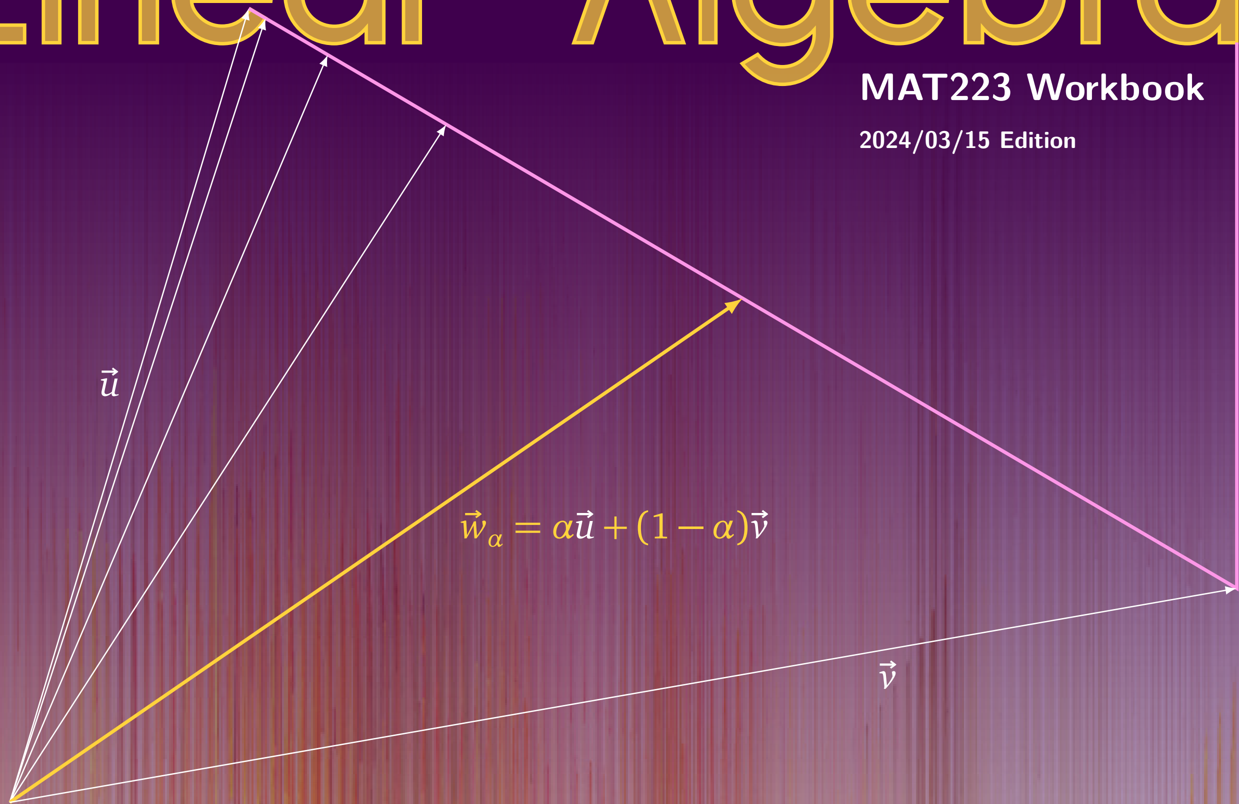
Linear Algebra



MAT223 Workbook

2024/03/15 Edition

Jason Siefken

Linear Algebra

© Jason Siefken, 2016–2022
Creative Commons By-Attribution Share-Alike

About this Book

For the student

This book is your introductory guide to linear algebra. It is divided into *modules*, and each module is further divided into *exposition*, *practice problems*, and *core exercises*.

The *exposition* is easy to find—it's the text that starts each module and explains the big ideas of linear algebra. The *practice problems* immediately follow the exposition and are there so you can practice with concepts you've learned. Following the practice problems are the *core exercises*. The core exercises build up, through examples, the concepts discussed in the exposition.

To optimally learn from this text, you should:

- Start each module by reading through the *exposition* to get familiar with the main ideas and linear algebra terminology.
- Work through the *core exercises* to develop an understanding and intuition behind the main ideas and their subtleties.
- Re-read the *exposition* and identify which concepts each core exercise connects with.
- Work through the *practice problems*. These will serve as a check on whether you've understood the main ideas well enough to apply them.

The core exercises. Most (but not all) core exercises will be worked through during lecture time, and there is space for you to work provided after each of the core exercises. The point of the core exercises is to develop the main ideas of linear algebra by exploring examples. When working on core exercises, think "it's the journey that matters not the destination". The answers are not the point! If you're struggling, keep with it. The concepts you struggle with you remember well, and if you look up the answer, you're likely to forget just a few minutes later.

So many definitions. A big part of linear algebra is learning precise and technical language.[1] There are many terms and definitions you need to learn, and by far the best way to successfully learn these terms is to understand where they come from, why they're needed, and practice using them. That is, don't try to memorize definitions word for word. Instead memorize the idea and *reconstruct* the definition; go through the core exercises and identify which definitions appear where; and explain linear algebra to others using these technical terms.

Contributing to the book. Did you find an error? Do you have a better way to explain a linear algebra concept? Please, contribute to this book! This book is open-source, and we welcome contributions and improvements. To contribute to/fix part of this book, make a *Pull Request* or open an *Issue* at `https://github.com/siefkenj/IBLLinearAlgebra`. If you contribute, you'll get your name added to the contributor list.

For the instructor

This book is designed for a one-semester introductory linear algebra course course with a focus on geometry (MAT223 at the University of Toronto). It has not been designed for an "intro to proofs"-style course, but could be adapted for one.

Unlike a traditional textbook that is grouped into chapters and sections by subject, this book is grouped into modules. Each module contains exposition about a subject, practice problems (for students to work on by themselves), and core exercises (for students to work on with your guidance). Modules group related

[1] Beyond three dimensions, things get very confusing very quickly. Having precise definitions allows us to make arguments that rely on logic instead of intuition; and logic works in all dimensions.

concepts, but the modules have been designed to facilitate learning linear algebra rather than to serve as a reference. For example, information about change-of-basis is spread across several non-consecutive modules; each time change-of-basis is readdressed, more detail is added.

Using the book. This book has been designed for use in large active-learning classrooms driven by a *think, pair-share*/small-group-discussion format. Specifically, the *core exercises* (these are the problems which aren't labeled "Practice Problems" and for which space is provided to write answers) are designed for use during class time.

A typical class day looks like:

1. **Student pre-reading.** Before class, students will read through the relevant module.

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

3. **Students work on problems.** Students work individually or in pairs/small groups on the prescribed core exercise. During this time the instructor moves around the room addressing questions that students may have and giving one-on-one coaching.

4. **Instructor intervention.** When most students have successfully solved the problem, the instructor refocuses the class by providing an explanation or soliciting explanations from students. 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 miss the point!).

   If students are having trouble, the instructor can give hints and additional guidance to ensure students' struggle is productive.

5. **Repeat step 3.**

Using this format, students are thinking (and happily so) most of the class. Further, after struggling with a question, students are especially primed to hear the insights of the instructor.

Conceptual lean. The *core exercises* are geared towards concepts instead of computation, though some core exercises focus on simple computation. They also have a geometric lean. Vectors are initially introduced with familiar coordinate notation, but eventually, coordinates are understood to be *representations* of vectors rather than "true" geometric vectors, and objects like the determinant are defined via oriented volumes rather than formulas involving matrix entries.

Specifically lacking are exercises focusing on the mechanical skills of row reduction and computing matrix inverses. Students must practice these skills, but they require little instructor intervention and so can be learned outside of lecture (which is why core exercises don't focus on these skills).

How to prepare. Running an active-learning classroom is less scripted than lecturing. The largest challenges are: (i) understanding where students are at, (ii) figuring out what to do given the current understanding of the students, and (iii) timing.

To prepare for a class day, you should:

1. **Strategize about learning objectives.** Figure out what the point of the day's lesson is and brain storm some examples that would illustrate that point.

2. **Work through the core exercises.**

3. **Reflect.** Reflect on how each core exercise addresses the day's goals. Compare with the examples you brainstormed and prepare follow-up questions that you can use in class to test for understanding.

4. **Schedule.** Write timestamps next to each core exercise indicating at what minute you hope to start each exercise. Give more time for the exercises that you judge as foundational, and be prepared to triage. It's appropriate to leave exercises or parts of exercises for homework, but change the order of exercises at your peril—they really do build on each other.

A typical 50 minute class is enough to get through 2–3 core exercises (depending on the difficulty), and class observations show that class time is split 50/50 between students working and instructor explanations.

License

Unless otherwise mentioned, pages of this document are licensed under the Creative Commons By-Attribution Share-Alike License. That means, you are free to use, copy, and modify this document provided that you provide attribution to the previous copyright holders and you release your derivative work under the same license. Full text of the license is at `http://creativecommons.org/licenses/by-sa/4.0/`

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 request, or opening an *issue* at `https://github.com/siefkenj/IBLLinearAlgebra`

Incorporated content. Content from other sources is reproduced here with permission and retains the Author's copyright. Please see the footnote of each page to verify the copyright.

Included in this text are tasks created by the Inquiry-Oriented Linear Algebra (IOLA) project. Details about these tasks can be found on their website `http://iola.math.vt.edu/`. Also included are some practice problems from Beezer's *A First Course in Linear Algebra* (marked with the symbol B next to the problem), and from Hefferon's *Linear Algebra* (marked with the symbol H next to the problem).

Contributing. You can report errors in the book or contribute to the book by filing an *Issue* or a *Pull Request* on the book's GitHub page: `https://github.com/siefkenj/IBLLinearAlgebra/`

Contributors

This book is a collaborative effort. The following people have contributed to its creation:

◦ Shukui Chen ◦ Hassan El-Sheikha ◦ Jesse Frohlich ◦ Sameul Khan ◦ Julia Kim ◦ Avery King ◦ Dan Le ◦ Xintong (Alucart) Li ◦ Ruo Ning (Nancy) Qiu ◦ Tianhao (Patrick) Wang ◦ Robert Wang ◦ Zack Wolske ◦

Dedication

This book is dedicated to Dr. Bob Burton—friend and mentor.

"Sometimes you have to walk the mystical path with practical feet."

Sets, Vectors & Notation

In this module you will learn

- The basics of sets and set-builder notation.
- The definition of vectors and how they relate to points.
- Column vector notation and how to represent vectors in drawings.
- How to compute linear combinations of vectors and use systems of linear equations to answer questions about linear combinations of vectors.

Sets

Modern mathematics makes heavy use of *sets*. A set is an unordered collection of distinct objects. We won't try and pin it down more than this—our intuition about collections of objects will suffice.[2] We write a set with curly-braces { and } and list the objects inside. For instance

$$\{1,2,3\}.$$

This would be read aloud as "the set containing the elements 1, 2, and 3". Things in a set are called *elements*, and the symbol $\in$ is used to specify that something is an element of a set. In contrast, $\notin$ is used to specify something is not an element of a set. For example,

$$3 \in \{1,2,3\} \qquad 4 \notin \{1,2,3\}.$$

Sets can contain mixtures of objects, including other sets. For example,

$$\{1,2,a,\{-70,\infty\},x\}$$

is a perfectly valid set.

It is tradition to use capital letters to name sets. So we might say $A=\{6,7,12\}$ or $X=\{7\}$. However there are some special sets which already have names/symbols associated with them. The *empty set* is the set containing no elements and is written $\emptyset$ or $\{\}$. Note that $\{\emptyset\}$ is *not* the empty set—it is the set containing the empty set! It is also traditional to call elements of a set *points* regardless of whether you consider them "point-like".

Operations on Sets

If the set A contains all the elements that the set B does, we call B a *subset* of A. Conversely, we call A a *superset* of B.

> **Subset & Superset.** The set B is a ***subset*** of the set A, written $B \subseteq A$, if for all $b \in B$ we also have $b \in A$. In this case, A is called a ***superset*** of B.[a]
>
> [a] Some mathematicians use the symbol $\subset$ instead of $\subseteq$.

Some simple examples are $\{1,2,3\} \subseteq \{1,2,3,4\}$ and $\{1,2,3\} \subseteq \{1,2,3\}$. There's something funny about that last example, though. Those two sets are not only subsets/supersets of each other, they're *equal*. As surprising as it seems, we actually need to define what it means for two sets to be equal.

> **Set Equality.** The sets A and B are ***equal***, written $A=B$, if $A \subseteq B$ and $B \subseteq A$.

Having a definition of equality to lean on will help us when we need to prove things about sets.

> **Example.** Let A be the set of numbers that can be expressed as $2n$ for some whole number n, and let B be the set of numbers that can be expressed as $m+1$ where m is an odd whole number. We will show $A=B$.
>
> First, let us show $A \subseteq B$. If $x \in A$, then $x=2n$ for some whole number n. Therefore
>
> $$x=2n=2(n-1)+1+1=m+1$$
>
> where $m=2(n-1)+1$ is, by definition, an odd number. Thus $x \in B$, which proves $A \subseteq B$.

[2] When you pursue more rigorous math, you rely on definitions to get yourself out of philosophical jams. For instance, with our definition of set, consider "the set of all sets that don't contain themselves". Such a set cannot exist! This is called *Russel's Paradox*, and shows that if we start talking about sets of sets, we may need more than intuition.

© *Jason Siefken, 2015–2022*

> Now we will show $B \subseteq A$. Let $x \in B$. By definition, $x = m+1$ for some odd m. By the definition of oddness, $m = 2k+1$ for some whole number k. Thus
>
> $$\begin{aligned} x = m+1 &= (2k+1)+1 = 2k+2 \\ &= 2(k+1) = 2n, \end{aligned}$$
>
> where $n = k+1$, and so $x \in A$. Since $A \subseteq B$ and $B \subseteq A$, by definition $A = B$.

Set-builder Notation

Specifying sets by listing all their elements can be a hassle, and if there are an infinite number of elements, it's impossible! Fortunately, *set-builder notation* solves these problems. If X is a set, we can define a 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 $Y = \{a \,:\, \text{some rule involving } a\}$.[3] You may equivalently use "|" instead of ":", writing $Y = \{a \,|\, \text{some rule involving } a\}$.

There are also some common operations we can do with two sets.

> **Unions & Intersections.** Let X and Y be sets. The ***union*** of X and Y and the ***intersection*** of X and Y are defined as follows.
>
> $$\text{(union)}\quad X \cup Y = \{a \,:\, a \in X \text{ or } a \in Y\}.$$
>
> $$\text{(intersection)}\quad X \cap Y = \{a \,:\, a \in X \text{ and } a \in Y\}.$$

For example, if $A = \{1,2,3\}$ and $B = \{-1,0,1,2\}$, then $A \cap B = \{1,2\}$ and $A \cup B = \{-1,0,1,2,3\}$. Set unions and intersections are *associative*, which means it doesn't matter how you apply parentheses to an expression involving just unions or just intersections. For example $(A \cup B) \cup C = A \cup (B \cup C)$, which means we can give an unambiguous meaning to an expression like $A \cup B \cup C$ (just put the parentheses wherever you like). But watch out, $(A \cup B) \cap C$ means something different than $A \cup (B \cap C)$!

Some common sets have special notation:

$$\begin{aligned} \emptyset &= \{\}, \text{ the empty set} \\ \mathbb{N} &= \{0,1,2,3,\dots\} = \{\text{natural numbers}\} \\ \mathbb{Z} &= \{\dots,-3,-2,-1,0,1,2,3,\dots\} = \{\text{integers}\} \\ \mathbb{Q} &= \{\text{rational numbers}\} \\ \mathbb{R} &= \{\text{real numbers}\} \\ \mathbb{R}^n &= \{\text{vectors in } n\text{-dimensional Euclidean space}\} \end{aligned}$$

Vectors & Scalars

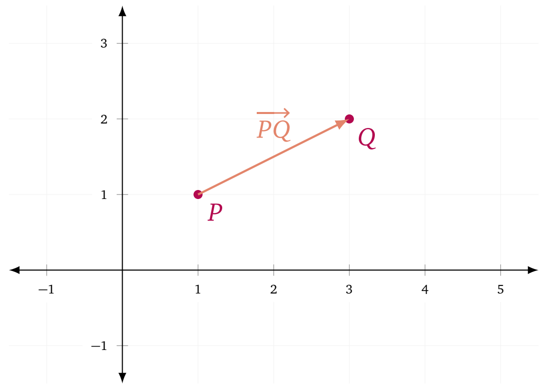
A *scalar* number (also referred to as a *scalar* or just an ordinary *number*) models a relationship between quantities. For example, a recipe might call for *six* times as much flour as sugar. In contrast, a *vector* models a relationship between points. For example, the store might be *2km East and 4km North* from your house. In this way, a vector may be thought of as a *displacement* with a *direction* and a *magnitude*.[4]

Given points $P = (1,1)$ and $Q = (3,2)$, we specify the *displacement* from P to Q as a vector $\overrightarrow{PQ}$ whose magnitude is $\sqrt{5}$ (as given by the Pythagorean theorem) and whose direction is specified by a directed line segment from P to Q.

[3] If you want to get technical, to make this notation unambiguous, you define a *universe of discourse*. That is, a set $\mathcal{U}$ containing every object you might want to talk about. Then $\{a \,:\, \text{some rule involving } a\}$ is short for $\{a \in \mathcal{U} \,:\, \text{some rule involving } a\}$.

[4] Though in this book we will treat vectors as geometric objects relating to Euclidean space, they are much more general. For instance, someone's internet browsing habits could be described by a vector—the topics they find most interesting might be the "direction" and the amount of time they browse might be the "magnitude."

© Jason Siefken, 2015–2022

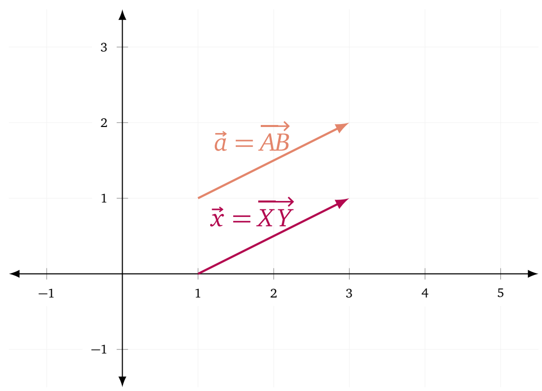


Vector Notation

There are many ways to represent vector quantities in writing. If we have two points, P and Q, we write $\overrightarrow{PQ}$ to represent the vector from P to Q. Absent points, a bold-faced letter (like **a**) or an arrow over a letter (like $\vec{a}$) are the most common vector notations. In this text, we will use $\vec{a}$ to represent a vector. The notation $\|\vec{a}\|$ represents the magnitude of the vector $\vec{a}$, which is sometimes called the *norm* or *length* of $\vec{a}$.

Graphically, we may represent vectors as directed line segments (a line segment with an arrow at one end), however we must take care to distinguish between the picture we draw and the "true" vector. For example, directed line segments always *start* somewhere, but a vector models a displacement and has no sense of "origin".

Consider the following: for the points $A = (1,1)$, $B = (3,2)$, $X = (1,0)$, and $Y = (3,1)$, define the vectors $\vec{a} = \overrightarrow{AB}$ and $\vec{x} = \overrightarrow{XY}$.



Are these the same or different vectors? As directed line segments, they are different because they are at different locations in space. However, both $\vec{a}$ and $\vec{x}$ have the same magnitude and direction. Thus, $\vec{a} = \vec{x}$ despite the fact that $A \neq X$.[5]

Takeaway. A vector is not the same as a line segment and a vector by itself has no "origin".

Vectors and Points

The distinction between vectors and points is sometimes nebulous because the two are so closely related. A *point* in Euclidean space specifies an absolute position whereas a vector specifies a displacement (i.e., a magnitude and direction). However, given a point P, one associates P with the vector $\vec{p} = \overrightarrow{OP}$, where O is the origin. Similarly, we associate the vector $\vec{v}$ with the point V so that $\overrightarrow{OV} = \vec{v}$. Thus, we have a way to unambiguously go back and forth between vectors and points.[6] As such, *we will treat vectors and points interchangeably*.

Takeaway. Vectors and points can and will be treated interchangeably.

Vector Arithmetic

Vectors provide a natural way to give directions. For example, suppose $\vec{e}_1$ points one kilometer eastwards and

[5] Some theories use *rooted vectors* instead of vectors as the fundamental object of study. A rooted vector represents a magnitude, direction, *and* a starting point. And, as rooted vectors, $\vec{a} \neq \vec{x}$ (from the example above). But for us, vectors will always be unrooted, even though our graphical representations of vectors might appear rooted.

[6] Mathematically, we say there is an *isomorphism* between vectors and points (once an origin is fixed, of course!).

© *Jason Siefken, 2015–2022*

$\vec{e}_2$ points one kilometer northwards. Now, if you were standing at the origin and wanted to move to a location 3 kilometers east and 2 kilometers north, you might say: “Walk 3 times the length of $\vec{e}_1$ in the $\vec{e}_1$ direction and 2 times the length of $\vec{e}_2$ in the $\vec{e}_2$ direction.” Mathematically, we express this as

$$3\vec{e}_1 + 2\vec{e}_2.$$

Of course, we’ve incidentally described a new vector. Let P be the point at 3-east and 2-north. Then

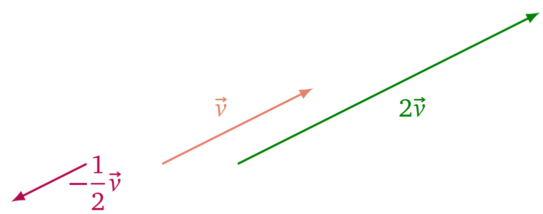
$$\overrightarrow{OP} = 3\vec{e}_1 + 2\vec{e}_2.$$

If the vector $\vec{r}$ points north but has a length of 10 kilometers, we have a similar formula:

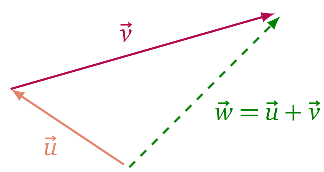
$$\overrightarrow{OP} = 3\vec{e}_1 + \frac{1}{5}\vec{r},$$

and we have the relationship $\vec{r} = 10\vec{e}_2$. Our notation here is very suggestive. Indeed, if we could make sense of “$\alpha\vec{v}$” (scalar multiplication) and “$\vec{v} + \vec{w}$” (vector addition) for any scalar α and any vectors $\vec{v}$ and $\vec{w}$, we could do algebra with vectors.

We will define scalar multiplication intuitively: For a vector $\vec{v}$ and a scalar $\alpha > 0$, the vector $\vec{w} = \alpha\vec{v}$ is the vector pointing in the same direction as $\vec{v}$ but with length scaled by α. That is, $\|\vec{w}\| = \alpha\|\vec{v}\|$. Similarly, $-\vec{v}$ is the vector of the same length as $\vec{v}$ but pointing in the exact opposite direction.



For two vectors $\vec{u}$ and $\vec{v}$, the sum $\vec{w} = \vec{u} + \vec{v}$ represents the displacement vector created by first displacing along $\vec{u}$ and then displacing along $\vec{v}$.



Takeaway. You add vectors tip to tail and you scale vectors by changing their length.

Now, there is one snag. What should $\vec{v} + (-\vec{v})$ be? Well, first we displace along $\vec{v}$ and then we displace in the exact opposite direction by the same amount. So, we have gone nowhere. This corresponds to a displacement with zero magnitude. But, what direction did we displace? Here we make a philosophical stand.

Zero Vector. The ***zero vector***, notated as $\vec{0}$, is the vector with no magnitude.

We will be pragmatic about the direction of the zero vector and say, *the zero vector does not have a well-defined direction*.[7] That means sometimes we consider the zero vector to point in every direction and sometimes we consider it to point in no directions. It depends on our mood—but we must never talk about *the* direction of the zero vector, since it’s not defined.

Formalizing, for vectors $\vec{u}$, $\vec{v}$, $\vec{w}$, and scalars α and β, the following laws are always satisfied:

$$(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w}) \quad \text{(Associativity)}$$
$$\vec{u} + \vec{v} = \vec{v} + \vec{u} \quad \text{(Commutativity)}$$
$$\alpha(\vec{u} + \vec{v}) = \alpha\vec{u} + \alpha\vec{v} \quad \text{(Distributivity)}$$

and

$$(\alpha\beta)\vec{v} = \alpha(\beta\vec{v}) \quad \text{(Associativity II)}$$
$$(\alpha + \beta)\vec{v} = \alpha\vec{v} + \beta\vec{v} \quad \text{(Distributivity II)}$$

[7] In the mathematically precise definition of vector, the idea of “magnitude” and “direction” are dropped. Instead, a set of vectors is defined to be a set over which you can reasonably define addition and scalar multiplication.

© Jason Siefken, 2015–2022

Indeed, if we intuitively think about vectors in flat (Euclidean) space, all of these properties are satisfied.[8] From now on, these properties of vector operations will be considered the *laws (or axioms) of vector arithmetic*.

We group scalar multiplication and vector addition under one name: *linear combinations*.

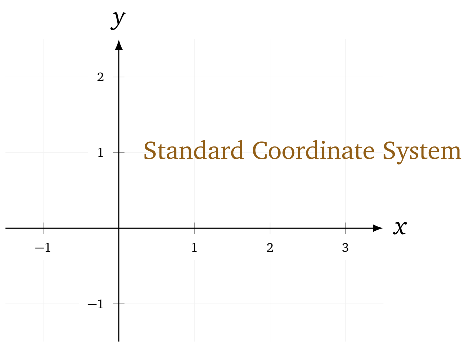
Linear Combination. A ***linear combination*** of the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ is a vector

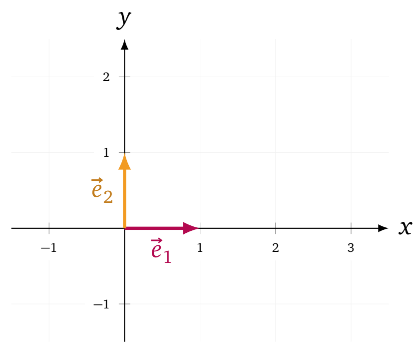
$$\vec{w} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n.$$

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

Coordinates and the Standard Basis

Consider the standard, flat, Euclidean plane (which is notated by $\mathbb{R}^2$). A coordinate system for $\mathbb{R}^2$ is a way to assign a unique pair of numbers to every point in $\mathbb{R}^2$. Though there are infinitely many coordinate systems we could choose for the plane, there is one standard one: the xy-coordinate system depicted below (which you're already familiar with).





In conjunction with the standard coordinate system, there are also *standard basis vectors*. The vector $\vec{e}_1$ always points one unit in the direction of the positive x-axis and $\vec{e}_2$ always points one unit in the direction of the positive y-axis.

Using the standard basis, we can represent every point (or vector) in the plane as a linear combination. If the point P has xy-coordinates (α, β), then $\overrightarrow{OP} = \alpha\vec{e}_1 + \beta\vec{e}_2$. Not only that, but this is the *only* way to represent the vector $\overrightarrow{OP}$ as a linear combination of $\vec{e}_1$ and $\vec{e}_2$.

Takeaway. Every vector in $\mathbb{R}^2$ can be written uniquely as a linear combination of the standard basis vectors.

For a vector $\vec{w} = \alpha\vec{e}_1 + \beta\vec{e}_2$, we call the pair (α, β) the *standard coordinates* of the vector $\vec{w}$. There are many equivalent notations used to represent a vector in coordinates.

(α, β)	parentheses
$\langle \alpha, \beta \rangle$	angle brackets
$\begin{bmatrix} \alpha & \beta \end{bmatrix}$	square brackets in a row (a row matrix/row vector)
$\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$	square brackets in a column (a column matrix/column vector)

Coordinates and vectors go hand in hand, and we will often write

$$\vec{v} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

as a shorthand for "$\vec{v} = \alpha\vec{e}_1 + \beta\vec{e}_2$".

Solving Problems with Coordinates

Coordinates allow for vector arithmetic to be carried out in a mechanical way. Suppose $\vec{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix}$.

[8] If we deviate from flat space, some of these rules are no longer respected. Consider moving 100 kilometers north then 100 kilometers east on a sphere. Is this the same as moving 100 kilometers east and then 100 kilometers north?

© Jason Siefken, 2015–2022

Then,

$$\vec u = \vec v \qquad \iff \qquad \begin{bmatrix} a\\ b\end{bmatrix} = \begin{bmatrix} x\\ y\end{bmatrix} \qquad \iff \qquad a = x \text{ and } b = y.$$

Further,

$$\vec u + \vec v = \begin{bmatrix} a\\ b\end{bmatrix} + \begin{bmatrix} x\\ y\end{bmatrix} = \begin{bmatrix} a+x\\ b+y\end{bmatrix} \qquad \text{and} \qquad t\vec u = t\begin{bmatrix} a\\ b\end{bmatrix} = \begin{bmatrix} ta\\ tb\end{bmatrix}$$

for any scalar t.

Using these rules, otherwise complicated questions about vectors can be reduced to simple algebra questions.[9]

Example. Let $\vec x = \vec e_1 - \vec e_2$, $\vec y = 3\vec e_1 - \vec e_2$, and $\vec r = 2\vec e_1 + 2\vec e_2$. Is $\vec r$ a linear combination of $\vec x$ and $\vec y$?

By definition, $\vec r$ is a linear combination of $\vec x$ and $\vec y$ if there exist scalars a and b such that

$$\vec r = a\vec x + b\vec y.$$

Rewriting everything in coordinates, we see this is equivalent to the equation

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

Therefore, we need to determine if the system of equations

$$\begin{cases} a + 3b = 2\\ -a - b = 2\end{cases}$$

has a solution. After solving, we find $a = -4$ and $b = 2$ is the only solution. Thus, $\vec r$ is a linear combination of $\vec x$ and $\vec y$. More specifically,

$$\vec r = -4\vec x + 2\vec y.$$

Higher Dimensions

We coordinatize three dimensional space (notated by $\mathbb{R}^3$) by constructing x, y, and z axes. Again, $\mathbb{R}^3$ has standard basis vectors $\vec e_1$, $\vec e_2$, and $\vec e_3$ which each point one unit along the x, y, and z axes, respectively.

Since we live in three dimensional space, its study has a long history, and many notations for the standard basis of three dimensional space are in use. This text will use $\vec e_1, \vec e_2, \vec e_3$, but other common notations include:

$$\begin{matrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}}\\ \hat\imath & \hat\jmath & \hat k\\ \mathbf{i} & \mathbf{j} & \mathbf{k}\\ \vec e_1 & \vec e_2 & \vec e_3\end{matrix}$$

Beyond three dimensions, drawing pictures becomes hard, but we can still use vectors. We use $\mathbb{R}^n$ to notate n-dimensional Euclidean space. The standard basis for $\mathbb{R}^n$ is $\vec e_1, \vec e_2, \ldots, \vec e_n$. Again, every vector in $\mathbb{R}^n$ can be written uniquely as a linear combination of the standard basis, and a coordinate representation of a vector in $\mathbb{R}^n$ is a list of n scalars.

Example. Let $\vec x, \vec y \in \mathbb{R}^3$ be given by $\vec x = 2\vec e_1 - \vec e_3$ and $\vec y = 6\vec e_2 + 3\vec e_3$. Compute $\vec z = \vec x + 2\vec y$.

$$\vec z = \vec x + 2\vec y = \begin{bmatrix} 2\\ 0\\ -1\end{bmatrix} + 2\begin{bmatrix} 0\\ 6\\ 3\end{bmatrix} = \begin{bmatrix} 2\\ 0\\ -1\end{bmatrix} + \begin{bmatrix} 0\\ 12\\ 6\end{bmatrix} = \begin{bmatrix} 2\\ 12\\ 5\end{bmatrix} = 2\vec e_1 + 12\vec e_2 + 5\vec e_3$$

Practice Problems

1 (a) Write the following vectors as column vectors.

i. $4\vec e_1 - 3\vec e_3 + 2\vec e_2 - 2\vec e_1 \in \mathbb{R}^3$.

ii. $\vec e_2 + \vec e_1 - 5\vec e_2 \in \mathbb{R}^2$.

(b) Write the following vectors as linear combinations of $\vec e_1$, $\vec e_2$, and $\vec e_3$.

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

[9] So simple, that computers are able to answer billions of such questions a second as you play your favorite video game!

© Jason Siefken, 2015–2022

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

2 Compute

$$3\begin{bmatrix}2\\-1\\1\\1\\0\end{bmatrix}+(-2)\begin{bmatrix}1\\2\\-7\\3\\0\end{bmatrix}+\begin{bmatrix}-3\\3\\9\\2\\2\end{bmatrix}$$

3[H] Decide if the vector is in the set. If it is, what value of the parameters produce that vector?

(a) $\begin{bmatrix}5\\-5\end{bmatrix}$ and the set

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

(b) $\begin{bmatrix}-1\\2\\1\end{bmatrix}$ and the set

$$\left\{\vec v\in\mathbb{R}^3 \,:\, \vec v=i\begin{bmatrix}-2\\1\\0\end{bmatrix}+j\begin{bmatrix}3\\0\\1\end{bmatrix} \text{ for some } i,j\in\mathbb{R}\right\}$$

(c) $\begin{bmatrix}3\\-1\end{bmatrix}$ and the set

$$\left\{\vec v\in\mathbb{R}^2 \,:\, \vec v=k\begin{bmatrix}-6\\2\end{bmatrix} \text{ for some } k\in\mathbb{R}\right\}$$

(d) $\begin{bmatrix}5\\4\end{bmatrix}$ and the set

$$\left\{\vec v\in\mathbb{R}^2 \,:\, \vec v=j\begin{bmatrix}5\\-4\end{bmatrix} \text{ for some } j\in\mathbb{R}\right\}$$

(e) $\begin{bmatrix}2\\1\\-1\end{bmatrix}$ and the set

$$\left\{\vec v\in\mathbb{R}^3 \,:\, \vec v=r\begin{bmatrix}1\\-1\\3\end{bmatrix}+\begin{bmatrix}0\\3\\-7\end{bmatrix} \text{ for some } r\in\mathbb{R}\right\}$$

(f) $\begin{bmatrix}1\\0\\1\end{bmatrix}$ and the set

$$\left\{\vec v\in\mathbb{R}^3 \,:\, \vec v=j\begin{bmatrix}2\\0\\1\end{bmatrix}+k\begin{bmatrix}-3\\-1\\1\end{bmatrix} \text{ for some } j,k\in\mathbb{R}\right\}$$

4 Draw the following subsets of $\mathbb{R}^2$ and then determine which are equal or subsets of each other.

(a) $A=\left\{\vec v\in\mathbb{R}^2 \,:\, \vec v=n\begin{bmatrix}2\\1\end{bmatrix} \text{ for some integer } n\in\mathbb{Z}\right\}$

(b) $B=\left\{\vec v\in\mathbb{R}^2 \,:\, \vec v=t\begin{bmatrix}4\\2\end{bmatrix} \text{ for some } t\in\mathbb{R}\right\}$

(c) $C=\left\{\vec v\in\mathbb{R}^2 \,:\, \vec v=n\begin{bmatrix}4\\2\end{bmatrix} \text{ for some integer } n\in\mathbb{Z}\right\}$

(d) $D=\left\{\vec v\in\mathbb{R}^2 \,:\, \vec v=t\begin{bmatrix}2\\1\end{bmatrix} \text{ for some } t\in\mathbb{R}\right\}$

5 Let $\vec a=\begin{bmatrix}1\\2\end{bmatrix}$, $\vec b=\begin{bmatrix}2\\4\end{bmatrix}$, $\vec c=\vec e_1+3\vec e_2$, and $\vec d=\vec a+\vec c$.

(a) Is $\vec e_1$ a linear combination of $\vec a$ and $\vec b$?

(b) Is $\vec d$ a linear combination of $\vec a$ and $\vec b$?

(c) Is $\vec p=\begin{bmatrix}1\\1\end{bmatrix}$ a linear combination of $\vec a$ and $\vec c$?

(d) Is $\vec q=\begin{bmatrix}-3\\3\end{bmatrix}$ a linear combination of $\vec a$, $\vec b$, $\vec c$, and $\vec d$?

6 Use set-builder notation to describe the following sets.

(a) The set of vectors in $\mathbb{R}^2$ whose coordinates are rational numbers.

(b) The set of vectors in $\mathbb{R}^2$ whose coordinates are irrational numbers.

(c) Let $P(\vec x)=-\vec x$. The set $\{P(\vec e_1),P(\vec e_2)\}$.

7 Which of the following statements are true about the set listed below? Justify your answers.

(a) $\mathcal{Y}$, the y-axis in $\mathbb{R}^3$.

i. $\mathcal{Y}$ is a finite set.

ii. Let

$$\mathcal{A}=\left\{\vec a\in\mathbb{R}^3 \,:\, \vec a=\beta\vec v \text{ for some } \vec v\in\mathcal{Y},\beta\in\mathbb{R}\right\},$$

then $\mathcal{A}\subseteq\mathcal{Y}$.

iii. For all vectors $\vec v\in\mathcal{Y}$, we have $\vec v\neq\vec 0$.

iv. For some vectors $\vec v\in\mathcal{Y}$, we have $\vec v\neq\vec 0$.

v. For all vectors $\vec v\in\mathcal{Y}$, there exists a vector $\vec x\in\mathcal{Y}$ such that $\vec x+\vec v=\vec e_2$.

vi. There exists a vector $\vec x\in\mathcal{Y}$ such that for all vectors $\vec v\in\mathcal{Y}$, we have $\vec x+\vec v=\vec e_2$.

(b) $\mathcal{S}$, the set of vectors in $\mathbb{R}^3$ whose coordinates are ±3.

i. $\mathcal{S}$ is a finite set.

ii. Let

$$\mathcal{A}=\left\{\vec a\in\mathbb{R}^3 \,:\, \vec a=\beta\vec v \text{ for some } \vec v\in\mathcal{S},\beta\in\mathbb{R}\right\},$$

then $\mathcal{A}\subseteq\mathcal{S}$.

iii. For all vectors $\vec v\in\mathcal{S}$, we have $\vec v\neq\vec 0$.

iv. For some vectors $\vec v\in\mathcal{S}$, we have $\vec v\neq\vec 0$.

v. For all vectors $\vec v\in\mathcal{S}$, there exists a vector $\vec x\in\mathcal{S}$ such that $\vec x+\vec v=\vec 0$.

vi. There exists a vector $\vec x\in\mathcal{S}$ such that for all vectors $\vec v\in\mathcal{S}$, we have $\vec x+\vec v=\vec 0$.

8 For each of the following statements, determine whether it is correct or not. If it is, prove it. Otherwise, give a counterexample.

(a) If $A\subseteq B$, then $A\cap B=A$.

(b) If $B\subseteq A$, then $A\cap B=A$.

(c) If $A\subseteq B$, then $A\cap B\neq B$.

(d) If $B\subseteq A$, then $A\cap B\neq B$.

(e) If $C\subseteq A\cap B$, then $C\subseteq A$.

(f) If $C\subseteq A\cup B$, then $C\subseteq A$.

(g) If $C\subseteq A\cup B$ and $C\subseteq B$, then $A\cap B\subseteq C$.

© Jason Siefken, 2015–2022

© Jason Siefken, 2015–2022

The Magic Carpet Ride

1 You are a young adventurer. Having spent most of your time in the mythical city of Oronto, you decide to leave 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 km East and 1 km 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 km East and 2 km 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 km East and 64 km 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?

© IOLA Team iola.math.vt.edu

The Magic Carpet Ride, Hide and Seek

2 You are a young adventurer. Having spent most of your time in the mythical city of Oronto, you decide to leave 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 km East and 1 km 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 km East and 2 km 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.

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

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

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

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

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

3

3.1 Which of the following statements are true?

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

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

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

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

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

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

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

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

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

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

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

© Jason Siefken, 2015–2022

4 Write the following in set-builder notation

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

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

Unions & Intersections

DEFINITION

Let X and Y be sets. The ***union*** of X and Y and the ***intersection*** of X and Y are defined as follows.

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

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

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

5.1 $X \cup Y$

5.2 $X \cap Y$

5.3 $X \cup Y \cup Z$

5.4 $X \cap Y \cap Z$

© Jason Siefken, 2015–2022

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

6.1 $V = \left\{ \vec{x} \in \mathbb{R}^2 : \vec{x} = \begin{bmatrix} 0 \\ t \end{bmatrix} \text{ for some } t \in \mathbb{R} \right\}$.

6.2 $H = \left\{ \vec{x} \in \mathbb{R}^2 : \vec{x} = \begin{bmatrix} t \\ 0 \end{bmatrix} \text{ for some } t \in \mathbb{R} \right\}$.

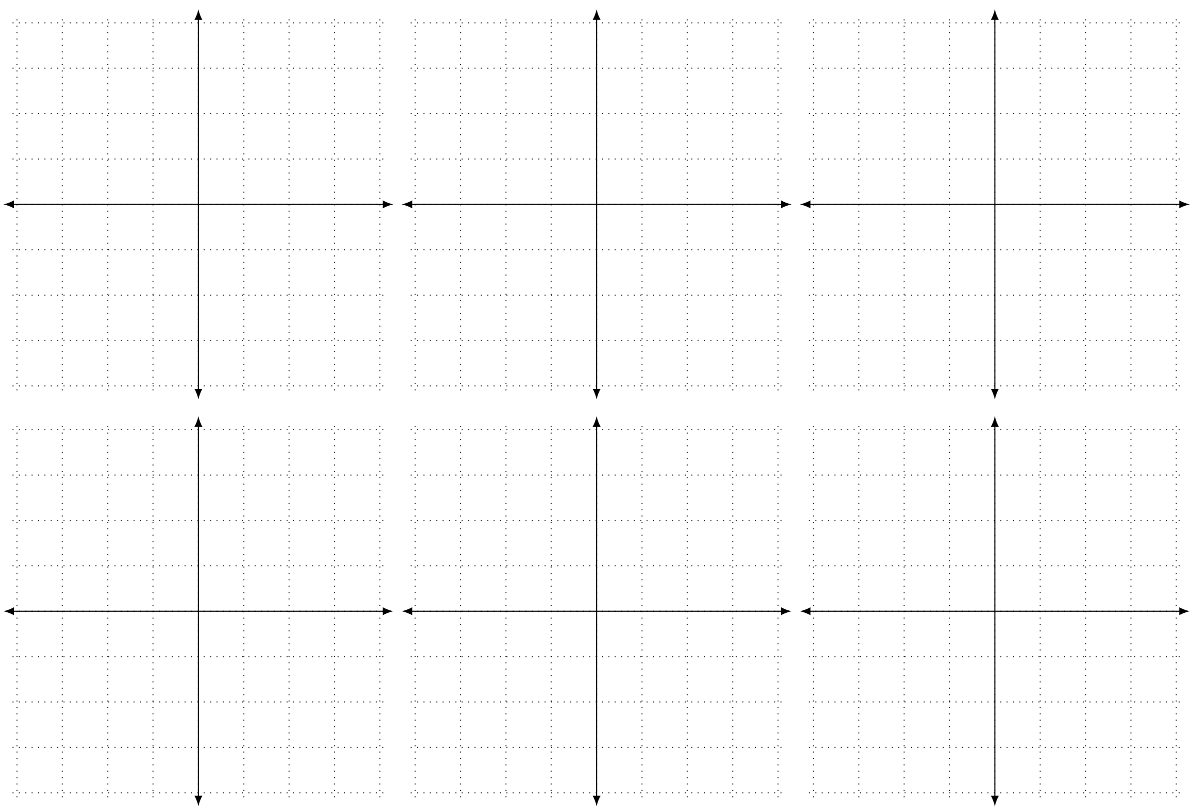
6.3 $D = \left\{ \vec{x} \in \mathbb{R}^2 : \vec{x} = t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ for some } t \in \mathbb{R} \right\}$.

6.4 $N = \left\{ \vec{x} \in \mathbb{R}^2 : \vec{x} = t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ for all } t \in \mathbb{R} \right\}$.

6.5 $V \cup H$.

6.6 $V \cap H$.

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



© Jason Siefken, 2015–2022

Vector Combinations

Linear Combination

DEFINITION

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

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

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

---

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

7.1 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 $\vec{v}_2$?

7.2 Is $\begin{bmatrix}3\\3\end{bmatrix}$ a linear combination of $\vec{v}_1$ and $\vec{v}_2$?

7.3 Is $\begin{bmatrix}0\\0\end{bmatrix}$ a linear combination of $\vec{v}_1$ and $\vec{v}_2$?

7.4 Is $\begin{bmatrix}4\\0\end{bmatrix}$ a linear combination of $\vec{v}_1$ and $\vec{v}_2$?

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

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

© Jason Siefken, 2015–2022

8 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 $\vec{m} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

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

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

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

© Jason Siefken, 2015–2022

© Jason Siefken, 2015–2022

Sets of Vectors, Lines & Planes

In this module you will learn

- How to draw a set of vectors making an appropriate choice of when to use line segments and when to use dots to represent vectors.
- The *vector form* of lines and planes, including how to determine the intersection of lines and planes in vector form.
- Restricted linear combinations and how to use them to represent common geometric objects (like line segments or polygons).

With a handle on vectors, we can now use them to describe some common geometric objects: lines and planes.

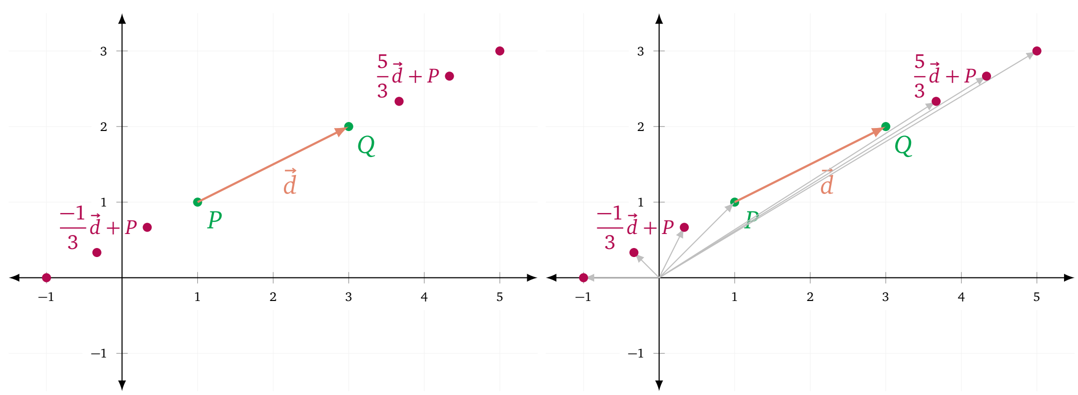
Lines

Consider for a moment the line ℓ through the points P and Q. When $P, Q \in \mathbb{R}^2$, we can describe ℓ with an equation of the form $y = mx + b$ (provided it isn't a vertical line), but if $P, Q \in \mathbb{R}^3$, it's much harder to describe ℓ with an equation. We can solve this problem by using vectors.

Let $\vec{d} = \overrightarrow{PQ}$ and consider the set of points (or vectors) $\vec{x}$ that can be expressed as

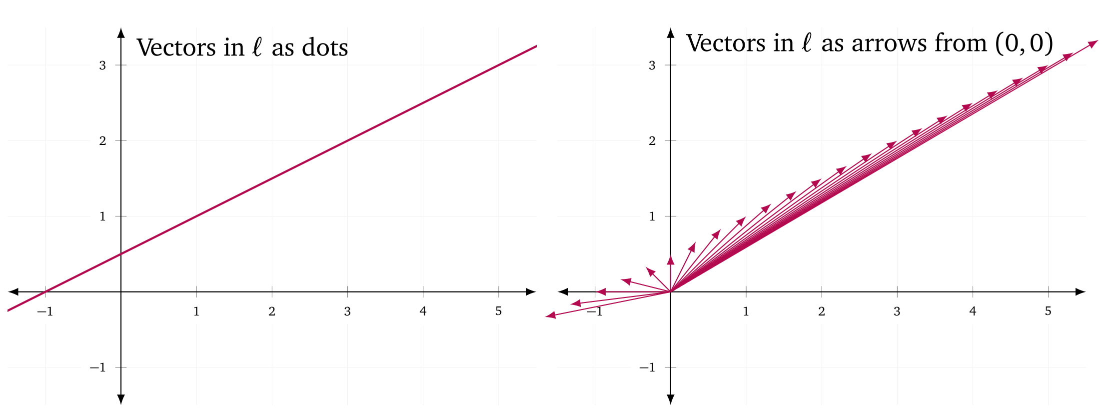
$$\vec{x} = t\vec{d} + P$$

for $t \in \mathbb{R}$. Geometrically, this is the set of all points we get by starting at P and displacing by some multiple of $\vec{d}$. This is a line!



We simultaneously interpret this line as a set of points (the points that make up the line) and as a set of vectors rooted at the origin (the vectors pointing from the origin to the line). Note that sometimes we draw vectors as directed line segments. Other times, we draw each vector by marking only its ending point because drawing each vector as a line segment would make it hard to see what is going on.

Which picture below do you think best represents ℓ?



© Jason Siefken, 2015–2022

Takeaway. When drawing a picture depicting several vectors, make an appropriate choice (arrows, dots, or a mix) so that the picture is clear.

The line ℓ described above can be written in set-builder notation as:

$$\ell = \{\vec{x} : \vec{x} = t\vec{d} + P \text{ for some } t \in \mathbb{R}\}.$$

Notice that in set-builder notation, we write "for some $t \in \mathbb{R}$." Make sure you understand why replacing "for some $t \in \mathbb{R}$" with "for all $t \in \mathbb{R}$" would be incorrect.

Writing lines with set-builder notation all the time can be overkill, so we will allow ourselves to describe lines in a shorthand called *vector form*.[10]

Vector Form of a Line. Let ℓ be a line and let $\vec{d}$ and $\vec{p}$ be vectors. If $\ell = \{\vec{x} : \vec{x} = t\vec{d} + \vec{p} \text{ for some } t \in \mathbb{R}\}$, we say the vector equation

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

is ℓ expressed in ***vector form***. The vector $\vec{d}$ is called a ***direction vector*** for ℓ.

We can also use coordinates when writing a line in vector form. For example,

$$\begin{bmatrix} x \\ y \end{bmatrix} = t\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

corresponds to the line passing through $\begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$ with $\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$ as a direction vector.

The "t" that appears in a vector form is called the *parameter variable*, and for this reason, some textbooks use the term *parametric form* in place of "vector form".

Writing a line in vector form requires only a point on the line and a direction for the line,[11] which makes converting from another form into vector form straightforward.

Example. Find vector form of the line $\ell \subseteq \mathbb{R}^2$ with equation $y = 2x + 3$.

First, we find two points on ℓ. By guess-and-check, we see $P = (0, 3)$ and $Q = (1, 5)$ are on ℓ. Thus, a direction vector for ℓ is given by

$$\vec{d} = \begin{bmatrix} 1 \\ 5 \end{bmatrix} - \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

We may now express ℓ in vector form as

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

or, using coordinates, as

$$\begin{bmatrix} x \\ y \end{bmatrix} = t\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

It's important to note that when we write a line in vector form, it is a *specific shorthand* notation. If we augment the notation, we no longer have written a line in "vector form".

Example. Let ℓ be a line, let $\vec{d}$ be a direction vector for ℓ, and let $\vec{p} \in \ell$ be a point on ℓ. Writing

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

or

$$\vec{x} = t\vec{d} + \vec{p} \quad \text{where} \quad t \in \mathbb{R}$$

specifies ℓ in vector form; both are shorthands for $\{\vec{x} : \vec{x} = t\vec{d} + \vec{p} \text{ for some } t \in \mathbb{R}\}$. But,

$$\vec{x} = t\vec{d} + \vec{p} \quad \text{for some} \quad t \in \mathbb{R}$$

and

$$\vec{x} = t\vec{d} + \vec{p} \quad \text{for all} \quad t \in \mathbb{R}$$

[10] $y = mx + b$ form of a line is also shorthand. The line ℓ described by the equation $y = mx + b$ is actually the set $\{(x, y) \in \mathbb{R}^2 : y = mx + b\}$.

[11] Notice that a direction vector *for* a line ℓ is different than a vector *in* a line ℓ.

© Jason Siefken, 2015–2022

are logical statements about the vectors $\vec{x}$, $\vec{d}$, and $\vec{p}$. These statements are either true or false; they do *not* specify ℓ in vector form.

Similarly, the statement

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

is mathematically nonsensical and does not specify ℓ in vector form. (On the left is a *set* and on the right is a *vector*!)

Takeaway. Vector form is a specific shorthand for a set. If "extra" words or symbols are added to the vector form, it stops being a shorthand.

But, why is vector form useful? For starters, every line can be expressed in vector form (you cannot write a vertical line in $y = mx + b$ form, and in $\mathbb{R}^3$, you would need two linear equations to represent a line). But, the most useful thing about expressing a line in vector form is that you can easily generate points on that line.

Suppose ℓ can be represented in vector form as $\vec{x} = t\vec{d} + \vec{p}$. Then, for every $t \in \mathbb{R}$, the vector $t\vec{d} + \vec{p} \in \ell$. Not only that, but as t ranges over $\mathbb{R}$, all points on ℓ are "traced out". Thus, we can find points on ℓ without having to "solve" any equations.

The downside to using vector form is that it is not unique. There are multiple direction vectors and multiple points for every line. Thus, merely by looking at the vector equation for two lines, it can be hard to tell if they're equal.

For example,

$$\begin{bmatrix} x \\ y \end{bmatrix} = t\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \qquad \begin{bmatrix} x \\ y \end{bmatrix} = t\begin{bmatrix} 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \qquad \text{and} \qquad \begin{bmatrix} x \\ y \end{bmatrix} = t\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

all represent the same line. In the second equation, the direction vector is parallel but scaled, and in the third equation, a different point on the line was chosen.

Recall that in vector form, the variable t is called the *parameter variable*. It is an instance of a *dummy variable*. In other words, t is a placeholder—just because "t" appears in two different vector forms, doesn't mean it's the same quantity.

To drive this point home, let's think about vector form in terms of the sets it specifies. Let $\vec{d}_1, \vec{d}_2 \neq \vec{0}$ and $\vec{p}_1, \vec{p}_2$ be vectors and define the lines

$$\ell_1 = \{\vec{x} \,:\, \vec{x} = t\vec{d}_1 + \vec{p}_1 \text{ for some } t \in \mathbb{R}\}$$

and

$$\ell_2 = \{\vec{x} \,:\, \vec{x} = t\vec{d}_2 + \vec{p}_2 \text{ for some } t \in \mathbb{R}\}.$$

These lines have vector forms $\vec{x} = t\vec{d}_1 + \vec{p}_1$ and $\vec{x} = t\vec{d}_2 + \vec{p}_2$. However, declaring that $\ell_1 = \ell_2$ if and only if $t\vec{d}_1 + \vec{p}_1 = t\vec{d}_2 + \vec{p}_2$ does *not* make sense. Instead, as per the definition, $\ell_1 = \ell_2$ if $\ell_1 \subseteq \ell_2$ and $\ell_2 \subseteq \ell_1$. If $\vec{x} \in \ell_1$ then $\vec{x} = t\vec{d}_1 + \vec{p}_1$ for some $t \in \mathbb{R}$. If $\vec{x} \in \ell_2$ then $\vec{x} = t\vec{d}_2 + \vec{p}_2$ for some *possibly different* $t \in \mathbb{R}$. This can get confusing really quickly. The easiest way to avoid confusion is to use different parameter variables when comparing different vector forms.

Example. Determine if the lines ℓ_1 and ℓ_2, given in vector form as

$$\vec{x} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \vec{x} = t\begin{bmatrix} 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 3 \end{bmatrix},$$

are the same line.

To determine this, we need to figure out if $\vec{x} \in \ell_1$ implies $\vec{x} \in \ell_2$ and if $\vec{x} \in \ell_2$ implies $\vec{x} \in \ell_1$.
If $\vec{x} \in \ell_1$, then $\vec{x} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ for some $t \in \mathbb{R}$. If $\vec{x} \in \ell_2$, then $\vec{x} = s\begin{bmatrix} 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ for some $s \in \mathbb{R}$. Thus if

$$t\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \vec{x} = s\begin{bmatrix} 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

© Jason Siefken, 2015–2022

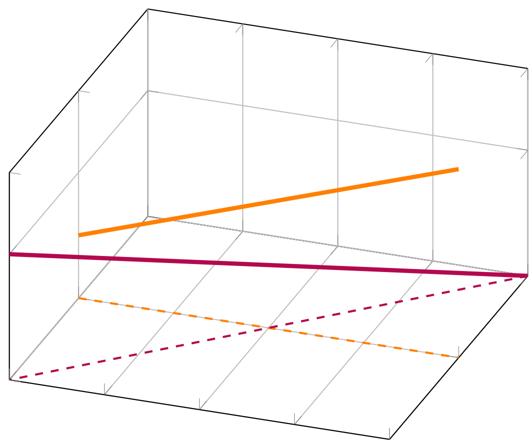
always has a solution, $\ell_1 = \ell_2$. Moving everything to one side, we see

$$\begin{aligned}\vec{0} = \begin{bmatrix}4\\3\end{bmatrix} - \begin{bmatrix}2\\1\end{bmatrix} + s\begin{bmatrix}2\\2\end{bmatrix} - t\begin{bmatrix}1\\1\end{bmatrix} &= \begin{bmatrix}2\\2\end{bmatrix} + s\begin{bmatrix}2\\2\end{bmatrix} - t\begin{bmatrix}1\\1\end{bmatrix}\\ &= (s+1)\begin{bmatrix}2\\2\end{bmatrix} - \frac{t}{2}\begin{bmatrix}2\\2\end{bmatrix}\\ &= (s+1-\frac{t}{2})\begin{bmatrix}2\\2\end{bmatrix}.\end{aligned}$$

This equation has a solution whenever $s+1-t/2=0$ has a solution. Since for every s, the equation $s+1-t/2=0$ has a solution, and for every t, the equation $s+1-t/2=0$ has a solution, we know $\ell_1 = \ell_2$.

Vector Form in Higher Dimensions

The geometry of lines in space ($\mathbb{R}^3$ and above) is a bit more complicated than that of lines in the plane. Lines in the plane either intersect or are parallel. In space, we have to be careful about what we mean by "parallel lines," since lines with entirely different directions can still fail to intersect.[12]



Example. Consider the lines described by

$$\begin{aligned}\vec{x} &= t(1,3,-2)+(1,2,1)\\ \vec{x} &= t(-2,-6,4)+(3,1,0).\end{aligned}$$

They have parallel directions since $(-2,-6,4) = -2(1,3,-2)$. Hence, in this case, we say the lines are ***parallel***. (How can we be sure the lines are not the same?)

Example. Consider the lines described by

$$\begin{aligned}\vec{x} &= t(1,3,-2)+(1,2,1)\\ \vec{x} &= t(0,2,3)+(0,3,9).\end{aligned}$$

They are not parallel because neither of the direction vectors is a multiple of the other. They may or may not intersect. (If they don't, we say the lines are ***skew***.) How can we find out? Mirroring our earlier approach, we can set their equations equal and see if we can solve for a point of intersection ***after ensuring we give their parametric variables different names***. We'll keep one parametric variable named t and name the other one s. Thus, we want

$$\vec{x} = t(1,3,-2)+(1,2,1) = s(0,2,3)+(0,3,9),$$

which after collecting terms yields

$$(t+1, 3t+2, -2t+1) = (0, 2s+3, 3s+9).$$

[12]Recall that in Euclidean geometry two lines are defined to be parallel if they coincide or never intersect.

© Jason Siefken, 2015–2022

Reading coordinate by coordinate, we get three equations

$$\begin{aligned} t+1 &= 0 \\ 3t+2 &= 2s+3 \\ -2t+1 &= 3s+9 \end{aligned}$$

in two unknowns s and t. This is an *overdetermined* system, and it may or may not have a solution. The first two equations yield $t=-1$ and $s=-2$. Putting these values in the last equation yields $(-2)(-1)+1=3(-2)+9$, which is indeed true. Hence, the equations are consistent, and the lines intersect. To find the point of intersection, put $t=-1$ in the equation for the vector equation of the first line (or $s=-2$ in that for the second) to obtain $(0,-1,3)$.

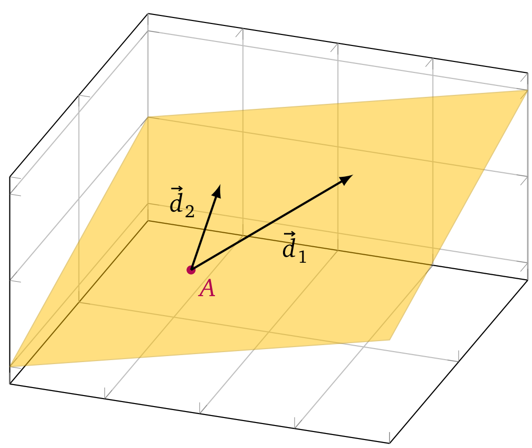
Planes

Any two distinct points define a line. To define a plane, we need three points. But there's a caveat: the three points cannot be on the same line, otherwise they'd define a line and not a plane. Let $A, B, C \in \mathbb{R}^3$ be three points that are not collinear and let $\mathcal{P}$ be the plane that passes through A, B, and C.

Just like lines, planes have direction vectors. For $\mathcal{P}$, both $\vec{d}_1 = \overrightarrow{AB}$ and $\vec{d}_2 = \overrightarrow{AC}$ are direction vectors. Of course, $\vec{d}_1$, $\vec{d}_2$ and their multiples are not the only direction vectors for $\mathcal{P}$. There are infinitely many more, including $\vec{d}_1+\vec{d}_2$, and $\vec{d}_1-7\vec{d}_2$, and so on. However, since a plane is a *two*-dimensional object, we only need two different direction vectors to describe it.

Like lines, planes have a vector form. Using the direction vectors $\vec{d}_1 = \overrightarrow{AB}$ and $\vec{d}_2 = \overrightarrow{AC}$, the plane $\mathcal{P}$ can be written in vector form as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t\vec{d}_1 + s\vec{d}_2 + A.$$



Vector Form of a Plane. 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 $\mathcal{P}$.

Example. Describe the plane $\mathcal{P} \subseteq \mathbb{R}^3$ with equation $z = 2x + y + 3$ in vector form.

To describe $\mathcal{P}$ in vector form, we need a point on $\mathcal{P}$ and two direction vectors for $\mathcal{P}$. By guess-and-check, we see the points

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

© Jason Siefken, 2015–2022

are all in $\mathcal{P}$. Thus

$$\vec{d}_1 = B - A = \begin{bmatrix}1\\0\\2\end{bmatrix} \qquad \text{and} \qquad \vec{d}_2 = C - A = \begin{bmatrix}0\\1\\1\end{bmatrix}$$

are both direction vectors for $\mathcal{P}$. Since these vectors are not parallel, we can express $\mathcal{P}$ in vector form as

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

Example. Find the line of intersection between $\mathcal{P}_1$ and $\mathcal{P}_2$ where the planes are given in vector form by

$$\overbrace{\vec{x} = t\begin{bmatrix}1\\1\\0\end{bmatrix} + s\begin{bmatrix}-1\\0\\1\end{bmatrix} + \begin{bmatrix}1\\2\\3\end{bmatrix}}^{\mathcal{P}_1} \qquad \text{and} \qquad \overbrace{\vec{x} = t\begin{bmatrix}-1\\0\\2\end{bmatrix} + s\begin{bmatrix}1\\2\\1\end{bmatrix} + \begin{bmatrix}0\\0\\3\end{bmatrix}}^{\mathcal{P}_2}.$$

Just like in the example for lines, we are looking for points $\vec{x}$ that are in both planes. To keep from getting mixed up, we'll use a, b, c, and d as parameter variables. Therefore, we are looking for solutions to

$$a\begin{bmatrix}1\\1\\0\end{bmatrix} + b\begin{bmatrix}-1\\0\\1\end{bmatrix} + \begin{bmatrix}1\\2\\3\end{bmatrix} = \vec{x} = c\begin{bmatrix}-1\\0\\2\end{bmatrix} + d\begin{bmatrix}1\\2\\1\end{bmatrix} + \begin{bmatrix}0\\0\\3\end{bmatrix}.$$

Collecting terms, this is equivalent to the system of equations

$$\begin{cases} a - b + \ \ c - \ \ d = -1 \\ a \qquad\qquad - 2d = -2 \\ \qquad b - 2c - \ \ d = 0 \end{cases}.$$

This system is underdetermined (there are four variables and three equations). If $\mathcal{P}_1$ and $\mathcal{P}_2$ indeed intersect in a line, we know there must be an infinite number of solutions to this system. After row reducing,[a] we see

$$\begin{bmatrix}a\\b\\c\\d\end{bmatrix} = \begin{bmatrix}r\\r/2-1\\-1\\r/2+1\end{bmatrix}$$

is a solution for every $r \in \mathbb{R}$. We can substitute these parameters into either of the original equations to get an equation for the line of intersection. Picking the second one, we see

$$\vec{x} = c\begin{bmatrix}-1\\0\\2\end{bmatrix} + d\begin{bmatrix}1\\2\\1\end{bmatrix} + \begin{bmatrix}0\\0\\3\end{bmatrix} = -\begin{bmatrix}-1\\0\\2\end{bmatrix} + (\frac{r}{2}+1)\begin{bmatrix}1\\2\\1\end{bmatrix} + \begin{bmatrix}0\\0\\3\end{bmatrix} = \frac{r}{2}\begin{bmatrix}1\\2\\1\end{bmatrix} + \begin{bmatrix}2\\2\\2\end{bmatrix}$$

is in both planes for every $r \in \mathbb{R}$. Therefore, we may express $\mathcal{P}_1 \cap \mathcal{P}_2$ in vector form as

$$\vec{x} = r\begin{bmatrix}1/2\\1\\1/2\end{bmatrix} + \begin{bmatrix}2\\2\\2\end{bmatrix}.$$

[a] See Appendix 1, System (14) and (15).

Restricted Linear Combinations

Using vectors, we can describe more than just lines and planes—we can describe all sorts of geometric objects.

Recall that when we write $\vec{x} = t\vec{d} + \vec{p}$ to describe the line ℓ, what we mean is

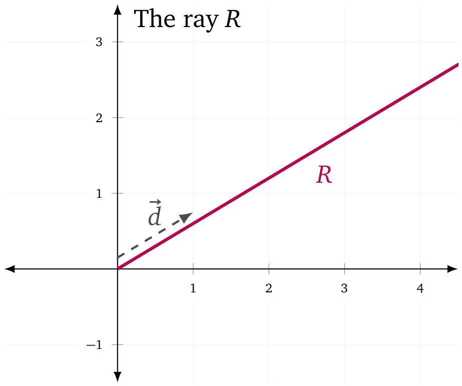
$$\ell = \{\vec{x} \,:\, \vec{x} = t\vec{d} + \vec{p} \text{ for some } t \in \mathbb{R}\}.$$

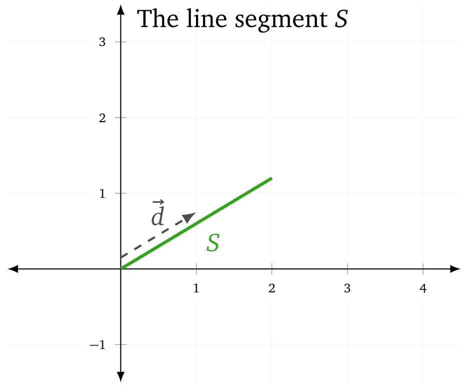
The line ℓ stretches off infinitely in both directions. But, what if we wanted to describe just a part of ℓ? We can

© Jason Siefken, 2015–2022

do this by placing additional restrictions on t. For example, consider the ray R and the line segment S:

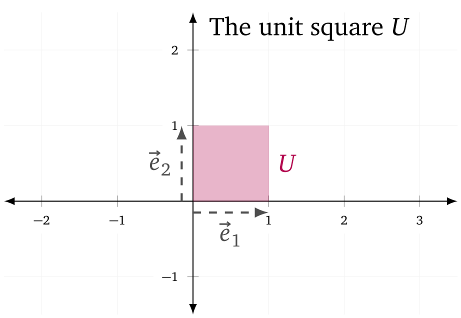
$$R = \{\vec{x} : \vec{x} = t\vec{d} + \vec{p} \text{ for some } t \geq 0\}$$
$$S = \{\vec{x} : \vec{x} = t\vec{d} + \vec{p} \text{ for some } t \in [0, 2]\}$$

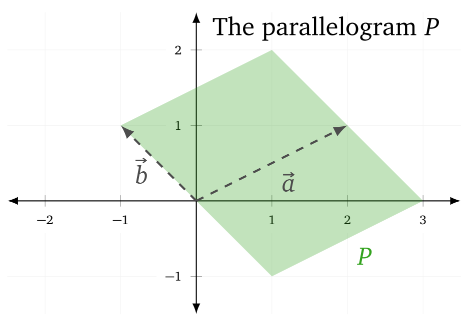




We can also make polygons by adding restrictions to the vector form of a plane. Let $\vec{a} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and consider the unit square U and the parallelogram P defined by

$$U = \{\vec{x} : \vec{x} = t\vec{e}_1 + s\vec{e}_2 \text{ for some } t, s \in [0, 1]\}$$
$$P = \{\vec{x} : \vec{x} = t\vec{a} + s\vec{b} \text{ for some } t \in [0, 1] \text{ and } s \in [-1, 1]\}$$





Each set so far is a set of linear combinations, and we have made different shapes by restricting the coefficients of those linear combinations. There are two ways of restricting linear combinations that arise often enough to get their own names.

Non-negative & Convex Linear Combinations.
Let $\vec{w} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n$. The vector $\vec{w}$ is called a ***non-negative*** linear combination of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ if

$$\alpha_1, \alpha_2, \ldots, \alpha_n \geq 0.$$

The vector $\vec{w}$ is called a ***convex*** linear combination of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ if

$$\alpha_1, \alpha_2, \ldots, \alpha_n \geq 0 \qquad \text{and} \qquad \alpha_1 + \alpha_2 + \cdots + \alpha_n = 1.$$

You can think of non-negative linear combinations as vectors you can arrive at by only displacing "forward". Convex linear combinations can be thought of as weighted averages of vectors (the average of $\vec{v}_1, \ldots, \vec{v}_n$ would be the convex linear combination with coefficients $\alpha_i = \frac{1}{n}$). A convex linear combination of two vectors gives a point on the line segment connecting them.

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

$$\begin{aligned} A &= \{\vec{x} : \vec{x} \text{ is a convex linear combination of } \vec{a} \text{ and } \vec{b}\} \\ &= \{\vec{x} : \vec{x} = \alpha\vec{a} + (1-\alpha)\vec{b} \text{ for some } \alpha \in [0, 1]\}. \end{aligned}$$

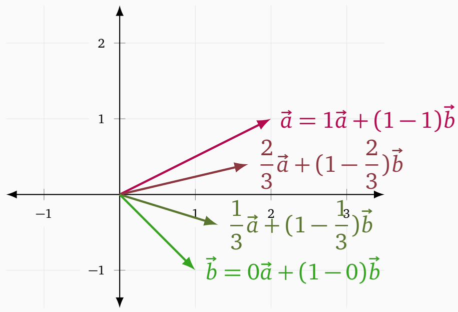
Draw A.

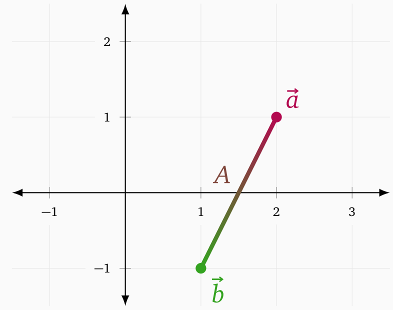
© Jason Siefken, 2015–2022

We know $\vec{x} = \alpha\vec{a} + (1-\alpha)\vec{b} \in A$ whenever $\alpha \in [0,1]$. If we rearrange the equation $\vec{x} = \alpha\vec{a} + (1-\alpha)\vec{b}$, we see

$$\vec{x} = \alpha\vec{a} - \alpha\vec{b} + \vec{b} = \alpha(\vec{a} - \vec{b}) + \vec{b},$$

which looks like the vector form of a line which passes through $\vec{b}$ with direction $\vec{a} - \vec{b}$. However, we have the additional restriction $\alpha \in [0,1]$, so A is only the part of that line which connects $\vec{a}$ and $\vec{b}$.





Since A is an infinite collection of vectors, it's better to draw vectors in A as dots rather than lines from the origin.

Practice Problems

1 Express the following lines in vector form.

(a) $\ell_1 \subseteq \mathbb{R}^2$ with equation $4x - 3y = -10$.

(b) $\ell_2 \subseteq \mathbb{R}^2$ which passes through the points $A = (1,1)$ and $B = (2,7)$.

(c) $\ell_3 \subseteq \mathbb{R}^2$ which passes through $\vec{0}$ and is parallel to the line with equation $4x - 3y = -10$.

(d) $\ell_4 \subseteq \mathbb{R}^3$ which passes through the points $A = (-1,-1,0)$ and $B = (2,3,5)$.

(e) $\ell_5 \subseteq \mathbb{R}^3$ which is contained in the yz-plane and where the coordinates of every point in ℓ_5 satisfy $x + 2y - 3z = 5$.

2 Express the following planes in vector form.

(a) $\mathcal{P}_1 \subseteq \mathbb{R}^3$ with equation $4x - z = 0$.

(b) $\mathcal{P}_2 \subseteq \mathbb{R}^3$ which passes through the points $A = (-1,-1,0)$, $B = (2,3,5)$, and $C = (3,3,3)$.

(c) $\mathcal{P}_3 \subseteq \mathbb{R}^3$ with equation $4x - 3y + z = -10$.

(d) $\mathcal{P}_4 \subseteq \mathbb{R}^3$ which is parallel to the yz-plane and passes through the point $X = (1,-1,1)$.

(e) $\mathbb{R}^2$.

(f) $\mathcal{P}_5 \subseteq \mathbb{R}^4$ which passes through $A = (1,-1,1,-1)$, and where the coordinates of every point in $\mathcal{P}_5$ satisfy the equations $x + y + 2z - w = 3$ and $x + y + z + w = 0$.

3 Let ℓ_1, ℓ_2, and ℓ_3 be described in vector form by

$$\overbrace{\vec{x} = t\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}1\\3\end{bmatrix}}^{\ell_1} \qquad \overbrace{\vec{x} = t\begin{bmatrix}1\\3\end{bmatrix} + \begin{bmatrix}1\\1\end{bmatrix}}^{\ell_2} \qquad \overbrace{\vec{x} = t\begin{bmatrix}2\\2\end{bmatrix} + \begin{bmatrix}2\\4\end{bmatrix}}^{\ell_3}.$$

(a) Determine which pairs of the lines ℓ_1, ℓ_2, and ℓ_3 intersect, coincide, or are parallel.

(b) What is $\ell_1 \cap \ell_2 \cap \ell_3$?

4 Let $\mathcal{P}_1 \subseteq \mathbb{R}^3$ be the plane with equation $x + 2y - z = 3$. Let $\mathcal{P}_2$ and ℓ be described in vector form by

$$\overbrace{\vec{x} = t\begin{bmatrix}1\\1\\1\end{bmatrix} + s\begin{bmatrix}0\\0\\2\end{bmatrix} + \begin{bmatrix}1\\3\\1\end{bmatrix}}^{\mathcal{P}_2}, \qquad \overbrace{\vec{x} = t\begin{bmatrix}1\\3\\1\end{bmatrix} + \begin{bmatrix}1\\1\\0\end{bmatrix}}^{\ell}.$$

(a) Find $\mathcal{P}_1 \cap \ell$.

(b) Find $\mathcal{P}_1 \cap \mathcal{P}_2$.

(c) Find $\mathcal{P}_2 \cap \ell$.

(d) Give an example of a plane $\mathcal{P}_3$ so that $\mathcal{P}_3 \cap \ell$ is empty.

(e) Does there exist a plane $\mathcal{P}_2'$ that is parallel to $\mathcal{P}_2$, but which does not intersect ℓ? Why or why not?

5 Let $\vec{a} = \begin{bmatrix}1\\1\end{bmatrix}$ and $\vec{b} = \begin{bmatrix}1\\-1\end{bmatrix}$. The goal of this question is to produce a drawing of the set of convex linear combinations of $\vec{a}$ and $\vec{b}$.

(a) Let A be the set of all non-negative linear combinations of $\vec{a}$ and $\vec{b}$. Draw A.

(b) Let ℓ be the set

$$\{\alpha\vec{a} + \beta\vec{b} \,:\, \alpha, \beta \in \mathbb{R} \text{ and } \alpha + \beta = 1\}$$

Rewrite ℓ in set-builder notation using only a single variable t. (Hint: Let t be α.)

(c) Justify why ℓ is a line, and write ℓ in vector form.

(d) Draw both A and ℓ on the same grid. On a separate grid, draw $A \cap \ell$.

(e) Write the $A \cap \ell$ in set-builder notation. How does $A \cap \ell$ relate to convex linear combinations?

(f) Determine the endpoints of $A \cap \ell$.

6 Let $\vec{a} = \begin{bmatrix}2\\0\end{bmatrix}$, $\vec{b} = \begin{bmatrix}0\\2\end{bmatrix}$, and $\vec{c} = \begin{bmatrix}-1\\-1\end{bmatrix}$. The goal of this question is to produce a drawing of the set of convex linear combinations of $\vec{a}$, $\vec{b}$, and $\vec{c}$. This requires an understanding of the previous question.

(a) Let $\vec{d} = \begin{bmatrix}1\\1\end{bmatrix}$. Write $\vec{d}$ as a convex linear combination of $\vec{a}$ and $\vec{b}$.

© Jason Siefken, 2015–2022

(b) Let $\vec{e} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Write $\vec{e}$ as a convex linear combination of $\vec{c}$ and $\vec{d}$.

(c) Substituting the answer to (6a) into the answer to part (6b), write $\vec{e}$ as a convex linear combination of $\vec{a}$, $\vec{b}$, and $\vec{c}$.

(d) Draw and label $\vec{a}$, $\vec{b}$, $\vec{c}$, $\vec{d}$, and $\vec{e}$ on the same grid.

(e) Draw the set of convex linear combinations of $\vec{a}$, $\vec{b}$, and $\vec{c}$. Justify your answer.

7 Let $\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\vec{y} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ and $\vec{z} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$. Draw the following subsets of $\mathbb{R}^2$.

(a) All non-negative linear combinations of $\vec{x}$ and $\vec{y}$.

(b) All non-negative linear combinations of $\vec{x}$ and $\vec{z}$.

(c) All convex linear combinations of $\vec{y}$ and $\vec{z}$.

(d) All convex linear combinations of $\vec{x}$ and $\vec{z}$.

(e) All convex linear combinations of $\vec{x}$, $\vec{y}$ and $\vec{z}$.

8 Describe the sets in (7c) and (7d) in set-builder notation.

9 Determine if the points $P = (-2, 0)$ and $Q = (0, -2)$ are convex linear combinations of the vectors $\vec{u} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} -5 \\ 8 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} -2 \\ -6 \end{bmatrix}$. First solve this question by drawing a picture. Then justify algebraically.

© Jason Siefken, 2015–2022 CC BY SA

© Jason Siefken, 2015–2022

Non-negative & Convex Linear Combinations

DEFINITION

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

$$\alpha_1, \alpha_2, \dots, \alpha_n \geq 0.$$

The vector $\vec{w}$ is called a ***convex*** linear combination of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ if

$$\alpha_1, \alpha_2, \dots, \alpha_n \geq 0 \qquad \text{and} \qquad \alpha_1 + \alpha_2 + \cdots + \alpha_n = 1.$$

---

9 Let

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

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

(a) linear combinations of $\vec{a}$ and $\vec{b}$?

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

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

9.2 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.

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

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

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

Otherwise, explain why it's not possible.

© Jason Siefken, 2015–2022

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

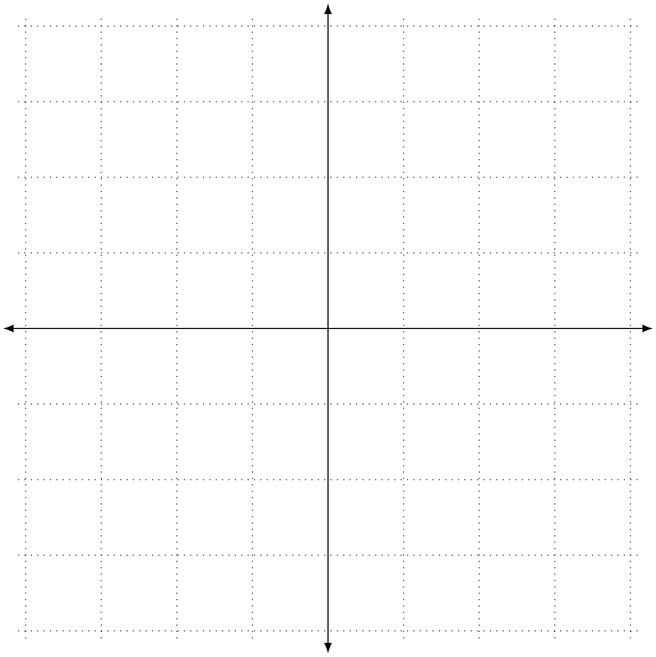
10.1 Describe L using set-builder notation.

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

10.3 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.

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

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



© Jason Siefken, 2015–2022

DEFINITION

Vector Form of a Line

Let ℓ be a line and let $\vec{d}$ and $\vec{p}$ be vectors. If $\ell = \{\vec{x} : \vec{x} = t\vec{d} + \vec{p} \text{ for some } t \in \mathbb{R}\}$, we say the vector equation

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

is ℓ expressed in ***vector form***. The vector $\vec{d}$ is called a ***direction vector*** for ℓ.

---

11 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 $z = y$.

11.1 Write ℓ in vector form. Is vector form of ℓ unique?

11.2 Write L in vector form.

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

© Jason Siefken, 2015–2022 CC BY SA

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

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

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

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

12.3 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?

© Jason Siefken, 2015–2022

DEFINITION

Vector Form of a Plane

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 $\mathcal{P}$.

13 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.

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

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

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

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

© Jason Siefken, 2015–2022

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

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

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

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

© Jason Siefken, 2015–2022

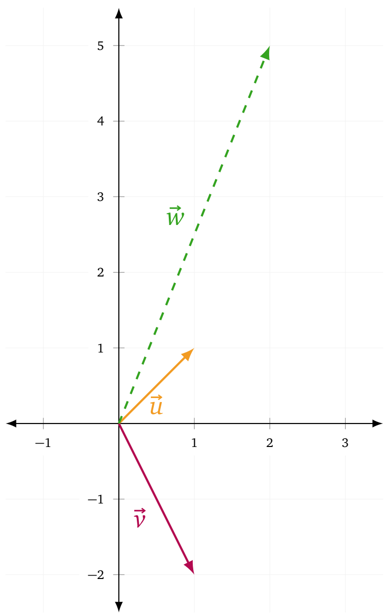
Spans, Translated Spans, and Linear Independence/Dependence

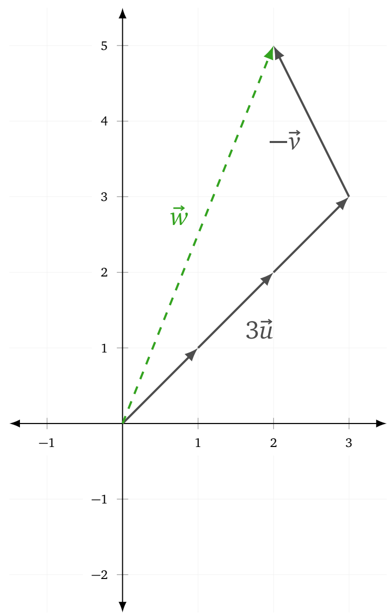
In this module you will learn

- The definition of span and how to visualize spans.
- How to express lines/planes/volumes through the origin as spans.
- How to express lines/planes/volumes *not* through the origin as *translated* spans using set addition.
- Geometric and algebraic definitions of linear independence and linear dependence.
- How to find linearly independent subsets.

Let $\vec{u} = \begin{bmatrix}1\\1\end{bmatrix}$ and $\vec{v} = \begin{bmatrix}1\\-2\end{bmatrix}$. Can the vector $\vec{w} = \begin{bmatrix}2\\5\end{bmatrix}$ be obtained as a linear combination of $\vec{u}$ and $\vec{v}$?

By drawing a picture, the answer appears to be *yes*.





Algebraically, we can use the definition of a *linear combination* to set up a system of equations. We know $\vec{w}$ can be expressed as a linear combination of $\vec{u}$ and $\vec{v}$ if and only if the vector equation

$$\vec{w} = \begin{bmatrix}2\\5\end{bmatrix} = \alpha\begin{bmatrix}1\\1\end{bmatrix} + \beta\begin{bmatrix}1\\-2\end{bmatrix} = \alpha\vec{u} + \beta\vec{v}$$

has a solution. By inspection, we see $\alpha = 3$ and $\beta = -1$ solve this equation.

After initial success, we might ask the following: *what are all the locations in $\mathbb{R}^2$ that can be obtained as a linear combination of $\vec{u}$ and $\vec{v}$?* Geometrically, it appears any location can be reached. To verify this algebraically, consider the vector equation

$$\vec{x} = \begin{bmatrix}x\\y\end{bmatrix} = \alpha\begin{bmatrix}1\\1\end{bmatrix} + \beta\begin{bmatrix}1\\-2\end{bmatrix} = \alpha\vec{u} + \beta\vec{v}. \tag{1}$$

Here $\vec{x}$ represents an arbitrary point in $\mathbb{R}^2$. If equation (1) always has a solution,[13] any vector in $\mathbb{R}^2$ can be obtained as a linear combination of $\vec{u}$ and $\vec{v}$.

We can solve this equation for α and β by considering the equations arising from the first and second coordinates. Namely,

$$\begin{aligned} x &= \alpha + \beta \\ y &= \alpha - 2\beta \end{aligned}$$

[13] The official terminology would be to say that the equations are always *consistent*.

© Jason Siefken, 2015–2022

Subtracting the second equation from the first, we get $x - y = 3\beta$ and so $\beta = (x-y)/3$. Plugging β into the first equation and solving, we get $\alpha = (2x+y)/3$. Thus, equation (1) *always* has the solution

$$\begin{aligned}\alpha &= \frac{1}{3}(2x+y)\\ \beta &= \frac{1}{3}(x-y).\end{aligned}$$

There is a formal term for the set of vectors that can be obtained as linear combinations of others: *span*.

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

$$\operatorname{span} V = \{\vec v : \vec v = \alpha_1\vec v_1 + \alpha_2\vec v_2 + \cdots + \alpha_n\vec v_n \text{ for some } \vec v_1, \vec v_2, \ldots, \vec v_n \in V \text{ and scalars } \alpha_1, \alpha_2, \ldots, \alpha_n\}.$$

Additionally, we define $\operatorname{span}\{\} = \{\vec 0\}$.

We just showed above that $\operatorname{span}\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-2\end{bmatrix}\right\} = \mathbb{R}^2$. Alternatively, we may use span as a verb and say the set $\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-2\end{bmatrix}\right\}$ *spans* $\mathbb{R}^2$.

Example. Let $\vec u = \begin{bmatrix}-1\\2\end{bmatrix}$ and $\vec v = \begin{bmatrix}1\\-2\end{bmatrix}$. Find $\operatorname{span}\{\vec u, \vec v\}$.

By the definition of span,

$$\operatorname{span}\{\vec u, \vec v\} = \{\vec x : \vec x = \alpha\vec u + \beta\vec v \text{ for some } \alpha, \beta \in \mathbb{R}\}.$$

We need to determine for which x and y the vector equation $\begin{bmatrix}x\\y\end{bmatrix} = \alpha\begin{bmatrix}-1\\2\end{bmatrix} + \beta\begin{bmatrix}1\\-2\end{bmatrix}$ is consistent. From the first and second coordinates, we get the system

$$\begin{aligned}x &= -\alpha + \beta\\ y &= 2\alpha - 2\beta\end{aligned}.$$

Adding 2 times the first equation to the second, we get $2x + y = 0$ and so $y = -2x$. Therefore, if $\begin{bmatrix}x\\y\end{bmatrix}$ makes the above system consistent, we must have

$$\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}t\\-2t\end{bmatrix} = t\vec v$$

for some t. Thus,

$$\operatorname{span}\{\vec u, \vec v\} = \{\vec x : \vec x = t\vec v \text{ for some } t\} = \operatorname{span}\{\vec v\},$$

which is a line through the origin with direction $\vec v$.

Example. Let $\vec a = \begin{bmatrix}1\\2\\1\end{bmatrix}$, $\vec b = \begin{bmatrix}0\\1\\0\end{bmatrix}$, and $\vec c = \begin{bmatrix}1\\1\\2\end{bmatrix}$. Show that $\mathbb{R}^3 = \operatorname{span}\{\vec a, \vec b, \vec c\}$.

If the equation

$$\vec x = \begin{bmatrix}x\\y\\z\end{bmatrix} = \alpha_1\begin{bmatrix}1\\2\\1\end{bmatrix} + \alpha_2\begin{bmatrix}0\\1\\0\end{bmatrix} + \alpha_3\begin{bmatrix}1\\1\\2\end{bmatrix} = \alpha_1\vec a + \alpha_2\vec b + \alpha_3\vec c$$

is always consistent, then any vector in $\mathbb{R}^3$ can be obtained as a linear combination of $\vec a$, $\vec b$, and $\vec c$. Reading off the coordinates, we get the system

$$\begin{aligned}x &= \alpha_1 + \alpha_3\\ y &= 2\alpha_1 + \alpha_2 + \alpha_3\\ z &= \alpha_1 + 2\alpha_3\end{aligned}.$$

© Jason Siefken, 2015–2022

Solving this system, we see

$$\begin{aligned}\alpha_1 &= 2x - z\\ \alpha_2 &= -3x + y + z\\ \alpha_3 &= -x + z\end{aligned}$$

is always a solution (no matter the values of x, y, and z). Therefore, $\operatorname{span}\{\vec{a}, \vec{b}, \vec{c}\} = \mathbb{R}^3$.

Representing Lines & Planes as Spans

If spans remind you of vector forms of lines and planes, your intuition is keen. Consider the line ℓ given in vector form by

$$\vec{x} = t\vec{d} + \vec{0}.$$

The line ℓ passes through the origin, and if we unravel its definition, we see

$$\ell = \{\vec{x} : \vec{x} = t\vec{d} + \vec{0} \text{ for some } t \in \mathbb{R}\} = \{\vec{x} : \vec{x} = t\vec{d} \text{ for some } t \in \mathbb{R}\} = \operatorname{span}\{\vec{d}\}.$$

Similarly, if $\mathcal{P}$ is a plane given in vector form by

$$\vec{x} = t\vec{d}_1 + s\vec{d}_2 + \vec{0},$$

then

$$\mathcal{P} = \{\vec{x} : \vec{x} = t\vec{d}_1 + s\vec{d}_2 \text{ for some } t, s \in \mathbb{R}\} = \operatorname{span}\{\vec{d}_1, \vec{d}_2\}.$$

If the "$\vec{p}$" in our vector form is $\vec{0}$, then that vector form actually defines a span. This means (if you accept that every line/plane through the origin has a vector form) that every line/plane through the origin can be written as a span. Conversely, if $X = \operatorname{span}\{\vec{v}_1, \dots, \vec{v}_n\}$ is a span, we know $\vec{0} = 0\vec{v}_1 + \cdots + 0\vec{v}_n \in X$, and so every span passes through the origin.

As it turns out, spans exactly describe points, lines, planes, and volumes[14] through the origin.

Example. The line $\ell_1 \subseteq \mathbb{R}^2$ is described by the equation $x + 2y = 0$ and the line $\ell_2 \subseteq \mathbb{R}^2$ is described by the equation $4x - 2y = 6$. If possible, describe ℓ_1 and ℓ_2 using spans.

We can express ℓ_1 in vector form by

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

and so

$$\ell_1 = \operatorname{span}\left\{\begin{bmatrix}2\\-1\end{bmatrix}\right\}.$$

However, ℓ_2 does not pass through $\vec{0}$, and so ℓ_2 cannot be written as a span.

Takeaway. Lines and planes through the origin, and only lines and planes through the origin, can be expressed as spans.

Set Addition

We're going to work around the fact that only objects which pass through the origin can be written as spans, but first let's take a detour and learn about *set addition*.

