Quantitative Engineering Analysis I

Sixth Edition

Fall 2020

Contents

Todo list

Part I

Faces: Linear Algebra Through Facial Recognition

Chapter 1

UPDATE UPDATE UPDATE: Week 1a: Facial Recognition

Schedule

1.1 Welcome and Orientation [10 mins]

Welcome to QEA Module One! In this module, we will explore linear algebra by applying fundamental ideas to data in general and image recognition in particular. By the end of this module you will have implemented a facial recognition method, and applied the ideas more broadly in a short project. Let's first imagine how a computer "sees" an image using numbers ...

1.2 Coding a face [30 mins]

First, we would like you to go to this link. Once you are there, please open the document "page1-X" that corresponds to your breakout room number. In that document you will find a smiley face that is unique to your breakout room. (**Note:** If your breakout room number is greater than 8, please subtract 8 from your room number and use that face.) Your first activity today will be to imagine converting this face into a form that a computer can understand. A grid is superimposed on the face for your reference.

Exercise 1.1

Design a method that enables a computer to (approximately) reproduce the face from a list of numbers and an algorithm that you define. The numbers can be grouped within the list, but your list should contain numbers only. An example of a list containing two groups of numbers is $[[0, 100], [14, 20]]$. (The numbers do not need to be grouped, for example $[4, 3, 77]$ would work.) You will create an algorithm (in other words, a very specific set of instructions) that tells the computer what to do with your list of numbers. When another group applies your algorithm to the list of numbers, they should be able to recreate the face. When you've defined your group's method,

- Go to the Google Slides presentation here and find the slide for your breakout room. The slides will look like Figure 1.2.

- Generate the list of numbers that represents your face using your method.
- Make a set of instructions (your algorithm) and write them in the text box area of your slide.

Solution 1.1

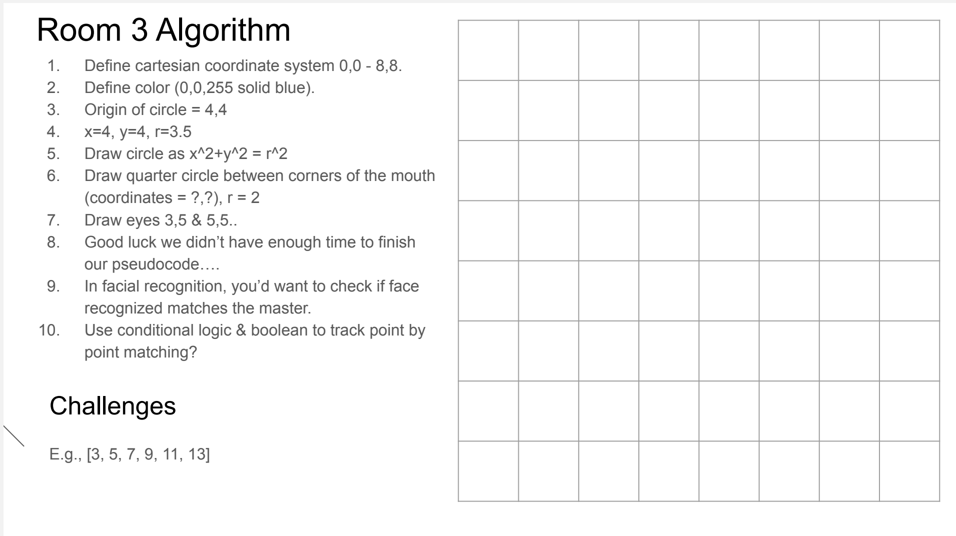


Figure 1.1: Example of an algorithm

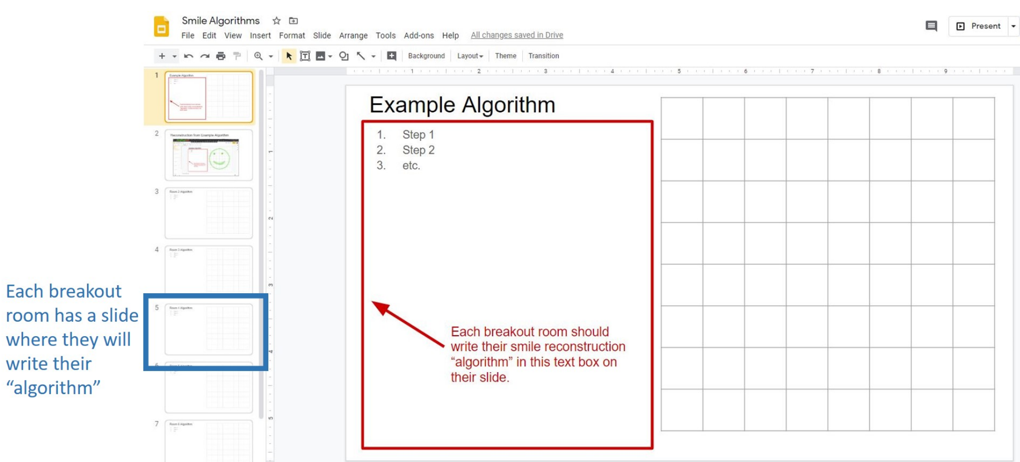


Figure 1.2: Algorithms developed by each breakout group should be written on the appropriate Google Slide within the text box.

1.3 Facial Reconstruction [20 mins]

Now that you have represented your face in algorithmic form, you will attempt to reconstruct a new face using another group's algorithm! While you are trying to reconstruct the face, think about the approach the other group took, and how it differs from your own.

Exercise 1.2

- You will try to recreate the face using the algorithm from the next (higher-numbered) breakout room (**note:** the highest number should wrap around to reconstruct room 1).
- In your breakout room, have one person share the google slide deck using the "Share Screen".
- Reconstruct the face collaboratively using the Zoom annotate tools, as shown in Figure 1.4. Using the annotate tools, draw directly on the 8x8 grid next to the algorithm text.
- Record any challenges you encounter directly on their algorithm text using the zoom annotate tools (text box, highlight, draw, etc.).
- When you have completed reconstructing the face and marking up the other group's algorithm using the annotate tools, save your share screen using the save button as shown in Figure 1.4. (Any marks you make within the Zoom annotation will not automatically save on the Google slide, so you have to take a screen shot to share your work.)
- Paste the picture you just saved of the reconstructed face in the next slide of the Google slides, titled "Reconstruction from Group X Algorithm." This way you can share your work!

Solution 1.2

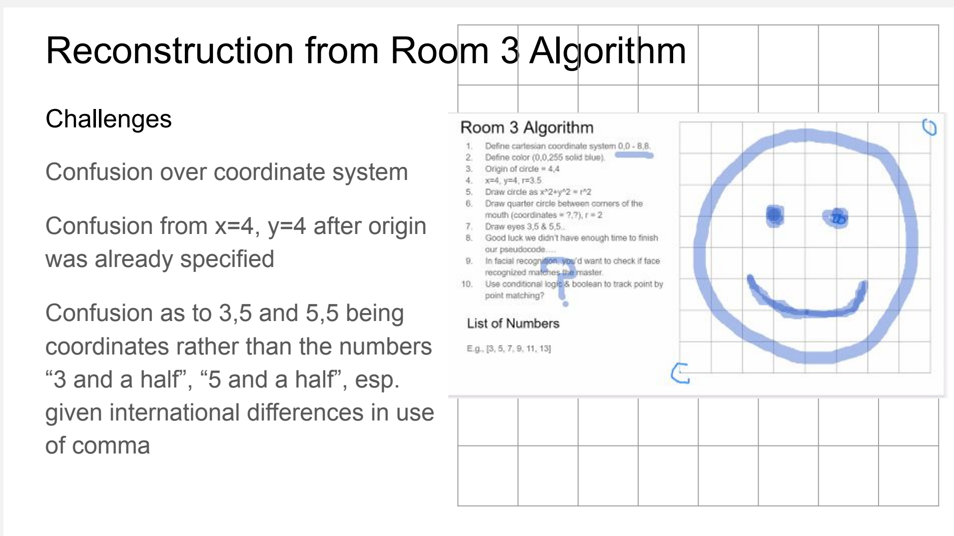


Figure 1.3: Example of a reconstruction

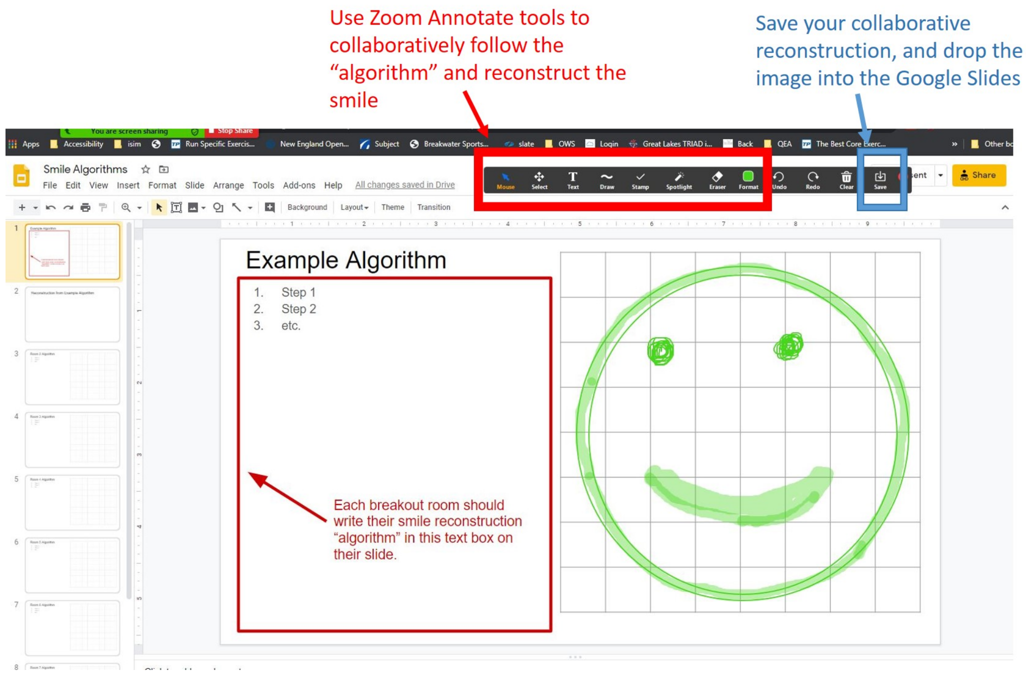


Figure 1.4: Use the annotate tools in Zoom to collaboratively reconstruct a face on the 8x8 grid using the algorithm instructions.

1.4 What did we learn? [15 mins]

Exercise 1.3

In your breakout room, please discuss the following prompts:

- Was the algorithm you crafted successful? What feedback did the other group leave for you?
- How did the other group's algorithm differ from yours? How successful were you in reconstructing their face?
- What components of the algorithm were essential (i.e. an origin, a cell numbering scheme, etc.)?
- Consider a photo of a human face. In what ways might your method contain inherent limits or biases?

1.5 Course Logistics and Survey [15 mins]

Chapter 2

Week 1b: Facespace

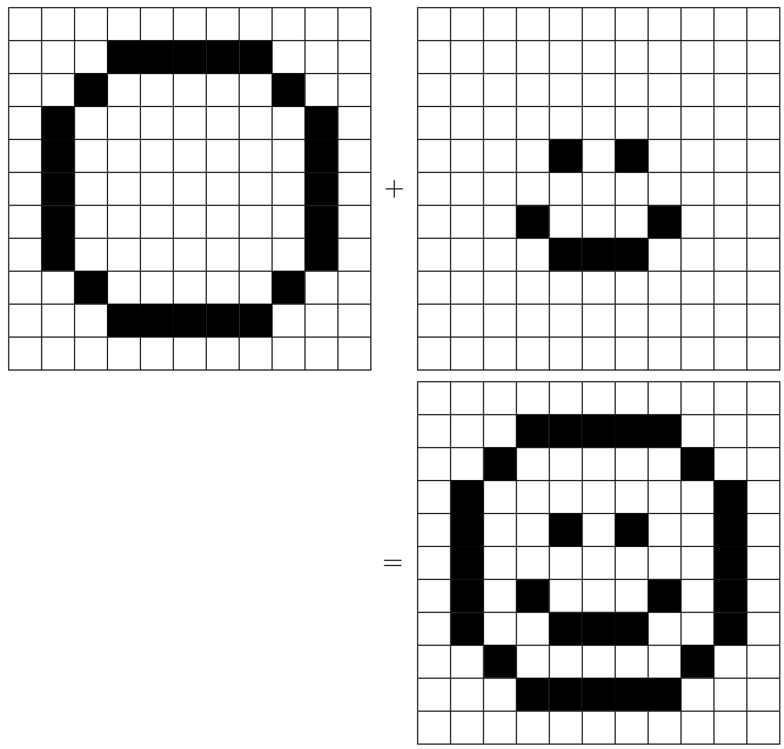
Schedule

Last class we thought about various ways to represent an image (e.g., a picture of a face). Today we're going to narrow in on a particular method of representing images: as a weighted sum of a set of building block images. You'll be working through exercises that will show how this type of representation works and why it is so powerful.

2.1 Pixel Arithmetic [10 mins]

Adding is one of the most basic operations in mathematics. While everyone here is familiar with the concept of adding numbers, we can generalize this idea to add together other sorts of entities. We can even think about what it means to add two images together.

As a simple example, let's add the following two images together (we'll explain more precisely how we are defining addition of images once you've seen the result).



Conceptually, this operation might seem straightforward. Adding two images results in an image that has a black pixel whenever either of the two images has a black pixel at a corresponding position.

More formally, we can think about black pixels as having a value of 255 and white pixels as having a value of 0 (gray pixels would have a value between these two values depending on how dark they are). (A scale from 0 to 255 seems like a weird choice, but there is a very good reason why this is the standard - digital storage, such as on your computer, uses binary (bit) - how many integers can you represent with an 8-bit number?) To add two images together, all we do is add the corresponding elements at a particular point in the grid! In this way addition on images works much the same as addition of a single number–the only difference is we perform the addition of single numbers multiple times for each position in the grid.

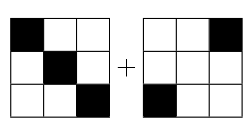
As an example, consider this addition problem.

Show an example of an image addition problem, conversion to numbers, and then back to an image

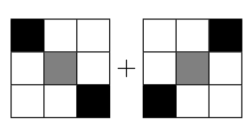
Exercise 2.1

With your group, work through the following pixel arithmetic problems on the board.

1.

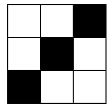
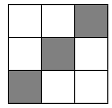


2.



Without too much of a leap, we can also multiply images by a number by multiplying each element in the image by that value. We can think of this multiplication operation as "scaling" the image.

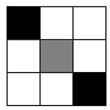
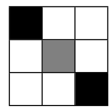
For example,

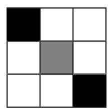
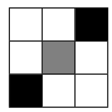
$0.5 \times$  $=$ 

Exercise 2.2

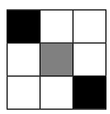
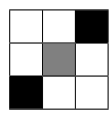
With your group, work through the following pixel arithmetic problems on the Zoom white board.

put a template for the answer to make it easier to do on the Zoom whiteboard

1. $0.5 \times$  $+ 0.5 \times$ 

2. $0.5 \times$  $+ 0.5 \times$ 

3. (Don't think about this one too hard. Just draw approximately what this would be)

   $0.9999 \times$  $+ 0.0001 \times$ 

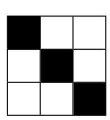
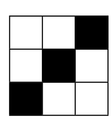
2.2 A Universal Set of Building Block Images [10 mins]

Now that we know how to add and scale images, let's think about how we might construct a set of building block images such that we can construct any image as a sum of scaled versions of these building blocks.

Exercise 2.3

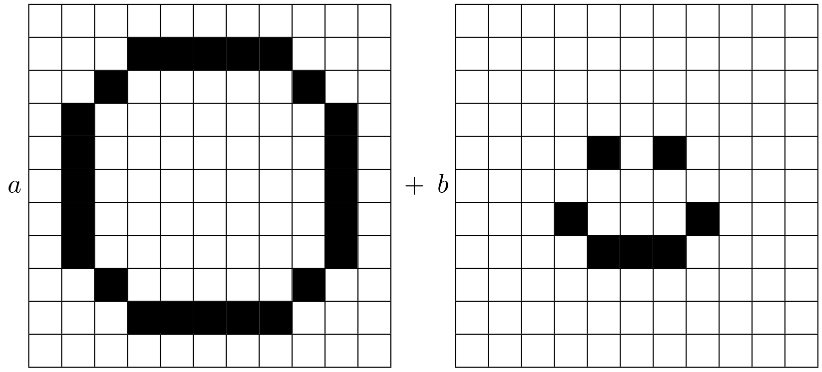
With your group, work through the following problems.

1. What is the range of images that could be constructed by summing over scaled versions of the following building block images? (c is a number between 0 and 1). Another way to think about this is, as you sweep the value of c from 0 to 1, how does the resultant sum of the two images change?

   $c \times$  $+ (1 - c) \times$ 

2. What is the range of images that could be constructed by summing over scaled versions of

the following building block images? (a and b are both numbers between 0 and 1). Instead of having one knob to turn (as in the previous exercise), you now have two.



In this case we can think of the values a and b as encoding of a particular smiley face. You will deduce the effect that both a and b have on the specific nature of the smiley face.

3. Building on the previous example, come up with your own way of representing a simple face like the one above as the sum of two or more scaled building block images. Be creative! It's up to you what sort of faces that your method is capable of representing.

You probably noticed from the previous three exercises that not all possible images can be constructed by adding scaled versions from a small set of building block images. Suppose you wanted to be able to represent *any* possible 3 pixel by 3 pixel image. While there are many possible ways to do this, for simplicity each of your building block images should only have a single black pixel (the rest should be white). At the board, define a set of building block images that lets you represent any possible 3 pixel by 3 pixel face in this manner. How many building block images did you need to represent all possible 3 pixel by 3 pixel faces? Are there any images that can't be represented as a sum of scaled versions from your building block images? How many building block images would you need if you wanted to encode all possible 5 pixel by 5 pixel faces? What about n pixels by n pixels?

change this from an exercise to explanatory text unless a good exercise can be found. Last time students found it so simple as to be confusing.

2.3 A Better Set of Building Blocks? [30 mins]

At the end of the previous section we showed how one can represent any possible image as a sum of scaled single-pixel images. This is a very powerful idea, but we can take it even farther. Before we continue, let's think about some of the ways in which this way of representing face images is not so great.

Exercise 2.4

Suppose you wish to represent 19 pixel by 19 pixel images of faces using the scheme you devised in the previous set of exercises (as a sum of scaled, single-pixel images). Here is an example of what such a face might look like.

1. If you think of the representation of each image as the scaling factor that you apply to each of your single-pixel images, how many numbers do you need to specify this one face image (you answered almost this exact question in the previous part, so don't overthink this).

   I think students found these confusing too. Need to look back at student comments;

2. How many numbers would you need to represent a 19 pixel by 19 pixel image of a flower? How many numbers would you need to represent a completely random 19 pixel by 19 images (one with no special structure)?

3. Suppose someone gives you one of the numbers needed to encode a particular face? Without looking at the face image itself, how much information (e.g., age, identity, sex, gender, etc.) could you determine about the face just from that one number?

As you probably noticed in the previous exercise, a major drawback of the encoding we worked out previously is that each scaling factor doesn't tell us all that much useful information about each face (and as a result we need a lot of these numbers to specify a particular face). It turns out that we can fix a lot of these shortcomings by carefully choosing our set of building block images. Reframing problems by choosing a different set of building blocks is one of the key ideas in this module.

One member of your group should go to this Google Slides presentation. Make a copy of the presentation, set the sharing so it is editable with a link, and copy the link into the Zoom chat window. Now, each group member should go to the link.

show screenshots of this process? Probably a walk-through video is best.

