

Give an example of a 2×2 matrix B that is non-commutative with A.

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

Try, $B = \begin{bmatrix} 1 & 6 \\ 6 & 1 \end{bmatrix} \begin{cases} AB = \begin{bmatrix} 7 & 7 \\ 0 & 0 \end{bmatrix} \\ BA = \begin{bmatrix} 1 & 1 \\ 6 & 6 \end{bmatrix} \end{cases}$

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}: \ AB \neq BA$$

The Transpose of a Matrix

A^T is the matrix whose columns are the rows of A.

Ex.

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 0 & 2 & 0 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 0 \\ 4 & 2 \\ 5 & 0 \end{bmatrix}$$

Properties of the Matrix Transpose

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(rA)^T = r(A^T)$
4. $(AB)^T = B^T A^T$

$$\underbrace{A}_{m\times n} \times \underbrace{B}_{n\times p} = \underbrace{AB}_{m\times p} \Rightarrow \left(\underbrace{AB}_{m\times p}^{T}\right)$$

$$\underbrace{B}_{p\times n}{}^T \times \underbrace{A}_{n\times m}{}^T = \left(\underbrace{AB}_{p\times m}^{T}\right)$$

AB: $\quad (AB)_{i,j} = \overrightarrow{Row}(A, i).\overrightarrow{Col}(B, j)$

$$\begin{aligned} (AB)^T \Rightarrow (AB)_{j,i} &= \overrightarrow{\text{Row}}(A, j).\overrightarrow{Col}(B, i) \\ &= \overrightarrow{\text{Row}}(B^T, i).\overrightarrow{Col}(A^T, j) \end{aligned}$$

Matrix Powers

$$x^2 = x \cdot x$$

$$A^2 = \underbrace{A}_{n\times m} \cdot \underbrace{A}_{n\times m} \Rightarrow n \text{ must be } m \text{ (square matrix)}$$

For any $n \times n$ matrix and possible integer k, A^k is the product of k copies of A.

$$A^k = AA \dots A$$

Ex. Compute C^2

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

$$C^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

$$\hookrightarrow C^8 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2^8 & 0 \\ 0 & 0 & 2^8 \end{bmatrix}$$

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \Rightarrow A^2 = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$

$$A = \begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix} \Rightarrow A^2 = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$$

$rot(^{\pi}/_{6})$ $rot(^{\pi}/_{3})$

Ex.

Define:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 8 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Which of these operations are defined, and what are the dimensions of the result?

1. $A + 3C$: Not Possible
2. $A(AB)^T$: $A \in \mathbb{R}^{2\times 2}$ $(AB)^T \in \mathbb{R}^{3\times 2}$: Not possible
3. $A + ABCB^T$: $A + ABCB^T \in \mathbb{R}^{2\times 2}$
4. $(AB)^2$: NP

   $AB \in \mathbb{R}^{2\times 3}$

   $(AB)^2 = ABAB$

   $\neq A^2B^2$

Additional Examples

$(a-b)(a+b) = a^2 - b^2$

$(a+b)^2 = a^2 + 2ab + b^2$

True or False:

1. For any I_n and any $A \in \mathbb{R}^{n\times m}$, $(I_n + A)(I_n - A) = I_n - A^2$

   $(I_n + A)(I_n - A) = {I_n}^2 + AI_n - I_nA - A^2$

   $= {I_n}^2 - A^2$: TRUE

2. For any A and B in $\mathbb{R}^{n\times n}$, $(A+B)^2 = A^2 + 2AB + B^2$

   $(A+B)^2 = (A+B)(A+B)$

   $= A^2 + AB + BA + B^2$

   $\neq 2AB$: FALSE

$$I_n = \begin{pmatrix} 1 & & (0) \\ & \ddots & \\ (0) & & 1 \end{pmatrix}$$

Unit 2

Saturday, November 13, 2021 8:49 PM

<u>Material Covered:</u>

Chapter 2: Matrix Algebra

- Section 2.2 : Inverse of a Matrix
- Section 2.3 : Invertible Matrices
- Section 2.4 : Partitioned Matrices
- Section 2.5 : Matrix Factorizations
- Section 2.6 : The Leontif Input-Output Model
- Section 2.7 : Computer Graphics
- Section 2.8 : Subspaces of $\mathbb{R}^n$
- Section 2.9 : Dimension and Rank

Chapter 3: Determinants

- Section 3.1 : Introduction to Determinants
- Section 3.2 : Properties of the Determinant
- Section 3.3 : Volume, Linear Transformations

Chapter 4: Vector Spaces

- Section 4.9 : Applications to Markov Chains

Chapter 5: Eigenvalues and Eigenvectors

- Section 5.1 : Eigenvectors and Eigenvalues
- Section 5.2 : The Characteristic Equation

Lecture 11

Friday, September 17, 2021 3:17 PM

Notes:

Section 2.2: Inverse of a Matrix

The Matrix Inverses

Definition:

$A \in \mathbb{R}^{n\times n}$ is invertible (or non-singular) if there is a $C \in \mathbb{R}^{n\times n}$ so that

$$AC = CA = I_n$$

If there is, we write $C = A^{-1}$

$$A \cdot A^{-1} = A^{-1} \cdot A = I_n$$

Uniqueness:

If $AC = CA = I_n$

And $AD = DA = I_n$

$CAD = CDA = C \ \hookleftarrow \times c$

The inverse of a 2×2 Matrix

There's a formula for computing the inverse of a 2×2 matrix.

Theorem:

The 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is non-singular if and only if $ad - bc \neq 0$, and then

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Ex.

State the inverse of the Matrix below

$$A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$$

$2 \times (-7) - (-3) \times 5 \neq 0$: A is invertible

$$A^{-1} = \frac{1}{2 \times (-7) - (-3) \times 5} = \begin{pmatrix} -7 & -5 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} -7 & -5 \\ 3 & 2 \end{pmatrix}$$

The Matrix Inverse

$$= I_n A\vec{x} = \vec{b} \Longrightarrow A^{-1}A\vec{x} = A^{-1}\vec{b}$$

Theorem:

$A \in \mathbb{R}^{n\times m}$ has an inverse if and only if for all $\vec{b} \in \mathbb{R}^n$, $A\vec{x} = \vec{b}$ has a unique solution. And, in this case $\vec{x} = A^{-1}\vec{b}$

Ex.

Solve the linear system

$3x_1 + 4x_2 = 7$

$5x_1 + 6x_2 = 7$

$$\begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix}\vec{x} = \begin{pmatrix} 7 \\ 7 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix}^{-1} = \frac{1}{(3 \times 6) - (5 \times 4)}\begin{pmatrix} 6 & -4 \\ -5 & 3 \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} 6 & -4 \\ -5 & 3 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 2.5 & -1.5 \end{pmatrix}$$

$$\therefore \vec{x} = A^{-1}\vec{b} = \begin{pmatrix} -3 & 2 \\ 2.5 & -1.5 \end{pmatrix}\begin{pmatrix} 7 \\ 7 \end{pmatrix} = \begin{pmatrix} -7 \\ 7 \end{pmatrix}$$

Properties of the Matrix Inverse

A and B are omvertob;e $n \times n$ matricies.

1. $(A^{-1})^{-1} = A$
2. $(AB)^{-1} = B^{-1}A^{-1}$ (Non-commutative)
3. $(A^T)^{-1} = (A^{-1})^T$

Ex.

True or False: $(ABC)^{-1} = C^{-1}B^{-1}A^{-1} \Longrightarrow$True

2. $\begin{cases} (B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I_n \\ (AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I_n \end{cases}$

3. $AA^{-1} = A^{-1}A = I_n$

   $(A^{-1})^T A^T = A^T(A^{-1})^T = I_n$

An Algorithm for Computing A^{-1}

If $A \in \mathbb{R}^{n\times n}$ and $n > 2$, how do we calculate A^{-1}?

Here's an algorithm we can use:

1. Row reduce the augmented matrix $(A|I_n)$
2. If reduction has form $(I_n|B)$ the A is invertible and $B = A^{-1}$. Otherwise, A is not invertible.

Ex.

Compute the inverse of $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 0 & 0 & 1 \end{bmatrix}$

$(A\vert I_n)$	$\left(\begin{array}{ccc\|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$ {solve "$AC = I_n$" → "$(A\vert I_n)$"}
$R_1 \leftarrow R_2$ ~	$\left(\begin{array}{ccc\|ccc} 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$
$R_1 \leftarrow R_1 - 3R_3$ ~ $R_2 \leftarrow R_2 - 2R_3$	$\left(\begin{array}{ccc\|ccc} 1 & 0 & 0 & 0 & 1 & -3 \\ 0 & 1 & 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$
$\therefore A =$	$\begin{pmatrix} 0 & 1 & -3 \\ 1 & 0 & -2 \\ 0 & 0 & 1 \end{pmatrix}$

Why Does This Work?

We can think of our algorithm as simulatenously solving n linear systems:

$$A\vec{x} = \overrightarrow{e_1}$$
$$A\vec{x} = \overrightarrow{e_2}$$
$$\vdots$$
$$A\vec{x} = \overrightarrow{e_n}$$

Each column of A^{-1} is $A^{-1}\overrightarrow{e_1} = \vec{x}$

Over the next few slides we explore another explanation for how ouralgorithm works. This other explanation uses elementary matrices.

Elementary Matrices

An elementary matrix, E, is one that differs byInby one row operation. Recall our elementary row operations:

1. swap rows
2. multiply a row by a non-zero scalar
3. add a multiple of one row to anotherWe can represent each operation by a matrix multiplication with an elementary matrix.

Ex.

- Swap $R_1 \leftrightarrow R_3$: $E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
- $R_2 \longleftarrow \pi R_2$ $E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

E: $R_3 \leftarrow R_3 + 5R_1$ $E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{pmatrix}$

$$E = E\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{pmatrix}$$

Ex.

Suppose

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

By inspection, what is E? How does it compare to I_3?

E: $R_2 \leftarrow R_2 + 2R_1$

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

Theorem

Returning to understanding why our algorithm works, we apply a sequence of row operations to A to obtain I_n:

$$(E_k \quad \cdots \quad E_3 \quad E_2 \quad E_1) \ \ A = I_n$$

Thus, $E_k \quad \cdots \quad E_3 \quad E_2 \quad E_1$ is the inverse matrix we seek.

Our algorithm for calculating the inverse of a matrix is the result of the following theorem.

Theorem

Matrix A is invertible if and only if it is row equivalent to the identity. In this case, the any sequence of elementary row operations that transforms A into I, applied I, generates A^{-1}.

$$(A|I_n)$$
$$(E_1A|E_1I_n)$$
$$(E_2E_1A|E_2E_1I_n)$$
$$(E_k \dots E_2E_1A|E_k \dots E_2E_1I_n)$$

Using The Inverse to Solve a Linear System

- We could use A^{-1} to solve a linear system

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

- We could calculate A^{-1} and then: $\vec{x} = A^{-1}\vec{b}$

- As our textbook points out, A^{-1} is seldom used: computing it can take a very long time, and is prone to numerical error.
- So why did we learn how to compute A^{-1}? Later on in this course,we use elementary matrices and properties of A^{-1} to derive results.
- A recurring theme of this course: just because we can do something a certain way, doesn't that we should.

Lecture 12

Monday, September 20, 2021 9:49 PM

Notes:

Section 2.3 Invertible Matrices

The Invertible Matrix Theorem

Theorem

Let A be an $n \times n$ matrix. These statements are all equivalent.

a) A is invertible.
b) A is row equivalent to I_n.
c) A has n pivotal columns. (All columns are pivotal.)
d) $A\vec{x} = \vec{0}$ has only the trivial solution.
e) The columns of A are linearly independent.
f) The linear transformation $\vec{x} \mapsto A\vec{x}$ is one-to-one.
g) The equation $A\vec{x} = \vec{v}$ has a solution for all $\vec{b} \in \mathbb{R}^n$.
h) The columns of A span $\mathbb{R}^n$.
i) The linear transformation $\vec{x} \mapsto A\vec{x}$ is onto.
j) There is a $n \times n$ matrix C so that $CA = I_n$. (A has a left inverse.)
k) There is a $n \times n$ matrix D so tha $AD = I_n$. (A has a right inverse.)
l) A^T is invertible.

Proofs:

If $CA = I_n$

If $A\vec{x} = A\vec{y}$: $cA\vec{x} = cA\vec{y}$

Thus, $\vec{x} \to Ax$ is one-to-one

If $AD = I_n$

Take: $\vec{b} \in \mathbb{R}^n$: $(AD)\vec{b} = I_n\vec{b}$

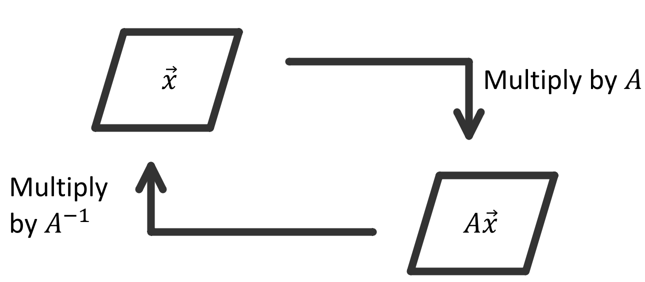
$A(D\vec{b}) = \vec{b}$

Thus, $\vec{x} \to Ax$ is onto

$CA = I_n \Longrightarrow A^T C^T = I_n$

Invertibility and Composition

The diagram below gives us another perspective on the role of A^{-1}.



The matrix inverse A^{-1} transforms Ax back to $\vec{x}$. This is because:

$A^{-1}(A\vec{x}) = (A^{-1}A)\vec{x} = \vec{x}$

The Invertible Matrix Theorem: Final Notes

Items j and k of the invertible matrix theorem (IMT) lead us directly to the following theorem.

Theorem

If A and B are $n \times n$ matricies and $AB = I_n$ then A and B are invertible and $B = A^{-1}$ and $A = B^{-1}$

- The IMT is a set of equivalent statements. They divide the set of all square matrices into two separate classes: invertible, and non-invertible.
- As we progress through this course, we will be able to add additional equivalent statements to the IMT (that deal with determinants, eigenvalues, etc).

If $AB = I_n$

By IMT: A is invertible

$A^{-1}(AB) = A^{-1}I_n$

Ex.1

Is this matrix invertible?

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

$\begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix}$	$R_2 \leftarrow R_2 - 3R_1$ ~ $R_3 \leftarrow R_3 + R_1$	$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 4 \\ 0 & -1 & -1 \end{bmatrix}$	$R_3 \leftarrow R_3 + R_2$ ~	$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix}$

Ex. 2

Singular = not invertible

If possible, fill in the missing elements of the matrices below with numbers so that each of the matrices are singular. If it is not possible to do so, state why.

$$\begin{pmatrix}1 & 0 & 1\\ 1 & 0 & 1\\ 0 & 0 & 1\end{pmatrix} \qquad \begin{pmatrix}1 & & 1\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix} \qquad \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 1\\ 0 & \boxed{1} & 1\end{pmatrix}$$

Not possible
2 pivot columns

Not possible
pivot in all columns

Possible
2 pivot columns

Matrix Completion Problems

- The previous example is an example of a matrix completion problem (MCP).
- MCPs are great questions for recitations, midterms, exams.
- The Netflix Problem is another example of an MCP.

Given a ratings matrix in which each entry (i, j) represents the rating of movie j by customer i if customer i has watched movie j, and is otherwise missing, predict the remaining matrix entries in order to make recommendations to customers on what to watch next.

Section 2.4: Partitioned Matricies

What is a partitioned matrix?
Ex.
This matrix:

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

Can also be written as:

$$\begin{bmatrix}\begin{bmatrix}3 & 1 & 4\\ 1 & 6 & 1\end{bmatrix} & \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\\ [0 \quad 0 \quad 0] & [4 \quad 2]\end{bmatrix}$$

We partitioned our matrix into four blocks, each of which has different dimensions.

Another Example of a Partitioned Matrix
Example: The reduced echelon form of a matrix. We can use a partitioned matrix to

$$\begin{bmatrix}1 & 0 & 0 & 0 & * & \cdots & 0\\ 0 & 1 & 0 & 0 & * & \cdots & 0\\ 0 & 0 & 1 & 0 & * & \cdots & 0\\ 0 & 0 & 0 & 1 & * & \cdots & 0\\ 0 & 0 & 0 & 0 & 0 & \cdots & 0\\ 0 & 0 & 0 & 0 & 0 & \cdots & 0\end{bmatrix}$$

This is useful when studying the null space of A, as we will see later in this course.

Row Column Method
Recall that a row vector times a column vector (of the right dimensions) is a scalar. For example,

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

This is the row column matrix multiplication method from Section 2.1.

Theorem
Let A be $m \times n$ and B be $n \times p$ matrix. Then, the (i, j) entry of AB is
$row_i A = col_j B$
This is the Row Column Method for matrix multiplication.

Partitioned matrices can be multiplied using this method, as if each block were a scalar (provided each block has appropriate dimensions).

$$\begin{pmatrix}A & B\\ C & D\end{pmatrix}\begin{pmatrix}E & F\\ G & H\end{pmatrix} = \begin{pmatrix}AE + BG & \cdots\cdots\\ \cdots\cdots & \cdots\cdots\end{pmatrix}$$

Example of Row Column Method
Recall, using our formula for a 2×2 matrix, $\begin{bmatrix}a & b\\ c & d\end{bmatrix}^{-1} = \begin{bmatrix}d & -b\\ -c & a\end{bmatrix}$

Ex.
Suppose $A \in \mathbb{R}^{n\times n}$, $V \in \mathbb{R}^{n\times n}$, and $C \in \mathbb{R}^{n\times n}$ are invertible matrices. Construct the inverse of $\begin{bmatrix}A & B\\ 0 & C\end{bmatrix}$.

$$\begin{pmatrix}A & B\\ 0 & C\end{pmatrix}\begin{pmatrix}W & X\\ Y & Z\end{pmatrix} = \begin{pmatrix}AW + BY & AX + BC\\ CY & CZ\end{pmatrix} = \begin{pmatrix}I_n & 0\\ 0 & I_n\end{pmatrix}$$

W, X, Y, Z ∈ $\mathbb{R}^{n\times n}$

$CY = 0$: $C^{-1}CY = C^{-1}0$
$CZ = I_n$: $Z = C^{-1}$
$AW + BY = I_n$: $W = A^{-1}$

$AX + BZ = 0$: $AX = -BC^{-1}$

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BC^{-1} \\ 0 & C^{-1} \end{pmatrix}$$

2×2 matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ac}\begin{bmatrix} c & -b \\ 0 & a \end{bmatrix}$$

Studio 8

Tuesday, September 21, 2021 12:30 PM

Worksheet 2.2 and 2.3, Invertible Matricies

1. Consider the sequence of row operations that reduce matrix A to the identity.

$$A = \underbrace{\begin{pmatrix} 0 & 4 & 0 \\ 1 & 0 & 0 \\ 0 & 8 & 1 \end{pmatrix}}_{A} \sim \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 8 & 1 \end{pmatrix}}_{E_1 A} \sim \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{E_2 E_1 A} \sim \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{E_3 E_2 E_1 A} = I_n$$

Construct the elementary matrices E_1, E_2, E_3.

$$E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}$$

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

2. Indicate whether the statements are true or false. A is an $n \times n$ matrix.
   a. If $A\vec{x} = A\vec{y}$ for some $\vec{x} \neq \vec{y}$, then A cannot be intertible.
      i. True
   b. If for some $\vec{b} \in \mathbb{R}^n$, $A\vec{x} = \vec{b}$ has more than one solution then A is invertible.
      i. False
   c. Every elementary matrix is invertible.
      i. True

3. Compute the inverse of the matrix, where $c \in \mathbb{R}$. For what values of c does the matrix have an inverse?

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 2 & 0 & 4 \\ 0 & -1 & c \end{bmatrix}$$

$\begin{bmatrix} 0 & 1 & 1 \\ 2 & 0 & 4 \\ 0 & -1 & c \end{bmatrix}$	$R_2 \leftrightarrow R_1$ $\sim$ $R_3 \leftrightarrow R_2$	$\begin{bmatrix} 2 & 0 & 4 \\ 0 & 1 & 1 \\ 0 & -1 & c \end{bmatrix}$	$R_3 \leftarrow R_3 + R_2$ $R_1 \leftarrow \frac{1}{2}R_1$ $R_1 \leftarrow R_1 + R_3$	$\begin{bmatrix} 1 & 0 & c+3 \\ 0 & 1 & 1 \\ 0 & 0 & c+1 \end{bmatrix}$	$R_3 \leftarrow \frac{1}{c+1}R_3$ $\sim$	$\begin{bmatrix} 1 & 0 & c+3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

$c \neq -1$

$$\therefore A^{-1} = E_3 E_2 E_1 A = \begin{pmatrix} -\frac{2}{c+1} & \frac{1}{2} & -\frac{2}{c+1} \\ 1 - \frac{1}{c+1} & 0 & -\frac{1}{c+1} \\ \frac{1}{c+1} & 0 & \frac{1}{c+1} \end{pmatrix}$$

4. Let A be an $n \times n$ matrix. Which statements guarantee that A is invertible?
   a. Every vector in $\mathbb{R}^n$ is in the span of the columns of A.
      i. True
   b. $A\vec{x} = \vec{b}$ has a solution for every $\vec{b} \in \mathbb{R}^n$.
      i. True
   c. Matrix A can be row reduced to the identitity matrix.
      i. True
   d. The range of the linear transform $\vec{x} \to A\vec{x}$ is $\mathbb{R}^n$.
      i. True

5. Two reasons that a matrix is not invertible are:
   a. One column is a multiple of another column.
   b. One column is the sum of other columns.

By inspection, identify which of the reasons above apply to these matrices.

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

$$C = \begin{bmatrix} 1 & 2 & 3 \\ 7 & 14 & 21 \\ 5 & 10 & 15 \end{bmatrix} \qquad D = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 2 \\ 3 & -1 & 2 \end{bmatrix}$$

Lecture 13

Wednesday, September 22, 2021 3:32 PM

Section 2.5: Matrix Factorization

Motivation

- Recall that we could solve $A\vec{x} = \vec{b}$ by using
  $\vec{x} = A^{-1}\vec{b}$
- This requires computation of the inverse of an $n \times n$ matrix, which is especially difficult for large n.
- Instead we could solve $A\vec{x} = \vec{b}$ with Gaussian Elimination, but this is not efficient for large n.
- There are more efficient and accurate methods for solving linear systems that rely on matrix factorizations.

Matrix Factorizations

- A matrix factorization, or matrix decomposition is a factorization of a matrix into a product of matrices.
- Factorizations can be useful for solving $A\vec{x} = \vec{b}$, or understanding the properties of a matrix.
- We explore a few matrix factorizations throughout this course.
- In this section, we factor a matrix into lower and into upper triangular matrices.

Triangular Matrices

- A rectangular matrix A is upper triangular if $a_i, j = 0$ for $i > j$

$$\begin{pmatrix} 1 & 5 & 0 \\ 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

- A rectangular matrix A is lower triangular if $a_i, j = 0$ for $i < j$

$$\begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}$$

Ask: Can you name a matrix that is both upper and lower triangular?

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad 0_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Diagonal matrix: $D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

The LU Factorization

Theorem

If A is an $m \times n$ matrix that can be row reduced to echelon form without row exchanges, then $A = LU$. L is a lower triangle $m \times n$ matrix with 1's on the diagonal, U is an echelon form of A.

Ex.

If $A \in \mathbb{R}^{3\times 2}$, the LU factorization has the form:

$$A = LU = \begin{pmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{pmatrix} \begin{pmatrix} * & * \\ 0 & * \\ 0 & 0 \end{pmatrix}$$

<u>Fact:</u> the LU factorization is unique

If $A = L_1U_1 = L_2U_2$

Then $L_1 = L_2$ and $U_1 = U_2$

Why We Can Compute the LU Factorization

Suppose A can be row reduced to echelon form U without interchanging rows. Then,

$E_p \cdots E_1 = U$

where the E_j are matricies that perform elementary row operations. They happen to be lower triangular and invertible, e.g.

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

Therefore,

$A = {E_1}^{-1} \cdots {E_p}^{-1} U = LU$

$R_3 \leftarrow R_3 - 2R_2$: $\quad E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}$

$E^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$

$$\begin{pmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{pmatrix}$$

What about swap?

$R_1 \longleftrightarrow R_3$: $\quad E = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$: Not lower Triangular

Using the LU Decomposition

Goal: given A and $\vec{b}$, solve $A\vec{x} = \vec{b}$ for $\vec{x}$.

Algorithm: construct $A = LU$, solve $A\vec{x} = LU\vec{x} = \vec{b}$ by:

1. Forward solve for $\vec{y}$ in $L\vec{y} = \vec{b}$.
2. Backwards solve for $\vec{x}$ in $U\vec{x} = \vec{y}$.

Ex.

Solve the linear system whose LU decomposition is given.

$$A = LU = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \vec{b} = \begin{pmatrix} 2 \\ 3 \\ 2 \\ 0 \end{pmatrix}$$

1. Forward: $L\vec{y} = \vec{b}$ $\qquad (\vec{y} = U\vec{x})$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 2 \\ 0 \end{pmatrix}$$

First row: y_1 = z
Second row: $y_1 + y_2 = 3$ → $y_2 = 1$
Third row: $2y_2 + y_3 = 2$ → $y_4 = 0$
Last row: $y_3 + y_4 = 0$ → $y_4 = 0$

↳ $\vec{y} = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$: $\qquad L\vec{y} = \vec{b}$

But we want $A\vec{x} = L(U\vec{x}) = \vec{b}$

1. Backward: $U\vec{x} = \vec{y}$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Last row: 0 = 0
Third row: $2x_3 = 6\ 0$ → $x_3 = 0$
Second row: $2x_2 + x_3 = 1$ → $x_2 = {}^1/_2$
First row: $x_1 = 2$

Thus: $\vec{x} = \begin{pmatrix} 2 \\ {}^1/_2 \\ 0 \end{pmatrix}$: $\qquad$ solution of $A\vec{x} = \vec{b}$

Indeed: $A\vec{x} = LU\vec{x} = L\vec{y} = \vec{b}$

An Algorithm for Computing LU

To compute the LU decomposition:

1. Reduce A to an echelon form U by a sequence of row replacement operations, if possible.
1. Place entries in L such that the same sequence of row operations reduces L to I.

Note that

- In MATH 1554, the only row replacement operation we can use is to replace a row with a multiple of a row above it.
- More advanced linear algebra courses address this limitation.

Ex.

Compute the LU factorization of A.

$$A = \begin{pmatrix} 4 & -3 & -1 & 5 \\ -16 & 12 & 2 & -17 \\ 8 & -6 & -12 & 22 \end{pmatrix}$$

A	E_1: $R_2 \leftarrow R_2 + 4R_1$ ~ E_2: $R_3 \leftarrow R_3 - 2R_1$	$\begin{pmatrix} 4 & -3 & -1 & 5 \\ -16 & 12 & 2 & -17 \\ 8 & -6 & -12 & 22 \end{pmatrix}$	~ E_3: $R_3 \leftarrow R_3 - 5R_2$	$\begin{pmatrix} 4 & -3 & -1 & 5 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & -3 \end{pmatrix} = U$

$$L = E_1^{-1}E_2^{-1}E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{pmatrix}$$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ \boxed{-4} & 1 & 0 \\ \boxed{2} & \boxed{5} & 1 \end{pmatrix}$$

Summary

- To solve $A\vec{x} = LU\vec{x} = \vec{b}$
  1. Forward solve for $\vec{y}$ in $L\vec{y} = \vec{b}$.
  2. Backwards solve for $\vec{x}$ in $U\vec{x} = \vec{y}$.
- To compute the LU decomposition:
  1. Reduce A to an echelon form U by a sequence of row replacement operations, if possible.
  2. Place entries in L such that the same sequence of row operations reduces L to I.
- The textbook offers a different explanation of how to construct the LU decomposition that students may find helpful.
- Another explanation on how to calculate the LU decomposition that students may find helpful is available from MIT Open Course Ware: www.youtube.com/watch?v=rhNKncraJMk

Studio 9

Thursday, September 23, 2021 12:36 PM

Worksheet 2.4 and 2.5, Partitioned Matrices and Matrix Factorizations

Worksheet Exercises

1. A and B are $n \times n$ invertible matricies, I_n is the $n \times n$ identity matrix, and 0 is the $n \times n$ matrix. Construct an expression for X in terms of A and B.

$$\begin{pmatrix} A & B \end{pmatrix}\begin{pmatrix} 0 & X & B \\ A & B & 0 \end{pmatrix}\begin{pmatrix} X \\ I_n \\ BA \end{pmatrix} = B^2 + BAX$$

$$\begin{pmatrix} A & B \end{pmatrix}\begin{pmatrix} 0 & X & B \\ A & B & 0 \end{pmatrix}\begin{pmatrix} X \\ I_n \\ BA \end{pmatrix} = \begin{pmatrix} AB & AX + B^2 & AB \end{pmatrix}\begin{pmatrix} X \\ I_n \\ BA \end{pmatrix} = (BAX + I_n(AB + B^2) + A^2B^2)$$

$$\therefore (BAX + I_n(AX + B^2) + A^2B^2) = B^2 + BAX \Longrightarrow I_n(AX + B^2) + A^2B^2 = B^2 \Longrightarrow AX + B^2 + A^2B^2 = B^2$$
$$\therefore AX + A^2B^2 = 0 \Longrightarrow X + AB^2 = 0 \Longrightarrow X = -AB^2$$

2. Compute the LU factorization for

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

A	E_1: $R_2 \leftarrow R_2 + R_1$ ~ E_2: $R_3 \leftarrow R_3 - R_2$	$\begin{bmatrix} -1 & 5 & 3 & 1 \\ 0 & -5 & 0 & 2 \\ 0 & -5 & 0 & 2 \end{bmatrix}$	~ E_3: $R_3 \leftarrow R_3 - R_2$	$\begin{bmatrix} -1 & 5 & 3 & 1 \\ 0 & -5 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$

$$L = E_1^{-1}E_2^{-1}E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ \boxed{-1} & 1 & 0 \\ \boxed{0} & \boxed{1} & 1 \end{pmatrix}$$

3. Compute the LU factorization of A and use it to solve for $A\vec{x} = \vec{b}$.

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -2/3 & 1 \end{bmatrix}\begin{bmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 0 \\ 4 \\ -4 \end{bmatrix}$$

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -2/3 & 1 \end{bmatrix}}_{L}\underbrace{\begin{bmatrix} 2 & -1 & 0 \\ 0 & 3/2 & -1 \\ 0 & 0 & 4/3 \end{bmatrix}}_{U}, \qquad \vec{b} = \begin{bmatrix} 0 \\ 4 \\ -4 \end{bmatrix}$$

Solve: $L\vec{y} = \vec{b}$ $\quad (\vec{y} = U\vec{x})$

$$\begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -2/3 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ -4 \end{bmatrix}$$

First row: $y_1 = 0$
Second row: $-1/2 y_1 + y_2 = 4 \rightarrow y_2 = 4$
Third row: $-2/3 y_2 + y_3 = -4 \rightarrow y_3 = -4/3$

$$\Rightarrow \vec{y} = \begin{bmatrix} 0 \\ 4 \\ -4/3 \end{bmatrix}: \qquad L\vec{y} = \vec{b}$$

Solve: $U\vec{x} = \vec{y}$

$$\begin{bmatrix} 2 & -1 & 0 \\ 0 & 3/2 & -1 \\ 0 & 0 & 4/3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ -4/3 \end{bmatrix}$$

Third row: $4/3 x_3 = -4/3 \rightarrow x_3 = -1$
Second row: $3/2 x_2 - x_3 = 4 \rightarrow x_2 = 2$
First row: $2x_1 - x_2 = 0 \rightarrow x_1 = 1$

Thus: $\vec{x} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$: $\qquad$ solution of $A\vec{x} = \vec{b}$

4. Written Explanation Exercise: What is the LU decompostiion good for? Your reasoning should involve computational efficiency.

LU decomposition is generally used for computational efficiency since the amount of steps it takes to find the inverse of A such that $A = LU$ is significantly less when compared to standard and more roundabout way of computing the inverse.

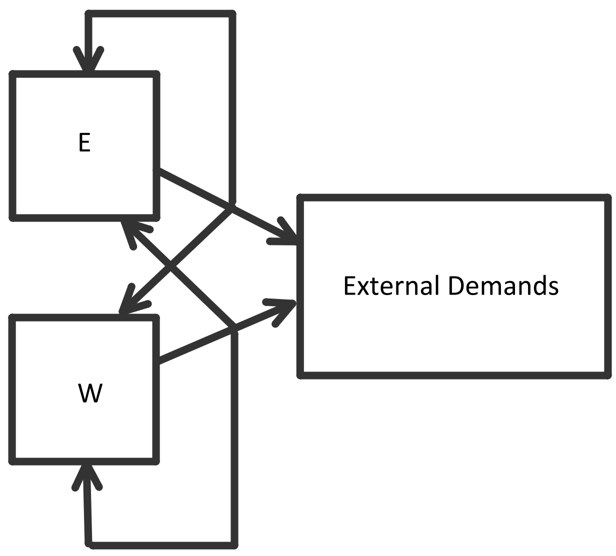
Lecture 14

Friday, September 24, 2021 3:29 PM

Notes:

Section 2.6: The Leontif Input-Output Model

Example: An Economy with Two Sectors



This economy contains two sectors.

1. electricity (E)
2. water (W)

The "external demands" is another part of the economy, which does notproduce E and W.

How might we represent this economy with a set of linear equations?

The Leontif Model: Internal Consumption

Suppose economy has N sectors, with outputs measured by $\vec{x} \in \mathbb{R}^N$.

$\vec{x} =$ output vector $x_i =$ element i of vector

$\vec{x} =$ number of units produced by sector i

The consumption matrix, C, describes how units are consumed by sectors to produce output. Two equivalent ways of defining C:

- Sector j requires a proportion of the units created by sector i. Call that $c_{i,j}x_i$
- Sector i sends a proportion of its units to sector j. Call that $c_{i,j}x_i$

Elements of C are $c_{i,j}$, with $c_{i,j} \in [0, 1]$ and

$C\vec{x} =$ units consumed

$\vec{x} - C\vec{x} =$ units left after internal consumption

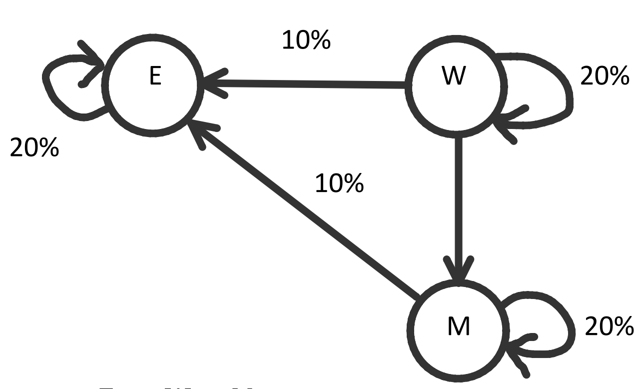
$C = \begin{array}{c} j \\ \left(\; + \; \right) \end{array} i$: $c_{i,j} =$ pattern of sector i needed to produce 1 unit of j.

Ex.