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

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

Set sums are very different than regular sums despite using the same symbol, "+".[15] However, they are very useful. Let $C = \{\vec{x} \in \mathbb{R}^2 : \|\vec{x}\| = 1\}$ be the unit circle centered at the origin, and consider the sets

$$X = C + \{\vec{e}_2\} \qquad Y = C + \{3\vec{e}_1, \vec{e}_2\} \qquad Z = C + \{\vec{0}, \vec{e}_1, \vec{e}_2\}.$$

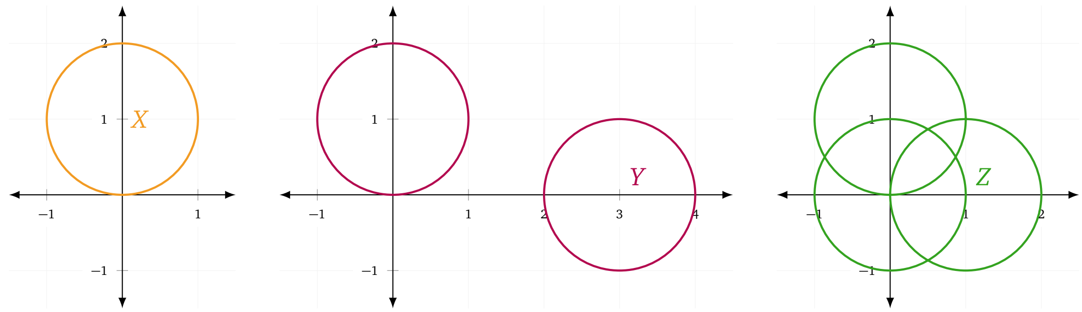
Rewriting, we see $X = \{\vec{x} + \vec{e}_2 : \|\vec{x}\| = 1\}$ is just C translated by $\vec{e}_2$. Similarly, $Y = \{\vec{x} + \vec{v} : \|\vec{x}\| = 1 \text{ and } \vec{v} = 3\vec{e}_1 \text{ or } \vec{v} = \vec{e}_2\} = (C + \{3\vec{e}_1\}) \cup (C + \{\vec{e}_2\})$, and so Y is the union of two translated copies of C.[16] Finally, Z is the union of three translated copies of C.

[14] We use the word *volume* to indicate the higher-dimensional analogue of a plane.

[15] For example, $A + \{\} = \{\}$, which might seem counterintuitive for an "addition" operation.

[16] If you want to stretch your mind, consider what $C + C$ is as a set.

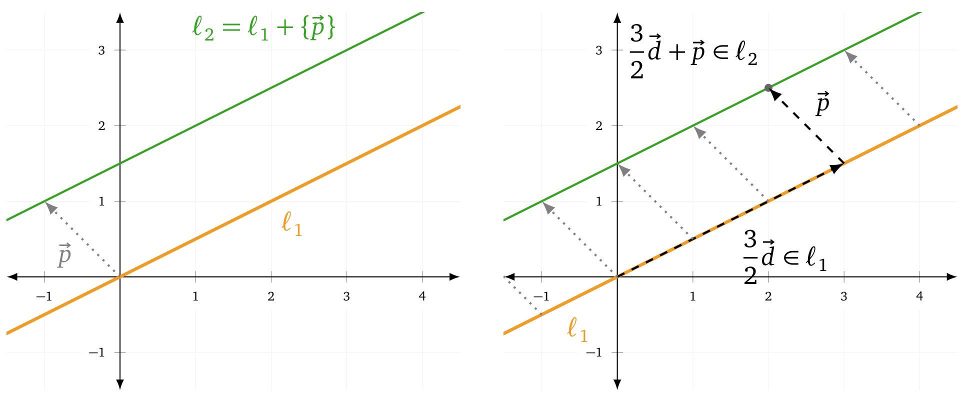
© Jason Siefken, 2015–2022



Translated Spans

Set addition allows us to easily create parallel lines and planes by translation. For example, consider the lines ℓ_1 and ℓ_2 given in vector form as $\vec{x} = t\vec{d}$ and $\vec{x} = t\vec{d} + \vec{p}$, respectively, where $\vec{d} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\vec{p} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. These lines differ from each other by a translation. That is, every point in ℓ_2 can be obtained by adding $\vec{p}$ to a corresponding point in ℓ_1. Using the idea of set addition, we can express this relationship by the equation

$$\ell_2 = \ell_1 + \{\vec{p}\}.$$



Note: it would be incorrect to write "$\ell_2 = \ell_1 + \vec{p}$". Because ℓ_1 is a set and $\vec{p}$ is not a set, "$\ell_1 + \vec{p}$" does not make mathematical sense.

Example. Recall $\ell_2 \subseteq \mathbb{R}^2$ is the line described by the equation $4x - 2y = 6$. Describe ℓ_2 as a translated span.

We can express ℓ_2 in vector form with the equation

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

Therefore,

$$\ell_2 = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\} + \left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}.$$

We can now see translated spans provide an alternative notation to vector form for specifying lines and planes. If Q is described in vector form by

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

then

$$Q = \operatorname{span}\{\vec{d}_1, \vec{d}_2\} + \{\vec{p}\}.$$

Takeaway. All lines and planes, whether through the origin or not, can be expressed as translated spans.

Linear Independence & Linear Dependence

Let

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

© Jason Siefken, 2015–2022

Since $\vec{w} = \vec{u} + \vec{v}$, we know that $\vec{w} \in \operatorname{span}\{\vec{u}, \vec{v}\}$. Geometrically, this is also clear because $\operatorname{span}\{\vec{u}, \vec{v}\}$ is the xy-plane in $\mathbb{R}^3$ and $\vec{w}$ lies on that plane.

What about $\operatorname{span}\{\vec{u}, \vec{v}, \vec{w}\}$? Intuitively, since $\vec{w}$ is already a linear combination of $\vec{u}$ and $\vec{v}$, we can't get anywhere *new* by taking linear combinations of $\vec{u}$, $\vec{v}$, *and* $\vec{w}$ compared to linear combinations of just $\vec{u}$ and $\vec{v}$. So $\operatorname{span}\{\vec{u}, \vec{v}\} = \operatorname{span}\{\vec{u}, \vec{v}, \vec{w}\}$.

Can we prove this from the definitions? Yes! Suppose $\vec{r} \in \operatorname{span}\{\vec{u}, \vec{v}, \vec{w}\}$. By definition,

$$\vec{r} = \alpha\vec{u} + \beta\vec{v} + \gamma\vec{w}$$

for some $\alpha, \beta, \gamma \in \mathbb{R}$. Since $\vec{w} = \vec{u} + \vec{v}$, we see

$$\vec{r} = \alpha\vec{u} + \beta\vec{v} + \gamma(\vec{u} + \vec{v}) = (\alpha + \gamma)\vec{u} + (\beta + \gamma)\vec{v} \in \operatorname{span}\{\vec{u}, \vec{v}\}.$$

Thus, $\operatorname{span}\{\vec{u}, \vec{v}, \vec{w}\} \subseteq \operatorname{span}\{\vec{u}, \vec{v}\}$. Conversely, if $\vec{s} \in \operatorname{span}\{\vec{u}, \vec{v}\}$, by definition,

$$\vec{s} = a\vec{u} + b\vec{v} = a\vec{u} + b\vec{v} + 0\vec{w}$$

for some $a, b \in \mathbb{R}$, and so $\vec{s} \in \operatorname{span}\{\vec{u}, \vec{v}, \vec{w}\}$. Thus $\operatorname{span}\{\vec{u}, \vec{v}\} \subseteq \operatorname{span}\{\vec{u}, \vec{v}, \vec{w}\}$. We conclude $\operatorname{span}\{\vec{u}, \vec{v}\} = \operatorname{span}\{\vec{u}, \vec{v}, \vec{w}\}$.

In this case, $\vec{w}$ was a redundant vector—it wasn't needed for the span. When a set contains a redundant vector, we call the set *linearly dependent*.

Linearly Dependent & Independent (Geometric).
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***.

We will also refer to sets of vectors (for example $\{\vec{v}_1, \dots, \vec{v}_n\}$) as being linearly independent or linearly dependent. For technical reasons, we didn't state the definition in terms of sets.[17]

The geometric definition of linear dependence says that the vectors $\vec{v}_1, \dots, \vec{v}_n$ are linearly dependent if you can remove at least one vector without changing the span. In other words, $\vec{v}_1, \dots, \vec{v}_n$ are linearly dependent *if there is a redundant vector*.

Example. Let $\vec{a} = \begin{bmatrix}1\\2\end{bmatrix}$, $\vec{b} = \begin{bmatrix}2\\3\end{bmatrix}$, $\vec{c} = \begin{bmatrix}4\\6\end{bmatrix}$, and $\vec{d} = \begin{bmatrix}4\\5\end{bmatrix}$. Determine whether $\{\vec{a}, \vec{b}, \vec{c}, \vec{d}\}$ is linearly independent or linearly dependent.

By inspection, we see $\vec{c} = 2\vec{b}$. Therefore,

$$\operatorname{span}\{\vec{a}, \vec{b}, \vec{c}, \vec{d}\} = \operatorname{span}\{\vec{a}, \vec{b}, \vec{d}\},$$

and so $\{\vec{a}, \vec{b}, \vec{c}, \vec{d}\}$ is linearly dependent.

Example. The planes $\mathcal{P}$ and $\mathcal{Q}$ are given in vector form by

$$\overbrace{\vec{x} = t\begin{bmatrix}1\\2\\1\end{bmatrix} + s\begin{bmatrix}2\\2\\1\end{bmatrix}}^{\mathcal{P}} \qquad \text{and} \qquad \overbrace{\vec{x} = t\begin{bmatrix}3\\4\\2\end{bmatrix} + s\begin{bmatrix}2\\2\\1\end{bmatrix}}^{\mathcal{Q}}.$$

Determine if $\mathcal{P}$ and $\mathcal{Q}$ are the same plane.

We could answer this question using techniques from Module 2, but for variety, let's see if we can answer the question using spans and linear dependence.

Let $\vec{a}_1 = \begin{bmatrix}1\\2\\1\end{bmatrix}$ and $\vec{a}_2 = \begin{bmatrix}2\\2\\1\end{bmatrix}$ be direction vectors for $\mathcal{P}$ and let $\vec{b}_1 = \begin{bmatrix}3\\4\\2\end{bmatrix}$ and $\vec{b}_2 = \begin{bmatrix}2\\2\\1\end{bmatrix}$ be direction vectors for $\mathcal{Q}$ and notice $\mathcal{P} = \operatorname{span}\{\vec{a}_1, \vec{a}_2\}$ and $\mathcal{Q} = \operatorname{span}\{\vec{b}_1, \vec{b}_2\}$.
By definition $\mathcal{P} = \mathcal{Q}$ if (i) every point in $\mathcal{P}$ is a point in $\mathcal{Q}$ and (ii) every point in $\mathcal{Q}$ is a point in $\mathcal{P}$.

[17] The issue is, every element of a set is unique. Clearly, the vectors $\vec{v}$ and $\vec{v}$ are linearly dependent, but $\{\vec{v}, \vec{v}\} = \{\vec{v}\}$, and so $\{\vec{v}, \vec{v}\}$ is technically a linearly independent set. This issue would be resolved by talking about *multisets* instead of sets, but it isn't worth the hassle.

© Jason Siefken, 2015–2022

Focusing on (i), let $\vec p = t\vec a_1 + s\vec a_2 \in \mathcal{P}$ be an arbitrary point in $\mathcal{P}$. We need to show $\vec p \in \mathcal{Q}$. Since $\{\vec b_1, \vec b_2\}$ is linearly independent and $\mathcal{Q} = \operatorname{span}\{\vec b_1, \vec b_2\}$, showing $\vec p \in \mathcal{Q}$ is equivalent to showing $\{\vec p, \vec b_1, \vec b_2\}$ is a linearly ***dependent*** set.
To see this, start by observing

$$\{\vec a_1, \vec b_1, \vec b_2\} \qquad \text{and} \qquad \{\vec a_2, \vec b_1, \vec b_2\}$$

are both linearly dependent sets: $\vec a_1 = \vec b_1 - \vec b_2 \in \operatorname{span}\{\vec b_1, \vec b_2\}$ and $\vec a_2 = \vec b_2 \in \operatorname{span}\{\vec b_1, \vec b_2\}$. Therefore, $\vec a_1 \in \mathcal{Q}$ and $\vec a_2 \in \mathcal{Q}$. Since both $\vec a_1$ and $\vec a_2$ are in $\mathcal{Q}$ and $\mathcal{Q} = \operatorname{span}\{\vec b_1, \vec b_2\}$ is itself a span, we know that every linear combination of $\vec a_1$ and $\vec a_2$ must be in $\mathcal{Q}$. In particular, $\vec p = t\vec a_1 + s\vec a_2 \in \mathcal{Q}$, which is what we wanted to show.
We can show (ii) similarly by observing that $\vec b_1 \in \mathcal{P}$ and $\vec b_2 \in \mathcal{P}$ and so any point $\vec q = t\vec b_1 + s\vec b_2 \in \mathcal{Q}$ must also be in $\mathcal{P}$.

We can also think of linear independence/dependence from an algebraic perspective. Suppose the vectors $\vec u$, $\vec v$, and $\vec w$ satisfy

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

The set $\{\vec u, \vec v, \vec w\}$ is linearly dependent since $\vec w \in \operatorname{span}\{\vec u, \vec v\}$, but equation (2) can be rearranged to get

$$\vec 0 = \vec u + \vec v - \vec w. \tag{3}$$

Here we have expressed $\vec 0$ as a linear combination of $\vec u$, $\vec v$, and $\vec w$. By itself, this is nothing special. After all, we know $\vec 0 = 0\vec u + 0\vec v + 0\vec w$ is a linear combination of $\vec u$, $\vec v$, and $\vec w$. However, the right side of equation (3) has non-zero coefficients, which makes the linear combination *non-trivial*.

Trivial Linear Combination.
The linear combination $\alpha_1\vec v_1 + \cdots + \alpha_n\vec v_n$ is called ***trivial*** if $\alpha_1 = \cdots = \alpha_n = 0$. If at least one $\alpha_i \neq 0$, the linear combination is called ***non-trivial***.

We can always write $\vec 0$ as a linear combination of vectors if we let all the coefficients be zero, but it turns out we can only write $\vec 0$ as a *non-trivial* linear combination of vectors if those vectors are linearly dependent. This is the inspiration for another definition of linear independence/dependence.

Linearly Dependent & Independent (Algebraic).
The vectors $\vec v_1, \vec v_2, \ldots, \vec v_n$ are ***linearly dependent*** if there is a non-trivial linear combination of $\vec v_1, \ldots, \vec v_n$ that equals the zero vector. Otherwise they are linearly independent.

The idea of a "redundant vector" coming from the geometric definition of linear dependence is easy to visualize, but it can be hard to prove things with—checking for linear independence with the geometric definition involves verifying for every vector that it is not in the span of the others. The algebraic definition on the other hand is less obvious, but the reasoning is easier. You only need to analyze solutions to one equation!

Example. Let $\vec u = \begin{bmatrix}1\\2\end{bmatrix}$, $\vec v = \begin{bmatrix}2\\3\end{bmatrix}$, and $\vec w = \begin{bmatrix}4\\5\end{bmatrix}$. Use the algebraic definition of linear independence to determine whether $\{\vec u, \vec v, \vec w\}$ is linearly independent or dependent.

We need to determine if there is a non-trivial solution to

$$x\vec u + y\vec v + z\vec w = \vec 0.$$

This vector equation is equivalent to the system of equations

$$\begin{cases} x + 2y + 4z = 0 \\ 2x + 3y + 5z = 0 \end{cases}.$$

Solving this system using row reduction, we see the complete solution set can be expressed as

$$\begin{bmatrix}x\\y\\z\end{bmatrix} = t\begin{bmatrix}2\\-3\\1\end{bmatrix}.$$

In particular, $(x, y, z) = (2, -3, 1)$ is a non-trivial solution to this system. Therefore $\{\vec u, \vec v, \vec w\}$ is linearly dependent.

Theorem. The geometric and algebraic definitions of linear independence are equivalent.

© Jason Siefken, 2015–2022

Proof. To show the two definitions are equivalent, we need to show that geometric $\implies$ algebraic and algebraic $\implies$ geometric.

(geometric $\implies$ algebraic) Suppose $\vec{v}_1, \dots, \vec{v}_n$ are linearly dependent by the geometric definition. That means that for some i, we have

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

Fix such an i. Then, by the definition of span we know

$$\vec{v}_i = \alpha_1\vec{v}_1 + \cdots + \alpha_{i-1}\vec{v}_{i-1} + \alpha_{i+1}\vec{v}_{i+1} + \cdots + \alpha_n\vec{v}_n,$$

and so

$$\vec{0} = \alpha_1\vec{v}_1 + \cdots + \alpha_{i-1}\vec{v}_{i-1} - \vec{v}_i + \alpha_{i+1}\vec{v}_{i+1} + \cdots + \alpha_n\vec{v}_n.$$

This must be a non-trivial linear combination because the coefficient of $\vec{v}_i$ is $-1 \neq 0$. Therefore, $\vec{v}_1, \dots, \vec{v}_n$ is linearly dependent by the algebraic definition.

(algebraic $\implies$ geometric) Suppose $\vec{v}_1, \dots, \vec{v}_n$ are linearly dependent by the algebraic definition. That means there exist $\alpha_1, \dots, \alpha_n$, not all zero, so that

$$\vec{0} = \alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n.$$

Fix i so that $\alpha_i \neq 0$ (why do we know there is such an i?). Rearranging we get

$$-\alpha_i\vec{v}_i = \alpha_1\vec{v}_1 + \cdots + \alpha_{i-1}\vec{v}_{i-1} + \alpha_{i+1}\vec{v}_{i+1} + \cdots + \alpha_n\vec{v}_n,$$

and since $\alpha_i \neq 0$, we can multiply both sides by $\frac{-1}{\alpha_i}$ to get

$$\vec{v}_i = \frac{-\alpha_1}{\alpha_i}\vec{v}_1 + \cdots + \frac{-\alpha_{i-1}}{\alpha_i}\vec{v}_{i-1} + \frac{-\alpha_{i+1}}{\alpha_i}\vec{v}_{i+1} + \cdots + \frac{-\alpha_n}{\alpha_i}\vec{v}_n.$$

This shows that

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

and so $\vec{v}_1, \dots, \vec{v}_n$ is linearly dependent by the geometric definition. ■

Linear Independence and Unique Solutions

The algebraic definition of linear independence can teach us something about solutions to systems of equations.

Recall the linearly dependent vectors

$$\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \vec{v} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \vec{w} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

which satisfy the non-trivial relationship $\vec{u} + \vec{v} - \vec{w} = \vec{0}$. Since $\vec{u} + \vec{v} - \vec{w} = \vec{0}$ is a non-trivial relationship giving $\vec{0}$, we can use it to generate others. For example,

$$\begin{aligned} 17(\vec{u} + \vec{v} - \vec{w}) &= 17\vec{u} + 17\vec{v} - 17\vec{w} = 17\vec{0} = \vec{0} \\ -3(\vec{u} + \vec{v} - \vec{w}) &= -3\vec{u} - 3\vec{v} + 3\vec{w} = -3\vec{0} = \vec{0} \\ &\vdots \end{aligned}$$

are all different non-trivial linear combinations that give $\vec{0}$. In other words, if the equation $\alpha\vec{u} + \beta\vec{v} + \gamma\vec{w} = \vec{0}$ has a non-trivial solution, it has *infinitely many* non-trivial solutions. Conversely, if the equation $\alpha\vec{u} + \beta\vec{v} + \gamma\vec{w} = \vec{0}$ has infinitely many solutions, one of them has to be non-trivial!

Equations where one side is $\vec{0}$ show up often and are called *homogeneous* equations.

Homogeneous System.
A system of linear equations or a vector equation in the variables $\alpha_1, \dots, \alpha_n$ is called ***homogeneous*** if it takes the form

$$\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n = \vec{0},$$

where the right side of the equation is $\vec{0}$.

This insight links linear independence and homogeneous systems together, and is encapsulated in the following theorem.

© Jason Siefken, 2015–2022

Theorem. The vectors $\vec{v}_1, \ldots, \vec{v}_n$ are linearly independent if and only if the homogeneous equation

$$\alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n = \vec{0}$$

has a unique solution.

This theorem has a practical application: suppose you wanted to decide if the vectors $\vec{a}$, $\vec{b}$, and $\vec{c}$ were linearly dependent. You could (i) find a non-trivial solution to $x\vec{a}+y\vec{b}+z\vec{c}=\vec{0}$, or (ii) merely show that $x\vec{a}+y\vec{b}+z\vec{c}=\vec{0}$ has more than one solution. Sometimes one is easier than the other.

Linear Independence and Vector Form

The equation

$$\vec{x} = t_1\vec{d}_1 + t_2\vec{d}_2$$

represents a plane in vector form whenever $\vec{d}_1$ and $\vec{d}_2$ are non-zero, non-parallel vectors. In other words, $\vec{x} = t_1\vec{d}_1 + t_2\vec{d}_2$ represents a plane whenever $\{\vec{d}_1, \vec{d}_2\}$ is linearly independent.

Does this reasoning work for lines too? The equation

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

represents a line in vector form precisely when $\vec{d} \neq \vec{0}$. And $\{\vec{d}\}$ is linearly independent exactly when $\vec{d} \neq 0$.

This reasoning generalizes to volumes. The equation

$$\vec{x} = t_1\vec{d}_1 + t_2\vec{d}_2 + t_3\vec{d}_3$$

represents a *volume* in vector form exactly when $\{\vec{d}_1, \vec{d}_2, \vec{d}_3\}$ is linearly independent. To see this, suppose $\{\vec{d}_1, \vec{d}_2, \vec{d}_3\}$ were linearly dependent. That means one or more vectors could be removed from $\{\vec{d}_1, \vec{d}_2, \vec{d}_3\}$ without changing its span. Therefore, if $\{\vec{d}_1, \vec{d}_2, \vec{d}_3\}$ is linearly dependent $\vec{x} = t_1\vec{d}_1 + t_2\vec{d}_2 + t_3\vec{d}_3$ at best represents a plane (though it could be a line or a point).

We now have a way of testing the validity of a vector-form representation of a line/plane/volume. Just check whether the chosen direction vectors are linearly independent!

Takeaway. When writing an object in vector form, the direction vectors must always be linearly independent.

Practice Problems

1. Let $A = \left\{\begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$.
   (a) Is A linearly independent or dependent?
   (b) Describe the span of A.
   (c) Can A be extended (i.e., can vectors be added to A) so that A spans all of $\mathbb{R}^3$?

2. For each set below, determine whether it spans a point, line, plane, volume, or other.
   (a) $\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$
   (b) $\left\{\begin{bmatrix}1\\3\end{bmatrix}, \begin{bmatrix}2\\6\end{bmatrix}\right\}$
   (c) $\left\{\begin{bmatrix}2\\4\end{bmatrix}, \begin{bmatrix}4\\2\end{bmatrix}\right\}$
   (d) $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}-1\\-2\end{bmatrix}\right\}$
   (e) $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}-1\\2\end{bmatrix}\right\}$
   (f) $\{\}$
   (g) $\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\3\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix}\right\}$
   (h) $\left\{\begin{bmatrix}5\\4\\-3\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}2\\2\\2\end{bmatrix}\right\}$
   (i) $\left\{\begin{bmatrix}1\\-2\\0\end{bmatrix}, \begin{bmatrix}4\\-5\\1\end{bmatrix}\right\}$
   (j) $\left\{\begin{bmatrix}1\\-2\\0\end{bmatrix}, \begin{bmatrix}4\\-5\\1\end{bmatrix}, \begin{bmatrix}5\\-7\\1\end{bmatrix}\right\}$

3. (a) For each set in question 2, determine whether it is linearly independent or dependent.
   (b) Is the set $\left\{\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}5\\6\\7\end{bmatrix}, \begin{bmatrix}9\\10\\11\end{bmatrix}, \begin{bmatrix}13\\14\\15\end{bmatrix}\right\}$ linearly independent or dependent?
   (c) Can you find a set of $n+1$ vectors in $\mathbb{R}^n$ that is linearly independent? Explain.

4. (a) If possible, express the following lines in $\mathbb{R}^2$ as spans. Otherwise, justify why the line cannot be expressed as a span.
      i. $x = 0$
      ii. $2x + 3y = 0$
      iii. $5x - 4y = 0$

© *Jason Siefken, 2015–2022*

iv. $-x-y=-1$

v. $9x-15y=8$

(b) For each line in question 4a that cannot be expressed as a span, express it as a translated span.

(c) Each equation below specifies a line or a plane in $\mathbb{R}^3$. If possible, express the specified line or plane as a span. Otherwise, justify why it cannot be expressed as a span.

i. $2x-y+z=4$

ii. $x+6y-z=0$

iii. $x+3z=0$

iv. $y=1$

v. $x=0$ and $z=0$

vi. $2x-y=2$ and $z=-1$

(d) For lines or planes in question 4c that cannot be expressed as spans, express as a translated span.

5 Determine if the following planes, expressed in vector form, are the same plane.

(a) $\vec{x}=t\begin{bmatrix}1\\2\end{bmatrix}+s\begin{bmatrix}2\\7\end{bmatrix}$ and $\vec{x}=t\begin{bmatrix}3\\5\end{bmatrix}+s\begin{bmatrix}8\\4\end{bmatrix}$.

(b) $\vec{x}=t\begin{bmatrix}2\\2\\3\end{bmatrix}+s\begin{bmatrix}1\\0\\5\end{bmatrix}$ and $\vec{x}=t\begin{bmatrix}1\\0\\5\end{bmatrix}+s\begin{bmatrix}4\\2\\13\end{bmatrix}$.

(c) $\vec{x}=t\begin{bmatrix}1\\2\\1\end{bmatrix}+s\begin{bmatrix}2\\2\\1\end{bmatrix}$ and $\vec{x}=t\begin{bmatrix}0\\1\\0\end{bmatrix}+s\begin{bmatrix}1\\2\\1\end{bmatrix}$.

6 Show that the set $\left\{\begin{bmatrix}2\\0\\7\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}6\\4\\11\end{bmatrix}\right\}$ is linearly dependent in two ways. First, using the geometric definition of linear dependence and then using the algebraic definition.

7 Choose vectors $\vec{p}$, $\vec{d}_1$, $\vec{d}_2$, $\vec{d}_3$ in $\mathbb{R}^4$ such that the vector equation $\vec{x}=t_1\vec{d}_1+t_2\vec{d}_2+t_3\vec{d}_3+\vec{p}$ specifies:

(a) A hyperplane passing through the origin.

(b) A plane not passing through the origin.

(c) A line passing through the origin.

(d) The point $(2,2,2,3)$.

8 Classify the sets $A=\{\}$ and $B=\{\vec{0}\}$ as linearly independent or dependent.

9 Let $S=\left\{\begin{bmatrix}1\\3\end{bmatrix},\begin{bmatrix}0\\-1\end{bmatrix}\right\}$ and let $T=\left\{\begin{bmatrix}-1\\-1\end{bmatrix},\begin{bmatrix}0\\2\end{bmatrix},\begin{bmatrix}0\\0\end{bmatrix}\right\}$. Draw the sets S, T, and $T+S$.

10 Let $S=\left\{\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}0\\-1\end{bmatrix},\begin{bmatrix}0\\0\end{bmatrix}\right\}$.

(a) Draw S, $S+S$, and $(S+S)+S$.

(b) Is $(S+S)+S=S+(S+S)$? Does the expression $S+S+S$ make sense?

(c) Draw $S+S+S+S+\cdots$.

11 Let $D\subseteq\mathbb{R}^2$ be the unit disk centered at the origin and let $L\subseteq\mathbb{R}^2$ be the line segment from $(0,0)$ to $(0,2)$.

(a) How many points are in D, L, and $D+L$?

(b) Draw $D+L$.

(c) Find the area of $D+L$.

(d) Suppose $S\subseteq\mathbb{R}^2$ makes a smiley face when drawn and the "thickness" of each line composing this smiley face is 0.01 units. Can you find a set A so that the set $S+A$ represents a smiley face where the lines have a thickness of 0.05? If so, give an example of such an A. Otherwise, explain why it is impossible.

12 Let $\vec{v}_1,\vec{v}_2,\vec{v}_3$ be vectors. For each of the following statements, justify whether the statement is true or false.

(a) If $\vec{v}_1$ can be written as a linear combination of $\vec{v}_2$ and $\vec{v}_3$, then $\{\vec{v}_1,\vec{v}_2,\vec{v}_3\}$ is linearly dependent.

(b) If $\{\vec{v}_1,\vec{v}_2,\vec{v}_3\}$ is linearly dependent, then $\vec{v}_1$ can be written as a linear combination of $\vec{v}_2$ and $\vec{v}_3$.

(c) If $\vec{v}_1=k\vec{v}_2$ for some real number k, then $\{\vec{v}_1,\vec{v}_2\}$ is linearly dependent.

(d) If $\vec{v}_1$ is not a scalar multiple of $\vec{v}_2$, then $\{\vec{v}_1,\vec{v}_2,\vec{v}_3\}$ is linearly independent.

(e) All spans contain $\vec{0}$.

© Jason Siefken, 2015–2022

© Jason Siefken, 2015–2022

Span

Span

DEFINITION

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

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

Additionally, we define $\operatorname{span}\{\} = \{\vec{0}\}$.

---

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

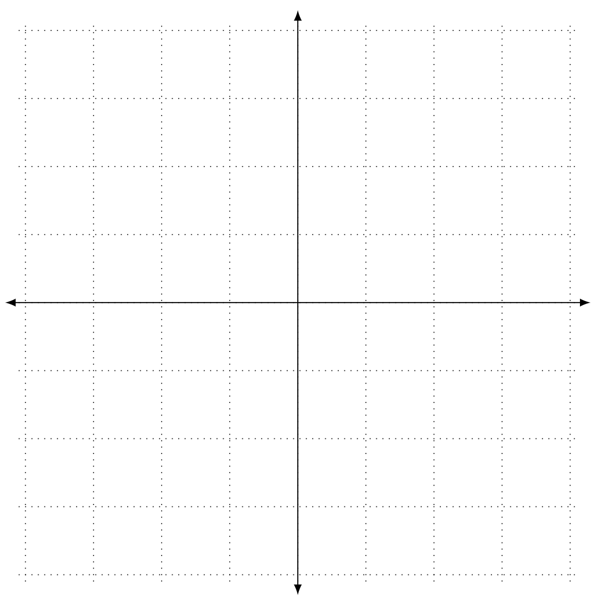
15.1 Draw $\operatorname{span}\{\vec{v}_1\}$.

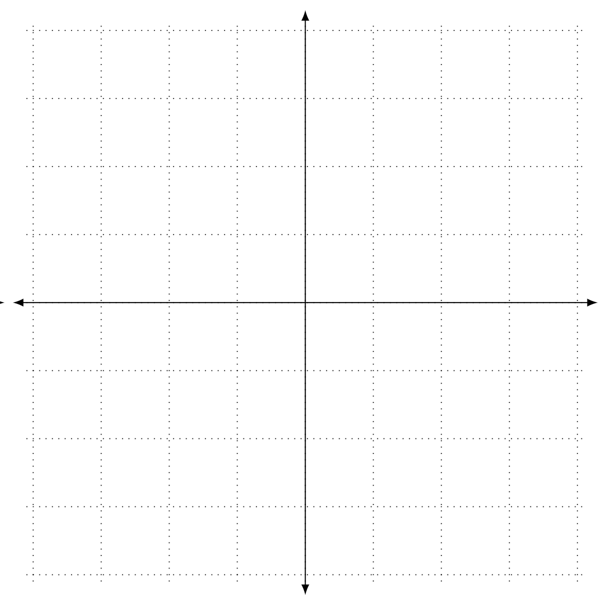
15.2 Draw $\operatorname{span}\{\vec{v}_2\}$.

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

15.4 Describe $\operatorname{span}\{\vec{v}_1, \vec{v}_3\}$.

15.5 Describe $\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.





© *Jason Siefken, 2015–2022*

16 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 $x - y = 4$.

16.1 If possible, describe ℓ_1 as a span. Otherwise explain why it's not possible.

16.2 If possible, describe ℓ_2 as a span. Otherwise explain why it's not possible.

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

© Jason Siefken, 2015–2022

Set Addition

DEF

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

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

---

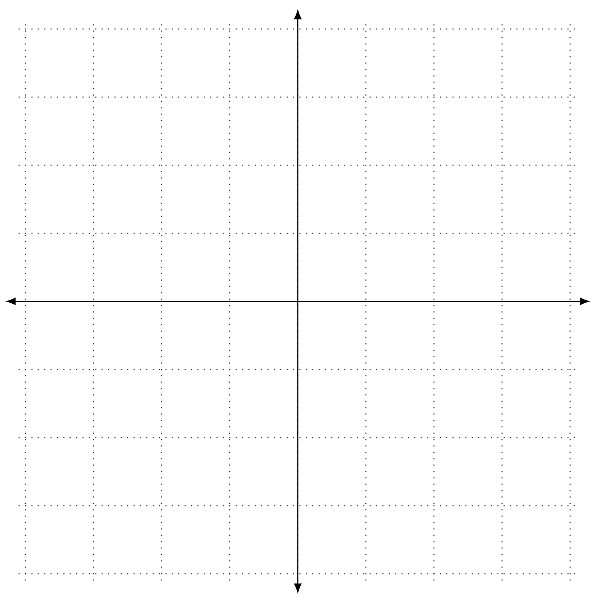
17 Let $A=\left\{\begin{bmatrix}1\\2\end{bmatrix}\right\}$, $B=\left\{\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}1\\-1\end{bmatrix}\right\}$, and $\ell=\operatorname{span}\left\{\begin{bmatrix}1\\-1\end{bmatrix}\right\}$.

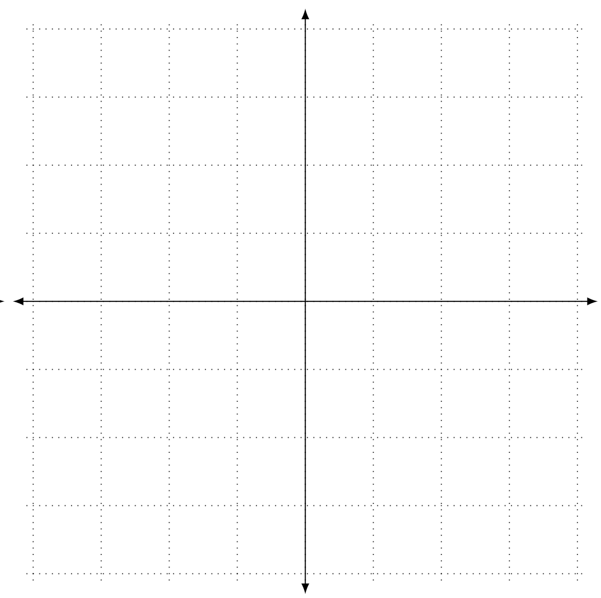
17.1 Draw A, B, and $A+B$ in the same picture.

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

17.3 Draw $\ell+A$.

17.4 Consider the line ℓ_2 given in vector form by $\vec{x}=t\begin{bmatrix}1\\-1\end{bmatrix}+\begin{bmatrix}1\\2\end{bmatrix}$. Can ℓ_2 be described using only a span? What about using a span and set addition?





© Jason Siefken, 2015–2022

18 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 $\operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 6 \\ 3 \\ 8 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 6 \end{bmatrix}\right\}$?

 © IOLA Team iola.math.vt.edu

DEFINITION

Linearly Dependent & Independent (Geometric)

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***.

---

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

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

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

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

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

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

© Jason Siefken, 2015–2022

Trivial Linear Combination

DEF

The linear combination $\alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n$ is called ***trivial*** if $\alpha_1 = \cdots = \alpha_n = 0$. If at least one $\alpha_i \neq 0$, the linear combination is called ***non-trivial***.

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

20.1 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?

20.2 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?

We now have an equivalent definition of linear dependence.

© Jason Siefken, 2015–2022

Linearly Dependent & Independent (Algebraic)

DEF

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

21

21.1 Explain how the geometric definition of linear dependence (original) implies this algebraic one (new).

21.2 Explain how this algebraic definition of linear dependence (new) implies the geometric one (original).

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).

© Jason Siefken, 2015–2022

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

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

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

Suppose that

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

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

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

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

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

© Jason Siefken, 2015–2022

23

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.

© IOLA Team iola.math.vt.edu

© Jason Siefken, 2015–2022

Dot Products & Normal Forms

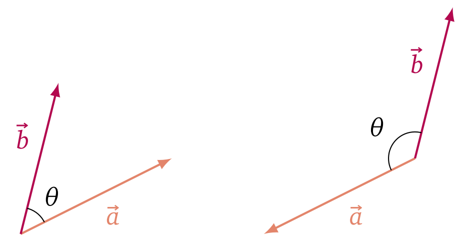
In this module you will learn

- Geometric and algebraic definitions of the dot product.
- How dot products relate to the length of a vector and the angle between two vectors.
- The *normal form* of lines, planes, and hyperplanes.

Let $\vec{a}$ and $\vec{b}$ be vectors rooted at the same point and let θ denote the *smaller* of the two angles between them (so $0 \leq \theta \leq \pi$). The *dot product* of $\vec{a}$ and $\vec{b}$ is defined to be

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

We will call this the *geometric definition of the dot product*.



The dot product is also sometimes called the *scalar product* because the result is a scalar.

Algebraically, we can define the dot product in terms of coordinates:

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

We will call this the *algebraic definition of the dot product*.

By switching between algebraic and geometric definitions, we can use the dot product to find quantities that are otherwise difficult to find.

Example. Find the angle between the vectors $\vec{v} = (1,2,3)$ and $\vec{w} = (1,1,-2)$.

From the algebraic definition of the dot product, we know

$$\vec{v} \cdot \vec{w} = 1(1) + 2(1) + 3(-2) = -3.$$

From the geometric definition, we know

$$\vec{v} \cdot \vec{w} = \|\vec{v}\|\|\vec{w}\| \cos\theta = \sqrt{14}\sqrt{6}\cos\theta = 2\sqrt{21}\cos\theta.$$

Equating the two definitions of $\vec{v} \cdot \vec{w}$, we see

$$\cos\theta = \frac{-3}{2\sqrt{21}}$$

and so $\theta = \arccos\left(\frac{-3}{2\sqrt{21}}\right)$.

The dot product has several interesting properties. Since the angle between $\vec{a}$ and itself is 0, the geometric definition of the dot product tells us

$$\vec{a} \cdot \vec{a} = \|\vec{a}\|\|\vec{a}\| \cos 0 = \|\vec{a}\|^2.$$

In other words,

$$\|\vec{a}\| = \sqrt{\vec{a} \cdot \vec{a}},$$

© Jason Siefken, 2015–2022

and so dot products can be used to compute the length of vectors.[18]

From the algebraic definition of the dot product, we can deduce several distributive laws. Namely, for any vectors $\vec a$, $\vec b$, and $\vec c$ and any scalar k we have

$$(\vec a+\vec b)\cdot\vec c=\vec a\cdot\vec c+\vec b\cdot\vec c \qquad \vec a\cdot(\vec b+\vec c)=\vec a\cdot\vec b+\vec a\cdot\vec c$$
$$(k\vec a)\cdot\vec b=k(\vec a\cdot\vec b)=\vec a\cdot(k\vec b)$$

and

$$\vec a\cdot\vec b=\vec b\cdot\vec a.$$

Orthogonality

Recall that for vectors $\vec a$ and $\vec b$, the relationship $\vec a\cdot\vec b=0$ can hold for two reasons: (i) either $\vec a=\vec 0$, $\vec b=\vec 0$, or both or (ii) $\vec a$ and $\vec b$ meet at 90°. Thus, the dot product can be used to tell if two vectors are perpendicular. There is some strangeness with the zero vector here, but it turns out this strangeness simplifies our lives mathematically.

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.

The definition of orthogonal encapsulates both the idea of two vectors forming a right angle and the idea of one of them being $\vec 0$.

Before we continue, let's pin down exactly what we mean by the *direction* of a vector. There are many ways we could define this term, but we'll go with the following.

Direction. The vector $\vec u$ points in the ***direction*** of the vector $\vec v$ if $\vec u=k\vec v$ for some scalar k. The vector $\vec u$ points in the ***positive direction*** of $\vec v$ if $\vec u=k\vec v$ for some positive scalar k.

The vector $2\vec e_1$ points in the direction of $\vec e_1$ since $\frac12(2\vec e_1)=\vec e_1$. Since $\frac12>0$, $2\vec e_1$ also points in the positive direction of $\vec e_1$. In contrast, $-\vec e_1$ points in the direction $\vec e_1$ but not the positive direction of $\vec e_1$.

When it comes to the relationship between two vectors, there are two extremes: they point in the same direction, or they are orthogonal. The dot product can be used to tell you which of these cases you're in, and more than that, it can tell you to what extent one vector points in the direction of another (even if they don't point in the same direction).

Example. Let $\vec a=\begin{bmatrix}1\\2\end{bmatrix}$, $\vec b=\begin{bmatrix}3\\3\end{bmatrix}$, $\vec c=\begin{bmatrix}2\\1\end{bmatrix}$, and $\vec v=\begin{bmatrix}3\\4\end{bmatrix}$. Which vector out of $\vec a$, $\vec b$, and $\vec c$ has a direction closest to the direction of $\vec v$?

We would like to know when θ, the angle between a pair of the given vectors, is smallest. This is equivalent to finding when $\cos\theta$ is closest to 1 (since $\cos 0=1$). By equating the geometric and algebraic definitions of the dot product, we know

$$\cos\theta=\frac{\vec p\cdot\vec q}{\|\vec p\|\|\vec q\|}.$$

Let α, β, and γ be the angles between the vector $\vec v$ and the vectors $\vec a$, $\vec b$, $\vec c$, respectively. Computing, we find

$$\cos\alpha=\frac{3+8}{5\sqrt5}=\frac{11\sqrt5}{25}\approx 0.9838699101$$
$$\cos\beta=\frac{9+12}{5\sqrt{18}}=\frac{7\sqrt2}{10}\approx 0.989949437$$
$$\cos\gamma=\frac{6+4}{5\sqrt5}=\frac{2\sqrt5}{5}\approx 0.894427191.$$

Since $\cos\beta$ is the closest to 1, we know $\vec b$ has a direction closest to that of $\vec v$.

Normal Form of Lines and Planes

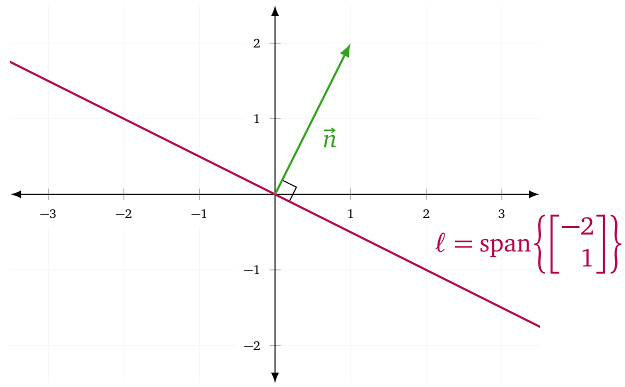
Let $\vec n=\begin{bmatrix}1\\2\end{bmatrix}$. If a vector $\vec v=\begin{bmatrix}v_1\\v_2\end{bmatrix}$ is orthogonal to $\vec n$, then

$$\vec n\cdot\vec v=v_1+2v_2=0,$$

[18] Oftentimes in non-geometric settings, the dot product between two vectors is defined first and then the length of $\vec a$ is actually *defined* to be $\sqrt{\vec a\cdot\vec a}$.

© Jason Siefken, 2015–2022

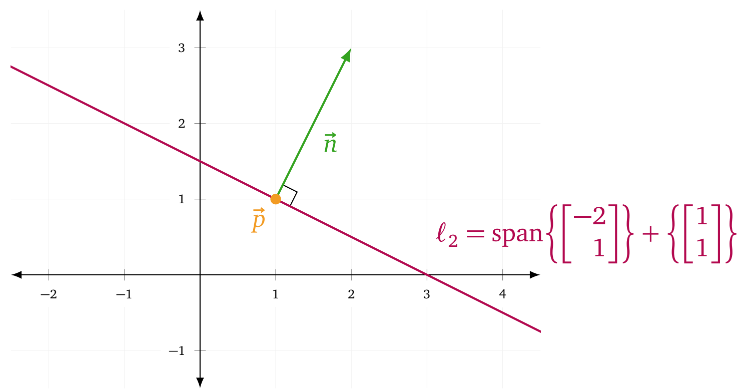
and so $v_1 = -2v_2$. In other words, $\vec{v}$ is orthogonal to $\vec{n}$ exactly when $\vec{v} \in \operatorname{span}\left\{\begin{bmatrix}-2\\1\end{bmatrix}\right\}$. What have we learned? The set of all vectors orthogonal to $\vec{n}$ forms a line $\ell = \operatorname{span}\left\{\begin{bmatrix}-2\\1\end{bmatrix}\right\}$. In this case, we call $\vec{n}$ a *normal vector* for ℓ.



Normal Vector. 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).

In $\mathbb{R}^2$, normal vectors provide yet another way to describe lines, including lines which don't pass through the origin.

Let $\vec{n} = \begin{bmatrix}1\\2\end{bmatrix}$ as before, and fix $\vec{p} = \begin{bmatrix}1\\1\end{bmatrix}$. If we draw the set of all vectors orthogonal to $\vec{n}$ but *root all the vectors at* $\vec{p}$, again we get a line, but this time the line passes through $\vec{p}$.



In fact, the line we get is $\ell_2 = \operatorname{span}\left\{\begin{bmatrix}-2\\1\end{bmatrix}\right\} + \left\{\begin{bmatrix}1\\1\end{bmatrix}\right\} = \ell + \{\vec{p}\}$, which is just ℓ (the parallel line through the origin) translated by $\vec{p}$.

Let's relate this to dot products and normal vectors. By definition, for every $\vec{v} \in \ell$, we have $\vec{n} \cdot \vec{v} = 0$. Since ℓ_2 is a translate of ℓ by $\vec{p}$, we deduce the relationship that for every $\vec{v} \in \ell_2$,

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

When a line is expressed as above, we say it is expressed in *normal form*.

Normal Form of a Line.
A line $\ell \subseteq \mathbb{R}^2$ is expressed in ***normal form*** if there exist vectors $\vec{n} \neq \vec{0}$ and $\vec{p}$ so that ℓ is the solution set to the equation

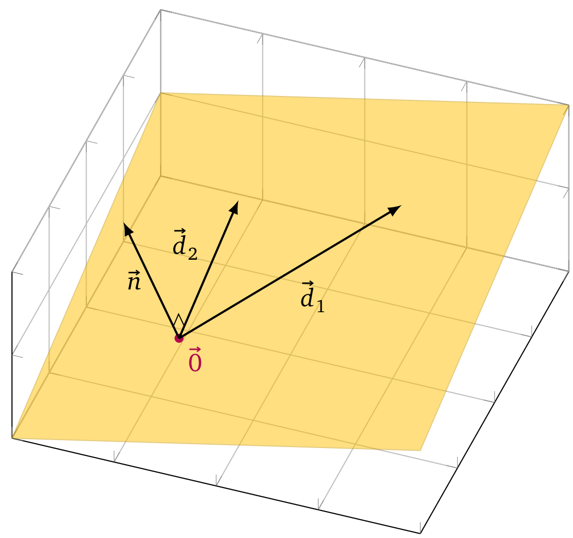
$$\vec{n} \cdot (\vec{x} - \vec{p}) = 0.$$

The equation $\vec{n} \cdot (\vec{x} - \vec{p}) = 0$ is called the ***normal form of*** ℓ.

Though the definition doesn't explicitly state it, if a line $\ell \in \mathbb{R}^2$ is expressed in normal form as $\vec{n} \cdot (\vec{x} - \vec{p}) = 0$, then $\vec{n}$ *is necessarily a normal vector for* ℓ. (Think about what the solution set to $\vec{n} \cdot (\vec{x} - \vec{p}) = 0$ would be if $\vec{n}$ happened to be $\vec{0}$!)

© Jason Siefken, 2015–2022

What about in $\mathbb{R}^3$? Fix a non-zero vector $\vec{n} \in \mathbb{R}^3$ and let $\mathcal{Q} \subseteq \mathbb{R}^3$ be the set of vectors orthogonal to $\vec{n}$. $\mathcal{Q}$ is a plane through the origin, and again, we call $\vec{n}$ a *normal vector* of the plane $\mathcal{Q}$.



In a similar way to the line, $\mathcal{Q}$ is the set of solutions to $\vec{n} \cdot \vec{x} = 0$. And, for any $\vec{p} \in \mathbb{R}^3$, the translated plane $\mathcal{Q} + \{\vec{p}\}$ is the solution set to

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

Thus, we see planes in $\mathbb{R}^3$ have a normal form just like lines in $\mathbb{R}^2$ do.

Example. Find vector form and normal form of the plane $\mathcal{P}$ passing through the points $A = (1, 0, 0)$, $B = (0, 1, 0)$ and $C = (0, 0, 1)$.

To find vector form of $\mathcal{P}$, we need a point on the plane and two direction vectors. We have three points on the plane, so we can obtain two direction vectors by subtracting these points in different ways. Let

$$\vec{d}_1 = \overrightarrow{AB} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \qquad \vec{d}_2 = \overrightarrow{AC} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Using the point A, we may now express $\mathcal{P}$ in vector form by

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

To write $\mathcal{P}$ in normal form, we need to find a normal vector for $\mathcal{P}$. By inspection, we see that $\vec{n} = (1, 1, 1)$ is a normal vector to $\mathcal{P}$. (If we weren't so insightful, we could also solve the system $\vec{n} \cdot \vec{d}_1 = 0$ and $\vec{n} \cdot \vec{d}_2 = 0$ to find a normal vector.) Now, we may express $\mathcal{P}$ in normal form as

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right) = 0.$$

In $\mathbb{R}^2$, only lines have a normal form, and in $\mathbb{R}^3$ only planes have a normal form. In general, we call objects in $\mathbb{R}^n$ which have a normal form *hyperplanes*.

Hyperplane. The set $X \subseteq \mathbb{R}^n$ is called a ***hyperplane*** if there exists $\vec{n} \neq \vec{0}$ and $\vec{p}$ so that X is the set of solutions to the equation

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

Hyperplanes always have dimension one less than the space they're contained in. So, hyperplanes in $\mathbb{R}^2$ are (one-dimensional) lines, hyperplanes in $\mathbb{R}^3$ are regular (two-dimensional) planes, and hyperplanes in $\mathbb{R}^4$ are (three-dimensional) volumes.

© Jason Siefken, 2015–2022

Hyperplanes and Linear Equations

Suppose $\vec{n}, \vec{p} \in \mathbb{R}^3$ and $\vec{n} \neq \vec{0}$. Then, solutions to

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

define a plane $\mathcal{P}$. But, $\vec{n} \cdot (\vec{x} - \vec{p}) = 0$ if and only if

$$\vec{n} \cdot \vec{x} = \vec{n} \cdot \vec{p} = \alpha.$$

Since $\vec{n}$ and $\vec{p}$ are fixed, α is a constant. Expanding using coordinates, we see

$$\vec{n} \cdot (\vec{x} - \vec{p}) = \vec{n} \cdot \vec{x} - \alpha = n_x x + n_y y + n_z z - \alpha = 0$$

and so $\mathcal{P}$ is the set of solutions to

$$n_x x + n_y y + n_z z = \alpha. \tag{4}$$

Equation (4) is sometimes called *scalar form* of a plane. For us, it will not be important to distinguish between scalar and normal form, but what is important is that we can use the row reduction algorithm to write the complete solution to (4), and this complete solution will necessarily be written in vector form.

> **Example.** Let $\mathcal{Q} \subseteq \mathbb{R}^3$ be the plane passing through $\vec{p}$ and with normal vector $\vec{n}$ where
>
> $$\vec{p} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad \text{and} \qquad \vec{n} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$
>
> Write $\mathcal{Q}$ in vector form.
>
> We know $\mathcal{Q}$ is the set of solutions to $\vec{n} \cdot (\vec{x} - \vec{p}) = 0$. In scalar form, this equation becomes
>
> $$\vec{n} \cdot (\vec{x} - \vec{p}) = \vec{n} \cdot \vec{x} - \vec{n} \cdot \vec{p} = x + y + z - 2 = 0.$$
>
> Rearranging, we see $\mathcal{Q}$ is the set of all solutions to
>
> $$x + y + z = 2.$$
>
> Using the row reduction algorithm to write the complete solution,[a] we get
>
> $$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}.$$
>
> [a] In some sense, this is overkill because the equation corresponds to the augmented matrix $\left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \end{array}\right]$, which is already row reduced.

Practice Problems

1 Compute the following dot products.

(a) $\begin{bmatrix} 9 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 10 \\ -3 \end{bmatrix}$

(b) $\begin{bmatrix} 1 \\ 36 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

(c) $\begin{bmatrix} 7 \\ 6 \\ -3 \end{bmatrix} \cdot \left(\begin{bmatrix} 5 \\ 11 \\ -1 \end{bmatrix} + \begin{bmatrix} -2 \\ -6 \\ -1 \end{bmatrix} \right)$

(d) $\begin{bmatrix} 1 \\ 3 \\ 0 \\ -5 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 2 \\ 1 \\ 2 \end{bmatrix}$

(e) $\left(\frac{1}{2} \begin{bmatrix} 2 \\ 5 \\ 4 \end{bmatrix} \right) \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$

2 Compute the length of the following vectors.

(a) $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$

(b) $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

(c) $4 \begin{bmatrix} 5 \\ -6 \\ 15 \\ 2 \end{bmatrix}$

3 For each pair of vectors listed below, determine if the angle between the vectors is greater than, less than, or equal to 90°.

© Jason Siefken, 2015–2022

(a) $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}-3\\4\end{bmatrix}$

(b) $\begin{bmatrix}1\\0\\1\end{bmatrix}$ and $\begin{bmatrix}-5\\4\\-3\end{bmatrix}$

(c) $\begin{bmatrix}1\\2\\3\end{bmatrix}$ and $\begin{bmatrix}-1\\-1\\2\end{bmatrix}$

4 For each vector, find two *unit* vectors orthogonal to it.

(a) $\begin{bmatrix}0\\1\end{bmatrix}$

(b) $\begin{bmatrix}1\\2\end{bmatrix}$

(c) $\begin{bmatrix}1\\3\\5\end{bmatrix}$

(d) $\begin{bmatrix}-13\\-4\\5\end{bmatrix}$

(e) $\begin{bmatrix}0\\1\\1\\\frac{1}{2}\end{bmatrix}$

5 Compute the distance between the following pairs of vectors.

(a) $\begin{bmatrix}-1\\1\end{bmatrix}$ and $\begin{bmatrix}-1\\-4\end{bmatrix}$

(b) $\begin{bmatrix}2\\-6\\5\end{bmatrix}$ and $\begin{bmatrix}-4\\7\\-3\end{bmatrix}$

(c) $\begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\begin{bmatrix}-1\\-1\\-1\end{bmatrix}$

(d) $\begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix}$ and $\begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}$

6 (a) Which vector out of $\begin{bmatrix}1\\0\end{bmatrix}$, $\begin{bmatrix}0\\1\end{bmatrix}$, and $\begin{bmatrix}4\\1\end{bmatrix}$ has a direction closest to that of $\begin{bmatrix}3\\5\end{bmatrix}$?

(b) Which vector out of $\begin{bmatrix}2\\3\\4\end{bmatrix}$, $\begin{bmatrix}1\\-1\\-1\end{bmatrix}$, and $\begin{bmatrix}-3\\0\\1\end{bmatrix}$ has a direction closest to that of $\begin{bmatrix}1\\0\\1\end{bmatrix}$?

7 For each plane specified, express the plane in both vector form and normal form.

(a) The plane $\mathcal{P} \subseteq \mathbb{R}^3$ passing through the points $A = (2,0,0)$, $B = (0,3,0)$ and $C = (0,0,-1)$.

(b) The plane $\mathcal{Q} \subseteq \mathbb{R}^3$ passing through the points $D = (1,1,1)$, $E = (1,-2,1)$ and $F = (0,12,0)$.

8 (a) Let $\mathcal{A} \subseteq \mathbb{R}^3$ be the plane passing through $\begin{bmatrix}0\\1\\1\end{bmatrix}$ and with normal vector $\begin{bmatrix}-1\\-1\\-1\end{bmatrix}$. Write $\mathcal{A}$ in vector form.

(b) Let $\mathcal{B} \subseteq \mathbb{R}^3$ be the plane passing through $\begin{bmatrix}1\\2\\3\end{bmatrix}$ and with normal vector $\begin{bmatrix}1\\-1\\0\end{bmatrix}$. Write $\mathcal{B}$ in vector form.

9 In this problem we will prove some algebraic properties of the dot product.

(a) Show by direct computation

$$\left(\begin{bmatrix}1\\2\end{bmatrix}+\begin{bmatrix}3\\4\end{bmatrix}\right)\cdot\begin{bmatrix}5\\6\end{bmatrix}=\begin{bmatrix}1\\2\end{bmatrix}\cdot\begin{bmatrix}5\\6\end{bmatrix}+\begin{bmatrix}3\\4\end{bmatrix}\cdot\begin{bmatrix}5\\6\end{bmatrix}$$

(b) For $\vec{x},\vec{y},\vec{z} \in \mathbb{R}^2$, justify whether or not it always holds that

$$(\vec{x}+\vec{y})\cdot\vec{z} = \vec{x}\cdot\vec{z}+\vec{y}\cdot\vec{z}.$$

Does the same conclusion hold true when $\vec{x},\vec{y},\vec{z} \in \mathbb{R}^n$?

(c) Show by direct computation

$$\left(6\begin{bmatrix}2\\3\end{bmatrix}\right)\cdot\begin{bmatrix}4\\5\end{bmatrix}=6\left(\begin{bmatrix}2\\3\end{bmatrix}\cdot\begin{bmatrix}4\\5\end{bmatrix}\right)$$

(d) For $\vec{x},\vec{y} \in \mathbb{R}^2$ and $k \in \mathbb{R}$, Justify whether or not it always holds that

$$(k\vec{x})\cdot\vec{y} = k(\vec{x}\cdot\vec{y}).$$

Does the same conclusion hold true when $\vec{x},\vec{y} \in \mathbb{R}^n$?

(e) The dot product is called *distributive*. Is this a good word to describe the dot product? Why?

10 Let $\vec{u},\vec{v} \in \mathbb{R}^n$. In this problem, we will prove $|\vec{u}\cdot\vec{v}| \leq \|\vec{u}\|\|\vec{v}\|$. This is called the *Cauchy-Schwarz inequality*.

(a) Assuming the geometric definition $\vec{u}\cdot\vec{v} = \|\vec{u}\|\|\vec{v}\|\cos\theta$ where θ is the angle between $\vec{u}$ and $\vec{v}$, prove the *Cauchy-Schwarz inequality*.

(b) The *Cauchy-Schwartz inequality* can also be proved using *only* the algebraic definition of the dot product. Keep in mind the following facts which come from the algebraic definition: Let $\vec{x},\vec{y},\vec{z} \in \mathbb{R}^n, k \in \mathbb{R}$. (i) $\|\vec{x}\| = \sqrt{\vec{x}\cdot\vec{x}}$; (ii) $(\vec{x}+\vec{y})\cdot\vec{z} = \vec{x}\cdot\vec{z}+\vec{y}\cdot\vec{z}$; (iii) $(k\vec{x})\cdot\vec{y} = k(\vec{x}\cdot\vec{y})$; (iv) $\vec{x}\cdot\vec{y} = \vec{y}\cdot\vec{x}$.

i. Explain why the result is immediate if one (or both) of $\vec{u},\vec{v}$ is the zero vector.

ii. Assume $\vec{u},\vec{v}$ are non-zero vectors. Consider the function $f:\mathbb{R}\to\mathbb{R}$ where $f(t) = \|t\vec{u}-\vec{v}\|^2$. Convince yourself that $f(t) \geq 0$ for all $t \in \mathbb{R}$.

iii. Simplify $f(t)$ into a quadratic formula so that $f(t) = at^2 - bt + c$ by determining its coefficients a,b,c in terms of $\vec{u},\vec{v}$.

iv. Prove $|\vec{u}\cdot\vec{v}| \leq \|\vec{u}\|\|\vec{v}\|$, the *Cauchy-Schwarz inequality*. Hint: Consider $f(\frac{b}{2a})$.

v. Prove $|\vec{u}\cdot\vec{v}| = \|\vec{u}\|\|\vec{v}\|$ if the vectors $\vec{u},\vec{v}$ are scalar multiples of each other.

11 Let $\vec{u},\vec{v} \in \mathbb{R}^n$. In this problem, we will prove $\|\vec{u}+\vec{v}\| \leq \|\vec{u}\|+\|\vec{v}\|$. This is called the *Triangle Inequality*.

© Jason Siefken, 2015–2022

(a) Pick your favorite vectors $\vec{u}, \vec{v}$ and draw a picture of $\vec{u}, \vec{v}$, and $\vec{u}+\vec{v}$. Root your vectors so that they form the edges of a triangle. Can you explain why the *Triangle Inequality* is true?

(b) Express $\|\vec{u}+\vec{v}\|$ in terms of dot products and square roots.

(c) Prove the *Triangle Inequality*.

12 Let $\mathcal{A} = \{\vec{v}_1, \dots, \vec{v}_k\} \subset \mathbb{R}^n$. A is a set of mutually orthogonal vectors if for all $i \neq j$, we have $\vec{v}_i \cdot \vec{v}_j = 0$.

(a) Suppose $\mathcal{A}$ is a set of mutually orthogonal vectors. Is $\mathcal{A}$ a linearly independent set? Why or Why not?

(b) Suppose $\mathcal{A}$ is a set of mutually orthogonal and non-zero vectors. Is $\mathcal{A}$ a linearly independent set? Why or Why not?

© Jason Siefken, 2015–2022

© Jason Siefken, 2015–2022

Dot Product

DEFINITION

Norm

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

DEFINITION

Dot Product

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

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

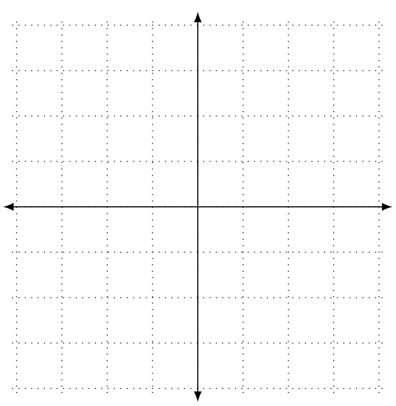
where θ is the angle between $\vec{a}$ and $\vec{b}$.

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

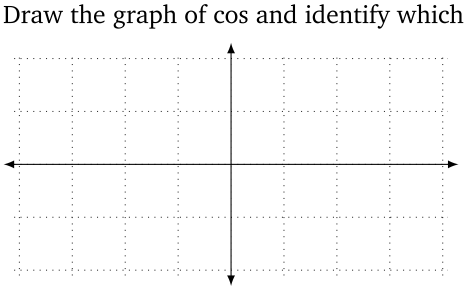
24.1 (a) Draw a picture of $\vec{a}$ and $\vec{b}$.

(b) Compute $\vec{a} \cdot \vec{b}$.

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



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



© Jason Siefken, 2015–2022

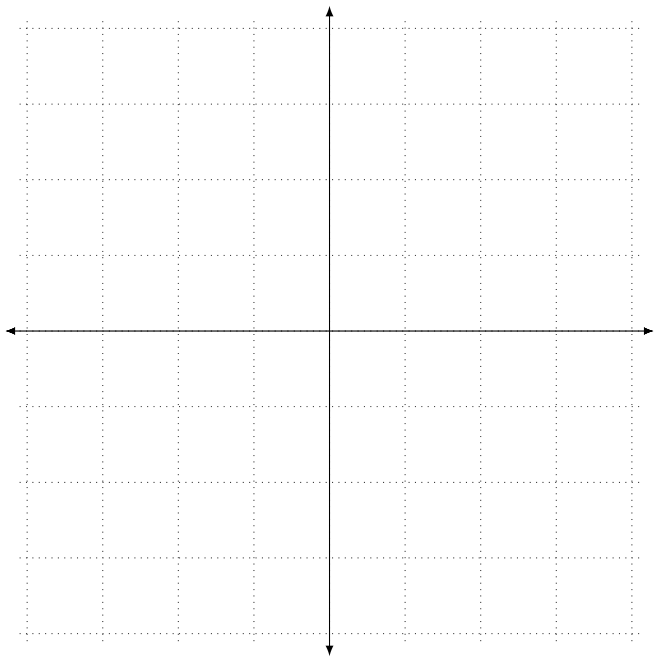
24.3 Draw a new picture of $\vec{a}$ and $\vec{b}$ and on that picture draw

(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.



24.4 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?)

(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 $\pi/2$.

(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?

© Jason Siefken, 2015–2022

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 $\|\vec{v}\| = 1$.

25

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

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

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

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

25.4 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 $\vec{u}$?

© Jason Siefken, 2015–2022

Orthogonal

DEF

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.

26

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

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

26.3 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 $\vec{y}$?

© Jason Siefken, 2015–2022

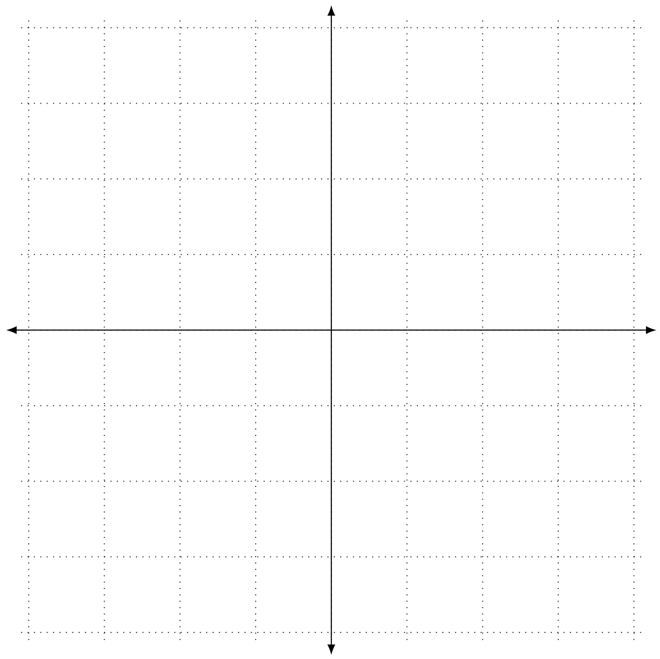
27

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

27.2 If $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $\vec{x}$ is orthogonal to $\vec{u}$, what is $\vec{x} \cdot \vec{u}$?

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

27.4 If possible, express A as a span.



© Jason Siefken, 2015–2022

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).

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

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

28.1 Find a vector $\vec{n}$ that is a normal vector for both ℓ_1 and ℓ_2.

28.2 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.

28.3 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 ℓ_1 and ℓ_2 in normal form.

28.4 Some textbooks would claim that ℓ_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?

© Jason Siefken, 2015–2022

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

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

29.2 Describe X using an equation.

29.3 Describe X as a span.

© Jason Siefken, 2015–2022

© Jason Siefken, 2015–2022

Projections & Vector Components

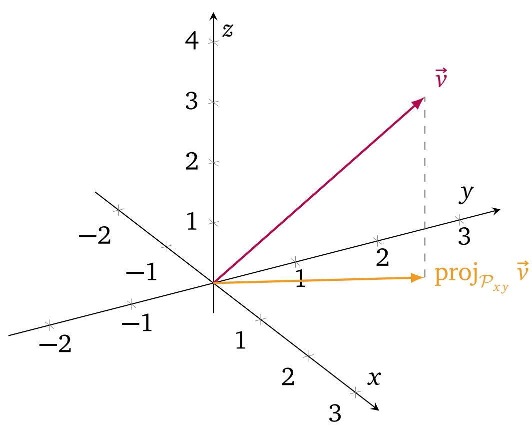
In this module you will learn

- The definition of the projection of a vector onto a set and the definition of the vector component of one vector in the direction of another.
- The relationship between projection, orthogonality, and vector components.
- How to project a vector onto a line.

Consider the following situation: you're designing a 3d video game, but your users only have 2d screens. Or, you have a 900-dimensional dataset, but you want to visualize it on a continuum (i.e., as a line). Each of these is an example of finding the best approximation to particular points given restrictions. In general, this operation is called a *projection*,[19] and in the world of linear algebra, it has a very particular meaning.

Projection. Let $X \subseteq \mathbb{R}^n$ be a set. The ***projection*** of the vector $\vec{v} \in \mathbb{R}^n$ onto X, written $\operatorname{proj}_X \vec{v}$, is the closest point in X to $\vec{v}$.

Let $\mathcal{P}_{xy} \subseteq \mathbb{R}^3$ be the xy-plane in $\mathbb{R}^3$ and let $\vec{v} = \begin{bmatrix}1\\2\\3\end{bmatrix}$. Intuitively, $\operatorname{proj}_{\mathcal{P}_{xy}} \vec{v}$ is the "shadow" that $\vec{v}$ would cast on $\mathcal{P}_{xy}$ if the sun were directly overhead. Upon drawing a picture, we conclude $\operatorname{proj}_{\mathcal{P}_{xy}} \vec{v} = \begin{bmatrix}1\\2\\0\end{bmatrix}$.



Continuing, let $\ell_y \subseteq \mathbb{R}^3$ be the y-axis in $\mathbb{R}^3$. It's a little bit harder to visualize what $\operatorname{proj}_{\ell_y} \vec{v}$ is, so let's appeal to some definitions.

By definition, every vector in ℓ_y takes the form $\vec{u}_t = \begin{bmatrix}0\\t\\0\end{bmatrix}$ for some $t \in \mathbb{R}$. The distance between $\vec{u}_t$ and $\vec{v}$ is

$$\|\vec{u}_t - \vec{v}\| = \left\| \begin{bmatrix}0\\t\\0\end{bmatrix} - \begin{bmatrix}1\\2\\3\end{bmatrix} \right\| = \sqrt{1^2 + (t-2)^2 + 3^2}.$$

Since $(t-2)^2$ is always positive, the quantity $\sqrt{1^2 + (t-2)^2 + 3^2}$ is minimized when $(t-2)^2 = 0$; that is, when $t = 2$. Thus, we see $\vec{u}_2$ is the closest vector in ℓ_y to $\vec{v}$ and so,

$$\operatorname{proj}_{\ell_y} \vec{v} = \vec{u}_2 = \begin{bmatrix}0\\2\\0\end{bmatrix}.$$

[19] What we define as a *projection* is sometimes called the *orthogonal projection* to distinguish it from other types of projections.

© *Jason Siefken, 2015–2022* CC BY SA

Example. Let $\ell \subseteq \mathbb{R}^2$ be the line given in vector form by $\vec{x} = t\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}3\\-2\end{bmatrix}$, and let $\vec{v} = \begin{bmatrix}-1\\-1\end{bmatrix}$. Use the definition of projection to find $\text{proj}_\ell \vec{v}$.

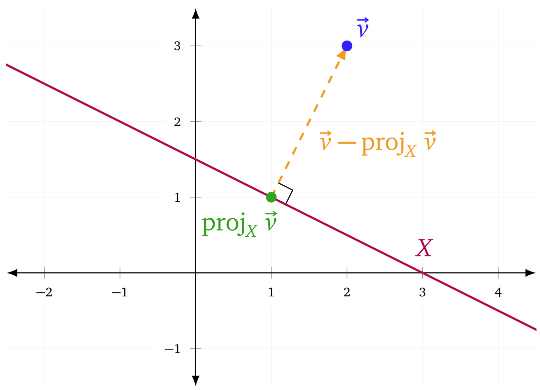
Let $\vec{u}_t = t\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}3\\-2\end{bmatrix} \in \ell$. By definition, the distance between $\vec{v}$ and $\vec{u}_t$ is given by

$$\|\vec{u}_t - \vec{v}\| = \left\|\left(t\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}3\\-2\end{bmatrix}\right) - \begin{bmatrix}-1\\-1\end{bmatrix}\right\| = \left\|\begin{bmatrix}t+4\\t-1\end{bmatrix}\right\| = \sqrt{2t^2+6t+17}.$$

The quantity $2t^2+6t+17$ is minimized when $t = -\frac{3}{2}$, and so the closest point in ℓ to $\vec{v}$ is $\vec{u}_{-3/2}$. Thus,

$$\text{proj}_\ell \vec{v} = -\frac{3}{2}\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}3\\-2\end{bmatrix} = \begin{bmatrix}3/2\\-7/2\end{bmatrix}.$$

Every example of a projection so far shares a geometric property. In the case of lines and planes, the vector from the projection to the original point is a normal vector for the line or plane (provided it's non-zero).



Stated precisely, if X is a line or plane and $\vec{v} \notin X$ is a vector, then $\vec{v} - \text{proj}_X \vec{v}$ is a normal vector for X. Using this fact, we can find projections onto lines and planes without needing to compute any distances!

Example. Let $\ell \subseteq \mathbb{R}^2$ be the line given in vector form by $\vec{x} = t\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}3\\-2\end{bmatrix}$, and let $\vec{v} = \begin{bmatrix}-1\\-1\end{bmatrix}$. Use the fact that $\vec{v} - \text{proj}_\ell \vec{v}$ is a normal vector to ℓ to find $\text{proj}_\ell \vec{v}$.

Since $\vec{v} - \text{proj}_\ell \vec{v}$ is a normal vector to ℓ, we know $\vec{v} - \text{proj}_\ell \vec{v}$ is orthogonal to $\vec{d} = \begin{bmatrix}1\\1\end{bmatrix}$. Let $\begin{bmatrix}x\\y\end{bmatrix} = \text{proj}_\ell \vec{v}$ for some unknown $x, y \in \mathbb{R}$. We now know

$$(\vec{v} - \text{proj}_\ell \vec{v}) \cdot \vec{d} = \left(\begin{bmatrix}-1\\-1\end{bmatrix} - \begin{bmatrix}x\\y\end{bmatrix}\right) \cdot \begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}-1-x\\-1-y\end{bmatrix} \cdot \begin{bmatrix}1\\1\end{bmatrix} = -2-x-y = 0.$$

That is,

$$x + y = -2. \tag{5}$$

Also, since $\text{proj}_\ell \vec{v} \in \ell$, we know

$$\text{proj}_\ell \vec{v} = \begin{bmatrix}x\\y\end{bmatrix} = t\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}3\\-2\end{bmatrix} = \begin{bmatrix}t+3\\t-2\end{bmatrix}.$$

From this, we have that $x - t = 3$ and $y - t = -2$. Combined with Equation (5), we have three equations and three unknowns which produce the following system of linear equations.

$$\begin{cases} \phantom{-t+{}} x + y = -2 \\ -t + x \phantom{{}+y} = 3 \\ -t \phantom{{}+x} + y = -2 \end{cases}$$

Solving this system, we conclude that $x = 3/2$ and $y = -7/2$ (we don't care about the value of t). Therefore $\text{proj}_\ell \vec{v} = \begin{bmatrix}3/2\\-7/2\end{bmatrix}$.

© Jason Siefken, 2015–2022

Takeaway. When projecting onto lines and planes, right angles appear in key places.

Projections Onto Other Sets

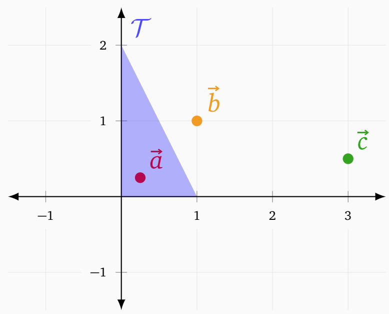
For projections onto lines and planes, we can use what we know about normal vectors to simplify our life. The same is true when projecting onto other sets, but we must always keep the definition in mind.

Example. Let $\mathcal{T} \subseteq \mathbb{R}^2$ be the filled in triangle with vertices $\begin{bmatrix}0\\0\end{bmatrix}$, $\begin{bmatrix}1\\0\end{bmatrix}$, and $\begin{bmatrix}0\\2\end{bmatrix}$, and let

$$\vec{a} = \begin{bmatrix}1/4\\1/4\end{bmatrix} \qquad \vec{b} = \begin{bmatrix}1\\1\end{bmatrix} \qquad \vec{c} = \begin{bmatrix}3\\1/2\end{bmatrix}.$$

Find $\text{proj}_{\mathcal{T}}\,\vec{a}$, $\text{proj}_{\mathcal{T}}\,\vec{b}$, and $\text{proj}_{\mathcal{T}}\,\vec{c}$.

We'll start by drawing a picture.



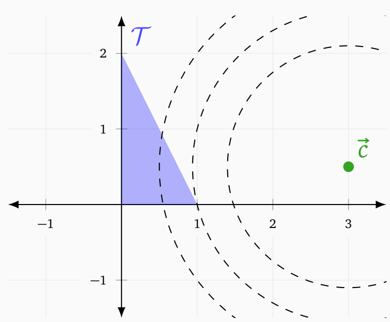
From the picture, we see that $\vec{a} \in \mathcal{T}$ and so

$$\text{proj}_{\mathcal{T}}\,\vec{a} = \vec{a}.$$

We also see that $\vec{b}$ is closest to the hypotenuse of $\mathcal{T}$, and so $\text{proj}_{\mathcal{T}}\,\vec{b}$ is the same as the projection of $\vec{b}$ onto the line $y = -2x + 2$. Computing, we find

$$\text{proj}_{\mathcal{T}}\,\vec{b} = \begin{bmatrix}3/5\\4/5\end{bmatrix}.$$

Finally, drawing concentric circles centered at $\vec{c}$, we see that the lower-right corner of $\mathcal{T}$ is the closest point in $\mathcal{T}$ to $\vec{c}$.



And so,

$$\text{proj}_{\mathcal{T}}\,\vec{c} = \begin{bmatrix}1\\0\end{bmatrix}.$$

Subtleties of Projections

You might be wondering, what is $\text{proj}_X\,\vec{v}$ if $\vec{v}$ is equidistant from *two* closest points in X? Or, what if X is an *open* set (for example, an open interval in $\mathbb{R}^1$)? Then there might not be a closest point in X to $\vec{v}$. In both these cases, we say $\text{proj}_X\,\vec{v}$ is *undefined*.

Formally, for a fixed set X, we consider $P(\vec{v}) = \text{proj}_X\,\vec{v}$ as a function that inputs and outputs vectors. And, as a function, P has a domain consisting of exactly the vectors $\vec{v}$ for which $P(\vec{v})$ is defined. As it happens, if X is

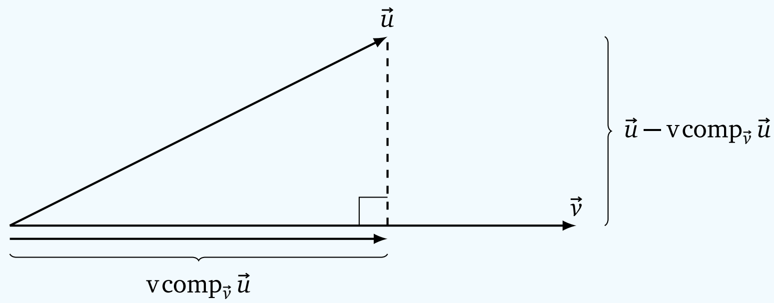
© Jason Siefken, 2015–2022

a line or a plane in $\mathbb{R}^n$, the domain of P is all of $\mathbb{R}^n$, and in this text, we will be sensible and only ask about projections that exist.

Vector Components

We've seen before that dot products can be used to measure how much one vector points in the direction of another. But, we can go further. Suppose $\vec{v} \neq \vec{0}$ and $\vec{u}$ are vectors. We might want to *decompose* $\vec{u}$ into the sum of two vectors, one which is in the direction of $\vec{v}$ and the other which is orthogonal to $\vec{v}$. The tool that does this is the *vector component*.

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



From the definition, it's obvious that

$$\vec{u} = \operatorname{vcomp}_{\vec{v}} \vec{u} + (\vec{u} - \operatorname{vcomp}_{\vec{v}} \vec{u})$$

is a decomposition of $\vec{u}$ into the sum of two vectors, one ($\operatorname{vcomp}_{\vec{v}} \vec{u}$) is in the direction of $\vec{v}$, and the other ($\vec{u} - \operatorname{vcomp}_{\vec{v}} \vec{u}$) is orthogonal to $\vec{v}$.

Example. Find the vector component of $\vec{a} = \begin{bmatrix}1\\2\end{bmatrix}$ in the direction of $\vec{b} = \begin{bmatrix}1\\1\end{bmatrix}$.

Since $\operatorname{vcomp}_{\vec{b}} \vec{a}$ is a vector in the direction of $\vec{b}$, we know

$$\operatorname{vcomp}_{\vec{b}} \vec{a} = k\vec{b}$$

for some $k \in \mathbb{R}$. Since $\vec{a} - \operatorname{vcomp}_{\vec{b}} \vec{a}$ is orthogonal to $\vec{b}$, we know

$$(\vec{a} - \operatorname{vcomp}_{\vec{b}} \vec{a}) \cdot \vec{b} = 0.$$

Combining these facts, we see

$$(\vec{a} - \operatorname{vcomp}_{\vec{b}} \vec{a}) \cdot \vec{b} = \left(\underbrace{\begin{bmatrix}1\\2\end{bmatrix}}_{\vec{a}} - \underbrace{\begin{bmatrix}k\\k\end{bmatrix}}_{k\vec{b}}\right) \cdot \underbrace{\begin{bmatrix}1\\1\end{bmatrix}}_{\vec{b}} = (1-k) + (2-k) = 3 - 2k = 0,$$

and so $k = 3/2$. Therefore

$$\operatorname{vcomp}_{\vec{b}} \vec{a} = k\vec{b} = \begin{bmatrix}3/2\\3/2\end{bmatrix}.$$

Since we'll be computing vector components often, let's try to find a formula for $\operatorname{vcomp}_{\vec{v}} \vec{u}$.

By definition, $\operatorname{vcomp}_{\vec{v}} \vec{u}$ is a vector in the direction of $\vec{v}$, so

$$\operatorname{vcomp}_{\vec{v}} \vec{u} = k\vec{v}.$$

Further, from the definition $\vec{u} - \operatorname{vcomp}_{\vec{v}} \vec{u}$ is orthogonal to $\vec{v}$, and so

$$\vec{v} \cdot (\vec{u} - \operatorname{vcomp}_{\vec{v}} \vec{u}) = \vec{v} \cdot (\vec{u} - k\vec{v}) = \vec{v} \cdot \vec{u} - k\vec{v} \cdot \vec{v} = 0.$$

Because $\vec{v} \neq \vec{0}$, we know $\vec{v} \cdot \vec{v} \neq 0$. Therefore, we may rearrange and solve for k to find

$$k = \frac{\vec{v} \cdot \vec{u}}{\vec{v} \cdot \vec{v}},$$

© Jason Siefken, 2015–2022

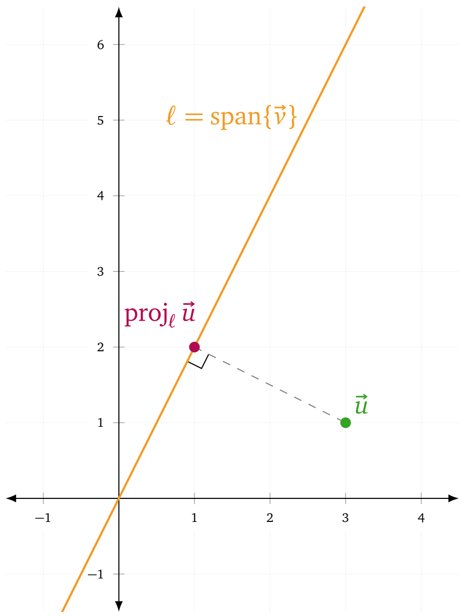
which means

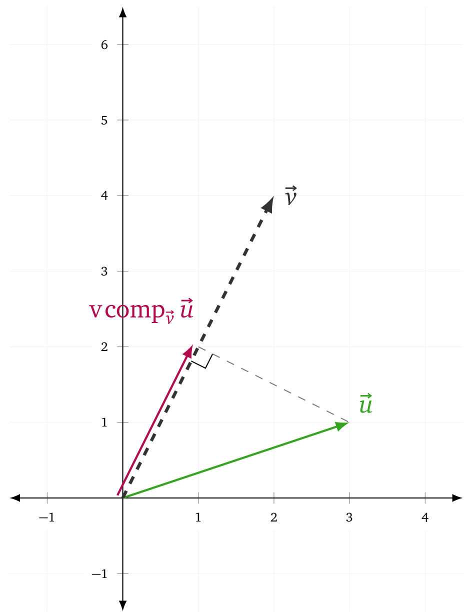
$$\operatorname{vcomp}_{\vec v}\vec u = \left(\frac{\vec v\cdot\vec u}{\vec v\cdot\vec v}\right)\vec v.$$

The Relationship Between Vector Components and Projections

Vector components and projections onto lines are closely related. So closely related that many textbooks use the single word *projection* to talk about both vector components and projections.[20] Let's take a moment to explore this relationship.

Let $\vec v = \begin{bmatrix}2\\4\end{bmatrix}$ and $\vec u = \begin{bmatrix}3\\1\end{bmatrix}$ and let $\ell = \operatorname{span}\{\vec v\}$. Drawing a picture of ℓ, $\vec u$, and $\operatorname{proj}_\ell \vec u$, we see that $\operatorname{proj}_\ell \vec u$ satisfies all the properties of $\operatorname{vcomp}_{\vec v}\vec u$.





Since $\ell = \operatorname{span}\{\vec v\}$ and $\operatorname{proj}_\ell \vec u \in \ell$, we know that $\operatorname{proj}_\ell \vec u$ is in the direction of $\vec v$. Further, using geometric arguments, we know $\vec u - \operatorname{proj}_\ell \vec u$ is a normal vector for ℓ and is therefore orthogonal to its direction vector $\vec v$! What's more, we didn't use anything in particular about $\vec u$ and $\vec v$ when making this argument (other than $\vec v \neq \vec 0$). This means, we have established a general fact.

Theorem. For vectors $\vec u$ and $\vec v \neq 0$, we have

$$\operatorname{proj}_{\operatorname{span}\{\vec v\}}\vec u = \operatorname{vcomp}_{\vec v}\vec u.$$

This is great news because vector components are easy to compute using dot products while projections are usually hard to compute.

Example. Compute the projection of $\vec a = \begin{bmatrix}3\\7\end{bmatrix}$ onto $L = \operatorname{span}\left\{\begin{bmatrix}1\\-4\end{bmatrix}\right\}$.

Let $\vec b = \begin{bmatrix}1\\-4\end{bmatrix}$. Since $L = \operatorname{span}\{\vec b\}$ and $\vec b \neq \vec 0$, by the theorem above, we have

$$\operatorname{proj}_L \vec a = \operatorname{vcomp}_{\vec b}\vec a = \left(\frac{\vec b\cdot\vec a}{\vec b\cdot\vec b}\right)\vec b = \frac{3-28}{1+16}\begin{bmatrix}1\\-4\end{bmatrix} = \begin{bmatrix}-25/17\\100/17\end{bmatrix}.$$

It's worth noting, however, that vector components are equal to projections *only in the case when you're projecting onto a span*. In general, projections and vector components are unrelated.

Example. Let $\vec a = \begin{bmatrix}3\\7\end{bmatrix}$, $\vec b = \begin{bmatrix}1\\-4\end{bmatrix}$, and let D be the line given in vector form by $\vec x = t\vec b + \vec a$. Show that $\operatorname{proj}_D \vec a \neq \operatorname{vcomp}_{\vec b}\vec a$.

[20] We will not.

© Jason Siefken, 2015–2022

By definition, $\operatorname{proj}_D \vec a$ is the closest point in D to $\vec a$. Since $\vec a \in D$, we must have

$$\operatorname{proj}_D \vec a = \vec a = \begin{bmatrix}3\\7\end{bmatrix}.$$

We already computed $\operatorname{vcomp}_{\vec b} \vec a = \begin{bmatrix}-25/17\\100/17\end{bmatrix}$ in the previous example, and so we see

$$\operatorname{vcomp}_{\vec b} \vec a = \begin{bmatrix}-25/17\\100/17\end{bmatrix} \neq \begin{bmatrix}3\\7\end{bmatrix} = \operatorname{proj}_D \vec a.$$

Takeaway. When projecting onto the span of a single vector, you can use vector components as a computational shortcut, but if the set isn't a span, you cannot.

Practice Problems

1 Let $T = \left\{\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}-1\\2\end{bmatrix}, \begin{bmatrix}1\\-2\end{bmatrix}\right\}$. Find $\operatorname{proj}_T \begin{bmatrix}3\\1\end{bmatrix}$.

2 Let $C = \left\{\vec v \in \mathbb{R}^2 : \|\vec v\| = 1\right\}$ be the unit circle in $\mathbb{R}^2$. Find $\operatorname{proj}_C \begin{bmatrix}2\\0\end{bmatrix}$. Justify your answer.

3 Let $\ell = \operatorname{span}\left\{\begin{bmatrix}2\\1\end{bmatrix}\right\}$, $L = \operatorname{span}\left\{\begin{bmatrix}2\\1\end{bmatrix}\right\} + \left\{\begin{bmatrix}4\\0\end{bmatrix}\right\}$, and let S be the set of convex linear combinations of $\begin{bmatrix}2\\1\end{bmatrix}$ and $\begin{bmatrix}4\\2\end{bmatrix}$.

For $\vec v = \begin{bmatrix}1\\0\end{bmatrix}$, find

(a) $\operatorname{proj}_\ell \vec v$.

(b) $\operatorname{proj}_L \vec v$.

(c) $\operatorname{proj}_S \vec v$.

4 Let T be the set of convex linear combinations of $\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}-1\\1\end{bmatrix}, \begin{bmatrix}-1\\-2\end{bmatrix}\right\}$. Find $\operatorname{proj}_T(\vec v)$, for

(a) $\vec v = \begin{bmatrix}3\\3\end{bmatrix}$

(b) $\vec v = \begin{bmatrix}0\\0\end{bmatrix}$

(c) $\vec v = \begin{bmatrix}1\\-2\end{bmatrix}$

(d) $\vec v = \begin{bmatrix}0\\-4\end{bmatrix}$

5 Explain in your own words how to find $\operatorname{proj}_\ell(\vec v)$ when $\ell = \operatorname{span}\{\vec d\}$ for some $\vec d \neq \vec 0$.

6 Let $\vec e_1 = \begin{bmatrix}1\\0\end{bmatrix}$, $\vec e_2 = \begin{bmatrix}0\\1\end{bmatrix}$, and $\vec u = \begin{bmatrix}2\\3\end{bmatrix}$.

(a) Draw $\vec e_1$, $\vec e_2$, $\vec u$, $\operatorname{vcomp}_{\vec e_1} \vec u$, and $\operatorname{vcomp}_{\vec e_2} \vec u$ on the same grid.

(b) Write down two characterizing properties for $\operatorname{vcomp}_{\vec e_2} \vec u$.

(c) Check that $\vec u - \operatorname{vcomp}_{\vec e_1} \vec u$ satisfies the above properties.

(d) $\operatorname{vcomp}_{\vec e_1} \vec u + \operatorname{vcomp}_{\vec e_2} \vec u = \vec u$. Does this always happen? Explain.

7 In this problem, we will find the projection of a vector onto a plane in $\mathbb{R}^3$. Let $\vec u = \begin{bmatrix}1\\2\\-2\end{bmatrix}$, $\vec v = \begin{bmatrix}0\\1\\1\end{bmatrix}$, $\vec a = \begin{bmatrix}6\\4\\-2\end{bmatrix}$, and let $\mathcal{P} = \operatorname{span}\{\vec u, \vec v\}$.

(a) Find $\operatorname{vcomp}_{\vec u}(\vec a)$ and $\operatorname{vcomp}_{\vec v}(\vec a)$.

(b) Show that $\vec a - \operatorname{vcomp}_{\vec u}(\vec a) - \operatorname{vcomp}_{\vec v}(\vec a)$ is a normal vector for $\mathcal{P}$.

(c) Use 7b to find $\operatorname{proj}_{\mathcal{P}}(\vec a)$.

© Jason Siefken, 2015–2022

Projections

Projection

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

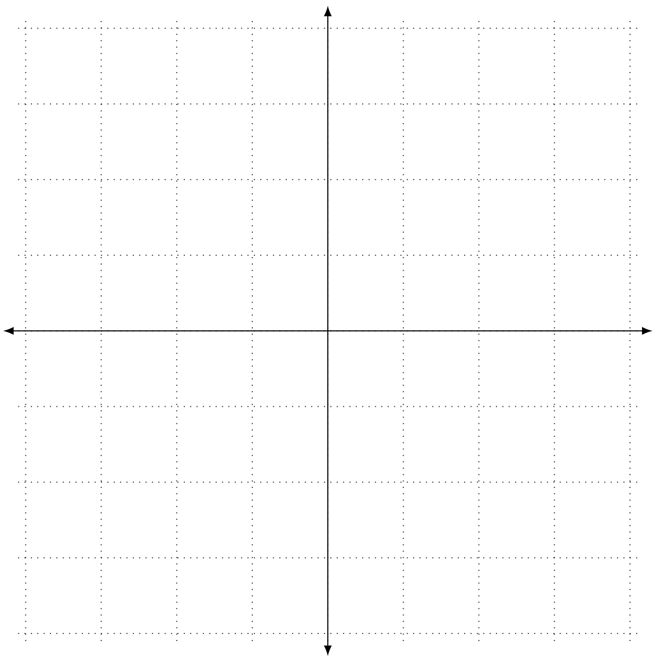
30 Let $\vec{a} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ and $\ell = \text{span}\{\vec{a}\}$.