What you see before you is a very carefully chosen set of building block images. You should notice that each column represents a different building block image and each row represents a different scaled version of that same building block. Today, we won't be going into detail about *how* we determined these particular building blocks. Instead, you will experiment with these building blocks in order to understand some of their properties.

- You can add these scaled building block images by overlapping them on the Google Slides presentation .

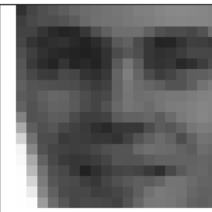
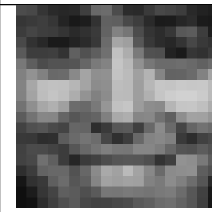
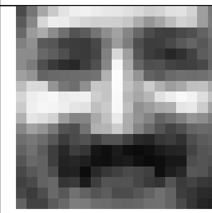
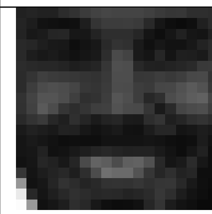
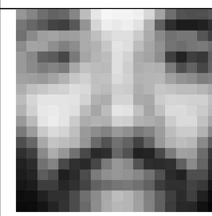
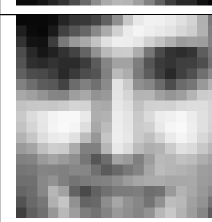
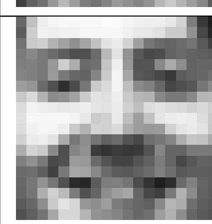
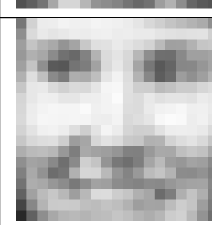
  post video

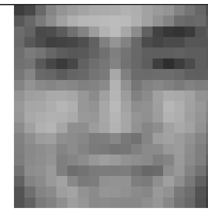
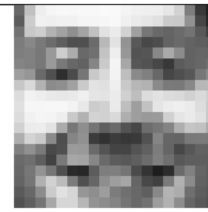
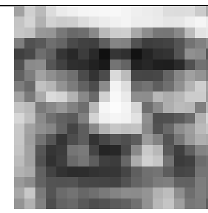
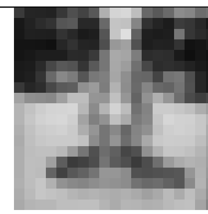
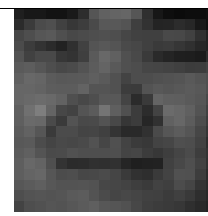
- Along with these building block images, we have determined optimal encodings for a number of different faces. Pick a few of these faces and try assembling them. Take turns so each member has a chance to try it.

- Have one of your group members choose a face and create its encoding (don't tell the rest of your group who you picked) The other members should try to guess which face it is.

  Note: that each column in the table corresponds to one of the building block images (column in the Slides presentation). Higher numbers in the table correspond to choosing the darker (more saturated) versions of each building block image. If a 0 appears for a particular building block, don't include that building block at all to construct a particular face.

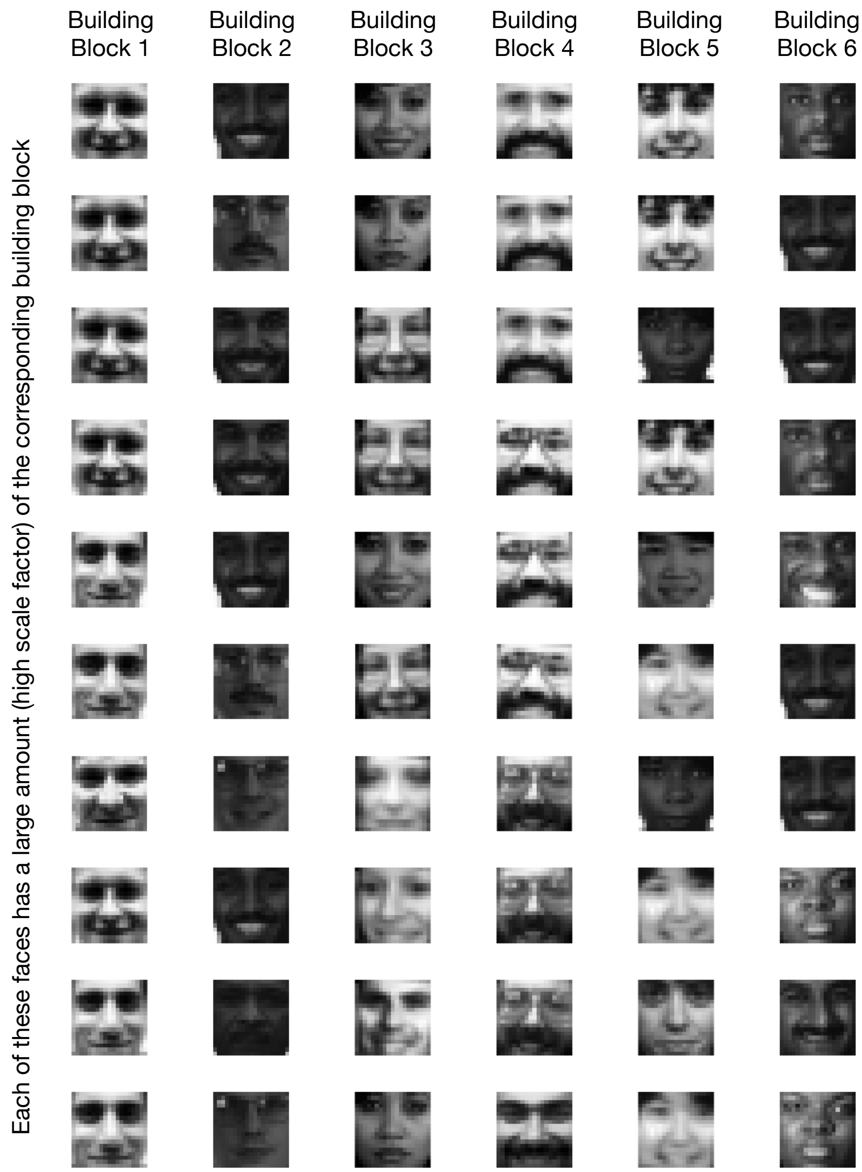
Intensity 1	**Intensity 2**	**Intensity 3**	**Intensity 4**	**Intensity 5**	**Intensity 6**	**face image**

3	3	0	0	2	2	
0	2	3	2	3	1	
3	0	1	4	1	1	
1	4	1	2	3	4	
0	0	3	3	2	1	
2	0	3	0	2	0	
2	0	1	3	0	1	
1	2	0	1	0	0	

1	1	3	0	1	2	
2	0	1	3	0	1	
3	2	0	2	0	1	
2	1	3	0	3	1	
0	3	0	1	3	3	
2	3	0	1	2	1	

- How many numbers do you now need to encode a 19 pixel by 19 pixel face? (For the purposes of this problem, don't count the numbers needed to encode each of the building blocks. You can assume those are already given to you.)
- Can you encode any possible face with this set of building blocks? If not, what seem to be the limitations?
- How well does this set of building blocks work for encoding these faces? Does it seem to work equally well across all faces? Which faces does it work well on (i.e., they can accurately be reconstructed from the building blocks) and which faces does it work poorly on?
- Looking at the building blocks themselves, what does each building block seem to represent? In other words, as you increases the amount of a particular building block, what features or qualities does that

impart on the resulting face. To help you think this through, below we have a grid of faces where each row corresponds with one of the six building block images and each of the faces in the row contains a large amount of that particular building block image in its encoding.



2.4 Towards an Optimal Basis [30 mins]

Exercise 2.5

In this question, we want you to think about process rather than particular techniques for solving this problem. If you have questions on what we mean by this, let us know.

Suppose someone has hired you as a consultant to create a method to encode 19 pixel by 19 pixel images of faces (similar to the ones you just experimented with) as a sum of scaled versions of just 10 building block images.

1. What questions would you want to ask the person that hired you in order to do a good job on this project? (i.e., what information do you need to know?)

2. What might be some qualities of a good set of building block images? (e.g., how would they look? what sort of dimensions of variability would they have?)

3. What sort of data might you need to collect in order to inform the set of building blocks you will ultimately deliver (this data could be images or it could be other quantitative or qualitative data)?

4. How might you determine whether your method is working (these could be quantitative measurements or qualitative observations of your system)?

5. Are there any other steps might you want to take to complete the project?

6. We will be digging into the various dimensions of the use of facial recognition technology in society later in this module, but for now we want to get you thinking about two particular components of that. Many face processing technologies work best on white males (e.g., check out the Gender Shades project). One possible explanation for this phenomenon is overt bias on the part of the creators of these technologies. Instead, for the sake of this exercise, let's suppose that the differences in performance are actually the result of subtle, unconscious bias in any number of decisions that the technology creators made during the design process. A second problem that plagues face processing algorithms is that they seem to work great when evaluated in the settings that the technology designers had in mind when they built the technology, but often work poorly when deployed in the real world. Looking back on the steps you listed above, flag steps that might have the potential to introduce bias into your system (e.g., having your system work better on one group of people than another or having it fail in a particular use case). It's okay if you don't know where bias might creep in, the purpose of this exercise is to get you asking questions rather than reaching conclusions.

Chapter 3

Homework 1: Introduction to Matrices

Schedule

Overview and Orientation

In this night assignment, we will learn some of the fundamental material about matrices and matrix operations.

Learning Objectives

Concepts

- Define a vector, a matrix and an array
- Describe the meaning of the dimensions of a vector, a matrix, and an array
- Give at least one interpretation of matrix-vector multiplication
- Calculate the product of a matrix-vector multiplication
- Understand dimensionality-requirements for matrix-vector multiplication and predict resulting dimensions
- Define and recognize the following special matrices: Identity, diagonal, square, rectangular, symmetric

MATLAB skills

- Determine the dimensions of a vector, matrix, or array variable
- Perform operations (addition, multiplication, transposition) on matrices
- Extract desired subarrays or matrices from arrays

Suggested Approach

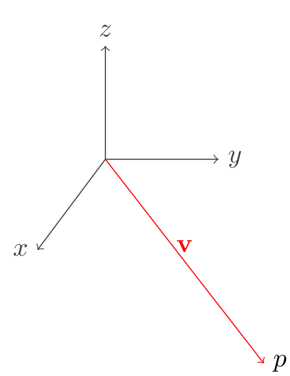
- First you should quickly scan through the assignment, see what is being asked, and assess the extent to which you already know how to do things. Spend no more than 30 minutes or so doing this.
- You should then read the assignment more closely, try out problems, and if appropriate, look at some of the other resources that are suggested. Don't spend more than 1 hour poking around at stuff online unless it is really being productive: it's easy to spend a lot of time there without accomplishing much.
- Then start doing the problems in earnest, and/or spend focused time with suggested resources.
- Once you've spent a total of 3-4 hours working on the assignment, you should check your progress. Are you on track to finish within about 7-8 hours? Do you feel confident that you can do the stuff that's left? If not, this is when you should ask for help. This means talk to a colleague, or talk to a ninja, or send an email to an instructor.
- You should turn in a PDF document with answers to all the numbered questions below. For the MATLAB assignments, please export your work to pdf. Please carefully label the problem number in your MATLAB script.

Resources to read and watch

There are lots of books about Linear Algebra and lots of useful videos on the web. Here are some specific recommendations:

- Introduction to Linear Algebra, by Strang
- Linear Algebra, by Lay
- Linear Algebra, by Cherney, Denton, Thomas, Waldron
- Homebrew videos
  - Matrices operating on vectors
  - Matrices operating on vectors (example)
  - Matrices operating on matrices
- Videos from others
  - Vectors, the very basics
  - 3Blue1Brown's YouTube series on Linear Algebra

3.1 Vectors



Consider the point $p = (1, 2, -1)$ in 3-dimensional space. We can associate a position vector $\mathbf{v}$ with this point, which is the vector from the origin to this point,

$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

Likewise, we can think of every vector as defining a point, if we assume that the vector emanates from the origin. So, for example, the vector

$$\mathbf{v} = \begin{bmatrix} 3 \\ -2 \\ 0 \\ 1 \end{bmatrix}$$

is identified with the point $(3, -2, 0, 1)$ in 4D. Often times we will mix and match these ideas and say things like: the vector (x, y, z). What we really mean when we say this is: the point (x, y, z) can be treated as the position vector

$$\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

The vector $\mathbf{v}$, as represented above, is called a column vector. We can also have row vectors such as the following

$$\mathbf{u} = \begin{bmatrix} p & q & r \end{bmatrix}.$$

The operation of converting a column vector to a row vector or vice-versa is called taking the *transpose* of the vector and is denoted with a superscript T. For example, the transpose of the row vector $\mathbf{u}$ from above is

$$\mathbf{u}^T = \begin{bmatrix} p \\ q \\ r \end{bmatrix} \tag{3.1}$$

and the transpose of the vector $\mathbf{v}$ from above is

$$\mathbf{v}^T = \begin{bmatrix} x & y & z \end{bmatrix}. \tag{3.2}$$

We can take the product of a row vector with a column vector using the following formula

$$\mathbf{uv} = \begin{bmatrix} p & q & r \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = px + qy + zr \tag{3.3}$$

If we start with two column vectors

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \text{ and } \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

of length n (i.e., they are n-dimensional), then we can take the *d*ot product

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n.$$

In some sense, the dot product is a measure of how aligned two vectors are. Here's the key formula:

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos\theta$$

where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$ and

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

is the length of the vector $\mathbf{v}$ in n-dimensional space.

Exercise 3.1

1. Assume $\mathbf{v}$ and $\mathbf{w}$ are two vectors of unit length, i.e., $\|\mathbf{v}\| = \|\mathbf{w}\| = 1$. Using the formula above, what angle between $\mathbf{v}$ and $\mathbf{w}$ maximizes the dot product? Using the formula above, what angle between $\mathbf{v}$ and $\mathbf{w}$ minimizes the dot product?

2. Compute $\mathbf{v} \cdot \mathbf{w}$ where

$$\mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ -4 \\ 6 \end{bmatrix}, \text{ and } \mathbf{w} = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 3 \end{bmatrix}$$

Solution 3.1

1. Using the formula, $\mathbf{v} \cdot \mathbf{w} = \cos(\theta)$. So, when $\theta = 0$, (i.e., the vectors point in the same direction) the dot product is maximized and when $\theta = \pi/2$ (i.e., the vectors are perpendicular) the dot product is minimized.

2. The dot product is

$$\mathbf{v} \cdot \mathbf{w} = -2 + 0 - 4 + 18 = 12$$

We'll learn more about the dot product as we go. For now, notice that the dot product equals the product of the transpose of one with the other

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^T \mathbf{w}. \tag{3.4}$$

Vectors can also be used to represent many things, such as data. Linear algebra provides a powerful set of tools to manipulate and analyze this data.

Exercise 3.2

For instance, you may have a three-dimensional vector $\mathbf{f}$ whose entries represent the numbers of different fruits you have in your refrigerator. For example, the first entry could be the number of oranges, the second the number of grapefruits and the third could be the number of apples. When organized in this manner, you can use products of row and column vectors to compute the number of different fruits there are. For instance, suppose that

$$\mathbf{f} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \tag{3.5}$$

i.e. you have 1 orange, 2 grapefruits, and 3 apples in your fridge.

1. Find a row vector $\mathbf{t}$ so that the product $\mathbf{tf}$ tells you the total number of fruits in your refrigerator.

2. Find a row vector $\mathbf{c}$ such that the product $\mathbf{cf}$ tells you the total number of *citrus* fruits in your refrigerator.

3. Suppose that in the genetically engineered future, all apples weigh 100 g, all grapefruits weigh 250 g and all oranges weigh 120 g. Find a row vector $\mathbf{w}$, such that the product $\mathbf{wf}$ tells you the total weight of fruits in your refrigerator.

Solution 3.2

1. Let $\mathbf{t} = \begin{bmatrix}1 & 1 & 1\end{bmatrix}$. Then $\mathbf{tf} = 1 + 2 + 3 = 6$, the total number of fruits in your refrigerator.
2. Let $\mathbf{c} = \begin{bmatrix}1 & 1 & 0\end{bmatrix}$. Then $\mathbf{cf} = 1 + 2 + 0 = 3$, the total number of citrus fruits in your refrigerator.
3. Let $\mathbf{w} = \begin{bmatrix}120 & 250 & 100\end{bmatrix}$. Then $\mathbf{wf} = 120 + 500 + 300 = 920$, the total weight of the fruits in your refrigerator.

If you wanted to know the vitamin C content of the fruits in your fridge, you could formulate a similar vector to compute it.

In the questions above, you took *linear combinations* of the entries of the vector $\mathbf{f}$ which gave you the desired quantity. *Linear algebra is the study of linear combinations.*

3.2 Matrices

Matrices are a set of numbers organized in a two-dimensional array. Matrices are a compact way to represent linear combinations. Matrices can also be used in a number of different ways, such as to represent data. When we multiply a matrix by a vector, it results in a new vector. Therefore, when we say "a matrix operates on a vector", we mean that the matrix multiplies the vector. Notation-wise, we use bold upper-case letters, e.g. $\mathbf{A}$, to represent a matrix and bold lower-case letters to represent a vector, e.g. $\mathbf{v}$.

For instance, you may define a two-dimensional matrix $\mathbf{G}$ with two rows and three columns as follows

$$\mathbf{G} = \begin{bmatrix}1 & 1 & 1\\1 & 1 & 0\end{bmatrix}. \tag{3.6}$$

Matrices and vectors come in different shapes and sizes and we refer to their shape and size by the number of rows and columns they have. A general matrix $\mathbf{A}$ has m rows and n columns, and we refer to this as an $m \times n$ matrix. Vectors are then examples of matrices: row vectors have a single row, i.e., they are $1 \times n$ matrices; and column vectors have a single column, i.e., they are $m \times 1$ matrices.

Matrices can only multiply vectors of a certain size and produce vectors of a certain size: an $m \times n$ matrix can only operate on a column vector of size $n \times 1$, and will produce an output vector which is a column vector of size $m \times 1$. (Likewise, matrices can only multiply other matrices of a certain size: an $m \times n$ matrix can only act on a matrix of size $n \times k$, and will produce an output matrix of size $m \times k$.) These basic properties will become clearer when we look at an example.

Consider the 3×2 matrix $\mathbf{A}$,

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

and the input vector $\mathbf{v}$

$$\mathbf{v} = \begin{bmatrix}-2\\1\end{bmatrix}.$$

The output vector $\mathbf{w}$ is computed as follows

$$\mathbf{w} = \begin{bmatrix}2 & 1\\3 & -1\\0 & 4\end{bmatrix}\begin{bmatrix}-2\\1\end{bmatrix} = \begin{bmatrix}(2)(-2) + (1)(1)\\(3)(-2) + (-1)(1)\\(0)(-2) + (4)(1)\end{bmatrix} = \begin{bmatrix}-3\\-7\\4\end{bmatrix}$$

There are two main ways to think about this multiplication. The most common view is to treat each entry of the new vector as a dot product between a row of the matrix and the column vector. So, for example, the first entry in the output vector is the dot product of two vectors

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

The second approach is to view the output vector as a linear combination of the columns of the matrix. The entries in the original vector are used as multiplication weights on each column of the matrix, i.e.

$$(-2)\begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} + (1)\begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} -3 \\ -7 \\ 4 \end{bmatrix}$$

We encourage you to use both approaches when you think about multiplication.

Exercise 3.3

Recall the matrix $\mathbf{G}$ defined in equation (3.6) and the vector $\mathbf{f}$ defined in Exercise 3.2, which kept track of the number of fruit of different types. What does the vector $\mathbf{Gf}$ represent?

Solution 3.3

The vector $\mathbf{Gf}$ is a 2×1 vector whose first entry represents the total number of fruits and second entry represents the number of citrus fruits.

Exercise 3.4

If a matrix multiplies a spatial vector, the resulting vector is *transformed* by the matrix, resulting in a new vector.