An economy contains three sectors, E, W, M. For every 100 units of output,

- E requires 20 units from E, 10 units from W, and 10 units from M
- W requires 0 units from E, 20 units from W, and 10 units from M
- M requires 0 units from E, 0 units from W, and 20 units from M

Construct the consumption matrix for this economy.



$$C = \begin{array}{c} \begin{array}{ccc} E & W & M \end{array} \\ \begin{pmatrix} 0.2 & 0 & 0 \\ 0.1 & 0.2 & 0 \\ 0.1 & 0.1 & 0.2 \end{pmatrix} \end{array} \begin{array}{c} \\ E \\ W \\ M \end{array}$$

$$C\vec{x} = \begin{pmatrix} 0.2xE & 0 & 0 \\ 0.1xE & 0.2xE & 0 \\ 0.1xE & 0.1xE & 0.2xE \end{pmatrix}$$

Solution: Creating C

Our consumption matrix is

$$C = \frac{1}{10}\begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 2 \end{pmatrix}$$

Note:

- total output for each sector is the sum along the outgoing edges for each sector, which generates rows of C
- elements of C represent percentages with no units, they have values between 0 and 1
- our output vector has units

The Leontif Model: Demand

There is also an external demand given by $\vec{d} \in \mathbb{R}^N$. We ask if there is an $\vec{x}$ such that

$$\vec{x} - C\vec{x} = \vec{d}$$

Solving for $\vec{x}$ yields

$$\vec{x} = (I - C)\vec{d}$$

This is the **Leontief Input-Output Model**.

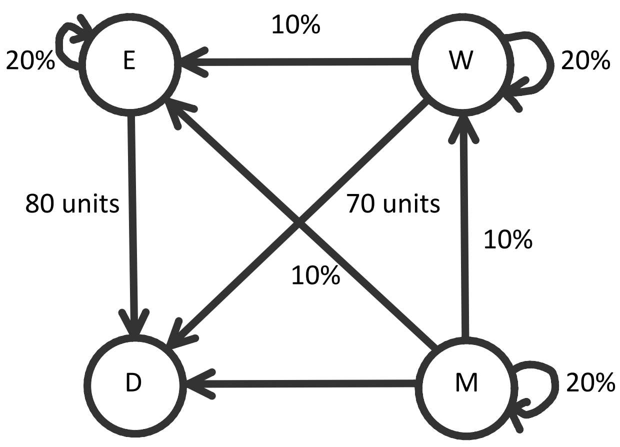
$$(I - C)\vec{x} = \vec{d}$$

If $(I - C)$ is invertible

Then $\vec{x} = (I - C)^{-1}\vec{d}$

Ex. 1 Revisited

Now suppose there is an external demand: what production level is required to satisfy a final demand of 80 units of E, 70 units of W, and 160 unites of M?



$$\vec{d} = \begin{pmatrix} 80 \\ 70 \\ 160 \end{pmatrix}$$

$$\therefore (I - C) = \frac{1}{10}\begin{pmatrix} 8 & 0 & 0 \\ -1 & 8 & 0 \\ -1 & -1 & 8 \end{pmatrix}$$

Solve $(I - C)\vec{x} = \vec{d}$

Solution:

The production level would be found by solving:

$$\vec{x} - C\vec{x} = \vec{d}$$
$$\vec{x} = (I - C)\vec{d}$$

$$\frac{1}{10}\begin{pmatrix} 8 & 0 & 0 \\ -1 & 8 & 0 \\ -1 & -1 & 8 \end{pmatrix}\vec{x} = \begin{pmatrix} 80 \\ 70 \\ 160 \end{pmatrix}$$

$$\begin{aligned} 8x_1 &= 800 \Longrightarrow x_1 = 100 \\ -x_1 + 8x_2 &= 700 \Longrightarrow x_2 = 100 \\ -x_1 \quad -x_2 + 8x_3 &= 1600 \Longrightarrow x_2 = 225 \end{aligned}$$

The output that balances demand with internal consumption is

$$\vec{x} = \begin{pmatrix} 100 \\ 100 \\ 225 \end{pmatrix}$$

The Importance of $(\boldsymbol{I} - \boldsymbol{C})^{-1}$

For the example above

$$(I - C)^{-1} \approx \begin{pmatrix} 1.25 & 0 & 0 \\ 0.15 & 1.25 & 0 \\ 0.18 & 0.17 & 1.25 \end{pmatrix}$$

The entries of $(I - C)^{-1} = B$ have this meaning: if the final demand vector $\vec{d}$ increases by one unit in the j^{th} place, the column vector b_j is the additional output required from other sectors.

So to meet an increase in demand for M by one unit, requires 1.25 of one additional units from M to meet internal consumption.

First/Odd demand: $\overrightarrow{d_0}$

$\overrightarrow{x_0} = (I - C)^{-1}\,\overrightarrow{d_0}$

New demand: $\overrightarrow{d_N} = \overrightarrow{d_0} + K\overrightarrow{e_1}$

$$\overrightarrow{x_N} = (I - C)^{-1}\overrightarrow{d_{w}} = \underbrace{(I - C)^{-1}\overrightarrow{d_0}}_{x_0} + \underbrace{K(I - C)^{-1} \quad \overrightarrow{e_1}}_{column\ i\ of\ (I-C)^{-1}}$$

Lecture 15

Monday, September 27, 2021 6:56 PM

Section 2.7: Computer Graphics

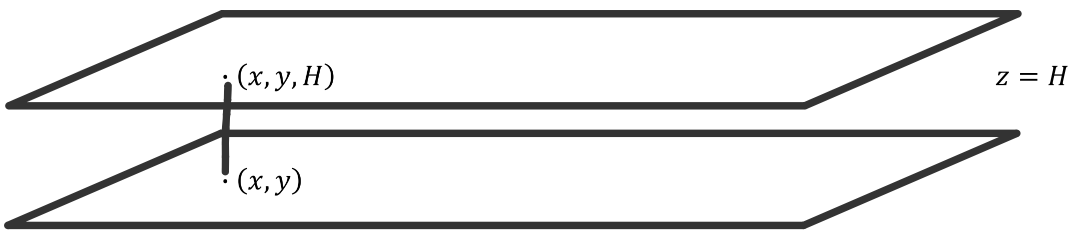
Homogeneous Coordinates
Translations of points in Rn does not correspond directly to a linear transform. Homogeneous coordinates are used model translations using matrix multiplication.

Homogeneous Coordinates in $\mathbb{R}^n$
Each point (x, y) in $\mathbb{R}^2$ can be identified with the point (x, y, H), $H \neq 0$, on the plane in $\mathbb{R}^3$ that lies H units above the xy-plane.

Note: we often we set $H = 1$.

Example: A translation of the form $(x, y) \to (x + h, y + k)$ can be represented as a matrix multiplication with homogeneous coordinates:

$$\begin{pmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x+h \\ y+k \\ 1 \end{pmatrix}$$



Homogeneous matrix of a 2D linear transform

$$\left(\begin{array}{cc|c} \multicolumn{2}{c|}{A} & 0 \\ & & 0 \\ \hline 0 & 0 & 1 \end{array}\right)$$, where A is a linear matrix in 2D

A Composite Transform with Homogeneous Coordinates
Triangle S is determined by three data points, $(1,1)$, $(2,4)$, $(3,1)$. Transform T rotates points by $\frac{\pi}{2}$ radians counterclockwise about the point $(0,1)$.

a) Represent the data with a matrix, D. Use homogeneous coordinates.
b) Use matrix multiplication to determine the image of S under T.
c) Sketch S and its image under T.

A. $D = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

B. First: translation of $\vec{u} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ $\quad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$

Second: rotation: $R = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Third: translation of $\vec{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ $\quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

↳ $T = BRA = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

Image of S under T is $TD = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & -3 & 0 \\ 2 & 3 & 4 \\ 1 & 1 & 1 \end{pmatrix}$

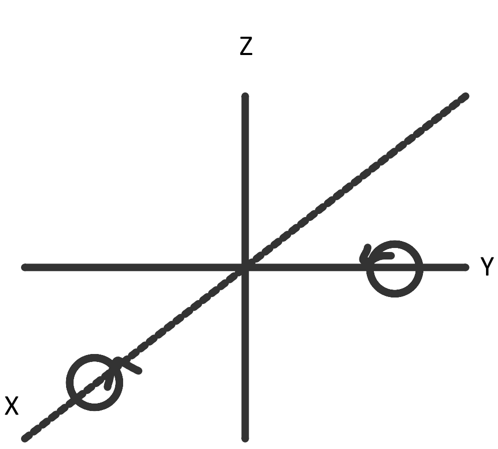
3D Homogenous Coordinates
Homogeneous coordinates in 3D are analogous to our 2D coordinates.

Homogeneous Coordinates in $\mathbb{R}^3$.
(X, Y, Z, H) are homogeneous coordinates for (x, y, z) in $\mathbb{R}^3$, $H \neq 0$, and

$$x = \frac{X}{H}, \qquad y = \frac{Y}{H}, \qquad z = \frac{Z}{H}$$

3D Transformation Matrices
Construct matrices for the following transformations.

a) A rotation in $\mathbb{R}^3$ about the y-axis by π radians.
b) A translation specified by the vector $\vec{p} = \begin{pmatrix} -2 \\ 3 \\ 4 \end{pmatrix}$



$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}$$

$$R_y(\theta) = \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix}$$

$$R_z(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Studio 10

Thursday, September 30, 2021 12:32 PM

Worksheet 2.6, 2.7, 2.8: The Leontif Input-Output Model, Computer Graphics, Subspaces of $\mathbb{R}^n$

Worksheet Exercises

1. An economy contains three sectors: X, Y, Z. For each unit of output,
   a. X requires .2 units from X, .1 units from Y, and .1 units from Z.
   b. Y requires 0 units from X, .2 units from Y, and .1 units from Z.
   c. Z requires 0 units from X, 0 units from Y, and .2 units from Z.

   Construct the consumption matrix for this economy. What production level is required to satisfy a final demand of 80 units of X, 60 units of Y, and 160 units of Z?

   $$C = \begin{pmatrix} .2 & 0 & 0 \\ .1 & .2 & 0 \\ .1 & .1 & .2 \end{pmatrix}$$

   $$\vec{x} = (I - C)^{-1} d = \begin{pmatrix} 100 \\ {}^{175}/_{2} \\ {}^{3575}/_{16} \end{pmatrix}$$

2. Rectangle S is determined by the data points, $(1,1), (3,1), (3,2), (1,2)$. Transform T reflects points through the line $y = 2 - x$
   a. Represent the data with a matrix, D. Use homogeneous coordinates.

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

   b. Use matrix multiplication to determine the image of S under T.

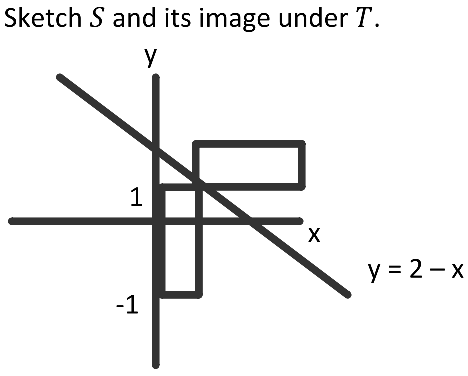
   $$T = T_2 R_f {T_2}^{-1}$$

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

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

   $$\therefore S \text{ under } T = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \\ 3 & -1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

   c. Sketch S and its image under T.



3. Transform $T_A = A\vec{x}$ rotatoes points in $\mathbb{R}^2$ about the point $(1,2)$. Construct a standard matrix for the transform using homogeneous coordinates. Leave your answer as a product of three matricies.

   $$T_A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

4. Construct the matrix for the transformation that performs a rotation in $\mathbb{R}^3$ about the x-axis by π radians.

   $$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

5. A has the reduced echelon form below. Construct a basis for Col$A$ and for Null$A$

   $$A = [\vec{a_1} \ \ \vec{a_2} \ \ \vec{a_3} \ \ \vec{a_4} \ \ \vec{a_5} \ \ \vec{a_6}] \sim \begin{bmatrix} 1 & 4 & 0 & 10 & 0 & 13 \\ 0 & 0 & 1 & -3 & 0 & -5 \\ 0 & 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Basis for Col$A$ $= \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$

Basis for Null$A$ $= \begin{bmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -10 \\ 0 \\ 3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -13 \\ 0 \\ 5 \\ 0 \\ -4 \\ 1 \end{bmatrix}$

Lecture 16

Wednesday, September 29, 2021 5:13 PM

Section 2.8: Subspaces of $\mathbb{R}^n$

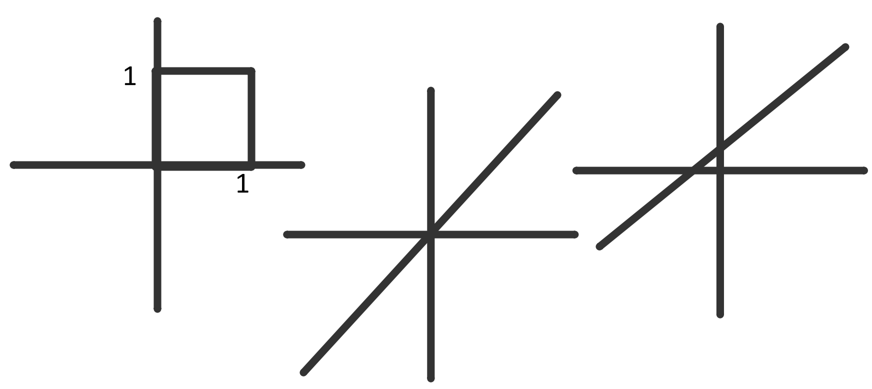
Subsets of $\mathbb{R}^n$
Definition
A subset of $\mathbb{R}^n$ is any collection of vectors that are in $\mathbb{R}^n$.

In $\mathbb{R}^2$: $\left\{\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}\pi\\e\end{pmatrix}\begin{pmatrix}\sqrt{3}/2\\1/2\end{pmatrix}\right\}$
unit circle $\{\vec{x} \in \mathbb{R}^2 | x_1 + x_2 = 1\}$

In $\mathbb{R}^3$: unit sphere $\{\vec{x} \in \mathbb{R}^3 | x_1 + x_2 + x_3 = 1\}$

Subspaces in $\mathbb{R}^n$
Definition
A subset H of $\mathbb{R}^n$ is a subspace if it is close under scalar multiplies and vector addition. That is: for any $c \in \mathbb{R}$ and for $\vec{u}, \vec{v} \in H$.
1. $c\vec{u} \in H$
2. $\vec{u} + \vec{v} \in H$

Note that condition 1 implies that the zero vector must be in H.
Ex.1: Which of the following subsets could be a subspace of $\mathbb{R}^2$



A) The unit square
NO
$(1/2 \quad 1/2) \in$ Unit Square

$7(1/2 \quad 1/2) \notin$ Unit Square

B) a line passing through the origin
YES

C) a line that doesn't pass through the origin
NO
$\vec{0} \notin \mathbb{R}^n$

Remark: $\{\vec{0}\}$ is the only bounded subspace of $\mathbb{R}^n$

The Column Space and the Null Space of a Matrix
Recall: for $\vec{v_1}, \dots \vec{v_p} \in \mathbb{R}^n$, that Span$\{\vec{v_1}, \dots \vec{v_p}\}$ is:
The set of all combinations of $\vec{v_1}, \dots \vec{v_p}$

This is a subspace, spanned by $\vec{v_1}, \dots \vec{v_p}$
Definition
Given an $m \times n$ matrix $A = [\vec{a_1} \quad \cdots \quad \vec{a_n}]$
1. The column space of A, Col A, is the subspace of $\mathbb{R}^m$ spaned by $\vec{a_1}, \dots \vec{a_n}$.
2. The null space of A, Null A, is the subspace of $\mathbb{R}^n$ spanned by the set of all vectors $\vec{x}$ that solve $A\vec{x} = \vec{0}$.

   $\vec{u}, \vec{v} \in$ Null $A, c \in \mathbb{R}$
   $A(c\vec{u}) = cA\vec{u} = \vec{0}$: $c\vec{u} \in$ Null A
   $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v}$: $\vec{u} + \vec{v} \in$ Null A

   $\Longrightarrow$Null A is a subspace.

Ex.

$$A = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} \quad \begin{matrix} Col(A) = Span\begin{pmatrix}1\\0\end{pmatrix} \\ Null(A) = Span\begin{pmatrix}0\\1\end{pmatrix}\end{matrix}$$

Ex.
Is $\vec{b}$ in the column space of A?

$$A = \begin{bmatrix}1 & -3 & -4\\-4 & 6 & -2\\-3 & 7 & 6\end{bmatrix} \sim \begin{bmatrix}1 & -3 & -4\\0 & -6 & -18\\0 & 0 & 0\end{bmatrix}, \qquad \vec{b} = \begin{pmatrix}3\\3\\-4\end{pmatrix}$$

$A\|\vec{b} = \left(\begin{array}{ccc\|c}1 & -3 & -4 & 3\\-4 & 6 & -2 & 3\\-3 & 7 & 6 & -4\end{array}\right)$	$R_2 \leftarrow R_2 + 4R_1$ ~ $R_3 \leftarrow R_3 + 3R_1$	$\left(\begin{array}{ccc\|c}1 & -3 & -4 & 3\\0 & 6 & -18 & 15\\0 & 2 & 6 & 5\end{array}\right)$	$R_3 \leftarrow R_3 - \frac{1}{3}R_1$ ~	$\left(\begin{array}{ccc\|c}1 & -3 & -4 & 3\\0 & 6 & -18 & 15\\0 & 0 & 0 & 0\end{array}\right)$

A$\vec{x} = \vec{b}$ is consistent: $\vec{b} \in Col\ A$

Remark: The third column of A is actually not needed here

Ex.2
Using the matrix on the previous slide: is $\vec{v}$ in the null space of A?

$$\vec{c} = \begin{pmatrix} -5\lambda \\ -3\lambda \\ \lambda \end{pmatrix}, \qquad \lambda \in \mathbb{R}$$

$$A\vec{v} = \begin{pmatrix} 1 & -3 & -4 \\ -4 & 6 & -2 \\ -3 & 7 & 6 \end{pmatrix} \begin{pmatrix} -5\lambda \\ -3\lambda \\ \lambda \end{pmatrix} = \lambda \begin{pmatrix} -5+9-4 \\ 18-18 \\ 0 \end{pmatrix} = \vec{0}$$

↳ $\vec{v} \in \text{Null}(A)$

Remark:

$$E\vec{v} = \begin{pmatrix} 1 & -3 & -4 \\ 0 & 6 & -18 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -5\lambda \\ -3\lambda \\ \lambda \end{pmatrix} = \lambda \begin{pmatrix} -5+9-4 \\ 18-18 \\ 0 \end{pmatrix} = \vec{0}$$

↳ $\vec{v} \in \text{Null}(E)$

Why: $A(\mid\vec{0}) \sim E(\mid\vec{0})$

→ $\mathrm{A}\vec{x} = \vec{0}$ and $\mathrm{E}\vec{x} = \vec{0}$ have the same solution set

Basis

Definition

A basis for a subspace H is a set of linearly independent vectors in H that span H.

Ex.

The set $H = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4 \,\middle|\, x_1 + 2x_2 + x_3 + 5x_4 = 0 \right\}$

a) H is a null space for what matrix λ?
b) Construct a basis for H.

a) $x_1 + 2x_2 + x_3 + 5x_4 = 0$

$$\underbrace{\begin{pmatrix} 1 & 2 & | & 5 \end{pmatrix}}_{A} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = 0$$

$H = \text{Null}(A)$

↳ This is a subspace of A

Studio 11

Thursday, September 30, 2021 12:31 PM

What is a subspace?

1. A subset S such that
   a. $\vec{0} \in S$
   b. If $\vec{a}, \vec{b} \in S$, then $\vec{a} + \vec{b} \in S$
   c. If $\vec{a} \in S$, then $c\vec{a} \in S$ for all $c \in S$
2. The span of some non-empty set of vectors
3. The column space of any matrix

Def: The dimension of a subspace S is the smallest # of vectors that span S.

$$Col\left(\underbrace{[\vec{u_1} \quad \cdots \quad \vec{u_n}]}_{A}\right) = Span(\vec{u_1}, \dots \vec{u_n})$$

$$Col(A) = \{Ax: \quad x \in \mathbb{R}^n\}$$
$$Null(A) = \{x: \quad Ax = 0\}$$

Worksheet 2.8, 2.9, Dimension and Rank

Worksheet Exercises

1. Construct a 3×3 matrix A with two pivotal columns, so that $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ is the null space of A.

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

   a. All 2×4 matricies have a non-trivial null space.
      True
   b. A 4×2 matrix with two pivot columns can have a non-trivial null space.
      False
   c. If the columns of a 6×6 matrix A are a basis ffor $\mathbb{R}^6$, the null space of A is the zero vector.
      True
2. A is a $n \times n$ matrix that has elements $a_{i,j}$ where

   $$a_{i,j} = \begin{cases} 0, when\ i + j\ is\ odd \\ 1, when\ i + j\ is\ even \end{cases}$$

   Suppose $n \geq 2$.

   a. What is the rank of A?
      2
   b. Give a basis for the column space of A.

   $$\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}$$

3. Which of the following, if any, are subspaces of $\mathbb{R}^3$? For those that are subspaces, what is the dimension of the subspace?
   a. $\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \middle| x_1 + x_2 = 4 \right\}$
      i. Not a subspace because $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ is not in the set
   b. $\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \middle| x_1 + x_2 + x_3 = 0, \quad x_1 + 2x_2 = 0 \right\}$
      i. 2
   c. $\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \middle| x_1 < x_2 < x_3 \right\}$
      i. Not a subspace because $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ is not in the set
   d. The null space of $A = \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 0 \end{pmatrix}$
      i. 1

Lecture 17

Friday, October 1, 2021 3:27 PM

<u>Notes:</u>

$$\vec{x}GH = \vec{x}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -2x_2 & -x_3 & -5x_4 \\ x_2 & & \\ & x_3 & \\ & & x_4 \end{pmatrix} = x_2\underbrace{\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}}_{\vec{v_1}} + x_3\underbrace{\begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}}_{\vec{v_2}} + x_4\underbrace{\begin{pmatrix} -5 \\ 0 \\ 0 \\ 1 \end{pmatrix}}_{\vec{v_3}}$$

$\Longrightarrow \{\vec{v_1},\ \vec{v_2},\ \vec{v_3}\}$ as a basis of H.

Ex.

Construct a basis for Null A and a basis for Col A

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

$\left\{\begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix},\ \begin{pmatrix} -1 \\ 2 \\ 5 \end{pmatrix}\right\}$ is a basis for Col A

$\vec{x} \in$ Null (A) if

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2x_2 \\ x_2 \\ 0 \\ x_4 \end{pmatrix} = x_2\begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$\left\{\begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix},\ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ is a basis for Null A

Additional Example

Let $V = \left\{\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 \middle| ab = 0\right\}$

1. Give an example of a vector in V. $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$
2. Give an example of a vector that is not in V. $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$
3. Is the zero vector in V? YES
4. Is V a subspace? $\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. No

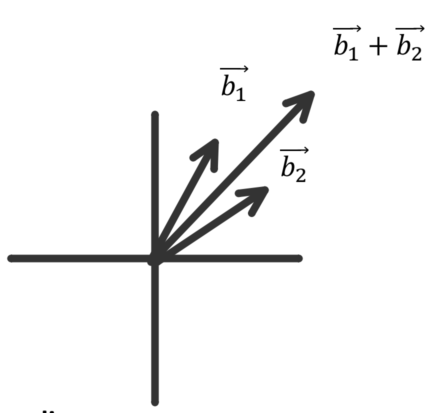
Section 2.9: Dimension and Rank

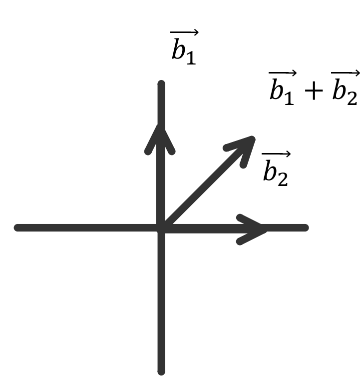
Choice of Basis

Key idea:

There are many possible choices of basis for a subspace. Our choice can give us dramatically different properties.

Example: sketch $\vec{b_1} + \vec{b_2}$ for the two different coordinate systems below.





Coordinates

Let $B = \{\vec{b_1}, \dots, \vec{b_p}\}$ be a basis for a subspace H. If $\vec{x}$ is in H, then the coordinates of $\vec{x}$ relative B are the weights (scalars) $c_1, \dots, c_p$ so that

$$\vec{x} = c_1\vec{b_1} + \cdots + c_p\vec{b_p}$$

And

$$[\vec{x}]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$$

Is the coordinate vector of $\vec{x}$ relative to B, or the B-coordinate vector of $\vec{x}$

Ex.1

Let $\vec{v_1} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\vec{v_2} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} 5 \\ 3 \\ 5 \end{bmatrix}$. Verify that $\vec{x}$ is in the span of $B = \{\vec{v_1},\ \vec{v_2}\}$, and calculate $[\vec{x}]_B$.

$\vec{x}$ is in the span of $\{\vec{v_1},\ \vec{v_2}\}$ if there exists c_1, c_2 such that $c_1\vec{v_1} + c_2\vec{v_2} = \vec{x}$

$$\left(\begin{array}{cc|c} 1 & 1 & 5 \\ 0 & 1 & 3 \\ 1 & 1 & 5 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 1 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right)$$

Thus, $\vec{x} = 2\vec{v_1} + 3\vec{v_2}$

$$[\vec{x}]_B = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

Dimension
Definition
The dimension (or cardinality) of a non-zero subspace H, dim H, is the number of vectors in a basis of H. We define $\dim\{0\} = 0$

Theorem
Any two choices of basis B_1 and B_2 of a non-zero subspace H have the same dimension.

Examples:
1. $\dim \mathbb{R}^n = n$
2. $H = \{(x_1, \dots, x_n) : x_1 + \cdots + x_n = 0\}$ has dimension $n - 1$
3. $\dim(Null\ A)$ is the number of free variables
4. $\dim(Col\ A)$ is the number of pivot variables

$$(1 \dots 1)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0, \quad \vec{x} = \begin{pmatrix} -x_2 & -x_3 & \cdots & -x_n \\ x_2 & & & \\ & x_3 & & \\ & & \ddots & \\ & & & x_n \end{pmatrix}$$

Proof: Assume $\#B_1 > \#x_2$

$$(v_1, \dots, v_m) \quad (v_1, \dots, v_n) \quad : n > m$$

$$\vec{v_1} = a_{11}\vec{v_1} + \cdots + a_{m1}\vec{v_m}$$
$$\vdots$$
$$\vec{v_n} = a_{11}\vec{v_1} + \cdots + a_{mn}\vec{v_n}$$

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$
$$\downarrow \qquad \downarrow$$
$$[\vec{v_1}]_{B_1} \qquad [\vec{v_n}]_{B_n}$$

Columns are linearly dependent
→ $\{\vec{v_1}, \dots, \vec{v_n}\}$ are linearly dependent
→ Not a basis

Rank
Definition
The rank of matrix A is the dimension of its column space.

Ex.2 Compute rank(A) and dim$(Null(A))$

$$\begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 4 & 7 & -4 & -3 & 9 \\ 6 & 9 & -5 & 2 & 4 \\ 0 & -9 & 6 & 5 & -6 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & 0 & 0 & 4 & -6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

3 pivot columns

Rank$(A) = 3$
dim$(Null(A)) = 2$

Rank, Basis, and Invertibility Theorems
Theorem (Rank Theorem)
If a matrix A has n columns, then Rank A + dim$(Nul\ A) = n$

Theorem (Basis Theorem)
Any two basis for a subspace have the same dimension

Theorem (Invertibility Theorem)
Let A be a $n \times n$ matrix. These conditions are equivalent
1. A is invertible.
2. The columns of A are a basis for $\mathbb{R}^n$.
3. Col A = $\dim(Col\ A) = n$.
4. Null A = $\{\vec{0}\}$.

Examples
If possible give an example of a 2×3 matrix A, that is in RREF and has the given properties.

a) Rank$(A) = 3$

   Not possible

b) Rank$(A) = 2$

$$A = \begin{pmatrix} 1 & 0 & * \\ 0 & 1 & * \end{pmatrix} \text{ or } \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

c) dim$(Null(A)) = 2$

$A = \begin{pmatrix} 1 & * & * \\ 0 & 0 & 0 \end{pmatrix}$ or $\begin{pmatrix} 0 & 1 & * \\ 0 & 0 & 0 \end{pmatrix}$ or $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

d) Null $A = \{\vec{0}\}$

Not possible

Lecture 18

Monday, October 4, 2021 3:20 PM

Notes:

Section 3.1: Introduction to Determinants

A Definition of the Determinant

Suppose A is $n \times n$ and has elements a_{ij}.

1. If $n = 1$, $A = [a_{1,1}]$, and has determinant $\det A = a_{1,1}$.
2. Inductive case: for $n > 1$,
   $$\det A = a_{1,1} \det A_{1,1} - a_{1,2} \det A_{1,2} + \cdots + (-1)^{1+n} a_{1,n} \det A_{1,n}$$
   where $A_{i,j}$ is the submatrix obtained by eliminating row i and columns j of A.

Example

$$A = \begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix} \Rightarrow A_{2,3} \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix}$$

Ex.1

$$\begin{pmatrix} \boxed{a} & - \\ | & d \end{pmatrix} \quad \begin{pmatrix} | & b \\ \boxed{c} & - \end{pmatrix}$$

Compute $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = a \det(A_{1,1}) - b \det(A_{1,3}) = ad - bc$$

Ex.2

Notation $\det(A) = |A|$

Compute $\det \begin{bmatrix} 1 & -5 & 0 \\ 2 & 4 & -1 \\ 0 & 2 & 0 \end{bmatrix} = \begin{vmatrix} 1 & -5 & 0 \\ 2 & 4 & -1 \\ 0 & 2 & 0 \end{vmatrix}$

$$\begin{vmatrix} \boxed{1} & - & - \\ | & 4 & -1 \\ | & 2 & 0 \end{vmatrix} \qquad \begin{vmatrix} - & \boxed{-5} & - \\ 2 & | & -1 \\ 0 & | & 0 \end{vmatrix} \qquad \begin{vmatrix} - & - & \boxed{0} \\ 2 & 4 & | \\ 0 & 2 & | \end{vmatrix}$$

$$\therefore |A| = 1 \times \det(A_{1,1}) - (-5)\det(A_{1,2}) + 0 \det(A_{1,3})$$

$$= \begin{vmatrix} 4 & -1 \\ 2 & 0 \end{vmatrix} + 5 \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix} + 0 = 2$$

Cofactor

Cofactors give us a more convenient notation for determinants.

Definition: Cofactor

The (i, j) cofactor of an $n \times n$ matrix A is
$$C_{i,j} = (-1)^{i+j} \det A_{i,j}$$

The pattern for the negative signs is

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix} \qquad \begin{aligned} \det(A) &= a_{1,1}\det(A_{1,1}) - a_{1,2}\det(A_{1,2}) + \cdots \\ &= a_{1,1}C_{1,1} + a_{1,2}C_{1,2} + a_{1,3}C_{1,3} + \cdots \end{aligned}$$

Theorem

The determinant of a matrix A can be computed down any row or column of the matrix. For instance, down the j^{th} column, the determinant is
$$\det(A) = a_{1,1}C_{1,1} + a_{1,2}C_{1,2} + a_{1,3}C_{1,3} + \cdots$$
This gives us a way to calculate determinants more efficiently

Ex.3

Compute the determinant of $\underbrace{\begin{bmatrix} 5 & 4 & 3 & 2 \\ 0 & 1 & 2 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & 3 \end{bmatrix}}_{A}$

$$|A| = 5C_{1,1} + 0C_{1,2} + 0C_{1,3} + 0C_{1,4}$$
$$= 5(-1)^{1+1} \begin{vmatrix} 1 & 2 & 0 \\ -1 & 1 & 0 \\ 1 & 1 & 3 \end{vmatrix}$$
$$= 5 \times 3 \times C_{3,3} = 15 \times (-1)^{3+3} \times \underbrace{\begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix}}_{1\times1-(-1)\times2=3} = 15 \times 1 \times 3 = 45$$

Triangular Matrices

Theorem

If A is a triangular matrix then
$$\det A = a_{1,1}a_{2,2}a_{3,3} \cdots a_{n,n}.$$

Ex.4
Compute the determinant of the matrix. Empty elements are zero.

$$\begin{bmatrix} 2 & 1 & & & & & \\ & 2 & 1 & & & & \\ & & 2 & 1 & & & \\ & & & 2 & 1 & & \\ & & & & 2 & 1 & \\ & & & & & 2 & 1 \\ & & & & & & 2 \end{bmatrix}$$

$|A| = 2^7 = 128$

Computational Efficiency
Note that computation of a co-factor expansion for an $N \times N$ matrix requires roughly $N!$ multiplications.
- A 10×10 matrix requires roughly $10! = 3.6$ million multiplications
- A 20×20 matrix requires $20! \approx 2.4 \times 10^{18}$ multiplications

Co-factor expansions may not be practical, but determinants are still useful.
- We will explore other methods for computing determinants that are more efficient.
- Determinants are very useful in multivariable calculus for solving certain integration problems.

Section 3.2: Properties of the Determinant

Row Operations
- We saw how determinants are difficult or impossible to compute with a cofactor expansion for large N.
- Row operations give us a more efficient way to compute determinants.

Theorem: Row Operations and the Determinant
Let A be a square matrix.
1. If a multiple of a row of A is added to another row produce B, then $\det B = \det A$.
2. If two rows are interchanged to produce B, then $\det B = -\det A$.
3. If one row of A is multiplied by a scalar k to produce B, then $\det B = -k \det A$.

