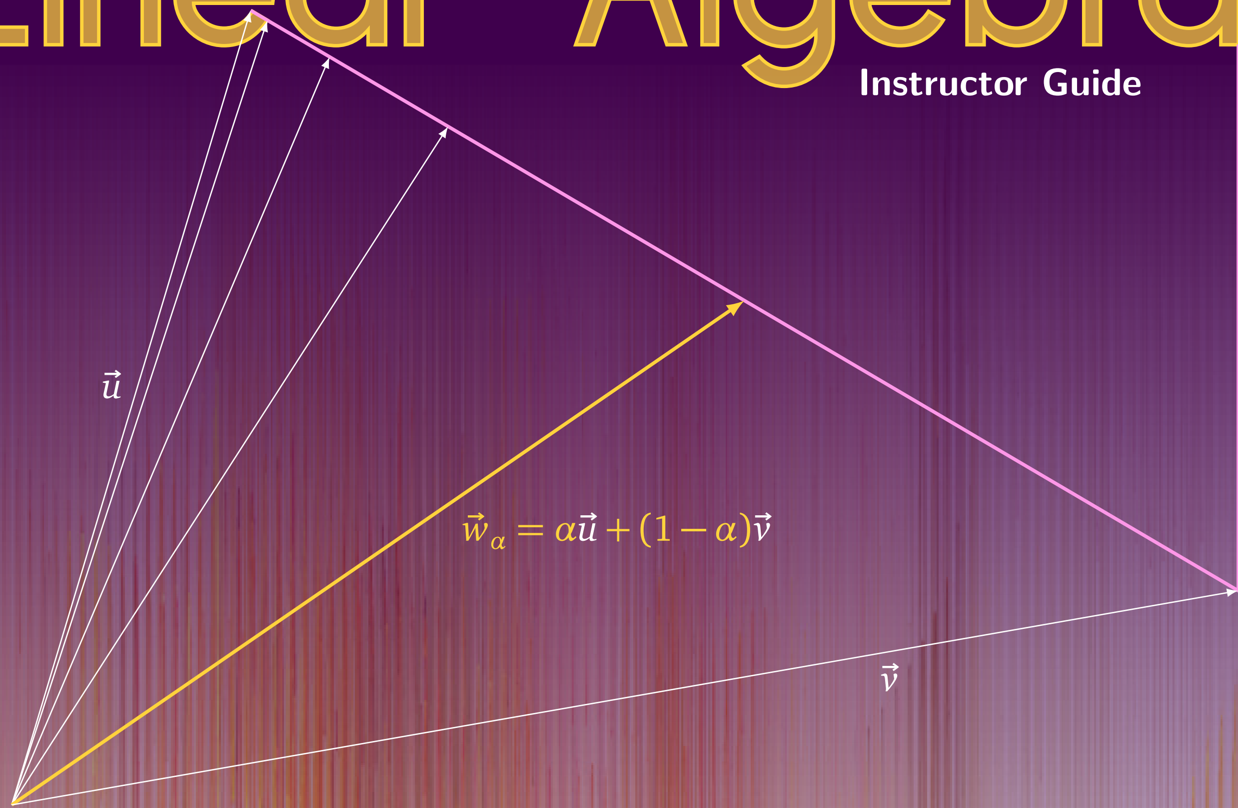
# Linear Algebra

## Instructor Guide




Jason Siefken

# Inquiry Based Linear Algebra



## About the Document

This document is a hybrid of many linear algebra resources, including those of the IOLA (Inquiry Oriented Linear Algebra) project, Jason Siefken's IBLLinearAlgebra project, and Asaki, Camfield, Moon, and Snipes' Radiograph and Tomography project.

This document is a mix of student projects, problem sets, and labs. A typical class day looks like:

1. **Introduction by instructor.** This may involve giving a definition, a broader context for the day's topics, or answering questions.

2. **Students work on problems.** Students work individually or in pairs on the prescribed problem. During this time the instructor moves around the room addressing questions that students may have and giving one-on-one coaching.

3. **Instructor intervention.** If most students have successfully solved the problem, the instructor regroups the class by providing a concise explanation so that everyone is ready to move to the next concept. This is also time for the instructor to ensure that everyone has understood the main point of the exercise (since it is sometimes easy to do some computation while being oblivious to the larger context).

   If students are having trouble, the instructor can give hints to the group, and additional guidance to ensure the students don't get frustrated to the point of giving up.

4. **Repeat step 2.**

Using this format, students are working (and happily so) most of the class. Further, they are especially primed to hear the insights of the instructor, having already invested substantially into each problem.

This problem-set is geared towards concepts instead of computation, though some problems focus on simple computation.



If you modify this document, you may add your name to the copyright list. Also, if you think your contributions would be helpful to others, consider making a pull requestion, or opening an *issue* at `https://github.com/siefkenj/IBLLinearAlgebra`

Lesson 1

# Linear Combinations

## Textbook

Section 1.1

## Objectives

- Internalize vectors as geometric objects representing displacements.
- Use column vector notation to write vectors.
- Relate points an vectors and be able to interpret a point as a vector and a vector as a point.
- Solve simple equations involving vectors.

## Motivation

Students have differing levels of experience with vectors. We want to establish a common notation for vectors and use vector notation along with algebra to solve simple questions. E.g., "How can I get to location $A$ given that I can only walk parallel to the lines $y = 4x$ and $y = -x$?"

We will use column vector notation and the idea of equating coordinates in order to solve problems.

**Notes/Misconceptions**

- We will use the language *component of $\vec{v}$ in the direction $\vec{u}$* in the future and it will be a *vector*. For this reason, try to refer to the entries of a column vector as *coordinates* or *entries* instead of components.
- Though we will almost exclusively use column vector notation in this course, students should be able to parse questions phrased in terms of row vectors.

# Task 1.1: The Magic Carpet Ride

You are a young traveler, leaving home for the first time. Your parents want to help you on your journey, so just before your departure, they give you two gifts. Specifically, they give you two forms of transportation: a hover board and a magic carpet. Your parents inform you that both the hover board and the magic carpet have restrictions in how they operate:

We denote the restriction on the hover board's movement by the vector $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$. By this we mean that if the hover board traveled "forward" for one hour, it would move along a "diagonal" path that would result in a displacement of 3 miles East and 1 mile North of its starting location.

We denote the restriction on the magic carpet's movement by the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. By this we mean that if the magic carpet traveled "forward" for one hour, it would move along a "diagonal" path that would result in a displacement of 1 mile East and 2 miles North of its starting location.

**Scenario One: The Maiden Voyage**

Your Uncle Cramer suggests that your first adventure should be to go visit the wise man, Old Man Gauss. Uncle Cramer tells you that Old Man Gauss lives in a cabin that is 107 miles East and 64 miles North of your home.

**Task:**

Investigate whether or not you can use the hover board and the magic carpet to get to Gauss's cabin. If so, how? If it is not possible to get to the cabin with these modes of transportation, why is that the case?

**Hands-on experience with vectors as displacements.**

- Internalize vectors as geometric objects representing displacements.
- Use column vector notation to write vectors.
- Use pre-existing knowledge of algebra to answer vector questions.

**Notes/Misconceptions**

- There are many ways to solve this problem. Some students might start with equations. After they use their equations to solve the problem, make them draw a picture and come up with a graphical solution.
- When the students start coming up with vector equations, give them the vocabulary of *linear combinations* and *column vector notation*.

Lesson 2

# Linear Combinations

## Textbook

Section 1.2

## Objectives

- Set up and solve vector equations $a\vec{v} + b\vec{u} = \vec{w}$. The solving method may be ad hoc.
- Use set notation and set operations/relations $\cup$, $\cap$, $\in$, $\subseteq$.
- Translate between set-builder notation and words in multiple ways.

## Motivation

We revisit questions about linear combinations more formally and generate a need for algebra. The algebra we do to solve vector equations will become algorithmic when we learn row reduction, but at the moment, any method is fine.

As we talk about more complex objects, we need precise ways to talk about groups of vectors. I.e., we need sets and set-builder notation. This preview of set-builder notation will take some of difficulty away when we define span as a set of vectors.

In this course we will be using formal and precise language. Part of this lesson is that there are multiple correct ways (and multiple incorrect ways) to use formal language. Gone are the days of "there's only one right answer and it is 4"!

**Notes/Misconceptions**

You will have a mix of MAT135/136 and MAT137 students. The MAT137 students will be doing logic and sets in their class. The MAT135 students won't. Make sure not to leave them behind!

# Task 1.2: The Magic Carpet Ride, Hide and Seek

You are a young traveler, leaving home for the first time. Your parents want to help you on your journey, so just before your departure, they give you two gifts. Specifically, they give you two forms of transportation: a hover board and a magic carpet. Your parents inform you that both the hover board and the magic carpet have restrictions in how they operate:

We denote the restriction on the hover board's movement by the vector $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$. By this we mean that if the hover board traveled "forward" for one hour, it would move along a "diagonal" path that would result in a displacement of 3 miles East and 1 mile North of its starting location.

We denote the restriction on the magic carpet's movement by the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. By this we mean that if the magic carpet traveled "forward" for one hour, it would move along a "diagonal" path that would result in a displacement of 1 mile East and 2 miles North of its starting location.

**Scenario Two: Hide-and-Seek**

Old Man Gauss wants to move to a cabin in a different location. You are not sure whether Gauss is just trying to test your wits at finding him or if he actually wants to hide somewhere that you can't visit him.

**Are there some locations that he can hide and you cannot reach him with these two modes of transportation?**

Describe the places that you can reach using a combination of the hover board and the magic carpet and those you cannot. Specify these geometrically and algebraically. Include a symbolic representation using vector notation. Also, include a convincing argument supporting your answer.

**Address an existential question involving vectors: "Is it possible to find a linear combination that does. . . ?"**

The goal of this problem is to
- Formalize geometric questions using the language of vectors.
- Find both geometric and algebraic arguments to support the same conclusion.
- Establish what a "negative multiple" of a vector should be.

**Notes/Misconceptions**

- Both *yes* and *no* are valid answers to this question depending on whether you are allowed to go backwards. Establish that "negative" multiples of a vector mean traveling backwards along that vector.
- This problem can be solved with algebra by finding a formula for the coefficients for an arbitrary position or with geometry, with arguments eventually hinging on the fact that non-parallel lines do not intersect.

# Sets and Set Notation

**Set**

DEFINITION

A ***set*** is a (possibly infinite) collection of items and is notated with curly braces (for example, $\{1, 2, 3\}$ is the set containing the numbers 1, 2, and 3). We call the items in a set ***elements***.

If $X$ is a set and $a$ is an element of $X$, we may write $a \in X$, which is read "$a$ is an element of $X$."

If $X$ is a set, a ***subset*** $Y$ of $X$ (written $Y \subseteq X$) is a set such that every element of $Y$ is an element of $X$. Two sets are called ***equal*** if they are subsets of each other (i.e., $X = Y$ if $X \subseteq Y$ and $Y \subseteq X$).

We can define a subset using ***set-builder notation***. That is, if $X$ is a set, we can define the subset

$$Y = \{a \in X : \text{some rule involving } a\},$$

which is read "$Y$ is the set of $a$ in $X$ **such that** some rule involving $a$ is true." If $X$ is intuitive, we may omit it and simply write $Y = \{a : \text{some rule involving } a\}$. You may equivalently use "|" instead of ":", writing $Y = \{a \mid \text{some rule involving } a\}$.

DEFINITION

Some common sets are

$$\mathbb{N} = \{\text{natural numbers}\} = \{\text{non-negative whole numbers}\}.$$

$$\mathbb{Z} = \{\text{integers}\} = \{\text{whole numbers, including negatives}\}.$$

$$\mathbb{R} = \{\text{real numbers}\}.$$

$$\mathbb{R}^n = \{\text{vectors in } n\text{-dimensional Euclidean space}\}.$$

1

1.1 Which of the following statements are true?

(a) $3 \in \{1, 2, 3\}$. True

(b) $1.5 \in \{1, 2, 3\}$. False

(c) $4 \in \{1, 2, 3\}$. False

(d) "b"$\in \{x : x \text{ is an English letter}\}$. True

(e) "ò"$\in \{x : x \text{ is an English letter}\}$. False

(f) $\{1, 2\} \subseteq \{1, 2, 3\}$. True

(g) For some $a \in \{1, 2, 3\}$, $a \geq 3$. True

(h) For any $a \in \{1, 2, 3\}$, $a \geq 3$. False

(i) $1 \subseteq \{1, 2, 3\}$. False

(j) $\{1, 2, 3\} = \{x \in \mathbb{R} : 1 \leq x \leq 3\}$. False

(k) $\{1, 2, 3\} = \{x \in \mathbb{Z} : 1 \leq x \leq 3\}$. True

**Practice reading sets and set-builder notation.**

The goal of this problem is to
- Become familiar with $\in$, $\subseteq$, and $=$ in the context of sets.
- Distinguish between $\in$ and $\subseteq$.
- Use quantifiers with sets.

**Notes/Misconceptions**

- Most are easy up through (h).
- Make students "fix" (i) so it becomes true.
- (j) and (k) are an opportunity to use the definition of set equality. Students don't realize that $=$'s has a definition.

2

Write the following in set-builder notation

2.1 The subset $A \subseteq \mathbb{R}$ of real numbers larger than $\sqrt{2}$.

$$\{x \in \mathbb{R} : x > \sqrt{2}\}.$$

2.2 The subset $B \subseteq \mathbb{R}^2$ of vectors whose first coordinate is twice the second.

$$\left\{\vec{v} \in \mathbb{R}^2 : \vec{v} = \begin{bmatrix} a \\ b \end{bmatrix} \text{ with } a = 2b\right\} \text{ or } \left\{\vec{v} \in \mathbb{R}^2 : \vec{v} = \begin{bmatrix} 2t \\ t \end{bmatrix} \text{ for some } t \in \mathbb{R}\right\}$$

$$\text{or } \left\{\begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a = 2b\right\}.$$

**Practice writing sets using set-builder notation.**

The goal of this problem is to
- Express English descriptions using math notation.
- Recognize there is more than one correct way to write formal math.
- Preview vector form of a line.

**Notes/Misconceptions**

- There are multiple correct ways to write each of these sets. It's a good opportunity to get man correct and incorrect sets up on the board for discussing.
- Don't worry about the geometry of $B$. That's coming in a later problem.

**Unions & Intersections**

DEFINITION

Two common set operations are ***unions*** and ***intersections***. Let $X$ and $Y$ be sets.

$$\text{(union)} \quad X \cup Y = \{a : a \in X \text{ or } a \in Y\}.$$

$$\text{(intersection)} \quad X \cap Y = \{a : a \in X \text{ and } a \in Y\}.$$

3 Let $X = \{1,2,3\}$ and $Y = \{2,3,4,5\}$ and $Z = \{4,5,6\}$. Compute

3.1 $X \cup Y$ $\{1,2,3,4,5\}$

3.2 $X \cap Y$ $\{2,3\}$

3.3 $X \cup Y \cup Z$ $\{1,2,3,4,5,6\}$

3.4 $X \cap Y \cap Z$ $\emptyset = \{\}$

Lesson 3

# Visualizing Sets, Formal Language of Linear Combinations

## Textbook

Section 1.2

## Objectives

- Draw pictures of formally-described subsets of $\mathbb{R}^2$.
- Graphically represent $\cup$ and $\cap$ for subsets of $\mathbb{R}^2$.
- Graphically represent linear combinations and then come up with algebraic arguments to support graphical intuition.

## Motivation

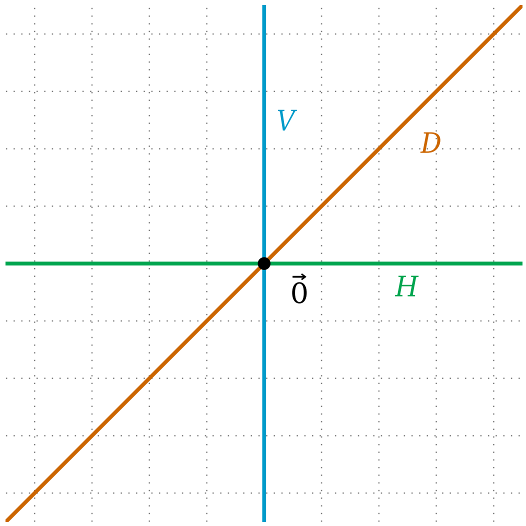
We want to build a bridge between the formal language of linear combinations and set-builder notation and geometric intuition. Where as last time the focus was on formal language, this time the focus is on linking geometry to formal descriptions.

4 Draw the following subsets of $\mathbb{R}^2$.

4.1 $V = \left\{ \vec{x} \in \mathbb{R}^2 : \vec{x} = \begin{bmatrix} 0 \\ t \end{bmatrix} \text{ for some } t \in \mathbb{R} \right\}$.

4.2 $H = \left\{ \vec{x} \in \mathbb{R}^2 : \vec{x} = \begin{bmatrix} t \\ 0 \end{bmatrix} \text{ for some } t \in \mathbb{R} \right\}$.

4.3 $D = \left\{ \vec{x} \in \mathbb{R}^2 : \vec{x} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ for some } t \in \mathbb{R} \right\}$.




4.4 $N = \left\{ \vec{x} \in \mathbb{R}^2 : \vec{x} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ for all } t \in \mathbb{R} \right\}$. $N = \{\}$.

4.5 $V \cup H$. $V \cup H$ looks like a "+" going through the origin.

4.6 $V \cap H$. $V \cap H = \{\vec{0}\}$ is just the origin.

4.7 Does $V \cup H = \mathbb{R}^2$?

No. $V \cup H$ does not contain $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ while $\mathbb{R}^2$ does contain $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

**Visualize sets of vectors.**

The goal of this problem is to
- Apply set-builder notation in the context of vectors.
- Distinguish between "for all" and "for some" in set builder notation.
- Practice unions and intersections.
- Practice thinking about set equality.

**Notes/Misconceptions**
- 1–3 will be easy.
- Have a discussion about when you should draw vectors as arrows vs. as points.
- 4 gets at a subtle point that will come up again when we define span.
- Many will miss 7. Writing a proof for this is good practice.

## Vector Combinations

**Linear Combination** (DEFINITION)

A ***linear combination*** of the vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ is a vector

$$\vec{w} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n.$$

The scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ are called the ***coefficients*** of the linear combination.

5 Let $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, and $\vec{w} = 2\vec{v}_1 + \vec{v}_2$.

5.1 Write $\vec{w}$ as a column vector. When $\vec{w}$ is written as a linear combination of $\vec{v}_1$ and $\vec{v}_2$, what are the coefficients of $\vec{v}_1$ and $\vec{v}_2$?

$\vec{w} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$; the coefficients are $(2, 1)$.

5.2 Is $\begin{bmatrix} 3 \\ 3 \end{bmatrix}$ a linear combination of $\vec{v}_1$ and $\vec{v}_2$? Yes. $\begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3\vec{v}_1 + 0\vec{v}_2$.

5.3 Is $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ a linear combination of $\vec{v}_1$ and $\vec{v}_2$? Yes. $\vec{0} = 0\vec{v}_1 + 0\vec{v}_2$.

5.4 Is $\begin{bmatrix} 4 \\ 0 \end{bmatrix}$ a linear combination of $\vec{v}_1$ and $\vec{v}_2$? Yes. $\begin{bmatrix} 4 \\ 0 \end{bmatrix} = 2\vec{v}_1 + 2\vec{v}_2$.

**Practice linear combinations.**

The goal of this problem is to
- Practice using the formal term *linear combination*.
- Foreshadow span.

**Notes/Misconceptions**
- In 2, the question should arise: "Is $3\vec{v}_1$ a linear combination of $\vec{v}_1$ *and* $\vec{v}_2$?" Address this.
- Refer to the magic carpet ride for 5. You don't need to do a full proof.

5.5 Can you find a vector in $\mathbb{R}^2$ that isn't a linear combination of $\vec{v}_1$ and $\vec{v}_2$?

No. $\begin{bmatrix}1\\0\end{bmatrix} = \frac{1}{2}\vec{v}_1 + \frac{1}{2}\vec{v}_2$ and $\begin{bmatrix}0\\1\end{bmatrix} = \frac{1}{2}\vec{v}_1 - \frac{1}{2}\vec{v}_2$. Therefore

$$\begin{bmatrix}a\\b\end{bmatrix} = a\begin{bmatrix}1\\0\end{bmatrix} + b\begin{bmatrix}0\\1\end{bmatrix} = a(\tfrac{1}{2}\vec{v}_1 + \tfrac{1}{2}\vec{v}_2) + b(\tfrac{1}{2}\vec{v}_1 - \tfrac{1}{2}\vec{v}_2) = (\tfrac{a+b}{2})\vec{v}_1 + (\tfrac{a-b}{2})\vec{v}_2.$$

Therefore any vector in $\mathbb{R}^2$ can be written as linear combinations of $\vec{v}_1$ and $\vec{v}_2$.

5.6 Can you find a vector in $\mathbb{R}^2$ that isn't a linear combination of $\vec{v}_1$?

Yes. All linear combinations of $\vec{v}_1$ have equal $x$ and $y$ coordinates, therefore $\vec{w} = \begin{bmatrix}2\\1\end{bmatrix}$ is not a linear combination of $\vec{v}_1$.

---

6 Recall the *Magic Carpet Ride* task where the hover board could travel in the direction $\vec{h} = \begin{bmatrix}3\\1\end{bmatrix}$ and the magic carpet could move in the direction $\vec{m} = \begin{bmatrix}1\\2\end{bmatrix}$.

**Practice formal writing.**

**Notes/Misconceptions**

- Make everyone *write*. They will think they can do it, but they will find it hard if they try.

6.1 Rephrase the sentence *"Gauss can be reached using just the magic carpet and the hover board"* using formal mathematical language.

Gauss's location can be written as a linear combination of $\vec{m}$ and $\vec{h}$.

6.2 Rephrase the sentence *"There is nowhere Gauss can hide where he is inaccessible by magic carpet and hover board"* using formal mathematical language.

Every vector in $\mathbb{R}^2$ can be written as a linear combination of $\vec{m}$ and $\vec{h}$.

6.3 Rephrase the sentence "*$\mathbb{R}^2$ is the set of all linear combinations of $\vec{h}$ and $\vec{m}$*" using formal mathematical language.

$\mathbb{R}^2 = \{\vec{v} : \vec{v} = t\vec{m} + s\vec{h} \text{ for some } t, s \in \mathbb{R}\}$.

Lesson 4

# Restricted Linear Combinations, Lines

## Textbook

Section 1.2

## Objectives

- Read and digest a new definition.
- Use pictures to explore a new concept.
- Convert from an equation-representation of a line to a set-representation.

## Motivation

Part of doing math in the world is reading and understanding other people's definitions. Most students will not have heard of non-negative linear combinations or convex linear combinations. This is a chance for them to read and try to understand these formal definitions. They will need to draw pictures to get an intuition about what these concepts mean.

These concepts are useful in their own right, and in particular, convex linear combinations can be used to describe line segments. Adding these definitions to a student's toolbox serves the goal of *being able to describe the world with mathematics*.

To that end, we start working with lines. Lines are something students have used since grade school, but they worked with them in $y = mx + b$ form which is only applicable in $\mathbb{R}^2$. We want to convert this representation into vector form and set-based descriptions which apply to all dimensions.

### Non-negative & Convex Linear Combinations

DEFINITION

The linear combination $\vec{w} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n$ is called a ***non-negative*** linear combination of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ if $\alpha_1, \alpha_2, \dots, \alpha_n \geq 0$.

If $\alpha_1, \alpha_2, \dots, \alpha_n \geq 0$ and $\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1$, then $\vec{w}$ is called a ***convex*** linear combination of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$.

7 Let

$$\vec{a} = \begin{bmatrix}1\\1\end{bmatrix} \quad \vec{b} = \begin{bmatrix}-1\\1\end{bmatrix} \quad \vec{c} = \begin{bmatrix}0\\1\end{bmatrix} \quad \vec{d} = \begin{bmatrix}0\\2\end{bmatrix} \quad \vec{e} = \begin{bmatrix}-1\\-1\end{bmatrix}.$$

7.1 Out of $\vec{a}$, $\vec{b}$, $\vec{c}$, $\vec{d}$, and $\vec{e}$, which vectors are

(a) linear combinations of $\vec{a}$ and $\vec{b}$? All of them, since any vector in $\mathbb{R}^2$ can be written as a linear combination of $\vec{a}$ and $\vec{b}$.

(b) non-negative linear combinations of $\vec{a}$ and $\vec{b}$? $\vec{a}$, $\vec{b}$, $\vec{c}$, $\vec{d}$.

(c) convex linear combinations of $\vec{a}$ and $\vec{b}$? $\vec{a}$, $\vec{b}$, $\vec{c}$.

7.2 If possible, find two vectors $\vec{u}$ and $\vec{v}$ so that

(a) $\vec{a}$ and $\vec{c}$ are non-negative linear combinations of $\vec{u}$ and $\vec{v}$ but $\vec{b}$ is not.

Let $\vec{u} = \vec{a}$ and $\vec{v} = \vec{c}$.

(b) $\vec{a}$ and $\vec{e}$ are non-negative linear combinations of $\vec{u}$ and $\vec{v}$.

Let $\vec{u} = \vec{a}$ and $\vec{v} = \vec{e}$.

(c) $\vec{a}$ and $\vec{b}$ are non-negative linear combinations of $\vec{u}$ and $\vec{v}$ but $\vec{d}$ is not.

Impossible. If $\vec{a}$ and $\vec{b}$ are non-negative linear combinations of $\vec{u}$ and $\vec{v}$, then every non-negative linear combination of $\vec{a}$ and $\vec{b}$ is also a non-negative linear combination of $\vec{u}$ and $\vec{v}$. And, we already concluded that $\vec{d}$ is a non-negative linear combination of $\vec{a}$ and $\vec{b}$.

(d) $\vec{a}$, $\vec{c}$, and $\vec{d}$ are convex linear combinations of $\vec{u}$ and $\vec{v}$.

Impossible. Convex linear combinations all lie on the same line segment, but $\vec{a}$, $\vec{c}$, and $\vec{d}$ are not collinear.

Otherwise, explain why it's not possible.

**Geometric meaning of *non-negative* and *convex* linear combinations.**

The goal of this problem is to

- Read and apply the definition of non-negative and convex linear combinations.
- Gain geometric intuition for non-negative and convex linear combinations.
- Learn how to describe line segments using convex linear combinations.

**Notes/Misconceptions**

- This question is about reading and applying; emphasize that before they start.
- The geometry won't be obvious. Ask them to *draw* specific linear combinations (e.g., $(1/2, 1/2)$) to get an idea.
- They know $\vec{a}$ and $\vec{b}$ span all vectors from problem 5.
- In part 1, they will forget $\vec{a}$ and $\vec{b}$ are linear combinations of themselves.
- Part 2 (b) highlights a degeneracy that will come up again when discussing linear independependence and dependence. Explain how the picture for non-negative linear combinations almost always looks one way, but this case is an exception.

## Lines and Planes

8 Let $L$ be the set of points $(x, y) \in \mathbb{R}^2$ such that $y = 2x + 1$.

8.1 Describe $L$ using set-builder notation.

$$\left\{ \vec{v} \in \mathbb{R}^2 \ : \ \vec{v} = \begin{bmatrix}t\\2t+1\end{bmatrix} \text{ for some } t \in \mathbb{R} \right\}$$

$$\text{or } \left\{ \begin{bmatrix}x\\y\end{bmatrix} \in \mathbb{R}^2 \ : \ y = 2x+1 \right\} \text{ or } \left\{ \begin{bmatrix}t\\2t+1\end{bmatrix} \in \mathbb{R}^2 \ : \ t \in \mathbb{R} \right\}$$

8.2 Draw $L$ as a subset of $\mathbb{R}^2$.

8.3 Add the vectors $\vec{a} = \begin{bmatrix}-1\\-1\end{bmatrix}$, $\vec{b} = \begin{bmatrix}1\\3\end{bmatrix}$ and $\vec{d} = \vec{b} - \vec{a}$ to your drawing.

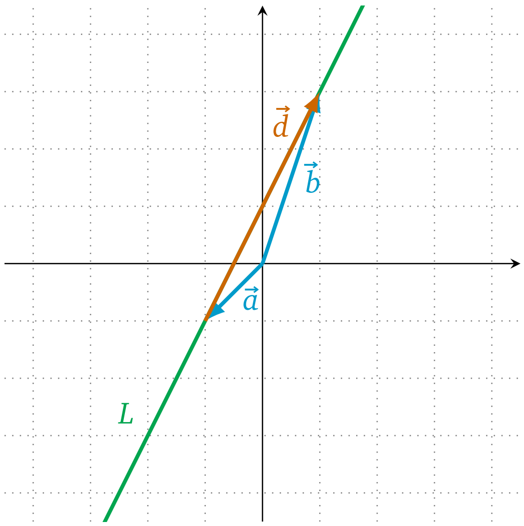
**Link prior knowledge to new notation/concepts.**

The goal of this problem is to

- Convert between $y = mx + b$ form of a line and the set-builder definition of the same line.
- Think about lines in terms of vectors rather than equations.

**Notes/Misconceptions**

- This question is foreshadowing for vector form of a line.
- In part 3, some will draw $\vec{d}$ from the origin and some will draw it on the line. Both are fine, but make sure they understand that $\vec{d} \notin L$ by the end of part 4.




8.4 Is $\vec{d} \in L$? Explain.

No. $\vec{d} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and so its entries don't satisfy $y = 2x + 1$.

8.5 For which $t \in \mathbb{R}$ is it true that $\vec{a} + t\vec{d} \in L$? Explain using your picture.

$\vec{a} + t\vec{d} \in L$ for any $t \in \mathbb{R}$. We can see this because if we start at the vector $\vec{a}$ and the displace by $t\vec{d}$, we will always be on the line $L$.

Lesson 5

# Vector Form of Lines, Intersecting Lines

## Textbook

Section 1.2

## Objectives

- Fluency with vector form of a line in $\mathbb{R}^2$ and $\mathbb{R}^3$.
- Recognize that vector form of a line is not unique.
- Find the intersection of two lines in vector form.

## Motivation

A single linear equation cannot describe a line in more than two dimensions. One way to describe a line that works in all dimensions is vector form, which is a shorthand for a particular set. Vector form has the upside that it makes it easy to produce points on a line, but it has the downside that it is not unique.

Vector form works because a line in any dimension can be defined by two points or, equivalently, a point and a direction. Though we don't yet have a systematic way to write solutions to a system of linear equations, if we have a system representing a line, all we need to do is guess two solutions to that system to find vector form of the line.

One thing vector form makes difficult is finding intersections, but intersections can be turned into just another algebra problem involving a system of equations.

**Notes/Misconceptions**

Giving a proper definition of vector form of a line is awkward and shouldn't be the focus. For vector form, that they "know it when they see it" and can "produce it" is good enough. (This is in contrast to other definitions which they must be able to correctly state).

**Notes/Misconceptions**

The biggest stumbling block for finding the intersection of two lines in vector form will be choosing different dummy variables before setting the lines equal.

**Vector Form of a Line**

DEFINITION

A line $\ell$ is written in ***vector form*** if it is expressed as

$$\vec{x} = t\vec{d} + \vec{p}$$

for some vector $\vec{d}$ and point $\vec{p}$. That is, $\ell = \{\vec{x} : \vec{x} = t\vec{d} + \vec{p} \text{ for some } t \in \mathbb{R}\}$. The vector $\vec{d}$ is called a ***direction vector*** for $\ell$.

---

9 Let $\ell \subseteq \mathbb{R}^2$ be the line with equation $2x + y = 3$, and let $L \subseteq \mathbb{R}^3$ be the line with equations $2x + y = 3$ and $z = y$.

9.1 Write $\ell$ in vector form. Is vector form of $\ell$ unique?

$\vec{x} = t\begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \end{bmatrix}$

The vector form is not unique, as any non-zero scalar multiple of $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ can serve as a direction vector. Additionally, any other point on the line can be used in place of $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$. For example, $\vec{x} = t\begin{bmatrix} -4 \\ 8 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is another vector form of $\ell$.

9.2 Write $L$ in vector form. $\vec{x} = t\begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \\ 3 \end{bmatrix}$. This is obtained by finding two points: one when $x = 0$ and one when $x = 1$ and subtracting them to find a direction vector for $L$.

9.3 Find another vector form for $L$ where both "$\vec{d}$" and "$\vec{p}$" are different from before.

$\vec{x} = t\begin{bmatrix} -3 \\ 6 \\ 6 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$

Again, any non-zero scalar multiple of the direction vector will work for $\vec{d}$, as will any other point on the line work for $\vec{p}$.

**Practice with vector form.**

The goal of this problem is to

- Express lines in $\mathbb{R}^2$ and $\mathbb{R}^3$ in vector form.
- Produce direction vectors by subtracting two points on a line.
- Recognize vector form is not unique.

**Notes/Misconceptions**

- If students get stuck on part 1, ask them to find a vector parallel to $\ell$. If they're still stuck, ask them to find a vector connecting two points on $\ell$.
- Many students will intuit part 1 but get stuck on part 2 because they can't draw it. Ask them to start by finding some points on $L$.

---

10 Let $A$, $B$, and $C$ be given in vector form by

$$\overbrace{\vec{x} = t\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}^{A} \qquad \overbrace{\vec{x} = t\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}}^{B} \qquad \overbrace{\vec{x} = t\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}^{C}.$$

10.1 Do the lines $A$ and $B$ intersect? Justify your conclusion.

Yes. $(0)\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = (-1)\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}.$

To find the intersection, if there is one, we must solve the vector equation:

$$t\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = s\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}.$$

One solution is when $t = 0$ and $s = -1$.

10.2 Do the lines $A$ and $C$ intersect? Justify your conclusion.

No. The vector equation

$$t\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = s\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

**Intersect lines in vector form.**

The goal of this problem is to

- Practice computing the intersection between lines in vector form.
- Recognize "$t$" as a dummy variable as used in vector form and that, when comparing lines in vector form, "$t$" needs to be replaced with non-dummy variables.

**Notes/Misconceptions**

- The most common mistake is to set two lines in vector form equal and use "$t$" as the variable in each one. Make sure to have a discussion about this.

has no solutions. This is equivalent to saying that the following system of equations has no solutions:

$$\begin{aligned} t &= 2s + 1 \\ 2t &= -s + 1 \\ 3t + 1 &= s + 1 \end{aligned}$$

The third equation tells us that $s = 3t$, which when substituted into the first equation forces $t = -\frac{1}{5}$ and therefore $s = -\frac{3}{5}$. However, these two numbers don't satisfy the second equation.

10.3 Let $\vec{p} \neq \vec{q}$ and suppose $X$ has vector form $\vec{x} = t\vec{d} + \vec{p}$ and $Y$ has vector form $\vec{x} = t\vec{d} + \vec{q}$. Is it possible that $X$ and $Y$ intersect?