1. Please draw the spatial vector

$$\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{3.7}$$

2. Please draw the vector $\mathbf{w} = \mathbf{Av}$, where $\mathbf{A}$ is

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{3.8}$$

3. What happened to $\mathbf{v}$ when you multiplied by $\mathbf{A}$?

4. Please draw the vector $\mathbf{u} = \mathbf{Bv}$, where $\mathbf{B}$ is

$$\mathbf{B} = \begin{bmatrix} \cos(30^\circ) & -\sin(30^\circ) \\ \sin(30^\circ) & \cos(30^\circ) \end{bmatrix} \tag{3.9}$$

5. What happened to $\mathbf{v}$ when you multiplied by $\mathbf{B}$?

6. Please draw the vector $\mathbf{t} = \mathbf{Rv}$, where $\mathbf{R}$ is

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{3.10}$$

7. What happened to $\mathbf{v}$ when you multiplied by $\mathbf{R}$?

8. Please draw a new spatial vector

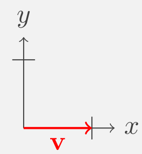
$$\mathbf{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{3.11}$$

9. Please draw the vector $\mathbf{s} = \mathbf{Rw}$

10. What does multiplying *any* vector by $\mathbf{R}$ do?

Solution 3.4

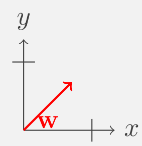
1. The vector $\mathbf{v}$ is



2. First, we compute

$$\mathbf{w} = \mathbf{A}\mathbf{v} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix},$$

which is visually represented as

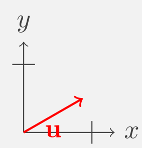


3. Multiplying $\mathbf{v}$ by $\mathbf{A}$ rotated the vector counterclockwise by 45 degrees.

4. First we compute

$$\mathbf{u} = \mathbf{B}\mathbf{v} = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$$

which is visually represented as

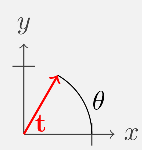


5. Multiplying $\mathbf{v}$ by $\mathbf{B}$ rotated the vector counterclockwise by 30 degrees.

6. First we compute

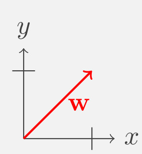
$$\mathbf{t} = \mathbf{R}\mathbf{v} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

which is visually represented as

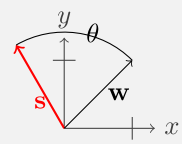


7. Multiplying $\mathbf{v}$ by $\mathbf{R}$ rotated the vector counterclockwise by θ degrees.

8. The vector $\mathbf{w}$ is



9. First we compute
$$\mathbf{s} = \mathbf{R}\mathbf{w} = \begin{bmatrix} \cos\theta - \sin\theta \\ \sin\theta + \cos\theta \end{bmatrix}$$
which is visually represented



10. Multiplying any vector by $\mathbf{R}$ rotates it by θ.

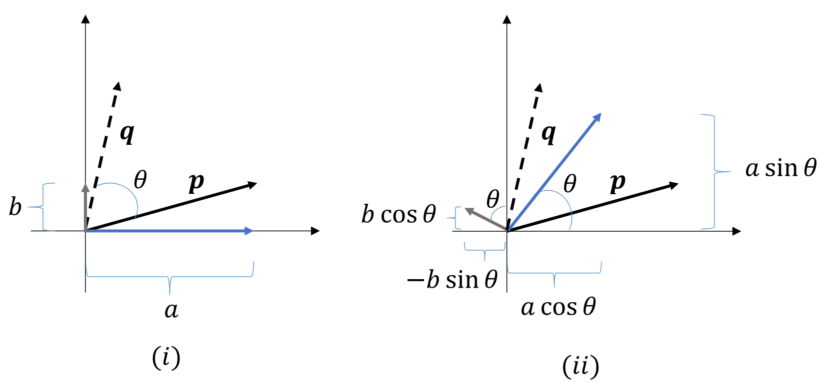


Figure 3.1: Rotation of vectors

You may have guessed that $\mathbf{R}$ defined above, rotates a vector counter-clockwise by θ. This is indeed true, and $\mathbf{R}$ is called a *rotation matrix* as it transforms vectors by rotating them. To understand why $\mathbf{R}$ is a rotation matrix, consider Figure 3.1 (i). Suppose that we wish to rotate the vector $\mathbf{p}$ counter-clockwise by θ, which will result in the vector $\mathbf{q}$. From the figure, we see that

$$\mathbf{p} = \begin{bmatrix} a \\ b \end{bmatrix}, \tag{3.12}$$

and $\mathbf{p}$ is the sum of the gray and blue vectors. If we now rotate the blue and gray vectors counter-clockwise by θ, we see that $\mathbf{q}$ is the sum of the rotated versions of the blue and gray vectors, as shown in Figure 3.1 (ii). By using trigonometry, we see that the blue vector in Figure 3.1 (ii) is

$$\begin{pmatrix} a\cos\theta \\ a\sin\theta \end{pmatrix} \tag{3.13}$$

and the gray vector in Figure 3.1 (ii) is

$$\begin{pmatrix} -b\sin\theta \\ b\cos\theta \end{pmatrix} \tag{3.14}$$

Therefore, $\mathbf{q}$ is given by

$$\mathbf{q} = \begin{pmatrix} a\cos\theta \\ a\sin\theta \end{pmatrix} + \begin{pmatrix} -b\sin\theta \\ b\cos\theta \end{pmatrix} = + \begin{pmatrix} a\cos\theta - b\sin\theta \\ a\sin\theta + b\cos\theta \end{pmatrix} = \mathbf{R}\mathbf{p}. \tag{3.15}$$

As we mentioned earlier, $m \times n$ matrices can multiply $n \times 1$ vectors and produce $m \times 1$ vectors. Consider a generic $m \times n$ matrix $\mathbf{A}$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \cdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

where the ij-th entry of this matrix, a_{ij} defined above, is the entry corresponding to the i-th row and j-th column. You can multiply an $n \times 1$ vector $\mathbf{v}$ by this matrix. Define the vector $\mathbf{v}$ as follows,

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

Now define another vector $\mathbf{w}$ which is the product of $\mathbf{A}$ and $\mathbf{v}$, i.e., $\mathbf{w} = \mathbf{A}\mathbf{v}$. If we define

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix}$$

then the i-th entry of $\mathbf{w}$, is given by the following sum

$$w_i = a_{i1}v_1 + a_{i2}v_2 \cdots a_{im}v_m = \sum_{j=1}^{n} a_{ij}v_j$$

Besides multiplication, a number of other operations can be done using matrices including addition, subtraction, inversion, transposition, etc. We will explore more of these and their associated properties now and later. All of these operations make matrices a very powerful tool in the study of many different systems which can be represented as linear transformations, or combinations.

3.3 Addition, subtraction, multiplication, and transpose of matrices

We can add matrices of the same size, and subtract them from one another. Both operations result in matrices of the same size and shape. The addition and subtraction operations are done element-wise. For instance the difference of the two matrices can be calculated as below

$$\mathbf{A} = \begin{bmatrix} 3 & 4 & 1 \\ 3 & 1 & 1 \end{bmatrix} \tag{3.16}$$

$$\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 1 \end{bmatrix} \tag{3.17}$$

$$\mathbf{A} - \mathbf{B} = \begin{bmatrix} (3-1) & (4-2) & (1-3) \\ (3-2) & (2-1) & (1-1) \end{bmatrix} \tag{3.18}$$

$$= \begin{bmatrix} 2 & 2 & -2 \\ 1 & -1 & 0 \end{bmatrix} \tag{3.19}$$

Multiplying a matrix by a scalar simply scales each entry of the matrix by the scale factor. For instance

$$3\mathbf{A} = \begin{bmatrix} 9 & 12 & 3 \\ 9 & 3 & 3 \end{bmatrix} \tag{3.20}$$

The transpose of a vector, denoted by the superscript ^{T} turns a column vector into a row vector, and vice versa. For matrices, the transpose replaces the rows with the columns (or vice-versa). For example,

$$\begin{bmatrix} 1 & 3 & 5 \\ 2 & 7 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 \\ 3 & 7 \\ 5 & 6 \end{bmatrix} \tag{3.21}$$

Since the columns are replaced with the rows, the shape of the matrix changes when you transpose it. The following property of transposes will be useful moving forward. Consider a matrix $\mathbf{A}$ and a vector $\mathbf{v}$. Then

$$(\mathbf{A}\mathbf{v})^T = \mathbf{v}^T\mathbf{A}^T \tag{3.22}$$

Exercise 3.5

Using $\mathbf{A}$ and $\mathbf{B}$ previously defined, evaluate $4\mathbf{A} - 5\mathbf{B}$

Solution 3.5

$$4\mathbf{A} - 5\mathbf{B} = \begin{bmatrix} 7 & 6 & -11 \\ 2 & -6 & -1 \end{bmatrix} \tag{3.23}$$

Exercise 3.6

If the matrix $\mathbf{A}$ has dimensions of 4×5, what are the dimensions of $\mathbf{A}^T$?

Solution 3.6

The dimensions of $\mathbf{A}^T$ are 5×4.

Exercise 3.7

If the matrix $\mathbf{A}$ is 4×5 (i.e., $\mathbf{A}$ has dimensions 4×5) and the vector $\mathbf{v}$ is 5×1, what are the dimensions of $\mathbf{A}\mathbf{v}$ and $(\mathbf{A}\mathbf{v})^T$?

Solution 3.7

$\mathbf{A}\mathbf{v}$ is 4×1 and $(\mathbf{A}\mathbf{v})^T$ is 1×4.

Matrices can be multiplied together to produce other matrices. In general, when you multiply a matrix $\mathbf{A}$ with another matrix $\mathbf{B}$, you need the matrix on the left side of the product to have the same number of columns as the number of rows in the matrix on the right side. In other words if $\mathbf{A}$ is $m \times n$, and $\mathbf{B}$ is $p \times q$, you need $n = p$ for the product $\mathbf{C} = \mathbf{A}\mathbf{B}$ to be defined. The product results in a new matrix $\mathbf{C}$ which is $m \times q$. The q columns of the product matrix $\mathbf{C}$ are precisely the q vectors that would result from multiplying $\mathbf{A}$ with the vectors formed by the columns of $\mathbf{B}$.

Consider the following matrices

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

The product of the two $\mathbf{C} = \mathbf{AB}$ is computed as follows

$$\mathbf{C} = \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} (2)(1) + (1)(-2) & (2)(5) + (1)(3) \\ (3)(1) + (-1)(-2) & (3)(5) + (-1)(3) \end{bmatrix} = \begin{bmatrix} 0 & 13 \\ 5 & 12 \end{bmatrix}$$

As a second example consider the matrices $\mathbf{A}$ and $\mathbf{B}$ defined below, and let the product $\mathbf{C} = \mathbf{AB}$.

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

$$\mathbf{C} = \begin{bmatrix} (1)(1) + (2)(2) & (1)(4) + (2)(3) \\ (3)(1) + (2)(2) & (3)(4) + (2)(3) \\ (4)(1) + (1)(2) & (4)(4) + (1)(3) \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 7 & 18 \\ 6 & 19 \end{bmatrix}$$

As mentioned above, one way of envisioning matrix multiplication is if we consider the columns of input matrix $\mathbf{B}$ as a set of column vectors, we can multiply these column vectors one at a time by the matrix $\mathbf{A}$, and the resulting vectors will be the corresponding columns of the output matrix $\mathbf{C}$, i.e.

$$\mathbf{AB} = \mathbf{A}[\mathbf{B}_1, \mathbf{B}_2, \ldots] = [\mathbf{AB}_1, \mathbf{AB}_2, \ldots]$$

where $\mathbf{B}_1$ is the first column of matrix $\mathbf{B}$ etc.

Consider the following matrices:

$$\mathbf{A} = \begin{bmatrix} -2 & 4 \\ 0 & 3 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 5 & -3 \\ -1 & -1 \end{bmatrix}$$

Exercise 3.8

Find the matrix product $\mathbf{AB}$.

Solution 3.8

$$\mathbf{AB} = \begin{bmatrix} -14 & 2 \\ -3 & -3 \end{bmatrix}$$

Exercise 3.9

Find the matrix product $\mathbf{BA}$

Solution 3.9

$$\mathbf{BA} = \begin{bmatrix} -10 & 11 \\ 2 & -7 \end{bmatrix}$$

Note that these two products are NOT equal. In general, matrix multiplication, unlike scalar multiplication, is NOT commutative. In other words, in general $\mathbf{AB} \neq \mathbf{BA}$. However, the distributive property IS valid for matrices: $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ so long as we keep the order of the multiplication the same

$(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA}+\mathbf{CA}$. Recall the definition of matrix addition: if two matrices are of the same size then they can be added and each entry of the new matrix is the sum of the entries of the original matrices, e.g.

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

In addition to matrices $\mathbf{A}$ and $\mathbf{B}$ defined above, consider the matrix

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

Exercise 3.10

Calculate $\mathbf{A}(\mathbf{B}+\mathbf{C})$.

Solution 3.10

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \begin{bmatrix} -16 & 12 \\ -12 & 3 \end{bmatrix}$$

Exercise 3.11

Calculate $\mathbf{AB}+\mathbf{AC}$. Is it equal to your previous answer?

Solution 3.11

It is the same answer, as expected, since you can distribute matrices.

Finally, since matrix multiplication is defined, there is no reason not to multiply a matrix by itself. This only works if it is a square matrix. (Think about why this is true.) Using $\mathbf{A}$ and $\mathbf{B}$ from above, evaluate the following expressions

Exercise 3.12

1. $\mathbf{A}^2$
2. $\mathbf{B}^3$

Solution 3.12

1. $$\mathbf{A}^2 = \begin{bmatrix} 4 & 4 \\ 0 & 9 \end{bmatrix}$$
2. $$\mathbf{B}^3 = \begin{bmatrix} 152 & -72 \\ -24 & 8 \end{bmatrix}$$

Exercise 3.13

There are lots of matrices that are special. Use a trusted linear algebra reference to define the following types of matrices, and provide an example of each:

1. Square Matrix
2. Rectangular Matrix
3. Diagonal Matrix
4. Identity Matrix
5. Symmetric Matrix

Solution 3.13

1. A square matrix is one that has size $n \times n$, e.g.,

$$\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}.$$

2. A rectangular matrix is one that has size $m \times n$ where n is not equal to m, e.g.,

$$\begin{bmatrix} * & * & * \\ * & * & * \end{bmatrix}.$$

3. A diagonal matrix is one whose only non-zero elements are on the diagonal from upper left to lower right, e.g.,

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

4. The identity matrix is a square matrix with all zeroes except along the diagonal from the upper left to lower right, where the entries are all 1, e.g.,

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

5. A matrix is symmetric if it is square and equal to its own transpose, i.e. $A = A^T$, e.g.,

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

When matrices operate on (i.e., multiply) spatial position vectors, the vector which results is another spatial position vector. The original spatial position has been 'transformed' into another position. In particular, there are specific matrices which accomplish specific desired transformations. These are used in many different disciplines.

The matrix

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.24}$$

when multiplying the vector

$$\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{3.25}$$

will reproduce the same vector, i.e. $\mathbf{Iv} = \mathbf{v}$. For this reason, the matrix $\mathbf{I}$ above is called an identity matrix. Identity matrices in higher dimensions are defined the same way, i.e., a 4-dimensional identity matrix is a 4×4 matrix with 1s on the diagonal and zeros everywhere else, i.e.,

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.26}$$

Exercise 3.14

1. Another important and simple operation is to be able to take a vector and scale (increase or decrease its length) it by an overall multiplicative factor while maintaining its direction. Consider the vector

   $$\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

   Thinking about how the identity matrix acts on this vector, propose a 3×3 matrix which scales this vector by a factor of 3 to the vector

   $$3\mathbf{v} = \begin{bmatrix} 3x \\ 3y \\ 3z \end{bmatrix}.$$

   In other words, find a 3×3 matrix $\mathbf{M}$ such that $\mathbf{Mv} = 3\mathbf{v}$ for any vector $\mathbf{v}$.

2. What if you want to scale the x component differently than the y component? Write down the 3×3 matrix which scales the x component by 3 and the y component by 5 and leaves the z component the same.

3. Write down the 3×3 matrix which scales the x component by a, the y component by b, and the z component by c.

Solution 3.14

1. $$\mathbf{M} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

2.
$$\mathbf{M} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3. We can generalize the result:
$$\mathbf{M} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

3.4 Matrix Operations in MATLAB

Exercise 3.15

In the command window, you can type in commands and press enter. Try the following commands and see what they do.

```
1+1
a=1+1
a
% you can start a comment with "%"
b=2;% this will appear as a variable in your workspace, but the semicolon ...
    suppresses the output
c=3,d=4,e=5;% use commas, semicolons, or shift+enter between commands that you ...
    want to execute together
1+2-(3*4/5)^6
clear a
a% should give you an error because a is not defined anymore
clear all
clc% only if you want to clear your workspace!
```

Now you will work on examples of matrices multiplying vectors to get yourselves comfortable with matrix operations in MATLAB. First, let's define the matrix $\mathbf{A}$ using MATLAB as follows

```
>> A = [2 1; 3 -1; 0 4]
```

Note that the semi-colon ends a row and begins a new row. To define the column vector $\mathbf{v}$ in MATLAB you can type the following command:

```
>>  v = [-2; 1]
```

whilst to define the row vector $\mathbf{u}$ in MATLAB you can type the following command

```
>>  u = [2 -3 1]
```

Notice that in this case each component of the vector is separated by a space - you could also separate them with a comma.

Exercise 3.16

Using the definitions for $\mathbf{A}$, $\mathbf{v}$, and $\mathbf{u}$ from above, please solve the following using MATLAB. Do the answers match what you expect? (Not all of these may be defined!)

1. A*v

2. `u*A`
3. `A*u`
4. `v*A`
5. `A(1:2,:)*v`
6. `u*A(:,2)`
7. `A(:, 2:4)*v`
8. `u*A(1,:)`

Solution 3.16

1. `[-3; -7; 4]`
2. `[-5 9]`
3. Does not work because the inner matrix dimensions must agree and here we have a 3×2 matrix multiplied by a 1×3 matrix
4. Does not work because the inner matrix dimensions must agree and here we have a 2×1 matrix multiplied by a 3×2 matrix
5. `[-3; -7]`
6. 9
7. Does not work because the index exceeds matrix dimensions. It is trying to access columns 2-4 of a two column matrix.
8. Does not work because the inner matrix dimensions must agree and here we have a 1×3 matrix multiplied by a 1×2 matrix.

3.5 Data in Matrices and Vectors

Most of the examples you saw up to now in this assignment involved vectors which represent spatial positions, and most of the matrices you encountered represent transformations of the spatial vectors. But, as you saw with the example involving fruits, vectors can also be used to store data. So can matrices.

For instance, you may have the following matrix

$$\begin{bmatrix} 41 & 35 & 37 & 43 \\ 49 & 40 & 48 & 61 \end{bmatrix} \tag{3.27}$$

whose first row represents the forecasted high temperature in Needham for the next 4 days (as of the day this was written) and the second row represents the forecasted high temperatures for Washington DC. By representing this data in matrix form, you can do a number of operations to help extract useful information from the data.

Exercise 3.17

For this exercise, you will work with historical temperature data for the cities of Boston, Providence, Washington DC and New York.

1. Download the file `temps.mat` from canvas and load the data in it into MATLAB using » `load temps.mat`. You should now have access to a matrix T which contains daily average temperatures from 1995 to 2015 for the cities of Boston, Providence, Washington DC and New York (we are not telling you in what order yet). By using MATLAB's `size` function, determine the dimensions of this matrix. Are the temperatures for each city contained in the rows or the columns of this matrix?

2. Extract the temperatures for each city into 4 different vectors `t1, t2, t3, t4`, and check that the dimensions of these vectors are as expected.