$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad |A| = ad - bc$

1. $R_2 \leftarrow R_2 + kR_1$
   $\begin{vmatrix} a & b \\ c + ka & d + kb \end{vmatrix} = a(d + kb) - b(c + ka) = |A|$
2. $R_1 \leftrightarrow R_2$
   $\begin{vmatrix} c & d \\ a & b \end{vmatrix} = cb - ad = -|A|$
3. $R_1 \leftarrow kR_1$
   $\begin{vmatrix} ka & kb \\ c & d \end{vmatrix} = kad - kbc = k|A|$

Question: $A \in \mathbb{R}^{3\times 3}$ $\qquad |A| = 3$

$|2A| = 8|A| = 24$

$|I_3| = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \qquad |3I_3| = \begin{vmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{vmatrix} = 27$

Studio 12

Tuesday, October 5, 2021 12:33 PM

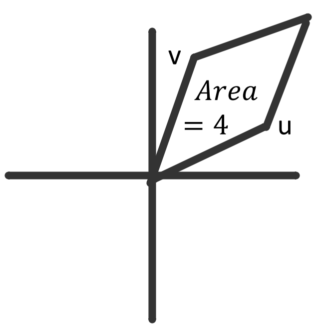
$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

$M_{2,3} = \begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix}$

<u>Determinants</u>

- Let A be a $n \times n$ matrix
- Let $M_{i,j}$ "(i,j) minor" be a matrix obtained by deleting the i^{th} row, and the j^{th} column
- Let C_i, j "(i,j) cofactor" is $(-1)^{i+j}\det(M_{i,j})$
- $\det(A) = a_{1,1}C_{1,1} + a_{1,2}C_{1,2} + \cdots + a_{1,n}C_{1,n}$
  $= a_{2,1}C_{2,1} + a_{2,2}C_{2,2} + \cdots + a_{n,2}C_{n,2}$

$A = [\vec{u} \quad \vec{v}]$



$\Longrightarrow |\det(A)| = 4$

Worksheet 3.1 to 3.3, Determinants

Worksheet Exercises

1. Discuss the computational efficiency of computing $\det(A)$ by cofactor expansion and by row operations. Which method is computationally better if A is a $n \times n$ and n is large? (Compare how many arithmetic operations it takes).
   For row operations, it would take N^3 steps to compute $\det(A)$. For cofactor expansion, you would need $N!$ Steps. Hence, when N is large, or even greater than ≈ 5.037, the row operations method would require less steps and would therefore be more computationaly efficient.
2. Use a determinant to identify all values of t and k such that the are the matricies are singular. Assume that t and k must be real numbers.
   a. $A = \begin{pmatrix} 3 & 5 \\ 5 & 3 \end{pmatrix} - tI_2$
      i. $\det A = (3-t)^2 - 25 = 0 \Rightarrow t^2 - 6t - 16 = 0 \Longrightarrow (t-8)(t+2) = 0 \Rightarrow t = -8, 2$
   b. $B = \begin{pmatrix} 0 & 1 & t \\ -3 & 10 & 0 \\ 0 & 5 & k \end{pmatrix}$
      i. $\det B = 3\begin{vmatrix} 1 & t \\ 5 & k \end{vmatrix} = 0 \Rightarrow 3k - 15t = 0 \Rightarrow k = 5t$
3. Let $[\vec{a} \quad \vec{b} \quad \vec{c} \quad \vec{d}]$ be a 4×4 matrix whose determinant is equal to 2. Compute the value of the determinant $[\vec{d} \quad \vec{b} \quad 3\vec{c} \quad \vec{a}]$.
   $[\vec{a} \quad \vec{b} \quad \vec{c} \quad \vec{d}] \Longrightarrow \det = 2$
   $[\vec{d} \quad \vec{b} \quad 3\vec{c} \quad \vec{a}] \Longrightarrow \det = 2(-1)(3) = -6$
4. R is the parallelogram determined by $\overrightarrow{p_1} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ and $\overrightarrow{p_2} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ If $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$, what is the area of the image of R under the map $\vec{x} \mapsto A\vec{x}$?
   a. $|\det R| = \left|\begin{vmatrix} 3 & 2 \\ 4 & 2 \end{vmatrix}\right| = |6 - 8| = 2$
   b. $\det A = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = 1 + 1 = 2$
   c. Therefore the area of the image of R under the map $\vec{x} \mapsto A\vec{x}$ is $2 \times 2 = 4$
5. $T_A = A\vec{x}$, where $A \in \mathbb{R}^{2\times 2}$, is a linear transformation that first rotates vectors in $\mathbb{R}^2$ counterclockwise by θ radians about the origin, then reflects them through the line $x_1 = x_2$. By inspection, what is the value of the determinant of A? You should compute its value to check your answer.
   a. $\det A = \det F \times \det R = (-1)(1) = -1$

Lecture 19

Wednesday, October 6, 2021   2:41 PM

Notes:

Ex.1

Compute $\begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix}$

$$|A| = \begin{matrix} R_2 \leftarrow R_2 + 2R_1 \\ \\ R_3 \leftarrow R_3 + R_1 \end{matrix} \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{vmatrix} R_2 \leftrightarrow R_1 \begin{vmatrix} 0 & 0 & -5 \\ 1 & -4 & -2 \\ 0 & 3 & 2 \end{vmatrix} = (-1) \times 1 \times 3 \times (-5) = 15$$

Invertibility

Important practical implication: If A is reduced to echelon form, by r interchanges of rows and columns, then

$$|A| = \begin{cases} (-1)^r \times (product\ of\ pivots), when\ A\ is\ invertible \\ 0, \qquad\qquad\qquad\qquad\quad , when\ A\ is\ singular \end{cases}$$

$$A \xrightarrow{E_1} \xrightarrow{E_2} \cdots \xrightarrow{E_k} U = \begin{pmatrix} x & \cdots & x \\ & \ddots & \vdots \\ & & x \end{pmatrix}$$

Ex.2 Compute the determinant

$$|A| = \begin{vmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 1 & 2 \end{vmatrix} \begin{matrix} R_4 \leftarrow R_4 + R_2 \\ \sim \\ R_2 \leftrightarrow R_1 \end{matrix} (-1) \begin{vmatrix} 2 & 5 & -7 & 3 \\ 0 & 1 & 2 & -1 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 5 \end{vmatrix} \begin{matrix} R_3 \leftarrow R_3 - 3R_2 \\ \sim \\ R_3 \leftrightarrow R_4 \end{matrix} (-1)(-1) \begin{vmatrix} 2 & 5 & -7 & 3 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -3 & 5 \\ 0 & 0 & 0 & 5 \end{vmatrix}$$

$$= 2 \times 1 \times (-3) \times 5 = -30$$

Properties of the Determinant

For any square matrices A and B, we can show the following.

1. $\det A = \det A^T$.
2. A is invertible if and only if $\det A \neq 0$.
3. $\det(AB) = \det A \cdot \det B$.

2. and 3. combined: if A is invertbile, then

$$|A^{-1}| = \left|\frac{1}{A}\right|$$

$$|A||A^{-1}| = |I_n| = 1$$

Proof of (1):

$$A \xrightarrow{E_1} \xrightarrow{E_2} \cdots \xrightarrow{E_k} U = \begin{pmatrix} x & \cdots & x \\ & \ddots & \vdots \\ & & x \end{pmatrix}$$

"form": $R_3 \leftarrow R_3 + 2R_1$

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}, \qquad E^T = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$|E| = 1 = |E^T|$$

"form": $R_3 \leftrightarrow R_1$

$$E = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = E^T$$

Finally $U = E_k \ldots E_2 E_1 A$

Using 3. $|U| = |E_k| \ldots |E_2||E_1||A|$

$U^T = {E_k}^T \ldots {E_2}^T {E_1}^T A^T$

$|U^T| = |{E_k}^T| \ldots |{E_2}^T||{E_1}^T||A^T|$

$\Longrightarrow |A| = |A^T|$ *we can do elementary row operations

Additional Example

Use a determinant to find all values of λ such that matrix C is not invertible.

$$C = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} - \lambda I_n$$

$$C - \lambda I_n = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} 5-\lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{pmatrix}$$

$C - \lambda I_n$ is not invertible if $|C - \lambda I_n| = 0$

$$\begin{vmatrix} 5-\lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = (5-\lambda)\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = (5-\lambda)(\lambda^2 - 1) = (5-\lambda)(1-\lambda)(\lambda+1);$$

Hence, $C - \lambda I_n$ is not invertible if $\lambda = \pm 1 \vee 5$

Additional Example
Determine the value of

$$\det A = \det\left(\overbrace{\begin{pmatrix} 0 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{pmatrix}}^{B}{}^{8}\right)$$

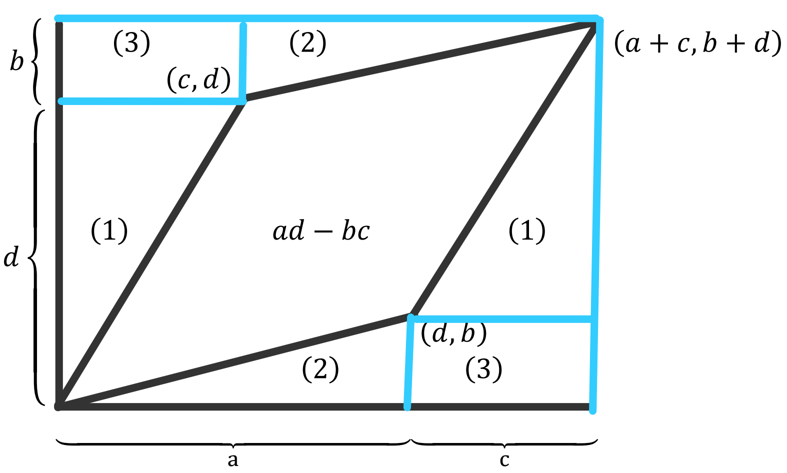
By the property of 3: $|AB| = |B||A|$
$\quad \rightarrow |B^8| = |B|^8$

$$|B| = \begin{vmatrix} 0 & 2 & 0 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{vmatrix} = -2\begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} = -2$$

$\therefore |A| = |B|^8 = (-2)^8 = 256$

Section 3.3: Volume, Linear Transformation

Determinants, Area and Volume
In $\mathbb{R}^2$, determinants give us the area of a parallelogram.



Area of parallelogram $= \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$

Area $= Area(Large\ Rectangle) - 2\big(Area(1) + Area(2) + Area(3)\big) = (a+c)(b+d) - 2\left(\frac{1}{2}cd + \frac{1}{2}ab + bc\right)$
$\quad = ad + ab + cd + bc - cd - ab - 2bc = ad - bc$
$\quad = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$

Studio 13

Thursday, October 7, 2021 12:31 PM

Practice Worksheet

1. Find the determinant of $\begin{bmatrix}1 & -2 & 0\\ 0 & 0 & 3\\ 0 & 1 & 0\end{bmatrix}$ via row reducation

   $$\det\begin{bmatrix}1 & -2 & 0\\ 0 & 0 & 3\\ 0 & 1 & 0\end{bmatrix} \sim -\det\begin{bmatrix}1 & -2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 3\end{bmatrix} = -(3) = -3$$

2. Let $A \in \mathbb{R}^{n\times n}$ be the matrix whose (i,j) entry is $\min\{i,j\}$. Find $\det A$
   a. Testing for when $n = 1,2,3,4$
   b. $\det[1] = 1$
   c. $\det\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix} = \det\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix} = 1$
   d. $\det\begin{bmatrix}1 & 1 & 1\\ 1 & 2 & 2\\ 1 & 2 & 3\end{bmatrix} = \det\begin{bmatrix}1 & 1 & 1\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix} = 1$
   e. $\det\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 2 & 2 & 2\\ 1 & 2 & 3 & 3\\ 1 & 2 & 3 & 4\end{bmatrix} = \det\begin{bmatrix}1 & 1 & 1 & 1\\ 0 & 1 & 1 & 1\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 1\end{bmatrix} = 1$
   f. $\therefore \det A = 1$
3. Find the area of the triangle $\begin{bmatrix}2\\3\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix}, \begin{bmatrix}1\\4\end{bmatrix}$
   a. $\begin{vmatrix}2 & 3 & 1\\ 3 & 4 & 4\end{vmatrix}$
   b. Substracting by $\begin{vmatrix}2\\3\end{vmatrix}$ so that the origin is at $(0,0)$
   c. $\begin{vmatrix}0\\0\end{vmatrix}\begin{vmatrix}1 & -1\\ 1 & 1\end{vmatrix}$
   d. Hence, $Area = \frac{1}{2}\det\begin{vmatrix}1 & -1\\ 1 & 1\end{vmatrix} = \frac{1}{2}(2) = 1$
4. Suppose $A \in \mathbb{R}^{2\times 2}$, the entries in A are integers, and $\det A = 1$. Then the entries of A^{-1} are also integers.
   a. True
5. $A \in \mathbb{R}^{n\times n}$ is one-to-one if and only if $\det A = 0$.
   a. False
6. A matrix $A \in \mathbb{R}^{2\times 2}$ maps regions of area 1 to regions of area 2 if and only if $\det A = 2$
   a. False
7. Suppose $A \in \mathbb{R}^{n\times n}$ has a 0 diagonal. Then $\det A = 0$.
   a. False
8. Suppose $A \in \mathbb{R}^{n\times n}$ and $Col(A)$ is spanned by $n - 1$ vectors. Then $\det A = 0$
   a. True
9. Complete the sentence "The more cheese, the more holes. The more holes, the less cheese. Therefore,
   a. Cheese $\propto$ # holes

Lecture 20

Friday, October 8, 2021 3:04 PM

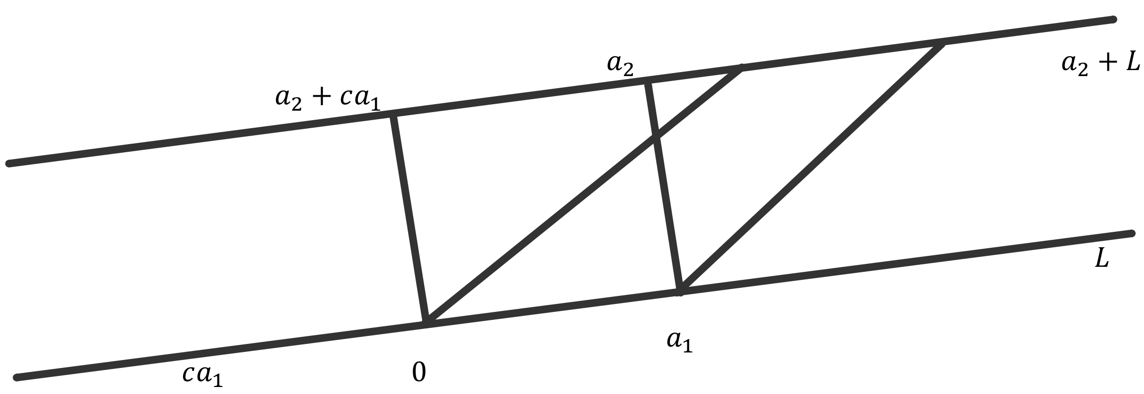
Notes:

Determinants as Area or Volume

Theorem

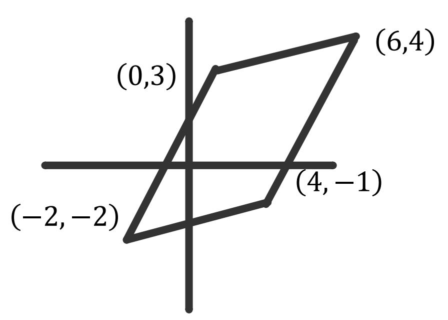
The volume of the parallel piped spanned by the columns of an $n \times n$ matrix A is $|\det A|$.

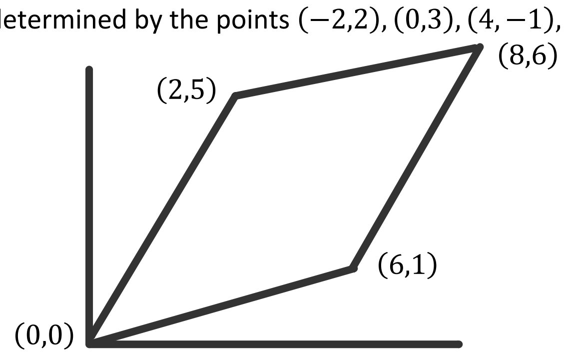
Key Geometric Fact (which works in any dimension). The area of the parallelogram spanned by two vectors $\vec{a}$, $c\vec{a} + \vec{b}$, for any scalar c.



Ex.1

Calculate the area of the parallelogram determined by the points $(-2,2), (0,3), (4,-1), (6,4)$





Area $= \left| |\vec{u} \quad \vec{v}| \right| = \left| \begin{vmatrix} 2 & 6 \\ 5 & 1 \end{vmatrix} \right| = |-28| = 28$

Linear Transformations

Theorem

If $T_A\colon \mathbb{R}^n \mapsto \mathbb{R}^n$, and S is some parallelogram in $\mathbb{R}^n$, then

$\quad \text{volume}(T_A(S)) = |\det(A)| \cdot \text{volume}(S)$

$B \in \mathbb{R}^{n\times n} \to S$: parallelogram spanned by the columns of B

$\text{Volume}\left(\overbrace{T_A(S)}^{AB}\right) = |\det A| \cdot \text{Volume}(S) = |\det A| \cdot |\det B|$

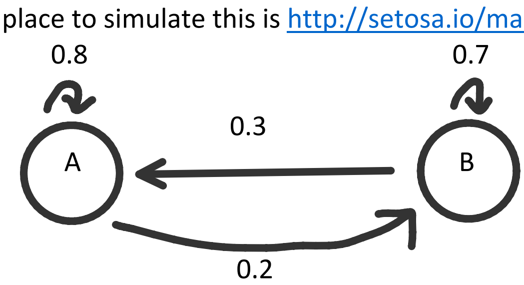
$\quad |\det AB| = |\det A| \cdot |\det B|$

Section 4.9: Applications to Markov Chains

Ex.1

- A small town has two libraries, A and B.
- After 1 month, among the books checked out of A.
  - 80% returned to A
  - 20% returned to B
- After 1 month, among the books checked out of B.
  - 30% returned to A
  - 70% returned to B

If both libraries have 1000 books today, how many books does each library have after 1 month? After one year? After n month? A place to simulate this is http://setosa.io/markov/index.html



Ex.1 Continued

The books are equally divided by between the two branches, denoted by $\vec{x_0} = \begin{bmatrix} .5 \\ .5 \end{bmatrix}$. What is the distribution after 1 month, call it $\vec{x_1}$? After two month?

$$\vec{x_1} = \begin{pmatrix} 0.8x_0A & 0.3x_0B \\ 0.2x_1A & 0.7x_1B \end{pmatrix} = \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix} \vec{x_0} = \begin{pmatrix} 0.55 \\ 0.45 \end{pmatrix}$$

$$\vec{x_2} = \begin{pmatrix} 0.8x_0A & 0.3x_0B \\ 0.2x_1A & 0.7x_1B \end{pmatrix} = P\vec{x_1} = P(P\vec{x_0}) = P^2\vec{x_0}$$

After k months, the distribution is $\vec{x_k}$, which is what in terms of $\vec{x_0}$?

After k months:

$\overrightarrow{x_k} = P^k \overrightarrow{x_0}$

Markov Chains
A few definitions:

- A probability vector is a vector, $\vec{x}$, with non-negative elements that sum to 1.
  - Ex: $\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1/2\\1/2\end{pmatrix}, \begin{pmatrix}1/3\\1/3\\1/3\end{pmatrix}, \begin{pmatrix}1/2\\0\\1/4\\1/4\end{pmatrix}, \ldots$
- A stochastic matrix is a square matrix, P, whose columns are probabiltiy vectors.
- A Markov chain is a sequence of probability vectors, $\overrightarrow{x_k}$, and a stochastic matrix P, such that:
  $\overrightarrow{x_{k+1}} = P\overrightarrow{x_k}, \qquad k = 0,1,2,\ldots$
- A steady-state vector for P is a probability vector $\vec{q}$ such that $P\vec{q} = \vec{q}$

Ex.2
Determine a steady-state vector for the stochastic matrix

$$\begin{pmatrix}.8 & .3\\ .2 & .7\end{pmatrix}$$

Steady-state vector: $P\vec{q} = \vec{q}$

$P\vec{q} - \vec{q} = 0$

$(P - I)\vec{q} = \vec{0}$

$\left(\begin{array}{cc|c}-0.2 & 0.3 & 0\\ 0.2 & -0.3 & 0\end{array}\right) \sim \left(\begin{array}{cc|c}-2 & 3 & 0\\ 0 & 0 & 0\end{array}\right)$: 1 free variable; line of solution

Example: $\vec{x} = \begin{pmatrix}3\\2\end{pmatrix}$ steady-state vector: $\frac{1}{5}\begin{pmatrix}3\\2\end{pmatrix} = \begin{pmatrix}0.6\\0.4\end{pmatrix}$

After a long time: 60% of books in A
40% of books in B

Convergence
We often want to know what happens to a process,

$\overrightarrow{x_{k+1}} = P\overrightarrow{x_k}, \qquad k = 0,1,2,\ldots$

as $k \to \infty$

Definition: a stochastic matrix P is regular if there is some k such that P^k only contains strictly positive entries.

Theorem
If P is a regular stochastic matrix, then P has a unique steady-state vector $\vec{q}$, and $\overrightarrow{x_{k+1}} = P\overrightarrow{x_k}$ converges to $\vec{q}$ as $k \to \infty$.

Ex: $P = \begin{pmatrix}0.8 & 0.3\\ 0.2 & 0.7\end{pmatrix} \to$ regular

$P = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} \to P^k$, for any $k \to$ not regular

$P = \begin{pmatrix}1 & 1/3\\ 0 & 2/3\end{pmatrix} \quad P^2 = \begin{pmatrix}1 & *\\ 0 & *\end{pmatrix} \ldots$ not regular

$P = \begin{pmatrix}1/3 & 0 & 1/4\\ 1/3 & 1 & 1/2\\ 1/3 & 0 & 1/4\end{pmatrix} = P^k = \begin{pmatrix}* & 0 & *\\ * & 1 & *\\ * & 0 & *\end{pmatrix} \to$ not regular

$P = \begin{pmatrix}1/2 & 1/3 & 1/5\\ 1/2 & 1/3 & 2/5\\ 0 & 1/3 & 2/5\end{pmatrix} \to$ regular claim: P^2 has no zero entry

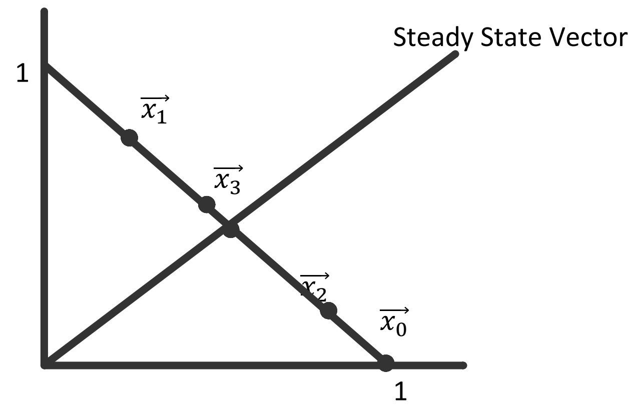
Lecture 21

Wednesday, October 13, 2021 3:33 PM

Notes:

Stochastic Vectors in the Plane

The stochastic vectors in the plane are the line segment below, and a stochastic matrix maps stochastic vectors to themselves. Iterates $P^k\overrightarrow{x_0}$ converge to the steady state.



$\overrightarrow{x_\infty} = \vec{q}$: Steady-state vector

$P^k \to [\overrightarrow{x_\infty} \quad \overrightarrow{x_\infty}]; k \to +\infty$

$P^k \underbrace{\to}_{k\to+\infty} (\vec{q} \quad \vec{q})$

From the previous theorem"

For any $\overrightarrow{x_0}$: $\overrightarrow{x_k}$: $P^k\overrightarrow{x_0} \underbrace{\to}_{k\to+\infty} \vec{q}$

Take $\overrightarrow{x_0} = \overrightarrow{R_1} = \begin{pmatrix}1\\0\end{pmatrix}$

$\overrightarrow{x_k} = \underbrace{P^k\overrightarrow{x_0}}_{P^k\overrightarrow{e_1}} \underbrace{\to}_{k\to+\infty} \underbrace{\vec{q}}_{1^{st}\ column\ of\ e_1}$

Do the same thing with $\overrightarrow{x_0} = \overrightarrow{e_2}$: 2^{nd} column of P^k

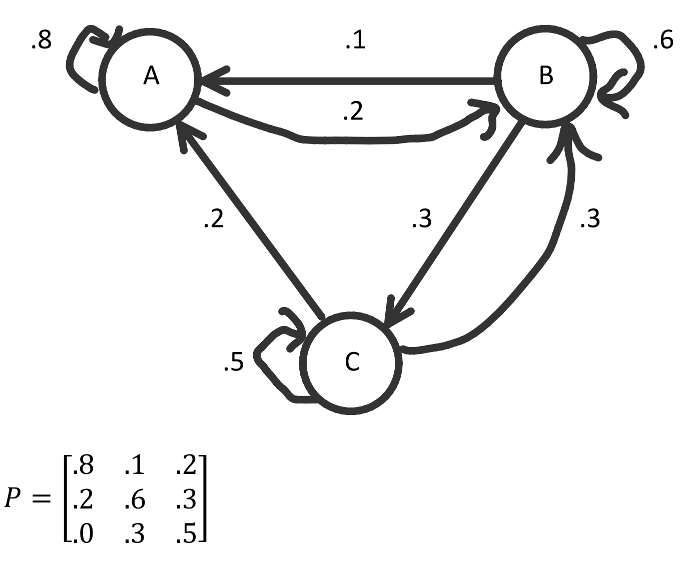
Ex.3

A car rental company has 3 rental locations, A, B, and C. Cars can be returned at any location. The table below gives the pattern of rental and returns for a given week.

	Rented from		
Returned to	A	B	C
A	.8	.1	.2
B	.2	.6	.3
C	.0	.3	.5

There are 10 cars at each location today.

a) Construct a stochastic matrix, P, for this problem.
b) What happens to the distribution of cars after a long time? You may assume that P is regular.



$$P = \begin{bmatrix} .8 & .1 & .2 \\ .2 & .6 & .3 \\ .0 & .3 & .5 \end{bmatrix}$$

Is P regular? YES

p^2 has no 0 entries

$P\vec{q} = \vec{q}$: $(P - I)\vec{q} = 0$

$$\left(\begin{array}{ccc|c} -0.2 & 0.1 & 0.2 & 0 \\ 0.2 & -0.4 & 0.3 & 0 \\ 0 & 0.3 & -0.5 & 0 \end{array}\right) \underset{\sim}{R_2 \leftarrow R_2 + R_1} \left(\begin{array}{ccc|c} -2 & 1 & 2 & 0 \\ 0 & -3 & 5 & 0 \\ 0 & 3 & -5 & 0 \end{array}\right) \underset{\sim}{R_3 \leftarrow R_3 + R_2} \left(\begin{array}{ccc|c} -2 & 1 & 2 & 0 \\ 0 & -3 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \underset{\sim}{R_1 \leftarrow 3R_1 + R_2}$$

$$\left(\begin{array}{ccc|c} -6 & 0 & 11 & 0 \\ 0 & -3 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

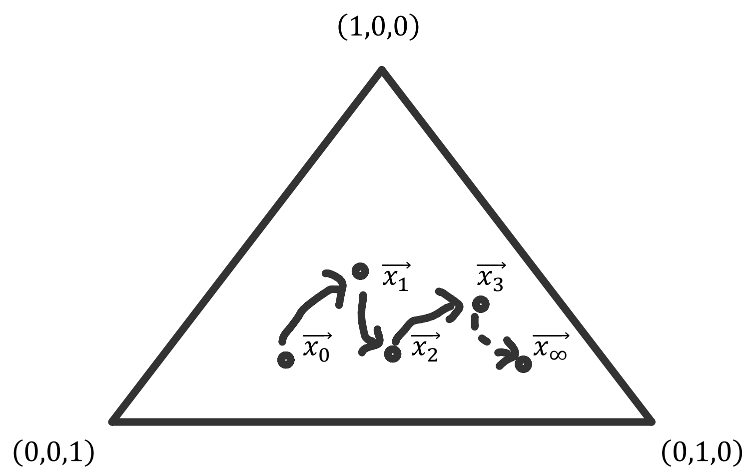
$$\begin{cases} x_1 = \dfrac{11}{6}x_2 \\ x_2 = \dfrac{5}{3}x_3 \end{cases}$$

Take $x_3 = 6$: $\vec{x} = \begin{pmatrix} 11 \\ 10 \\ 6 \end{pmatrix}$

Then: $\boxed{\vec{q} = \frac{1}{27}\begin{pmatrix} 11 \\ 10 \\ 6 \end{pmatrix}}$

Solution to $P\vec{x} = \vec{x}$: $\vec{x} = c\begin{pmatrix} 11 \\ 10 \\ 6 \end{pmatrix}$, $c \in \mathbb{R}$

The stochastic vectors in $\mathbb{R}^3$, are vectors $\begin{bmatrix} s \\ t \\ 1-s-t \end{bmatrix}$, where $0 \le s, t \le 1$ and $s + t \le 1$. P 'contracts' stochastic vectors to x_∞.



Section 5.1: Eigenvectors and Eigenvalues

Eigenvectors and Eigenvalues

If $A \in \mathbb{R}^{n\times n}$, and there is a $\vec{v} \neq 0$ in $\mathbb{R}^n$, and

$A\vec{v} = \lambda\vec{v}$

Then $\vec{v}$ is an eigenvector for A, and $\lambda \in \mathbb{C}$ is the corresponding eigenvalue.

Note that

- We will only consider square matrices.
- If $\lambda \in \mathbb{R}$, then
  - when $\lambda > 0$, $A\vec{v}$ and $\vec{v}$ point in the same direction
  - when $\lambda < 0$, $A\vec{v}$ and $\vec{v}$ point in opposite directions
- Even when all entries of A and $\vec{v}$ are real, λ can be complex (a rotation of the plane has no real eigenvalues.)
- We explore complex eigenvalues in Section 5.5.

Ex.1

Which of the following are eigenvectors of $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$? What are the corresponding eigenvalues?

a) $\vec{v_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} A\vec{v_1} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2\vec{v_1}$
   a. $\vec{v_1}$ is an eigenvector the eigenvalue 2.
b) $\vec{v_2} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} A\vec{v_2} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0\vec{v_2}$
   a. $\vec{v_2}$ is an eigenvector the eigenvalue 0.
c) $\vec{v_3} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} A\vec{v_3} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
   a. $\vec{v_3}$ is not an eigenvector. (it is $\vec{0}$)
d) $\vec{v_4} = \begin{pmatrix} k \\ k \end{pmatrix} A\vec{v_4} = \begin{pmatrix} 2k \\ 2k \end{pmatrix} = 2\vec{v_4}$
   a. $A\vec{v_4} = A(k\vec{v_1}) = kA\vec{v_1} = k(2\vec{v_1}) = 2\vec{v_4}$
   b. If $\vec{v}$ is an eigenvector for λ, so is $k\vec{v}$ for any $k \neq 0$.
e) $\vec{v_5} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} A\vec{v_5} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} =? \lambda\begin{pmatrix} 2 \\ 0 \end{pmatrix}$
   a. $\vec{v_5}$ is not an eigen vector
      i. $(\Delta\vec{v_5} = \vec{v_1} + \vec{v_2})$

Ex.2

Confirm that $\lambda = 3$ is an eigenvalue of $A = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix}$

We look for $\vec{v} \neq 0$ such that $A\vec{v} = \lambda\vec{v}$

$A\vec{v} = 3\vec{v}$

$(A - 3I)\vec{v} = \vec{0}$

$A - 3I = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix} - \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} -1 & -4 \\ -1 & -4 \end{pmatrix}$ singular

Then there exists $\vec{v} \neq 0$ such that $(\mathrm{A} - 3\mathrm{I})\vec{v} = \vec{0}$

Meaning: $A\vec{v} = 3\vec{v}$

Eigenspace

Definition

Suppose $A \in \mathbb{R}^{n\times n}$. The eigenvectors for a given λ span a subspace of $\mathbb{R}^n$ called the λ-eigenspace of A.

Note: the λ-eigenspace for matrix A is $\mathrm{Nul}(A - \lambda I) = \{eigenvectors\} \cup \{\vec{0}\}$

Ex.3

Construct a basis for the eigenspaces for the matrix whose eigenvalues are given, and sketch the eigenvectors.

$\begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix} \quad \lambda = -1, 2$

$A\vec{x} = \lambda\vec{x} \Leftrightarrow (A - \lambda I)\vec{x} = \vec{0}$

$$\Longleftrightarrow \vec{x} \in \underbrace{\mathrm{Nul}(\vec{x} - \lambda I)}_{\lambda\text{−eigenspace}}$$

Theorems

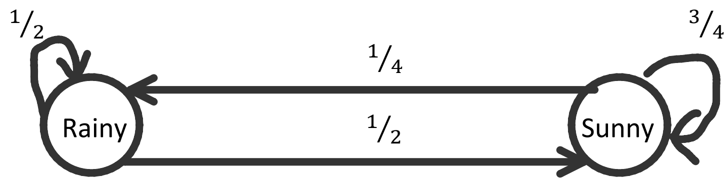
Proofs for the most these theorems are in Section 5.1. If time permits, we will explain or prove all/most of these theorems in lecture.

1. The diagonal elements of a triangular matrix are its eigenvalues.
2. A invertible $\Longleftrightarrow$ 0 is not an eigenvalue of A.
3. Stochastic matrices have an eigenvalue equal to 1.
4. If $\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_k}$ are eigenvectors that correspond to distinct eigenvalues, then $\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_k}$ are linearly independent.

Studio 14

Thursday, October 14, 2021 12:25 PM

Def: a matrix is regular if P^n has all positive entries for some n.



- What's the probability of being sunny in many many days? $^2/_3$

  Transition matrix $P = \begin{array}{c c} & \begin{array}{cc} R & S \end{array} \\ \begin{array}{c} R \\ S \end{array} & \begin{bmatrix} ^1/_2 & ^1/_4 \\ ^1/_2 & ^3/_4 \end{bmatrix} \end{array}$
- What's the probability of it being sunny in 2 days given that it is rainy today?
  P^2_{12}
- Steady state: a vector x such that $Px = x$
  $\begin{bmatrix} -^1/_2 & ^1/_4 \\ ^1/_2 & -^1/_4 \end{bmatrix} \sim \begin{bmatrix} 1 & -^1/_2 \\ 0 & 0 \end{bmatrix}$
  $\Rightarrow x_1 = \frac{1}{2} x_2 \Rightarrow x_2 = 2x_1$
  $x_1 + x_2 = 1$
  $\Rightarrow 3x_1 = 1 \Rightarrow x_1 = ^1/_3 \Rightarrow x_2 = ^2/_3$

Steady-state vector: $\begin{pmatrix} ^1/_3 \\ ^2/_3 \end{pmatrix}$

Worksheet 4.9, Applications to Markov Chains

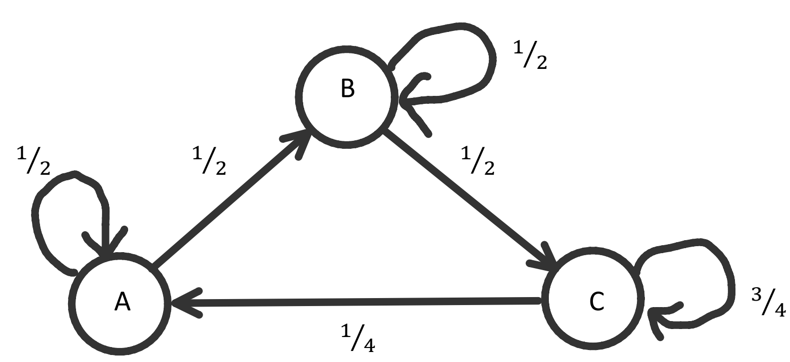
Worksheet Exercises

1. Determine whether P and Q are a regular stochastic matricies.

   $P = \begin{pmatrix} .8 & 0 \\ .2 & 1 \end{pmatrix} \qquad Q = \frac{1}{4}\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 3 & 1 & 2 \end{pmatrix}$

   P: $\begin{pmatrix} .8 & 0 \\ .2 & 1 \end{pmatrix}\begin{pmatrix} .8 & 0 \\ .2 & 1 \end{pmatrix} = \begin{pmatrix} .64 & \boxed{0} \\ .35 & 1 \end{pmatrix}$ NO

   Q: $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 3 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 4 \\ 6 & 3 & 6 \\ 9 & 9 & 6 \end{pmatrix}$ YES
2. Consider the Markov chain below



   What is the transition matrix? Calculate the steady-state vector.

   Transition Matrix:
   $\begin{bmatrix} ^1/_2 & 0 & ^1/_4 \\ ^1/_2 & ^1/_2 & 0 \\ 0 & ^1/_2 & ^3/_4 \end{bmatrix}$

   Steady-state vector calculation:
   $(P - I)\vec{x} = \vec{0}$
   $\begin{bmatrix} -^1/_2 & 0 & ^1/_4 & 0 \\ ^1/_2 & -^1/_2 & 0 & 0 \\ 0 & ^1/_2 & -^1/_4 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} -^1/_2 & 0 & ^1/_4 & 0 \\ 0 & -^1/_2 & ^1/_4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

   $-^1/_2 x_1 + ^1/_4 x_3 \Rightarrow x_1 = ^1/_2 x_3$
   $-^1/_2 x_2 + ^1/_4 x_3 \Rightarrow x_2 = ^1/_2 x_3$

   $\vec{x} = x_3 \begin{bmatrix} ^1/_2 \\ ^1/_2 \\ 1 \end{bmatrix}$

   $x_3 = 2$: $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$

   $\boxed{\therefore \vec{q} = \begin{bmatrix} ^1/_4 \\ ^1/_4 \\ ^1/_2 \end{bmatrix}}$

3. Suppose there are two cities, X and Y. Every year,

- 70% of people from X stay in X, the remaining 30% move to Y.
- 40% of people from Y stay in Y, the remaining 60% move to X.