Yes. If $\vec{q} = \vec{p} + a\vec{d}$ for $a \neq 0$, then $X$ and $Y$ will actually be the same line, since in this case

$$\vec{x} = t\vec{d} + \vec{q} = t\vec{d} + (\vec{p} + a\vec{d}) = (t + a)\vec{d} + \vec{p}.$$

For example, the following two vector equations represent the same line.

$$\vec{x} = t\begin{bmatrix}1\\1\\1\end{bmatrix} + \begin{bmatrix}0\\0\\0\end{bmatrix} \qquad \text{and} \qquad \vec{x} = t\begin{bmatrix}1\\1\\1\end{bmatrix} + \begin{bmatrix}7\\7\\7\end{bmatrix}.$$

Lesson 6

# Planes, Span

## Textbook

Section 1.2

## Objectives

- Describe a plane in vector form.
- Visualize spans.
- Recognize the dimension of span($X$) is not necessarily how many vectors are in $X$.
- Define *span*.

## Motivation

Planes are just like lines but one dimension higher. Vector form of a plane is just like vector form of a plane with all the advantages and disadvantages. But, we now have *two* direction vectors.

Spans are similar to lines and planes; span$\{\vec{a}, \vec{b}\}$ looks a lot like vector form of the plane $\vec{x} = t\vec{a} + s\vec{b}$. Except, span$\{\vec{a}, \vec{b}\}$ may not always be a plane. We haven't defined linear independence and linear dependence yet, but we will continue to foreshadow it by seeing that the dimension of the span of a set is not always the size of that set.

Knowing definitions is an essential part of solving math problems. Span is the first definition that students will think they "know" but won't be able to write down.

**Vector Form of a Plane**

DEFINITION

A plane $\mathcal{P}$ is written in ***vector form*** if it is expressed as

$$\vec{x} = t\vec{d}_1 + s\vec{d}_2 + \vec{p}$$

for some vectors $\vec{d}_1$ and $\vec{d}_2$ and point $\vec{p}$. That is, $\mathcal{P} = \{\vec{x} : \vec{x} = t\vec{d}_1 + s\vec{d}_2 + \vec{p} \text{ for some } t, s \in \mathbb{R}\}$. The vectors $\vec{d}_1$ and $\vec{d}_2$ are called ***direction vectors*** for $\mathcal{P}$.

---

11 Recall the intersecting lines $A$ and $B$ given in vector form by

$$\overbrace{\vec{x} = t\begin{bmatrix}1\\2\\3\end{bmatrix} + \begin{bmatrix}0\\0\\1\end{bmatrix}}^{A} \qquad \overbrace{\vec{x} = t\begin{bmatrix}-1\\1\\1\end{bmatrix} + \begin{bmatrix}-1\\1\\2\end{bmatrix}}^{B}.$$

Let $\mathcal{P}$ the plane that contains the lines $A$ and $B$.

11.1 Find two direction vectors for $\mathcal{P}$.

Two possible answers are:

$$\vec{d}_1 = \begin{bmatrix}1\\2\\3\end{bmatrix} \qquad \text{and} \qquad \vec{d}_2 = \begin{bmatrix}-1\\1\\1\end{bmatrix}.$$

These are the two direction vectors we already know are in the plane—the ones from the two lines:

Note that neither of these is a multiple of the other, so they really are two unique direction vectors in $\mathcal{P}$.

11.2 Write $\mathcal{P}$ in vector form.

$$\vec{x} = t\vec{d}_1 + s\vec{d}_2 + \vec{p} = t\begin{bmatrix}1\\2\\3\end{bmatrix} + s\begin{bmatrix}-1\\1\\1\end{bmatrix} + \begin{bmatrix}0\\0\\1\end{bmatrix}.$$

We already have two direction vectors, so we just needed a point on the plane. We used the point $\vec{p} = \begin{bmatrix}0\\0\\1\end{bmatrix}$ that we already know is on line $A$.

11.3 Describe how vector form of a plane relates to linear combinations.

The vector form of a plane says that a vector $\vec{x}$ is on the plane exactly when it is equal to some linear combination of $\vec{d}_1$ and $\vec{d}_2$, plus $\vec{p}$.

Another way of saying the same thing is that the vector $\vec{x}$ is on the plane exactly when $\vec{x} - \vec{p}$ is equal to some linear combination of $\vec{d}_1$ and $\vec{d}_2$.

11.4 Write $\mathcal{P}$ in vector form using different direction vectors and a different point.

One possible answer:

$$\vec{x} = t\begin{bmatrix}-1\\-2\\-3\end{bmatrix} + s\begin{bmatrix}-7\\7\\7\end{bmatrix} + \begin{bmatrix}-1\\1\\2\end{bmatrix}.$$

As with the equations of lines from before, we can use any non-zero scalar multiple of either direction vector and get the same plane. We also used the point $\vec{q} = \begin{bmatrix}-1\\1\\2\end{bmatrix}$ that we already knew is on line $B$.

**Apply vector form of a plane.**

The goal of this problem is to
- Use direction vectors for lines given in vector form.
- Think about planes in terms of vectors rather than equations.
- Combine direction vectors in a plane to produce new direction vectors.

**Notes/Misconceptions**

- Students may think they need to find the intersection of the lines to serve as their "$\vec{p}$". They don't. All they need is a point on the plane!

**Connect vector form and scalar form of a plane.**

The goal of this problem is to
- Produce direction vectors for a plane defined by an equation.
- Generalize the procedure for finding direction vectors that was used for lines.

**Notes/Misconceptions**

- This problem is scaffolded and should be straight forward.
- You have the opportunity to discuss if $\vec{x} = t\vec{d} + s(-\vec{d}) + \vec{p}$ is a valid vector form of $\mathcal{P}$. The two direction vectors are different, after all.

12 Let $\mathcal{Q} \subseteq \mathbb{R}^3$ be a plane with equation $x+y+z=1$.

12.1 Find three points in $\mathcal{Q}$.

There are many choices here, of course. Three natural ones are:

$$\vec{p}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} \qquad \vec{p}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix} \qquad \vec{p}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$$

12.2 Find two direction vectors for $\mathcal{Q}$.

Now that we have three points on the plane, we can use the direction vectors joining any two pairs of them. For example:

$$\vec{d}_1 = \vec{p}_1 - \vec{p}_2 = \begin{bmatrix}1\\-1\\0\end{bmatrix} \qquad \vec{d}_2 = \vec{p}_1 - \vec{p}_3 = \begin{bmatrix}1\\0\\-1\end{bmatrix}.$$

12.3 Write $\mathcal{Q}$ in vector form.

Using the point $\vec{p}_1$ from above, one possible answer is:

$$\vec{x} = t\vec{d}_1 + s\vec{d}_2 + \vec{p}_1 = t\begin{bmatrix}1\\-1\\0\end{bmatrix} + s\begin{bmatrix}1\\0\\-1\end{bmatrix} + \begin{bmatrix}1\\0\\0\end{bmatrix}.$$

## Span

**Span**

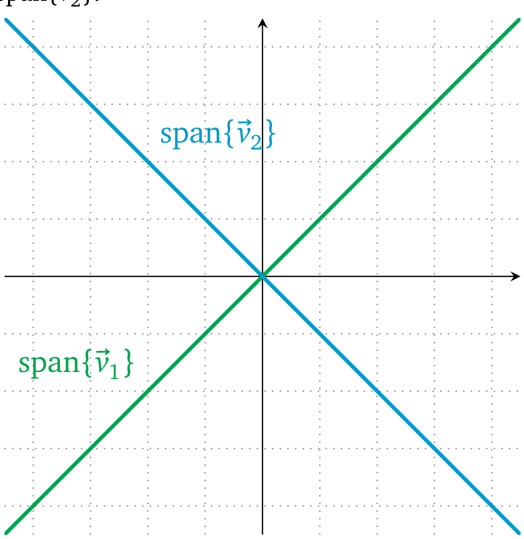
DEF The ***span*** of a set of vectors $V$ is the set of all linear combinations of vectors in $V$. That is,

$$\operatorname{span} V = \{\vec{v} : \vec{v} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n \text{ for some } \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in V \text{ and scalars } \alpha_1, \alpha_2, \ldots, \alpha_n\}.$$

13 Let $\vec{v}_1 = \begin{bmatrix}1\\1\end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix}1\\-1\end{bmatrix}$, and $\vec{v}_3 = \begin{bmatrix}2\\2\end{bmatrix}$.

13.1 Draw $\operatorname{span}\{\vec{v}_1\}$.

13.2 Draw $\operatorname{span}\{\vec{v}_2\}$.




13.3 Describe $\operatorname{span}\{\vec{v}_1, \vec{v}_2\}$.

$\operatorname{span}\{\vec{v}_1, \vec{v}_2\} = \mathbb{R}^2$

We can see this since for any $\begin{bmatrix}x\\y\end{bmatrix} \in \mathbb{R}^2$,

$$\begin{bmatrix}x\\y\end{bmatrix} = \frac{x}{2}\left(\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}1\\-1\end{bmatrix}\right) + \frac{y}{2}\left(\begin{bmatrix}1\\1\end{bmatrix} - \begin{bmatrix}1\\-1\end{bmatrix}\right) = \frac{x+y}{2}\vec{v}_1 + \frac{x-y}{2}\vec{v}_2$$

**Notes/Misconceptions**

- There's an opportunity for another "for some" vs. "for all" discussion involving the definition of span. Many students won't understand why "the set of all linear combinations" should be written with a "for some $\vec{v}_1 \ldots$".

**Apply the definition of *span*.**

The goal of this problem is to

- Practice applying a new definition in a familiar context ($\mathbb{R}^2$).
- Recognize spans as lines and planes through the origin.

**Notes/Misconceptions**

- These vectors have been used in other problems, so the bulk of the argument for part 3 has already been made.
- Parts 4 & 5 foreshadow linear dependence. Emphasize that you cannot tell the size (dimension) of $\operatorname{span} X$ by knowing the number of vectors in $X$.

13.4 Describe span$\{\vec{v}_1, \vec{v}_3\}$. span$\{\vec{v}_1, \vec{v}_3\}$ = span$\{\vec{v}_1\}$, a line through the origin with direction vector $\vec{v}_1$.

13.5 Describe span$\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. span$\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ = span$\{\vec{v}_1, \vec{v}_2\} = \mathbb{R}^2$

Lesson 7

# Span, Translated Span

## Textbook

Section 1.2

## Objectives

- Explain why spans always go through the origin.
- Express lines or planes through the origin as spans.
- Express lines or planes not through the origin as translated spans.

## Motivation

Translated spans link vector form of lines and planes with sets and spans. Soon we will have the vocabulary of linear independence and be able to talk about independent direction vectors of a plane, but right now just connecting the concepts and notation is enough.

14 Let $\ell_1 \subseteq \mathbb{R}^2$ be the line with equation $x - y = 0$ and $\ell_2 \subseteq \mathbb{R}^2$ the line with equation $x - y = 4$.

14.1 If possible, describe $\ell_1$ as a span. Otherwise explain why it's not possible.

$\ell_1 = \operatorname{span}\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$, since $\begin{bmatrix}x\\y\end{bmatrix} \in \ell_1$ if and only if $x = y$, which in turn is true if and only if it is a scalar multiple of $\begin{bmatrix}1\\1\end{bmatrix}$.

14.2 If possible, describe $\ell_2$ as a span. Otherwise explain why it's not possible.

This is not possible. $\begin{bmatrix}0\\0\end{bmatrix}$ is an element of the span of *any* set of vectors, since we can use all zeroes as the scalars in a linear combination, but $\begin{bmatrix}0\\0\end{bmatrix} \notin \ell_2$.

14.3 Does the expression $\operatorname{span}(\ell_1)$ make sense? If so, what is it? How about $\operatorname{span}(\ell_2)$?

Both of these expressions do make sense. One can compute the span of any set of vectors, and these lines are just special set of points in $\mathbb{R}^2$ which we are already used to thinking of as vectors.

$\operatorname{span}(\ell_1) = \ell_1$, since all of the vectors on the line $\ell_1$ are already multiples of $\begin{bmatrix}1\\1\end{bmatrix}$, as we discovered earlier.

$\operatorname{span}(\ell_2)$ equals all of $\mathbb{R}^2$. It's easy to see that the vectors $v = \begin{bmatrix}4\\0\end{bmatrix}$ and $w = \begin{bmatrix}0\\-4\end{bmatrix}$ are both on $\ell_2$, and the span of these two vectors alone is all of $\mathbb{R}^2$.

**Connect geometric figures to spans.**

The goal of this problem is to

- Identify a relationship between lines and spans.
- Describe a line through the origin as a span.
- Identify when a line cannot be described as a span.
- Apply the definition of $\operatorname{span} X$ even when $X$ is infinite.

**Notes/Misconceptions**

- The lines are not written in $y = mx + b$ form on purpose. We avoid this form in linear algebra since it cannot describe all lines.
- Part 3 will really stretch their minds. Students at this point are not used to applying definitions. They will have a conception of $\operatorname{span} X$ where $X$ is finite and will forget the definition because this conception is "good enough". This question forces them to think back to the definition.

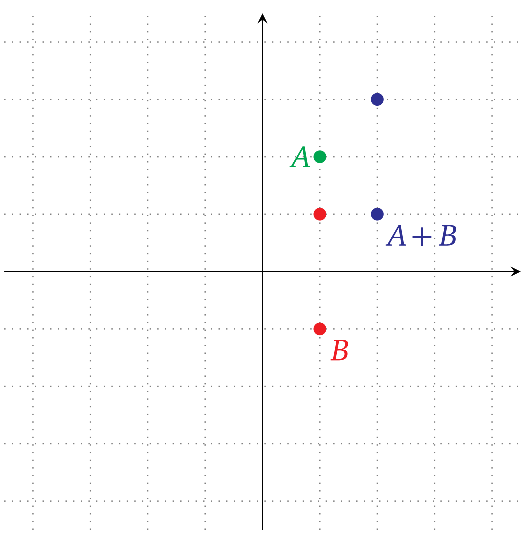
**Set Addition** (DEF)

If $A$ and $B$ are sets of vectors, then the ***set sum*** of $A$ and $B$, denoted $A + B$, is

$$A + B = \{\vec{x} \,:\, \vec{x} = \vec{a} + \vec{b} \text{ for some } \vec{a} \in A \text{ and } \vec{b} \in B\}.$$

15 Let $A = \left\{\begin{bmatrix}1\\2\end{bmatrix}\right\}$, $B = \left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}\right\}$, and $\ell = \operatorname{span}\left\{\begin{bmatrix}1\\-1\end{bmatrix}\right\}$.

15.1 Draw $A$, $B$, and $A + B$ in the same picture.




15.2 Is $A + B$ the same as $B + A$?

Yes. Since $A$ and $B$ are such small sets we could just compute all the vectors in $A + B$ and $B + A$ and see that they're equal. However, we know that real numbers can be added up in any order, and the coordinates of an element of $A + B$ or $B + A$ are simply sums of the corresponding coordinates of elements of $A$ and $B$.

15.3 Draw $\ell + A$.

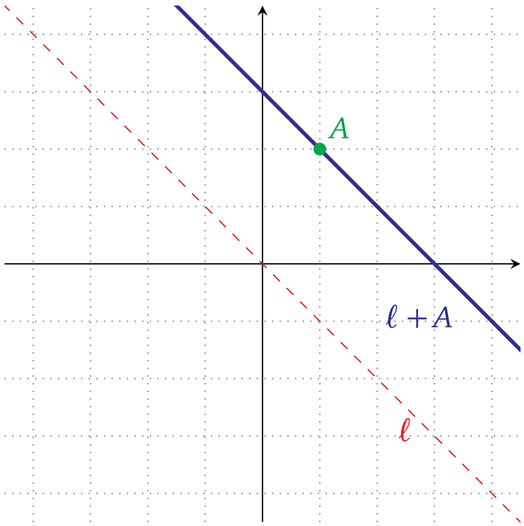
**Describing geometry using sets.**

The goal of this problem is to

- Practice applying a new definition in a familiar context ($\mathbb{R}^2$).
- Gain an intuitive understanding of set addition.
- Describe lines that don't pass through $\vec{0}$ using a combination of set addition and spans.

**Notes/Misconceptions**

- Set addition will be brand new to most students, even those with a linear algebra background.
- Special care must be taken to differentiate $\vec{a} + \vec{b}$ and $\{\vec{a}\} + \{\vec{b}\}$.
- We will use set addition mainly as a mathematical notation to describe translated spans. In this case the summands will be an infinite set and a singleton. There is no need to explore the sum of two infinite sets.




15.4 Consider the line $\ell_2$ given in vector form by $\vec{x} = t\begin{bmatrix}1\\-1\end{bmatrix} + \begin{bmatrix}1\\2\end{bmatrix}$. Can $\ell_2$ be described using only a span? What about using a span and set addition?

$\ell_2$ cannot be described using only a span, for the same reason as the line $\ell_2$ in Problem 14.2 couldn't be. We know that the origin $\begin{bmatrix}0\\0\end{bmatrix}$ must be an element of any span, but it is not a point on $\ell_2$.

$\ell_2$ can be described as a span plus a set addition though. Specifically, $\ell_2 = \ell + A$.

Lesson 8

# Linear Independence & Dependence

## Textbook

Section 1.2

## Objectives

- Define linear independence/dependence using spans.
- Pick linearly independent subsets with the same span by inspection.
- Explain why having a "closed loop" or trivial linear combination means a set is linearly dependent.

## Motivation

Linear independence/dependence is one of the biggest concepts in linear algebra. Linear independence/dependence tells us whether a set has redundant information in it with respect to spans. The idea of a having redundant information vs. not comes up all the time in the world (sometimes it's a plus, sometimes it's not).

Knowing a set is independent tells us what its span will look like (in terms of what dimension it will be). It is also an abstract concept that has both a "geometric" definition and an "algebraic" one. Geometrically, a set is linearly dependent if you can remove a vector without the span changing. Algebraically a set is linearly dependent if there is a non-trivial linear combination giving the zero vector. This lesson focuses on the geometric definition (with the algebraic definition coming next).

Though the algebraic definition is easier to work with in proofs, the geometric definition provides intuition about how to visualize linearly dependent sets.

**Notes/Misconceptions**

Don't define a linearly dependent *set*, define a linearly dependent *list*. Otherwise you cannot talk about $\begin{bmatrix}1\\1\end{bmatrix}$ and $\begin{bmatrix}1\\1\end{bmatrix}$ be linearly dependent since sets don't contain duplicates.

# Task 1.3: The Magic Carpet, Getting Back Home

Suppose you are now in a three-dimensional world for the carpet ride problem, and you have three modes of transportation:

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{v}_2 = \begin{bmatrix} 6 \\ 3 \\ 8 \end{bmatrix} \qquad \vec{v}_3 = \begin{bmatrix} 4 \\ 1 \\ 6 \end{bmatrix}$$

You are only allowed to use each mode of transportation **once** (in the forward or backward direction) for a fixed amount of time ($c_1$ on $\vec{v}_1$, $c_2$ on $\vec{v}_2$, $c_3$ on $\vec{v}_3$).

1. Find the amounts of time on each mode of transportation ($c_1$, $c_2$, and $c_3$, respectively) needed to go on a journey that starts and ends at home *or* explain why it is not possible to do so.
2. Is there more than one way to make a journey that meets the requirements described above? (In other words, are there different combinations of times you can spend on the modes of transportation so that you can get back home?) If so, how?
3. Is there anywhere in this 3D world that Gauss could hide from you? If so, where? If not, why not?
4. What is $\text{span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 6 \\ 3 \\ 8 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 6 \end{bmatrix}\right\}$?

**Span in higher dimensions.**

The goal of this problem is to

- Examine subtleties that exist in three dimensions that are missing in two dimensions.
- Apply linear algebra tools to answer open-ended questions.

**Notes/Misconceptions**

This problem is set up to prime theorems about linearly dependent vectors. In particular, it

- gives an example of a non-trivial linear combination giving $\vec{0}$;
- shows that if there is one non-trivial linear combination giving $\vec{0}$, there are others;
- shows that 3 non-parallel vectors in $\mathbb{R}^3$ need not span $\mathbb{R}^3$.

The problem also allows linking to previous linear algebra ideas. A system of equations can be used to find *all* non-trivial solutions, and showing a particular system of equations is inconsistent will show that the span is not $\mathbb{R}^3$.

## Linearly Dependent & Independent (Geometric)

DEFINITION

We say the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are ***linearly dependent*** if for at least one $i$,

$$\vec{v}_i \in \operatorname{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{i-1}, \vec{v}_{i+1}, \dots, \vec{v}_n\}.$$

Otherwise, they are called ***linearly independent***.

**Geometric definition of linear independence/dependence.**

**Notes/Misconceptions**

- This definition is conceptually simple but notationally hard.
- The definition is phrased in terms of a list of vectors (instead of a set) to avoid issues with the fact that sets cannot have repeated elements. (e.g., if $\vec{a} \neq \vec{0}$, then the set $\{\vec{a}, \vec{a}\} = \{\vec{a}\}$ is linearly independent, whereas the list of vectors $\vec{a}, \vec{a}$ is linearly dependent.)
- Many students will not realize that $\vec{v}_i$ is being "left out" of the span.
- Students might assume, for example, that $\vec{v}_1$ could always be removed from the span. This misconception is targeted in a later problem.
- Have students rephrase this definition in plain language.

---

16

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

16.1 Describe $\operatorname{span}\{\vec{u}, \vec{v}, \vec{w}\}$.

The $xy$-plane in $\mathbb{R}^3$. That is, the set of all vectors in $\mathbb{R}^3$ with $z$-coordinate equal to zero.

16.2 Is $\{\vec{u}, \vec{v}, \vec{w}\}$ linearly independent? Why or why not?

No. $\vec{w} = \vec{u} + \vec{v}$, and so $\vec{w} \in \operatorname{span}\{\vec{u}, \vec{v}\}$.

Let $X = \{\vec{u}, \vec{v}, \vec{w}\}$.

16.3 Give a subset $Y \subseteq X$ so that $\operatorname{span} Y = \operatorname{span} X$ and $Y$ is linearly independent.

$Y = \{\vec{u}, \vec{v}\}$ is one example that works.

16.4 Give a subset $Z \subseteq X$ so that $\operatorname{span} Z = \operatorname{span} X$ and $Z$ is linearly independent and $Z \neq Y$.

$Z = \{\vec{u}, \vec{w}\}$ and $Z = \{\vec{v}, \vec{w}\}$ both have the same span as $Y$ above.

**Apply the (geometric) definition of linear independence/dependence.**

The goal of this problem is to

- Develop a mental picture linking linear dependence and "redundant" vectors.
- Practice applying a new definition.
- Find multiple linearly independent subsets of a linearly dependent set.

**Notes/Misconceptions**

- Students won't find this problem hard.
- There may be some confusion about what "describe" means.

## Trivial Linear Combination

DEF

We say a linear combination $a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n$ is ***trivial*** if $a_1 = a_2 = \cdots = a_n = 0$.

---

17

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

17.1 Consider the linearly dependent set $\{\vec{u}, \vec{v}, \vec{w}\}$ (where $\vec{u}, \vec{v}, \vec{w}$ are defined as above). Can you write $\vec{0}$ as a non-trivial linear combination of vectors in this set? $\vec{0} = \vec{u} + \vec{v} - \vec{w}$.

17.2 Consider the linearly independent set $\{\vec{u}, \vec{v}\}$. Can you write $\vec{0}$ as a non-trivial linear combination of vectors in this set?

No. Suppose

$$a_1\vec{u} + a_2\vec{v} = \vec{0}$$

was a non-trivial linear combination. Then at least one of $a_1$ or $a_2$ is non-zero. If $a_1$ is non-zero, then

$$\vec{u} = -\frac{a_2}{a_1}\vec{v}$$

and so $\vec{u} \in \operatorname{span}\{\vec{v}\}$. If $a_2$ is non-zero, then

$$\vec{v} = -\frac{a_1}{a_2}\vec{u}.$$

and so $\vec{v} \in \operatorname{span}\{\vec{u}\}$. In either case, $\{\vec{u}, \vec{v}\}$ would be linearly dependent.

**Link trivial/non-trivial linear combinations to linear independence/dependence.**

**Notes/Misconceptions**

- Part 2 is a chance to practice writing arguments.
- Many students will mistakenly assume that non-trivial means *all* coefficients are non-zero and use this in their argument. Bring out this misconception.

Lesson 9

# Linear Independence & Dependence—Equivalent Definitions

## Textbook

Section 1.2

## Objectives

- Define linear independence/dependence in terms of trivial linear combinations.
- Explain how the geometric and algebraic definitions of linear independence/dependence relate.
- Explain the connection between a vector equation having multiple solutions and those vectors being linearly independent/dependent.
- Identify the largest linearly independent set that could exist in $\mathbb{R}^n$.

## Motivation

We've done geometry, now let's do algebra. The geometric and algebraic definitions are equivalent, but they suggest different consequences. The geometric definition of linear independence tells us about the dimension of a span. The algebraic definition tells us about the number of solutions to a vector equation.

We now have an equivalent definition of linear dependence.

**Linearly Dependent & Independent (Algebraic)**

DEF The vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are ***linearly dependent*** if there is a non-trivial linear combination of $\vec{v}_1, \dots, \vec{v}_n$ that equals the zero vector.

**Notes/Misconceptions**

- The algebraic definition of linear independence/dependence is good for proofs but hard to intuit.

---

18

18.1 Explain how this algebraic definition (new) implies the geometric one (original).

Suppose the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ is linearly dependent in this new sense. That means there are scalars $a_1, a_2, \dots, a_n$, at least one of which is non-zero, such that

$$a_1\vec{v}_1 + \cdots + a_n\vec{v}_n = 0.$$

Suppose $a_i \neq 0$. Then

$$\vec{v}_i = \tfrac{a_1}{a_i}\vec{v}_1 + \cdots + \tfrac{a_{i-1}}{a_i}\vec{v}_{i-1} + \tfrac{a_{i+1}}{a_i}\vec{v}_{i+1} + \cdots + \tfrac{a_n}{a_i}\vec{v}_n.$$

This means $\vec{v}_i \in \operatorname{span}\{\vec{v}_1, \dots, \vec{v}_{i-1}, \vec{v}_{i+1}, \dots, \vec{v}_n\}$, which is precisely the geometric definition of linear dependence.

18.2 Explain how the geometric definition (original) implies this algebraic one (new).

Suppose that $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are linearly dependent according to the geometric definition. Fix $i$ so that $\vec{v}_i \in \operatorname{span}\{\vec{v}_1, \dots, \vec{v}_{i-1}, \vec{v}_{i+1}, \dots, \vec{v}_n\}$.

By the definition of span, we know that

$$\vec{v}_i = \beta_1\vec{v}_1 + \cdots + \beta_{i-1}\vec{v}_{i-1} + \beta_{i+1}\vec{v}_{i+1} + \cdots + \beta_n\vec{v}_n.$$

Thus

$$\vec{0} = -\vec{v}_i + \beta_1\vec{v}_1 + \cdots + \beta_{i-1}\vec{v}_{i-1} + \beta_{i+1}\vec{v}_{i+1} + \cdots + \beta_n\vec{v}_n,$$

and this is a non-trivial linear combination since the coefficient of $\vec{v}_i$ is $-1 \neq 0$.

Since we have geometric def $\implies$ algebraic def, and algebraic def $\implies$ geometric def ($\implies$ should be read aloud as 'implies'), the two definitions are *equivalent* (which we write as algebraic def $\iff$ geometric def).

**Link algebraic and geometric definitions of linear independence/dependence.**

The goal of this problem is to

- Understand how the algebraic and geometric definitions of linear independence/dependence relate.
- Practice writing mathematical arguments.

**Notes/Misconceptions**

- You could easily spend an entire class working with students to get well-written proofs. Be mindful of time. Since this is not a proofs-based course, we want to focus on the intuition here.

---

19

Suppose for some unknown $\vec{u}, \vec{v}, \vec{w}$, and $\vec{a}$,

$$\vec{a} = 3\vec{u} + 2\vec{v} + \vec{w} \qquad \text{and} \qquad \vec{a} = 2\vec{u} + \vec{v} - \vec{w}.$$

19.1 Could the set $\{\vec{u}, \vec{v}, \vec{w}\}$ be linearly independent?

No. If both equations are true, they would combine to show

$$3\vec{u} + 2\vec{v} + \vec{w} = 2\vec{u} + \vec{v} - \vec{w}.$$

Collecting all the terms on the left side, we get:

$$\vec{u} + \vec{v} + 2\vec{w} = \vec{0},$$

which is a non-trivial linear combination of vectors in the given set equalling the zero vector.

Suppose that

$$\vec{a} = \vec{u} + 6\vec{r} - \vec{s}$$

is the *only* way to write $\vec{a}$ using $\vec{u}, \vec{r}, \vec{s}$.

19.2 Is $\{\vec{u}, \vec{r}, \vec{s}\}$ linearly independent?

Yes. If it were not, there would exist scalars $a_1, a_2, a_3$, not all of which are zero, such that:

$$a_1\vec{u} + a_2\vec{r} + a_3\vec{s} = \vec{0}.$$

But then

$$\vec{u} + 6\vec{r} - \vec{s} + (a_1\vec{u} + a_2\vec{r} + a_3\vec{s})$$

would be another way to write $\vec{a}$ using only the same three vectors.

**Linear dependence and infinite solutions.**

The goal of this problem is to

- Connect linear dependence with infinite solutions.
- Connect linear independence with unique solutions.

**Notes/Misconceptions**

- Part 1 won't give the students much trouble.
- Part 2 will be hard. The idea that you could add $\vec{0}$ but change the coefficients of a linear combination is hard to thin of and even harder to write correctly.

19.3 Is $\{\vec{u}, \vec{r}\}$ linearly independent?

Yes. If it were not, we would necessarily have $\vec{u} = \beta\vec{r}$ for some scalar $\beta$. But then

$$(\beta + 6)\vec{r} - \vec{s}$$

would be another way to write $\vec{a}$ using only the same three vectors.

19.4 Is $\{\vec{u}, \vec{v}, \vec{w}, \vec{r}\}$ linearly independent?

No. We know from earlier that $\vec{u} + \vec{v} + 2\vec{w} = \vec{0}$, and so $\vec{u} + \vec{v} + 2\vec{w} + 0\vec{r} = \vec{0}$ is a non-trivial linear combination of the vectors in this set that equals the zero vector.

## Task 1.4: Linear Independence and Dependence, Creating Examples

1. Fill in the following chart keeping track of the strategies you used to generate examples.

| | Linearly independent | Linearly dependent |
|---|---|---|
| A set of 2 vectors in $\mathbb{R}^2$ | | |
| A set of 3 vectors in $\mathbb{R}^2$ | | |
| A set of 2 vectors in $\mathbb{R}^3$ | | |
| A set of 3 vectors in $\mathbb{R}^3$ | | |
| A set of 4 vectors in $\mathbb{R}^3$ | | |

2. Write at least two generalizations that can be made from these examples and the strategies you used to create them.

**Notes/Misconceptions**

- Make sure "more than $n$ vectors in $\mathbb{R}^n$ is linearly dependent" comes out.

Lesson 10

# Dot Product, Orthogonality

## Textbook

Section 1.3

## Objectives

- Compute the dot product of two vectors.
- Compute the length of a vector.
- Find the distance between two vectors.
- Define what it means for vectors to be orthogonal.
- Interpret the sign of the dot product geometrically.
- Create a unit vector in the direction of another.

## Motivation

Studying $\mathbb{R}^n$ we're in a natural inner product space with lengths and angles. The dot product allows us to get at lengths and angles. It will also give an alternative way to compute matrix products (dot product with rows instead of linear combination of columns).

Most importantly, the dot product tells us how much two vectors point in the same direction as well as when they're orthogonal.

# Dot Product

## Norm

DEFINITION

The ***norm*** of a vector $\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ is the length/magnitude of $\vec{v}$. It is written $\|\vec{v}\|$ and can be computed from the Pythagorean formula

$$\|\vec{v}\| = \sqrt{v_1^2 + \cdots + v_n^2}.$$

## Dot Product

DEFINITION

If $\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ are two vectors in $n$-dimensional space, then the ***dot product*** of $\vec{a}$ an $\vec{b}$ is

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

Equivalently, the dot product is defined by the geometric formula

$$\vec{a} \cdot \vec{b} = \|\vec{a}\|\|\vec{b}\| \cos\theta$$

where $\theta$ is the angle between $\vec{a}$ and $\vec{b}$.

**Notes/Misconceptions**

- It'd be great to say $\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$, but that would make the geometric definition of the dot product circular.

**Notes/Misconceptions**

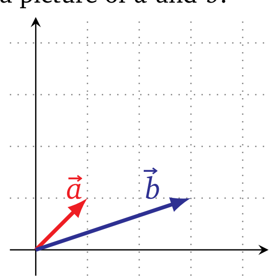
- The two definitions are useful in different contexts.
- Students will gravitate towards the algebraic definition because it has an easier formula. Emphasize that if they don't know both definitions, there will be problems they can't solve.

---

20

Let $\vec{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, and $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$.

20.1 (a) Draw a picture of $\vec{a}$ and $\vec{b}$.




(b) Compute $\vec{a} \cdot \vec{b}$. $\vec{a} \cdot \vec{b} = (1)(3) + (1)(1) = 4$.

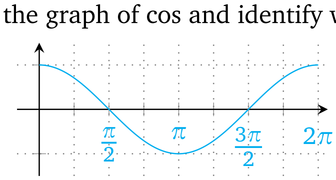
(c) Find $\|\vec{a}\|$ and $\|\vec{b}\|$ and use your knowledge of the multiple ways to compute the dot product to find $\theta$, the angle between $\vec{a}$ and $\vec{b}$. Label $\theta$ on your picture.

$\|\vec{a}\| = \sqrt{(1)^2 + (1)^2} = \sqrt{2}$ and $\|\vec{b}\| = \sqrt{(3)^2 + (1)^2} = \sqrt{10}$.

Using the two definitions of the dot product we have:

$$\begin{aligned} \vec{a} \cdot \vec{b} &= \|\vec{a}\|\|\vec{b}\| \cos\theta \\ \implies 4 &= (\sqrt{2})(\sqrt{10}) \cos\theta \\ \implies \theta &= \arccos\left(\frac{2}{\sqrt{5}}\right) \end{aligned}$$

20.2 Draw the graph of cos and identify which angles make cos negative, zero, or positive.




Cosine is positive for angles in the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, as well as all shifts of this interval by a multiple of $2\pi$ in either direction.

**Practicing dot products.**

The goal of this problem is to

- Use both the algebraic and geometric definitions of the dot product as appropriate to compute dot products.
- Gain an intuition that positive dot product means "pointing in similar directions", negative dot product means "pointing in opposite directions", and zero dot product means "pointing in orthogonal directions".

**Notes/Misconceptions**

- For part 3(d), some students will suggest a vector that "points out of the page". This is a great idea! Except that such a vector would have the wrong number of components, so the dot products wouldn't be defined.
- For part 4, students will try to solve an equation. Encourage them to guess-and-check instead. It will be much faster.

cos is positive for angles in the interval $\left(\frac{\pi}{2}, \frac{3\pi}{2}\right)$, as well as all shifts of this interval by a multiple of $2\pi$ in either direction.