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

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

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

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



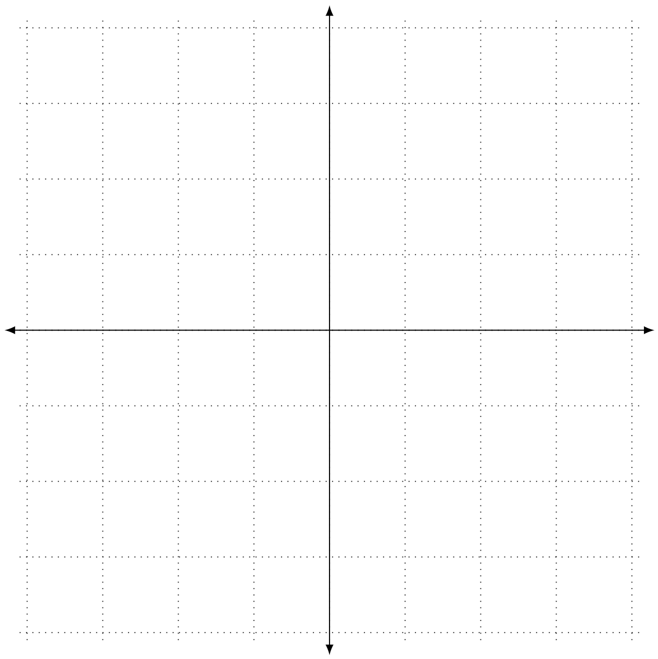
© Jason Siefken, 2015–2022

31 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 $\vec{c} = \begin{bmatrix}1\\3\end{bmatrix}$.

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

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

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

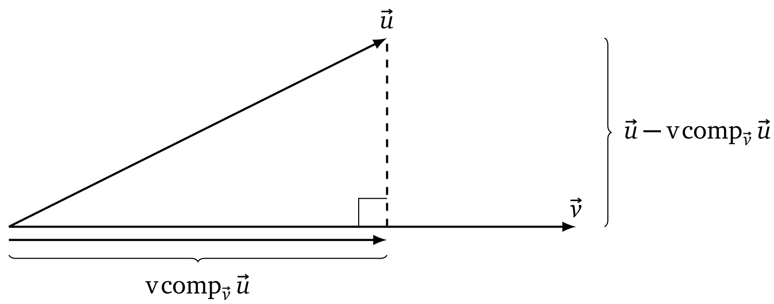


© Jason Siefken, 2015–2022

Vector Components

DEFINITION

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



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

32.1 List two conditions that $\operatorname{vcomp}_{\vec{b}} \vec{a}$ must satisfy.

32.2 Find a formula for $\operatorname{vcomp}_{\vec{b}} \vec{a}$.

© Jason Siefken, 2015–2022

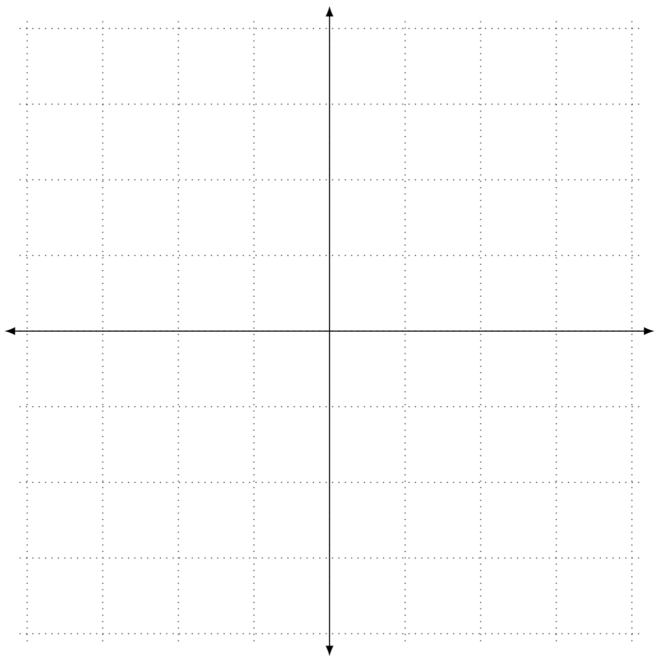
33 Let $\vec{d}=\begin{bmatrix}3\\3\end{bmatrix}$ and $\vec{u}=\begin{bmatrix}1\\2\end{bmatrix}$.

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

33.2 How do $\operatorname{proj}_{\operatorname{span}\{\vec{d}\}}\vec{u}$ and $\operatorname{vcomp}_{\vec{d}}\vec{u}$ relate?

33.3 Compute $\operatorname{proj}_{\operatorname{span}\{\vec{d}\}}\vec{u}$ and $\operatorname{vcomp}_{\vec{d}}\vec{u}$.

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



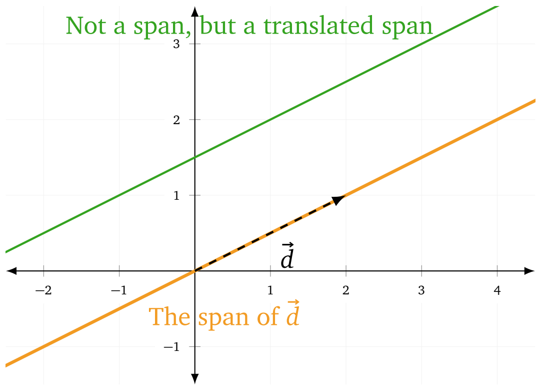
© *Jason Siefken, 2015–2022*

Subspaces & Bases

In this module you will learn

- Formal and intuitive definitions of subspaces.
- The relationship between subspaces and spans.
- How to prove whether or not a set is a subspace.
- How to find a basis for and the dimension of a subspace.

Lines or planes through the origin can be written as spans of their direction vectors. However, a line or plane that doesn't pass through the origin cannot be written as a span—it must be expressed as a *translated* span.



There's something special about sets that can be expressed as (untranslated) spans. In particular, since a linear combination of linear combinations is still a linear combination, a span is *closed* with respect to linear combinations. That is, by taking linear combinations of vectors in a span, you cannot escape the span. In general, sets that have this property are called *subspaces*.

Subspace. A non-empty subset $V \subseteq \mathbb{R}^n$ is called a ***subspace*** if for all $\vec{u}, \vec{v} \in V$ and all scalars k we have

(i) $\vec{u} + \vec{v} \in V$; and

(ii) $k\vec{u} \in V$.

In the definition of a subspace, property (i) is called being *closed with respect to vector addition* and property (ii) is called being *closed with respect to scalar multiplication*.

Subspaces generalize the idea of *flat spaces through the origin*. They include lines, planes, volumes and more.

Example. Let $\mathcal{V} \subseteq \mathbb{R}^2$ be the complete solution to $x + 2y = 0$. Show that $\mathcal{V}$ is a subspace.

Let $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ be in $\mathcal{V}$, and let k be a scalar.
By definition, we have

$$\begin{aligned} u_1 + 2u_2 &= 0 \\ v_1 + 2v_2 &= 0 \end{aligned}.$$

We will show that $\mathcal{V}$ is nonempty and that (i) $\vec{u} + \vec{v} \in \mathcal{V}$; and (ii) $k\vec{u} \in \mathcal{V}$.
First we will show (i). Observe that

$$\vec{u} + \vec{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}$$

and the coordinates of $\vec{u} + \vec{v}$ satisfy

$$(u_1 + v_1) + 2(u_2 + v_2) = (u_1 + 2u_2) + (v_1 + 2v_2) = 0 + 0 = 0.$$

Since the coordinates of $\vec{u} + \vec{v}$ satisfy the equation $x + 2y = 0$, we have shown that $\vec{u} + \vec{v} \in \mathcal{V}$.
Next we will show (ii). Observe that

$$k\vec{u} = \begin{bmatrix} ku_1 \\ ku_2 \end{bmatrix}$$

© Jason Siefken, 2015–2022

and the coordinates of $k\vec{u}$ satisfy

$$(ku_1)+2(ku_2)=k(u_1+2u_2)=k0=0.$$

And so, we have shown that $k\vec{u}\in\mathcal{V}$.

Finally, since $\vec{0}=\begin{bmatrix}0\\0\end{bmatrix}$ satisfies $x+2y=0$, we conclude that $\vec{0}\in\mathcal{V}$ and so $\mathcal{V}$ is non-empty.

Thus, by the definition, we have shown that $\mathcal{V}$ is a subspace.

Example. Let $\mathcal{W}\subseteq\mathbb{R}^2$ be the line expressed in vector form as

$$\vec{x}=t\begin{bmatrix}1\\2\end{bmatrix}+\begin{bmatrix}1\\1\end{bmatrix}.$$

Determine whether $\mathcal{W}$ is a subspace.

$\mathcal{W}$ is *not* a subspace. To see this, notice that $\vec{v}=\begin{bmatrix}1\\1\end{bmatrix}\in\mathcal{W}$, but $0\vec{v}=\vec{0}\notin\mathcal{W}$. Therefore, $\mathcal{W}$ is not closed under scalar multiplication and so it cannot be a subspace.

As mentioned earlier, subspaces and spans are deeply connected. This connection is given by the following theorem.

Theorem (Subspace-Span). Every subspace is a span and every span is a subspace. More precisely, $\mathcal{V}\subseteq\mathbb{R}^n$ is a subspace if and only if $\mathcal{V}=\operatorname{span}\mathcal{X}$ for some set $\mathcal{X}$.

Proof. We will start by showing every span is a subspace. Fix $\mathcal{X}\subseteq\mathbb{R}^n$ and let $\mathcal{V}=\operatorname{span}\mathcal{X}$. First note that if $\mathcal{X}\neq\{\}$, then $\mathcal{V}$ is non-empty because $\mathcal{X}\subseteq\mathcal{V}$ and if $\mathcal{X}=\{\}$, then $\mathcal{V}=\{\vec{0}\}$, and so is still non-empty.

Fix $\vec{v},\vec{u}\in\mathcal{V}$. By definition there are $\vec{x}_1,\vec{x}_2,\ldots,\vec{y}_1,\vec{y}_2,\ldots\in\mathcal{X}$ and scalars $\alpha_1,\alpha_2,\ldots,\beta_1,\beta_2,\ldots$ so that

$$\vec{v}=\sum\alpha_i\vec{x}_i\qquad\vec{u}=\sum\beta_i\vec{y}_i.$$

To verify property (i), observe that

$$\vec{u}+\vec{v}=\sum\alpha_i\vec{x}_i+\sum\beta_i\vec{y}_i$$

is also a linear combination of vectors in $\mathcal{X}$ (because all $\vec{x}_i$ and $\vec{y}_i$ are in $\mathcal{X}$), and so $\vec{u}+\vec{v}\in\operatorname{span}\mathcal{X}=\mathcal{V}$.

To verify property (ii), observe that for any scalar α,

$$\alpha\vec{v}=\alpha\sum\alpha_i\vec{x}_i=\sum(\alpha\alpha_i)\vec{x}_i\in\operatorname{span}\mathcal{X}=\mathcal{V}.$$

Since $\mathcal{V}$ is non-empty and satisfies both properties (i) and (ii), it is a subspace.

Now we will prove that every subspace is a span. Let $\mathcal{V}$ be a subspace and consider $\mathcal{V}'=\operatorname{span}\mathcal{V}$. Since taking a span may only enlarge a set, we know $\mathcal{V}\subseteq\mathcal{V}'$. If we establish that $\mathcal{V}'\subseteq\mathcal{V}$, then $\mathcal{V}=\mathcal{V}'=\operatorname{span}\mathcal{V}$, which would complete the proof.

Fix $\vec{x}\in\mathcal{V}'$. By definition, there are $\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_n\in\mathcal{V}$ and scalars $\alpha_1,\alpha_2,\ldots,\alpha_n$ so that

$$\vec{x}=\sum\alpha_i\vec{v}_i.$$

Observe that $\alpha_i\vec{v}_i\in\mathcal{V}$ for all i, since $\mathcal{V}$ is closed under scalar multiplication. It follows that $\alpha_1\vec{v}_1+\alpha_2\vec{v}_2\in\mathcal{V}$, because $\mathcal{V}$ is closed under sums. Continuing, $(\alpha_1\vec{v}_1+\alpha_2\vec{v}_2)+\alpha_3\vec{v}_3\in\mathcal{V}$ because $\mathcal{V}$ is closed under sums. Applying the principle of finite induction, we see

$$\vec{x}=\sum\alpha_i\vec{v}_i=\Big(((\alpha_1\vec{v}_1+\alpha_2\vec{v}_2)+\alpha_3\vec{v}_3)+\cdots+\alpha_{n-1}\vec{v}_{n-1}\Big)+\alpha_n\vec{v}_n\in\mathcal{V}.$$

Thus $\mathcal{V}'\subseteq\mathcal{V}$, which completes the proof. ∎

The previous theorem is saying that spans and subspaces are two ways of talking about the same thing. Spans provide a *constructive* definition of lines/planes/volumes/etc. through the origin. That is, when you describe a line/plane/etc. through the origin as a span, you're saying "this is a line/plane/etc. through the origin because every point in it is a linear combination of *these specific vectors*". In contrast, subspaces provide a *categorical* definition of lines/planes/etc. through the origin. When you describe a line/plane/etc. through the origin as a subspace, you're saying "this is a line/plane/etc. through the origin because these *properties* are satisfied".[21]

[21] Categorical definitions are useful when working with objects where it's hard to pin down exactly what the elements inside are.

© Jason Siefken, 2015–2022

Takeaway. Spans and subspaces are two different ways of talking about the same objects: points/lines/planes/etc. through the origin.

Special Subspaces

When thinking about $\mathbb{R}^n$, there are two special subspaces that are always available. The first is $\mathbb{R}^n$ itself. $\mathbb{R}^n$ is obviously non-empty, and linear combinations of vectors in $\mathbb{R}^n$ remain in $\mathbb{R}^n$. The second is the *trivial subspace*, $\{\vec{0}\}$.

Trivial Subspace. The subset $\{\vec{0}\} \subseteq \mathbb{R}^n$ is called the ***trivial subspace***.

Theorem. The trivial subspace is a subspace.

Proof. First note that $\{\vec{0}\}$ is non-empty since $\vec{0} \in \{\vec{0}\}$. Now, since $\vec{0}$ is the only vector in $\{\vec{0}\}$, properties (i) and (ii) follow quickly:

$$\vec{0} + \vec{0} = \vec{0} \in \{\vec{0}\}$$

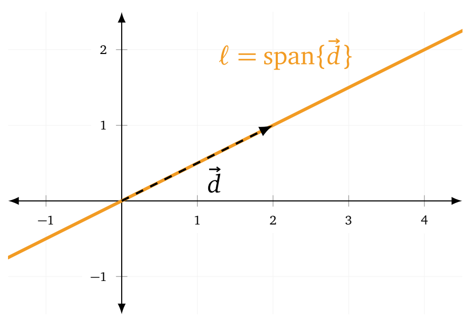
and

$$\alpha\vec{0} = \vec{0} \in \{\vec{0}\}.$$

■

Bases

Let $\vec{d} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and consider $\ell = \operatorname{span}\{\vec{d}\}$.



We know that ℓ is a subspace, and we defined ℓ as the span of $\{\vec{d}\}$, but we didn't have to define ℓ that way. We could have, for instance, defined $\ell = \operatorname{span}\{\vec{d}, -2\vec{d}, \frac{1}{2}\vec{d}\}$. However, $\operatorname{span}\{\vec{d}\}$ is a simpler way to describe ℓ than $\operatorname{span}\{\vec{d}, -2\vec{d}, \frac{1}{2}\vec{d}\}$. This property is general: the simplest descriptions of a line involve the span of only one vector.

Analogously, let $\mathcal{P} = \operatorname{span}\{\vec{d}_1, \vec{d}_2\}$ be the plane through the origin with direction vectors $\vec{d}_1$ and $\vec{d}_2$. There are many ways to write $\mathcal{P}$ as a span, but the simplest ones involve exactly two vectors. The idea of a *basis* comes from trying to find the simplest description of a subspace.

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

In short, a basis for a subspace is a linearly independent set that spans that subspace.

Example. Let $\ell = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ -4 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}\right\}$. Find two different bases for ℓ.

We are looking for a set of linearly independent vectors that spans ℓ. Notice that $\begin{bmatrix} 1 \\ 2 \end{bmatrix} = -\frac{1}{2}\begin{bmatrix} -2 \\ -4 \end{bmatrix} = 2\begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$. Therefore,

$$\operatorname{span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\} = \operatorname{span}\left\{\begin{bmatrix} -2 \\ -4 \end{bmatrix}\right\} = \operatorname{span}\left\{\begin{bmatrix} 1/2 \\ 1 \end{bmatrix}\right\} = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ -4 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}\right\} = \ell.$$

Because $\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$ is linearly independent and spans ℓ, we have that $\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$ is a basis for ℓ. Similarly, $\left\{\begin{bmatrix} 1/2 \\ 1 \end{bmatrix}\right\}$ is another basis for ℓ.

© Jason Siefken, 2015–2022

Unpacking the definition of basis a bit more, we can see that a basis for a subspace is a set of vectors that is *just the right size* to describe everything in the subspace. It's not too big—because it is linearly independent, there are no redundancies. It's not too small—because we require it to span the subspace.[22]

There are several facts everyone should know about bases:

1. Bases are not unique. Every subspace (except the trivial subspace) has multiple bases.
2. Given a basis for a subspace, every vector in the subspace can be written as a *unique* linear combination of vectors in that basis.
3. Any two bases for the same subspace have the same number of elements.

You can prove the first fact by observing that if $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \ldots\}$ is a basis with at least one element,[23] then $\{2\vec{b}_1, 2\vec{b}_2, \ldots\}$ is a different basis. The second fact is a consequence of all bases being linearly independent. The third fact is less obvious and takes some legwork to prove, so we will accept it as is.

Dimension

Let $\mathcal{V}$ be a subspace. Though there are many bases for $\mathcal{V}$, they all have the same number of vectors in them. And, this number says something fundamental about $\mathcal{V}$: it tells us the maximum number of linearly independent vectors that can simultaneously exist in $\mathcal{V}$. We call this number the *dimension* of $\mathcal{V}$.

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

This definition agrees with our intuition about lines and planes: the dimension of a line through $\vec{0}$ is 1, and the dimension of a plane through $\vec{0}$ is 2. It even tells us the dimension of the single point $\{\vec{0}\}$ is 0.[24]

Example. Find the dimension of $\mathbb{R}^2$.

Since $\{\vec{e}_1, \vec{e}_2\}$ is a basis for $\mathbb{R}^2$, we know $\mathbb{R}^2$ is two dimensional.

Example. Let $\ell = \operatorname{span}\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}-2\\-4\end{bmatrix}, \begin{bmatrix}1/2\\1\end{bmatrix}\right\}$. Find the dimension of the subspace ℓ.

This is the same subspace from the earlier example where we found $\left\{\begin{bmatrix}1\\2\end{bmatrix}\right\}$ and $\left\{\begin{bmatrix}1/2\\1\end{bmatrix}\right\}$ were bases for ℓ. Both these bases contain one element, and so ℓ is a one dimensional subspace.

Example. Let $A = \{(x_1, x_2, x_3, x_4) : x_1 + 2x_2 - x_3 = 0 \text{ and } x_1 + 6x_4 = 0\}$. Find a basis for and the dimension of A.

A is the complete solution to the system

$$\begin{cases} x_1 + 2x_2 - x_3 & = 0 \\ x_1 \qquad\qquad + 6x_4 & = 0 \end{cases},$$

which can be expressed in vector form as

$$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} = t\begin{bmatrix}0\\1/2\\1\\0\end{bmatrix} + s\begin{bmatrix}-6\\3\\0\\1\end{bmatrix}.$$

Therefore $A = \operatorname{span}\left\{\begin{bmatrix}0\\1/2\\1\\0\end{bmatrix}, \begin{bmatrix}-6\\3\\0\\1\end{bmatrix}\right\}$. Since $\left\{\begin{bmatrix}0\\1/2\\1\\0\end{bmatrix}, \begin{bmatrix}-6\\3\\0\\1\end{bmatrix}\right\}$ is a linearly independent spanning set with two elements, A is two dimensional.

Like $\mathbb{R}^2$ and $\mathbb{R}^3$, whenever we discuss $\mathbb{R}^n$, we always have a standard basis that comes along for the ride.

[22] If you're into British fairy tales, you might call a basis a *Goldilocks set*.
[23] The empty set is a basis for the trivial subspace.
[24] The dimension of a line, plane, or point not through the origin is defined to be the dimension of the subspace obtained when it is translated to the origin.

© Jason Siefken, 2015–2022

Standard Basis. The ***standard basis*** for $\mathbb{R}^n$ is the set $\{\vec{e}_1, \ldots, \vec{e}_n\}$ where

$$\vec{e}_1 = \begin{bmatrix}1\\0\\0\\\vdots\end{bmatrix} \qquad \vec{e}_2 = \begin{bmatrix}0\\1\\0\\\vdots\end{bmatrix} \qquad \vec{e}_3 = \begin{bmatrix}0\\0\\1\\\vdots\end{bmatrix} \qquad \cdots.$$

That is $\vec{e}_i$ is the vector with a 1 in its ith coordinate and zeros elsewhere.

Note: the notation $\vec{e}_i$ is context specific. If we say $\vec{e}_i \in \mathbb{R}^2$, then $\vec{e}_i$ must have exactly two coordinates. If we say $\vec{e}_i \in \mathbb{R}^{45}$, then $\vec{e}_i$ must have 45 coordinates.

Practice Problems

1 For each of the following sets, prove whether or not they are a subspace.

(a) $\mathcal{T} \subseteq \mathbb{R}^2$, where $\mathcal{T}$ is the complete solution to $3x - y = 0$.

(b) $\mathcal{U} \subseteq \mathbb{R}^2$, where $\mathcal{U}$ is the complete solution to $\frac{1}{2}x - 6y = 0$.

(c) $\mathcal{V} \subseteq \mathbb{R}^2$, where $\mathcal{V}$ is the complete solution to $x - 5y - 1 = 0$.

(d) $\mathcal{X} \subseteq \mathbb{R}^3$, where $\mathcal{X}$ is the complete solution to $5x - \pi y + (\ln 2)z = 0$.

(e) $\mathcal{Q} \subseteq \mathbb{R}^n$, where $\mathcal{Q}$ is the complete solution to $a_1x_1 + a_2x_2 + \ldots + a_{n-1}x_{n-1} + a_nx_n = 0$ where $a_1, \ldots, a_n \in \mathbb{R}$.

2 For each of the following sets, prove whether or not it is a subspace.

(a) $\mathcal{A} \subseteq \mathbb{R}^2$, where $\mathcal{A}$ is specified in vector form by $\vec{x} = t\begin{bmatrix}5\\-7\end{bmatrix} + \begin{bmatrix}1\\2\end{bmatrix}$.

(b) $\mathcal{B} \subseteq \mathbb{R}^2$, where $\mathcal{B}$ is specified in vector form by $\vec{x} = t\begin{bmatrix}-3\\4\end{bmatrix}$.

(c) $\mathcal{C} \subseteq \mathbb{R}^3$, where $\mathcal{C}$ is specified in vector form by $\vec{x} = t\begin{bmatrix}1\\0\\5\end{bmatrix}$.

(d) $\mathcal{D} \subseteq \mathbb{R}^3$, where $\mathcal{D}$ is specified in vector form by $\vec{x} = t\begin{bmatrix}2\\3\\4\end{bmatrix} + s\begin{bmatrix}10\\20\\131\end{bmatrix} + \begin{bmatrix}0\\0\\6\end{bmatrix}$.

(e) $\mathcal{E} \subseteq \mathbb{R}^3$, where $\mathcal{E}$ is specified in vector form by $\vec{x} = t\begin{bmatrix}5\\7\\1\end{bmatrix} + s\begin{bmatrix}2\\-2\\1\end{bmatrix}$.

3 Use the definition of subspace to prove each span below is a subspace.

(a) $\operatorname{span}\left\{\begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}1\\2\end{bmatrix}\right\}$

(b) $\operatorname{span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}2\\0\\0\end{bmatrix}\right\}$

4 A non-empty subset $\mathcal{V} \subseteq \mathbb{R}^n$ is called a subspace if for all $\vec{u}, \vec{v} \in \mathcal{V}$ and all scalars k we have (i) $\vec{u} + \vec{v} \in \mathcal{V}$ and (ii) $k\vec{u} \in \mathcal{V}$. For each set below, list which of property (i), property (ii), or non-emptiness fails. Justify your answer.

(a) $\{(x, y, z) : x + y + z = 4\}$

(b) $\{\}$

(c) $\{(x, y) : x = y^2\}$

(d) $\{(x_1, x_2) : x_1 \geq 0\}$

(e) $\{(x, y) : x^2 + y^2 = 0\}$

5 For each subset below, determine whether or not it is a subspace. If it is a subspace, find (i) its dimension and (ii) a basis for it.

(a) $\operatorname{span}\left\{\begin{bmatrix}2\\3\end{bmatrix}, \begin{bmatrix}-4\\-6\end{bmatrix}, \begin{bmatrix}1\\3/2\end{bmatrix}\right\}$

(b) $\operatorname{span}\left\{\begin{bmatrix}1\\0\\-2\end{bmatrix}, \begin{bmatrix}0\\2\\5\end{bmatrix}, \begin{bmatrix}1\\2\\3\end{bmatrix}\right\}$

(c) The plane given in vector form by

$$\vec{x} = t\begin{bmatrix}6\\1\\1\end{bmatrix} + s\begin{bmatrix}6\\0\\6\end{bmatrix}$$

(d) The line given in vector form by

$$\vec{x} = t\begin{bmatrix}2\\2\\3\end{bmatrix}$$

(e) The complete solution to

$$\begin{bmatrix}1\\-2\\3\end{bmatrix} \cdot \left(\begin{bmatrix}x\\y\\z\end{bmatrix} - \begin{bmatrix}2\\2\\\frac{2}{3}\end{bmatrix}\right) = 0$$

(f) The complete solution to

$$\begin{bmatrix}1\\3\\3\\7\end{bmatrix} \cdot \left(\begin{bmatrix}x\\y\\z\\w\end{bmatrix} - \begin{bmatrix}0\\0\\0\\0\end{bmatrix}\right) = 0$$

6 Which of the following are bases for $\mathbb{R}^3$?

(a) $\left\{\begin{bmatrix}2\\6\\1\end{bmatrix}, \begin{bmatrix}4\\2\\1\end{bmatrix}, \begin{bmatrix}6\\8\\2\end{bmatrix}\right\}$

(b) $\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\-1\\0\end{bmatrix}, \begin{bmatrix}-2\\1\\1\end{bmatrix}\right\}$

(c) $\left\{\begin{bmatrix}2\\3\\5\end{bmatrix}, \begin{bmatrix}5\\-4\\2\end{bmatrix}\right\}$

© Jason Siefken, 2015–2022

(d) $\left\{ \begin{bmatrix} 2 \\ 5 \\ -6 \end{bmatrix}, \begin{bmatrix} 4 \\ 11 \\ -12 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -3 \end{bmatrix} \right\}$

7 For each statement, determine if it is true or false. Justify your answer by referring to a definition or a theorem.

(a) All spans are subspaces.

(b) All subspaces can be expressed as spans.

(c) All translated spans are subspaces.

(d) The empty set is a subspace.

(e) The set $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$ is a subspace.

8 Give two examples of subspaces of $\mathbb{R}^4$ that are (i) 1 dimensional, (ii) 3 dimensional. Can you give an example of a subspace that is 0 dimensional?

9 Let $\mathcal{G} \subseteq \mathbb{R}^n$ be a subspace. Give upper and lower bounds for the dimension of $\mathcal{G}$.

© Jason Siefken, 2015–2022

Subspaces and Bases

Subspace (DEFINITION)

A non-empty subset $V \subseteq \mathbb{R}^n$ is called a ***subspace*** if for all $\vec{u}, \vec{v} \in V$ and all scalars k we have

(i) $\vec{u} + \vec{v} \in V$; and

(ii) $k\vec{u} \in V$.

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

---

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

34.1 $A = \left\{ \vec{x} \in \mathbb{R}^2 \, : \, \vec{x} = \begin{bmatrix} a \\ 0 \end{bmatrix} \text{ for some } a \in \mathbb{Z} \right\}$.

34.2 $B = \left\{ \vec{x} \in \mathbb{R}^2 \, : \, \vec{x} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$.

34.3 $C = \left\{ \vec{x} \in \mathbb{R}^2 \, : \, \vec{x} = \begin{bmatrix} 0 \\ t \end{bmatrix} \text{ for some } t \in \mathbb{R} \right\}$.

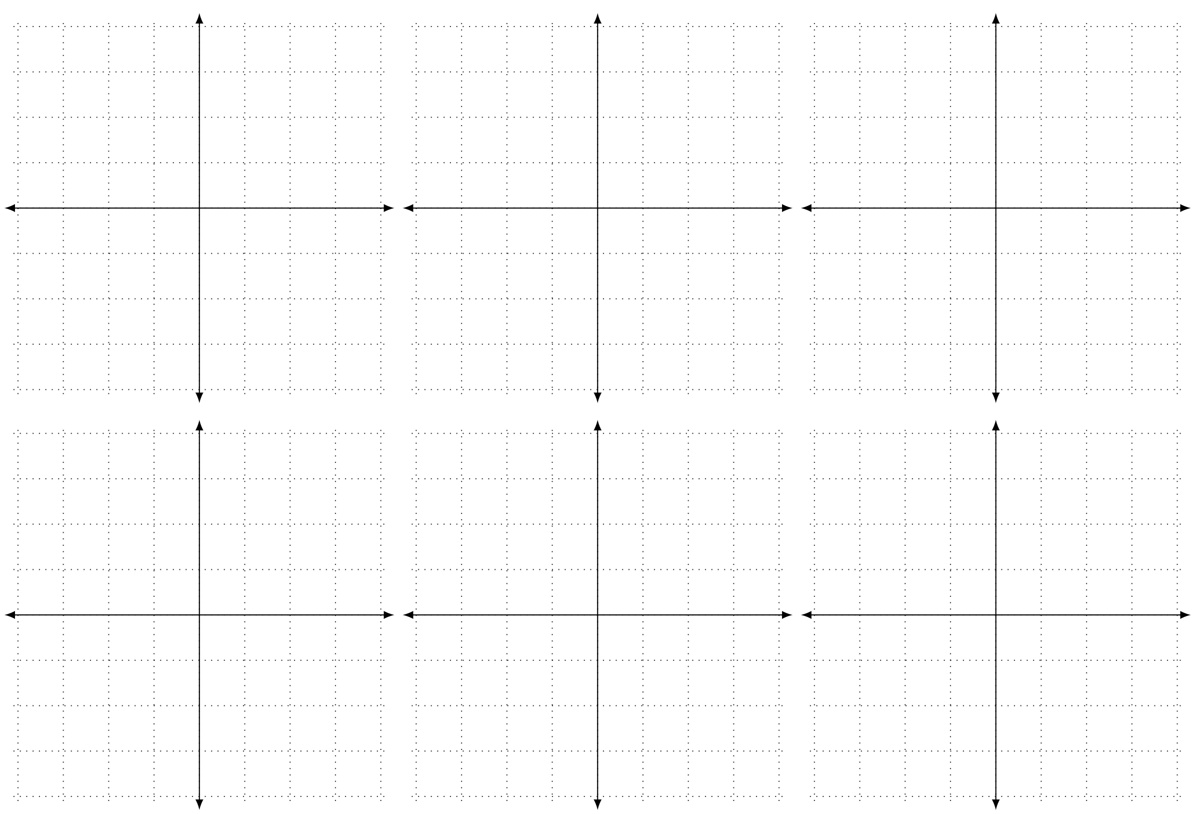
34.4 $D = \left\{ \vec{x} \in \mathbb{R}^2 \, : \, \vec{x} = \begin{bmatrix} 0 \\ t \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ for some } t \in \mathbb{R} \right\}$.

34.5 $E = \left\{ \vec{x} \in \mathbb{R}^2 \, : \, \vec{x} = \begin{bmatrix} 0 \\ t \end{bmatrix} \text{ or } \vec{x} = \begin{bmatrix} t \\ 0 \end{bmatrix} \text{ for some } t \in \mathbb{R} \right\}$.

34.6 $F = \left\{ \vec{x} \in \mathbb{R}^2 \, : \, \vec{x} = t\begin{bmatrix} 3 \\ 1 \end{bmatrix} \text{ for some } t \in \mathbb{R} \right\}$.

34.7 $G = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$.

34.8 $H = \operatorname{span}\{\vec{u}, \vec{v}\}$ for some unknown vectors $\vec{u}, \vec{v} \in \mathbb{R}^2$.



© Jason Siefken, 2015–2022

DEF **Basis**

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

DEF **Dimension**

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

---

35 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 $V = \operatorname{span}\{\vec{u}, \vec{v}, \vec{w}\}$.

35.1 Describe V.

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

35.3 Give a basis for V.

35.4 Give another basis for V.

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

35.6 What is the dimension of V?

© Jason Siefken, 2015–2022

36 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 $Q=\operatorname{span}\{\vec{b},\vec{c}\}$.

36.1 Give a basis for and the dimension of P.

36.2 Give a basis for and the dimension of Q.

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

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

© Jason Siefken, 2015–2022

 © Jason Siefken, 2015–2022

Matrix Representations of Systems of Linear Equations

In this module you will learn

- How to represent a system of linear equations as a matrix equation.
- Multiple ways to interpret solutions of systems of linear equations.
- How linear independence/dependence relates to solutions to matrix equations.
- How to use matrix equations to find normal vectors to lines or planes.

Matrix-vector multiplication gives a compact way to represent systems of linear equations.

Consider the system

$$\begin{cases} x+2y-2z=-15 \\ 2x+\ \ y-5z=-21 \\ x-4y+\ \ z=18 \end{cases}, \tag{6}$$

which is equivalent to the vector equation

$$\begin{bmatrix} x+2y-2z \\ 2x+y-5z \\ x-4y+z \end{bmatrix} = \begin{bmatrix} -15 \\ -21 \\ 18 \end{bmatrix}.$$

We can rewrite (6) using matrix-vector multiplication:

$$\underbrace{\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & -5 \\ 1 & -4 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -15 \\ -21 \\ 18 \end{bmatrix}.$$

The matrix A on the left is called the *coefficient matrix* because it is made up of the coefficients from equation (6).

By using coefficient matrices, every system of linear equations can be rewritten as a single matrix equation of the form

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

where A is a coefficient matrix, $\vec{x}$ is a column vector of variables, and $\vec{b}$ is a column vector of constants.

Example. Consider the one equation system

$$\left\{ x-4y+z=5 \right. \tag{7}$$

and the two-equation system

$$\begin{cases} x-4y+z=5 \\ \qquad y-z=9 \end{cases}. \tag{8}$$

Rewrite each system as a single matrix equation.

We can rewrite (7) as

$$\begin{bmatrix} 1 & -4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \end{bmatrix}.$$

Multiplying out to verify, we see,

$$\begin{bmatrix} 1 & -4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-4y+z \end{bmatrix} = \begin{bmatrix} 5 \end{bmatrix},$$

which is indeed equivalent to (7).

© Jason Siefken, 2015–2022 CC BY SA

Similarly, we can rewrite (8) as

$$\begin{bmatrix}1 & -4 & 1\\ 0 & 1 & -1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}5\\ 9\end{bmatrix}.$$

Multiplying out to verify, we see,

$$\begin{bmatrix}1 & -4 & 1\\ 0 & 1 & -1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}x-4y+z\\ 0x+y-z\end{bmatrix}=\begin{bmatrix}5\\ 9\end{bmatrix},$$

which is equivalent to (8).

Interpretations of Matrix Equations

The solution set to a system of linear equations, like

$$\begin{cases} x+2y-2z=-15\\ 2x+\ \ y-5z=-21\\ x-4y+\ \ z=18\end{cases},\tag{9}$$

can be interpreted as the intersection of three planes (or hyperplanes if there were more variables). Each equation (each row) specifies a plane, and the solution set is the intersection of all of these planes. Rewriting a system of equations in matrix form gives two additional ways to interpret the solution set.

The *Column* Picture

Using the column interpretation of matrix-vector multiplication, we see that system (9) is equivalent to

$$\begin{bmatrix}1 & 2 & -2\\ 2 & 1 & -5\\ 1 & -4 & 1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=x\begin{bmatrix}1\\ 2\\ 1\end{bmatrix}+y\begin{bmatrix}2\\ 1\\ -4\end{bmatrix}+z\begin{bmatrix}-2\\ -5\\ 1\end{bmatrix}=\begin{bmatrix}-15\\ -21\\ 18\end{bmatrix}.$$

We now see that asking, "*What coefficients allow* $\begin{bmatrix}1\\ 2\\ 1\end{bmatrix}$, $\begin{bmatrix}2\\ 1\\ -4\end{bmatrix}$, *and* $\begin{bmatrix}-2\\ -5\\ 1\end{bmatrix}$ *to form* $\begin{bmatrix}-15\\ -21\\ 18\end{bmatrix}$ *as a linear combination?*" is equivalent to asking, "*What are the solutions to system* (9)*?*" Here, $\begin{bmatrix}1\\ 2\\ 1\end{bmatrix}$, $\begin{bmatrix}2\\ 1\\ -4\end{bmatrix}$, and $\begin{bmatrix}-2\\ -5\\ 1\end{bmatrix}$ are the columns of the coefficient matrix.

The *Row* Picture

The row interpretation gives another perspective. Let $\vec r_1$, $\vec r_2$, and $\vec r_3$ be the rows of the coefficient matrix for system (9). Then, system (9) is equivalent to

$$\begin{bmatrix}1 & 2 & -2\\ 2 & 1 & -5\\ 1 & -4 & 1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}\vec r_1\\ \hline \vec r_2\\ \hline \vec r_3\end{bmatrix}\vec x=\begin{bmatrix}\vec r_1\cdot\vec x\\ \vec r_2\cdot\vec x\\ \vec r_3\cdot\vec x\end{bmatrix}=\begin{bmatrix}-15\\ -21\\ 18\end{bmatrix}.$$

In other words, we can interpret solutions to system (9) as vectors whose dot product with $\vec r_1$ is -15, whose dot product with $\vec r_2$ is -21, and whose dot product with $\vec r_3$ is 18. Given that the dot product has a geometric interpretation, this perspective is powerful (especially when the right side of the equation is all zeros!).

Interpreting Homogeneous Systems

Consider the homogeneous system/matrix equation

$$\begin{aligned} x+2y-2z&=0\\ 2x+\ \ y-5z&=0\\ x-4y+\ \ z&=0\end{aligned}\quad\iff\quad\underbrace{\begin{bmatrix}1 & 2 & -2\\ 2 & 1 & -5\\ 1 & -4 & 1\end{bmatrix}}_{A}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}.\tag{10}$$

Now, the column interpretation of system (10) is: *what linear combinations of the column vectors of A give $\vec 0$?* This directly relates to the question of whether the column vectors of A are linearly independent.

© Jason Siefken, 2015–2022

Let $\vec r_1$, $\vec r_2$, and $\vec r_3$ be the rows of A. The row interpretation of system (10) asks: *what vectors are simultaneously orthogonal to $\vec r_1$, $\vec r_2$, and $\vec r_3$?*

Takeaway. There are three ways to interpret solutions to a matrix equation $A\vec x = \vec b$: (i) the intersection of hyperplanes specified by the rows; (ii) what linear combinations of the columns of A give $\vec b$; (iii) what vectors yield the entries of $\vec b$ when dot producted with the rows of A.

Example. Find all vectors orthogonal to $\vec a = \begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\vec b = \begin{bmatrix}1\\2\\1\end{bmatrix}$.

To find all vectors orthogonal to $\vec a$ and $\vec b$ we need to find vectors $\vec x$ satisfying $\vec a \cdot \vec x = 0$ and $\vec b \cdot \vec x = 0$. This is equivalent to solving the matrix equation

$$\left[\begin{array}{c}\vec a\\\hline \vec b\end{array}\right]\vec x = \begin{bmatrix}\vec a\cdot\vec x\\ \vec b\cdot\vec x\end{bmatrix} = \underbrace{\begin{bmatrix}1&1&1\\1&2&1\end{bmatrix}}_{A}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}.$$

By row reducing A, we get

$$\operatorname{rref}(A) = \begin{bmatrix}1&0&1\\0&1&0\end{bmatrix},$$

and so the complete solution expressed in vector form is

$$\vec x = t\begin{bmatrix}-1\\0\\1\end{bmatrix}.$$

The row picture is particularly applicable when trying to find normal vectors.

Example. Let $\mathcal{Q}$ be the hyperplane specified in vector form by

$$\vec x = t\begin{bmatrix}1\\1\\-1\\1\end{bmatrix} + s\begin{bmatrix}0\\1\\0\\1\end{bmatrix} + r\begin{bmatrix}2\\0\\0\\0\end{bmatrix} + \begin{bmatrix}1\\2\\3\\4\end{bmatrix}.$$

Find a normal vector for $\mathcal{Q}$ and write $\mathcal{Q}$ in normal form.

Like the above example, since normal vectors for $\mathcal{Q}$ need to be orthogonal to $\vec d_1 = \begin{bmatrix}1\\1\\-1\\1\end{bmatrix}$, $\vec d_2 = \begin{bmatrix}0\\1\\0\\1\end{bmatrix}$, and $\vec d_3 = \begin{bmatrix}2\\0\\0\\0\end{bmatrix}$, we can find the normal vectors by solving

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

By row reducing A, we get

$$\operatorname{rref}(A) = \begin{bmatrix}1&0&0&0\\0&1&0&1\\0&0&1&0\end{bmatrix},$$

and so we get that the complete solution expressed in vector form is

$$\vec x = t\begin{bmatrix}0\\-1\\0\\1\end{bmatrix}.$$

© Jason Siefken, 2015–2022

Therefore, any non-zero multiple of $\begin{bmatrix}0\\-1\\0\\1\end{bmatrix}$ is a normal vector for $\mathcal{Q}$. For example, $\vec{n}=\begin{bmatrix}0\\-1\\0\\1\end{bmatrix}$ is a normal vector for $\mathcal{Q}$, and $\mathcal{Q}$ can be written in normal form as

$$\begin{bmatrix}0\\-1\\0\\1\end{bmatrix}\cdot\left(\begin{bmatrix}x\\y\\z\\w\end{bmatrix}-\begin{bmatrix}1\\2\\3\\4\end{bmatrix}\right)=0.$$

Practice Problems

1 Each system of equations below concerns the variables x, y, and z. Rewrite each system as a single matrix equation.

(a) $\begin{cases} x-y+z=1\\ 2x-y+z=2\\ 3x+y-z=3\end{cases}$

(b) $\{x+z=6$

(c) $\begin{cases} 5x-9y+2z=0\\ \quad\ -y \qquad =1\end{cases}$

2 Find all vectors orthogonal to:

(a) $\begin{bmatrix}1\\2\\3\end{bmatrix}$ and $\begin{bmatrix}2\\2\\3\end{bmatrix}$

(b) $\begin{bmatrix}0\\5\\6\end{bmatrix}$ and $\begin{bmatrix}1\\10\\2\end{bmatrix}$

(c) $\begin{bmatrix}1\\0\\0\end{bmatrix}$, $\begin{bmatrix}0\\1\\0\end{bmatrix}$, and $\begin{bmatrix}0\\0\\1\end{bmatrix}$

(d) $\begin{bmatrix}2\\6\\-1\end{bmatrix}$

3 Express each plane or hyperplane in normal form.

(a) $\vec{x}=t\begin{bmatrix}0\\2\\2\end{bmatrix}+s\begin{bmatrix}1\\1\\1\end{bmatrix}+\begin{bmatrix}1\\0\\3\end{bmatrix}$

(b) $\vec{x}=t\begin{bmatrix}1\\6\\8\end{bmatrix}+s\begin{bmatrix}2\\0\\2\end{bmatrix}+\begin{bmatrix}0\\0\\9\end{bmatrix}$

(c) $\vec{x}=t\begin{bmatrix}1\\5\\15\\20\end{bmatrix}+s\begin{bmatrix}3\\0\\35\\59\end{bmatrix}+r\begin{bmatrix}1\\4\\0\\18\end{bmatrix}+\begin{bmatrix}1\\6\\0\\0\end{bmatrix}$

4 Let $\mathcal{P}=\{(x,y,z)\,:\,2x+4y-4z=7\}$, let $\mathcal{Q}$ be the plane specified in vector form by

$$\vec{x}=t\begin{bmatrix}-1\\2\\0\end{bmatrix}+s\begin{bmatrix}5\\0\\2\end{bmatrix}+\begin{bmatrix}0\\7\\1\end{bmatrix},$$

and let $\mathcal{R}$ be the plane specified in normal form by

$$\begin{bmatrix}2\\-8\\2\end{bmatrix}\cdot\left(\begin{bmatrix}x\\y\\z\end{bmatrix}-\begin{bmatrix}1\\7\\0\end{bmatrix}\right)=0.$$

Find $\mathcal{P}\cap\mathcal{Q}\cap\mathcal{R}$.

© Jason Siefken, 2015–2022

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

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

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

37.3 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 *intersecting lines* (hint: think about systems of equations).

37.4 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 *linear combinations* (hint: think about the columns of A).

 © *Jason Siefken, 2015–2022*

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

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

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

© Jason Siefken, 2015–2022

39 Consider the system represented by

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

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

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

© *Jason Siefken, 2015–2022*

40 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 $\vec{r} \in \mathbb{R}^2$.

40.1 How many rows does M have?

40.2 Find such an M.

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

40.4 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?

© Jason Siefken, 2015–2022

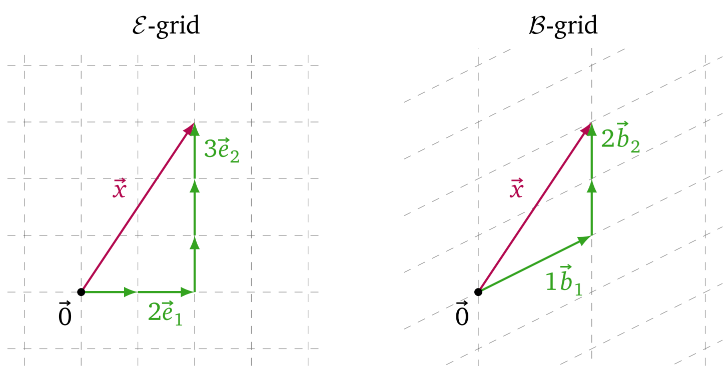
Coordinates & Change of Basis I

In this module you will learn

- Notation for representing a vector in multiple bases.
- The distinction between a vector and its representation.
- How to compute multiple representations of a vector.
- The definition of an *oriented* basis.

Recall that when we write $\vec x = \begin{bmatrix}2\\3\end{bmatrix}$, what we actually mean is $\vec x = 2\vec e_1 + 3\vec e_2$. The numbers 2 and 3 are called the coordinates of the vector $\vec x$ with respect to the standard basis. However, in general, subspaces have many bases, and so it is possible to represent a single vector in *many different ways* as coordinates with respect to different bases.

Let $\vec x = \begin{bmatrix}2\\3\end{bmatrix}$, let $\mathcal{E} = \{\vec e_1, \vec e_2\}$ be the standard basis for $\mathbb{R}^2$, and let $\mathcal{B} = \{\vec b_1, \vec b_2\}$, where $\vec b_1 = \begin{bmatrix}2\\1\end{bmatrix}$ and $\vec b_2 = \begin{bmatrix}0\\1\end{bmatrix}$, be another basis for $\mathbb{R}^2$. The coordinates of $\vec x$ with respect to $\mathcal{E}$ are $(2,3)$, but the coordinates of $\vec x$ with respect to $\mathcal{B}$ are $(1,2)$.



The coordinates $(2,3)$ and $(1,2)$ represent $\vec x$ equally well, and when solving problems, we should pick the coordinates that make our problem the easiest.[25] However, now that we are representing vectors in multiple bases, we need a way to keep track of what coordinates correspond to which basis.

Representation in a Basis.
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 $\alpha_1, \dots, \alpha_n$.

Example. 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 $\vec c_1 = \vec e_1 + \vec e_2$, and $\vec c_2 = 3\vec e_2$ be another basis for $\mathbb{R}^2$. Given that $\vec v = 2\vec e_1 - \vec e_2$, find $[\vec v]_{\mathcal{E}}$ and $[\vec v]_{\mathcal{C}}$.

[25] For example, maybe in one choice of coordinates, we can avoid all fractions in our calculations—this could be good if you're programming a computer that rounds decimals.

© Jason Siefken, 2015–2022

Since $\vec{v} = 2\vec{e}_1 - \vec{e}_2$, we know
$$[\vec{v}]_{\mathcal{E}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$
To find $[\vec{v}]_{\mathcal{C}}$, we need to write $\vec{v}$ as a linear combination of $\vec{c}_1$ and $\vec{c}_2$. Suppose
$$\vec{v} = x\vec{c}_1 + y\vec{c}_2$$
for some unknown scalars x and y. On the one hand,
$$\vec{v} = 2\vec{e}_1 - \vec{e}_2,$$
and on the other hand,
$$\vec{v} = x\vec{c}_1 + y\vec{c}_2 = x(\vec{e}_1 + \vec{e}_2) + 3y\vec{e}_2 = x\vec{e}_1 + (x+3y)\vec{e}_2.$$
Combining these two equations, we have
$$2\vec{e}_1 - \vec{e}_2 = x\vec{e}_1 + (x+3y)\vec{e}_2,$$
and so
$$(x-2)\vec{e}_1 + (x+3y+1)\vec{e}_2 = \vec{0}.$$
Since $\vec{e}_1$ and $\vec{e}_2$ are linearly independent, the only way for the above equation to be satisfied is if $x - 2 = 0$ ***and*** $x + 3y + 1 = 0$. Thus, we need to solve the system
$$\begin{cases} x = 2 \\ x + 3y = -1 \end{cases}.$$
After solving, we see $\vec{v} = 2\vec{c}_1 - \vec{c}_2$, and so
$$[\vec{v}]_{\mathcal{C}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

Notation Conventions

We need to revisit some past notation. Up to this point, we have been writing $\vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ to mean $\vec{x} = 2\vec{e}_1 + 3\vec{e}_2$. However, given the representation-in-a-basis notation, we should be writing
$$\vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}_{\mathcal{E}},$$
where $\mathcal{E}$ is the standard basis for $\mathbb{R}^2$.[26] We should write $\begin{bmatrix} 2 \\ 3 \end{bmatrix}_{\mathcal{E}}$ because the coordinates $(2, 3)$ refer to *different* vectors for *different* bases. However, most of the time we are only thinking about the standard basis. So, the convention we will follow is:

- If a problem involves only one basis, we may write $\begin{bmatrix} x \\ y \end{bmatrix}$ to mean $\begin{bmatrix} x \\ y \end{bmatrix}_{\mathcal{E}}$ where $\mathcal{E}$ is the standard basis.
- If there are multiple bases in a problem, we will always write $\begin{bmatrix} x \\ y \end{bmatrix}_{\mathcal{X}}$ to specify a vector in coordinates relative to a particular basis $\mathcal{X}$.

Takeaway. If a problem only involves the standard basis, we may use the notation we always have. If a problem involves multiple bases, we must ***always*** use representation-in-a-basis notation.

[26]One might wonder if we've just made a circular definition. In Module 6, $\vec{e}_1 \in \mathbb{R}^2$ was defined to be $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. But with our notation, this is the same as $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{\mathcal{E}}$, which is itself true by definition! To get around this, we need to declare the existence of the standard basis some other way. The physics solution is to define $\vec{e}_1$, $\vec{e}_2$, etc. as physical vectors in space. The abstract mathematics solution is to declare that $\vec{e}_1$, $\vec{e}_2$, etc. exist and are linearly independent and say nothing more.

© Jason Siefken, 2015–2022

True Vectors vs. Representations

[27]

The Belgian surrealist René Magritte painted the work above, which is subtitled, "This is not a pipe". Why? Because, of course, it is not a pipe. It is a painting of a pipe! In this work, Magritte points out a distinction that will soon become very important to us—the distinction between an object and a representation of that object.

Let $\vec{x} = 2\vec{e}_1 + 3\vec{e}_2 \in \mathbb{R}^2$. The vector $\vec{x}$ is a *real-life geometrical thing*, and to emphasize this, we will call $\vec{x}$ a *true* vector. In contrast, when we write the column matrix $[\vec{x}]_{\mathcal{E}} = \begin{bmatrix}2\\3\end{bmatrix}$, we are writing a *list of numbers*. The list of numbers $\begin{bmatrix}2\\3\end{bmatrix}$ has no meaning until we give it a meaning by assigning it a basis. For example, by writing $\begin{bmatrix}2\\3\end{bmatrix}_{\mathcal{E}}$, we declare that the numbers 2 and 3 are the coefficients of $\vec{e}_1$ and $\vec{e}_2$. By writing $\begin{bmatrix}2\\3\end{bmatrix}_{\mathcal{B}}$ where $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$, we declare that the numbers 2 and 3 are the coefficients of $\vec{b}_1$ and $\vec{b}_2$. Since a list of numbers without a basis has no meaning, we must acknowledge

$$\vec{x} \neq [\vec{x}]_{\mathcal{E}} = \begin{bmatrix}2\\3\end{bmatrix},$$

since the left side of the equation is a *true vector* and the right side is a *list of numbers*. Similarly, we must acknowledge

$$[\vec{x}]_{\mathcal{E}} \neq \begin{bmatrix}2\\3\end{bmatrix}_{\mathcal{E}} = \vec{x},$$

since the left side is a *list of numbers* and the right side is a *true vector*.

To help keep the notation straight in your head, for a basis $\mathcal{X}$, remember the rule

$$[\text{true vector}]_{\mathcal{X}} = \text{list of numbers} \qquad \text{and} \qquad [\text{list of numbers}]_{\mathcal{X}} = \text{true vector}.$$

It's easy to get confused when answering questions that involve multiple bases; precision will make these problems much easier.

Orientation of a Basis

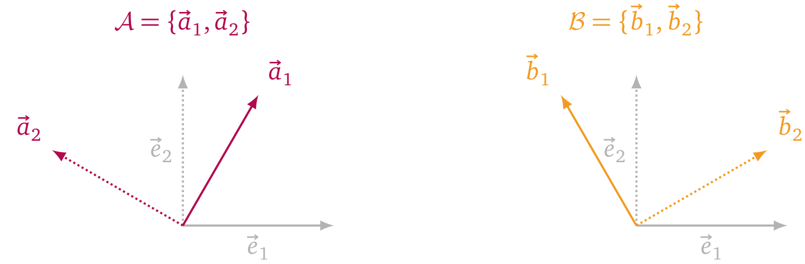
How can you tell the difference between a hand and a foot? They're similar in structure[28]—a hand has five fingers and a foot has five toes—but they're different in shape—fingers are much longer than toes and the thumb sticks off the hand at a different angle than the big toe sticks off the foot.

How about a harder question: how can you tell the difference between a left hand and a right hand? Any length or angle measurement you make on an (idealized) left hand or right hand will be identical. But, we know they're different because they differ in *orientation*.[29]

We'll build up to the definition of orientation in stages. Consider the ordered bases $\mathcal{E}$, $\mathcal{A}$, and $\mathcal{B}$ shown below.

[27] Image taken from Wikipedia: `https://en.wikipedia.org/wiki/File:MagrittePipe.jpg`
[28] We might say hands and feet are *topologically* equivalent.
[29] Other words for orientation include *chirality* and *handedness*.

© Jason Siefken, 2015–2022

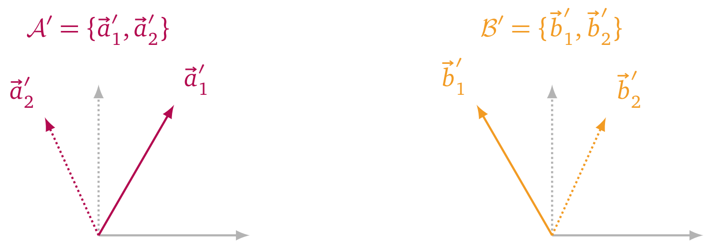


The $\mathcal{A}$ basis can be rotated to get the $\mathcal{E}$ basis while maintaining the proper order of the basis vectors (i.e., $\vec a_1 \mapsto \vec e_1$ and $\vec a_2 \mapsto \vec e_2$), but it is impossible to rotate the $\mathcal{B}$ basis to get the $\mathcal{E}$ basis while maintaining the proper order. In this case, we say that $\mathcal{E}$ and $\mathcal{A}$ have the same orientation and $\mathcal{E}$ and $\mathcal{B}$ have opposite orientations. Even though the lengths and angles between all vectors in the $\mathcal{A}$ basis and the $\mathcal{B}$ basis are the same, we can distinguish the $\mathcal{A}$ and $\mathcal{B}$ bases because they have different *orientations*.

Orientations of bases come in exactly two flavors: *right-handed* (or *positively oriented*) and *left-handed* (or *negatively oriented*). By convention, the standard basis is called right-handed.

Orthonormal bases—bases consisting of unit vectors that are orthogonal to each other—are called right-handed if they can be rotated to align with the standard basis, otherwise they are called left-handed. In this way, the right-hand–left-hand analogy should be clear: two right hands or two left hands can be rotated to align with each other, but a left hand and a right can never be rotated to alignment.

However, not all bases are orthonormal! Consider the bases $\mathcal{E}$, $\mathcal{A}'$, $\mathcal{B}'$.

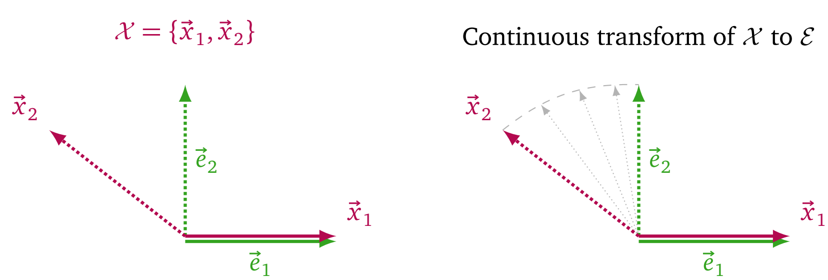


The bases $\mathcal{A}'$ and $\mathcal{B}'$ differ only slightly from $\mathcal{A}$ and $\mathcal{B}$. Neither can be *rotated* to obtain $\mathcal{E}$, however we'd still like to say $\mathcal{A}'$ is right-handed and $\mathcal{B}'$ is left-handed. The following, fully general definition, allows us to do so.

> **Orientation of a Basis.** 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***.

The term *continuously transformed* can be given a precise definition,[30] but it will be enough for us to imagine that a continuous transform between two bases is equivalent to a "movie" where one basis smoothly and without jumps transforms into the other.

Let's consider some examples. Let $\mathcal{X}=\{\vec x_1,\vec x_2\}$ as depicted below. We could imagine $\vec x_1,\vec x_2$ continuously transforming to $\vec e_1,\vec e_2$ by $\vec x_1$ staying in place and $\vec x_2$ smoothly moving along the dotted line.

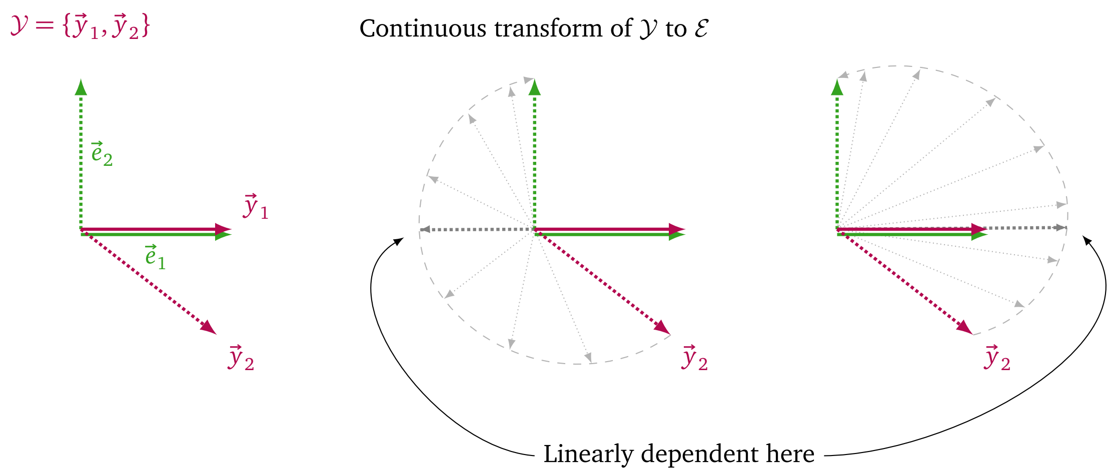


Because at every step along this motion, the set of $\vec x_1$ and the transformed $\vec x_2$ is linearly independent, $\mathcal{X}$ is *positively* oriented.

[30]Because you crave precision, here it is: the basis $\vec a_1,\dots,\vec a_n$ can be *continuously transformed* to the basis $\vec b_1,\dots,\vec b_n$ if there exists a continuous function $\Phi:[0,1]\to\{n\text{-tuples of vectors}\}$ such that $\Phi(0)=(\vec a_1,\dots,\vec a_n)$ and $\Phi(1)=(\vec b_1,\dots,\vec b_n)$. Here, continuity is defined in the multi-variable calculus sense.

© Jason Siefken, 2015–2022

Let $\mathcal{Y} = \{\vec{y}_1, \vec{y}_2\}$ as depicted below. We are in a similar situation, except this time, somewhere along $\vec{y}_2$'s path, the set of $\vec{y}_1$ and the transformed $\vec{y}_2$ becomes linearly dependent.

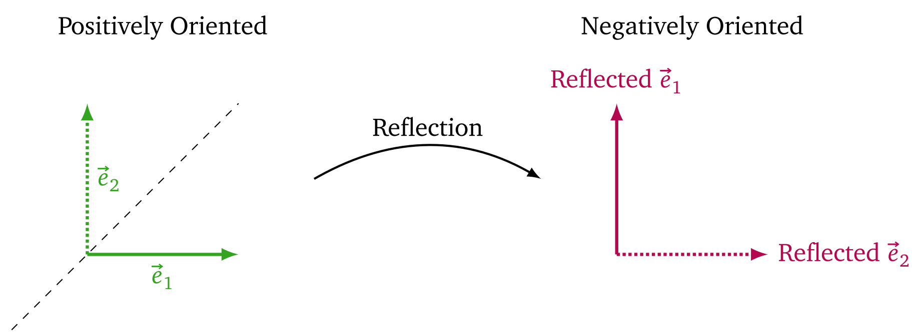


Maybe that was just bad luck and we might be able to transform along a different path and stay linearly independent. It turns out, we are doomed to fail, because $\mathcal{Y}$ is *negatively* oriented.

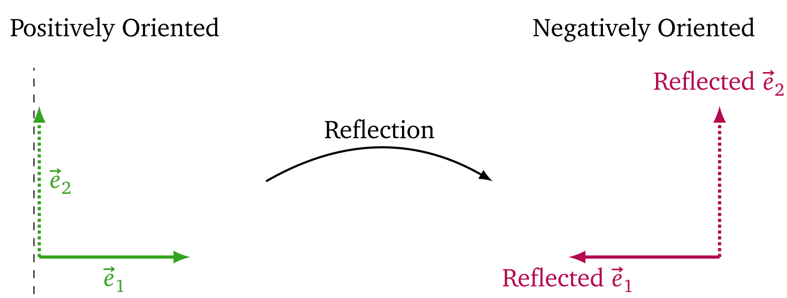
Using the definition of the orientation of a basis to answer questions is difficult because to determine that a basis is negatively oriented, you need to make a determination about *every possible* way to continuously transform a basis to the standard basis. This is hard enough in $\mathbb{R}^2$ and gets much harder in $\mathbb{R}^3$. Fortunately, we will encounter computational tools that will allow us to numerically determine the orientation of a basis, but, for now, the idea is what's important.

Reversing Orientation

Reflections reverse orientation and can manifest in two ways.[31] Consider the reflection of $\mathcal{E} = \{\vec{e}_1, \vec{e}_2\}$ across the line $y = x$.



This reflection sends $\{\vec{e}_1, \vec{e}_2\} \mapsto \{\vec{e}_2, \vec{e}_1\}$. Alternatively, reflection across the line $x = 0$ sends $\{\vec{e}_1, \vec{e}_2\} \mapsto \{-\vec{e}_1, \vec{e}_2\}$.



Both $\{\vec{e}_2, \vec{e}_1\}$ and $\{-\vec{e}_1, \vec{e}_2\}$, as ordered bases, are negatively oriented. This is indicative of a general theorem.

[31] Think back to hands. The left and right hands *are* reflections of each other.

© Jason Siefken, 2015–2022

Theorem. Let $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ be an ordered basis. The ordered basis obtained from $\mathcal{B}$ by replacing $\vec{b}_i$ with $-\vec{b}_i$ and the ordered basis obtained from $\mathcal{B}$ by swapping the order of $\vec{b}_i$ and $\vec{b}_j$ (with $i \neq j$) have the opposite orientation as $\mathcal{B}$.

Practice Problems

1 (a) Let $\vec{u} = \vec{e}_1 + 8\vec{e}_2$, $\vec{v} = -\vec{e}_1 + 3\vec{e}_2$, and $\vec{w} = 2\vec{e}_1$.

i. Find $[\vec{u}]_{\mathcal{E}}$, $[\vec{v}]_{\mathcal{E}}$ and $[\vec{w}]_{\mathcal{E}}$, where $\mathcal{E}$ is the standard basis for $\mathbb{R}^2$.

ii. Let $\mathcal{A} = \{3\vec{e}_1 + 2\vec{e}_2, 4\vec{e}_1 - \vec{e}_2\}$. Find $[\vec{u}]_{\mathcal{A}}$, $[\vec{v}]_{\mathcal{A}}$ and $[\vec{w}]_{\mathcal{A}}$.

iii. Let $\mathcal{B} = \{11\vec{e}_2, \vec{e}_1 + \frac{5}{2}\vec{e}_2\}$. Find $[\vec{u}]_{\mathcal{B}}$, $[\vec{v}]_{\mathcal{B}}$ and $[\vec{w}]_{\mathcal{B}}$.

(b) Let $\vec{q} = 11\vec{e}_2 - 4\vec{e}_3$, $\vec{r} = 5\vec{e}_1 - 12\vec{e}_2 + 8\vec{e}_3$, and $\vec{s} = \vec{e}_1 - 5\vec{e}_2 + 2\vec{e}_3$.

i. Find $[\vec{q}]_{\mathcal{E}}$, $[\vec{r}]_{\mathcal{E}}$ and $[\vec{s}]_{\mathcal{E}}$ where $\mathcal{E}$ is the standard basis for $\mathbb{R}^3$.

ii. Let $\mathcal{D} = \{\vec{e}_1 + 2\vec{e}_2, -3\vec{e}_1 + 5\vec{e}_2 - 4\vec{e}_3, -8\vec{e}_1 + 4\vec{e}_2 + 11\vec{e}_3\}$. Find $[\vec{q}]_{\mathcal{D}}$, $[\vec{r}]_{\mathcal{D}}$ and $[\vec{s}]_{\mathcal{D}}$.

iii. Let $\mathcal{F} = \{\vec{e}_1 + 4\vec{e}_2 + 4\vec{e}_3, -3\vec{e}_1 + 20\vec{e}_2, 21\vec{e}_2 + 16\vec{e}_3\}$. Find $[\vec{q}]_{\mathcal{F}}$, $[\vec{r}]_{\mathcal{F}}$ and $[\vec{s}]_{\mathcal{F}}$.

(c) How does negating one vector in a basis change the orientation?

(d) How does swapping the order of two different vectors in a basis change the orientation?

2 (a) Let $[\vec{a}]_{\mathcal{E}} = \begin{bmatrix} 5 \\ -12 \end{bmatrix}$ where $\mathcal{E}$ is the standard basis for $\mathbb{R}^2$. Find a basis $\mathcal{M}$ for $\mathbb{R}^2$ such that $[\vec{a}]_{\mathcal{M}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

(b) Let $[\vec{b}]_{\mathcal{E}} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ where $\mathcal{E}$ is the standard basis for $\mathbb{R}^3$. Find a basis $\mathcal{N}$ for $\mathbb{R}^3$ such that $[\vec{b}]_{\mathcal{N}} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$.

3 Determine the orientation of each of the following bases for $\mathbb{R}^2$.

(a) $\left\{\begin{bmatrix} -5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \end{bmatrix}\right\}$

(b) $\left\{\begin{bmatrix} 2 \\ -6 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \end{bmatrix}\right\}$

(c) $\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$

(d) $\left\{\begin{bmatrix} 6 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}\right\}$

4 (a) Determine the orientation of the basis $\mathcal{V} = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ for $\mathbb{R}^3$ where $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $\vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$.

(b) Consider the basis $\mathcal{V}' = \{\vec{v}_1, 3\vec{v}_2, \vec{v}_3\}$ for $\mathbb{R}^3$. What is the orientation of this basis?

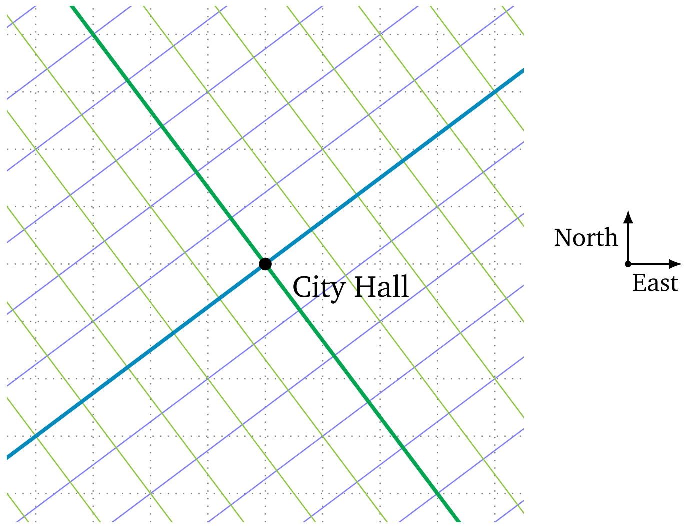
(c) Consider the basis $\mathcal{V}'' = \{-4\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ for $\mathbb{R}^3$. What is the orientation of this basis?

5 (a) Give two examples of positively oriented bases for $\mathbb{R}^2$ and briefly explain how you know their orientation.

(b) Give two examples of negatively oriented bases for $\mathbb{R}^2$ and briefly explain how you know their orientation.

© Jason Siefken, 2015–2022

41 The mythical 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.

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

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

41.3 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 $\operatorname{span}\{\vec{e}_1, \vec{e}_2\} = \operatorname{span}\{\vec{d}_1, \vec{d}_2\}$?

41.4 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.

© Jason Siefken, 2015–2022

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 $\vec{v} = \alpha_1 \vec{b}_1 + \dots + \alpha_n \vec{b}_n$.

Conversely,

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}_{\mathcal{B}} = \alpha_1 \vec{b}_1 + \dots + \alpha_n \vec{b}_n$$

is notation for the linear combination of $\vec{b}_1, \dots, \vec{b}_n$ with coefficients $\alpha_1, \dots, \alpha_n$.

---

42 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 $\mathbb{R}^2$.

42.1 Express $\vec{c}_1$ and $\vec{c}_2$ as a linear combination of $\vec{e}_1$ and $\vec{e}_2$.

42.2 Express $\vec{e}_1$ and $\vec{e}_2$ as a linear combination of $\vec{c}_1$ and $\vec{c}_2$.

42.3 Let $\vec{v} = 2\vec{e}_1 + 2\vec{e}_2$. Find $[\vec{v}]_{\mathcal{E}}$ and $[\vec{v}]_{\mathcal{C}}$.

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

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

42.6 What is YX?

© Jason Siefken, 2015–2022

Orientation of a Basis

DEF

The ordered basis $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ 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***.

---

43 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 θ be the angle between $\vec{u}_\theta$ and $\vec{e}_1$ measured counter-clockwise starting at $\vec{e}_1$.

43.1 For which θ is $\{\vec{e}_1, \vec{u}_\theta\}$ a linearly independent set?

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

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

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

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

© Jason Siefken, 2015–2022

© Jason Siefken, 2015–2022

Linear Transformations

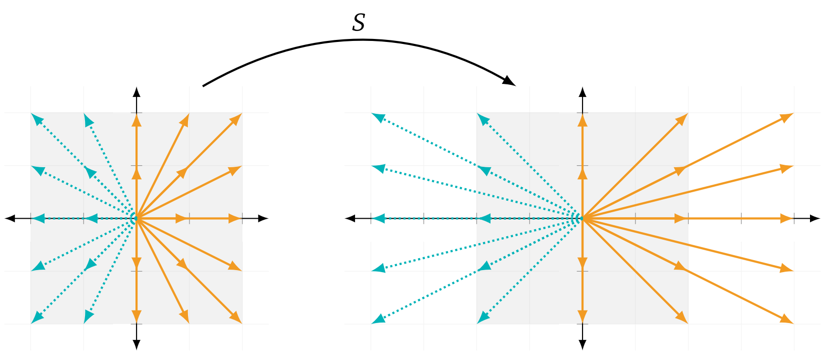
In this module you will learn

- The definition of a linear transformation.
- The definition of the image of a set under a transformation.
- How to prove whether a transformation is linear or not.
- How to find a matrix for a linear transformation.
- The difference between a matrix and a linear transformation.

Now that we have a handle on the basics of vectors, we can start thinking about *transformations*. Transformation (or map) is another word for a function, and transformations show up any time you need to describe vectors changing. For example, the transformation

$$S:\mathbb{R}^2\to\mathbb{R}^2 \qquad \text{defined by} \qquad \begin{bmatrix}x\\y\end{bmatrix}\mapsto\begin{bmatrix}2x\\y\end{bmatrix}$$

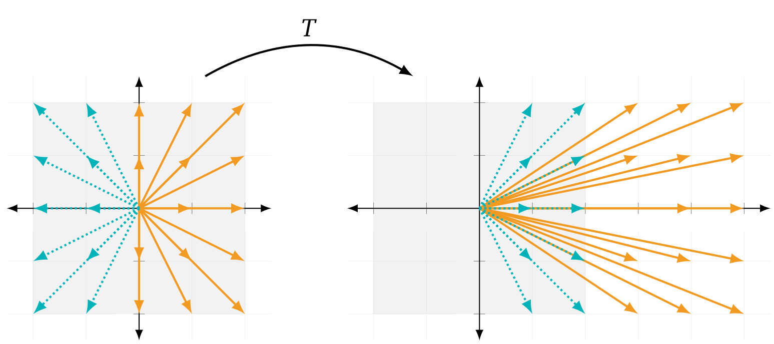
stretches all vectors in the $\vec{e}_1$ direction by a factor of 2.



The transformation

$$T:\mathbb{R}^2\to\mathbb{R}^2 \qquad \text{defined by} \qquad \begin{bmatrix}x\\y\end{bmatrix}\mapsto\begin{bmatrix}x+3\\y\end{bmatrix}$$

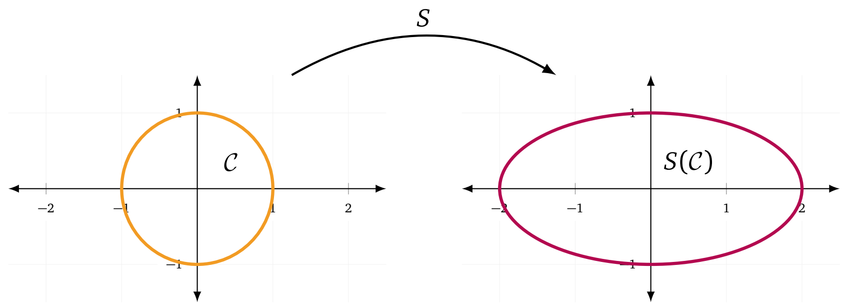
translates all vectors 3 units in the $\vec{e}_1$ direction.



Images of Sets

Recall the transformation $S:\mathbb{R}^2\to\mathbb{R}^2$ defined by $\begin{bmatrix}x\\y\end{bmatrix}\mapsto\begin{bmatrix}2x\\y\end{bmatrix}$. If we had a bunch of vectors in the plane, applying S would stretch those vectors in the $\vec{e}_1$ direction by a factor of 2. For example, let C be the circle of radius 1 centered at $\vec{0}$. Applying S to all the vectors that make up C produces an ellipse.

© Jason Siefken, 2015–2022



The operation of applying a transformation to a specific set of vectors and seeing what results is called taking the *image* of a set.

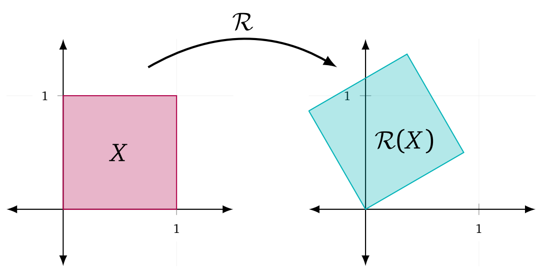
Image of a Set. Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a transformation and let $X \subseteq \mathbb{R}^n$ be a set. The ***image of the set X under L***, denoted $L(X)$, is the set

$$L(X) = \{\vec{y} \in \mathbb{R}^m \ : \ \vec{y} = L(\vec{x}) \text{ for some } \vec{x} \in X\}.$$

In plain language, the image of a set X under a transformation L is the set of all outputs of L when the inputs come from X.

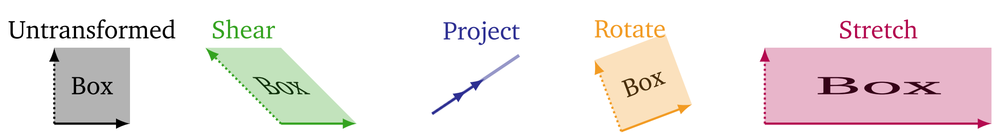
If you think of sets in $\mathbb{R}^n$ as black-and-white “pictures” (a point is black if it’s in the set and white if it’s not), then the image of a set under a transformation is the output after applying the transformation to the “picture”.

Images allow one to describe complicated geometric figures in terms of an original figure and a transformation. For example, let $\mathcal{R} : \mathbb{R}^2 \to \mathbb{R}^2$ be rotation counter clockwise by 30° and let $X = \{x\vec{e}_1 + y\vec{e}_2 \ : \ x, y \in [0, 1]\}$ be the filled-in unit square. Then, $\mathcal{R}(X)$ is the filled-in unit square that meets the x-axis at an angle of 30°. Try describing that using set builder notation!



Linear Transformations

Linear algebra’s main focus is the study of a special category of transformations: the *linear* transformations. Linear transformations include rotations, dilations (stretches), shears, and more.



Linear transformations are an important type of transformation because (i) we have a complete theory of linear transformations (non-linear transformations are notoriously difficult to understand), and (ii) many non-linear transformations can be approximated by linear ones.[32] All this is to say that linear transformations are worthy of our study.

Without further ado, let’s define what it means for a transformation to be linear.

Linear Transformation. 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 α.

[32] Just like in one-variable calculus where if you zoom into a function at a point its graph looks like a line, if you zoom into a (non-linear) transformation at a point, it looks like a linear one.

© Jason Siefken, 2015–2022

In plain language, the transformation T is linear, or has the property of *linearity*, if it distributes over addition and scalar multiplication. In other words, *T distributes over linear combinations*.

Example. Let $S : \mathbb{R}^2 \to \mathbb{R}^2$ and $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$\begin{bmatrix} x \\ y \end{bmatrix} \overset{S}{\mapsto} \begin{bmatrix} 2x \\ y \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} x \\ y \end{bmatrix} \overset{T}{\mapsto} \begin{bmatrix} x \\ y+4 \end{bmatrix}.$$

For each of S and T, determine whether the transformation is linear.

Let $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ be vectors, and let α be a scalar.
We first consider $\mathcal{S}$. We need to verify that $\mathcal{S}(\vec{u}+\vec{v}) = \mathcal{S}(\vec{u}) + \mathcal{S}(\vec{v})$ and $\mathcal{S}(\alpha\vec{u}) = \alpha\mathcal{S}(\vec{u})$.
Computing, we see

$$\mathcal{S}(\vec{u}+\vec{v}) = \mathcal{S}\left(\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \end{bmatrix}\right) = \begin{bmatrix} 2u_1+2v_1 \\ u_2+v_2 \end{bmatrix} = \begin{bmatrix} 2u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 2v_1 \\ v_2 \end{bmatrix} = \mathcal{S}(\vec{u}) + \mathcal{S}(\vec{v})$$

and

$$\mathcal{S}(\alpha\vec{u}) = \begin{bmatrix} 2\alpha u_1 \\ \alpha u_2 \end{bmatrix} = \mathcal{S}\begin{bmatrix} \alpha u_1 \\ \alpha u_2 \end{bmatrix} = \alpha \begin{bmatrix} 2u_1 \\ u_2 \end{bmatrix} = \alpha\mathcal{S}(\vec{u}),$$

and so $\mathcal{S}$ satisfies all the properties of a linear transformation.

Next we consider $\mathcal{T}$. Notice that $\mathcal{T}(\vec{u}+\vec{v}) = \begin{bmatrix} u_1+v_1 \\ u_2+v_2+4 \end{bmatrix}$ doesn't look like $\mathcal{T}(\vec{u}) + \mathcal{T}(\vec{v}) = \begin{bmatrix} u_1+v_1 \\ u_2+v_2+8 \end{bmatrix}$.
Therefore, we will guess that $\mathcal{T}$ is not linear and look for a counter example.
Using $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, we see

$$\mathcal{T}(\vec{e}_1+\vec{e}_2) = \mathcal{T}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 5 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 4 \end{bmatrix} + \begin{bmatrix} 0 \\ 5 \end{bmatrix} = \mathcal{T}(\vec{e}_1) + \mathcal{T}(\vec{e}_2).$$

Since at least one required property of a linear transformation is violated, $\mathcal{T}$ cannot be a linear transformation.

Function Notation vs. Linear Transformation Notation

Linear transformations are just special types of functions. In calculus, it is traditional to use lower case letters for a function and parenthesis "(" and ")" around the input to the function.

$$\underbrace{f : \mathbb{R} \to \mathbb{R}}_{\text{a function named } f} \qquad \underbrace{f(x)}_{f \text{ evaluated at } x}$$

For (linear) transformations, it is traditional to use capital letters to describe the function/transformation and parenthesis around the input are optional.

$$\underbrace{T : \mathbb{R}^n \to \mathbb{R}^m}_{\text{a transformation named } T} \qquad \underbrace{T(\vec{x})}_{T \text{ evaluated at } \vec{x}} \qquad \underbrace{T\vec{x}}_{\text{also } T \text{ evaluated at } \vec{x}}$$

Since sets are also traditionally written using capital letters, sometimes a font variant is used to when writing the transformation or the set. For example, we might use a regular X to denote a set and a calligraphic $\mathcal{T}$ to describe a transformation.

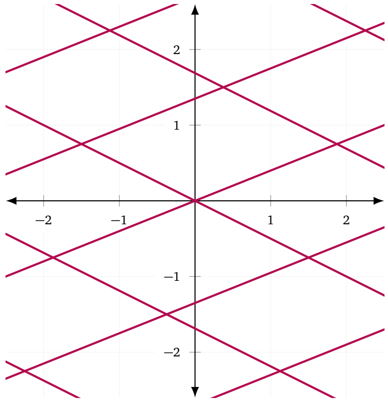
Another difference you might not be used to is that, in linear algebra, we make a careful distinction between a function and its output. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. In calculus, you might consider the phrases "the function f" and "the function $f(x)$" to both make sense. In linear algebra, the first phrase is valid and the second is *not*. By writing $f(x)$, we are indicating "the output of the function f when x is input". So, properly we should say "the *number* $f(x)$".

This distinction might seem pedantic now, but by keeping our functions as functions and our numbers/vectors as numbers/vectors, we can avoid some major confusion in the future.

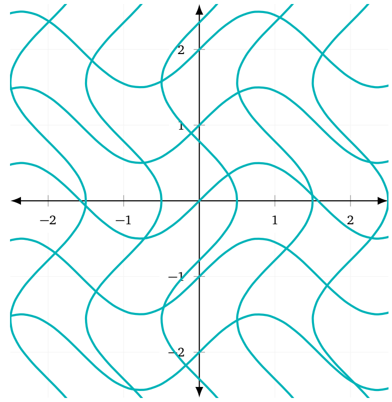
The "look" of a Linear Transformation

Images under linear transformations have a certain look to them. Based just on the word *linear* you can probably guess which figure below represents the image of a grid under a linear transformation.

© Jason Siefken, 2015–2022



Linear



Non-linear

Let's prove some basic facts about linear transformations.

Theorem. If $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation, then $T(\vec 0)=\vec 0$.

Proof. Suppose $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation and $\vec v\in\mathbb{R}^n$. We know that $0\vec v=\vec 0$, so by linearity we have

$$T(\vec 0)=T(0\vec v)=0T(\vec v)=\vec 0.$$

■

Theorem. If $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation, then T takes lines to lines (or points).

Proof. Suppose $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation and let $\ell\subseteq\mathbb{R}^n$ be the line given in vector form by $\vec x=t\vec d+\vec p$. We want to prove that $T(\ell)$, the image of ℓ under the transformation T, is a line or a point.

By definition, every point in ℓ takes the form $t\vec d+\vec p$ for some scalar t. Therefore, every point in $T(\ell)$ takes the form $T(t\vec d+\vec p)$ for some scalar t. But, T is a linear transformation, so

$$T(t\vec d+\vec p)=tT(\vec d)+T(\vec p).$$

If $T(\vec d)\neq\vec 0$, then $\vec x=tT(\vec d)+T(\vec p)$ describes a line in vector form and so $T(\ell)$ is a line. If $T(\vec d)=\vec 0$, then $T(\ell)=\{t\vec 0+T(\vec p)\,:\,t\text{ is a scalar}\}=\{T(\vec p)\}$ is a point. ■

Theorem. If $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation, then T takes parallel lines to parallel lines (or points).

Proof. Suppose $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation and let ℓ_1 and ℓ_2 be parallel lines. Then, we may describe ℓ_1 in vector form as $\vec x=t\vec d+\vec p_1$ and we may describe ℓ_2 in vector form as $\vec x=t\vec d+\vec p_2$. Note that since the lines are parallel, the direction vectors are the same.

Now, $T(\ell_1)$ can be described in vector form by

$$\vec x=tT(\vec d)+T(\vec p_1)$$

and $T(\ell_2)$ can be described in vector form by

$$\vec x=tT(\vec d)+T(\vec p_2).$$

Written this way and provided $T(\ell_1)$ and $T(\ell_2)$ are actually lines, we immediately see that $T(\ell_1)$ and $T(\ell_2)$ have the same direction vectors and hence are parallel.

If $T(\ell_1)$ is instead a point, then we must have $T(\vec d)=\vec 0$, and so $T(\ell_2)$ must also be a point. ■

Theorem. If $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation, then T takes subspaces to subspaces.

Proof. Let $T:\mathbb{R}^n\to\mathbb{R}^m$ be a linear transformation and let $V\subseteq\mathbb{R}^n$ be a subspace. We need to show that $T(V)$ satisfies the properties of a subspace.

Since V is non-empty, we know $T(V)$ is non-empty.

Let $\vec x,\vec y\in T(V)$. By definition, there are vectors $\vec u,\vec v\in V$ so that

$$\vec x=T(\vec u)\qquad\text{and}\qquad\vec y=T(\vec v).$$

© Jason Siefken, 2015–2022

Since T is linear, we know

$$\vec{x}+\vec{y}=T(\vec{u})+T(\vec{v})=T(\vec{u}+\vec{v}).$$

Because V is a subspace, we know $\vec{u}+\vec{v}\in V$ and so we conclude $\vec{x}+\vec{y}=T(\vec{u}+\vec{v})\in T(V)$.

Similarly, for any scalar α we have

$$\alpha\vec{x}=\alpha T(\vec{u})=T(\alpha\vec{u}).$$

Since V is a subspace, $\alpha\vec{u}\in V$ and so $\alpha\vec{x}=T(\alpha\vec{u})\in T(V)$. ■

Linear Transformations and Proofs

When proving things in math, you have all of logic at your disposal, and that freedom can be combined with creativity to show some truly amazing things. But, for better or for worse, proving whether or not a transformation is linear usually doesn't require substantial creativity.

Let $T:\mathbb{R}^n\to\mathbb{R}^n$ be defined by $T(\vec{v})=2\vec{v}$. To show that T is linear, we need to show that for *all* inputs $\vec{x}$ and $\vec{y}$ and for *all* scalars α we have

$$T(\vec{x}+\vec{y})=T(\vec{x})+T(\vec{y})\qquad\text{and}\qquad T(\alpha\vec{x})=\alpha T(\vec{x}).$$

But, there are an infinite number of choices for $\vec{x}$, $\vec{y}$, and α. How can we argue about all of them at once?

Consider the following proof that T is linear.

Proof. Let $\vec{x},\vec{y}\in\mathbb{R}^n$ and let α be a scalar. By applying the definition of T, we see

$$T(\vec{x}+\vec{y})=2(\vec{x}+\vec{y})=2\vec{x}+2\vec{y}=T(\vec{x})+T(\vec{y}).$$

Similarly,

$$T(\alpha\vec{x})=2(\alpha\vec{x})=\alpha(2\vec{x})=\alpha T(\vec{x}).$$

Since T satisfies the two properties of a linear transformation, T is a linear transformation. ■

This proof starts out with "let $\vec{x},\vec{y}\in\mathbb{R}^n$ and let α be a scalar". In what follows, the only properties of $\vec{x}$ and $\vec{y}$ we use come from the fact that they're in $\mathbb{R}^n$ (the domain of T) and the only fact about α we use is that it's a scalar. Because of this, $\vec{x}$, and $\vec{y}$ are considered *arbitrary* vectors and α is an *arbitrary* scalar. Put another way, the argument that followed would work for every single pair of vectors $\vec{x},\vec{y}\in\mathbb{R}^n$ and for every scalar α. Thus, by fixing arbitrary vectors at the start of our proof, we are (i) able to argue about all vectors at once while (ii) having named vectors that we can actually use in equations.

Takeaway. Starting a linearity proof with "***let*** $\vec{x},\vec{y}\in\mathbb{R}^n$ ***and let*** α ***be a scalar***" allows you to argue about all vectors and scalars simultaneously.

The proof given above is very typical, and almost every proof of the linearity of a function $T:\mathbb{R}^n\to\mathbb{R}^m$ will look something like

Proof. Let $\vec{x},\vec{y}\in\mathbb{R}^n$ and let α be a scalar. By applying the definition of T, we see

$$T(\vec{x}+\vec{y})=\text{application(s) of the definition}=T(\vec{x})+T(\vec{y}).$$

Similarly,

$$T(\alpha\vec{x})=\text{application(s) of the definition}=\alpha T(\vec{x}).$$

Since T satisfies the two properties of a linear transformation, T is a linear transformation. ■

This isn't to say that proving whether or not a transformation is linear is *easy*, but all the cleverness and insight required appears in the "application(s) of the definition" parts.

What about showing a transformation is *not* linear? Here we don't need to show something true for all vectors and all scalars. We only need to show something is false for *one* pair of vectors or *one* pair of a vector and a scalar.

When proving a transformation is not linear, we can pick one of the properties of linearity (distribution over vector addition or distribution over scalar multiplication) and a *single example* where that property fails.[33]

[33] It's often tempting to argue that the properties of linearity fail for all inputs, but this is a dangerous path! For instance, if $T(\vec{0})=\vec{0}$, then $T(\vec{a})=T(\vec{a}+\vec{0})=T(\vec{a})+T(\vec{0})=T(\vec{a})$ *regardless* of whether T is linear or not.

© Jason Siefken, 2015–2022

Example. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be defined by $T(\vec{x}) = \vec{x} + \vec{e}_1$. Show that T is *not* linear.

Proof. We will show that T does not distribute with respect to scalar multiplication. Observe that

$$T(2\vec{0}) = T(\vec{0}) = \vec{e}_1 \neq 2\vec{e}_1 = 2T(\vec{0}).$$

Therefore, T cannot be a linear transformation. ∎

Matrix Transformations

We already know two ways to interpret matrix multiplication—linear combinations of the columns and dot products with the rows—and we're about to have a third.

Let $M = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$. For a vector $\vec{v} \in \mathbb{R}^2$, $M\vec{v}$ is another vector in $\mathbb{R}^2$. In this way, we can think of multiplication by M as a transformation on $\mathbb{R}^2$. Define

$$T : \mathbb{R}^2 \to \mathbb{R}^2 \qquad \text{by} \qquad T(\vec{x}) = M\vec{x}.$$

Because T is defined by a matrix, we call T a *matrix transformation*. It turns out all matrix transformations are linear transformations and most linear transformations are matrix transformations.[34]

When it comes to specifying linear transformations, matrices are heroes, providing a compact notation (just like they did for systems of linear equations). For example, we could say, "The linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ that doubles the x-coordinate and triples the y-coordinate", or we could say, "The matrix transformation given by $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$".

When talking about matrices and linear transformations, we must keep in mind that they are not the same thing. A matrix is a box of numbers and has no meaning until we give it meaning. A linear transformation is a function that inputs vectors and outputs vectors. We can *specify* a linear transformation using a matrix, but a matrix by itself is *not* a linear transformation.[35]

Takeaway. Matrices and linear transformations are closely related, but they aren't the same thing.