The initial populations of X and Y are 100 and 200, respectively.

a. What is the stochastic matrix that represents this situation?

$P = \begin{bmatrix} 0.7 & 0.6 \\ 0.3 & 0.4 \end{bmatrix}$

b. After a long period of time, what is the population in city X?

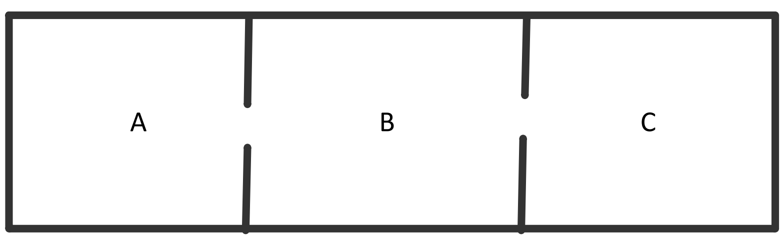
$(P - I)\vec{x} = \vec{0}$

$\begin{bmatrix} 0.7 & 0.6 \\ 0.3 & 0.4 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -0.3 & 0.6 \\ 0.3 & -0.6 \end{bmatrix} \sim \begin{bmatrix} -0.3 & 0.6 \\ 0 & 0 \end{bmatrix} \sim \begin{bmatrix} -1 & 2 \\ 0 & 0 \end{bmatrix} \Rightarrow t\begin{bmatrix} 1 \\ 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$

∴ population of $x_1 = 2/3\,(\text{total}) = 2/3\,(300) = 200$

4. **Written Explanation Exercise** Let P be a stochastic $n \times n$ matrix with positive entries. Give two methods of finding the steady state solution
   - Solve $[P - I \quad | \; 0]$
   - Steady state is ≈any column of P^k for larger k $(\text{only true for regular})$
5. A mouse lives in a maze that has at least three rooms. Each room is connected to at least one other room (in other words, every room is connected). At every hour, the mouse moves from the room where it is in, to one of the rooms it is connected to, with equal probability.

a. Design any mouse maze and its corresponding transition matrix P.



b. Is your P regular stochastic?

No

c. In the long run, is there a room that the mouse is more likely to be in at a given time? If so, which room?

Note: this problem is related to the PageRank problem that we explore later in this class

$P = \begin{bmatrix} 0 & 1/2 & 0 \\ 1 & 0 & 1 \\ 0 & 1/2 & 0 \end{bmatrix}$

$(P - I)\vec{x} = \vec{0} \Rightarrow \begin{bmatrix} -1 & 1/2 & 0 \\ 1 & -1 & 1 \\ 0 & 1/2 & -1 \end{bmatrix} \sim \begin{bmatrix} 0 & -1/2 & 1 \\ 1 & 0 & -1 \\ 0 & 1/2 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{cases} x_1 = x_3 \\ x_2 = 2x_3 \\ x_1 + x_2 + x_3 = 1 \end{cases} \Rightarrow \begin{bmatrix} 1/4 \\ 1/2 \\ 1/4 \end{bmatrix}$

Lecture 22

Friday, October 15, 2021 3:13 PM

Notes:

For $\lambda_1 = -1$:

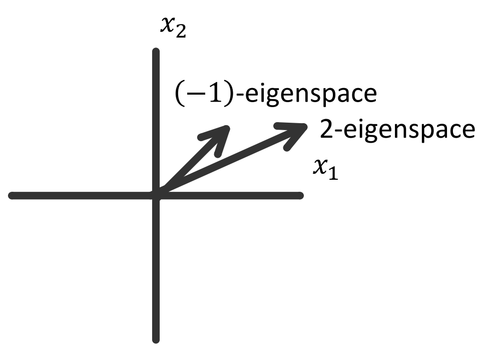
$$A - \lambda_1 I = A + I = \begin{pmatrix} 6 & -6 \\ 3 & -3 \end{pmatrix}$$

$\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$ is a basis for the (-1)-eigenspace

For $\lambda_1 = 2$:

$$A - \lambda_1 I = A + I = \begin{pmatrix} 3 & -6 \\ 3 & -6 \end{pmatrix}$$

$\left\{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right\}$ is a basis for the 2-eigenspace



1. $T = \begin{pmatrix} d_1 & * & * \\ & \ddots & * \\ (0) & & d_n \end{pmatrix}$ $\quad (T - \lambda I) = \begin{pmatrix} d_1 - \lambda & * & * \\ & \ddots & * \\ (0) & & d_n - \lambda \end{pmatrix}$

   $(T - \lambda I)$ is singular $\Leftrightarrow \lambda \in \{d_1, d_2, \dots, d_n\}$
2. Invertible $\Leftrightarrow A\vec{x} = \vec{0}$ how only the trivial solution $\vec{x} = \vec{0}$

   $\Leftrightarrow$ 0 is not an eigenvalue
3. Proof by induction:

   If $k = 2$: $\vec{v_1} \to \lambda_1, \vec{v_2} \to \lambda_2$: $\lambda_1 \neq \lambda_2$

   If $c_1\vec{v_1} + c_2\vec{v_2} = \vec{0}$ and $c_1 \neq 0$

   Then $\vec{v_1} = -\frac{c_1}{c_2}\vec{v_2}$

   $\underbrace{A(c_1\vec{v_1} + c_2\vec{v_2})}_{\vec{0}} = c_1 A\vec{v_1} + c_2 A\vec{v_2} = c_1\lambda_1\vec{v_1} + c_2\lambda_2\vec{v_2} = -c_2\lambda_1\vec{v_1} + c_2\lambda_2\vec{v_2} = c_2\underbrace{(\lambda_2 - \lambda_1)}_{\neq \vec{0}} = \underbrace{\vec{v_2}}_{\neq \vec{0}}$

   $\Rightarrow c_2 = 0 \Rightarrow c_1 = 0$: $\vec{v_1}, \vec{v_2}$ are linearly independent

Warning!

We can't determine the eigenvalues of a matrix from its reduced form.

Row reductions change the eigenvalues of a matrix.

Ex: suppose $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. The eigenvalues are $\lambda = 2,0$ because

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

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

- But the reduced echelon form of A is: $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$
- The reduced echelon form is triangular, and its eigenvalues are: 0 and 1 (triangular matrix)

Section 5.2: The Characteristic Equation

The Characteristic Polynomial

Recall

λ is an eigenvalue of $A \Leftrightarrow (A - \lambda I)$ is not invertible

Therefore, to calculate the eigenvalues of A, we can solve

$\det(A - \lambda I) = 0$

The quantity $\det(A - \lambda I)$ is the characteristic polynomial of A.

The quantity $\det(A - \lambda I) = 0$ is the characteristic equation of A.

The roots of the characteristic polynomial are the eigenvalues of A.

Ex.1

The characteristic polynomial of $A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$ is:

$$|A - \lambda I| = \begin{bmatrix} 5 - \lambda & 2 \\ 2 & 1 - \lambda \end{bmatrix} = (5 - \lambda)(1 - \lambda) - 4 = \lambda^2 - 6\lambda + 1$$

So the eigenvalues of A are:

$$\lambda = \frac{6 \pm \sqrt{36 - 5}}{2} = 3 \pm \frac{\sqrt{32}}{2} = 3 \pm 2\sqrt{2}$$

Definition

A is a matrix

Trace(A) = sum of the elements of the diagonals

Characteristic Polynomial of 2×2 Matricies

Express the characteristic equation of

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

In terms of its determinant. What is the equation when M is singular?

$$|A-\lambda I| = \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} = (a-\lambda)(d-\lambda) - bc = \lambda^2 + (a+d)\lambda + ad - bc = \lambda^2 - \mathrm{Trace}(M)\lambda + \det(M)$$

If the eigenvalues are λ_1 and λ_2

$$|M-\lambda I| == (\lambda-\lambda_1)(\lambda-\lambda_2) = \lambda^2 + (\lambda_1+\lambda_2)\lambda + \lambda_1\lambda_2$$

<u>Theorem:</u> A and A^T have the same eigenvalues

<u>Proof:</u>

$$|A-\lambda I| = |(A-\lambda I)^T| = |A^T - \lambda I^T| = |A^T - \lambda I|$$

Application: If P is stochastic 1 is an eigenvalue of P

P: the sum of each column is 1

P^T: the sum of each row is 1

$P^T \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$: 1 is an eigenvalue of P^T and therefore also for P.

Lecture 23

Monday, October 18, 2021 6:45 PM

Notes:

if M is singular: $\det M = 0$

$$|M - \lambda I| = \lambda^2 - \text{Trace}(M)\lambda$$
$$= \lambda(\lambda - \text{Trace}(M))$$

Algebraic Multiplicity

Definition:

The algebraic multiplicity of an eigenvalue is its multiplicity as a root of the characteristic polynomial.

if $|A - \lambda I| = (\lambda - 1)^3(\lambda + 2)^2(\lambda - 7)$

Eigenvalues	Algebraic Multiplicity
1	3
−2	2
7	1

Ex.

Compute the algebraic multiplicities of the eigenvalues for the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Eigenvalues	Algebraic Multiplicity
−1	1
0	2
1	1

Geometric Multiplicity

Definition

The geometric multiplicity of an eigenvalue λ is the dimension of $\text{Null}(A - \lambda I)$

1. Geometric multicity is always at least 1. It can be smaller then algebraic multiplicity
2. Here is the basic example

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2$$

$\lambda = 0$ is the only eigenvalue. Its algebraic multiplicity is 2, but the geometric multiplicity is 1.

$$\vec{x} \in \text{Null}(A - 0I) = \text{Null}(A)$$
$$\underbrace{A\vec{x}}_{\binom{x_2}{0}} = \binom{0}{0} \Rightarrow \vec{x} = x_1\binom{1}{0}$$

Ex.

Give an example of a 4×4 matrix with $\lambda = 0$ the only eigenvalue, but the geometric multiplicity of $\lambda = 0$ is one.

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$\rightarrow$ 1 free variable, $\dim(\text{Null}(A)) = 1$

Recall: Long-Term Behavior of Markov Chains

Recall:

- We often want to know what happens to a Markov Chain
  $$\overrightarrow{x_{k+1}} = P\overrightarrow{x_k}, \qquad k = 0,1,2,\dots$$
  as $k \to \infty$
- If P is regular, then there is a unique steady-state vector

Now lets ask:

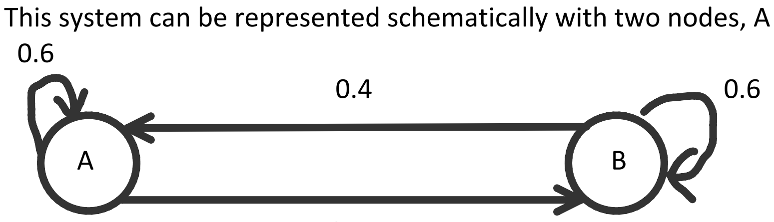
- If we don't know whether P is regular, what else might we do to describe the long-term behavior of the system?
- What can eigenvalues tell us about the behavior of these systems?

Example: Eigenvalues and Markov Chains

Consider the Markov Chain

$$\overrightarrow{x_{k+1}} = P\overrightarrow{x_k} = \begin{pmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{pmatrix}\overrightarrow{x_k}, \qquad k = 0,1,2,3,\dots, \qquad \overrightarrow{x_0} = \binom{1}{0}$$

This system can be represented schematically with two nodes, A and B:



Goal: use eigenvalues to describe the long-term behavior of our system.

What are the eigenvalues of P?

$$P = \begin{pmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{pmatrix}$$

P is stochasic: 1

$\lambda_1 + \lambda_2 = 1.2$: 0.2

What are the corresponding eigenvalues of P?

$\lambda_1 = 1$: $(P - I \quad | \quad 0) = \left(\begin{matrix} -0.4 & 0.4 \\ 0.4 & -0.4 \end{matrix} \,\middle|\, \begin{matrix} 0 \\ 0 \end{matrix}\right)$

$\rightarrow \overrightarrow{v_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$\lambda_2 = 0.2$: $(P - 0.2I \quad | \quad 0) = \left(\begin{matrix} 0.4 & 0.4 \\ 0.4 & 0.4 \end{matrix} \,\middle|\, \begin{matrix} 0 \\ 0 \end{matrix}\right)$

$\rightarrow \overrightarrow{v_2} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Note: $\{\overrightarrow{v_1}, \overrightarrow{v_2}\}$ are a basis of $\mathbb{R}^2$

Use the eigenvalues and eigenvectors of P to analyze the long-term behavior of the system. In other words, determine what $\overrightarrow{x_k}$ tends to as $k \rightarrow \infty$.

Idea: express $\overrightarrow{x_0}$ in the basis $\left\{ \underbrace{\overrightarrow{v_1}}_{\lambda_1 = 1}, \underbrace{\overrightarrow{v_2}}_{\lambda_2 = 0.2} \right\}$:

$$\begin{aligned}
\overrightarrow{x_0} &= c_1\overrightarrow{v_1} + c_2\overrightarrow{v_2} \\
\overrightarrow{x_1} &= P\overrightarrow{x_0} = P(c_1\overrightarrow{v_1} + c_2\overrightarrow{v_2}) \\
&= c_1P_1\overrightarrow{v_1} + c_2P_2\overrightarrow{v_2} \\
&= c_1\lambda_1\overrightarrow{v_1} + c_2\lambda_2\overrightarrow{v_2} \\
\overrightarrow{x_2} &= P^2\overrightarrow{x_0} = P(c_1\lambda_1\overrightarrow{v_1} + c_2\lambda_2\overrightarrow{v_2}) \\
&= c_1\lambda_1P_1\overrightarrow{v_1} + c_2\lambda_2P_2\overrightarrow{v_2} \\
&= c_1{\lambda_2}^2\overrightarrow{v_1} + c_2{\lambda_2}^2\overrightarrow{v_2} \\
\overrightarrow{x_k} &= P^k = c_1{\lambda_2}^k\overrightarrow{v_1} + c_2{\lambda_2}^k\overrightarrow{v_2}
\end{aligned}$$

but $\begin{cases} {\lambda_1}^k = 1^k = 1 \\ {\lambda_2}^k = (0.2)^k \overset{k\rightarrow\infty}{\rightrightarrows} 0 \end{cases}$

Thus:

$\lim_{k\rightarrow\infty} \overrightarrow{x_k} = \underbrace{c_1\overrightarrow{v_1}}_{\text{our unique steady-state vector}}$

if $\overrightarrow{x_0} = \begin{pmatrix} P \\ 1-P \end{pmatrix}$ where $0 \leq p \leq 1$

$\overrightarrow{x_0} = c_1\overrightarrow{v_1} + c_2\overrightarrow{v_2}$

$\begin{pmatrix} P \\ 1-P \end{pmatrix} = c_1\begin{pmatrix} 1 \\ 1 \end{pmatrix} = c_2\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} c_1 + c_2 \\ c_1 - c_2 \end{pmatrix}$

$\begin{cases} c_1 + c_2 = P \\ c_1 - c_2 = 1 - P \end{cases}$

$\rightarrow c_1 = {}^1\!/_2, \qquad c_2 = P - {}^1\!/_2$

Thus:

$c_1\overrightarrow{v_1} = \begin{pmatrix} {}^1\!/_2 \\ {}^1\!/_2 \end{pmatrix}$

Similar Matrices

Definition

Two $n \times n$ matrices A and B are similar if there is a matrix P so that $A = PBP^{-1}$.

Theorem

If A and B are similar, then they have the same characteristic polynomial.

If time permits, we will explain or prove this theorem in lecture. Note:

- Two matrices, A and B, do not need to be similar to have the same eigenvalues. For example,

  $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = B$

  $|A - \lambda I| = \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2 = |B - \lambda I|$

  but $PBP^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

  $\rightarrow A$ and B are not similar

<u>Proof</u>

$|A - \lambda I| = |PBP^{-1} - P(\lambda I)P^{-1}| = |P(B - \lambda I)P^{-1}| = |P||B - \lambda I||P^{-1}| = |B - \lambda I|$

Additional Examples

1. True or false
   a. If A is similar to the identity matrix, then A is equal to the identity matrix.
      i. If $A = (PI)P^{-1} = PP^{-1} = I$
      ii. True
   b. A row replacement operation on a matrix does not change its eigenvalues.
      i. $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$
      ii. False
2. For what values of k does the matrix have one real eigenvalue with algebraic multiplicity 2?
   a. $\begin{pmatrix} -3 & k \\ 2 & -6 \end{pmatrix}$
      i. "Highway": $A = \begin{pmatrix} -3 & k \\ 2 & -6 \end{pmatrix}$

         if $\lambda_1 = \lambda_2$: $\lambda_1 + \lambda_2 = \text{Trace}(A) = -9$

         then: $\det(A) = 18 - 2k = \lambda_1\lambda_2 = {}^{81}\!/_4 \Rightarrow k = -{}^9\!/_8$

Studio 15

Tuesday, October 19, 2021 12:27 PM

Def. λ is an eigenvalue of A.
if $A\vec{x} = \lambda\vec{x}$ for some $\vec{x} \neq \vec{0}$.
Here, x is called an eigenvector.
(x, λ) is an eigenpair.

How to find eigenvalues?
λ eigenvalue $\Leftrightarrow A\vec{x} = \lambda\vec{x}$ for some $\vec{x} \neq \vec{0}$.
$\Leftrightarrow (A - \lambda I)\vec{x} = \vec{0}$ some $\vec{x} \neq \vec{0}$.
$\Leftrightarrow A - \lambda I$ not invertible
$\Leftrightarrow \det(A - \lambda I) = 0$.

The eigenvalues of A are given by zeros of $p(\lambda) = \det(A - \lambda I)$, otherwise known as the characteristic polynomial of A.

Ex.
$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
$p(\lambda) = \det\left(\begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix}\right) = \lambda^2 - 1 = (\lambda + 1)(\lambda - 1) \Rightarrow$ eigenvalues are 1 and -1 $(\mathrm{alg.\,mult1})$

Say λ is an eigenvalue. Let $S = \{\vec{x}: A\vec{x} = \lambda\vec{x}\}$. Then S is a subspace. The geometric multiplicity of $\lambda = \dim(S)$.

Find $\vec{x}$ such that $\underbrace{A\vec{x} = \lambda\vec{x}}_{(A-\lambda I)\vec{x}=\vec{0}}$

Worksheet 5.1 and 5.2: Eigenvectors and Eigenvalues, The Characteristic Equation

Worksheet Exercises
1. If possible, give an example of:
   a. A 2×2 matrix, $A \in \mathbb{R}^{2\times2}$, whose eigenvalues have non-zero imaginary components.
      i. $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$
   b. A non-zero 2×2 matrix, $A \in \mathbb{R}^{2\times2}$, that is not triangular, but has a zero eigenvalue.
      i. $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$
2. Determine whether $\vec{u}$ and $\vec{v}$ are eigenvectors of A. If so, what are their eigenvalues? Do not find the characteristic polynomial.
   $A = \begin{pmatrix} -3 & -3 & 2 \\ 6 & 4 & 0 \\ 5 & 3 & 0 \end{pmatrix}, \quad \vec{u} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$
   $A\vec{u} = \begin{pmatrix} -3 & -3 & 2 \\ 6 & 4 & 0 \\ 5 & 3 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ -2 \end{pmatrix}$
   Eigenvalue $= -2$
   $A\vec{v} = \begin{pmatrix} -3 & -3 & 2 \\ 6 & 4 & 0 \\ 5 & 3 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -4 \\ 10 \\ 8 \end{pmatrix}$
   No eigenvalue
3. T is a linear transformation in $\mathbb{R}^2$. Without constructing A, identify one eigenvalue of A.
   a. T reflects points across the line $x_1 = -x_2$.
      $\lambda = 1$
   b. T projects points onto one of the coordinate axes.
      $\lambda = 0$
4. For what values of k (if any) does A have one real eigenvalue of algebraic multiplicly 2?
   $A = \begin{pmatrix} -4 & k \\ 2 & -2 \end{pmatrix}$
   $\det(A - \lambda I) = \begin{vmatrix} -4-\lambda & k \\ 2 & -2-\lambda \end{vmatrix} = (-4-\lambda)(-2-\lambda) - 2k = 8 + 4\lambda + 2\lambda + \lambda^2 - 2k = \lambda^2 + 6\lambda + 8 - 2k \Rightarrow k = -{}^1\!/_2$
5. $\mathrm{tr}(A)$ is the sum of the elements on the main diagonal of A. If $\mathrm{tr}(A) = 2$, $\det(A) = I$, and $A \in \mathbb{R}^{2\times2}$, compute the eigenvalues of A. *Hint*: *let* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$
   a. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $p(\lambda) = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + ad - bc = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$
6. Written Explanation Exercise If $Av = \lambda v$ with $v \neq 0$ annd A is invertible, can you find an eigenvalue/eigenvector of A^{-1}? Can A has a zero eigenvalue?
   a. $A\vec{v} = \lambda\vec{v} \quad (\vec{v} \neq 0)$
      $A^{-1}A = A^{-1}\lambda\vec{v} \Rightarrow \vec{v} = \lambda A\vec{v} \Rightarrow A^{-1}\vec{v} = \dfrac{1}{\lambda}\vec{v}$

Lecture 24

Wednesday, October 20, 2021 3:00 PM

<u>Review Session:</u>

If $A = PBP^{-1}$ and $B\vec{v} = \lambda\vec{v}$,

$A\vec{w} = \lambda\vec{w}$ such that $\vec{w} = P\vec{v}$

Unit 3

Saturday, November 13, 2021 8:49 PM

<u>Material Covered:</u>

Chapter 5: Eigenvalues and Eigenvectors
- Section 5.3 : Diagonalization
- Section 5.5 : Complex Eigenvalues

Chapter 10 : Finite-State Markov Chains
- Section 10.2 : The Steady-State Vector and Page Rank

Chapter 6: Orthogonality and Least Squares
- Section 6.1 : Inner Product, Length, and Orthogonality
- Section 6.2 : Orthogonal Sets
- Section 6.3 : Orthogonal Projections
- Section 6.4 : The Gram-Schmidt Process
- Section 6.5 : Least-Squares Problems
- Section 6.6 : Applications to Linear Models

Lecture 25

Friday, October 22, 2021 4:01 PM

Notes:

Section 5.3: Diagonalization

Diagonal Matrices

A matrix is diagonal if the only non-zero elements, if any, are on the main diagonal.

The following are all diagonal matrices.

$$\begin{bmatrix}2 & 0\\0 & 2\end{bmatrix},\quad [2],\quad I_n,\quad \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$$

We'll only be working with diagonal square matrices in this course.

$$D = \begin{pmatrix}1 & 0 & 0\\0 & 0 & 0\\0 & 0 & 5\end{pmatrix}$$

Powers of Diagonal Matrices

If A is diagonal, then A^k is easy to compute. For example,

$$A = \begin{pmatrix}3 & 0\\0 & 0.5\end{pmatrix}$$

$$A^2 = \begin{pmatrix}3 & 0\\0 & 0.5\end{pmatrix}\begin{pmatrix}3 & 0\\0 & 0.5\end{pmatrix} = \begin{pmatrix}3^2 & 0\\0 & (0.5)^2\end{pmatrix}$$

$$A^k = \begin{pmatrix}3^k & 0\\0 & (0.5)^k\end{pmatrix}$$

But what if A is not diagonal?

Diagonalization

Suppose $A \in \mathbb{R}^{n\times n}$. We say that A is diagonalizable if it is similar to a diagonal matrix, D. That is, we can write

$$A = PDP^{-1}$$

$$A^2 = PDP^{-1}PDP^{-1} = PD^2P^{-1}$$

$$\vdots$$

$$A^k = PD^kP^{-1}$$

Theorem:

If A is diagonalizable $\Leftrightarrow$ A has n linearly independent eigenvectors.

Note: the symbol $\Leftrightarrow$ means "if and only if".

Also note that $A = PDP^{-1}$ if and only if

$$A = [\overrightarrow{v_1}\quad \overrightarrow{v_2}\quad \cdots\quad \overrightarrow{v_n}]\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}[\overrightarrow{v_1}\quad \overrightarrow{v_2}\quad \cdots\quad \overrightarrow{v_n}]^{-1}$$

Where $\overrightarrow{v_1}, \dots, \overrightarrow{v_n}$ are linearly independent eigenvectors, and $\lambda_1, \dots, \lambda_n$ are the corresponding eigenvalues (in order)

Proof:

If A has n linearly independent eigenvectors $\overrightarrow{v_1}, \dots, \overrightarrow{v_n}$ for the eigen values $\lambda_1, \dots, \lambda_n$

→ define $P = (\overrightarrow{v_1}, \dots, \overrightarrow{v_n})$: P is invertible

$$AP = (A\overrightarrow{v_1}, \dots, A\overrightarrow{v_n}) = (\lambda_1\overrightarrow{v_1}, \dots, \lambda_n\overrightarrow{v_n})$$

$$PD = (\overrightarrow{v_1}, \dots, \overrightarrow{v_n})\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix} = (\lambda_1\overrightarrow{v_1}, \dots, \lambda_n\overrightarrow{v_n}) \qquad D = \begin{pmatrix}\lambda_1 & & (0)\\ & \ddots & \\ (0) & & \lambda_n\end{pmatrix}$$

$$AP = PD,\ A = PDP^{-1}$$

Ex.1

Diagonalize if possible

$$A = \begin{pmatrix}2 & 6\\0 & -1\end{pmatrix}$$

Eigenvalue: $2, -1$

$$\lambda_1 = 2: (A - 2I \quad | \quad 0) = \left(\begin{array}{cc|c}0 & 6 & 0\\0 & -3 & 0\end{array}\right)$$

Take $\overrightarrow{v_1} = \begin{pmatrix}1\\0\end{pmatrix}$

$$\lambda_2 = -1: (A + I \quad | \quad 0) = \left(\begin{array}{cc|c}3 & 6 & 0\\0 & 0 & 0\end{array}\right)$$

Take $\overrightarrow{v_2} = \begin{pmatrix}-2\\1\end{pmatrix}$

Define: $P = \begin{pmatrix}1 & -2\\0 & 1\end{pmatrix}$, $D = \begin{pmatrix}2 & 0\\0 & -1\end{pmatrix}$

Then $A = PDP^{-1}$

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

Ex.2

Diagonalize if possible

$$B = \begin{pmatrix}3 & 1\\0 & 3\end{pmatrix}$$

Eigenvalue: 3

$$(B - 3I \quad | \quad 0) = \left(\begin{array}{cc|c}0 & 1 & 0\\0 & 0 & 0\end{array}\right) \rightarrow \text{1 free variable}$$

$\dim \operatorname{Null}(B - 3I) = 1$

B is not diagonizable

Distinct Eigenvalues

Theorem

If A is $n \times n$ and has n disctinct eigenvalues, then A is diagonalizable.

Why does this theorem hold?

5.1: eigenvectors for distinct eigenvalues are linearly independent.

Is it necessary for an $n \times n$ matrix to have n distinct eigenvalues for it to be diagonalizable? NO!

$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ is diagonizable

Non-Distinct Eigenvalues

Theorem. Suppose

- A is $n \times n$
- A has distinct eigenvalues $\lambda_1, \dots, \lambda_n$, $k \leq n$
- a_i = algebraic multiplicity of λ_i
- d_i = dimension of λ_i eigenspace ("geometric multiplicity")

Then

1. $d_i < a_i$ for all i
2. A is diagonalizable $\Leftrightarrow \Sigma d_i = n \Leftrightarrow d_i = a_i$ for all $d_i i$
3. A is diagonalizable $\Leftrightarrow$ the eigenvectors, for all eigenvalues, together form a basis for $\mathbb{R}^n$.

Ex.

	EV	A	D
$A = \begin{pmatrix} 2 & 6 \\ 0 & -1 \end{pmatrix}$	2	1	1
	-1	1	1
$B = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$	3	2	1

Ex.3

The eigenvalues of A are $\lambda = 3,1$. If possible, construct P and D such that $AP = PD$

$$A = \begin{pmatrix} 7 & 4 & 16 \\ 2 & 5 & 8 \\ -2 & -2 & -5 \end{pmatrix}$$

$$\lambda_1 = 1: (A - I \quad | \quad 0) = \left(\begin{array}{ccc|c} 6 & 4 & 16 & 0 \\ 2 & 4 & 8 & 0 \\ -2 & -2 & -6 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 3 & 2 & 8 & 0 \\ 1 & 2 & 4 & 0 \\ 1 & 1 & 3 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\mathrm{Null}(A - I) = \mathrm{Span}\begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 3: (A - 3I \quad | \quad 0) = \left(\begin{array}{ccc|c} 4 & 2 & -2 & 0 \\ 2 & 2 & 8 & 0 \\ -2 & -2 & -8 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$\vec{x} \in \mathrm{Null}(A - 3I)$ if

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2 & -4x_3 \\ x_2 & \\ & x_3 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}$$

$$\text{Now: } P = \begin{pmatrix} -2 & -1 & -4 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}: \quad A = PDP^{-1}$$

Additional Example

Note that

$$\overrightarrow{x_k} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \overrightarrow{x_{k-1}}, \quad \overrightarrow{x_0} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad k = 1,2,3, \dots$$

Generates a well-known sequence of numbers

Use a diagonalization to find a matrix equation that gives the n^{th} number in this sequence

$$\overrightarrow{x_0} = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\overrightarrow{x_1} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = \begin{pmatrix} F_1 \\ F_1 + F_2 \end{pmatrix} = \begin{pmatrix} F_2 \\ F_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\overrightarrow{x_2} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \overrightarrow{x_3} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} \overrightarrow{x_3} = \begin{pmatrix} 5 \\ 8 \end{pmatrix} \dots$$

$$\to F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right)$$

$$\therefore \overrightarrow{x_n} = PD^n P^{-1} \overrightarrow{x_0}$$

Lecture 26

Monday, October 25, 2021 3:56 PM

Notes:

Chapter 5.5: Complex Eigenvalues

Imaginary Numbers
Recall: When calculating roots of polynomials, we can encounter square roots of negative numbers. For example:

$$x^2 + 1 = 0$$

The roots of this equation are:

$$\pm\sqrt{-1}$$

We usually write $\sqrt{-1}$ as i (for "imaginary").

Addition and Multiplication
The imaginary (or complex) numbers are denoted by $\mathbb{C}$, where

$$\mathbb{C} = \{a + bi | a, b \text{ in } \mathbb{R}\}$$

We can identify $\mathbb{C}$ with $\mathbb{R}^2$: $a + bi \leftrightarrow (a, b)$

We can add and multiply complex numbers as follows:
$(2 - 3i) + (-1 + i) = 1 - 2i$
$(2 - 3i)(-1 + i) = -2 + 2i + 3i - 3i^2 = 1 + 5i$

Complex Conjugate, Absolute Value, Polar Form
We can conjugate complex numbers: $\overline{a + bi} = a - bi$

The absolute value of complex number: $|a + bi| = \sqrt{a^2 + b^2}$

We can write complex numbers in polar form: $a + bi = r(\cos(\emptyset) + i\sin(\emptyset))$

Complex Conjugate Properties
If x and y are complex numbers, $\vec{v} \in \mathbb{C}^n$, it can be shown that:

- $\overline{(x + y)} = \bar{x} + \bar{y}$
- $\overline{A\vec{v}} = A\overline{\vec{v}}$
- $\text{lm}(x\bar{x}) = 0$

$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{C}^2$

$$\overline{A\vec{v}} = \begin{pmatrix} \overline{v_1 + 2v_2} \\ \overline{3v_1 + 4v_2} \end{pmatrix} = \begin{pmatrix} \overline{v_1} + \overline{2v_2} \\ \overline{3v_1} + \overline{4v_2} \end{pmatrix} = A\begin{pmatrix} \overline{v_1} \\ \overline{v_2} \end{pmatrix}$$

$x\bar{x} = (a + bi)(a - bi) = a^2 + b^2$

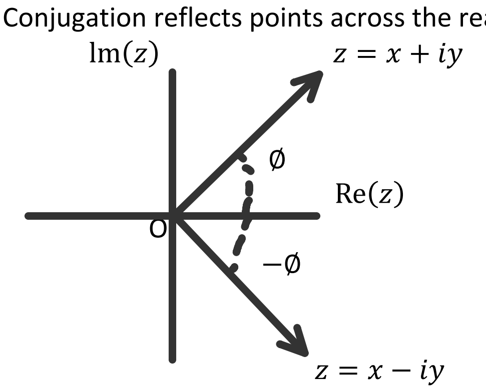
Ex. True or false: if x and y are complex numbers, then

$$\overline{(xy)} = \bar{x}\,\bar{y}$$

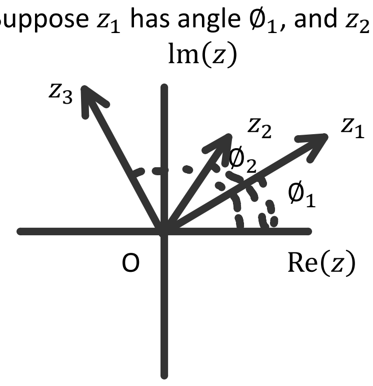
True: $xy = (a + bi)(c + di) = ac - bd + (ad - bc)i$
$\overline{xy} = ac - bd - (ad - bc)i$
$\bar{x}\,\bar{y} = (a + bi)(c - di) = ac - bd - (ad + bc)i$

Applications: $\overline{x^2} = \bar{x}\,\bar{x} = \overline{x^2} \dots \overline{x^n} = \bar{x}^n$
↳ P real polynomial: $\overline{P(x)} = P(\bar{x})$

Polar Form and the Complex Conjugate
Conjugation reflects points across the real axis



Euler's Formula
Suppose z_1 has angle $\emptyset_1$, and z_2 has angle $\emptyset_2$.



The product $z_1 z_2$ has angle $\emptyset_1 + \emptyset_2$ and modulus $|z||w|$. Easy to remember using Euler's formula.

$$z = |z|e^{i\emptyset}$$

The product $z_1 z_2$ is:

$$z_3 = z_1 z_2 = |z_1|e^{i\emptyset_1}|z_2|e^{i\emptyset_2} = |z_1||z_2|e^{i(\emptyset_1+\emptyset_2)}$$

Complex Number and Polynomials

Theorem: Fundamental Theorem of Algebra

Every polynomial of degree n has exactly n complex roots, counting multiplicity.

Theorem

1. If $\lambda \in \mathbb{C}$ is a root of a real polynomial $p(x)$, then the conjugate $\overline{\lambda}$ is also a root of $p(x)$.
2. If λ is an eigenvalue of real matrix A with eigenvector $\vec{v}$, then $\overline{\lambda}$ is an eigenvalue of A with eigenvector $\overline{\vec{v}}$.

   a. $p(\lambda) = 0$
      $\overline{p(\lambda)} = \overline{0} = 0$
   b. $\overline{A\vec{v}} = \overline{\lambda\vec{v}} = A\overline{\vec{v}} = \overline{\lambda}\overline{\vec{v}}$

Ex.

Four of the eigenvalues of a 7×7 matrix are $-2, 4+i, -4-i$, and i.

What are the other eigenvalues?