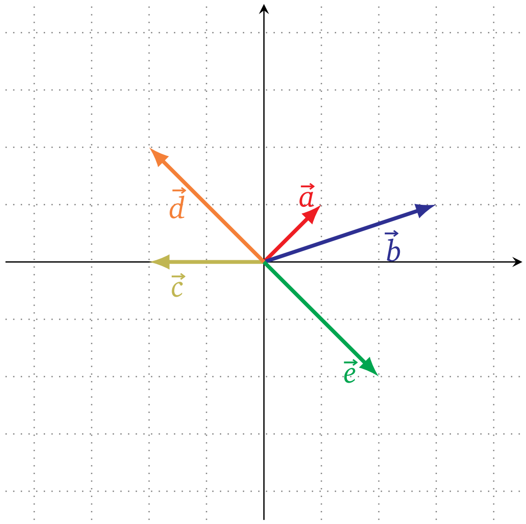
20.3 Draw a new picture of $\vec{a}$ and $\vec{b}$ and on that picture draw

(a) a vector $\vec{c}$ where $\vec{c} \cdot \vec{a}$ is negative.

(b) a vector $\vec{d}$ where $\vec{d} \cdot \vec{a} = 0$ and $\vec{d} \cdot \vec{b} < 0$.

(c) a vector $\vec{e}$ where $\vec{e} \cdot \vec{a} = 0$ and $\vec{e} \cdot \vec{b} > 0$.

(d) Could you find a vector $\vec{f}$ where $\vec{f} \cdot \vec{a} = 0$ and $\vec{f} \cdot \vec{b} = 0$? Explain why or why not.




(d) $\vec{f} = \vec{0}$ is the only possibility. For any vector $\vec{f} = \begin{bmatrix} x \\ y \end{bmatrix}$, we can compute:

$$\vec{f} \cdot \vec{a} = x + y \qquad \text{and} \qquad \vec{f} \cdot \vec{b} = 3x + y.$$

If these both equal zero, the first equation says that $y = -x$, and in turn the second one says $x = 0$ (and so $y = 0$ as well).

20.4 Recall the vector $\vec{u}$ whose coordinates are given at the beginning of this problem.

(a) Write down a vector $\vec{v}$ so that the angle between $\vec{u}$ and $\vec{v}$ is $\pi/2$. (Hint, how does this relate to the dot product?)

$\vec{v} = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}$ is one such vector.

Since $\cos(\pi/2) = 0$, from the second definition of the dot product above we know we are looking for a $\vec{v}$ such that $\vec{u} \cdot \vec{v} = 0$. Using the first definition of the dot product, we can see that the $\vec{v}$ given above is one possibility.

(b) Write down another vector $\vec{w}$ (in a different direction from $\vec{v}$) so that the angle between $\vec{w}$ and $\vec{u}$ is $\pi/2$.

$\vec{w} = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$ is a possible answer.

$\vec{u} \cdot \vec{w} = 0$, and $\vec{w}$ is clearly not parallel to $\vec{v}$ from above.

(c) Can you write down other vectors different than both $\vec{v}$ and $\vec{w}$ that still form an angle of $\pi/2$ with $\vec{u}$? How many such vectors are there?

Yes. $\begin{bmatrix} 0 \\ 2 \\ -4 \end{bmatrix}$ is one possibility.

There are actually infinitely many such vectors; any linear combination of $\vec{w}$ and $\vec{v}$ will work.

To see this, note that any such vector $\vec{x}$ is of the form

$$\vec{x} = t \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} + s \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} t - s \\ t + s \\ -3t - s \end{bmatrix},$$

for scalars $t$ and $s$. We can then compute

$$\vec{u}\cdot\vec{x}=(1)(t-s)+(2)(t+s)+(1)(-3t-s)=0,$$

and so any such vector $\vec{x}$ forms an angle of $\pi/2$ with $\vec{u}$.

THM

For a vector $\vec{v}\in\mathbb{R}^n$, the formula

$$\|\vec{v}\|=\sqrt{\vec{v}\cdot\vec{v}}$$

always holds.

DEF

**Distance**

The ***distance*** between two vectors $\vec{u}$ and $\vec{v}$ is $\|\vec{u}-\vec{v}\|$.

DEF

**Unit Vector**

A vector $\vec{v}$ is called a ***unit vector*** if $\|\vec{v}\|=1$.

21 Let $\vec{u}=\begin{bmatrix}1\\2\\1\end{bmatrix}$ and $\vec{v}=\begin{bmatrix}1\\1\\3\end{bmatrix}$.

21.1 Find the distance between $\vec{u}$ and $\vec{v}$.

$\vec{u}-\vec{v}=\begin{bmatrix}0\\1\\-2\end{bmatrix}$, and so $\|\vec{u}-\vec{v}\|=\sqrt{5}$.

21.2 Find a unit vector in the direction of $\vec{u}$.

$\frac{1}{\sqrt{6}}\vec{u}=\begin{bmatrix}\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\\\frac{1}{\sqrt{6}}\end{bmatrix}$.

$\|\vec{u}\|=\sqrt{6}$, and so if we multiply $\vec{u}$ by $\frac{1}{\sqrt{6}}$, the length of the resulting vector will be 1.

21.3 Does there exist a *unit vector* $\vec{x}$ that is distance 1 from $\vec{u}$?

No. $\|\vec{u}\|=\sqrt{6}$, and so the shortest length that a vector whose distance from $\vec{u}$ is 1 can have is $\sqrt{6}-1$, which is greater than 1.

21.4 Suppose $\vec{y}$ is a unit vector and the distance between $\vec{y}$ and $\vec{u}$ is 2. What is the angle between $\vec{y}$ and $\vec{u}$?

The angle between $\vec{u}$ and $\vec{y}$ is $\arccos\left(\frac{3}{2\sqrt{6}}\right)$.

By assumption, $2=\|\vec{u}-\vec{y}\|$, and so

$$\begin{aligned}4&=\|\vec{u}-\vec{y}\|^2\\&=(\vec{u}-\vec{y})\cdot(\vec{u}-\vec{y})\\&=\vec{u}\cdot\vec{u}-2(\vec{u}\cdot\vec{y})+\vec{y}\cdot\vec{y}\\&=\|\vec{u}\|^2-2\vec{u}\cdot\vec{y}+\|\vec{y}\|^2\\&=6-2\vec{u}\cdot\vec{y}+1.\end{aligned}$$

Then we rearrange to find that $\vec{u}\cdot\vec{y}=\frac{3}{2}$.

Using this in the second definition of the dot product, we see:

$$\frac{3}{2}=\left(\sqrt{6}\right)(1)\cos\theta,$$

where $\theta$ is the angle between $\vec{u}$ and $\vec{y}$.

DEF

**Orthogonal**

Two vectors $\vec{u}$ and $\vec{v}$ are ***orthogonal*** to each other if $\vec{u}\cdot\vec{v}=0$. The word orthogonal is synonymous with the word perpendicular.

**Practice using norms.**

The goal of this problem is to

- Practice finding the distance between two vectors.
- Produce a unit vector pointing in the same direction as another vector.
- Intuitively apply the triangle inequality: $\|\vec{a}+\vec{b}\|\leq\|\vec{a}\|+\|\vec{b}\|$.

**Notes/Misconceptions**

- Part 1 will be easy.
- Many will have trouble with part 2. Encourage them to write definitions! $\vec{a}$ is in the direction of $\vec{u}$ if $\vec{a}=k\vec{u}$ and $\vec{a}$ is a unit vector if $\|\vec{a}\|=1$. Now, solve for $k$.

  They will forget that $\sqrt{k^2}=|k|$ and so there are two solutions unless you insist on $k\geq 0$. You could mention "in the direction of" vs. the more specific "in the positive direction of" if you want to have a detailed discussion.
- Have them draw a picture for part 3.
- Part 4 will be very hard. Only go into it if you have plenty of time.

**Apply the definition of *orthogonal*.**

The goal of this problem is to

- Gain an intuitive understanding of *orthogonal vectors*.
- Produce orthogonal vectors via guess-and-check.
- Apply the Pythagorean theorem to orthogonal vectors to find lengths.

**Notes/Misconceptions**

- It won't occur to most students that $\vec{0}$ is orthogonal to everything. Make sure this come up.
- Guessing-and-checking is a valid and useful mathematical technique. Student aren't comfortable with this method because there's no algorithm for it (and so it doesn't seem reliable and repeatable). We should change their attitude. Most problems in their lives won't be solved with repeatable algorithms (unless they work on an assembly line).

22.1 Find two vectors orthogonal to $\vec{a} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$. Can you find two such vectors that are not parallel?

Two such vectors are $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -6 \\ -2 \end{bmatrix}$.

It is impossible for two non-parallel vectors to both be orthogonal to $\vec{a}$. If $\vec{b} = \begin{bmatrix} x \\ y \end{bmatrix}$ is orthogonal to $\vec{a}$, then we must have that $x - 3y = 0$, or in other words that $x = 3y$. Any $\vec{b}$ satisfying this is a multiple of $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

22.2 Find two vectors orthogonal to $\vec{b} = \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}$. Can you find two such vectors that are not parallel?

Two such vectors are $\begin{bmatrix} 7 \\ 1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$.

These two vectors are not parallel.

22.3 Suppose $\vec{x}$ and $\vec{y}$ are orthogonal to each other and $\|\vec{x}\| = 5$ and $\|\vec{y}\| = 3$. What is the distance between $\vec{x}$ and $\vec{y}$?

The distance between them must be $\sqrt{34}$.

One way to see this is with Pythagoras' theorem. Two perpendicular line segments of lengths 3 and 5 form the two shorter sides of a right angle triangle, and so the length of the third side is $\sqrt{5^2 + 3^2} = \sqrt{34}$.

An equivalent way to see this is to use what we know about dot products to calculate $\|\vec{x} - \vec{y}\|$ as follows:

$$\|\vec{x} - \vec{y}\| = \sqrt{(\vec{x} - \vec{y}) \cdot (\vec{x} - \vec{y})} = \sqrt{\|\vec{x}\|^2 - 2(\vec{x} \cdot \vec{y}) + \|\vec{y}\|^2} = \sqrt{5^2 + 2(0) + 3^2},$$

where in the last step we've used the fact that $\vec{x}$ and $\vec{y}$ are orthogonal, so $\vec{x} \cdot \vec{y} = 0$.

Lesson 11

# Normal Form of Lines and Planes

## Textbook

Section 1.3

## Objectives

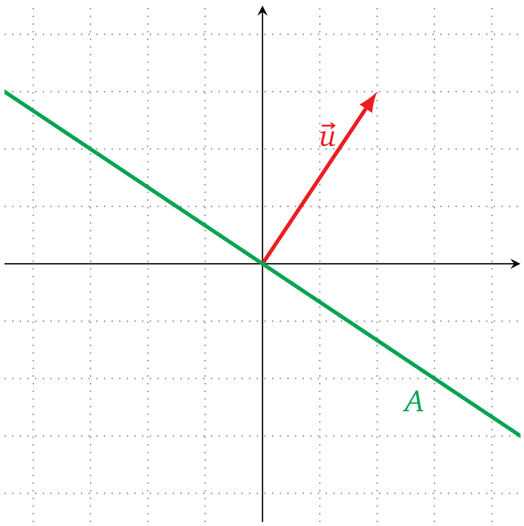
- Describe lines and planes in normal form.

## Motivation

Physics often describes surfaces in terms of normal and tangential components. Normal form of lines and planes is one way to get at this decomposition. Further, thinking about lines and planes in terms of right angles will help when visualizing orthogonal projections.

23

23.1 Draw $\vec{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and *all* vectors orthogonal to it. Call this set $A$.




23.2 If $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $\vec{x}$ is orthogonal to $\vec{u}$, what is $\vec{x} \cdot \vec{u}$? $\vec{x} \cdot \vec{u} = 0$, by the definition of orthogonality.

23.3 Expand the dot product $\vec{u} \cdot \vec{x}$ to get an equation for $A$.

$A$ is the line with vector equation $\vec{x} = t \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

If $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix} \in A$, then $\vec{x} \cdot \vec{u} = 2x + 3y = 0$.

23.4 If possible, express $A$ as a span. $A = \operatorname{span}\left\{\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right\}$.

**Generate lines using orthogonality.**

The goal of this problem is to

- Visually see how the set of all vectors orthogonal to a given vector forms a line.
- Given a line defined as the set of all vectors orthogonal to a given vector, express the line using an equation or span.

**Notes/Misconceptions**

- This problem won't be hard, so don't spend too much time on it.
- For part 1, students might insist on drawing arrowheads and tails on their vectors. This is an opportunity to discuss when you should draw arrowheads/tails and when you shouldn't.

**Normal Vector**

DEF A ***normal vector*** to a line (or plane or hyperplane) is a non-zero vector that is orthogonal to all direction vectors for the line (or plane or hyperplane).

24

Let $\vec{d} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\vec{p} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, and define the lines

$$\ell_1 = \operatorname{span}\{\vec{d}\} \qquad \text{and} \qquad \ell_2 = \operatorname{span}\{\vec{d}\} + \{\vec{p}\}.$$

24.1 Find a vector $\vec{n}$ that is a normal vector for both $\ell_1$ and $\ell_2$.

$\vec{n} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ is one possibility.

This vector is orthogonal to $\vec{d}$, which is a direction vector for both lines.

24.2 Let $\vec{v} \in \ell_1$ and $\vec{u} \in \ell_2$. What is $\vec{n} \cdot \vec{v}$? What about $\vec{n} \cdot (\vec{u} - \vec{p})$? Explain using a picture.

$\vec{n} \cdot \vec{v} = \vec{n} \cdot (\vec{u} - \vec{p}) = 0$.

This is because any $\vec{v} \in \ell_1$ is a multiple of $\vec{d}$, which is orthogonal to $\vec{n}$. Similarly, for any $\vec{u} \in \ell_2$, the vector $\vec{u} - \vec{p}$ is a direction vector for $\ell_2$, and so it is orthogonal to $\vec{n}$.

$\vec{n} \cdot \vec{u} = 3$, since any such $\vec{u}$ is of the form $\vec{u} = \vec{p} + t\vec{d}$ for some scalar $t$, and so

$$\vec{n} \cdot \vec{u} = \vec{n} \cdot (\vec{p} + t\vec{d}) = \vec{n} \cdot \vec{p} + t(\vec{n} \cdot \vec{d}) = 3 + t(0) = 3.$$

24.3 A line is expressed in *normal form* if it is represented by an equation of the form $\vec{n} \cdot (\vec{x} - \vec{q}) = 0$ for some $\vec{n}$ and $\vec{q}$. Express $\ell_1$ and $\ell_2$ in normal form.

**Express lines in normal form.**

The goal of this problem is to

- Express lines, including lines that don't pass through $\vec{0}$, in normal form.
- See the $\vec{q}$ in $\vec{n} \cdot (\vec{x} - \vec{q}) = 0$ is similar to the $\vec{p}$ in the vector form $\vec{x} = t\vec{d} + \vec{p}$ in that it accommodates lines that don't pass through $\vec{0}$.
- Use the dot product to represent a line compactly.

**Notes/Misconceptions**

- In part 2 there may be some lingering confusion about the difference between a vector and a point. In particular, $\vec{u} - \vec{p}$ is a direction vector for $\ell_2$, but is not *contained in* $\ell_2$.
- In part 3, have a discussion about the role that $\vec{q}$ plays. Namely, that it "translates", in a similar way that $\vec{p}$ in the vector form $\vec{x} = t\vec{d} + \vec{p}$ translates.

  However, be careful about the discussion here. Students may find it very confusing that we "subtract $\vec{q}$" to translate in normal form, whereas in vector form we add $\vec{p}$. Have a good explanation prepared if you go down this route.

A normal form of $\ell_1$ is $\begin{bmatrix}2\\-1\end{bmatrix}\cdot\vec{x}=0$.

A normal form of $\ell_2$ is $\begin{bmatrix}2\\-1\end{bmatrix}\cdot\left(\vec{x}-\begin{bmatrix}1\\-1\end{bmatrix}\right)=0$. In the previous part we saw that $\vec{n}\cdot\vec{x}=\vec{n}\cdot\vec{p}$ for all $\vec{x}\in\ell_2$, or in other words $\vec{n}\cdot(\vec{x}-\vec{p})=0$.

24.4 Some textbooks would claim that $\ell_2$ could be expressed in normal form as $\begin{bmatrix}2\\-1\end{bmatrix}\cdot\vec{x}=3$. How does this relate to the $\vec{n}\cdot(\vec{x}-\vec{p})=0$ normal form? Where does the 3 come from?

Let $\vec{n}=\begin{bmatrix}2\\-1\end{bmatrix}$ and let $\vec{x}\in\ell_2$. From the previous part, we know

$$0=\vec{n}\cdot(\vec{x}-\vec{p})=\vec{n}\cdot\vec{x}-\vec{n}\cdot\vec{p}=\vec{n}\cdot\vec{x}-3.$$

Therefore

$$\vec{n}\cdot\vec{x}=3.$$

---

25 Let $\vec{n}=\begin{bmatrix}1\\1\\1\end{bmatrix}$.

25.1 Use set-builder notation to write down the set, $X$, of all vectors orthogonal to $\vec{n}$. Describe this set geometrically.

$X=\left\{\vec{x}\in\mathbb{R}^3\ :\ \vec{x}\cdot\vec{n}=0\right\}$.

Geometrically, this is a plane through the origin and perpendicular to $\vec{n}$.

25.2 Describe $X$ using an equation. $x+y+z=0$.

25.3 Describe $X$ as a span. $X=\operatorname{span}\left\{\begin{bmatrix}1\\-1\\0\end{bmatrix},\begin{bmatrix}1\\0\\-1\end{bmatrix}\right\}$ is one way to do this.

**Planes in normal form.**

The goal of this problem is to

- Observe that the set of all vectors orthogonal to another in $\mathbb{R}^3$ is a plane.
- Translate descriptions of sets into precise mathematical statements using set-builder notation.
- Express a plane in multiple ways.

**Notes/Misconceptions**

- Students have already worked with this plane in problem 12, but they won't remember. Emphasize part 1 and go quickly through the other parts, referring to problem 12 if needed.

Lesson 12

# Projections

## Textbook

Section 1.4

## Objectives

- Project a vector onto lines and finite sets.
- Find the components of one vector in terms of another.

## Motivation

Projection of a vector onto a set, defined as the closet point in the set to the vector, is a general operation used outside of linear algebra. However, in the land of linear algebra, we have exact formulas for the projection. Projections are a chance to explore a seemingly simple definition and see it relate to sets, lines, normal form, and vector form.

$\mathrm{comp}_{\vec{v}}\,\vec{u}$ is the component of a vector in the direction of another, which is sometimes called the projection of $\vec{u}$ onto $\vec{v}$. It relates to how much one vector points in the direction of another and provides a decomposition of vectors in terms of orthogonal components.

**Notes/Misconceptions**

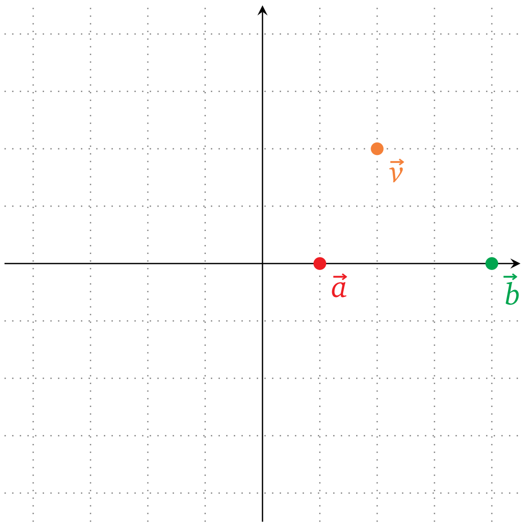
In this class, we don't write $\mathrm{proj}_{\vec{v}}\,\vec{u}$, i.e., the projection of one vector onto another. We instead call this $\mathrm{comp}_{\vec{v}}\,\vec{u}$. We do this so as not to confuse $\mathrm{proj}_{\{\vec{v}\}}\,\vec{u}$ and $\mathrm{comp}_{\vec{v}}\,\vec{u}$. One is projection onto a singleton. The other is $\frac{\vec{u}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v}$.

**Projection**

DEF Let $X$ be a set. The ***projection*** of the vector $\vec{v}$ onto $X$, written $\text{proj}_X \vec{v}$, is the closest point in $X$ to $\vec{v}$.

26 Let $\vec{a} = \begin{bmatrix}1\\0\end{bmatrix}$, $\vec{b} = \begin{bmatrix}4\\0\end{bmatrix}$, $\vec{v} = \begin{bmatrix}2\\2\end{bmatrix}$ and $\ell = \text{span}\{\vec{a}\}$.

26.1 Draw $\vec{a}$, $\vec{b}$, and $\vec{v}$ in the same picture.




26.2 Find $\text{proj}_{\{\vec{b}\}} \vec{v}$, $\text{proj}_{\{\vec{a},\vec{b}\}} \vec{v}$.

$\text{proj}_{\{\vec{b}\}} \vec{v} = \vec{b}$. Since there is only one point in $\{\vec{b}\}$, it must be the closest point to $\vec{v}$.

$\text{proj}_{\{\vec{a},\vec{b}\}} \vec{v} = \vec{a}$. We can simply compute $\|\vec{v}-\vec{a}\| = \sqrt{5}$ and $\|\vec{v}-\vec{b}\| = \sqrt{8}$, so $\vec{a}$ is closer to $\vec{v}$.

26.3 Find $\text{proj}_\ell \vec{v}$. (Recall that a quadratic $at^2 + bt + c$ has a minimum at $t = -\frac{b}{2a}$).

$\text{proj}_\ell \vec{v} = 2\vec{a} = \begin{bmatrix}2\\0\end{bmatrix}$.

Any point in $\ell$ is of the form $t\vec{a}$ for some scalar $t$. The distance between such a point and $\vec{v}$ is

$$\|\vec{v} - t\vec{a}\| = \sqrt{\|v\|^2 - 2t(\vec{v}\cdot\vec{a}) + t^2\|a\|^2} = \sqrt{8 - 4t + t^2}$$

The quadratic inside the square root has a minimum at $t = 2$, so $2\vec{a}$ is the closest point in the line to $\vec{v}$.

26.4 Is $\vec{v} - \text{proj}_\ell \vec{v}$ a normal vector for $\ell$? Why or why not?

Yes.

By the previous part, $\vec{v} - \text{proj}_\ell \vec{v} = \vec{v} - 2\vec{a} = \begin{bmatrix}0\\2\end{bmatrix}$. This vector is orthogonal to $\vec{a}$, and therefore to $\ell$.

27 Let $K$ be the line given in vector form by $\vec{x} = t\begin{bmatrix}1\\2\end{bmatrix} + \begin{bmatrix}1\\0\end{bmatrix}$ and let $\vec{c} = \begin{bmatrix}1\\3\end{bmatrix}$.

27.1 Make a sketch with $\vec{c}$, $K$, and $\text{proj}_K \vec{c}$ (you don't need to compute $\text{proj}_K \vec{c}$ exactly).

**Apply the definition of projection.**

The goal of this problem is to

- Use the definition of projection to compute projections onto finite sets and lines.
- Pick an appropriate representation of a line to solve a projection problem.

**Notes/Misconceptions**

- Projections will be new and strange. Some students will be familiar with the notation $\text{proj}_{\vec{b}} \vec{a}$ and will confuse this with $\text{proj}_{\{\vec{b}\}} \vec{a}$. Make them apply the definition of projection.
- When visualizing projections, it is often better to draw vectors as dots instead of line segments. Students will have a cluttered diagram if they try to draw vectors with "tails".
- General projections have no formula. This might make students unhappy, but it's an important point to emphasize.
- The goal of part 3 is to have students pick a from the many representations of $\ell$ a suitable one. Vector avoids multivariable calculus.
- Some students might already know about a relationship between orthogonality and a closest point. If you don't want to optimize a quadratic, you may ask students to do part 4 first with a picture and then do part 3 by inspection.

**Project onto lines.**

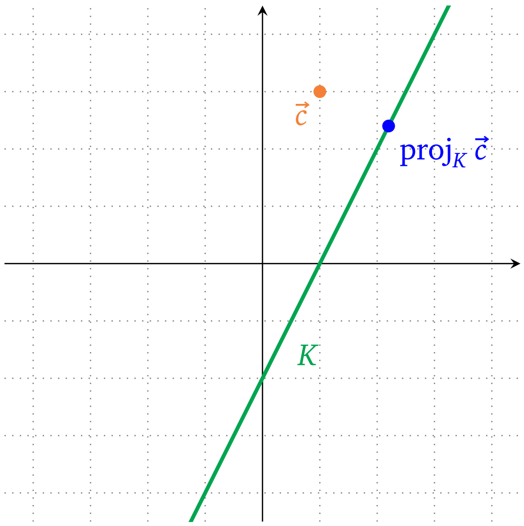
The goal of this problem is to

- Use orthogonality to compute the projection onto a line.
- Project onto lines that don't pass through $\vec{0}$.

**Notes/Misconceptions**

- This is not a calculus class, so we want to avoid doing projections via calculus. This problem forces the use of orthogonality to solve a projection question.
- When drawing pictures for part 2, there will be a question of where you should root your vectors. Rooting your vectors on $K$ makes the picture clearer.




27.2 What should $(\vec{c} - \text{proj}_K \vec{c}) \cdot \begin{bmatrix}1\\2\end{bmatrix}$ be? Explain.

$(\vec{c} - \text{proj}_K \vec{c}) \cdot \begin{bmatrix}1\\2\end{bmatrix} = 0.$

From our picture we can see that $c - \text{proj}_K \vec{c}$ is perpendicular to the line $K$, and so the dot product of this vector with any direction vector for $K$ should be zero.

27.3 Use your formula from the previous part to find $\text{proj}_K \vec{c}$ *without* computing any distances.

$\text{proj}_K \vec{c} = \frac{1}{5}\begin{bmatrix}11\\12\end{bmatrix}$

If $\text{proj}_K \vec{c} = \begin{bmatrix}x\\y\end{bmatrix}$, the formula from the previous part tells us

$$\left(\begin{bmatrix}1\\3\end{bmatrix} - \begin{bmatrix}x\\y\end{bmatrix}\right) \cdot \begin{bmatrix}1\\2\end{bmatrix} = 1 - x + 6 - 2y = 0 \qquad \iff \qquad x + 2y = 7$$

So we need a point on $K$ that satisfies this equation. In other words, we need

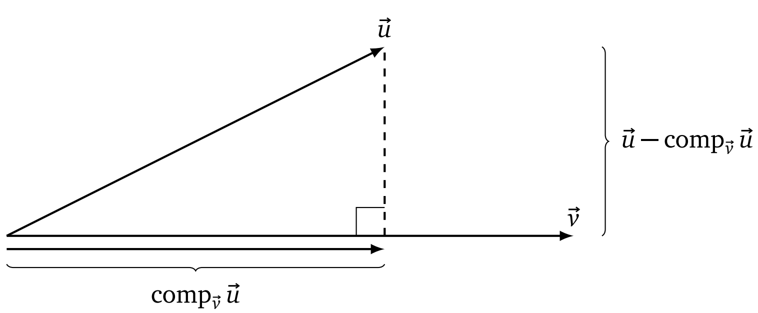
$$(t+1) + 2(2t) = 7 \qquad \implies \qquad t = \frac{6}{5}.$$

The point on $K$ for this value of $t$ is $\frac{1}{5}\begin{bmatrix}11\\12\end{bmatrix}$.

**Component**

DEFINITION

Let $\vec{u}$ and $\vec{v} \neq \vec{0}$ be vectors. The ***component of $\vec{u}$ in the $\vec{v}$ direction***, written $\text{comp}_{\vec{v}} \vec{u}$, is the vector in the direction of $\vec{v}$ so that $\vec{u} - \text{comp}_{\vec{v}} \vec{u}$ is orthogonal to $\vec{v}$.




**Notes/Misconceptions**

- This is called $\text{proj}_{\vec{v}} \vec{u}$ in some courses. This operation is renamed *component* ... to avoid possible notational confusion with $\text{proj}_{\{\vec{v}\}} \vec{u}$.

**Component of a vector in the direction of another.**

The goal of this problem is to

- Read and apply a new definition.
- Use orthogonality to obtain a formula for components in terms of dot products.

**Notes/Misconceptions**

- For part 1, ask students to codify their conditions with formulas. This will be *hard*. Students will not know how to read the definition, They will get one property easily, but a second property will escape them.
- Part 2 is an exercise in applying the two formulas from part 1.

28 Let $\vec{a}, \vec{b} \in \mathbb{R}^3$ be unknown vectors.

28.1 List two conditions that $\text{comp}_{\vec{b}} \vec{a}$ must satisfy.

$\text{comp}_{\vec{b}} \vec{a}$ must be a scalar multiple of $\vec{b}$.

$\vec{a} - \text{comp}_{\vec{b}} \vec{a}$ must be orthogonal to $\vec{b}$, or in other words $(\vec{a} - \text{comp}_{\vec{b}} \vec{a}) \cdot \vec{b} = 0$.

28.2 Find a formula for $\text{comp}_{\vec{b}} \vec{a}$.

$\text{comp}_{\vec{b}}\,\vec{a} = \dfrac{\vec{a}\cdot\vec{b}}{\vec{b}\cdot\vec{b}}\vec{b}.$

From the previous part, we should have $\text{comp}_{\vec{b}}\,\vec{a} = t\vec{b}$ for some scalar $t$, and $(\vec{a} - \text{comp}_{\vec{b}}\,\vec{a})\cdot\vec{b} = 0$.

Combining these, we get:

$$0 = (\vec{a} - t\vec{b})\cdot\vec{b} = \vec{a}\cdot\vec{b} - t\vec{b}\cdot\vec{b} = \vec{a}\cdot\vec{b} - t(\vec{b}\cdot\vec{b}).$$

Solving for $t$, we get $t = \dfrac{\vec{a}\cdot\vec{b}}{\vec{b}\cdot\vec{b}}$.

Lesson 13

# Projections, Subspaces

## Textbook

Sections 1.2, 1.4

## Objectives

- Identify $\mathrm{proj}_{\mathrm{span}\{\vec v\}}\vec u$ with $\mathrm{comp}_{\vec v}\vec u$.
- Identify $\mathrm{comp}_{\vec v}\vec u$ and $\mathrm{comp}_{\alpha\vec v}\vec u$ for all $\alpha \neq 0$, including negative $\alpha$.
- Define subspace.
- Distinguish subspaces and non-subspaces of $\mathbb{R}^2$.

## Motivation

Spans are a constructive way to describe lines, planes, and other flat objects. Subspaces are a categorical way of defining flat objects. Instead of explaining how to find the vectors in a set, we list their properties. This is a really powerful idea that facilitates abstraction.

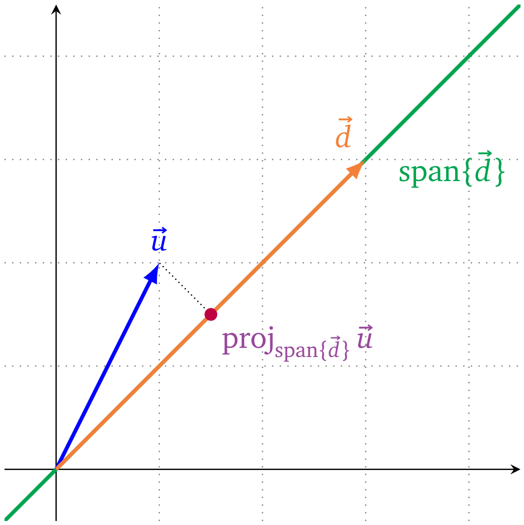
Since we do not do abstract vector spaces in this course, subspaces are the first place (unless you count projections) students will encounter a set defined by its properties. Subspaces are suitable for a first-encounter because 1) the properties are simple and familiar and 2) subspaces of $\mathbb{R}^n$ have a concrete geometric interpretation.

**Notes/Misconceptions**

- Philosophically, a subspace should be defined as a non-empty set closed under linear combinations. However, defining it as closed under addition and scalar multiplication gives students new to proofs something explicit to hang on to when attempting a proof.
- Some people define a subspace as a set containing $\vec 0$ and satisfying closure. We define a subspace as a non-empty set satisfying closure. We won't be trying to trick students by asking if an empty set is a subspace, so don't belabor the point.
- Proving a set is a subspace is one of the few proofs that we will hold students accountable for. That is, it is one of the few things we expect students to be able to write down completely correctly using formal mathematical language. They will need practice to be able to do this!

**29** Let $\vec{d} = \begin{bmatrix}3\\3\end{bmatrix}$ and $\vec{u} = \begin{bmatrix}1\\2\end{bmatrix}$.

29.1 Draw $\vec{d}$, $\vec{u}$, $\operatorname{span}\{\vec{d}\}$, and $\operatorname{proj}_{\operatorname{span}\{\vec{d}\}} \vec{u}$ in the same picture.




29.2 How do $\operatorname{proj}_{\operatorname{span}\{\vec{d}\}} \vec{u}$ and $\operatorname{comp}_{\vec{d}} \vec{u}$ relate? They are equal.

29.3 Compute $\operatorname{proj}_{\operatorname{span}\{\vec{d}\}} \vec{u}$ and $\operatorname{comp}_{\vec{d}} \vec{u}$.

Using our formula from the previous problem

$$\operatorname{proj}_{\operatorname{span}\{\vec{d}\}} \vec{u} = \operatorname{comp}_{\vec{d}} \vec{u} = \frac{\vec{u}\cdot\vec{d}}{\|\vec{d}\|^2}\vec{d} = \tfrac{9}{18}\vec{d} = \tfrac{1}{2}\begin{bmatrix}3\\3\end{bmatrix}.$$

29.4 Compute $\operatorname{comp}_{-\vec{d}} \vec{u}$. Is this the same as or different from $\operatorname{comp}_{\vec{d}} \vec{u}$? Explain.

$$\operatorname{comp}_{-\vec{d}} \vec{u} = \frac{\vec{u}\cdot(-\vec{d})}{\|-\vec{d}\|^2}(-\vec{d}) = \tfrac{-9}{18}(-\vec{d}) = \tfrac{1}{2}\begin{bmatrix}3\\3\end{bmatrix} = \operatorname{comp}_{\vec{d}} \vec{u}.$$

We expect them to be equal since $\vec{d}$ and $-\vec{d}$ are in the same direction as one another.

**Relate components and projections.**

The goal of this problem is to

- Find a connection between components and projections onto spans.
- Recognize that $\operatorname{comp}_{\vec{u}} \vec{v} = \operatorname{comp}_{-\vec{u}} \vec{v}$.

**Notes/Misconceptions**

- Part 4 will be counterintuitive. Students may think $\operatorname{comp}_{-\vec{d}} \vec{u} = -\operatorname{comp}_{\vec{d}} \vec{u}$. Referring back to part 2 and noticing $\operatorname{span}\{\vec{d}\} = \operatorname{span}\{-\vec{d}\}$ should be enlightening.

## Subspaces and Bases

**DEFINITION — Subspace**

A ***subspace*** $V \subseteq \mathbb{R}^n$ is a non-empty subset such that

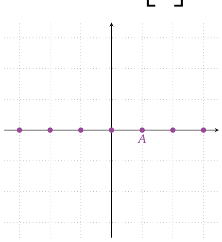
(i) $\vec{u}, \vec{v} \in V$ implies $\vec{u} + \vec{v} \in V$.

(ii) $\vec{u} \in V$ implies $k\vec{u} \in V$ for all scalars $k$.

Subspaces give a mathematically precise definition of a “flat space through the origin.”

**30** For each set, draw it and explain whether or not it is a subspace of $\mathbb{R}^2$.

30.1 $A = \left\{\vec{x} \in \mathbb{R}^2 : \vec{x} = \begin{bmatrix}a\\0\end{bmatrix} \text{ for some } a \in \mathbb{Z}\right\}$.