3. Find the average temperature of each city using MATLAB's `mean` function, and guess, based on geography, which of the vectors corresponds to the temperature for which city.

4. What are the maximum and minimum temperatures for Boston in the 20 years for which you have data? As you might expect there are MATLAB functions called `max` and `min`.

5. On the same axes, plot graphs for the daily temperatures for the four cities for the last year for which you have data. Use MATLAB's `legend, xlabel, ylabel` functions to label the graphs.

6. Suppose that a genie told you that you can guess the temperature of New York, which we call T_n, using the temperatures of Boston, Providence, and Washington DC, which we respectively call T_b, T_p and T_w. From the matrix T, extract a 3×365 matrix of daily temperatures for the last year (for which you have data) in Boston, Providence and Washington DC.

7. The genie says that a good approximation for the temperature on a given day in New York is given by

$$T_n \approx 0.2235 T_b + 0.4193 T_p + 0.3856 T_w. \tag{3.28}$$

   Formulate a matrix equation which uses the matrix from the previous part and the formula from the genie to guess the daily temperature in New York for the last year. Apply this equation in MATLAB.

8. On the same axes, plot your prediction for the temperature in New York from the previous part, and the true temperature data which you extract from T. Is the prediction close?

In the course of this module, you will learn how to come up with the coefficients we provided here using historical data. (No, we don't actually have a genie.)

Solution 3.17

1. After loading the temperatures you can see that they are stored inside a matrix called T which has 4 rows and 7670 columns, so presumably the temperature for each city is stored in a row.

2. We can extract the first temperature by typing the following » `t1 = T(1,:)` - this simply grabs all of the elements in the first row, so that `t1` should be a row vector of size 1 by 7670. We create the other vectors in a similar way.

3. We can take the mean of the first city by typing » `mean(t1)` and we get 51.7667. The other means respectively are 51.9140, 58.4365, and 55.9451. A little bit of geography suggests that the cities are ordered as follows: Boston, Providence, DC, New York.

4. We can compute the maximum by typing » `max(t1)` and we get 90.7. The minimum is 0.7.

5. We are only supposed to grab the last year (365 days) so for Boston we would type » `plot(t1(end-364:end))`, or we could use the actual size of the vector.

6. Boston, Providence, and DC are stored in the first three rows. We'll extract their data and store it in a new matrix S by typing » `S = T(1:3,end-364:end)`, which grabs the first three rows and the last 365 entries.

7. We can define Tn by typing » `Tn = 0.2235*S(1,:) + 0.4193*S(2,:) + 0.3856*S(3,:)` since the city temperatures are stored in the each of the three rows of the matrix S.

8. Graphically they look pretty good. We can also examine the data a little more closely by looking at the difference between the predicted temperature and the actual temperature - it fluctuates with a mean of roughly 8.8115e-04, a maximum of 7.5334, and a minimum of -6.8966. Compared to the actual temperatures this implies that the prediction is never any worse than roughly 10%.

3.6 Conceptual Quiz

Please figure out the answer to these questions and mark your answer in Canvas. You can retake the quiz, as needed.

1. $\mathbf{A}$ is a 3×4 matrix and $\mathbf{B}$ is a 4×2 matrix. What is the size of $\mathbf{AB}$?

   A. 2×3
   B. 3×1
   C. 3×2
   D. The product is not defined.

2. What is the result of the following matrix product

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

   A.
$$\begin{bmatrix} -5 & -7 \\ 18 & 14 \\ 24 & 23 \end{bmatrix}$$

   B.
$$\begin{bmatrix} -5 & -7 \\ 18 & 14 \\ 29 & 23 \end{bmatrix}$$

   C.
$$\begin{bmatrix} -5 & 18 & 24 \\ -7 & 14 & 23 \end{bmatrix}$$

   D.
$$\begin{bmatrix} -5 & -7 & 3 \\ 18 & 14 & 6 \\ 24 & 23 & 9 \end{bmatrix}$$

3. Match the following items (* means any number):

1. Rectangular Matrix	A.	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
2. Diagonal Matrix	B.	$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$
3. Identity Matrix	C.	$\begin{bmatrix} * & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{bmatrix}$
4. Symmetric Matrix	D.	$\begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix}$

4. Which of the following matrices will scale the length of any 2-D vector by $\frac{1}{2}$?

   A.
   $$\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

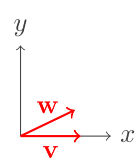
   B.
   $$\begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

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

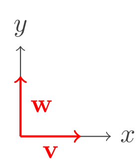
   D.
   $$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

5. All of the following vectors are unit length. In which picture is $\mathbf{v} \cdot \mathbf{w}$ the largest?

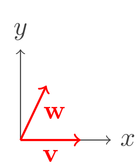
   A.



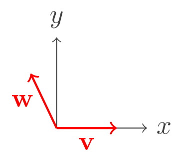
   B.



   C.



D.



Chapter 4

Week 2a: Matrix Transformations

Schedule

4.1 Debrief [15 minutes]

- With your table-mates, identify a list of key concepts/take home messages/things you learned in the assignment. Try to group them in categories like "Concepts", "Technical Details", "Matlab", etc.
- Try to resolve your confusions with your table-mates and by talking to an instructor.

4.2 Synthesis [30 minutes]

Exercise 4.1

These are fundamental ideas about matrices and it is important to complete these. They should be done by hand.

1. What is the difference between a scalar, a vector, a matrix, and an array?
2. What are the rules for adding matrices?
3. When can two matrices be multiplied, and what is the size of the output?
4. What is the distributive property for matrix multiplication?
5. What is the associative property for matrix multiplication?
6. What is the commutative property for matrix multiplication?

Solution 4.1

1. Scalars, vectors, and matrices are examples of arrays. A 0-dimensional array can be thought of as a scalar. A 1-dimensional array is a vector. A 2-dimensional array is a matrix.

2. The matrices have to be the same size and addition is element-wise.
3. The matrices have to compatible (inner dimensions agree), and the output is dictated by the outer dimensions, i.e. $(n \times m)(r \times s) = (n \times s)$.
4. Distributive property: $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
5. Associative property: $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$
6. Commutative property: Two matrices commute if $\mathbf{AB} = \mathbf{BA}$ but this is not always true.

Exercise 4.2

These are synthesis problems. It would be helpful to complete these. They should be done by hand.

1. Consider the matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Show that $\mathbf{A}^2$ commutes with $\mathbf{A}$.
2. Use the distribution law to expand $(\mathbf{A}+\mathbf{B})^2$ assuming that $\mathbf{A}$ and $\mathbf{B}$ are matrices of appropriate size. How does this compare to the situation for real numbers?
3. Show that $\mathbf{D} = \begin{bmatrix} 4 & -2 \\ 3 & -3 \end{bmatrix}$ satisfies the matrix equation $\mathbf{D}^2 - \mathbf{D} - 6\mathbf{I} = \mathbf{0}$.

Solution 4.2

1. You need to show that $\mathbf{A}^2\mathbf{A} = \mathbf{A}\mathbf{A}^2$ for this particular matrix. You can do it by multiplying.
2. Using the distributive property you can see that $(\mathbf{A}+\mathbf{B})^2 = (\mathbf{A}+\mathbf{B})(\mathbf{A}+\mathbf{B}) = \mathbf{A}^2 + \mathbf{AB} + \mathbf{BA} + \mathbf{B}^2$
3. If you plug $\mathbf{D}$ and $\mathbf{D}^2$ into the equation you should find that the result is a zero matrix.

Exercise 4.3

These are challenge problems. Pick one of them to wrestle with. It is not important to complete these. They should be done by hand.

1. The matrix exponential is defined by the power series

$$\exp \mathbf{A} = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!}$$

   Assume $\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$. Find a formula for $\exp \mathbf{A}$.
2. The real number 0 has just one square root: 0. Show, however, that the 2×2 zero matrix has infinitely many square roots by finding all 2×2 matrices $\mathbf{A}$ such that $\mathbf{A}^2 = \mathbf{0}$.
3. Use induction to prove that $\mathbf{A}^n$ commutes with $\mathbf{A}$ for any square matrix $\mathbf{A}$ and positive integer n.

Solution 4.3

1. The matrix exponential is defined by the power series $\exp \mathbf{A} = \mathbf{I} + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \ldots$. Notice that this $\mathbf{A}$ is diagonal and $\mathbf{A}^2 = \begin{bmatrix} 2^2 & 0 \\ 0 & 3^2 \end{bmatrix}$ and the exponential becomes $\exp \mathbf{A} = \begin{bmatrix} 1 + 2 + 2^2/2! + \ldots & 0 \\ 0 & 1 + 3 + 3^2/2! + \ldots \end{bmatrix}$. If you have seen power series before then you will recognise that $\exp \mathbf{A} = \begin{bmatrix} \exp 2 & 0 \\ 0 & \exp 3 \end{bmatrix}$.
2. You can define a general two by two matrix $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, find $\mathbf{A}^2$, set each of the entries equal to zero and find constraints on the entries a, b, c, d.
3. You need to show that $\mathbf{A}^n\mathbf{A} = \mathbf{A}\mathbf{A}^n$ for any square matrix $\mathbf{A}$ and any positive integer n by induction. First you show it is true for $n = 1$ and $n = 2$. Then assume it is true for some $n = k$, and prove that it must be true for $n = k + 1$. You use the fact that $\mathbf{A}$ commutes with itself and the associative property, i.e $\mathbf{A}^2\mathbf{A} = (\mathbf{A}\mathbf{A})\mathbf{A} = \mathbf{A}(\mathbf{A}\mathbf{A}) = \mathbf{A}\mathbf{A}^2$.

4.3 2D Rotation Matrices [45 minutes]

We're going to think about how to use rotation matrices to rotate a geometrical object. In doing so we will solidify fundamental concepts around matrix multiplication and start to explore the notion of "inverse". For clarity we will first work in 2D. Recall that the rotation matrix $\mathbf{R}(\theta)$:

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

will rotate an object counterclockwise **about the origin** through an angle of θ.

Exercise 4.4

This is a hands-on, conceptual problem involving the multiplication of 2D rotation matrices.

1. Place an object on your table, and imagine that the origin of an xy-coordinate system is at the center of your object with $+z$ pointing upwards.
2. Rotate it counterclockwise by 30 degrees, and then again by another 60 degrees. What is it's orientation now? How would you get there in one rotation instead? What does this suggest about the multiplication of rotation matrices?
3. What happens if you first rotate it by 60 degrees, and then by 30 degrees? What does this suggest about the commutative property of 2D rotation matrices?

Solution 4.4

1. Okay, I placed my book on the table.
2. You could get there by rotating once by 90 degrees. This suggests that the product of two rotation matrices of angles θ_1 and θ_2 is a rotation matrix of $\theta_1 + \theta_2$, i.e. $\mathbf{R}(\theta_1)\mathbf{R}(\theta_2) = \mathbf{R}(\theta_1 + \theta_2)$.

3. You end up in the same orientation so it doesn't matter the order. This suggests that the order of multiplication doesn't matter so that two rotation matrices must commute.

Exercise 4.5

This is an algebra problem involving the multiplication of 2D rotation matrices.

1. Use some algebra to show that 2D rotation matrices commute, i.e. $\mathbf{R}(\theta_1)\mathbf{R}(\theta_2) = \mathbf{R}(\theta_2)\mathbf{R}(\theta_1)$.
2. Use some algebra to show that $\mathbf{R}(\theta_1)\mathbf{R}(\theta_2) = \mathbf{R}(\theta_1 + \theta_2)$. You will need to look up some trig identities.

Solution 4.5

1. You could multiply out two rotation matrices with angle θ_1 and θ_2 in the two different orders and you will observe that the output is the same because real numbers commute, i.e. $\cos\theta_1 \cos\theta_2 = \cos\theta_2 \cos\theta_1$.
2. If you multiply two matrices together you will get the following expression in the first row and first column, $\cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2$. You will find a trig identity which reduces this to $\cos(\theta_1 + \theta_2)$. Similar reductions take place for the other elements.

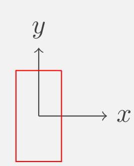
Exercise 4.6

Now, consider a rectangle of width 2 and height 4, centered at the origin. For clarity, this means that the corners of the rectangle have coordinates $(1, 2)$, $(-1, 2)$, $(-1, -2)$, and $(1, -2)$.

1. Plot these four points by hand and connect them with lines to complete the rectangle.
2. Now, using the appropriate rotation matrix, transform each of the corner points by a rotation through 30 degrees counterclockwise (recall that the sin and cos of 30 degrees can be expressed exactly). Compute and plot the resulting points by hand and connect them with lines. Does the resulting figure look like you'd expect?

Solution 4.6

1. The rectangle is



2. The rotation matrix is

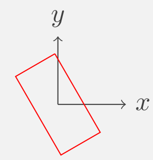
$$\mathbf{R} = \begin{bmatrix} \cos 30 & -\sin 30 \\ \sin 30 & \cos 30 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}.$$

Applying this to each point, we get

$$\mathbf{R}\begin{bmatrix}1\\2\end{bmatrix}=\begin{bmatrix}\frac{\sqrt{3}-2}{2}\\\frac{1+2\sqrt{3}}{2}\end{bmatrix},\ \mathbf{R}\begin{bmatrix}1\\-2\end{bmatrix}=\begin{bmatrix}\frac{\sqrt{3}+2}{2}\\\frac{1-2\sqrt{3}}{2}\end{bmatrix},$$

$$\mathbf{R}\begin{bmatrix}-1\\2\end{bmatrix}=\begin{bmatrix}\frac{-\sqrt{3}-2}{2}\\\frac{-1+\sqrt{3}}{2}\end{bmatrix},\ \mathbf{R}\begin{bmatrix}-1\\-2\end{bmatrix}=\begin{bmatrix}\frac{-\sqrt{3}+2}{2}\\\frac{-1-\sqrt{3}}{2}\end{bmatrix}.$$

And the rotated figure looks like,



Exercise 4.7

Now, let's do it in MATLAB.

1. Create and plot the original 4 points:$(1,2)$, $(-1,2)$, $(-1,-2)$, and $(1,-2)$. Then create the matrix that rotates them by 30 degrees counterclockwise, transform each of the four original points using the rotation matrix, and plot the resulting points. Does this look right? *Reminder:* `plot(1,2,'x')` *puts a mark at the point (1,2). Matlab: the functions cos and sin expect radians, while cosd and sind expect degrees.*

2. Operating on individual points with the rotation matrix is cool, but we can be much more efficient by operating on all 4 points at the same time. Write down the matrix whose columns represent the four corners of the rectangle. Then write down the matrix multiplication problem we can solve to transform the rectangle from above all at once. Create these matrices in MATLAB to perform the rotation in a single operation. Plot the resulting matrix to confirm your transformation! *Some MATLAB tips:* `plot(X,Y)` *creates a line plot of the values in the vector Y versus those in the vector X. So if you wanted to plot a line from the origin (0,0) to the point (1,2), you would do this: plot([0 1],[0 2]).* The command `axis([-xlim xlim -ylim ylim])` sets the axes of the current plot to run from `-xlim` to `xlim` and from `-ylim` to `ylim`

3. What is the area of the rectangle before and after the rotation?

4. What matrix should you use to undo this rotation? Define it in MATLAB and check.

5. Show on the board that the product of this matrix with the original rotation matrix is the identity matrix. For clarity, let's give this matrix the symbol $\mathbf{R}^{-1}$. It is the matrix that inverts the original operation and is known as the *inverse* of the matrix $\mathbf{R}$.

Solution 4.7

1. There are lots of ways to do this point by point. Here is an example of how to transform the bottom right point:

```
>> BR = [1;-2]
```

```
>> plot(BR(1,:),BR(2,:),'b*')
>> rotmatrix = [cosd(30) -sind(30); sind(30) cosd(30)]
>> nBR = rotmatrix*BR
>> plot(nBR(1,:),nBR(2,:),'r*')
```

2. There are lots of ways to do this. Here is an example where we include the first point twice so that the points can easily be connected with lines:

```
>> pts = [1 -1 -1 1 1;2 2 -2 -2 2]
>> npts = rotmatrix*pts
>> plot(pts(1,:),pts(2,:),'b'), hold on
>> plot(pts(1,:),pts(2,:),'r')
>> axis([-3 3 -3 3])
>> axis equal
```

3. The area of the rectangle is the same before and after rotation: 8 square units.

4. To undo this rotation you could simply rotate it by 30 degrees clockwise, using the matrix

$$\mathbf{R}^{-1} = \begin{bmatrix} \cos 30 & \sin 30 \\ -\sin 30 & \cos 30 \end{bmatrix}.$$

5. The product of $\mathbf{R}^{-1}$ and $\mathbf{R}$ is

$$\mathbf{R}^{-1}\mathbf{R} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \cos^2\theta + \sin^2\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

where we have used the trig identity $\cos^2\theta + \sin^2\theta = 1$.

Chapter 5

Week 2b: Matrix Transformations

Schedule

do we need some sort or reorientation?

5.1 3D Rotations [45 minutes]

We can extend the idea of 2D rotations to 3D rotations. The simplest approach is to think of 3D rotations as a composition of rotations about different axes. First let's define the rotation matrices for counterclockwise rotations of angle θ about the x, y and z axes respectively.

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \tag{5.1}$$

$$\mathbf{R}_y = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \tag{5.2}$$

$$\mathbf{R}_z = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.3}$$

For example, to first rotate a vector $\mathbf{v}$ counterclockwise by θ about the x axis followed by counterclockwise by ϕ about the z axis, you need to do the following

$$\begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \mathbf{v} \tag{5.4}$$

We will next look at some sequence of physical rotations and relate them to these rotation matrices.

Exercise 5.1

Hold a closed book in front of you, with the top of the book towards the ceiling ($+z = (0, 0, 1)$ direction) and the cover of the book pointed towards you ($+x = (1, 0, 0)$ direction), which leaves the opening side of the book pointing towards your right ($+y = (0, 1, 0)$) and the spine toward the left.

1. Rotate the book by 90 degrees counter-clockwise about the x-axis, then from this position, rotate the book by 90 degrees counter-clockwise about the z-axis. Which direction is the cover of the book facing now?

2. Return to the starting position. Now rotate the book by 90 degrees counter-clockwise about the z axis, and then from this position, rotate the book by 90 degrees counter-clockwise about the x axis. Which direction is the cover of the book facing now? Is it the same as in part a?

3. An operation "commutes" if changing the order of operation doesn't change the result. Do 3D rotations commute?

4. The cover of the book is originally pointed towards $(1, 0, 0)$. Multiply this vector with the appropriate sequence of rotation matrices from above to reproduce your motions from part 1. Do you end up with the correct final cover direction?

5. Multiply the $(1, 0, 0)$ vector with the appropriate sequence of rotation matrices to reproduce the motions from part 2. Do you end up with the correct final cover direction?

6. Multiply the result of the previous part by the appropriate sequence of rotation matrices to return to the original $(1, 0, 0)$ vector.

7. From either of your answers to part 4 or part 5, try, instead of operating on the $(1, 0, 0)$ vector sequentially with one rotation matrix and then the other, take the product of the two rotation matrices first, and then multiply $(1, 0, 0)$ with the resultant matrix. Does this reproduce your answer?

8. Based on your answers to the previous parts, show that $(\mathbf{R}_z\mathbf{R}_x)^{-1} = \mathbf{R}_x^{-1}\mathbf{R}_z^{-1}$. This is a general property of matrix inverses – it works for all square, invertible matrices, not just rotation matrices!

Solution 5.1

1. The cover is now facing toward the $+y$ axis (the positive part of the y axis).

2. The cover is now facing the $+z$ axis. This is different than in part a.

3. Since the answers for the first two parts are different, 3D rotations do not commute.

4. Let $\mathbf{v}$ be the vector that represents the initial direction of the cover of the book,

$$\mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

   Rotation by 90 degrees counterclockwise around the x axis is given by

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

   so that the new vector becomes

$$\mathbf{R}_x\mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Rotation by 90 degrees counterclockwise around the z axis is given by

$$\mathbf{R}_z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

so that the new vector becomes

$$\mathbf{R}_z \mathbf{R}_x \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

which is the correct final direction.