So what are some correct ways to specify a linear transformation using a matrix? For a matrix M, the following are correct.

- The transformation T defined by $T(\vec{x}) = M\vec{x}$.
- The transformation given by multiplication by M.
- The transformation induced by M.
- The matrix transformation given by M.
- The linear transformation whose matrix is M.

Finding a Matrix for a Linear Transformation

Every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ has a matrix, and we can use basic algebra to find an appropriate matrix.

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Since T inputs vectors with n coordinates and outputs vectors with m coordinates, we know any matrix for T must be $m \times n$. The process of finding a matrix for T can now be summarized as follows: (i) create an $m \times n$ matrix of variables, (ii) use known input-output pairs for T to set up a system of equations involving the unknown variables, (iii) solve for the variables.

Example. Let $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $\mathcal{T}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + y \\ x \end{bmatrix}$. Find a matrix, M, for $\mathcal{T}$.

Because $\mathcal{T}$ is a transformation for $\mathbb{R}^2 \to \mathbb{R}^2$, M will be a 2×2 matrix. Let

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

[34] If you believe in the axiom of choice and you allow infinitely sized matrices, every linear transformation can be expressed as a matrix transformation.

[35] Consider the function defined by $f(x) = 2x$. You would never say that the function f is 2!

© Jason Siefken, 2015–2022

We now need to use input-output pairs to "calibrate" M. We know

$$\mathcal{T}\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}3\\1\end{bmatrix}\qquad\text{and}\qquad\mathcal{T}\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}.$$

Since M is a matrix for $\mathcal{T}$, we know $\mathcal{T}\vec{x}=M\vec{x}$ for all $\vec{x}$, and so

$$M\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}a+b\\c+d\end{bmatrix}=\begin{bmatrix}3\\1\end{bmatrix}$$

and

$$M\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}b\\d\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}.$$

This gives us the system of equations

$$\begin{cases}a+b\qquad\quad=3\\ \qquad\quad c+d=1\\ \quad b\qquad\quad\;=1\\ \qquad\qquad\; d=0\end{cases},$$

and solving this system tells us

$$M=\begin{bmatrix}a&b\\c&d\end{bmatrix}=\begin{bmatrix}2&1\\1&0\end{bmatrix}.$$

Practice Problems

1 For each transformation listed below, prove whether or not it is a linear transformation.

(a) $\mathcal{A}:\mathbb{R}^2\to\mathbb{R}^2$ defined by $\mathcal{A}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}-x\\y\end{bmatrix}$.

(b) $\mathcal{B}:\mathbb{R}^2\to\mathbb{R}^2$ defined by $\mathcal{B}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}-x-1\\y\end{bmatrix}$.

(c) $\text{id}:\mathbb{R}^2\to\mathbb{R}^2$, the function that leaves its input unchanged.

(d) $\mathcal{C}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{C}$ sends all vectors above the x-axis to $\vec{0}$ and all vectors on or below the x-axis to $-\vec{e}_2$.

2 Draw the image of the unit circle under each transformation listed in 1.

3 Let $M=\begin{bmatrix}1&-2&3\\-4&5&-6\end{bmatrix}$ and let T_M be the corresponding matrix transformation.

(a) Determine the domain and codomain of T_M.

(b) Calculate $T_M\left(\begin{bmatrix}2\\-1\\1\end{bmatrix}\right)$.

(c) Find the image of the standard basis vectors of the domain under T_M.

4 Find a matrix for each transformation below, or explain why no such matrix exists.

(a) $\mathcal{S}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{S}$ is the transformation that doubles every vector.

(b) $\mathcal{R}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{R}$ is rotation clockwise by $135°$.

(c) $\mathcal{T}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{T}$ translates every vector by $3\vec{e}_1$.

(d) $\mathcal{P}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{P}$ is projection onto the y-axis.

(e) $\mathcal{F}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{F}$ is reflection over the line $y=x$.

5 Let $S:\mathbb{R}^n\to\mathbb{R}^m$ and $T:\mathbb{R}^n\to\mathbb{R}^m$ be linear transformations, and define the transformation $R:\mathbb{R}^n\to\mathbb{R}^m$ by $R(\vec{x})=S(\vec{x})+T(\vec{x})$. Show that R is also linear.

6 For a fixed vector $\vec{a}\in\mathbb{R}^3$, define the function $D_{\vec{a}}$ by $D_{\vec{a}}(\vec{x})=\vec{a}\cdot\vec{x}$.

(a) Identify the domain and codomain of $D_{\vec{a}}$.

(b) Show that when $\vec{a}=\vec{e}_1$, then $D_{\vec{a}}$ is a linear transformation.

(c) Is $D_{\vec{a}}$ a linear transformation for all $\vec{a}$? Prove your claim.

(d) Find a matrix for $D_{\vec{a}}$ or explain why no such matrix exists.

7 Let $T:\mathbb{R}^2\to\mathbb{R}^2$ be a transformation with the property that $T(V)$ is a subspace whenever V is a subspace. Is this enough information to conclude that T is a linear transformation? Justify your answer.

8 For each statement below, determine whether it is true or false. Justify your answer.

(a) Every transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be represented by a matrix.

(b) The image of a subspace under a linear transformation is not a subspace.

(c) A transformation that takes every vector in the domain to $\vec{0}$ is not linear.

(d) Every matrix is a linear transformation.

(e) Parallel lines stay parallel under a linear transformation.

9 Let $\mathcal{T}:\mathbb{R}^2\to\mathbb{R}^2$ be the transformation that doubles the length of its input. The following statements about $\mathcal{T}$ are either incorrect or incomplete. Fix each statement so that it is correct and complete.

(a) $\mathcal{T}=\begin{bmatrix}2&0\\0&2\end{bmatrix}$.

© Jason Siefken, 2015–2022

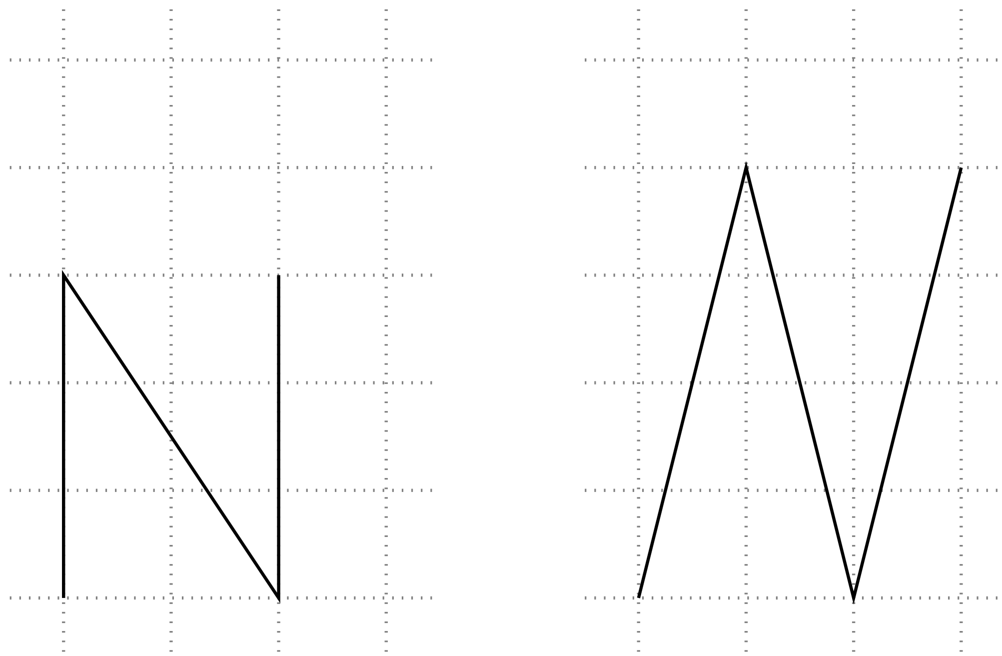
(b) Since $\mathcal{T}\vec{x} = 2\vec{x}$ for every $\vec{x}$, we can say $\mathcal{T} = 2$.

(c) $\mathcal{T}$ is a linear transformation because $2(\vec{x} + \vec{y}) = 2\vec{x} + 2\vec{y}$.

© *Jason Siefken, 2015–2022*

Italicizing N

44 The citizens of Oronto want to erect a sign welcoming visitors to the city. They've commissioned letters to be built, but at the last council meeting, they decided they wanted italicised letters instead of regular ones. Can you help them?



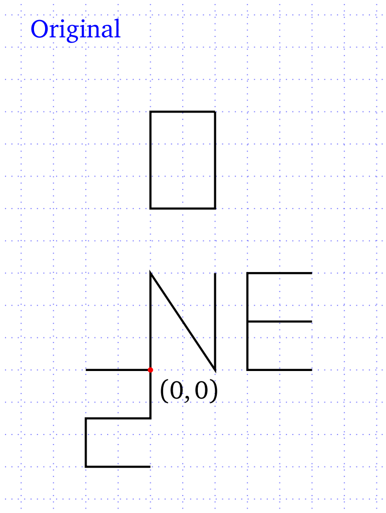
Suppose that the "N" on the left is written in regular 12-point font. Find a matrix A that will transform the "N" into the letter on the right which is written in an *italic* 16-point font.

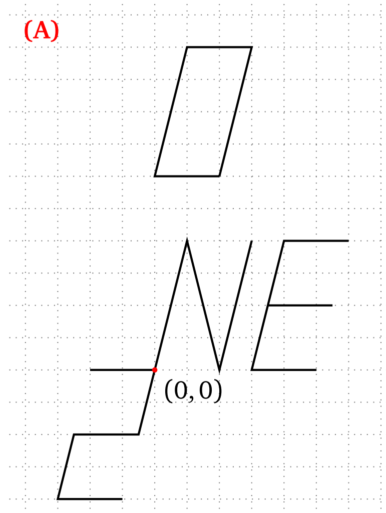
Work with your group to write out your solution and approach. Make a list of any assumptions you notice your group making or any questions for further pursuit.

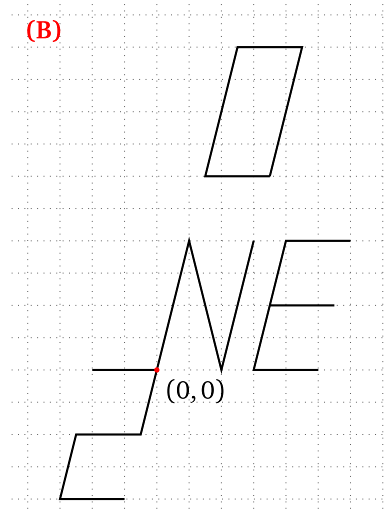
© IOLA Team iola.math.vt.edu

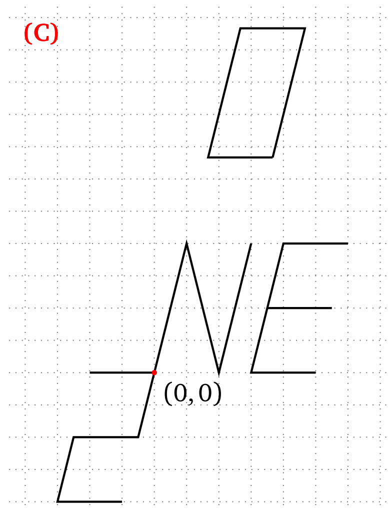
Beyond the N

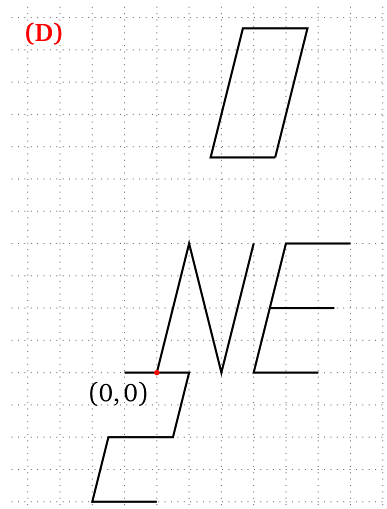
45 Some council members were wondering how letters placed in other locations in the plane would be transformed under $A = \begin{bmatrix} 1 & 1/3 \\ 0 & 4/3 \end{bmatrix}$. If other letters are placed around the "N," the council members argued over four different possible results for the transformed 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?











© IOLA Team iola.math.vt.edu

46 $\mathcal{R}:\mathbb{R}^2\to\mathbb{R}^2$ is the transformation that rotates vectors counter-clockwise by 90°.

46.1 Compute $\mathcal{R}\begin{bmatrix}1\\0\end{bmatrix}$ and $\mathcal{R}\begin{bmatrix}0\\1\end{bmatrix}$.

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

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

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

© Jason Siefken, 2015–2022

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 α.

47

47.1 Classify the following as linear transformations or not.

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

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

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

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

© Jason Siefken, 2015–2022

Image of a Set

DEF

Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a transformation and let $X \subseteq \mathbb{R}^n$ be a set. The ***image of the set X under L***, denoted $L(X)$, is the set

$$L(X) = \{\vec{y} \in \mathbb{R}^m \,:\, \vec{y} = L(\vec{x}) \text{ for some } \vec{x} \in X\}.$$

---

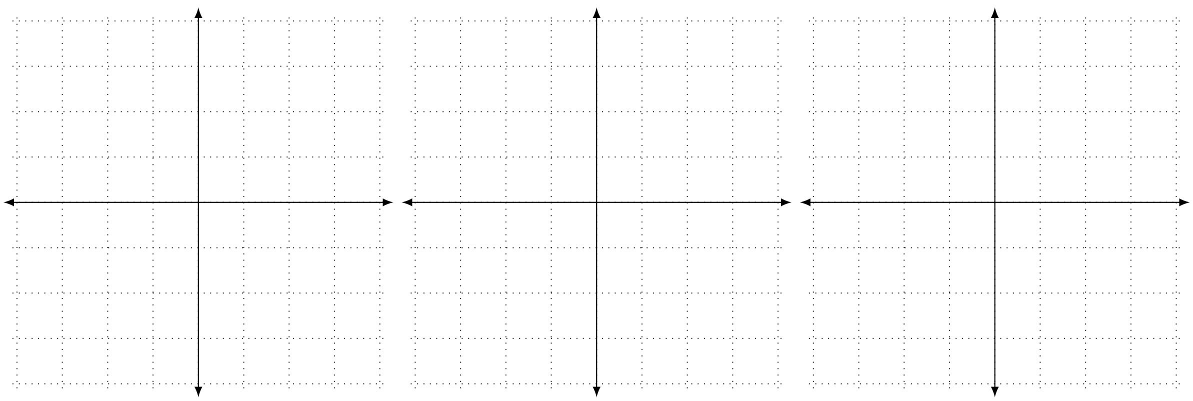
48 Let $S = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1 \right\} \subseteq \mathbb{R}^2$ be the filled-in unit square and let $C = \{\vec{0}, \vec{e}_1, \vec{e}_2, \vec{e}_1 + \vec{e}_2\} \subseteq \mathbb{R}^2$ be the corners of the unit square.

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

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

48.3 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?

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



© *Jason Siefken, 2015–2022*

© Jason Siefken, 2015–2022

The Composition of Linear Transformations

In this module you will learn

- How to break a complicated transformation into the composition of simpler ones.
- How the composition of linear transformations relates to matrix multiplication.

In life, we encounter situations where we do one thing after another. For example, you put on your socks and then your shoes. We might call this whole operation (of first putting on your socks and then your shoes) "getting dressed", and it is an example of *function composition*.

Composition of Functions. Let $f : A \to B$ and $g : B \to C$. The ***composition*** of g and f, notated $g \circ f$, is the function $h : A \to C$ defined by
$$h(x) = g \circ f(x) = g\big(f(x)\big).$$

We can formalize the shoes-socks example with mathematical notation.

$$\text{getting dressed} = (\text{putting on shoes}) \circ (\text{putting on socks})$$

Or, if X represented putting on socks, S represented putting on shoes, and D represented getting dressed, $D = S \circ X$.

This real-life example has utility when talking to children. Getting dressed is a complicated operation, but by breaking it up into simpler operations, even a young child can understand the process. In this vein, we can understand complicated linear transformations by breaking them up into the composition of simpler ones.

For example, define

$$\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^2 \qquad \text{by} \qquad \mathcal{T}(\vec{x}) = \begin{bmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \vec{x}.$$

It's hard to understand what $\mathcal{T}$ does just by looking at inputs and outputs. However, if we were keen enough to notice that
$$\mathcal{T} = \mathcal{S} \circ \mathcal{R}$$
where $\mathcal{R}$ was rotation counter-clockwise by $45°$ and $\mathcal{S}$ was a stretch in the $\vec{e}_1$ direction by a factor of 2, suddenly $\mathcal{T}$ wouldn't be such a mystery.

How does one figure out the "simple transformations" that when composed give the full transformation? In truth, there are many, many methods and there are whole books written about how to find these decompositions efficiently. We will encounter two algorithms for this,[36] but for now our methods will be ad hoc.

Example. Let $\mathcal{U} : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation given by $M = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & 0 \end{bmatrix}$, let $\mathcal{R} : \mathbb{R}^2 \to \mathbb{R}^2$ be rotation counter clockwise by $45°$, and let $\mathcal{P} : \mathbb{R}^2 \to \mathbb{R}^2$ be projection onto the x-axis. Write $\mathcal{U}$ as the composition (in some order) of $\mathcal{R}$ and $\mathcal{P}$.

We will use $\vec{e}_1$ and $\vec{e}_2$ to determine whether $\mathcal{U}$ is $\mathcal{R} \circ \mathcal{P}$ or $\mathcal{P} \circ \mathcal{R}$.
Computing,
$$\mathcal{U}(\vec{e}_1) = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix} \qquad \text{and} \qquad \mathcal{U}(\vec{e}_2) = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}.$$

For $\mathcal{R} \circ \mathcal{P}$:

$$\mathcal{R} \circ \mathcal{P}(\vec{e}_1) = \mathcal{R}(\mathcal{P}(\vec{e}_1)) = \mathcal{R}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$
$$\mathcal{R} \circ \mathcal{P}(\vec{e}_2) = \mathcal{R}(\mathcal{P}(\vec{e}_2)) = \mathcal{R}\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

[36] Phrased in terms of matrices instead of linear transformations, the decompositions we will study are called: (i) decomposition into *elementary* matrices, and (ii) *diagonalization*.

Module 10

© Jason Siefken, 2015–2022

For $\mathcal{P}\circ\mathcal{R}$:

$$\mathcal{P}\circ\mathcal{R}(\vec{e}_1)=\mathcal{P}(\mathcal{R}(\vec{e}_1))=\mathcal{P}\begin{bmatrix}\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\end{bmatrix}=\begin{bmatrix}\frac{\sqrt{2}}{2}\\0\end{bmatrix}$$
$$\mathcal{P}\circ\mathcal{R}(\vec{e}_2)=\mathcal{P}(\mathcal{R}(\vec{e}_2))=\mathcal{P}\begin{bmatrix}-\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\end{bmatrix}=\begin{bmatrix}-\frac{\sqrt{2}}{2}\\0\end{bmatrix}.$$

Since $\mathcal{P}\circ\mathcal{R}$ agrees with $\mathcal{U}$ on the standard basis (i.e., $\mathcal{P}\circ\mathcal{R}$ and $\mathcal{U}$ output the same vectors when $\vec{e}_1$ and $\vec{e}_2$ are input), they must agree for all vectors. Therefore $\mathcal{U}=\mathcal{P}\circ\mathcal{R}$.

Compositions and Matrix Products

Let $\mathcal{A}:\mathbb{R}^2\to\mathbb{R}^2$ and $\mathcal{B}:\mathbb{R}^2\to\mathbb{R}^2$ be matrix transformations with matrices

$$M_A=\begin{bmatrix}1&2\\0&2\end{bmatrix}\qquad\text{and}\qquad M_B=\begin{bmatrix}-1&-1\\-2&0\end{bmatrix}.$$

(Make sure you understand why $\mathcal{A}\neq M_A$ before continuing!)

Define $\mathcal{T}=\mathcal{A}\circ\mathcal{B}$. Since $\mathcal{T}:\mathbb{R}^2\to\mathbb{R}^2$ is a linear transformation, we know $\mathcal{T}$ has a matrix M_T which is 2×2. We can find M_T by the usual methods. First, compute some input-output pairs for $\mathcal{T}$.

$$\mathcal{T}\begin{bmatrix}1\\0\end{bmatrix}=\mathcal{A}\left(\mathcal{B}\begin{bmatrix}1\\0\end{bmatrix}\right)=\mathcal{A}\begin{bmatrix}-1\\-2\end{bmatrix}=\begin{bmatrix}-5\\-4\end{bmatrix}\qquad\text{and}\qquad\mathcal{T}\begin{bmatrix}0\\1\end{bmatrix}=\mathcal{A}\left(\mathcal{B}\begin{bmatrix}0\\1\end{bmatrix}\right)=\mathcal{A}\begin{bmatrix}-1\\0\end{bmatrix}=\begin{bmatrix}-1\\0\end{bmatrix}$$

Letting $M_T=\begin{bmatrix}a&b\\c&d\end{bmatrix}$, we use the input-output pairs to see

$$\begin{bmatrix}a\\c\end{bmatrix}=M_T\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}-5\\-4\end{bmatrix}\qquad\text{and}\qquad\begin{bmatrix}b\\d\end{bmatrix}=M_T\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}-1\\0\end{bmatrix},$$

and so

$$M_T=\begin{bmatrix}-5&-1\\-4&0\end{bmatrix}.$$

We found M_T, the matrix for $\mathcal{T}$, using traditional techniques, but could we have used M_A and M_B to somehow find M_T? As it turns out, yes, we could have!

By definition,

$$\mathcal{A}\vec{x}=M_A\vec{x}\qquad\text{and}\qquad\mathcal{B}\vec{x}=M_B\vec{x},$$

since $\mathcal{A}$ and $\mathcal{B}$ are matrix transformations. Therefore,

$$\mathcal{A}(\mathcal{B}\vec{x})=M_A(M_B\vec{x}).$$

But, matrix multiplication is *associative*,[37] and so

$$M_A(M_B\vec{x})=(M_AM_B)\vec{x}.$$

Thus M_AM_B must be a matrix for $\mathcal{A}\circ\mathcal{B}=\mathcal{T}$. Indeed, computing the matrix product, we see

$$M_AM_B=\begin{bmatrix}1&2\\0&2\end{bmatrix}\begin{bmatrix}-1&-1\\-2&0\end{bmatrix}=\begin{bmatrix}-5&-1\\-4&0\end{bmatrix}=M_T.$$

The fact that matrix multiplication corresponds to function composition is no coincidence. It is the very reason matrix multiplication is defined the way it is. This is reiterated in the following theorem.

Theorem. If $\mathcal{P}:\mathbb{R}^a\to\mathbb{R}^b$ and $\mathcal{Q}:\mathbb{R}^c\to\mathbb{R}^a$ are matrix transformations with matrices M_P and M_Q, then $\mathcal{P}\circ\mathcal{Q}$ is a matrix transformation whose matrix is given by the matrix product M_PM_Q.

It should now be clear why the order of matrix multiplication matters. The order of function composition matters (you must put on your socks before your shoes!), and since matrix multiplication corresponds to function composition, the order of matrix multiplication must matter.

[37] If an operation is associative, it means that where you put the parenthesis doesn't matter.

© Jason Siefken, 2015–2022

Practice Problems

1 (a) Let $\mathcal{U}:\mathbb{R}^2\to\mathbb{R}^2$ be the matrix transformation given by $\begin{bmatrix}0 & 0\\ -\sqrt{2}/2 & -\sqrt{2}/2\end{bmatrix}$. Further, let $\mathcal{P}:\mathbb{R}^2\to\mathbb{R}^2$ be the projection onto the y-axis, and let $\mathcal{R}:\mathbb{R}^2\to\mathbb{R}^2$ be the rotation clockwise by 135°.

i. Find a matrix for $\mathcal{R}\circ\mathcal{P}$.

ii. Find a matrix for $\mathcal{P}\circ\mathcal{R}$.

iii. Write $\mathcal{U}$ as the composition (in some order) of $\mathcal{R}$ and $\mathcal{P}$.

(b) Let $\mathcal{V}:\mathbb{R}^2\to\mathbb{R}^2$ be the matrix transformation given by $\begin{bmatrix}0 & 2\\ 2 & 0\end{bmatrix}$. Further, let $\mathcal{S}:\mathbb{R}^2\to\mathbb{R}^2$ be the transformation that doubles every vector, and let $\mathcal{F}:\mathbb{R}^2\to\mathbb{R}^2$ be the transformation reflecting over the line $y=x$.

i. Find a matrix for $\mathcal{F}\circ\mathcal{S}$.

ii. Find a matrix for $\mathcal{S}\circ\mathcal{F}$.

iii. Write $\mathcal{V}$ as the composition (in some order) of $\mathcal{F}$ and $\mathcal{S}$.

2 Let $\mathcal{A}:\mathbb{R}^2\to\mathbb{R}^2$ and $\mathcal{B}:\mathbb{R}^2\to\mathbb{R}^2$ be matrix transformations with matrices

$$M_{\mathcal{A}}=\begin{bmatrix}2 & 2\\ 1 & 3\end{bmatrix}\quad\text{and}\quad M_{\mathcal{B}}=\begin{bmatrix}3 & 2\\ 0 & 4\end{bmatrix}$$

and let $M_{\mathcal{T}}$ be the matrix for $\mathcal{T}=\mathcal{A}\circ\mathcal{B}$.

(a) Find $M_{\mathcal{T}}$ by computing input-output pairs for $\mathcal{T}$.

(b) Find $M_{\mathcal{T}}$ by using matrix multiplication applied to $M_{\mathcal{A}}$ and $M_{\mathcal{B}}$.

3 Let $\mathcal{A}:\mathbb{R}^3\to\mathbb{R}^2$ and $\mathcal{B}:\mathbb{R}^2\to\mathbb{R}^1$ be matrix transformations with matrices

$$M_{\mathcal{A}}=\begin{bmatrix}2 & 1 & 0\\ 2 & 3 & 0\end{bmatrix}\quad and\quad M_{\mathcal{B}}=\begin{bmatrix}3 & 2\end{bmatrix}$$

and let $M_{\mathcal{T}}$ be the matrix for $\mathcal{T}=\mathcal{B}\circ\mathcal{A}$.

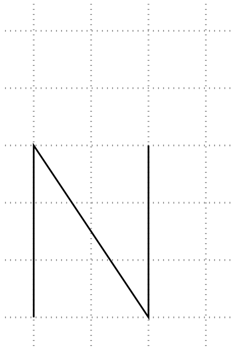
(a) Find $M_{\mathcal{T}}$ by computing input-output pairs for $\mathcal{T}$.

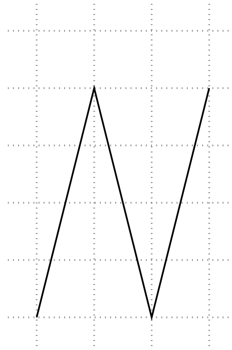
(b) Find $M_{\mathcal{T}}$ by using matrix multiplication applied to $M_{\mathcal{A}}$ and $M_{\mathcal{B}}$.

© Jason Siefken, 2015–2022

© Jason Siefken, 2015–2022

49





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 italicizing the N 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.

 © IOLA Team iola.math.vt.edu

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

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

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

50.3 Write down matrices A and B for $\mathcal{P} \circ \mathcal{R}$ and $\mathcal{R} \circ \mathcal{P}$.

50.4 How do the matrices A and B relate to the matrices P and R?

© Jason Siefken, 2015–2022

Range & Nullspace of a Linear Transformation

In this module you will learn

- The definition of the range and null space of a linear transformation.
- How to precisely notate the matrix for a linear transformation.
- The fundamental subspaces corresponding to a matrix (row space, column space, null space) and how they relate to the range and null space of a linear transformation.
- How to find a basis for the fundamental subspaces of a matrix.
- The definition of rank and the rank-nullity theorem.

Associated with every linear transformation are two specially named subspaces: the range and the null space.

Range

Range. The ***range*** (or ***image***) of a linear transformation $T : V \to W$ is the set of vectors that T can output. That is,

$$\mathrm{range}(T) = \{\vec{y} \in W \ : \ \vec{y} = T\vec{x} \text{ for some } \vec{x} \in V\}.$$

The range of a linear transformation has the exact same definition as the range of a function—it's the set of all outputs. In other words, the range of a linear transformation is the *image* of the entire domain with respect to that linear transformation.[38] However, unlike the range of an arbitrary function, the range of a linear transformation is always a subspace.

Theorem. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then $\mathrm{range}(T) \subseteq \mathbb{R}^m$ is a subspace.

Proof. Since $\mathrm{range}(T) = T(\mathbb{R}^n)$ and $\mathbb{R}^n$ is non-empty, we know that $\mathrm{range}(T)$ is non-empty. Therefore, to show that $\mathrm{range}(T)$ is a subspace, what remains to be shown is (i) that it's closed under vector addition, and (ii) that it is closed under scalar multiplication.

(i) Let $\vec{x}, \vec{y} \in \mathrm{range}(T)$. By definition, there exist $\vec{u}, \vec{v} \in \mathbb{R}^n$ such that $\vec{x} = T(\vec{u})$ and $\vec{y} = T(\vec{v})$. Since T is linear,

$$\vec{x} + \vec{y} = T(\vec{u}) + T(\vec{v}) = T(\vec{u} + \vec{v}),$$

and so $\vec{x} + \vec{y} \in \mathrm{range}(T)$.

(ii) Let $\vec{x} \in \mathrm{range}(T)$ and let α be a scalar. By definition, there exists $\vec{u} \in \mathbb{R}^n$ such that $\vec{x} = T(\vec{u})$, and so by the linearity of T,

$$\alpha\vec{x} = \alpha T(\vec{u}) = T(\alpha\vec{u}).$$

Therefore $\alpha\vec{x} \in \mathrm{range}(T)$.

■

When analyzing subspaces, we are often interested in how big they are. That information is captured by a number—the *dimension* of the subspace. For transformations, we also have a notion of how "big" they are, which is captured in a number called the *rank*.

Rank of a Linear Transformation. For a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, the ***rank*** of T, denoted $\mathrm{rank}(T)$, is the dimension of the range of T.

The rank of a linear transformation can be used to measure its complexity or compressibility. A rank 0 transformation must send all vectors to $\vec{0}$. A rank 1 transformation must send all vectors to a line, etc.. So, by knowing just a single number—the rank—you can judge how complicated the set of outputs of a linear transformation will be.

Example. Let $\mathcal{P}$ be the plane given by $x + y + z = 0$, and let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be projection onto $\mathcal{P}$. Find $\mathrm{range}(T)$ and $\mathrm{rank}(T)$.

[38] Some people say "the *image* of T" as a short way of saying "the image of the entire domain of T under T". Used in this sense $\mathrm{Image}(T) = \mathrm{range}(T)$.

© Jason Siefken, 2015–2022

First we will find range(T). Since T is a projection onto $\mathcal{P}$, we know range(T) $\subseteq \mathcal{P}$. Because $T(\vec{p}) = \vec{p}$ for all $\vec{p} \in \mathcal{P}$, we know $\mathcal{P} \subseteq$ range(T), and so

$$\mathrm{range}(T) = \mathcal{P}.$$

Since $\mathcal{P}$ is a plane, we know $\dim(\mathcal{P}) = 2 = \dim(\mathrm{range}(T)) = \mathrm{rank}(T)$.

Null Space

The second special subspace is called the *null space*.

Null Space. The ***null space*** (or ***kernel***) of a linear transformation $T : V \to W$ is the set of vectors that get mapped to the zero vector under T. That is,

$$\mathrm{null}(T) = \{\vec{x} \in V \ : \ T\vec{x} = \vec{0}\}.$$

We've seen null spaces before. In the context of matrices when we asked questions like, "Are these column vectors linearly independent?" Now that we understand linear transformation and subspaces, we can consider this question anew.

Just like the range of a linear transformation, the null space of a linear transformation is always a subspace.

Theorem. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then null(T) $\subseteq \mathbb{R}^n$ is a subspace.

Proof. Since T is linear, $T(\vec{0}) = \vec{0}$ and so $\vec{0} \in$ null(T) which shows that null(T) is non-empty. Therefore, to show that null(T) is a subspace, we only need to show (i) that it's closed under vector addition, and (ii) that it is closed under scalar multiplication.

(i) Let $\vec{x}, \vec{y} \in$ null(T). By definition, $T(\vec{x}) = T(\vec{y}) = \vec{0}$. By linearity we see

$$T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y}) = \vec{0} + \vec{0} = \vec{0},$$

and so $\vec{x} + \vec{y} \in$ null(T).

(ii) Let $\vec{x} \in$ null(T) and let α be a scalar. By definition, $T(\vec{x}) = \vec{0}$, and so by the linearity of T,

$$T(\alpha\vec{x}) = \alpha T(\vec{x}) = \alpha\vec{0} = \vec{0}.$$

Therefore $\alpha\vec{x} \in$ null(T).

■

Akin to the rank–range connection, there is a special number called the *nullity* which specifies the dimension of the null space.

Nullity. For a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, the ***nullity*** of T, denoted nullity(T), is the dimension of the null space of T.

Example. Let $\mathcal{P}$ be the plane given by $x + y + z = 0$, and let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be projection onto $\mathcal{P}$. Find null(T) and nullity(T).

First we will find null(T). Since T is a projection onto $\mathcal{P}$ (and because $\mathcal{P}$ passes through $\vec{0}$), we know every normal vector for $\mathcal{P}$ will get sent to $\vec{0}$ when T is applied. And, besides $\vec{0}$ itself, these are the only vectors that get sent to $\vec{0}$. Therefore

$$\mathrm{null}(T) = \{\text{normal vectors}\} \cup \{\vec{0}\} = \mathrm{span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}\right\}.$$

Since null(T) is a line, we know nullity(T) = 1.

Fundamental Subspaces of a Matrix

Every linear transformation has a range and a null space. Analogously, every matrix is associated with three fundamental subspaces.

© Jason Siefken, 2015–2022

Fundamental Subspaces. Associated with any matrix M are three fundamental subspaces: the ***row space*** of M, denoted $\mathrm{row}(M)$, is the span of the rows of M; the ***column space*** of M, denoted $\mathrm{col}(M)$, is the span of the columns of M; and the ***null space*** of M, denoted $\mathrm{null}(M)$, is the set of solutions to $M\vec{x} = \vec{0}$.

Computationally, it's much easier to find the row space/column space/null space of a matrix than it is to find the range/null space of a linear transformation because we can turn matrix questions into systems of linear equations.

Example. Find the null space of $M = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix}$.

To find the null space of M, we need to solve the homogeneous matrix equation $M\vec{x} = \vec{0}$. Row reducing, we see

$$\mathrm{rref}(M) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix},$$

and so the z column is a free variable column. Therefore, the complete solution can be expressed in vector form as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix},$$

and so

$$\mathrm{null}(M) = \mathrm{span}\left\{ \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} \right\}.$$

The column space and row space are just as easy to compute, since it just involves picking a basis from the existing row or column vectors.

Example. Let $M = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix}$. Find a basis for the row space and the column space of M.

First the column space. We need to pick a basis for $\mathrm{span}\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 5 \\ -2 \end{bmatrix} \right\}$, which is the same thing as picking a maximal linearly independent subset of $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 5 \\ -2 \end{bmatrix} \right\}$.
Putting these vectors as columns in a matrix and row reducing, we see

$$\mathrm{rref}\left(\begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix} \right) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

The first and second columns are the only pivot columns and so the first and second ***original*** vectors form a maximal linearly independent subset. Thus,

$$\mathrm{col}(M) = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix} \right\} = \mathbb{R}^2 \qquad \text{and a basis is} \qquad \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix} \right\}.$$

To find the row space, we need to pick a basis for $\mathrm{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ -2 \end{bmatrix} \right\}$. Repeating a similar procedure, we see

$$\mathrm{rref}\left(\begin{bmatrix} 1 & 2 \\ 2 & -2 \\ 5 & -2 \end{bmatrix} \right) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

and so $\left\{ \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ -2 \end{bmatrix} \right\}$ is linearly independent. Therefore

$$\mathrm{row}(M) = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ -2 \end{bmatrix} \right\} \qquad \text{and a basis is} \qquad \left\{ \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ -2 \end{bmatrix} \right\}.$$

© Jason Siefken, 2015–2022

When talking about fundamental subspaces, we often switch between talking about column vectors and row vectors belonging to a matrix. The operation of swapping rows for columns is called the *transpose*.

Transpose.
Let M be an $n \times m$ matrix defined by

$$M = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nm} \end{bmatrix}.$$

The ***transpose*** of M, notated M^T, is the $m \times n$ matrix produced by swapping the rows and columns of M. That is

$$M^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ a_{13} & a_{23} & \cdots & a_{n3} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{bmatrix}.$$

Using the transpose, we can make statements like

$$\mathrm{col}(M) = \mathrm{row}(M^T) \qquad \text{and} \qquad \mathrm{row}(M) = \mathrm{col}(M^T).$$

In addition, it helps us state the following theorem.

Theorem (Row-Col Dimension). For a matrix A, the dimension of the row space equals the dimension of the column space.

Proof. For this proof, we will rely on what we know about the row reduction algorithm and what the reduced row echelon form of a matrix tells us.

Claim 1: $\mathrm{row}(\mathrm{rref}(A)) \subseteq \mathrm{row}(A)$. To see this, observe that to get $\mathrm{rref}(A)$, we take linear combinations of the rows of A. Therefore, it must be that the span of the rows of $\mathrm{rref}(A)$ is contained in the span of the rows of A.

Claim 2: $\mathrm{row}(\mathrm{rref}(A)) = \mathrm{row}(A)$. To see this, observe that every elementary row operation is reversible. Therefore every row in A can be obtained as a linear combination of rows in $\mathrm{rref}(A)$ (by just reversing the steps). Thus the row vectors of $\mathrm{rref}(A)$ and the row vectors of A must have the same span.

Claim 3: The non-zero rows of $\mathrm{rref}(A)$ form a basis for $\mathrm{row}(A)$. We already know that the non-zero rows of $\mathrm{rref}(A)$ span $\mathrm{row}(A)$, so we only need to argue that they are linearly independent. However, this follows immediately from the fact that $\mathrm{rref}(A)$ is in reduced row echelon form. Above and below every pivot in $\mathrm{rref}(A)$ are zeros. Therefore, a row in $\mathrm{rref}(A)$ with a pivot cannot be written as a linear combination of any other row. Since every non-zero row has a pivot, this proves the claim.

Now, note the following two facts.

1. The columns of A corresponding to pivot columns of $\mathrm{rref}(A)$ form a basis for $\mathrm{col}(A)$.

2. The non-zero rows of $\mathrm{rref}(A)$ form a basis for $\mathrm{row}(A)$.

To complete the proof, note that every pivot of $\mathrm{rref}(A)$ lies in exactly one row and one column. Therefore, the number of basis vectors in $\mathrm{row}(A)$ is the same as the number of basis vectors in $\mathrm{col}(A)$. ∎

Equations, Null Spaces, and Geometry

Let $M = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix}$. Using the typical row-reduction steps, we know that the complete solution to $M\vec{x} = \vec{0}$ (i.e., the null space of M) can be expressed in vector form as

$$\vec{x} = t \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}.$$

© Jason Siefken, 2015–2022

Similarly, the complete solution to $M\vec{x}=\vec{b}$ where $\vec{b}=\begin{bmatrix}1\\2\end{bmatrix}$ can be expressed in vector form as

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

The set of solutions to $M\vec{x}=\vec{0}$ and $M\vec{x}=\vec{b}$ look very similar. In fact,

$$\{\text{solutions to } M\vec{x}=\vec{b}\}=\{\text{solutions to } M\vec{x}=\vec{0}\}+\{\vec{p}\}\qquad\text{where}\qquad \vec{p}=\begin{bmatrix}1\\0\\0\end{bmatrix}.$$

Or, phrased another way, the solution set to $M\vec{x}=\vec{b}$ is

$$\operatorname{null}(M)+\{\vec{p}\}.$$

In the context of what we already know about lines and translated spans, this makes perfect sense. We know that the solution set to $M\vec{x}=\vec{b}$ is a line (which doesn't pass through the origin) and may therefore be written as a translated span $\operatorname{span}\{\vec{d}\}+\{\vec{p}\}$. Here $\vec{d}$ is a direction vector for the line and $\vec{p}$ is a point on the line.

Because $\vec{p}\in\operatorname{span}\{\vec{d}\}+\{\vec{p}\}$, we call $\vec{p}$ a *particular solution* to $M\vec{x}=\vec{b}$. Using a similar argument, we can show that for any matrix A, and any vector $\vec{b}$, the set of all solutions to $A\vec{x}=\vec{b}$ (provided there are any) can be expressed as

$$V+\{\vec{p}\}$$

where V is a subspace and $\vec{p}$ is a particular solution. In fact, we can do better. We can say $V=\operatorname{null}(A)$.

Theorem. Let A be a matrix, $\vec{b}$ be a vector, and let $\vec{p}$ be a particular solution to $A\vec{x}=\vec{b}$. Then, the set of all solutions to $A\vec{x}=\vec{b}$ is

$$\operatorname{null}(A)+\{\vec{p}\}.$$

Proof. Let $S=\{\text{all solutions to } A\vec{x}=\vec{b}\}$ and assume $\vec{p}\in S$. We will show $S=\operatorname{null}(A)+\{\vec{p}\}$.

First we will show $\operatorname{null}(A)+\{\vec{p}\}\subseteq S$. Let $\vec{v}\in\operatorname{null}(A)+\{\vec{p}\}$. By definition, $\vec{v}=\vec{n}+\vec{p}$ for some $\vec{n}\in\operatorname{null}(A)$. Now, by linearity of matrix multiplication and the definition of the null space,

$$A\vec{v}=A(\vec{n}+\vec{p})=A\vec{n}+A\vec{p}=\vec{0}+\vec{b}=\vec{b},$$

and so $\vec{v}\in S$.

Next we will show $S\subseteq\operatorname{null}(A)+\{\vec{p}\}$. First observe that for any $\vec{u},\vec{v}\in S$ we have

$$A(\vec{u}-\vec{v})=A\vec{u}-A\vec{v}=\vec{b}-\vec{b}=\vec{0},$$

and so $\vec{u}-\vec{v}\in\operatorname{null}(A)$.

Fix $\vec{w}\in S$. By our previous observation, $\vec{w}-\vec{p}\in\operatorname{null}(A)$. Therefore

$$\vec{w}=(\vec{w}-\vec{p})+\vec{p}\in\operatorname{null}(A)+\{\vec{p}\},$$

which completes the proof. ■

Takeaway. To write the complete solution to $A\vec{x}=\vec{b}$, all you need is the null space of A and a particular solution to $A\vec{x}=\vec{b}$.

Null spaces are also closely connected with row spaces. Let $\mathcal{P}\subseteq\mathbb{R}^3$ be the plane with equation $x+2y+2z=0$. We can rewrite this equation as a matrix equation and as the equation of a plane in normal form.

$$\underbrace{\begin{bmatrix}1&2&2\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\vec{0}}_{\text{a matrix equation}}\qquad\qquad\underbrace{\begin{bmatrix}1\\2\\2\end{bmatrix}\cdot\begin{bmatrix}x\\y\\z\end{bmatrix}=0}_{\text{normal form}}$$

© Jason Siefken, 2015–2022

Now we see that $\mathcal{P} = \mathrm{null}\left(\begin{bmatrix}1 & 2 & 2\end{bmatrix}\right)$ and that every non-zero vector in $\mathrm{row}\left(\begin{bmatrix}1 & 2 & 2\end{bmatrix}\right)$ is a normal vector for $\mathcal{P}$. In other words, $\mathrm{null}\left(\begin{bmatrix}1 & 2 & 2\end{bmatrix}\right)$ is orthogonal to $\mathrm{row}\left(\begin{bmatrix}1 & 2 & 2\end{bmatrix}\right)$.

This is no coincidence. Let M be a matrix and let $\vec{r}_1, \dots, \vec{r}_n$ be the rows of M. By definition,

$$M\vec{x} = \begin{bmatrix}\vec{r}_1 \cdot \vec{x} \\ \vdots \\ \vec{r}_n \cdot \vec{x}\end{bmatrix},$$

and so solutions to $M\vec{x} = \vec{0}$ are precisely the vectors which are orthogonal to every row of M. In other words, $\mathrm{null}(M)$ consists of all vectors orthogonal to the rows of M. Conversely, $\mathrm{row}(M)$ consists of all vectors orthogonal to everything in $\mathrm{null}(M)$. We can use this fact to approach questions in a new way.

Example. Let $\vec{a} = \begin{bmatrix}1 \\ 2 \\ 5\end{bmatrix}$ and $\vec{b} = \begin{bmatrix}2 \\ -2 \\ -2\end{bmatrix}$. Find the set of all vectors orthogonal to both $\vec{a}$ and $\vec{b}$.

Let $M = \begin{bmatrix}1 & 2 & 5 \\ 2 & -2 & -2\end{bmatrix}$ be the matrix whose rows are $\vec{a}$ and $\vec{b}$. Since $\mathrm{null}(M)$ consists of all vectors orthogonal to $\mathrm{row}(M)$, the set we are looking for is $\mathrm{null}(M)$. Computing via row reduction, we find

$$\mathrm{null}(M) = \mathrm{span}\left\{\begin{bmatrix}-1 \\ -2 \\ 1\end{bmatrix}\right\}.$$

Transformations and Matrices

Matrices are connected to systems of linear equations via matrix equations (like $A\vec{x} = \vec{b}$) and to linear transformations through matrix transformations (like $\mathcal{T}(\vec{x}) = M\vec{x}$). This means that we can think about systems of equations in terms of linear transformations and we can gain insight about linear transformations by looking at systems of equations!

In preparation for this, let's reconsider matrix transformations and be pedantic about our notation.

Let $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let M be its corresponding matrix. $\mathcal{T}$ is a function that inputs and outputs *vectors*. M is a box of numbers, which has no meaning by itself, but we know how to multiply M by lists of numbers (or other boxes of numbers). Therefore, strictly speaking, the expression "$M\vec{x}$" doesn't make sense. The quantity "$\vec{x}$" is a vector, but we only know how to multiply M by lists of numbers.

Ah! But we know how to turn $\vec{x}$ into a list of numbers. Just pick a basis! The expression

$$M[\vec{x}]_{\mathcal{E}}$$

makes perfect sense since $[\vec{x}]_{\mathcal{E}}$ is a list of numbers. Continuing to be pedantic, we know $\mathcal{T}(\vec{x}) \neq M[\vec{x}]_{\mathcal{E}}$ since the left side is a vector and the right side is a list of numbers. We can fix this by either turning the right side into a vector or the left side into a list of numbers. Doing this, we see the precise relationship between a linear transformation $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ and its matrix M is

$$[\mathcal{T}(\vec{x})]_{\mathcal{E}} = M[\vec{x}]_{\mathcal{E}}.$$

If we have a matrix M, by picking a basis (usually the standard basis), we can define a linear transformation by first taking the input vector and rewriting it in the basis, next multiplying by the matrix, and finally taking the list of numbers and using them as coefficients for a linear combination involving the basis vectors. This is what we actually mean when we say that a matrix *induces* a linear transformation.

Induced Transformation.
Let M be an $n \times m$ matrix. We say M ***induces*** a linear transformation $\mathcal{T}_M : \mathbb{R}^m \to \mathbb{R}^n$ defined by

$$[\mathcal{T}_M \vec{v}]_{\mathcal{E}'} = M[\vec{v}]_{\mathcal{E}},$$

where $\mathcal{E}$ is the standard basis for $\mathbb{R}^m$ and $\mathcal{E}'$ is the standard basis for $\mathbb{R}^n$.

Previously, we would write "$\mathcal{T}(\vec{x}) = M\vec{x}$" which hides the fact that when we relate a matrix and a linear transformation, there is a basis hidden in the background. And, like before, when we're only considering a single basis, we can be sloppy with our notation and write things like "$M\vec{x}$", but when there are multiple bases or when

© Jason Siefken, 2015–2022

we're trying to be extra precise, we must make sure our boxes/lists of numbers and our transformations/vectors stay separate.

Example. Let $\mathcal{T}$ be the transformation induced by the matrix $M = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix}$, and let $\vec{v} = 3\vec{e}_1 - 3\vec{e}_3$. Compute $\mathcal{T}(\vec{v})$.

Since $\mathcal{T}$ is induced by $M = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix}$, by definition,

$$[\mathcal{T}_M \vec{v}]_{\mathcal{E}'} = M[\vec{v}]_{\mathcal{E}} = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix}[\vec{v}]_{\mathcal{E}}.$$

Further, since $\vec{v} = 3\vec{e}_1 - 3\vec{e}_3$, by definition we have $[\vec{v}]_{\mathcal{E}} = \begin{bmatrix} 3 \\ 0 \\ -3 \end{bmatrix}$. Therefore,

$$[\mathcal{T}_M \vec{v}]_{\mathcal{E}'} = \begin{bmatrix} 1 & 2 & 5 \\ 2 & -2 & -2 \end{bmatrix}\begin{bmatrix} 3 \\ 0 \\ -3 \end{bmatrix} = \begin{bmatrix} -12 \\ 12 \end{bmatrix}.$$

In other words, $\mathcal{T}(\vec{v}) = \begin{bmatrix} -12 \\ 12 \end{bmatrix}_{\mathcal{E}'} = -12\vec{e}_1 + 12\vec{e}_2$.

Using induced transformations, we can extend linear-transformation definitions to matrix definitions. In particular, we can define the rank and nullity of a matrix.

Rank of a Matrix. Let M be a matrix. The ***rank*** of M, denoted rank(M), is the rank of the linear transformation induced by M.

Nullity of a Matrix. Let M be a matrix. The ***nullity*** of M, denoted nullity(M), is the nullity of the linear transformation induced by M.

Range vs. Column Space & Null Space vs. Null Space

Let $M = \begin{bmatrix} C_1 & C_2 & \cdots & C_m \end{bmatrix}$ be an $m \times m$ matrix with columns $C_1, \ldots, C_m$, and let $\mathcal{T}$ be the transformation induced by M. The column space of M is the set of all linear combinations of the columns of M. But, let's be precise. The columns of M are lists of numbers, so to talk about the column space of M, we need to turn them into vectors. Fortunately, we have a nice notation for that. Since C_i is a list of numbers, $[C_i]_{\mathcal{E}}$ is a (true) vector, and

$$\mathrm{col}(M) = \mathrm{span}\{[C_1]_{\mathcal{E}}, [C_2]_{\mathcal{E}}, \ldots, [C_m]_{\mathcal{E}}\}.$$

Can we connect this to the range of $\mathcal{T}$? Well, by the definition of matrix multiplication, we know that

$$M\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = M[\vec{e}_1]_{\mathcal{E}} = C_1$$

and in general $M[\vec{e}_i]_{\mathcal{E}} = C_i$. By the definition of induced transformation, we know

$$[\mathcal{T}(\vec{e}_i)]_{\mathcal{E}} = M[\vec{e}_i]_{\mathcal{E}} = C_i,$$

and so

$$\mathcal{T}(\vec{e}_i) = [C_i]_{\mathcal{E}}.$$

Every input to $\mathcal{T}$ can be written as a linear combination of $\vec{e}_i$'s (because $\mathcal{E}$ is a basis) and so, because $\mathcal{T}$ is linear, every output of $\mathcal{T}$ can be written as a linear combination of $[C_i]_{\mathcal{E}}$'s. In other words,

$$\mathrm{range}(\mathcal{T}) = \mathrm{col}(M).$$

This means that when trying to answer a question about the range of a linear transformation, we could think about the column space of its matrix instead (or vice versa).

© Jason Siefken, 2015–2022

Example. Let $\mathcal{T} : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by

$$\mathcal{T}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x - z \\ 4x - 2z \end{bmatrix}.$$

Find range$(\mathcal{T})$ and rank$(\mathcal{T})$.

Let M be a matrix for $\mathcal{T}$. We know range$(\mathcal{T}) = $ col(M) and rank$(\mathcal{T}) = \dim(\text{range}(\mathcal{T})) = \dim(\text{col}(M))$. By inspection, we see that

$$M = \begin{bmatrix} 2 & 0 & -1 \\ 4 & 0 & -2 \end{bmatrix}.$$

Again, by inspection, we see that $\left\{\begin{bmatrix} 2 \\ 4 \end{bmatrix}\right\}$ is a basis for col(M) and col(M) is one dimensional. Therefore,

$$\text{range}(\mathcal{T}) = \text{span}\left\{\begin{bmatrix} 2 \\ 4 \end{bmatrix}\right\} \qquad \text{and} \qquad \text{rank}(\mathcal{T}) = 1.$$

There is an alternative definition of the rank of a matrix which commonly appears. We'll state it as a theorem.

Theorem. Let M be a matrix. The rank of M is equal to the number of pivots in rref(M).

Proof. We know that rank$(M) = \dim(\text{range}(\mathcal{T}_M)) = \dim(\text{col}(M))$ where $\mathcal{T}_M$ is the transformation induced by M. Further, a basis for col(M) consists of a maximal linearly independent subset of the columns of M. To find such a subset, we row reduce M and look at the columns of M that correspond to pivot columns of rref(M).

When all is said and done, the number of elements in a basis for col(M) will be the number of pivots in rref(M), which is the same as rank(M). ■

Takeaway. If $\mathcal{T}$ is a linear transformation and M is a corresponding matrix, range$(\mathcal{T}) = $ col(M), and answering questions about M answers questions about $\mathcal{T}$.

Just like the range–column-space relationship, we also have a null-space–null-space relationship. More specifically, if $\mathcal{T}$ is a linear transformation with matrix M, then null$(\mathcal{T}) = $ null(M). From this fact, we deduce the following theorem.

Theorem. Let $\mathcal{T}$ be a linear transformation and let M be a matrix for $\mathcal{T}$. Then nullity$(\mathcal{T})$ is equal to the number of free variable columns in rref(M).

Proof. We know nullity$(\mathcal{T}) = \dim(\text{null}(\mathcal{T})) = \dim(\text{null}(M))$. Further, we know that the complete solution to $M\vec{x} = \vec{0}$ will take the form

$$\vec{x} = t_1\vec{d}_1 + \cdots + t_k\vec{d}_k$$

where k is the number of free variable columns in rref(M). The algorithm for writing the complete solution to $M\vec{x} = \vec{0}$ ensures that $\{\vec{d}_1, \dots, \vec{d}_k\}$ is a basis for null(M), and so nullity$(\mathcal{T}) = k$, which completes the proof. ■

Combining these facts, we can reformulate the Row-Col Dimension theorem as a theorem about ranks.

Theorem. For a matrix A, we have rank$(A) = $ rank(A^T).

Proof. By the Row-Col Dimension theorem, we know $\dim(\text{col}(A)) = \dim(\text{row}(A))$. By the definition of the transpose, we know row$(A) = $ col(A^T). Therefore,

$$\dim(\text{col}(A)) = \dim(\text{col}(A^T)),$$

which is another way of saying rank$(A) = $ rank(A^T). ■

The Rank-Nullity Theorem

The rank and the nullity of a linear transformation/matrix are connected by a powerful theorem.

Theorem (Rank-nullity Theorem for Matrices). For a matrix A,

$$\text{rank}(A) + \text{nullity}(A) = \# \text{ of columns in } A.$$

© Jason Siefken, 2015–2022

The rank-nullity theorem's statement is simple, but it is surprisingly useful. For example, consider the question: how many normal directions does a plane have in $\mathbb{R}^3$ or in $\mathbb{R}^4$ or in $\mathbb{R}^5$?

We already know the answer in $\mathbb{R}^3$: a plane has a two-dimensional set of direction vectors and a one-dimensional set (a line) of normal vectors.[39] But, we can verify this fact using the rank-nullity theorem.

Let

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

and let $\mathcal{P} = \mathrm{null}(M)$. We know $\mathcal{P}$ is a plane with equation $x + 2y + 2z = 0$ and therefore is two dimensional. Further, non-zero vectors in the row space of M are normal vectors for $\mathcal{P}$. Since $\mathrm{null}(M)$ is two-dimensional and M has three columns, the rank-nullity theorem tells us that $\mathrm{rank}(M) = 1$. Therefore $\dim(\mathrm{col}(M)) = \dim(\mathrm{row}(M)) = 1$. We conclude the set of normal vectors to $\mathcal{P} = \mathrm{null}(M)$ is a line (if we include $\vec{0}$).

By contrast, let $\mathcal{Q} \subseteq \mathbb{R}^4$ be the plane in $\mathbb{R}^4$ given in vector form by

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

How many normal vectors does $\mathcal{Q}$ have? Well, the matrix $A = \begin{bmatrix}1 & 2 & 2 & 2\\ -1 & 1 & -1 & 1\end{bmatrix}$ has rank 2 and therefore has nullity 2. This means there exist two linearly independent normal directions for $\mathcal{Q}$.[40]

There is an equivalent rank-nullity theorem for linear transformations.

Theorem (Rank-nullity Theorem for Linear Transformations). Let $\mathcal{T}$ be a linear transformation. Then

$$\mathrm{rank}(\mathcal{T}) + \mathrm{nullity}(\mathcal{T}) = \dim(\text{domain of } \mathcal{T}).$$

Just like the rank-nullity theorem for matrices, the rank-nullity theorem for linear transformations can give insights about linear transformations that would be otherwise hard to see.

Practice Problems

1 For the following matrices, find their null space, column space, and row space.

(a) $M_1 = \begin{bmatrix}1 & 2 & 1\\ 3 & 1 & -2\\ 8 & 6 & -2\end{bmatrix}$.

(b) $M_2 = \begin{bmatrix}0 & 2 & 1\\ 3 & 2 & 5\end{bmatrix}$.

(c) $M_3 = \begin{bmatrix}1 & 2\\ 3 & 1\\ 4 & 0\end{bmatrix}$.

(d) $M_4 = \begin{bmatrix}1 & -2 & 0 & -1\\ 3 & 5 & -1 & 0\\ 2 & 3 & -2 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$.

2 Let $\mathcal{P}$ be the plane given by $3x + 4y + 5z = 0$, and let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be projection onto $\mathcal{P}$.

(a) Find $\mathrm{range}(T)$ and $\mathrm{rank}(T)$.

(b) Find $\mathrm{null}(T)$ and $\mathrm{nullity}(T)$.

3 Find the range and null space of the following linear transformations.

(a) $\mathcal{P} : \mathbb{R}^2 \to \mathbb{R}^2$, where $\mathcal{P}$ is projection on to the line $y = x$.

(b) Let $\theta \in \mathbb{R}$ and let $\mathcal{R} : \mathbb{R}^2 \to \mathbb{R}^2$ to be the transformation which rotates all vectors by counter-clockwise by θ radians.

(c) $\mathcal{F} : \mathbb{R}^2 \to \mathbb{R}^2$, where $\mathcal{F}$ reflects over the x-axis.

(d) $\mathcal{M} : \mathbb{R}^3 \to \mathbb{R}^3$ where $\mathcal{M}$ is the matrix transformation given by $\begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{bmatrix}$.

(e) $\mathcal{Q} : \mathbb{R}^3 \to \mathbb{R}^1$ defined by $\mathcal{Q}\begin{bmatrix}x\\y\\z\end{bmatrix} = x + z$.

4 (a) Let $\mathcal{T}$ be the transformation induced by the matrix $\begin{bmatrix}7 & 5\\ -2 & -2\end{bmatrix}$, and $\vec{v} = 3\vec{e}_1 - 3\vec{e}_2$. Compute $\mathcal{T}\vec{v}$ and $[\mathcal{T}\vec{v}]_{\mathcal{E}}$.

(b) Let $\mathcal{T}$ be the transformation induced by the matrix $\begin{bmatrix}3 & 7 & 5\\ 1 & -2 & -2\end{bmatrix}$, and $\vec{v} = 2\vec{e}_1 + 0\vec{e}_2 + 4\vec{e}_3$. Compute $\mathcal{T}\vec{v}$ and $[\mathcal{T}\vec{v}]_{\mathcal{E}}$.

5 For each statement below, determine whether it is true or false. Justify your answer.

(a) Let A be an arbitrary matrix. Then $\mathrm{col}(A) = \mathrm{col}(A^T)$.

(b) Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a transformation (not necessarily linear). If $\mathrm{null}(T) = \{\vec{x} \in \mathbb{R}^m : T(\vec{x}) = \vec{0}\}$ is a subspace, then T is linear.

[39] Technically, the set of normal vectors for a plane in $\mathbb{R}^3$ is a line without $\vec{0}$, since $\vec{0}$ is never considered a normal vector.

[40] Notice that we are using the complementary argument to the example in $\mathbb{R}^3$. For this example, the plane is the *row space* and the set of normal vectors is the null space (ignoring $\vec{0}$, of course).

© *Jason Siefken, 2015–2022*

(c) Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation. Then $\text{nullity}(T) \geq n$.

(d) Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation induced by a matrix M. If $\text{rank}(T) = n$, then $\text{nullity}(M) = 0$.

6 Give an example of a 3×4 matrix M with the specified rank, or explain why one cannot exist.

(a) $\text{rank}(M) = 0$

(b) $\text{rank}(M) = 1$

(c) $\text{rank}(M) = 3$

(d) $\text{rank}(M) = 4$

7 Let $\mathcal{P} : \mathbb{R}^3 \to \mathbb{R}^3$ be the projection onto the xy-plane.

(a) Find a matrix $M_{\mathcal{P}}$ for the transformation.

(b) Find the range of $\mathcal{P}$.

(c) Find the column space of $M_{\mathcal{P}}$. Are there any similarities to your answer in the previous part?

(d) Find the null space of $\mathcal{P}$ and $M_{\mathcal{P}}$. Are there similarities between the null space of a linear transformation and its associated matrix?

© Jason Siefken, 2015–2022

DEF

Range

The ***range*** (or ***image***) of a linear transformation $T : V \to W$ is the set of vectors that T can output. That is,

$$\text{range}(T) = \{\vec{y} \in W \ : \ \vec{y} = T\vec{x} \text{ for some } \vec{x} \in V\}.$$

DEFINITION

Null Space

The ***null space*** (or ***kernel***) of a linear transformation $T : V \to W$ is the set of vectors that get mapped to the zero vector under T. That is,

$$\text{null}(T) = \{\vec{x} \in V \ : \ T\vec{x} = \vec{0}\}.$$

---

51 Let $\mathcal{P} : \mathbb{R}^2 \to \mathbb{R}^2$ be projection onto $\text{span}\{\vec{u}\}$ where $\vec{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ (like before).

51.1 What is the range of $\mathcal{P}$?

51.2 What is the null space of $\mathcal{P}$?

© Jason Siefken, 2015–2022

52 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be an arbitrary linear transformation.

52.1 Show that the null space of T is a subspace.

52.2 Show that the range of T is a subspace.

© Jason Siefken, 2015–2022

DEFINITION

Induced Transformation

Let M be an $n \times m$ matrix. We say M ***induces*** a linear transformation $\mathcal{T}_M : \mathbb{R}^m \to \mathbb{R}^n$ defined by

$$[\mathcal{T}_M \vec{v}]_{\mathcal{E}'} = M[\vec{v}]_{\mathcal{E}},$$

where $\mathcal{E}$ is the standard basis for $\mathbb{R}^m$ and $\mathcal{E}'$ is the standard basis for $\mathbb{R}^n$.

---

53 Let $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, let $\vec{v} = \vec{e}_1 + \vec{e}_2 \in \mathbb{R}^2$, and let $\mathcal{T}_M$ be the transformation induced by M.

53.1 What is the difference between "$M\vec{v}$" and "$M[\vec{v}]_{\mathcal{E}}$"?

53.2 What is $[\mathcal{T}_M \vec{e}_1]_{\mathcal{E}}$?

53.3 Can you relate the columns of M to the range of $\mathcal{T}_M$?

© Jason Siefken, 2015–2022

Fundamental Subspaces

DEF Associated with any matrix M are three fundamental subspaces: the ***row space*** of M, denoted $\mathrm{row}(M)$, is the span of the rows of M; the ***column space*** of M, denoted $\mathrm{col}(M)$, is the span of the columns of M; and the ***null space*** of M, denoted $\mathrm{null}(M)$, is the set of solutions to $M\vec{x}=\vec{0}$.

---

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

54.1 Describe the row space of A.

54.2 Describe the column space of A.

54.3 Is the row space of A the same as the column space of A?

54.4 Describe the set of all vectors orthogonal to the rows of A.

54.5 Describe the null space of A.

54.6 Describe the range and null space of T_A, the transformation induced by A.

© Jason Siefken, 2015–2022

55

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \qquad C = \mathrm{rref}(B) = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}$$

55.1 How does the row space of B relate to the row space of C?

55.2 How does the null space of B relate to the null space of C?

55.3 Compute the null space of B.

© *Jason Siefken, 2015–2022*

56

$$P = \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix} \qquad Q = \mathrm{rref}(P) = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$$

56.1 How does the column space of P relate to the column space of Q?

56.2 Describe the column space of P and the column space of Q.

© Jason Siefken, 2015–2022

Rank

DEF

For a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, the ***rank*** of T, denoted $\operatorname{rank}(T)$, is the dimension of the range of T.

For an $m \times n$ matrix M, the ***rank*** of M, denoted $\operatorname{rank}(M)$, is the dimension of the column space of M.

---

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

57.1 Describe $\operatorname{range}(\mathcal{P})$ and $\operatorname{range}(\mathcal{R})$.

57.2 What is the rank of $\mathcal{P}$ and the rank of $\mathcal{R}$?

57.3 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?

57.4 Make a conjecture about how the rank of a transformation and the rank of its corresponding matrix relate. Can you justify your claim?

© Jason Siefken, 2015–2022

58

58.1 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) $\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$.

59

Consider the homogeneous system

$$\begin{array}{rrrl} x & +2y & +z & =0\\ x & +2y & +3z & =0\\ -x & -2y & +z & =0 \end{array} \tag{11}$$

and the non-augmented matrix of coefficients $A=\begin{bmatrix}1 & 2 & 1\\ 1 & 2 & 3\\ -1 & -2 & 1\end{bmatrix}$.

59.1 What is rank(A)?

59.2 Give the general solution to system (11).

59.3 Are the column vectors of A linearly independent?

59.4 Give a non-homogeneous system with the same coefficients as (11) that has

(a) infinitely many solutions

(b) no solutions.

© Jason Siefken, 2015–2022

60

60.1 The rank of a 3×4 matrix A is 3. Are the column vectors of A linearly independent?

60.2 The rank of a 4×3 matrix B is 3. Are the column vectors of B linearly independent?

© Jason Siefken, 2015–2022

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

$$\operatorname{rank}(A) + \operatorname{nullity}(A) = \#\text{ of columns in } A.$$

61

61.1 Is there a version of the rank-nullity theorem that applies to linear transformations instead of matrices? If so, state it.

62

The vectors $\vec{u}, \vec{v} \in \mathbb{R}^9$ are linearly independent and $\vec{w} = 2\vec{u} - \vec{v}$. Define $A = [\vec{u}|\vec{v}|\vec{w}]$.

62.1 What is the rank and nullity of A?

62.2 What is the rank and nullity of A^T?

© Jason Siefken, 2015–2022

Inverse Functions & Inverse Matrices

Module 12

In this module you will learn

- The definition of an inverse function and an inverse matrix.
- How to decompose a matrix into the product of elementary matrices and how to use elementary matrices to compute inverses.
- How the order of matrix multiplication matters.
- How row-reduction and matrix inverses relate.

We should think of transformations or functions as machines that perform some manipulation of their input and then give an output. This perspective allows us to divide functions into two natural categories: those that can be undone and those that cannot. The official term for a function that can be undone is an *invertible* function.

Invertible Functions

The simplest function is the *identity function*.

Identity Function.
Let X be a set. The ***identity function*** with domain and codomain X, notated $\mathrm{id} : X \to X$, is the function satisfying
$$\mathrm{id}(x) = x$$
for all $x \in X$.

The identity function is the function that does nothing to its input.[41] When doing precise mathematics, we often prove a function or composition of functions does nothing to its input by showing it is *equal* to the identity function.[42]

In plain terms, a function is invertible if it can be undone. More precisely a function is invertible if there exists an inverse function that when composed with the original function produces the identity function and vice versa.

Inverse Function.
Let $f : X \to Y$ be a function. We say f is ***invertible*** if there exists a function $g : Y \to X$ so that $f \circ g = \mathrm{id}$ and $g \circ f = \mathrm{id}$. In this case, we call g an ***inverse*** of f and write
$$f^{-1} = g.$$

Let's consider an example. You have some money in your pockets. Let $l : \{\text{nickels in left pocket}\} \to \mathbb{N}$ be the function that adds up the value of all the nickels in your left pocket. Let $r : \{\text{nickels in either pocket}\} \to \mathbb{N}$ be the function that adds up the value of all the nickels in both of your pockets. In this case, l would be invertible—if you know that $l(\#\text{ nickels}) = 25$, you must have had 5 nickels in your left pocket. We can write down a formula for l^{-1} as
$$l^{-1}(n) = \frac{n}{5}.$$

However, r is not invertible. If $r(\#\text{ nickels}) = 25$, you *might* have had 5 nickels in your left pocket, but you might have 3 nickels in your left pocket and 2 in your right. We just don't know, so no inverse to r can exist.

What we've just learned is that for a function to be invertible, it must be *one-to-one*.

One-to-one.
Let $f : X \to Y$ be a function. We say f is ***one-to-one*** (or ***injective***) if distinct inputs to f produce distinct outputs. That is $f(x) = f(y)$ implies $x = y$.

Whenever a function f is one-to-one, there exists a function g so that $g \circ f = \mathrm{id}$. However, this is not enough to declare that f is invertible[43] because we also need $f \circ g = \mathrm{id}$. To ensure this, we need f to be *onto*.

[41] Technically, for every set there exists a unique identity function with that set as the domain/codomain, but we won't belabor this point.
[42] This is similar to saying that we know $\vec{x} = \vec{y}$ if and only if $\vec{x} - \vec{y} = \vec{0}$.
[43] In this situation, we say that f is *left*-invertible.

© Jason Siefken, 2015–2022

Onto.
Let $f : X \to Y$ be a function. We say f is ***onto*** (or ***surjective***) if every point in the codomain of f gets mapped to. That is $\mathrm{range}(f) = Y$.

Every invertible function is both one-to-one and onto, and every one-to-one and onto function is invertible. And, as we will learn, this has implications for the rank and nullity of linear transformations.

Invertibility and Linear Transformations

Let's now focus on linear transformations. We know that a linear transformation $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^m$ is invertible if and only if it is one-to-one and onto.

If $\mathcal{T}$ is one-to-one, that means that distinct inputs to $\mathcal{T}$ yield distinct outputs. In other words, the solution to $\mathcal{T}(\vec{x}) = \vec{b}$ is always unique. But, the set of all solutions to $\mathcal{T}(\vec{x}) = \vec{b}$ can be expressed as

$$\mathrm{null}(\mathcal{T}) + \{\vec{p}\}.$$

Therefore, $\mathcal{T}$ is one-to-one if and only if $\mathrm{nullity}(\mathcal{T}) = 0$. If $\mathcal{T}$ is onto, then $\mathrm{range}(\mathcal{T}) = \mathbb{R}^m$ and so $\mathrm{rank}(\mathcal{T}) = m$.

Now, suppose $\mathcal{T}$ is one-to-one and onto. By the rank-nullity theorem,

$$\mathrm{rank}(\mathcal{T}) + \mathrm{nullity}(\mathcal{T}) = 0 + m = m = n = \dim(\text{domain of } \mathcal{T}),$$

and so $\mathcal{T}$ has the same domain and codomain (at least a domain and codomain of the same dimension).

Using the rank-nullity theorem, we can start developing a list of properties that are equivalent to invertibility of a linear transformation.

- $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^m$ is invertible if and only if $\mathrm{nullity}(\mathcal{T}) = 0$ and $\mathrm{rank}(\mathcal{T}) = m$.
- $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^m$ is invertible if and only if $m = n$ and $\mathrm{nullity}(\mathcal{T}) = 0$.
- $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^m$ is invertible if and only if $m = n$ and $\mathrm{rank}(\mathcal{T}) = m$.

Example. Let $\mathcal{P} : \mathbb{R}^2 \to \mathbb{R}^2$ be projection onto the x-axis and let $\mathcal{R} : \mathbb{R}^2 \to \mathbb{R}^2$ be rotation counter-clockwise by $15°$. Classify each of $\mathcal{P}$ and $\mathcal{R}$ as invertible or not.

Notice that $\mathcal{P}(\vec{e}_2) = \mathcal{P}(2\vec{e}_2) = \vec{0}$, therefore $\mathcal{P}$ is not one-to-one and so is not invertible.

Let $\mathcal{Q} : \mathbb{R}^2 \to \mathbb{R}^2$ be rotation clockwise by $15°$. $\mathcal{R}$ and $\mathcal{Q}$ will undo each other. Phrased mathematically,

$$\mathcal{R} \circ \mathcal{Q} = \mathrm{id} \qquad \text{and} \qquad \mathcal{Q} \circ \mathcal{R} = \mathrm{id}.$$

Therefore, $\mathcal{Q}$ is an inverse of $\mathcal{R}$, and so $\mathcal{R}$ is invertible.

One important fact about linear transformations is that if a linear transformation is invertible, its inverse is also a linear transformation.

Theorem. Let $\mathcal{T}$ be an invertible linear transformation. Then $\mathcal{T}^{-1}$ is also a linear transformation.

Proof. Let $\mathcal{T}$ be an invertible linear transformation and let $\mathcal{T}^{-1}$ be its inverse. We need to show that (i) $\mathcal{T}^{-1}$ distributes over addition and (ii) $\mathcal{T}^{-1}$ distributes over scalar multiplication.

(i) First observe that since $\mathcal{T} \circ \mathcal{T}^{-1} = \mathrm{id}$ and because $\mathcal{T}$ is linear, we have

$$\vec{a} + \vec{b} = \mathcal{T} \circ \mathcal{T}^{-1}\vec{a} + \mathcal{T} \circ \mathcal{T}^{-1}\vec{b} = \mathcal{T}(\mathcal{T}^{-1}\vec{a} + \mathcal{T}^{-1}\vec{b}).$$

Since $\mathcal{T}^{-1} \circ \mathcal{T} = \mathrm{id}$, by using the fact that $\vec{a} + \vec{b} = \mathcal{T}(\mathcal{T}^{-1}\vec{a} + \mathcal{T}^{-1}\vec{b})$ we know

$$\mathcal{T}^{-1}(\vec{a} + \vec{b}) = \mathcal{T}^{-1}\Big(\mathcal{T}(\mathcal{T}^{-1}\vec{a} + \mathcal{T}^{-1}\vec{b})\Big) = \mathcal{T}^{-1}\vec{a} + \mathcal{T}^{-1}\vec{b}.$$

(ii) Similar to the proof of (i), we see

$$\mathcal{T}^{-1}(\alpha\vec{a}) = \mathcal{T}^{-1}\Big(\alpha(\mathcal{T} \circ \mathcal{T}^{-1}\vec{a})\Big) = \mathcal{T}^{-1} \circ \mathcal{T}(\alpha\mathcal{T}^{-1}\vec{a}) = \alpha\mathcal{T}^{-1}\vec{a}.$$

■

Invertibility and Matrices

In the world of matrices, the *identity matrix* takes the place of the identity function.

© *Jason Siefken, 2015–2022*

Identity Matrix.
An ***identity matrix*** is a square matrix with ones on the diagonal and zeros everywhere else. The $n \times n$ identity matrix is denoted $I_{n\times n}$, or just I when its size is implied.

We can now define what it means for a matrix to be invertible.[44]

Matrix Inverse.
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 A^{-1}.

Example. Determine whether the matrices $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ and $B = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$ are inverses of each other.

$$AB = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}\begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$
$$BA = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

Therefore, A and B are inverses of each other.

Example. Determine whether the matrices $A = \begin{bmatrix} 2 & 5 & 0 \\ -3 & -7 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} -7 & -5 \\ 3 & 2 \\ 1 & 1 \end{bmatrix}$ are inverses of each other.

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

but

$$BA = \begin{bmatrix} -7 & -5 \\ 3 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 5 & 0 \\ -3 & -7 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -2 & 0 \end{bmatrix} \neq I.$$

Therefore, A and B are not inverses of each other.

Since every matrix induces a linear transformation, we can use the facts we know about invertible linear transformations to produce facts about invertible matrices. In particular:

- An $n \times m$ matrix A is invertible if and only if $\text{nullity}(A) = 0$ and $\text{rank}(A) = n$.
- An $n \times n$ matrix A is invertible if and only if $\text{nullity}(A) = 0$.
- An $n \times n$ matrix A is invertible if and only if $\text{rank}(A) = n$.

Matrix Algebra

The linear equation $ax = b$ has solution $x = \frac{b}{a}$ whenever $a \neq 0$. We arrive at this solution by dividing both sides of the equation by a. Does a similar process exist for solving the matrix equation $A\vec{x} = \vec{b}$? It sure does!

Unfortunately, we cannot divide by a matrix, but to solve $A\vec{x} = \vec{b}$, we don't need to "divide" by a matrix, we just need to eliminate A from the left side. This could be accomplished by using an inverse.

Suppose A is invertible, then

$$A\vec{x} = \vec{b} \qquad \implies \qquad A^{-1}A\vec{x} = A^{-1}\vec{b} \qquad \implies \qquad \vec{x} = A^{-1}\vec{b}.$$

Thus, if we have the inverse of a matrix handy, we can use it to solve a system of equations.

Example. Use the fact that $\begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}^{-1} = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$ to solve the system $\begin{cases} 2x + 5y = 2 \\ -3x - 7y = 1 \end{cases}$.

[44] This should look very similar to what it means for a function to be invertible.

© Jason Siefken, 2015–2022

The system can be rewritten as

$$\begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Multiplying both sides by $\begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}^{-1}$ gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}^{-1}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -19 \\ 8 \end{bmatrix}.$$

It's important to note that, unlike in the case with regular scalars, the order of matrix multiplication matters. So, whereas with scalars you could get away with something like

$$ax = b \qquad \implies \qquad \frac{1}{a}ax = b\frac{1}{a} \qquad \implies \qquad x = \frac{b}{a},$$

with matrices $A\vec{x} = \vec{b}$ *does not imply* $A^{-1}A\vec{x} = \vec{b}A^{-1}$. In fact, if $\vec{b}$ is a column vector, the expression $\vec{b}A^{-1}$ is almost always undefined!

Finding a Matrix Inverse

Whereas before we only knew how to solve a matrix equation $A\vec{x} = \vec{b}$ using row reduction, we now know how to use A^{-1} to solve the same system. In fact, A^{-1} is the exact matrix so that $\vec{x} = A^{-1}\vec{b}$ is the solution to $A\vec{x} = \vec{b}$. Therefore, by picking different $\vec{b}$'s and solving for $\vec{x}$, we can find A^{-1}.

Example. Let $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$. Find A^{-1}.

We know $A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ will be a 2×2 matrix, and we know $\vec{x} = A^{-1}\vec{b}$ will always be the unique solution to $A\vec{x} = \vec{b}$. Therefore, we can find A^{-1} by finding $\vec{x}, \vec{b}$ pairs that satisfy $A\vec{x} = \vec{b}$.
Using row reduction, we see

$$A\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{has solution} \quad \vec{x} = \begin{bmatrix} -7 \\ 3 \end{bmatrix}, \qquad \text{and} \qquad A\vec{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{has solution} \quad \vec{x} = \begin{bmatrix} -5 \\ 2 \end{bmatrix}.$$

Therefore

$$A^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} -7 \\ 3 \end{bmatrix} \qquad \text{and} \qquad A^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} -5 \\ 2 \end{bmatrix},$$

and so

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

Elementary Matrices

Finding the inverse of a matrix can be a lot of work. However if you already know how to undo what the matrix does, finding the inverse might not be so hard. For example, if R_{30} is the matrix that rotates vectors in $\mathbb{R}^2$ counter-clockwise by 30°, its inverse must be R_{-30}, the matrix that rotates vectors in $\mathbb{R}^2$ *clockwise* by 30°.

Like before when we analyzed linear transformations by breaking them up into compositions of simpler linear transformations, another strategy for finding an inverse matrix is to break a matrix into simpler ones whose inverses we can just write down.

Some of the simplest matrices around are the *elementary matrices*.

Elementary Matrix.
A matrix is called an ***elementary matrix*** if it is an identity matrix with a single elementary row operation applied.

Examples of elementary matrices include

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

These matrices are obtained from the row operations "multiply the last row by −5", "add 7 times the last row to

© *Jason Siefken, 2015–2022*

the first", and "swap the first two rows".

Elementary matrices are useful because multiplying by an elementary matrix performs the corresponding elementary row operation! See for yourself:

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&-5\end{bmatrix}\begin{bmatrix}a&b&c\\d&e&f\\g&h&i\end{bmatrix}=\begin{bmatrix}a&b&c\\d&e&f\\-5g&-5h&-5i\end{bmatrix}$$

$$\begin{bmatrix}1&0&7\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}a&b&c\\d&e&f\\g&h&i\end{bmatrix}=\begin{bmatrix}a+7g&b+7h&c+7i\\d&e&f\\g&h&i\end{bmatrix}$$

$$\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}a&b&c\\d&e&f\\g&h&i\end{bmatrix}=\begin{bmatrix}d&e&f\\a&b&c\\g&h&i\end{bmatrix}$$

As a refresher, the elementary row operations are:

- multiply a row by a non-zero constant;
- add a multiple of one row to another; and
- swap two rows.

Each one of these operations can be undone, and so every elementary matrix is invertible. What's more, the inverse is another elementary matrix that is easy to write down.

Example. Find the inverse of $E=\begin{bmatrix}1&0&7\\0&1&0\\0&0&1\end{bmatrix}$.

Since E corresponds to the row operation "add 7 times the last row to the first", E^{-1} must correspond to the row operation "subtract 7 times the last row from the first". Therefore,

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

Elementary Matrices and Inverses

For a matrix M to be invertible, we know that M must be square and $\operatorname{nullity}(M)=0$. That means, M is invertible if and only if $\operatorname{rref}(M)=I$. In other words, M is invertible if there is a sequence of elementary row operations that turn M into I. Each one of these row operations can be represented by an elementary matrix, which gives us the following theorem.

Theorem. A matrix M is invertible if and only if there are elementary matrices $E_1,\ldots,E_k$ so that

$$E_k\cdots E_2E_1M=I.$$

Now, suppose M is invertible and let $E_1,\ldots,E_k$ be elementary matrices so that $E_k\cdots E_2E_1M=I$. We now know

$$E_k\cdots E_2E_1M=\underbrace{(E_k\cdots E_2E_1)}_{Q}M=QM=I.$$

If we can argue that $MQ=I$, then Q will be the inverse of M!

Theorem. If A is a square matrix and $AB=I$ for some matrix B, then $BA=I$.

Proof. Suppose A is a square matrix and that $AB=I$. Since $AB=I$, B must also be square. Since $\operatorname{null}(B)\subseteq\operatorname{null}(AB)$, we know $\operatorname{nullity}(B)\le\operatorname{nullity}(AB)=\operatorname{nullity}(I)=0$, and so B is invertible (since it's a square matrix whose nullity is 0). Let B^{-1} be the inverse of B. Observe now that

$$A=AI=A(BB^{-1})=(AB)B^{-1}=IB^{-1}=B^{-1},$$

and so $A=B^{-1}$. Finally, substituting B^{-1} for A shows

$$BA=BB^{-1}=I.$$

© Jason Siefken, 2015–2022

■

In light of this theorem, we now have a new algorithm for finding the inverse of a matrix—find elementary matrices that turn the matrix into the identity matrix and multiply those elementary matrices together to find the inverse.

Example. Let $A=\begin{bmatrix}1&2&0\\0&4&0\\0&-1&1\end{bmatrix}$. Find A^{-1} using elementary matrices.

We can row-reduce A with the following steps:

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

The elementary matrices corresponding to these steps are

$$E_1=\begin{bmatrix}1&0&0\\0&\frac{1}{4}&0\\0&0&1\end{bmatrix}\qquad E_2=\begin{bmatrix}1&-2&0\\0&1&0\\0&0&1\end{bmatrix}\qquad E_3=\begin{bmatrix}1&0&0\\0&1&0\\0&1&1\end{bmatrix}.$$

We now have

$$E_3E_2E_1A=I,$$

and so

$$A^{-1}=E_3E_2E_1=\begin{bmatrix}1&0&0\\0&1&0\\0&1&1\end{bmatrix}\begin{bmatrix}1&-2&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&\frac{1}{4}&0\\0&0&1\end{bmatrix}=\begin{bmatrix}1&-\frac{1}{2}&0\\0&\frac{1}{4}&0\\0&\frac{1}{4}&1\end{bmatrix}.$$

Decomposition into Elementary Matrices

If A is an invertible matrix, then the double-inverse of A (i.e., $(A^{-1})^{-1}$) is A itself.[45] This is easily proved. By definition, $(A^{-1})^{-1}$ is a matrix B so that $BA^{-1}=I$ and $A^{-1}B=I$. But $B=A$ satisfies this condition!

Now, suppose M is an invertible matrix. Then, there exists a sequence of elementary matrices $E_1,\dots,E_k$ so that $E_k\cdots E_2E_1M=I$ and

$$M^{-1}=E_k\cdots E_2E_1.$$

Therefore

$$M=(M^{-1})^{-1}=(E_k\cdots E_2E_1)^{-1}.$$

Thinking carefully about what $(E_k\cdots E_2E_1)^{-1}$ should be, we see that

$$(E_1^{-1}E_2^{-1}\cdots E_k^{-1})E_k\cdots E_2E_1=I\qquad\text{and}\qquad E_k\cdots E_2E_1(E_1^{-1}E_2^{-1}\cdots E_k^{-1})=I,$$

and so

$$M=(E_k\cdots E_2E_1)^{-1}=E_1^{-1}E_2^{-1}\cdots E_k^{-1}.$$

(Notice the order of matrix multiplication reversed!) Each E_i^{-1} is also an elementary matrix, and so we have just shown that every invertible matrix can be written as the product of elementary matrices. This is actually a double-sided implication (if and only if).

Theorem. A matrix M is invertible if and only if it can be written as the product of elementary matrices.

Proof. Suppose M is invertible. Then, there exists a sequence of elementary matrices $E_1,\dots,E_k$ so that $E_k\cdots E_1M=I$. It follows that

$$M=E_1^{-1}E_2^{-1}\cdots E_k^{-1}$$

is the product of elementary matrices. Conversely, since the product of invertible matrices is invertible and every elementary matrix is invertible, the product of elementary matrices must be invertible. Therefore, if M is not invertible, it *cannot* be written as the product of elementary matrices. ■

[45] Formally we say that the operation of taking a matrix inverse is an *involution*.

© Jason Siefken, 2015–2022

Practice Problems

1 Determine whether the following linear transformations are one-to-one, onto, or both. As well, determine whether or not they are invertible. Justify your answers.

(a) $\mathcal{S}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{S}$ is the linear transformation that doubles every vector.

(b) $\mathcal{R}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{R}$ the linear transformation that rotates every vector clockwise by $72°$.

(c) $\mathcal{P}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{P}$ the linear transformation that projects every vector onto the y-axis.

(d) $\mathcal{F}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{F}$ is the linear transformation that reflects every vector over the line $y=x$.

(e) $\mathcal{T}:\mathbb{R}^3\to\mathbb{R}^3$, where $\mathcal{T}$ is the linear transformation induced by the matrix $M_{\mathcal{T}}=\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}$.

(f) $\mathcal{U}:\mathbb{R}^3\to\mathbb{R}^2$, where $\mathcal{U}$ is the linear transformation induced by the matrix $M_{\mathcal{U}}=\begin{bmatrix}1&2&3\\3&4&5\end{bmatrix}$.

2 Invert the following matrices or explain why they are not invertible.

(a) $M_1=\begin{bmatrix}2&3\\1&-1\end{bmatrix}$

(b) $M_2=\begin{bmatrix}1&0\\1&0\end{bmatrix}$

(c) $M_3=\begin{bmatrix}0&2&1\\1&0&1\\-2&3&0\end{bmatrix}$

(d) $M_4=\begin{bmatrix}2&0&1&8\\1&-5&2&2\\3&-1&0&7\end{bmatrix}$

(e) $M_5=\begin{bmatrix}0&-3&1&2\\1&0&-1&1\\2&-1&0&0\\0&0&1&3\end{bmatrix}$

3 Solve the following systems in two ways: (i) by using row reduction, and (ii) by using inverse matrices.

(a) $\begin{cases}2x+\ \ y=5\\3x+7y=3\end{cases}$

(b) $\begin{cases}2x+2y+3z=4\\2x+2y+\ \ z=0\\4x+5y+6z=2\end{cases}$

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

(a) Express A^{-1} as the product of elementary matrices.

(b) Express A as the product of elementary matrices.

5 For each statement below, determine whether it is true or false. Justify your answer.

(a) For an arbitrary linear transformation $T:\mathbb{R}^m\to\mathbb{R}^n$, if $n\neq m$, then the linear transformation is not invertible.

(b) The matrix $M=\begin{bmatrix}1&0&0\\1&1&0\end{bmatrix}$ is an elementary matrix.

(c) Every elementary matrix is invertible.

(d) The product of elementary matrices is sometimes an elementary matrix.

(e) The product of elementary matrices is always an elementary matrix.

(f) A matrix that induces an invertible linear transformation is necessarily invertible.

(g) A transformation that is one-to-one and onto is always invertible.

(h) For two matricies A and B, if $AB=I$, then A and B are invertible.

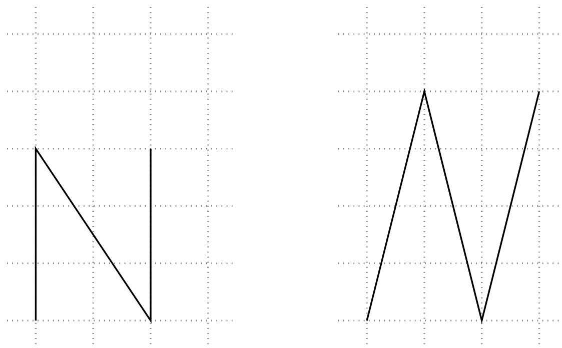
© Jason Siefken, 2015–2022

© Jason Siefken, 2015–2022

Getting back N

63 "We've made a terrible mistake," a council member says. "Can we go back to the regular N?"

Recall the original Italicising N task.



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.

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.*

The Oronto city council has asked you to *unitalicise* the N. Your new task is to 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.

© IOLA Team iola.math.vt.edu

64 64.1 Apply the row operation $R_3 \mapsto R_3 + 2R_1$ to the 3×3 identity matrix and call the result E_1.

64.2 Apply the row operation $R_3 \mapsto R_3 - 2R_1$ to the 3×3 identity matrix and call the result E_2.

Elementary Matrix

DEF A matrix is called an ***elementary matrix*** if it is an identity matrix with a single elementary row operation applied.

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

64.3 Compute E_1A and E_2A. How do the resulting matrices relate to row operations?

64.4 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.

64.5 Without computing, what should E_2E_1 be? What about E_1E_2? Now compute and verify.

© Jason Siefken, 2015–2022

Matrix Inverse

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 A^{-1}.

65 Consider the matrices

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

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

65.1 Which pairs of matrices above are inverses of each other?

© Jason Siefken, 2015–2022

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

66.1 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.

66.2 What is $E_2 E_1 B$?

66.3 Find B^{-1}.

66.4 Can you outline a procedure for finding the inverse of a matrix using elementary matrices?

© Jason Siefken, 2015–2022

67

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & 4 \\ 1 & 3 & -3 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad C = [A|\vec{b}] \qquad A^{-1} = \begin{bmatrix} 9 & -3/2 & -5 \\ -5 & 1 & 3 \\ -2 & 1/2 & 1 \end{bmatrix}$$

67.1 What is $A^{-1}A$?

67.2 What is $\mathrm{rref}(A)$?

67.3 What is $\mathrm{rref}(C)$? (Hint, there is no need to actually do row reduction!)

67.4 Solve the system $A\vec{x} = \vec{b}$.

© Jason Siefken, 2015–2022

68

68.1 For two square matrices X, Y, should $(XY)^{-1} = X^{-1}Y^{-1}$?

68.2 If M is a matrix corresponding to a non-invertible linear transformation T, could M be invertible?

© Jason Siefken, 2015–2022

Change of Basis II

In this module you will learn

- How to create *change-of-basis* matrices.
- How to write a linear transformation in multiple bases.

Given a basis $\mathcal{A}$ for $\mathbb{R}^n$, every vector $\vec{x}\in\mathbb{R}^n$ uniquely corresponds to the list of numbers $[\vec{x}]_\mathcal{A}$ (its coordinates with respect to $\mathcal{A}$), and the operation of writing a vector in a basis is *invertible*.