$A$ is not a subspace, since for example $\begin{bmatrix}1\\0\end{bmatrix} \in A$ but $\frac{1}{2}\begin{bmatrix}1\\0\end{bmatrix} \notin A$.

30.2 $B = \left\{\vec{x} \in \mathbb{R}^2 : \vec{x} \neq \begin{bmatrix}0\\0\end{bmatrix}\right\}$.

**Visualizing subspaces.**

The goal of this problem is to

- Read and apply the definition of subspace.
- Identify from a picture whether or not a set is a subspace.
- Write formal arguments showing whether or not certain sets are subspaces.

**Notes/Misconceptions**

Every part of this problem highlights a misconception or builds an abstraction.

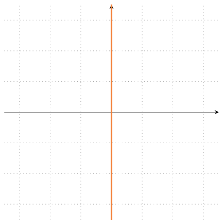
1. Satisfies (i) but not (ii).
2. Looks fairly flat, but violates (i) and (ii) in a specify way.
3. First subspace.
4. Looks flat, but violates (i) and (ii).
5. Satisfies (ii) but not (i).
6. Second subspace.
7. Alternate notation which requires unpacking a definition.
8. Abstractly defined subspace requires an abstract proof.

- Proving a set is a subset is one of the few *proofs* we hold students accountable for, and they need practice. Remind them that to prove something is not a subspace they only need to prove that one condition is violated, and this can be done with an example. However, to prove something is a subspace, we need to make an argument about every single vector.
- Subspace proofs can follow a template. “Let $\vec{u}, \vec{v} \in X$ and let $\alpha$ be a scalar. By definition ..., and so ...”. Let students try to write a proof on their own. Then, give them the template. Applying the template is harder than it seems, so emphasize for every subspace proof how it fits the template.

$B$ is not a subspace, since for example $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}-1\\0\end{bmatrix}$ are both in $B$, but their sum is $\begin{bmatrix}0\\0\end{bmatrix}$ which is not in $B$.

30.3 $C = \left\{\vec{x} \in \mathbb{R}^2 : \vec{x} = \begin{bmatrix}0\\t\end{bmatrix} \text{ for some } t \in \mathbb{R}\right\}$.

$C$ is a subspace.

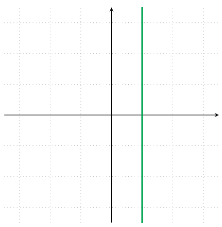


(i) Let $\vec{u}, \vec{v} \in C$. Then $\vec{u} = \begin{bmatrix}0\\t\end{bmatrix}$ and $\vec{v} = \begin{bmatrix}0\\s\end{bmatrix}$ for some $s, t \in \mathbb{R}$.

But then $\vec{u} + \vec{v} = \begin{bmatrix}0\\s+t\end{bmatrix} \in C$.

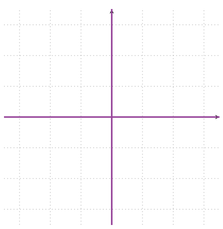
(ii) Let $\vec{u} = \begin{bmatrix}0\\t\end{bmatrix} \in C$. For any scalar $\alpha$ we have $\alpha\vec{u} = \begin{bmatrix}0\\\alpha t\end{bmatrix} \in C$.

30.4 $D = \left\{\vec{x} \in \mathbb{R}^2 : \vec{x} = \begin{bmatrix}0\\t\end{bmatrix} + \begin{bmatrix}1\\1\end{bmatrix} \text{ for some } t \in \mathbb{R}\right\}$.



$D$ is not a subspace, since for example $\begin{bmatrix}1\\1\end{bmatrix} \in D$, but $0\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix} \notin D$.

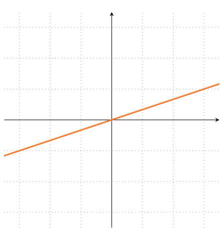
30.5 $E = \left\{\vec{x} \in \mathbb{R}^2 : \vec{x} = \begin{bmatrix}0\\t\end{bmatrix} \text{ or } \vec{x} = \begin{bmatrix}t\\0\end{bmatrix} \text{ for some } t \in \mathbb{R}\right\}$.



$E$ is not a subspace, since for example $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$ are both in $E$, but their sum is $\begin{bmatrix}1\\1\end{bmatrix}$ which is not in $E$.

30.6 $F = \left\{\vec{x} \in \mathbb{R}^2 : \vec{x} = t\begin{bmatrix}3\\1\end{bmatrix} \text{ for some } t \in \mathbb{R}\right\}$.

$F$ is a subspace.



(i) Let $\vec{u}, \vec{v} \in F$. Then $\vec{u} = t\begin{bmatrix}3\\1\end{bmatrix}$ and $\vec{v} = s\begin{bmatrix}3\\1\end{bmatrix}$ for some $s, t \in \mathbb{R}$.

But then $\vec{u} + \vec{v} = (s+t)\begin{bmatrix}3\\1\end{bmatrix} \in F$.

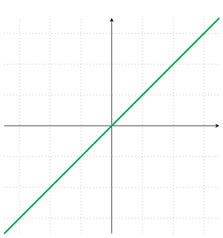
(ii) Let $\vec{u} = t\begin{bmatrix}3\\1\end{bmatrix} \in F$. For any scalar $\alpha$ we have $\alpha\vec{u} = (\alpha t)\begin{bmatrix}3\\1\end{bmatrix} \in F$.

30.7 $G = \operatorname{span}\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$.

$G$ is a subspace.

By definition of a span, $G = \left\{\vec{x} \in \mathbb{R}^2 : \vec{x} = t\begin{bmatrix}1\\1\end{bmatrix} \text{ for some } t \in \mathbb{R}\right\}$.

The proof that $G$ is a subspace now proceeds similarly to the proof for $F$ above.



(i) Let $\vec{u}, \vec{v} \in G$. Then $\vec{u} = t\begin{bmatrix}1\\1\end{bmatrix}$ and $\vec{v} = s\begin{bmatrix}1\\1\end{bmatrix}$ for some $s, t \in \mathbb{R}$.

But then $\vec{u} + \vec{v} = (s+t)\begin{bmatrix}1\\1\end{bmatrix} \in G$.

(ii) Let $\vec{u} = t\begin{bmatrix}1\\1\end{bmatrix} \in G$. For any scalar $\alpha$ we have $\alpha\vec{u} = (\alpha t)\begin{bmatrix}1\\1\end{bmatrix} \in G$.

30.8 $H = \operatorname{span}\{\vec{u}, \vec{v}\}$ for some unknown vectors $\vec{u}, \vec{v} \in \mathbb{R}^2$.

$H$ is a subspace.

(i) Let $\vec{x}, \vec{y} \in H$. Then $\vec{x} = \alpha_1\vec{u} + \alpha_2\vec{v}$ and $\vec{y} = \beta_1\vec{u} + \beta_2\vec{v}$ for some scalars $\alpha_1, \alpha_2, \beta_1, \beta_2$. But then

$$\vec{x} + \vec{y} = \alpha_1\vec{u} + \alpha_2\vec{v} + \beta_1\vec{u} + \beta_2\vec{v} = (\alpha_1 + \beta_1)\vec{u} + (\alpha_2 + \beta_2)\vec{v} \in H.$$

(ii) Let $\vec{x} = \alpha_1\vec{u} + \alpha_2\vec{v} \in H$. For any scalar $\beta$ we have $\beta\vec{u} = (\beta\alpha_1)\vec{u} + (\beta\alpha_2)\vec{v} \in H$.

Lesson 14

# Basis, Dimension

## Textbook

Sections 1.2, 4.3

## Objectives

- Define Basis.
- Define Dimension.
- Find a basis for a subspace.
- Find the dimension of a subspace.

## Motivation

Bases are sets of just enough vectors to describe every vector in a subspace. An additional consequence of a basis is that every vector can be *uniquely* represented as a linear combination of basis vectors. Using this fact we will be able to consider objects in multiple different coordinate systems. However, now is the time to get familiar with what a basis is and how to find one.

Dimension ties the abstract notion of subspace to our intuition about Euclidean space. We already know a plane in $\mathbb{R}^3$ is two dimensional, but now we know where that number *two* comes from.

**Basis**

DEF A ***basis*** for a subspace $V$ is a linearly independent set of vectors, $\mathcal{B}$, so that $\operatorname{span}\mathcal{B} = V$.

**Dimension**

DEF The ***dimension*** of a subspace $V$ is the number of elements in a basis for $V$.

---

31 Let $\vec{u} = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $\vec{v} = \begin{bmatrix}0\\1\\0\end{bmatrix}$, $\vec{w} = \begin{bmatrix}1\\1\\0\end{bmatrix}$, and $V = \operatorname{span}\{\vec{u}, \vec{v}, \vec{w}\}$.

31.1 Describe $V$. $V$ is the $xy$-plane in $\mathbb{R}^3$.

31.2 Is $\{\vec{u}, \vec{v}, \vec{w}\}$ a basis for $V$? Why or why not?

No. The set $\{\vec{u}, \vec{v}, \vec{w}\}$ is linearly dependent since $\vec{w} = \vec{u} + \vec{v}$.

31.3 Give a basis for $V$. $\{\vec{u}, \vec{v}\}$.

31.4 Give another basis for $V$. $\{\vec{u}, \vec{w}\}$ or $\{\vec{v}, \vec{w}\}$.

31.5 Is $\operatorname{span}\{\vec{u}, \vec{v}\}$ a basis for $V$? Why or why not?

No. $\operatorname{span}\{\vec{u}, \vec{v}\}$ is an infinite set of vectors which includes $\vec{0}$, so it cannot be linearly independent and therefore isn't a basis.

31.6 What is the dimension of $V$?

A basis for $V$ has two vectors so it is two-dimensional. We also know this because $V$ is the $xy$-plane in $\mathbb{R}^3$ and all planes are two-dimensional.

**Apply the definitions of basis and dimension to a simple example.**

The goal of this problem is to learn

- To apply the definition of basis and dimension.
- Intuition that a plane is two dimensional.
- A basis is not unique, but always has the same size (this is not proved).
- Spans are never bases—you must not confuse a subspace with its basis!

**Notes/Misconceptions**

- Students will claim $V$ is $\mathbb{R}^2$ and fail to distinguish $\mathbb{R}^2$ and the $xy$-plane in $\mathbb{R}^3$.
- Parts 2, 3, 4, 6 will be easy; don't belabor them.
- Students will fail to distinguish $\operatorname{span}\{\vec{u}, \vec{v}\}$ from $\{\vec{u}, \vec{v}\}$. Make sure this distinction comes out.

---

32 Let $\vec{a} = \begin{bmatrix}1\\2\\3\end{bmatrix}$, $\vec{b} = \begin{bmatrix}4\\5\\6\end{bmatrix}$, $\vec{c} = \begin{bmatrix}7\\8\\8\end{bmatrix}$ (notice these vectors are linearly independent) and let $P = \operatorname{span}\{\vec{a}, \vec{b}\}$ and $Q = \operatorname{span}\{\vec{b}, \vec{c}\}$.

32.1 Give a basis for and the dimension of $P$.

$\{\vec{a}, \vec{b}\}$ is a basis for $P$, and so its dimension is 2.

32.2 Give a basis for and the dimension of $Q$.

$\{\vec{b}, \vec{c}\}$ is a basis for $Q$, and so its dimension is 2.

32.3 Is $P \cap Q$ a subspace? If so, give a basis for it and its dimension.

Yes. $\{\vec{b}\}$ is a basis for $P \cap Q$, and so its dimension is 1.

$P$ and $Q$ are both planes and are not parallel (since $\vec{a}, \vec{b}, \vec{c}$ are linearly independent). The intersection of any two non-parallel planes in $\mathbb{R}^3$ is a line. We know that $\vec{0}$ and $\vec{b}$ are on this line, and therefore the line is $\operatorname{span}\{\vec{b}\}$

32.4 Is $P \cup Q$ a subspace? If so, give a basis for it and its dimension.

No. For example $\vec{a}$ and $\vec{c}$ are both in $P \cup Q$, but $\vec{a} + \vec{c} \notin P \cup Q$.

Proof: A vector is in $P \cup Q$ if it is in $P$ or $Q$, so we must show that $\vec{a} + \vec{c} \notin P$ and $\vec{a} + \vec{c} \notin Q$

$\vec{a} + \vec{c} \notin P$ since if it were, we would also have $(\vec{a} + \vec{c}) - \vec{a} = \vec{c} \in P$. We know this is impossible since the vectors $\vec{a}, \vec{b}, \vec{c}$ are linearly independent, and so $\vec{c}$ does not equal a linear combination of $\vec{a}$ and $\vec{b}$.

An analogous argument shows that $\vec{a} + \vec{c} \notin Q$.

**The relationship between subspaces, bases, unions, and intersections.**

The goal of this problem is to learn

- Recognize intersections of subspaces as subspaces.
- Recognize the union of subspaces need not be a subspace.
- Visualize planes in $\mathbb{R}^3$ to solve problems without computations.

**Notes/Misconceptions**

- Remind students that if they're stuck, they can always create a system of equations to help answer their question.
- For part 3, students could row-reduce if they're stuck, but the nicer way is to notice $\vec{b}$ is common to both spans and so, since the planes are not parallel, their intersection must be the line $\operatorname{span}\{\vec{b}\}$.
- Part 4 is hard to prove, but you can make a hand-wavy argument without too much trouble. Don't worry about getting a rock-solid proof for this part.

Lesson 15

# Matrices

## Textbook

Section 3.1

## Objectives

- Write a system of linear equations as a matrix equation.
- Write a matrix equation as a system of linear equations.
- Pose familiar problems (e.g., “find a normal vector”, or “do these planes intersect?”) as matrix-equation questions.

## Motivation

Matrices will soon become a powerful tool to study linear transformations. However, we will start out viewing them as a notation to represent systems of linear equations. The fact that matrix-vector multiplication has two interpretations, as a linear combination of columns or as a dot product with rows, already connects geometry and angles with questions about linear combinations.

33 Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 3 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$.

33.1 Compute the product $A\vec{x}$.

$$A\vec{x} = \begin{bmatrix} x + 2y \\ 3x + 3y \end{bmatrix}.$$

33.2 Write down a system of equations that corresponds to the matrix equation $A\vec{x} = \vec{b}$.

$$\begin{aligned} x + 2y &= -2 \\ 3x + 3y &= -1 \end{aligned}$$

33.3 Let $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ be a solution to $A\vec{x} = \vec{b}$. Explain what $x_0$ and $y_0$ mean in terms of *linear combinations* (hint: think about the columns of $A$).

$x_0$ and $y_0$, when used as scalars in a linear combination of the columns of $A$, make the vector $\vec{b}$. In other words:

$$x_0 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + y_0 \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}.$$

33.4 Let $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ be a solution to $A\vec{x} = \vec{b}$. Explain what $x_0$ and $y_0$ mean in terms of *intersecting lines* (hint: think about systems of equations).

The lines represented by the equations $x + 2y = -2$ and $3x + 3y = -1$ from the system of equations above intersect at the point $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$.

**Relate matrix equations and systems of linear equations.**

The goal of this problem is to
- Use matrix-vector multiplication to represent a system of equations with compact notation.
- View a matrix equation as a statement about (i) linear combinations of column vectors and (ii) a system of equations coming from the rows.

**Notes/Misconceptions**
- Students know two different interpretations of matrix multiplication from the homework.

34 Let $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, $\vec{w} = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$.

34.1 How could you determine if $\{\vec{u}, \vec{v}, \vec{w}\}$ was a linearly independent set?

The set is linearly independent if and only if no non-trivial linear combination of the vectors $\vec{u}, \vec{v}, \vec{w}$ equals $\vec{0}$. That is, if $x, y, z$ are scalars such that $x\vec{u} + y\vec{v} + z\vec{w} = \vec{0}$, then $x = y = z = 0$.

In other words, the only solution of the following system of equations is $x = y = z = 0$.

$$\begin{aligned} x + 4y + 7z &= 0 \\ 2x + 5y + 8z &= 0 \\ 3x + 6y + 9z &= 0 \end{aligned}$$

34.2 Can your method be rephrased in terms of a matrix equation? Explain.

The system of linear equations above can be represented by the matrix equation $A\vec{x} = \vec{0}$, where $A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$ and $\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

So another way to say the above is that the set is linearly independent if and only if the only solution to the equation $A\vec{x} = \vec{0}$ is $\vec{x} = \vec{0}$.

**Rephrase previous questions using matrix equations.**

The goal of this problem is to
- Rephrase the question of linear independence as the special matrix equation $A\vec{x} = \vec{0}$.

**Notes/Misconceptions**
- We haven't formally introduced the term *homogeneous system* yet. Now is a good time. The equation $A\vec{x} = \vec{0}$ will be used again when talking about null spaces.

35 Consider the system represented by

$$\begin{bmatrix} 1 & -3 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \vec{b}.$$

**Interpret matrix equations.**

The goal of this problem is to
- Use knowledge about systems of linear equations to answer questions about matrix equations.

**Notes/Misconceptions**
- This is mainly an exercise in applying existing knowledge about interpreting RREF to matrix equations.

35.1 If $\vec{b} = \begin{bmatrix}1\\2\\3\end{bmatrix}$, is the set of solutions to this system a point, line, plane, or other?

This system has no solutions, since, if we expand the matrix equation into a system of equations, the third equation would be $0 = 3$, which is impossible.

35.2 If $\vec{b} = \begin{bmatrix}1\\1\\0\end{bmatrix}$, is the set of solutions to this system a point, line, plane, or other?

A line. The system would be

$$\begin{aligned} x - 3y &= 1 \\ z &= 1 \\ 0 &= 0 \end{aligned}$$

A vector $\vec{x} = \begin{bmatrix}x\\y\\z\end{bmatrix}$ that satisfies this system must have $z = 1$, and by the first equation in the system any value of $x$ determines the value of $y$, and vice versa. In other words the system has one free variable, and so its set of solutions is a line.

---

36 Let $\vec{d}_1 = \begin{bmatrix}1\\1\\2\end{bmatrix}$ and $\vec{d}_2 = \begin{bmatrix}-1\\1\\0\end{bmatrix}$. Let $\mathcal{P}$ be the plane given in vector form by $\vec{x} = t\vec{d}_1 + s\vec{d}_2$. Further, suppose $M$ is a matrix so that $M\vec{r} \in \mathcal{P}$ for any $\vec{r} \in \mathbb{R}^2$.

36.1 How many rows does $M$ have?

Three. It must have three rows in order for $M\vec{r}$ to be an element of $\mathbb{R}^3$.

36.2 Find such an $M$.

$M = \begin{bmatrix}1 & -1\\1 & 1\\2 & 0\end{bmatrix}$ is one possible answer, since if $\vec{r} = \begin{bmatrix}a\\b\end{bmatrix}$, then $M\vec{r} = a\vec{d}_1 + b\vec{d}_2$.

Another less interesting answer is the $3 \times 2$ zero matrix.

36.3 Find necessary and sufficient conditions (phrased as equations) for $\vec{n}$ to be a normal vector for $\mathcal{P}$.

$\vec{n}$ is normal to $\mathcal{P}$ if and only if $\vec{n} \neq \vec{0}$, $\vec{n} \cdot \vec{d}_1 = 0$, *and* $\vec{n} \cdot \vec{d}_2 = 0$

36.4 Find a matrix $K$ so that non-zero solutions to $K\vec{x} = \vec{0}$ are normal vectors for $\mathcal{P}$. How do $K$ and $M$ relate?

$K = \begin{bmatrix}1 & 1 & 2\\-1 & 1 & 0\end{bmatrix}$. $K$ and $M$ are transposes of one another.

The conditions $\vec{n} \cdot \vec{d}_1 = 0$ and $\vec{n} \cdot \vec{d}_2 = 0$ from the previous part translate to the following system of equations:

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

This system of equations can be represented by the matrix equation

$$\begin{bmatrix}1 & 1 & 2\\-1 & 1 & 0\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}.$$

**Apply matrix equations to planes.**

The goal of this problem is to
- Rephrase properties of a plane in terms of matrix equations.
- Be able to describe one application of the transpose.

**Notes/Misconceptions**
- We are using the range of a matrix operator to describe a plane and the null space of a matrix to define normal vectors.
- Part 2 admits some trivial answers, but the interesting one is when the range is $\mathcal{P}$.
- For part 3, students might miss the $\vec{n} \neq \vec{0}$ condition.
- The matrices $K$ and $M$ do not have to be transposes of each other, but arrange it so they are.

Lesson 16

# Change of Basis I

## Textbook

Section 4.4

## Objectives

- Write a vector in multiple bases.
- Explain what the notation $[\vec{v}]_{\mathcal{B}}$ means.
- Explain what the notation $\begin{bmatrix} a \\ b \\ c \end{bmatrix}_{\mathcal{B}}$ means.

## Motivation

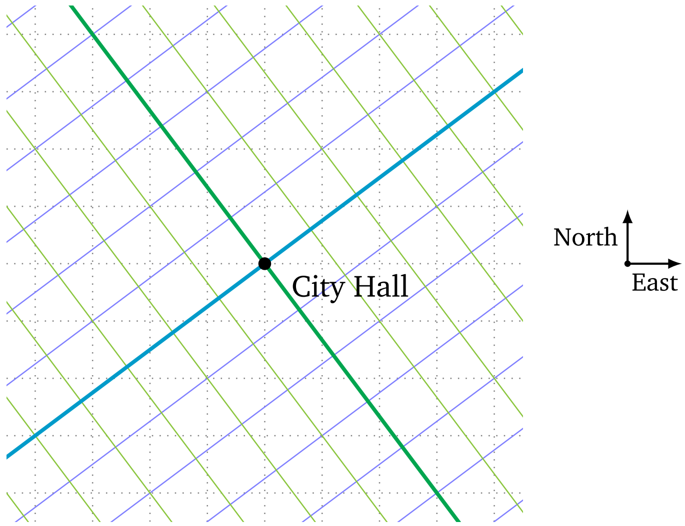
One of the most useful ideas in linear algebra is that you can represent a vector, a geometric object, with a list of numbers. This is done by picking a basis. So far we've implicitly used the standard basis, but now we're going to use other bases.

Now lists of numbers can mean many different things and can be identified with vectors in many ways, so we need some notation to keep things straight. It's important now to distinguish when something is a list of numbers (a matrix) and when it is a vector. This distinction will arise again when we talk about linear transformations and their matrix representations.

**Notes/Misconceptions**

So far, we have written $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ to mean $\begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{E}}$ where $\mathcal{E}$ is the standard basis. We will continue to do this as convenient, but if multiple bases are ever involved, we will be careful to specify the basis.

37 The fictional town of Oronto is not aligned with the usual compass directions. The streets are laid out as follows:




Instead, every street is parallel to the vector $\vec{d}_1 = \frac{1}{5}\begin{bmatrix} 4 \text{ east} \\ 3 \text{ north} \end{bmatrix}$ or $\vec{d}_2 = \frac{1}{5}\begin{bmatrix} -3 \text{ east} \\ 4 \text{ north} \end{bmatrix}$. The center of town is City Hall at $\vec{0} = \begin{bmatrix} 0 \text{ east} \\ 0 \text{ north} \end{bmatrix}$.

Locations in Oronto are typically specified in *street coordinates*. That is, as a pair $(a, b)$ where $a$ is how far you walk along streets in the $\vec{d}_1$ direction and $b$ is how far you walk in the $\vec{d}_2$ direction, provided you start at city hall.

37.1 The points $A = (2, 1)$ and $B = (3, -1)$ are given in street coordinates. Find their east-north coordinates.

$A = (1, 2)$ and $B = (3, 1)$ in east-north coordinates.

We obtain $A$ for example by finding the vector $2\vec{d}_1 + \vec{d}_2$.

37.2 The points $X = (4, 3)$ and $Y = (1, 7)$ are given in east-north coordinates. Find their street coordinates. $X = (5, 0)$ and $Y = (5, 5)$ in street coordinates.

37.3 Define $\vec{e}_1 = \begin{bmatrix} 1 \text{ east} \\ 0 \text{ north} \end{bmatrix}$ and $\vec{e}_2 = \begin{bmatrix} 0 \text{ east} \\ 1 \text{ north} \end{bmatrix}$. Does $\operatorname{span}\{\vec{e}_1, \vec{e}_2\} = \operatorname{span}\{\vec{d}_1, \vec{d}_2\}$?

Yes. Both of these sets spans all of $\mathbb{R}^2$.

37.4 Notice that $Y = 5\vec{d}_1 + 5\vec{d}_2 = \vec{e}_1 + 7\vec{e}_2$. Is the point $Y$ better represented by the pair $(5, 5)$ or by the pair $(1, 7)$? Explain.

It is equally well represented by either pair. For example, the street coordinates might be more useful for a resident of Oronto, while the east-north coordinates might be more useful for someone looking at Oronto on a world map.

**Motivate change of basis.**

The goal of this problem is to

- Describe points in multiple bases when given a visual description of the basis or when given the basis vectors numerically.
- Recognize ambiguity when faced with the question, "Which basis is better?"

**Notes/Misconceptions**

- The Oronto streets are set up like 3,4,5 right triangle and all the numbers are easy. This question can be answered from the picture or algebraically.
- Part 4 is important. A convincing argument is that one basis is not intrinsically "better" than the other is that if the compass rose were not shown, you couldn't tell from the picture which way was north (since the streets are orthogonal to each other just like north and east are)!

  Emphasize that the "best" representation of a point will depend on what question you're trying to answer.

## Representation in a Basis

DEFINITION

Let $\mathcal{B}=\{\vec b_1,\dots,\vec b_n\}$ be a basis for a subspace $V$ and let $\vec v\in V$. The ***representation of $\vec v$ in the $\mathcal{B}$ basis***, notated $[\vec v]_{\mathcal{B}}$, is the column matrix

$$[\vec v]_{\mathcal{B}}=\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix}.$$

where $\alpha_1,\dots,\alpha_n$ uniquely satisfy $\vec v=\alpha_1\vec b_1+\cdots+\alpha_n\vec b_n$.

Conversely,

$$\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix}_{\mathcal{B}}=\alpha_1\vec b_1+\cdots+\alpha_n\vec b_n$$

is notation for the linear combination of $\vec b_1,\dots,\vec b_n$ with coefficients $\alpha_1,\dots,\alpha_n$.

38 Let $\mathcal{E}=\{\vec e_1,\vec e_2\}$ be the standard basis for $\mathbb{R}^2$ and let $\mathcal{C}=\{\vec c_1,\vec c_2\}$ where $\vec c_1=\begin{bmatrix}2\\1\end{bmatrix}_{\mathcal{E}}$ and $\vec c_2=\begin{bmatrix}5\\3\end{bmatrix}_{\mathcal{E}}$ be another basis for $\mathbb{R}^2$.

38.1 Express $\vec c_1$ and $\vec c_2$ as a linear combination of $\vec e_1$ and $\vec e_2$. $\vec c_1=2\vec e_1+\vec e_2$ and $\vec c_2=5\vec e_1+3\vec e_2$.

38.2 Express $\vec e_1$ and $\vec e_2$ as a linear combination of $\vec c_1$ and $\vec c_2$. $\vec e_1=3\vec c_1-\vec c_2$ and $\vec e_2=-5\vec c_1+2\vec c_2$.

38.3 Let $\vec v=2\vec e_1+2\vec e_2$. Find $[\vec v]_{\mathcal{E}}$ and $[\vec v]_{\mathcal{C}}$.

$[\vec v]_{\mathcal{E}}=\begin{bmatrix}2\\2\end{bmatrix}$ and $[\vec v]_{\mathcal{C}}=\begin{bmatrix}-4\\2\end{bmatrix}$.

The second one is since

$$\vec v=2\vec e_1+2\vec e_2=2(3\vec c_1-\vec c_2)+2(-5\vec c_1+2\vec c_2)=-4\vec c_1+2\vec c_2.$$

38.4 Can you find a matrix $X$ so that $X[\vec w]_{\mathcal{C}}=[\vec w]_{\mathcal{E}}$ for any $\vec w$?

$X=\begin{bmatrix}2&5\\1&3\end{bmatrix}$ is such a matrix.

We know $X$ must be a $2\times 2$ matrix, so suppose $X=\begin{bmatrix}a&b\\c&b\end{bmatrix}$ for some $a,b,c,d\in\mathbb{R}$.

From the first part above, we know

$$\begin{bmatrix}1\\0\end{bmatrix}_{\mathcal{C}}=\begin{bmatrix}2\\1\end{bmatrix}_{\mathcal{E}}\quad\text{and}\quad\begin{bmatrix}0\\1\end{bmatrix}_{\mathcal{C}}=\begin{bmatrix}5\\3\end{bmatrix}_{\mathcal{E}},$$

and so we need $X$ to satisfy

$$X\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}2\\1\end{bmatrix}\quad\text{and}\quad X\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}5\\3\end{bmatrix}.$$

But $X\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}a\\c\end{bmatrix}$ and $X\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}b\\d\end{bmatrix}$, so we can now immediately solve for $a,b,c,d$ to find that $X$ must be the matrix $\begin{bmatrix}2&5\\1&3\end{bmatrix}$.

38.5 Can you find a matrix $Y$ so that $Y[\vec w]_{\mathcal{E}}=[\vec w]_{\mathcal{C}}$ for any $\vec w$?

$Y=\begin{bmatrix}3&-5\\-1&2\end{bmatrix}$ is such a matrix.

Using similar reasoning to the previous part, we know $Y$ must be a $2\times 2$ matrix, so suppose $Y=\begin{bmatrix}a&b\\c&b\end{bmatrix}$ for some $a,b,c,d\in\mathbb{R}$.

From the second part above, we know

$$\begin{bmatrix}1\\0\end{bmatrix}_{\mathcal{E}}=\begin{bmatrix}3\\-1\end{bmatrix}_{\mathcal{C}}\quad\text{and}\quad\begin{bmatrix}0\\1\end{bmatrix}_{\mathcal{E}}=\begin{bmatrix}-5\\2\end{bmatrix}_{\mathcal{C}},$$

**Change of basis notation.**

The goal of this problem is to

- Practice using change-of-basis notation.
- Compute representations of vectors in different bases.
- Find a matrix that computes a change of basis.

**Notes/Misconceptions**

- Up to this point, we have been sloppy, writing $\begin{bmatrix}1\\2\end{bmatrix}$ when we mean $\begin{bmatrix}1\\2\end{bmatrix}_{\mathcal{C}}$. For this problem, be precise. Unscripted matrices are boxes of numbers with no meaning other than what we give them. Subscripted matrices are actual vectors.

  We will be sloppy again in the future—if we interpret $\begin{bmatrix}1\\2\end{bmatrix}$ as a vector, we will always assume it is $\begin{bmatrix}1\\2\end{bmatrix}_{\mathcal{E}}$ unless specified otherwise. But, at the start of change-of-basis, we will be careful.
- For part 5, we do not know about inverse matrices yet. This question should be answered from first principles.
- It will be hard for students to get their head around this notation. $\begin{bmatrix}1\\2\end{bmatrix}$ and $[\vec v]_{\mathcal{X}}$ are boxes of numbers while $\vec v$ and $\begin{bmatrix}1\\2\end{bmatrix}_{\mathcal{X}}$ are vectors.

and so we need $Y$ to satisfy

$$Y\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}3\\-1\end{bmatrix} \quad\text{and}\quad Y\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}-5\\2\end{bmatrix},$$

But $Y\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}a\\c\end{bmatrix}$ and $Y\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}b\\d\end{bmatrix}$, so we can now immediately solve for $a, b, c, d$ to find that $Y$ must be the matrix $\begin{bmatrix}3&-5\\-1&2\end{bmatrix}$.

38.6 What is $YX$?

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

Lesson 17

# Orientation, Matrix Transformations

## Textbook

Section 3.2, 3.3

## Objectives

- Identify the orientation of ordered bases in $\mathbb{R}^2$.
- Given a set of input and output vectors for a linear transformation $L : \mathbb{R}^2 \to \mathbb{R}^2$, find a matrix for the transformation.
- Given a picture $X$ and its image under a linear transformation, find a matrix for the transformation.

## Motivation

Orientation is a topic that comes up in physics and explains the sign of the determinant. We define orientation with an existential statement about whether or not certain homeomorphisms exist. The goal is not to prove anything rigorously about orientation, but to get students to make pictures for themselves of vectors moving.

The idea is simple: $n-1$ vectors span a space that partitions $\mathbb{R}^n$ in two. Add a vector in the top partition (appropriately ordered) and you get a positive orientation; add to the bottom and you get a negative orientation. There's no way to get from one to the other without passing through the hyperplane. We focus on $\mathbb{R}^2$ so that the pictures are easy to draw. Eventually we will compute orientation from the determinant, but it's nice to have a grounding in where it comes from.

While we're thinking dynamically, we can start thinking about transformation. We already know how to multiply a matrix and a vector and interpret it in two different ways. Now we will add a third: multiplication by a given matrix is a transformation from vectors to vectors.

Most of our study of matrix transformations will be of transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$, even though non-square matrices can describe other transformations. For now we stick with pictures of $\mathbb{R}^2$ since they are easy to draw. Then we will generalize to linear transformations.

**Orientation of a Basis**

DEFINITION

The ordered basis $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ is ***right-handed*** or ***positively oriented*** if it can be continuously transformed to the standard basis (with $\vec{b}_i \mapsto \vec{e}_i$) while remaining linearly independent throughout the transformation. Otherwise, $\mathcal{B}$ is called ***left-handed*** or ***negatively oriented***.

39 Let $\{\vec{e}_1, \vec{e}_2\}$ be the standard basis for $\mathbb{R}^2$ and let $\vec{u}_\theta$ be a unit vector. Let $\theta$ be the angle between $\vec{u}_\theta$ and $\vec{e}_1$ measured counter-clockwise starting at $\vec{e}_1$.

39.1 For which $\theta$ is $\{\vec{e}_1, \vec{u}_\theta\}$ a linearly independent set? Every $\theta$ that is not a multiple of $\pi$.

39.2 For which $\theta$ can $\{\vec{e}_1, \vec{u}_\theta\}$ be continuously transformed into $\{\vec{e}_1, \vec{e}_2\}$ and remain linearly independent the whole time?

Every $\theta \in (0, \pi)$.

For $\theta \in (\pi, 2\pi)$, a continuous transformation of $\vec{u}_\theta$ to $\vec{e}_2$ would have to cross the $x$-axis, at which point $\{\vec{e}_1, \vec{u}_\theta\}$ would cease to be linearly independent.

39.3 For which $\theta$ is $\{\vec{e}_1, \vec{u}_\theta\}$ right-handed? Left-handed?

It is right-handed for $\theta \in (0, \pi)$, and left handed for $\theta \in (\pi, 2\pi)$.

39.4 For which $\theta$ is $\{\vec{u}_\theta, \vec{e}_1\}$ (in that order) right-handed? Left-handed?

It is right-handed for $\theta \in (\pi, 2\pi)$, and left handed for $\theta \in (0, \pi)$.

39.5 Is $\{2\vec{e}_1, 3\vec{e}_2\}$ right-handed or left-handed? What about $\{2\vec{e}_1, -3\vec{e}_2\}$?

$\{2\vec{e}_1, 3\vec{e}_2\}$ is right-handed and $\{2\vec{e}_1, -3\vec{e}_2\}$ is left-handed.