5. Using the matrices from above,

$$\mathbf{R}_x \mathbf{R}_z \mathbf{v} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

6. To rotate 90 degrees clockwise around the x axis we use the matrix

$$\mathbf{R}_x^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

and to rotate 90 degrees clockwise around the z axis we use the matrix

$$\mathbf{R}_z^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then we can return the vector $(0, 0, 1)$ to its original position $(1, 0, 0)$ by

$$\mathbf{R}_z^{-1} \mathbf{R}_x^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

7. We can multiply the rotation matrices together and perform a single matrix multiplication. For part d, the relevant matrix product is

$$\mathbf{R}_z \mathbf{R}_x = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

and we see that

$$\mathbf{R}_z \mathbf{R}_x \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

as expected.

8. We can see from the previous parts that

$$(\mathbf{R}_z \mathbf{R}_x)^{-1} = \mathbf{R}_x^{-1} \mathbf{R}_z^{-1}.$$

In other words, when you take the inverse, the order of operations must swap!

5.2 Reflection and Shearing [30 minutes]

In this activity we will meet reflection and shearing matrices, which will allow us to explore transformation matrices in general.

5.2.1 Reflection

Exercise 5.2

What do the following *reflection* matrices do? Think about it first, draw some sketches and then test your hypothesis in MATLAB using the rectangle with vertices $(0,0)$, $(2,0)$, $(2,1)$, and $(0,1)$. How much does the area of your basic rectangle change, if at all? What is the inverse of each?

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

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

3.
$$\begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$$

Solution 5.2

1. This matrix reflects everything over the y-axis. In the figure below, the original blue rectangle becomes the orange rectangle. The area of the rectangle stays the same.

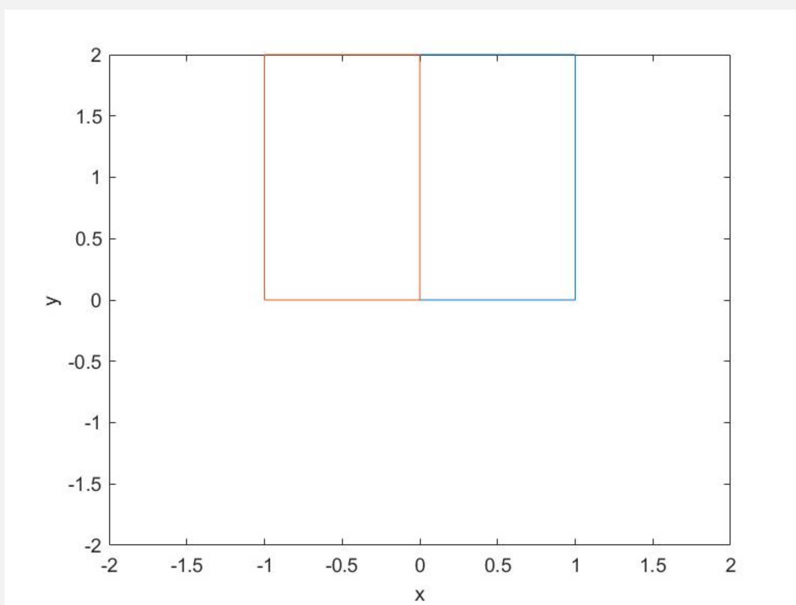


Figure 5.1: Reflection over y-axis.

2. This matrix reflects everything over the x-axis. In the figure below, the original blue rectangle

becomes the orange rectangle. The area of the rectangle stays the same.

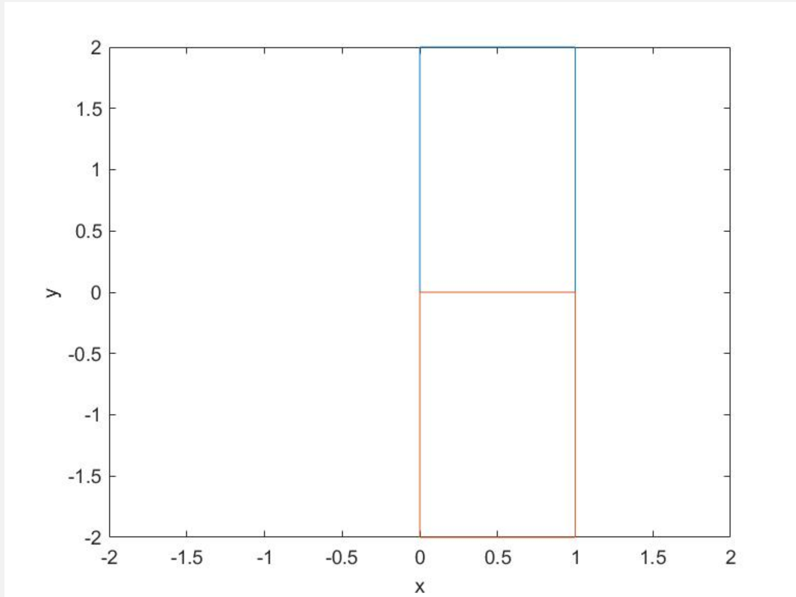


Figure 5.2: Reflection over x-axis.

3. For example, let $\theta = 30$ degrees. Then the rectangle is reflected along the line that is 30 degrees counterclockwise from the x-axis. In the figure below, the original blue rectangle becomes the orange rectangle.

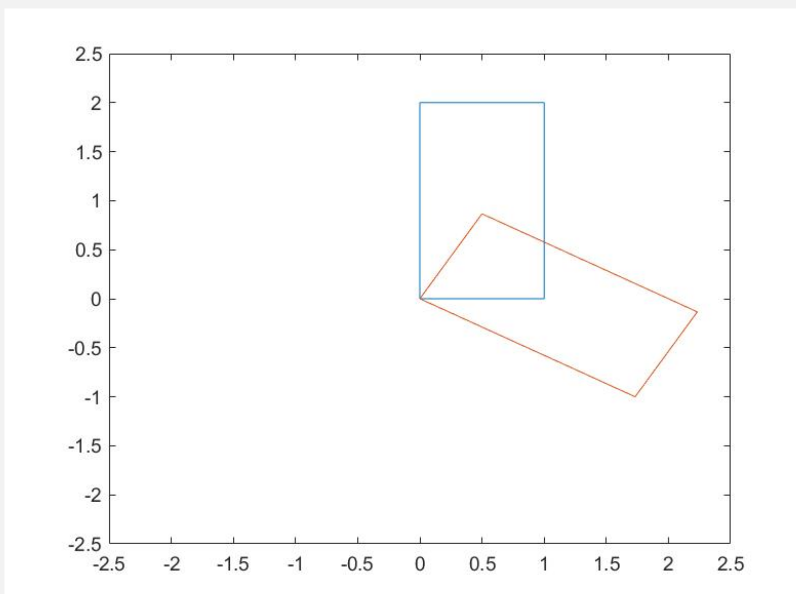


Figure 5.3: Reflection over 30 degree line.

Notice that, if we plug in $\theta = 90$, we get the matrix from part 1, which reflects over the x-axis (i.e., 90 degree line) and, if we plug in $\theta = 0$, we get the matrix from part 2, which reflects over the y-axis (i.e., the 0 degree line).

5.2.2 Shearing

Exercise 5.3

What do the following *shearing* matrices do? Think about it first, draw some sketches and then test your hypothesis in MATLAB with the rectangle with vertices $(0,0)$, $(2,0)$, $(2,1)$, and $(0,1)$. How much does the area of your basic rectangle change, if at all? What is the inverse of each?

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

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

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

4. $$\begin{bmatrix} 1 & 0 \\ 2k & 1 \end{bmatrix}$$

Solution 5.3

1. This shearing matrix pulls the points along horizontal lines and the strength of the pull is proportional to the y coordinate. In the figure below, the blue rectangle is sheared to become the orange rectangle:

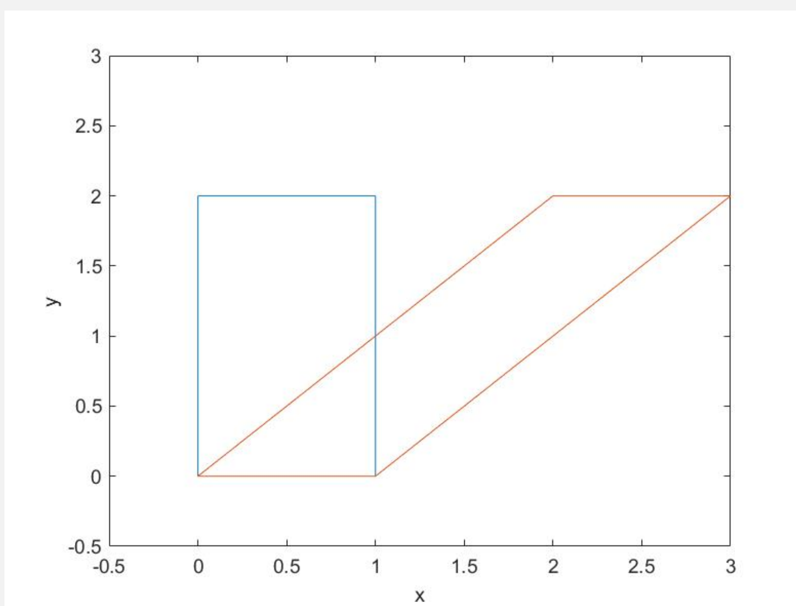


Figure 5.4: Shearing in x direction.

The area of the rectangle does not change. The inverse is

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

2. This shearing matrix pulls the points along vertical lines and the strength of the pull is proportional to the x coordinate. In the figure below, the blue rectangle is sheared to become the orange rectangle:

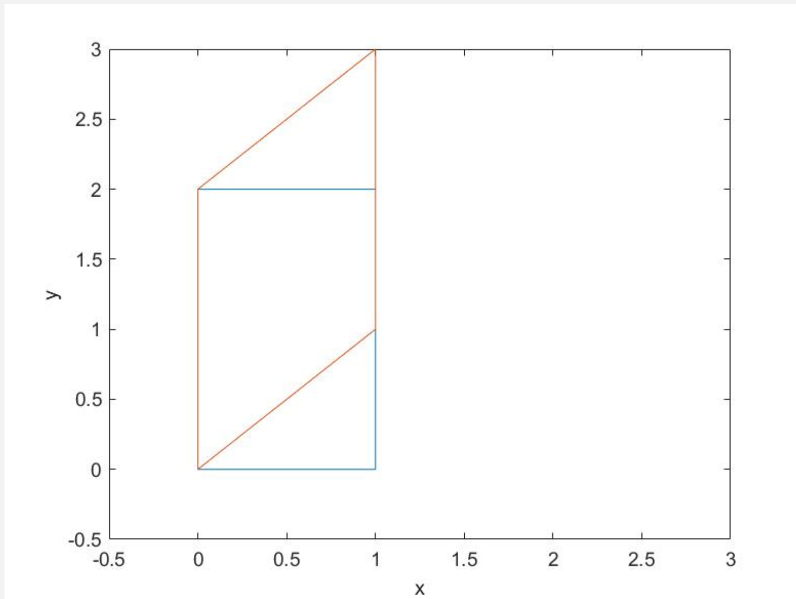


Figure 5.5: Shearing in y direction.

The area of the rectangle does not change. The inverse is

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

3. This shearing matrix pulls the points along horizontal lines and the strength of the pull is proportional to the y coordinate and the constant k (the bigger the k, the stronger the pull). In the figure below, with $k = 2$, the blue rectangle is sheared to become the orange rectangle:

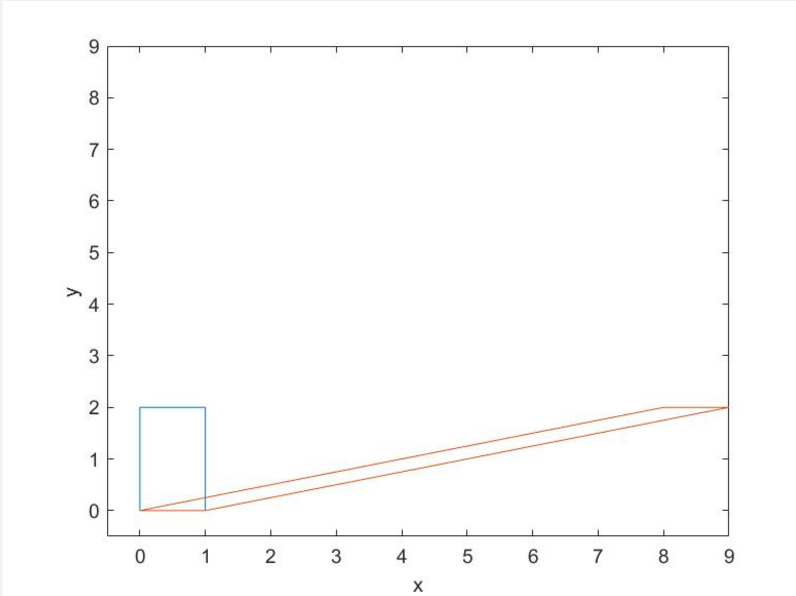


Figure 5.6: Shearing in x direction with $k = 2$.

The area of the rectangle does not change. The inverse is

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

4. This shearing matrix pulls the points along vertical lines and the strength of the pull is proportional to the x coordinate and the constant k (the bigger the k, the stronger the pull). In the figure below, with $k = 2$, the blue rectangle is sheared to become the orange rectangle:

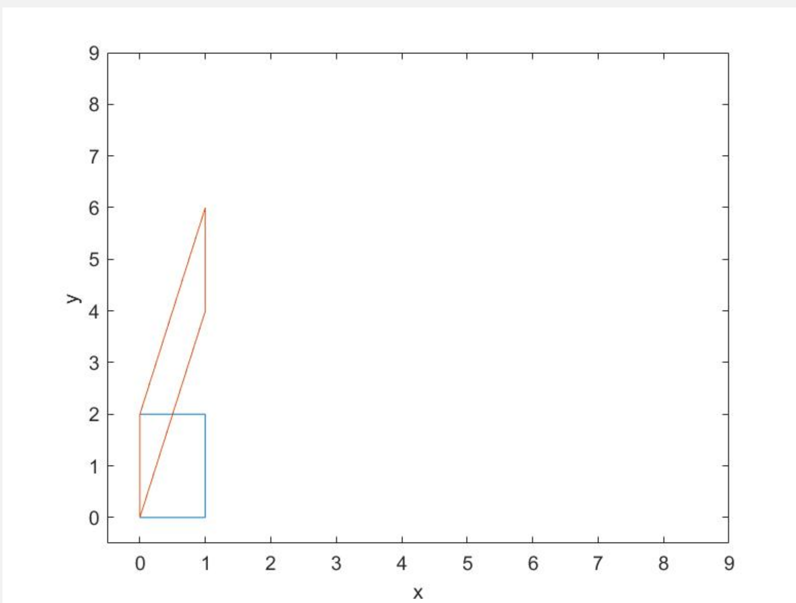


Figure 5.7: Shearing in y direction with $k = 2$.

The area of the rectangle does not change. The inverse is

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

5.3 Review and Preview [15 minutes]

Instructors' Notes: John's Review:

- Matrices are holders of data, e.g. coordinates of points
- Matrices are transformation operators, e.g. rotation, reflection, shearing.
- Matrices have algebraic properties:
  - $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
  - $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
  - $\mathbf{AB} \neq \mathbf{BA}$ (don't always commute)
- 2D Rotation Matrix (does commute)

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

- 3D Rotation Matrices (do not commute about different axes)
- Order of operations: $\mathbf{ABv}$ implies that $\mathbf{B}$ acts on $\mathbf{v}$, and then $\mathbf{A}$ acts on the result.
- Inverse Matrix: undoes a transformation
  - $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$
  - $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

Chapter 6

Night 2: Matrix Operations

Overview and Orientation

Learning Objectives

Concepts

- Compute the determinant of a 2×2 matrix
- Know the relationship between the determinant of a matrix and whether the matrix is invertible
- Find the inverse of a 2×2 matrix by hand
- Use computational tools to find the inverse of an $n \times n$ matrix
- Design a 2 or 3-dimensional matrix that will scale a vector by given amounts in the x, y or z direction
- Design a 3-dimensional matrix that will translate a 2-D vector by given amounts in x and y

MATLAB skills

- Represent a set of points in 2-D space (i.e., pairs of x, y values) as column vectors
- Transform a set of 2-D points (i.e., the outline of a shape) using a matrix to rotate and translate the original
- Multiply matrices and find their inverses
- Compute the determinant of a matrix

6.0.1 Suggested Approach

See Night 1 assignment for our general suggested approach to night assignments and a list of linear algebra resources.

6.1 Determinant of a Matrix

The determinant of a square matrix is a property of the matrix which indicates many important things, including whether a matrix is invertible or not. We will see more of this when we see matrix inverses shortly.

The determinant of a matrix $\mathbf{G}$ is denoted a few different ways.

$$\det(\mathbf{G}) = |\mathbf{G}| \tag{6.1}$$

Consider a generic 2×2 matrix $\mathbf{G}$:

$$\mathbf{G} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

The formula for the determinant of a 2×2 matrix is quite straightforward:

$$\det(\mathbf{G}) = ad - bc \tag{6.2}$$

For example, for the following 2×2 matrix,

$$\begin{aligned} \det\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\right) &= \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \\ &= (1)(4) - (2)(3) = -2 \end{aligned} \tag{6.3}$$

Exercise 6.1

Return to the transformation matrices in the day assignment and calculate the determinant for the following:

1. The generic 2×2 rotation matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

2. The matrix which reflects over the y axis **(we'll be diving into the reflection matrices in Week 2b, so if you haven't seen this yet, you can skip this problem or just turn the crank using the formula)**.

we'll be meeting this next class

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

3. The matrix which shears in the horizontal direction **(we'll be diving into shearing matrices in Week 2b, so if you haven't seen this yet, you can skip this problem or just turn the crank using the formula)**.

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

Solution 6.1

1. The determinant is 1. (Recall that $\cos^2\theta + \sin^2\theta = 1$.)
2. The determinant is -1.
3. The determinant is 1.

Exercise 6.2

1. What do the following matrices do? Think about it first, draw some sketches and then test your hypothesis in MATLAB. How much does the area of your basic rectangle change, if at all?