Eigenvalues:

-2

$4+i \rightarrow 4-i$

$-4-i \rightarrow -4+i$

$2 \rightarrow -i$

Ex.3

The matrix that rotates vectors by $\emptyset = \pi/_4$ radians about the origin, and then scales (or dilates) vectors by $r = \sqrt{2}$, il

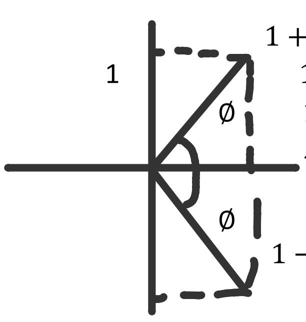
$$A = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\begin{bmatrix} \cos(\emptyset) & -\sin(\emptyset) \\ \sin(\emptyset) & \cos(\emptyset) \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

What are the eigenvalues of A? Express them in polar form.

$$A = r\begin{pmatrix} \cos(\emptyset) & -\sin(\emptyset) \\ \sin(\emptyset) & \cos(\emptyset) \end{pmatrix}$$

$|A - \lambda I| = \lambda^2 - 2\lambda + 2$

$$\lambda_{1,2} = \frac{2 \pm \sqrt{4-8}}{2} = 1 \pm i$$



$1 + i = \sqrt{z}e^{i\pi/_4}$

$1 - i = \sqrt{z}e^{-i\pi/_4}$

$z = |z|e^{i\emptyset}$

Ex.

The matrix in the previous example is a special case of this matrix:

$$C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

Calculate the eigenvalues of C and express them in polar form.

$|C - \lambda I| = \lambda^2 - 2a\lambda + a^2 + b^2$

$$\lambda_{1,2} = \frac{2 \pm \sqrt{4a^2 - 4(a^2+b^2)}}{2} = a \pm \sqrt{-b^2} = a \pm ib$$

$$\lambda = \sqrt{a^2+b^2}e^{\pm\tan^{-1}\left(\frac{b}{a}\right)}$$

$$C = \sqrt{a^2+b^2}\begin{pmatrix} a/\sqrt{a^2+b^2} & -b/\sqrt{a^2+b^2} \\ b/\sqrt{a^2+b^2} & a/\sqrt{a^2+b^2} \end{pmatrix} = r\begin{pmatrix} \cos(\emptyset) & -\sin(\emptyset) \\ \sin(\emptyset) & \cos(\emptyset) \end{pmatrix}$$

Diagonalization

Let A be a real 2×2 matrix with a complex eigenvalue $\lambda = a - bi$ (where $b \neq 0$) and associated eigenvector $\vec{v}$. Then we may construct the diagonalization

$$A = PCP^{-1}$$

where

$P = (\mathrm{Re}(\vec{v}) \quad \mathrm{Im}(\vec{v}))$ and $C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$

Note that following.

- C is referred to as a rotation dilation matrix, because it is the composition of a rotation by $\emptyset$ and dilation by r.
- The proof for why the columns of P are always linearly independent is a bit long, it goes beyond the scope of this course.

Studio 16

Tuesday, October 26, 2021 12:48 PM

Worksheet 5.3, Diagonalization

Worksheet Exercises

1. Recall from lecture: matrix A is diagonalizable if it can be written as $A = PDP^{-1}$
    a. P is a matrix whose columns are linearly independent eigenvector of A
    b. D is a diagonal matrix
    c. The elements on the main diagonal of D are eigenvalues of A
    d. A diagonal matrix is a matrix that in which nondiagonal entries are 0.
    e. The geometric multiplicity of an eigenvalue is:
        i. $\dim(\mathrm{Null}(A - \lambda I)) \rightarrow$ how many linearly independent eigenvectors exist
    f. A matrix can be diagonalized when the geometric multiplicities of all the eigenvalues:
        i. Is equal to the algebraic multiplicity of eigenvalues
2. If possible, construct P and D so that $A = PDP^{-1}$. Eigenvalues of A are given.
    a. $A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}, \lambda = 2,2,5$

$A - 2I = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

Suppose $\vec{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ is an eigenvector for $\lambda = 2$

$A\vec{v} = 2\vec{v}$

$(A - 2I)\vec{v} = A\vec{v} - 2\vec{v} = 0$

$\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_3\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$\therefore$ For $\lambda = 2$,

$\vec{v_1} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \qquad \vec{v_2}\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$

$A - 5I = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$

Suppose $\vec{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ is an eigenvector for $\lambda = 5$

$A\vec{v} = 5\vec{v}$

$(A - 5I)\vec{v} = A\vec{v} - 5\vec{v} = 0$

$\left(\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}\right)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} + x_2\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + x_3\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$\therefore$ For $\lambda = 5$,

$\vec{v_3} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}^{-1}$

    b. $A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}, \lambda = -2,5$

$A + 2I = \begin{pmatrix} 3 & 3 \\ 4 & 5 \end{pmatrix}$

Suppose $\vec{v} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ is an eigenvector for $\lambda = -2$

$A\vec{v} = -2\vec{v}$

$(A - 2I)\vec{v} = A\vec{v} + 2\vec{v} = 0$

$\left(\begin{bmatrix} 3 \\ 4 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix}\right)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1\begin{bmatrix} 3 \\ 4 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$\therefore$ NOT DIAGONALIZABLE

    c. $A = \begin{pmatrix} 3 & 2 \\ 0 & 3 \end{pmatrix}$
3. If possible, give an example of:
    a. A singular 2×2 matrix in echelon form that can be diagonalized.
    b. A singular 2×2 matrix in echelon form that cannot be diagonalized.
    c. A invertible 2×2 matrix in echelon form that can be diagonalized.
    d. A invertible 2×2 matrix in echelon form that cannot be diagonalized.
4. Indicate whether the statements are true or false.
    If A is diagonalizable, then so is A^2.
    If A^2 is diagonalizable, then so is A.
5. **Written Explanation Exercise** Given an example of an upper triangular 4×4 matrix A such that 0 is its only eigenvalue and such that its eigenspace is 3-dimensional. Explain why the eigenspace has dimension 3.

Worksheet 5.5, Complex Eigenvalues

Worksheet Exercises

1. Indicate whether the statements are true or false.
    a. There exists a real 2×2 matrix with the eigenvalues i and $2i$.
        i. False
    b. Every real 3×3 matrix must have a real eigenvalue.
        i. True
2. A is a composition of a rotation and a scaling. Give the angle of rotation, $\emptyset$, and the scale factor, r
    a. $A = \begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix} = r\begin{pmatrix} \cos(\emptyset) & -\sin(\emptyset) \\ \sin(\emptyset) & \cos(\emptyset) \end{pmatrix} = \begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix}$

b. $\emptyset = {}^{\pi}/_{6}, r = 2$

3. Let $A = \begin{pmatrix} 4 & -1 \\ 2 & 6 \end{pmatrix}$. Find an invertible matrix P and a rotation-dilation matrix C such that $A = PCP^{-1}$
   a. $P^{-1}CP = A = \begin{pmatrix} 4 & -1 \\ 2 & 6 \end{pmatrix}$
   b. Solve λ
      i. $\lambda = 5 \pm i$
      ii. Choose $\lambda = 5 - i$
      iii. $C = \begin{pmatrix} \mathrm{Re}(\lambda) & \mathrm{Im}(\lambda) \\ \mathrm{Im}(\lambda) & \mathrm{Re}(\lambda) \end{pmatrix} = \begin{pmatrix} 5 & 1 \\ -1 & 5 \end{pmatrix}$
      iv. $\vec{v} = \begin{bmatrix} 1 \\ -1 + i \end{bmatrix}$
      v. $P = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$
4. Matrix A is a 2×2 matrix that satisfies the equality
   a. $A^2 + 2A = -6I_2$

   I_2 is the 2×2 identity matrix. Compute the eigenvalues of A.
5. **Written Explanation Exercise** Can a 7×7 have 2 real eigenvalues and 5 non-real eigenvalues? If A is an $n \times n$ matrix and n matrix and n is odd, why does A have a real eigenvalues?

Lecture 27

Wednesday, October 27, 2021

Ex.
If possible, construct matrices P and C such that $AP = PC$

$$A = \begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix}$$

$|A - \lambda I| = \lambda^2 - 4\lambda + 5$, eigenvalues: $\lambda = \frac{4 \pm \sqrt{16-20}}{2} = 2 \pm i$

Take $2 - i$, $a - bi$: $a = 2, b = 1$ $\quad {}^*C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$

$\vec{v}$ eigenvector for λ: $(A - \lambda I \quad | \quad 0) = \left(\begin{array}{cc|c} -1+i & -2 & 0 \\ 1 & 1+i & 0 \end{array}\right)$

First row: $(-1+i)x_1 = 2x_2$

For example, take $x_1 = 2$: $x_2 = -1 + i$

$\vec{v} = \begin{pmatrix} 2 \\ -1+i \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} + i \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

For $p = \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix}$ and $C = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$

Then, $A = PCP^{-1}$

Section 10.2: The Steady-State Vector and PageRank

Steady State Vectors
Recall the car rental problem from our Section 4.9 lecture.

Problem
A car rental company has 3 rental locations, A, B, and C.

Rented From

		A	B	C
	A	.8	.1	.2
Returned To	B	.2	.6	.3
	C	.0	.3	.5

There are 10 cars at each location today, what happens to the distribution of cars after a long time?

Long Term Behavior
Can use the transition matrix, P, to find the distribution of cars after 1 week:

$\overrightarrow{x_1} = P\overrightarrow{x_0}$

The distribution of cars after 2 weeks is:

$\overrightarrow{x_2} = P\overrightarrow{x_1} = PP\overrightarrow{x_0}$

The distribution of cars after n weeks is:

$\overrightarrow{x_n} = P^n\overrightarrow{x_0}$

Long Term Behavior
To investigate the long-term behavior of a system that has a regular transition matrix P, we could:

1. compute the steady-state vector, $\vec{q}$, by solving $\vec{q} = P\vec{q}$.
2. compute $P^n\overrightarrow{x_0}$ for large n.
3. compute P^n for large n, each column of the resulting matrix is the steady-state

Theorem 1
If P is a regular m ×m transition matrix with m ≥ 2, then the following statements are all true.

1. There is a stochastic matrix Π such that
$$\lim_{n\to\infty} P^n = \Pi$$
2. Each column of Π is the same probability vector $\vec{q}$.
3. For any initial probability vector $\overrightarrow{x_0}$,
$$\lim_{n\to\infty} P^n\,\overrightarrow{x_0} = \vec{q}$$
4. P has a unique eigenvector, $\vec{q}$, which has eigenvalue $\lambda = 1$.
5. The eigenvalues of P satisfy $|\lambda| \le 1$.

We will apply this theorem when solving PageRank problems.

$\overrightarrow{v_1}, \dots, \overrightarrow{v_n}$ basis of eigenvectors

for $\lambda_1, \dots, \lambda_n$

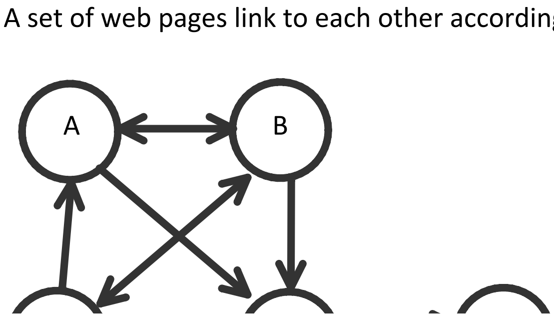
$\overrightarrow{x_0} = c_1\overrightarrow{v_1} + \dots + c_n\overrightarrow{v_n}$

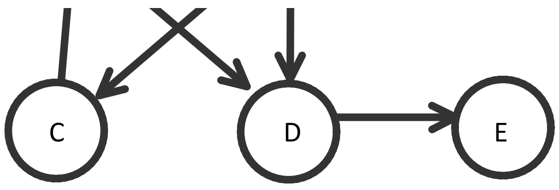
$P^k\overrightarrow{x_0} = c_1{\lambda_1}^k\overrightarrow{v_1} \underbrace{+ \dots + c_n{\lambda_n}^k\overrightarrow{v_n}}_{|\lambda_j|<1}$

For $j = 2$

Thus: ${\lambda_j}^k \underbrace{\longrightarrow}_{k\to+\infty} 0$

Ex. 1
A set of web pages link to each other according to this diagram.





Page A has links to pages B and D.
Page B has links to pages A, C and D.
We make two assumptions:

a) A user on a page in this web is equally likely to go to any of the pages that their page links to.
b) If a user is on a page that does not link to other pages, the user stays at that page.

Use these assumptions to construct a Markov chain that represents how users navigate the above web.

Solution:

$$P = \begin{pmatrix} 0 & 1/3 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/3 & 0 & 0 & 0 \\ 1/2 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Transition Matrix, Importance, and PageRank

- The square matrix we constructed in the previous example is a transition matrix. It describes how users transition between pages in the web.
- The steady-state vector, $\vec{q}$, for the Markov-chain, can characterize the long-term behavior of users in a given web.
- If $\vec{q}$ is unique, the importance of a page in a web is given by its corresponding entry in $\vec{q}$.
- The PageRank is the ranking assigned to each page based on its importance. The highest ranked page has PageRank 1, the second PageRank 2, and so on.
- Two pages with same importance receive the same PageRank (some other method would be needed to resolve ties)

Is the transition matrix in Example 1 a regular matrix? NO

If column k is just points to row k, then the matrix is non-regular

$$\overrightarrow{x_0} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \to \overrightarrow{x_1} = P \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \Rightarrow \overrightarrow{x_n} = \underbrace{P^n \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}}_{5^{\text{th}} \text{ column of } P^n} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \overrightarrow{e_s}$$

Adjustment 1
If a user reaches a page that does not link to other pages, the user will choose any page in the web, with equal probability, and move to that page.

Let's denote this modified transition matrix as P_*. Our transition matrix in Example 1 becomes:

$$P = \begin{pmatrix} 0 & 1/3 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/3 & 0 & 0 & 0 \\ 1/2 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} = P_* = \begin{pmatrix} 0 & 1/3 & 1/2 & 0 & 1/5 \\ 1/2 & 0 & 1/2 & 0 & 1/5 \\ 0 & 1/3 & 0 & 0 & 1/5 \\ 1/2 & 1/3 & 0 & 0 & 1/5 \\ 0 & 0 & 0 & 1 & 1/5 \end{pmatrix}$$

Adjustment 2
A user at any page will navigate to any page among those that their page links to with equal probability p, and to any page in the web with equal probability $1-p$. The transition matrix becomes

$$G = {}_pP_* + (1-p)k$$

All the elements of the $n \times n$ matrix k are equal to $1/n$.

p is referred to as the damping factor, Google is said to use $p = 0.85$.
With adjustments 1 and 2, our the Google matrix is:

$$\overset{\text{stochastic}}{\vec{G}} = \underbrace{{}_pP_*}_{\text{probability p to follow the links}} + \underbrace{\overset{k=\begin{pmatrix} 1/n & \cdots & 1/n \\ \vdots & \ddots & \vdots \\ 1/n & \cdots & 1/n \end{pmatrix}}{(1-p)k}}_{\text{probability } (1-p) \text{ to go wherever you want}}$$

$$G = \begin{pmatrix} 0.030 & 0.31\overline{3} & 0.425 & 0.030 & 0.200 \\ 0.425 & 0.030 & 0.425 & 0.030 & 0.200 \\ 0.030 & 0.31\overline{3} & 0.030 & 0.030 & 0.200 \\ 0.425 & 0.31\overline{3} & 0.030 & 0.030 & 0.200 \\ 0.030 & 0.030 & 0.030 & 0.880 & 0.200 \end{pmatrix}$$

Computing Page Rank

- Because G is stochastic, for any initial probability vector $\overrightarrow{x_0}$,
- $\lim_{n \to \infty} G^n \overrightarrow{x_0} - \vec{q}$
- We can obtain steady-state evaluating $G^n \overrightarrow{x_0}$ for large n, by solving $G\vec{q} = \vec{q}$, or by evaluating $\overrightarrow{x_n} = G\overrightarrow{x_{n-1}}$ for large n.
- Elements of the steady-state vector give the importance of each page in the web, which can be used to determine PageRank.
- Largest element in steady-state vector corresponds to page with PageRank 1, second largest with PageRank 2, and

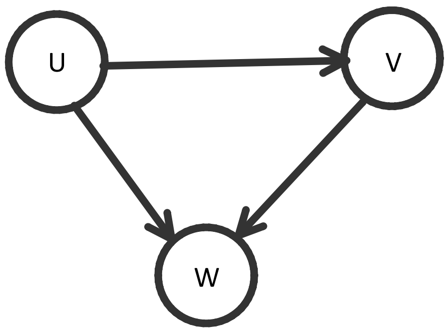
so on.

On an exam,

- problems that require a calculator will not be on your exam
- you may construct your G matrix using factions instead of decimal expansions

Studio 17

Thursday, October 28, 2021 12:34 PM



$G = \underbrace{P^*}_{p\text{transition matrix}} + (1-p)\underbrace{k}_{\text{regulaization}}$

$$P^* = \begin{bmatrix} 0 & 0 & 1/3 \\ 1/2 & 0 & 1/3 \\ 1/2 & 1 & 1/3 \end{bmatrix}$$

$$k = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

Solve $(G - I \quad | \quad 0)$

Find a solution v to whose entries add to 1, v is a steady state

Say v is the steady.

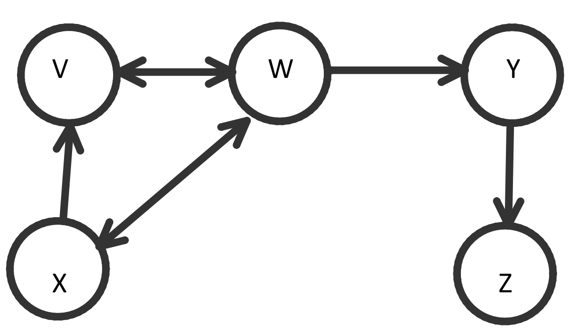
E.G. $v = \begin{bmatrix} .2 \\ .5 \\ .3 \end{bmatrix} \Rightarrow v > w > u$

(not the actual SS)

Worksheet 10.2, The Steady-State Vector and Page Rank

Worksheet Exercises

1. A set of web pages link to each other according to this diagram.



   a. Create the transition matrix, P, for this web.
      i. $P = \begin{pmatrix} 0 & 1/3 & 1/2 & 0 & 1/5 \\ 1 & 0 & 1/2 & 0 & 1/5 \\ 0 & 1/3 & 0 & 0 & 1/5 \\ 0 & 1/3 & 0 & 0 & 1/5 \\ 0 & 0 & 0 & 1 & 1/5 \end{pmatrix}$
   b. Construct the Google Matrix for this web, G. Use damping factor $p = 0.85$.
      i. $G = 0.85P_p + 0.95 \begin{pmatrix} 1/5 & \cdots & 1/5 \\ \vdots & \ddots & \vdots \\ 1/5 & \cdots & 1/5 \end{pmatrix}$, $\quad \vec{v} = \begin{pmatrix} 0.211 \\ 0.109 \\ 0.148 \\ 0.148 \\ 0.182 \end{pmatrix}$
   c. During an exam, to determine the page ranks of each page in the web a you would be given the steady-state vector G. Because you are not taking an exam right now, compute the steady-state vector and page ranks of each page on the web. You can use software.

   Hint: For a web with only two pages that are linked to each other, we can compute the steady state using MATLAB or Octave using these commands.

   Pstar $= 1/2\,[0\ 1;\ 1\ 0]$
   K $= 1/2 * \text{ones}(2)$
   G $=_p * \text{Pstar} + (1-p) * K$
   $G^{\uparrow}(100)$

   There are many free online Octave compilers.
2. Suppose p and q are real numbers on the open interval $(0,1)$, and
   $A = \begin{pmatrix} p & 1-q \\ 1-p & q \end{pmatrix}$
   a. Is A stochastic? Is A regular?
      i. $p = 0.2, q = 0.7$ (random #s to solution)
      ii. $\begin{pmatrix} 0.2 & 0.3 \\ 0.8 & 0.7 \end{pmatrix} \rightarrow$ all positive entries; hence, A is stochastic and regular

b. By inspection, what is one eigenvalue of A?
   i. Since A is stochastic and regular that means it has a unique steady-state which implies $\lambda = 1$

c. Compute the steady-state vector of A.
   i. $(A - I \quad | \quad 0) \Rightarrow \left(\begin{matrix} p-1 & 1-q \\ 1-p & q-1 \end{matrix} \;\middle|\; \begin{matrix} 0 \\ 0 \end{matrix}\right) \sim \left(\begin{matrix} 1 & {}^{1-q}/_{p-1} \\ 0 & 0 \end{matrix} \;\middle|\; \begin{matrix} 0 \\ 0 \end{matrix}\right)$

   $\Rightarrow t\begin{bmatrix} -{}^{1-q}/_{p-1} \\ 1 \end{bmatrix} \sim \begin{bmatrix} {}^{(-1+q)}/_{(-2+q+p)} \\ {}^{(p-1)}/_{(-2+q+p)} \end{bmatrix}$

d. Compute the limit $\lim_{n\to\infty} A^n$
   i. $\to [\vec{v} \quad \vec{v}] \Rightarrow \begin{bmatrix} {}^{(-1+q)}/_{(q+p-2)} & {}^{(-1+q)}/_{(q+p-2)} \\ {}^{(p-1)}/_{(q+p-2)} & {}^{(p-1)}/_{(q+p-2)} \end{bmatrix}$

3. Consider the dynamical system $\overrightarrow{x_k} = A\overrightarrow{x_{k-1}}$, $k = 1,2,3, \ldots,$ where

$$A = \begin{pmatrix} .5 & .25 & .25 \\ .25 & .5 & .25 \\ .25 & .25 & .5 \end{pmatrix}, \overrightarrow{x_0} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

The eigenvalues of A are 1 and ${}^1/_4$. Analyze the long-term behavior of the system. In other words, determine what $\overrightarrow{x_k}$ tends to as $k \to \infty$

$$x_k = A^k x_0$$

$$A = PDP^{-1}$$

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & & \\ & {}^1/_4 & \\ & & {}^1/_4 \end{bmatrix}$$

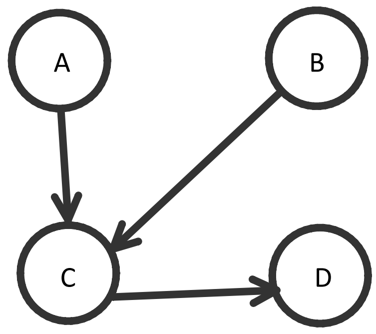
$$P^{-1} = {}^1/_3 \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

$$x_k = A^k x_0 = x_k = A^k x_0 x_0 = PD^kP^{-1}x_0 \to P\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} P^{-1}x_0 = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} {}^1/_3 \\ {}^1/_3 \\ {}^1/_3 \end{bmatrix}$$

Lecture 28

Friday, October 29, 2021 3:31 PM

<u>Notes:</u>
Construct the Google Matrix for the web below. Which page do you think will have the highest PageRank? How would your result depend on the damping factor p? Use software to explore the questions.



1. Compute: P

$$G = 0.85P_* + 0.15\begin{pmatrix} 1/4 & \\ 1/4 & \cdots \\ 1/4 & \\ 1/4 & \end{pmatrix}$$

2. Use a computer to find $\vec{q}$ for G

$$P_* = \begin{pmatrix} 0 & 0 & 0 & 1/4 \\ 0 & 0 & 0 & 1/4 \\ 1 & 1 & 0 & 1/4 \\ 0 & 0 & 1 & 1/4 \end{pmatrix}$$

2. Columns of G^n

$$\vec{q} \cong \begin{pmatrix} 0.13 \\ 0.13 \\ 0.34 \\ 0.41 \end{pmatrix}, \vec{q} \cong \frac{1}{1599}\begin{pmatrix} 200 \\ 200 \\ 540 \\ 654 \end{pmatrix}$$

PageRank

1. D
1) C
2) A and B

Section 6.1: Inner Product, Length, and Orthogonality

<u>Remark</u>

$(AB)_{i,j} = \overrightarrow{\text{Row}}(A,i).\overrightarrow{\text{Col}}(B,j)$

The Dot Product

The dot product between two vectors, $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$, is defined as

$$\vec{u}\cdot\vec{v} = \vec{u}^T\vec{v} = [u_1 \quad u_2 \quad \cdots \quad u_n]\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

Ex.1: For what values of k is $\vec{u}\cdot\vec{v} = 0$?

$$\vec{u} = \begin{pmatrix} -1 \\ 3 \\ k \\ 2 \end{pmatrix}, \qquad \vec{v} = \begin{pmatrix} 4 \\ 2 \\ 1 \\ -3 \end{pmatrix}$$

$\vec{u}\cdot\vec{v} = -4 + 6 + k - 6 = k - 4$

$\vec{u}\cdot\vec{v} = 0 \Leftrightarrow k = 4$

Properties of the Dot Product

The dot product is a special form of matrix multiplication, so it inherits linear properties.

Theorem (Basic Identities of Dot Product)

Let $\vec{u}, \vec{v}, \vec{w}$ be three vectors in $\mathbb{R}^n$, and $c \in \mathbb{R}$.

1) (Symmetry) $\vec{u}\cdot\vec{w} = \vec{w}\cdot\vec{u}$
2) (Linear in each vector) $(\vec{v}+\vec{w})\cdot\vec{u} = \vec{v}\cdot\vec{u} + \vec{w}\cdot\vec{u}$
3) (Scalars) $(c\vec{u})\cdot\vec{w} = c(\vec{u}\cdot\vec{w})$
4) (Positivity) $\vec{u}\cdot\vec{u} \geq 0$, and the dot product equals 0 iff $\vec{u} = \vec{0}$

$$\vec{u}\cdot\vec{u} = {u_1}^2 + {u_2}^2 + \cdots + {u_n}^2$$

The Length of a Vector

Definition

The length of a vector $\vec{u} \in \mathbb{R}^n$ is

$$\|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}} = \sqrt{{u_1}^2 + {u_2}^2 + \cdots + {u_n}^2}$$

Ex. The length of $\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$ is $\sqrt{1^2 + 3^2 + 2^2} = \sqrt{14}$

Ex.

Let $\vec{u}, \vec{v}$ be two vectors in $\mathbb{R}^n$ with $\|\vec{u}\| = 5$, $\|\vec{v}\| = \sqrt{3}$, and $\vec{u}\cdot\vec{v} = -1$.

Compute the value of $\|\vec{u} + \vec{v}\|$

$$\begin{aligned}\|\vec{u}+\vec{v}\|^2 &= (\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v})\\ &= \vec{u}\cdot\vec{u}+\vec{u}\cdot\vec{v}+\vec{v}\cdot\vec{v}\\ &= \|\vec{u}\|^2+2\vec{u}\cdot\vec{v}+\|\vec{v}\|^2\\ &= 25-2+3\\ &= 26\\ &\to \|\vec{u}+\vec{v}\|=\sqrt{26}\end{aligned}$$

$$\|c\vec{u}\| = \sqrt{c^2{u_1}^2+\cdots+c^2{u_n}^2} = \sqrt{c^2}\sqrt{{u_1}^2+\cdots+{u_n}^2} = |c|\|\vec{u}\|$$

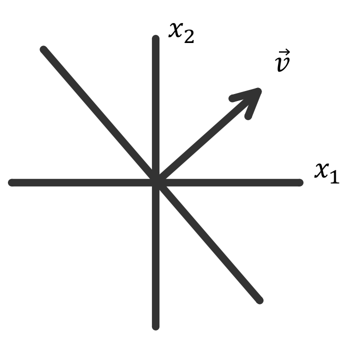
Length of Vectors and Unit Vectors
Note: for any vector $\vec{v}$ and scalar c, the length of $c\vec{v}$ is
$\|c\vec{v}\| = |c|\|\vec{v}\|$
Definition
If $\vec{v}\in\mathbb{R}^n$ has length one, we say it is a unit vector
For example, each of the following vectors are until vectors.

$$\vec{e_1}=\begin{pmatrix}1\\0\end{pmatrix}\qquad \vec{y}=\frac{1}{\sqrt{5}}\begin{pmatrix}1\\2\end{pmatrix}\qquad \vec{v}=\frac{1}{\sqrt{3}}\begin{pmatrix}1\\0\\1\\1\end{pmatrix}$$

Distance in $\mathbb{R}^n$
Definition
For $\vec{u},\vec{v}\in\mathbb{R}^n$, the distance between $\vec{u}$ and $\vec{v}$ is given by the formula
$\|\vec{u}-\vec{v}\|$
Example: Compute the distance from $\vec{u}=\begin{pmatrix}7\\1\end{pmatrix}$ and $\vec{v}=\begin{pmatrix}3\\2\end{pmatrix}$
Here: $\vec{u}-\vec{v}=\begin{pmatrix}4\\-1\end{pmatrix}$
$\|\vec{u}-\vec{v}\| = \sqrt{16+1}=\sqrt{17}$

Orthogonality
Definition (Orthogonal Vectors)
Two vectors $\vec{u}$ and $\vec{w}$ ar e orthogonal if $\vec{u}\cdot\vec{w}=0$. This is equivalent to:
$\|\vec{u}-\vec{w}\|^2=\|\vec{u}\|^2+\|\vec{w}\|^2$
Note: The zero vector in $\mathbb{R}^n$ is orthogonal to every vector in $\mathbb{R}^n$. But we usually only mean non-zero vectors.

Ex.2
Sketch the subspace spanned by the set of all vectors $\vec{v}$ that are orthogonal to $\vec{v}=\begin{pmatrix}3\\2\end{pmatrix}$



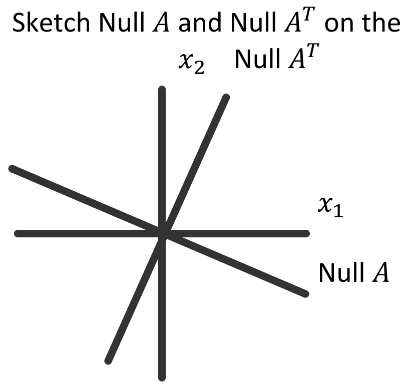
$\vec{u}=\begin{pmatrix}x_1\\x_2\end{pmatrix}$
$\vec{u}\cdot\vec{v}=0 \Leftrightarrow 3x_1+2x_2=0$
$\updownarrow$
$\vec{v}+\vec{u}=0 \Leftrightarrow \vec{u}\in \mathrm{Null}(\vec{v})$

Orthogonal Compliments
Definitions
Let W be a subspace of $\mathbb{R}^n$. Vector $\vec{z}\in\mathbb{R}^n$ is orthogonal to W if $\vec{z}$ is orthogonal to every vector in W.
He set of all vectors orthogonal to W is a subspace, the orthogonal compliment of W, or W^T or 'W prep.'
$W^T=\{\vec{z}\in\mathbb{R}^n :\ \vec{z}\cdot\vec{w}=0\ \forall\ \vec{w}\in W\}$

Ex.3
Example: suppose $A=\begin{pmatrix}1&3\\2&6\end{pmatrix}$

- Col A is the span of $\vec{a_1}=\begin{pmatrix}1\\2\end{pmatrix}$
- Col A^T is the span of $\vec{z}=\begin{pmatrix}2\\-1\end{pmatrix}$

Sketch Null A and Null A^T on the grid below.



$\mathrm{Null}\ (A)=\mathrm{Span}\begin{pmatrix}3\\-1\end{pmatrix}$
$\mathrm{Null}\ (A)^{\perp}=\mathrm{Span}\begin{pmatrix}1\\3\end{pmatrix}$

Ex.4
Line L is a subspace of $\mathbb{R}^3$ spanned by $\vec{v}=\begin{pmatrix}2\\-1\\2\end{pmatrix}$. Then the space $L^{\perp}$ is a place. Construct an equation of the

plane $L^{\perp}$.

$$\vec{u} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$
$$\vec{u} \cdot \vec{v} = 0 \Leftrightarrow x_1 - x_2 + 2x_3 = 0$$

Row A

Definition

Row A is a the space spanned by the rows of matrix A.

We can show that

- $\dim(\mathrm{Row}(A)) = \dim(\mathrm{Col}(A))$
- A basis for Row A is a pivot rows of A

Note that $\mathrm{Row}(A) = \mathrm{Col}(A^T)$, but in general Row A and Col A are not related to each other

Ex.5

Describe the Null (A) is terms of an orthogonal subspace.

A vector $\vec{x}$ is in Null A if and only if

1. $A\vec{x} = \vec{0}$
2. This means that $\vec{x}$ is orthogonal to each row of A
3. Row A is orthogonal to Null A
4. The dimension of Row A plus the dimension of Null A equals n (# columns)

Rank Theorem

$$\underbrace{\dim(\mathrm{Col}(A))}_{\dim(\mathrm{Row}(A))} + \dim(\mathrm{Null}(A)) = n$$

Theorem (The Four Subspaces)

For any $A \in \mathbb{R}^{m\times n}$, the orthogonal complement of Row A is Null A, and the orthogonal complement of Col A is Null A^T

We know: Row (A) $\perp$ Null (A) in $\mathbb{R}^n$

Apply to A^T: $\underbrace{\text{Row } (A^T)}_{\text{Col A}}$ $\perp$ Null (A^T) in $\mathbb{R}^n$

Lecture 29

Monday, November 1, 2021 5:24 PM

<u>Notes:</u>

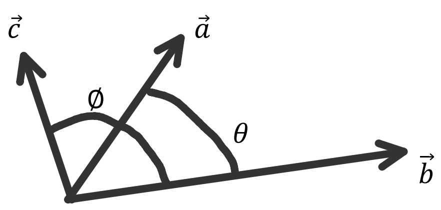
Angles

Theorem

$\vec{a}\cdot\vec{b} = |\vec{a}||\vec{b}|\cos\theta$. Thus, if $\vec{a}\cdot\vec{b}=0$, then:

- $\vec{a}$ and/or $\vec{b}$ are 0 vectors, or $\vec{a}$ and $\vec{b}$ are orthogonal vectors.

For example, consider the vectors below.



"easy case": $\vec{b}$ is in the direction of the x_1-axis

$$a = (\|\vec{a}\|\cos\theta, \|\vec{a}\|\sin\theta)$$
$$b = \left(\|\vec{b}\|, 0\right)$$

Thus: $\vec{a}\cdot\vec{b} = \|\vec{a}\|\|\vec{b}\|\cos\theta + 0$

Section 6.2: Orthogonal Sets

Orthogonal Vector Sets

Definition

A set of vectors $\{\vec{u_1},\dots,\vec{u_p}\}$ are an orthogonal set of vectors if for each $j\neq k$, $\vec{u_j}\perp\vec{u_k}$.

Ex: Fill in the missing entries to make $\{\vec{u_1},\vec{u_2},\vec{u_3}\}$ and orthogonal set of vectors.

$$\vec{u_1}=\begin{bmatrix}4\\0\\1\end{bmatrix},\qquad \vec{u_2}=\begin{bmatrix}-2\\0\\\boxed{8}\end{bmatrix},\qquad \vec{u_3}=\begin{bmatrix}0\\\boxed{*}\\\boxed{0}\end{bmatrix}$$

$$\vec{u_2}=\begin{pmatrix}-2\\0\\x\end{pmatrix}\to \vec{u_1}\cdot\vec{u_2}=-8+x$$
$$\vec{u_3}=\begin{pmatrix}0\\x\\z\end{pmatrix}\to \begin{matrix}\vec{u_1}\cdot\vec{u_3}=2\\ \vec{u_2}\cdot\vec{u_3}=8z\end{matrix}$$

$\to y$ is free.