If we have two bases, $\mathcal{A}$ and $\mathcal{B}$, for $\mathbb{R}^n$, we have two equally valid ways of representing a vector in coordinates.

$$\begin{tikzcd} & \vec{x}\in\mathbb{R}^n \arrow[dl, leftrightarrow, "\text{Represent in basis } \mathcal{A}"'] \arrow[dr, leftrightarrow, "\text{Represent in basis } \mathcal{B}"] & \\ {[\vec{x}]_\mathcal{A}} & & {[\vec{x}]_\mathcal{B}} \end{tikzcd}$$

Not only that, but there must be a function that converts between $[\vec{x}]_\mathcal{A}$ and $[\vec{x}]_\mathcal{B}$. The function works as follows: input the list of numbers $[\vec{x}]_\mathcal{A}$, use those numbers as coefficients of the $\mathcal{A}$ basis vectors to get the true vector $\vec{x}$, and then find the coordinates of that vector with respect to the $\mathcal{B}$ basis.

Example. Let $\mathcal{A}=\{\vec{a}_1,\vec{a}_2\}$ where $\vec{a}_1=\begin{bmatrix}1\\1\end{bmatrix}_\mathcal{E}$ and $\vec{a}_2=\begin{bmatrix}1\\-1\end{bmatrix}_\mathcal{E}$ and let $\mathcal{B}=\{\vec{b}_1,\vec{b}_2\}$ where $\vec{b}_1=\begin{bmatrix}2\\1\end{bmatrix}_\mathcal{E}$ and $\vec{b}_2=\begin{bmatrix}5\\3\end{bmatrix}_\mathcal{E}$ be bases for $\mathbb{R}^2$. Given that $[\vec{x}]_\mathcal{A}=\begin{bmatrix}2\\-3\end{bmatrix}$, find $[\vec{x}]_\mathcal{B}$.

By definition,
$$\vec{x}=2\vec{a}_1-3\vec{a}_2=2(\vec{e}_1+\vec{e}_2)-3(\vec{e}_1-\vec{e}_2)=-\vec{e}_1+5\vec{e}_2.$$
We need to rewrite $\vec{x}$ as a linear combination of $\vec{b}_1=2\vec{e}_1+\vec{e}_2$ and $\vec{b}_2=5\vec{e}_1+3\vec{e}_2$. That is, we need to solve the equation

$$\vec{x}=-\vec{e}_1+5\vec{e}_2=\alpha(2\vec{e}_1+\vec{e}_2)+\beta(5\vec{e}_1+3\vec{e}_2)=(2\alpha+5\beta)\vec{e}_1+(\alpha+3\beta)\vec{e}_2.$$

Equating the coefficients of $\vec{e}_1$ and $\vec{e}_2$, we get

$$\begin{cases}2\alpha+5\beta=-1\\ \alpha+3\beta=5\end{cases},$$

which has a unique solution $(\alpha,\beta)=(-28,11)$. We conclude

$$[\vec{x}]_\mathcal{B}=\begin{bmatrix}-28\\11\end{bmatrix}.$$

For a basis $\mathcal{A}$, the invertible function that takes a vector $\vec{x}$ and generates the coordinates $[\vec{x}]_\mathcal{A}$ is a linear function. Therefore, for bases $\mathcal{A}$ and $\mathcal{B}$, the function that converts $[\vec{x}]_\mathcal{A}$ to $[\vec{x}]_\mathcal{B}$ must have a matrix. This matrix is called the *change of basis* matrix.

Change of Basis Matrix. Let $\mathcal{A}$ and $\mathcal{B}$ be bases for $\mathbb{R}^n$. The matrix M is called a ***change of basis*** matrix (which converts from $\mathcal{A}$ to $\mathcal{B}$) if for all $\vec{x}\in\mathbb{R}^n$

$$M[\vec{x}]_\mathcal{A}=[\vec{x}]_\mathcal{B}.$$

Notationally, $[\mathcal{B}\leftarrow\mathcal{A}]$ stands for the change of basis matrix converting from $\mathcal{A}$ to $\mathcal{B}$, and we may write $M=[\mathcal{B}\leftarrow\mathcal{A}]$.

Example. Let $\mathcal{A}=\{\vec{a}_1,\vec{a}_2\}$ where $\vec{a}_1=\begin{bmatrix}1\\1\end{bmatrix}_\mathcal{E}$ and $\vec{a}_2=\begin{bmatrix}1\\-1\end{bmatrix}_\mathcal{E}$ and let $\mathcal{B}=\{\vec{b}_1,\vec{b}_2\}$ where $\vec{b}_1=\begin{bmatrix}2\\1\end{bmatrix}_\mathcal{E}$ and

© Jason Siefken, 2015–2022

$\vec b_2 = \begin{bmatrix}5\\3\end{bmatrix}_{\mathcal{E}}$ be bases for $\mathbb{R}^2$. Find the change of basis matrix $[\mathcal{B}\leftarrow\mathcal{A}]$.

We know $[\mathcal{B}\leftarrow\mathcal{A}]$ will be a 2×2 matrix and that

$$[\mathcal{B}\leftarrow\mathcal{A}][\vec a_1]_{\mathcal{A}} = [\vec a_1]_{\mathcal{B}} \qquad \text{and} \qquad [\mathcal{B}\leftarrow\mathcal{A}][\vec a_2]_{\mathcal{A}} = [\vec a_2]_{\mathcal{B}}.$$

Therefore, we need to compute $[\vec a_1]_{\mathcal{B}}$ and $[\vec a_2]_{\mathcal{B}}$. Repeating the procedure from the previous example, we find

$$[\vec a_1]_{\mathcal{B}} = \begin{bmatrix}-2\\1\end{bmatrix} \qquad \text{and} \qquad [\vec a_2]_{\mathcal{B}} = \begin{bmatrix}8\\-3\end{bmatrix},$$

and so

$$[\mathcal{B}\leftarrow\mathcal{A}] = \begin{bmatrix}-2 & 8\\ 1 & -3\end{bmatrix}.$$

We can now enhance our diagram from earlier.

$$\begin{tikzcd}
 & \vec x \in \mathbb{R}^n \arrow[dl, leftrightarrow, "\text{Represent in } \mathcal{A} \text{ basis}"'] \arrow[dr, leftrightarrow, "\text{Represent in } \mathcal{B} \text{ basis}"] & \\
{[\vec x]_{\mathcal{A}}} \arrow[rr, dashed, shift right, "{[\mathcal{B}\leftarrow\mathcal{A}]}"'] & & {[\vec x]_{\mathcal{B}}} \arrow[ll, dashed, shift right, "{[\mathcal{A}\leftarrow\mathcal{B}]}"']
\end{tikzcd}$$

The notation $[\mathcal{B}\leftarrow\mathcal{A}]$ for the matrix that changes from the $\mathcal{A}$ basis to the $\mathcal{B}$ basis is suggestive. Suppose we have another basis $\mathcal{C}$. We can obtain $[\mathcal{C}\leftarrow\mathcal{A}]$ by multiplying $[\mathcal{B}\leftarrow\mathcal{A}]$ on the left by $[\mathcal{C}\leftarrow\mathcal{B}]$. That is,

$$[\mathcal{C}\leftarrow\mathcal{A}] = [\mathcal{C}\leftarrow\mathcal{B}][\mathcal{B}\leftarrow\mathcal{A}].$$

The backwards arrow "$\leftarrow$" in the change-of-basis matrix notation comes because when we multiply a vector and a matrix, the matrix is always to the *left* of the vector. So,

$$[\vec x]_{\mathcal{C}} = [\mathcal{C}\leftarrow\mathcal{A}][\vec x]_{\mathcal{A}} = [\mathcal{C}\leftarrow\mathcal{B}][\mathcal{B}\leftarrow\mathcal{A}][\vec x]_{\mathcal{A}}.$$

As such, the notation for the change of basis matrix chains, allowing you to figure out what's going on without too much trouble.

Change of Basis Matrix in Detail

Let $\mathcal{A}$ and $\mathcal{B}$ be bases for $\mathbb{R}^n$ and $M = [\mathcal{B}\leftarrow\mathcal{A}]$ be the matrix that changes from the $\mathcal{A}$ to the $\mathcal{B}$ basis. Since we can change vectors back from $\mathcal{B}$ to $\mathcal{A}$, we know M is invertible and

$$M^{-1} = [\mathcal{A}\leftarrow\mathcal{B}].$$

Just playing with notation, we see

$$M^{-1}M = [\mathcal{A}\leftarrow\mathcal{B}][\mathcal{B}\leftarrow\mathcal{A}] = [\mathcal{A}\leftarrow\mathcal{A}] = I \qquad MM^{-1} = [\mathcal{B}\leftarrow\mathcal{A}][\mathcal{A}\leftarrow\mathcal{B}] = [\mathcal{B}\leftarrow\mathcal{B}] = I,$$

which makes sense. The matrices $[\mathcal{A}\leftarrow\mathcal{A}]$ and $[\mathcal{B}\leftarrow\mathcal{B}]$ take vectors and rewrite them in the same basis, which is to say, they do nothing to the vectors.

The argument above shows that every change of basis matrix is invertible. The converse is also true.

Theorem. An $n\times n$ matrix is invertible if and only if it is a change of basis matrix.

Proof. Suppose $M = [\mathcal{B}\leftarrow\mathcal{A}]$ is a change-of-basis matrix. Then

$$M^{-1} = [\mathcal{A}\leftarrow\mathcal{B}],$$

and so M is invertible.

Alternatively, suppose $M = [C_1|C_2|\cdots|C_n]$ is an invertible $n\times n$ matrix with columns $C_1, \ldots, C_n$. Let $\vec c_i = [C_i]_{\mathcal{E}}$. That is, $\vec c_i$ is the vector which comes from interpreting C_i as coordinates with respect to the standard basis.

Since M is invertible, $\operatorname{rref}(M) = I$, and so $\{\vec c_1, \ldots, \vec c_n\}$ is a linearly independent set of n vectors. Therefore $\mathcal{C} = \{\vec c_1, \ldots, \vec c_n\}$ is a basis for $\mathbb{R}^n$. Now, observe

$$M[\vec c_i]_{\mathcal{C}} = C_i = [\vec c_i]_{\mathcal{E}}$$

© Jason Siefken, 2015–2022

for $i = 1, \dots, n$, and so $M = [\mathcal{E} \leftarrow \mathcal{C}]$ is a change-of-basis matrix. ■

The proof of the above theorem highlights something interesting. Let $\mathcal{A} = \{\vec{a}_1, \dots, \vec{a}_n\}$ be a basis for $\mathbb{R}^n$. It is always the case that

$$[\vec{a}_i]_{\mathcal{A}} = \begin{bmatrix} \vdots \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

has a 1 in the ith position and zeros elsewhere. Now, let $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ be another basis for $\mathbb{R}^n$ and define the matrix $M = \big[[\vec{a}_1]_{\mathcal{B}} \quad [\vec{a}_2]_{\mathcal{B}} \quad \cdots \quad [\vec{a}_n]_{\mathcal{B}}\big]$ to be the matrix with columns $[\vec{a}_1]_{\mathcal{B}}, \dots, [\vec{a}_n]_{\mathcal{B}}$. Since multiplying a matrix by $[\vec{a}_i]_{\mathcal{A}}$ will pick out the ith column, we have that

$$M[\vec{a}_i]_{\mathcal{A}} = [\vec{a}_i]_{\mathcal{B}}.$$

In other words,

$$M = [\mathcal{B} \leftarrow \mathcal{A}].$$

Transformations and Bases

A linear transformation $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ always has a matrix associated with it. This matrix is defined as the matrix M so that

$$[\mathcal{T}\vec{x}]_{\mathcal{E}} = M[\vec{x}]_{\mathcal{E}}.$$

But, what if we swapped out $\mathcal{E}$ for a different basis?

Linear Transformation in a Basis. Let $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let $\mathcal{B}$ be a basis for $\mathbb{R}^n$. The ***matrix for $\mathcal{T}$ with respect to $\mathcal{B}$***, notated $[\mathcal{T}]_{\mathcal{B}}$, is the $n \times n$ matrix satisfying

$$[\mathcal{T}\vec{x}]_{\mathcal{B}} = [\mathcal{T}]_{\mathcal{B}}[\vec{x}]_{\mathcal{B}}.$$

In this case, we say the matrix $[\mathcal{T}]_{\mathcal{B}}$ is the representation of $\mathcal{T}$ in the $\mathcal{B}$ basis.

Just like there are many ways to write down coordinates for a vector—one per choice of basis—there are many ways to write down a matrix for a linear transformation. Up to this point, when we've said "M is a matrix for $\mathcal{T}$", what we meant is "$M = [\mathcal{T}]_{\mathcal{E}}$". And, like with vectors, if we talk about a matrix for a linear transformation without specifying the basis, we mean the matrix for the transformation with respect to the standard basis.

Example. Let $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$ where $\vec{b}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}_{\mathcal{E}}$ and $\vec{b}_2 = \begin{bmatrix} 5 \\ -7 \end{bmatrix}_{\mathcal{E}}$ be a basis for $\mathbb{R}^2$ and let $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation that stretches in the $\vec{e}_1$ direction by a factor of 2. Find $[\mathcal{T}]_{\mathcal{E}}$ and $[\mathcal{T}]_{\mathcal{B}}$.

Since $\mathcal{T}\vec{e}_1 = 2\vec{e}_1$ and $\mathcal{T}\vec{e}_2 = \vec{e}_2$, We know

$$[\mathcal{T}]_{\mathcal{E}}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \qquad \text{and} \qquad [\mathcal{T}]_{\mathcal{E}}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and so

$$[\mathcal{T}]_{\mathcal{E}} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

We can find $[\mathcal{T}]_{\mathcal{B}}$ in two ways: directly from the definition, or by using change of basis matrices. First, we will work directly from the definition.

To find $[\mathcal{T}]_{\mathcal{B}}$, we need to figure out what $\mathcal{T}$ does to $\vec{b}_1$ and $\vec{b}_2$. However, since $\mathcal{T}$ is described in term of $\vec{e}_1$ and $\vec{e}_2$, it might be easier to express $\vec{e}_1$ and $\vec{e}_2$ in the $\mathcal{B}$ basis, and then analyze $\mathcal{T}$.

Computing,

$$\begin{aligned} [\vec{e}_1]_{\mathcal{B}} &= [\mathcal{B} \leftarrow \mathcal{E}]\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -7 \\ 3 \end{bmatrix} \\ [\vec{e}_2]_{\mathcal{B}} &= [\mathcal{B} \leftarrow \mathcal{E}]\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ 2 \end{bmatrix} \end{aligned}.$$

© Jason Siefken, 2015–2022

Now we know

$$[\mathcal{T}]_{\mathcal{B}}[\vec{e}_1]_{\mathcal{B}} = [\mathcal{T}\vec{e}_1]_{\mathcal{B}} = [2\vec{e}_1]_{\mathcal{B}} = \begin{bmatrix}-14\\6\end{bmatrix}$$
$$[\mathcal{T}]_{\mathcal{B}}[\vec{e}_2]_{\mathcal{B}} = [\mathcal{T}\vec{e}_2]_{\mathcal{B}} = [\vec{e}_2]_{\mathcal{B}} = \begin{bmatrix}-5\\2\end{bmatrix}.$$

Since $[\mathcal{T}]_{\mathcal{B}}$ is a 2×2 matrix, we can use what we know to solve for its entries, finding

$$[\mathcal{T}]_{\mathcal{B}} = \begin{bmatrix}-13 & -35\\6 & 16\end{bmatrix}.$$

Let's try finding $[\mathcal{T}]_{\mathcal{B}}$ using change of basis matrices. We already know $[\mathcal{T}]_{\mathcal{E}}$, and so

$$[\mathcal{T}]_{\mathcal{B}} = [\mathcal{B}\leftarrow\mathcal{E}][\mathcal{T}]_{\mathcal{E}}[\mathcal{E}\leftarrow\mathcal{B}].$$

Further, we know

$$[\mathcal{E}\leftarrow\mathcal{B}] = \begin{bmatrix}2 & 5\\-3 & -7\end{bmatrix} \qquad\text{and}\qquad [\mathcal{B}\leftarrow\mathcal{E}] = \begin{bmatrix}2 & 5\\-3 & -7\end{bmatrix}^{-1} = \begin{bmatrix}-7 & -5\\3 & 2\end{bmatrix}.$$

Putting it all together,

$$[\mathcal{T}]_{\mathcal{B}} = \begin{bmatrix}-7 & -5\\3 & 2\end{bmatrix}\begin{bmatrix}2 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}2 & 5\\-3 & -7\end{bmatrix} = \begin{bmatrix}-13 & -35\\6 & 16\end{bmatrix}.$$

Similar Matrices

Just like some bases are better than others to represent particular vectors, some bases are better than others to represent a particular linear transformation.

Example. Let $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$ where $\vec{b}_1 = \begin{bmatrix}2\\-3\end{bmatrix}_{\mathcal{E}}$ and $\vec{b}_2 = \begin{bmatrix}5\\-7\end{bmatrix}_{\mathcal{E}}$ be a basis for $\mathbb{R}^2$ and let $\mathcal{S}:\mathbb{R}^2\to\mathbb{R}^2$ be the transformation that stretches in the $\vec{b}_1 = 2\vec{e}_1 - 3\vec{e}_2$ direction by a factor of 2 and reflects vectors in the $\vec{b}_2 = 5\vec{e}_1 - 7\vec{e}_2$ direction. Find $[\mathcal{S}]_{\mathcal{E}}$ and $[\mathcal{S}]_{\mathcal{B}}$.

In this example, $\mathcal{S}$ is described in terms of the $\mathcal{B}$ basis. We know

$$\mathcal{S}\vec{b}_1 = 2\vec{b}_1 \qquad\text{and}\qquad \mathcal{S}\vec{b}_2 = -\vec{b}_2.$$

Therefore,

$$[\mathcal{S}]_{\mathcal{B}} = \begin{bmatrix}2 & 0\\0 & -1\end{bmatrix}.$$

To find $[\mathcal{S}]_{\mathcal{E}}$, we will use change of basis matrices. Notice that

$$[\mathcal{S}]_{\mathcal{E}} = [\mathcal{E}\leftarrow\mathcal{B}][\mathcal{S}]_{\mathcal{B}}[\mathcal{B}\leftarrow\mathcal{E}],$$

and that

$$[\mathcal{E}\leftarrow\mathcal{B}] = \begin{bmatrix}2 & 5\\-3 & -7\end{bmatrix} \qquad\text{and}\qquad [\mathcal{B}\leftarrow\mathcal{E}] = \begin{bmatrix}2 & 5\\-3 & -7\end{bmatrix}^{-1} = \begin{bmatrix}-7 & -5\\3 & 2\end{bmatrix}.$$

Therefore

$$[\mathcal{S}]_{\mathcal{E}} = \begin{bmatrix}2 & 5\\-3 & -7\end{bmatrix}\begin{bmatrix}2 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}-7 & -5\\3 & 2\end{bmatrix} = \begin{bmatrix}-43 & -30\\63 & 44\end{bmatrix}.$$

In the example above, $[\mathcal{S}]_{\mathcal{B}}$ is a much nicer matrix than $[\mathcal{S}]_{\mathcal{E}}$. However, the two matrices relate to each other. After all,

$$[\mathcal{S}]_{\mathcal{B}} = [\mathcal{B}\leftarrow\mathcal{E}][\mathcal{S}]_{\mathcal{E}}[\mathcal{E}\leftarrow\mathcal{B}].$$

In this case, we call these matrices *similar*.[46]

Similar Matrices. The 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

[46] Another commonly used term is *conjugate*.

© Jason Siefken, 2015–2022

matrix X so that
$$A = XBX^{-1}.$$

The X in the definition of similar matrices is always a change-of-basis matrix.

When studying a linear transformation, you can pick any basis to represent it in and study the resulting matrix. Different choices of basis will give you different perspectives on the linear transformation. In what's to follow, we will work to find the "best" basis in which to study a given linear transformation.[47]

Practice Problems

1 Let
$$\mathcal{A} = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}_{\mathcal{E}}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}_{\mathcal{E}} \right\} \quad \text{and} \quad \mathcal{B} = \left\{ \begin{bmatrix} 3 \\ -1 \end{bmatrix}_{\mathcal{E}}, \begin{bmatrix} -2 \\ 3 \end{bmatrix}_{\mathcal{E}} \right\}.$$
be bases for $\mathbb{R}^2$. Define $\vec{x} \in \mathbb{R}^2$ by $[\vec{x}]_{\mathcal{A}} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

(a) Find $[\vec{x}]_{\mathcal{E}}$ and $[\vec{x}]_{\mathcal{B}}$.

(b) Find the change of basis matrices $[\mathcal{E} \leftarrow \mathcal{A}]$, $[\mathcal{A} \leftarrow \mathcal{E}]$, $[\mathcal{B} \leftarrow \mathcal{A}]$, and $[\mathcal{A} \leftarrow \mathcal{B}]$.

2 Let
$$\mathcal{A} = \left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}_{\mathcal{E}}, \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}_{\mathcal{E}}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}_{\mathcal{E}} \right\} \quad \text{and} \quad \mathcal{B} = \{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$$
be bases for $\mathbb{R}^3$ where
$$\vec{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{A}} \qquad \vec{b}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}_{\mathcal{A}} \qquad \vec{b}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}_{\mathcal{A}}.$$

(a) Find the representation of $\vec{b}_1$, $\vec{b}_2$, and $\vec{b}_3$ in the standard basis.

(b) Find the change of basis matrices $[\mathcal{A} \leftarrow \mathcal{E}]$ and $[\mathcal{E} \leftarrow \mathcal{B}]$.

(c) Use $[\mathcal{A} \leftarrow \mathcal{E}]$ and $[\mathcal{E} \leftarrow \mathcal{B}]$ to compute $[\mathcal{A} \leftarrow \mathcal{B}]$.

3 Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}_{\mathcal{E}}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{E}} \right\}$. For each linear transformation $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^2$ defined below, compute $[\mathcal{T}]_{\mathcal{E}}$ and $[\mathcal{T}]_{\mathcal{B}}$.

(a) Let $\mathcal{T}$ be the transformation that rotates every vector counter clockwise by $90°$.

(b) Let $\mathcal{T}$ be the transformation that projects every vector onto the y-axis.

(c) Let $\mathcal{T}$ be the transformation that doubles every vector.

(d) Let $\mathcal{T}$ be the transformation that reflects every vector over the line $y = x$.

4 For each statement below, determine whether it is true or false. Justify your answer.

(a) Any invertible $n \times n$ matrix can be viewed as a change of basis matrix.

(b) Any $n \times n$ matrix is similar to itself.

(c) Let A be an $m \times n$ matrix. If $m \neq n$, then there is no matrix that is similar to A.

(d) Any invertible $n \times n$ matrix A is similar to A^{-1} since $AA^{-1} = I$.

[47] If you cannot wait, the "best" basis will turn out to be the *eigen basis* (provided it exists).

© Jason Siefken, 2015–2022

© Jason Siefken, 2015–2022

69 Let $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$ where $\vec{b}_1 = \begin{bmatrix}1\\1\end{bmatrix}_{\mathcal{E}}$, $\vec{b}_2 = \begin{bmatrix}1\\-1\end{bmatrix}_{\mathcal{E}}$ and let $X = [\vec{b}_1|\vec{b}_2]$ be the matrix whose columns are $[\vec{b}_1]_{\mathcal{E}}$ and $[\vec{b}_2]_{\mathcal{E}}$.

69.1 Write down X.

69.2 Compute $[\vec{e}_1]_{\mathcal{B}}$ and $[\vec{e}_2]_{\mathcal{B}}$.

69.3 Compute $X[\vec{e}_1]_{\mathcal{B}}$ and $X[\vec{e}_2]_{\mathcal{B}}$. What do you notice?

69.4 Find the matrix X^{-1}. How does X^{-1} relate to change of basis?

© Jason Siefken, 2015–2022

70 Let $\mathcal{E} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ be the standard basis for $\mathbb{R}^n$ and let $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$ be another basis for $\mathbb{R}^n$. Define the matrix $X = [\vec{b}_1|\vec{b}_2|\cdots|\vec{b}_n]$ to be the matrix whose columns are the $\vec{b}_i$ vectors written in the standard basis. Notice that X converts vectors from the $\mathcal{B}$ basis into the standard basis. In other words,

$$X[\vec{v}]_{\mathcal{B}} = [\vec{v}]_{\mathcal{E}}.$$

70.1 Should X^{-1} exist? Explain.

70.2 Consider the equation

$$X^{-1}[\vec{v}]_{?} = [\vec{v}]_{?}.$$

Can you fill in the "?" symbols so that the equation makes sense?

70.3 What is $[\vec{b}_1]_{\mathcal{B}}$? How about $[\vec{b}_2]_{\mathcal{B}}$? Can you generalize to $[\vec{b}_i]_{\mathcal{B}}$?

© Jason Siefken, 2015–2022

71 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.

71.1 Compute $[\vec{c}_1]_{\mathcal{C}}$ and $[\vec{c}_2]_{\mathcal{C}}$.

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.

71.2 Compute $T\begin{bmatrix}2\\1\end{bmatrix}_{\mathcal{E}}$ and $T\begin{bmatrix}5\\3\end{bmatrix}_{\mathcal{E}}$.

71.3 Compute $[T\vec{c}_1]_{\mathcal{C}}$ and $[T\vec{c}_2]_{\mathcal{C}}$.

71.4 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 $\begin{bmatrix}?\\?\end{bmatrix}_{\mathcal{C}}$).

71.5 Find $[T]_{\mathcal{C}}$, the matrix for T in the $\mathcal{C}$ basis.

71.6 Find $[T]_{\mathcal{E}}$, the matrix for T in the standard basis.

DEFINITION

Similar Matrices

The 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

$$A = XBX^{-1}.$$

© Jason Siefken, 2015–2022

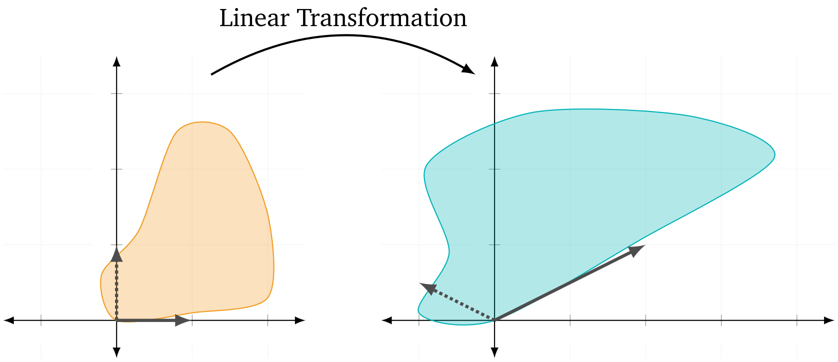
© Jason Siefken, 2015–2022

Determinants

In this module you will learn

- The definition of the determinant of a linear transformation and of a matrix.
- How to interpret the determinant as a change-of-volume factor.
- How to relate the determinant of $S \circ T$ to the determinant of S and of T.
- How to compute the determinants of elementary matrices and how to compute determinants of large matrices using row reduction.

Linear transformations transform vectors, but they also change *sets*.



It turns out to be particularly useful to track by how much a linear transformation changes area/volume. This number (which is associated with a linear transformation with the same domain and codomain) is called the *determinant*.[48]

Volumes

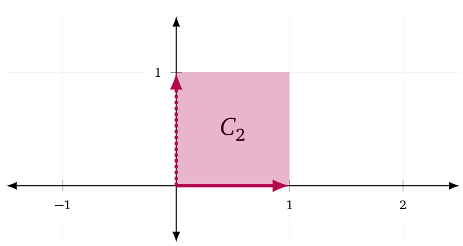
In this module, most examples will be in $\mathbb{R}^2$ because they're easier to draw. The definitions given will extend to $\mathbb{R}^n$ for any n, however we need to establish some conventions to properly express these ideas in English. In English, we say that a two-dimensional figure has an *area* and a three-and-up dimensional figure has a *volume*. In this section, *we will use the term volume to also mean area* where appropriate.

To measure how volume changes, we need to compare input volumes and output volumes. The easiest volume to compute is that of the *unit n-cube*, which has a special notation.

Unit n-cube. 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

$$C_n = \left\{ \vec{x} \in \mathbb{R}^n : \vec{x} = \sum_{i=1}^{n} \alpha_i \vec{e}_i \text{ for some } \alpha_1, \dots, \alpha_n \in [0,1] \right\} = [0,1]^n.$$

C_2 should look familiar as the unit square in $\mathbb{R}^2$ with lower-left corner at the origin.



C_n always has volume 1,[49] and by analyzing the image of C_n under a linear transformation, we can see by how much a given transformation changes volume.

[48] This number is *almost* the determinant. The only difference is that the determinant might have a ± in front.

[49] The fact that the volume of C_n is 1 is actually by definition.

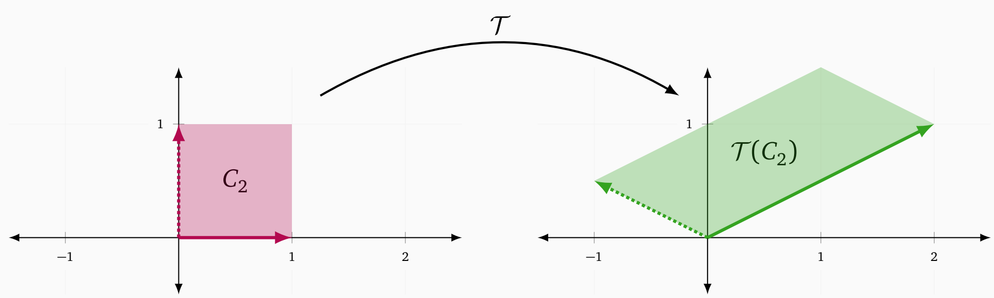
© Jason Siefken, 2015–2022

Example. Let $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $\mathcal{T}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x - y \\ x + \frac{1}{2}y \end{bmatrix}$. Find the volume of $\mathcal{T}(C_2)$.

Recall that C_2 is the unit square in $\mathbb{R}^2$ with sides given by $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Applying the linear transformation $\mathcal{T}$ to $\vec{e}_1$ and $\vec{e}_2$, we obtain

$$\mathcal{T}(\vec{e}_1) = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \mathcal{T}(\vec{e}_2) = \begin{bmatrix} -1 \\ \frac{1}{2} \end{bmatrix}.$$

Plotting $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ \frac{1}{2} \end{bmatrix}$, we see $\mathcal{T}(C_2)$ is a parallelogram with base $\sqrt{5}$ and height $\frac{2\sqrt{5}}{5}$.



Therefore, the volume of $\mathcal{T}(C_2)$ is 2.

Let Vol(X) stand for the volume of the set X. Given a linear transformation $\mathcal{S} : \mathbb{R}^n \to \mathbb{R}^n$, we can define a number

$$\operatorname{Vol Change}(\mathcal{S}) = \frac{\operatorname{Vol}(\mathcal{S}(C_n))}{\operatorname{Vol}(C_n)} = \frac{\operatorname{Vol}(\mathcal{S}(C_n))}{1} = \operatorname{Vol}(\mathcal{S}(C_n)).$$

A priori, Vol Change$(\mathcal{S})$ only describes how $\mathcal{S}$ changes the volume of C_n. However, because $\mathcal{S}$ is a linear transformation, Vol Change$(\mathcal{S})$ actually describes how $\mathcal{S}$ changes the volume of any figure.

Theorem. Let $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let $X \subseteq \mathbb{R}^n$ be a subset with volume α. Then the volume of $\mathcal{T}(X)$ is $\alpha \cdot \operatorname{Vol Change}(\mathcal{T})$.

A full proof of the above theorem requires calculus and limits, but the linear algebra ideas are based on the following theorems.

Theorem. Suppose $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation, $X \subseteq \mathbb{R}^n$ is a subset, and the volume of $\mathcal{T}(X)$ is α. Then for any $\vec{p} \in \mathbb{R}^n$, the volume of $\mathcal{T}(X + \{\vec{p}\})$ is α.

Proof. Fix $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$, $X \subseteq \mathbb{R}^n$, and $\vec{p} \in \mathbb{R}^n$. Combining linearity with the definition of set addition, we see

$$\mathcal{T}(X + \{\vec{p}\}) = \mathcal{T}(X) + \mathcal{T}(\{\vec{p}\}) = \mathcal{T}(X) + \{\mathcal{T}(\vec{p})\}$$

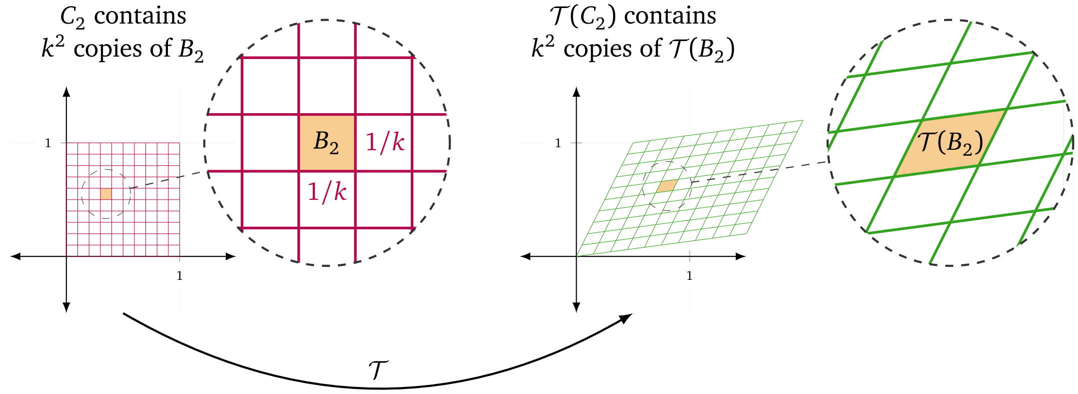
and so $\mathcal{T}(X + \{\vec{p}\})$ is just a translation of $\mathcal{T}(X)$. Since translations don't change volume, $\mathcal{T}(X + \{\vec{p}\})$ and $\mathcal{T}(X)$ must have the same volume. ■

Theorem. Fix k and let B_n be C_n scaled to have side lengths $\frac{1}{k}$ and let $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. Then

$$\operatorname{Vol Change}(\mathcal{T}) = \frac{\operatorname{Vol}(\mathcal{T}(B_n))}{\operatorname{Vol}(B_n)}.$$

Rather than giving a formal proof of the above theorem, let's make a motivating picture.

© Jason Siefken, 2015–2022

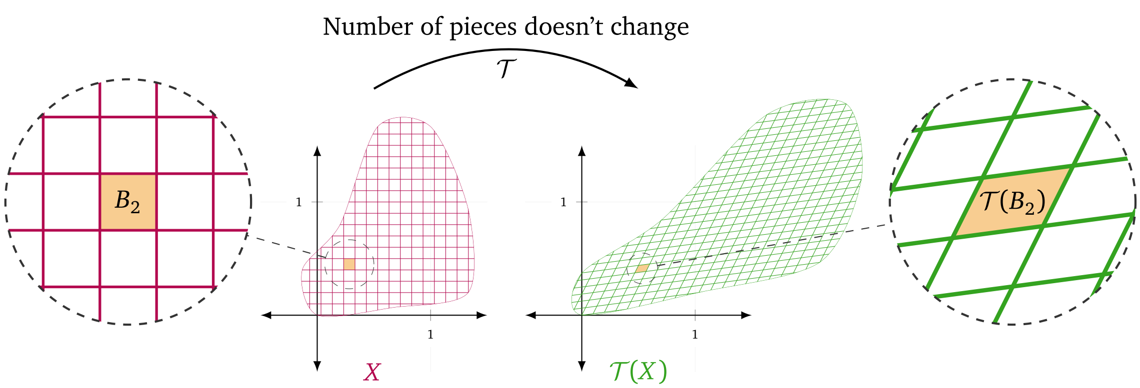


The argument now goes: there are k^n copies of B_n in C_n and k^n copies of $\mathcal{T}(B_n)$ in $T(C_n)$. Thus,

$$\operatorname{Vol}\operatorname{Change}(\mathcal{T}) = \frac{\operatorname{Vol}(\mathcal{T}(C_n))}{\operatorname{Vol}(\mathcal{C}_n)} = \frac{k^n \operatorname{Vol}(\mathcal{T}(B_n))}{k^n \operatorname{Vol}(\mathcal{B}_n)} = \frac{\operatorname{Vol}(\mathcal{T}(B_n))}{\operatorname{Vol}(\mathcal{B}_n)}.$$

Now we can finally show that for a linear transformation $\mathcal{T}:\mathbb{R}^n \to \mathbb{R}^n$, the number "Vol Change($\mathcal{T}$)" actually corresponds to how much $\mathcal{T}$ changes the volume of any figure by.

The argument goes as follows: for a figure $X \subseteq \mathbb{R}^n$, we can fill it with shrunken and translated copies, B_n, of C_n. The same number of copies of $\mathcal{T}(B_n)$ fit inside $\mathcal{T}(X)$ as do B_n's fit inside X. Therefore, the change in volume between $\mathcal{T}(X)$ and X must be the same as the change in volume between $\mathcal{T}(B_n)$ and B_n, which is Vol Change($\mathcal{T}$).



The Determinant

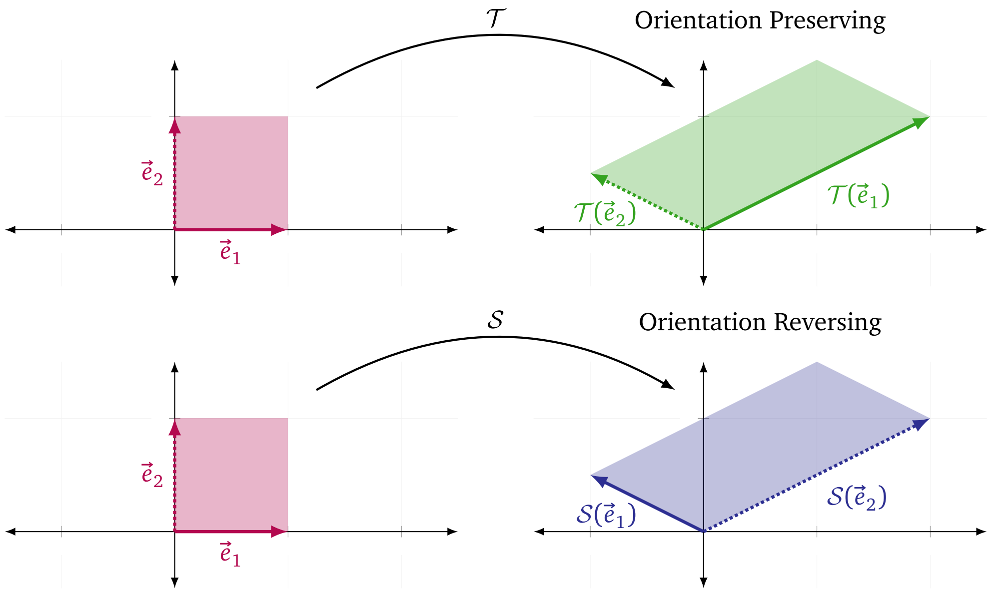
The determinant of a linear transformation $\mathcal{T}:\mathbb{R}^n \to \mathbb{R}^n$ is *almost* the same as Vol Change($\mathcal{T}$), but with one twist: *orientation*.

> **Determinant.** The ***determinant*** of a linear transformation $\mathcal{T}:\mathbb{R}^n \to \mathbb{R}^n$, denoted $\det(\mathcal{T})$ or $|\mathcal{T}|$, 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.

We need to understand what the term *oriented volume* means. We've previously defined the orientation of a basis, and we can use the orientation of a basis to define whether a linear transformation is *orientation preserving* or *orientation reversing*.

> **Orientation Preserving Linear Transformation.** Let $\mathcal{T}:\mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. We say $\mathcal{T}$ is ***orientation preserving*** if the ordered basis $\{\mathcal{T}(\vec{e}_1),\dots,\mathcal{T}(\vec{e}_n)\}$ is positively oriented and we say $\mathcal{T}$ is ***orientation reversing*** if the ordered basis $\{\mathcal{T}(\vec{e}_1),\dots,\mathcal{T}(\vec{e}_n)\}$ is negatively oriented. If $\{\mathcal{T}(\vec{e}_1),\dots,\mathcal{T}(\vec{e}_n)\}$ is not a basis for $\mathbb{R}^n$, then $\mathcal{T}$ is neither orientation preserving nor orientation reversing.

© Jason Siefken, 2015–2022



In the figure above, $\mathcal{T}$ is orientation preserving and $\mathcal{S}$ is orientation reversing.

For an arbitrary linear transformation $\mathcal{Q} : \mathbb{R}^n \to \mathbb{R}^n$ and a set $X \subseteq \mathbb{R}^n$, we define the *oriented volume* of $\mathcal{Q}(X)$ to be $+\operatorname{Vol}\mathcal{Q}(X)$ if $\mathcal{Q}$ is orientation preserving and $-\operatorname{Vol}\mathcal{Q}(X)$ if $\mathcal{Q}$ is orientation reversing.

Example. Let $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $\mathcal{T}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + y \\ -x + \frac{1}{2}y \end{bmatrix}$. Find $\det(\mathcal{T})$.

This is the same $\mathcal{T}$ as from the previous example where we computed $\operatorname{Vol}\mathcal{T}(C_2) = 2$. Since $\mathcal{T}$ is orientation preserving, we conclude that $\det(\mathcal{T}) = 2$.

Example. Let $\mathcal{S} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $\mathcal{S}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x + y \\ x + y \end{bmatrix}$. Find $\det(\mathcal{S})$.

By drawing a picture, we see that $\mathcal{S}(C_2)$ is a square and $\operatorname{Vol}\mathcal{S}(C_2) = 2$. However, $\mathcal{S}(\vec{e}_1) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\mathcal{S}(\vec{e}_2) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ form a negatively oriented basis, and so $\mathcal{S}$ is orientation reversing. Therefore, $\det(\mathcal{S}) = -\operatorname{Vol}\mathcal{S}(C_2) = -2$.

Example. Let $\mathcal{P} : \mathbb{R}^2 \to \mathbb{R}^2$ be projection onto the line with equation $x + 2y = 0$. Find $\det(\mathcal{P})$.

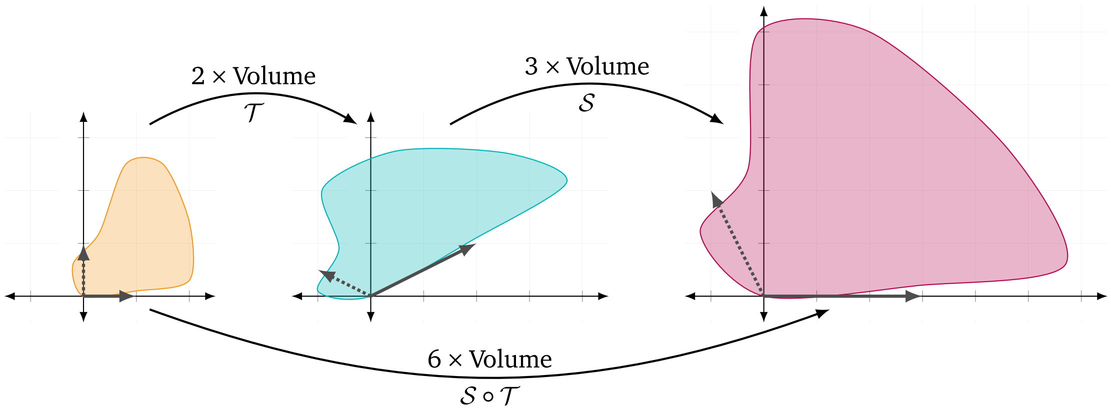
Because $\mathcal{P}$ projects everything to a line, we know $\mathcal{P}(C_2)$ must be a line segment and therefore has volume zero. Thus $\det(\mathcal{P}) = 0$.

Determinants of Composition

Volume changes are naturally multiplicative. If a linear transformation $\mathcal{T}$ changes volume by a factor of α and $\mathcal{S}$ changes volume by a factor of β, then $\mathcal{S} \circ \mathcal{T}$ changes volume by a factor of $\beta\alpha$. Thus, determinants must also be multiplicative.[50]

[50] To fully argue this, we need to show that the composition of two orientation-reversing transformations is orientation preserving.

© *Jason Siefken, 2015–2022*



Theorem. Let $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ and $\mathcal{S} : \mathbb{R}^n \to \mathbb{R}^n$ be linear transformations. Then

$$\det(\mathcal{S} \circ \mathcal{T}) = \det(\mathcal{S}) \det(\mathcal{T}).$$

This means that we can compute the determinant of a complicated transformation by breaking it up into simpler ones and computing the determinant of each piece.

Determinants of Matrices

The determinant of a matrix is defined as the determinant of its induced transformation. That means, the determinant is multiplicative with respect to matrix multiplication (because it's multiplicative with respect to function composition).

Theorem. Let A and B be $n \times n$ matrices. Then

$$\det(AB) = \det(A) \det(B).$$

We will derive an algorithm for finding the determinant of a matrix by considering the determinant of elementary matrices. But first, consider the following theorem.

Theorem (Volume Theorem I). For a square matrix M, $\det(M)$ is the oriented volume of the parallelepiped[a] given by the column vectors.

[a] A parallelepiped is the n-dimensional analog of a parallelogram.

Proof. Let M be an $n \times n$ matrix and let $\mathcal{T}_M$ be its induced transformation. We know the sides of $\mathcal{T}_M(C_n)$ are given by $\{\mathcal{T}_M(\vec{e}_1), \dots, \mathcal{T}_M(\vec{e}_n)\}$. And, by definition,

$$[\mathcal{T}_M(\vec{e}_i)]_{\mathcal{E}} = M[\vec{e}_i]_{\mathcal{E}} = \text{ } i\text{th column of } M.$$

Therefore $\mathcal{T}_M(C_n)$ is the parallelepiped whose sides are given by the columns of M. ∎

This means we can think about the determinant of a matrix by considering its columns. Now we are ready to consider the determinants of the elementary matrices!

There are three types of elementary matrices corresponding to the three elementary row operations. For each one, we need to understand how the induced transformation changes volume.

Multiply a row by a non-zero constant α. Let E_m be such an elementary matrix. Scaling one row of I is equivalent to scaling one column of I, and so the columns of E_m specify a parallelepiped that is scaled by α in one direction.

For example, if

$$E_m = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha \end{bmatrix} \qquad \text{then} \qquad \{\vec{e}_1, \vec{e}_2, \vec{e}_3\} \mapsto \{\vec{e}_1, \vec{e}_2, \alpha\vec{e}_3\}.$$

Thus $\det(E_m) = \alpha$.

Swap two rows. Let E_s be such an elementary matrix. Swapping two rows of I is equivalent to swapping two columns of I, so E_s is I with two columns swapped. This reverses the orientation of the basis given by the columns.

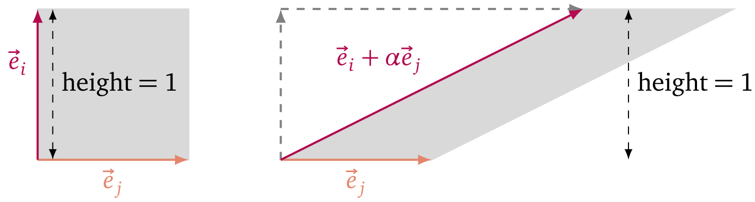
© Jason Siefken, 2015–2022

For example, if

$$E_s = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{then} \qquad \{\vec{e}_1, \vec{e}_2, \vec{e}_3\} \mapsto \{\vec{e}_2, \vec{e}_1, \vec{e}_3\}.$$

Thus $\det(E_s) = -1$.

Add a multiple of one row to another. Let E_a be such an elementary matrix. The columns of E_a are the same as the columns of I except that one column where $\vec{e}_i$ is replaced with $\vec{e}_i + \alpha\vec{e}_j$. This has the effect of *shearing* C_n in the $\vec{e}_j$ direction.



Since C_n is sheared in a direction parallel to one of its other sides, its volume is not changed. Thus $\det(E_a) = 1$.

Takeaway. The determinants of elementary matrices are all easy to compute and the determinant of the most-used type of elementary matrix is 1.

Now, by decomposing a matrix into the product of elementary matrices, we can use the multiplicative property of the determinant (and the formulas for the determinants of the different types of elementary matrices) to compute the determinant of an invertible matrix.

Example. Use elementary matrices to find the determinant of $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

We can row-reduce A with the following steps.

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

The elementary matrices corresponding to these steps are

$$E_1 = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \qquad \text{and} \qquad E_3 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix},$$

and so $E_3E_2E_1A = I$. Therefore

$$A = E_1^{-1}E_2^{-1}E_3^{-1}I = E_1^{-1}E_2^{-1}E_3^{-1} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

Using the fact that the determinant is multiplicative, we get

$$\begin{aligned} \det(A) &= \det\left(\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}\right)\det\left(\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}\right)\det\left(\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\right) \\ &= (1)(-2)(1) = -2. \end{aligned}$$

Determinants and Invertibility

We can use elementary matrices to compute the determinant of any invertible matrix by decomposing it into the product of elementary matrices. But, what about non-invertible matrices?

Let M be an $n \times n$ matrix that is *not* invertible. Then, we must have $\operatorname{nullity}(M) > 0$ and $\dim(\operatorname{col}(M)) = \operatorname{rank}(M) < n$. Geometrically, this means there is at least one line of vectors, $\operatorname{null}(M)$, that gets collapsed to $\vec{0}$, and the column space of M must be "flattened" (i.e., it has lost a dimension). Therefore, the volume of the parallelepiped given by the columns of M must be zero, and so $\det(M) = 0$.

Based on this argument, we have the following theorem.

© Jason Siefken, 2015–2022

Theorem. Let A be an $n \times n$ matrix. A is invertible if and only if $\det(A) \neq 0$.

Proof. If A is invertible, $A = E_1 \cdots E_k$, where $E_1, \dots, E_k$ are elementary matrices, and so

$$\det(A) = \det(E_1 \cdots E_k) = \det(E_1) \cdots \det(E_k).$$

All elementary matrices have non-zero determinants, and so $\det(A) \neq 0$.

Conversely, if A is not invertible, $\operatorname{rank}(A) < n$, which means the parallelepiped given by the columns of A is "flattened" and has zero volume. ■

We now have another way to tell if a matrix is invertible! But, for an invertible matrix A, how do $\det(A)$ and $\det(A^{-1})$ relate? Well, by definition

$$AA^{-1} = I,$$

and so

$$\det(AA^{-1}) = \det(A)\det(A^{-1}) = \det(I) = 1,$$

which gives

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Determinants and Transposes

Somewhat mysteriously, we have the following theorem.

Theorem (Volume Theorem II). The determinant of a square matrix A is equal to the oriented volume of the parallelepiped given by the rows of A.

Volume Theorem II can be concisely stated as $\det(A) = \det(A^T)$, and joins other strange transpose-related facts (like $\operatorname{rank}(A) = \operatorname{rank}(A^T)$).

We can prove Volume Theorem II using elementary matrices.

Proof. Suppose A is not invertible. Then, neither is A^T and so $\det(A) = \det(A^T) = 0$.

Suppose A is invertible and $A = E_1 \cdots E_k$ where $E_1, \dots, E_k$ are elementary matrices. We then have

$$A^T = E_k^T \cdots E_1^T,$$

which follows from the fact that the transpose reverses the order of matrix multiplication (i.e., $(XY)^T = Y^T X^T$). However, for each E_i, we may observe that E_i^T is another elementary matrix of the same type and with the same determinant. Therefore,

$$\begin{aligned}\det(A^T) = \det(E_k^T \cdots E_1^T) &= \det(E_k^T) \cdots \det(E_i^T) \\ &= \det(E_k) \cdots \det(E_1) \\ &= \det(E_1) \cdots \det(E_k) = \det(E_1 \cdots E_k) = \det(A).\end{aligned}$$

The key observations for this proof are that (i) $\det(E_i^T) = \det(E_i)$ and (ii) since the $\det(E_i)$'s are just scalars, the order in which they are multiplied doesn't matter. ■

Practice Problems

1 Let $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $\mathcal{T}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x - y \\ x - \frac{1}{4}y \end{bmatrix}$. Find the volume of $\mathcal{T}(C_2)$.

2 Let $\mathcal{S} : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $\mathcal{S}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x + y + z \\ x - \frac{1}{2}y \\ z \end{bmatrix}$. Find the volume of $\mathcal{S}(C_3)$.

3 Let $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $\mathcal{T}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2y \\ -x - y \end{bmatrix}$.

(a) Draw $\mathcal{E}$ and $\mathcal{T}(\mathcal{E})$ and then determine whether $\mathcal{T}$ is orientation preserving or orientation reversing.

(b) Find $\det(\mathcal{T})$.

4 For each linear transformation defined below, find its determinant.

(a) $\mathcal{S} : \mathbb{R}^2 \to \mathbb{R}^2$, where $\mathcal{S}$ shortens every vector by a factor of $\frac{2}{3}$.

(b) $\mathcal{R} : \mathbb{R}^2 \to \mathbb{R}^2$, where $\mathcal{R}$ is rotation counter-clockwise by $90°$.

(c) $\mathcal{F} : \mathbb{R}^2 \to \mathbb{R}^2$, where $\mathcal{F}$ is reflection across the line $y = -x$.

(d) $\mathcal{G} : \mathbb{R}^2 \to \mathbb{R}^2$, where $\mathcal{G}(\vec{x}) = \mathcal{P}(\vec{x}) + \mathcal{Q}(\vec{x})$ and where $\mathcal{P}$ is projection onto the line $y = x$ and $\mathcal{Q}$ is projection onto the line $y = -\frac{1}{2}x$.

© Jason Siefken, 2015–2022

(e) $\mathcal{T}:\mathbb{R}^3\to\mathbb{R}^3$, where $\mathcal{T}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}x-y+z\\z+x-\frac{1}{3}y\\z\end{bmatrix}$.

(f) $\mathcal{J}:\mathbb{R}^3\to\mathbb{R}^3$, where $\mathcal{J}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\\x+y+z\end{bmatrix}$.

(g) $\mathcal{K}\circ\mathcal{H}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{H}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}x+2y\\-x-y\end{bmatrix}$, and $\mathcal{K}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}-x-2y\\x+y\end{bmatrix}$.

5 Let $A=\begin{bmatrix}2&3\\1&5\end{bmatrix}$.

(a) Use elementary matrices to find $\det(A)$.

(b) Draw a picture of the parallelogram given by the rows of A.

(c) Draw a picture of the parallelogram given by the columns of A.

(d) How do the areas of the parallelograms drawn in parts 5b and 5c relate?

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

(a) Use elementary matrices to find $\det(A)$.

(b) Find $\det(A^{-1})$.

(c) Find $\det(A^T)$, and compare your answer with 6a. Are they the same? Explain.

7 Let A be an $n\times n$ matrix that can be decomposed into the product of elementary matrices.

(a) What is $\operatorname{rank}(A)$? Justify your answer.

(b) What is $\operatorname{null}(A^{-1})$? Justify your answer.

8 Anna and Ella are studying the relationship between determinant and volume. In particular, they are studying $\mathcal{S}:\mathbb{R}^3\to\mathbb{R}^3$ defined by $\mathcal{S}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}4x\\2z\\0\end{bmatrix}$, and $\mathcal{T}:\mathbb{R}^3\to\mathbb{R}^2$ defined by $\mathcal{T}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}2x\\8z\end{bmatrix}$.

For each conversation below, (a) evaluate Anna and Ella's arguments as *correct*, *mostly correct*, or *incorrect*; (b) point out where each argument makes correct/incorrect statements; (c) give a correct numerical value for the determinant or explain why it doesn't exist.

(a) *Anna says:*

Since the image of C_3 under $\mathcal{S}$ is the parallelepiped generated by $\begin{bmatrix}4\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix}, and \begin{bmatrix}0\\2\\0\end{bmatrix}$, which is 2-dimensional parallelogram, the volume of $\mathcal{S}(C_3)$ is just the area of this parallelogram, which is 8. Thus, $\det(\mathcal{S})=8$.

Ella says:

$\det(\mathcal{S})$ is undefined, because $\mathcal{S}$ is not invertible.

(b) *Anna says:*

Since the image of C_3 under $\mathcal{T}$ is the parallelepiped generated by $\begin{bmatrix}2\\0\end{bmatrix},\begin{bmatrix}0\\0\end{bmatrix}$, and $\begin{bmatrix}0\\8\end{bmatrix}$, which is a parallelogram in $\mathbb{R}^2$, the signed volume of $\mathcal{T}(C_3)$ is just the signed area of this parallelogram, which is 16. Thus, $\det(\mathcal{T})=16$.

Ella says:

$\det(\mathcal{T})$ is undefined, because $\det(\mathcal{T})$ is only defined when the domain and codomain of $\mathcal{T}$ are the same.

© Jason Siefken, 2015–2022

DEFINITION

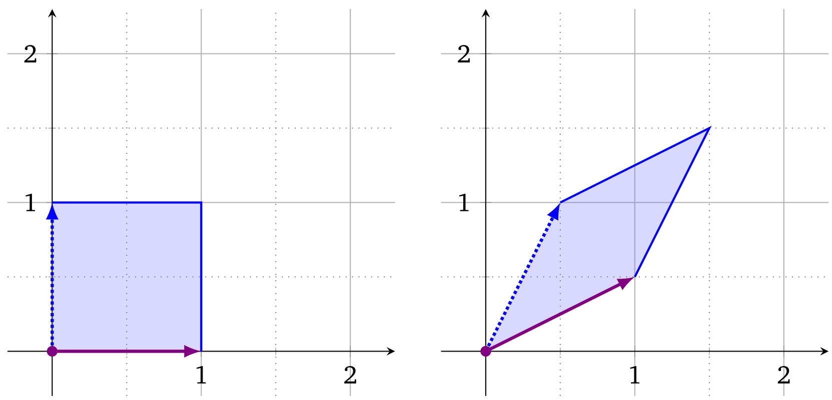
Unit n-cube

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.

72 The picture shows what the linear transformation T does to the unit square (i.e., the unit 2-cube).



72.1 What is $T\begin{bmatrix}1\\0\end{bmatrix}$, $T\begin{bmatrix}0\\1\end{bmatrix}$, $T\begin{bmatrix}1\\1\end{bmatrix}$?

72.2 Write down a matrix for T.

72.3 What is the volume of the image of the unit square (i.e., the volume of $T(C_2)$)? You may use trigonometry.

 © Jason Siefken, 2015–2022

Determinant

DEF The ***determinant*** of a linear transformation $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$, denoted $\det(\mathcal{T})$ or $|\mathcal{T}|$, 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.

---

73 We know the following about the linear transformation A:

$$A\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}2\\0\end{bmatrix} \qquad \text{and} \qquad A\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix}.$$

73.1 Draw C_2 and $A(C_2)$, the image of the unit square under A.

73.2 Compute the area of $A(C_2)$.

73.3 Compute $\det(A)$.

---

74 Suppose R is a rotation counter-clockwise by 30°.

74.1 Draw C_2 and $R(C_2)$.

74.2 Compute the area of $R(C_2)$.

74.3 Compute $\det(R)$.

© Jason Siefken, 2015–2022

75 We know the following about the linear transformation F:

$$F\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}0\\1\end{bmatrix} \qquad \text{and} \qquad F\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}.$$

75.1 What is $\det(F)$?

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

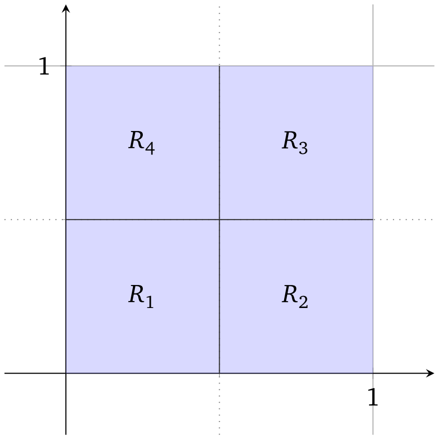
THM For a square matrix M, $\det(M)$ is the oriented volume of the parallelepiped (n-dimensional parallelogram) given by the row vectors of M.

76 76.1 Explain Volume Theorem I using the definition of determinant.

76.2 Based on Volume Theorems I and II, how should $\det(M)$ and $\det(M^T)$ relate for a square matrix M?

© Jason Siefken, 2015–2022

77



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.

$$M\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ y \end{bmatrix}$$

$$T : \mathbb{R}^2 \to \mathbb{R}^2 \text{ has determinant } 2$$

$$S : \mathbb{R}^2 \to \mathbb{R}^2 \text{ has determinant } 3$$

77.1 Find the volumes (areas) of R_1, R_2, R_3, R_4, and R.

77.2 Compute the oriented volume of $M(R_1)$, $M(R_2)$, and $M(R)$.

77.3 Do you have enough information to compute the oriented volume of $T(R_2)$? What about the oriented volume of $T(R + \{\vec{e}_2\})$?

77.4 What is the oriented volume of $S \circ T(R)$? What is $\det(S \circ T)$?

© Jason Siefken, 2015–2022

- 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.

78.1 What is $\det(E_f)$?

78.2 What is $\det(E_m)$?

78.3 What is $\det(E_a)$?

78.4 What is $\det(E_f E_m)$?

78.5 What is $\det(4I_{3\times 3})$?

78.6 What is $\det(W)$ where $W = E_f E_a E_f E_m E_m$?

© Jason Siefken, 2015–2022

79

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

79.1 What is $\det(U)$?

79.2 V is a square matrix and $\mathrm{rref}(V)$ has a row of zeros. What is $\det(V)$?

© Jason Siefken, 2015–2022

80

80.1 V is a square matrix whose columns are linearly dependent. What is $\det(V)$?

80.2 P is projection onto $\operatorname{span}\left\{\begin{bmatrix}-1\\-1\end{bmatrix}\right\}$. What is $\det(P)$?

© Jason Siefken, 2015–2022

81 Suppose you know $\det(X) = 4$.

81.1 What is $\det(X^{-1})$?

81.2 Derive a relationship between $\det(Y)$ and $\det(Y^{-1})$ for an arbitrary matrix Y.

81.3 Suppose Y is not invertible. What is $\det(Y)$?

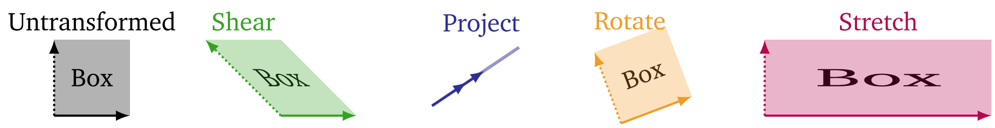
© Jason Siefken, 2015–2022

Eigenvalues and Eigenvectors

In this module you will learn

- The definition of eigenvalues and eigenvectors.
- That eigenvectors give a particularly nice basis in which to study a linear transformation.
- How the characteristic polynomial relates to eigenvalues.

From here on out, we will only be considering linear transformations with the same domain and codomain (i.e., transformations $\mathcal{T}:\mathbb{R}^n\to\mathbb{R}^n$). Why? Because that will allow us to *compare* input and output vectors. By comparing inputs and outputs, we may describe a linear transformation as a stretch, twist, shear, rotation, projection, or some combination of all of these operations.



It's the stretched vectors that we're most interested in now. If $\mathcal{T}$ stretches the vector $\vec v$, then $\mathcal{T}$, in that direction, can be described by $\vec v\mapsto \alpha\vec v$, which is an easy-to-understand linear transformation. The "stretch" directions for a linear transformation have a special name—*eigen directions*—and the vectors that are stretched are called *eigenvectors*.

Eigenvector. Let X be a linear transformation or a matrix. 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 λ. We call λ the ***eigenvalue*** of X corresponding to the eigenvector $\vec v$.

The word *eigen* is German for characteristic, representative, or intrinsic, and we will see that eigenvectors provide one of the best contexts in which to understand a linear transformation.

Example. Let $\mathcal{P}:\mathbb{R}^2\to\mathbb{R}^2$ be projection onto the line ℓ given by $y=x$. Find the eigenvectors and eigenvalues of $\mathcal{P}$.

We are looking for vectors $\vec v\neq\vec 0$ such that $\mathcal{P}\vec v=\lambda\vec v$ for some λ. Since $\mathcal{P}(\ell)=\ell$, we know for any $\vec v\in\ell$

$$\mathcal{P}(\vec v)=1\vec v=\vec v.$$

Therefore, any non-zero multiple of $\begin{bmatrix}1\\1\end{bmatrix}$ is an eigenvector for $\mathcal{P}$ with corresponding eigenvalue 1.

By considering the null space of $\mathcal{P}$, we see, for example,

$$\mathcal{P}\begin{bmatrix}1\\-1\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}=0\begin{bmatrix}1\\-1\end{bmatrix},$$

and so $\begin{bmatrix}1\\-1\end{bmatrix}$ and all its non-zero multiples are eigenvectors of $\mathcal{P}$ with corresponding eigenvalue 0.

Finding Eigenvectors

Sometimes you can find the eigenvectors/values of a linear transformation just by thinking about it. For example, for reflections, projections, and dilations, the eigen directions are geometrically clear. However, for an arbitrary matrix transformation, it may not be obvious.

Our goal now will be to see if we can leverage linear algebra knowledge to find eigenvectors/values. So that we don't have to switch back and forth between thinking about linear transformations and thinking about matrices, let's just think about matrices for now.

Let M be a square matrix. The vector $\vec v\neq\vec 0$ is an eigenvector for M if and only if there exists a scalar λ so that

$$M\vec v=\lambda\vec v. \tag{12}$$

© Jason Siefken, 2015–2022

Put another way, $\vec{v} \neq \vec{0}$ is an eigenvector for M if and only if

$$M\vec{v} - \lambda\vec{v} = (M - \lambda I)\vec{v} = \vec{0}.$$

The middle equation provides a key insight. The operation $\vec{v} \mapsto M\vec{v} - \lambda\vec{v}$ can be achieved by multiplying $\vec{v}$ by the single matrix $E_\lambda = M - \lambda I$.

Now we have that $\vec{v} \neq \vec{0}$ is an eigenvector for M if and only if

$$E_\lambda\vec{v} = (M - \lambda I)\vec{v} = M\vec{v} - \lambda\vec{v} = \vec{0},$$

or, phrased another way, $\vec{v}$ is a non-zero vector satisfying $\vec{v} \in \mathrm{null}(E_\lambda)$.

We've reduced the problem of finding eigenvectors/values of M to finding the null space of E_λ, a related matrix.

Characteristic Polynomial

Let M be an $n \times n$ matrix and define $E_\lambda = M - \lambda I$. Every eigenvector for M must be in the null space of E_λ for some λ. However, because eigenvectors must be non-zero, the only chance we have of finding an eigenvector is if $\mathrm{null}(E_\lambda) \neq \{\vec{0}\}$. In other words, we would like to know when $\mathrm{null}(E_\lambda)$ is *non-trivial*.

We're well equipped to answer this question. Because E_λ is an $n \times n$ matrix, we know E_λ has a non-trivial null space if and only if E_λ is not invertible which is true if and only if $\det(E_\lambda) = 0$. Every λ defines a different E_λ where eigenvectors could be hiding. By viewing $\det(E_\lambda)$ *as a function of* λ, we can use our mathematical knowledge of single-variable functions to figure out when $\det(E_\lambda) = 0$.

The quantity $\det(E_\lambda)$, viewed as a function of λ, has a special name—it's called the *characteristic polynomial*.[51]

Characteristic Polynomial.
For a matrix A, the ***characteristic polynomial*** of A is

$$\mathrm{char}(A) = \det(A - \lambda I).$$

Example. Find the characteristic polynomial of $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

By the definition of the characteristic polynomial of A, we have

$$\begin{aligned} \mathrm{char}(A) &= \det(A - \lambda I) \\ &= \det\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = \det\left(\begin{bmatrix} 1-\lambda & 2 \\ 3 & 4-\lambda \end{bmatrix}\right) \\ &= (1-\lambda)(4-\lambda) - 6 = \lambda^2 - 5\lambda - 2. \end{aligned}$$

For an $n \times n$ matrix A, $\mathrm{char}(A)$ has some nice properties.

- $\mathrm{char}(A)$ is a polynomial.[52]
- $\mathrm{char}(A)$ has degree n.
- The coefficient of the λ^n term in $\mathrm{char}(A)$ is ± 1; $+1$ if n is even and -1 if n is odd.
- $\mathrm{char}(A)$ evaluated at $\lambda = 0$ is $\det(A)$.
- The roots of $\mathrm{char}(A)$ are precisely the eigenvalues of A.

We will just accept these properties as facts, but each of them can be proved with the tools we've developed.

Using the Characteristic Polynomial to find Eigenvalues

With the characteristic polynomial in hand, finding eigenvectors/values becomes easier.

Example. Find the eigenvectors/values of $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$.

[51] This time the term is traditionally given the English name, rather than being called the *eigenpolynomial*.
[52] A priori, it's not obvious that $\det(A - \lambda I)$ should be a polynomial as opposed to some other type of function.

© *Jason Siefken, 2015–2022*

Like the previous example, we first compute char(A).

$$\begin{aligned}\text{char}(A) &= \det\left(\begin{bmatrix}1-\lambda & 2\\ 3 & 2-\lambda\end{bmatrix}\right)\\ &= (1-\lambda)(2-\lambda)-6 = \lambda^2-3\lambda-4 = (4-\lambda)(-1-\lambda)\end{aligned}$$

Next, we solve for when char(A) $=0$ to find eigenvalues, which are $\lambda_1=-1$ and $\lambda_2=4$.
We know non-zero vectors in null$(A-\lambda_1 I)$ are eigenvectors with eigenvalue -1. Computing,

$$\text{null}(A-\lambda_1 I) = \text{null}\left(\begin{bmatrix}2 & 2\\ 3 & 3\end{bmatrix}\right) = \text{span}\left\{\begin{bmatrix}1\\-1\end{bmatrix}\right\},$$

And so the eigenvectors of A corresponding to eigenvalue $\lambda_1=-1$ are the non-zero multiples of $\begin{bmatrix}1\\-1\end{bmatrix}$.
Similarly, for $\lambda_2=4$, we compute

$$\text{null}(A-\lambda_2 I) = \text{null}\left(\begin{bmatrix}-3 & 2\\ 3 & -2\end{bmatrix}\right) = \text{span}\left\{\begin{bmatrix}2\\3\end{bmatrix}\right\},$$

and so the eigenvectors for A with eigenvalue 4 are the non-zero multiples of $\begin{bmatrix}2\\3\end{bmatrix}$.

Using the characteristic polynomial, we can show that every eigenvalue for a matrix is a root of some polynomial (the characteristic polynomial). In general, finding roots of polynomials is a hard problem,[53] and it's not one we will focus on. However, it's handy to have the quadratic formula in your back pocket for factoring particularly stubborn polynomials.

Example. Find the eigenvectors/values of $A=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}$.

First, we find the roots of char(A) by setting it to 0.

$$\begin{aligned}\text{char}(A) &= \det\left(\begin{bmatrix}1-\lambda & 2\\ 3 & 4-\lambda\end{bmatrix}\right)\\ &= (1-\lambda)(4-\lambda)-6 = \lambda^2-5\lambda-2 = 0\end{aligned}$$

By the quadratic formula,[a] we find that

$$\lambda_1 = \frac{5-\sqrt{33}}{2} \qquad \lambda_2 = \frac{5+\sqrt{33}}{2}$$

are the roots of char(A).
Following the procedure outlined above, we need to find null$(A-\lambda_1 I)$ and null$(A-\lambda_2 I)$.
We will start by row reducing $A-\lambda_1 I$.

$$\begin{aligned}\begin{bmatrix}1-\frac{5-\sqrt{33}}{2} & 2\\ 3 & 4-\frac{5-\sqrt{33}}{2}\end{bmatrix} &\to \begin{bmatrix}\frac{-3+\sqrt{33}}{2} & 2\\ 3 & \frac{3+\sqrt{33}}{2}\end{bmatrix} \to \begin{bmatrix}1 & \frac{4}{-3+\sqrt{33}}\\ 1 & \frac{3+\sqrt{33}}{6}\end{bmatrix}\\ &\to \begin{bmatrix}1 & \frac{4(3+\sqrt{33})}{(-3+\sqrt{33})(3+\sqrt{33})}\\ 1 & \frac{3+\sqrt{33}}{6}\end{bmatrix} \to \begin{bmatrix}1 & \frac{3+\sqrt{33}}{6}\\ 1 & \frac{3+\sqrt{33}}{6}\end{bmatrix}\\ &\to \begin{bmatrix}1 & \frac{3+\sqrt{33}}{6}\\ 0 & 0\end{bmatrix}\end{aligned}$$

Thus, we conclude that the eigenvectors with eigenvalue $\frac{5-\sqrt{33}}{2}$ are the non-zero multiples of $\begin{bmatrix}\frac{3+\sqrt{33}}{6}\\ -1\end{bmatrix}$.
Similarly, the eigenvectors with eigenvalue $\frac{5+\sqrt{33}}{2}$ are the non-zero multiples of $\begin{bmatrix}\frac{3-\sqrt{33}}{6}\\ -1\end{bmatrix}$.

[a] Recall that the roots of ax^2+bx+c are given by $\frac{-b\pm\sqrt{b^2-4ac}}{2a}$.

[53] In fact, numerically approximating eigenvalues turns out to be easier than finding roots of a polynomial, so many numerical root finding algorithms actually create a matrix with an appropriate characteristic polynomial and use numerical linear algebra to approximate its roots.

© Jason Siefken, 2015–2022

Transformations without Eigenvectors

Are there linear transformations without eigenvectors? Well, it depends on exactly what you mean. Let $\mathcal{R}:\mathbb{R}^2\to\mathbb{R}^2$ be rotation counter-clockwise by $90°$. Are there any non-zero vectors that don't change direction when $\mathcal{R}$ is applied? Certainly not.

Let's examine further. We know $M_R=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$ is a matrix for $\mathcal{R}$, and

$$\operatorname{char}(M_R)=\lambda^2+1.$$

The polynomial λ^2+1 has no real roots, which means that M_R (and $\mathcal{R}$) have no real eigenvalues. However, λ^2+1 does have *complex* roots of $\pm i$. So far, we've always thought of scalars as real numbers, but if we allow complex numbers as scalars and view $\mathcal{R}$ as a transformation from $\mathbb{C}^2\to\mathbb{C}^2$, it would have eigenvalues and eigenvectors.

Complex numbers play an invaluable role in advanced linear algebra and applications of linear algebra to physics. We will leave the following theorem as food for thought.[54]

Theorem. If A is a square matrix, then A always has an eigenvalue provided complex eigenvalues are permitted.

Practice Problems

1 For each linear transformation defined below, find its eigenvectors and eigenvalues. If it has no eigenvectors/values, explain why not.

(a) $\mathcal{S}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{S}$ stretches every vector by the factor of 3.

(b) $\mathcal{R}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{R}$ rotates every vector clockwise by $\frac{\pi}{4}$.

(c) $\mathcal{P}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{P}$ projects every vector onto the line ℓ given by $y=-x$.

(d) $\mathcal{F}:\mathbb{R}^2\to\mathbb{R}^2$, where $\mathcal{F}$ reflects every vector over the line ℓ given by $y=-x$.

(e) $T:\mathbb{R}^3\to\mathbb{R}^3$, where T is a linear transformation induced by the matrix $\begin{bmatrix}1&2&3\\3&4&5\\5&6&7\end{bmatrix}$.

(f) $U:\mathbb{R}^3\to\mathbb{R}^2$, where U is a linear transformation induced by the matrix $\begin{bmatrix}1&2&3\\3&4&5\end{bmatrix}$.

2 Let $A=\begin{bmatrix}a&b\\c&d\end{bmatrix}$, where $a,b,c,d\in\mathbb{R}$.

(a) Find the characteristic polynomial of A.

(b) Find conditions on a,b,c,d so that A has (i) two distinct real eigenvalues, (ii) exactly one real eigenvalue, (iii) no real eigenvalues.

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

(a) Find the eigenvalues of B.

(b) Find the eigenvalues of B^T.

(c) A vector $\vec v\neq\vec 0$ is called a *left-eigenvector* for B if $\vec vB=\lambda\vec v$ for some scalar λ (Here we consider $\vec v$ a *row* vector). Find all left eigenvectors for B.

4 For each statement below, determine whether it is true or false. Justify you answer.

(a) Zero cannot be an eigenvalue of any matrix.

(b) $\vec 0$ cannot be an eigenvector of any matrix.

(c) A 2×2 matrix always has a real eigenvalue.

(d) A 3×3 matrix always has a real eigenvalue.

(e) A 3×2 matrix always has a real eigenvalue.

(f) The matrix $M=\begin{bmatrix}3&3&3&3\\3&3&3&3\\3&3&3&3\\3&3&3&3\end{bmatrix}$ has exactly one eigenvalue.

(g) An invertible square matrix can never have zero as an eigenvalue.

(h) A non-invertible square matrix *always* has zero as an eigenvalue.

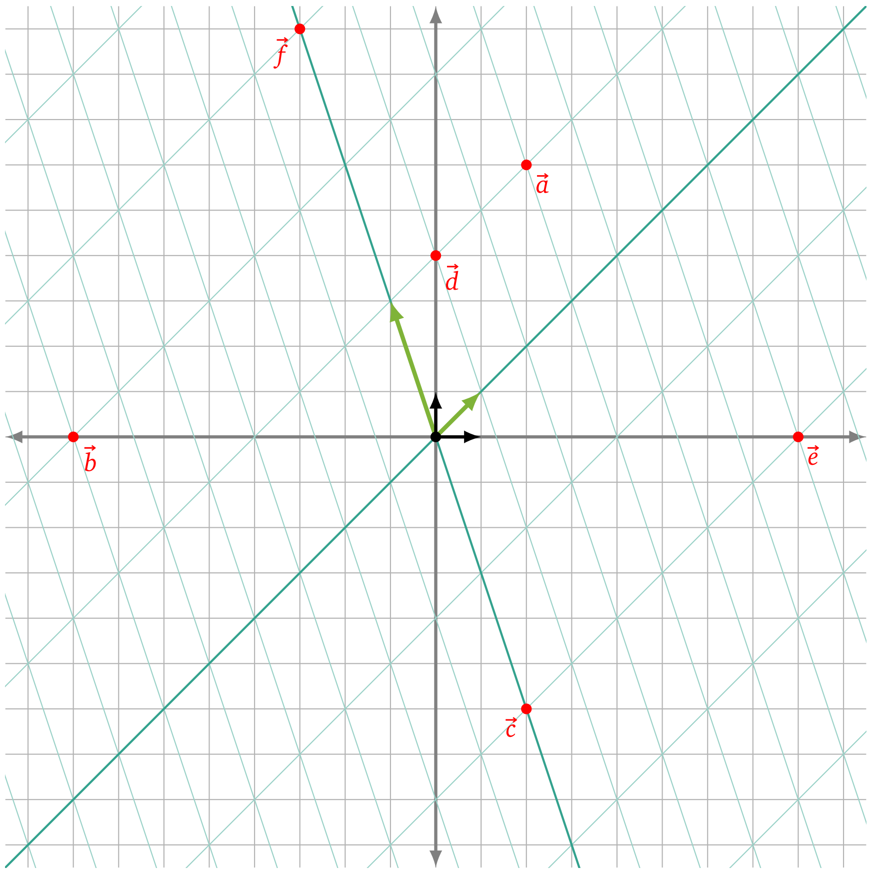
[54] The theorem is a direct corollary of the fundamental theorem of algebra.

© *Jason Siefken, 2015–2022*

The Green and the Black

82 The subway system of Oronto is laid out in a skewed grid. All tracks run parallel to one of the green lines shown. Compass directions are given by the black lines.

While studying the subway map, you decide to pick two bases to help: 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. The city commission is considering renumbering all the stops along the $y = -3x$ direction. You deduce that the commission's proposal can be modeled by a linear transformation.

   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? What coordinate system would you propose the city commission use to describe their plans?

© IOLA Team iola.math.vt.edu

Eigenvectors

Eigenvector

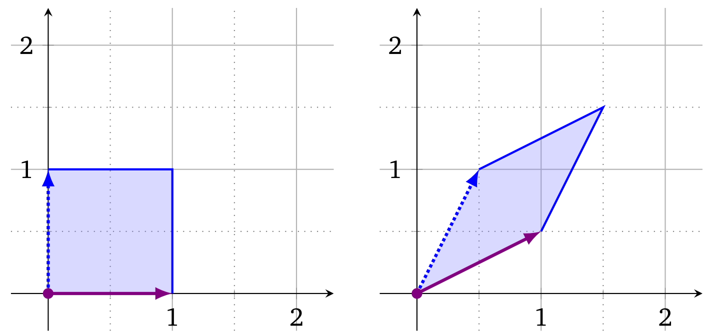
DEFINITION

Let X be a linear transformation or a matrix. 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 λ. We call λ the ***eigenvalue*** of X corresponding to the eigenvector $\vec{v}$.

83 The picture shows what the linear transformation T does to the unit square (i.e., the unit 2-cube).



83.1 Give an eigenvector for T. What is the eigenvalue?

83.2 Can you give another?

© Jason Siefken, 2015–2022

84 For some matrix A,

$$A\begin{bmatrix}3\\3\\1\end{bmatrix}=\begin{bmatrix}2\\2\\2/3\end{bmatrix}\qquad\text{and}\qquad B=A-\frac{2}{3}I.$$

84.1 Give an eigenvector and a corresponding eigenvalue for A.

84.2 What is $B\begin{bmatrix}3\\3\\1\end{bmatrix}$?

84.3 What is the dimension of $\operatorname{null}(B)$?

84.4 What is $\det(B)$?

© Jason Siefken, 2015–2022

85 Let $C = \begin{bmatrix} -1 & 2 \\ 1 & 0 \end{bmatrix}$ and $E_\lambda = C - \lambda I$.

85.1 For what values of λ does E_λ have a non-trivial null space?

85.2 What are the eigenvalues of C?

85.3 Find the eigenvectors of C.

© Jason Siefken, 2015–2022

DEFINITION

Characteristic Polynomial

For a matrix A, the ***characteristic polynomial*** of A is

$$\mathrm{char}(A) = \det(A - \lambda I).$$

---

86 Let $D = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$.

86.1 Compute $\mathrm{char}(D)$.

86.2 Find the eigenvalues of D.

© Jason Siefken, 2015–2022

87 Suppose $\text{char}(E) = -\lambda(2-\lambda)(-3-\lambda)$ for some unknown 3×3 matrix E.

87.1 What are the eigenvalues of E?

87.2 Is E invertible?

87.3 What can you say about $\text{nullity}(E)$, $\text{nullity}(E-3I)$, $\text{nullity}(E+3I)$?

© Jason Siefken, 2015–2022

Diagonalization

In this module you will learn

- How to diagonalize a matrix.
- When a matrix can and cannot be diagonalized.

Suppose $\mathcal{T}$ is a linear transformation and $\vec{v}_1$ and $\vec{v}_2$ are eigenvectors with eigenvalues λ_1 and λ_2. With this setup, for any $\vec{a} \in \operatorname{span}\{\vec{v}_1, \vec{v}_2\}$, we can compute $\mathcal{T}(\vec{a})$ with minimal effort.

Let's get specific. Define $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^2$ to be the linear transformation with matrix $M = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$. Let $\vec{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, and notice that $\vec{v}_1$ is an eigenvector for $\mathcal{T}$ with eigenvalue -1 and that $\vec{v}_2$ is an eigenvector for $\mathcal{T}$ with eigenvalue 4. Let $\vec{a} = \vec{v}_1 + \vec{v}_2$.

Now,

$$\mathcal{T}(\vec{a}) = \mathcal{T}(\vec{v}_1 + \vec{v}_2) = \mathcal{T}(\vec{v}_1) + \mathcal{T}(\vec{v}_2) = -\vec{v}_1 + 4\vec{v}_2.$$

We didn't need to refer to the entries of M to compute $\mathcal{T}(\vec{a})$.

Exploring further, let $\mathcal{V} = \{\vec{v}_1, \vec{v}_2\}$ and notice that $\mathcal{V}$ is a basis for $\mathbb{R}^2$. By definition $[\vec{a}]_{\mathcal{V}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and so we just computed

$$\mathcal{T}\begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{V}} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}_{\mathcal{V}}.$$

When represented in the $\mathcal{V}$ basis, computing $\mathcal{T}$ is easy. In general,

$$\mathcal{T}(\alpha\vec{v}_1 + \beta\vec{v}_2) = \alpha\mathcal{T}(\vec{v}_1) + \beta\mathcal{T}(\vec{v}_2) = -\alpha\vec{v}_1 + 4\beta\vec{v}_2,$$

and so

$$\mathcal{T}\begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{\mathcal{V}} = \begin{bmatrix} -\alpha \\ 4\beta \end{bmatrix}_{\mathcal{V}}.$$

In other words, $\mathcal{T}$, when acting on vectors written in the $\mathcal{V}$ basis, just multiplies each coordinate by an eigenvalue. This is enough information to determine the matrix for $\mathcal{T}$ in the $\mathcal{V}$ basis:

$$[\mathcal{T}]_{\mathcal{V}} = \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}.$$

The matrix representations $[\mathcal{T}]_{\mathcal{E}} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ and $[\mathcal{T}]_{\mathcal{V}} = \begin{bmatrix} -1 & 0 \\ 0 & 4 \end{bmatrix}$ are equally valid, but writing $\mathcal{T}$ in the $\mathcal{V}$ basis gives a very simple matrix!

Diagonalization

Recall that two matrices are similar if they represent the same transformation but in possibly different bases. The process of *diagonalizing* a matrix A is that of finding a diagonal matrix that is similar to A, and you can bet that this process is closely related to eigenvectors/values.

Let $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and suppose that $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ is a basis so that

$$[\mathcal{T}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{bmatrix}$$

is a diagonal matrix. This means that $\vec{b}_1, \ldots, \vec{b}_n$ are eigenvectors for $\mathcal{T}$! The proof goes as follows:

$$[\mathcal{T}]_{\mathcal{B}}[\vec{b}_1]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \alpha_1[\vec{b}_1]_{\mathcal{B}} = [\alpha_1\vec{b}_1]_{\mathcal{B}},$$

© Jason Siefken, 2015–2022

and in general

$$[\mathcal{T}]_{\mathcal{B}}[\vec{b}_i]_{\mathcal{B}} = \alpha_i[\vec{b}_i]_{\mathcal{B}} = [\alpha_i\vec{b}_i]_{\mathcal{B}}.$$

Therefore, for $i = 1, \dots, n$, we have

$$\mathcal{T}\vec{b}_i = \alpha_i\vec{b}_i.$$

Since $\mathcal{B}$ is a basis, $\vec{b}_i \neq \vec{0}$ for any i, and so each $\vec{b}_i$ is an eigenvector for $\mathcal{T}$ with corresponding eigenvalue α_i.

We've just shown that if a linear transformation $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ can be represented by a diagonal matrix, then there must be a basis for $\mathbb{R}^n$ consisting of eigenvectors for $\mathcal{T}$. The converse is also true.

Suppose again that $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation and that $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ is a basis of eigenvectors for $\mathcal{T}$ with corresponding eigenvalues $\alpha_1, \dots, \alpha_n$. By definition,

$$\mathcal{T}(\vec{b}_i) = \alpha_i\vec{b}_i,$$

and so

$$\mathcal{T}\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} \alpha_1 k_1 \\ \alpha_2 k_2 \\ \vdots \\ \alpha_n k_n \end{bmatrix}_{\mathcal{B}} \qquad \text{which is equivalent to} \qquad [\mathcal{T}]_{\mathcal{B}}\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} = \begin{bmatrix} \alpha_1 k_1 \\ \alpha_2 k_2 \\ \vdots \\ \alpha_n k_n \end{bmatrix}.$$

The only matrix that does this is

$$[\mathcal{T}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{bmatrix},$$

which is a diagonal matrix.

What we've shown is summarized by the following theorem.

> **Theorem.** A linear transformation $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ can be represented by a diagonal matrix if and only if there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors for $\mathcal{T}$. If $\mathcal{B}$ is such a basis, then $[\mathcal{T}]_{\mathcal{B}}$ is a diagonal matrix.

Now that we have a handle on representing a linear transformation by a diagonal matrix, let's tackle the problem of diagonalizing a matrix itself.

> **Diagonalizable.** A matrix is ***diagonalizable*** if it is similar to a diagonal matrix.

Suppose A is an $n \times n$ matrix. A induces some transformation $\mathcal{T}_A : \mathbb{R}^n \to \mathbb{R}^n$. By definition, this means $A = [\mathcal{T}_A]_{\mathcal{E}}$. The matrix B is similar to A if there is some basis $\mathcal{V}$ so that $B = [\mathcal{T}_A]_{\mathcal{V}}$. Using change-of-basis matrices, we see

$$A = [\mathcal{E} \leftarrow \mathcal{V}][\mathcal{T}_A]_{\mathcal{V}}[\mathcal{V} \leftarrow \mathcal{E}] = [\mathcal{E} \leftarrow \mathcal{V}]B[\mathcal{V} \leftarrow \mathcal{E}].$$

In other words, A and B are similar if there is some invertible change-of-basis matrix P so

$$A = PBP^{-1}.$$

Based on our earlier discussion, B will be a diagonal matrix if and only if P is the change-of-basis matrix for a basis of eigenvectors. In this case, we know B will be the diagonal matrix with eigenvalues along the diagonal (in the proper order).

> **Example.** Let $A = \begin{bmatrix} 1 & 2 & 5 \\ -11 & 14 & 5 \\ -3 & 2 & 9 \end{bmatrix}$ be a matrix and notice that $\vec{v}_1 = \begin{bmatrix} 5 \\ 5 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, and $\vec{v}_3 = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ are eigenvectors for A. Diagonalize A.
>
> First, we find the eigenvalues that correspond to the eigenvectors $\vec{v}_1, \vec{v}_2$, and $\vec{v}_3$. Computing,
>
> $$A\vec{v}_1 = \begin{bmatrix} 20 \\ 20 \\ 4 \end{bmatrix} = 4\vec{v}_1, \qquad A\vec{v}_2 = \begin{bmatrix} 8 \\ 8 \\ 8 \end{bmatrix} = 8\vec{v}_2, \qquad \text{and} \qquad A\vec{v}_3 = \begin{bmatrix} 12 \\ 36 \\ 12 \end{bmatrix} = 12\vec{v}_3,$$
>
> and so the eigenvalue corresponding to $\vec{v}_1$ is 4, to $\vec{v}_2$ is 8, and to $\vec{v}_3$ is 12.

© Jason Siefken, 2015–2022

The change-of-basis matrix which converts from the $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ to the standard basis is

$$P = \begin{bmatrix} 5 & 1 & 1 \\ 5 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix},$$

and

$$P^{-1} = \begin{bmatrix} \frac{1}{4} & 0 & -\frac{1}{4} \\ \frac{1}{4} & -\frac{1}{2} & \frac{5}{4} \\ -\frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}.$$

Define D to be the 3×3 matrix with the eigenvalues of A along the diagonal (in the order, $4, 8, 12$). That is, the matrix A written in the basis of eigenvectors is

$$D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 12 \end{bmatrix}.$$

We now know

$$A = PDP^{-1} = \begin{bmatrix} 5 & 1 & 1 \\ 5 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 12 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 & -\frac{1}{4} \\ \frac{1}{4} & -\frac{1}{2} & \frac{5}{4} \\ -\frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix},$$

and that D is the diagonalized form of A.

Non-diagonalizable Matrices

Is every matrix diagonalizable? Unfortunately the world is not that sweet. But, we have a tool to tell if a matrix is diagonalizable—checking to see if there is a basis of eigenvectors.

Example. Is the matrix $R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ diagonalizable?

Computing, $\mathrm{char}(R) = \lambda^2 + 1$ has no real roots. Therefore, R has no real eigenvalues. Consequently, R has no real eigenvectors, and so R is not diagonalizable.[a]

[a] If we allow complex eigenvalues, then R *is* diagonalizable and is similar to the matrix $\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$. So, to be more precise, we might say R is not *real* diagonalizable.

Example. Is the matrix $D = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$ diagonalizable?

For every vector $\vec{v} \in \mathbb{R}^2$, we have $D\vec{v} = 5\vec{v}$, and so every non-zero vector in $\mathbb{R}^2$ is an eigenvector for D. Thus, $\mathcal{E} = \{\vec{e}_1, \vec{e}_2\}$ is a basis of eigenvectors for $\mathbb{R}^2$, and so D is diagonalizable.[a]

[a] Of course, every square matrix is similar to itself and D is already diagonal, so of course it's diagonalizable.

Example. Is the matrix $J = \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix}$ diagonalizable?

Computing, $\mathrm{char}(J) = (5 - \lambda)^2$ which has a double root at 5. Therefore, 5 is the only eigenvalue of J. The eigenvectors of J all lie in

$$\mathrm{null}(J - 5I) = \mathrm{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}.$$

Since this is a one dimensional space, there is no basis for $\mathbb{R}^2$ consisting of eigenvectors for J. Therefore, J is not diagonalizable.

Example. Is the matrix $K = \begin{bmatrix} 5 & 1 \\ 0 & 2 \end{bmatrix}$ diagonalizable?

Computing, $\mathrm{char}(K) = (5 - \lambda)(2 - \lambda)$ which has roots at 5 and 2. Therefore, 5 and 2 are the eigenvalues of

© Jason Siefken, 2015–2022

K. The eigenvectors of K lie in one of

$$\operatorname{null}(K-5I)=\operatorname{span}\left\{\begin{bmatrix}1\\0\end{bmatrix}\right\} \qquad \text{or} \qquad \operatorname{null}(K-2I)=\operatorname{span}\left\{\begin{bmatrix}-1\\3\end{bmatrix}\right\}.$$

Picking one eigenvector from each null space, we have that $\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}-1\\3\end{bmatrix}\right\}$ is a basis for $\mathbb{R}^2$ consisting of eigenvectors of K. Thus, K is diagonalizable.