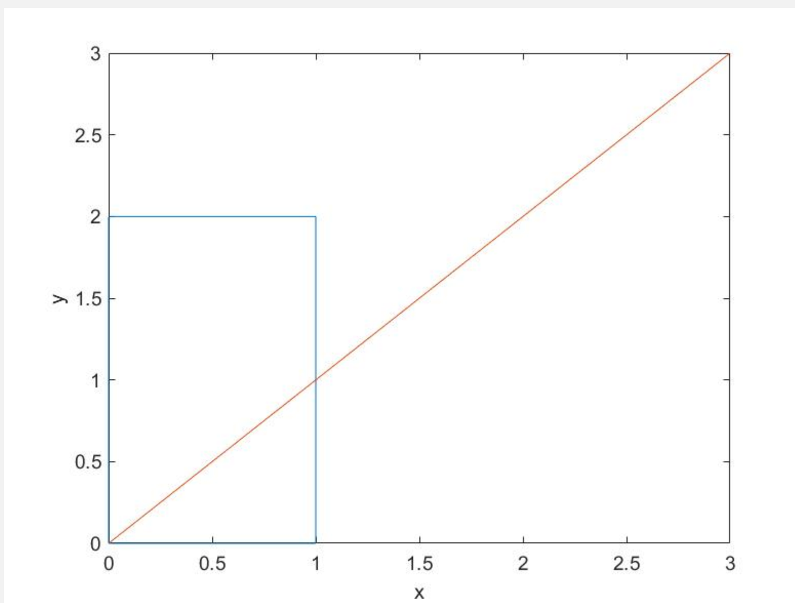
   (a)
$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

   (b)
$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

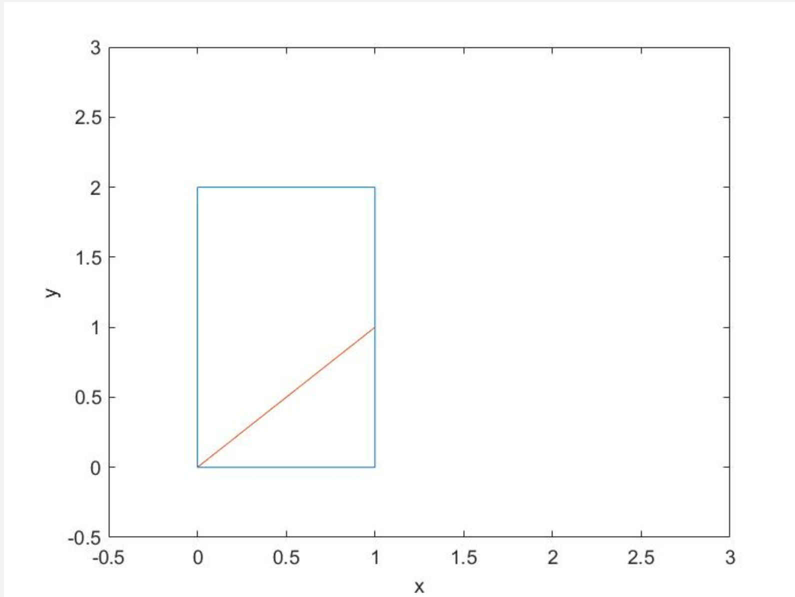
2. Is it possible to "undo" the matrices above? Why or why not?

Solution 6.2

1. Each of the figures below shows the basic blue rectangle and the orange rectangle, which is the result of applying the transformation.



(a)

(b) 

2. It is not possible to undo these matrix transformations. Since everything is squished onto the same line, we would not be able to distinguish the original vectors.

Notice that, in the above matrices, the first row is a constant multiple of the second row. In other words, the matrix looks like $\begin{bmatrix} a & b \\ ca & cb \end{bmatrix}$ for some constant c. If we apply a matrix of this form to a point in 2D space reprsented by the vector $\begin{bmatrix} x \\ y \end{bmatrix}$, then the result will be $\begin{bmatrix} z \\ cz \end{bmatrix}$, where $z = ax + by$. In other words, the resulting point will always fall on the line $y = cx$.

Exercise 6.3

1. What are the determinants of the two matrices from the previous exercise, Exercise 6.2?
2. Generalizing from Exercise 6.1 and Exercise 6.2, what's the relationship between the determinant of a matrix and the result of transforming a rectangle by that matrix?

Solution 6.3

I don't think we ever had this

Finding the determinant of an $n \times n$ matrix, where $n > 2$, is a bit more computationally intensive. If you want to learn how to do the procedure by hand, check out this Khan Academy video. For this course, we simply recommend you use the `det` function in MATLAB.

6.2 Matrix Inverses

6.2.1 Inverse of 2×2 Matrices

In class you worked with rotation matrices and transformations that were compositions of simpler rotations, and you learned how to invert them. When you multiply a vector by any matrix (not just ones that are associated with simple spatial transformations), you transform the original vector into a new vector. More generally (than rotations), you can *often* undo the linear transformation (just like you did with the rotation matrix). Undoing this linear transformation is a linear transformation itself! Therefore the act of undoing a linear transformation can be formulated with a matrix multiply.

Exercise 6.4

Consider the following matrices and vector. (Don't try to interpret these as intuitive geometrical operations; we're just using them to explore the determinant.) Work out the following problems in MATLAB.

$$\mathbf{P} = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix} \tag{6.4}$$

$$\mathbf{Q} = \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -2 & 1 \end{bmatrix} \tag{6.5}$$

$$\mathbf{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \tag{6.6}$$

1. Find $\mathbf{w} = \mathbf{P}\mathbf{u}$.
2. Find $\mathbf{Q}\mathbf{w}$. How is this related to $\mathbf{u}$?
3. Find $\mathbf{Q}\mathbf{P}$. Does the answer look familiar?
4. Find $\mathbf{P}\mathbf{Q}$.
5. Find the determinant of $\mathbf{P}$. In MATLAB, you can compute the determinant of any (not just 2×2) matrix using the `det` function.
6. Find the determinant of $\mathbf{Q}$.

Solution 6.4

1. $$\mathbf{w} = \mathbf{P}\mathbf{u} = \begin{bmatrix} 7 \\ 17 \end{bmatrix}$$
2. $$\mathbf{Q}\mathbf{w} = \mathbf{Q}\mathbf{P}\mathbf{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$
3. $$\mathbf{Q}\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

   which is the identity matrix
4. The determinate of $\mathbf{P}$ is 2.
5. The determinate of $\mathbf{Q}$ is $\frac{1}{2}$.

A matrix $\mathbf{B}$ is said to be the inverse of the matrix $\mathbf{A}$ if, and only if, $\mathbf{BA} = \mathbf{I}$ and $\mathbf{AB} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. For 2×2 matrices, the inverse (if it exists) is given by the following

$$\mathbf{G} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{6.7}$$

$$\mathbf{G}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{6.8}$$

The last equation should indicate to you that the inverse of the matrix $\mathbf{G}^{-1}$ is only defined if $ad - bc \neq 0$. Sweet mother of linear algebra, $ad - bc$ is our buddy the determinant. More generally, any square matrix can be inverted if and only if its determinant is non-zero.

Now let's practice calculating inverses, some of their properties, and how we may use them.

Exercise 6.5

All matrices $\mathbf{A}$ and $\mathbf{B}$ which have inverses have the following properties

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$
$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$$

1. Using the above properties, please compute the following by hand.

   (a) If

   $$\mathbf{P} = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix} \tag{6.9}$$

   $$\mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \tag{6.10}$$

   $$\tag{6.11}$$

   find $(\mathbf{PB})^{-1}$. Recall that you already know the inverse of $\mathbf{P}$ from earlier.

   (b) For $\mathbf{P}$ as defined above, find

   $$(\mathbf{P}^T)^{-1} \tag{6.12}$$

2. Use the inverse formula to calculate the inverses for the first three matrices in Exercise 6.1. Confirm your answers by multiplying the inverse with the original matrix.

Solution 6.5

1. (a)

   $$(\mathbf{PB})^{-1} = \begin{bmatrix} -17/2 & 7/2 \\ 5 & -2 \end{bmatrix}$$

   (b)

   $$(\mathbf{P}^T)^{-1} = \begin{bmatrix} 3/2 & -2 \\ -1/2 & 1 \end{bmatrix}$$

2.

$$\left(\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \right)^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

$$\left(\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \right)^{-1} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\left(\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}\right)^{-1} = \begin{bmatrix}1 & -1\\0 & 1\end{bmatrix}$$

We don't have solutions for the enumerated list

Note that solving matrix-vector equations like above can be done without explicitly computing the matrix inverse which is computationally expensive. (A nod to our future friend, left matrix divide or backslash divide.)

6.2.2 Inverse of $n \times n$ Matrices

For higher-dimensional matrices, e.g. $n \times n$ matrices for $n > 2$, the matrix inverse is defined in the same way. Suppose you have an $n \times n$ matrix $\mathbf{A}$ and an $n \times n$ matrix $\mathbf{B}$. Then $\mathbf{B}$ is the inverse of $\mathbf{A}$ if and only if $\mathbf{BA} = \mathbf{I}$ and $\mathbf{AB} = \mathbf{I}$. The following are some properties of inverses of matrices

- Only square matrices are invertible, i.e., only square matrices have inverses.
- A matrix has an inverse only if its determinant is non-zero.

There are a number of different procedures to compute the inverse of higher-dimensional matrices, but we will not be going into the details of their computation here. You can look them up if you are interested, or need to in the future. In MATLAB, you can compute the inverse of a matrix using the `inv` function.

Exercise 6.6

1. Consider the example with the fruits that you worked out earlier. Now, in addition to apples and oranges, suppose you also had an unknown number of pears which each weigh 3 oz, and cost $3. Additionally, suppose that the total weight of the fruits is 45 oz, and you paid a total of $21 for the fruit.
   (a) If possible find the numbers of oranges, apples and pears. If not, please explain why.
   (b) Suppose that you additionally know that you have a total of 14 fruits. Can you formulate and solve a matrix-vector equation to find out the numbers of oranges, apples and pears you have?
   (c) What is the determinant of the matrix you have set up to solve this?
2. The fruit vendors bought the pricing algorithm from Uber. Oranges are still $2, pears are now only $1.50, and (due to an influx of teachers) apples are now surging at $1.50 each. Their weights stay the same. You return to the market, and again purchase 14 fruits, which have the same total weight and total cost.
   (a) Can you formulate and solve a matrix-vector equation to find out the numbers of oranges, apples and pears you have?
   (b) What is the determinant of the matrix you have set up to solve this?

Solution 6.6

1. (a) It's not possible to find the numbers of oranges, apples, and pears. We have the equation

$$\begin{bmatrix} 2 & 1 & 3 \\ 4 & 3 & 3 \end{bmatrix} \begin{bmatrix} n_0 \\ n_a \\ n_p \end{bmatrix} = \begin{bmatrix} 21 \\ 45 \end{bmatrix},$$

but we cannot take the inverse of a 2×3 (non-square) matrix.

(b) Now we have the equation

$$\begin{bmatrix} 2 & 1 & 3 \\ 4 & 3 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} n_0 \\ n_a \\ n_p \end{bmatrix} = \begin{bmatrix} 21 \\ 45 \\ 14 \end{bmatrix}.$$

So by taking the inverse of the 3×3 matrix we find that $n_0 = 3, n_a = 9$ and $n_p = 2$.

(c) The determinant of the matrix is 2.

2. (a) The equation becomes

$$\begin{bmatrix} 2 & \frac{3}{2} & \frac{3}{2} \\ 4 & 3 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} n_0 \\ n_a \\ n_p \end{bmatrix} = \begin{bmatrix} 21 \\ 45 \\ 14 \end{bmatrix}.$$

But the matrix is not invertible, so we cannot solve for the number of fruit.

(b) The determinant of the matrix is 0.

6.3 Transformation Matrices, Continued

6.3.1 Scaling

Returning to two dimensions. In the Night 1 assignment, you also learned about scaling matrices. Recall that the scaling matrix $\mathbf{S}$ scales the x-component by s_1 and the y-component by s_2

$$\mathbf{S} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix}.$$

Let's assume for the moment that $s_1 = 2$ and $s_2 = 1/3$. Working with the rectangles defined in class whose corners have coordinates $(1, 2)$, $(1, -2)$, $(-1, 2)$, and $(-1, -2)$ complete the following activities:

Exercise 6.7

1. Predict what would happen if you operate on the rectangle with $\mathbf{S}$.
2. Write a MATLAB script to carry out this operation and check your prediction.
3. How does the area of the rectangle change?
4. What matrix should you use to *undo* this scaling? Show that the product of this matrix with the original scaling matrix is the *identity* matrix.
5. Define it in MATLAB and check. Again, this is the *inverse* matrix and we give it the symbol $\mathbf{S}^{-1}$.
6. In MATLAB, change the value of s_2 to 1 and find the product of the new $\mathbf{S}$ and your rectangle. How does the area of the rectangle change? Change the value of s_2 back to 1/3.

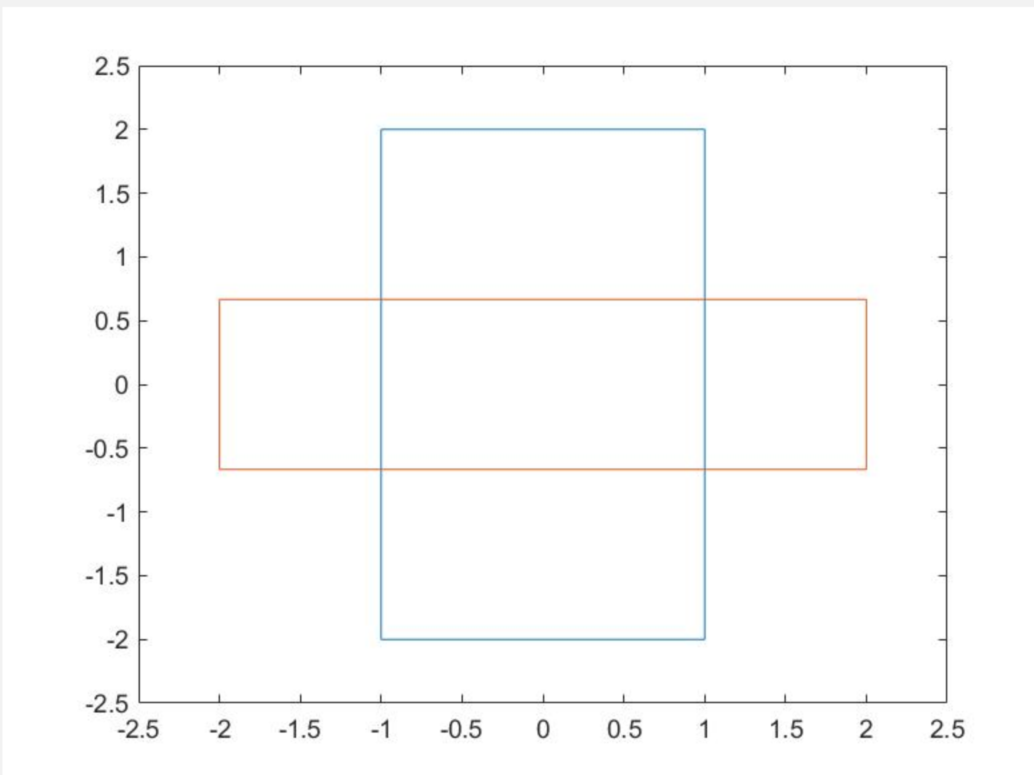
7. Predict what would happen if you operate on the original rectangle with $\mathbf{SR}$, where $\mathbf{R}$ is the rotation matrix. How about $\mathbf{RS}$? Implement both of these in MATLAB and check.

8. How would you *undo* each of these operations ($\mathbf{SR}$ and $\mathbf{RS}$)? How is the inverse of the product related to the individual inverses, i.e. what is the relationship between $(\mathbf{SR})^{-1}$ and $\mathbf{S}^{-1}$ and $\mathbf{R}^{-1}$? What about $(\mathbf{RS})^{-1}$?

Solution 6.7

1. The length of the rectangle would double in the x direction and be reduced to 1/3 the length in the y direction.

2. First we define the corners of the rectangle as the columns in a matrix
   » `points=[1 1 -1 -1; 2 -2 -2 2]`
   and we define the scaling matrix
   » `S=[2 0; 0 1/3]`. Then we simply multiply them
   » `scaledpoint=S*points`.
   Plotting them, here is the original rectangle in blue and the scaled rectangle in orange



3. The area is reduced from 8 units^2 to 5.33 units^2, or 2/3 of the original area.

4. To undo the process we use the inverse of the $\mathbf{S}$ matrix, or $\mathbf{S}^{-1}$ would be used.

$$\mathbf{S}^{-1} = \begin{bmatrix} 0.5 & 0 \\ 0 & 3 \end{bmatrix}.$$

   You should check that $\mathbf{S}^{-1}\mathbf{S} = \mathbf{S}\mathbf{S}^{-1} = \mathbf{I}$.

5. We define the inverse matrix » `Sinv=[0.5 0; 0 3]` and check that » `S*Sinv` and `Sinv*S` both produce the identity matrix.

6. The area of the rectangle doubles.

7. When the original rectangle is operated on with

$$\mathbf{SR}$$

, the resulting image will be a horizontally stretched parallelogram. When the original rectangle is operated on with $\mathbf{RS}$, the resulting image will be the scaled rectangle from the previous exercise only rotated 60 degrees counter-clockwise.

8. $(\mathbf{SR})^{-1} = \mathbf{R}^{-1}\mathbf{S}^{-1}$ or $(\mathbf{RS})^{-1} = \mathbf{S}^{-1}\mathbf{R}^{-1}$

6.3.2 Translation

It would be really useful if, in addition to scaling and rotating our objects, we could translate them. Let's start by thinking about vectors and then we will figure out how to represent translation as a matrix operation.

Consider an initial vector $\mathbf{v}$ and a translation vector $\mathbf{t}$. The new translated vector is simply $\mathbf{v}+\mathbf{t}$. For example, if you start with the initial vector $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ and translate it using the vector $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ then the new vector is just $\begin{bmatrix} x+2 \\ y+3 \end{bmatrix}$. More generally, if the translation vector is $\begin{bmatrix} t_x \\ t_y \end{bmatrix}$ then the new vector will be $\begin{bmatrix} x+t_x \\ y+t_y \end{bmatrix}$.

Wouldn't it be handy if we could define translation as a matrix operation? Yes, indeed it would be, we hear you say. Here is the standard method: add another entry to the original vector, and set it equal to 1, i.e., $\mathbf{v} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$. Now define the translation matrix as

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}.$$

Exercise 6.8

1. Show that $\mathbf{Tv}$ accomplishes the process of translation (if you ignore the third entry in the new vector). What is the final vector?

2. Predict what would happen if you operate on our old friend the rectangle with the translation matrix defined by $t_x = 2$ and $t_y = 3$.

3. Write a MATLAB script to carry out this operation and check your prediction. How has the area of your rectangle changed?

4. What matrix should you use to *undo* this translation? Show on paper that the product of this matrix with the original translation matrix is the *identity* matrix. Define it in MATLAB and check. Again, this is the *inverse* matrix and we give it the symbol $\mathbf{T}^{-1}$.

5. Choose a rotation matrix $\mathbf{R}$. Predict what would happen if you operate on the original rectangle with $\mathbf{TR}$. How about $\mathbf{RT}$? Implement both of these in MATLAB and check. How would you undo each of these operations? (You will first have to adjust your definition of $\mathbf{R}$ so that it is the correct size.)

6. Predict what would happen if you operate on the original rectangle with $\mathbf{STR}$. How about $\mathbf{TRS}$? How would you *undo* each of these operations? (You will first have to adjust your definition of $\mathbf{S}$ so that it is the correct size.)

7. How would you generalize translation to 3D?

Solution 6.8

1. $\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x + t_x \\ y + t_y \\ 1 \end{bmatrix}$

2. The rectangle would be moved 2 to the right and 3 up.

3. The area of the rectangle does not change.

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

5. If the original rectangle is operated on by $\mathbf{TR}$, the rectangle would first be rotated with respect to the origin and than translated. If the original rectangle is operated on by $\mathbf{TR}$, the rectangle would first be translated and then rotated. As rotation happens with respect to the origin, the 2 operations will not result in the same rectangle.

   To undo the operation $\mathbf{TR}$, the resulting figure should be operated on by $\mathbf{R}^{-1}\mathbf{T}^{-1}$. To undo the operation $\mathbf{RT}$, the resulting figure should be operated on by $\mathbf{T}^{-1}\mathbf{R}^{-1}$.

6. If the original rectangle is operated on with $\mathbf{STR}$, the resulting image will be of the rectangle rotated 60 degrees around the origin, translated 2 to the right and 3 up and then scaled by $\mathbf{S}$. If the original rectangle is operated on with $\mathbf{TRS}$, the resulting image will be the scaled rectangle rotated 60 degrees around the origin and then translated 2 to the right and 3 up.

   To undo $\mathbf{STR}$, the resulting figure should be operated on by $\mathbf{R}^{-1}\mathbf{T}^{-1}\mathbf{S}^{-1}$. To undo $\mathbf{TRS}$, the resulting figure should be operated on by $\mathbf{S}^{-1}\mathbf{R}^{-1}\mathbf{T}^{-1}$.