Linear Independence

Let $\{\vec{u_1},\dots,\vec{u_p}\}$ be an orthogonal set of vectors. Then, for scalars $c_1,\dots c_p$.

$$\|c_1\vec{u_1}+\dots+c_p\vec{u_p}\|^2 = {c_1}^2\vec{u_1}^2+\dots+{c_p}^2\vec{u_p}^2.$$

In particular, if all the vectors $\vec{u_r}$ are non-zero, the set of vectors $\{\vec{u_1},\dots,\vec{u_p}\}$ are linearly independent.

$$\|c_1\vec{u_1}+\dots+c_p\vec{u_p}\|^2 = \left(\sum_{i=1}^{p}c_i\vec{u_i}\right)\left(\sum_{i=1}^{p}c_i\vec{u_i}\right)$$

But $\vec{u_i}\cdot\vec{u_j}=\begin{cases}\|\vec{u_i}\|^2 & if\ i=j\\ 0 & if\ i\neq j\end{cases}$

$$\to \left\|\sum_{i=1}^{p}c_i\vec{u_i}\right\|^2 = {c_1}^2\|\vec{u_1}\|^2+\dots+{c_p}^2\|\vec{u_p}\|^2$$

If none of the vector $\vec{u_1}$ is $\vec{0}$:

Assume: $c_1\vec{u_1}+\dots+c_p\vec{u_p}=\vec{0}$

We have: $\|c_1\vec{u_1}+\dots+c_p\vec{u_p}\|^2={c_1}^2\vec{u_1}^2+\dots+{c_p}^2\vec{u_p}^2$

$\to c_1=c_2=\dots=c_p=0$

$\to \vec{u_1},\dots,\vec{u_p}$ are linearly independent

Orthogonal Bases

Basis $\{\vec{u_1},\dots,\vec{u_p}\}$ of W

$\vec{w}\in W$: $\vec{w}=c_1\vec{u_1}+\dots+c_p\vec{u_p}$

Then $(\vec{u_1},\dots,\vec{u_p})\begin{pmatrix}c_1\\\vdots\\c_n\end{pmatrix}=\vec{w}$

Theorem (Expansion in Orthogonal Basis)

Let $\{\vec{u_1},\dots,\vec{u_p}\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. Then, for any vector $\vec{w}\in W$.

$$\vec{w}=c_1\vec{u_1}+\dots+c_p\vec{u_p}.$$

Above, the scalars are $c_q=\frac{\vec{w}\cdot\vec{u_q}}{\vec{u_q}\cdot\vec{u_q}}$.

For example, any vector $\vec{w}\in\mathbb{R}^3$ can be written as a linear combination of $\{\vec{e_1},\vec{e_2},\vec{e_3}\}$, or some orthogonal basis $\{\vec{u_1},\vec{u_2},\vec{u_3}\}$.

$\vec{w}=c_1\vec{u_1}+\dots+c_p\vec{u_p}$

$1\le q\le p$: $\vec{w}\cdot\vec{u_q}=c_1\vec{u_1}\vec{u_q}+\dots+c_p\vec{u_p}\vec{u_q}=c_q\vec{u_q}\cdot\vec{u_q}=\boxed{\frac{\vec{w}\cdot\vec{u_q}}{\vec{u_q}\cdot\vec{u_q}}}$

Ex.

$$\vec{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{u} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{s} = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}$$

Let W be the subspace of $\mathbb{R}^3$ that is orthogonal to $\vec{x}$.

a) Check that an orthogonal basis for W is given by $\vec{u}$ and $\vec{x}$

$\vec{u} \cdot \vec{x} = 1 - 2 + 1 = 0\text{: } \vec{u} \in \overrightarrow{W}$

$\vec{v} \cdot \vec{x} = -1 + 0 + 1 = 0\text{: } \vec{v} \in \overrightarrow{W}$

$\vec{u} \cdot \vec{v} = -1 + 0 + 1 = 0\text{: } \vec{u}, \vec{v}$ linearly independent

$\rightarrow \{\vec{u}, \vec{v}\}$ isan orthogonal basis of W.

b) Compute the expansion $\vec{s}$ in basis W.

$$\vec{s} = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}$$

$\vec{s} \cdot \vec{x} = 3 - 4 + 1 = 0\text{: } \vec{s} \in \overrightarrow{W}$

$\rightarrow \ \vec{s} = c_1 \vec{u} + c_2 \vec{v}$

$$c_1 = \frac{\vec{s} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} = \frac{3 + 8 + 1}{1 + 4 + 1} = \frac{12}{6} = 2$$

$$c_2 = \frac{\vec{s} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} = \frac{-3 + 1}{1 + 1} = \frac{-2}{2} = -1$$

$\rightarrow \boxed{\vec{s} = 2\vec{u} - \vec{v}}$

Projections

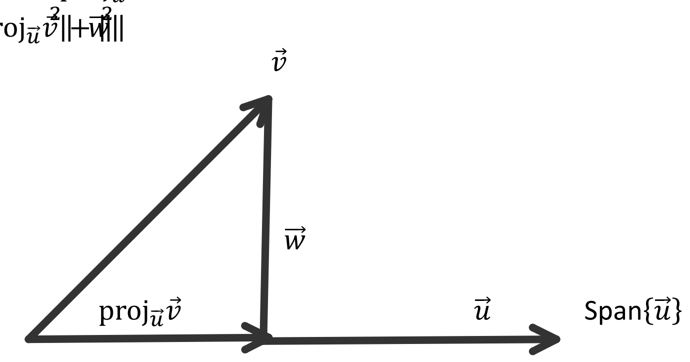
Let $\vec{u}$ be a non-zero vector, and let $\vec{v}$ be some other vector. The orthogonal prokection of $\vec{v}$ onto the direction of $\vec{u}$ is the vector in the span of $\vec{u}$ that is closest to $\vec{v}$.

$$\boxed{\text{proj}_{\vec{u}} \vec{v} = \frac{\vec{v} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \vec{u}}$$

The vector $\vec{w} = \vec{v} - \text{proj}_{\vec{u}} \vec{v}$ is orthogonal to $\vec{u}$, so that

$\vec{v} = \text{proj}_{\vec{u}} \vec{v} + \vec{w}$

$\|\vec{v}\|^2 = \|\text{proj}_{\vec{u}} \vec{v}\|^2 + \|\vec{w}\|^2$



$\text{proj}_{\vec{u}} \vec{v} =$ Span$\{\vec{u}\}$

$\rightarrow \text{proj}_{\vec{u}} \vec{v} = k\vec{u}, k \in \mathbb{R}$

$\vec{u} \cdot \vec{v} = \vec{u} \cdot (\text{proj}_{\vec{u}} \vec{v} + \vec{w})$

$= k\vec{u} \cdot \vec{u} + \vec{u} \cdot \vec{w}$

$\rightarrow \dfrac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}}$

Ex.

Let L be spanned by $\vec{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

1. Calculate the projection of $\vec{y} = (-3, 5, 6, -4)$ onto line L.

$$\text{proj}_{\vec{u}} \vec{y} = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \vec{u} = \frac{-3 + 5 + 6 - 4}{1 + 1 + 1} \vec{u} = \vec{u}$$

2. How close is $\vec{y}$ to the line L?

$$\|\vec{y} - \text{proj}_{\vec{u}} \vec{x}\| = \|\vec{y} - \vec{u}\| = \left\| \begin{pmatrix} -4 \\ 4 \\ 5 \\ -5 \end{pmatrix} \right\| = \sqrt{16 + 16 + 25 + 25} = \sqrt{82}$$

Definition

Definition (Orthonormal Basis)

An orthonormal basis for a subspace W is an orthogonal basis $\{\overrightarrow{u_1}, \dots, \overrightarrow{u_p}\}$ in which every vector $\overrightarrow{u_q}$ has unit length. In this case, for each $\vec{w} \in W$.

$$\vec{w} = (\vec{w} \cdot \overrightarrow{u_1})\overrightarrow{u_1} + \dots + (\vec{w} \cdot \overrightarrow{u_p})\overrightarrow{u_p}$$

$$\|\vec{w}\| = \sqrt{(\vec{w} \cdot \overrightarrow{u_1})^2 + \dots + (\vec{w} \cdot \overrightarrow{u_p})^2}$$

$$\overrightarrow{c_q} = \frac{\vec{w} \cdot \overrightarrow{u_q}}{\overrightarrow{u_q} \cdot \overrightarrow{u_q}} = \vec{w} \cdot \overrightarrow{u_q}$$

Orthonormal (orthogonal + normal)

Studio 18

Tuesday, November 2, 2021 12:34 PM

Def. The dot product of $\vec{u}, \vec{v}$ is $\vec{u} \cdot \vec{v} = u_1 v_1 + \cdots + u_n v_n$

Ex. $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} = 1(-3) + 2(1) + 1(0)$

Def. $\vec{u}$ is orthogonal to $\vec{v}$ if $\vec{u} \cdot \vec{v} = 0$

Def. The length $\vec{u}$ is $\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}}$

Note: $\vec{u} \cdot \vec{u} = {u_1}^2 + \cdots + {u_n}^2$

Def. $\{u_1, \ldots, u_n\}$ is orthogonal il

$u_i \cdot u_j = 0$ for all $i \neq j$

$\|u_i\| = 1$ for all i

Ex. $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ orthonormal

Prop. The projection of $\vec{v}$ onto $\vec{u}$ is $\text{proj}_{\vec{u}} \vec{v} = \frac{\vec{v} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \vec{u}$

The closest vector to $\vec{v}$ in Span$\{u\}$ is $\text{proj}_{\vec{u}} \vec{v}$

Worksheet 6.1, Inner Product, Length, and Orthogonality

Worksheet Exercises

1. Fill in the blanks.
   a. The distance between the vector $\vec{u} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ and the line spanned by $\vec{w} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is $\text{proj}_{\vec{w}} \vec{u} = \frac{\vec{u} \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \vec{w} = \frac{2}{1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \left\| \begin{pmatrix} 0 \\ 3 \end{pmatrix} \right\| =$
   $\sqrt{(0)^2 + (3)^2} = 3$
   b. If W is the plane spanned by the vectors $\vec{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, a basis of $W^\perp$ is given by $\vec{w} = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$.
   $\vec{w} = \text{Col} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}, \vec{w}^\perp = \text{Null} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$
   c. If $V = \{\vec{x} \in \mathbb{R}^3 | x_1 + x_2 = x_3\}$, then $\dim V = 2$, and $\dim V^\perp = 1$.
   $V = \text{Null}\begin{bmatrix} 1 & 1 & -1 \end{bmatrix}$
   $\Rightarrow \dim(V) = 2$
   $\Rightarrow \dim(V^\perp) = \dim(\text{Row}\begin{bmatrix} 1 & 1 & -1 \end{bmatrix}) = 1$
   $\dim(V) + \dim(V^\perp) = n$
2. W is the set of all vectors of the form $\begin{pmatrix} x \\ y \\ x+y \end{pmatrix}$. Which of the vectors are in $W^\perp$?
   $\vec{u} = \begin{pmatrix} 8 \\ -5 \\ 8 \end{pmatrix}, \boxed{\vec{v} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}}, \vec{w} = \begin{pmatrix} 3 \\ 3 \\ -3 \end{pmatrix}$
3. True or False
   a. If $\vec{x} \in \text{Null}(A)$, then $\vec{x}$ is orthogonal to the rows of matrix A.
      i. True
   b. If $\vec{u}$ and $\vec{v}$ are non-zero orthogonal vectors, then they are linearly independent.
      i. True

Worksheet 6.2, Orthogonal Sets

Worksheet Exercises

1. Indicate whether the statements are true or false
   a. If the columns of an $n \times n$ matrix A are orthonormal, then the linear mapping $\vec{x} \to A\vec{x}$ preserves lengths.
      i. True
   b. If P is a stochastic matrix, then the columns of P have unit length.
      i. False
2. Write $\vec{y}$ as the sum of a vector parallel to $\vec{u}$ and a vector perpendicular to $\vec{u}$.
   $\vec{y} = \begin{pmatrix} -1 \\ -5 \\ 10 \end{pmatrix}, \vec{u} = \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix}$
3. Find the coordinates for $\vec{v}$ in the subspace spanned by the orthogonal vectors $\vec{u_1}$ and $\vec{u_2}$.
   $\vec{v} = \begin{pmatrix} 0 \\ -5 \\ -3 \end{pmatrix}, \vec{u_1} = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}, \vec{u_2} = \begin{pmatrix} 6 \\ -7 \\ -10 \end{pmatrix}$
   $\text{proj}_{\vec{u_1}} \vec{v} = \frac{\vec{v} \cdot \vec{u_1}}{\vec{u_1} \cdot \vec{u_1}} \vec{u_1} = \frac{4}{9} u_1$
   $\text{proj}_{\vec{u_2}} \vec{v} = \frac{\vec{v} \cdot \vec{u_2}}{\vec{u_2} \cdot \vec{u_2}} \vec{u_2} = \frac{65}{185} u_2$
   Coordinates: $\left(\frac{4}{9}, \frac{65}{185}\right)$
4. Give examples of the following.
   a. A matrix, A, in RREF, such that $\dim\left(\text{Row}(A)^\perp\right) = 1$ and $\dim\left(\text{Col}(A)^\perp\right) = 2$.
   $\dim\left(\text{Row}(A)^\perp\right) = 1 \Rightarrow A$ has 1 free variable
   $\dim\left(\text{Col}(A)^\perp\right) = 2 \Rightarrow A^\perp$ has 2 free variables
   $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$
   b. Two linearly independent vectors $\mathbb{R}^3$, $\vec{u}$ and $\vec{v}$, such that $\vec{u} \cdot \vec{x} = \vec{v} \cdot \vec{x} = 0$, where $\vec{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$
   $\vec{u} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \vec{v} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$

c. A 3×3 matrix in RREF, A, such that $\left(\mathrm{Null}(A)\right)^{\perp}$ is spanned by $\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$.

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

d. A non-zero vector, $\vec{w}$, whose projection $\mathrm{Col}(A)$ is $\vec{w}$, where $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$.

$$\vec{w} = \begin{pmatrix} \square \\ \square \\ \square \end{pmatrix}$$

Lecture 30

Wednesday, November 3, 2021 3:35 PM

Notes:

Example

The subspace W is a subspace of $\mathbb{R}^3$ perpendicular to $x = (1,1,1)$. Calculate the missing coefficients in the orthonormal basis for W.

$$u = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\-1\end{bmatrix} \qquad v = \frac{1}{\sqrt{6}}\begin{bmatrix}1\\-2\\1\end{bmatrix}$$

$$\vec{v} = \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} \qquad \begin{matrix}\vec{v}\cdot\vec{u} = 0 \Leftrightarrow x_1 = x_3\\ \vec{v}\cdot\vec{x} = 0 \Leftrightarrow x_1 + x_2 + x_3 = 0\end{matrix}$$

Orthogonal Matrices

An orthogonal matrix is a square matrix whose columns are orthonormal

Theorem

An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I_n$.

Can U have orthonormal columns if $n > m$? NO (needs to be linearly independent)

Proof:

U matrix with columns $\overrightarrow{u_1}, \dots, \overrightarrow{u_n}$

$$U^T U = \begin{pmatrix}\overrightarrow{u_1}\cdot\overrightarrow{u_1} & \overrightarrow{u_1}\cdot\overrightarrow{u_2} & \cdots & \overrightarrow{u_1}\cdot\overrightarrow{u_n}\\ \overrightarrow{u_2}\cdot\overrightarrow{u_1} & \overrightarrow{u_2}\cdot\overrightarrow{u_2} & \cdots & \overrightarrow{u_2}\cdot\overrightarrow{u_n}\\ & & \ddots & \vdots\\ & & & \overrightarrow{u_n}\cdot\overrightarrow{u_n}\end{pmatrix}$$

$$U^T U = I_n \Leftrightarrow \begin{cases}\overrightarrow{u_i}\cdot\overrightarrow{u_i} = 1 \text{ for } 1 \le i \le n\\ \overrightarrow{u_i}\cdot\overrightarrow{u_k} = 0 \text{ for i} \ne k\end{cases}$$

→ If U is square $U^{-1} = U^T$

Theorem

Assume $m \times n$ matrix U has orthonormal columns. Then,

1. (Preserves length) $\|U\vec{x}\| = \|\vec{x}\|$
2. (Preserves angle) $(U\vec{x})\cdot(U\vec{y}) = \vec{x}\cdot\vec{y}$
3. (Preserves orthogonality) $U\vec{x}\cdot U\vec{y} = 0 \Leftrightarrow \vec{x}\cdot\vec{y} = 0$

Proof:

2. $(U\vec{x})\cdot(U\vec{y}) = (U\vec{x})^T\cdot(U\vec{y}) = \vec{x}^T\underbrace{U^T U}_{I_n}\vec{y} = \vec{x}^T\vec{y} = \vec{x}\cdot\vec{y}$

   From 2 we have,

   $\|U\vec{x}\|^2 = U\vec{x}\cdot U\vec{x} = \vec{x}\cdot\vec{x} = \|\vec{x}\|^2$

Example

Compute the length of the vector below.

$$\begin{bmatrix}1/2 & 2/\sqrt{14}\\ 1/2 & 1/\sqrt{14}\\ 1/2 & -3/\sqrt{14}\\ 1/2 & 0\end{bmatrix}\begin{bmatrix}\sqrt{2}\\-3\end{bmatrix}$$

$\begin{pmatrix}1/2\\1/2\\1/2\\1/2\end{pmatrix}$ and $\begin{pmatrix}2/\sqrt{14}\\1/\sqrt{14}\\-3/\sqrt{14}\\0\end{pmatrix}$ are orthonormal

$$\to \left\|\begin{bmatrix}1/2 & 2/\sqrt{14}\\ 1/2 & 1/\sqrt{14}\\ 1/2 & -3/\sqrt{14}\\ 1/2 & 0\end{bmatrix}\begin{bmatrix}\sqrt{2}\\-3\end{bmatrix}\right\| = \left\|\begin{bmatrix}\sqrt{2}\\-3\end{bmatrix}\right\| = \sqrt{11}$$

Section 6.3: Orthogonal Projections

Ex.1

Let $\overrightarrow{u_1}, \dots, \overrightarrow{u_5}$ be an orthonormal basis for $\mathbb{R}^5$. Let $W = \text{Span}(\overrightarrow{u_1}, \overrightarrow{u_2})$. For a vector $\vec{y} \in \mathbb{R}^5$, write $\vec{y} = \hat{y} + w^\perp$, where $\hat{y} \in W$ and $w^\perp \in W^T$.

$\vec{y} \in \mathbb{R}^5$ and $\{\overrightarrow{u_1}, \dots, \overrightarrow{u_5}\}$ basis of $\mathbb{R}^5$

$$\vec{y} = c_1\overrightarrow{u_1} + c_2\overrightarrow{u_2} + c_3\overrightarrow{u_3} + c_4\overrightarrow{u_4} + c_5\overrightarrow{u_5}$$

$$c_q = \vec{y}\cdot\overrightarrow{u_q} \text{ for } 1 \le q \le 5$$

$$\vec{y} = \underbrace{c_1\overrightarrow{u_1} + c_2\overrightarrow{u_2}}_{\in W} + \underbrace{c_3\overrightarrow{u_3} + c_4\overrightarrow{u_4} + c_5\overrightarrow{u_5}}_{\in w^T}$$

$$\vec{y} = \hat{y} + w^T$$

$$\hat{y} = c_1\overrightarrow{u_1} + c_2\overrightarrow{u_2} \in W$$

$$w^T = c_3\overrightarrow{u_3} + c_4\overrightarrow{u_4} + c_5\overrightarrow{u_5} \in w^T$$

Remark:

If $\vec{y} \in W$: $\vec{y} = \hat{y}$, $w^\perp = \vec{0}$

If $\vec{y} \in w^{\perp}$: $\hat{y} = \vec{0}$, $\vec{y} = w^{\perp}$

Orthogonal Decomposition Theorem
Theorem
Let W be a subspace of $\mathbb{R}^n$. Then, each vector $\vec{y} \in \mathbb{R}^n$ has the unique decomposition

$$\vec{y} = \hat{y} + w^{\perp},\ \hat{y} \in W,\ w^{\perp} \in W^{\perp}.$$

And, if $\overrightarrow{u_1}, \dots, \overrightarrow{u_p}$ is any orthogonal basis for W.

$$\hat{y} = \frac{\vec{y} \cdot \overrightarrow{u_1}}{\overrightarrow{u_1} \cdot \overrightarrow{u_1}}\overrightarrow{u_1} + \cdots + \frac{\vec{y} \cdot \overrightarrow{u_p}}{\overrightarrow{u_p} \cdot \overrightarrow{u_p}}\overrightarrow{u_p}$$

We say that $\hat{y}$ is the orthogonal projection of $\vec{y}$ onto W.

Explanation
We can write

$$\hat{y} = \sum_{i=1}^{P} \frac{\vec{y} \cdot \overrightarrow{u_i}}{\overrightarrow{u_i} \cdot \overrightarrow{u_i}}\overrightarrow{u_i} \in W = \mathrm{Span}(\overrightarrow{u_1}, \dots, \overrightarrow{u_p})\ \ \overrightarrow{u_q} \cdot \overrightarrow{u_i} = 0 \Leftrightarrow i \neq q$$

Then, $w^{\perp} = \vec{y} - \hat{y}$ because

$$w^{\perp} \cdot \overrightarrow{u_q} \cdot (\vec{y} - \hat{y}) = \overrightarrow{u_q} \cdot \vec{y} - \hat{y} \cdot \left(\sum_{i=1}^{P} \frac{\vec{y} \cdot \overrightarrow{u_i}}{\overrightarrow{u_i} \cdot \overrightarrow{u_i}}\overrightarrow{u_i}\right)$$

$\rightarrow w^{\perp}$ is orthogonal to $\overrightarrow{u_1}, \dots, \overrightarrow{u_p}$

$$w^{\perp} \in W^{\perp}$$

Ex.2a

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

Construct the decomposition $\vec{y} = \hat{y} + w^{\perp}$ where $\hat{y}$ is the orthogonal projection of $\vec{y}$ onto the subspace $W = \mathrm{Span}\{\overrightarrow{u_1}, \overrightarrow{u_2}\}$.

$$\hat{y} = \frac{\vec{y} \cdot \overrightarrow{u_1}}{\overrightarrow{u_1} \cdot \overrightarrow{u_1}}\overrightarrow{u_1} + \frac{\vec{y} \cdot \overrightarrow{u_2}}{\overrightarrow{u_2} \cdot \overrightarrow{u_2}}\overrightarrow{u_2} = \frac{8}{8}\overrightarrow{u_1} + \frac{3}{1}\overrightarrow{u_2} = \overrightarrow{u_1} + 3\overrightarrow{u_2} \in W = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}$$

$$\rightarrow w^{\perp} = \vec{y} - \hat{y} = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}$$

Check: $w^{\perp} \cdot \overrightarrow{u_1} = w^{\perp} \cdot \overrightarrow{u_2} = 0 \in W^{\perp}$

Best Approximation Theorem
Theorem
Let W be a subspace of $\mathbb{R}^n$, $\vec{y} \in \mathbb{R}^n$, and $\hat{y}$ is the orthogonal projection of $\vec{y}$ onto W. Then for any $\vec{w} \neq \vec{y} \in W$, we have

$$\|\vec{y} - \hat{y}\| < \|\vec{y} - \vec{w}\|$$

That is, $\hat{y}$ is the unique vector in W that is closest to $\vec{y}$.

Proof
The orthogonal projection of $\vec{y}$ onto W is the closest point in W to $\vec{y}$

$$\vec{v} \in W \quad \vec{y} - \hat{y} = w^{\perp} \in W^{\perp}$$
$$\vec{v} \neq \hat{y} \quad \vec{v} - \hat{y} \in W$$

Pythagorean Theorem:

$$\|\vec{y} - \vec{v}\|^2 = \|\vec{y} - \hat{y}\|^2 + \underbrace{\|\vec{v} - \hat{y}\|^2}_{\neq 0}$$

$$\|\vec{y} - \vec{v}\|^2 > \|\vec{y} - \hat{y}\|^2 \Rightarrow \|\vec{y} - \vec{v}\| > \|\vec{y} - \hat{y}\|$$

Ex.2b

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

What is the distance between $\vec{y}$ and subspace $W = \mathrm{Span}\{\overrightarrow{u_1}, \overrightarrow{u_2}\}$? Note that these are the same vectors in **Ex.2a**

The distance between $\vec{y}$ and subspace W is $\|w^{\perp}\| = \sqrt{8} = 2\sqrt{2}$

$A\vec{x} = \vec{b}$ is consistent iff $\vec{b} \in \mathrm{Col}(A)$

Studio 19 & 20

Thursday, November 4, 2021 12:36 PM

Gram-Schmidt:

- Process for converting basis to an orthonormal one
- Idea: successively subtract projection of current vector onto previous ones
- Start with basis $\{\overrightarrow{v_1}, \dots, \overrightarrow{v_n}\}$

$$\overrightarrow{v_1} = \overrightarrow{x_1}$$

$$\overrightarrow{v_2} = \overrightarrow{x_2} - \frac{\overrightarrow{x_2} \cdot \overrightarrow{v_1}}{\overrightarrow{v_1} \cdot \overrightarrow{v_1}} \overrightarrow{v_1}$$

$$\overrightarrow{v_3} = \overrightarrow{x_3} - \frac{\overrightarrow{x_3} \cdot \overrightarrow{v_1}}{\overrightarrow{v_1} \cdot \overrightarrow{v_1}} \overrightarrow{v_1} - \frac{\overrightarrow{x_3} \cdot \overrightarrow{v_2}}{\overrightarrow{v_2} \cdot \overrightarrow{v_2}} \overrightarrow{v_2}$$

$$\vdots$$

$$\overrightarrow{v_n} = \overrightarrow{x_n} - \frac{\overrightarrow{x_n} \cdot \overrightarrow{v_1}}{\overrightarrow{v_1} \cdot \overrightarrow{v_1}} \overrightarrow{v_1} - \cdots - \frac{\overrightarrow{x_n} \cdot \overrightarrow{v_{p-1}}}{\overrightarrow{v_{p-1}} \cdot \overrightarrow{v_{p-1}}} \overrightarrow{v_{p-1}}$$

Worksheet 6.3 and 6.4: Orthogonal Projections, The Gram-Schmidt Process

1. $\vec{y} = \begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix}$, $\overrightarrow{u_1} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $\overrightarrow{u_2} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$
   a. Determine whether $\overrightarrow{u_1}$ and $\overrightarrow{u_2}$
      i. Are linearly independent
         1) True
      ii. Are mutually orthogonal
         1) True
      iii. Are orthonormal
         1) False
      iv. Span $\mathbb{R}^3$
         1) False
   b. Is $\vec{y}$ in $W = \mathrm{Span}(\overrightarrow{u_1}, \overrightarrow{u_2})$
      i. False
   c. Compute the vector, $\hat{y} \in W$, that most closely appromixmates $\vec{y}$.
      i. $\hat{y} = \mathrm{proj}_{\vec{W}} \vec{y} = \mathrm{proj}_{\overrightarrow{u_1}} \vec{y} + \mathrm{proj}_{\overrightarrow{u_2}} \vec{y} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$
   d. Construct a vector, $\vec{z}$, that is in $W^{\perp}$
      i. $\begin{pmatrix} 0 \\ 0 \\ 4 \end{pmatrix} = \vec{y} - \hat{y}$
2. Compute the QR decomposition of $A = \begin{pmatrix} 1 & 5 \\ 3 & 1 \\ -2 & 4 \end{pmatrix}$
   a. $A = QR$, Q has orthonomal columns; R upper triangular
   b. To find Q, run G-S on columns of A to get $Q = \{\overrightarrow{v_1}, \dots, \overrightarrow{v_n}\}$
   c. $\overrightarrow{v_1} = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} \Rightarrow \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$
   d. $\overrightarrow{v_2} = \begin{pmatrix} 5 \\ 1 \\ 4 \end{pmatrix} - \frac{0}{14} \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ 4 \end{pmatrix} \Rightarrow \frac{1}{\sqrt{42}} \begin{pmatrix} 5 \\ 1 \\ 4 \end{pmatrix}$
   e. $Q = \begin{pmatrix} 1/\sqrt{14} & 5/\sqrt{42} \\ 3/\sqrt{14} & 1/\sqrt{42} \\ -2/\sqrt{14} & 4/\sqrt{42} \end{pmatrix}$
   f. $R = Q^T A = \begin{pmatrix} 1/\sqrt{14} & 3/\sqrt{14} & -2/\sqrt{14} \\ 5/\sqrt{42} & 1/\sqrt{42} & 4/\sqrt{42} \end{pmatrix} \begin{pmatrix} 1 & 5 \\ 3 & 1 \\ -2 & 4 \end{pmatrix} = \begin{pmatrix} 14/\sqrt{14} & 0 \\ 0 & 5/\sqrt{10} \end{pmatrix}$
3. $\{\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}\}$ is an orthogonal basis for subspace V. Classify each set as a bassi for V, an orthogonal basis for V, or not a basis for V.
   a. $\{3\overrightarrow{v_3}, 2\overrightarrow{v_1}, \overrightarrow{v_2}\}$
      i. $(3v_3) \cdot (2v_1) = 6\left(\underbrace{v_3 \cdot v_1}_{=0}\right) = 0$
      ii. $\Rightarrow$ Orthogonal basis
   b. $\{(\overrightarrow{v_1} + \overrightarrow{v_2}), (\overrightarrow{v_1} - \overrightarrow{v_2}), \overrightarrow{v_3}\}$
      i. $(\overrightarrow{v_1} + \overrightarrow{v_2}) \cdot (\overrightarrow{v_1} - \overrightarrow{v_2}) = v_1 \cdot v_1 - v_1 \cdot v_2 + v_2 \cdot v_1 - v_2 \cdot v_2 = \|u_1\|^2 + \|u_2\|^2 \neq 0$
      ii. $\Rightarrow$ Not necessary an orthogonal basis
   c. $\{(\overrightarrow{v_1} + \overrightarrow{v_2}), (\overrightarrow{v_1} - \overrightarrow{v_2}), (\overrightarrow{v_3} - \overrightarrow{v_1})\}$
      i. $(\overrightarrow{v_1} + \overrightarrow{v_2}) \cdot (\overrightarrow{v_1} - \overrightarrow{v_2}) = v_1 \cdot v_1 - v_1 \cdot v_2 + v_2 \cdot v_1 - v_2 \cdot v_2 = \|u_1\|^2 + \|u_2\|^2 \neq 0$
      ii. $\Rightarrow$ Not necessary an orthogonal basis
4. Indicate whether he statement are true or false.
   a. If $\vec{y}$ is in subspace W, the orthogonal projection of $\vec{y}$ onto W is $\vec{y}$
      i. True
   b. If c is orthogonal to $\vec{v}$ and $\vec{w}$, then $\vec{x}$ is also orthogonal to $\vec{v} - \vec{w}$
      i. True
5. If possible, give an example of:
   a. Two linearly independent vectors that are orthogonal to $\begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$.
      i. $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$
   b. A subspace of $\mathbb{R}^3$, S, such that $\dim(S^{\perp}) = 2$
      i. Span$\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}$
6. **Written Explanation Exercise** Let $u_1, \dots, u_k$ be an orthonormal family of vectors in $\mathbb{R}^n$. Explain why applying the Gram-Schmidt process to the pivotal columns of the $n \times (n+k)$ matrix $A = [u_1 \dots u_k e_1 \dots e_n]$ gives an orthonormal basis of $\mathbb{R}^n$ that contains $u_1, \dots, u_k$.

Lecture 31

Friday, November 5, 2021 3:30 PM

Notes:

Section 6.4: The Gram-Schmidt Process

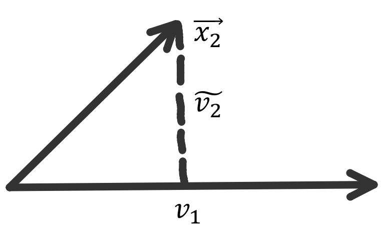
Ex.1

The vectors below span a subspace W of $\mathbb{R}^4$. Construct an orthogonal basis for W.

$$\overrightarrow{x_1} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \quad \overrightarrow{x_2} = \begin{bmatrix}0\\1\\1\\1\end{bmatrix}, \quad \overrightarrow{x_3} = \begin{bmatrix}0\\0\\1\\1\end{bmatrix}$$

Idea: $W_1 = \mathrm{Span}(\overrightarrow{x_1})$
$W_2 = \mathrm{Span}(\overrightarrow{x_1}, \overrightarrow{x_2})$
$W_2 = \mathrm{Span}(\overrightarrow{x_1}, \overrightarrow{x_2}, \overrightarrow{x_3}) = W$

1. Orthogonal basis for W_1
   $\rightarrow \overrightarrow{v_1} = \overrightarrow{x_1}$
2. Orthogonal basis for W_2
   $$\widetilde{v_2} = \overrightarrow{x_2} - \mathrm{proj}_{\overrightarrow{W_1}}\overrightarrow{x_2}$$
   $$= \overrightarrow{x_2} - \frac{\overrightarrow{x_2}\cdot\overrightarrow{v_1}}{\overrightarrow{v_1}\cdot\overrightarrow{v_1}}\overrightarrow{v_1} \in W_2$$
   $$= \begin{pmatrix}0\\1\\1\\1\end{pmatrix} - \frac{3}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} = \frac{1}{4}\begin{pmatrix}-3\\1\\1\\1\end{pmatrix}$$



Trick: take $\overrightarrow{v_2} = \begin{pmatrix}-3\\1\\1\\1\end{pmatrix}$

3. Orthogonal basis for W_3
   $$\widetilde{v_3} = \overrightarrow{x_3} - \mathrm{proj}_{W_2}\overrightarrow{x_3} = \overrightarrow{x_3} - \frac{\overrightarrow{x_3}\cdot\overrightarrow{v_1}}{\overrightarrow{v_1}\cdot\overrightarrow{v_1}}\overrightarrow{v_1} - \frac{\overrightarrow{x_3}\cdot\overrightarrow{v_2}}{\overrightarrow{v_2}\cdot\overrightarrow{v_2}}\overrightarrow{v_2} = \frac{1}{3}\begin{pmatrix}0\\-2\\1\\1\end{pmatrix}$$

Trick: take $\overrightarrow{v_3} = \begin{pmatrix}0\\-2\\1\\1\end{pmatrix}$

$\rightarrow \{\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}\}$ is an orthogonal basis of W

The Gram-Schmidt Process

Given a basis $\{\overrightarrow{x_1}, \dots, \overrightarrow{x_p}\}$ for a subspace W of $\mathbb{R}^n$, iteratively define

$$\overrightarrow{v_1} = \overrightarrow{x_1}$$
$$\overrightarrow{v_2} = \overrightarrow{x_2} - \frac{\overrightarrow{x_2}\cdot\overrightarrow{v_1}}{\overrightarrow{v_1}\cdot\overrightarrow{v_1}}\overrightarrow{v_1}$$
$$\overrightarrow{v_3} = \overrightarrow{x_3} - \frac{\overrightarrow{x_3}\cdot\overrightarrow{v_1}}{\overrightarrow{v_1}\cdot\overrightarrow{v_1}}\overrightarrow{v_1} - \frac{\overrightarrow{x_3}\cdot\overrightarrow{v_2}}{\overrightarrow{v_2}\cdot\overrightarrow{v_2}}\overrightarrow{v_2}$$
$$\vdots$$
$$\overrightarrow{v_n} = \overrightarrow{x_n} - \frac{\overrightarrow{x_n}\cdot\overrightarrow{v_1}}{\overrightarrow{v_1}\cdot\overrightarrow{v_1}}\overrightarrow{v_1} - \dots - \frac{\overrightarrow{x_n}\cdot\overrightarrow{v_{p-1}}}{\overrightarrow{v_{p-1}}\cdot\overrightarrow{v_{p-1}}}\overrightarrow{v_{p-1}}$$

Proof

$\rightarrow$ See example 1

Geometric Interpretation

Suppose $\overrightarrow{x_1}, \overrightarrow{x_2}, \overrightarrow{x_3}$ are linearly independent vectors in $\mathbb{R}^3$. We wish to construct an orthogonal basis for the space that they span.

We construct vectors $\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}$, which form our orthogonal basis. $W_1 = \mathrm{Span}(\overrightarrow{v_1})$, $W_2 = \mathrm{Span}(\overrightarrow{v_1}, \overrightarrow{v_2})$

Orthonormal Bases

Definition

A set of vectors form an orthonormal basis if the vectors are mutually orthogonal and have unit length.