**Visually understand orientation.**

The goal of this problem is to

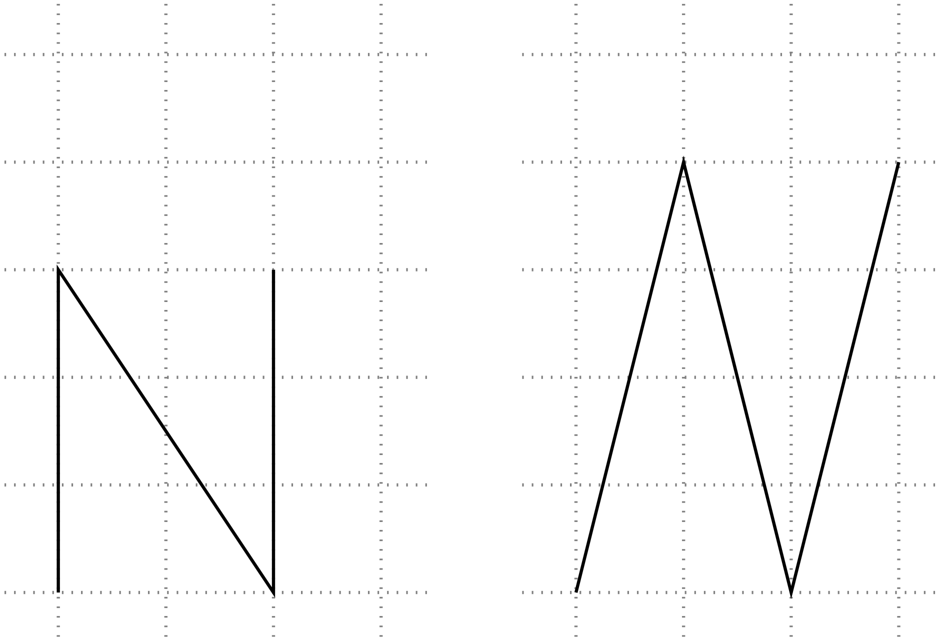
- Determine the orientation of a basis from a picture.
- Recognize the order of vectors in a basis relates to the orientation of that basis.

**Notes/Misconceptions**

- Orientation in $\mathbb{R}^2$ is graphically easy, but conceptually very abstract. The definition references "continuously transformed". Do not try to make this precise. It is sufficient to have a visual representation.
- For part 2, have them draw some examples of $\{\vec{e}_1, \vec{u}_\theta\}$ and $\{\vec{e}_1, \vec{e}_2\}$ on the same set of axes and a dotted line showing how one can transform into the other.
- Most students will picture transformations that preserve the length of $\vec{u}_\theta$, but that need not be the case.

  The (hidden) quantifier order in the definition may be a stumbling block for students. Some might claim that since there exists a transformation that takes $\vec{u}_\theta$ through $\vec{0}$ before its destination, $\{\vec{e}_1, \vec{u}_\theta\}$ must always be left-handed. Don't spend time on this unless students bring it up.

# Task 2.1: Italicizing N



Suppose that the "N" on the left is written in regular 12-point font. Find a matrix $A$ that will transform the "N" into the letter on the right which is written in an *italic* 16-point font.

Work with your group to write out your solution and approach. Make a list of any assumptions you notice your group making or any questions for further pursuit.

**Notes/Misconceptions**

- Students might need some hints. Good starting hints are: "Discuss what size the inputs and outputs to your matrix are and then what size the matrix must be." and "If you're stuck, try making a matrix of variables and then figuring out what the variables are."
- Many groups will miss the "12"-point and "16"-point designations, assuming that each grid line indicates one unit.
- Groups will pick different vectors: some will pick vectors rooted in the lower-left and others will pick vectors that go "along" the $N$. Have a discussion about how these relate via linear combinations and why you get the same matrix either way.
- Most groups will only deal with the "corners" of the $N$ and ignore the fact that the $N$ is made of up line segments. Using the language of convex linear combinations and linearity of matrix multiplication, we can come up with a proof that generalizes the corners to the whole $N$.

Lesson 18

# Linear Transformations I

## Textbook

Section 3.2

## Objectives

- Use the fact that linear transformations take lines to lines and preserve ratios to classify transformations as linear or not.
- Use the formal definition of a linear transformation to prove whether or not a transformation is linear.

## Motivation

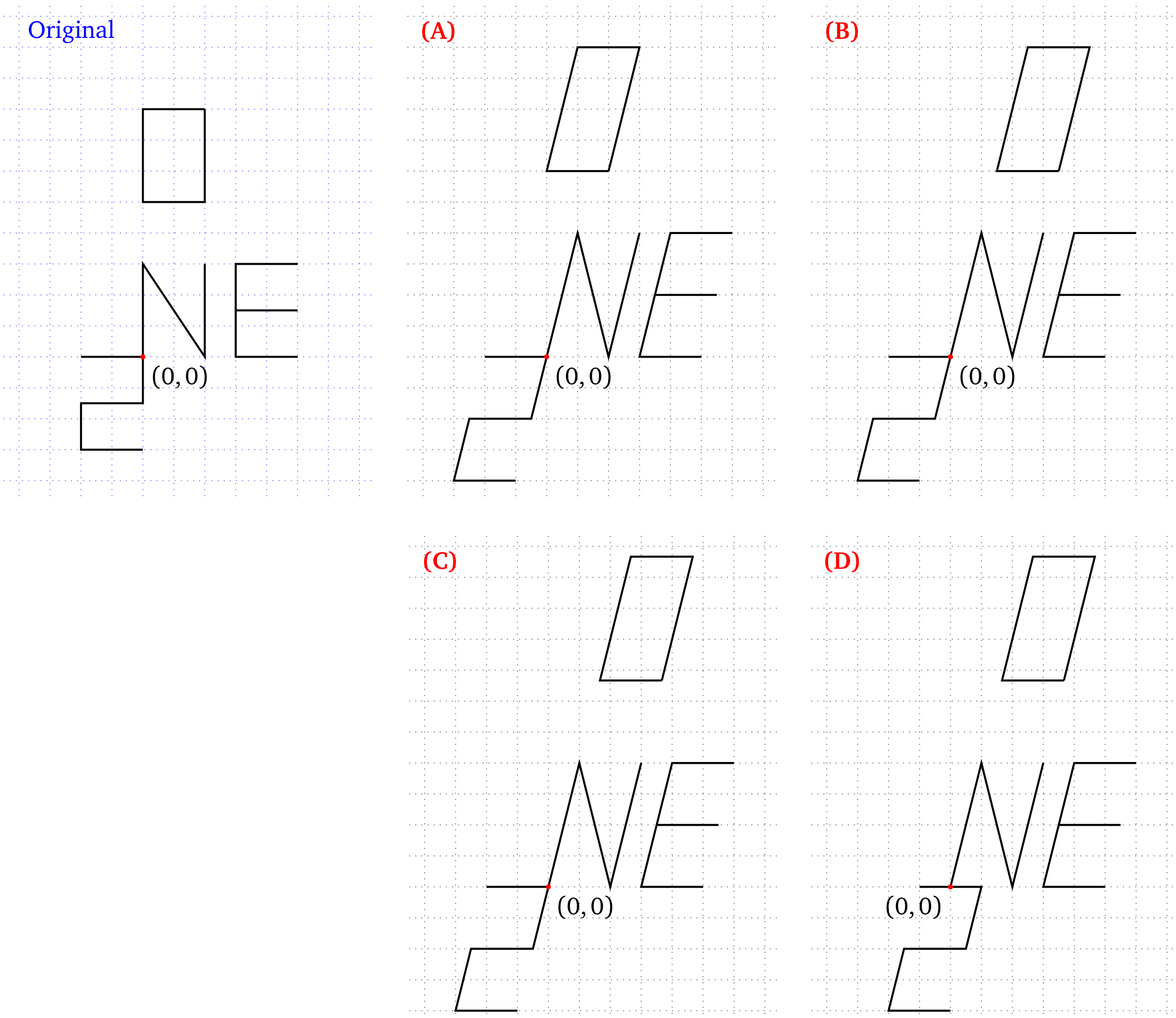
Linear transformations are one of the fundamental objects of study in linear algebra. It's time to see what they look like and what they can do.

We want to develop both a visual and algebraic intuition for linear transformations. Visually, a linear transformation takes straight lines to straight lines and preserves "ratios" between vectors (making this statement precise is hard). Algebraically, linear transformations are exactly the functions that factor in and out of linear combinations, making them easy to work with.

Though few things in our world are linear, calculus shows us how to approximate non-linear functions with linear ones. As a consequence, linear transformations show up all over the place, and so it's worth spending time to understand them.

# Task 2.2: Beyond the N

A few students were wondering how letters placed in other locations in the plane would be transformed under $A = \begin{bmatrix} 1 & 1/3 \\ 0 & 4/3 \end{bmatrix}$. If other letters are placed around the "N," the students argued over four different possible results for the transformed letters. Which choice below, if any, is correct, and why? If none of the four options are correct, what would the correct option be, and why?




**Notes/Misconceptions**

- Every picture focuses on a different property of linear transformations. (D) will be the easiest to argue against. Be warned, the numbers don't work out so nicely, and the correct answer is not aligned to the grid.

40 $\mathcal{R}:\mathbb{R}^2\to\mathbb{R}^2$ is the transformation that rotates vectors counter-clockwise by 90°.

40.1 Compute $\mathcal{R}\begin{bmatrix}1\\0\end{bmatrix}$ and $\mathcal{R}\begin{bmatrix}0\\1\end{bmatrix}$. $\mathcal{R}\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}0\\1\end{bmatrix}$ and $\mathcal{R}\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}-1\\0\end{bmatrix}$.

40.2 Compute $\mathcal{R}\begin{bmatrix}1\\1\end{bmatrix}$. How does this relate to $\mathcal{R}\begin{bmatrix}1\\0\end{bmatrix}$ and $\mathcal{R}\begin{bmatrix}0\\1\end{bmatrix}$?

$$\mathcal{R}\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}-1\\1\end{bmatrix}=\mathcal{R}\begin{bmatrix}1\\0\end{bmatrix}+\mathcal{R}\begin{bmatrix}0\\1\end{bmatrix}.$$

40.3 What is $\mathcal{R}\left(a\begin{bmatrix}1\\0\end{bmatrix}+b\begin{bmatrix}0\\1\end{bmatrix}\right)$?

$$\mathcal{R}\left(a\begin{bmatrix}1\\0\end{bmatrix}+b\begin{bmatrix}0\\1\end{bmatrix}\right)=a\begin{bmatrix}0\\1\end{bmatrix}+b\begin{bmatrix}-1\\0\end{bmatrix}.$$

Rotating a vector and then multiplying by a scalar gives the same result as multiplying first then rotating. Similarly, adding two vectors and then rotating their sum gives the same result as rotating them and then adding.

40.4 Write down a matrix $R$ so that $R\vec{v}$ is $\vec{v}$ rotated counter-clockwise by 90°.

$R=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$ is such a matrix.

**Linear Transformation**

DEFINITION

Let $V$ and $W$ be subspaces. A function $T:V\to W$ is called a ***linear transformation*** if

$$T(\vec{u}+\vec{v})=T\vec{u}+T\vec{v} \qquad \text{and} \qquad T(\alpha\vec{v})=\alpha T\vec{v}$$

for all vectors $\vec{u},\vec{v}\in V$ and all scalars $\alpha$.

41 41.1 Classify the following as linear transformations or not.

(a) $\mathcal{R}$ from before (rotation counter-clockwise by 90°).

A linear transformation. We proved this in the previous problem.

(b) $W:\mathbb{R}^2\to\mathbb{R}^2$ where $W\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}x^2\\y\end{bmatrix}$.

Not a linear transformation, since for example $W\left(2\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}4\\0\end{bmatrix}\neq\begin{bmatrix}2\\0\end{bmatrix}=2W\begin{bmatrix}1\\0\end{bmatrix}$.

(c) $T:\mathbb{R}^2\to\mathbb{R}^2$ where $T\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}x+2\\y\end{bmatrix}$.

Not a linear transformation, since for example $T\left(2\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}4\\0\end{bmatrix}\neq 2T\begin{bmatrix}1\\0\end{bmatrix}$.

(d) $\mathcal{P}:\mathbb{R}^2\to\mathbb{R}^2$ where $\mathcal{P}\begin{bmatrix}x\\y\end{bmatrix}=\operatorname{comp}_{\vec{u}}\begin{bmatrix}x\\y\end{bmatrix}$ and $\vec{u}=\begin{bmatrix}2\\3\end{bmatrix}$.

A linear transformation.
We found a general formula for $\operatorname{comp}_{\vec{u}}$ in a previous exercise:

$$\operatorname{comp}_{\vec{u}}\vec{x}=\frac{\vec{u}\cdot\vec{x}}{\vec{u}\cdot\vec{u}}\vec{u}=\frac{\vec{u}\cdot\vec{x}}{13}\vec{u}.$$

For any two vectors $\vec{x}$ and $\vec{y}$, we have

$$\begin{aligned}\operatorname{comp}_{\vec{u}}(\vec{x}+\vec{y})&=\frac{\vec{u}\cdot(\vec{x}+\vec{y})}{13}\vec{u}\\&=\frac{\vec{u}\cdot\vec{x}}{13}\vec{u}+\frac{\vec{u}\cdot\vec{y}}{13}\vec{u}\\&=\operatorname{comp}_{\vec{u}}\vec{x}+\operatorname{comp}_{\vec{u}}\vec{y}.\end{aligned}$$

For any $\vec{x}$ and scalar $\alpha$, we have

$$\operatorname{comp}_{\vec{u}}(\alpha\vec{x})=\frac{\vec{u}\cdot(\alpha\vec{x})}{13}\vec{u}=\frac{\alpha(\vec{u}\cdot\vec{x})}{13}\vec{u}=\alpha\operatorname{comp}_{\vec{u}}\vec{x}.$$

**Apply geometric transformations to vectors.**

The goal of this problem is to
- Given a transformation described in words, compute the result of the transformation applied to particular vectors.
- Use linear combinations to compute the result of rotations applied to unknown vectors.
- Distinguish between a general transformation and a matrix transformation.

**Notes/Misconceptions**

- Expect at least one student to say $\mathcal{R}\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}1&0\end{bmatrix}$.
- In part 2, emphasize the relationship between linear combinations in the input and linear combinations in the output.
- I part 4, emphasize that $\mathcal{R}$ is *not* the matrix $R$. That multiplication by $R$ is just one way to write $\mathcal{R}$. (For analogy, the squaring function is *not* "$x^2$", nor is it "$x\cdot x$". Those are both sequences of symbols that describe how to square a number $x$, but the squaring function itself is an abstract *function*.)

**Apply the definition of a linear transformation to examples.**

The goal of this problem is to
- Distinguish between a linear transformation and a non-linear transformation.
- Provide a proof of whether a transformation is linear or not.

**Notes/Misconceptions**

- When talking about transformations, it is common to drop the parentheses around the function argument. E.g., $T(\vec{x})\equiv T\vec{x}$. Point this out to students.
- Part (a) is hard to write down a proof for. Don't spend so much time on it.
- Spend time on the remaining parts having students write proofs. Proofs of whether something is linear follow a template. Motivate the template as: *Start with the definitions, then write what you want. Next, write what you know.* By that point, the problem is almost finished.

Lesson 19

# Linear Transformations II, Composition of Linear Transformations

## Textbook

Section 3.2, 3.3

## Objectives

- Draw the image of a set under a (not necessarily linear) transformation.
- Recognize composition of functions as not commutative.
- Decompose complicated transformations in terms of simpler ones.

## Motivation

Linear transformations are a big topic, and we're continuing our study.

We start by introducing the technical term *image* so that we can describe a transformation changing a bunch of points at once. This term will be used again when we start looking at determinants, which measure how transformations change volumes.

Linear transformations are nice because the composition of two linear transformations is another linear transformation. That means, given a complicated transformation, we can search for a sequence of simple transformations that will have the same effect. This is where many of the matrix decompositions (which we aren't covering in this course) come from.

Geometrically, it's easier to reason about simple transformations in a sequence than a complicated transformation. Algebraically, we will be relying on composition of linear functions and its relationship to matrices when we decompose a matrix into a product of elementary matrices (which in turn will allow us to think about the determinant algebraically).

DEFINITION

**Image of a Set**

Let $L : V \to W$ be a transformation and let $X \subseteq V$ be a set. The ***image of the set $X$ under $L$***, denoted $L(X)$, is the set

$$L(X) = \{\vec{x} \in W \,:\, \vec{x} = L(\vec{y}) \text{ for some } \vec{y} \in X\}.$$

42 Let $S = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : 0 \le x, y \le 1 \right\} \subseteq \mathbb{R}^2$ be the filled-in unit square and let $C = \{\vec{0}, \vec{e}_1, \vec{e}_2, \vec{e}_1 + \vec{e}_2\} \subseteq \mathbb{R}^2$ be the corners of the unit square.

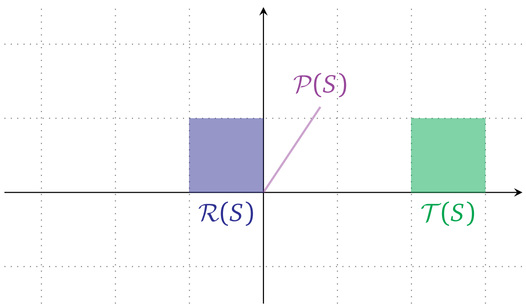
42.1 Find $\mathcal{R}(C)$, $W(C)$, and $T(C)$ (where $\mathcal{R}$, $W$, and $T$ are from the previous question).

$\mathcal{R}(C) = \{\vec{0}, \vec{e}_2, -\vec{e}_1, -\vec{e}_1 + \vec{e}_2\}$.

$W(C) = C$.

$T(C) = \left\{ \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\}$.

42.2 Draw $\mathcal{R}(S)$, $T(S)$, and $\mathcal{P}(S)$ (where $\mathcal{R}$, $T$, and $\mathcal{P}$ are from the previous question).




42.3 Let $\ell = \{\text{all convex combinations of } \vec{a} \text{ and } \vec{b}\}$ be a line segment with endpoints $\vec{a}$ and $\vec{b}$ and let $A$ be a linear transformation. Must $A(\ell)$ be a line segment? What are its endpoints?

$A(\ell)$ must be a line segment, with endpoints $A(\vec{a})$ and $A(\vec{b})$.

For any scalars $\alpha_1$ and $\alpha_2$, by the linearity of $A$ we have: $A(\alpha_1\vec{a} + \alpha_2\vec{b}) = \alpha_1 A(\vec{a}) + \alpha_2 A(\vec{b})$.

If $\alpha_1 + \alpha_2 = 1$, then the linear combination on the right is also convex, and so $A(\ell)$ is the set of convex combinations of $A(\vec{a})$ and $A(\vec{b})$. This is precisely the straight line segment joining $A(\vec{a})$ and $A(\vec{b})$.

Note that if $A(\vec{a}) = A(\vec{b})$ (for example, if $A$ is the zero transformation), then $A(\ell)$ will consist of the single point, which we think of as a "degenerate" line segment in this situation.

42.4 Explain how images of sets relate to the *Italicizing N* task.

The task asked us to find a linear transformation such that the image of the regular "N" is the italicized "N".

By the previous exercise, we now know it suffices to find a linear transformation that sends the four endpoints of line segments on the regular "N" to the corresponding four endpoints on the italicized "N".

**Work with *Images*.**

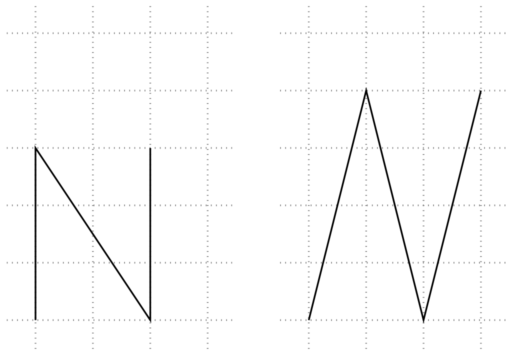
The goal of this problem is to

- Compute images of sets under transformations.
- Develop geometric intuition for transformations of $\mathbb{R}^n$ in terms of inputs and outputs.
- Relate *images* to graphical problems like italicising $N$.

**Notes/Misconceptions**

- The idea of images applies to linear and non-linear functions. Though soon we will only consider linear functions, now we will work in generality.

# Task 2.3: Pat and Jamie



Suppose that the “N” on the left is written in regular 12-point font. Find a matrix $A$ that will transform the “N” into the letter on the right which is written in an *italic* 16-point font.

Two students—Pat and Jamie—explained their approach to the Italicizing N task as follows:

> *In order to find the matrix A, we are going to find a matrix that makes the “N” taller, find a matrix that italicizes the taller “N,” and a combination of those two matrices will give the desired matrix A.*

1. Do you think Pat and Jamie’s approach allowed them to find $A$? If so, do you think they found the same matrix that you did during Italicising N?

2. Try Pat and Jamie’s approach. Either (a) come up with a matrix $A$ using their approach, or (b) explain why their approach does not work.

**Decompose a transformation into a composition of simpler transformations.**

The goal of this problem is to

- Decompose a transformation into simpler ones.
- Produce examples showing matrix multiplication is not commutative.

**Notes/Misconceptions**

- Carefully plan how to name your matrices. For example, $T$ for the matrix that makes it *taller* and $L$ for the matrix that *leans* the $N$.
- Some students will have the question, “Do we lean the taller $N$ or the original $N$?” Make sure this discussion point comes out.

Lesson 20

# Composition & Matrix Transformations, Range, Nullspace

## Textbook

Section 3.4

## Objectives

- Recognize matrix multiplication as an operation related to function composition.
- Define *range* and *null space* for both a matrix and a transformation.
- Compute the range and null space of linear transformations described with a formula or geometrically.

## Motivation

We've studied composition of linear transformations in detail, and we have explored matrix transformations. Now it is time to see how the two relate. That is, given two matrix transformations, the matrix for their composition can be obtained by multiplying the matrices in the proper order.

Every linear transformation has a range and a null space. The dimension of the range will give the rank of the transformation, and the dimension of the null space will tell how far a transformation is from being one-to-one. These two subspaces come up a lot when posing and solving problems, so we'd like to get some familiarity with them, both computationally and visually.

43 Define $\mathcal{P}$ to be projection onto $\operatorname{span}\{\vec{u}\}$ where $\vec{u}=\begin{bmatrix}2\\3\end{bmatrix}$, and let $\mathcal{R}$ be rotation counter-clockwise by 90°.

**Connect function composition and matrix multiplication.**

The goal of this problem is to

- Distinguish between matrices and linear transformations.
- Explain the relationship between matrix multiplication and composition of linear transformations.

**Notes/Misconceptions**

- Many students won't distinguish between $\mathcal{P}$ and $P$. Nor will the understand that the fact that $PR$ is the matrix for $P\circ R$ is a theorem—they will think it's obvious.
- For part 3, many students will multiply matrices (based on intuition), but couldn't answer the question without matrix multiplication. Encourage them to return to the definition of function composition and the procedure for finding a matrix for a linear transformation.

43.1 Find a matrix $P$ so that $P\vec{x}=\mathcal{P}(\vec{x})$ for all $\vec{x}\in\mathbb{R}^2$.

$P=\frac{1}{13}\begin{bmatrix}4&6\\6&9\end{bmatrix}$ is such a matrix.

The matrix $P$ corresponding to $\mathcal{P}$ is a $2\times 2$ matrix, so suppose $P=\begin{bmatrix}a&b\\c&d\end{bmatrix}$ for some $a,b,c,d\in\mathbb{R}$. Then we know that if $\{\vec{e}_1,\vec{e}_2\}$ is the standard basis for $\mathbb{R}^2$,

$$P(\vec{e}_1)=\begin{bmatrix}a\\c\end{bmatrix}\quad\text{and}\quad P(\vec{e}_2)=\begin{bmatrix}b\\d\end{bmatrix}.$$

We know from an earlier exercise that $\mathcal{P}(\vec{x})=\dfrac{\vec{x}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\vec{u}$. Therefore, the first column of $P$ is

$$\begin{bmatrix}a\\c\end{bmatrix}=\mathcal{P}(\vec{e}_1)=\frac{2}{13}\vec{u}=\frac{1}{13}\begin{bmatrix}4\\6\end{bmatrix}$$

and the second column of $P$ is

$$\begin{bmatrix}b\\d\end{bmatrix}=\mathcal{P}(\vec{e}_2)=\frac{3}{13}\vec{u}=\frac{1}{13}\begin{bmatrix}6\\9\end{bmatrix}.$$

43.2 Find a matrix $R$ so that $R\vec{x}=\mathcal{R}(\vec{x})$ for all $\vec{x}\in\mathbb{R}^2$.

$R=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$ is such a matrix.

Using the same reasoning as the previous part, we can compute

$$\mathcal{R}(\vec{e}_1)=\vec{e}_2=\begin{bmatrix}0\\1\end{bmatrix}\quad\text{and}\quad\mathcal{R}(\vec{e}_2)=-\vec{e}_1=\begin{bmatrix}-1\\0\end{bmatrix}.$$

Therefore, the matrix $R$ for $\mathcal{R}$ is the matrix with the two vectors above as its respective columns.

43.3 Write down matrices $A$ and $B$ for $\mathcal{P}\circ\mathcal{R}$ and $\mathcal{R}\circ\mathcal{P}$.

$A=\frac{1}{13}\begin{bmatrix}6&-4\\9&-6\end{bmatrix}$ and $B=\frac{1}{13}\begin{bmatrix}-6&-9\\4&6\end{bmatrix}$ are two such matrices.

Using the same reasoning as above, we can compute

$$(\mathcal{P}\circ\mathcal{R})(\vec{e}_1)=\mathcal{P}(\mathcal{R}(\vec{e}_1))=\mathcal{P}(\vec{e}_2)=\frac{1}{13}\begin{bmatrix}6\\9\end{bmatrix}\qquad\text{and}\qquad(\mathcal{P}\circ\mathcal{R})(\vec{e}_2)=\mathcal{P}(\mathcal{R}(\vec{e}_2))=\mathcal{P}(-\vec{e}_1)=\frac{1}{13}\begin{bmatrix}-4\\-6\end{bmatrix}.$$

Therefore, the matrix $A$ for $\mathcal{P}\circ\mathcal{R}$ is the matrix with the two vectors above as its respective columns.

Similarly, for $\mathcal{R}\circ\mathcal{P}$, we can compute:

$$(\mathcal{R}\circ\mathcal{P})(\vec{e}_1)=\mathcal{R}(\mathcal{P}(\vec{e}_1))=\mathcal{R}\left(\frac{1}{13}\begin{bmatrix}4\\6\end{bmatrix}\right)=\frac{1}{13}\begin{bmatrix}-6\\4\end{bmatrix}$$
$$(\mathcal{R}\circ\mathcal{P})(\vec{e}_2)=\mathcal{R}(\mathcal{P}(\vec{e}_2))=\mathcal{R}\left(\frac{1}{13}\begin{bmatrix}6\\9\end{bmatrix}\right)=\frac{1}{13}\begin{bmatrix}-9\\6\end{bmatrix}.$$

Therefore, the matrix $B$ for $\mathcal{R}\circ\mathcal{P}$ is the matrix with these two vectors as its respective columns.

43.4 How do the matrices $A$ and $B$ relate to the matrices $P$ and $R$?

$A=PR$ and $B=RP$.

We can compute these matrix products to see this, but from the previous parts, we know that for any vector $\vec{x}$

$$A\vec{x}=(\mathcal{P}\circ\mathcal{R})(\vec{x})=\mathcal{P}(\mathcal{R}(\vec{x}))=\mathcal{P}(R\vec{x})=PR\vec{x}$$

and

$$B\vec{x} = (\mathcal{R} \circ \mathcal{P})(\vec{x}) = \mathcal{R}(\mathcal{P}(\vec{x}) = \mathcal{R}(P\vec{x}) = RP\vec{x}.$$

Using $\vec{x} = \vec{e}_1$ shows that first column of $A$ must equal the first column of $PR$, and using $\vec{x} = \vec{e}_2$ shows that the second column of $A$ must equal the second column of $PR$, and therefore $A = PR$. For the same reason, we must also have $B = RP$.

**Range**

DEF

The ***range*** (or ***image***) of a linear transformation $T : V \to W$ is the set of vectors that $T$ can output. That is,

$$\text{range}(T) = \{\vec{y} \in W \ : \ \vec{y} = T\vec{x} \text{ for some } \vec{x} \in V\}.$$

**Null Space**

DEFINITION

The ***null space*** (or ***kernel***) of a linear transformation $T : V \to W$ is the set of vectors that get mapped to zero under $T$. That is,

$$\text{null}(T) = \{\vec{x} \in V \ : \ T\vec{x} = \vec{0}\}.$$

---

44 Let $\mathcal{P} : \mathbb{R}^2 \to \mathbb{R}^2$ be projection onto span$\{\vec{u}\}$ where $\vec{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ (like before).

44.1 What is the range of $\mathcal{P}$?

range$(\mathcal{P}) = $ span$\{\vec{u}\}$.

$\mathcal{P}(\vec{x})$ is by definition the vector in span$\{\vec{u}\}$ that is closest to $\vec{x}$, so in particular $\mathcal{P}(\vec{x}) \in$ span$\{\vec{u}\}$ for all $\vec{x} \in \mathbb{R}^2$. Therefore range$(\mathcal{P}) \subseteq$ span$\{\vec{u}\}$.

On the other hand, $\mathcal{P}(\alpha\vec{u}) = \alpha\mathcal{P}(\vec{u}) = \alpha\vec{u}$ for any scalar $\alpha$, and so range$(\mathcal{P}) = $ span$\{\vec{u}\}$.

44.2 What is the null space of $\mathcal{P}$?

$$\text{null}(\mathcal{P}) = \text{span}\left\{\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right\}.$$

A vector $\vec{x}$ projects to $\vec{0}$ if and only if $\vec{x}$ is on the line perpendicular to span$\{\vec{u}\}$ passing through the origin.

**Understanding ranges and null spaces.**

The goal of this problem is to

- Read and apply the definition of range and null space.
- Geometrically visualize the range and null space of a projection.

**Notes/Misconceptions**

- The definition of range is *much* harder for students because there is an extra quantifier.
- Encourage students to draw pictures. Projections are nice because you can illustrate a set and its image in the same drawing without the drawing becoming too cluttered.

---

45 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be an arbitrary linear transformation.

45.1 Show that the null space of $T$ is a subspace.

(i) Let $\vec{u}, \vec{v} \in$ null$(T)$. Applying the linearity of $T$ we see $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}) = \vec{0} + \vec{0} = \vec{0}$, and so $\vec{u} + \vec{v} \in$ null$(T)$.

(ii) Let $\vec{u} \in$ null$(T)$ and let $\alpha$ be any scalar. Again using the linearity of $T$ we see $T(\alpha\vec{u}) = \alpha T(\vec{u}) = \alpha\vec{0} = \vec{0}$, and so $\alpha\vec{u} \in$ null$(T)$.

45.2 Show that the range of $T$ is a subspace.

(i) Let $\vec{y}, \vec{z} \in$ range$(T)$. Then there exist $\vec{u}, \vec{v} \in \mathbb{R}^n$ such that $T(\vec{u}) = \vec{y}$ and $T(\vec{v}) = \vec{z}$. Then $\vec{y} + \vec{z} = T(\vec{u}) + T(\vec{v}) = T(\vec{u} + \vec{v})$, since $T$ is linear, and so $\vec{y} + \vec{z} \in$ range$(T)$.

(ii) Let $\vec{y} \in$ range$(T)$ and let $\alpha$ be any scalar. Then there exists $\vec{u} \in \mathbb{R}^n$ such that $T(\vec{u}) = \vec{y}$, and $\alpha\vec{y} = \alpha T(\vec{u}) = T(\alpha\vec{u})$, since $T$ is linear, and so $\alpha\vec{y} \in$ range$(T)$.

**Practicing proofs.**

The goal of this problem is to

- Practice proving an abstract set (the range or the null space) is a subspace.

**Notes/Misconceptions**

- These proofs have the same template as the previous subspace proofs but are more abstract and so will present a new challenge to students. When they are stuck, emphasize the technique: *write the definitions; write what you want; write what you know.* Most will not have internalized this procedure!

Lesson 21

# Fundamental Subspaces

## Textbook

Section 3.4

## Objectives

- Recognize a matrix as the representation of a linear transformation in a basis.
- Distinguish between matrices and linear transformations.
- Compute the fundamental subspaces of a matrix and relate them to the range and null space of the corresponding linear transformation.
- Identify which subspaces of a matrix are changed when row operations are performed.

## Motivation

Just like column vectors were really a shorthand for describing linear combinations of standard basis vectors, matrices are a shorthand for describing linear transformations. Most of the time we play fast and loose with notation and treat matrices as linear transformations. However there are times, like when talking about change of basis, when it is important to distinguish between a transformation and its representation. In preparation for this, we are going to spend some time studying the relationship between linear transformations and matrices.

Once we find a matrix for a linear transformation, we have a natural set of row vectors and column vectors (coming from the matrix). These span the row space and the column space, and together with the null space give the fundamental subspaces of a matrix. The column space of a matrix corresponds to the range of a transformation and the null space of a matrix corresponds to the null space of a transformation. The row space is the odd one out, but is included since it is the orthogonal complement to the null space (which will allow us to tie together normal form and vector form of planes through the origin).

In this setting, we also start studying row reduction as a "transformation". It changes the column space but not the null space. When we study elementary matrices and inverses, we will see row reduction in terms of matrix multiplication.

### Induced Transformation

DEFINITION

Let $M$ be an $n \times m$ matrix. We say $M$ ***induces*** a linear transformation $T_M : \mathbb{R}^m \to \mathbb{R}^n$ defined by

$$[T_M \vec{v}]_{\mathcal{E}'} = M[\vec{v}]_{\mathcal{E}},$$

where $\mathcal{E}$ is the standard basis for $\mathbb{R}^m$ and $\mathcal{E}'$ is the standard basis for $\mathbb{R}^n$.

**Notes/Misconceptions**

- Students will ask if $\mathcal{E}$ and $\mathcal{E}'$ are the same thing.

---

46 Let $M$ be a $2 \times 2$ matrix and let $\vec{v} \in \mathbb{R}^2$. Further, let $T_M$ be the transformation induced by $M$.

46.1 What is the difference between "$M\vec{v}$" and "$M[\vec{v}]_{\mathcal{E}}$"?

"$M\vec{v}$" is ambiguous notation, as it is only defined if $\vec{v}$ is a specific column vector. There are infinitely many different bases of $\mathbb{R}^2$, and so a given vector $\vec{v}$ has infinitely many different representations as a column vector, each in a different basis.

"$M[\vec{v}]_{\mathcal{E}}$" is unambiguous, as $[\vec{v}]_{\mathcal{E}}$ is an explicit representation of $\vec{v}$ in a particular basis.

46.2 What is $[T_M \vec{e}_1]_{\mathcal{E}}$?

It is the first column of $M$.

By definition, $[T_M \vec{e}_1]_{\mathcal{E}} = M[\vec{e}_1]_{\mathcal{E}} = M\begin{bmatrix}1\\0\end{bmatrix}$, which equals the first column of $M$.

46.3 Can you relate the columns of $M$ to the range of $T_M$?

The range of $T_M$ equals the span of the columns of $M$.