Takeaway. Not all matrices are diagonalizable, but you can check if an $n\times n$ matrix is diagonalizable by determining whether there is a basis of eigenvectors for $\mathbb{R}^n$.

Geometric and Algebraic Multiplicities

When analyzing linear transformations or matrices, we're often interested in studying the subspaces where vectors are stretched by only one eigenvalue. These are called the *eigenspaces*.

Eigenspace. Let A be an $n\times n$ matrix with eigenvalues $\lambda_1,\ldots,\lambda_m$. The ***eigenspace*** of A corresponding to the eigenvalue λ_i is the null space of $A-\lambda_i I$. That is, it is the space spanned by all eigenvectors that have the eigenvalue λ_i.
The ***geometric multiplicity*** of an eigenvalue λ_i is the dimension of the corresponding eigenspace. The ***algebraic multiplicity*** of λ_i is the number of times λ_i occurs as a root of the characteristic polynomial of A (i.e., the number of times $x-\lambda_i$ occurs as a factor).

Now is the time when linear algebra and regular algebra (the solving of non-linear equations) combine. We know, every root of the characteristic polynomial of a matrix gives an eigenvalue for that matrix. Since the degree of the characteristic polynomial of an $n\times n$ matrix is always n, the fundamental theorem of algebra tells us exactly how many roots to expect.

Recall that the *multiplicity* of a root of a polynomial is the power of that root in the factored polynomial. So, for example $p(x)=(4-x)^3(5-x)$ has a root of 4 with multiplicity 3 and a root of 5 with multiplicity 1.

Example. Let $R=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$ and find the geometric and algebraic multiplicity of each eigenvalue of R.

Computing, $\operatorname{char}(R)=\lambda^2+1$ which has no real roots. Therefore, R has no real eigenvalues.[a]

[a] If we allow complex eigenvalues, then the eigenvalues i and $-i$ both have geometric and algebraic multiplicity of 1.

Example. Let $D=\begin{bmatrix}5&0\\0&5\end{bmatrix}$ and find the geometric and algebraic multiplicity of each eigenvalue of D.

Computing, $\operatorname{char}(D)=(5-\lambda)^2$, so 5 is an eigenvalue of D with algebraic multiplicity 2. The eigenspace of D corresponding to 5 is $\mathbb{R}^2$. Thus, the geometric multiplicity of 5 is 2.

Example. Let $J=\begin{bmatrix}5&1\\0&5\end{bmatrix}$ and find the geometric and algebraic multiplicity of each eigenvalue of J.

Computing, $\operatorname{char}(J)=(5-\lambda)^2$, so 5 is an eigenvalue of J with algebraic multiplicity 2. The eigenspace of J corresponding to 5 is $\operatorname{span}\left\{\begin{bmatrix}1\\0\end{bmatrix}\right\}$. Thus, the geometric multiplicity of 5 is 1.

Example. Let $K=\begin{bmatrix}5&1\\0&2\end{bmatrix}$ and find the geometric and algebraic multiplicity of each eigenvalue of K.

Computing, $\operatorname{char}(K)=(5-\lambda)(2-\lambda)$, so 5 and 2 are eigenvalues of K, both with algebraic multiplicity 1. The eigenspace of K corresponding to 5 is $\operatorname{span}\left\{\begin{bmatrix}1\\0\end{bmatrix}\right\}$ and the eigenspace corresponding to 2 is $\operatorname{span}\left\{\begin{bmatrix}-1\\3\end{bmatrix}\right\}$. Thus, both 5 and 2 have a geometric multiplicity of 1.

Consider the following two theorems.

© Jason Siefken, 2015–2022

Theorem (Fundamental Theorem of Algebra). Let p be a polynomial of degree n. Then, if complex roots are allowed, the sum of the multiplicities of the roots of p is n.

Theorem. Let λ be an eigenvalue of the matrix A. Then

$$\text{geometric mult}(\lambda) \leq \text{algebraic mult}(\lambda).$$

We can now deduce the following.

Theorem. An $n \times n$ matrix A is diagonalizable if and only if the sum of its geometric multiplicities is equal to n. Further, provided complex eigenvalues are permitted, A is diagonalizable if and only if its geometric multiplicities are equal to its corresponding algebraic multiplicities.

Proof. Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_k$. Let $E_1, \ldots, E_k$ be bases for the eigenspaces corresponding to $\lambda_1, \ldots, \lambda_k$. We will start by showing $E = E_1 \cup \cdots \cup E_k$ is a linearly independent set using the following two lemmas.

No New Eigenvalue Lemma. Suppose that $\vec{v}_1, \ldots, \vec{v}_k$ are linearly independent eigenvectors of a matrix A, and let $\lambda_1, \ldots, \lambda_k$ be the corresponding eigenvalues. Then, any eigenvector for A contained in $\text{span}\{\vec{v}_1, \ldots, \vec{v}_k\}$ must have one of $\lambda_1, \ldots, \lambda_k$ as its eigenvalue.

The proof goes as follows. Suppose $\vec{v} = \sum_{i \leq k} \alpha_i \vec{v}_i$ is an eigenvector for A with eigenvalue λ. We now compute $A\vec{v}$ in two different ways: once by using the fact that $\vec{v}$ is an eigenvector, and again by using the fact that $\vec{v}$ is a linear combination of other eigenvectors. Observe

$$A\vec{v} = \lambda\vec{v} = \lambda\left(\sum_{i\leq k} \alpha_i \vec{v}_i\right) = \sum_{i \leq k} \alpha_i \lambda \vec{v}_i$$

and

$$A\vec{v} = A\left(\sum_{i\leq k} \alpha_i \vec{v}_i\right) = \sum_{i \leq k} \alpha_i A\vec{v}_i = \sum_{i\leq k} \alpha_i \lambda_i \vec{v}_i.$$

We now have

$$\vec{0} = A\vec{v} - A\vec{v} = \sum_{i \leq k} \alpha_i \lambda \vec{v}_i - \sum_{i\leq k} \alpha_i \lambda_i \vec{v}_i = \sum_{i \leq k} \alpha_i (\lambda - \lambda_i)\vec{v}_i.$$

Because $\vec{v}_1, \ldots, \vec{v}_k$ are linearly independent, we know $\alpha_i(\lambda - \lambda_i) = 0$ for all $i \leq k$. Further, because $\vec{v}$ is non-zero (it's an eigenvector), we know at least one α_i is non-zero. Therefore $\lambda - \lambda_i = 0$ for at least one i. In other words, $\lambda = \lambda_i$ for at least one i, which is what we set out to show.[55]

Basis Extension Lemma. Let $P = \{\vec{p}_1, \ldots, \vec{p}_a\}$ and $Q = \{\vec{q}_1, \ldots, \vec{q}_b\}$ be linearly independent sets, and suppose $P \cup \{\vec{q}\}$ is linearly independent for all non-zero $\vec{q} \in \text{span}\, Q$. Then $P \cup Q$ is linearly independent.

To show this, suppose $\vec{0} = \alpha_1\vec{p}_1 + \cdots + \alpha_a\vec{p}_a + \beta_1\vec{q}_1 + \cdots + \beta_b\vec{q}_b$ is a linear combination of vectors in $P \cup Q$. Let $\vec{q} = \beta_1\vec{q}_1 + \cdots + \beta_b\vec{q}_b$. First, note that $\vec{q}$ must be the zero vector. If not, $\vec{0} = \alpha_1\vec{p}_1 + \cdots + \alpha_a\vec{p}_a + \vec{q}$ is a non-trivial linear combination of vectors in $P \cup \{\vec{q}\}$, which contradicts the assumption that $P \cup \{\vec{q}\}$ is linearly independent. Since we've established $\vec{0} = \vec{q} = \beta_1\vec{q}_1 + \cdots + \beta_b\vec{q}_b$, we conclude $\beta_1 = \cdots = \beta_b = 0$ because Q is linearly independent. It follows that since $\vec{0} = \alpha_1\vec{p}_1 + \cdots + \alpha_a\vec{p}_a + \vec{q} = \alpha_1\vec{p}_1 + \cdots + \alpha_a\vec{p}_a + \vec{0}$, we must have that $\alpha_1 = \cdots = \alpha_a = 0$ because P is linearly independent. This shows that the only way to express $\vec{0}$ as a linear combination of vectors in $P \cup Q$ is as the trivial linear combination, and so $P \cup Q$ is linearly independent.

Now we can put our lemmas to good use. We will use induction to show that $E = E_1 \cup \cdots \cup E_k$ is linearly independent. By assumption E_1 is linearly independent. Now, suppose $U = E_1 \cup \cdots \cup E_j$ is linearly independent. By construction, every non-zero vector $\vec{v} \in \text{span}\, E_{j+1}$ is an eigenvector for A with eigenvalue λ_{j+1}. Therefore, since $\lambda_{j+1} \neq \lambda_i$ for $1 \leq i \leq j$, we may apply the *No New Eigenvalue Lemma* to see that $\vec{v} \notin \text{span}\, U$. It follows that $U \cup \{\vec{v}\}$ is linearly independent. Since E_{j+1} is itself linearly independent, we may now apply the *Basis Extension Lemma* to deduce that $U \cup E_{j+1}$ is linearly independent. This shows that $E = E_1 \cup \cdots \cup E_k$ is linearly independent.

To conclude notice that by construction, geometric mult$(\lambda_i) = |E_i|$. Since $E = E_1 \cup \cdots \cup E_k$ is linearly independent, the E_i's must be disjoint and so $\sum \text{geometric mult}(\lambda_i) = \sum |E_i| = |E|$. If $\sum \text{geometric mult}(\lambda_i) = n$, then $E \subseteq \mathbb{R}^n$ is a linearly independent set of n vectors and so is a basis for $\mathbb{R}^n$. Finally, because we have a basis for $\mathbb{R}^n$ consisting of eigenvectors for A, we know A is diagonalizable.

[55] You may notice that we've proved something stronger than we needed: if an eigenvector is a linear combination of linearly independent eigenvectors, the only non-zero coefficients of that linear combination must belong to eigenvectors with the same eigenvalue.

© Jason Siefken, 2015–2022

Conversely, if there is a basis E for $\mathbb{R}^n$ consisting of eigenvectors, we must have a linearly independent set of n eigenvectors. Grouping these eigenvectors by eigenvalue, an application of the *No New Eigenvalue Lemma* shows that each group must actually be a basis for its eigenspace. Thus, the sum of the geometric multiplicities must be n.

Finally, if complex eigenvalues are allowed, the algebraic multiplicities sum to n. Since the algebraic multiplicities bound the geometric multiplicities, the only way for the geometric multiplicities to sum to n is if corresponding geometric and algebraic multiplicities are equal.

■

Practice Problems

1 For each of the matrices below, find the geometric and algebraic multiplicity of each eigenvalue.

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

(b) $B = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$

(c) $C = \begin{bmatrix} 3 & 0 \\ 3 & 0 \end{bmatrix}$

(d) $D = \begin{bmatrix} 0 & 3/2 & 4 \\ 0 & 1 & 0 \\ -1 & 1 & 4 \end{bmatrix}$

(e) $E = \begin{bmatrix} 2 & 1/2 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 2 \end{bmatrix}$

2 For each matrix from question 1, diagonalize the matrix if possible. Otherwise explain why the matrix cannot be diagonalized.

3 Give an example of a 4×4 matrix with 2 and 7 as its only eigenvalues.

4 Can the geometric multiplicity of an eigenvalue ever be 0? Explain.

5 (a) Show that if $\vec{v}_1$ and $\vec{v}_2$ are eigenvectors for a matrix M corresponding to different eigenvalues, then $\vec{v}_1$ and $\vec{v}_2$ are linearly independent.

(b) If possible, give an example of a non-diagonalizable 3×3 matrix where 1 and -1 are the only eigenvalues.

(c) If possible, give an example of a non-diagonalizable 2×2 matrix where 1 and -1 are the only eigenvalues.

© Jason Siefken, 2015–2022

88 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.

88.1 Find the eigenvalues of T_A.

88.2 Find the characteristic polynomial of T_A.

88.3 Compute $T_A\vec{w}$ where $\vec{w} = 2\vec{v}_1 - \vec{v}_2$.

88.4 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.

Notice that $\mathcal{V} = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is a basis for $\mathbb{R}^3$.

88.5 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.

© Jason Siefken, 2015–2022

89 Recall from Problem 88 that

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \qquad \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \qquad \vec{v}_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

and $\mathcal{V} = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. Let T_A be the transformation induced by A and let $P = [\vec{v}_1|\vec{v}_2|\vec{v}_3]$ be the matrix with columns $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$ (written in the standard basis).

89.1 Describe in words what P and P^{-1} do in terms of change-of-basis.

89.2 If you were asked to compute $T_A\vec{y}$ for some $\vec{y} \in \mathbb{R}^3$, which basis would you prefer to do your computations in? Explain.

89.3 Given a vector $\vec{y} \in \mathbb{R}^3$ written in the standard basis, is there a way to compute $T_A\vec{y}$ without using the matrix A? (You may use P and P^{-1}, just not A.) Explain.

89.4 Can you find a matrix D so that

$$PDP^{-1} = A?$$

89.5 $[\vec{x}]_{\mathcal{V}} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$. Compute $T_A^{100}\vec{x}$. Express your answer in both the $\mathcal{V}$ basis and the standard basis.

© Jason Siefken, 2015–2022

Diagonalizable

DEF A matrix is ***diagonalizable*** if it is similar to a diagonal matrix.

---

90 Let B be an $n \times n$ matrix and let T_B be the induced transformation. Suppose T_B has eigenvectors $\vec{v}_1, \ldots, \vec{v}_n$ which form a basis for $\mathbb{R}^n$, and let $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues.

90.1 How do the eigenvalues and eigenvectors of B and T_B relate?

90.2 Is B diagonalizable (i.e., similar to a diagonal matrix)? If so, explain how to obtain its diagonalized form.

90.3 What if one of the eigenvalues of T_B is zero? Would B be diagonalizable?

90.4 What if the eigenvectors of T_B did not form a basis for $\mathbb{R}^n$. Would B be diagonalizable?

© Jason Siefken, 2015–2022

Eigenspace

DEFINITION

Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \dots, \lambda_m$. The ***eigenspace*** of A corresponding to the eigenvalue λ_i is the null space of $A - \lambda_i I$. That is, it is the space spanned by all eigenvectors that have the eigenvalue λ_i.

The ***geometric multiplicity*** of an eigenvalue λ_i is the dimension of the corresponding eigenspace. The ***algebraic multiplicity*** of λ_i is the number of times λ_i occurs as a root of the characteristic polynomial of A (i.e., the number of times $x - \lambda_i$ occurs as a factor).

---

91 Let $F = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ and $G = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$.

91.1 Is F diagonalizable? Why or why not?

91.2 Is G diagonalizable? Why or why not?

91.3 What are the geometric and algebraic multiplicities of each eigenvalue of F? What about the multiplicities for each eigenvalue of G?

91.4 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?

© Jason Siefken, 2015–2022

Systems of Linear Equations I

In this appendix you will learn

- What a system of linear equations is.
- What the solution set to a system of equations is, and what it means for a system of equations to be consistent or inconsistent.
- How augmented matrices can be used to solve systems of linear equations.
- How to apply row reduction to find a unique solution to a system of linear equations and to determine if a system of linear equations is consistent or inconsistent.

An *equation* encodes a relationship between quantities. For example, writing

$$\underbrace{\text{Slices of cake}}_{C} = \underbrace{\text{Slices you ate}}_{M} + \underbrace{\text{Slices your brother ate}}_{B}$$

specifies a precise relationship between the quantities C, M, and B. Without more information, C, M, and B could be almost anything. As such, we call C, M, and B *variables* or *unknowns*. However, the relationship *between* them is precisely defined.

Additional relationships give rise to additional equations, which we express concisely as a *system of equations*, that is, a list of equations. For example, suppose you know the cake had six pieces and your brother ate twice as many pieces as you. We might now write the system

$$\begin{aligned} C &= M + B \\ B &= 2M \\ C &= 6 \end{aligned}$$

which should be interpreted as: "the relationship $C = M + B$ holds *and* the relationship $B = 2M$ holds *and* the relationship $C = 6$ holds." All this information, taken together, is enough to deduce the unknowns: $M = 2$, $B = 4$, and $C = 6$.

Systems of equations naturally appear in linear algebra through vector equations. Suppose $\vec{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. You might wonder if $\vec{w}$ was a linear combination of $\vec{u}$ and $\vec{v}$. The answer is *yes* if and only if the vector equation

$$\vec{w} = a\vec{u} + b\vec{v}$$

has a solution for some a and b. Written in coordinates, this equation is equivalent to

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

Equating coordinates, a system of equations appears:

$$\begin{cases} a + 2b = 1 \\ 2a + 3b = 1 \end{cases}$$

Every vector equation, by way of coordinates, corresponds to a system of equations. And, fortunately for us, there is an *algorithm* to find all solutions to these systems.[56]

Systems of Linear Equations

There's no guarantee that a general equation, like $x^4 - e^x + 7 = 0$, has a solution, and it might be impossible to decide if an arbitrary equation has a solution, let alone what the solutions are![57] However, for *linear* equations and systems of linear equations we can *always* tell whether there is a solution and what the solution(s) are.

[56] Saying there is an *algorithm* for "X" means that there is a specific set of (non-random) rules that *always* accomplishes "X". As a consequence, doing "X" never requires special insight. For example, there *is* an algorithm for multiplying numbers, but there *is not* an algorithm for factoring polynomials of degree 5 or greater.

[57] Fermat's Last Theorem famously claimed that $a^n + b^n = c^n$ has no positive integer solutions for $n \geq 3$. However, it took 350 years of human ingenuity before anyone was able rigorously back up the claim.

© Jason Siefken, 2015–2022

Linear Equation. A ***linear equation*** in the variables $x_1, \dots, x_n$ is one that can be expressed as

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = c$$

for constants $a_1, \dots, a_n$ and c. A ***system of linear equations*** is a system of equations consisting of one or more linear equations.

Every vector equation corresponds to an *equivalent* system of linear equations and vice versa, where equivalent means "expresses the same relationships between variables".

Example. Write down the vector equation corresponding to the system of linear equations $\begin{cases} 2x+3y+z=2 \\ y-z=-1 \end{cases}$ and the system of linear equations corresponding to the vector equation $t\vec{w}+\vec{u}=r\vec{v}$ where $\vec{w}=\begin{bmatrix}1\\-1\end{bmatrix}$, $\vec{u}=\begin{bmatrix}2\\3\end{bmatrix}$, and $\vec{v}=\begin{bmatrix}4\\4\end{bmatrix}$.

The system $\begin{cases} 2x+3y+z=2 \\ 0x+y-z=-1 \end{cases}$ corresponds to the vector equation

$$x\begin{bmatrix}2\\0\end{bmatrix}+y\begin{bmatrix}3\\1\end{bmatrix}+z\begin{bmatrix}1\\-1\end{bmatrix}=\begin{bmatrix}2\\-1\end{bmatrix}.$$

As for the vector equation $t\vec{w}+\vec{u}=r\vec{v}$, rewriting each vector in coordinates gives us a corresponding system of linear equations:

$$t\vec{w}+\vec{u}=r\vec{v} \to t\begin{bmatrix}1\\-1\end{bmatrix}+\begin{bmatrix}2\\3\end{bmatrix}=r\begin{bmatrix}4\\4\end{bmatrix}$$
$$\begin{bmatrix}t+2\\-t+3\end{bmatrix}=\begin{bmatrix}4r\\4r\end{bmatrix}\to\begin{cases}-4r+t=-2\\-4r-t=-3\end{cases}.$$

Takeaway. Every vector equation corresponds to a system of linear equations and every system of linear equations corresponds to a vector equation.

Solution Sets

Before looking at how to solve systems of linear equations, let's agree on some terminology.

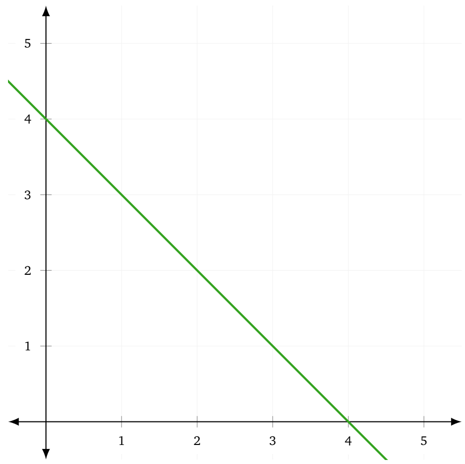
A *solution* to an equation is a particular choice of values for the variables that satisfy (i.e. make true) the equation. For example

$$x+y=4 \tag{13}$$

has a solution $x=y=2$. However, $x=y=2$ is just one of *many* possible solutions; we also have $x=4$ and $y=0$ or $x=-2$ and $y=6$. The *solution set*, also called the *complete solution*, to an equation (or system of equations) is the set of all possible solutions. For example, the solution set to Equation (13) is $S=\{(x,y) : y=4-x\}$. In this case, S contains infinitely many solutions, including $(x,y)=(2,2)$, the first solution we found.

Solution sets look a lot like sets of vectors: the set $S=\{(x,y) : y=4-x\}$ could be thought of as a subset of $\mathbb{R}^2$ where we identify a solution $x=a$ and $y=b$ with the column vector $\begin{bmatrix}a\\b\end{bmatrix}$. Drawing S as a subset of $\mathbb{R}^2$, we see a familiar picture.

© Jason Siefken, 2015–2022



It's the graph of the line given in $y = mx + b$ form by $y = -x + 4$. In other words, *via solution sets, equations and systems of equations can represent geometric objects.*

Consistent & Inconsistent Systems

Consider the following equations (as separate equations, not as a system):

$$x^2 = 0 \qquad \text{with solution set} \qquad S_x \subseteq \mathbb{R},$$

$$y^2 = 4 \qquad \text{with solution set} \qquad S_y \subseteq \mathbb{R},$$

and

$$z^2 = -1 \qquad \text{with solution set} \qquad S_z \subseteq \mathbb{R}.$$

$S_x = \{0\}$ consists of a single number. $S_y = \{2, -2\}$ consists of two numbers, and $S_z = \{\}$ consists of no numbers.[58] In this case, we would call the first two equations *consistent* and the third equation *inconsistent*.

> **Consistent & Inconsistent.** An equation or system of equations is called ***consistent*** if it has at least one solution. That is, its solution set is non-empty. Otherwise, an equation or system of equations is called ***inconsistent***.

Why the word *consistent*? This comes from the term *logically consistent* which means "able to be true". An equation is an assertion that the left hand side equals the right hand side. If that can never happen, the assertion is not logically consistent.

This terminology becomes more clear with systems. Consider the system

$$\begin{cases} x - y = 0 \\ x - y = 1 \end{cases}.$$

From the first equation, we deduce $y = x$. From the second equation, we deduce $x = 1 + y$. Since $x = x$, we know that $y = x = 1 + y$ and therefore $y = 1 + y$. However, this is never true! There is a logical inconsistency between the two equations. In isolation they're fine, but taken together they're not.

Equivalent Systems

Two systems of equations are logically equivalent if they express the same relationships between their variables. For example, the equations $x = 2y$ and $\frac{1}{2}x = y$ express the exact same relationship between the variables x and y. This can be formalized using solution sets.

> **Equivalent Systems.** Two equations or systems of equations are called ***equivalent*** if they have the same solution sets.

Again, $x = 2y$ and $\frac{1}{2}x = y$ both have the same solution set (a line through the origin of slope $\frac{1}{2}$), and so they are equivalent.

Philosophical note: the process of "doing algebra" can be viewed as the process of *manipulating equations/systems into easier to understand equivalent equations/systems*. When you're asked to algebraically solve $x^2 - 4 = 0$.

[58] We're not allowing complex numbers at the moment.

© Jason Siefken, 2015–2022

You might first factor to get the equivalent equation $(x-2)(x+2)=0$. Then, since non-zero numbers cannot multiply to give zero, we know $x-2=0$ or $x+2=0$, which in turn is equivalent to $x=\pm 2$. It's always been about equivalent systems![59]

Row Reduction

Consider the vector equation

$$t\vec{u}+s\vec{v}+r\vec{w}=\vec{p} \qquad \text{where} \qquad \vec{u}=\begin{bmatrix}1\\2\\1\end{bmatrix},\ \vec{v}=\begin{bmatrix}2\\1\\-4\end{bmatrix},\ \vec{w}=\begin{bmatrix}-2\\-5\\1\end{bmatrix},\ \vec{p}=\begin{bmatrix}-15\\-21\\18\end{bmatrix}.$$

By expanding in terms of coordinates, we get an equivalent system of linear equations.

$$\begin{cases} t+2s-2r=-15 & \text{row}_1\\ 2t+s-5r=-21 & \text{row}_2\\ t-4s+r=18 & \text{row}_3 \end{cases} \tag{14}$$

The most general way to solve any system is by *substitution*. For System (14), we could solve the first equation for t, substitute the result in the remaining equations, solve the next equation for s, etc.. However, because System (14) is a *linear* system, there's an alternate method: *row reduction*.[60]

Study the following manipulations of System (14) and convince yourself that each operation produces a system equivalent to the one it came from.

$$\begin{cases} t+2s-2r=-15\\ 2t+s-5r=-21\\ t-4s+r=18 \end{cases} \xrightarrow{\text{row}_3\mapsto\text{row}_3-\text{row}_1} \begin{cases} t+2s-2r=-15\\ 2t+s-5r=-21\\ -6s+3r=33 \end{cases}$$

$$\xrightarrow{\text{row}_2\mapsto\text{row}_2-2\text{row}_1} \begin{cases} t+2s-2r=-15\\ -3s-r=9\\ -6s+3r=33 \end{cases}$$

$$\xrightarrow{\text{row}_3\mapsto\text{row}_3-2\text{row}_2} \begin{cases} t+2s-2r=-15\\ -3s-r=9\\ 5r=15 \end{cases} \tag{15}$$

From the final system, System (15), it's easy to see that $r=3$. From there, we can substitute $r=3$ into the second row of System (15) to find $s=-4$ and we can substitute both r and s into the first row of System (15) to find $t=-1$.

By adding and subtracting rows, we "reduced" the number of variables from some equations until they were easy to solve. As an added benefit, every system along the way to System (15) was nicely organized and formatted. In fact, the systems are so well organized that we can save time by not writing the variables and keeping track of the numbers in an *augmented matrix*.[61] That is, instead of writing

$$\begin{cases} t+2s-2r=-15\\ 2t+s-5r=-21\\ t-4s+r=18 \end{cases}$$

we will write

$$\left[\begin{array}{ccc|c} 1 & 2 & -2 & -15\\ 2 & 1 & -5 & -21\\ 1 & -4 & 1 & 18 \end{array}\right].$$

We call the matrix an *augmented matrix* to stress that it contains two types of information: the *coefficients* of the variables t, s, and r and the *constants* on the right hand side of the equations. An (optional) vertical line separates the two types of numbers.

[59] Technically, up to this point we've been talking about *conjunctive* systems, which means that a solution must hold for all equations of a system. The system $x=\pm 2$ is a *disjunctive* system, which means a solution only needs to hold for *one* of the equations ($x=2$ or $x=-2$), but the idea is the same.

[60] Row reduction is sometimes referred to as *Gaussian elimination*, *Gauss-Jordan elimination*, or just *elimination*; though there are subtle differences between Gaussian and Gauss-Jordan elimination, they aren't important, and we'll refer to all similar methods as *row reduction*.

[61] A *matrix* is a box of numbers. An *augmented matrix* is a matrix with extra information associated with it.

© Jason Siefken, 2015–2022

Now, the process of row reduction looks like this:

$$\begin{cases} t+2s-2r=-15 \\ 2t+\ s-5r=-21 \\ t-4s+\ r=18 \end{cases} \to \left[\begin{array}{ccc|c} 1 & 2 & -2 & -15 \\ 2 & 1 & -5 & -21 \\ 1 & -4 & 1 & 18 \end{array}\right] \xrightarrow{\text{row}_3 \mapsto \text{row}_3 - \text{row}_1} \left[\begin{array}{ccc|c} 1 & 2 & -2 & -15 \\ 2 & 1 & -5 & -21 \\ 0 & -6 & 3 & 33 \end{array}\right]$$

$$\xrightarrow{\text{row}_2 \mapsto \text{row}_2 - 2\text{row}_1} \left[\begin{array}{ccc|c} 1 & 2 & -2 & -15 \\ 0 & -3 & -1 & 9 \\ 0 & -6 & 3 & 33 \end{array}\right]$$

$$\xrightarrow{\text{row}_3 \mapsto \text{row}_3 - 2\text{row}_2} \left[\begin{array}{ccc|c} 1 & 2 & -2 & -15 \\ 0 & -3 & -1 & 9 \\ 0 & 0 & 5 & 15 \end{array}\right] \to \begin{cases} t+2s-2r=-15 \\ -3s-\ r=9 \\ 5r=15 \end{cases}$$

The operations are identical, but we write augmented matrices instead of equations.

Takeaway. Augmented matrices are a notational tool that makes the process of doing row reduction more efficient.

The Rules of Row Reduction

So far, we've been able to row reduce systems by adding a multiple of one row to another,[62] but to fully solve *any* system, we need additional operations.[63]

Elementary Row Operations. The three ***elementary row operations***, which can be performed on a matrix or system of equations, are

- swapping two rows (written $\text{row}_i \leftrightarrow \text{row}_j$),
- multiplying a row by a non-zero scalar (written $\text{row}_i \mapsto k\,\text{row}_i$), and
- adding a multiple of one row to another (written $\text{row}_i \mapsto \text{row}_i + k\,\text{row}_j$).

Notice that each elementary row operation can be undone. For example, if you perform $\text{row}_i \mapsto k\,\text{row}_i$, you can undo it with $\text{row}_i \mapsto \frac{1}{k}\text{row}_i$. Therefore, applying an elementary row operation to a system is guaranteed to produce an equivalent system.

The strategy for solving a system is now summarized as:

1. Rewrite the system as an augmented matrix.
2. Use elementary row operations to zero-out the lower triangle of the augmented matrix.
3. Convert the matrix back to a system of equations.
4. Read off the solution (substituting when necessary).

Example. Find a solution to the following system:

$$\begin{cases} a+3b+2c=1 \\ 2a+7b+5c=2 \\ -a-4b\qquad=11 \end{cases}.$$

To do so, we rewrite the system as an augmented matrix then row reduce.

$$\begin{cases} a+3b+2c=1 \\ 2a+7b+5c=2 \\ -a-4b\qquad=11 \end{cases} \to \left[\begin{array}{ccc|c} 1 & 3 & 2 & 1 \\ 2 & 7 & 5 & 2 \\ -1 & -4 & 0 & 11 \end{array}\right] \xrightarrow{\text{row}_2 \mapsto \text{row}_2 - 2\text{row}_1} \left[\begin{array}{ccc|c} 1 & 3 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ -1 & -4 & 0 & 11 \end{array}\right]$$

$$\xrightarrow{\text{row}_3 \mapsto \text{row}_3 + \text{row}_1 + \text{row}_2} \left[\begin{array}{ccc|c} 1 & 3 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 3 & 12 \end{array}\right] \to \begin{cases} a+3b+2c=1 \\ b+\ c=0 \\ 3c=12 \end{cases}$$

[62] Technically, we subtracted, but that's just adding a negative.

[63] If you're clever, you can think up alternatives to the elementary row operations that work just as well, but there's good reason to favor the three elementary row operations. We'll see them when discussing matrix decompositions.

© Jason Siefken, 2015–2022

The third row reveals that $c = 4$; substituting into the second row, we find $b = -4$. Now we can substitute $b = -4$ and $c = 4$ into the first row and we obtain $a = 5$.

Thus, the solution is $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 5 \\ -4 \\ 4 \end{bmatrix}$. Since this is the only solution to the system, the solution set is $\left\{ \begin{bmatrix} 5 \\ -4 \\ 4 \end{bmatrix} \right\}$.

Example. Solve the system

$$\begin{cases} 3t + s + 13r = -2 \\ t \quad\;\; + 5r = 1 \\ -t + s - 6r = -8 \\ t + s + 4r = -6 \end{cases}.$$

Again, we row reduce the corresponding augmented matrix to find an equivalent system from which we can more easily compute the solution.

$$\begin{cases} 3t + s + 13r = -2 \\ t \quad\;\; + 5r = 1 \\ -t + s - 6r = -8 \\ t + s + 4r = -6 \end{cases} \to \left[\begin{array}{rrr|r} 3 & 1 & 13 & -2 \\ 1 & 0 & 5 & 1 \\ -1 & 1 & -6 & -8 \\ 1 & 1 & 4 & -6 \end{array}\right] \xrightarrow{\text{row}_1 \leftrightarrow \text{row}_2} \left[\begin{array}{rrr|r} 1 & 0 & 5 & 1 \\ 3 & 1 & 13 & -2 \\ -1 & 1 & -6 & -8 \\ 1 & 1 & 4 & -6 \end{array}\right]$$

$$\xrightarrow{\text{row}_2 \mapsto \text{row}_2 - 3\text{row}_1} \left[\begin{array}{rrr|r} 1 & 0 & 5 & 1 \\ 0 & 1 & -2 & -5 \\ -1 & 1 & -6 & -8 \\ 1 & 1 & 4 & -6 \end{array}\right]$$

$$\xrightarrow{\text{row}_3 \mapsto \text{row}_3 + \text{row}_1 - \text{row}_2} \left[\begin{array}{rrr|r} 1 & 0 & 5 & 1 \\ 0 & 1 & -2 & -5 \\ 0 & 0 & 1 & -2 \\ 1 & 1 & 4 & -6 \end{array}\right]$$

$$\xrightarrow{\text{row}_4 \mapsto \text{row}_4 - \text{row}_1 - \text{row}_2} \left[\begin{array}{rrr|r} 1 & 0 & 5 & 1 \\ 0 & 1 & -2 & -5 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 1 & -2 \end{array}\right]$$

$$\xrightarrow{\text{row}_4 \mapsto \text{row}_4 - \text{row}_3} \left[\begin{array}{rrr|r} 1 & 0 & 5 & 1 \\ 0 & 1 & -2 & -5 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right] \to \begin{cases} t \quad + 5r = 1 \\ \quad s - 2r = -5 \\ \qquad\quad r = -2 \\ \qquad\quad 0 = 0 \end{cases}$$

Our equivalent system reveals $r = -2$, which we can substitute back into the first and second rows to find that $t = 11$ and $s = -9$.

As a vector, the solution is $\begin{bmatrix} t \\ s \\ r \end{bmatrix} = \begin{bmatrix} 11 \\ -9 \\ -2 \end{bmatrix}$ and so the solution set is $\left\{ \begin{bmatrix} 11 \\ -9 \\ -2 \end{bmatrix} \right\}$.

In the examples so far, we've stopped row reducing when the equations became simple enough to solve by inspection. However, we could continue row reducing until the system is as simple as possible.

Example. Solve the system

$$\begin{cases} a + 3b + 2c = 1 \\ 2a + 7b + 5c = 2 \\ -a - 4b \qquad = 11 \end{cases}$$

Notice that we solved this system using a combination of row reduction and substitution in a previous example. This time, let us use only row reduction.
The augmented matrix for this system is

$$\left[\begin{array}{rrr|r} 1 & 3 & 2 & 1 \\ 2 & 7 & 5 & 2 \\ -1 & -4 & 0 & 11 \end{array}\right].$$

© Jason Siefken, 2015–2022

Based on the work from the previous example, we know it can be reduced to

$$\left[\begin{array}{ccc|c} 1 & 3 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 3 & 12 \end{array}\right].$$

Now let us continue row reducing.

$$\left[\begin{array}{ccc|c} 1 & 3 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 3 & 12 \end{array}\right] \xrightarrow{\text{row}_3 \mapsto \frac{1}{3}\text{row}_3} \left[\begin{array}{ccc|c} 1 & 3 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

$$\xrightarrow{\text{row}_2 \mapsto \text{row}_2 - \text{row}_3} \left[\begin{array}{ccc|c} 1 & 3 & 2 & 1 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 4 \end{array}\right]$$

$$\xrightarrow{\text{row}_1 \mapsto \text{row}_1 - 3\text{row}_2 - 2\text{row}_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 4 \end{array}\right] \to \begin{cases} a & = 5 \\ b & = -4 \\ c & = 4 \end{cases}$$

The solution is $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 5 \\ -4 \\ 4 \end{bmatrix}$ and the solution set is $\left\{ \begin{bmatrix} 5 \\ -4 \\ 4 \end{bmatrix} \right\}$, which is the same as we got before.

What happens when you apply row reduction to an inconsistent system? Let's see. Consider the system

$$\begin{cases} x + \ \ y = 1 \\ 4x + 4y = 7 \end{cases}. \tag{16}$$

Before continuing, convince yourself that this system is inconsistent. The augmented matrix for System (16) is

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 4 & 4 & 7 \end{array}\right].$$

We apply the row operation $\text{row}_2 \mapsto \text{row}_2 - 4\text{row}_1$ to get

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 0 & 3 \end{array}\right],$$

which corresponds to the system

$$\begin{cases} x + \ \ y = 1 \\ 0x + 0y = 3 \end{cases}.$$

But, the last equation says $0x + 0y = 3$, which is not true for any choice of x and y! Thus, we see applying row reduction to an inconsistent system reveals its inconsistency.

Example. Find a solution to the following system:

$$\begin{cases} x \qquad \ \ + \ \ z = 4 \\ x + \ \ y + 2z = -8 \\ x + 3y + 4z = -18 \end{cases}.$$

© Jason Siefken, 2015–2022

As usual we rewrite the system as an augmented matrix and then row reduce.

$$\begin{cases} x \quad\quad + \ z = 4 \\ x + \ \ y + 2z = -8 \\ x + 3y + 4z = -18 \end{cases} \to \left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 1 & 1 & 2 & -8 \\ 1 & 3 & 4 & -18 \end{array}\right] \xrightarrow{\text{row}_3 \mapsto \text{row}_3 - \text{row}_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 1 & 1 & 2 & -8 \\ 0 & 2 & 2 & -10 \end{array}\right]$$

$$\xrightarrow{\text{row}_2 \mapsto \text{row}_2 - \text{row}_1} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 1 & -12 \\ 0 & 2 & 2 & -10 \end{array}\right]$$

$$\xrightarrow{\text{row}_3 \mapsto \text{row}_3 - 2\text{row}_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 1 & -12 \\ 0 & 0 & 0 & 14 \end{array}\right] \to \begin{cases} x \quad\quad + \ z = 4 \\ \quad\quad y + \ z = -12 \\ 0x + 0y + 0z = 14 \end{cases}$$

The equation $0x + 0y + 0z = 14$ is never true and so the system is inconsistent. Since there are no values of x, y, and z that satisfy the system, the solution set is $\{\}$, the empty set.

Practice Problems

1 For each equation given below, determine if it is a linear equation. If not, explain what makes it nonlinear.

(a) $\cos(4)x_1 + \mathrm{e}y_2 + \pi z_3 = \mathrm{e}^{\pi}$

(b) $4x_1 + 2x_2 + 5x_4 = 4x_2 + 4x_5 + 5$

(c) $5x + 2y + 8z = \cos(y)$

(d) $12x + 3xy + 5z = 2$

(e) $\cos(4)x + \sin(4)y = \tan(4)x$

(f) $\frac{x}{y} = 1$

2 Convert each vector equation given below to a system of linear equations.

(a) $x\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} 4 \\ 6 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -5 \\ 2 \end{bmatrix}$

(b) $x\begin{bmatrix} 7 \\ 16 \end{bmatrix} + y\begin{bmatrix} 8 \\ 13 \end{bmatrix} = \begin{bmatrix} 11 \\ 30 \end{bmatrix}$

(c) $\vec{u} + t\vec{u} - s(\vec{v} + \vec{w}) = \vec{0}$ where $\vec{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

3 Convert each system of linear equations given below to a vector equation.

(a) $\begin{cases} \quad\quad 4x_2 + 2x_3 = 0 \\ x_1 \quad\quad + 2x_3 = 0 \\ \quad\quad 9x_2 + 2x_3 = 1 \end{cases}$

(b) $\begin{cases} 0x + 0y + 0z = 0 \\ \ \ x + \ \ y + \ \ z = 3 \end{cases}$

4 Consider the vector equation $x\begin{bmatrix} 2 \\ 4 \end{bmatrix} + y\begin{bmatrix} 8 \\ 16 \end{bmatrix} = \vec{b}$ where $\vec{b}$ is unknown.

(a) Show that if $\vec{b} = \begin{bmatrix} 7 \\ 14 \end{bmatrix}$, the system is consistent.

(b) Are there other vectors $\vec{b}$ that make the system consistent? If so, how many? Justify your answer.

(c) Show that if $\vec{b} = \begin{bmatrix} 5 \\ 12 \end{bmatrix}$, the system is inconsistent.

(d) Are there other vectors $\vec{b}$ that make the system inconsistent? If so, how many? Justify your answer.

5 On Kokoro's farm, there is a cage with 35 animals, some of which are chickens and some of which are rabbits. Kokoro counted the total number of legs in the cage and found that there were 94 legs in all (notably, each chicken has exactly two legs and each rabbit has four legs). Kokoro decides to use this information to figure out how many chickens and how many rabbits there are.[64]

(a) Set up a system of linear equations that you could solve to answer Kokoro's question.

(b) Is the system consistent? If so, answer Kokoro's question.

(c) Kokoro wants to set up three other cages. For each described cage below, explain using complete English sentences, whether such a configuration is possible. Justify your answers using linear algebra.

i. Kokoro wants to set up a cage with *cats* and *dogs* (notably, each cat has exactly four legs and each dog has four legs) so that there are 35 animals in total, and the total number of legs is 94.

ii. Kokoro wants to set up a cage with *cats* and *dogs* so that there are 35 animals in total, and the total number of legs is 140.

iii. Kokoro wants to set up a cage with *chickens* and *rabbits* so that there are 42 animals in total, and the total number of legs is 77.

6 For each statement below, determine whether it is true or false. Justify your answer.

(a) A system of linear equations of 4 variables with 3 equations is always consistent.

(b) Any system of linear equation with $0x_1 + 0x_2 + \cdots + 0x_n = 0$ being one of the equations must be consistent.

(c) There are $m, c \in \mathbb{R}$ so that the y-axis is the solution set to the equation $y = mx + c$.

(d) There are $m, c \in \mathbb{R}$ so that the x-axis is the solution set to the equation $y = mx + c$.

[64] This problem based on a classical Chinese problem from the ancient Chinese treatise *Mathematical Classic of Master Sun* (or *Sunzi Suanjing*) written during 3rd to 5th centuries A.D.

© *Jason Siefken, 2015–2022*

(e) There are $m_1, m_2, c \in \mathbb{R}$ so that the x-axis (in $\mathbb{R}^3$) is the solution set to the equation $z = m_1 x + m_2 y + c$.

(f) A system of exactly one equation can have an empty solution set.

© Jason Siefken, 2015–2022

 © Jason Siefken, 2015–2022

Systems of Linear Equations II

Appendix 2

In this appendix you will learn

- How to put a matrix into reduced row echelon form.
- How to use free variables to write down the complete solution to a system of linear equations.
- That a system of linear equations has 0, 1, or infinitely many solutions.
- How solution sets to systems of linear equations relate to intersecting hyperplanes.

Consider the system

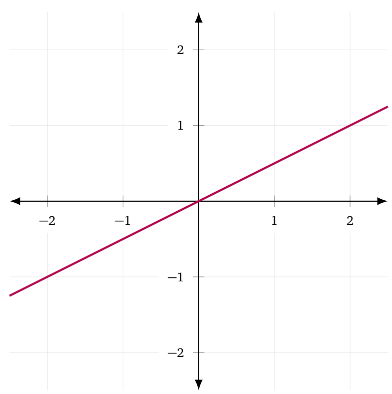
$$\begin{cases} x - 2y = 0 \\ 2x - 4y = 0 \end{cases}. \tag{17}$$

Notice that every solution to the first equation is also a solution to the second equation. Applying row reduction, we get the system

$$\begin{cases} x - 2y = 0 \\ 0x + 0y = 0 \end{cases},$$

but that second equation, $0x + 0y = 0$, is funny. It is always true, no matter the choice of x and y. It adds no new information! In retrospect, it might be obvious that both equations from System (17) contain the same information making one equation redundant.

System (17) is an example of an *underdetermined* system of equations, meaning there is not enough information to uniquely determine the value of each variable. Its solution set is a line, which we can find by graphing.



From this picture, we could describe the complete solution to System (17) in *vector form* by

$$\vec{x} = t\begin{bmatrix}2\\1\end{bmatrix}.$$

But, what about a more complicated system? The system

$$\begin{cases} x + y + z = 1 \\ \quad\;\; y - z = 2 \end{cases}$$

is also underdetermined. It has a complete solution described by

$$\begin{bmatrix}x\\y\\z\end{bmatrix} = t\begin{bmatrix}-2\\1\\1\end{bmatrix} + \begin{bmatrix}-1\\2\\0\end{bmatrix},$$

but this is much harder to find by graphing. Fortunately, we won't have to graph. Row reduction, combined with the notion of *free variables*, will provide a solution.

© Jason Siefken, 2015–2022

Reduced Row Echelon Form

Before we tackle complete solutions for underdetermined systems, we need to talk about *reduced row echelon form*,[65] which is abbreviated *rref*. The reduced row echelon form of a matrix is the simplest (in terms of reading off solutions) form a matrix can be turned into via elementary row operations.

Reduced Row Echelon Form (RREF). A matrix X is in ***reduced row echelon form*** if the following conditions hold:

- The first non-zero entry in every row is a 1; these entries are called ***pivots*** or ***leading ones***.
- Above and below each leading one are zeros.
- The leading ones form an echelon (staircase) pattern. That is, if row i has a leading one, every leading one appearing in row $j>i$ also appears to the *right* of the leading one in row i.
- All rows of zeros occur at the bottom of X.

Columns of a reduced row echelon form matrix that contain pivots are called ***pivot columns***.[a]

[a] If a matrix is augmented, we usually do not refer to the augmented column as a pivot column, even if it contains a pivot.

Example. Which of the follow matrices are in reduced row echelon form? For those that are, identify which columns are pivot columns. For those that are not, what condition(s) fail?

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

A is not in reduced row echelon form because the second row of A is a row of zeros but does not occur at the bottom.

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

B is not in reduced row echelon form for two reasons: (i) the first non-zero entry in the third row is not a 1, and (ii) the entry below the leading one in the second row is not zero.

$$B=\begin{bmatrix}1&0&0&8\\0&1&3&7\\0&2&1&4\end{bmatrix}$$

C is in reduced row echelon form and the first, second, third columns are the pivot columns of C.

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

D is not in reduced row echelon form for two reasons: (i) the entries above the leading one in the third row are not all zeros, and (ii) the leading one in the second row appears to the left of the leading one in the first row.

$$D=\begin{bmatrix}0&1&3&6\\1&0&0&9\\0&0&1&4\end{bmatrix}$$

We've encountered the reduced row echelon form of a matrix already in the examples of Appendix 1. Recall the system

$$\begin{cases} t+2s-2r=-15\\ 2t+\ \ s-5r=-21\\ t-4s+\ \ r=18\end{cases}\qquad\text{with augmented matrix}\qquad X=\left[\begin{array}{ccc|c}1&2&-2&-15\\2&1&-5&-21\\1&-4&1&18\end{array}\right].$$

[65] Reduced row echelon form is alternatively called *row reduced echelon form*; whether you say "reduced row" or "row reduced" makes no difference to the math!

© Jason Siefken, 2015–2022

The matrix X could be row reduced to

$$X' = \left[\begin{array}{ccc|c} 1 & 2 & -2 & -15 \\ 0 & -3 & -1 & 9 \\ 0 & 0 & 5 & 15 \end{array}\right],$$

which was suitable for solving the system. However, X' is not in reduced row echelon form (the leading non-zero entries must all be ones). We can further row reduce X' to

$$X'' = \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

X'' is the *reduced row echelon form* of X, and reading off the solution to the original system from X'' is as simple as can be!

Every matrix, M, has a unique reduced row echelon form, written rref(M), which can be obtained from M by applying elementary row operations. There are many ways to compute the reduced row echelon form of a matrix, but the following algorithm always works.

Row Reduction Algorithm. Let M be a matrix.

1. If M takes the form $M = [\vec{0}|M']$ (that is, its first column is all zeros), apply the algorithm to M'.
2. If not, perform a row-swap (if needed) so the upper-left entry of M is non-zero.
3. Let α be the upper-left entry of M. Perform the row operation $\text{row}_1 \mapsto \frac{1}{\alpha}\text{row}_1$. The upper-left entry of M is now 1 and is called a ***pivot***.
4. Use row operations of the form $\text{row}_i \mapsto \text{row}_i + \beta\,\text{row}_1$ to zero every entry below the pivot.
5. Now, M has the form
$$M = \left[\begin{array}{c|c} 1 & ?\,? \\ \hline \vec{0} & M' \end{array}\right],$$
where M' is a *submatrix* of M. Apply the algorithm to M'.

The resulting matrix is in ***pre-reduced row echelon form***. To put the matrix in ***reduced row echelon form***, additionally apply step 6.

6. Use the row operations of the form $\text{row}_i \mapsto \text{row}_i + \beta\,\text{row}_j$ to zero above each pivot.

Though there might be more efficient ways, and you might encounter ugly fractions, the row reduction algorithm will *always* convert a matrix to its reduced row echelon form.

Example. Apply the row-reduction algorithm to the matrix

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

First notice that M starts with a column of zeros, so we will focus on the right side of M. We will draw a line to separate it.

$$M = \left[\begin{array}{c|cccc} 0 & 0 & 0 & -2 & -2 \\ 0 & 1 & 2 & 3 & 2 \\ 0 & 2 & 4 & 5 & 3 \end{array}\right]$$

Next, we perform a row swap to bring a non-zero entry to the upper left.

$$\left[\begin{array}{c|cccc} 0 & 0 & 0 & -2 & -2 \\ 0 & 1 & 2 & 3 & 2 \\ 0 & 2 & 4 & 5 & 3 \end{array}\right] \xrightarrow{\text{row}_1 \leftrightarrow \text{row}_2} \left[\begin{array}{c|cccc} 0 & 1 & 2 & 3 & 2 \\ 0 & 0 & 0 & -2 & -2 \\ 0 & 2 & 4 & 5 & 3 \end{array}\right]$$

© Jason Siefken, 2015–2022

The upper-left entry is already a 1, so we can use it to zero all entries below.

$$\left[\begin{array}{c|cccc} 0 & 1 & 2 & 3 & 2 \\ 0 & 0 & 0 & -2 & -2 \\ 0 & 2 & 4 & 5 & 3 \end{array}\right] \xrightarrow{\text{row}_3 \mapsto \text{row}_3 - 2\text{row}_1} \left[\begin{array}{c|cccc} 0 & 1 & 2 & 3 & 2 \\ 0 & 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & -1 & -1 \end{array}\right]$$

Now we work on the submatrix.

$$\left[\begin{array}{cc|ccc} 0 & 1 & 2 & 3 & 2 \\ \hline 0 & 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & -1 & -1 \end{array}\right]$$

Again, the submatrix has a first column of zeros, so we pass to a sub-submatrix.

$$\left[\begin{array}{ccc|cc} 0 & 1 & 2 & 3 & 2 \\ \hline 0 & 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & -1 & -1 \end{array}\right]$$

Now we turn the upper left entry into a 1 and use that pivot to zero all entries below.

$$\left[\begin{array}{ccc|cc} 0 & 1 & 2 & 3 & 2 \\ \hline 0 & 0 & 0 & -2 & -2 \\ 0 & 0 & 0 & -1 & -1 \end{array}\right] \xrightarrow{\text{row}_2 \mapsto \frac{-1}{2}\text{row}_2} \left[\begin{array}{ccc|cc} 0 & 1 & 2 & 3 & 2 \\ \hline 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 & -1 \end{array}\right] \xrightarrow{\text{row}_3 \mapsto \text{row}_3 + \text{row}_2} \left[\begin{array}{ccc|cc} 0 & 1 & 2 & 3 & 2 \\ \hline 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The matrix is now in pre-reduced row echelon form. To put it in reduced row echelon form, we zero above each pivot.

$$\begin{bmatrix} 0 & 1 & 2 & 3 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{row}_1 \mapsto \text{row}_1 - 3\text{row}_2} \begin{bmatrix} 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

All matrices, whether augmented or not, have a reduced row echelon form. Correctly applying the row reduction algorithm takes practice, but being able to row reduce a matrix is the analogue of "knowing your multiplication tables" for linear algebra.

Free Variables & Complete Solutions

By now we are very familiar with the system

$$\begin{cases} x + 2y - 2z = -15 \\ 2x + y - 5z = -21 \\ x - 4y + z = 18 \end{cases},$$

which has a unique solution $(x, y, z) = (-1, -4, 3)$. We can compute this by row reducing the associated augmented matrix:

$$\text{rref}\left(\left[\begin{array}{ccc|c} 1 & 2 & -2 & -15 \\ 2 & 1 & -5 & -21 \\ 1 & -4 & 1 & 18 \end{array}\right]\right) = \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 3 \end{array}\right],$$

which corresponds to the system

$$\begin{cases} x = -1 \\ y = -4 \\ z = 3 \end{cases},$$

from which the solution is immediate. But what happens when there isn't a unique solution?

Consider the system

$$\begin{cases} x + 3y = 2 \\ 2x + 6y = 4 \end{cases}. \tag{18}$$

When using an augmented matrix to solve this system, we run into an issue.

$$\left[\begin{array}{cc|c} 1 & 3 & 2 \\ 2 & 6 & 4 \end{array}\right] \xrightarrow{\text{row reduces to}} \left[\begin{array}{cc|c} 1 & 3 & 2 \\ 0 & 0 & 0 \end{array}\right]$$

© Jason Siefken, 2015–2022

From the reduced row echelon form, we're left with the equation $x+3y=2$, which isn't exactly a *solution*. Effectively, the original system had only one equation's worth of information, so we cannot solve for both x and y based on the original system. To get ourselves out of this pickle, we will use a notational trick: introduce the arbitrary equation $y=t$.[66] Now, because we've already done row-reduction, we see

$$\begin{cases} x+3y=2 \\ 2x+6y=4 \\ y=t \end{cases} \xrightarrow{\text{row reduces to}} \begin{cases} x+3y=2 \\ y=t \end{cases}.$$

Here we've omitted the equation $0=0$ since it adds no information. Now, we can solve for x and y in terms of t.

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

Notice that t here stands for an arbitrary real number. Any choice of t produces a valid solution to the original system (go ahead, pick some values for t and see what happens). We call t a *parameter* and y a *free variable*.[67] Notice further that

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

is vector form of the line $x+3y=2$.

Though you can usually make many choices about which variables are free variables, one choice always works: *pick all variables corresponding to non-pivot columns to be free variables*. For this reason, we refer to non-pivot non-augmented columns of a row-reduced matrix as *free variable columns*.

Example. Use row reduction to find the complete solution to $\begin{cases} x+y+z=1 \\ y-z=2 \end{cases}$

The corresponding augmented matrix for the system is

$$A=\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 2 \end{array}\right].$$

A is already in pre-reduced row echelon form, so we only need to zero above each pivot.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 2 \end{array}\right] \xrightarrow{\text{row}_1\mapsto\text{row}_1-\text{row}_2} \left[\begin{array}{ccc|c} 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 2 \end{array}\right]=\operatorname{rref}(A).$$

The third column of $\operatorname{rref}(A)$ is a free variable column, so we introduce the arbitrary equation $z=t$ and solve the system in terms of t:

$$\begin{cases} x \qquad +2z=-1 \\ \quad y-\ \ z=2 \\ \qquad\quad\ \ z=t \end{cases}.$$

Written in vector form, the complete solution is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} -1-2t \\ 2+t \\ t \end{bmatrix}=t\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}+\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix},$$

and written as a set, the solution set is

$$\left\{\vec{x}\in\mathbb{R}^3 : \vec{x}=t\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}+\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} \text{ for some } t\in\mathbb{R}\right\}.$$

Consider the (somewhat strange) system of one equation

$$\left\{0x+0y+z=1\right..$$

[66] This equation is called *arbitrary* because it introduces no new information about the original variables. The restrictions on x and y aren't changed by introducing the fact $y=t$.

[67] We call y *free* because we may pick it to be anything we want and still produce a solution to the system.

© Jason Siefken, 2015–2022

The solution set for this system is the xy-plane in $\mathbb{R}^3$ shifted up by one unit. We can use row reduction and free variables to see this.

The system corresponds to the augmented matrix

$$\left[\begin{array}{ccc|c} 0 & 0 & 1 & 1 \end{array}\right]$$

which is already in reduced row echelon form. It's third column is the only pivot column, making columns 1 and 2 free variable columns (remember, we don't count augmented columns as free variable columns). Thus, we introduce two arbitrary equations, $x = t$ and $y = s$, and solve the new system

$$\begin{cases} 0x + 0y + z = 1 \\ x \qquad\qquad\;\; = t \\ \qquad\; y \qquad = s \end{cases}$$

for (x, y, z), which gives

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} t \\ s \\ 1 \end{bmatrix} = t\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Using row reduction and free variables, we can find complete solutions to very complicated systems. The process is straight-forward enough that even a computer can do it![68]

Example. Consider the system of equations in the variables x, y, z, and w:

$$\begin{cases} \qquad\qquad -2w = -2 \\ \;\; y + 2z + 3w = 2 \\ 2y + 4z + 5w = 3 \end{cases}$$

Find the solution set for this system.

The augmented matrix corresponding to this system is

$$M = \left[\begin{array}{cccc|c} 0 & 0 & 0 & -2 & -2 \\ 0 & 1 & 2 & 3 & 2 \\ 0 & 2 & 4 & 5 & 3 \end{array}\right],$$

which we've row reduced in a previous example:

$$\mathrm{rref}(M) = \left[\begin{array}{cccc|c} 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Here, columns 1 and 3 are free variable columns, so we introduce the equations $x = t$ and $z = s$. Now, solving the system

$$\begin{cases} \quad y + 2z \qquad = -1 \\ \qquad\qquad\; w = 1 \\ x \qquad\qquad\;\; = t \\ \qquad\;\; z \qquad = s \end{cases}$$

for (x, y, z, w), gives

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} t \\ -1-2s \\ s \\ 1 \end{bmatrix} = t\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}.$$

Thus, the solution set to the system is

$$\left\{ \vec{x} \in \mathbb{R}^4 \;:\; \vec{x} = t\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} \text{ for some } t, s \in \mathbb{R} \right\}.$$

[68]Computers usually don't follow the algorithm outlined here because they have to deal with *rounding error*. But, there is a modification of the row reduction algorithm called row reduction with *partial pivoting* which solves some issues with rounding error.

© *Jason Siefken, 2015–2022*

Free Variables & Inconsistent Systems

If you need a free variable/parameter to describe the complete solution to a system of linear equations, the system necessarily has an infinite number of solutions—one coming from every choice of value for your free variable/parameter. However, one still needs to be careful when deciding *from an augmented matrix* whether a system of linear equations has an infinite number of solutions.

Consider the augmented matrices A and B, which are given in reduced row echelon form.

$$A=\left[\begin{array}{cc|c}1&2&-1\\0&0&0\end{array}\right]\qquad B=\left[\begin{array}{cc|c}1&2&0\\0&0&1\end{array}\right]$$

Both matrices lack a pivot in their second column. However, A corresponds to a system with an infinite solution set, while B corresponds to an inconsistent system with an empty solution set. We can debate whether it is appropriate to say that B has a free variable column,[69] but one thing is clear: when evaluating the number of solutions to a system, you must pay attention to whether or not the system is consistent.

Putting everything together, we can fully classify the number of solutions to a system of linear equations based on pivots/free variables.

Consistent	Pivots	Number of Solutions
False	At least one column doesn't have a pivot	0
True	Every column has a pivot	1
True	At least one column doesn't have a pivot	Infinitely many

This information is so important, we will also codify it in a theorem.

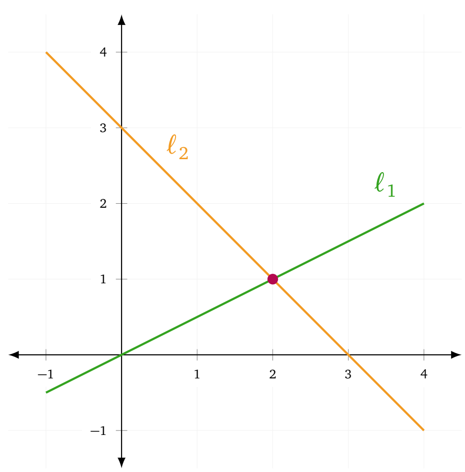
Theorem. A system of linear equations has either 0 solutions, 1 solution, or infinitely many solutions.

The Geometry of Systems of Equations

Consider the system of equations

$$\begin{cases} x-2y=0 & \text{row}_1\\ x+\ \ y=3 & \text{row}_2\end{cases}\tag{19}$$

The only values of x and y that satisfy both equations is $(x,y)=(2,1)$. However, each row, viewed in isolation, specifies a line in $\mathbb{R}^2$. Call the line coming from the first row ℓ_1 and the line coming from the second row ℓ_2.



These two lines intersect exactly at the point $\begin{bmatrix}2\\1\end{bmatrix}$. And, of course they should. By definition, a solution to a system of equations satisfies *all* equations. In other words, a solution to System (19) is a point that lies in both ℓ_1 and ℓ_2. In other words, solutions lie in $\ell_1\cap\ell_2$.

Takeaway. Geometrically, a solution to a system of equations is the intersection of objects specified by the individual equations.

[69] On the one hand, the second column fits the description. On the other hand, you cannot make any choices when picking a solution, since there are no solutions.

© Jason Siefken, 2015–2022

This perspective sheds some light on inconsistent systems. The system

$$\begin{cases} x - 2y = 0 & \text{row}_1 \\ 2x - 4y = 2 & \text{row}_2 \end{cases}$$

is inconsistent. And, when we graph the lines specified by the rows, we see that they are parallel and never intersect. Thus, the solution set is empty.

Planes & Hyperplanes

Consider the solution set to a single linear equation viewed in isolation. For example, in the three-variable case, we might consider

$$x + 2y - z = 3.$$

The solution set to this equation is a *plane*. Why? For starters, writing down the complete solution involves picking two free variables. Suppose we pick $y = t$ and $z = s$. Then, before we even do a calculation, we know the complete solution will be described in vector form by

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

where $\vec{d}_1$, $\vec{d}_2$, and $\vec{p}$ come from doing the usual computations. But, that is vector form of a plane!

In general, a single equation in n variables requires $n-1$ free variables to describe its complete solution. The only exception is the trivial equation, $0x_1 + \cdots + 0x_n = 0$, which requires n free variables. For the sake of brevity, from now on we will assume that *a linear equation in n variables* means *a non-trivial linear equation in n variables*.

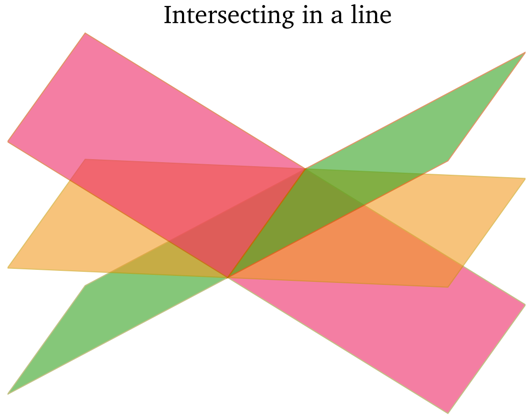
Applying this knowledge, we can construct a table for systems consisting of a single linear equation.

Number of Variables	Number of Free Variables	Complete Solution
2	1	Line in $\mathbb{R}^2$
3	2	Plane in $\mathbb{R}^3$
4	3	Volume in $\mathbb{R}^4$

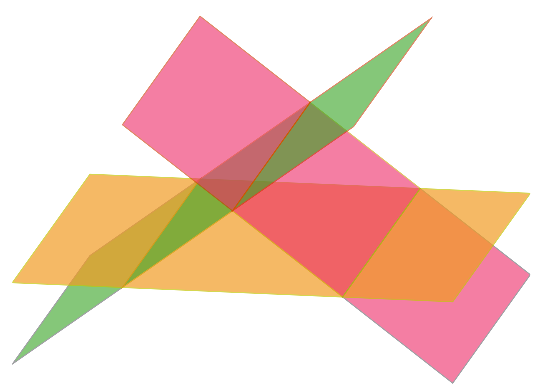
Notice that the dimension of the solution set (a line being one dimensional, a plane being two dimensional, and a volume being three dimensional) is always one less than the dimension of the ambient space ($\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{R}^4$).[70] Such sets are called *hyperplanes* because they are flat and plane-like. However, unlike a plane, the dimension of a *hyperplane* need not be two.

With our newfound geometric intuition, we can understand solutions to systems of linear equations in a different way. The solution set to a system of linear equations of two variables must be the result of intersecting lines. Therefore, the only options are: a point, a line, or the empty set. The solution to a system of linear equations of three variables is similarly restricted. It can only be: a point, a line, a plane, or the empty set.

Intersecting in a line



Empty intersection



In higher dimensions, the story is the same: solution sets are formed by intersecting hyperplanes and we can use algebra to precisely describe these sets of intersection.

[70]Another way to describe these sets would be to say that they have *co-dimension* 1.

© *Jason Siefken, 2015–2022*

Practice Problems

1 Find the complete solution to the following systems.

(a) $\begin{cases} 4x+6y+3z-10w=6 \\ 5x+2y+z-7w=2 \\ -6x+2y+z+4w=2 \end{cases}$

(b) $\begin{cases} 2x+2y+z=-1 \\ y-4z+2w=3 \\ x-y-3z-4w=5 \end{cases}$

(c) $\begin{cases} x+y-2z=-5 \\ -4x+y+5z=3 \end{cases}$

(d) $\begin{cases} 3x-2y=-4 \\ x+y+3z=3 \\ -4x+y-3z=1 \end{cases}$

(e) $\begin{cases} x-y+2z=-1 \\ 2x+y+4z=1 \\ 3x-4y+3z=-2 \end{cases}$

(f) $\begin{cases} 2x+z=8 \\ x+y+z=4 \\ x+3y+2z=4 \\ 3x+2y+4z=9 \end{cases}$

2 For each system of linear equations given below: (i) write down its augmented matrix, (ii) use row reduction algorithm to determine if it is consistent or not, (iii) for each consistent system, give the complete solution.

(a) $\begin{cases} -10x_1-4x_2+4x_3=28 \\ 3x_1+x_2-x_3=-8 \\ x_1+x_2-\frac{1}{2}x_3=-3 \end{cases}$

(b) $\begin{cases} 3x_1-2x_2+4x_3=54 \\ 5x_1-3x_2+6x_3=88 \\ x_1=-3 \end{cases}$

(c) $\{x+2y=5$

(d) $\begin{cases} 4x=6 \\ 2x=3 \end{cases}$

(e) $\begin{cases} x_1+2x_2+4x_3-3x_4=0 \\ 3x_1+5x_2+6x_3-4x_4=1 \\ 4x_1+5x_2-2x_3+3x_4=3 \end{cases}$

(f) $\begin{cases} x_1-x_2+5x_3+x_4=1 \\ x_1+x_2-2x_3+3x_4=3 \\ 3x_1-x_2+8x_3+x_4=5 \\ x_1+3x_2-9x_3+7x_4=5 \end{cases}$

(g) $\{0x+0y+0z=0$

3 (a) Let $\vec{v}_1=\begin{bmatrix}1\\1\\-2\\4\end{bmatrix}$, $\vec{v}_2=\begin{bmatrix}1\\4\\0\\2\end{bmatrix}$ and $\vec{v}_3=\begin{bmatrix}-2\\-2\\4\\-8\end{bmatrix}$.

Set up and solve a system of linear equations whose solution will determine if the vectors $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ are linearly independent.

(b) Let $\vec{v}_1=\begin{bmatrix}1\\2\\3\end{bmatrix}$, $\vec{v}_2=\begin{bmatrix}-2\\1\\0\end{bmatrix}$ and $\vec{v}_3=\begin{bmatrix}2\\7\\1\end{bmatrix}$.

Set up and solve a system of linear equations whose solution will determine if the vectors $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ span $\mathbb{R}^3$.

(c) Let ℓ_1 and ℓ_2 be described in vector form by

$$\overbrace{\vec{x}=t\begin{bmatrix}1\\3\end{bmatrix}+\begin{bmatrix}1\\1\end{bmatrix}}^{\ell_1} \quad \overbrace{\vec{x}=t\begin{bmatrix}2\\1\end{bmatrix}+\begin{bmatrix}3\\4\end{bmatrix}}^{\ell_2}.$$

Set up and solve a system of linear equations whose solution will determine if the lines ℓ_1 and ℓ_2 intersect.

(d) Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be described in vector form by

$$\mathcal{P}_1: \quad \vec{x}=t\begin{bmatrix}1\\-1\\0\end{bmatrix}+s\begin{bmatrix}-1\\-1\\2\end{bmatrix},$$

$$\mathcal{P}_2: \quad \vec{x}=t\begin{bmatrix}1\\-1\\1\end{bmatrix}+s\begin{bmatrix}-1\\3\\-2\end{bmatrix}+\begin{bmatrix}0\\1\\-1\end{bmatrix}.$$

Set up and solve a system of linear equations whose solution will determine if the planes $\mathcal{P}_1$ and $\mathcal{P}_2$ intersect.

4 Presented below some students' arguments for question 3. Evaluate whether their reasoning is totally correct, mostly correct, or incorrect. If their reasoning is not totally correct, point out what mistake(s) they made and how they might be fixed.

(a) i. Consider the vector equation

$$x\vec{v}_1+y\vec{v}_2+z\vec{v}_3=\vec{0}$$

where $x,y,z\in\mathbb{R}$.

Since $(x,y,z)=(0,0,0)$ is a solution to the equation, the equation has the trivial solution. Therefore, the vectors $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ are linearly independent.

(a) ii. Consider the vector equation

$$x\vec{v}_1+y\vec{v}_2+z\vec{v}_3=\vec{0}$$

where $x,y,z\in\mathbb{R}$.

Notice that $(x,y,z)=(-2,0,-1)$ is a solution to the equation. Since $y=0$ in this solution, it is a trivial solution and therefore the vectors $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ are linearly independent.

(c) i. The lines ℓ_1 and ℓ_2 intersect when their x and y-coordinates are equal. Equating x and y-coordinates gives

$$\begin{cases} t+1=2t+3 \\ 3t+1=t+4 \end{cases}.$$

This system is equivalent to

$$\begin{cases} t=-2 \\ 2t=3 \end{cases}.$$

Since this system is inconsistent, the lines ℓ_1 and ℓ_2 do not intersect.

© Jason Siefken, 2015–2022

(c) ii. The lines ℓ_1 and ℓ_2 intersect when their x and y-coordinates are equal. Equating x and y-coordinates gives

$$\begin{cases} t+1=2s+3 \\ 3t+1=s+4 \end{cases}.$$

This system is equivalent to

$$\begin{cases} t-2s=2 \\ 3t-\ \ s=3 \end{cases},$$

and the solution is

$$\begin{bmatrix} t \\ s \end{bmatrix} = \begin{bmatrix} 4/5 \\ -3/5 \end{bmatrix}.$$

Therefore the lines ℓ_1 and ℓ_2 intersect at $\begin{bmatrix} 4/5 \\ 3/5 \end{bmatrix}$.

(d) i. Notice that

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

and

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

So, $\vec{x} = \begin{bmatrix} 1/2 \\ -1/2 \\ 0 \end{bmatrix}$ is a point on $\mathcal{P}_1$ and $\mathcal{P}_2$. Therefore the planes $\mathcal{P}_1$ and $\mathcal{P}_2$ intersect.

(d) ii. Notice that $\vec{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is a point in $\mathcal{P}_2$, but this point is not in $\mathcal{P}_1$. Therefore the planes do not intersect.

© Jason Siefken, 2015–2022

Matrices & Matrix Operations

In this appendix you will learn:

- What a matrix is and how to describe a matrix in terms of its shape and entries.
- Special types of matrices, including diagonal and triangular matrices.
- How to multiply a matrix by a vector and a matrix by a matrix as well as multiple ways to interpret matrix-vector and matrix-matrix multiplication.
- Basic properties of matrix-matrix multiplication: non-commutative, associated, distributive.

A *matrix* is a box (rectangular array) of numbers, usually surrounded by brackets.[71] For example,

$$\begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{bmatrix} \qquad \begin{bmatrix}x\\ y\\ z\end{bmatrix} \qquad \begin{pmatrix}7.5 & -2\\ -3 & 0\end{pmatrix} \qquad \begin{matrix}1 & 2\\ 3 & 4\\ 5 & 6\end{matrix}$$

are matrices. In this book, we always write matrices with square brackets "$[\cdots]$". We have already used matrices in two ways: to keep track of the coefficients when solving a system of linear equations and to describe vectors (as a column of numbers). In both of these cases, matrices were a notational tool used to keep related numbers together. But, as we will soon see, matrices are also mathematical objects that you can do arithmetic with.

Matrix Notation

A matrix can be described by its *shape*[72] (the number of rows and columns in the matrix) and its *entries* (the numbers inside the matrix). Traditionally, matrices are labeled with capital letters and their entries are labeled with lower-case letters.

Consider the matrix A with m rows and n columns:

$$A = \begin{array}{c c} & \overleftrightarrow{\quad n \text{ columns} \quad} \\ m \text{ rows} \updownarrow & \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n}\\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn}\end{bmatrix}\end{array}$$

We call A an "m *by* n matrix", or in notation, an "$m \times n$ matrix".[73] The entries of A are indexed by their $(row, column)$ coordinates. So, the $(1,1)$ entry of A is a_{11}, the $(2,1)$ entry of A is a_{21}, etc.. When subscripting a matrix entry, it is tradition to omit a separator between the row and column index. That is, we write a_{ij} instead of $a_{i,j}$ or $a_{(i,j)}$.[74]

> **Example.** Let $A = \begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{bmatrix}$. Find the shape of A as well as the $(1,3)$ entry of A.
>
> A has two rows and three columns, so A is a 2×3 matrix. The $(1,3)$ entry of A is the number in the first row and third column of A, which is 3.

Since a matrix is completely determined by its shape and entries, we can define a matrix via a formal. For example, define B to be the 2×3 matrix whose (i,j) entry, b_{ij}, satisfies the formula $b_{ij} = i + j$. In this case

$$B = \begin{bmatrix}b_{11} & b_{12} & b_{13}\\ b_{21} & b_{22} & b_{23}\end{bmatrix} = \begin{bmatrix}2 & 3 & 4\\ 3 & 4 & 5\end{bmatrix}.$$

The shorthand $B = [b_{ij}]$ means that "B is a matrix whose (i,j) entry is b_{ij}". Using this shorthand, we could alternatively say $B = [b_{ij}]$ is a 2×3 matrix satisfying $b_{ij} = i + j$.

---

[71] The word *matrix* comes from the Latin word for womb; you can think of a matrix as "holding numbers together".

[72] Other terms for the *shape* of a matrix include the "*size* of a matrix" and the "*dimensions* of a matrix". But, be careful not to confuse this usage of "dimension" with the term "dimension" in the context of subspaces.

[73] In this context, "$\times$" is read as "by"

[74] There's nothing wrong with including a separator. It's just not common practice.

© Jason Siefken, 2015–2022

Example. Let $C = [c_{ij}]$ be a 3×3 matrix satisfying $c_{ij} = i - j$. Write down C.

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}.$$

Basic Terms

A matrix has three special parts: the diagonal, the upper triangle, and the lower triangle.

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{bmatrix}}_{\text{Diagonal}} \quad \underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{bmatrix}}_{\text{Upper Triangle}} \quad \underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{bmatrix}}_{\text{Lower Triangle}}$$

Formally, we define the diagonal and upper/lower triangle of a matrix in terms of the row and column coordinates.

Diagonal. The ***diagonal*** of an $m \times n$ matrix $A = [a_{ij}]$ consists of the entries a_{ij} satisfying $i = j$.

Upper & Lower Triangle. Let $A = [a_{ij}]$ be an $m \times n$ matrix. The ***upper triangle*** of A consists the entries a_{ij} satisfying $j \geq i$. The ***lower triangle*** of A consists of the entries a_{ij} satisfying $j \leq i$.

Special Matrices

There are several special matrices that come up often.

Triangular Matrices. A matrix is called ***upper triangular*** if all non-zero entries lie in the upper triangle of the matrix and a matrix is called ***lower triangular*** if all non-zero entries lie in the lower triangle. A matrix is called ***triangular*** if it is either upper or lower triangular.

Square Matrix. A matrix is called ***square*** if it has the same number of rows as columns.

Diagonal Matrix. A square matrix is called ***diagonal*** the only non-zero entries in the matrix appear on the diagonal.

Symmetric Matrix. The square matrix $A = [a_{ij}]$ is called ***symmetric*** if its entries satisfy $a_{ij} = a_{ji}$. Alternatively, if the entries of A satisfy $a_{ij} = -a_{ji}$, then A is called ***skew-symmetric*** or ***anti-symmetric***.

Zero Matrix. A matrix is called a ***zero matrix*** if all its entries are zero.

Identity Matrix.
An ***identity matrix*** is a square matrix with ones on the diagonal and zeros everywhere else. The $n \times n$ identity matrix is denoted $I_{n \times n}$, or just I when its size is implied.

Example. Identify the diagonal of $A = \begin{bmatrix} -2 & 3 \\ 5 & 6 \\ 7 & 7 \end{bmatrix}$.

The diagonal of A consists of the entries in A whose row coordinate is equal to the column coordinate. So, the diagonal of A consists of -2 and 6.

Example. Apply a single row operation to $B = \begin{bmatrix} -2 & 3 \\ 0 & 6 \\ 0 & 12 \end{bmatrix}$ to make it an upper triangular matrix.

To make B an upper triangular matrix, we need all entries below the diagonal to be zero. More specifically, we need to get rid of the 12 in the lower-right corner of B. By applying the row operation $\text{row}_3 \mapsto \text{row}_3 - 2\,\text{row}_2$

© Jason Siefken, 2015–2022

to B, we get the upper triangular matrix

$$B' = \begin{bmatrix} -2 & 3 \\ 0 & 6 \\ 0 & 0 \end{bmatrix}$$

Example. If possible, produce a matrix that is both upper and lower triangular.

For a matrix to be upper triangular, all entries below the diagonal must be zero. For a matrix to be lower triangular, all entries above the diagonal need to be zero. Therefore, if a matrix is both upper and lower triangular, the only non-zero entries of the matrix must be on the diagonal. It follows that

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{etc.}$$

are valid examples.

Matrix Arithmetic

In Module 1, we saw that vectors were an extension of numbers that allowed us to describe directions. Similarly, we can view matrices as a more general type of "number". Matrices can do everything numbers can, everything vectors can, and more!

Basic Operations

The rules for addition and scalar multiplication of matrices are what you expect: to add two matrices, add the corresponding entries, and to scalar multiply a matrix, distribute the scalar to each entry.

For example,

$$2\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} -1 & 2 & 0 \\ 3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1-1 & 2+2 & 3+0 \\ 4+3 & 5+0 & 6+1 \end{bmatrix} = \begin{bmatrix} 0 & 4 & 3 \\ 7 & 5 & 7 \end{bmatrix}.$$

While any matrix can be scalar multiplied by any scalar, matrix addition only makes sense for *compatible* matrices. That is, you can only add together two matrices of the same shape.

Takeaway. If two matrices are of the same shape, you can add them by adding entries "straight across"; you can multiply a matrix by a scalar by distributing the scalar to each entry of the matrix.

Example. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, let $B = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$, and let $C = \begin{bmatrix} -1 & 0 & -1 \\ 2 & 1 & 2 \end{bmatrix}$. Compute $2A + B$ and $A + 3C$, if possible.

First, note that A is a 2×3 matrix and so it can only be added to another 2×3 matrix. Since B is a 2×2 matrix, $2A + B$ is not defined. But, C is a 2×3 matrix, so $A + 3C$ is defined. Computing,

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

$$= \begin{bmatrix} -2 & 2 & 0 \\ 10 & 8 & 12 \end{bmatrix}.$$

Matrix-Vector Multiplication

Matrices and vectors interact via *matrix-vector* multiplication. There are two equivalent ways to think about matrix-vector multiplication: in terms of columns (the *column picture*) and in terms of rows (the *row picture*).

© Jason Siefken, 2015–2022

Column Picture

Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \vec c_1 & \vec c_2 & \cdots & \vec c_n \\ | & | & & | \end{bmatrix}$$

be an $m \times n$ matrix and let

$$\vec c_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} \qquad \vec c_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} \qquad \cdots \qquad \vec c_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

be the vectors corresponding to the columns of A. Further, let $\vec x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ be a vector.

We define the matrix-vector product $A\vec x$ to be the linear combination of $\vec c_1, \vec c_2, \ldots, \vec c_n$ with coefficients x_1, x_2, etc.. That is,

$$A\vec x = \begin{bmatrix} | & | & & | \\ \vec c_1 & \vec c_2 & \cdots & \vec c_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec c_1 + x_2\vec c_2 + \cdots + x_n\vec c_n.$$

Example. Let $B = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$ and let $\vec v = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$. Compute $B\vec v$ using the column picture.

The column vectors of B are $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$, so

$$B\vec v = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \end{bmatrix} = 4\begin{bmatrix} 1 \\ -2 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 1 \end{bmatrix}.$$

The column picture of matrix-vector multiplication hints that matrix-vector multiplication can be used to encode sophisticated problems involving linear combinations (see Module 7 for details).

Row Picture

Alternatively, let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \text{——} & \vec r_1 & \text{——} \\ \text{——} & \vec r_2 & \text{——} \\ & \vdots & \\ \text{——} & \vec r_m & \text{——} \end{bmatrix}$$

be an $m \times n$ matrix and let

$$\vec r_1 = \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{bmatrix} \qquad \vec r_2 = \begin{bmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2n} \end{bmatrix} \qquad \cdots \qquad \vec r_m = \begin{bmatrix} a_{m1} \\ a_{m2} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

be vectors corresponding to the rows of A. Note that we are writing the *row vectors of A in column vector form*.

Further, let $\vec x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ be a vector.

© Jason Siefken, 2015–2022

We alternatively define the matrix-vector product $A\vec{x}$ as the vector whose coordinates are the dot products of the rows of A and the vector $\vec{x}$. That is,

$$A\vec{x} = \begin{bmatrix} \text{——} & \vec{r}_1 & \text{——} \\ \text{——} & \vec{r}_2 & \text{——} \\ & \vdots & \\ \text{——} & \vec{r}_m & \text{——} \end{bmatrix} \vec{x} = \begin{bmatrix} \vec{r}_1 \cdot \vec{x} \\ \vec{r}_2 \cdot \vec{x} \\ \vdots \\ \vec{r}_m \cdot \vec{x} \end{bmatrix}.$$

Example. Let $B = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$ and let $\vec{v} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$. Compute $B\vec{v}$ using the row picture. Verify that the result matches with what you get from the column picture.

The row vectors of B are $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 3 \end{bmatrix}$, so

$$B\vec{v} = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 3 \end{bmatrix} \\ \begin{bmatrix} -2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 3 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} (1)(4) + (2)(3) \\ (-2)(4) + (3)(3) \end{bmatrix} = \begin{bmatrix} 10 \\ 1 \end{bmatrix}.$$

This is the same vector we got in the previous example using the column picture!

Since the row picture of matrix-vector multiplication involves dot products, which in turn relate to angles and geometry, the row picture hints that matrix-vector multiplication can be used to encode sophisticated problems involving the angles between multiple vectors (see Module 7 for more).

Compatibility

Matrix-vector multiplication is only possible when the shape of the matrix is compatible with the size of the vector. That is the number of *columns* of the matrix must match the number of *coordinates* in the vector. (Try some examples using the row an column picture to make sure you agree.)

The result of a matrix-vector product is always a vector, but the number of coordinates in the output vector can change. For example, if M is a 2×3 matrix, the product $M\vec{v}$ is only defined if $\vec{v} \in \mathbb{R}^3$. However, the resulting vector $\vec{w} = M\vec{v}$ is in $\mathbb{R}^2$ (try an example and verify for yourself). This means matrix-vector multiplication can be used to move vectors between different spaces!

Takeaway. Let A be an $m \times n$ matrix and $\vec{x}$ be a vector. The matrix-vector product $A\vec{x}$ is only defined if $\vec{x}$ has n coordinates. In that case, the result is a vector with m coordinates.

Matrix-Matrix Multiplication

In many circumstances, we can also multiply two matrices with each other. To do so, we repeatedly apply matrix-vector multiplication. Let C and A be matrices and let $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$ be the columns of A. Then,

$$CA = C \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_k \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ C\vec{a}_1 & C\vec{a}_2 & \cdots & C\vec{a}_k \\ | & | & & | \end{bmatrix}.$$

Here, we "distributed" C into the matrix A, creating a new matrix whose columns are $C\vec{a}_1, C\vec{a}_2, \ldots$. Using the row picture to expand each $C\vec{a}_i$, we arrive at an explicit formula. Let $\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_m$ be the rows of C. Then,

$$CA = \begin{bmatrix} \text{——} & \vec{r}_1 & \text{——} \\ \text{——} & \vec{r}_2 & \text{——} \\ & \vdots & \\ \text{——} & \vec{r}_m & \text{——} \end{bmatrix} \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_k \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \vec{r}_1 \cdot \vec{a}_1 & \vec{r}_1 \cdot \vec{a}_2 & \cdots & \vec{r}_1 \cdot \vec{a}_k \\ \vec{r}_2 \cdot \vec{a}_1 & \vec{r}_2 \cdot \vec{a}_2 & \cdots & \vec{r}_2 \cdot \vec{a}_k \\ \vdots & \vdots & \ddots & \vdots \\ \vec{r}_m \cdot \vec{a}_1 & \vec{r}_m \cdot \vec{a}_2 & \cdots & \vec{r}_m \cdot \vec{a}_k \end{bmatrix}.$$

Example. Let $X = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 0 \end{bmatrix}$ and $Y = \begin{bmatrix} 2 & 3 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$. Compute XY and YX.

© Jason Siefken, 2015–2022

Computing XY entry by entry, we get the $(1,1)$ entry is $\begin{bmatrix}1\\2\\3\end{bmatrix}\cdot\begin{bmatrix}2\\1\\1\end{bmatrix}=7$, the $(2,1)$ entry is $\begin{bmatrix}0\\-1\\0\end{bmatrix}\cdot\begin{bmatrix}2\\1\\1\end{bmatrix}=-1$, and so on. Computing all the entries we get

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

Computing YX entry by entry, we get the $(1,1)$ entry is $\begin{bmatrix}2\\3\end{bmatrix}\cdot\begin{bmatrix}1\\0\end{bmatrix}=2$, the $(2,1)$ entry is $\begin{bmatrix}1\\1\end{bmatrix}\cdot\begin{bmatrix}1\\0\end{bmatrix}=1$, and so on. Computing all the entries we get

$$YX=\begin{bmatrix}2&3\\1&1\\1&0\end{bmatrix}\begin{bmatrix}1&2&3\\0&-1&0\end{bmatrix}=\begin{bmatrix}2&1&6\\1&1&3\\1&2&3\end{bmatrix}.$$

From the previous example, we see that multiplying matrices in different orders can produce different results. Formally we say that matrix multiplication is not *commutative* (in contrast, scalars can be multiplied in any order). This non-commutativity holds even for square matrices.[75] For example,

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

Further, for a matrix-matrix multiplication to be possible, the shapes of each matrix must be compatible. Using our knowledge of matrix-vector multiplication, we can deduce that if the matrix-matrix product CA makes sense, then the number of *columns* of C must match the number of *rows* of A.

Writing the shape of two matrices side-by-side allows for a quick compatibility check.

$$\text{rows of } C\times\underbrace{\text{columns of } C\quad\text{rows of } A}_{\text{must be equal for } CA \text{ to exist}}\times\text{columns of } A$$

A successful matrix-matrix multiplication will always result in a matrix with the number of rows of the first and the number of columns of the second.

$$\underbrace{\text{rows of } C\times\text{columns of } C\quad\text{rows of } A\times\text{columns of } A}_{\text{successful product will be a rows of } C\times\text{columns of } A \text{ matrix}}$$

Example. Let A be a 2×3 matrix, let B be a 3×4 matrix, and let C be a 1×3 matrix. Determine the shape of the matrices resulting from all possible products of A, B, and C.

For the product of two matrices to exist, the number of columns of the first matrix must equal the number of rows in the second. Therefore, the only matrix products that are possible are AB and CB.
AB is the product of a 2×3 matrix with a 3×4 matrix, and so will be a 2×4 matrix.
CB is the product of a 1×3 matrix with a 3×4 matrix, and so will be a 1×4 matrix.

Matrix Algebra

Let A, B, and C be $n\times n$ matrices and let α be a scalar. We can now write algebraic expressions like

$$A(B+\alpha C).$$

Since the matrices are all $n\times n$, such expressions are always defined and the results are again $n\times n$ matrices. We can *almost* treat arithmetic with $n\times n$ matrices like arithmetic with numbers, save the fact that changing the order of multiplication might change the result. Many familiar properties of arithmetic carry over to matrices. For example, matrix multiplication is both *associative* and *distributive*. That is,

$$(AB)C=A(BC)\qquad\text{and}\qquad A(B+C)=AB+AC\qquad\text{and}\qquad(A+B)C=AC+BC.$$

We're already familiar with the special matrices I, the identity matrix, and $\mathbf{0}$, the zero matrix. In terms of matrix algebra, these behave like the numbers 1 and 0. That is,

$$IA=AI=A\qquad\text{and}\qquad\mathbf{0}A=A\mathbf{0}=\mathbf{0}$$

[75] Of course, it is *possible* that $AB=BA$ for matrices A and B. It just doesn't happen very often.

© Jason Siefken, 2015–2022

for any compatible square matrix A.

To kick it up a level, when working with square matrices, we can define *polynomials* of matrices. Using familiar exponent notation, $A^2 = AA$, we can formulate questions like

> Does the equation $A^2 = -I$ have a 2×2 matrix solution?

Famously, the equation $x^2 = -1$ has no real solutions, but $A^2 = -I$ actually does have real 2×2 matrix solutions (see if you can find one)! In this text we will only scratch the surface of what can be done with matrix algebra, but it's powerful stuff.[76]

> Matrix algebra behaves a lot like regular algebra except that the order of multiplication matters and matrices must always have compatible sizes.

More Notation

Linear algebra has many different products: scalar multiplication, dot products, matrix-vector products, and matrix-matrix products, to name a few. To distinguish between these different products, we use different notations.

For matrix-vector and matrix-matrix products, we use *adjacency* to represent multiplication. That is, we write

$$A\vec{v} \qquad \text{and} \qquad AB$$

to indicate a product. Specifically, we do *not* use the symbols "$\cdot$" or "$\times$" to represent matrix-vector or matrix-matrix products (these symbols are reserved for the dot product and cross product, respectively).

Practice Problems

1 For each description below, if possible, create a matrix matching the description. Otherwise, explain why such a matrix doesn't exist.

(a) A 2×2 diagonal matrix whose entries sum to -1.

(b) A 2×2 symmetric matrix whose entries sum to -1.

(c) A 4×2 symmetric matrix whose entries sum to -1.

(d) A 3×3 skew-symmetric matrix whose entries sum to -1.

(e) A 1×4 matrix $A = [a_{ij}]$ whose entries satisfy $a_{ij} = \sqrt{i+j}$.

2 Consider the following matrices $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 2 \\ 1 & 6 & 5 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, $C = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$, and $D = \begin{bmatrix} 0 & -2 & 1 \end{bmatrix}$. For each of the following, (i) determine if the operation is defined, and (ii) compute the result using both the column picture and row picture of multiplication (if applicable).

(a) AC

(b) $2A+B$

(c) $A-B$

(d) CA

(e) AB

(f) BA

(g) DC

(h) CD

3 In general, matrix multiplication is non-commutative. However, some types of matrices are special.

Let A and B be 2×2 diagonal matrices and let X and Y be $n\times n$ diagonal matrices.

(a) Show by direct computation that $AB = BA$.

(b) Show that both XY and YX also diagonal matrices.

(c) Is it true that $XY = YX$ no matter n? Explain.

4 Classify the following statements as true or false.

(a) A matrix in reduced row echelon form is an upper triangular matrix.

(b) A diagonal matrix is in reduced row echelon form.

(c) Every zero matrix is also square.

(d) A zero matrix is neither upper or lower triangular.

(e) A matrix that is both upper triangular and lower triangular must be diagonal.

(f) Using row operations every lower triangular matrix can be converted into an upper triangular matrix.

(g) The product of two lower triangular matrices is a lower triangular matrix (provided the product is defined).

5 Let R be a $1\times n$ matrix and let C be an $n\times 1$ matrix.

(a) Is the product RC defined? If so, what is its shape?