Example

The two vectors below form an orthogonal basis for a subspace W. Obtain an orthonormal basis for W.

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

$\left\{\frac{1}{\sqrt{13}}\begin{bmatrix}3\\2\\0\end{bmatrix}, \frac{1}{\sqrt{14}}\begin{bmatrix}-2\\3\\1\end{bmatrix}\right\}$ is an orthonormal basis for W

QR Factorization

Theorem

Any $m \times n$ matrix A with linearly independent columns has the QR factorization

$$A = QR$$

Where

1. Q is $m \times n$, its columns are an orthonormal basis for $\mathrm{Col}(A)$
2. R is $n \times n$, upper triangular, with positive entries on its diagonal, and the length of the j^{th} column of R is equal to the length of the j^{th} column of A.

Proof

$A = (\overrightarrow{a_1}, \dots, \overrightarrow{a_n}) : \overrightarrow{a_1}, \dots, \overrightarrow{a_n}$ linearly independent
$Q = (\overrightarrow{q_1}, \dots, \overrightarrow{q_n}) : \overrightarrow{q_1}, \dots, \overrightarrow{q_n}$ orthonormal basis for $\mathrm{Col}(A)$
Obtained by Gram-Schmidt

Section 6.3: decomposition of a matrix in an orthonormal basis:

$\overrightarrow{a_1} = (\overrightarrow{a_1} \cdot \overrightarrow{q_1}) \cdot \overrightarrow{q_1} = \underbrace{r_{11}}_{=0} \cdot \overrightarrow{q_1}$

$\overrightarrow{a_2} = (\overrightarrow{a_2} \cdot \overrightarrow{q_1}) \cdot \overrightarrow{q_1} + (\overrightarrow{a_2} \cdot \overrightarrow{q_2}) \cdot \overrightarrow{q_2} += r_{12} \cdot \overrightarrow{q_1} + \underbrace{r_{22}}_{=0} \cdot \overrightarrow{q_2}$

$\vdots$ if $r_{22} < 0$: change $\overrightarrow{q_2}$ change $-\overrightarrow{q_2}$ in Q

$\overrightarrow{a_n} = (\overrightarrow{a_n} \cdot \overrightarrow{q_1}) \cdot \overrightarrow{q_1} + \cdots + (\overrightarrow{a_n} \cdot \overrightarrow{q_n}) \cdot \overrightarrow{q_n} = r_{1n} \cdot \overrightarrow{q_1} + \cdots + \underbrace{r_{nn}}_{\neq 0} \cdot \overrightarrow{q_n}$

Define: $R = \begin{pmatrix} r_{11} & r_{22} & \cdots & r_{1n} \\ & \ddots & & \vdots \\ & (0) & \ddots & \vdots \\ & & & r_{nn} \end{pmatrix} = (\overrightarrow{r_1} \dots \overrightarrow{r_n})$

$QR = \begin{pmatrix} \underbrace{Q\overrightarrow{r_1}}_{Q_x \begin{pmatrix} r_{11} \\ 0 \\ 0 \end{pmatrix} = r_{11} \cdot \overrightarrow{q_1} = \overrightarrow{q_1}} & \dots & \underbrace{Q\overrightarrow{r_n}}_{Q_x \begin{pmatrix} r_{1n} \\ \vdots \\ r_{nn} \end{pmatrix} = r_{1n} \cdot \overrightarrow{q_1} \underbrace{+ \cdots +}_{a_n} r_{nn} \overrightarrow{q_1}} \end{pmatrix}$

$\rightarrow A = QR$

Lecture 32

Monday, November 8, 2021 3:37 PM

Notes:

Ex.

Construct the QR decomposition for $A = \begin{bmatrix} 3 & -2 \\ 2 & 3 \\ 0 & 1 \end{bmatrix} = \underbrace{(\vec{a_1} \quad \vec{a_2})}_{\vec{a_1}\cdot\vec{a_2}=0}$

Why?: $A\vec{x} = \vec{b}$

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

$\underbrace{\rightarrow R\vec{x} = Q^T\vec{b}}_{\text{triangular system}}$

$$Q = \begin{pmatrix} 3/\sqrt{13} & -2/\sqrt{14} \\ 2/\sqrt{13} & 3/\sqrt{14} \\ 0 & 1/\sqrt{14} \end{pmatrix}$$

$$R = Q^T A \begin{pmatrix} 3/\sqrt{13} & 2/\sqrt{13} & 0 \\ -2/\sqrt{14} & 3/\sqrt{14} & 3/\sqrt{14} \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 2 & 3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{13} & 0 \\ 0 & \sqrt{14} \end{pmatrix}$$

Further Example

Construct the QR decomposition for $A = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 0 & 2 \\ -1 & 1 & -3 \\ 1 & 0 & 1 \end{bmatrix}$

* orthogonal basis of $\mathrm{Col}(A)$: $\vec{v_1}, \vec{v_2}, \vec{v_3}$

$$\vec{v_1} = \vec{a_1} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

$$\vec{v_2} = \vec{a_2} - \frac{\vec{a_2}\cdot\vec{v_1}}{\vec{v_1}\cdot\vec{v_1}}\vec{v_1} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \frac{-2}{3}\begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/3 \\ 0 \\ 1/3 \\ 2/3 \end{pmatrix}, \text{take } \vec{v_2} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix}$$

$$\vec{v_3} = \vec{a_3} - \frac{\vec{a_3}\cdot\vec{v_1}}{\vec{v_1}\cdot\vec{v_1}}\vec{v_1} - \frac{\vec{a_3}\cdot\vec{v_2}}{\vec{v_2}\cdot\vec{v_2}}\vec{v_2} = \begin{pmatrix} -1 \\ 2 \\ -3 \\ 1 \end{pmatrix} - \frac{3}{3}\begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix} - 0\vec{v_2} = \begin{pmatrix} -2 \\ 2 \\ -2 \\ 0 \end{pmatrix}, \text{take } \vec{v_3} = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

$$\rightarrow Q = \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{6} & -1/\sqrt{3} \\ 0 & 0 & 1/\sqrt{3} \\ -1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{3} \\ 1/\sqrt{3} & 2/\sqrt{6} & 0 \end{pmatrix}$$

$$\rightarrow R = Q^T\vec{b} = \begin{pmatrix} 1/\sqrt{3} & 0 & -1/\sqrt{3} & 1/\sqrt{3} \\ -1/\sqrt{6} & 0 & 1/\sqrt{6} & 2/\sqrt{6} \\ -1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & 2 \\ -1 & 1 & -3 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{3} & -2/\sqrt{3} & \sqrt{3} \\ 0 & 2/\sqrt{6} & 0 \\ 0 & 0 & 6/\sqrt{3} \end{pmatrix}$$

Check: $\|\vec{r_2}\| = \sqrt{\frac{4}{3} + \frac{4}{6}} = \sqrt{2} = \|\vec{a_2}\|$

$\|\vec{r_3}\| = \sqrt{3 + \frac{36}{3}} = \sqrt{15} = \|\vec{a_3}\|$

If you find Q and

$$R = \begin{pmatrix} 1 & * & * \\ 0 & -2 & * \\ 0 & 0 & 3 \end{pmatrix}$$

$$A = QR = Q\overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{\tilde{C}:C_2 \leftarrow -C_2}\overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{\tilde{R}:R_2 \leftarrow -R_2} R$$

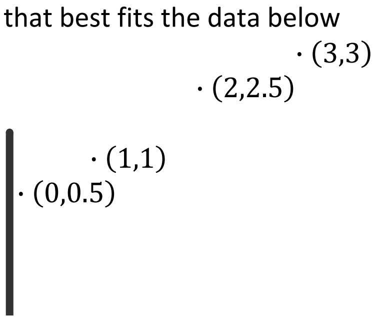
Section 6.5: Least-Square Problems

Inconsistent Systems

Suppose we want to construct a line of the form

$y = mx + b$

that best fits the data below



$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} 0.5 \\ 1 \\ 2.5 \\ 3 \end{bmatrix}$$

Can we ‘solve’ this inconsistent system?

If $y = mx + b$
First Part: $(x, y) = (0,0.5)$: $0.5 = b$
Second Part: $(x, y) = (1,1)$: $1 = m + b \Rightarrow m = 0.5$
Third Part: $(x, y) = (2,2.5)$: $2.5 = \underbrace{2m + b}_{=1.5} \Rightarrow NO$

The Least Squares Solution to a Linear System
Definition: Least Squares Solution
Let A be a $m \times n$ matrix. A least squares solution to $A\vec{x} = \vec{b}$ is the solution $\hat{x}$ for which

$$\|\vec{b} - A\hat{x}\| \leq \|\vec{b} - A\vec{x}\|$$

for all $\vec{x} \in \mathbb{R}^n$.

If $\vec{x}$ is a solution:
$$\|\vec{b} - A\vec{x}\| = \|\vec{0}\| = 0$$
Here: if the system is inconsistent:
$$\|\vec{b} - A\vec{x}\| > 0 \quad \forall \vec{x} \in \mathbb{R}^n$$
Least-Square solution: $\hat{x}$:
$$\|\vec{b} - A\hat{x}\| \leq \|\vec{b} - A\vec{x}\| \forall \vec{x} \in \mathbb{R}^n$$

Lecture 33

Wednesday, November 10, 2021 10:18 PM

<u>Notes:</u>

A Geometric Interpretation

The vector $\vec{b}$ is closer to $A\vec{x}$ than to $A\vec{x}$ for all other $\vec{x} \in \mathrm{Col}(A)$.

1. If $\vec{b} \in \mathrm{Col}(A)$, then $\hat{x}$ is a solution to $A\vec{x} = \vec{b}$
2. Seek $\hat{x}$ so that $A\hat{x}$ is as close to $\vec{b}$ as possible. That is, $\hat{x}$ should solve $A\hat{x} = \hat{b}$ where $\hat{b}$ is the orthogonal projection of $\vec{b}$ onto $\mathrm{Col}(A)$

The Normal Equations

Theorem (Normal Equations for Least Squares)

The least squares solutions to $A\vec{x} = \vec{b}$ coincide with the solutions to

$$\underbrace{A^T A\vec{x} = A^T\vec{b}}_{\text{Normal Equations}}$$

Derivation

1. $\hat{x}$ is the least squares solution, is equivalent to $\vec{b} - A\hat{x}$ is orthogonal to $\mathrm{Col}(A)$ and $\left(\mathrm{Col}(A)\right)^{\perp} = \mathrm{Null}(A)$
2. A vector $\vec{v}$ is $\left(\mathrm{Null}(A)\right)^{\perp}$ if and only if $A^T\vec{v} = \vec{0}$
3. So we obtain the Normal Equations:

   $\hat{x}$ least square solution

   Iff $\vec{b} - A\hat{x} \in \mathrm{Null}(A^T)$

   Iff $A^T(\vec{b} - A\hat{x}) = 0$

   Iff $\underbrace{A^T A\hat{x} = A^T\vec{b}}_{\text{Normal Equations}}$

Ex.

Compute the least squares solution to $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Solution:

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix}$$

$$A^T\vec{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

The normal equations $A^T A\hat{x} = A^T\vec{b}$ because

$$\begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix}\hat{x} = \begin{pmatrix} 19 \\ 11 \end{pmatrix}$$

$$B^{-1} = \frac{1}{85 - 1}\begin{pmatrix} 5 & -1 \\ -1 & 17 \end{pmatrix}$$

$$\rightarrow \vec{x} = \frac{1}{84}\begin{pmatrix} 5 & -1 \\ -1 & 17 \end{pmatrix}\begin{pmatrix} 19 \\ 11 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Theorem

Theorem (Unique Solutions for Least Squares)

Let A be any $m \times n$ matrix. These statements are equivalent.

1. The equation $A\vec{x} = \vec{b}$ has a unique least-squares solution for each $\vec{b} \in \mathbb{R}^m$.
2. The columns of A are linearly independent.
3. The matrix $A^T A$ is invertible.

And, if these statemetns hold, the least square solution is

$$\hat{x} = (A^T A)^{-1} A^T\vec{b}$$

Use heuristic: $A^T A$ plays the role of 'length-squared' of the matrix A.

Ex.2

Compute the least squares solution to $A\vec{x} = \vec{b}$

$$A = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

Hint: the columns of A are orthogonal.

$A = (\overrightarrow{a_1} \quad \overrightarrow{a_2})$: linearly independent (unique least-square solution), easy to compute $\hat{b}$

$$\hat{b} = \frac{\vec{b} \cdot \overrightarrow{a_1}}{\overrightarrow{a_1} \cdot \overrightarrow{a_1}}\overrightarrow{a_1} + \frac{\vec{b} \cdot \overrightarrow{a_2}}{\overrightarrow{a_2} \cdot \overrightarrow{a_2}}\overrightarrow{a_2} = \frac{8}{4}\overrightarrow{a_1} + \frac{45}{90}\overrightarrow{a_2} = 2\overrightarrow{a_1} + \frac{1}{2}\overrightarrow{a_2}$$

$$\rightarrow \hat{x} = \begin{pmatrix} 2 \\ 1/2 \end{pmatrix}$$

Lecture 34

Friday, November 12, 2021 3:23 PM

Notes:

Theorem (Least Squares and QR decomposition. Then for each $\vec{b} \in \mathbb{R}^m$ the equation $A\vec{x} = \vec{b}$ has the unique least squares solution

$$R\hat{x} = Q^T\vec{b}.$$

(Remember, R is upper triangular, so the equation above is solved by back-substitution.)

$A = QR$

Normal Equations: $A^TA\vec{x} = A^T\vec{b}$

$R^TQ^TQR\vec{x} = R^TQ^T\vec{b}$

$R^TR\vec{x} = R^TQ^T\vec{b}$

$R\vec{x} = Q^T\vec{b}$

Ex.3 Compute the least squares solution to $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$$

Solution. The QR decomposition of A is

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

$$Q^T\vec{b} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}\begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \\ 4 \end{bmatrix}$$

And then we solve by backend substitution $R\vec{x} = Q^T\vec{b}$

$$\begin{bmatrix} 2 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \\ 4 \end{bmatrix} \rightarrow \begin{matrix} x_3 = 2 \\ x_2 = -6 \\ x_1 = 10 \end{matrix} \rightarrow \vec{x} = \begin{pmatrix} 10 \\ -6 \\ 2 \end{pmatrix}$$

Section 6.6: Applications to Linear Models

The Least Squares Line

Graph below gives an approximate linear relationship between x and y.

1. Black circles are data.
2. Blue line is the least squares line.
3. Lengths of red lines are the residuals

The least squares line minimizes the sum of squares of the residuals

Ex.1 Compute the least squares line $y = \beta_0 + \beta_1 x$ that best fits the data

X	2	5	7	8
Y	1	1	4	3

We want to solve

$$\begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix}\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 4 \\ 3 \end{bmatrix}$$

This is a least-squares problem: $X\vec{\beta} = \vec{y}$

The normal equations are $X^TX = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$

$$X^T\vec{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 59 \end{bmatrix}$$

So the least-squares solution is given by:

$$\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 59 \end{bmatrix}$$

$$y = \beta_0 + \beta_1 x = -\frac{5}{21} + \frac{19}{42}x$$

As we may have guessed, β_0 is negative, and β_1 is positive.

Least Squares Fitting for Other Curves

We con consider the least squares fitting for the form

$y = c_0 + c_1 f_1(x) + c_2 + c_2 f_2(x) + \cdots + c_k + c_k f_k(x)$

If functions f_i are known, this is a linear problem in the c_i variables.

Ex.

Consider the data in the table below

X	1	0	0	1
Y	2	1	0	6

Determine the coefficients c_1 and c_2 for the curve $y = c_1 x + c_2 x^2$ that best fits the data.

$x\vec{\beta} = \vec{y}$

$$\begin{pmatrix} 1 & x_1 & \cos(x_1) & \sin(x_1) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & \cos(x_n) & \sin(x_n) \end{pmatrix}\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Normal Equations

$$X^T X = \begin{pmatrix} -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$X^T \vec{y} = \begin{pmatrix} -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 0 \\ 6 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}$$

$$\left(\begin{array}{cc|c} 2 & 0 & 4 \\ 0 & 2 & 8 \end{array}\right) \to c_1 = 2, c_2 = 4$$

Projection method:

$$y = \mathrm{Proj}_{\mathrm{Col}(A)} \vec{y} = \frac{\vec{y} \cdot \overrightarrow{x_1}}{x_1 \cdot x_1} x_1 + \frac{\vec{y} \cdot \overrightarrow{x_2}}{x_2 \cdot x_2} x_2$$

$$= \frac{4}{2} \overrightarrow{x_1} + \frac{8}{2} \overrightarrow{x_2} = X \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

$$\to \vec{c} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

Studio 21

Thursday, November 11, 2021 12:31 PM

Least Squares:

Want to solve $A\vec{x}=\vec{b}$ but this system is inconsistent. Instead, find $\vec{x}$ to minimize $\|A\vec{x}-\vec{b}\|$

Def. $\hat{x}$ is a least squares solution to $A\vec{x}=\vec{b}$ if $\|A\hat{x}-\vec{b}\| \leq \|A\vec{x}-\vec{b}\|$

If A has linearly independent columns, then $\hat{x}=(A^TA)^{-1}A^T\vec{b}$

Suppose A has independent columns. $A=QR$, $Q=[\overrightarrow{u_1}\ldots\overrightarrow{u_n}]$ which forms $\{\overrightarrow{u_1},\ldots,\overrightarrow{u_n}\}$ orthonormal basis for $\mathrm{Col}(A)$

$$\Rightarrow \hat{b}=\mathrm{Proj}_{\mathrm{Col}(A)}\vec{b}=\mathrm{Proj}_{\overrightarrow{u_1}}\vec{b}+\cdots\mathrm{Proj}_{\overrightarrow{u_n}}\vec{b}=(\overrightarrow{u_1}\cdot\vec{b})\overrightarrow{u_1}+\cdots+(\overrightarrow{u_n}\cdot\vec{b})\overrightarrow{u_n}=Q\begin{bmatrix}\overrightarrow{u_1}\cdot\vec{b}\\ \vdots\\ \overrightarrow{u_n}\cdot\vec{b}\end{bmatrix}=QQ^T\vec{b}$$

$$A\hat{x}=QQ^T\vec{b}\Rightarrow A^TA\hat{x}=A^TQQ^T\vec{b}=R^TQ^TQQ^T\vec{b}=A^T\vec{b}\Rightarrow\hat{x}=(A^TA)^{-1}A^T\vec{b}$$

Worksheet 6.5 and 6.6: Least-Squares Problems, Applications to Linear Models

Worksheet Exercises

1. Fill in the blanks. These questions concern that least squares solution $\hat{x}$ to $A\vec{x}=\vec{b}$.
   a. If $A=QR$, then $A^TA=R^TR$.
   b. If the columns of A are linearly independent, then $\hat{x}=(A^TA)^{-1}A^T\vec{b}$
   c. If $\vec{b}$ is in the column space of A, then $A\hat{x}=\vec{b}$.
   d. If $A=QR$, then R is invertible then $\hat{x}=R^{-1}Q^T\vec{b}$
2. These questions concern the least squares solution $\hat{x}$ to $A\vec{x}=\vec{b}$. Indicate whether the statements are true or false.
   a. The solution $\hat{x}$ is chosen so that $A\hat{x}$ is close as possible to $\vec{b}$.
      i. True
   b. If $\vec{y}\neq\hat{x}$ then $\|A\hat{x}-\vec{b}\| < \|A\vec{y}-\vec{b}\|$
      i. False
      ii. (A dependent columns)
   c. If the columns of A are linearly independent, then the least squares solution is unique.
      i. True
3. Use the QR decomposition to calculate the least squares solution to $A\vec{x}=\vec{b}$.
   a. $A=QR=\begin{pmatrix}2/3 & -1/3\\ 2/3 & 2/3\\ 1/3 & -2/3\end{pmatrix}\begin{pmatrix}3&5\\0&1\end{pmatrix}\qquad \vec{b}=\begin{pmatrix}7\\3\\1\end{pmatrix}$

   $$\Rightarrow R\hat{x}=Q^T\vec{b}\Rightarrow\begin{pmatrix}3&5\\0&1\end{pmatrix}\hat{x}=\begin{pmatrix}2/3 & 2/3 & 1/3\\ -1/3 & 2/3 & -2/3\end{pmatrix}\begin{pmatrix}7\\3\\1\end{pmatrix}\Rightarrow\begin{pmatrix}3&5\\0&1\end{pmatrix}\hat{x}=\begin{pmatrix}7\\-1\end{pmatrix}\Rightarrow x_1=4, x_2=-1\Rightarrow\begin{pmatrix}4\\-1\end{pmatrix}$$
4. **Written Explanation Exercise** Explain step by step how to find the best fit line for a collection of n data points $\begin{bmatrix}x_1\\y_1\end{bmatrix},\ldots,\begin{bmatrix}x_k\\y_k\end{bmatrix}$ in $\mathbb{R}^2$. Why is the best fit line unqiue?

   Line of best fit: $y=ax+b$, $a\begin{bmatrix}x_1\\ \vdots\\ x_k\end{bmatrix}+b\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}y_1\\ \vdots\\ y_k\end{bmatrix}\Rightarrow\begin{bmatrix}x_1 & 1\\ \vdots & \vdots\\ x_k & 1\end{bmatrix}\begin{bmatrix}a\\b\end{bmatrix}=\begin{bmatrix}y_1\\ \vdots\\ y_k\end{bmatrix}$
5. Four points in $\mathbb{R}^3$ with coordinates (x,y,z) are give in the table below.
   a. $x \quad -1\ 0\ 2\ 1$
   b. $y \quad 2\ 1\ 0\ 2$
   c. $z \quad 9\ 1\ 0\ -1$

   Determine the coefficients c_1 and c_2 for the plane $z=c_1x+c_2y$ that best fits the data. *Hint: normal eqautions*.

Studio 22

Tuesday, November 16, 2021 12:33 PM

<u>T or F</u>

- The range of $T(x) = Ax$ is $\text{Row}(A)$
  - False! range of $T(x) = Ax$ is $\text{Col}(A)$

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

$A\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} \in \text{Range}(T)$

But, $A\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 2 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 3 \\ 0 \end{bmatrix} \in \text{Col}(A)$

- $\dim(\text{Row}(A)) = \dim(\text{Col}(A))$
  - True (both are the # of pivots)
- Given a subspace S and a vector b, there is a unique $x \in S$ that minimizes $\|x - b\|$
  - True (the minimizer is just $\text{Proj}_S b$)
- If A is $n \times n$ and invertible, then A is diagonalizable.
  - False $\left(\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \text{counter example}\right)$
    - $p(\lambda) = \det(A - \lambda I) = (1 - \lambda)^2 \Rightarrow 1 \text{ is an eigenvalue of alg. mult. } 2$
    - Is $\text{geom}(1) = 2$?
      - $A - I \sim \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \Rightarrow \text{geom}(1) = 1.$
      - Since $\text{geom}(1) < \text{alg}(1)$, A is not diagonalizable
- A is $n \times n$ and has n distinct eigenvalues, then A is diagonalizable
  - True $(\text{alg}(\lambda) = 1 \text{ for all } \lambda. \textit{Since } 1 \leq \text{geom}(\lambda) \leq \text{alg}(\lambda),\ \text{geom}(\lambda) = 1 \text{ for all } \lambda)$

Def. A is diagonalizable if $A = PDP^{-1}$ where D is diagonal.

Review Worksheet Exam 3

1. State True or False:
   a. A matrix $A \in \mathbb{R}^{4\times4}$ which has eigenvalues 0, 1, 2, 3 is diagonalizable.
      i. True
   b. There exists a matrix $A \in \mathbb{R}^{3\times3}$ with eigenvalues $i, i + 1, 1$.
      i. False
   c. If u, v are orthogonal vectors then $\|2u - 3v\| = \|2u + 3v\|$
      i. True
   d. A least squares solution $\hat{x}$ for the system $Ax = b$ satisfies $A\hat{x} = b$ if and only if $b \in \text{Col}(A)$.
      i. True
2. Use Gram Schmidt to find an orthonormal basis for the space spanned by the following vectors:

   $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix}$

   $\left\{\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{30}}\begin{pmatrix} 5 \\ -2 \\ -1 \end{pmatrix}, \quad \frac{1}{\sqrt{5}}\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}\right\}$

3. Find the best fit line $y = mx + b$ for the following data:

X	1	2	3	4
Y	3	1	-1	5

4. Find matrices P and D such that P is invertible, D is diagonal, and $A = PDP^{-1}$, where $A = \begin{pmatrix} 4 & -3 & -3 \\ 3 & -2 & -3 \\ -1 & 1 & 2 \end{pmatrix}$. The characteristic polynomial of A is $p_A(t) = -(t-1)^2(t-2)$.

5. Find matrices P and C such that P is invertible, C is a rotation-dilation matrix, and $A = PCP^{-1}$, where $A = \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}$

Final Exam

Monday, November 15, 2021   7:53 PM

<u>Material Covered:</u>

Chapter 7: Symmetric Matrices and Quadratic Forms

- Section 7.1 : Diagonalization of Symmetric Matrices
- Section 7.2 : Quadratic Forms
- Section 7.3 : Constrained Optimization
- Section 7.4 : The Singular Value Decomposition

Lecture 35

Monday, November 15, 2021 7:53 PM

Notes:

Section 7.1: Diagonalization of Symmetrix Matrices

Symmetric Matrices
Definition
Matrix A is symmetric if $A^T = A$
Ex. Which of the following matrices are symmetric? Symbols * and * represent real numbers.

$A = [*]$
$B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
$C = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$
$D = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$
$E = \begin{bmatrix} 4 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$
$F = \begin{bmatrix} 4 & 2 & 0 & 1 \\ 2 & 0 & 7 & 4 \\ 0 & 7 & 6 & 0 \\ 1 & 4 & 0 & 3 \end{bmatrix}$

$A^T A$ is Symmetric
A very common example: For any matrix A with columns $a_1, .., a_n$

$$A^T A = \begin{bmatrix} - & {a_1}^T & - \\ - & {a_2}^T & - \\ \vdots & \vdots & \vdots \\ - & {a_n}^T & - \end{bmatrix} \begin{bmatrix} | & | & \cdots & | \\ {a_1}^T & {a_2}^T & \cdots & {a_n}^T \\ | & | & \cdots & | \end{bmatrix} = \begin{bmatrix} {a_1}^T a_1 & {a_1}^T a_2 & \cdots & {a_1}^T a_n \\ {a_2}^T a_1 & {a_2}^T a_2 & \cdots & {a_2}^T a_n \\ \vdots & \vdots & \ddots & \vdots \\ {a_n}^T a_1 & {a_n}^T a_2 & \cdots & {a_n}^T a_n \end{bmatrix}$$

$${a_1}^T a_2 = a_1 \cdot a_2 = a_2 \cdot a_1 = {a_2}^T a_1$$
$$(A^T A)^T = A^T (A^T)^T = A^T A$$

Symmetric Matrices and their Eigenspaces
Theorem
A is a symmetric matrix, with eigenvectors $\vec{v_1}$ and $\vec{v_2}$ corresponding to two distinct eigenvalues. Then $\vec{v_1}$ and $\vec{v_2}$ are orthogonal.

More generally, eigenspaces associated to distinct eigenvalues are orthogonal subspaces.

Proof:
$\vec{v_1}$ eigenvector for λ_1 with $\lambda_1 \neq \lambda_2$
$\vec{v_2}$ eigenvector for λ_2
By the previous fact:
$A\vec{v_1} \cdot \vec{v_2} = \vec{v_1} \cdot A\vec{v_2}, \qquad \lambda = A$
Thus: $(\lambda_1 - \lambda_2)\vec{v_1} \cdot \vec{v_2} = 0$

Ex.1
Diagonalize A using an orthogonal matrix. Eigenvalues of A are given.

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \lambda = -1,1$$

If A is symmetric and diagonalizable:
P: matrix of orthonormal eigenvectors
$\rightarrow A = PDP^{-1} = PDP^T$

$$\lambda_1 = -1: A + I = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\rightarrow \vec{v_1} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 1: A - I = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix} \sim \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\rightarrow \vec{v_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Here: $\vec{v_2} \cdot \vec{v_1} = \vec{v_3} \cdot \vec{v_1} = 0$

$$P = \begin{pmatrix} -{}^{1}/_{\sqrt{2}} & 0 & {}^{1}/_{\sqrt{2}} \\ 0 & 1 & 0 \\ {}^{1}/_{\sqrt{2}} & 0 & {}^{1}/_{\sqrt{2}} \end{pmatrix}, \qquad D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Spectral Theorem
Recall: If P is an orthogonal $n \times n$ matrix, then $P^{-1} = P^T$, which implies $A = PDP^T$ is diagonalizable and symmetric.
Theorem: Spectral Theorem
An $n \times n$ symmetric matrix A has the following properties.

1. All eigenvalues of A are real.

2. The dimension of each eigenspace is full, that it's dimension is equal to it's algebraic multiplicity.
3. The eigenspaces are mutually orthogonal.
4. A can be diagonalized $A = PDP^T$, where D Is diagonal and P is orthogonal.

Proof:

1. Assume that $\lambda_1 \neq \mathbb{R}$
   $\overrightarrow{v_1}$ eigenvector for λ_1
   $\overrightarrow{v_2}$ eigenvector for I_1

For symmetric matrices:

$A\overrightarrow{v_1} \cdot \overrightarrow{v_2} = \overrightarrow{v_1} \cdot A\overrightarrow{v_2}, \qquad \lambda = A$

Thus: $\overrightarrow{v_1} \cdot \overrightarrow{v_2} = 0$

If $\overrightarrow{v_1} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$: $\overrightarrow{v_1} \cdot \overrightarrow{v_1} = \begin{pmatrix} |z_1|^2 \\ \vdots \\ |z_n|^2 \end{pmatrix}$

Spectral Decomposition of a Matrix

Spectral Decomposition

Suppose A can be orthogonally diagonalized as

$$A = PDP^T = [\overrightarrow{u_1} \quad \cdots \quad \overrightarrow{u_n}] \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} = \begin{bmatrix} \overrightarrow{u_1}^T \\ \vdots \\ \overrightarrow{u_n}^T \end{bmatrix}$$

Then A has the decomposition

$$A = \lambda_1 \overrightarrow{u_1}\overrightarrow{u_1}^T + \cdots + \lambda_n \overrightarrow{u_n}\overrightarrow{u_n}^T = \sum_{i=1}^{n} \lambda_i \overrightarrow{u_i}\overrightarrow{u_i}^T$$

Each term in the sum, $\lambda_i \overrightarrow{u_i}\overrightarrow{u_i}^T$, is an $n \times$ matrix with rank 1

$$\overrightarrow{u_1} = \begin{bmatrix} \overrightarrow{u_1} \\ \vdots \\ \overrightarrow{u_n} \end{bmatrix} \quad \overrightarrow{u_1}\overrightarrow{u_1}^T = [u_1\overrightarrow{u_1} \quad \cdots \quad u_n\overrightarrow{u_n}]$$

$$A = [\overrightarrow{u_1} \cdots \overrightarrow{u_n}] \begin{bmatrix} \lambda_1 \overrightarrow{u_1}^T \\ \vdots \\ \lambda_n \overrightarrow{u_n}^T \end{bmatrix} = \begin{bmatrix} u_{11} & u_{21} & \cdots & u_{n1} \\ \vdots & \vdots & & \vdots \\ u_{1n} & u_{2n} & & u_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 \overrightarrow{u_{11}} & \cdots & \lambda_1 \overrightarrow{u_{1n}} \\ \vdots & & \vdots \\ \lambda_n \overrightarrow{u_{n1}} & \cdots & \lambda_n \overrightarrow{u_{nn}} \end{bmatrix} = \lambda_1 \overrightarrow{u_1}\overrightarrow{u_1}^T + \cdots + \lambda_n \overrightarrow{u_n}\overrightarrow{u_n}^T$$

Ex.2

Construct a spectral decomposition for A whose orthogonal diagonalization is given.

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} PDP^T = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

$$\overrightarrow{u_1} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

$\lambda_1 = 4$

$$\overrightarrow{u_1} = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

$\lambda_2 = 2$

$$A = \lambda_1 \overrightarrow{u_1}\overrightarrow{u_1}^T + \lambda_2 \overrightarrow{u_2}\overrightarrow{u_2}^T$$

$$\overrightarrow{u_1}\overrightarrow{u_1}^T = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

$$\overrightarrow{u_2}\overrightarrow{u_2}^T = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

$$A = 4 \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} + 2 \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

Lecture 36

Friday, November 19, 2021 3:30 PM

Notes:

Section 7.2: Quadratic Forms

Quadratic Forms

Definition

A quadratic form is a function $Q: \mathbb{R}^n \to \mathbb{R}$, given by

$$Q(\vec{x}) = \vec{x}^T A\vec{x} = [x_1 \quad x_2 \quad \cdots \quad x_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix}$$

Matrix A is $n \times n$ and symmetric.

In the above, $\vec{x}$ is a vector of variables

$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}: Q(\vec{x}) = (x_1 \quad x_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = {x_1}^2 + {x_2}^2$

Ex.1

Compute the quadratic form $\vec{x}^T A\vec{x}$ for the matricies below

For A:

$$Q(\vec{x}) = (x_1 \quad x_2) \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1 \quad x_2) \begin{pmatrix} 4x_1 \\ 3x_2 \end{pmatrix} = 4{x_1}^2 + 3{x_2}^2$$

For B:

$$Q(\vec{x}) = (x_1 \quad x_2) \begin{pmatrix} 4 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1 \quad x_2) \begin{pmatrix} 4x_1 + x_2 \\ x_1 - 3x_2 \end{pmatrix} = 4{x_1}^2 + x_1x_2 + x_1x_2 - 3{x_2}^2 = 4{x_1}^2 + 2x_1x_2 - 3{x_2}^2$$

Cross-Term: the coefficients

Ex.1 - Surface Plots

The surfaces for Example 1 are shown below.

Students are not expected to be able to sketch quadratic surfaces, but it is helpful to see what they look like.

Ex.2

Write Q in the form $\vec{x}^T A\vec{x}$ for $\vec{x} \in \mathbb{R}^3$

$$Q(\vec{x}) = 5{x_1}^2 - 2{x_2}^2 + 3{x_3}^2 + 6x_1x_3 - 12x_2x_3$$

$$A = \begin{pmatrix} 5 & 0 & 3 \\ 0 & -1 & -6 \\ 3 & -6 & 3 \end{pmatrix}$$

Change of Variable

If $\vec{x}$ is a variable vector in $\mathbb{R}^n$, then a change of variable can be represented as:

$\vec{x} = P\vec{y}$, or $\vec{y} = P^{-1}\vec{x}$

With this change of variable, the quadratic form $\vec{x}^T A\vec{x}$ becomes:

$$Q = \vec{x}^T A\vec{x} = (P\vec{y})^T AP\vec{y} = \underbrace{\vec{y}^T \underbrace{P^T AP}_{\text{symmetric}} \vec{y}}_{\text{quadratic form in } \vec{y}}$$

Idea: if A is symmetric, there exists P orthogonal and D diagonal such that $A = PDP^T$

$\to D = P^T AP$

Ex.3

Make a change of variable $\vec{x} = P\vec{y}$ that transforms $Q = \vec{x}^T A\vec{x}$ so that it does not have cross terms. The orthogonal decomposition of A is given.