7.
$$\begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

6.3.3 Putting it all together: Dancing Animals

In this activity you will animate a circus act. (No real or imaginary animals will be injured in this performance.) Here is what we would like you to do:

Exercise 6.9

1. Decide on an animal.

2. Decide on a circus act that consists of a set of translations, rotations (think back to Day 2),

shearings, and/or scalings in some order. Storyboard this idea and imagine the resulting animation.

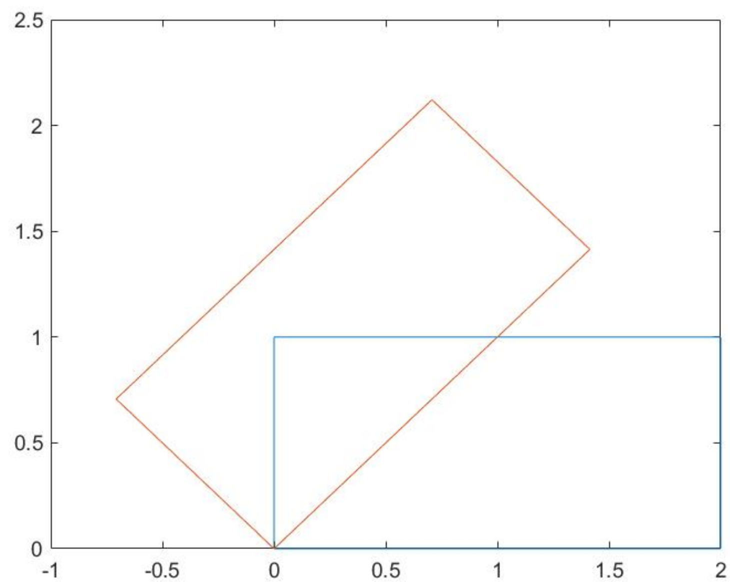
3. Propose a set of points that defines the outline and relevant features of your animal. You may find `ginput` useful. Define the points in MATLAB and plot your animal.

4. Create a script that makes your animal dance (in 2-D, unless you really want to go 3-D). You may want to make use of the `pause` and `drawnow` commands.

5. Now use your sequence of operations and animate your animal! In class you will have the opportunity to show off your dancing animal!

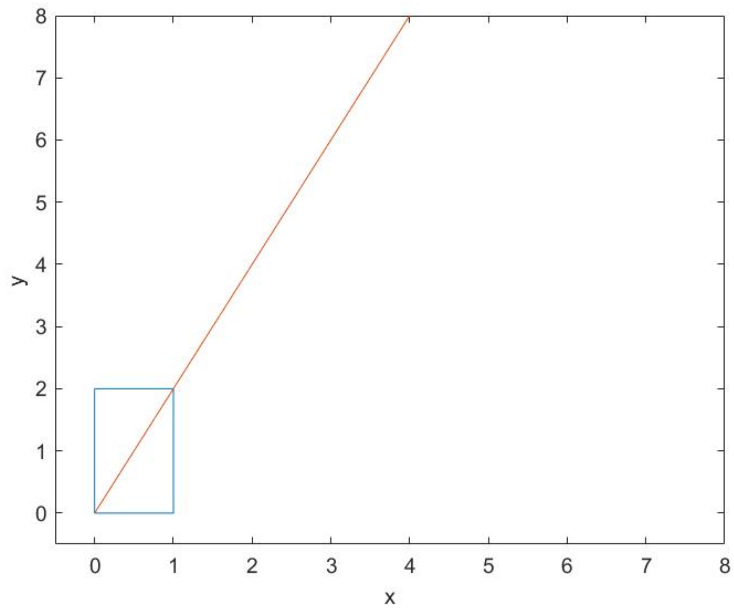
6.4 Conceptual Quiz (submit via Canvas)

1. The orange shape is the result of applying a matrix $\mathbf{M}$ to the blue rectangle.



What is the determinant of $\mathbf{M}$?

2. The orange shape is the result of applying a matrix $\mathbf{M}$ to the blue rectangle.



What is the determinant of $\mathbf{M}$?

3. The determinant is multiplicative, i.e, $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$. Let $\mathbf{M}$ be a matrix such that $\det(\mathbf{M}) = \frac{1}{3}$. What's $\det(\mathbf{M}^{-1})$? (Hint: $\det(\mathbf{I}) = 1$.)

4. Let R be a rectangle with area 1. Apply the scaling matrix $\mathbf{S} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix}$. What is the area of $\mathbf{S}R$?

   A. $\frac{s_1 s_2}{2}$
   B. 1
   C. $s_1 s_2$
   D. $s_1 + s_2$

5. True or false: Any shearing matrix $\mathbf{S}$ and any rotation matrix $\mathbf{R}$ commute, i.e., $\mathbf{RS} = \mathbf{SR}$.

Chapter 7

Week 3a: Linear Independence, Span, Basis, and Decomposition

Schedule

7.1 Debrief and Dancing Animal Demos [30 mins]

- Please discuss your overnight work with your breakout-room mates, create a set of key concepts, and a set of ideas that you are still confused by.
- Be prepared to demo your dancing animal to your breakout room.

7.2 Synthesis [20 mins]

Exercise 7.1

You should do all of these.

1. Assume the matrix $\mathbf{D}$ represents a geometrical object. What is the correct matrix expression if we want to rotate it first ($\mathbf{R}$), then scale it ($\mathbf{S}$), and finally translate ($\mathbf{T}$) it?

   A. $\mathbf{DRST}$
   B. $\mathbf{TSRD}$
   C. $\mathbf{RSTD}$
   D. $\mathbf{DTSR}$

2. What would be the correct expression in order to undo the transformation in the previous problem?

3. $\mathbf{A}$ and $\mathbf{B}$ are square, invertible matrices of the same size. Which of the following are **always** true (no matter the entries in $\mathbf{A}$ and B?

A. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$
B. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
C. $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$
D. $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$
E. $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$
F. $\mathbf{AB} = \mathbf{BA}$
G. $\det(\mathbf{AB}) = \det(\mathbf{A}) + \det(\mathbf{B})$
H. $(\mathbf{AB})^T = \mathbf{A}^T\mathbf{B}^T$
I. $(\mathbf{AB})^{-1} = \mathbf{A}^{-1}\mathbf{B}^{-1}$

7.3 Mini Lecture Linear Independence, Span, Basis [20 mins]

7.4 Linear Independence [20 mins]

A set of non-zero vectors is linearly independent if it is not possible to scale and sum them to make the all zeros vector, except when the scale factors are all zero.

If 3-dimensional vectors $\mathbf{x}_1, \mathbf{x}_2\mathbf{x}_3$ are linearly independent, it means that it is *not* possible to find scale factors c_1, c_2, c_3 so that

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + c_3\mathbf{x}_3 = \mathbf{0} \tag{7.1}$$

except when c_1, c_2, c_3 are all zero.

This property also implies that if you have n linearly independent, n-dimensional vectors, you can express any other n-dimensional vector by scaling and summing those linearly independent vectors.

If 3-dimensional vectors $\mathbf{x}_1, \mathbf{x}_2\mathbf{x}_3$ are linearly independent, it means that for any 3-dimensional vector $\mathbf{x}_d$, it is possible to find scale factors d_1, d_2, d_3 so that

$$d_1\mathbf{x}_1 + d_2\mathbf{x}_2 + d_3\mathbf{x}_3 = \mathbf{x}_d\,. \tag{7.2}$$

Exercise 7.2

1. Determine which of the following sets of vectors are linearly independent.

   (a) $\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\2\end{bmatrix}$

   (b) $\begin{bmatrix}1\\3\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}$

   (c) $\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}3\\4\\3\end{bmatrix}$

   (d) $\mathbf{p}$, $\mathbf{q}$, $\mathbf{r}$ and $\mathbf{s}$, where the vectors are all 3-dimensional.

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

Solution 7.2

1. (a) They are linearly independent since they span $\mathbf{R}^3$.
   (b) They are linearly dependent since the first vector is equal to the second vector plus two times the third vector.
   (c) They are linearly dependent since the third vector is equal to the first vector plus two times the second vector.
   (d) They are linearly dependent. You can have a maximum of n linearly independent vectors in $\mathbf{R}^n$.
   (e) They are linearly independent since they do not lie on the same line.

Chapter 8

Week 3b: Linear Independence, Span, Basis

In the last class, you were introduced to the idea of linear independence, span and basis. Today, we will dig deeper into these ideas. Let's review the activities from the end of last class.

8.1 Synthesis [20 mins]

Exercise 8.1

1. In words, describe what it means for a set of vectors to be linearly independent.

2. Determine which of the following sets of vectors are linearly independent.

   (a) $\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\2\end{bmatrix}$

   (b) $\begin{bmatrix}1\\3\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}$

   (c) $\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}3\\4\\3\end{bmatrix}$

   (d) $\mathbf{p}$, $\mathbf{q}$, $\mathbf{r}$ and $\mathbf{s}$, where the vectors are all 3-dimensional.

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

Solution 8.1

1. A set of vectors is linearly independent if it is not possible to scale and sum these vectors to result in the all-zeros vector, except if all the scale factors are zero.

2. (a) They are linearly independent since they span $\mathbf{R}^3$.

   (b) They are linearly dependent since the first vector is equal to the second vector plus two times the third vector.

(c) They are linearly dependent since the third vector is equal to the first vector plus two times the second vector.

(d) They are linearly dependent. You can have a maximum of n linearly independent vectors in $\mathbf{R}^n$.

(e) They are linearly independent since they do not lie on the same line.

8.2 Linear Independence, Span and Basis [70 mins]

8.2.1 Deeper Dive Into Linear Independence and Span [20 mins]

Exercise 8.2

Consider two column vectors

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{a}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \tag{8.1}$$

Both these vectors lie on the xy-plane since their z components are zero. Define a new vector $\mathbf{a}_3 = c_1\mathbf{a}_1 + c_2\mathbf{a}_2$, where c_1 and c_2 are arbitrary variables. Therefore $\mathbf{a}_3$ is a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$.

1. Does $\mathbf{a}_3$ also lie on the xy-plane?
2. Next, define a 3×3 matrix $\mathbf{A}$ whose columns are $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$. Show that the product of $\mathbf{A}$ and any 3×1 vector always lies on the xy-plane.

Solution 8.2

1. Yes, a linear combination of two vectors which lie in the xy-plane will also lie in the xy-plane.
2. Let $\mathbf{A}$ be the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & c_1 + c_2 \\ 1 & 2 & c_1 + 2c_2 \\ 0 & 0 & 0 \end{bmatrix}$$

and let $\mathbf{v}$ be an arbitrary 3×1 vector

$$\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Then the product

$$\mathbf{A}\mathbf{v} = \begin{bmatrix} x + y + (c_1 + c_2)z \\ x + 2y + (c_1 + 2c_2)z \\ 0 \end{bmatrix}$$

lies in the xy-plane

Exercise 8.3

Next, we will do a similar problem, but in MATLAB. Consider the following matrix:

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & 4 \\ 1 & 1 & 3 \end{bmatrix} \tag{8.2}$$

The third column of this matrix equals the second column plus twice the first column. Hence these three vectors lie on some plane (not the xy-plane as in the previous part).

1. Open up MATLAB and using the `quiver3` command together with `hold on`, please plot the vectors corresponding to the three columns of $\mathbf{B}$. Note that typing » `quiver3(0,0,0,1,1,1);` in MATLAB, plots an arrow from the origin to the point (1,1,1), i.e. it plots the vector corresponding to the first column of $\mathbf{B}$. Typing » `hold on` in MATLAB results in subsequent calls to quiver appearing on the same axes, without erasing previous arrows.
2. Using the "rotate 3D" function on the MATLAB figure window, rotate the figure around so that it appears as if all three arrows overlap. This should indicate that the vectors lie on a plane.
3. Using `det` compute the determinant of matrix $\mathbf{B}$. Does this make sense?

Solution 8.3

1. Type the following into MATLAB:
```
» quiver3(0,0,0,1,1,1)
» hold on
» quiver3(0,0,0,1,2,1)
» quiver3(0,0,0,3,4,3)
```
2.
3. The determinant of $\mathbf{B}$ is zero. Recall that a matrix is not invertible if and only if the determinant is zero. This matrix is not invertible since it collapses all vectors to a plane.

The fundamental property here is that the columns of the $\mathbf{A}$ and $\mathbf{B}$ matrices are not *linearly independent*. We shall next define the idea of linearly independent vectors more formally.

- A finite set $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ of vectors in $\mathbf{R}^n$ is said to be *linearly dependent* if there exist scalars $c_1, c_2, \ldots, c_m$ which are not all zero, such that

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_m\mathbf{x}_m = \mathbf{0}.$$

Note that $\mathbf{R}^n$ here refers to the set of all n-dimensional vectors that are made up of real numbers. (For example, $\mathbf{R}^1$ is the real line and $\mathbf{R}^2$ is the plane.) For any value of n, $\mathbf{R}^n$ is an example of a *vector space* - we will meet different examples of vector spaces in the future. We can also express this equation using a matrix $\mathbf{A}$, whose columns are $\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_m$.

$$\begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_m \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = \mathbf{0}\,. \tag{8.3}$$

If a non-zero solution exists to $\mathbf{Ac} = \mathbf{0}$ then the set of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ is linearly dependent. In the case of a square matrix ($n = m$), the vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_m$ are linearly dependent if and only if the $\det(\mathbf{A}) = 0$. Otherwise, the only way to satisfy the equation above is if $c_1 = c_2 = \cdots = c_m = 0$. Figure 8.1 illustrates two examples of three vectors that are in 3D space, but are linearly dependent, since in each case, all three vectors are on a plane.

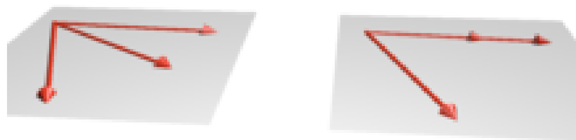


Figure 8.1: Linearly dependent vectors in $\mathbf{R}^3$. (from Wikimedia Commons).

- The set of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ is *linearly independent* if it is not linearly dependent. In other words, the set of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ is linearly independent if

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_m\mathbf{x}_m = \mathbf{0} \tag{8.4}$$

only when $c_1 = c_2 = \cdots = c_m = 0$. In other words, if the only solution to $\mathbf{Ac} = \mathbf{0}$ is $\mathbf{c} = \mathbf{0}$, the set of vectors made up of the columns of $\mathbf{A}$ is linearly independent. For a square matrix this means the set is linearly independent if and only if $\det(\mathbf{A}) \neq 0$.

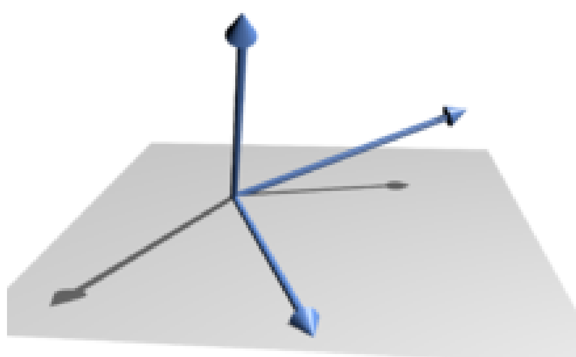


Figure 8.2: Linearly independent vectors in $\mathbf{R}^3$. (from Wikimedia Commons).

- The *span* of S is the set of all linear combinations of its vectors. In other words, the span of the set S is the set of all possible vectors of the form

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_m\mathbf{x}_m$$

The *span* is usually denoted by $span(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$.

- A finite set $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ of vectors is said to form a basis of a vector space V, if the vectors in S are linearly independent, and every point in V can be expressed as a linear combination of the vectors in the set S. Hence, if a set of vectors S is linearly independent those vectors form a *basis* of the set which is the span of those vectors.

Let's solidify our understanding of linear dependence, bases and span by working on a few problems by hand.

Exercise 8.4

1. In words, describe the span of the vectors $\begin{bmatrix}1\\1\end{bmatrix}$ and $\begin{bmatrix}1\\-1\end{bmatrix}$.
2. In words, describe the span of the vectors $\begin{bmatrix}1\\1\\0\end{bmatrix}$, $\begin{bmatrix}2\\3\\0\end{bmatrix}$ and $\begin{bmatrix}1\\-1\\0\end{bmatrix}$ which are all in 3-dimensional Euclidean space.

Solution 8.4

1. The span of these two vectors is all over $\mathbf{R}^2$, i.e., a plane.
2. The span of these three vectors is the xy-plane in $\mathbf{R}^3$.

8.2.2 Orthogonality [30 mins]

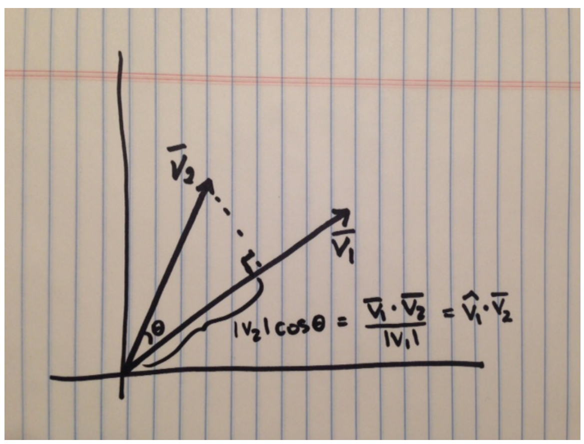


Figure 8.3: Projection

By trigonometry, if we have two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ which have an angle of θ between them, the component of $\mathbf{v}_2$ which lies along the direction of $\mathbf{v}_1$ is $|\mathbf{v}_2|\cos\theta$. Since the dot product of the two vectors can be expressed as $|\mathbf{v}_1||\mathbf{v}_2|\cos\theta$, this component (referred to as the projection) can be written as $\mathbf{v}_1 \cdot \mathbf{v}_2/|\mathbf{v}_1|$. If the projection is zero, the vectors are *orthogonal*, and $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. If the vectors are unit length, in addition to being normal, the vectors are said to be *orthonormal*. Additionally, if a basis set is made up of orthonormal vectors, it is known as an orthonormal basis.

A square matrix with columns of unit vectors which are orthogonal to each other is known as an orthogonal matrix. An orthogonal marix $\mathbf{A}$ has the property that $\mathbf{A}^T = \mathbf{A}^{-1}$.

Exercise 8.5

Which of the following pairs of vectors are orthogonal or orthonormal?

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

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

3. $\begin{bmatrix}\frac{2}{\sqrt{13}}\\\frac{-2}{\sqrt{13}}\end{bmatrix}, \begin{bmatrix}\frac{-3}{\sqrt{13}}\\\frac{3}{\sqrt{13}}\end{bmatrix}$

Solution 8.5

1. The dot product of these two vectors is non-zero, so they are not orthogonal.

2. The dot product of these two vectors is zero, so they are orthogonal.

3. The dot product of these two vectors is zero, so they are orthogonal. Furthermore, each vector is unit length, so they are orthonormal.

8.2.3 Decomposition [20 mins]

Suppose we have a set (collection) of m basis vectors $\{\mathbf{v}_i\}$ which are normalized ($|\mathbf{v}_i| = 1$), mutually orthogonal ($\mathbf{v}_i^T \mathbf{v}_j = 0$ unless $i = j$) and span our space (every point can be written as some linear combination of the vectors $\{\mathbf{v}_i\}$). How do we actually find the linear combination which is equal to a given vector in our space?

Let's say we have a vector $\mathbf{w}$ which we are interested in expressing as a linear combination of our set of orthonormal vectors $\{\mathbf{v}_i\}$. We can write this linear combination as

$$\mathbf{w} = \sum_{i=1}^{m} c_i \mathbf{v}_i \tag{8.5}$$

and our problem is now to find the coefficients c_i in this expression.

The obvious option is to pack the basis vectors $\mathbf{v}_i$ into the columns of a matrix $\mathbf{A}$, and find solutions of

$$\mathbf{A}\mathbf{c} = \mathbf{w}$$

Since the columns of $\mathbf{A}$ are formed from basis vectors they are linearly independent and a non-zero solution exists and can be determined by the usual methods.