By the previous part, the first column of $M$ is in the range of $T_M$. By a similar argument, the second column of $M$ is also in the range of $T_M$, since it equals $[T_M \vec{e}_2]_{\mathcal{E}}$. Therefore the span of the columns of $M$ is a subset of the range of $T_M$.

On the other hand, if $\vec{v}_1 = \begin{bmatrix}a\\c\end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix}b\\d\end{bmatrix}$ are the columns of $M$ and $\vec{x} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2$ is an element of $\operatorname{span}\{\vec{v}_1, \vec{v}_2\}$, then

$$[\vec{x}]_{\mathcal{E}} = \alpha_1\begin{bmatrix}a\\c\end{bmatrix} + \alpha_2\begin{bmatrix}b\\d\end{bmatrix} = \alpha_1[T_M\vec{e}_1]_{\mathcal{E}} + \alpha_2[T_M\vec{e}_2]_{\mathcal{E}} = [T_M(\alpha_1\vec{e}_1 + \alpha_2\vec{e}_2)]_{\mathcal{E}}.$$

Therefore $\vec{x}$ is in the range of $T_M$.

**Formalizing the connection between matrices and linear transformations.**

The goal of this problem is to
- Distinguish between linear transformations and matrices.
- Explain how to relate matrices and linear transformations.
- Practice using formal language and notation, avoiding category errors.

**Notes/Misconceptions**

- At this point, students have a notion that linear transformations and matrices are connected, and some will know that matrices and linear transformations are not the same thing. However, most will not know how to create correct sentences that distinguish between matrices and linear transformations.

  This problem forces the use of precise language and terminology (e.g., that of change of basis) to discuss a matrix and its induced transformation.
- Throughout most of the course, we are sloppy in distinguishing $\begin{bmatrix}1\\2\end{bmatrix}$ and $\begin{bmatrix}1\\2\end{bmatrix}_{\mathcal{E}}$. This allows us to write $T\left(\begin{bmatrix}1\\2\end{bmatrix}\right)$ when $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation (instead of the correct $T\left(\begin{bmatrix}1\\2\end{bmatrix}_{\mathcal{E}}\right)$).

  Acknowledge this—we will continue to be sloppy some times, but when push comes to shove, we must be able to be precise.
- Saying "$T$ is the transformation induced by $M$" is slightly faster than saying "$T$ is the transformation given by multiplication by $M$ when vectors are written in the standard basis", but both are formal and correct. We might also say, "Consider the matrix transformation $M : \mathbb{R}^m \to \mathbb{R}^n$." When we say this, we mean that we are using the letter $M$ to represent both the matrix and the induced transformation.

### Fundamental Subspaces

DEFINITION

Associated with any matrix $M$ are three fundamental subspaces: the ***row space*** of $M$, denoted $\operatorname{row}(M)$, is the span of the rows of $M$; the ***column space*** of $M$, denoted $\operatorname{col}(M)$, is the span of the columns of $M$; and the ***null space*** of $M$, denoted $\operatorname{null}(M)$, is the set of solutions to $M\vec{x} = \vec{0}$.

---

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

47.1 Describe the row space of $A$.

$\operatorname{row}(A) = \operatorname{span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$, which is the $xy$-plane in $\mathbb{R}^3$.

47.2 Describe the column space of $A$.

$$\operatorname{col}(A) = \operatorname{span}\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}0\\0\end{bmatrix}\right\} = \operatorname{span}\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right\} = \mathbb{R}^2.$$

47.3 Is the row space of $A$ the same as the column space of $A$?

No.

Although they are both two dimensional spaces, $\operatorname{row}(A)$ is a subspace of $\mathbb{R}^3$ and all vectors in it have three coordinates (with the third always being zero), while $\operatorname{col}(A)$ is a subspace of $\mathbb{R}^2$ and all vectors in it have two coordinates. Therefore, these two spaces are different.

47.4 Describe the set of all vectors perpendicular to the rows of $A$.

The $z$-axis in $\mathbb{R}^3$.

A vector $\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ is perpendicular to the rows of $A$ if and only if its dot product with both rows is zero. That is

$$\vec{x} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = x = 0 \quad \text{and} \quad \vec{x} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = y = 0.$$

$\vec{x}$ satisfies these equations if and only if $\vec{x} = \begin{bmatrix} 0 \\ 0 \\ t \end{bmatrix}$ for some real number $t$, or in other words if $\vec{x}$ is on the $z$-axis.

47.5 Describe the null space of $A$.

The $z$-axis in $\mathbb{R}^3$.

A vector $\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ is in null($A$) if and only if

$$A\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \vec{0}.$$

These are the same conditions as in the previous part, so the set of vectors satisfying this is the $z$-axis.

47.6 Describe the range and null space of $T_A$, the transformation induced by $A$.

range($T_A$) = col($A$) = $\mathbb{R}^2$ and null($T_A$) = null($A$), which is the $z$-axis in $\mathbb{R}^3$.

By Problem 46.3, the range of an induced transformation equals the span of the columns of the matrix. In other words, range($T_A$) = col($A$).

Next, by definition $\vec{v} \in \text{null}(T_A)$ when $[T_A\vec{v}]_{\mathcal{E}} = A[\vec{v}]_{\mathcal{E}} = \vec{0}$. In other words, $\vec{v} \in \text{null}(T_A)$ if and only if $[\vec{v}]_{\mathcal{E}} \in \text{null}(A)$. We know from the previous part that null($A$) is the $z$-axis in $\mathbb{R}^3$.

**Fundamental subspaces of a matrix.**

The goal of this problem is to
- Compute row and column spaces of a matrix.
- Recognize that row and column spaces may be unrelated.
- Geometrically relate the row space to the null space.
- Connect the fundamental subspaces of a matrix to the range and null space of a transformation.

**Notes/Misconceptions**
- In part 3, a lot of students will say they're both $\mathbb{R}^2$.
- In part 4, students might claim $\mathbb{R}^3 \backslash \{xy\text{-plane}\}$. These students are still thinking about rooted vectors that can be moved anywhere.
- In part 5, emphasize the new geometric interpretation of null space.

---

## 48

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \qquad C = \text{rref}(B) = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}$$

48.1 How does the row space of $B$ relate to the row space of $C$?

They are equal.

Row operations replace rows with linear combinations of rows. Therefore, since $C$ is the matrix $B$ after the application of some row operations, row($C$) $\subseteq$ row($B$).

Since row operations are all reversible, we also know that $B$ can be obtained from $C$ by applying row operations, so row($B$) $\subseteq$ row($C$).

Therefore, row($B$) = row($C$).

48.2 How does the null space of $B$ relate to the null space of $C$?

They are equal.

A vector is in null($B$) or null($C$) if and only if it is orthogonal to all vectors in row($B$) or all vectors in row($C$), respectively. But row($B$) = row($C$) by the previous part, so their null spaces must also be equal.

48.3 Compute the null space of $B$.

$$\text{null}(C) = \text{span}\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}.$$

We compute null($C$), since it equals null($B$) by the previous part.

**Fundamental subspaces and row reduction.**

The goal of this problem is to
- Explain why row reduction doesn't change the row space or the null space.

**Notes/Misconceptions**
- In part 2, encourage students to think both geometrically and algebraically. How do the solutions to $B\vec{x} = \vec{0}$ and $C\vec{x} = \vec{0}$ relate? What is the orthogonal complement to their row spaces?

$\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ is in null($C$) if and only if $C\vec{x} = \begin{bmatrix} x - z \\ y + 2z \end{bmatrix} = \vec{0}$. The complete solution to this matrix equation is

$$\text{null}(C) = \left\{ \begin{bmatrix} t \\ -2t \\ t \end{bmatrix} \in \mathbb{R}^3 : t \in \mathbb{R} \right\} = \text{span}\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}.$$

49

**Fundamental subspaces and row reduction.**

The goal of this problem is to

- Recognize that row reduction may change the column space of a matrix.

**Notes/Misconceptions**

- This question should be straightforward for students.

$$P = \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix} \qquad Q = \text{rref}(P) = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$$

49.1 How does the column space of $P$ relate to the column space of $Q$?

They are not equal, but have the same dimension.

49.2 Describe the column space of $P$ and the column space of $Q$.

$\text{col}(P) = \text{span}\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right\} = \text{span}\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$, which is the $y$-axis in $\mathbb{R}^2$.

$\text{col}(Q) = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix} \right\} = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$, which is the $x$-axis in $\mathbb{R}^2$.

Lesson 22

# Rank

## Textbook

Section 2.2, 3.4

## Objectives

- Compute the rank of a linear transformation or a matrix.
- Relate rank, linear independence of column vectors, and number of solutions.

## Motivation

Rank is a number that tells us a lot about a linear transformation or a matrix. In particular, if a linear transformation $T$ is full rank, we know it is one-to-one and onto and that there are unique solutions to $T(\vec{x}) = \vec{b}$. In contrast, if the rank is not full, we know the transformation either fails to be one-to-one or fails to be onto. By considering the domain of $T$, we can decide which property fails.

The language of rank will help us reason about dimension, and for us the main application will be the rank-nullity theorem. We will not use the ideas of rank to their full potential, but in future courses, students will hear the term "rank" thrown around a lot.

**Notes/Misconceptions**

- Students will gravitate towards rank for matrices because it's definition is "easier". Further, the book emphasizes this definition more. However, when viewed as a fact about transformations, it won't matter which basis you write down a transformation in. For this reason, we study the rank of a transformation directly.

**Rank**

DEF

For a linear transformation $T : V \to W$, the ***rank*** of $T$, denoted $\operatorname{rank}(T)$, is the dimension of the range of $T$.

For an $n \times m$ matrix $M$, the ***rank*** of $M$, denoted $\operatorname{rank}(M)$, is the number of pivots in $\operatorname{rref}(M)$.

---

50 Let $\mathcal{P}$ be projection onto $\operatorname{span}\{\vec{u}\}$ where $\vec{u} = \begin{bmatrix}2\\3\end{bmatrix}$, and let $\mathcal{R}$ be rotation counter-clockwise by $90°$.

50.1 Describe $\operatorname{range}(\mathcal{P})$ and $\operatorname{range}(\mathcal{R})$.

$\operatorname{range}(\mathcal{P}) = \operatorname{span}\{\vec{u}\}$, and $\operatorname{range}(\mathcal{R}) = \mathbb{R}^2$.

For $\mathcal{P}$, by the definition of projection $\mathcal{P}(\vec{x})$ is the vector in $\operatorname{span}\{\vec{u}\}$ that is closest to $\vec{x}$, so in particular $\mathcal{P}(\vec{x}) \in \operatorname{span}\{\vec{u}\}$ for all $\vec{x} \in \mathbb{R}^2$. Therefore $\operatorname{range}(\mathcal{P}) \subseteq \operatorname{span}\{\vec{u}\}$.

On the other hand, $\mathcal{P}(\alpha\vec{u}) = \alpha\mathcal{P}(\vec{u}) = \alpha\vec{u}$ for any scalar $\alpha$, and so $\operatorname{range}(\mathcal{P}) = \operatorname{span}\{\vec{u}\}$.

For $\mathcal{Q}$, we have that any vector $\vec{x} \in \mathbb{R}^2$, $\vec{x} = \mathcal{Q}(\vec{y})$, where $\vec{y}$ is the rotation of $\vec{x}$ clockwise by $90°$. Therefore $\operatorname{range}(\mathcal{Q}) = \mathbb{R}^2$.

50.2 What is the rank of $\mathcal{P}$ and the rank of $\mathcal{R}$?

$\operatorname{rank}(\mathcal{P}) = 1$ and $\operatorname{rank}(\mathcal{R}) = 2$.

By the previous part, we know $\operatorname{range}(\mathcal{P})$ is 1-dimensional and $\operatorname{range}(\mathcal{Q})$ is 2-dimensional.

50.3 Let $P$ and $R$ be the matrices corresponding to $\mathcal{P}$ and $\mathcal{R}$. What is the rank of $P$ and the rank of $R$?

$\operatorname{rank}(P) = 1$ and $\operatorname{rank}(R) = 2$.

By Problem 43, $P = \frac{1}{13}\begin{bmatrix}4 & 6\\6 & 9\end{bmatrix}$ and $R = \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}$ are the matrices corresponding to $\mathcal{P}$ and $\mathcal{R}$. Then we compute:

$$\operatorname{rref}(P) = \begin{bmatrix}1 & \frac{3}{2}\\0 & 0\end{bmatrix} \quad \text{and} \quad \operatorname{rref}(R) = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}.$$

These matrices have 1 and 2 pivots, respectively.

50.4 Make a conjecture about how the rank of a transformation and the rank of its corresponding matrix relate. Can you justify your claim?

They are equal.

By Problem 46.3, the range of a transformation is equal to the column space of its corresponding matrix, and therefore the dimensions of these two spaces are equal. In other words, the rank of a transformation is equal to the dimension of the column space of its corresponding matrix.

We already know that the dimension of the column space of a matrix is equal to the number of pivots in its reduced row echelon form, and that is by definition the rank of the matrix.

**Rank of linear transformations.**

The goal of this problem is to
- Apply the definition of *rank* to compute the rank of a linear transformation.
- Use geometric intuition to compute the rank of a linear transformation.
- Relate the rank of a linear transformation to the rank of its matrix.

**Notes/Misconceptions**
- Don't labor over proofs in this question.

---

51

51.1 Determine the rank of (a) $\begin{bmatrix}1 & 1\\2 & 2\end{bmatrix}$ (b) $\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}$ (c) $\begin{bmatrix}1 & 1 & 0\\0 & 0 & 1\end{bmatrix}$ (d) $\begin{bmatrix}3\\3\\2\end{bmatrix}$ (e) $\begin{bmatrix}1 & 0 & 1\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$.

For each part, we compute the reduced row echelon form of the matrix and count the number of pivots.

(a) $\operatorname{rank}\left(\begin{bmatrix}1 & 1\\2 & 2\end{bmatrix}\right) = 1$, since $\operatorname{rref}\left(\begin{bmatrix}1 & 1\\2 & 2\end{bmatrix}\right) = \begin{bmatrix}1 & 1\\0 & 0\end{bmatrix}$ has one pivot.

(b) $\operatorname{rank}\left(\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\right) = 2$, since $\operatorname{rref}\left(\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\right) = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$ has two pivots

(c) $\operatorname{rank}\left(\begin{bmatrix}1 & 1 & 0\\0 & 0 & 1\end{bmatrix}\right) = 2$. This matrix is already in reduced row echelon form, and has two pivots.

**Rank of matrices.**

The goal of this problem is to
- Use the definition of rank to compute the rank of matrices.

**Notes/Misconceptions**
- Part (d) will be the only hard one; don't spend much time on this question.

(d) rank$\left(\begin{bmatrix}3\\3\\2\end{bmatrix}\right)=1$, since rref$\left(\begin{bmatrix}3\\3\\2\end{bmatrix}\right)=\begin{bmatrix}1\\0\\0\end{bmatrix}$ has one pivot.

(e) rank$\left(\begin{bmatrix}1&0&1\\0&1&0\\0&0&1\end{bmatrix}\right)=3$, since rref$\left(\begin{bmatrix}1&0&1\\0&1&0\\0&0&1\end{bmatrix}\right)=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$ has three pivots.

---

52 Consider the homogeneous system

$$\begin{array}{llll} x & +2y & +z & =0\\ x & +2y & +3z & =0\\ -x & -2y & +z & =0\end{array} \tag{1}$$

and the non-augmented matrix of coefficients $A=\begin{bmatrix}1&2&1\\1&2&3\\-1&-2&1\end{bmatrix}$.

**Connect rank to existing concepts.**

The goal of this problem is to

- Connect rank($A$) to the number of solutions to $A\vec{x}=\vec{0}$.
- Connect rank($A$) to linear independence or dependence of the columns of $A$.

**Notes/Misconceptions**

- This problem is a warmup to the abstract one that follows.

52.1 What is rank($A$)?

rank($A$) = 2, since rref($A$) $=\begin{bmatrix}1&2&0\\0&0&1\\0&0&0\end{bmatrix}$ as two pivots.

52.2 Give the general solution to system (1).

$\vec{x}$ is a solution to the system if $\vec{x}=t\begin{bmatrix}-2\\1\\0\end{bmatrix}$ for some real number $t$.

If $\vec{x}=\begin{bmatrix}x\\y\\z\end{bmatrix}$ is a solution to the system, then we must have $\begin{bmatrix}1&2&0\\0&0&1\\0&0&0\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}x+2y\\z\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$, from which it follows that $z=0$ and $x=-2y$. In other words, any scalar multiple of $\begin{bmatrix}-2\\1\\0\end{bmatrix}$ is a solution.

52.3 Are the column vectors of $A$ linearly independent?

No. The second column is two times the first column.

52.4 Give a non-homogeneous system with the same coefficients as (1) that has

(a) infinitely many solutions

(b) no solutions.

(a)

$$\begin{array}{llll} x & +2y & +z & =1\\ x & +2y & +3z & =1\\ -x & -2y & +z & =-1\end{array}$$

(b)

$$\begin{array}{llll} x & +2y & +z & =0\\ x & +2y & +3z & =0\\ -x & -2y & +z & =1\end{array}$$

---

53 53.1 The rank of a $3\times 4$ matrix $A$ is 3. Are the column vectors of $A$ linearly independent?

No. A $3\times 4$ matrix has four columns, each of which are vectors in $\mathbb{R}^3$. It is not possible for four different vectors in $\mathbb{R}^3$ to be linearly independent.

53.2 The rank of a $4\times 3$ matrix $B$ is 3. Are the column vectors of $B$ linearly independent?

Yes. Since rank($B$) = 3, there are three pivots in rref($B$). Pivot positions in rref($B$) indicate a maximal linearly independent subset of the columns of $B$. Since there are three columns in $B$ and three pivots, the three columns of $B$ must be linearly independent.

**Connect the rank of a matrix to the linear independence/dependence of its columns.**

The goal of this problem is to

- Determine the linear independence/dependence of the columns of a matrix based on its size and rank.

Lesson 23

# Rank-nullity Theorem, Inverses I

## Textbook

Section 3.4, 3.5

## Objectives

- Relate the rank-nullity theorem for matrices and the rank-nullity theorem for transformations.
- Use the rank-nullity theorem to compute ranks and nullities.
- View an inverse transformation as one that "undoes" the transformation.
- Recognize the rank of $A$ and $A^T$ are equal.

## Motivation

The rank-nullity theorem is our big application of rank. It states that whatever left-over dimensions of the domain are not accounted for by the rank end up in the null space. This is valuable to know because a matrix equation $A\vec{x} = \vec{b}$ has infinitely many solutions if and only if $A\vec{x} = \vec{b}$ is consistent and $A$ has a non-trivial null space. Using the rank-nullity theorem, we can connect the rank of a matrix to whether a matrix equation has infinitely many solutions. And, pulling this back into the world of transformations, we can start reasoning about the number of solutions to an abstract equation $T(\vec{x}) = \vec{b}$ where $T$ is a linear transformation.

We also start talking about inverses. Invertible transformations always have full rank, but our goal here is to start thinking about transformations that "undo" other transformations. We are building up to things like an intuitive understanding of $(T \circ S)^{-1} = S^{-1} \circ T^{-1}$.

**Notes/Misconceptions**

- The fact that $\text{rank}(A) = \text{rank}(A^T)$ is easy to see by reasoning about the dimensions of the row space and column space of $A$ via row-reduction and pivots. This is where RREF really shines.

## Rank-nullity Theorem

THEOREM

The ***nullity*** of a matrix is the dimension of the null space.

The rank-nullity theorem for a matrix $A$ states

$$\operatorname{rank}(A) + \operatorname{nullity}(A) = \#\text{ of columns in } A.$$

---

54 54.1 Is there a version of the rank-nullity theorem that applies to linear transformations instead of matrices? If so, state it.

Yes. If $T : V \to W$ is a linear transformation, then $\operatorname{rank}(T) + \dim(\operatorname{null}(T)) = \dim(V)$.

If $A$ is the matrix corresponding to $T$, then $\operatorname{rank}(T) = \operatorname{rank}(A)$ by Problem 50.4.

$\operatorname{null}(T) = \operatorname{null}(A)$ by Problem 47.6, since $T = T_A$, and so $\dim(\operatorname{null}(T)) = \operatorname{nullity}(A)$.

Finally, the number of columns of $A$ is equal to the dimension of the domain of $T$.

**Relate linear transformation concepts with matrix concepts.**

The goal of this problem is to

- Rephrase the rank-nullity theorem as stated for matrices as the rank-nullity theorem for linear transformations.

**Notes/Misconceptions**

- The rank-nullity theorem for matrices is easy to argue for using the reduced row echelon form. The rank-nullity theorem for linear transformations is much harder to prove (unless you exploit the links between transformations and their corresponding matrices).

---

55 The vectors $\vec{u}, \vec{v} \in \mathbb{R}^9$ are linearly independent and $\vec{w} = 2\vec{u} - \vec{v}$. Define $A = [\vec{u}|\vec{v}|\vec{w}]$.

55.1 What is the rank and nullity of $A^T$?

$\operatorname{rank}(A^T) = 2$ and $\operatorname{nullity}(A^T) = 7$.

$A^T$ is the matrix with rows $\vec{u}, \vec{v}$, and $\vec{w}$. Since $\vec{w} = 2\vec{u} - \vec{v}$, the third row of $A^T$ can be reduced to a row of zeros by the row operation $R_3 \mapsto R_3 - 2R_1 + R_2$. Neither of the first two rows can be reduced to rows of zeros since they are linearly independent. Therefore $\operatorname{rref}(A^T)$ has two pivots, meaning $\operatorname{rank}(A^T) = 2$.

The rank-nullity theorem then says that $2 + \operatorname{nullity}(A^T) = 9$, and so $\operatorname{nullity}(A^T) = 7$

55.2 What is the rank and nullity of $A$?

$\operatorname{rank}(A) = 2$ and $\operatorname{nullity}(A) = 1$.

We know that $\operatorname{rank}(A)$ equals the number of pivots in $\operatorname{rref}(A)$, which in turn equals the dimension of $\operatorname{col}(A)$. Since $A$ has two linearly independent columns, $\dim(\operatorname{col}(A)) = 2$.

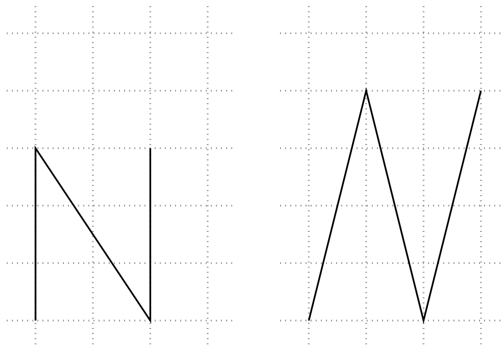
Again, the rank-nullity theorem then says that $2 + \operatorname{nullity}(A) = 3$, and so $\operatorname{nullity}(A) = 1$

**Apply the rank-nullity theorem.**

The goal of this problem is to

- Apply the rank-nullity theorem to compute the rank or nullity of unknown matrices.

## Task 2.4: Getting back N



Suppose that the "N" on the left is written in regular 12-point font. Find a matrix $A$ that will transform the "N" into the letter on the right which is written in an *italic* 16-point font.

Two students—Pat and Jamie—explained their approach to the Italicizing N task as follows:

> *In order to find the matrix A, we are going to find a matrix that makes the "N" taller, find a matrix that italicizes the taller "N," and a combination of those two matrices will give the desired matrix A.*

Consider the new task: find a matrix $C$ that transforms the "N" on the right to the "N" on the left.

1. Use any method you like to find $C$.
2. Use a method similar to Pat and Jamie's method, only use it to find $C$ instead of $A$.

**Notes/Misconceptions**

- This is the time to think about composition and "undoing" transformations. In particular, the order that you undo them in matters!
- Plan the names of your transformations carefully. One option is to use the letters $T$ for matrix that makes it taller, $L$ for the matrix that leans, $U$ for the matrix that unleans, and $S$ for the matrix that shortens.

Lesson 24

# Inverses II, Elementary Matrices

## Textbook

Section 3.5, 3.6

## Objectives

- Define the *inverse* of a matrix.
- Define an *elementary matrix*.
- Use elementary matrices to compute inverses.

## Motivation

A common mathematical technique is to decompose a complicated problem into lots of simpler ones. In this lesson, the complicated problem is that of finding the inverse of a matrix. We will do this by recasting row reduction in terms of elementary matrices and using the rules of matrix arithmetic to produce the inverse of a matrix. This is not the most efficient way to produce the inverse, but we do it this way for three reasons:

- Elementary matrices bring row reduction out of the realm of "algorithms on entries" and into the realm of matrices and function composition.
- Decomposition into elementary matrices foreshadows further matrix decompositions (like $LU$ or $QR$, etc.), even though we won't be working with those decompositions in this course.
- Decomposition into elementary matrices will be used to compute general determinants.

As far as inverses are concerned, they will allow us to recast solutions to systems of equations as "inverse problems".

**Notes/Misconceptions**

- Most students won't know the definition of the inverse of a function. They think of the inverse of a function in terms of "the reflection of the graph across $y = x$". This is a big problem for us.

56 56.1 Apply the row operation $R_3 \mapsto R_3 + 2R_1$ to the $3 \times 3$ identity matrix and call the result $E_1$.

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} \xrightarrow{R_3 \mapsto R_3+2R_1} \begin{bmatrix}1&0&0\\0&1&0\\2&0&1\end{bmatrix} = E_1.$$

56.2 Apply the row operation $R_3 \mapsto R_3 - 2R_1$ to the $3 \times 3$ identity matrix and call the result $E_2$.

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} \xrightarrow{R_3 \mapsto R_3-2R_1} \begin{bmatrix}1&0&0\\0&1&0\\-2&0&1\end{bmatrix} = E_2.$$

DEF An ***elementary matrix*** is the identity matrix with a single row operation applied.

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

56.3 Compute $E_1A$ and $E_2A$. How do the resulting matrices relate to row operations?

$$E_1A = \begin{bmatrix}1&2&3\\4&5&6\\9&12&15\end{bmatrix} \text{ and } E_2A = \begin{bmatrix}1&2&3\\4&5&6\\5&4&3\end{bmatrix}.$$

$E_1A$ is the result applying the row operation $R_3 \mapsto R_3 + 2R_1$ to $A$, and similarly $E_2A$ is the result of applying the row operation $R_3 \mapsto R_3 - 2R_1$ to $A$.

56.4 Without computing, what should the result of applying the row operation $R_3 \mapsto R_3 - 2R_1$ to $E_1$ be? Compute and verify.

It should be the identity matrix, since the row operation $R_3 \mapsto R_3 - 2R_1$ should undo the operation $R_3 \mapsto R_3 + 2R_1$.

56.5 Without computing, what should $E_2E_1$ be? What about $E_1E_2$? Now compute and verify.

They should both be the identity matrix.

The solution to part 3 above lead us to believe that applying $E_1$ to a matrix has the effect of applying the row operation $R_3 \mapsto R_3 + 2R_1$ to it. Applying that row operation to $E_2$ would produce the identity matrix, so we expect that $E_1E_2$ should equal the identity matrix.

Similar reasoning leads us to believe that $E_2E_1$ should also equal the identity matrix.

Indeed, we can compute

$$E_1E_2 = \begin{bmatrix}1&0&0\\0&1&0\\2&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\-2&0&1\end{bmatrix} = \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} = \begin{bmatrix}1&0&0\\0&1&0\\-2&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\2&0&1\end{bmatrix} = E_2E_1.$$

DEF The ***inverse*** of a matrix $A$ is a matrix $B$ such that $AB = I$ and $BA = I$. In this case, $B$ is called the inverse of $A$ and is notated by $A^{-1}$.

57 Consider the matrices

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

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

**Elementary matrices.**

The goal of this problem is to
- Define *elementary matrices*.
- Relate elementary matrices to row reduction.
- Use the "reversibility" of elementary row operations to create inverses to elementary matrices.

**Notes/Misconceptions**
- One punchline is that elementary matrices allow do the same thing as elementary row operations. This means we can understand row reduction in terms of matrix multiplication instead of as an algorithm that operates on matrix entries.
- Parts 4 and 5 foreshadow inverses. They're meant to start building the intuition that inverses "undo" operations.

**Notes/Misconceptions**
- The definition is written as $AB = I$ and $BA = I$ instead of $AB = BA = I$ so that a simple dimension argument cannot rule out invertibility.

**Apply the definition of *inverse matrix*.**

The goal of this problem is to
- Use the definition of *inverse matrix* to identify whether two matrices are inverses of each other.

**Notes/Misconceptions**
- In a large class, students will find all pairs, but many will miss $F,F$ as a pair.
- Make sure to have a discussion of $B,C$, since $BC = I$ but $CB \neq I$. The fact you can multiply both ways and get $I$ is a really important part of the definition!

57.1 Which pairs of matrices above are inverses of each other?

$A$ and $D$ are inverses of each other, and $F$ is its own inverse.

---

58

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

58.1 Use two row operations to reduce $B$ to $I_{2\times 2}$ and write an elementary matrix $E_1$ corresponding to the first operation and $E_2$ corresponding to the second.

$$\begin{bmatrix}1&4\\0&2\end{bmatrix}\xrightarrow{R_2\mapsto\frac{1}{2}R_2}\begin{bmatrix}1&4\\0&1\end{bmatrix}\xrightarrow{R_1\mapsto R_1-4R_2}\begin{bmatrix}1&0\\0&1\end{bmatrix}.$$

The two elementary matrices are $E_1=\begin{bmatrix}1&0\\0&\frac{1}{2}\end{bmatrix}$ and $E_2=\begin{bmatrix}1&-4\\0&1\end{bmatrix}$.

58.2 What is $E_2E_1B$?

$$E_2E_1B=\begin{bmatrix}1&-4\\0&1\end{bmatrix}\begin{bmatrix}1&0\\0&\frac{1}{2}\end{bmatrix}\begin{bmatrix}1&4\\0&2\end{bmatrix}=\begin{bmatrix}1&0\\0&1\end{bmatrix}.$$

58.3 Find $B^{-1}$.

$$B^{-1}=E_2E_1=\begin{bmatrix}1&-2\\0&\frac{1}{2}\end{bmatrix}=\frac{1}{2}\begin{bmatrix}2&-4\\0&1\end{bmatrix}.$$

By the previous part we already know that $(E_2E_1)B=I$. We can also check that $B(E_2E_1)=I$, meaning $E_2E_1$ is the inverse of $B$.

58.4 Can you outline a procedure for finding the inverse of a matrix using elementary matrices?

Suppose $A$ is a matrix that can be row reduced to the identity. Let $E_1,E_2,\ldots,E_n$ be the elementary matrices corresponding to the sequence of row operations that reduces $A$ to $I$. Then as we have seen, we have $E_nE_{n-1}\cdots E_2E_1A=I$.

Thus $E_nE_{n-1}\cdots E_2E_1$ is the inverse of $A$.

**Compute inverses.**

The goal of this problem is to

- Use elementary matrices to compute matrix inverses.
- Decompose an invertible matrix into the product of elementary matrices.

**Notes/Misconceptions**

- There are two different ways to row reduce $B$ using two elementary row operations. They give different pairs of elementary matrices, but produce the same inverse. This is a good discussion point.
- In part 3, many will forget to check that $B(E_2E_1)=I$. They don't know the theorem that for a square matrix a left inverse is also a right inverse, so they must check this!
- In part 4, emphasize the order in which the elementary matrices must be multiplied.
- You can use the two different decompositions of $B^{-1}$ as an opportunity to show that order matters.

Lesson 25

# Applications of Inverses I

## Textbook

Section 3.5

## Objectives

- Recognize applying an inverse as "undoing" a transformation.
- Use inverses to solve matrix equations.
- Relate inverses to row reduction.
- Correctly reproduce a formula for $(XY)^{-1}$.
- Apply inverses to change of basis.

## Motivation

Inverses give a new way to solve systems of equations. Using row reduction, solving $A\vec{x} = \vec{b}_i$ for $i = 1, \ldots, 7$ would involve row-reducing 7 times. But, if we compute $A^{-1}$ we can do matrix multiplication instead of row reduction!

59

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & 4 \\ 1 & 3 & -3 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad C = [A|\vec{b}] \qquad A^{-1} = \begin{bmatrix} 9 & -3/2 & -5 \\ -5 & 1 & 3 \\ -2 & 1/2 & 1 \end{bmatrix}$$

59.1 What is $A^{-1}A$?

$A^{-1}A = I$. This is true by the definition of an inverse, but we can also verify it by hand.

59.2 What is rref($A$)?

$$\text{rref}(A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I.$$

59.3 What is rref($C$)? (Hint, there is no need to actually do row reduction!)

$$\text{rref}(C) = \left[I\middle|A^{-1}\vec{b}\right] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & -9 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & 2 \end{array}\right].$$

We know that the reduced row echelon form of $C$ must be of the form $[I|\vec{c}]$ for some $\vec{c}$, and we know that multiplying on the left by $A^{-1}$ is equivalent to applying the sequence of row operations that reduces $A$ to $\text{rref}(A) = I$. So the same sequence of row operations applied to $\vec{b}$, the last column of $C$, will produce the vector $\vec{c} = A^{-1}\vec{b} = \begin{bmatrix} -9 \\ 6 \\ 2 \end{bmatrix}$.

59.4 Solve the system $A\vec{x} = \vec{b}$.

The system has one solution: $\vec{x} = A^{-1}\vec{b} = \begin{bmatrix} -9 \\ 6 \\ 2 \end{bmatrix}$.

We can read this solution from the reduced row echelon form of the augmented matrix $C$ representing this system. We can also multiply both sides of the equation on the left by $A^{-1}$:

$$A\vec{x} = \vec{x} \implies A^{-1}A\vec{x} = A^{-1}\vec{b} \implies \vec{x} = A^{-1}\vec{b}.$$

**Solve systems with inverses.**

The goal of this problem is to

- Relate inverse matrices to the previous methods for solving equations, row reduction.
- Symbolically write the solution to a matrix equation using inverses.

**Notes/Misconceptions**

- Many will not know how to answer part 3. Ask them to think $A^{-1}$ as representing a series of elementary row operations (because it is a product of elementary matrices).
- In part 4, emphasize that order matters and that $\vec{b}A^{-1}$ doesn't even make sense as a product.

60

60.1 For two square matrices $X, Y$, should $(XY)^{-1} = X^{-1}Y^{-1}$?

No.

By the definition of an inverse we need $(XY)^{-1}(XY) = I$, so that multiplying by $(XY)^{-1}$ undoes multiplication by $XY$. To do this, we must first undo multiplication by $X$, then undo multiplication by $Y$. That is, we must first multiply by $X^{-1}$ then multiply by $Y^{-1}$.

In other words, we expect that $(XY)^{-1} = Y^{-1}X^{-1}$. We can then verify this by computing

$$(XY)(Y^{-1}X^{-1}) = XYY^{-1}X^{-1} = XIX^{-1} = XX^{-1} = I$$

and

$$(Y^{-1}X^{-1})(XY) = Y^{-1}X^{-1}XY = Y^{-1}IY = Y^{-1}Y = I.$$

60.2 If $M$ is a matrix corresponding to a non-invertible linear transformation $T$, could $M$ be invertible?

No.

Suppose $M^{-1}$ exists. Then $M^{-1}M = MM^{-1} = I$. Let $S$ be the linear transformation induced by $M^{-1}$. Since $M$ is the matrix for $T$ we must have $S \circ T = T \circ S = \text{id}$. But then $S$ would be the inverse of $T$, which is impossible.