(b) Is the product CR defined? If so, what is its shape?

(c) Let $\vec{r}$ be the (only) row vector in R and let $\vec{c}$ be the (only) column vector in C. Are $\vec{r}\cdot\vec{c}$ and RC the same? Explain.

(d) Let $\vec{x},\vec{y}\in\mathbb{R}^n$ and let $R_{\vec{x}}$ be the $1\times n$ matrix with $\vec{x}$ as a row vector and let $C_{\vec{y}}$ be the $n\times 1$ matrix with $\vec{y}$ as a column vector. The *inner product* of $\vec{x}$ and $\vec{y}$ is defined to be $R_{\vec{x}}C_{\vec{y}}$. The *outer product* of $\vec{x}$ and $\vec{y}$ is defined to be $C_{\vec{y}}R_{\vec{x}}$.

i. How does the inner product of $\vec{x}$ and $\vec{y}$ relate to the dot product of $\vec{x}$ and $\vec{y}$?

[76] Galois theory and representation theory both heavily rely on matrix algebra.

© Jason Siefken, 2015–2022

ii. Let Q be the outer product of $\vec{x}$ and $\vec{y}$. What does the reduced row echelon form of Q look like?

iii. Let Q be the outer product of $\vec{x}$ and $\vec{y}$. Show that the columns of Q are always linearly dependent when $n \geq 2$.

6 A 3×3 matrix is called a *Heisenberg matrix* if it takes the form $\begin{bmatrix}1 & a & c\\ 0 & 1 & b\\ 0 & 0 & 1\end{bmatrix}$ for some $a, b, c \in \mathbb{R}$.

(a) Show that if A and B are Heisenberg matrices, then so are AB and BA.

(b) If A and B are Heisenberg matrices, is it always the case that $AB = BA$? Give a proof or a counter example.

(c) Let $X = \begin{bmatrix}1 & a & c\\ 0 & 1 & b\\ 0 & 0 & 1\end{bmatrix}$ and let $Y = \begin{bmatrix}1 & -1 & ab-c\\ 0 & 1 & -b\\ 0 & 0 & 1\end{bmatrix}$. Show that $XY = I_{3\times 3}$.

7 Let X be a matrix of the form $\begin{bmatrix}a & -b\\ b & a\end{bmatrix}$ for some $a, b \in \mathbb{R}$.

(a) Show that X^2 has the same form as X.

(b) Is there a solution to the matrix equation $X^2 = I_{2\times 2}$? If so, how many?

(c) Is there a solution to the matrix equation $X^2 = -I_{2\times 2}$? If so, how many?

(d) Let Y be an arbitrary 2×2 matrix. How many solutions are there to the equation $Y^2 = I_{2\times 2}$?

(e) Do you agree with the statement "every positive real number has exactly two square roots"? Do you agree with the statement "every diagonal matrix with positive entries on the diagonal has exactly two square roots"? Explain.

© Jason Siefken, 2015–2022

Formulas for 2×2 and 3×3 Determinants

In this appendix you will learn:

- A practical formula for 2×2 determinants.
- How to calculate 3×3 determinants using diagonal method.
- How to calculate 3×3 determinants using cofactor expansion method.

Module 14 discusses the theory of determinants and gives a general algorithm for computing determinants by using elementary matrices. But, since 2×2 and 3×3 matrices arise so often in day-to-day life,[77] it is worth learning some special-purpose formulas for computing the determinants of 2×2 and 3×3 matrices.

It should be noted that these formulas are *special*. Though there do exist formulas for determinants of $n \times n$ matrices, *they are exponentially more complex than the formulas for 2×2 and 3×3 matrices*. As such, determinants of large matrices are usually computed using row reduction/elementary matrices and not formulas.[78]

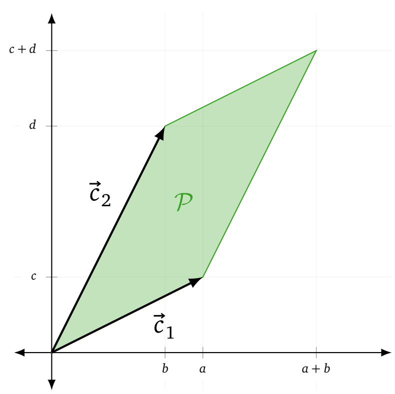
Computing 2×2 Determinants

For a 2×2 matrix, we can calculate its determinant directly from its entries.

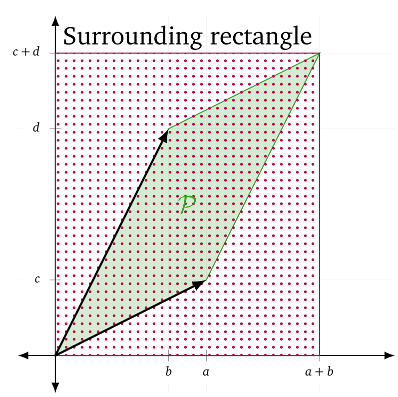
Theorem. Let $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then,

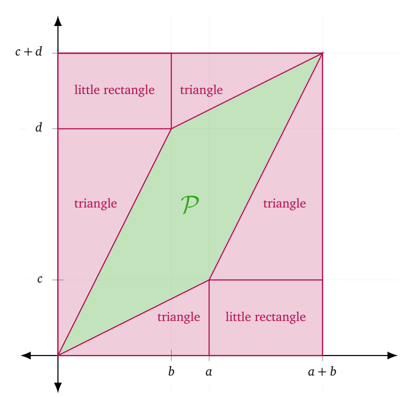
$$\det(M) = ad - bc.$$

The 2×2 determinant formula can be deduced from Volume Theorem I. Let $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and let $\vec{c}_1 = \begin{bmatrix} a \\ c \end{bmatrix}$ and $\vec{c}_2 = \begin{bmatrix} b \\ d \end{bmatrix}$ be the columns of M. We need to compute the area of the parallelogram $\mathcal{P}$, with sides $\vec{c}_1$ and $\vec{c}_2$.



We can compute the area of $\mathcal{P}$ by computing the area of a rectangle that contains $\mathcal{P}$ and subtracting off any area that we "over counted".





[77] The day-to-day life of a mathematics student, at least!

[78] General determinant formulas are primarily useful as theoretical tools for writing proofs.

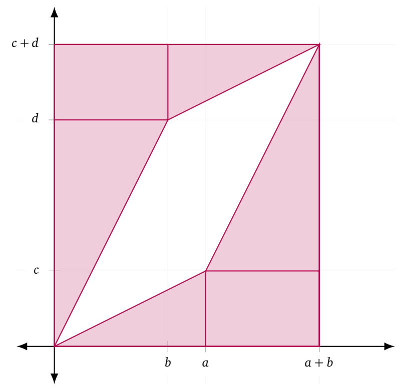
© Jason Siefken, 2015–2022

Thus,

$$\mathrm{Vol}(\mathcal{P}) = \text{area of big rectangle} - \text{area of little rectangles} - \text{area of triangles}.$$

Using the coordinates for $\vec{c}_1$ and $\vec{c}_2$, we get

$$\mathrm{Vol}(\mathcal{P}) = \underbrace{(a+b)(d+c)}_{\text{area of big rectangle}} - \underbrace{2bc}_{\text{area of little rectangles}} \underbrace{-2\frac{ac}{2} + 2\frac{bd}{2}}_{\text{area of triangles}} = ad - bc. \tag{20}$$



Equation (20) is beautiful and simple, but its derivation should give you pause. Volume Theorem I refers to *oriented* volume and we didn't make any reference to orientation in our figures! Indeed, we played tricks with pictures. We drew $\vec{c}_1$ and $\vec{c}_2$ in a right-handed orientation in the first quadrant, even though the vectors $\begin{bmatrix} a \\ c \end{bmatrix}$ and $\begin{bmatrix} b \\ d \end{bmatrix}$ could be in *any* quadrant (and one or both could even be the zero vector)! To fully justify Equation (20), we need to consider cases based on all the possible ways $\vec{c}_1$ and $\vec{c}_2$ can form a parallelogram. However, it turns out that every case gives the same answer: $\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$.

Example. Directly compute the determinant of $M = \begin{bmatrix} 1 & 6 \\ 2 & 7 \end{bmatrix}$ using the 2×2 formula. Then, find the determinant of M after decomposing it into the product of elementary matrices.

Using the 2×2 formula, we get

$$\det(M) = (1)(7) - (2)(6) = -5.$$

Alternatively, row reducing and keeping track of the elementary matrices for each step, we see

$$\underbrace{\begin{bmatrix} 1 & -6 \\ 0 & 1 \end{bmatrix}}_{E_3} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{5} \end{bmatrix}}_{E_2} \underbrace{\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}}_{E_1} M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and so

$$M = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -5 \end{bmatrix}\begin{bmatrix} 1 & 6 \\ 0 & 1 \end{bmatrix} = E_1^{-1}E_2^{-1}E_3^{-1}.$$

E_1^{-1} and E_3^{-1} both have determinant 1, and E_2^{-1} has determinant -5. Thus,

$$\det(M) = \det(E_1^{-1})\det(E_2^{-1})\det(E_3^{-1}) = (1)(-5)(1) = -5,$$

which is exactly what we got using the formula.

Computing 3×3 Determinants

The formula for a 3×3 matrix is more complicated than the 2×2 formula.

© Jason Siefken, 2015–2022

Theorem. Let $M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$. Then

$$\det(M) = aei + bfg + cdh - gec - hfa - idb.$$

Fortunately, there is a clever mnemonic for remembering this formula called *the Rule of Sarrus* or *the diagonal trick*.

Rule of Sarrus

Let $M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$. To compute the determinant of M using the Rule of Sarrus, apply the following four steps.

Step 1. Augment M with copies of its *first two columns*.

$$\left[\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}\right]$$

Step 2. Multiply together and then *add* the entries along the three diagonals of the new matrix. These are called the *diagonal products*.

$$\left[\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}\right]$$

$$\text{sum of diagonal products} = aei + bfg + cdh.$$

Step 3. Multiply together and then *subtract* the entries along the three anti-diagonals. These are called the *anti-diagonal products*.

$$\left[\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}\right]$$

$$\text{difference of anti-diagonal products} = -gec - hfa - idb$$

Step 4. Add the diagonal products and subtract the anti-diagonal products to get the determinant.

$$\det(M) = aei + bfg + cdh - gec - hfa - idb.$$

Example. Use the diagonal trick to compute $\det\left(\begin{bmatrix} 1 & 4 & 0 \\ -2 & 3 & 1 \\ 0 & 2 & 1 \end{bmatrix}\right)$.

$$\left[\begin{array}{ccc|cc} 1 & 4 & 0 & 1 & 4 \\ -2 & 3 & 1 & -2 & 3 \\ 0 & 2 & 1 & 0 & 2 \end{array}\right]$$

$$\text{sum of diagonal products} = (1)(3)(1) + (4)(1)(0) + (0)(-2)(2) = 3 + 0 + 0.$$

$$\left[\begin{array}{ccc|cc} 1 & 4 & 0 & 1 & 4 \\ -2 & 3 & 1 & -2 & 3 \\ 0 & 2 & 1 & 0 & 2 \end{array}\right]$$

$$\text{difference of anti-diagonal products} = -(0)(3)(0) - (2)(1)(1) - (1)(-2)(4) = -0 - 2 - (-8).$$

Thus,

$$\det\left(\begin{bmatrix} 1 & 4 & 0 \\ -2 & 3 & 1 \\ 0 & 2 & 1 \end{bmatrix}\right) = 3 + 0 + 0 \quad - 0 - 2 - (-8) = 9.$$

© Jason Siefken, 2015–2022

It may be tempting to apply the Rule of Sarrus to 4×4 and larger matrices, but *don't do it!* There is a formula for 4×4 determinants, but *it's not given by the Rule of Sarrus*.[79]

Like the 2×2 formula for determinants, we can derive the 3×3 formula directly from the definition. However, it takes quite a bit more work.[80]

Determinant Formulas and Orientation

Determinants and orientation are connected and our determinant formulas (if we accept them as true) give us an alternative way to determine the orientation of a basis.

Let $\mathcal{B}=\{\vec b_1,\vec b_2\}$ be an ordered basis for $\mathbb{R}^2$, and let $M=[\vec b_1|\vec b_2]$ be the matrix whose columns are $\vec b_1$ and $\vec b_2$. Since $\mathcal{B}$ is linearly independent, we know that $\det(M)\neq 0$. Further, applying the definition of the determinant, we know

$$\det(M)>0$$

means that $\mathcal{B}$ is a right-handed basis and $\det(M)<0$ means $\mathcal{B}$ is a left-handed basis.

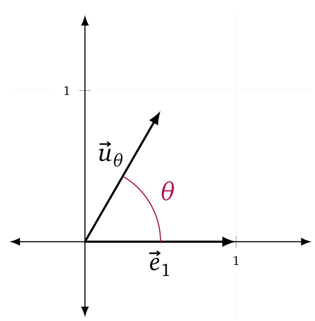
Example. Use a determinant to decide whether the ordered basis $\left\{\begin{bmatrix}1\\2\end{bmatrix},\begin{bmatrix}-3\\2\end{bmatrix}\right\}$ is left-handed or right-handed.

Let $A=\begin{bmatrix}1&-3\\2&2\end{bmatrix}$ be the matrix whose columns are the elements of the given ordered basis.
Using the formula for 2×2 determinants gives us

$$\det(A)=(1)(2)-(2)(-3)=8>0$$

and so we conclude $\left\{\begin{bmatrix}1\\2\end{bmatrix},\begin{bmatrix}-3\\2\end{bmatrix}\right\}$ is a right-handed basis.

Recall the ordered basis $\mathcal{Q}=\{\vec e_1,\vec u_\theta\}$ where $\vec u_\theta=\begin{bmatrix}\cos\theta\\\sin\theta\end{bmatrix}$ is the unit vector which forms an angle of θ with the positive x-axis.



Visually, we can see that $\mathcal{Q}$ should be right-handed when $\theta\in(0,\pi)$, left handed when $\theta\in(\pi,2\pi)$ and $\mathcal{Q}$ is not a basis when $\theta=0$ or $\theta=\pi$.

But what does the determinant say?

Computing the determinant of the matrix $Q=[\vec e_1|\vec u_\theta]$ directly using the 2×2 determinant formula, we get

$$\det(Q)=\det([\vec e_1|\vec u_\theta])=\det\left(\begin{bmatrix}1&\cos\theta\\0&\sin\theta\end{bmatrix}\right)=\sin\theta.$$

Notice that $\det(Q)=\sin\theta>0$ when $\theta\in(0,\pi)$, $\det(Q)=\sin\theta<0$ when $\theta\in(\pi,2\pi)$ and $\det(Q)=\sin\theta=0$ when $\theta\in\{0,\pi\}$.

The determinant supports our intuition.

[79]Because your curiosity is never ending, here's the formula. For a matrix 4×4 matrix $A=[a_{ij}]$, we have $\det(A)=a_{11}a_{22}a_{33}a_{44}-a_{11}a_{22}a_{34}a_{43}-a_{11}a_{23}a_{32}a_{44}+a_{11}a_{23}a_{34}a_{42}+a_{11}a_{24}a_{32}a_{43}-a_{11}a_{24}a_{33}a_{42}-a_{12}a_{21}a_{33}a_{44}+a_{12}a_{21}a_{34}a_{43}+a_{12}a_{23}a_{31}a_{44}-a_{12}a_{23}a_{34}a_{41}-a_{12}a_{24}a_{31}a_{43}+a_{12}a_{24}a_{33}a_{41}+a_{13}a_{21}a_{32}a_{44}-a_{13}a_{21}a_{34}a_{42}-a_{13}a_{22}a_{31}a_{44}+a_{13}a_{22}a_{34}a_{41}+a_{13}a_{24}a_{31}a_{42}-a_{13}a_{24}a_{32}a_{41}-a_{14}a_{21}a_{32}a_{43}+a_{14}a_{21}a_{33}a_{42}+a_{14}a_{22}a_{31}a_{43}-a_{14}a_{22}a_{33}a_{41}-a_{14}a_{23}a_{31}a_{42}+a_{14}a_{23}a_{32}a_{41}$. This formula involves 24 products. The 5×5 formula involves 120 products and the 6×6 formula involves 720 products. It only gets worse from there.

[80]If you're interested in proving the 3×3 determinant formula, try using the elementary matrix approach rather than computing the volume of a parallelepiped directly.

© Jason Siefken, 2015–2022

Practice Problems

1 For each matrix given below, calculate its determinant using both row reduction/elementary matrices and the 2×2 determinant formula.

(a) $\begin{bmatrix} 1 & 0 \\ 2 & 4 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 5 \\ 1 & 5 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

2 For each matrix given below, calculate its determinant using both row reduction/elementary matrices and the 3×3 determinant formula.

(a) $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 2 \\ 1 & 6 & 5 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & -4 & 1 \\ 2 & 6 & 5 \\ 2 & 2 & 3 \end{bmatrix}$

(c) $\begin{bmatrix} -1 & 0 & -8 \\ 1 & -3 & -8 \\ 1 & -2 & -1 \end{bmatrix}$

(d) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

(e) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

3 For each ordered set given below, use a determinant to decide whether it is a right-handed basis, a left-handed basis, or not a basis.

(a) $\left\{\begin{bmatrix} 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$

(b) $\left\{\begin{bmatrix} 1 \\ -3 \end{bmatrix}, \begin{bmatrix} -2 \\ 6 \end{bmatrix}\right\}$

(c) $\left\{\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}\right\}$

(d) $\left\{\begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}$

(e) $\left\{\begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 6 \end{bmatrix}\right\}$

4 Find all values of $a, b \in \mathbb{R}$ so that the ordered set $\left\{\begin{bmatrix} a^2 \\ ab \end{bmatrix}, \begin{bmatrix} ab \\ b \end{bmatrix}\right\}$ is (a) a right-handed basis, (b) a left-handed basis, (c) not a basis.

5 Let $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The *adjugate matrix* (sometimes called the *classical adjoint*) of M, notated M^{adj}, is the matrix given by $M^{\text{adj}} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Prove that if M is invertible, then $M^{-1} = \dfrac{M^{\text{adj}}}{\det(M)}$.

6 For each statement below, determine whether it is true or false. Justify your answer.

(a) A 2×2 matrix M has determinant 1 if and only if $M = I_{2\times 2}$.

(b) A 3×3 matrix M has determinant 1 if and only if $\operatorname{Vol Change}(\mathcal{T}_M)$ is equal to 1, where $\mathcal{T}_M$ is the transformation given by $\mathcal{T}_M(\vec{x}) = M\vec{x}$.

(c) For vectors $\vec{a}, \vec{b} \in \mathbb{R}^2$, it is always the case that $\det([\vec{a}|\vec{b}]) = -\det([\vec{b}|\vec{a}])$.

(d) For a 2×2 or 3×3 matrix M, multiplying a single entry of M by 4 will change $\det(M)$ by a factor of 4.

(e) For a square matrix A, it is always the case that $\det(A^T A) \geq 0$.

© Jason Siefken, 2015–2022

© Jason Siefken, 2015–2022

Practice Problem Solutions

Solutions for Module 1

1 (a) i. $\begin{bmatrix} 2 \\ 2 \\ -3 \end{bmatrix}$

ii. $\begin{bmatrix} 1 \\ -4 \end{bmatrix}$

(b) i. $\vec{e}_1 - 2\vec{e}_2 + 3\vec{e}_3$

ii. $3\vec{e}_2$

2 $\begin{bmatrix} 1 \\ -4 \\ 26 \\ -1 \\ 2 \end{bmatrix}$

3 (a) Yes, take $k = 5$.

(b) Yes, take $i = 2, j = 1$.

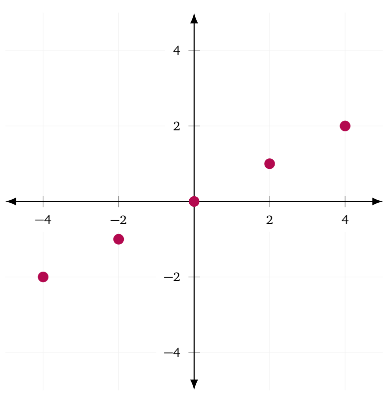
(c) Yes, take $k = -\frac{1}{2}$.

(d) No.

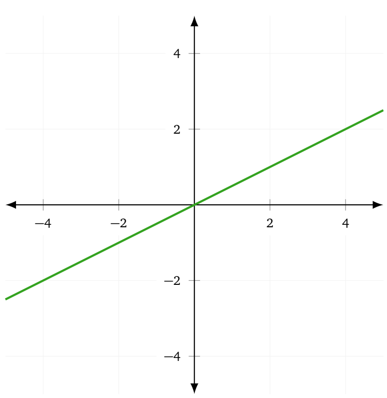
(e) Yes, take $r = 2$.

(f) No.

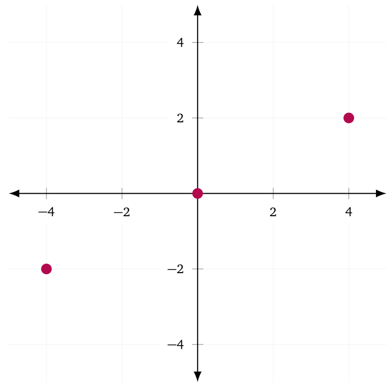
4 (a)



(b)



(c)



(d) The set D and B are equal.

We have $C \subseteq A$, $C \subseteq B$, $C \subseteq D$, $A \subseteq B$, $A \subseteq D$, and $B = D$.

5 (a) No.

(b) No.

(c) Yes.

(d) Yes.

6 (a) $\left\{ \vec{v} \in \mathbb{R}^2 : \vec{v} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \text{ for some } \alpha, \beta \in \mathbb{Q} \right\}$

(b) $\left\{ \vec{v} \in \mathbb{R}^2 : \vec{v} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \text{ for some } \alpha, \beta \in \mathbb{R} \setminus \mathbb{Q} \right\}$

(c) $\{ \vec{v} \in \mathbb{R}^2 : \vec{v} = -\vec{e}_1 \text{ or } \vec{v} = -\vec{e}_2 \}$

7 (a) i. False.
ii. True.
iii. False.
iv. True.
v. True.
vi. False.

(b) i. True.
ii. False.
iii. True.
iv. True.
v. True.
vi. False.

8 (a) Correct.

(b) Incorrect.

(c) Incorrect.

(d) Incorrect.

(e) Correct.

(f) Incorrect.

(g) Incorrect.

Solutions for Module 2

1 (a) $\ell_1 : \vec{x} = t \begin{bmatrix} 3 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 6 \end{bmatrix}$

© Jason Siefken, 2015–2022

(b) $\ell_2 : \vec{x} = t\begin{bmatrix}1\\6\end{bmatrix} + \begin{bmatrix}1\\1\end{bmatrix}$

(c) $\ell_3 : \vec{x} = t\begin{bmatrix}3\\4\end{bmatrix}$

(d) $\ell_4 : \vec{x} = t\begin{bmatrix}3\\4\\5\end{bmatrix} + \begin{bmatrix}2\\3\\5\end{bmatrix}$

(e) $\ell_5 : \vec{x} = t\begin{bmatrix}0\\3\\2\end{bmatrix} + \begin{bmatrix}0\\1\\-1\end{bmatrix}$

2 (a) $\mathcal{P}_1 : \vec{x} = t\begin{bmatrix}1\\0\\4\end{bmatrix} + s\begin{bmatrix}0\\1\\0\end{bmatrix}$

(b) $\mathcal{P}_2 : \vec{x} = t\begin{bmatrix}3\\4\\5\end{bmatrix} + s\begin{bmatrix}4\\4\\3\end{bmatrix} + \begin{bmatrix}2\\3\\5\end{bmatrix}$

(c) $\mathcal{P}_3 : \vec{x} = t\begin{bmatrix}1\\0\\-4\end{bmatrix} + s\begin{bmatrix}0\\1\\3\end{bmatrix} + \begin{bmatrix}0\\3\\-1\end{bmatrix}$

(d) $\mathcal{P}_4 : \vec{x} = t\begin{bmatrix}0\\1\\0\end{bmatrix} + s\begin{bmatrix}0\\0\\1\end{bmatrix} + \begin{bmatrix}1\\-1\\1\end{bmatrix}$

(e) $\mathbb{R}^2 : \vec{x} = t\begin{bmatrix}1\\0\end{bmatrix} + s\begin{bmatrix}0\\1\end{bmatrix}$

(f) $\mathcal{P}_5 : \vec{x} = t\begin{bmatrix}-1\\1\\0\\0\end{bmatrix} + s\begin{bmatrix}-3\\0\\2\\1\end{bmatrix} + \begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}$

3 (a) The lines ℓ_1 and ℓ_3 coincide. The line ℓ_2 intersects with both ℓ_1 and ℓ_3.

(b) $\ell_1 \cap \ell_2 \cap \ell_3 = \left\{\begin{bmatrix}2\\4\end{bmatrix}\right\}$

4 (a) $\mathcal{P}_1 \cap \ell = \left\{\begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$

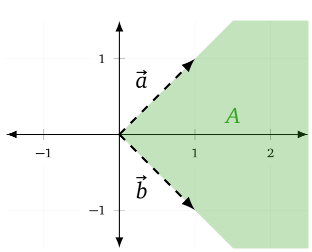
(b) $\mathcal{P}_1 \cap \mathcal{P}_2 : \vec{x} = t\begin{bmatrix}1\\1\\3\end{bmatrix} + \begin{bmatrix}-1/3\\5/3\\0\end{bmatrix}$

(c) $\mathcal{P}_2 \cap \ell = \left\{\begin{bmatrix}2\\4\\1\end{bmatrix}\right\}$

(d) $\mathcal{P}_3 : \vec{x} = t\begin{bmatrix}1\\3\\1\end{bmatrix} + s\begin{bmatrix}1\\0\\0\end{bmatrix}$

(e) Suppose $\mathcal{P}_2'$ is a plane that is parallel to $\mathcal{P}_2$. Notice that $\mathcal{P}_2'$ can be expressed with the equation $x-y=c$ for some $c \in \mathbb{R}$.
Finding such a plane $\mathcal{P}_2'$ which does not intersect ℓ is now equivalent to finding a number $c \in \mathbb{R}$ so that $(1+t)-(1+3t) \neq c$ for all $t \in \mathbb{R}$. But, $t = -1/2c$ solves the equation $(1+t)-(1+3t) = c$, so any plane $\mathcal{P}_2'$ that is parallel to $\mathcal{P}_2$ must intersect ℓ.

5 (a)



(b) ℓ is given by the set

$$\{\vec{x} \in \mathbb{R}^2 : \vec{x} = t\vec{a} + (1-t)\vec{b} \text{ for some } t \in \mathbb{R}\}.$$

(c) The above set can be rewritten as

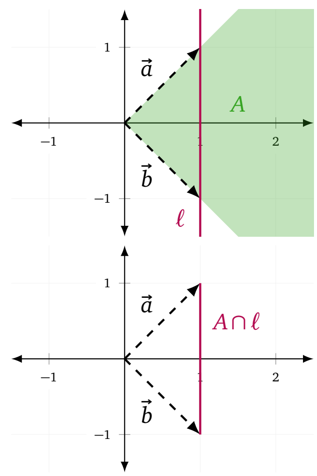
$$\{\vec{x} \in \mathbb{R}^2 : \vec{x} = t(\vec{a} - \vec{b}) + \vec{b} \text{ for some } t \in \mathbb{R}\}.$$

This is exactly the line given in vector form by

$$\vec{x} = t(\vec{a} - \vec{b}) + \vec{b}.$$

Since $\vec{a} - \vec{b} = \begin{bmatrix}0\\2\end{bmatrix}$, ℓ is the vertical line containing $\vec{b}$ (and $\vec{a}$).

(d)



(e) $A \cap \ell$ is the set

$$\{\alpha\vec{a} + \beta\vec{b} : \alpha, \beta \geq 0 \text{ and } \alpha + \beta = 1\}.$$

This is the set of convex linear combinations of $\vec{a}$ and $\vec{b}$.

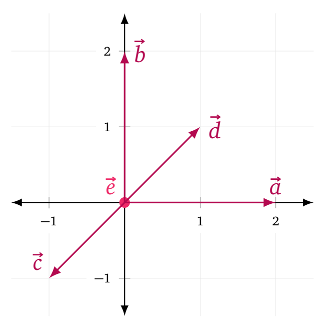
(f) The endpoints of $A \cap \ell$ are $\vec{a}$ and $\vec{b}$.

6 (a) $\vec{d} = \frac{1}{2}\vec{a} + \frac{1}{2}\vec{b}$

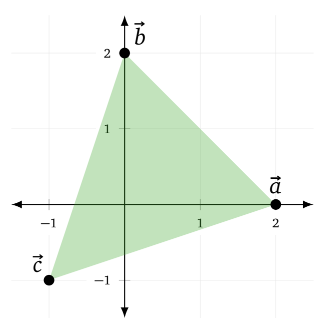
(b) $\vec{e} = \frac{1}{2}\vec{c} + \frac{1}{2}\vec{d}$

(c) $\vec{e} = \frac{1}{2}\vec{c} + \frac{1}{2}(\frac{1}{2}\vec{a} + \frac{1}{2}\vec{b}) = \frac{1}{2}\vec{c} + \frac{1}{4}\vec{a} + \frac{1}{4}\vec{b}$

(d)



(e)



© Jason Siefken, 2015–2022

The set is the filled-in triangle with vertices given by $\vec{a}$, $\vec{b}$, and $\vec{c}$. To see this, notice that the set of convex linear combinations of $\vec{a}$, $\vec{b}$, and $\vec{c}$ is the set of convex linear combinations of $\vec{c}$ and any convex linear combination of $\vec{a}$ and $\vec{b}$. Indeed,

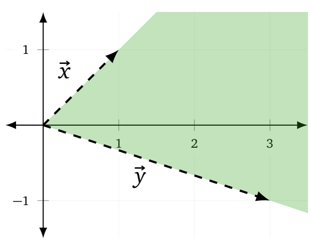
$$\alpha_1\vec{a}+\alpha_2\vec{b}+\alpha_3\vec{c} = (1-\alpha_3)\left(\frac{\alpha_1}{1-\alpha_3}\vec{a}+\frac{\alpha_2}{1-\alpha_3}\vec{b}\right)+\alpha_3\vec{c},$$

and $\alpha_1\vec{a}+\alpha_2\vec{b}+\alpha_3\vec{c}$ is a convex linear combination of $\vec{a}$, $\vec{b}$, and $\vec{c}$ exactly when

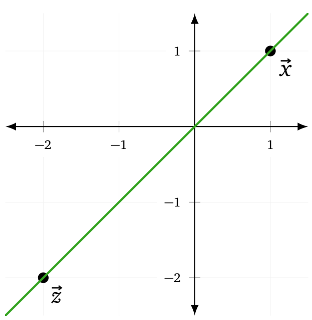
$$\frac{\alpha_1}{1-\alpha_3}\vec{a}+\frac{\alpha_2}{1-\alpha_3}\vec{b}$$

is a convex linear combination of $\vec{a}$ and $\vec{b}$ (verify this!). By the previous part, the set of convex linear combinations of $\vec{a}$ and $\vec{b}$ is the line segment between $\vec{a}$ and $\vec{b}$. Call this line segment S. Now we know the set of convex linear combinations of $\vec{a}$, $\vec{b}$, and $\vec{c}$ is the union of every line segment from $\vec{c}$ and a vector in S. This is the filled-in triangle with vertices given by $\vec{a}$, $\vec{b}$, and $\vec{c}$.

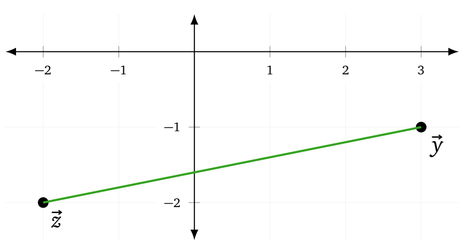
7 (a)



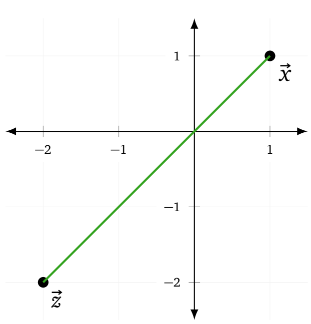
(b)



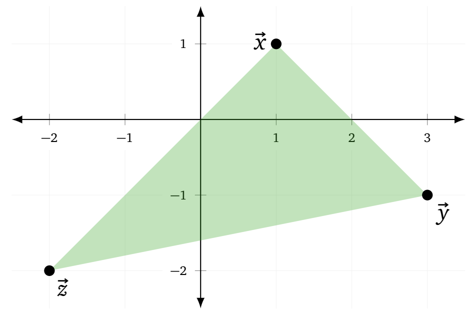
(c)



(d)



(e)



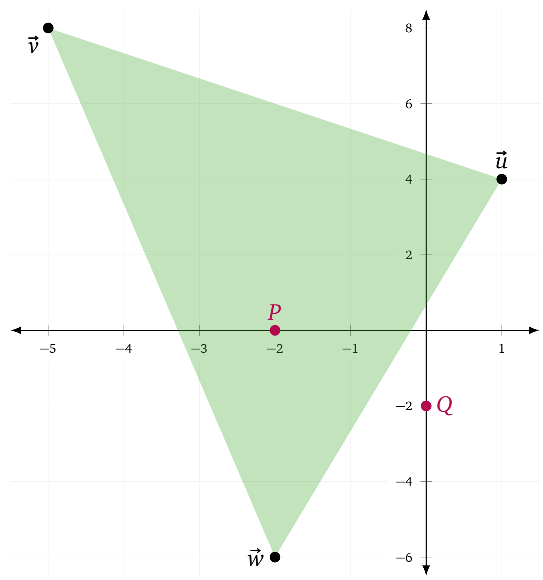
8 (7c)

$$\left\{\vec{v}\in\mathbb{R}^2 : \vec{v}=t\begin{bmatrix}-5\\-1\end{bmatrix}+\begin{bmatrix}3\\-1\end{bmatrix} \text{ for some } 0\le t\le 1\right\}.$$

(7d)

$$\left\{\vec{w}\in\mathbb{R}^2 : \vec{w}=t\begin{bmatrix}-3\\-3\end{bmatrix}+\begin{bmatrix}1\\1\end{bmatrix} \text{ for some } 0\le t\le 1\right\}.$$

9



Geometrically, the set of convex linear combinations of $\vec{u}$, $\vec{v}$, and $\vec{w}$ is a filled-in triangle with vertices at $\vec{u},\vec{v}$ and $\vec{w}$. The point P lies inside this triangle, while Q does not.

To argue algebraically, suppose $P=t_1\vec{u}+t_2\vec{v}+t_3\vec{w}$ and $t_1+t_2+t_3=1$. From these assumptions, we can set up a system of equations which has a unique solution

$$P=\frac{1}{4}\vec{u}+\frac{1}{4}\vec{v}+\frac{1}{2}\vec{w},$$

and so P is a convex linear combination of $\vec{u}$, $\vec{v}$, and $\vec{w}$. The same procedure with Q gives a unique solution with coefficients $t_1=\frac{5}{9}, t_2=-\frac{1}{9}, t_3=\frac{5}{9}$, and so Q is not a convex linear combination of $\vec{u}$, $\vec{v}$, and $\vec{w}$.

Solutions for Module 3

1 (a) A is linearly dependent.

(b) span(A) is the plane $\left\{\begin{bmatrix}x\\y\\0\end{bmatrix}\in\mathbb{R}^3 : x,y\in\mathbb{R}\right\}$.

© Jason Siefken, 2015–2022

(c) Yes. If $A' = A \cup \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$, then $\operatorname{span}(A') = \mathbb{R}^3$.

2 (a) Line
  (b) Line
  (c) Plane
  (d) Line
  (e) Plane
  (f) Point
  (g) Plane
  (h) Volume
  (i) Plane
  (j) Plane

3 (a) i. Linearly Independent
    ii. Linearly Dependent
    iii. Linearly Independent
    iv. Lineraly Dependent
    v. Linearly Independent
    vi. Linearly Independent
    vii. Linearly Dependent
    viii. Linearly Independent
    ix. Linearly Independent
    x. Linearly Dependent
  (b) Linearly Dependent
  (c) No. The solutions to the vector equation $\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + \ldots + \alpha_{n+1} \vec{x}_{n+1} = 0$ for $\alpha_1, \alpha_2, \ldots, \alpha_{n+1} \in \mathbb{R}$ are the solutions to a system of n equations in $n+1$ variables. This system is consistent since $\alpha_1 = \alpha_2 = \ldots = \alpha_{n+1} = 0$ is a solution. The row reduced echelon form of the corresponding augmented matrix has at least one free variable column since there are more columns than rows. Hence there are infinitely many solutions, and in particular there exists a non-trivial solution to the above vector equation.

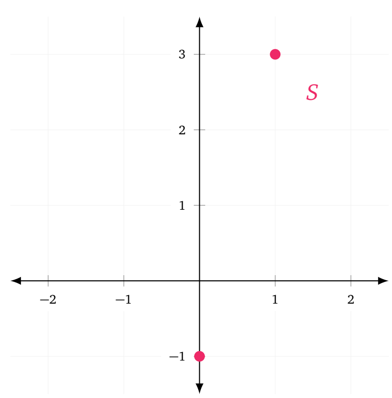
4 (a) i. $\operatorname{span}\left\{ \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}$
    ii. $\operatorname{span}\left\{ \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right\}$
    iii. $\operatorname{span}\left\{ \begin{bmatrix} 4 \\ 5 \end{bmatrix} \right\}$
    iv. Not possible since $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is not on the line.
    v. Not possible since $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is not on the line.
  (b) 4(a)iv: $\operatorname{span}\left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\} + \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$
    4(a)v: $\operatorname{span}\left\{ \begin{bmatrix} 5 \\ 3 \end{bmatrix} \right\} + \left\{ \begin{bmatrix} 8/9 \\ 0 \end{bmatrix} \right\}$
  (c) i. Not possible since $\vec{0}$ is not in the plane.
    ii. $\operatorname{span}\left\{ \begin{bmatrix} 6 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$
    iii. $\operatorname{span}\left\{ \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$
    iv. Impossible since $\vec{0}$ is not in the plane.
    v. $\operatorname{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$
    vi. Not possible since $\vec{0}$ is not in the plane.
  (d) 4(c)i: $\operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\} + \left\{ \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \right\}$
    4(c)iv: $\operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} + \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$
    4(c)vi: $\operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right\} + \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$

5 (a) Same plane ($\mathbb{R}^2$).
  (b) Same plane ($2\begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 13 \end{bmatrix}$).
  (c) Different plane ($\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$ is not in the second plane).

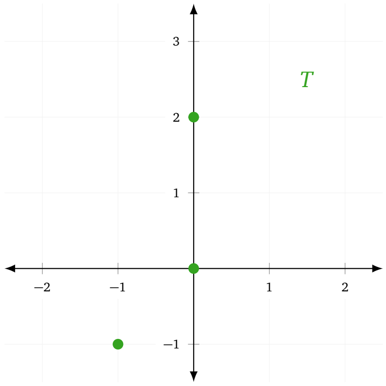
6 The set is linearly dependent since: $\begin{bmatrix} 6 \\ 4 \\ 11 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 7 \end{bmatrix} + 4\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ (geometric) or $\begin{bmatrix} 2 \\ 0 \\ 7 \end{bmatrix} + 4\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + (-1)\begin{bmatrix} 6 \\ 4 \\ 11 \end{bmatrix} = \vec{0}$ (algebraic).

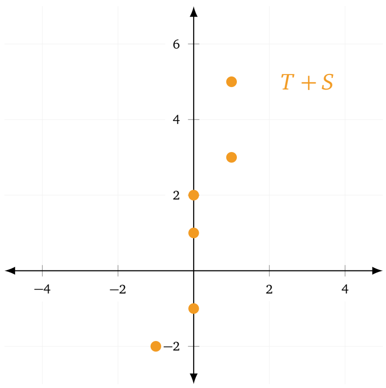
7 (a) $\vec{d}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \vec{d}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \vec{d}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \vec{p} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$
  (b) $\vec{d}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, and $\vec{d}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \vec{d}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \vec{p} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$
  (c) $\vec{d}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \vec{d}_2 = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \vec{d}_3 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \vec{p} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$
  (d) $\vec{d}_1 = \vec{d}_2 = \vec{d}_3 = \vec{0}, \vec{p} = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 3 \end{bmatrix}$

8 The set A is linearly independent. Since A contains no vectors, there is no way to write $\vec{0}$ as a non-trivial linear combination of vectors in A. However, B is linearly dependent, since $\vec{0} = 7\vec{0}$ is a non-trivial linear combination of vectors in B that gives the zero vector.

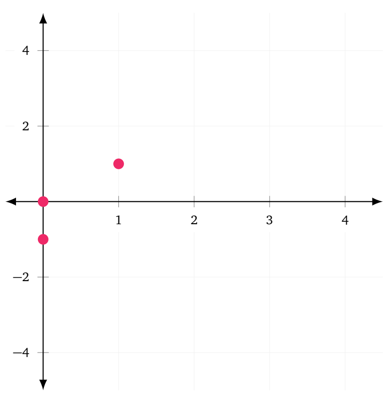
9


© Jason Siefken, 2015–2022

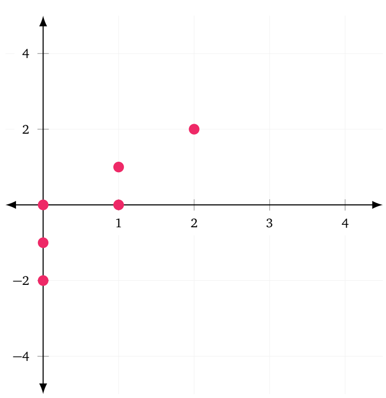




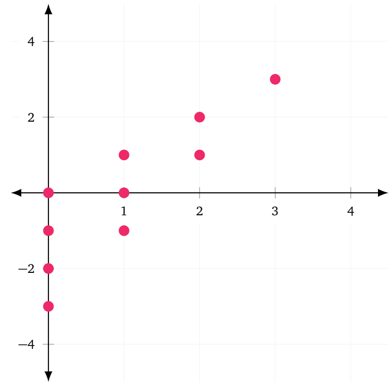
10 (a) The set S:



The set $S+S$:

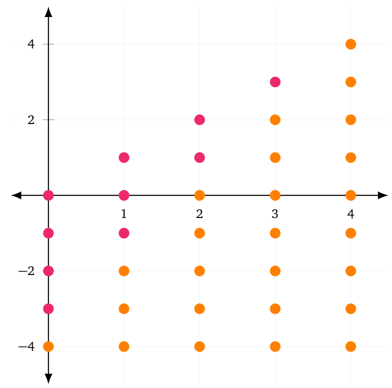


The set $(S+S)+S$:



(b) The set $(S+S)+S$ and $S+(S+S)$ are equal since vector addition is associative. This means that $S+S+S$ is well-defined and we can drop the parentheses, so the expression $S+S+S$ makes sense.

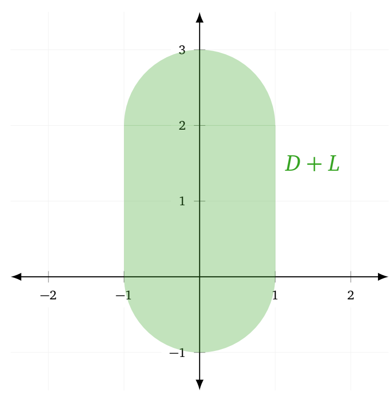
(c)



Notice that we obtain the orange points when we add an additional S to the set sum. Since we are considering the infinite sum $S+S+S+S+\cdots$, the process continues infinitely.

11 (a) There are infinitely many points in each of D, L, and $D+L$.

(b)



(c) $D+L$ can be decomposed into 2 unit radius half-circles and a square with side length 2. The area of $D+L$ is then $\pi+4$.

(d) We can take A to be a circle of radius 0.02 units.

12 (a) True. This follows from the geometric definition of linear dependence.

(b) False. $\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$ is linearly dependent, but $\begin{bmatrix}1\\0\end{bmatrix}$ is not a linear combination of $\begin{bmatrix}0\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$.

(c) True. $\vec{v}_1$ is a linear combination of $\vec{v}_2$ and so $\{\vec{v}_1,\vec{v}_2\}$ is linearly dependent by the definition of linear dependence.

© Jason Siefken, 2015–2022

(d) False. $\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$ is linearly dependent, but $\begin{bmatrix}1\\0\end{bmatrix}$ is not a scalar multiple of $\begin{bmatrix}0\\1\end{bmatrix}$.

(e) True. The linear combination of any finite set with all coefficients zero is $\vec{0}$.

Solutions for Module 4

1 (a) 78 (b) 2 (c) 57 (d) 12 (e) −1

2 (a) 2 (b) $\sqrt{14}$ (c) $4\sqrt{290}$

3 (a) Greater than 90° (b) Greater than 90° (c) Less than 90°

4 (a) $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}-1\\0\end{bmatrix}$

(b) $\dfrac{1}{\sqrt{5}}\begin{bmatrix}2\\-1\end{bmatrix}$ and $\dfrac{1}{\sqrt{5}}\begin{bmatrix}-2\\1\end{bmatrix}$

(c) $\dfrac{1}{\sqrt{10}}\begin{bmatrix}3\\-1\\0\end{bmatrix}$ and $\dfrac{1}{\sqrt{26}}\begin{bmatrix}5\\0\\-1\end{bmatrix}$

(d) $\dfrac{1}{\sqrt{6}}\begin{bmatrix}1\\-2\\1\end{bmatrix}$ and $\dfrac{1}{\sqrt{41}}\begin{bmatrix}0\\5\\4\end{bmatrix}$

(e) $\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$ and $\dfrac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\-1\\0\end{bmatrix}$

5 (a) 5 (b) $\sqrt{269}$ (c) $2\sqrt{3}$ (d) $\sqrt{5}$

6 (a) $\begin{bmatrix}0\\1\end{bmatrix}$ (b) $\begin{bmatrix}2\\3\\4\end{bmatrix}$

7 (a) $\vec{x} = t\begin{bmatrix}2\\0\\1\end{bmatrix} + s\begin{bmatrix}0\\3\\1\end{bmatrix} + \begin{bmatrix}2\\0\\0\end{bmatrix}$ and $\begin{bmatrix}3\\2\\-6\end{bmatrix}\cdot\left(\vec{x} - \begin{bmatrix}2\\0\\0\end{bmatrix}\right) = 0$

(b) $\vec{x} = t\begin{bmatrix}1\\1\\1\end{bmatrix} + s\begin{bmatrix}0\\1\\0\end{bmatrix}$; $\begin{bmatrix}1\\0\\-1\end{bmatrix}\cdot\vec{x} = 0$

8 (a) $\vec{x} = t\begin{bmatrix}-1\\1\\0\end{bmatrix} + s\begin{bmatrix}0\\1\\-1\end{bmatrix} + \begin{bmatrix}0\\1\\1\end{bmatrix}$ (b) $\vec{x} = t\begin{bmatrix}1\\1\\0\end{bmatrix} + s\begin{bmatrix}0\\0\\1\end{bmatrix} + \begin{bmatrix}1\\2\\3\end{bmatrix}$

9 (a) We get the value 56 on both sides.

(b) Writing the vectors in terms of the coefficients, we get

$$(x_1+y_1)z_1 + (x_2+y_2)z_2 = x_1z_1 + y_1z_1 + x_2z_2 + y_2z_2$$

which is always true.

In the general case, the left side will be the sum of $(x_i + y_i)z_i$ and the right side will be the sum of $x_iz_i + y_iz_i$. Since these two terms are always equal, the two sums are equal. So yes, the same conclusion hold true in all dimensions.

(c) We get the value 138 on both sides.

(d) Writing it in terms of the coefficients, we get

$$(kx_1)y_1 + (kx_2)y_2 = k(x_1y_1 + x_2y_2)$$

which is always true.

In the general case, the left side will be the sum of kx_iy_i and the right side will be the product of k and the sum of x_iy_i. Distributing the product of k onto the sum, we get that these two results are equal. So yes, the same conclusion hold true in all dimensions.

(e) The dot product distributes onto sums, just like the typical multiplication of real numbers.

12 (a) $\mathcal{A}$ is linearly dependent if $\vec{0} \in \mathcal{A}$, and is linearly independent otherwise.

(b) Suppose

$$\alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n = \vec{0}$$

for some $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$.

For any $i = 1, \ldots, n$, take the dot product of $\vec{v}_i$ with both sides of this equation, we have $\alpha_i\|\vec{v}_i\|^2 = 0$. Since $\vec{v}_i$ is a non-zero vector, this implies that $\alpha_i = 0$.

Then, $\alpha_1 = \cdots = \alpha_n = 0$ and therefore $\mathcal{A}$ is linearly independent.

Solutions for Module 5

1 The distances from $\begin{bmatrix}3\\1\end{bmatrix}$ to $\begin{bmatrix}0\\0\end{bmatrix}$, $\begin{bmatrix}-1\\2\end{bmatrix}$, and $\begin{bmatrix}1\\-2\end{bmatrix}$ are $\sqrt{10}$, $\sqrt{17}$, and $\sqrt{13}$, respectively. So $\operatorname{proj}_T\begin{bmatrix}3\\1\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$.

2 Suppose $\vec{v} = \begin{bmatrix}x\\y\end{bmatrix} \in C$. We would like to minimize $\left\|\begin{bmatrix}x\\y\end{bmatrix} - \begin{bmatrix}2\\0\end{bmatrix}\right\|$, or equivalently $\left\|\begin{bmatrix}x\\y\end{bmatrix} - \begin{bmatrix}2\\0\end{bmatrix}\right\|^2$. This expression can be rewritten as

$$\left\|\begin{bmatrix}x\\y\end{bmatrix} - \begin{bmatrix}2\\0\end{bmatrix}\right\|^2 = (x-2)^2 + y^2 = (x^2+y^2) - 4x + 4$$

$$= \|\vec{v}\|^2 - 4x + 4 = 5 - 4x.$$

Since $x \leq 1$, the above expression is minimized when $x = 1$ (and thus $y = 0$). That is,

$$\operatorname{proj}_C\begin{bmatrix}2\\0\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}.$$

3 (a) Let $\vec{u} = \begin{bmatrix}2\\1\end{bmatrix}$. Then $\operatorname{proj}_\ell \vec{v} = t\vec{u}$ for some $t \in \mathbb{R}$ which minimizes

$$\|\vec{v} - t\vec{u}\|^2 = \|\vec{u}\|^2\, t^2 - (2\vec{u}\cdot\vec{v})t + \|\vec{v}\|^2.$$

This quantity is minimized when $t = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|^2} = \frac{2}{5}$, so

$$\operatorname{proj}_\ell \vec{v} = \frac{2}{5}\vec{u} = \frac{1}{5}\begin{bmatrix}4\\2\end{bmatrix}.$$

© Jason Siefken, 2015–2022

(b) Let $\vec{u}=\begin{bmatrix}2\\1\end{bmatrix}$. Then $\mathrm{proj}_L\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}4\\0\end{bmatrix}+t\vec{u}$ for some $t\in\mathbb{R}$ which minimizes

$$\left\|\begin{bmatrix}1\\0\end{bmatrix}-\begin{bmatrix}4\\0\end{bmatrix}-t\vec{u}\right\|^2=(-3-2t)^2+(0-t)^2$$

$$=9+12t+5t^2.$$

The quantity $9+12t+5t^2=5(t+\frac{6}{5})^2+\frac{9}{5}$ is minimized when $t=-\frac{6}{5}$, so

$$\mathrm{proj}_L\vec{v}=\begin{bmatrix}4\\0\end{bmatrix}-\frac{6}{5}\vec{u}=\frac{1}{5}\begin{bmatrix}8\\-6\end{bmatrix}$$

(c) The set S is equal to $\{t\vec{u}:1\le t\le 2\}$ (check this), and so $S\subseteq\ell$. We found $\mathrm{proj}_\ell\vec{v}=\begin{bmatrix}4/5\\2/5\end{bmatrix}$ which is not in S. Therefore, $\mathrm{proj}_S\vec{v}$ must be one of the endpoints of S. Checking both endpoints, we conclude $\mathrm{proj}_S\vec{v}=\begin{bmatrix}2\\1\end{bmatrix}$.

4 Geometrically, T is a filled in triangle with vertices $\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}-1\\1\end{bmatrix}$ and $\begin{bmatrix}-1\\-2\end{bmatrix}$. So for points outside the triangle, the closest point in T will be on the nearest side of T, and so we can project onto T by projecting onto line segments.

(a) $\mathrm{proj}_T\begin{bmatrix}3\\3\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix}$; since $\begin{bmatrix}3\\3\end{bmatrix}$ is above T, the closest point will be on the line segment $y=1,-1\le x\le 1$.

(b) $\mathrm{proj}_T\begin{bmatrix}0\\0\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$, since $\vec{0}\in T$.

(c) $\mathrm{proj}_T\begin{bmatrix}1\\-2\end{bmatrix}=\frac{1}{13}\begin{bmatrix}-5\\-14\end{bmatrix}$; let ℓ be the line segment $\vec{x}=\begin{bmatrix}1\\1\end{bmatrix}+t\begin{bmatrix}-2\\-3\end{bmatrix},0\le t\le 1$. Then $\mathrm{proj}_T\begin{bmatrix}1\\-2\end{bmatrix}=\mathrm{proj}_\ell\begin{bmatrix}1\\-2\end{bmatrix}$, and then either by minimizing the length function, or drawing a perpendicular line to ℓ, we find the closest point is when $t=\frac{9}{13}$.

(d) $\mathrm{proj}_T\begin{bmatrix}0\\-4\end{bmatrix}=\begin{bmatrix}-1\\-2\end{bmatrix}$, as in (c), but now the minimizer is at $t>1$, so the constraints $0\le t\le 1$ force us to take the closest point on the line segment.

7 (a) Using the formula for vector components, we have

$$\mathrm{vcomp}_{\vec{u}}(\vec{a})=\frac{\vec{a}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\vec{u}=\frac{18}{9}\vec{u}=\begin{bmatrix}2\\4\\-4\end{bmatrix}$$

$$\mathrm{vcomp}_{\vec{v}}(\vec{a})=\frac{\vec{a}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\vec{v}=\frac{2}{2}\vec{v}=\begin{bmatrix}0\\1\\1\end{bmatrix}.$$

(b) Let $\vec{b}=\vec{a}-\mathrm{vcomp}_{\vec{u}}(\vec{a})-\mathrm{vcomp}_{\vec{v}}(\vec{a})$. Directly computing, we have $\vec{b}=\begin{bmatrix}4\\-1\\1\end{bmatrix}$.

To show $\vec{b}$ is orthogonal to P, we need to check that

$$\vec{b}\cdot\vec{u}=4-2-2=0$$

and

$$\vec{b}\cdot\vec{v}=0-1+1=0.$$

Hence $\vec{b}$ is a normal vector to P. (Note that this only worked because $\vec{u}\cdot\vec{v}=0$. Subtracting each vector component from $\vec{a}$ will not produce a normal vector in general.)

(c) Since $\vec{b}$ is orthogonal to P and $\vec{a}-\vec{b}$ is a linear combination of $\vec{u}$ and $\vec{v}$, the vector $\vec{a}-\vec{b}=\begin{bmatrix}2\\5\\-3\end{bmatrix}$ is the closest point to $\vec{a}$ on P.

Solutions for Module 6

1 (a) $\mathcal{T}$ is a subspace: it is non-empty since $\begin{bmatrix}0\\0\end{bmatrix}\in\mathcal{T}$.

i. Let $\begin{bmatrix}x_1\\y_1\end{bmatrix},\begin{bmatrix}x_2\\y_2\end{bmatrix}\in\mathcal{T}$. Then $3x_1-y_1=0=3x_2-y_2$. Hence $3(x_1+x_2)-(y_1+y_2)=(3x_1-y_1)+(3x_2-y_1)=0+0=0$ and $\begin{bmatrix}x_1\\y_1\end{bmatrix}+\begin{bmatrix}x_2\\y_2\end{bmatrix}\in\mathcal{T}$.

ii. Let $\begin{bmatrix}x\\y\end{bmatrix}\in\mathcal{T}$. Then $3x-y=0$. Hence for all $\alpha\in\mathbb{R}$, $3\alpha x-\alpha y=\alpha(3x-y)=\alpha\cdot 0=0$ and $\alpha\begin{bmatrix}x\\y\end{bmatrix}\in\mathcal{T}$.

(b) $\mathcal{U}$ is a subspace: it is non-empty since $\begin{bmatrix}0\\0\end{bmatrix}\in\mathcal{U}$.

i. Let $\begin{bmatrix}x_1\\y_1\end{bmatrix},\begin{bmatrix}x_2\\y_2\end{bmatrix}\in\mathcal{U}$. Then $\frac{1}{2}x_1-6y_1=0=\frac{1}{2}x_2-6y_2$. Hence $\frac{1}{2}(x_1+x_2)-6(y_1+y_2)=(\frac{1}{2}x_1-6y_1)+(\frac{1}{2}x_2-6y_2)=0+0=0$ and $\begin{bmatrix}x_1\\y_1\end{bmatrix}+\begin{bmatrix}x_2\\y_2\end{bmatrix}\in\mathcal{U}$.

ii. Let $\begin{bmatrix}x\\y\end{bmatrix}\in\mathcal{U}$. Then $\frac{1}{2}x-6y=0$. Hence for all $\alpha\in\mathbb{R}$, $\frac{1}{2}\alpha x-6\alpha y=\alpha(\frac{1}{2}x-6y)=\alpha\cdot 0=0$ and $\alpha\begin{bmatrix}x\\y\end{bmatrix}\in\mathcal{U}$.

(c) $\mathcal{V}$ is not a subspace, since for example $\begin{bmatrix}6\\1\end{bmatrix}\in\mathcal{V}$, but $0\cdot\begin{bmatrix}6\\1\end{bmatrix}=\vec{0}\notin\mathcal{V}$.

(d) $\mathcal{X}$ is a subspace: it is non-empty since $\begin{bmatrix}0\\0\\0\end{bmatrix}\in\mathcal{X}$.

i. Let $\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix},\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}\in\mathcal{X}$. Then $5x_1-\pi y_1+(\ln 2)z_1=0=5x_2-\pi y_2+(\ln 2)z_2$. Hence $5(x_1+x_2)-\pi(y_1+y_2)+(\ln 2)(z_1+z_2)=(5x_1-\pi y_1+(\ln 2)z_1)+(5x_2-\pi y_2+(\ln 2)z_2)=0+0=0$ and $\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix}+\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}\in\mathcal{X}$.

ii. Let $\begin{bmatrix}x\\y\\z\end{bmatrix}\in\mathcal{X}$. Then $5x-\pi y+(\ln 2)z=0$. Hence for all $\alpha\in\mathbb{R}$, $5\alpha x-\pi\alpha y+(\ln 2)\alpha z=\alpha(5x-\pi y+(\ln 2)z)=\alpha\cdot 0=0$ and $\alpha\begin{bmatrix}x\\y\\z\end{bmatrix}\in\mathcal{X}$.

© Jason Siefken, 2015–2022

(e) $\mathcal{Q}$ is a subspace: it is non-empty since $\vec{0} \in \mathcal{Q}$.

i. Let $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix} \in \mathcal{Q}$. Then $a_1x_1+\ldots+a_nx_n = 0 = a_1x_1'+\ldots+a_nx_n'$. Hence $a_1(x_1+x_1')+\ldots+a_n(x_n+x_n') = (a_1x_1+\ldots+a_nx_n)+(a_1x_1'+\ldots+a_nx_n') = 0+0 = 0$ and $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix} \in \mathcal{Q}$.

ii. Let $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathcal{Q}$. Then $a_1x_1+\ldots+a_nx_n = 0$. Hence for any scalar $\alpha \in \mathbb{R}$, $a_1\alpha x_1+\ldots+a_n\alpha x_n = \alpha(a_1x_1+\ldots+a_nx_n) = \alpha\cdot 0 = 0$ and $\alpha\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathcal{Q}$.

2 (a) $\mathcal{A}$ is not a subspace, since for example $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \mathcal{A}$, but $0\cdot\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \vec{0} \notin \mathcal{A}$.

(b) $\mathcal{B}$ is a subspace: it is non-empty since $\begin{bmatrix} -3 \\ 4 \end{bmatrix} \in \mathcal{B}$.

i. Let $\vec{u},\vec{v} \in \mathcal{B}$. Then $\vec{u} = t_1\begin{bmatrix} -3 \\ 4 \end{bmatrix}$ and $\vec{v} = t_2\begin{bmatrix} -3 \\ 4 \end{bmatrix}$ for some $t_1,t_2 \in \mathbb{R}$. But then $\vec{u}+\vec{v} = (t_1+t_2)\begin{bmatrix} -3 \\ 4 \end{bmatrix} \in \mathcal{B}$.

ii. Let $\vec{u} = t\begin{bmatrix} -3 \\ 4 \end{bmatrix} \in \mathcal{B}$. For any scalar $\alpha \in \mathbb{R}$, we have $\alpha\vec{u} = (\alpha t)\begin{bmatrix} -3 \\ 4 \end{bmatrix} \in \mathcal{B}$.

(c) $\mathcal{C}$ is a subspace: it is non-empty since $\begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} \in \mathcal{C}$.

i. Let $\vec{u},\vec{v} \in \mathcal{C}$. Then $\vec{u} = t_1\begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$ and $\vec{v} = t_2\begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$ for some $t_1,t_2 \in \mathbb{R}$. But then $\vec{u}+\vec{v} = (t_1+t_2)\begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} \in \mathcal{C}$.

ii. Let $\vec{u} = t\begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} \in \mathcal{C}$. For any scalar $\alpha \in \mathbb{R}$, we have $\alpha\vec{u} = (\alpha t)\begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} \in \mathcal{C}$.

(d) $\mathcal{D}$ is not a subspace, since for example $\begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix} \in \mathcal{D}$, but $0\cdot\begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix} = \vec{0} \notin \mathcal{D}$. ($\begin{bmatrix} 0 \\ 0 \\ -6 \end{bmatrix}$ cannot be written as a linear combination of $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ and $\begin{bmatrix} 10 \\ 20 \\ 131 \end{bmatrix}$.)

(e) $\mathcal{E}$ is a subspace: it is non-empty since $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \in \mathcal{E}$.

i. Let $\vec{u},\vec{v} \in \mathcal{E}$. Then $\vec{u} = t_1\begin{bmatrix} 5 \\ 7 \\ 1 \end{bmatrix} + s_1\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$ and $\vec{v} = t_2\begin{bmatrix} 5 \\ 7 \\ 1 \end{bmatrix} + s_2\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$ for some $t_1,t_2,s_1,s_2 \in \mathbb{R}$. But then $\vec{u}+\vec{v} = (t_1+t_2)\begin{bmatrix} 5 \\ 7 \\ 1 \end{bmatrix} + (s_1+s_2)\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \in \mathcal{E}$.

ii. Let $\vec{u} = t\begin{bmatrix} 5 \\ 7 \\ 1 \end{bmatrix} + s\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \in \mathcal{E}$. For any scalar $\alpha \in \mathbb{R}$, we have $\alpha\vec{u} = (\alpha t)\begin{bmatrix} 5 \\ 7 \\ 1 \end{bmatrix} + (\alpha s)\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \in \mathcal{E}$.

3 (a) Let $\mathcal{A}$ be $\operatorname{span}\left\{\begin{bmatrix} 0 \\ 1 \end{bmatrix},\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$. $\mathcal{A}$ is a subspace: it is non-empty since $\vec{0} \in \mathcal{A}$.

i. Let $\vec{u},\vec{v} \in \mathcal{A}$. Then $\vec{u} = t_1\begin{bmatrix} 0 \\ 1 \end{bmatrix} + s_1\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\vec{v} = t_2\begin{bmatrix} 0 \\ 1 \end{bmatrix} + s_2\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ for some $t_1,t_2,s_1,s_2 \in \mathbb{R}$. But then $\vec{u}+\vec{v} = (t_1+t_2)\begin{bmatrix} 0 \\ 1 \end{bmatrix} + (s_1+s_2)\begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \mathcal{A}$.

ii. Let $\vec{u} = t\begin{bmatrix} 0 \\ 1 \end{bmatrix} + s\begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \mathcal{A}$. For any scalar $\alpha \in \mathbb{R}$, we have $\alpha\vec{u} = (\alpha t)\begin{bmatrix} 0 \\ 1 \end{bmatrix} + (\alpha s)\begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \mathcal{A}$.

(b) Let $\mathcal{B}$ be $\operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}\right\}$. Observe that since $\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, this is the same thing as saying $\mathcal{B} = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\}$.

$\mathcal{B}$ is a subspace: it is non-empty since $\vec{0} \in \mathcal{B}$.

i. Let $\vec{u},\vec{v} \in \mathcal{B}$. Then $\vec{u} = t_1\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + s_1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\vec{v} = t_2\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + s_2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ for some $t_1,t_2,s_1,s_2 \in \mathbb{R}$. But then $\vec{u}+\vec{v} = (t_1+t_2)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + (s_1+s_2)\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \in \mathcal{B}$.

ii. Let $\vec{u} = t\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + s\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \in \mathcal{B}$. For any scalar $\alpha \in \mathbb{R}$, we have $\alpha\vec{u} = (\alpha t)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + (\alpha s)\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \in \mathcal{B}$.

© Jason Siefken, 2015–2022

4 (a) Property (i) fails, since $\begin{bmatrix}4\\0\\0\end{bmatrix}$ is in the set, but $\begin{bmatrix}4\\0\\0\end{bmatrix}+\begin{bmatrix}4\\0\\0\end{bmatrix}=\begin{bmatrix}8\\0\\0\end{bmatrix}$ is not.

Property (ii) fails, since $\begin{bmatrix}4\\0\\0\end{bmatrix}$ is in the set, but $0\cdot\begin{bmatrix}4\\0\\0\end{bmatrix}=\vec{0}$ is not.

(b) Non-emptiness fails, since it is the empty set.

(c) Property (i) fails, since $\begin{bmatrix}1\\1\end{bmatrix}$ is in the set ($1=1^2$), but $\begin{bmatrix}1\\1\end{bmatrix}+\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}2\\2\end{bmatrix}$ is not ($2\neq 2^2$).

Property (ii) fails, since $\begin{bmatrix}1\\1\end{bmatrix}$ is in the set, but $2\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}2\\2\end{bmatrix}$ is not.

(d) Property (ii) fails, since $\begin{bmatrix}1\\0\end{bmatrix}$ is in the set, but $-\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}-1\\0\end{bmatrix}$ is not.

(e) None of the properties fail. The set is $\{\vec{0}\}$, the trivial subspace.

5 (a) Yes, it is a subspace. Its dimension is 1 and a basis for it is $\left\{\begin{bmatrix}2\\3\end{bmatrix}\right\}$.

(b) Yes, it is a subspace. Its dimension is 2 and a basis for it is $\left\{\begin{bmatrix}1\\0\\-2\end{bmatrix},\begin{bmatrix}0\\2\\5\end{bmatrix}\right\}$.

(c) Yes, it is a subspace. Its dimension is 2 and a basis for it is $\left\{\begin{bmatrix}6\\1\\1\end{bmatrix},\begin{bmatrix}6\\0\\6\end{bmatrix}\right\}$.

(d) Yes, it is a subspace. Its dimension is 1 and a basis for it is $\left\{\begin{bmatrix}2\\2\\3\end{bmatrix}\right\}$.

(e) Yes, it is a subspace. Its dimension is 2 and a basis for it is $\left\{\begin{bmatrix}6\\6\\2\end{bmatrix},\begin{bmatrix}2\\1\\0\end{bmatrix}\right\}$.

(f) Yes, it is a subspace. Its dimension is 3 and a basis for it is $\left\{\begin{bmatrix}3\\-1\\0\\0\end{bmatrix},\begin{bmatrix}3\\0\\-1\\0\end{bmatrix},\begin{bmatrix}7\\0\\0\\-1\end{bmatrix}\right\}$.

6 (a) Not a basis, as it is not linearly independent: the third vector is equal to the sum of the other two.

(b) Not a basis, since four vectors in $\mathbb{R}^3$ cannot be linearly independent.

(c) Not a basis, as two vectors cannot span all of $\mathbb{R}^3$. You need at least three vectors.

(d) It is a basis.

7 (a) True by the Subspace-Span Theorem.

(b) True by the Subspace-Span Theorem.

(c) False. Translated spans cannot all be expressed as spans (since they do not need to contain $\vec{0}$, but spans do), but the Subspace-Span Theorem says that all subspaces can be expressed as spans.

(d) False. By the definition of a subspace, they cannot be the empty set.

(e) False. By the definition of a subspace, since $\begin{bmatrix}1\\2\end{bmatrix}$ and $\begin{bmatrix}2\\3\end{bmatrix}$ are in the set, $\begin{bmatrix}1\\2\end{bmatrix}+\begin{bmatrix}2\\3\end{bmatrix}=\begin{bmatrix}3\\5\end{bmatrix}$ should also be in the set. But it isn't.

8 (i) $\operatorname{span}\left\{\begin{bmatrix}1\\2\\3\\4\end{bmatrix}\right\}$ and the line with vector form $\vec{x}=t\begin{bmatrix}2\\2\\3\\0\end{bmatrix}$.

(ii) $\operatorname{span}\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\\0\end{bmatrix}\right\}$ and the volume with equation $x_1+x_2+x_3+x_4=0$.

The set $\{\vec{0}\}$ is the only 0-dimensional subspace.

9 As we've seen in problem 8, the dimension of $\mathcal{G}$ can be as small as 0. This is our lower bound.

As for the upper bound, we know that $n+1$ vectors in $\mathbb{R}^n$ cannot be linearly independent. So a basis for $\mathcal{G}$ can have at most n vectors. Hence n is an upper bound for the dimension of $\mathcal{G}$.

Solutions for Module 7

1 (a) $\begin{bmatrix}1&-1&1\\2&-1&1\\3&1&-1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}1\\2\\3\end{bmatrix}$

(b) $\begin{bmatrix}1&0&1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}6\end{bmatrix}$

(c) $\begin{bmatrix}5&-9&2\\0&-1&0\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\1\end{bmatrix}$

2 (a) If $\vec{x}=\begin{bmatrix}x\\y\\z\end{bmatrix}$ is orthogonal to $\begin{bmatrix}1\\2\\3\end{bmatrix}$ and $\begin{bmatrix}2\\2\\3\end{bmatrix}$, then

$$\vec{x}\cdot\begin{bmatrix}1\\2\\3\end{bmatrix}=0 \implies x+2y+3z=0$$

$$\vec{x}\cdot\begin{bmatrix}2\\2\\3\end{bmatrix}=0 \implies 2x+2y+3z=0.$$

This means we need to solve the system

$$\underbrace{\begin{bmatrix}1&2&3\\2&2&3\end{bmatrix}}_{A}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}.$$

© Jason Siefken, 2015–2022

Row reducing A yields

$$\operatorname{rref}(A)=\begin{bmatrix}1&0&0\\0&1&3/2\end{bmatrix},$$

so third column corresponds to a free variable. Let $z=t$, then $x=0$ and $y=-\frac{3t}{2}$. The complete solution expressed in vector form is

$$\vec{x}=t\begin{bmatrix}0\\-3/2\\1\end{bmatrix}\quad\text{or}\quad\vec{x}=t\begin{bmatrix}0\\-3\\2\end{bmatrix}.$$

(b) Proceeding as in 0a, we need to solve the matrix equation

$$\underbrace{\begin{bmatrix}0&5&6\\1&10&2\end{bmatrix}}_{B}\underbrace{\begin{bmatrix}x\\y\\z\end{bmatrix}}_{\vec{x}}=\begin{bmatrix}0\\0\end{bmatrix}.$$

We obtain

$$\operatorname{rref}(B)=\begin{bmatrix}1&0&-10\\0&1&6/5\end{bmatrix},$$

so the complete solution is

$$\vec{x}=t\begin{bmatrix}10\\-6/5\\1\end{bmatrix}\quad\text{or}\quad\vec{x}=t\begin{bmatrix}50\\-6\\5\end{bmatrix}.$$

(c) We need to solve the matrix equation

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

The only solution is $\vec{x}=\vec{0}$.

(d) We need to solve the matrix equation

$$\underbrace{\begin{bmatrix}2&6&-1\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}x\\y\\z\end{bmatrix}}_{\vec{x}}=\begin{bmatrix}0\end{bmatrix}.$$

Row reducing, we obtain

$$\operatorname{rref}(A)=\begin{bmatrix}1&3&-1/2\end{bmatrix}.$$

Both the second and third columns correspond to free variables. Let $y=s$ and $z=t$, then $x=-3s+\frac{t}{2}$, so we have

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

Therefore, the complete solution is

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

3 *Key fact to remember*: a normal vector is orthogonal to the direction vectors.

(a) If $\vec{n}=\begin{bmatrix}n_x\\n_y\\n_z\end{bmatrix}$ is a normal vector, then

$$\vec{n}\cdot\begin{bmatrix}0\\2\\2\end{bmatrix}=0\quad\Longrightarrow\quad 2n_y+2n_z=0$$

$$\vec{n}\cdot\begin{bmatrix}1\\1\\1\end{bmatrix}=0\quad\Longrightarrow\quad n_x+n_y+n_z=0.$$

Expressed in matrix form, this system becomes

$$\underbrace{\begin{bmatrix}0&2&2\\1&1&1\end{bmatrix}}_{A}\begin{bmatrix}n_x\\n_y\\n_z\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}.$$

Row reduction yields

$$\operatorname{rref}(A)=\begin{bmatrix}1&0&0\\0&1&1\end{bmatrix},$$

so complete solution in vector form is

$$\vec{n}=t\begin{bmatrix}0\\-1\\1\end{bmatrix}.$$

This means that any non-zero multiple of $\begin{bmatrix}0\\-1\\1\end{bmatrix}$ is a normal vector for this plane, so a normal form of the plane is

$$\begin{bmatrix}0\\-1\\1\end{bmatrix}\cdot\left(\begin{bmatrix}x\\y\\z\end{bmatrix}-\begin{bmatrix}1\\0\\3\end{bmatrix}\right)=0.$$

(b) We first wish to find all vectors $\vec{n}$ orthogonal to $\begin{bmatrix}1\\6\\8\end{bmatrix}$ and $\begin{bmatrix}2\\0\\2\end{bmatrix}$. To do this, we solve the system

$$\begin{bmatrix}1&6&8\\2&0&2\end{bmatrix}\begin{bmatrix}n_x\\n_y\\n_z\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}.$$

Row reducing, we have

$$\begin{bmatrix}1&6&8\\2&0&2\end{bmatrix}\to\begin{bmatrix}1&6&8\\0&-12&-14\end{bmatrix}\to\begin{bmatrix}1&6&8\\0&6&7\end{bmatrix}\to\begin{bmatrix}1&0&1\\0&6&7\end{bmatrix}$$

which has nullspace equal to $\operatorname{span}\left\{\begin{bmatrix}6\\7\\-6\end{bmatrix}\right\}$.

Since $\begin{bmatrix}0\\0\\9\end{bmatrix}$ is a point on the plane, we then get that a normal form of the plane is

$$\begin{bmatrix}6\\7\\-6\end{bmatrix}\cdot\left(\begin{bmatrix}x\\y\\z\end{bmatrix}-\begin{bmatrix}0\\0\\9\end{bmatrix}\right)=0.$$

© Jason Siefken, 2015–2022

(c) We first wish to find all vectors $\vec{n}$ orthogonal to $\begin{bmatrix}1\\5\\15\\20\end{bmatrix}$, $\begin{bmatrix}3\\0\\35\\59\end{bmatrix}$, and $\begin{bmatrix}1\\4\\0\\18\end{bmatrix}$. To do this we solve the system

$$\begin{bmatrix}1&5&15&20\\3&0&35&59\\1&4&0&18\end{bmatrix}\begin{bmatrix}n_x\\n_y\\n_z\\n_w\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Row reducing, we have

$$\begin{aligned}
\begin{bmatrix}1&5&15&20\\3&0&35&59\\1&4&0&18\end{bmatrix} &\to \begin{bmatrix}1&5&15&20\\3&0&35&59\\0&-1&-15&-2\end{bmatrix}\\
&\to \begin{bmatrix}1&5&15&20\\0&-15&-10&-1\\0&-1&-15&-2\end{bmatrix}\\
&\to \begin{bmatrix}1&5&15&20\\0&15&10&1\\0&1&15&2\end{bmatrix}\\
&\to \begin{bmatrix}1&5&15&20\\0&1&15&2\\0&15&10&1\end{bmatrix}\\
&\to \begin{bmatrix}1&5&15&20\\0&1&15&2\\0&0&-215&-29\end{bmatrix}\\
&\to \begin{bmatrix}1&0&-60&10\\0&1&15&2\\0&0&-215&-29\end{bmatrix}\\
&\to \begin{bmatrix}1&0&-60&10\\0&1&15&2\\0&0&215&29\end{bmatrix}\\
&\to \begin{bmatrix}1&0&-60&10\\0&1&0&-\frac{1}{43}\\0&0&215&29\end{bmatrix}\\
&\to \begin{bmatrix}1&0&0&\frac{778}{43}\\0&1&0&-\frac{1}{43}\\0&0&215&29\end{bmatrix}\\
&\to \begin{bmatrix}1&0&0&\frac{778}{43}\\0&1&0&-\frac{1}{43}\\0&0&1&\frac{29}{215}\end{bmatrix}
\end{aligned}$$

which has nullspace equal to $\operatorname{span}\left\{\begin{bmatrix}3890\\-5\\29\\-215\end{bmatrix}\right\}$.

Thus, since $\begin{bmatrix}1\\6\\0\\0\end{bmatrix}$ is a point on the hyperplane, we then get that the normal form of the hyperplane is

$$\begin{bmatrix}3890\\-5\\29\\-215\end{bmatrix}\cdot\left(\begin{bmatrix}x\\y\\z\\w\end{bmatrix}-\begin{bmatrix}1\\6\\0\\0\end{bmatrix}\right)=0.$$

4 We first express all the planes in the same form (any form works but we choose to use the equation form). $\mathcal{P}$ is already expressed in this form: $2x+4y-4z=7$. We are given $\mathcal{Q}$ in the form

$$\vec{x}=t\begin{bmatrix}-1\\2\\0\end{bmatrix}+s\begin{bmatrix}5\\0\\2\end{bmatrix}+\begin{bmatrix}0\\7\\1\end{bmatrix}.$$

Writing this into vector form yields

$$\begin{aligned}x&=-t+5s\\y&=2t+7\\z&=2s+1\end{aligned}$$

then substituting the above values yields

$$\begin{aligned}2x&=-2t+5(2s)\\2x&=-(y-7)+5(z-1)\\2x+y-5z&=2.\end{aligned}$$

Thus, having eliminated t and s from our original equations we find that the equation form of $\mathcal{Q}$ is $2x+y-5z=2$. We are given $\mathcal{R}$ in the form

$$\begin{bmatrix}2\\-8\\2\end{bmatrix}\cdot\left(\begin{bmatrix}x\\y\\z\end{bmatrix}-\begin{bmatrix}1\\7\\0\end{bmatrix}\right)=0$$

which, expanding the dot product, yields

$$\begin{aligned}\begin{bmatrix}2\\-8\\2\end{bmatrix}\cdot\begin{bmatrix}x\\y\\z\end{bmatrix}&=\begin{bmatrix}2\\-8\\2\end{bmatrix}\cdot\begin{bmatrix}1\\7\\0\end{bmatrix}\\2x-8y+2z&=2-7(8)\\2x-8y+2z&=-54.\end{aligned}$$

Thus the equation form of $\mathcal{R}$ is $2x-8y+2z=-54$. Having written all three planes in equation form, finding $\mathcal{P}\cap\mathcal{Q}\cap\mathcal{R}$ is the same as finding the solution set to the matrix equation

$$\begin{bmatrix}2&4&-4\\2&1&-5\\2&-8&2\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}7\\2\\-54\end{bmatrix}.$$

Row-reducing on both sides yields

$$\begin{bmatrix}2&4&-4\\2&1&-5\\2&-8&2\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}7\\2\\-54\end{bmatrix}$$

© Jason Siefken, 2015–2022

$$\to \begin{bmatrix}2&4&-4\\0&-3&-1\\2&-8&2\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}7\\-5\\-54\end{bmatrix}$$

$$\to \begin{bmatrix}2&4&-4\\0&-3&-1\\0&-12&6\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}7\\-5\\-61\end{bmatrix}$$

$$\to \begin{bmatrix}2&0&-\frac{16}{3}\\0&-3&-1\\0&-12&6\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}\frac{1}{3}\\-5\\-61\end{bmatrix}$$

$$\to \begin{bmatrix}2&0&-\frac{16}{3}\\0&-3&-1\\0&0&10\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}\frac{1}{3}\\-5\\-41\end{bmatrix}$$

$$\to \begin{bmatrix}2&0&-\frac{16}{3}\\0&-3&-1\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}\frac{1}{3}\\-5\\-\frac{41}{10}\end{bmatrix}$$

$$\to \begin{bmatrix}2&0&0\\0&-3&-1\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}-\frac{323}{15}\\-5\\-\frac{41}{10}\end{bmatrix}$$

$$\to \begin{bmatrix}2&0&0\\0&-3&0\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}-\frac{323}{15}\\-\frac{91}{10}\\-\frac{41}{10}\end{bmatrix}$$

$$\to \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}-\frac{323}{30}\\\frac{91}{30}\\-\frac{41}{10}\end{bmatrix}$$

so that $\mathcal{P}\cap\mathcal{Q}\cap\mathcal{R}$ consists of exactly one point.

$$\mathcal{P}\cap\mathcal{Q}\cap\mathcal{R}=\left\{\begin{bmatrix}-\frac{323}{30}\\\frac{91}{30}\\-\frac{41}{10}\end{bmatrix}\right\}.$$

Solutions for Module 8

1 (a) i. $[\vec{u}]_{\mathcal{E}}=\begin{bmatrix}1\\8\end{bmatrix}$, $[\vec{v}]_{\mathcal{E}}=\begin{bmatrix}-1\\3\end{bmatrix}$, and $[\vec{w}]_{\mathcal{E}}=\begin{bmatrix}2\\0\end{bmatrix}$.

ii. $[\vec{u}]_{\mathcal{A}}=\begin{bmatrix}3\\-2\end{bmatrix}$, $[\vec{v}]_{\mathcal{A}}=\begin{bmatrix}1\\-1\end{bmatrix}$, and $[\vec{w}]_{\mathcal{A}}=\begin{bmatrix}2/11\\4/11\end{bmatrix}$.

iii. $[\vec{u}]_{\mathcal{B}}=\begin{bmatrix}1/2\\1\end{bmatrix}$, $[\vec{v}]_{\mathcal{B}}=\begin{bmatrix}1/2\\-1\end{bmatrix}$, and $[\vec{w}]_{\mathcal{B}}=\begin{bmatrix}-5/11\\2\end{bmatrix}$.

(b) i. $[\vec{q}]_{\mathcal{E}}=\begin{bmatrix}0\\11\\-4\end{bmatrix}$, $[\vec{r}]_{\mathcal{E}}=\begin{bmatrix}5\\-12\\8\end{bmatrix}$, and $[\vec{s}]_{\mathcal{E}}=\begin{bmatrix}1\\-5\\2\end{bmatrix}$.

ii. We are given $\mathcal{D}=\left\{\begin{bmatrix}1\\2\\0\end{bmatrix}_{\mathcal{E}},\begin{bmatrix}-3\\5\\-4\end{bmatrix}_{\mathcal{E}},\begin{bmatrix}-8\\4\\11\end{bmatrix}_{\mathcal{E}}\right\}$.

To find $[\vec{q}]_{\mathcal{D}}$, we need to find scalars x,y,z such that

$$x\begin{bmatrix}1\\2\\0\end{bmatrix}+y\begin{bmatrix}-3\\5\\-4\end{bmatrix}+z\begin{bmatrix}-8\\4\\11\end{bmatrix}=\begin{bmatrix}0\\11\\-4\end{bmatrix},$$

which gives rise to a system of equations whose augmented matrix is

$$\left[\begin{array}{ccc|c}1&-3&-8&0\\2&5&4&11\\0&-4&11&-4\end{array}\right].$$

This can be row reduced to

$$\left[\begin{array}{ccc|c}1&0&0&3\\0&1&0&1\\0&0&1&0\end{array}\right].$$

Therefore, $[\vec{q}]_{\mathcal{D}}=\begin{bmatrix}3\\1\\0\end{bmatrix}$. Similarly we can find

$$[\vec{r}]_{\mathcal{D}}=\begin{bmatrix}-1\\-2\\0\end{bmatrix},\quad [\vec{s}]_{\mathcal{D}}=\begin{bmatrix}-66/67\\-39/67\\-2/67\end{bmatrix}.$$

iii. We are given $\mathcal{F}=\left\{\begin{bmatrix}1\\4\\4\end{bmatrix}_{\mathcal{E}},\begin{bmatrix}-3\\20\\0\end{bmatrix}_{\mathcal{E}},\begin{bmatrix}0\\21\\16\end{bmatrix}_{\mathcal{E}}\right\}$.

Therefore, $[\vec{q}]_{\mathcal{F}}=\begin{bmatrix}3\\1\\-1\end{bmatrix}$, $[\vec{r}]_{\mathcal{F}}=\begin{bmatrix}2\\-1\\0\end{bmatrix}$, and $[\vec{s}]_{\mathcal{F}}=\begin{bmatrix}-23/130\\-51/130\\11/65\end{bmatrix}$.

2 (a) Let $\mathcal{M}=\{\vec{u},\vec{v}\}$. Then $[\vec{a}]_{\mathcal{M}}=\begin{bmatrix}1\\0\end{bmatrix}\Rightarrow\vec{a}=1\vec{u}+0\vec{v}\Rightarrow\vec{a}=\vec{u}$. So as long as we choose the first vector in $\mathcal{M}$ as $\vec{a}$, any linearly independent second vector will do. Thus take $\mathcal{M}=\{5\vec{e}_1-12\vec{e}_2,\vec{e}_1\}$

(b) The observation here is that the numbers that appear in the coordinates in both $[\vec{b}]_{\mathcal{E}}$ and $[\vec{b}]_{\mathcal{N}}$ are same but they are just shuffled. Thus we can take $\mathcal{N}$ to be the corresponding permutation of $\mathcal{E}$. Take $\mathcal{N}=\{\vec{e}_3,\vec{e}_2,\vec{e}_1\}$

3 (a) Positively oriented. (Rotate both vectors 180 degrees clockwise and scale)

(b) Negatively oriented.

(c) Negatively oriented. ($\begin{bmatrix}1\\2\end{bmatrix}$ is to be transformed to $\vec{e}_1$. If we rotate it clockwise at some stage it points at the same direction as $\begin{bmatrix}2\\1\end{bmatrix}$, thus making it linearly dependent. If we try to rotate it counter clockwise, at some stage it also points at the same (negative) direction as $\begin{bmatrix}2\\1\end{bmatrix}$, where it becomes linerly dependent.)

(d) Positively oriented.

4 (a) Negatively oriented. (only $\vec{v}_3$ needs to be taken care of. To transform it to $\vec{e}_3$, no matter what we do, at some stage we have to cross the xy-plane. At that stage it becomes linearly dependent.)

(b) Negatively oriented. ($3\vec{v}_2$ can be scaled to $\vec{v}_2$. Then we are back at situation (a).)

(c) Positively oriented. (This is a little bit tricky. Note that $\vec{v}_2$ is already $\vec{e}_2$. So we can leave that alone and limit our operations to the xz-plane. In the xz-plane rotate $-4\vec{v}_1$ and $\vec{v}_3$ counter clockwise 180 degrees. Then $-4\vec{v}_1$ transforms to $4\vec{e}_1$, which then can be scaled to $\vec{e}_1$, and $\vec{v}_3$ transforms to $\vec{e}_3$.)

© Jason Siefken, 2015–2022

Solutions for Module 9

1 (a) Yes. $\mathcal{A}$ is a linear transformation. Let $\vec{u}=\begin{bmatrix}u_1\\u_2\end{bmatrix}$ and $\vec{v}=\begin{bmatrix}v_1\\v_2\end{bmatrix}$ be two vectors and α be a scalar. $\mathcal{A}(\vec{u}+\vec{v})=\mathcal{A}\begin{bmatrix}u_1+v_1\\u_2+v_2\end{bmatrix}=\begin{bmatrix}-u_1-v_1\\u_2+v_2\end{bmatrix}=\begin{bmatrix}-u_1\\u_2\end{bmatrix}+\begin{bmatrix}-v_1\\v_2\end{bmatrix}=\mathcal{A}\begin{bmatrix}u_1\\u_2\end{bmatrix}+\mathcal{A}\begin{bmatrix}v_1\\v_2\end{bmatrix}=\mathcal{A}(\vec{u})+\mathcal{A}(\vec{v})$ and $\mathcal{A}(\alpha\vec{u})=\mathcal{A}\begin{bmatrix}\alpha u_1\\\alpha u_2\end{bmatrix}=\begin{bmatrix}-\alpha u_1\\\alpha u_2\end{bmatrix}=\alpha\begin{bmatrix}-u_1\\u_2\end{bmatrix}=\alpha\mathcal{A}\begin{bmatrix}u_1\\u_2\end{bmatrix}=\alpha\mathcal{A}(\vec{u})$ So $\mathcal{A}$ satisfies all the properties of a linear transformation.

(b) No. $\mathcal{B}\begin{bmatrix}0\\0\end{bmatrix}=\begin{bmatrix}-1\\0\end{bmatrix}\neq\begin{bmatrix}0\\0\end{bmatrix}$

(c) Yes. Let $\vec{u}=\begin{bmatrix}u_1\\u_2\end{bmatrix}$ and $\vec{v}=\begin{bmatrix}v_1\\v_2\end{bmatrix}$ be two vectors and α be a scalar. $\mathbf{id}(\vec{u}+\vec{v})=\mathbf{id}\begin{bmatrix}u_1+v_1\\u_2+v_2\end{bmatrix}=\begin{bmatrix}u_1+v_1\\u_2+v_2\end{bmatrix}=\begin{bmatrix}u_1\\u_2\end{bmatrix}+\begin{bmatrix}v_1\\v_2\end{bmatrix}=\mathbf{id}\begin{bmatrix}u_1\\u_2\end{bmatrix}+\mathbf{id}\begin{bmatrix}v_1\\v_2\end{bmatrix}=\mathbf{id}(\vec{u})+\mathbf{id}(\vec{v})$ and $\mathbf{id}(\alpha\vec{u})=\mathbf{id}\begin{bmatrix}\alpha u_1\\\alpha u_2\end{bmatrix}=\begin{bmatrix}\alpha u_1\\\alpha u_2\end{bmatrix}=\alpha\begin{bmatrix}u_1\\u_2\end{bmatrix}=\alpha\,\mathbf{id}\begin{bmatrix}u_1\\u_2\end{bmatrix}=\alpha\,\mathbf{id}(\vec{u})$ So $\mathbf{id}$ satisfies all the properties of a linear transformation.

(d) No. As $\vec{0}$ is on the x-axis, $\mathcal{C}(\vec{0})=-\vec{e}_2\neq\vec{0}$.

2 (a) Notice $\mathcal{A}$ is just reflection about y-axis. If you reflect all points on the circle about the y-axis you get back the unit circle.

(b) $\mathcal{B}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}-x-1\\y\end{bmatrix}=\begin{bmatrix}-x\\y\end{bmatrix}+\begin{bmatrix}-1\\0\end{bmatrix}$. So the transformation $\mathcal{B}$ reflects over the y-axis and then translates. We already know the first operation keeps the circle as it is. So the final image is just a translated circle, that is unit circle with center at (-1,0).

(c) id leaves all points unchanged. So the image of the unit circle under id is still the unit circle.

(d) The unit circle under $\mathcal{C}$ is the set $\{(0,0),(0,-1)\}$.

3 (a) Recall that an $m\times n$ matrix represents a transformation from $\mathbb{R}^n\to\mathbb{R}^m$. The given matrix M is 2×3, so $T_M:\mathbb{R}^3\to\mathbb{R}^2$. Thus the domain of T_M is $\mathbb{R}^3$ and codomain is $\mathbb{R}^2$.

(b) $T_M\left(\begin{bmatrix}2\\-1\\1\end{bmatrix}\right)=\begin{bmatrix}1&-2&3\\-4&5&-6\end{bmatrix}\begin{bmatrix}2\\-1\\1\end{bmatrix}=\begin{bmatrix}7\\-19\end{bmatrix}$

(c) $T_M(\vec{e}_1)=\begin{bmatrix}1\\-4\end{bmatrix}$, $T_M(\vec{e}_2)=\begin{bmatrix}-2\\5\end{bmatrix}$, $T_M(\vec{e}_3)=\begin{bmatrix}3\\-6\end{bmatrix}$

4 (a) $\begin{bmatrix}2&0\\0&2\end{bmatrix}$

(b) To determine the image of $\vec{e}_1,\vec{e}_2$, we use trigonometry. Note $\mathcal{R}\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}-\cos(45^o)\\-\sin(45^o)\end{bmatrix}=\begin{bmatrix}-1/\sqrt{2}\\-1/\sqrt{2}\end{bmatrix}$ and $\mathcal{R}\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}\cos(45^o)\\-\sin(45^o)\end{bmatrix}=\begin{bmatrix}1/\sqrt{2}\\-1/\sqrt{2}\end{bmatrix}$. To see this draw pictures, draw angles and remember that $\vec{e}_1$ and $\vec{e}_2$ are unit vectors. Thus the matrix of $\mathcal{R}$ is $\begin{bmatrix}-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\end{bmatrix}$.