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} = PDP^{-1}: \ Q(\vec{x}) = 3{x_1}^2 + 4x_1x_2 + 6{x_2}^2$$

$$P = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$$

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 7 \end{pmatrix}$$

$$Q(\vec{x}) = \vec{x}^T A\vec{x} \Rightarrow A = PDP^T = \vec{y}^T D\vec{y} = 2{y_1}^2 + 7{y_2}^2$$

(No cross-terms)

Geometry

Suppose $Q(\vec{x}) = \vec{x}^T A\vec{x}$, where $A \in \mathbb{R}^{n\times n}$ is symmetric. Then the set of $\vec{x}$ that satisfies

$C = \vec{x}^T A\vec{x}$

defines a curve or surface in $\mathbb{R}^n$

Previous Example:

$Q(\vec{x}) = 3{x_1}^2 + 4x_1x_2 + 6{x_2}^2$

$3{x_1}^2 + 4x_1x_2 + 6{x_2}^2 = 8, \qquad 2{y_1}^2 + 7{y_2}^2 = 8$

$\Rightarrow y_1 = 2, \qquad y_2 = \sqrt{\frac{8}{7}}$

Principle Axes Theorem

Theorem

If A is a symmetric matrix then there exists an orthogonal change of variable $\vec{x} = P\vec{y}$ that transforms $\vec{x}^T A\vec{x}$ to $\vec{x}^T D\vec{x}$ with no cross-product terms.

Proof:

$A = PDP^{-1}PDP^T$
(P orthogonal, D diagonal)
$Q = \vec{x}^T A\vec{x} = \vec{y}^T P^T AP\vec{y} = \vec{y}^T D\vec{y} = \lambda_1 {y_1}^2 + \lambda_2 {y_2}^2 + \cdots + \lambda_n {y_n}^2$

Ex.5

Compute the quadratic form $Q(\vec{x}) = \vec{x}^T A\vec{x}$ for $A = \begin{pmatrix} 5 & 2 \\ 2 & 8 \end{pmatrix}$ and find a change of variable that removes the cross-product term.

$Q(\vec{x}) = 5{x_1}^2 + 4x_1x_2 + 8{x_2}^2$

$A = \begin{pmatrix} 5 & 2 \\ 2 & 8 \end{pmatrix} \lambda = 4{,}9$

$|A - \lambda I| = \lambda^2 - 13\lambda + 36$

$$\lambda_{1,2} = \frac{13 \pm \sqrt{169 - 144}}{2} = 4 \text{ or } 9$$

$\lambda_1 = 4$:

$A - 4I = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \quad \vec{v_1} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$

$\lambda_1 = 9$:

$\vec{v_2} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ since they are orthogonal

$P = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \quad \vec{x} = P\vec{y}$

$\rightarrow Q = 4{y_1}^2 + 9{y_2}^2$

Classifying Quadratic Forms

$Q = {x_1}^2 + {x_2}^2$
$Q = -{x_1}^2 - {x_2}^2$

Definition

A quadratic form Q is

1. Positive definite if $Q > 0$ for all $\vec{x} \neq \vec{0}$.
2. Negative definite if $Q < 0$ for all $\vec{x} \neq \vec{0}$.
3. Positive semidefinite if $Q \geq 0$ for all $\vec{x}$.
4. Negative semidefinite if $Q \leq 0$ for all $\vec{x}$.
5. Indefinite if Q can be positive or negative.

Quadratic Forms and Eigenvalues

Theorem

If A is a symmetric matrix with eigenvalues λ_i, then $Q = \vec{x}^T A\vec{x}$ is

1. Positive definite iff $\lambda_i > 0$ for all i
   a. Semidefinite ≥ 0
2. Negative definite iff $\lambda_i < 0$ for all i
   a. Semidefinite ≤ 0
3. Indefinite iff $\lambda_i > 0, \lambda_j < 0$ for some i, j

Proof:

If $A = PDP^T$
$\vec{x} = P\vec{y}$
$Q = \lambda_1 {y_1}^2 + \lambda_2 {y_2}^2 + \cdots + \lambda_n {y_n}^2$

Ex.6

We can now return to our motivating question: does this inequality hold for all x, y?

$x^2 - 6xy + 9y^2 \geq 0$

$A = \begin{pmatrix} 1 & -3 \\ -3 & 9 \end{pmatrix}$

Eigenvalues: 0, 10
Q is positive semidefinite

$A = \begin{pmatrix} 2 & 4 \\ 4 & 5 \end{pmatrix} |A| = -6 = \lambda_1\lambda_2 \Rightarrow$ one is positive and one is negative
Q is indefinite

$6{x_1}^2 + 3x_1x_2 + 7{x_2}^2 \geq 0?$

$A = \begin{pmatrix} 6 & 3/2 \\ 3/2 & 7 \end{pmatrix}$

$\Rightarrow$ Positive definite

Lecture 37

Monday, November 22, 2021 3:28 PM

Notes:

Section 7.3: Constrained Optimization

Ex.1

The surface of a unit sphere in $\mathbb{R}^3$ is given by

$1 = {x_1}^2 + {x_2}^2 + {x_3}^2 = \|\vec{x}\|$

Q is a qunaitity we want to optimize

$Q(\vec{x}) = 9{x_1}^2 + 4{x_2}^2 + 3{x_3}^2$

Find the largest and smallest values of Q on the surface of the sphere.

$3({x_1}^2 + {x_2}^2 + {x_3}^2) \leq Q(\vec{x}) \leq 9({x_1}^2 + {x_2}^2 + {x_3}^2)$

Where $({x_1}^2 + {x_2}^2 + {x_3}^2) = 1$

Here:

The minimum is 3, attained at $\vec{x} = \pm\vec{e_3}$

The maximum is 9, attained at $\vec{x} = \pm\vec{e_1}$

A Constrained Optimization Problem

Suppose we wish to find the maximum or minimum values of

$Q(\vec{x}) = \vec{x}^T A\vec{x}$

Subject to

$\|\vec{x}\| = 1$

That is, we want to find

$m = \min\{Q(\vec{x}) \,\|\vec{x}\| = 1\}$

$M = \max\{Q(\vec{x}) \,\|\vec{x}\| = 1\}$

This is an example of a constrained optimization problem. Note that we may also want to know were these extreme values are obtained.

Constrained Optimization and Eigenvalues

Theorem

If $Q(\vec{x}) = \vec{x}^T A\vec{x}$, A is a real $n \times n$ symmetric matrix, with eigenvalues

$\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n$

and associated normalized eigenvectors

$\vec{u_1}, \vec{u_2}, \dots, \vec{u_n}$

Then, subject to the constraint $\|\vec{x}\| = 1$,

- The maximum value of $Q(\vec{x}) = \lambda_1$, attained at $\vec{x} = \pm\vec{u_1}$.
- The minimum value of $Q(\vec{x}) = \lambda_n$, attained at $\vec{x} = \pm\vec{u_n}$.

Proof:

$P = (\vec{u_1}, \dots, \vec{u_n})$ orthogonal

Define $\vec{y}$ by $\vec{x} = P\vec{y}$

$Q = \lambda_1{y_1}^2 + \lambda_2{y_2}^2 + \cdots + \lambda_n{y_n}^2$

Here $1 = \|\vec{x}\| = \|P\vec{y}\| = \|\vec{y}\|$ since P has orthonormal columns

As in the previous example:

Minimum is λ_n, attained for $\vec{y} = \pm\vec{e_n}$

Then $\vec{x} = \pm P\vec{e_n} = \pm\vec{u_n}$

Maximum is λ_1, attained for $\vec{y} = \pm\vec{e_1}$

Then $\vec{x} = \pm P\vec{e_1} = \pm\vec{u_1}$

Ex.2

Calculate the maximum and minimum values of $\mathrm{Q}(\vec{x}) = \vec{x}^T A\vec{x}$, $\vec{x} \in \mathbb{R}^3$, subject to $\|\vec{x}\| = 1$, and identify points where these values are obtained.

$Q(\vec{x}) = {x_1}^2 + 2x_2x_3$

The symmetric matrix A associated to Q is

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$|A - \lambda I| = (I - \lambda)\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = (1-\lambda)(\lambda^2 - 1)(1-\lambda)(\lambda+1)(\lambda-1)$$

Eigenvalues: 1,1, -1

Eigenvectors:

$$\lambda_1 = \lambda_2 = 1\colon\ (A - I) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \sim \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\rightarrow \vec{u_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \vec{u_2} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$\lambda_3 = -1$: by orthogonality:

$$\rightarrow \vec{u_3} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

- The minimum value of -1, attained at $\vec{x} = \pm\vec{u_3}$.
- The maximum value of 1, attained at $\vec{x} = \pm\vec{u_1}$.
  - Or $\vec{x} = \pm\vec{u_2}$
  - Or $\vec{x} = a\vec{u_1} + b\vec{u_2}$ with $a^2 + b^2 = 1$
  - $\|\vec{x}\| = a^2 + b^2$
    - It is a circle!

Geom. Multiplicity $= 1$, line

Geom. Multiplicity $= 2$, circle

An Orthogonality Constraint

Theorem

Suppose
$Q(\vec{x}) = \vec{x}^T A\vec{x}$, A is a real $n \times n$ symmetric matrix, with eigenvalues

$\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n$

and associated normalized eigenvectors

$\overrightarrow{u_1}, \overrightarrow{u_2}, \dots, \overrightarrow{u_n}$

Then, subject to the constraint $\|\vec{x}\| = 1$, $\vec{x} \cdot \overrightarrow{u_1} = 0$,

- The maximum value of $Q(\vec{x}) = \lambda_2$, attained at $\vec{x} = \pm\overrightarrow{u_*}$.
- The minimum value of $Q(\vec{x}) = \lambda_n$, attained at $\vec{x} = \pm\overrightarrow{u_n}$.

Note that λ_2 is the second largest eigenvalue of A.
$P = (\overrightarrow{u_1}, \dots, \overrightarrow{u_n})$ orthogonal

Define $\vec{y}$ by $\vec{x} = P\vec{y}$

$Q = \lambda_1 {y_1}^2 + \lambda_2 {y_2}^2 + \cdots + \lambda_n {y_n}^2$
Here: $0 = \vec{x} \cdot \overrightarrow{u_1} = (P\vec{y}) \cdot (P\overrightarrow{e_1}) = \vec{y} \cdot \overrightarrow{e_1} = \overrightarrow{y_1}$

Since P has orthogonal columns, preserve dot product

$Q = \lambda_2 {y_2}^2 + \cdots + \lambda_n {y_n}^2$

Minimum: λ_n for $\vec{y} = \overrightarrow{e_n} (\vec{x} = \pm\overrightarrow{u_n})$

Maximum: λ_2 for $\vec{y} = \overrightarrow{e_2} (\vec{x} = \pm\overrightarrow{u_1})$

Ex.3
Calculate the maximum and minimum values of $Q(\vec{x}) = \vec{x}^T A\vec{x}$, $\vec{x} \in \mathbb{R}^3$, subject to $\|\vec{x}\| = 1$ and $\vec{x} \cdot \overrightarrow{u_1} = 0$, and identify points where these values are obtained.

$$Q(\vec{x}) = {x_1}^2 + 2x_2x_3, \qquad \overrightarrow{u_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

From example 2:

Maximum is 1, attained for $\vec{x} = \pm\overrightarrow{u_2}$

Ex.4
Calculate the maximum and minimum values of $Q(\vec{x}) = \vec{x}^T A\vec{x}$, $\vec{x} \in \mathbb{R}^3$, subject to $\|\vec{x}\| = 5$, and identify points where these values are obtained.

$Q(\vec{x}) = {x_1}^2 + 2x_2x_3$

Eigenvalues: 1, 1, -1

Max $Q(\vec{x}) = 25$
$\|\vec{x}\| = 5$
2 points of view to see this:

1) Define $\vec{y}$ by $\vec{x} = P\vec{y}$, $Q = \lambda_1 {y_1}^2 + \lambda_2 {y_2}^2 + \cdots + \lambda_n {y_n}^2$

$\lambda_n({y_1}^2 + {y_2}^2 + {y_3}^2) \leq Q \leq \lambda_1({y_1}^2 + {y_2}^2 + {y_3}^2)$

$= \|\vec{y}\| = \|\vec{x}\| = 25$

1) $\|\vec{x}\| = 5$, $\tilde{x} = \frac{\vec{x}}{\|\vec{x}\|} = \frac{\vec{x}}{5}$

$\vec{x} = 5\tilde{x}, \|\tilde{x}\| = 1$
$Q(\vec{x}) = Q(5\tilde{x}) = 25Q(\tilde{x})$
$\rightarrow \max Q(\vec{x}) = 25 \max Q(\tilde{x})$

$\|\vec{x}\| = 5, \|\tilde{x}\| = 1$

Lecture 38

Monday, November 29, 2021 3:26 PM

<u>Notes:</u>

Section 7.4: The Singular Value Decomposition

Ex.1
The linear transform whose standard matrix is
$$A = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}2\sqrt{2} & 0\\ 0 & \sqrt{2}\end{pmatrix} = \begin{pmatrix}2 & -1\\ 2 & 1\end{pmatrix}$$

Maps the unit circle in $\mathbb{R}^2$ to an ellipse, as shown below. Identify the unit vector $\vec{x}$ in which $\|A\vec{x}\|$ is maximized and compute this length.
Goal: $\max\limits_{\|\vec{x}\|=1} \|A\vec{x}\|$
Remark: $\max\|A\vec{x}\|$ and $\max\|A\vec{x}\|^2$ occur at the same place for $\vec{x}$
$\|A\vec{x}\|^2 = (A\vec{x})\cdot(A\vec{x}) = (A\vec{x})^T A\vec{x} = \vec{x}A^T A\vec{x} = \vec{x}^T A^T A\vec{x}$: it's quadratic form, so we can just use the eigenvalues of $A^T A$

Ex.1 - Solution
$A = \begin{pmatrix}2 & -1\\ 2 & 1\end{pmatrix}$
$A^T A = \begin{pmatrix}2 & 2\\ -1 & 1\end{pmatrix}\begin{pmatrix}2 & -1\\ 2 & 1\end{pmatrix} = \begin{pmatrix}8 & 0\\ 0 & 2\end{pmatrix}$

Eigenvalues/Eigenvectors:
- $8, \begin{pmatrix}1\\0\end{pmatrix}$
- $2, \begin{pmatrix}0\\1\end{pmatrix}$
- $\rightarrow \max\limits_{\|\vec{x}\|=1}\|A\vec{x}\|^2 = 8$ attained at $\vec{x} = \pm\vec{e_1}$

$\Rightarrow \max\limits_{\|\vec{x}\|=1}\|A\vec{x}\| = \sqrt{8} = 2\sqrt{2}$ attained at $\vec{x} = \pm\vec{e_1}$

Singular Values
The matrix $A^T A$ is always symmetric, with non-negative eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n \geq 0$. Let $\{\vec{v_1}, \dots, \vec{v_n}\}$ be the associated orthonormal eigenvectors. Then
$$\|A\vec{v_j}\|^2 = (A\vec{v_j})^T(A\vec{v_j}) = \vec{v_j}^T A^T A\vec{v_j} = \lambda_j \vec{v_j}\cdot\vec{v_j} = \lambda_j$$
If the A has rank r, then $\{A\vec{v_1}, \dots, A\vec{v_r}\}$ is an orthogonal basis for $\mathrm{Col}(A)$:
For $1 \leq j < k \leq r$:
$$A\vec{v_j}\cdot A\vec{v_k} = (A\vec{v_j})^T A\vec{v_k} = \vec{v_j}^T A^T A\vec{v_k} = \lambda_k \vec{v_j}\cdot\vec{v_k} = 0$$
Definition: $\sigma_1 = \sqrt{\lambda_1} \geq \sigma_2 = \sqrt{\lambda_2} \cdots \geq \sigma_n = \sqrt{\lambda_n}$ are the singular values of A.

To sum up:
We have: $\|A\vec{v_j}\| = \sqrt{\lambda_j} = \sigma_j$
- If $i \neq j$: $A\vec{v_i}\cdot A\vec{v_j} = 0$

$A\vec{v_1} \dots A\vec{v_n} \in \mathrm{Col}(A)$
$\vec{x} \in \mathbb{R}^n$: $\vec{x} = c_1\vec{v_1} + \cdots + c_n\vec{v_n}$
- $A\vec{x} = c_1 A\vec{v_1} + \cdots + c_n A\vec{v_n}$

$\Rightarrow \mathrm{Span}(A\vec{v_1}, \dots, A\vec{v_n}) = \mathrm{Col}(A)$
- Say, $\dim\mathrm{Col}(A) = r$

$\mathrm{Span}(A\vec{v_1}, \dots, A\vec{v_n}) = \mathrm{Span}(A\vec{v_1}, \dots, A\vec{v_r})$
- $\rightarrow \{A\vec{v_1}, \dots, A\vec{v_r}\}$ orthogonal basis of $\mathrm{Col}(A)$

The SVD
$$\underbrace{A}_{m\times n} = \underbrace{U}_{m\times m}\ \underbrace{\Sigma}_{m\times n}\ \underbrace{V^T}_{n\times n}$$

Theorem: Singular Value Decomposition
A $m \times m$ matrix with rank r and non-zero singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$ has a decomposition $U\Sigma V^T$ where
$$\Sigma = \begin{bmatrix}D & 0\\ 0 & 0\end{bmatrix}_{m\times n} = \begin{bmatrix}\sigma_1 & 0 & \cdots & 0 & \\ 0 & \sigma_2 & \cdots & \vdots & 0\\ \vdots & \vdots & \ddots & & \\ 0 & 0 & \cdots & \sigma_r & \\ & 0 & & & 0\end{bmatrix}$$
U is an $m \times n$ orthogonal matrix, and V is an $n \times n$ orthogonal matrix.

$A \in \mathbb{R}^{m\times n}$
$\rightarrow A^T A \in \mathbb{R}^{n\times n}$ has eigenvlaues $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n \geq$
$\{\vec{v_1}, \dots, \vec{v_n}\}$ orthonormal basis of eigenvectors
$\{A\vec{v_1}, \dots, A\vec{v_r}\}$ orthogonal basis of $\mathrm{Col}(A)$
Define: $\vec{u_i} = \frac{1}{\|A\vec{v_i}\|}A\vec{v_i} = \frac{1}{\sigma_i}A\vec{v_i}$ for $1 \leq i \leq r$
Define $U = \Big(\underbrace{\vec{u_1}, \dots, \vec{u_r}}_{\text{orthonormal basis of Col}(A)}\quad \underbrace{\vec{u_{r+1}}, \dots, \vec{u_{r+n}}}_{\text{orthonormalbasis for Col}(A)^\perp = \mathrm{Null}(A^T)}\Big)$
- $V = (\vec{v_1}, \dots, \vec{v_n})$

$$\Sigma = \begin{pmatrix}\vec{v_1} & & & & \\ & \ddots & & & \\ & & \vec{v_r} & & \\ & & & 0 & \\ & & & & \ddots & \\ & & & & & 0\end{pmatrix}$$
We want to prove that:
$A = U\Sigma V^T \Leftrightarrow AV = U\Sigma$

$AV = (A\overrightarrow{v_1}, \dots, A\overrightarrow{v_n}) = (\overrightarrow{v_1 u_1} \quad \cdots \quad \overrightarrow{v_r u_r} \quad \vec{0} \quad \cdots \quad \vec{0})$

$$U\Sigma = (\overrightarrow{u_1}, \dots, \overrightarrow{u_m}) \begin{pmatrix} \overrightarrow{v_1} & & & & & \\ & \ddots & & & & \\ & & \overrightarrow{v_r} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} = (\overrightarrow{v_1 u_1} \quad \cdots \quad \overrightarrow{v_r u_r} \quad \vec{0} \quad \cdots \quad \vec{0})$$

$\rightarrow AV = U\Sigma$

$\Leftrightarrow A = U\Sigma V^{-1} = U\Sigma V^T$

$M = U \cdot \Sigma \cdot V^*$

Algorithm to find the SVD of A

Suppose A is $m \times n$ has rank $r \leq n$.

1. Compute the squared singular values of $A^T A$, ${\sigma_i}^2$
2. Compute the unit singular vector of $A^T A$, $\overrightarrow{v_i}$, use them to form V.
3. Compute an orthonormal basis for $\text{Col}(A)$ using
   $\overrightarrow{u_i} = \frac{1}{\sigma_i} A\overrightarrow{v_i}$, $i = 1,2, \dots r$

Extend the set $\{\overrightarrow{u_i}\}$ to form an orthonormal basis for $\mathbb{R}^m$, use the bassi for form U.

Ex.2

Write down the singular value decomposition for

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

Ex.3

Construct the singular value decomposition of

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

(It has rank 1.)

Studio 23

Tuesday, November 30, 2021 12:32 PM

Constrained Optimization
$Q(\vec{x}) = {x_1}^2 + 2{x_2}^2$
What is max $Q(\vec{x})$ where $\vec{x} \in \mathbb{R}^2$? (unconstrained optimization)
$\max Q(\vec{x}) = \infty$

What is max $Q(\vec{x})$ where $\|\vec{x}\| = 1$? (constrained optimization)

In general, $Q(\vec{x}) = \vec{x}^T A\vec{x}$, A is symmetric
$\max Q(\vec{x}) = \lambda_{\max} = \max\{\lambda: \lambda \text{ eigenvalue of } A\}$
And $\lambda_{\max} = Q(\overrightarrow{x_{\max}})$ where $\|\overrightarrow{x_{\max}}\| = 1$ and $A\overrightarrow{x_{\max}} = \lambda_{\max}\overrightarrow{x_{\max}}$

$Q(\vec{x}) = \vec{x}^T A\vec{x} = \vec{x}^T \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \vec{x}$
$\quad \rightarrow \lambda_{\max} = 2$
$\underbrace{\max Q(\vec{x})}_{\|\vec{x}\|=1} = 2$ attained at $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$

Worksheet 7.3 Constrained Optimization
Worksheet Exercises

1. Indicate whether the statements are true or false.
    a. The largest value of a positive definite quadratic form $\vec{x}^T A\vec{x}$ is the largest eigenvalue of A.
        i. False
    b. The largest value of a positive definite quadratic form $\vec{x}^T A\vec{x}$ subject to $\|\vec{x}\| = 1$ is the largest value on the diagonal of A.
        i. False
2. Calculate the maximum and minimum values of the quadratic form $\vec{x}^T A\vec{x}$ subject to $\|\vec{x}\| = 1$. Identify where this maximum is obtained.
    a. $Q(\vec{x}) = 4{x_1}^2 + {x_2}^2 + 4x_1x_2 + 3{x_3}^2$, $\vec{x} \in \mathbb{R}^3$
    b. $A = \begin{bmatrix} 4 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, $A - \lambda I = \begin{bmatrix} 4-\lambda & 2 & 0 \\ 2 & 1-\lambda & 0 \\ 0 & 0 & 3-\lambda \end{bmatrix} = (3-\lambda)(\lambda^2 - 5\lambda) = 0,3,5$. Max = 5, Min = 0.
    c. $A - 5I = \text{Null}\left(\begin{bmatrix} -1 & 2 & 0 \\ 2 & -4 & 0 \\ 0 & 0 & -2 \end{bmatrix}\right) = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \Rightarrow \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = x_1, \underbrace{\max Q(\vec{x})}_{\|\vec{x}\|=1} = Q(x_1)$
    d. $A - 0I = \text{Null}\left(\begin{bmatrix} 4 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}\right) = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \Rightarrow \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = x_2, \underbrace{\min Q(\vec{x})}_{\|\vec{x}\|=1} = Q(x_2)$
3. Calculate the maximum and minimum values of the quadratic form Q subject to $\|\vec{x}\| = 1$ and $\vec{x} \cdot \vec{u} = 0$.
    a. $Q(\vec{x}) = 4{x_1}^2 + {x_2}^2 + 4x_1x_2 + 3{x_3}^2$, $\vec{u} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$
    b. $A - 3I = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow x_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$
    c. $\underbrace{\underbrace{\max Q(\vec{x})}_{\|\vec{x}\|=1}}_{\vec{x}\cdot\vec{u}=0} = Q(x_3)$
    d. $\underbrace{\underbrace{\min Q(\vec{x})}_{\|\vec{x}\|=1}}_{\vec{x}\cdot\vec{u}=0} = Q(x_2)$
4. If possible, give example of the following.
    a. A quadratic form $Q: \mathbb{R}^3 \mapsto \mathbb{R}$, that has the maximum value 12, subject to the constraint that $\|\vec{x}\| = 1$.
        i. $Q(\vec{x}) = {x_1}^2 + 12{x_2}^2 + 4{x_3}^2$
    b. A quadratic form $Q: \mathbb{R}^3 \mapsto \mathbb{R}$, that has the maximum value 4 at two distinct locations, subject to the constraint that $\|\vec{x}\| = 1$.
        i. $Q(\vec{x}) = 4{x_1}^2 + 2{x_2}^2 + 4{x_3}^2$

Lecture 39

Wednesday, December 1, 2021 3:30 PM

Notes:

$M\vec{x} = U\Sigma V^T\vec{x}$

Ex.2
Write down the singular value decomposition for

$$\begin{bmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = U\Sigma V^T$$

1st: Σ

$$A^TA = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}$$

$\rightarrow \sigma_1 = 3, \quad \sigma_2 = 2$

$$\Rightarrow \Sigma = \begin{pmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

2nd: V

$\lambda_1 = 9$: $A^TA - 9I = \begin{pmatrix} -5 & 0 \\ 0 & 0 \end{pmatrix} \vec{v_1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$\lambda_2 = 4$: $\vec{v_2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

3rd: U

$U = (\vec{u_1} \quad \vec{u_2} \quad \vec{u_3} \quad \vec{u_4})$

$$\vec{u_1} = \frac{1}{\sigma_1}A\vec{v_1} = \frac{1}{3}\begin{pmatrix} 0 \\ -3 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

$$\vec{u_2} = \frac{1}{\sigma_2}A\vec{v_2} = \frac{1}{2}\begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Here we can take:

$$\vec{u_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \vec{u_4} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

$$U = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and: $A = U\Sigma V^T$
with U, Σ, V defined above.

Ex.3
Construct the singular value decomposition of

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

(It has rank 1.)

1st: Σ

$$A^TA = \begin{pmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix}$$

$\rightarrow \sigma_1 = 3\sqrt{2}$

$$\Rightarrow \Sigma = \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

2nd: V

$\lambda_1 = 18$: $A^TA - 18I = \begin{pmatrix} -9 & -9 \\ -9 & -9 \end{pmatrix} \vec{v_1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$\lambda_2 = 0$: by orthogonality: $\vec{v_1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$$\rightarrow V = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

3rd: U

$U = (\vec{u_1} \quad \vec{u_2} \quad \vec{u_3})$

$$\vec{u_1} = \frac{1}{\sigma_1}A\vec{v_1} = \frac{1}{3\sqrt{2}}\begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 2 \\ -4 \\ 4 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$$

Here we know $\vec{u_2}, \vec{u_3}$ are orthogonal to $\vec{u_1}$:

$\vec{x} \in \mathrm{Null}(1 \quad -2 \quad 2)$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 - 2x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

<u>Problem:</u> $\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ are not orthogonal

$\rightarrow$ Gram-Schmidt

$$\vec{u_2} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

$$\vec{u_3} = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} - \frac{\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}}{\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} + \frac{4}{5} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

$$\rightarrow \vec{u_3} = \frac{1}{5} \begin{pmatrix} -2 \\ 4 \\ 5 \end{pmatrix}$$

$$\rightarrow U = \begin{pmatrix} 1/3 & 2/3 & -2/3\sqrt{5} \\ -2/3 & 1/\sqrt{5} & 4/3\sqrt{5} \\ 2/3 & 0 & 5/3\sqrt{5} \end{pmatrix}$$

$\Rightarrow A = U\Sigma V^T$

with U, Σ, V defined above.

Applications of the SVD

The SVD has been applied to many modern applications in CS, engineering, and mathematics (our textbook mentions the first four).

- Estimating the rank and condition number of a matrix
- Constructing bases for the four fundamental spaces
- Computing the pseudoinverse of a matrix
- Linear least squares problems
- Non-linear least-squares

Normal Equations: $A^T A\vec{x} = A^T \vec{b}$

if $A = U\Sigma V^T$

$$A^T A = (U\Sigma V^T)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V \underbrace{\Sigma^T \Sigma}_{\begin{pmatrix} \vec{v_1}^2 & & \\ & \ddots & \\ & & \vec{v_n}^2 \end{pmatrix}} V^T\text{: orthogonal diagonalization of } A^T A$$

The Condition Number of a Matrix

If A is an invertible $n \times n$ matrix, the ratio

$$\frac{\sigma_1}{\sigma_n}$$

is the condition number of A.

Note that:

- The condition number of a matrix describes the sensitivity of a solution to $A\vec{x} = \vec{b}$ is to errors in A.
- We could define the condition number for a rectangular matrix, but that would go beyond the scope of this course.

Ex.4

For $A = U\Sigma V^*$ determines the rank of A, and orthonormal bases for $\mathrm{Null}(A)$ and $\mathrm{Col}(A^{\perp})$

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

$$\Sigma = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$V^T = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \sqrt{0.2} & 0 & 0 & 0 & \sqrt{0.8} \\ 0 & 0 & 0 & 1 & 0 \\ -\sqrt{0.8} & 0 & 0 & 0 & \sqrt{0.2} \end{bmatrix}$$

$\mathrm{Rank}(A) = 3 \ (r_1, r_2, r_3 \neq 0)$

$\{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$ basis for $\mathrm{Col}(A)$

$\{\vec{u_4}\} = \left\{ \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} \right\}$ is an orthogonal bassi for $\mathrm{Col}(A^{\perp}) = \mathrm{Null}(A^T)$

Rank Theorem: $\dim \mathrm{Null}(A) = 2$

We know: $A\vec{v_1} \neq \vec{v_1}$

$\Rightarrow \{\vec{v_4}, \vec{v_5}\}$ is an orthonormal basis for $\mathrm{Null}(A)$

$$\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\sqrt{0.8} \\ 0 \\ 0 \\ \sqrt{0.2} \end{pmatrix} \right\}$$

The Four Fundamental Spaces

1. $A\overrightarrow{v_s} = \sigma_s \overrightarrow{u_s}$
2. $\overrightarrow{v_1}, \dots, \overrightarrow{v_r}$ is an orthonormal basis for $\mathrm{Row}(A)$
3. $\overrightarrow{u_1}, \dots, \overrightarrow{u_r}$ is an orthonormal basis for $\mathrm{Col}(A)$
4. $\overrightarrow{v_{r+1}}, \dots, \overrightarrow{v_n}$ is an orthonormal basis for $\mathrm{Null}(A)$
5. $\overrightarrow{u_{r+1}}, \dots, \overrightarrow{u_n}$ is an orthonormal basis for $\mathrm{Null}(A^T)$

$$U = \left(\underbrace{\overrightarrow{u_1} \quad \cdots \quad \overrightarrow{u_r}}_{\text{orthonormal basis for Col}(A)} \quad \underbrace{\overrightarrow{u_{r+1}} \quad \cdots \quad \overrightarrow{u_m}}_{\text{orthonormal basis for Null}(A^T)} \right)$$

$$V = \left(\underbrace{\overrightarrow{v_1} \quad \cdots \quad \overrightarrow{v_r}}_{\text{orthonormal basis for Row}(A)} \quad \underbrace{\overrightarrow{v_{r+1}} \quad \cdots \quad \overrightarrow{v_n}}_{\text{orthonormal basis for Null}(A)} \right)$$

The Spectral Decomposition of a Matrix

The SVD can also be used to construct the spectral decomposition for any matrix with rank r

$$A = \sum_{s=1}^{r} \sigma_s \overrightarrow{u_s} \overrightarrow{u_s}^T$$

Where $\overrightarrow{u_s}, \overrightarrow{u_s}$ are the s^th columns of U and V respectively.
For the case when $A = A^T$, we obtain the same spectral decomposition that we encountered in Section 7.2.

→ Check that it works with Examples 2 and 3.

Studio 24

Thursday, December 2, 2021 12:36 PM

<u>SVD</u>

$A = U\Sigma V^T$

U, V: orthogonal

Σ: diagonal with non − increasing entries

$A: m \times n$

$\Sigma: m \times m$

$V: n \times n$

if $A = U\Sigma V^T$

$$A^TA = (U\Sigma V^T)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V \underbrace{\Sigma^T\Sigma}_{\begin{pmatrix}\vec{v_1}^2 & & \\ & \ddots & \\ & & \vec{v_n}^2\end{pmatrix}} V^T: \text{orthogonal diagonalization of } A^TA$$

$\Rightarrow$ columns of V are the ortthonormal set of eigenvectors (in decreasing order of eigenvalues)

Worksheet 7.4, The Singular Value Decomposition

Worksheet Exercises

1. Indicate whether the statements are true or false.
   a. Every matrix has a singular value decomposition.
      i. True
   b. If A is symmetric, then its factorization $A = UDU^T$ is also its SVD.
      i. True if entries in D are non-increasing $(U = V)$
   c. The maximum value of $\|A\vec{x}\|$ subject to $\|\vec{x}\| = 1$ is σ_1.
      i. $\|A\vec{x}\|^2 = \|U\Sigma V^T\vec{x}\|^2$
      $= \vec{x}^T A^T A\vec{x} = Q(\vec{x})$
      $\max Q(\vec{x}) = \lambda_{\max}\ (A^TA) = {\sigma_1}^2$, for $\|\vec{x}\| = 1$
      $\Rightarrow \max\|A\vec{x}\| = \sigma_1$, for $\|\vec{x}\| = 1$
      ii. True
2. Construct the SVD of
$$A = \begin{bmatrix} 4 & -2 \\ 2 & -1 \\ 0 & 0 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 20 & -10 \\ -10 & 5 \end{bmatrix}$$

$p(\lambda) = (20-\lambda)(5-\lambda) - 100 = \lambda^2 - 25\lambda = \lambda(\lambda - 25)$

$\lambda = 0$: $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ eigenvector

$\lambda = 25$: $A^TA - 25I = \begin{bmatrix} -5 & -10 \\ -10 & -20 \end{bmatrix}$

$\begin{bmatrix} -2 \\ 1 \end{bmatrix}$ eigenvector

$$V = \frac{1}{\sqrt{5}}\begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix}, \qquad \sigma_1 = 5, \qquad \sigma_2 = 0$$

$$\Rightarrow \Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$u_1 = \frac{1}{\sigma_1} Av_1 = \frac{1}{5}\begin{bmatrix} -\frac{10}{\sqrt{5}} \\ -\frac{5}{\sqrt{5}} \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \\ 0 \end{bmatrix}$$

$$u_2 = \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \\ 0 \end{bmatrix}, \qquad u_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \Rightarrow U = [u_1 \quad u_2 \quad u_3]$$

3. Find a unit vector $\vec{x}$ for which $A\vec{x}$ has maximum length

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

$$A^TA = \begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$$

$p(\lambda) = (8-\lambda)(5-\lambda) - 4 = \lambda^2 - 13\lambda + 36 = (\lambda - 9)(\lambda - 4)$

$\Rightarrow \lambda_{\max}\ (A^TA) = 9$

$$A^TA - 9I = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} \Rightarrow \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\left\|\begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\| = \frac{1}{\sqrt{5}}\left\|\begin{bmatrix} 3 \\ 6 \end{bmatrix}\right\| = \frac{\sqrt{45}}{\sqrt{5}} = 3 = \sigma_1$$

4. By inspection, construct an SVD of the diagonal matrix

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

$$A^TA = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$$

Eigenvalue of 4: $\pm\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

Eigenvalue of 9: $\pm\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$\Rightarrow$ 4 SVDs

The SVD of a matrix is not unique: how many different SVDs can you create from the matrix above?

5. **Written Explanation Exercise** Let A be an $m \times n$ matrix of rank r. If r is much smaller than m and n, explain how the following version of the singular value decomposition

$$A = \sigma_1 u_1 {u_1}^T + \cdots + \sigma_r u_r {u_r}^T$$

gives an efficient way to store A (this is called data compression).

Size of A: $m \times n$

Size of SVD decomposition: $r(m+n)$