**Inverses and composition.**

The goal of this problem is to

- Create a correct formula for $(XY)^{-1}$ and explain it algebraically or in terms of function composition.
- Relate invertibility of a matrix and its induced transformation.

**Notes/Misconceptions**

- Depending on time, you can do a "soft" argument for part 2.

## More Change of Basis

**Inverses and change of basis.**

The goal of this problem is to

- Use inverses to answer change-of-basis questions.
- Explain why the inverse of a change-of-basis matrix is another change of basis matrix.

**Notes/Misconceptions**

- It's been a while since change of basis. Students might be rusty.
- This question was foreshadowed in the initial change-of-basis exercises.

61 Let $\mathcal{B}=\{\vec b_1,\vec b_2\}$ where $\vec b_1=\begin{bmatrix}1\\1\end{bmatrix}$, $\vec b_2=\begin{bmatrix}1\\-1\end{bmatrix}$ and let $X=[\vec b_1|\vec b_2]$ be the matrix whose columns are $\vec b_1$ and $\vec b_2$.

61.1 Compute $[\vec e_1]_{\mathcal{B}}$ and $[\vec e_2]_{\mathcal{B}}$.

$[\vec e_1]_{\mathcal{B}}=\frac{1}{2}\begin{bmatrix}1\\1\end{bmatrix}$ and $[\vec e_2]_{\mathcal{B}}=\frac{1}{2}\begin{bmatrix}1\\-1\end{bmatrix}$.

This is because $\vec e_1=\frac{1}{2}(\vec b_1+\vec b_2)$ and $\vec e_2=\frac{1}{2}(\vec b_1-\vec b_2)$

61.2 Compute $X[\vec e_1]_{\mathcal{B}}$ and $X[\vec e_2]_{\mathcal{B}}$. What do you notice?

$X[\vec e_1]_{\mathcal{B}}=\frac{1}{2}\begin{bmatrix}1&1\\1&-1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1\\1\end{bmatrix}+\frac{1}{2}\begin{bmatrix}1\\-1\end{bmatrix}=\frac{1}{2}\vec b_1+\frac{1}{2}\vec b_2=\begin{bmatrix}1\\0\end{bmatrix}$ and

$X[\vec e_2]_{\mathcal{B}}=\frac{1}{2}\begin{bmatrix}1&1\\1&-1\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1\\1\end{bmatrix}-\frac{1}{2}\begin{bmatrix}1\\-1\end{bmatrix}=\frac{1}{2}\vec b_1-\frac{1}{2}\vec b_2=\begin{bmatrix}0\\1\end{bmatrix}$

We notice that multiplying by $X$ turns the representations of these two vectors in the basis $\mathcal{B}$ into representations in the standard basis.

61.3 Find the matrix $X^{-1}$. How does $X^{-1}$ relate to change of basis?

$X^{-1}=\frac{1}{2}\begin{bmatrix}1&1\\1&-1\end{bmatrix}$.

$X^{-1}$ should undo what $X$ does. In the previous part we saw that $X$ takes vectors represented in $\mathcal{B}$ and represents them in the standard basis. So $X^{-1}$ should do the reverse, and take vectors represented in the standard basis and represent them in the basis $\mathcal{B}$.

Lesson 26

# Applications of Inverses II, Change of Basis

## Textbook

Section 4.4

## Objectives

- Create and use change-of-basis matrices.
- Describe in words what a change-of-basis matrix and its inverse does.
- Given a linear transformation, explain how to get from a matrix representation in a basis $\mathcal{A}$ to a matrix representation in a basis $\mathcal{B}$.

## Motivation

Change of basis is one of *the* big ideas in linear algebra. The universe has no privileged basis. Some bases are nicer than others, they may be orthonormal, for example, but other than that, all bases are created equal. Therefore, we must understand the relationship *between* bases. We need facility in converting from one basis to another. And, lo and behold, converting between representations in a basis is a linear transformation!

Just as coordinates represent vectors with respect to a basis, matrices represent linear transformations with respect to a basis. And, just as the same vector can have two different representations in terms of coordinates, the same transformation can have different representations in terms of a basis. If two matrices represent the same transformation, we call them *similar*. The question of "given two matrices, how can you tell if they're similar?" is answered in the next course, MAT224, with the study of Jordan forms.

62 Let $\mathcal{E}=\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n\}$ be the standard basis for $\mathbb{R}^n$. Given a basis $\mathcal{B}=\{\vec{b}_1,\vec{b}_2,\ldots,\vec{b}_n\}$ for $\mathbb{R}^n$, the matrix $X=[\vec{b}_1|\vec{b}_2|\cdots|\vec{b}_n]$ converts vectors from the $\mathcal{B}$ basis into the standard basis. In other words,

$$X[\vec{v}]_{\mathcal{B}}=[\vec{v}]_{\mathcal{E}}.$$

62.1 Should $X^{-1}$ exist? Explain.

Yes. $X$ converts vectors from the $\mathcal{B}$ basis to the standard basis, and this process can be undone. $X^{-1}$ is the matrix that does this.

62.2 Consider the equation

$$X^{-1}[\vec{v}]_?=[\vec{v}]_?.$$

Can you fill in the "?" symbols so that the equation makes sense?

$X^{-1}[\vec{v}]_{\mathcal{E}}=[\vec{v}]_{\mathcal{B}}$.

As we said in the previous part $X^{-1}$ should undo what $X$ does, meaning it should convert vectors from the standard basis into the $\mathcal{B}$ basis.

62.3 What is $[\vec{b}_1]_{\mathcal{B}}$? How about $[\vec{b}_2]_{\mathcal{B}}$? Can you generalize to $[\vec{b}_i]_{\mathcal{B}}$?

$[\vec{b}_1]_{\mathcal{B}}=\begin{bmatrix}1\\0\\0\\\vdots\\0\end{bmatrix}$, and $[\vec{b}_2]_{\mathcal{B}}=\begin{bmatrix}0\\1\\0\\\vdots\\0\end{bmatrix}$, where each of these vectors have $n$ coordinates.

In general, $[\vec{b}_i]_{\mathcal{B}}$ should be the column vector with zeroes in all coordinates except for a 1 in the $i^{\text{th}}$ coordinate.

**Inverses and change of basis in arbitrary dimensions.**

The goal of this problem is to

- Recognize $[\vec{b}_1|\cdots|\vec{b}_n]$ as a change-of-basis matrix.
- Explain why changing basis is an invertible operation.
- Explain how the representation of the standard basis vectors as columns of 0's and one 1 is a result of representing a vector in its own basis and not something special about the standard basis.

**Notes/Misconceptions**

- This problem abstracts the previous and should come easily.
- In part 1, The existence of $X^{-1}$ can be argued by arguing that change of basis is invertible, or that $X$ has $n$ linearly independent columns and so is a square matrix with full rank. Ideally, both these arguments come out.
- In part 3, many students will assert "$[\vec{b}_1]_{\mathcal{B}}=\vec{e}_1$". We are often loose with our notation, writing $\begin{bmatrix}1\\0\end{bmatrix}$ when we mean $\begin{bmatrix}1\\0\end{bmatrix}_{\mathcal{E}}$. However, in this case, $[\vec{b}_1]_{\mathcal{B}}$ is explicitly a list of numbers and we're not allowed to be sloppy with notation.
- A student might ask how $[\vec{b}_1|\cdots|\vec{b}_n]$ is a matrix if the $\vec{b}_i$ are vectors and not lists of numbers. In this case, we're sloppy and actually mean $\Big[[\vec{b}_1]_{\mathcal{E}}|\cdots|[\vec{b}_n]_{\mathcal{E}}\Big]$. However, that is notational overload for most students. Don't bring it up unless they do.

63 Let $\vec{c}_1=\begin{bmatrix}2\\1\end{bmatrix}_{\mathcal{E}}$, $\vec{c}_2=\begin{bmatrix}5\\3\end{bmatrix}_{\mathcal{E}}$, $\mathcal{C}=\{\vec{c}_1,\vec{c}_2\}$, and $A=\begin{bmatrix}2&5\\1&3\end{bmatrix}$. Note that $A^{-1}=\begin{bmatrix}3&-5\\-1&2\end{bmatrix}$ and that $A$ changes vectors from the $\mathcal{C}$ basis to the standard basis and $A^{-1}$ changes vectors from the standard basis to the $\mathcal{C}$ basis.

63.1 Compute $[\vec{c}_1]_{\mathcal{C}}$ and $[\vec{c}_2]_{\mathcal{C}}$. $[\vec{c}_1]_{\mathcal{C}}=\begin{bmatrix}1\\0\end{bmatrix}$ and $[\vec{c}_2]_{\mathcal{C}}=\begin{bmatrix}0\\1\end{bmatrix}$.

Let $T:\mathbb{R}^2\to\mathbb{R}^2$ be the linear transformation that stretches in the $\vec{c}_1$ direction by a factor of 2 and doesn't stretch in the $\vec{c}_2$ direction at all.

63.2 Compute $T\begin{bmatrix}2\\1\end{bmatrix}_{\mathcal{E}}$ and $T\begin{bmatrix}5\\3\end{bmatrix}_{\mathcal{E}}$. $T\begin{bmatrix}2\\1\end{bmatrix}_{\mathcal{E}}=T\vec{c}_1=2\vec{c}_1=\begin{bmatrix}4\\2\end{bmatrix}_{\mathcal{E}}$ and $T\begin{bmatrix}5\\3\end{bmatrix}_{\mathcal{E}}=T\vec{c}_2=\vec{c}_2=\begin{bmatrix}5\\3\end{bmatrix}_{\mathcal{E}}$.

63.3 Compute $[T\vec{c}_1]_{\mathcal{C}}$ and $[T\vec{c}_2]_{\mathcal{C}}$. $[T\vec{c}_1]_{\mathcal{C}}=[2\vec{c}_1]_{\mathcal{C}}=\begin{bmatrix}2\\0\end{bmatrix}$ and $[T\vec{c}_2]_{\mathcal{C}}=[\vec{c}_2]_{\mathcal{C}}=\begin{bmatrix}0\\1\end{bmatrix}$.

63.4 Compute the result of $T\begin{bmatrix}\alpha\\\beta\end{bmatrix}_{\mathcal{C}}$ and express the result in the $\mathcal{C}$ basis (i.e., as a vector of the form $\begin{bmatrix}?\\?\end{bmatrix}_{\mathcal{C}}$).

$T\begin{bmatrix}\alpha\\\beta\end{bmatrix}_{\mathcal{C}}=\begin{bmatrix}2\alpha\\\beta\end{bmatrix}_{\mathcal{C}}$.

If $\vec{v}$ is a vector such that $[\vec{v}]_{\mathcal{C}}=\begin{bmatrix}\alpha\\\beta\end{bmatrix}$, then $\vec{v}=\alpha\vec{c}_1+\beta\vec{c}_2$. Since $T$ is linear, we can then compute

$$T\vec{v}=T(\alpha\vec{c}_1+\beta\vec{c}_2)=\alpha T(\vec{c}_1)+\beta T(\vec{c}_2)=2\alpha\vec{c}_1+\beta\vec{c}_2=\begin{bmatrix}2\alpha\\\beta\end{bmatrix}_{\mathcal{C}}.$$

63.5 Find $[T]_{\mathcal{C}}$, the matrix for $T$ in the $\mathcal{C}$ basis.

$[T]_{\mathcal{C}}=\begin{bmatrix}2&0\\0&1\end{bmatrix}$.

From the results of the previous parts, we know that we must have $[T]_{\mathcal{C}}\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}2\\0\end{bmatrix}$ and $[T]_{\mathcal{C}}\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}0\\1\end{bmatrix}$, so these must be the first and second columns of $[T]_{\mathcal{C}}$, respectively.

63.6 Find $[T]_{\mathcal{E}}$, the matrix for $T$ in the standard basis.

$[T]_{\mathcal{E}}=\begin{bmatrix}7&-10\\3&-4\end{bmatrix}$

There are two methods to determine this.

Method 1: Since $\vec{e}_1=3\vec{c}_1-\vec{c}_2$ and $\vec{e}_2=-5\vec{c}_1+2\vec{c}_2$, we compute

$$[T\vec{e}_1]_{\mathcal{E}}=[T(3\vec{c}_1-\vec{c}_2)]_{\mathcal{E}}=3[T(\vec{c}_1)]_{\mathcal{E}}-[T(\vec{c}_2)]_{\mathcal{E}}=3\begin{bmatrix}4\\2\end{bmatrix}-\begin{bmatrix}5\\3\end{bmatrix}=\begin{bmatrix}7\\3\end{bmatrix}$$

and

$$[T\vec{e}_2]_{\mathcal{E}}=[T(-5\vec{c}_1+2\vec{c}_2)]_{\mathcal{E}}=-5[T(\vec{c}_1)]_{\mathcal{E}}+2[T(\vec{c}_2)]_{\mathcal{E}}=-5\begin{bmatrix}4\\2\end{bmatrix}+2\begin{bmatrix}5\\3\end{bmatrix}=\begin{bmatrix}-10\\-4\end{bmatrix}.$$

These two vectors are the respective columns of $[T]_{\mathcal{E}}$, as usual.

Method 2: Since $A$ changes vectors from the $\mathcal{C}$ basis to the standard basis and $A^{-1}$ changes vectors from the standard basis to the $\mathcal{C}$ basis, we know $[T]_{\mathcal{E}}=A[T]_{\mathcal{C}}A^{-1}$. Using $[T]_{\mathcal{C}}$ from the previous part, we compute

$$[T]_{\mathcal{E}}=\begin{bmatrix}2&5\\1&3\end{bmatrix}\begin{bmatrix}2&0\\0&1\end{bmatrix}\begin{bmatrix}3&-5\\-1&2\end{bmatrix}=\begin{bmatrix}7&-10\\3&-4\end{bmatrix}.$$

**Representations of transformations.**

The goal of this problem is to

- Represent a transformation as a matrix in different bases.
- Recognize that some bases give *nicer* matrix representations than others.
- Connect the definition of similar matrices to change-of-basis.

**Notes/Misconceptions**

- This problem is coming back again after we do eigenvectors.
- Since we have vectors in multiple bases and transformations in multiple bases, it's especially important to be consistent with vector/list of numbers notation.
- Students have been primed for this problem for a long time. It won't be *so* hard, but they'll need the instructors help to connect the pieces.
- Part 6 can be solved in two ways: (i) find $[T]_{\mathcal{E}}$ by "calibrating" using vectors written in the $\mathcal{E}$ basis, and (ii) using $[T]_{\mathcal{C}}$ and the change of basis matrices $X$ and $X^{-1}$. Make sure both of these come out. You'll have more wow-factor if they come out in order (i) (the one they already know) then (ii) (the one that gives a new perspective).

DEFINITION

**Similar Matrices**

A matrices $A$ and $B$ are called ***similar matrices***, denoted $A\sim B$, if $A$ and $B$ represent the same linear transformation but in possibly different bases. Equivalently, $A\sim B$ if there is an invertible matrix $X$ so that

$$A=XBX^{-1}.$$

Lesson 27

# Determinants

## Textbook

Section 5.4

## Objectives

- Define the determinant as an oriented volume.
- Given a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$, compute the determinant from the definition.
- Correctly extend the definition of orientation of a basis to the sign of the determinant.

## Motivation

Determinants are geometric and combinatoric objects. We leave the combinatoric perspective for another course and focus only on the geometric perspective. We do this because:

- Seeing determinants as change of volumes explains the use of the Jacobian in multivariable calculus.
- Determinants of linear transformations, by definition, don't depend on a choice of basis.
- The product rule for determinants is a consequence of composition of functions.
- $\det(T) = 0 \implies T$ is not invertible has a straightforward explanation: If the volume becomes zero, a dimension was "crushed" and therefore $T$ has a non-trivial null space.

The disadvantage of the purely geometric approach is that $\det(A) = \det(A^T)$ becomes a mysterious theorem (though this can be explained with elementary matrices).

**Notes/Misconceptions**

- From the online homework, students already know how to compute $2 \times 2$ and $3 \times 3$ determinants.
- At this point in the class, everything is coming together. In particular, orientation of a basis, images of sets, null spaces, and invertibility.

**Unit $n$-cube**

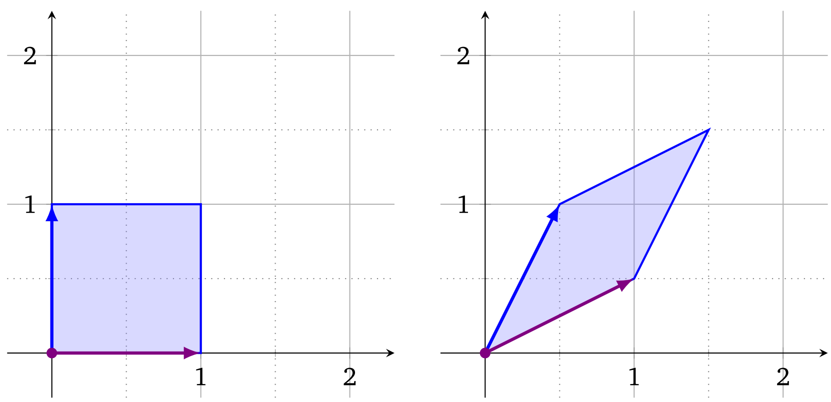
DEFINITION

The ***unit $n$-cube*** is the $n$-dimensional cube with sides given by the standard basis vectors and lower-left corner located at the origin. That is

$$C_n = \left\{ \vec{x} \in \mathbb{R}^n : \vec{x} = \sum_{i=1}^{n} \alpha_i \vec{e}_i \text{ for some } \alpha_1, \ldots, \alpha_n \in [0,1] \right\} = [0,1]^n.$$

The sides of the unit $n$-cube are always length 1 and its volume is always 1.

64 The picture shows what the linear transformation $T$ does to the unit square (i.e., the unit 2-cube).




**Volumes of images.**

The goal of this problem is to
- Apply the definitions of *unit n-cube* and *image of a set*.
- Use tools from outside of linear algebra class to compute the area of a polygon.
- Be comfortable using the word "volume" in $\mathbb{R}^2$.

**Notes/Misconceptions**
- Most of this problem is review.
- Students will struggle computing the area of the rhombus. By dividing the figure into right triangles you can compute its area quickly.

64.1 What is $T\begin{bmatrix}1\\0\end{bmatrix}$, $T\begin{bmatrix}0\\1\end{bmatrix}$, $T\begin{bmatrix}1\\1\end{bmatrix}$?

$T\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}1\\\frac{1}{2}\end{bmatrix}$, $T\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}\frac{1}{2}\\1\end{bmatrix}$, and $T\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}\frac{3}{2}\\\frac{3}{2}\end{bmatrix}$?

We can see first two directly in the picture.

Using the linearity of $T$, we can compute

$$T\begin{bmatrix}1\\1\end{bmatrix} = T\left(\begin{bmatrix}1\\0\end{bmatrix}+\begin{bmatrix}0\\1\end{bmatrix}\right) = T\begin{bmatrix}1\\0\end{bmatrix}+T\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}1\\\frac{1}{2}\end{bmatrix}+\begin{bmatrix}\frac{1}{2}\\1\end{bmatrix} = \begin{bmatrix}\frac{3}{2}\\\frac{3}{2}\end{bmatrix}.$$

**Notes/Misconceptions**
- This is the first time orientation has shown up since we saw it last. Spend some time explaining how orientation of a basis relates to orientation of $T(C_n)$.

  It's helpful to draw an analogy with "area under the curve" and "integral" from calculus. Rectangles with positive integral have sides in the $\vec{e}_1, \vec{e}_2$ directions (right-handed basis), and rectangles with negative integral have sides in the $\vec{e}_1, -\vec{e}_2$ directions (left-handed basis).

64.2 Write down a matrix for $T$.

The matrix for $T$ in the standard basis is $\begin{bmatrix}1 & \frac{1}{2}\\ \frac{1}{2} & 1\end{bmatrix}$.

64.3 What is the volume of the image of the unit square (i.e., the volume of $T(C_2)$)? You may use trigonometry.

The volume is $\frac{3}{4}$.

**Determinant**

DEF

The ***determinant*** of a linear transformation $X : \mathbb{R}^n \to \mathbb{R}^n$ is the oriented volume of the image of the unit $n$-cube. The determinant of a square matrix is the determinant of its induced transformation.

65 We know the following about the transformation $A$:

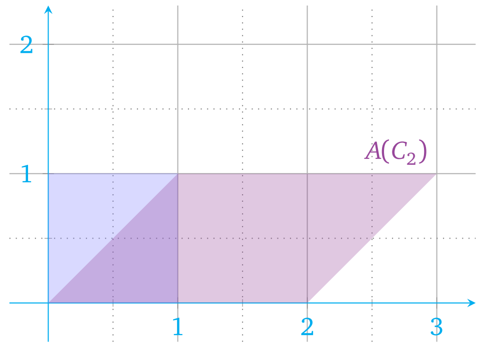
$$A\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}2\\0\end{bmatrix} \quad \text{and} \quad A\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix}.$$

**Apply the definition of determinant.**

The goal of this problem is to
- Compute a determinant from the definition.
- Practice finding the area of a parallelogram.

**Notes/Misconceptions**
- This example is very hands-on with minimal complications.

65.1 Draw $C_2$ and $A(C_2)$, the image of the unit square under $A$.




65.2 Compute the area of $A(C_2)$. The area of this parallelogram is 2.

65.3 Compute $\det(A)$.

$\det(A) = 2$.

The parallelogram with sides $\begin{bmatrix}2\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$ is positively oriented, so $\det(A) = +2$.

---

66 Suppose $R$ is a rotation counter-clockwise by 30°.

**Apply the definition of determinant.**

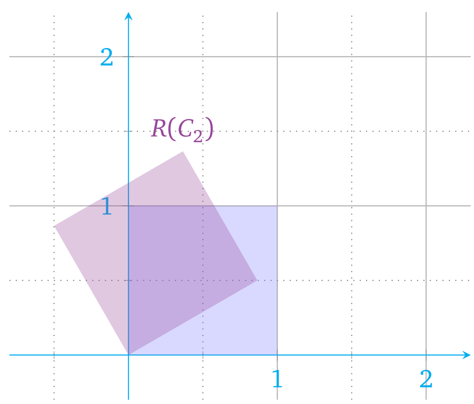
The goal of this problem is to

- Compute a determinant from the definition by applying geometric reasoning.

**Notes/Misconceptions**

- This example is difficult to "compute" without thinking about it.

66.1 Draw $C_2$ and $R(C_2)$.




66.2 Compute the area of $R(C_2)$.

The area is 1.

$R$ rotates the entire unit square, which does not change its area.

66.3 Compute $\det(R)$.

Since $R$ preserves orientation, $\det(R)$ must be positive. Since $R$ does not change the area of the unit square, $\det(R) = +1$.

---

67 We know the following about the transformation $F$:

$$F\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix} \quad \text{and} \quad F\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}.$$

**Apply the definition of determinant.**

The goal of this problem is to

- Compute a determinant from the definition when orientation is reversed.

67.1 What is $\det(F)$?

$\det(F) = -1$.

$F$ does not change the area of the unit square, but reverses its orientation, so $\det(F) = -1$.

**Volume Theorem I** (THM)

For a square matrix $M$, $\det(M)$ is the oriented volume of the parallelepiped ($n$-dimensional parallelogram) given by the column vectors of $M$.

**Volume Theorem II** (THM)

For a square matrix $M$, $\det(M)$ is the oriented volume of the parallelepiped ($n$-dimensional parallelogram) given by the row vectors of $M$.

**Relate determinants of transformations and matrices.**

The goal of this problem is to

- Relate the image of the unit cube under a transformation $T$ to the columns of $T$'s matrix representation.
- Relate the determinant of a matrix and its transpose.

**Notes/Misconceptions**

- The explanation for part 1 harks back to the relationship between the range of a transformation and the column space of its matrix.

68 68.1 Explain Volume Theorem I using the definition of determinant.

If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a transformation with standard matrix $M = [\vec{c}_1|\cdots|\vec{c}_n]$, then $T(\vec{e}_i)$ (represented in the standard basis) will be $c_i$, the $i$th column of $M$. The image of the unit cube will be a parallelepiped with sides $T(\vec{e}_1) = \vec{c}_1, \ldots, T(\vec{e}_n) = \vec{c}_n$, and so $\det(T)$ will be the oriented volume of the parallelepiped with sides given by $\vec{c}_1, \ldots, \vec{c}_n$.

68.2 Based on Volume Theorems I and II, how should $\det(M)$ and $\det(M^T)$ relate for a square matrix $M$?

$\det(M) = \det(M^T)$. Since the transpose switches columns for rows, this is an immediate consequence of Volume Theorems I and II.

Lesson 28

# Determinants and Compositions

## Textbook

Section 5.4, 5.2

## Objectives

- Compute the volume/area of a figure after multiple linear transformations have been applied.
- Explain the multiplicative property $\det(A \circ B) = \det(A)\det(B)$ in terms of function composition and change of area.
- Compute the determinant of each type of elementary matrix.
- Quickly compute the determinant of triangular matrices.
- Compute the determinant of an arbitrary matrix by decomposing it into the product of elementary matrices.
- Explain the relationship $\det(X^{-1}) = 1/\det(X)$.

## Motivation

Determinants, especially higher-dimensional ones, are difficult to compute from the definition. However, since $\det(A \circ B) = \det(A)\det(B)$, we can compute the determinant of complicated transformations/matrices by decomposing them into simple ones.

Elementary matrices come in three types, one for each type of elementary operation. Two of the elementary matrices change the determinant by a multiplicative factor, and the last one (the workhorse of row reduction) doesn't change the determinant at all. Given a decomposition of a matrix as the product of elementary matrices, it is easy to compute the determinant. However, if we just want to compute the determinant, it's easier than that. By row reducing while keeping track of row swaps and when we normalize rows, we can efficiently compute the determinant of any matrix.

Elementary matrices also nail down two theorems:

(1) $\det(A) = 0 \iff A$ is not invertible.

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

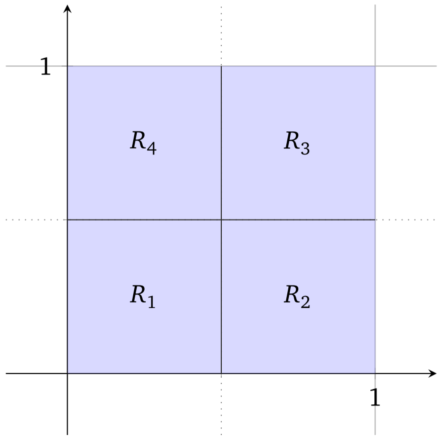
The first theorem is proved because, for a matrix $A$, we have $A = X\,\mathrm{rref}(A)$ where $X$ is the product of elementary matrices. If $\mathrm{rref}(A) \neq I$, the matrix cannot be invertible and $\det(A) = 0$. If $\mathrm{rref}(A) = I$, then $\det(A) \neq 0$.

The second theorem follows by noticing that if $E$ is an elementary matrix, $E^T$ is an elementary matrix of the same type and it changes volume by the same amount. Thus, by decomposing $A$ as a product of elementary matrices (assume it is invertible), we see $\det(A) = \det(A^T)$.

**Notes/Misconceptions**

- Students will have memorized the formula for $2 \times 2$ and $3 \times 3$ determinants from their homework.

69




Let $R = R_1 \cup R_2 \cup R_3 \cup R_4$. You know the following about the linear transformations $M$, $T$, and $S$.

$$M\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ y \end{bmatrix}$$

$T : \mathbb{R}^2 \to \mathbb{R}^2$ has determinant 2

$S : \mathbb{R}^2 \to \mathbb{R}^2$ has determinant 3

69.1 Find the volumes (areas) of $R_1$, $R_2$, $R_3$, $R_4$, and $R$.

The volumes of $R_1$, $R_2$, $R_3$, and $R_4$ are $\frac{1}{4}$. The volume of $R$ is 1.

69.2 Compute the oriented volume of $M(R_1)$, $M(R_2)$, and $M(R)$.

The oriented volumes of $M(R_1)$ and $M(R_2)$ are $\frac{1}{2}$ and the oriented volume of $M(R)$ is 2.

69.3 Do you have enough information to compute the oriented volume of $T(R_2)$? What about the oriented volume of $T(R + \{\vec{e}_2\})$?

Yes. The oriented volume of $T(R_2) = \frac{1}{2}$ and the oriented volume of $T(R + \{\vec{e}_2\}) = 2$.

We don't have enough information to determine what $T(R_2)$ looks like, but we do know (i) $T(R_2)$ will be a parallelogram, (ii) $T(R)$ has oriented volume 2, and (iii) $T(R)$ is made of four translated copies of $T(R_2)$. From this we deduce that the oriented volume of $T(R_2) = \frac{1}{4}\big(\text{oriented volume of } T(R)\big) = (\frac{1}{4})(2)$.

To find the oriented volume of $T(R + \{\vec{e}_2\})$, we use linearity to observe

$$T(R + \{\vec{e}_2\}) = T(R) + T(\{\vec{e}_2\}) = T(R) + \{T\vec{e}_2\}.$$

This shows that $T(R + \{\vec{e}_2\})$ is just a translation of $T(R)$ and therefore has the same oriented volume.

69.4 What is the oriented volume of $S \circ T(R)$? What is $\det(S \circ T)$?

They are both equal to 6.

$S \circ T(R) = S(T(R))$. We already know $T(R)$ has a volume of 2, and so $S(T(R))$ has a volume of 6, since $S$ scales the volumes of all regions by 3. The oriented volume of $S \circ T(R)$ is the determinant of $S \circ T$ by definition.

**Determinants and areas.**

The goal of this problem is to

- Use determinants to compute areas/volumes of images of arbitrary sets.
- See determinants as a "change of area/volume" *factor*.
- Explain the multiplicative property of determinants in terms of area/volume changes.

**Notes/Misconceptions**

- So far, determinants have only been used to compute the volume of $T(C_n)$. However, they can be used to compute the volume of $T(X)$ given that you know the volume of $X$.

  The argument comes from calculus: if we can show it works for translated and scaled copies of $C_n$, it works for any reasonable shape.

  The pieces of this argument are there, but it isn't put together. Put this together for the students in part 4.
- When concluding part 4, make sure the formula $\det(T \circ M) = \det(T)\det(M)$ comes up.

70

- $E_f$ is $I_{3\times 3}$ with the first two rows swapped.
- $E_m$ is $I_{3\times 3}$ with the third row multiplied by 6.
- $E_a$ is $I_{3\times 3}$ with $R_1 \mapsto R_1 + 2R_2$ applied.

70.1 What is $\det(E_f)$?

$\det(E_f) = -1$.

$\det(I_{3\times 3}) = 1$, and swapping one pair of rows of a matrix changes the sign of its determinant.

70.2 What is $\det(E_m)$?

$\det(E_m) = 6$.

Multiplying one row of a matrix by a constant multiplies its determinant by the same constant.

70.3 What is $\det(E_a)$?

**Determinants of elementary matrices.**

The goal of this problem is to

- Memorize the determinant of each type of elementary matrix.
- Justify why the determinant of an elementary matrix of type "add a multiple of one row to another" is always 1.
- Outline a method to compute determinants of arbitrary matrices.

**Notes/Misconceptions**

- Part 1 and 2 will be easy.
- Part 3 makes sense if it is explained right. Carefully plan out what pictures you will use to explain this. It's helpful to draw the parallelepiped with the $z$-axis sticking straight out of the board, effectively reducing it to the 2d case.
- The punchline from part 6 is that you can now compute the determinant of any matrix by decomposing it into the product of elementary matrices. If you haven't explicitly talked about this decomposition before, now is the time.

$\det(E_a) = 1$.

Adding a multiple of one row of a matrix to another row has no effect on its determinant.

70.4 What is $\det(E_f E_m)$? $\det(E_f E_m) = \det(E_f)\det(E_m) = (-1)(6) = -6$.

70.5 What is $\det(4I_{3\times 3})$? $\det(4I_{3\times 3}) = 4^3 = 64$.

70.6 What is $\det(W)$ where $W = E_f E_a E_f E_m E_m$?

$\det(W) = \det(E_f)\det(E_a)\det(E_f)\det(E_m)\det(E_m) = (-1)(1)(-1)(6)(6) = 36$.

---

71

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

71.1 What is $\det(U)$?

$\det(U) = -12$.

71.2 $V$ is a square matrix and $\mathrm{rref}(V)$ has a row of zeros. What is $\det(V)$?

$\det(V) = 0$.

**Reasoning about determinants via elementary matrices.**

The goal of this problem is to

- Develop a shortcut for computing determinants of triangular matrices.
- Reason about the determinant of a matrix when given its reduced row echelon form.

**Notes/Misconceptions**

- In part 1, give the hint, "imagine that you decomposed $U$ into the product of elementary matrices. What types of elementary matrices would you need?"
- For part 1, a good way to to explain it is by explaining that reducing $U$ to a diagonal matrix won't affect the determinant. Now reason geometrically about a diagonal matrix.
- For part 2, argue algebraically: there is a matrix $R$ so $V = R\,\mathrm{rref}(V)$. Therefore $\det(V) = \det(R\,\mathrm{rref}(V)) = \det(V)\det(\mathrm{rref}(V))$, but $\mathrm{rref}(V)$ has a zero on the diagonal! The geometric perspective will come in the next problem.

---

72

72.1 $V$ is a square matrix whose columns are linearly dependent. What is $\det(V)$?

$\det(V) = 0$.

72.2 $P$ is projection onto $\mathrm{span}\left\{\begin{bmatrix} -1 \\ -1 \end{bmatrix}\right\}$. What is $\det(P)$?

$\det(P) = 0$.

The image of the unit square under $P$ is a line segment, which has zero volume.

**Determinants of singular matrices.**

The goal of this problem is to

- Reason geometrically about why a transformation that isn't one-to-one has a zero determinant.

**Notes/Misconceptions**

- For part 1, emphasize the geometry first: if the columns are linearly dependent they will specify a "flattened" parallelepiped, which must have zero volume.

  Then, connect this to the previous question where $\mathrm{rref}(V)$ had a row of zeros.
- For part 2, reemphasize the "flattening" picture. You can also mention that if the transformation is not invertible, the image of $C_n$ will be flattened.

---

73

Suppose you know $\det(X) = 4$.

73.1 What is $\det(X^{-1})$?

$\det(X^{-1}) = \frac{1}{4}$.

We know that $XX^{-1} = I$. Therefore we must have that $\det(XX^{-1}) = \det(X)\det(X^{-1}) = \det(I) = 1$, and so $\det(X^{-1}) = \frac{1}{4}$.

73.2 Derive a relationship between $\det(Y)$ and $\det(Y^{-1})$ for an arbitrary matrix $Y$.

$\det(Y^{-1}) = \frac{1}{\det(Y)}$.

Using the same reasoning as the previous part, we know that $YY^{-1} = I$. Therefore we must have $\det(Y)\det(Y^{-1}) = \det(YY^{-1}) = \det(I) = 1$, and so $\det(Y^{-1}) = \frac{1}{\det(Y)}$.

73.3 Suppose $Y$ is not invertible. What is $\det(Y)$?

$\det(Y) = 0$.

If $Y$ is not invertible, it has linearly dependent columns. Therefore the parallelepiped formed by the columns of $Y$ will be "flattened" and have zero volume.