(c) No such matrix exists as this is not a linear transformation.

(d) $\begin{bmatrix}0&0\\0&1\end{bmatrix}$

(e) $\begin{bmatrix}0&1\\1&0\end{bmatrix}$

5 Let $\vec{u}$ and $\vec{v}$ be two vectors in $\mathbb{R}^n$ and c be a scalar. Then we have, $R(\vec{u}+\vec{v})=S(\vec{u}+\vec{v})+T(\vec{u}+\vec{v})$. Since given that both S,T are linear transformations, we get $S(\vec{u}+\vec{v})+T(\vec{u}+\vec{v})=S(\vec{u})+S(\vec{v})+T(\vec{u})+T(\vec{v})=(S(\vec{u})+T(\vec{u}))+(S(\vec{v})+T(\vec{v}))=R(\vec{u})+R(\vec{v})$. Hence we have, $R(\vec{u}+\vec{v})=R(\vec{u})+R(\vec{v})$. Similarly, $R(c\vec{u})=S(c\vec{u})+T(c\vec{u})=cS(\vec{u})+cT(\vec{u})=cR(\vec{u})$. Thus R is also linear.

6 (a) Since dot product is only defined for vectors in the same space, if $\vec{a}\in\mathbb{R}^3$, $\vec{x}$ must also be in $\mathbb{R}^3$. So the domain of $D_{\vec{a}}$ is $\mathbb{R}^3$.

(b) Let $\vec{u}=\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}$ and $\vec{v}=\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}$ be two vectors and α be a scalar. Then

$$\begin{aligned}D_{\vec{e}_1}(\vec{u}+\vec{v})&=\begin{bmatrix}1\\0\\0\end{bmatrix}\cdot\begin{bmatrix}u_1+v_1\\u_2+v_2\\u_3+v_3\end{bmatrix}=u_1+v_1\\&=\begin{bmatrix}1\\0\\0\end{bmatrix}\cdot\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}+\begin{bmatrix}1\\0\\0\end{bmatrix}\cdot\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\\&=D_{\vec{e}_1}(\vec{u})+D_{\vec{e}_1}(\vec{v})\end{aligned}$$

and

$$\begin{aligned}D_{\vec{e}_1}(\alpha\vec{u})&=\begin{bmatrix}1\\0\\0\end{bmatrix}\cdot\begin{bmatrix}\alpha u_1\\\alpha u_2\\\alpha u_3\end{bmatrix}=\alpha u_1\\&=\alpha\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\cdot\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}\right)\\&=\alpha D_{\vec{e}_1}(\vec{u})\end{aligned}$$

(c) Yes, $D_{\vec{a}}$ is a linear transformation for all vectors $\vec{a}$. The properties of linear transformation can be proved from properties of dot product. Let $\vec{u}$ and $\vec{v}$ be two vectors and α be a scalar. Then $D_{\vec{a}}(\vec{u}+\vec{v})=\vec{a}\cdot(\vec{u}+\vec{v})=\vec{a}\cdot\vec{u}+\vec{a}\cdot\vec{v}=D_{\vec{a}}(\vec{u})+D_{\vec{a}}(\vec{v})$ and

$$D_{\vec{a}}(\alpha\vec{u})=\vec{a}\cdot(\alpha\vec{u})=\alpha\,\vec{a}\cdot\vec{u}=\alpha D_{\vec{a}}(\vec{u})$$

(d) $D_{\vec{a}}:\mathbb{R}^3\to\mathbb{R}$. Thus the matrix, say M_D, will be 1×3. If $\vec{a}=\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}$ Then, $M_D=\begin{bmatrix}a_1&a_2&a_3\end{bmatrix}$

7 No. Here is a (counter) example of a transformation which is not linear : $T:\mathbb{R}^2\to\mathbb{R}^2$ is such that T keeps the coordinates of the points on x-axis unchanged and sends every other point to $\vec{0}$. That is, T is given by $T\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$ if $y\neq 0$, and $T\begin{bmatrix}x\\0\end{bmatrix}=\begin{bmatrix}x\\0\end{bmatrix}$.

Then T satisfies the given condition : $T(\mathbb{R}^2)$ is the x-axis which is a subspace. $T(\vec{0})=\vec{0}\Rightarrow\{T(\vec{0})\}=\{\vec{0}\}$, so T

© Jason Siefken, 2015–2022

sends the zero subspace to the zero subspace. Other subspaces of $\mathbb{R}^2$ are lines through the origin. T sends all such lines except for the x-axis to the subspace $\{\vec{0}\}$. T sends the subspace x-axis to x-axis. Thus $T(V)$ is a subspace whenever V is a subspace. But clearly T is not linear. Take for example : $T\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$ by construction. But $\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}+\begin{bmatrix}0\\1\end{bmatrix}$ and $\left(T\begin{bmatrix}1\\0\end{bmatrix}+T\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}1\\0\end{bmatrix}+\begin{bmatrix}0\\0\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}\neq T\begin{bmatrix}1\\1\end{bmatrix}$

8 (a) False. Take $T:\mathbb{R}^2\to\mathbb{R}^2$ given by $T\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}x-1\\0\end{bmatrix}$. If there were a matrix $\begin{bmatrix}a&b\\c&d\end{bmatrix}$ representing this transformation, then consider $T\begin{bmatrix}1\\0\end{bmatrix}$ and $T\begin{bmatrix}2\\0\end{bmatrix}$. $T\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$ which implies that $a=0,c=0$. But $T\begin{bmatrix}2\\0\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix}\begin{bmatrix}2\\0\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}$ which implies $a=\frac{1}{2},c=0$ — a contradiction! So no such matrix exists. The correct statement is the following: Every "linear" transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be represented by a matrix.

(b) False. This follows directly from the theorem given in this chapter that if $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation, then T takes subspaces to subspaces.

(c) False. The transformation is linear. Check the properties. Note: this transformation is sometimes called "the zero transformation.'

(d) False. A matrix is just a box of numbers, it has no meaning unless we give it meaning. We can specify a linear transformation by using a matrix, but a matrix by itself is not a linear transformation.

(e) False. There is a theorem in this module which explains this point. Suppose $T:\mathbb{R}^2\to\mathbb{R}^2$ is the tranformation which takes all vectors to the zero vector, ℓ_1 is given by the equation $y=x+1$ and ℓ_2 is given by the equation $y=x+2$. Then T is a linear transformation but $T(\ell_1)=T(\ell_2)=\{(0,0)\}$. Hence it doesn't make sense to say ℓ_1 and ℓ_2 are parallel under T.

9 (a) $\mathcal{T}$ can be defined by $\mathcal{T}\vec{x}=\begin{bmatrix}2&0\\0&2\end{bmatrix}\vec{x}$.

(b) Since $\mathcal{T}\vec{x}=2\vec{x}$ for every $\vec{x}\in\mathbb{R}^2$ we can define $\mathcal{T}$ as the function which multiplies its input by the scalar two.

(c) $\mathcal{T}$ is a linear transformation because $\mathcal{T}(\vec{x}+\vec{y})=2(\vec{x}+\vec{y})=2\vec{x}+2\vec{y}=\mathcal{T}\vec{x}+\mathcal{T}\vec{y}$ and $\mathcal{T}(\alpha\vec{x})=2\alpha\vec{x}=\alpha 2\vec{x}=\alpha\mathcal{T}\vec{x}$ for all $\vec{x},\vec{y}\in\mathbb{R}^2$ and all scalars α.

Solutions for Module 10

1 (a)

$$(\mathcal{R}\circ\mathcal{P})\left(\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}0\\0\end{bmatrix},\quad(\mathcal{R}\circ\mathcal{P})\left(\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\end{bmatrix}$$

$$(\mathcal{P}\circ\mathcal{R})\left(\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}0\\-\frac{1}{\sqrt{2}}\end{bmatrix},\quad(\mathcal{P}\circ\mathcal{R})\left(\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}0\\-\frac{1}{\sqrt{2}}\end{bmatrix}$$

We can use the effect of the transformation on the standard basis to compute the matrix.

i.

$$M_{\mathcal{R}\circ\mathcal{P}}=\begin{bmatrix}\mathcal{R}(\mathcal{P}(\vec{e}_1)) & \mathcal{R}(\mathcal{P}(\vec{e}_2))\end{bmatrix}=\begin{bmatrix}0&\frac{1}{\sqrt{2}}\\0&-\frac{1}{\sqrt{2}}\end{bmatrix}$$

ii.

$$M_{\mathcal{P}\circ\mathcal{R}}=\begin{bmatrix}\mathcal{P}(\mathcal{R}(\vec{e}_1)) & \mathcal{P}(\mathcal{R}(\vec{e}_2))\end{bmatrix}=\begin{bmatrix}0&0\\-\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\end{bmatrix}$$

iii. Since $M_{\mathcal{U}}=M_{\mathcal{P}\circ\mathcal{R}}$ and $\frac{1}{\sqrt{2}}=\frac{\sqrt{2}}{2}$, we can assert that $\mathcal{U}=\mathcal{P}\circ\mathcal{R}$.

(b)

$$(\mathcal{F}\circ\mathcal{S})\left(\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}0\\2\end{bmatrix},\quad(\mathcal{F}\circ\mathcal{S})\left(\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}2\\0\end{bmatrix}$$

$$(\mathcal{S}\circ\mathcal{F})\left(\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}0\\2\end{bmatrix},\quad(\mathcal{S}\circ\mathcal{F})\left(\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}2\\0\end{bmatrix}$$

We can use the effect of the transformation on the standard basis to compute the matrix.

i. $M_{\mathcal{F}\circ\mathcal{S}}=\begin{bmatrix}\mathcal{F}(\mathcal{S}(\vec{e}_1)) & \mathcal{F}(\mathcal{S}(\vec{e}_2))\end{bmatrix}=\begin{bmatrix}0&2\\2&0\end{bmatrix}$

ii. $M_{\mathcal{S}\circ\mathcal{F}}=\begin{bmatrix}\mathcal{S}(\mathcal{F}(\vec{e}_1)) & \mathcal{S}(\mathcal{F}(\vec{e}_2))\end{bmatrix}=\begin{bmatrix}0&2\\2&0\end{bmatrix}$

iii. $\mathcal{S}\circ\mathcal{F}=\mathcal{V}=\mathcal{F}\circ\mathcal{S}$.

2 (a)

$$\begin{aligned}M_{\mathcal{T}}\begin{bmatrix}1\\0\end{bmatrix}&=M_{\mathcal{A}}M_{\mathcal{B}}\begin{bmatrix}1\\0\end{bmatrix}=M_{\mathcal{A}}\begin{bmatrix}3&2\\0&4\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}\\&=\begin{bmatrix}2&2\\1&3\end{bmatrix}\begin{bmatrix}3\\0\end{bmatrix}=\begin{bmatrix}6\\3\end{bmatrix}\\M_{\mathcal{T}}\begin{bmatrix}0\\1\end{bmatrix}&=M_{\mathcal{A}}M_{\mathcal{B}}\begin{bmatrix}0\\1\end{bmatrix}=M_{\mathcal{A}}\begin{bmatrix}3&2\\0&4\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix}\\&=\begin{bmatrix}2&2\\1&3\end{bmatrix}\begin{bmatrix}2\\4\end{bmatrix}=\begin{bmatrix}12\\14\end{bmatrix}\end{aligned}$$

Therefore, $M_{\mathcal{T}}=\begin{bmatrix}6&12\\3&14\end{bmatrix}$.

(b) $M_{\mathcal{T}}=M_{\mathcal{A}}M_{\mathcal{B}}=\begin{bmatrix}2&2\\1&3\end{bmatrix}\begin{bmatrix}3&2\\0&4\end{bmatrix}=\begin{bmatrix}6&12\\3&14\end{bmatrix}$

3 (a) $\mathcal{T}=\mathcal{B}\circ\mathcal{A}$, thus $\mathcal{T}:\mathbb{R}^3\to\mathbb{R}$.

$$\mathcal{T}\begin{bmatrix}1\\0\\0\end{bmatrix}=(\mathcal{B}\circ\mathcal{A})\begin{bmatrix}1\\0\\0\end{bmatrix}=\mathcal{B}\begin{bmatrix}2\\2\end{bmatrix}=10$$

$$\mathcal{T}\begin{bmatrix}0\\1\\0\end{bmatrix}=(\mathcal{B}\circ\mathcal{A})\begin{bmatrix}0\\1\\0\end{bmatrix}=\mathcal{B}\begin{bmatrix}1\\3\end{bmatrix}=9$$

 © Jason Siefken, 2015–2022

$$\mathcal{T}\begin{bmatrix}0\\0\\1\end{bmatrix} = (\mathcal{B}\circ\mathcal{A})\begin{bmatrix}0\\0\\1\end{bmatrix} = \mathcal{B}\begin{bmatrix}0\\0\end{bmatrix} = 0.$$

Therefore $M_{\mathcal{T}} = \begin{bmatrix}10 & 9 & 0\end{bmatrix}$

(b) $M_{\mathcal{T}} = M_{\mathcal{B}}M_{\mathcal{A}} = \begin{bmatrix}3 & 2\end{bmatrix}\begin{bmatrix}2 & 1 & 0\\2 & 3 & 0\end{bmatrix} = \begin{bmatrix}10 & 9 & 0\end{bmatrix}$

Solutions for Module 11

1 (a) $M_1 = \begin{bmatrix}1 & 2 & 1\\3 & 1 & -2\\8 & 6 & -2\end{bmatrix}$; $\text{rref}(M_1) = \begin{bmatrix}1 & 0 & -1\\0 & 1 & 1\\0 & 0 & 0\end{bmatrix}$.

$\text{null}(M_1) = \{\vec{x}\in\mathbb{R}^3 : M_1\vec{x} = \vec{0}\}$; therefore, we need to solve $[M_1|\vec{0}]$.

$$\text{null}(M_1) = \left\{\vec{x}\in\mathbb{R}^3 : \vec{x} = s\begin{bmatrix}1\\-1\\1\end{bmatrix} \text{ for some } s\in\mathbb{R}\right\} = \text{span}\left\{\begin{bmatrix}1\\-1\\1\end{bmatrix}\right\}$$

$$\text{col}(M_1) = \text{span}\left\{\begin{bmatrix}1\\3\\8\end{bmatrix},\begin{bmatrix}2\\1\\6\end{bmatrix},\begin{bmatrix}1\\-2\\-2\end{bmatrix}\right\} = \text{span}\left\{\begin{bmatrix}1\\3\\8\end{bmatrix},\begin{bmatrix}2\\1\\6\end{bmatrix}\right\}$$

$$\text{row}(M_1) = \text{span}\left\{\begin{bmatrix}1\\2\\1\end{bmatrix},\begin{bmatrix}3\\1\\-2\end{bmatrix},\begin{bmatrix}8\\6\\-2\end{bmatrix}\right\} = \text{span}\left\{\begin{bmatrix}1\\2\\1\end{bmatrix},\begin{bmatrix}3\\1\\-2\end{bmatrix}\right\} = \text{span}\left\{\begin{bmatrix}1\\0\\-1\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$$

(b) $\text{rref}(M_2) = \begin{bmatrix}1 & 0 & 4/3\\0 & 1 & 1/2\end{bmatrix}$.

$$\text{null}(M_2) = \text{span}\left\{\begin{bmatrix}4/3\\1/2\\-1\end{bmatrix}\right\}$$

$$\text{col}(M_2) = \text{span}\left\{\begin{bmatrix}0\\3\end{bmatrix},\begin{bmatrix}2\\2\end{bmatrix}\right\} = \mathbb{R}^2$$

$$\text{row}(M_2) = \text{span}\left\{\begin{bmatrix}0\\2\\1\end{bmatrix},\begin{bmatrix}3\\2\\5\end{bmatrix}\right\}$$

(c) $\text{rref}(M_3) = \begin{bmatrix}1 & 0\\0 & 1\\0 & 0\end{bmatrix}$.

$$\text{null}(M_3) = \left\{\begin{bmatrix}0\\0\end{bmatrix}\right\}$$

$$\text{col}(M_3) = \text{span}\left\{\begin{bmatrix}1\\3\\4\end{bmatrix},\begin{bmatrix}2\\1\\0\end{bmatrix}\right\}$$

$$\text{row}(M_3) = \text{span}\left\{\begin{bmatrix}1\\2\end{bmatrix},\begin{bmatrix}3\\1\end{bmatrix}\right\} = \mathbb{R}^2$$

(d) $\text{rref}(M_4) = I_{4\times 4}$.

$$\text{null}(M_4) = \left\{\begin{bmatrix}0\\0\\0\\0\end{bmatrix}\right\}$$

$$\text{col}(M_4) = \mathbb{R}^4$$

$$\text{row}(M_4) = \mathbb{R}^4$$

2 (a) Since T is the projection onto $\mathcal{P}$, $T(\vec{x}) = \vec{x}$ for any $\vec{x}\in\mathcal{P}$ and thus $\mathcal{P}\subseteq\text{range}(T)$.

Let $\vec{y}\in\text{range}(T)$. This means that $\vec{y} = T(\vec{x})$ for some $\vec{x}\in\mathbb{R}^3$. By definition of a projection, $\vec{y}$ is the closest point in $\mathcal{P}$ to $\vec{x}$, so $\vec{y}\in\mathcal{P}$ and thus $\text{range}(T)\subseteq\mathcal{P}$.

This shows that $\text{range}(T) = \mathcal{P}$.

Since $\text{range}(T) = \mathcal{P}$ is a plane, we have $\text{rank}(T) = \dim(\text{range}(T)) = \dim(\mathcal{P}) = 2$.

(b) By the rank-nullity theorem,

$$\text{rank}(T) + \text{nullity}(T) = \dim(\text{domain of } T).$$

Thus,

$$\text{nullity}(T) = \dim(\text{domain of } T) - \text{rank}(T) = 3 - 2 = 1.$$

Since $\vec{n} = \begin{bmatrix}3\\4\\5\end{bmatrix}$ is normal to $\mathcal{P}$, we have $T(\vec{n}) = \vec{0}$ and thus $\vec{n}\in\text{null}(T)$. Since $\text{nullity}(T) = 1$ and $\vec{n}$ is a non-zero vector in $\text{null}(T)$, we have

$$\text{null}(T) = \text{span}\left\{\begin{bmatrix}3\\4\\5\end{bmatrix}\right\}.$$

3 (a) Similar to Problem 2.(a),

$$\text{range}(P) = \{\text{line given by the equation } y = x\} = \text{span}\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$$

$$\text{null}(P) = \{\text{line given by the equation } y = -x\} = \text{span}\left\{\begin{bmatrix}1\\-1\end{bmatrix}\right\}$$

(b) Notice that if we have a non-zero vector, after rotation the vector still remains non-zero as rotation does not change magnitude of vectors. Thus, $\text{null}(R) = \left\{\begin{bmatrix}0\\0\end{bmatrix}\right\}$. From the rank-nullity theorem, $\text{rank}(R) = 2$ implies that $\text{range}(R) = \mathbb{R}^2$.

(c) Since $F\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x\\-y\end{bmatrix}$, if $F\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$, then $x = y = 0$. Thus, $\text{null}(F) = \left\{\begin{bmatrix}0\\0\end{bmatrix}\right\}$. From the rank-nullity theorem, $\text{rank}(F) = 2$ implies that $\text{range}(F) = \mathbb{R}^2$.

© Jason Siefken, 2015–2022

(d) Let M be the matrix of the transformation $\mathcal{M}$. $\mathrm{rref}(M) = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ So, $\mathrm{range}(M) = \mathrm{col}(M) = \mathrm{span}\left\{\begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}\right\}$ and $\mathrm{null}(M) = \mathrm{span}\left\{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}\right\}$.

(e) Notice that $\mathrm{range}(Q) \subseteq \mathbb{R}$. For any $t \in \mathbb{R}^1$, we have $Q\begin{bmatrix} t \\ 0 \\ 0 \end{bmatrix} = t + 0 = t$ and thus $t \in \mathrm{range}(Q)$. Hence, $\mathrm{range}(Q) = \mathbb{R}^1$. For $\vec{x} \in \mathbb{R}^3$, $Q(\vec{x}) = Q\begin{bmatrix} x \\ y \\ z \end{bmatrix} = x + z = 0$. Thus, $\mathrm{null}(Q)$ is the plane given by the equation $x + z = 0$ Alternatively, $\mathrm{null}(Q) = \mathrm{span}\left\{\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$

4 (a) $[\vec{v}]_{\mathcal{E}} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$ So, $[T\vec{v}]_{\mathcal{E}} = \begin{bmatrix} 7 & 5 \\ -2 & -2 \end{bmatrix}\begin{bmatrix} 3 \\ -3 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$
Thus $T\vec{v} = 6\vec{e}_1$.

(b) $[\vec{v}]_{\mathcal{E}} = \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}$ So, $[T\vec{v}]_{\mathcal{E}} = \begin{bmatrix} 3 & 7 & 5 \\ 1 & -2 & -2 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 26 \\ -6 \end{bmatrix}$
Thus $T\vec{v} = 26\vec{e}_1 - 6\vec{e}_2$ [Here $\vec{e}_1, \vec{e}_2 \in \mathbb{R}^2$]

5 (a) False. $\mathrm{range}(A)$ and $\mathrm{range}(A^T)$ need not even be in the same space.

For example, take $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, A^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

Then $\mathrm{range}(A) = \{xy-\text{plane in } \mathbb{R}^3\}$, whereas $\mathrm{range}(A^T) = \mathbb{R}^2$

(b) False. Consider, $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x^2 \\ y \end{bmatrix}$$

Then, $\mathrm{null}(T) = \left\{\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\}$ which is a subspace. But T is not linear.

(c) False. From rank-nullity theorem,

$$\mathrm{nullity}(T) \le \dim(\text{ domain of } T) \quad = m \quad \text{(in this case)}$$

So, for any $n > m$ this is false. Consider for example

$$T : \mathbb{R}^1 \to \mathbb{R}^2 \text{ given by } \quad T(x) = \begin{bmatrix} x \\ 0 \end{bmatrix}; \ \mathrm{nullity}(T) = 0$$

(d) False. This is false whenever $m > n$ (follows from rank-nullity theorem) Take for example: $T : \mathbb{R}^3 \to \mathbb{R}^2$ induced by the matrix $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$; we have that $\mathrm{rank}(T) = \dim(\mathrm{col} A) = 2$, and $\mathrm{nullity}(T) = \mathrm{nullity}(A) = 1$.

6 (a) $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

(d) A 3×4 matrix with rank 4 cannot exist. Examples of justifications are: 1) The column space of a 3×4 matrix is a subspace of $\mathbb{R}^3$ and so cannot be four dimensional; 2) A matrix with rank 4 has 4 pivots, but a 3×4 matrix can have at most 3 pivots, which is less than 4.

7 (a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

(b) The range of $\mathcal{P}$ is the xy-plane.

(c) The column space of $M_{\mathcal{P}}$ is $\mathrm{span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$, which is the xy-plane. The column space of a matrix is the same as the range of its induced transformation.

(d) The null space of $\mathcal{P}$ and $M_{\mathcal{P}}$ is $\mathrm{span}\left\{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$, which is the z-axis. The null space of a matrix is the same as the null space of its induced transformation.

Solutions for Module 12

1 (a) $\mathcal{S}$ is both one-to-one and onto. No two distinct vectors can be doubled to become the same vector, thus the transformation is one-to-one. Every vector is double of the vector that is half of itself, thus the transformation is onto. $\mathcal{S}$ is invertible, its inverse is the transformation that halves every vector.

(b) $\mathcal{R}$ is both one-to-one and onto. No two distinct vectors can be rotated by 72° clockwise to become the same vector, thus the transformation is one-to-one. Every vector is the 72° clockwise rotation of the vector that is the 72° counter-clockwise rotation of itself, thus the transformation is onto. $\mathcal{R}$ is invertible, its inverse is the transformation that rotates every vector counter-clockwise by 72°.

(c) $\mathcal{P}$ is neither one-to-one nor onto. Every vector on the x-axis is sent to the origin so there are distinct inputs that do not map to distinct outputs, thus the transformation is not one-to-one. Vectors that do not lie on the y-axis are not in the range of the transformation, thus the transformation is not onto. $\mathcal{P}$ is not invertible.

(d) $\mathcal{F}$ is both one-to-one and onto. No two distinct vectors can be reflected to become the same vector, thus the transformation is one-to-one. Every vector is the reflection of the vector that is the reflection of itself, thus the transformation is onto. $\mathcal{F}$ is invertible, its inverse is itself—the reflection about the line $y = x$.

© Jason Siefken, 2015–2022

(e) $\mathcal{T}$ is neither one-to-one nor onto. The RREF of $M_{\mathcal{T}}$ is $\begin{bmatrix}1&0&-1\\0&1&0\\0&0&0\end{bmatrix}$. $\operatorname{rank}(\mathcal{T})=\operatorname{rank}(M_{\mathcal{T}})=2$. by the rank-nullity theorem the nullity is 1. An entire line of vectors is getting mapped to $\vec{0}$ so there are distinct inputs that do not map to distinct outputs, thus the transformation is not one-to-one. The range of $\mathcal{T}$ = the column space of $M_{\mathcal{T}}$ which is not all of $\mathbb{R}^3$, Thus the transformation is not onto. $\mathcal{T}$ is not invertible.

(f) $\mathcal{U}$ is onto but is not one-to-one. Because $\mathcal{U}$ maps from $\mathbb{R}^3$ to $\mathbb{R}^2$, the rank is at most 2, so by the Rank-Nullity Theorem, the nullity must be at least 1. More than one vector is getting mapped to $\vec{0}$ so there are distinct inputs that do not map to distinct outputs, thus the transformation is not one-to-one. Since $\operatorname{rref}(M_{\mathcal{U}})$ is $\begin{bmatrix}1&0&-1\\0&1&2\end{bmatrix}$. $\operatorname{rank}(\mathcal{U})=\operatorname{rank}(M_{\mathcal{U}})=2$, which is the dimension of the codomain, thus the transformation is onto. $\mathcal{U}$ is not invertible.

2 (a) To find the inverse we row reduce M_1, applying each of the row operations to the identity matrix.

$$\left[\begin{array}{cc|cc}2&3&1&0\\1&-1&0&1\end{array}\right]$$

$$\left[\begin{array}{cc|cc}1&\frac{3}{2}&\frac{1}{2}&0\\1&-1&0&1\end{array}\right]$$

$$\left[\begin{array}{cc|cc}1&\frac{3}{2}&\frac{1}{2}&0\\0&-\frac{5}{2}&-\frac{1}{2}&1\end{array}\right]$$

$$\left[\begin{array}{cc|cc}1&\frac{3}{2}&\frac{1}{2}&0\\0&1&\frac{1}{5}&-\frac{2}{5}\end{array}\right]$$

$$\left[\begin{array}{cc|cc}1&0&\frac{1}{5}&\frac{3}{5}\\0&1&\frac{1}{5}&-\frac{2}{5}\end{array}\right]$$

So $M_1^{-1}=\begin{bmatrix}\frac{1}{5}&\frac{3}{5}\\\frac{1}{5}&-\frac{2}{5}\end{bmatrix}$.

(b) M_2 does not row reduce to the identity matrix so it is not invertible.

(c) $M_3^{-1}=\begin{bmatrix}3&-3&-2\\2&-2&-1\\-3&4&2\end{bmatrix}$

(d) M_4 is not invertible because it is not a square matrix.

(e) $M_5^{-1}=\frac{1}{23}\begin{bmatrix}-4&-1&12&3\\-8&-2&1&6\\-3&-18&9&8\\1&6&-3&5\end{bmatrix}$

5 (a) True. If $m<n$ then T cannot be onto. If $m>n$ then T cannot be one-to-one. In both cases, T cannot be invertible. We may conclude that the dimension of the domain and codomain must be equal in order for a transformation to be invertible.

(b) False. Elementary matricies are one row operation away from the identity matrix. The identity matrix of any size is always square. There is no identity matrix such that performing one row operation on it yields M. We may conclude that elementary matricies are always square.

(c) True. The inverse of an elementary matrix E is another elementary matrix E^{-1} which corresponds to the row operation that turns E into the identity matrix. E^{-1} is the “opposite" row operation.

(d) True. Let E_1 correspond to multiplying row 1 by 4. Let E_2 correspond to multiplying row 1 by 2. The product E_2E_1 corresponds to multiplying row 1 by 8, which is also a single row operation and thus has a corresponding elementary matrix.

(e) False. Let E_1 correspond to multiplying row 1 by 4. Let E_2 correspond to multiplying row 2 by 2. The product E_2E_1 corresponds to multiplying row 1 by 4 and row 2 by 2, which is not a single row operation and thus does not have a corresponding elementary matrix.

(f) True.

(g) True.

(h) False. Let $A=\begin{bmatrix}1&0&0\\0&1&0\end{bmatrix}$. Let $B=\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}$. $AB=I$. $BA=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}\neq I$. A and B are not invertible. It is required that $AB=BA=I$ in order for matrices A and B to be invertible.

Solutions for Module 13

1 (a) Note that by definition

$$A=\{2\vec{e}_1+\vec{e}_2,\vec{e}_1-2\vec{e}_2\}.$$

Since $[\vec{x}]_{\mathcal{A}}=\begin{bmatrix}1\\-1\end{bmatrix}$, it follows that

$$\vec{x}=(2\vec{e}_1+\vec{e}_2)-(\vec{e}_1-2\vec{e}_2)=\vec{e}_1+3\vec{e}_2.$$

It thus follows that $[\vec{x}]_{\mathcal{E}}=\begin{bmatrix}1\\3\end{bmatrix}$.

To find $[\vec{x}]_{\mathcal{B}}$, we must first express the two elements of $\mathcal{A}$ as linear combinations of the two elements of $\mathcal{B}$. This involves solving two systems of linear equations:

$$2\vec{e}_1+\vec{e}_2=x_1(3\vec{e}_1-\vec{e}_2)+x_2(-2\vec{e}_1+3\vec{e}_2)$$

and

$$\vec{e}_1-2\vec{e}_2=y_1(3\vec{e}_1-\vec{e}_2)+y_2(-2\vec{e}_1+3\vec{e}_2).$$

Written as column vectors, these are

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

and

$$\begin{bmatrix}1\\-2\end{bmatrix}=y_1\begin{bmatrix}3\\-1\end{bmatrix}+y_2\begin{bmatrix}-2\\3\end{bmatrix}.$$

Solving the systems and applying the definition, we obtain

$$[\vec{a}_1]_{\mathcal{B}}=\begin{bmatrix}x_1\\x_2\end{bmatrix}=\frac{1}{7}\begin{bmatrix}8\\5\end{bmatrix},\quad [\vec{a}_2]_{\mathcal{B}}=\begin{bmatrix}y_1\\y_2\end{bmatrix}=\frac{1}{7}\begin{bmatrix}-1\\-5\end{bmatrix}.$$

Now we have

$$[\vec{x}]_{\mathcal{B}}=[\vec{a}_1]_{\mathcal{B}}-[\vec{a}_2]_{\mathcal{B}}=\frac{1}{7}\begin{bmatrix}8\\5\end{bmatrix}-\frac{1}{7}\begin{bmatrix}-1\\-5\end{bmatrix}=\frac{1}{7}\begin{bmatrix}9\\10\end{bmatrix}.$$

© Jason Siefken, 2015–2022

(b) By definition, $[\mathcal{E}\leftarrow\mathcal{A}]$ is the matrix M satisfying

$$M[\vec{x}]_{\mathcal{A}}=[\vec{x}]_{\mathcal{E}}$$

for all vectors $\vec{x}$. In particular, it must satisfy

$$M[\vec{a}_1]_{\mathcal{A}}=[\vec{a}_1]_{\mathcal{E}},\quad M[\vec{a}_2]_{\mathcal{A}}=[\vec{a}_2]_{\mathcal{E}}.$$

It thus follows that M satisfies

$$M\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}2\\1\end{bmatrix},\quad M\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}1\\-2\end{bmatrix}.$$

Solving for M gives

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

To find $[\mathcal{A}\leftarrow\mathcal{E}]$, we simply need to compute M^{-1}. Following the explicit inverse formula for 2×2 matrices, we see that

$$[\mathcal{A}\leftarrow\mathcal{E}]=M^{-1}=\frac{1}{5}\begin{bmatrix}2&1\\1&-2\end{bmatrix}.$$

To compute $[\mathcal{B}\leftarrow\mathcal{A}]$, we note that

$$[\mathcal{B}\leftarrow\mathcal{A}]=[\mathcal{B}\leftarrow\mathcal{E}][\mathcal{E}\leftarrow\mathcal{A}].$$

We compute $[\mathcal{B}\leftarrow\mathcal{E}]$ readily as

$$[\mathcal{E}\leftarrow\mathcal{B}]=\begin{bmatrix}3&-2\\-1&3\end{bmatrix}.$$

Thus,

$$[\mathcal{B}\leftarrow\mathcal{E}]=[\mathcal{E}\leftarrow\mathcal{B}]^{-1}=\frac{1}{7}\begin{bmatrix}3&2\\1&3\end{bmatrix}.$$

Therefore,

$$[\mathcal{B}\leftarrow\mathcal{A}]=\frac{1}{7}\begin{bmatrix}3&2\\1&3\end{bmatrix}\begin{bmatrix}2&1\\1&-2\end{bmatrix}=\frac{1}{7}\begin{bmatrix}8&-1\\5&-5\end{bmatrix}.$$

Taking inverse gives $[\mathcal{A}\leftarrow\mathcal{B}]$.

2 (a) By definition, we have

$$\begin{aligned}[\vec{b}_1]_{\mathcal{E}}&=[\vec{a}_1]_{\mathcal{E}}=[2\vec{e}_1+\vec{e}_2]_{\mathcal{E}}\\&=\begin{bmatrix}2\\1\\0\end{bmatrix},\\ [\vec{b}_2]_{\mathcal{E}}&=[\vec{a}_1]_{\mathcal{E}}+[\vec{a}_2]_{\mathcal{E}}\\&=[(2\vec{e}_1+\vec{e}_2)+(\vec{e}_1-2\vec{e}_2)]_{\mathcal{E}}\\&=[3\vec{e}_1-\vec{e}_2]_{\mathcal{E}}\\&=\begin{bmatrix}3\\-1\\0\end{bmatrix},\\ [\vec{b}_3]_{\mathcal{E}}&=[\vec{a}_1]_{\mathcal{E}}+[\vec{a}_2]_{\mathcal{E}}+[\vec{a}_3]_{\mathcal{E}}\\&=[(2\vec{e}_1+\vec{e}_2)+(\vec{e}_1-2\vec{e}_2)+(\vec{e}_3)]_{\mathcal{E}}\\&=[3\vec{e}_1-\vec{e}_2+\vec{e}_3]_{\mathcal{E}}\\&=\begin{bmatrix}3\\-1\\1\end{bmatrix}.\end{aligned}$$

(b) We have that

$$[\mathcal{E}\leftarrow\mathcal{A}]=\begin{bmatrix}2&1&0\\1&-2&0\\0&0&1\end{bmatrix}.$$

By part (a), we have

$$[\mathcal{E}\leftarrow\mathcal{B}]=\begin{bmatrix}2&3&3\\1&-1&-1\\0&0&1\end{bmatrix}.$$

Inverting the former, we see that

$$[\mathcal{A}\leftarrow\mathcal{E}]=\frac{1}{5}\begin{bmatrix}2&1&0\\1&-2&0\\0&0&5\end{bmatrix}.$$

(c) We have

$$\begin{aligned}[\mathcal{A}\leftarrow\mathcal{B}]&=[\mathcal{A}\leftarrow\mathcal{E}][\mathcal{E}\leftarrow\mathcal{B}]\\&=\frac{1}{5}\begin{bmatrix}2&1&0\\1&-2&0\\0&0&5\end{bmatrix}\begin{bmatrix}2&3&3\\1&-1&-1\\0&0&1\end{bmatrix}\\&=\begin{bmatrix}1&1&1\\0&1&1\\0&0&1\end{bmatrix}.\end{aligned}$$

3 (a) $[\mathcal{T}]_{\mathcal{E}}=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$ and $[\mathcal{T}]_{\mathcal{B}}=\begin{bmatrix}-1&-2\\1&1\end{bmatrix}$

(b) $[\mathcal{T}]_{\mathcal{E}}=\begin{bmatrix}0&0\\0&1\end{bmatrix}$

To compute $[\mathcal{T}]_{\mathcal{B}}$, note that $\mathcal{T}\vec{b}_1=\mathcal{T}\vec{e}_1=\vec{0}$ and $\mathcal{T}\vec{b}_2=\mathcal{T}(\vec{e}_1+\vec{e}_2)=\vec{e}_2$. Further, $\vec{e}_2=-\vec{b}_1+\vec{b}_2$. Therefore, $[\mathcal{T}]_{\mathcal{B}}=\begin{bmatrix}0&-1\\0&1\end{bmatrix}$.

(c) $[\mathcal{T}]_{\mathcal{E}}=[\mathcal{T}]_{\mathcal{B}}=\begin{bmatrix}2&0\\0&2\end{bmatrix}$

(d) Note that $\mathcal{T}\vec{e}_1=\vec{e}_2$ and $\mathcal{T}\vec{e}_2=\vec{e}_1$, so $[\mathcal{T}]_{\mathcal{E}}=\begin{bmatrix}0&1\\1&0\end{bmatrix}$. To compute $[\mathcal{T}]_{\mathcal{B}}$, we simply observe that $\vec{b}_1=\vec{e}_1$ and $\vec{b}_2=\vec{e}_1+\vec{e}_2$. Thus, $\mathcal{T}\vec{b}_1=\vec{e}_2=-\vec{b}_1+\vec{b}_2$ and $\mathcal{T}\vec{b}_2=\vec{e}_1+\vec{e}_2=\vec{b}_2$. Therefore, $[\mathcal{T}]_{\mathcal{B}}=\begin{bmatrix}-1&0\\1&1\end{bmatrix}$.

4 (a) True. A square matrix M is invertible if and only if it is a change of basis matrix.

(b) True. Any square matrix M satsifies $M=IMI^{-1}$.

(c) True. The definition of similarity requires the existence of an invertible matrix P, i.e. a square matrix, and a matrix B such that $B=PAP^{-1}$. If A is not a square matrix, then either PA or AP^{-1} is not defined, so such a matrix P cannot exist.

(d) False. For example, $A=\begin{bmatrix}2&0\\0&-1\end{bmatrix}$ is not similar to its inverse.

Solutions for Module 14

© Jason Siefken, 2015–2022

1 The volume of $\mathcal{T}(C_2)$ is equal to the absolute value of the determinant of $\mathcal{T}$. We have that

$$[\mathcal{T}]_{\mathcal{E}} = \begin{bmatrix} 3 & -1 \\ 1 & -1/4 \end{bmatrix},$$

so $\det \mathcal{T} = -3/4 + 1 = 1/4$. Since this number is positive, it is also the desired volume.

2 We start by computing the determinant of $\mathcal{S}$. The determinant of $\mathcal{S}$ can be computed from $[\mathcal{S}]_{\mathcal{E}}$, which is given by

$$[\mathcal{S}]_{\mathcal{E}} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Since determinant is preserved by row operations of the form "add a multiple of one row to another", we can partially row reduce $[\mathcal{S}]_{\mathcal{E}}$ (using only that row operation) without changing the determinant. Thus, the determinant of $[\mathcal{S}]_{\mathcal{E}}$ is the same as the determinant of

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

This matrix is triangular, so the determinant is just the product of the entries on the diagonal. Therefore, $\det[\mathcal{S}]_{\mathcal{E}} = -2$. But volume is non-negative, so the volume of $\mathcal{S}(C_3)$ is 2.

3 (a) Computing, we see $\mathcal{T}(\vec{e}_1) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\mathcal{T}(\vec{e}_2) = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. Drawing these two vectors, we see that $\mathcal{T}(\vec{e}_1), \mathcal{T}(\vec{e}_2)$ can be continuously transformed back into $\vec{e}_1, \vec{e}_2$ while staying linearly independent the whole time. Therefore $\mathcal{T}$ is orientation preserving.

(b) $\det \mathcal{T}$ is equal to the determinant of the matrix

$$[\mathcal{T}]_{\mathcal{E}} = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix},$$

which is 1.

4 (a) The matrix for $\mathcal{S}$ in any basis is $\begin{bmatrix} 2/3 & 0 \\ 0 & 2/3 \end{bmatrix}$, so the determinant is $4/9$.

(b) $\mathcal{R}$ does not change volume or orientation so its determinant is 1.

(c) $\mathcal{F}$ does not change volume but it reverses orientation so its determinant is -1.

(d) Though the determinants of $\mathcal{P}$ and $\mathcal{Q}$ are both 0, the determinant of $\mathcal{G}$ is not zero! We can compute the standard matrix for $\mathcal{G}$ by noticing $\mathcal{G}(\vec{e}_1) = \begin{bmatrix} 13/10 \\ 1/10 \end{bmatrix}$ and $\mathcal{G}(\vec{e}_2) = \begin{bmatrix} 1/10 \\ 7/10 \end{bmatrix}$. Therefore

$$[\mathcal{G}]_{\mathcal{E}} = \begin{bmatrix} 13/10 & 1/10 \\ 1/10 & 7/10 \end{bmatrix}$$

and so $\det \mathcal{G} = 9/10$.

(e) The matrix $[\mathcal{T}]_{\mathcal{E}}$ is given by

$$[\mathcal{T}]_{\mathcal{E}} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1/3 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Subtracting the first row from the second gives the matrix

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

which has the same determinant as $[\mathcal{T}]_{\mathcal{E}}$, and since this matrix is upper triangular, its determinant is simply $2/3$.

(f) The map $\mathcal{J}$ maps every vector in $\mathbb{R}^3$ into $\operatorname{span}\left\{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$, hence $\mathcal{J}$ is not invertible. Therefore, $\det \mathcal{J} = 0$.

(g) The determinant of the composition of the two maps is just the product of the determinants of the two maps. The matrices for $\mathcal{K}$ and $\mathcal{H}$ (with respect to $\mathcal{E}$) are

$$\begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}$$

and each has determinant one, so the determinant of $\mathcal{K} \circ \mathcal{H}$ is also 1.

5 (a) Put $E_1 = \begin{bmatrix} 1 & 0 \\ -1/2 & 1 \end{bmatrix}$. Then

$$E_1 A = \begin{bmatrix} 2 & 3 \\ 0 & 7/2 \end{bmatrix}.$$

Put $E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 2/7 \end{bmatrix}$. Then

$$E_2 E_1 A = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}.$$

Put $E_3 = \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}$. Then

$$E_3 E_2 E_1 A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

Finally, put $E_4 = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix}$. Then

$$E_4 E_3 E_2 E_1 A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Therefore,

$$\det A = \det E_1^{-1} \det E_2^{-1} \det E_3^{-1} \det E_4^{-1} = 7.$$

(b)

(c)

(d) They have the same area.

6 (a) By row reducing and keeping track of our steps, we see that

$$E_5 E_4 E_3 E_2 E_1 A = I$$

where

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/3 \end{bmatrix} \qquad E_4 = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

© Jason Siefken, 2015–2022

$$E_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

Therefore $A = E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1}E_5^{-1}$. By thinking about the relationship between elementary matrices and determinants, we see that $\det E_1^{-1} = \det E_4^{-1} = \det E_5^{-1} = 1$ and that $\det E_2^{-1} = 2$ and $\det E_3^{-1} = 3$. Therefore $\det A = 6$.

(b) $\det(A^{-1}) = 1/\det(A) = 1/6$.

(c) We have

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

Note that

$$(E_1A)^T = A^TE_1^T = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 0 & 1 & 3 \end{bmatrix}.$$

This matrix is *lower* triangular, so the determinant is equal to the product of the diagonal entries, which is still 6. Further $\det(E_1^T) = 1$, and so $\det(A^T) = 6$.

7 (a) The rank of A is equal to n, since each elementary matrix has non-zero determinant and since A can be expressed as a product of elementary matrices, it also has non-zero determinant.

(b) The nullspace of A^{-1} is trivial (i.e. equal to $\{\vec{0}\}$), since A^{-1} is invertible.

8 (a) *Anna's argument is incorrect.*

Reason: Since $\mathcal{S}$ is a linear transformation on $\mathbb{R}^3$, its determinant is given by the signed change of *3-dimensional volume*. Anna's argument is incorrect because she considered the 2-dimensional volume of $\mathcal{S}(C_3)$.

Ella's argument is incorrect.

Reason: The determinant is defined for all linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$, no matter whether it is invertible or not.

Finally, $\det(\mathcal{S}) = 0$, because since $\mathcal{S}(C_3)$ is a *2-dimensional* object in $\mathbb{R}^3$, its *3-dimensional volume* is 0. Therefore, $\text{Vol Change}(\mathcal{S}) = 0$, and we conclude that $\det(\mathcal{S}) = 0$.

(b) *Anna's argument is incorrect.*

Reason: The determinant function is only defined for linear transformations with same domain and codomain.

Ella's argument is correct.

Finally, $\det(\mathcal{T})$ is undefined, because the domain and codomain of $\mathcal{T}$ are not the same.

Solutions for Module 15

1 (a) Every non-zero vector in $\mathbb{R}^2$ is an eigenvector with eigenvalue 3.

(b) $\text{char}(R)$ has no real root, so R has no real eigenvalue or eigenvectors.

(c) There are two eigenvalues. 0 is an eigenvalue with eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and 1 is an eigenvalue with eigenvector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

(d) There are two eigenvalues. -1 is an eigenvalue with eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and 1 is an eigenvalue with eigenvector $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

(e) $\text{char}(T) = -\lambda(\lambda^2 - 12\lambda - 12)$. Then, we have three eigevalues. 0 is an eigenvalue with eigenvector $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, $6+4\sqrt{3}$ is an eigenvalue with eigenvector $\begin{bmatrix} 2 \\ 2+\sqrt{3} \\ 2\sqrt{3} \end{bmatrix}$, and $6-4\sqrt{3}$ is an eigenvalue with eigenvector $\begin{bmatrix} 2 \\ 2-\sqrt{3} \\ 2-2\sqrt{3} \end{bmatrix}$.

(f) U is induced by a 2×3 matrix, and eigenvalues/eigenvectors are only defined for linear maps from $\mathbb{R}^n$ to itself. So, U has no eigenvalues or eigenvectors.

2 (a) By definition,

$$\begin{aligned} \text{char}(A) = \det(A - \lambda I) &= \det \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} \\ &= \lambda^2 - (a+d)\lambda + ad - bc. \end{aligned}$$

(b) By the quadratic formula, the discriminant Δ of $\text{char}(A)$ is $\Delta = (a+d)^2 - 4(ad-bc) = (a-d)^2 + 4bc$. So, A has two distinct real eigenvalues if $(a-d)^2 + 4bc > 0$, one real eigenvalue if $(a-d)^2 + 4bc = 0$, and no real eigenvalues if $(a-d)^2 + 4bc < 0$.

3 (a) $\text{char}(B) = \det \begin{bmatrix} 1-\lambda & 2 \\ 0 & 4-\lambda \end{bmatrix} = (1-\lambda)(4-\lambda)$. So, B has eigenvalues 1 and 4.

(b) $\text{char}(B) = \text{char}(B^T)$, so B^T also has eigenvalues 1 and 4.

(c) The equation $\vec{v}^TB = \lambda\vec{v}^T$ holds if and only if $B^T\vec{v} = \lambda\vec{v}$. Here, we interpret $\vec{v}$ as a column vector and $\vec{v}^T$ as a row vector.

We observe that B^T has eigenvector $\begin{bmatrix} -3 \\ 2 \end{bmatrix}$ with eigenvalue 1 and an eigenvector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ with eigenvalue 4. Hence B has left eigenvectors $\begin{bmatrix} -3 & 2 \end{bmatrix}$ with eigenvalue 1 and left eigenvector $\begin{bmatrix} 0 & 1 \end{bmatrix}$ with eigenvalue 4. It follows that all non-zero scalar multiples of $\begin{bmatrix} -3 & 2 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \end{bmatrix}$ are also left eigenvectors.

4 (a) False. 0 is an eigenvalue of $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

(b) True. An eigenvector is a nonzero vector by definition.

(c) False. $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ has no real eigenvalue.

(d) True. Its characteristic polynomial has degree 3 and hence has at least one real root.

(e) False. Eigenvalues are not defined for a non-square matrix.

© Jason Siefken, 2015–2022

(f) False. $\begin{bmatrix}3&3&3&3\\3&3&3&3\\3&3&3&3\\3&3&3&3\end{bmatrix}$ has 0 as an eigenvalue with eigenvector $\begin{bmatrix}1\\-1\\0\\0\end{bmatrix}$. It also has eigenvalue 12 with eigenvector $\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$.

(g) True. Any eigenvector with eigenvalue 0 lies in the null space of the matrix, which implies that the null space has dimension at least one.

(h) True. Suppose A was a non-invertible square matrix. Then, we must have nullity$(A) > 0$, and so null(A) contains at least one non-zero vector, $\vec{v}$. By definition $A\vec{v} = \vec{0} = 0\vec{v}$, and so $\vec{v}$ is an eigenvector for A with eigenvalue 0.

Solutions for Module 16

1 (a) i. $\lambda = 1$, Algebraic: 1, Geometric: 1
ii. $\lambda = 2$, Algebraic: 1, Geometric: 1

(b) i. $\lambda = 3$, Algebraic: 2, Geometric: 2

(c) i. $\lambda = 0$, Algebraic: 1, Geometric: 1
ii. $\lambda = 3$, Algebraic: 1, Geometric: 1

(d) i. $\lambda = 1$, Algebraic: 1, Geometric: 1
ii. $\lambda = 2$, Algebraic: 2, Geometric: 1

(e) i. $\lambda = 2$, Algebraic: 2, Geometric: 2
ii. $\lambda = 1$, Algebraic: 1, Geometric: 1

2 (a) $\begin{bmatrix}2&0\\0&1\end{bmatrix}$

(b) $\begin{bmatrix}3&0\\0&3\end{bmatrix}$

(c) $\begin{bmatrix}3&0\\0&0\end{bmatrix}$

(d) Not diagonalizable.

(e) $\begin{bmatrix}2&0&0\\0&2&0\\0&0&1\end{bmatrix}$

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

4 No. If λ is an eigenvalue of a matrix A, then $det(A-\lambda I) = 0$ and therefore $A - \lambda I$ is not invertible. Specifically, nullity$(A - \lambda I) \geq 1$ and hence there exists at least one eigenvector for the eigenvalue λ. Therefore the geometric multiplicity of λ is at least one.

5 (a) Let $\vec{v}_1$ and $\vec{v}_2$ be eigenvectors for a matrix M corresponding to distinct eigenvalues λ_1 and λ_2 respectively. Let $a, b \in \mathbb{R}$ be such that $a\vec{v}_1 + b\vec{v}_2 = \vec{0}$. Multiplying both sides by $M - \lambda_1 I$, we get

$$b(\lambda_2 - \lambda_1)\vec{v}_2 = \vec{0}.$$

Since $\vec{v}_2$ is an eigenvector, it is nonzero. Hence, either $b = 0$ or $\lambda_2 - \lambda_1 = 0$. Since $\lambda_1 \neq \lambda_2$ we know $\lambda_2 - \lambda_1 \neq 0$ and so $b = 0$.

We have deduced that $a\vec{v}_1 = \vec{0}$. However, since $\vec{v}_1$ is nonzero (because it is an eigenvector), we must have that $a = 0$. This means that $\vec{v}_1$ and $\vec{v}_2$ are linearly independent.

(b) $\begin{bmatrix}1&1&0\\0&1&0\\0&0&-1\end{bmatrix}$

(c) This is impossible. Suppose that for some matrix M, $\vec{v}_1$ is an eigenvector corresponding to 1 and $\vec{v}_2$ is an eigenvector corresponding to -1. By 5a, $\vec{v}_1$ and $\vec{v}_2$ are linearly independent and thus form a basis for $\mathbb{R}^2$. Since, $\mathbb{R}^2$ has a basis consisting of eignvectors of M, we know M is diagonalizable.

Solutions for Appendix 1

1 (a) Linear equation.

(b) Linear equation.

(c) Not a linear equation because of the $\cos(y)$ term.

(d) Not a linear equation because of the $3xy$ term.

(e) Linear equation.

(f) Not a linear equation because of the $\frac{x}{y}$ term. Note that it is *almost* equivalent to the equation $x = y$, but they are not equivalent because $x = 0, y = 0$ is a solution to the latter equation but not the former.

2 (a) $\begin{cases} x \quad\quad + 4z = 2 \\ -x + y + 6z = -5 \\ \quad\quad\quad\quad z = 2 \end{cases}$

(b) $\begin{cases} 7x + \ \ 8y = 11 \\ 16x + 13y = 30 \end{cases}$

(c) $\begin{cases} -5s + t = -1 \\ -3s + t = -1 \end{cases}$

3 (a) $x_1\begin{bmatrix}0\\1\\0\end{bmatrix} + x_2\begin{bmatrix}4\\0\\9\end{bmatrix} + x_3\begin{bmatrix}2\\2\\2\end{bmatrix} = \begin{bmatrix}0\\0\\1\end{bmatrix}$

(b) $x\begin{bmatrix}0\\1\end{bmatrix} + y\begin{bmatrix}0\\1\end{bmatrix} + z\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\3\end{bmatrix}$

4 (a) If $\vec{b} = \begin{bmatrix}7\\14\end{bmatrix}$, then the vector equation becomes

$$x\begin{bmatrix}2\\4\end{bmatrix} + y\begin{bmatrix}8\\16\end{bmatrix} = \begin{bmatrix}7\\14\end{bmatrix}.$$

Converting it to a system of linear equations and row reducing we get

$$\begin{cases} 2x + \ \ 8y = 7 \\ 4x + 16y = 14 \end{cases} \to \begin{cases} x + 4y = 3.5 \\ 0x + 0y = 0 \end{cases}.$$

The solution to this system is then

$$\begin{cases} x = 3.5 - 4t \\ y = t \end{cases} \quad (t \in \mathbb{R}).$$

This system is consistent.

© Jason Siefken, 2015–2022

(b) There are vectors $\vec{b}$ that makes the system consistent. For instance, any vector $\vec{b} = \vec{b} = \begin{bmatrix} t \\ 2t \end{bmatrix}$ where $t \in \mathbb{R}$ makes the system consistent. Since there are infinitely many real numbers, we conclude that there are infinitely many vectors $\vec{b}$ that makes the system consistent.

(c) If $\vec{b} = \begin{bmatrix} 5 \\ 14 \end{bmatrix}$, then the vector equation becomes

$$x\begin{bmatrix} 2 \\ 4 \end{bmatrix} + y\begin{bmatrix} 8 \\ 16 \end{bmatrix} = \begin{bmatrix} 5 \\ 12 \end{bmatrix}.$$

Converting it to a system of linear equations and row reducing we get

$$\begin{cases} 2x + 8y = 5 \\ 4x + 16y = 12 \end{cases} \to \begin{cases} x + 4y = 2.5 \\ 0x + 0y = 2 \end{cases}.$$

This system is inconsistent.

(d) There are vectors $\vec{b}$ that makes the system inconsistent. For instance, $\begin{bmatrix} 10 \\ 24 \end{bmatrix}$ is is such a vector. In general, any vector $\vec{b}$ with $\vec{b} = \begin{bmatrix} 5t \\ 12t \end{bmatrix}$ where $t \in \mathbb{R}$ ($t \neq 0$) makes the system inconsistent. Since there are infinitely many real numbers, we conclude that there are infinitely many vectors $\vec{b}$ that makes the system inconsistent.

5 (a) Let x be the number of chickens, and let y be the number of rabbits. Using the information given in the problem, we have

$$\begin{cases} x + y = 35 \\ 2x + 4y = 94 \end{cases}.$$

(b) Row reducing

$$\begin{cases} x + y = 35 \\ 2x + 4y = 94 \end{cases},$$

we get

$$\begin{cases} x + y = 35 \\ \phantom{x + {}} y = 12 \end{cases}.$$

This shows that the system is consistent. The solution to this system is $x = 23, y = 12$. Thus, there are 23 chickens and 12 rabbits in the farm.

(c) Before discussing each configuration, we point out that a configuration is possible if there exists a natural number solution to the system of linear equations associated with the configuration.

i. For the first configuration, let x be the number of cats, and let y be the number of dogs. Using the information given in the problem, we have

$$\begin{cases} x + y = 35 \\ 4x + 4y = 94 \end{cases}.$$

Row reducing this system, we get

$$\begin{cases} x + y = 35 \\ 0x + 0y = -46 \end{cases}.$$

This system is inconsistent, which means there's no solution to this system. Therefore, Kokoro's first configuration is not possible.

ii. For the second configuration, let x be the number of cats, and let y be the number of dogs. Using the information given in the problem, we have

$$\begin{cases} x + y = 35 \\ 4x + 4y = 140 \end{cases}.$$

Row reducing this system, we get

$$\begin{cases} x + y = 35 \\ 0x + 0y = 0 \end{cases}.$$

This system is consistent, and the complete solution is given by

$$\begin{cases} x \phantom{{}= y} = 35 - t \\ y = t \end{cases} (t \in \mathbb{R}).$$

Take $t = 1$, and we get a natural number solution $x = 34, y = 1$. (In fact, there is more than one natural number solution.) Therefore, Kokoro's second configuration is possible.

iii. For the third configuration, let x be the number of chickens, and let y be the number of rabbits. Using the information given in the problem, we have

$$\begin{cases} x + y = 42 \\ 2x + 4y = 77 \end{cases}.$$

Row reducing this system, we get

$$\begin{cases} x + y = 42 \\ \phantom{x + {}} y = -\frac{7}{2} \end{cases}.$$

This system is consistent and the unique solution is $x = \frac{91}{2}, y = -\frac{7}{2}$ However, there cannot be 91/2 of a chicken, so Kokoro's third configuration is not possible.

6 (a) False. A counterexample is given by

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 1 \\ x_1 + x_2 + x_3 + x_4 = 2 \\ x_1 + x_2 + x_3 + x_4 = 3 \end{cases}.$$

(b) False. A counterexample is given by

$$\begin{cases} 0x_1 + 0x_2 = 0 \\ 0x_1 + 0x_2 = 1 \end{cases}.$$

(c) False. Assume the y-axis can be represented as the complete solution to $y = mx + c$ for some m, c. Since $(x, y) = (0, 0)$ and $(x, y) = (0, 1)$ are both on the y axis, we know $0 = 0m + c$ and $1 = 0m + c$. This gives $0 = 1$, which is false. Therefore, there's no $m, c \in \mathbb{R}$ so that the y-axis is the solution set to the equation $y = mx + c$.

(d) True. Take $m = 0$, $c = 0$. The equation then becomes $y = 0$. A complete solution to this equation is given by $\begin{bmatrix} t \\ 0 \end{bmatrix}$ ($t \in \mathbb{R}$), which is exactly the x-axis.

(e) False. The x-axis in $\mathbb{R}^3$ can be described as $\left\{ \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^3 : x \in \mathbb{R} \right\}$. Assume the x-axis can be represented as the complete solution to $z = m_1x + m_2y + c$ for some m_1, m_2, c. Since $(x, y, z) = (0, 0, 0)$ is on the x axis, we know $c = 0$. Since $(x, y, z) =$

© Jason Siefken, 2015–2022

$(1,0,0)$ is on the x axis, we know that $m_1 = 0$. The equation then becomes $z = m_2 y$. However, for each choice of m_2, $x = 0, y = 1, z = m_2$ is a solution to the system which does not lie in the x-axis. Therefore, there's no there's no $m_1, m_2, c \in \mathbb{R}$ so that the x-axis is the solution set to the equation $z = m_1 x + m_2 y + c$.

(f) True. An example is given by

$$\left\{0x + 0y = 1\right. .$$

Solutions for Appendix 2

1 (a) Let

$$X = \left[\begin{array}{cccc|c} 4 & 6 & 3 & -10 & 6 \\ 5 & 2 & 1 & -7 & 2 \\ -6 & 2 & 1 & 4 & 2 \end{array}\right]$$

be the augmented matrix corresponding to the system.

By row reduction,

$$\operatorname{rref}(X) = \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 1/2 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

The third and fourth column of rref(X) are free variable columns, so we introduce the arbitrary equations $z = t$ and $w = s$ and solve the following system in terms of t and s:

$$\begin{cases} x \qquad\qquad\quad - w = 0 \\ \quad y + (1/2)z - w = 1 \\ \qquad\qquad z \quad\; = t \\ \qquad\qquad\quad w = s \end{cases}.$$

Written in vector form, the complete solution is

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} s \\ 1 - (1/2)t + s \\ t \\ s \end{bmatrix} = t \begin{bmatrix} 0 \\ -1/2 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

(b) Let

$$X = \left[\begin{array}{cccc|c} 2 & 2 & 1 & 0 & -1 \\ 0 & 1 & -4 & 2 & 3 \\ 1 & -1 & -3 & -4 & 5 \end{array}\right]$$

be the augmented matrix corresponding to the system.

By row reduction,

$$\operatorname{rref}(X) = \left[\begin{array}{cccc|c} 1 & 0 & 0 & -2 & 1 \\ 0 & 1 & 0 & 2 & -1 \\ 0 & 0 & 1 & 0 & -1 \end{array}\right].$$

The fourth column of rref(X) is a free variable column, so we introduce the arbitrary equation $w = t$ and solve the following system in terms of t:

$$\begin{cases} x \qquad\quad - 2w = 1 \\ \quad y \quad\; + 2w = -1 \\ \qquad z \qquad = -1 \\ \qquad\quad w = t \end{cases}.$$

Written in vector form, the complete solution is

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

(c) Let

$$X = \left[\begin{array}{ccc|c} 1 & 1 & -2 & -5 \\ -4 & 1 & 5 & 3 \end{array}\right]$$

be the augmented matrix corresponding to the system.

By row reduction,

$$\operatorname{rref}(X) = \left[\begin{array}{ccc|c} 1 & 0 & -7/5 & -8/5 \\ 0 & 1 & -3/5 & -17/5 \end{array}\right].$$

The third column of rref(X) is a free variable column, so we introduce the arbitrary equation $z = t$ and solve the following system in terms of t:

$$\begin{cases} x \qquad - 7/5z = -8/5 \\ \quad y - 3/5z = -17/5 \\ \qquad\quad z = t \end{cases}.$$

Written in vector form, the complete solution is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -8/5 + 7/5t \\ -17/5 + 3/5t \\ t \end{bmatrix} = t \begin{bmatrix} 7/5 \\ 3/5 \\ 1 \end{bmatrix} + \begin{bmatrix} -8/5 \\ -17/5 \\ 0 \end{bmatrix}.$$

(d) Let

$$X = \left[\begin{array}{ccc|c} 3 & -2 & 0 & -4 \\ 1 & 1 & 3 & 3 \\ -4 & 1 & -3 & 1 \end{array}\right]$$

be the augmented matrix corresponding to the system.

By row reduction,

$$\operatorname{rref}(X) = \left[\begin{array}{ccc|c} 1 & 0 & 6/5 & 2/5 \\ 0 & 1 & 9/5 & 13/5 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The third column of rref(X) is a free variable column, so we introduce the arbitrary equation $z = t$ and solve the following system in terms of t:

$$\begin{cases} x \qquad + 6/5z = 2/5 \\ \quad y + 9/5z = 13/5 \\ \qquad\quad z = t \end{cases}.$$

Written in vector form, the complete solution is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2/5 - 6/5t \\ 13/5 - 9/5t \\ t \end{bmatrix} = t \begin{bmatrix} -6/5 \\ -9/5 \\ 1 \end{bmatrix} + \begin{bmatrix} 2/5 \\ 13/5 \\ 0 \end{bmatrix}.$$

(e) Let

$$X = \left[\begin{array}{ccc|c} 1 & -1 & 2 & -1 \\ 2 & 1 & 4 & 1 \\ 3 & -4 & 3 & -2 \end{array}\right]$$

be the augmented matrix corresponding to the system.

By row reduction,

$$\operatorname{rref}(X) = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 4/3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2/3 \end{array}\right].$$

© Jason Siefken, 2015–2022

Written in vector form, the complete solution is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4/3 \\ 1 \\ -2/3 \end{bmatrix}.$$

(f) Let

$$X = \left[\begin{array}{ccc|c} 2 & 0 & 1 & 8 \\ 1 & 1 & 1 & 4 \\ 1 & 3 & 2 & 4 \\ 3 & 2 & 4 & 9 \end{array}\right]$$

be the augmented matrix corresponding to the system.

By row reduction,

$$\mathrm{rref}(X) = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Written in vector form, the complete solution is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}.$$

2 (a) i.

$$\left[\begin{array}{ccc|c} -10 & -4 & 4 & 28 \\ 3 & 1 & -1 & -8 \\ 1 & 1 & -1/2 & -3 \end{array}\right]$$

(a) ii.

$$\left[\begin{array}{ccc|c} -10 & -4 & 4 & 28 \\ 3 & 1 & -1 & -8 \\ 1 & 1 & -1/2 & -3 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 1 & -1/2 & -3 \\ 3 & 1 & -1 & -8 \\ -10 & -4 & 4 & 28 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 1 & -1/2 & -3 \\ 0 & -2 & 1/2 & 1 \\ 0 & 6 & -1 & -2 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 1 & -1/2 & -3 \\ 0 & 1 & -1/4 & -1/2 \\ 0 & 6 & -1 & -2 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 1 & -1/2 & -3 \\ 0 & 1 & -1/4 & -1/2 \\ 0 & 0 & 1/2 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 1 & -1/2 & -3 \\ 0 & 1 & -1/4 & -1/2 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

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

(a) iii. This system of linear equations is consistent. Its complete solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 2 \end{bmatrix}.$$

(b) i.

$$\left[\begin{array}{ccc|c} 3 & -2 & 4 & 54 \\ 5 & -3 & 6 & 88 \\ 1 & 0 & 0 & -3 \end{array}\right]$$

(b) ii.

$$\left[\begin{array}{ccc|c} 3 & -2 & 4 & 54 \\ 5 & -3 & 6 & 88 \\ 1 & 0 & 0 & -3 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 0 & 0 & -3 \\ 5 & -3 & 6 & 88 \\ 3 & -2 & 4 & 54 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 0 & 0 & -3 \\ 0 & -3 & 6 & 103 \\ 0 & -2 & 4 & 63 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 0 & 0 & -3 \\ 0 & 1 & -2 & -103/3 \\ 0 & -2 & 4 & 63 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 0 & 0 & -3 \\ 0 & 1 & -2 & -103/3 \\ 0 & 0 & 0 & -17/3 \end{array}\right]$$

(b) iii. This system of linear equations is inconsistent.

(c) i.

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

(c) ii. The augmented matrix of this system of linear equations is already in reduced row echelon form.

(c) iii. This system of linear equations is consistent. Its complete solution is

$$\begin{bmatrix} x \\ y \end{bmatrix} = t\begin{bmatrix} -2 \\ 1 \end{bmatrix} + \begin{bmatrix} 5 \\ 0 \end{bmatrix}.$$

(d) i.

$$\left[\begin{array}{c|c} 4 & 6 \\ 2 & 3 \end{array}\right]$$

(d) ii.

$$\left[\begin{array}{c|c} 4 & 6 \\ 2 & 3 \end{array}\right] \to \left[\begin{array}{c|c} 1 & 3/2 \\ 2 & 3 \end{array}\right] \to \left[\begin{array}{c|c} 1 & 3/2 \\ 0 & 0 \end{array}\right]$$

(d) iii. This system of linear equations is consistent. Its complete solution is $x = 3/2$.

(e) i.

$$\left[\begin{array}{cccc|c} 1 & 2 & 4 & -3 & 0 \\ 3 & 5 & 6 & -4 & 1 \\ 4 & 5 & -2 & 3 & 3 \end{array}\right]$$

(e) ii.

$$\left[\begin{array}{cccc|c} 1 & 2 & 4 & -3 & 0 \\ 3 & 5 & 6 & -4 & 1 \\ 4 & 5 & -2 & 3 & 3 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & 2 & 4 & -3 & 0 \\ 0 & -1 & -6 & 5 & 1 \\ 0 & -3 & -18 & 15 & 3 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & 2 & 4 & -3 & 0 \\ 0 & 1 & 6 & -5 & -1 \\ 0 & -3 & -18 & 15 & 3 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & 2 & 4 & -3 & 0 \\ 0 & 1 & 6 & -5 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & 0 & -8 & 7 & 2 \\ 0 & 1 & 6 & -5 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

© Jason Siefken, 2015–2022

(e) iii. This system of linear equations is consistent. Its complete solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = t \begin{bmatrix} 8 \\ -6 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -7 \\ 5 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \\ 0 \\ 0 \end{bmatrix}.$$

(f) i.

$$\left[\begin{array}{cccc|c} 1 & -1 & 5 & 1 & 1 \\ 1 & 1 & -2 & 3 & 3 \\ 3 & -1 & 8 & 1 & 5 \\ 1 & 3 & -9 & 7 & 5 \end{array}\right]$$

(f) ii.

$$\left[\begin{array}{cccc|c} 1 & -1 & 5 & 1 & 1 \\ 1 & 1 & -2 & 3 & 3 \\ 3 & -1 & 8 & 1 & 5 \\ 1 & 3 & -9 & 7 & 5 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & -1 & 5 & 1 & 1 \\ 0 & & -2 & 3 & 3 \\ 3 & -1 & 8 & 1 & 5 \\ 1 & 3 & -9 & 7 & 5 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & -1 & 5 & 1 & 1 \\ 0 & 2 & -7 & 2 & 2 \\ 0 & 2 & -7 & -2 & 2 \\ 0 & 4 & -14 & 6 & 4 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & -1 & 5 & 1 & 1 \\ 0 & 1 & -7/2 & 1 & 1 \\ 0 & 2 & -7 & -2 & 2 \\ 0 & 4 & -14 & 6 & 4 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & -1 & 5 & 1 & 1 \\ 0 & 1 & -7/2 & 1 & 1 \\ 0 & 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 2 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & -1 & 5 & 1 & 1 \\ 0 & 1 & -7/2 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & -1 & 5 & 1 & 1 \\ 0 & 1 & -7/2 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{cccc|c} 1 & 0 & 3/2 & 0 & 2 \\ 0 & 1 & -7/2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

(f) iii. This system of linear equations is consistent. Its complete solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = t \begin{bmatrix} -3/2 \\ 7/2 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

(g) i.

$$\left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \end{array}\right]$$

(g) ii. The augmented matrix of this system of linear equations is already in reduced row echelon form.

(g) iii. This system of linear equations is consistent. Its complete solution is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + r \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

3 (a) The vectors $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ are linearly independent if

$$x\vec{v}_1 + y\vec{v}_2 + z\vec{v}_3 = \vec{0}$$

only has the trivial solution.

This vector equation is equivalent to the system of linear equations

$$\begin{cases} x + \ \ y - 2z = 0 \\ x + 4y - 2z = 0 \\ -2x \qquad + 4z = 0 \\ 4x + 2y - 8z = 0 \end{cases}.$$

The complete solution to this system is

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

In particular, $(x, y, z) = (2, 0, 1)$ is a non-trivial solution to this system, so the vectors $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ are linearly dependent.

(b) By definition, the vectors $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$ span $\mathbb{R}^3$ if every vector can be written as a linear combination of $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$. In other words, the equation

$$x\vec{v}_1 + y\vec{v}_2 + z\vec{v}_3 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

is consistent for all choices of a, b, and c. This vector equation is equivalent to the system of linear equations

$$\begin{cases} x - 2y + 2z = a \\ 2x + \ \ y + 7z = b \\ 3x \qquad + \ \ z = c \end{cases}.$$

Row reducing, we notice that every column is a pivot column and so the system is always consistent. Therefore, $\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \mathbb{R}^3$.

(c) The lines ℓ_1 and ℓ_2 intersect when their x and y-coordinates. We first set the parameter variable of ℓ_1 to t and the parameter variable of ℓ_2 to s. Then, equating the coordinates gives the system of linear equations

$$\begin{cases} t - 2s = 2 \\ 3t - \ \ s = 3 \end{cases}.$$

The solution to this system is

$$\begin{bmatrix} t \\ s \end{bmatrix} = \begin{bmatrix} 4/5 \\ -3/5 \end{bmatrix}.$$

Since $\vec{x} = \begin{bmatrix} 9/5 \\ 17/5 \end{bmatrix}$ when $t = 4/5$ (or $s = -3/5$), the intersection of ℓ_1 and ℓ_2 is the point $\begin{bmatrix} 9/5 \\ 17/5 \end{bmatrix}$.

(d) The planes $\mathcal{P}_1$ and $\mathcal{P}_2$ intersect when their coordinates are equal. Relabeling the parameter variables for $\mathcal{P}_2$ as q and r and equating both vector forms, we get the following system of linear equations:

$$\begin{cases} t - \ \ s - q + \ \ r = 0 \\ -t - \ \ s + q - 3r = 1 \\ \qquad 2s - q + 2r = -1 \end{cases}.$$

© Jason Siefken, 2015–2022

The complete solution to this system is

$$\begin{bmatrix} t\\ s\\ q\\ r\end{bmatrix} = u\begin{bmatrix}-2\\-1\\0\\1\end{bmatrix} + \begin{bmatrix}-1/2\\-1/2\\0\\0\end{bmatrix}.$$

Thus there are an infinite number of points in $\mathcal{P}_1 \cap \mathcal{P}_2$.

To find these points, we substitute $q=0$ and $r=u$ into the vector form of $\mathcal{P}_2$. This shows us that $\mathcal{P}_1 \cap \mathcal{P}_2$ can be expressed in vector form by

$$\begin{bmatrix} x\\ y\\ z\end{bmatrix} = u\begin{bmatrix}-1\\3\\-2\end{bmatrix} + \begin{bmatrix}0\\1\\-1\end{bmatrix}.$$

4 (a) The reasoning is incorrect. The solution $(x,y,z)=(0,0,0)$ is the trivial solution to the vector equation, and it is always a solution to the homogeneous equation

$$\alpha_1\vec{v}_1+\cdots+\alpha_k\vec{v}_k=\vec{0}$$

no matter what $\vec{v}_1,\ldots,\vec{v}_k$ are.

To determine if a set of vectors is linearly independent, we need to find out whether the trivial solution is the *only* solution to the vector equation. That is, there does not exist any non-trivial solution to the vector equation.

(b) The reasoning is incorrect. A trivial solution is the solution where *all* the variables equal zero, so the solution $(x,y,z)=(-2,0,-1)$ is not a trivial solution.

(c) The reasoning is incorrect. When equating coordinates of two different vector forms, the parameter variables needs to be set to different letters.

A valid system of linear equations is

$$\begin{cases} t-2s=2\\ 3t-\ \ s=3\end{cases}.$$

Here we have set the parameter t in the vector form of ℓ_2 to s.

(d) The reasoning is incorrect. The solution $\begin{bmatrix}4/5\\-3/5\end{bmatrix}$ to the system of linear equations gives the value of t and s at the intersection. To find the intersection of ℓ_1 and ℓ_2, the value $t=4/5$ or $s=-3/5$ needs to be plugged into the vector form of ℓ_1 or ℓ_2.

(e) The reasoning is correct. Since $\vec{x}=\begin{bmatrix}1/2\\-1/2\\0\end{bmatrix}$ is a point on both $\mathcal{P}_1$ and $\mathcal{P}_2$, it is in the intersection of $\mathcal{P}_1$ and $\mathcal{P}_2$, so the planes $\mathcal{P}_1$ and $\mathcal{P}_2$ intersect. However, we can not determine if the intersection is a line or a plane based on only one point, so we need to set up and solve an appropriate system of linear equations.

(f) The reasoning is incorrect. Finding one point that is in $\mathcal{P}_2$ but not in $\mathcal{P}_1$ shows that $\mathcal{P}_1$ does not intersect $\mathcal{P}_2$ *at that point*, but does not rule out the possibility that $\mathcal{P}_1$ intersects $\mathcal{P}_2$ at a *different* point.

Solutions for Appendix 3

1 Note: there are many matrices possible.

(a) $\begin{bmatrix}1&0\\0&-2\end{bmatrix}$

(b) $\begin{bmatrix}1&1\\1&-4\end{bmatrix}$ or $\begin{bmatrix}1&0\\0&-2\end{bmatrix}$

(c) Impossible. A symmetric matrix must be square.

(d) Impossible. Let $A=[a_{ij}]$ be a skew symmetric matrix. By definition, $a_{ij}=-a_{ji}$. In particular, the diagonal entries satisfy $a_{kk}=-a_{kk}$ and so must be zero. For every other entry a_{ij} with $i\neq j$, there exists a corresponding entry $-a_{ji}$. Therefore the sum of all entries must be zero.

(e) $\begin{bmatrix}\sqrt{2}&\sqrt{3}&2&\sqrt{5}\end{bmatrix}$

2 (a) $\begin{bmatrix}1\\5\\5\end{bmatrix}$

(b) $\begin{bmatrix}4&1&1\\2&2&4\\2&12&13\end{bmatrix}$

(c) $\begin{bmatrix}-1&-1&-1\\1&-2&2\\1&6&2\end{bmatrix}$

(d) Not defined

(e) $\begin{bmatrix}2&1&1\\2&1&7\\2&13&16\end{bmatrix}$

(f) $\begin{bmatrix}4&6&7\\2&0&4\\3&18&15\end{bmatrix}$

(g) $\begin{bmatrix}4\end{bmatrix}$

(h) $\begin{bmatrix}0&-2&1\\0&2&-1\\0&-4&2\end{bmatrix}$

4 (a) True

(b) False. Consider $\begin{bmatrix}0&0\\0&1\end{bmatrix}$.

(c) False. Zero matrices can be of any size.

(d) False. All zero matrices are both upper and lower triangular.

(e) False. Diagonal matrices must be square; upper/lower triangular matrices can be of any size.

(f) True

(g) True

5 (a) Yes. 1×1

(b) Yes. $n\times n$

(c) They are related but not the same. $\vec{r}\cdot\vec{c}$ is a scalar and RC is a 1×1 matrix.

(d) i. The inner product of $\vec{x}$ and $\vec{y}$ is $[\vec{x}\cdot\vec{y}]$. That is, it is the 1×1 matrix with entry $\vec{x}\cdot\vec{y}$.

ii. The reduced row echelon form of Q looks like a (possibly) non-zero row followed by rows of zeros.

iii. Let $y_1,\ldots,y_n\in\mathbb{R}$ be the entries in $\vec{y}$. Then the columns of Q are $y_1\vec{x},\ldots,y_n\vec{x}$. These columns are all scalar multiples of the same vector, $\vec{x}$, and so are linearly dependent.

© Jason Siefken, 2015–2022

6 (a) Multiplying

$$\begin{bmatrix}1&a&b\\0&1&c\\0&0&1\end{bmatrix}\begin{bmatrix}1&x&y\\0&1&z\\0&0&1\end{bmatrix}=\begin{bmatrix}1&a+x&c+ay+z\\0&1&b+y\\0&0&1\end{bmatrix}$$

and so two Heisenberg matrices aways multiply together to form another Heisenberg matrix.

(b) It may be that $AB \neq BA$. For example

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

(c) By multiplying out, we see $XY = I$ (and $YX = I$).

7 (a) Multiplying out, we see

$$\begin{bmatrix}a&-b\\b&a\end{bmatrix}\begin{bmatrix}x&-y\\y&x\end{bmatrix}=\begin{bmatrix}ax-by&-(ay+bx)\\ay+bx&ax-by\end{bmatrix}$$

has the required form.

(b) Multiplying, we see

$$X^2=\begin{bmatrix}a&-b\\b&a\end{bmatrix}^2=\begin{bmatrix}a^2-b^2&-2ab\\2ab&a^2-b^2\end{bmatrix}=\begin{bmatrix}1&0\\0&1\end{bmatrix}$$

implies that $a=0$ or $b=0$. If $a=0$, then $-b^2=1$, which is impossible. Therefore $b=0$. This means $a^2=1$ which has solutions $a=\pm1$. Therefore there are exactly two solutions to $X^2=I$.

(c) Multiplying, we see

$$X^2=\begin{bmatrix}a&-b\\b&a\end{bmatrix}^2=\begin{bmatrix}a^2-b^2&-2ab\\2ab&a^2-b^2\end{bmatrix}=\begin{bmatrix}-1&0\\0&-1\end{bmatrix}$$

implies that $a=0$ or $b=0$. If $a=b$, then $a^2=-1$, which is impossible. Therefore $a=0$. This means $-b^2=-1$ which has solutions $b=\pm1$. Therefore there are exactly two solutions to $X^2=-I$.

(d) There are infinitely many solutions to $Y^2=I$. For example $\begin{bmatrix}0&t\\1/t&0\end{bmatrix}^2=I$ for any non-zero t.

(e) Yes to the first, no to the second. Matrices are more general than numbers!

Solutions for Appendix 4

1 (a) 4

(b) 0

(c) 1

(d) 0

2 (a) −12

(b) −16

(c) 21

(d) 12

(e) 0

3 (a) Since $\det\left(\begin{bmatrix}1&1\\-3&2\end{bmatrix}\right)=5>0$, the ordered set $\left\{\begin{bmatrix}1\\-3\end{bmatrix},\begin{bmatrix}1\\2\end{bmatrix}\right\}$ is a right-handed basis.

(b) Since $\det\left(\begin{bmatrix}1&-2\\-3&6\end{bmatrix}\right)=0$, the set $\left\{\begin{bmatrix}1\\-3\end{bmatrix},\begin{bmatrix}-2\\6\end{bmatrix}\right\}$ is not linearly independent and so is not a basis.

(c) Since $\det\left(\begin{bmatrix}1&1&1\\0&2&-1\\3&5&1\end{bmatrix}\right)=-2<0$, the ordered set $\left\{\begin{bmatrix}1\\0\\3\end{bmatrix},\begin{bmatrix}1\\2\\5\end{bmatrix},\begin{bmatrix}1\\-1\\1\end{bmatrix}\right\}$ is a left-handed basis.

(d) Since $\det\left(\begin{bmatrix}1&1&1\\4&2&1\\9&3&1\end{bmatrix}\right)=-2<0$, the ordered set $\left\{\begin{bmatrix}1\\4\\9\end{bmatrix},\begin{bmatrix}1\\2\\3\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$ is a left-handed basis.

(e) Since $\det\left(\begin{bmatrix}4&4&2\\2&2&1\\4&0&6\end{bmatrix}\right)=0$, the set $\left\{\begin{bmatrix}4\\2\\4\end{bmatrix},\begin{bmatrix}4\\2\\0\end{bmatrix},\begin{bmatrix}2\\1\\6\end{bmatrix}\right\}$ is not linearly independent and so is not a basis.

4 Before answering, note that $\det\left(\begin{bmatrix}a^2&ab\\ab&b\end{bmatrix}\right)=a^2b-a^2b^2=a^2(b-b^2)$.

(a) If $\left\{\begin{bmatrix}a^2\\ab\end{bmatrix},\begin{bmatrix}ab\\b\end{bmatrix}\right\}$ is a right-handed basis, then $\det\left(\begin{bmatrix}a^2&ab\\ab&b\end{bmatrix}\right)=a^2(b-b^2)>0$. This implies that a^2 and $b-b^2$ are both nonzero and have the same sign. Since $a^2\geq0$, we must have $a^2>0$ and $b-b^2>0$. The roots of $b-b^2$ are $b=0$ and $b=1$, so $b-b^2>0$ implies $0<b<1$. Therefore, $\left\{\begin{bmatrix}a^2\\ab\end{bmatrix},\begin{bmatrix}ab\\b\end{bmatrix}\right\}$ is a right-handed basis when $a\neq0$ and $0<b<1$.

(b) If $\left\{\begin{bmatrix}a^2\\ab\end{bmatrix},\begin{bmatrix}ab\\b\end{bmatrix}\right\}$ is a left-handed basis, then $\det\left(\begin{bmatrix}a^2&ab\\ab&b\end{bmatrix}\right)=a^2(b-b^2)<0$. This implies that a^2 and $b-b^2$ are both nonzero and have different signs. Since $a^2\geq0$, this implies that $a^2>0$ and $b-b^2<0$. Roots of $b-b^2$ are $b=0$ and $b=1$, so $b-b^2<0$ implies $b<0$ or $b>1$. Therefore, $\left\{\begin{bmatrix}a^2\\ab\end{bmatrix},\begin{bmatrix}ab\\b\end{bmatrix}\right\}$ is a left-handed basis when $a\neq0$ and either $b<0$ or $b>1$.

(c) If $\left\{\begin{bmatrix}a^2\\ab\end{bmatrix},\begin{bmatrix}ab\\b\end{bmatrix}\right\}$ is not a basis, then $\det\left(\begin{bmatrix}a^2&ab\\ab&b\end{bmatrix}\right)=a^2(b-b^2)=0$. This implies that $a^2=0$ or $b-b^2=0$. Therefore, $\left\{\begin{bmatrix}a^2\\ab\end{bmatrix},\begin{bmatrix}ab\\b\end{bmatrix}\right\}$ is not a basis if one of the following conditions holds: $a=0$, $b=1$, or $b=0$.

5 Since

$$\frac{M^{\mathrm{adj}}}{\det M}M=\frac{1}{ad-bc}\begin{bmatrix}da-bc&db-bd\\ac-ca&ad-cb\end{bmatrix}$$

$$=\begin{bmatrix}1&0\\0&1\end{bmatrix}=I_{2\times2}$$

© Jason Siefken, 2015–2022

and

$$M\frac{M^{\text{adj}}}{\det M} = \frac{1}{ad-bc}\begin{bmatrix} ad-bc & ba-ab \\ cd-dc & da-cb \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_{2\times 2},$$

we conclude that

$$M^{-1} = \frac{M^{\text{adj}}}{\det(M)}.$$

6 (a) False. A counterexample is $M = \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$.

(b) False. A counterexample is $M = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$.

M does not change volume, but it does reverse orientation.

(c) True. Note that $[\vec{a}|\vec{b}]$ is just $[\vec{b}|\vec{a}]$ with its columns swapped. The oriented volume of the parallelogram generated by $\vec{a}$ and $\vec{b}$ is equal to the negative of the oriented volume of the parallelogram generated by $(\vec{b},\vec{a})$. Using Volume Theorem I, we have $\det([\vec{a}|\vec{b}]) = -\det([\vec{b}|\vec{a}])$.

(d) False. A counterexample is $M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. We multiply the $(1,2)$-entry by 4 to get another matrix M'. Note that M' is still $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, and $\det(M) = 1 = \det(M') \neq 4\det(M)$.

(e) True. Note that $\det(A^TA) = \det(A^T)\det(A) = \det(A)\det(A) = \det(A)^2 \geq 0$.

© Jason Siefken, 2015–2022

Indices

Index of Symbols

C_n, 171, 179
$L(X)$ (Image of a set), 108, 119
M^T, 130
$[\mathcal{T}]_{\mathcal{B}}$, 163
$[\vec{v}]_{\mathcal{B}}$, 97, 104
$|\mathcal{T}|$, 173, 180
$[\mathcal{B}\leftarrow\mathcal{A}]$, 161
$\mathrm{col}(M)$, 129, 140
id, 147
$\mathbb{N}$, 2
$\mathrm{null}(M)$, 129, 140
$\mathrm{nullity}(M)$, 133
$\mathrm{proj}_X\,\vec{v}$, 69, 75
$\mathbb{Q}$, 2
$\mathbb{R}$, 2
$\mathbb{R}^n$, 2
$\mathrm{rank}(M)$, 133
$\mathrm{rank}(T)$, 127
$\mathrm{row}(M)$, 129, 140
$\mathrm{span}\,V$, 34, 43
$\mathrm{vcomp}_{\vec{v}}\,\vec{u}$, 72, 77
$\mathbb{Z}$, 2
$\cap$, 2, 12
$\mathrm{char}(A)$, 188, 195
$\cup$, 2, 12
$\emptyset$, 1
$\in$, 1
$\|\cdot\|$, 3
$\notin$, 1
$\det(\mathcal{T})$, 173, 180
$\sim$, 164, 169
$\subseteq$, 1
$\vec{0}$, 4
$\vec{a}\cdot\vec{b}$, 53
$g\circ f$, 121

Index of Terms

Basis, 81, 86
- change-of-basis, 198
- left-handed, 100, 105
- negatively oriented, 100, 105
- orientation of, 99, 238
- positively oriented, 100, 105
- right-handed, 100, 105, 236
- standard basis, 5, 83

Change of basis matrix, 161
Characteristic polynomial, 188, 195
Coordinates, 5

Determinant, 173, 180
- 2×2 formula, 235
- 3×3 formula, 237

Diagonalizable, 199
Dimension, 127
Displacement, 2
Dot product, 53
- algebraic definition, 53
- geometric definition, 53

Eigenspace, 200, 206
Eigenvalue, 187, 192
- algebraic multiplicity, 200, 206
- geometric multiplicity, 200, 206

Eigenvector, 187, 192, 198

Function
- composition, 121
- identity function, 147
- inverse of, 147
- invertible function, 147

Heisenberg matrix, 234
Hyperplane, 56

Injective function, 147
Inner product, 233
Isomorphism, 3

Line
- direction vector for, 18, 29
- normal form, 55
- normal vector, 55, 66
- parallel, 20
- skew, 20
- vector form of, 18, 29

Linear combination, 5, 14
- coefficient of, 5, 14
- convex, 23, 27
- non-negative, 23, 27
- trivial, 38, 48

Linear transformation, 108, 118, 190
- composition of, 174
- induced, 132, 139
- notation, 109
- nullity, 128
- orientation preserving, 173
- orientation reversing, 173
- proving linearity, 111
- range (image), 127, 137
- rank, 127, 143
- representation in a basis, 112, 163

Linear transformations
- null space (kernel), 128, 137

Linearity, 109
Linearly dependent, 37, 47
Linearly independent, 37, 47

Matrix
- coefficient matrix, 89
- column space, 129, 140
- definition, 227
- diagonal matrix, 198
- diagonalizable, 198, 205
- elementary matrix, 150, 156, 175, 235
- fundmamental subspaces of, 129, 140
- identity matrix, 149, 228
- inverse of, 149, 157
- invertible, 151, 162, 177
- linear transformation induced by, 132, 139
- matrix multiplication, 122
- matrix transformation, 112
- null space, 129, 140
- nullity, 133, 143
- of a linear transformation, 163
- rank, 133
- reduced row echelon form, 130
- row space, 129, 140
- similar matrices, 197
- transpose of, 130

Matrix equations
- the column picture, 90
- the row picture, 90

© Jason Siefken, 2015–2022

Normal vector, 56
Null space, 188

One-to-one function, 147
Onto function, 148
Orthogonal, 54, 64
Outer product, 233

Parameter variable, 18
Plane
direction vector for, 21, 31
normal vector, 55, 66
scalar form of, 57
vector form of, 21, 31
Point
in Euclidean Space, 3
of a set, 1
Projection, 69, 75

Rank-nullity theorem
for linear transformations, 135, 148
for matrices, 134
Russel's Paradox, 1

Scalar, 2
Scalar product, 53
Set
element of, 1
empty, 1
equality, 1
image of, 108, 119
intersection, 2, 12
union, 2, 12
Set sum, 35, 45
Set-builder notation, 2
Span, 34, 43
translated, 36
Standard basis, 5
Subspace, 79, 85, 127
dimension, 82, 86
trivial subspace, 81
Surjective function, 148
System of linear equations
coefficient matrix, 89
equivalent, 212
homogeneous, 39
overdetermined system, 21

True vector, 99

Unit n-cube (C_n), 171, 179

Vector, 2
direction of, 2, 54
magnitude of, 2, 3
norm of, 3
positive direction of, 54
representation in a basis, 97, 104
rooted, 3
Vector component, 72, 77
Vector operations
addition, 4
laws of, 4
scalar multiplication, 4
Volume, 35, 171
theorem I, 175, 235
theorem II, 177

Zero vector ($\vec{0}$), 4
Zero vector ($\vec{0}$), 110

Index of Definitions

Determinant, 173, 180

Eigenspace, 200, 206
Eigenvalue, 187, 192
algebraic multiplicity, 200, 206
geometric multiplicity, 200, 206
Eigenvector, 187, 192
Equivalent systems, 209

Injective function, 147

Linear transformation
orientation preserving, 173
orientation reversing, 173
representation in a basis, 163

Matrix
diagonal, 228
diagonal of, 228
diagonalizible, 198, 205
elementary matrix, 150, 156
identity matrix, 149, 228
inverse of, 149, 157
lower triangle of, 228
lower triangular, 228
of a linear transformation, 163
similar matrices, 164, 169
skew-symmetric, 228
square, 228
symmetric, 228
triangular, 228
upper triangle of, 228
upper triangular, 228
zero matrix, 228

One-to-one function, 147
Onto function, 148

Surjective function, 148

Unit n-cube (C_n), 171, 179

© Jason Siefken, 2015–2022