This is consistent with our previous findings. For a square matrix $Y$, $\det(Y^{-1}) = \frac{1}{\det(Y)}$. This formula always works, except when $\det(Y) = 0$.

**Determinants and invertibility.**

The goal of this problem is to

- Produce the determinant of $X^{-1}$ give the determinant of $X$.
- Explain why a transformation $T$ is invertible if and only if $\det(T) \neq 0$.

Lesson 29

# Eigenstuff I

## Textbook

Section 6.1

## Objectives

- Compute linear transformations presented in different bases.
- Recognize a preferred basis (the eigen basis) for transformations which can be described as stretching in particular directions.

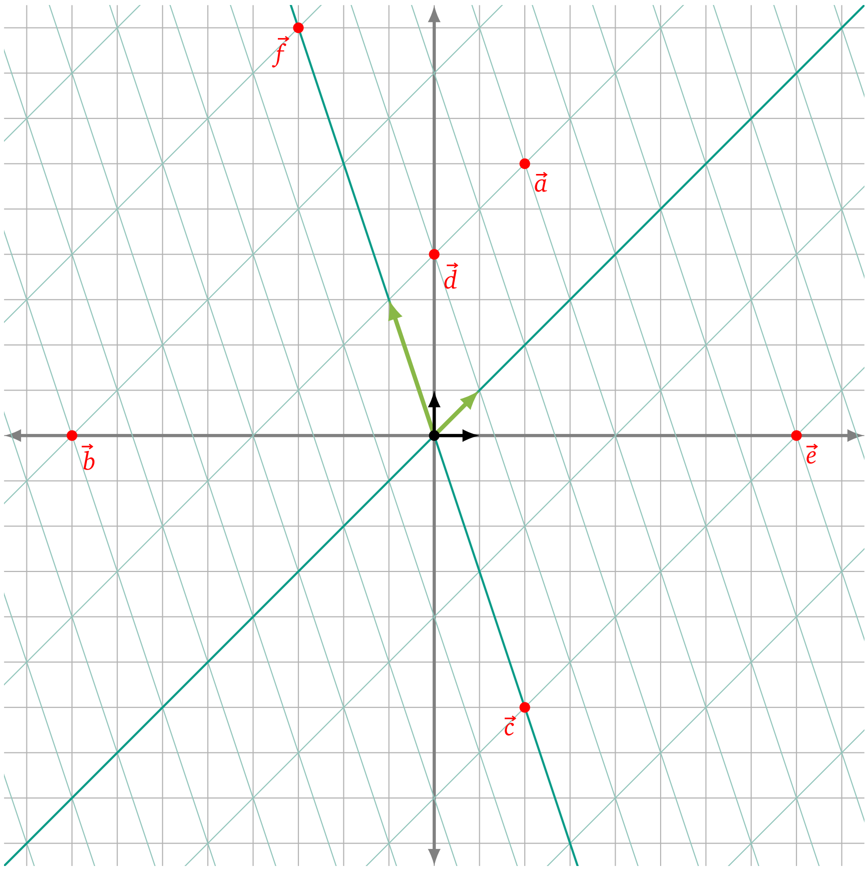
## Motivation

Eigenvectors and diagonalization are the capstones of this course. However, they are a major stumbling block for students because they can easily get lost in the algebraic procedure of finding eigenvectors/values and not have any grounding of their meaning.

We will start out with a very concrete example where we see that describing a transformation in terms of stretch directions gives us a preferred basis. We will later realize this basis as the eigen basis.

# Task 3.1: The Green and the Black

Consider the following two bases for $\mathbb{R}^2$: the green basis $\mathcal{G} = \{\vec{g}_1, \vec{g}_2\}$ and the black basis $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$.




1. Write each point above in both the green and the black bases.
2. Find a change-of-basis matrix $X$ that converts vectors from a green basis representation to a black basis representation. Find another matrix $Y$ that converts vectors from a black basis representation to a green basis representation.
3. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation that stretches in the $y = -3x$ direction by a factor of 2 and leaves vectors in the $y = x$ direction fixed.

   Describe what happens to the vectors $\vec{u}$, $\vec{v}$, and $\vec{w}$ when $T$ is applied given that

$$[\vec{u}]_{\mathcal{G}} = \begin{bmatrix} 6 \\ 1 \end{bmatrix} \qquad [\vec{v}]_{\mathcal{G}} = \begin{bmatrix} 4 \\ -3 \end{bmatrix} \qquad [\vec{w}]_{\mathcal{B}} = \begin{bmatrix} -8 \\ -7 \end{bmatrix}.$$

4. When working with the transformation $T$, which basis do you prefer vectors be represented in?

**Notes/Misconceptions**

This problem is very similar to a previous problem except that the vectors are more complicated and most of the information is presented graphically instead of numerically. It is also longer and has less scaffolding.

- In part 2, of course $Y = X^{-1}$. Some students will compute it that way and others will find $Y$ from first principles.
- In part 4, we have an opportunity to practice communicating a value judgement. Ask students to prepare a 20 second explanation that they might give in a business meeting where an excavation team was trying to decide what basis to record measurements in.
- Though the problem doesn't explicitly ask it, you have the opportunity to discuss inverse matrices (in terms of change of basis) and similar matrices (giving different, but equivalent, representations of $T$).

Lesson 30

# Eigenstuff II

## Textbook

Section 6.1

## Objectives

- Define eigenvector/value.
- Compute eigenvectors/values given a geometric description of a linear transformation.
- Argue that $A-\lambda I$ has a non-trivial null space $\iff$ $\lambda$ is an eigenvector for $A$.

## Motivation

We want students to be able to explain how the procedure for finding eigenvectors/values relates to the definition of eigenvectors/values. Therefore, we proceed slowly, exploring properties of eigenvectors/values one step at a time.

As we've already seen, if $\vec{v}$ is an eigenvector for $A$, then $A\vec{v}$ is easy to compute. It's less obvious to students that $(A+tI)\vec{v}$ is equally easy to compute. The algorithm for finding eigenvectors/values relies on (i) our ability to compute the null space of $A-\lambda I$, and (ii) our ability to algorithmically determine if $A-\lambda I$ is invertible. In this lesson we establish both needs, but don't yet practice an algorithm.

## Eigenvectors

**Eigenvector**

DEFINITION

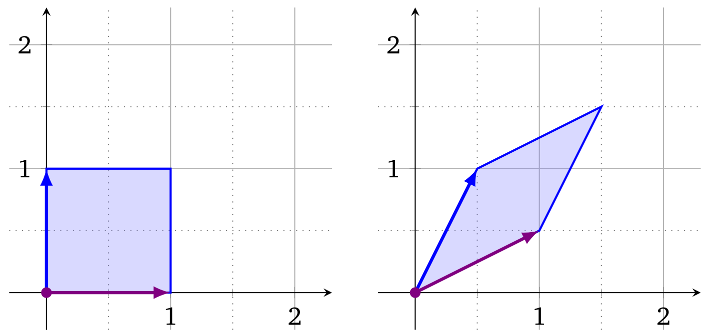
Let $X$ be a linear transformation. An ***eigenvector*** for $X$ is a non-zero vector that doesn't change directions when $X$ is applied. That is, $\vec{v} \neq \vec{0}$ is an eigenvector for $X$ if

$$X\vec{v} = \lambda\vec{v}$$

for some scalar $\lambda$. We call $\lambda$ the ***eigenvalue*** of $X$ corresponding to the eigenvector $\vec{v}$.

---

74 The picture shows what the linear transformation $T$ does to the unit square (i.e., the unit 2-cube).




74.1 Give an eigenvector for $T$. What is the eigenvalue?

$\vec{v} = \begin{bmatrix}1\\1\end{bmatrix}$ is an eigenvector for $T$, with corresponding eigenvalue $\frac{3}{2}$.

We can see from the image that $T\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}\frac{3}{2}\\\frac{3}{2}\end{bmatrix} = \frac{3}{2}\begin{bmatrix}1\\1\end{bmatrix}$.

74.2 Can you give another?

Any scalar multiple of $\begin{bmatrix}1\\1\end{bmatrix}$ is also an eigenvector for $T$.

For any scalar $\alpha$, we have $T\left(\alpha\begin{bmatrix}1\\1\end{bmatrix}\right) = \alpha T\begin{bmatrix}1\\1\end{bmatrix} = \alpha\frac{3}{2}\begin{bmatrix}1\\1\end{bmatrix}$, meaning $\alpha\begin{bmatrix}1\\1\end{bmatrix}$ is an eigenvector for $T$ with eigenvalue $\frac{3}{2}\alpha$.

More interestingly, since $T\begin{bmatrix}-1\\1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}-1\\1\end{bmatrix}$, $\begin{bmatrix}-1\\1\end{bmatrix}$ is an eigenvector with corresponding eigenvalue $\frac{1}{2}$.

**Apply the definition of eigenvector/value geometrically.**

The goal of this problem is to
- Find eigenvectors/values from transformations defined geometrically.
- Produce new eigenvectors from existing ones by scaling.

**Notes/Misconceptions**

- Most students will miss the eigenvector $\begin{bmatrix}-1\\1\end{bmatrix}$, but they may succeed in finding multiple eigenvectors by scaling the obvious one. Scaling is a great idea! But, after they figure out how to produce multiple eigenvectors by scaling, ask them if they can find another linearly independent eigenvector.

---

75 For some matrix $A$,

$$A\begin{bmatrix}3\\3\\1\end{bmatrix} = \begin{bmatrix}2\\2\\2/3\end{bmatrix} \qquad \text{and} \qquad B = A - \tfrac{2}{3}I.$$

75.1 Give an eigenvector and a corresponding eigenvalue for $A$.

$\begin{bmatrix}3\\3\\1\end{bmatrix}$ is an eigenvector for $A$, with corresdponding eigenvalue $\frac{2}{3}$.

75.2 What is $B\begin{bmatrix}3\\3\\1\end{bmatrix}$?

$$B\begin{bmatrix}3\\3\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

**Apply the definition of eigenvector/value algebraically.**

The goal of this problem is to
- Identify numerically whether a vector is an eigenvector.
- Use numerical evidence to compute an eigenvalue.
- Reason about the matrix $A-\lambda I$ given $\lambda$ is an eigenvalue.

**Notes/Misconceptions**

- In part 1, students might need a hint like, "try testing the only vector that you're given information about."
- Part 2 will be surprisingly hard for students.
- Part 3 has no exact answer. Many students will claim nullity$(B) = 3$. Make sure you discuss why you can only draw a limited conclusion.

We compute

$$B\begin{bmatrix}3\\3\\1\end{bmatrix}=(A-\frac{2}{3}I)\begin{bmatrix}3\\3\\1\end{bmatrix}=A\begin{bmatrix}3\\3\\1\end{bmatrix}-\frac{2}{3}I\begin{bmatrix}3\\3\\1\end{bmatrix}=\begin{bmatrix}2\\2\\2/3\end{bmatrix}-\begin{bmatrix}2\\2\\2/3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

75.3 What is the dimension of null($B$)?

The most we can say is that nullity($B$) $\geq 1$.

We know $\begin{bmatrix}3\\3\\1\end{bmatrix}\in$ null($B$) by the previous part, and so the dimension of null($B$) is at least 1. It could be larger, but we do not have enough information to say for sure.

75.4 What is det($B$)? $\det(B)=0$.

---

76 Let $C=\begin{bmatrix}-1&2\\1&0\end{bmatrix}$ and $E_\lambda=C-\lambda I$.

76.1 For what values of $\lambda$ does $E_\lambda$ have a non-trivial null space?

$\lambda=-2$ and $\lambda=1$.

$E_\lambda$ has a non-trivial null space exactly when its determinant is zero. We compute:

$$\det(E_\lambda)=\det\left(\begin{bmatrix}-1-\lambda&2\\1&-\lambda\end{bmatrix}\right)=(-1-\lambda)(-\lambda)-(1)(2)=\lambda^2+\lambda-2=(\lambda-1)(\lambda+2).$$

This equals zero exactly when $\lambda=-2$ or $\lambda=1$.

76.2 What are the eigenvalues of $C$?

$-2$ and $1$.

The scalar $\lambda$ is an eigenvalue of $C$ if and only if $C\vec{v}=\lambda\vec{v}$ for some $\vec{v}\neq\vec{0}$. Thus

$$\vec{0}=C\vec{v}-\lambda\vec{v}=(C-\lambda I)\vec{v}=E_\lambda\vec{v},$$

and so $E_\lambda$ has a non-trivial null space if and only if $\lambda$ is an eigenvalue of $C$.

76.3 Find the eigenvectors of $C$.

$\begin{bmatrix}-2\\1\end{bmatrix}$ and $\begin{bmatrix}1\\1\end{bmatrix}$ (along with all non-zero scalar multiples of these).

We know from the previous part that finding an eigenvector with corresponding eigenvalue $-2$ amounts to finding the non-zero vectors in the null space of $E_{-2}=\begin{bmatrix}1&2\\1&2\end{bmatrix}$. Computing,

$$\mathrm{null}(E_{-2})=\mathrm{span}\left\{\begin{bmatrix}-2\\1\end{bmatrix}\right\}.$$

Similarly, the eigenvectors with corresponding eigenvalue 1 are the non-zero vectors in the null space of $E_1=\begin{bmatrix}-2&2\\1&-1\end{bmatrix}$, and we compute that $\mathrm{null}(E_1)=\mathrm{span}\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$.

**Explore the matrix $C-\lambda I$.**

The goal of this problem is to

- Relate the matrix $C-\lambda I$ to the problem of finding eigenvectors/values.
- Relate the equation $C\vec{x}=\lambda\vec{x}$ to the null space of the matrix $E_\lambda=C-\lambda I$.
- Use the determinant to determine when a parameterized family of matrices is invertible or not.
- Numerically compute eigenvalues/vectors without extra geometric information.

**Notes/Misconceptions**

- This example combines lots of linear algebra tools. Make sure to draw connections with our previous work so that students see that everything we did was for a reason.
- In part 1, students might need a little push along the lines of: "what can the determinant tell you about a null space?"
- In part 2, it's time to make the connection explicit: $C\vec{v}=\lambda\vec{v}$ if and only if $\vec{v}$ is in the null space of $C-\lambda I$. And in this case, $\lambda$ is the eigenvalue.
- In part 3, students might need a little push: "remember, you know how to compute the null space of any matrix (using row reduction)!"

Lesson 31

# Characteristic Polynomial, Diagonalization I

## Textbook

Section 6.2

## Objectives

- Define *characteristic polynomial*
- Explain how the characteristic polynomial relates to eigenvalues.
- Use the characteristic polynomial to determine if a matrix is invertible.
- Compute eigenvalues given eigenvectors.
- Compute the image of a vector written in an eigen basis.

## Motivation

We are building up to diagonalization. One of the tools along the way is the characteristic polynomial, which connects algebra (in particular the fundamental theorem of algebra) to matrices. It is a deep and mysterious connection and is not technically needed to study eigenvalues and eigenvectors[1], but characteristic polynomials are used in many fields and provide a quick way to find eigenvalues.

We are not studying characteristic polynomials outright. Instead, they are a clever object that allows us to determine when a matrix $A-\lambda I$ has a non-trivial null space. (Alas, the study of minimal polynomials must wait for a different course.) *Our end-goal is to see that the eigen basis is useful.*

**Notes/Misconceptions**

- We define $\text{char}(A) = \det(A-\lambda I)$. Some define $\text{char}(A) = \det(\lambda I - A)$.

  The up side of the second definition is that the characteristic polynomial is always monic. The up side of our definition is that $\text{char}(A)(0) = \det(A)$, and the down side is that the leading coefficient of $\text{char}(A)$ is $\pm\lambda^n$ depending on whether $n$ is even or odd.

  When writing a factored characteristic polynomial, write

  $$(\lambda_1 - x)(\lambda_2 - x)\cdots$$

  instead of

  $$\pm(x-\lambda_1)(x-\lambda_2)\cdots$$

  to avoid issues.

[1] For example, see *Linear Algebra Done Right*'s approach.

## Characteristic Polynomial

DEF

For a matrix $A$, the ***characteristic polynomial*** of $A$ is

$$\mathrm{char}(A) = \det(A - \lambda I).$$

77 Let $D = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$.

77.1 Compute $\mathrm{char}(D)$.

$\mathrm{char}(D) = (-2-\lambda)(3-\lambda)$.

We compute:

$$\det(D-\lambda I) = \det\left(\begin{bmatrix} 1-\lambda & 2 \\ 3 & -\lambda \end{bmatrix}\right) = (1-\lambda)(-\lambda)-(2)(3) = \lambda^2-\lambda-6 = (-2-\lambda)(3-\lambda).$$

77.2 Find the eigenvalues of $D$.

The eigenvalues of $D$ are $-2$ and $3$. The eigenvalues of $D$ are the roots of $\mathrm{char}(D)$.

78 Suppose $\mathrm{char}(E) = -\lambda(2-\lambda)(-3-\lambda)$ for some unknown $3 \times 3$ matrix $E$.

78.1 What are the eigenvalues of $E$?

0, 2, and $-3$.

The eigenvalues of $E$ are the roots of $\mathrm{char}(E)$.

78.2 Is $E$ invertible?

No.

Since 0 is an eigenvalue of $E$, there must be a non-zero vector $\vec{v}$ such that $E\vec{v} = 0\vec{v} = \vec{0}$. This means $\mathrm{nullity}(E) > 0$, which implies $E$ is not invertible.

78.3 What can you say about $\mathrm{nullity}(E)$, $\mathrm{nullity}(E-3I)$, $\mathrm{nullity}(E+3I)$?

$\mathrm{nullity}(E) = 1$, $\mathrm{nullity}(E-3I) = 0$, $\mathrm{nullity}(E+3I) = 1$

Notice that evaluating the characteristic polynomial at $\lambda$ give the determinant of $E - \lambda I$. From this, we can determine the invertibility of any matrix of the form $E - \lambda I$.

Since $E$ is not invertible, $\mathrm{nullity}(E) \geq 1$. Since $E-3I$ is invertible, $\mathrm{nullity}(E-3I) = 0$, and since $E+3I$ is not invertible, $\mathrm{nullity}(E+3I) \geq 1$.

To pin down the nullities of $E$ and $E+3I$ exactly takes more work. $E$ has three distinct eigenvalues and so $E$ must have three linearly independent eigenvectors $\vec{v}_0, \vec{v}_2, \vec{v}_{-3}$. Further, $\mathrm{span}\{\vec{v}_i\} \subseteq \mathrm{null}(E-\lambda_i I)$. Since $\vec{v}_0, \vec{v}_1, \vec{v}_{-3} \in \mathbb{R}^3$, it must be the case that $\mathrm{null}(E-\lambda_i I)$ is one dimensional for $i \in \{0, 2, -3\}$.

79 Consider

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \qquad \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \qquad \vec{v}_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

and notice that $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are eigenvectors for $A$. Let $T_A$ be the transformation induced by $A$.

79.1 Find the eigenvalues of $A$.

The eigenvalues of $A$ are 2, $-1$, and 1.

We compute that $A\vec{v}_1 = 2\vec{v}_1$, $A\vec{v}_2 = -\vec{v}_2$, and $A\vec{v}_3 = \vec{v}_3$, so 2, $-1$, and 1 are eigenvalues of $A$. By the last part of the previous problem, there are no other eigenvalues.

79.2 Find the characteristic polynomial of $A$.

$\mathrm{char}(A) = (2-\lambda)(1-\lambda)(-1-\lambda) = -(\lambda-2)(\lambda+1)(\lambda-1)$.

Sine we know the roots of the characteristic polynomial are the eigenvalues and we know $\mathrm{char}(A)$ is a cubic, we can immediately write down $\mathrm{char}(A) = (2-\lambda)(1-\lambda)(-1-\lambda)$ without computing a determinant.

79.3 Compute $A\vec{w}$ where $w = 2\vec{v}_1 - \vec{v}_2$.

**Notes/Misconceptions**

- $\mathrm{char}(A)$ could be defined as $\det(\lambda I - A)$ but if done this way, the constant term of the polynomial is not equal to the determinant. The downside of our definition is that $\mathrm{char}(A)$ is not always monic.

**Apply the definition of the characteristic polynomial.**

The goal of this problem is to

- Compute a characteristic polynomial by applying the definition.
- Relate the characteristic polynomial to eigenvalues.

**Notes/Misconceptions**

- This problem is largely a sanity check to see if students remember what was done in the previous problem.

**Getting information from the characteristic polynomial.**

The goal of this problem is to

- Use the characteristic polynomial to determine eigenvalues and invertibility of matrices.
- Relate the characteristic polynomial to determinants.

**Notes/Misconceptions**

- Part 2 will seem obvious to students, but only in retrospect. They will forget that the characteristic polynomial relates to the determinant!
- Part 3 requires carefully keeping track of signs. The definition of characteristic polynomial is $\det(A - \lambda I)$, not $\det(A + \lambda I)$.

  This part also requires a lot of sophistication to pin down. It's easy to argue that the appropriate nullities will be $\geq 1$ and $< 3$, but arguing that they are equal to 1 (and not 2) takes work. Don't bother completing the argument pinning down the exact nullities (but feel free to tell them the exact nullities).

$A\vec{w} = 4\vec{v}_1 + \vec{v}_2$.

Using the computations we did in the first part above, we find

$$A\vec{w} = A(2\vec{v}_1 - \vec{v}_2) = 2A\vec{v}_1 - A\vec{v}_2 = 2(2\vec{v}_1) - (-\vec{v}_2) = 4\vec{v}_1 + \vec{v}_2.$$

79.4 Compute $T_A\vec{u}$ where $\vec{u} = a\vec{v}_1 + b\vec{v}_2 + c\vec{v}_3$ for unknown scalar coefficients $a, b, c$.

$T_A\vec{u} = 2a\vec{v}_1 - b\vec{v}_2 + c\vec{v}_3$.

Using the same reasoning as the previous part, we compute

$$T_A\vec{u} = aA\vec{v}_1 + bA\vec{v}_2 + cA\vec{v}_3 = 2a\vec{v}_1 - b\vec{v}_2 + c\vec{v}_3.$$

Notice that $\mathcal{V} = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is a basis for $\mathbb{R}^3$.

79.5 If $[\vec{x}]_{\mathcal{V}} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$ is $\vec{x}$ written in the $\mathcal{V}$ basis, compute $T_A\vec{x}$ in the $\mathcal{V}$ basis.

$[T_A\vec{x}]_{\mathcal{V}} = \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}$.

If $[\vec{x}]_{\mathcal{V}} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$, then $\vec{x} = \vec{v}_1 + 3\vec{v}_2 + 4\vec{v}_3$. Using the previous part, we then have that $T_A\vec{x} = 2\vec{v}_1 - 3\vec{v}_2 + 4\vec{v}_3$, so $[T_A\vec{x}]_{\mathcal{V}} = \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}$.

**Eigen bases.**

The goal of this problem is to

- Compute eigenvalues when given eigenvectors.
- Compute a characteristic polynomial without using a determinant when given eigenvalues.
- Compute the result of a transformation when vectors are written in an eigen basis.

**Notes/Misconceptions**

- In part 2, many students will waste their time computing a determinant instead of using the information they already have. For those that use the eigenvalues directly, many will miss the fact that the leading coefficient of $\mathrm{char}(A)$ is $-1$. This point is not important.
- Parts 4 and 5 should look very familiar to students. We have done them again and again. The moral is, some bases are better than others, and the eigen basis is awesome!

Lesson 32

# Diagonalization II

## Textbook

Section 6.2

## Objectives

- Explain how diagonalization can be used to compute large matrix powers.
- Explain a link between eigenvectors, eigenvalues, and diagonalization.

## Motivation

We finally have all the tools we need to diagonalize a matrix and to explain why diagonalization is useful. Let's do it!

However, not all matrices are diagonalizable. We need to carefully outline that we need a basis of eigenvectors to diagonalize a matrix, and that this condition is an if and only if.

80 The transformation $P^{-1}$ takes vectors in the standard basis and outputs vectors in their $\mathcal{V}$-basis representation. Here, $A$, $T_A$, and $\mathcal{V}$ come from Problem 79.

80.1 Describe in words what $P$ does.

$P$ undoes what $P^{-1}$ does, which is to say that it takes vectors in the $\mathcal{V}$ basis and outputs vectors in their representation in the standard basis.

80.2 Describe how you can use $P$ and $P^{-1}$ to easily compute $T_A\vec{y}$ for any $\vec{y} \in \mathbb{R}^3$.

Computing $[T_A]_{\mathcal{V}}[\vec{y}]_{\mathcal{V}}$ is easy, since $[T_A]_{\mathcal{V}}$ just multiplies each coordinate of $[\vec{y}]_{\mathcal{V}}$ by a scalar. We know that $P^{-1}[\vec{y}]_{\mathcal{E}} = [\vec{y}]_{\mathcal{V}}$ and that $P[\vec{x}]_{\mathcal{V}} = [\vec{x}]_{\mathcal{E}}$ and so given any vector $\vec{v}$ represented by $[\vec{v}]_{\mathcal{E}}$ in the standard basis, we have

$$A[\vec{v}]_{\mathcal{E}} = P[T_A]_{\mathcal{V}}P^{-1}[\vec{v}]_{\mathcal{E}}$$

since

$$P[T_A]_{\mathcal{V}}P^{-1}[\vec{v}]_{\mathcal{E}} = P[T_A]_{\mathcal{V}}[\vec{v}]_{\mathcal{V}} = P[T_A\vec{v}]_{\mathcal{V}} = [T_A\vec{v}]_{\mathcal{E}} = A[\vec{v}]_{\mathcal{E}}.$$

80.3 Can you find a matrix $D$ so that

$$PDP^{-1} = A?$$

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

$D = [A]_{\mathcal{V}}$, so by the previous part we have that that for any vector $\vec{v}$

$$A[\vec{v}]_{\mathcal{E}} = PDP^{-1}[\vec{v}]_{\mathcal{E}}.$$

80.4 $[\vec{x}]_{\mathcal{V}} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$. Compute $A^{100}\vec{x}$.

$$A^{100}\vec{x} = \begin{bmatrix} 2^{100} \\ 3 \\ 4 \end{bmatrix}_{\mathcal{V}}.$$

By the previous problem, we know how $A$ acts on vectors represented in the $\mathcal{V}$ basis: it multiplies the first coordinate by 2, the second by $-1$, and leaves the third coordinate unchanged. So we compute

$$A^{100}\begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} = A^{99}\begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix} = A^{98}\begin{bmatrix} 2^2 \\ 3 \\ 4 \end{bmatrix} = A^{97}\begin{bmatrix} 2^3 \\ -3 \\ 4 \end{bmatrix} = \ldots$$

To express $A^{100}\vec{x}$ in the standard basis, we use $P$.

$$[A^{100}\vec{x}]_{\mathcal{E}} = P[A^{100}\vec{x}]_{\mathcal{V}} = \begin{bmatrix} 2^{100} - 1 \\ 2^{100} + 7 \\ 2^{100} - 6 \end{bmatrix}.$$

**Diagonalizable**

DEF A matrix is ***diagonalizable*** if it is similar to a diagonal matrix.

81 For an $n \times n$ matrix $T$, suppose its eigenvectors $\{\vec{v}_1, \ldots \vec{v}_n\}$ form a basis for $\mathbb{R}^n$. Let $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues.

81.1 Is $T$ diagonalizable (i.e., similar to a diagonal matrix)? If so, explain how to obtain its diagonalized form.

**Diagonalizing matrices.**

The goal of this problem is to
- Diagonalize a matrix.
- Explain diagonalization in terms of change of basis.
- Use diagonalization to compute large matrix powers.

**Notes/Misconceptions**
- In part 2, students will struggle to articulate their ideas.
- In part 4, students are regularly impressed seeing the $P$'s and $P^{-1}$'s cancel in $PDP^{-1}PDP^{-1}\cdots$.

**Eigenvectors and diagonalization.**

The goal of this problem is to
- Explain how the existence of a basis of eigenvectors implies diagonalizability.
- Explain how if there isn't a basis of eigenvectors a matrix is not diagonalizable.
- Not confuse diagonalizability and invertibility.

**Notes/Misconceptions**
- In part 2, many students will answer "no" on autopilot. Make sure they think about this. It's a common error that students conflate diagonalizability with invertibility.
- Part 3 is hard to explain. Prepare your explanation carefully.

Yes.

Let $\mathcal{V} = \{\vec{v}_1, \ldots \vec{v}_n\}$ be the basis consisting of eigenvectors of $T$. Then by definition, $T\vec{v}_i = \lambda_i \vec{v}_i$ for each $i$.

Let $P$ be the matrix that takes vectors represented in the $\mathcal{V}$ basis and outputs their representations in the standard basis $\mathcal{E}$. Then, for example, we should have that

$$P^{-1}TP \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = P^{-1}TP[\vec{v}_1]_{\mathcal{V}} = P^{-1}T[\vec{v}_1]_{\mathcal{E}} = \lambda_1 P^{-1}[\vec{v}_1]_{\mathcal{E}} = \lambda_1[\vec{v}_1]_{\mathcal{V}} = \begin{bmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Therefore, the first column of $P^{-1}TP$ is $\begin{bmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$. By similar reasoning, the $i^{\text{th}}$ column of $P^{-1}TP$ consists of all zeroes except for $\lambda_i$ in the $i^{\text{th}}$ position. In other words, $T$ is similar to the diagonal matrix $D$ with $\lambda_1, \lambda_2, \ldots, \lambda_n$ along the diagonal, in that order.

81.2 What if one of the eigenvalues of $T$ is zero? Is $T$ diagonalizable?

Yes.

The argument in the previous part does not depend on any of the eigenvalues being non-zero.

81.3 What if the eigenvectors of $T$ did not form a basis for $\mathbb{R}^n$. Would $T$ be diagonalizable?

No.

The argument we used in the first part definitely would not work.

Consider the converse, and assume $T$ is similar to diagonal matrix $D$. That is, suppose there is an invertible matrix $P$ such that $T = PDP^{-1}$. Then, if $\vec{v}_1$ is the first column of $P$ and $\lambda_1$ is the first entry on the diagonal of $D$, we would have

$$T\vec{v}_1 = PDP^{-1}\vec{v}_1 = PD \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \lambda_1 P \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \lambda_1 \vec{v}_1,$$

meaning that $\vec{v}_1$ is an eigenvector of $T$. Similarly, all of the columns of $P$ would be eigenvectors of $T$, with eigenvalues equal to the corresponding entry on the diagonal of $D$. Since $P$ is invertible its columns must be linearly independent, and therefore the $n$ columns of $P$ would form a basis of $\mathbb{R}^n$ consisting of eigenvectors of $T$.

Lesson 33

# Diagonalization III

## Textbook

Section 6.2

## Objectives

- Define *eigenspace*, *geometric multiplicity*, and *algebraic multiplicity*.
- State the theorem a matrix is diagonalizable iff the algebraic multiplicities match the geometric multiplicities and sum to the dimension of the space.
- Show a particular matrix is not diagonalizable.
- Produce a non-diagonalizable matrix.

## Motivation

Diagonalization is great, but not all matrices can be diagonalized. This is because the transformation might not have a basis of eigenvectors. The sum of the geometric multiplicities gives the dimension of the span of all the eigenvectors. If it is not $n$ (for an $n \times n$ matrix), the transformation cannot be diagonalizable. This is the start of the theory of Jordan forms. We aren't developing this theory, but we do want to understand the limitations of diagonalization.

**Notes/Misconceptions**

- Half of the students in this class have never seen complex numbers before. We aren't holding students accountable for knowing how to use complex numbers, but the fundamental theorem of algebra is worth mentioning.

### Eigenspace

DEFINITION

Let $A$ be a matrix with eigenvalues $\{\lambda_1, \dots, \lambda_m\}$. The ***eigenspace*** of $A$ corresponding to the eigenvalue $\lambda_i$ is the null space of $A - \lambda_i I$. That is, it is the space spanned by all eigenvectors that have the eigenvalue $\lambda_i$.

The ***geometric multiplicity*** of an eigenvalue $\lambda_i$ is the dimension of the eigenspace corresponding to $\lambda_i$. The ***algebraic multiplicity*** of $\lambda_i$ is the number of times $\lambda_i$ occurs as a root of the characteristic polynomial of $A$ (i.e., the number of times $x - \lambda_i$ occurs as a factor).

82 Let $F = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ and $G = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$.

82.1 Is $F$ diagonalizable? Why or why not?

No.

$F$ is diagonalizable if and only if there is a basis of $\mathbb{R}^2$ consisting of eigenvectors of $F$, so we begin by computing all eigenvectors of $F$. $\mathrm{char}(F) = (3-\lambda)^2$, so the only eigenvalue of $F$ is 3, meaning that the eigenvectors of $F$ are precisely the non-zero vectors in $\mathrm{null}(F - 3I)$. We check that

$$\mathrm{null}(F - 3I) = \mathrm{null}\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \mathrm{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\},$$

which is one-dimensional, and so there cannot be a basis of $\mathbb{R}^2$ consisting of eigenvectors of $F$.

82.2 Is $G$ diagonalizable? Why or why not?

Yes. $G$ is already diagonal (and is necessarily similar to itself).

82.3 What is the geometric and algebraic multiplicity of each eigenvalue of $F$? What about the multiplicities for each eigenvalue of $G$?

The only eigenvalue for $F$ is 3. Its geometric multiplicity is 1, and its algebraic multiplicity is 2.

The only eigenvalue for $G$ is 3. Its geometric and algebraic multiplicity is 2.

82.4 Suppose $A$ is a matrix where the geometric multiplicity of one of its eigenvalues is smaller than the algebraic multiplicity of the same eigenvalue. Is $A$ diagonalizable? What if all the geometric and algebraic multiplicities match?

If one of the geometric multiplicities is smaller than the corresponding algebraic multiplicity, $A$ cannot be diagonalizable.

Since the characteristic polynomial of an $n \times n$ matrix has degree $n$, it has at most $n$ real roots. Since each root is an eigenvalue, we have

$$\sum \text{algebraic multiplicities } \leq n.$$

If one of the geometric multiplicities is smaller than the algebraic multiplicities, we have

$$\sum \text{geometric multiplicities } < n,$$

and so there cannot be a basis for $\mathbb{R}^n$ consisting of eigenvectors.

For the converse statement, we need the fundamental theorem of algebra: *a degree $n$ polynomial has exactly $n$ complex roots, counting multiplicity.*

If we allow eigenvalues to be complex numbers, then

$$\sum \text{algebraic multiplicities } = n,$$

and so if all geometric and algebraic multiplicities are equal, we have

$$\sum \text{geometric multiplicities } = n.$$

Thus, there would be a basis of eigenvectors.

**Non-diagonalizable matrices.**

The goal of this problem is to

- Explain why not all matrices are diagonalizable.
- Memorize an example of a non-diagonalizable matrix.

**Notes/Misconceptions**

- Part 1 asks students to do all the diagonalization steps without any scaffolding. This will take time and may take some prodding.
- The explanation for part 3 depends on whether you allow complex numbers. If you allow complex numbers, then $\sum$ algebraic multiplicities $=$ dim of space. If not, it's $\leq$ dim of space. In either case, geometric multiplicity $<$ algebraic multiplicity implies not diagonalizable. But only in the case of an algebraically closed field do you get converse.
- If you like, you can bring up an example of a rotation that has non-real eigenvalues, and so isn't real-diagonalizable for a different reason.