However, our basis vectors form an orthogonal set (collection) which permits a more direct calculation. Consider a particular vector $\mathbf{v}_k$ in our basis set, and let's take the dot product between $\mathbf{v}_k$ and our vector $\mathbf{w}$:

$$\mathbf{v}_k^T \mathbf{w} = \mathbf{v}_k^T \sum_{i=1}^{m} c_i \mathbf{v}_i \tag{8.6}$$

Distributing the dot product into the summation we have:

$$\mathbf{v}_k^T \mathbf{w} = \sum_{i=1}^{m} c_i \mathbf{v}_k^T \mathbf{v}_i \tag{8.7}$$

But from orthogonality we know that the dot product of any two different vectors in our orthonormal set is zero, so all terms in the sum where $k \neq i$ are zero. This leads to the following simplification

$$\mathbf{v}_k^T \mathbf{w} = c_k \mathbf{v}_k^T \mathbf{v}_k \tag{8.8}$$

In addition, since our set of vectors is normalized, we know that $\mathbf{v}_k^T\mathbf{v}_k = 1$, leaving us with

$$\mathbf{v}_k^T\mathbf{w} = c_k \tag{8.9}$$

This gives us a very nice, simple way of decomposing a vector into a linear combination of the vectors within our basis set. The dot product of each basis vector with our target vector will result in the coefficient of that term in the linear decomposition.

Exercise 8.6

1. There are many (in general, an infinite number) of bases for a given set V. Hence, we can describe elements in the set V as linear combinations of vectors from different bases. Consider the following two basis sets which form bases for 2-dimensional space.

   - $\mathbf{v}_1 = \begin{bmatrix}1\\0\end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix}0\\1\end{bmatrix}$ and
   - $\mathbf{u}_1 = \begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix}-\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix}$

   Express the vector $\mathbf{w} = \begin{bmatrix}2\\3\end{bmatrix}$ as a linear combination of the first basis set (i.e., a sum of scaled versions of each vector in the basis set). Repeat for the second. Please make two different drawings of $\begin{bmatrix}2\\3\end{bmatrix}$, one expressed as a sum of scaled vectors in the first basis set and another for the vectors from the second basis set. Please label the lengths of each vector in the set.

2. Suppose that you wish to write the vector $\mathbf{w} = \begin{bmatrix}1\\2\\4\end{bmatrix}$ as a linear combination of the vectors

   $$\mathbf{v}_1 = \begin{bmatrix}1\\1\\1\end{bmatrix},\ \mathbf{v}_2 = \begin{bmatrix}3\\1\\2\end{bmatrix} \text{ and } \mathbf{v}_3 = \begin{bmatrix}1\\2\\2\end{bmatrix}.$$

   Please write a matrix equation to find the coefficients of the linear combination, and solve for the coefficients using MATLAB if possible.

3. Representing vectors using different bases is a very powerful technique that we will keep coming back to in this class (in both semesters). Vectors described in different bases can give us insight that may not be so obvious when viewed in the original basis. Representing vectors in different bases can also be used for dimensionality reduction, which is an important technique that is used to speed up computations and compress data in a number of different fields. Here we will consider a problem of lossy data compression using a change of basis. Lossy compression refers to methods of representing data more efficiently, but with a loss of accuracy. Examples of lossy data compression include jpg images, and mp3 audio files. If care is taken in lossy compression, the effects of the data loss can be kept at acceptable levels (this is of course subjective and dependent on the application). We will start with a toy example and then move to more complicated ones in subsequent homework problems. Consider a set of four 2-dimensional data variables stored in the following vectors:

   $$\mathbf{d}_1 = \begin{bmatrix}2.2\\1.2\end{bmatrix}, \mathbf{d}_2 = \begin{bmatrix}1\\0.6\end{bmatrix}, \mathbf{d}_3 = \begin{bmatrix}1.5\\0.7\end{bmatrix}, \mathbf{d}_4 = \begin{bmatrix}1.7\\0.8\end{bmatrix} \tag{8.10}$$

   (a) In MATLAB, plot the data using points (without lines connecting them) by typing `plot([2.2 1 1.5 1.7],[1.2 0.6 0.7 0.8], 'o');` You will find that these points lie close to the line through the origin with slope 1/2.

(b) Define a unit vector that points in the direction $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and call it $\mathbf{u}_1$. Find another unit vector that is orthogonal to $\mathbf{u}_1$ and call it $\mathbf{u}_2$. These vectors form a basis in 2 dimensional space.

(c) Rather than storing the original data, we are now going to express the original data in terms of the new basis that we have defined. To do that, write $\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3$ and $\mathbf{d}_4$, as a linear combination of $\mathbf{u}_1$ and $\mathbf{u}_2$. You can use MATLAB here to find the coefficients.

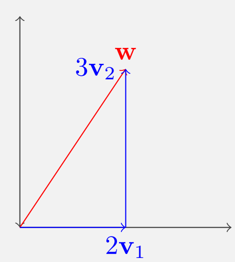
(d) In this toy example, we are going to "compress" our data by only keeping the coefficients corresponding to $\mathbf{u}_1$. i.e. we will discard the coefficient corresponding to $\mathbf{u}_2$. Suppose that we wish to recover approximations to $\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \mathbf{d}_4$, from the four coefficients. These approximations, which you should denote by $\tilde{\mathbf{d}}_1, \cdots \tilde{\mathbf{d}}_4$, are all scaled versions of $\mathbf{u}_1$. In your axes from part a, please plot the points corresponding to $\tilde{\mathbf{d}}_1, \cdots \tilde{\mathbf{d}}_4$. Do you think they make good approximations?

(e) We can describe how well our compressed data represents our original data. One way to do this is to calculate the difference between our original and compressed data, and call this error vector $\mathbf{f_i} = \mathbf{d}_i - \tilde{\mathbf{d}}_i$. Now, compute the size of this error using $norm(\mathbf{f_i})$ for $i = 1, 2, 3, 4$. Then, summarize the error by finding the root-mean-square (RMS) error between your approximations and the true data points. The RMS function squares the errors, takes the mean, and then takes the square root. This quantity is a single number that can be used to measure how well or poorly your compressed data represents your original data. You may find MATLAB's `norm` and `rms` functions helpful here.

This toy example illustrates that we can sometime be more efficient (albeit at the cost of some accuracy) in representing (or computing) data when it is expressed in certain bases.

Solution 8.6

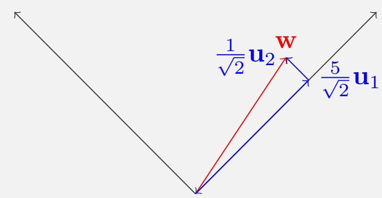
1. It's clear that $2\mathbf{v}_1 + 3\mathbf{v}_2 = \mathbf{w}$. We visualize this as



To write $\mathbf{w}$ as a linear combination of the basis vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ requires a bit more work. We can set up the matrix equation

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

and solve to learn that $\frac{5}{\sqrt{2}}\mathbf{u}_1 + \frac{1}{\sqrt{2}}\mathbf{u}_2 = \mathbf{w}$. We can visualize this as



2. First, we create a matrix in MATLAB whose columns are the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$,
   » `V=[1 3 1; 1 1 2; 1 2 2]`
   and the vector $\mathbf{w}$,
   » `w=[1; 2; 4]`.
   Let $\mathbf{c}$ be the vector of coefficients. We have the equation $\mathbf{V}\mathbf{c} = \mathbf{w}$, so to solve for $\mathbf{c}$ we compute $\mathbf{c} = \mathbf{V}^{-1}\mathbf{w}$. In MATLAB, we use » `inv(V)*w`. This tells us that $\mathbf{w} = -10\mathbf{v}_1 + 2\mathbf{v}_2 + 5\mathbf{v}_3$.

3. (a)

   (b) We define » `u1=[2; 1]` and » `u2=[-1; 2]`. There are other choices for $\mathbf{u}_2$, but they are all constant multiples of this choice, e.g., » `u2=[-2;4]`.

   (c) Create a 2×2 matrix with $\mathbf{u}_1$ and $\mathbf{u}_2$ as the columns,
   » `U=[2 -1; 1 2]`
   and a 2×4 matrix the vectors $\mathbf{d}_i$ as the columns
   » `D=[2.2 1 1.5 1.7; 1.2 0.6 0.7 0.8]`.
   Then compute
   » `inv(U)*D`
   to get the matrix of coefficients. This tells us that

$$\mathbf{d}_1 = 1.12\mathbf{u}_1 + 0.04\mathbf{u}_2,\ \mathbf{d}_2 = 0.52\mathbf{u}_1 + 0.04\mathbf{u}_2,$$

$$\mathbf{d}_3 = 0.74\mathbf{u}_1 - 0.02\mathbf{u}_2,\ \text{and } \mathbf{d}_1 = 0.84\mathbf{u}_1 - 0.02\mathbf{u}_2.$$

Chapter 9

Night 3: Independence, Span, Bases, and Linear Systems of Algebraic Equations

Learning Objectives

Concepts

- Determine for a system of 3 or fewer unknowns whether it has a unique solution, no solution or infinite solutions.
- Create a set of linear equations from a narrative about how the unknown variables are related to given data.
- Represent a system of linear equations with matrix, vector notation
- Solve a linear system of equations

MATLAB skills

- Compute the determinant of a matrix
- Solve systems of linear equations of the form $\mathbf{A}\mathbf{x} = \mathbf{b}$ using all three methods: inverse matrix, linsolve, or backslash operator.

9.0.1 Suggested Approach

See Night 1 for suggested approaches to the assignment and list of resources.

9.1 Linear Independence and Bases

Exercise 9.1

1. Is this set of vectors linearly independent?

$$\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}4\\5\\3\end{bmatrix} \tag{9.1}$$

2. Please express the vector $\begin{bmatrix}3\\6\\4\end{bmatrix}$ as a linear combination of the vectors $\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}$

3. Show that the following vectors form an orthogonal bases for $\mathbf{R}^4$

$$\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}, \begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\-\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}, \begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}, \begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\\frac{1}{2}\\-\frac{1}{2}\end{bmatrix} \tag{9.2}$$

4. Suppose that

$$\begin{bmatrix}5\\-2\\0\\1\end{bmatrix} = c_1 \begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}, c_2 \begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\-\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}, c_3 \begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}, c_4 \begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\\frac{1}{2}\\-\frac{1}{2}\end{bmatrix} \tag{9.3}$$

   Please find c_1, c_2, c_3 and c_4. As an aside, these vectors form what is called a Walsh code (which can be expanded to higher dimensions). Walsh codes are used in wireless communications so signals from multiple users can be added together (e.g. at the antenna of a cell tower), and then separated by different mobiles.

5. Construct a 4×4 matrix $\mathbf{A}$ whose columns are the vectors from the previous part. Show that c_1, c_2, c_3, c_4 can be found by solving

$$\mathbf{A} \begin{bmatrix}c_1\\c_2\\c_3\\c_4\end{bmatrix} = \begin{bmatrix}5\\-2\\0\\1\end{bmatrix} \tag{9.4}$$

Solution 9.1

1. No, because

$$3\begin{bmatrix}1\\1\\1\end{bmatrix} + \begin{bmatrix}1\\2\\0\end{bmatrix} - \begin{bmatrix}4\\5\\3\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix} \tag{9.5}$$

2.

$$\begin{bmatrix}3\\6\\4\end{bmatrix} = 2\begin{bmatrix}1\\1\\0\end{bmatrix} + 3\begin{bmatrix}0\\1\\1\end{bmatrix} + \begin{bmatrix}1\\1\\1\end{bmatrix} \tag{9.6}$$

3.

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

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

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

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

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

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

Note that when you expand out eacn product of the row and column vectors above, you get a sum of two $\frac{1}{4}$ terms and two $-\frac{1}{4}$ terms, resulting in zero. Therefore, the vectors are all orthogonal to each other.

4. We have already proven that the vectors in this question are mutually orthogonal in the previous part. Additionally, note that they are all of unit length, and they span $\mathbf{R}^4$, because they are 4 mutually orthogonal, 4-dimensional vectors. To find c_1, c_2, c_3, c_4, we can do the following Suppose that

$$c_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 5 \\ -2 \\ 0 \\ 1 \end{bmatrix} = 1$$

$$, c_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 5 \\ -2 \\ 0 \\ 1 \end{bmatrix} = 2$$

$$c_3 = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 5 \\ -2 \\ 0 \\ 1 \end{bmatrix} = 3$$

$$c_4 = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 5 \\ -2 \\ 0 \\ 1 \end{bmatrix} = 4 \tag{9.7}$$

9.2 Determinants and Invertibility

You have already encountered the determinant in class: the determinant of a square matrix is a property of the matrix which among other things indicates whether a matrix is invertible or not: if the determinant of a square matrix is zero, it is non-invertible. As a reminder:

The determinant of a matrix $\mathbf{G}$ is denoted a few different ways.

$$\det(\mathbf{G}) = |\mathbf{G}| \tag{9.8}$$

For a generic 2×2 matrix $\mathbf{G}$

$$\mathbf{G} = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

the formula for the determinant is quite straightforward:

$$\det(\mathbf{G}) = ad - bc \tag{9.9}$$

For example, for the following 2×2 matrix,

$$\begin{aligned} \det\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\right) &= \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \\ &= (1)(4) - (2)(3) = -2 \end{aligned} \tag{9.10}$$

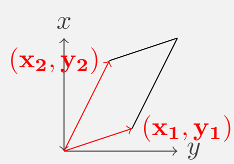
You already considered the determinant of some transformation matrices, now let's consider what the determinant is really telling us about a general matrix.

Exercise 9.2

1. Let $\mathbf{A}$ be a 2×2 matrix

$$\mathbf{A} = \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix}.$$

We can think of the columns of $\mathbf{A}$ as two vectors beginning at the origin and ending at the points (x_1, y_1) and (x_2, y_2), respectively. These vectors form a parallelogram, as shown here:

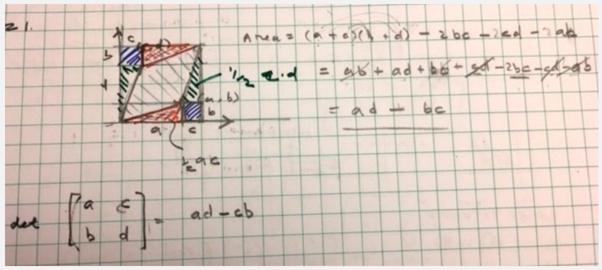


Show that the magnitude (i.e., absolute value) of $\det(\mathbf{A})$ is equal to the area of a parallelogram formed by the column vectors of the matrix $\mathbf{A}$.

2. What is the determinant of $\mathbf{A}$ if its column vectors are on the same line? Graphically, what happens to the parallelogram?

Solution 9.2

1.

2. 

3. The determinant is equal to 0, or det($\mathbf{A}$)=0.

From this, you should get the feeling for the fact that the determinant is a measure of how co-linear the columns of $\mathbf{A}$ are: or in other words, how linearly independent the two columns are. The determinant therefore lets us know quickly if a linear system of algebraic equations has a solution, as illustrated in the following example.

Exercise 9.3

Consider the following matrix whose columns lie on the same line: the second column is simply twice the first column.

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \tag{9.11}$$

1. What is det($\mathbf{A}$)?
2. Find all the solutions to $\mathbf{Ax} = \mathbf{0}$.
3. For which vectors $\mathbf{b}$ does $\mathbf{Ax} = \mathbf{b}$ have a solution? Why are there only certain $\mathbf{b}$ vectors that lead to solutions to $\mathbf{Ax} = \mathbf{b}$?

Solution 9.3

1. det($\mathbf{A}$)=(1)(4)-(2)(2)=0
2. There are infinitely many solutions of the form $-x_1 = 2x_2$.
3. Solutions are of the form $\mathbf{b} = \begin{bmatrix} k \\ 2k \end{bmatrix}$ where k is a constant.

While the formula for the determinant of a 2×2 matrix is quite straightforward, the procedures for computing the determinant of larger matrices is more difficult, but they are well known and well documented. Fortunately, MATLAB has the `det` function which computes the determinant.

9.3 Linear Systems of Algebraic Equations: Formulation and Definition

In previous classes, you've encountered a bunch of exercises where you had to operate on a vector to find another vector:

$$\mathbf{Ax} = \mathbf{b}, \tag{9.12}$$

where $\mathbf{A}$ and $\mathbf{x}$ were known, and your job was to find $\mathbf{b}$. While this is fun and, as you saw above in the rectangle exercise, can be useful, there is another related problem which is easily as important. It involves the same equation, but now you know $\mathbf{A}$ and $\mathbf{b}$ and need to find the vector $\mathbf{x}$. As we will discuss here, this problem captures the concept of a Linear System of Algebraic Equations.

One key idea in building models is the step of abstraction: going from some real-world situation to an abstracted model for the system (e.g., a set of differential equations). There are two important aspects of building such a model: first, deciding what to include or ignore, and second, deciding how to mathematically represent those things you choose to include.

One particularly common kind of mathematical framing is a set of linear algebraic equations, which can be represented by a matrix equation. A general system of m linear algebraic equations in n unknown variables $x_1, x_2, \ldots, x_n$ takes the form

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n &= b_2 \\
\ldots &= \ldots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots + a_{mn}x_n &= b_m
\end{aligned}$$

where $a_{11}, a_{12}, \ldots, a_{mn}$ are known as coefficients and $b_1, b_2, b_3, \ldots, b_m$ are constants. We can write this using matrices and vectors in the form

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

where $\mathbf{A}$ is the $m \times n$ *coefficient matrix*, $\mathbf{x}$ is the $n \times 1$ unknown vector, and $\mathbf{b}$ is a $m \times 1$ constant vector which is known. In other words,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \; \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \; \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Note that "linear" here means linear in terms of the unknown variables, e.g., if x is an unknown there are only terms like ax, and no terms like $\sin(x)$, x^2, $1/x$, etc. It is often the case that you might have *coefficients* that appear to be non-linear; for example, in solving physics problems, you might have coefficients that depended on trig functions of angles, such as $(L\cos\theta)F_x$, which is is linear in F_x but not linear in θ). Be careful to be clear about what you're solving for when you decide whether something is linear or non-linear.

9.4 Using Matrix Inverses to Solve Linear Systems

Over the last week, you have worked with rotation matrices, and transformations that were compositions of simpler rotations, and learned how to invert them. When you multiply a vector by any matrix (not just ones that are associated with simple spatial transformations), you transform the original vector $\mathbf{x}$ into a new vector $\mathbf{b}$.

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

More generally (than rotations), you can *often* undo the linear transformation (just like you did with the rotation matrix). Undoing this linear transformation is a linear transformation itself! Therefore the act of undoing a linear transformation can be formulated with a matrix multiply.

$$\mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$
$$\Rightarrow \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

This reduces our linear system of algebraic equations problem to the problem of finding the inverse of our matrix $\mathbf{A}$. Note this is only possible if $\mathbf{A}$ is square and *invertible*.

When solving a system of equations, at least half of the battle is typically getting your system abstracted to the point that it can be thought of as a system of linear equations. The following are a set of problems. You don't need to solve these problems – you just need to formulate them as linear algebra problems.

9.4.1 An Investment Example

In this section we will focus on deciding whether and how you can abstract the system to a mathematical model that can be written as a matrix equation.

Exercise 9.4

Suppose that the following table describes the stock holdings of three of the QEA instructors. Also suppose that on a given day the value of the Apple, IBM and General Mill's stock are \$100, \$50 and \$20 respectively.

	Apple	IBM	General Mills
Paul	100	100	100
Siddhartan	100	200	0
John	50	50	200

1. *Here's your first linear algebra formulation question:* What is the total value of the holdings for each professor on the day in question? Can you formulate this as a matrix expression? If so, what is it? If not, why not?

2. Now, suppose that you do not know how many shares of each stock are owned by the instructors. However, you know that the total value of the stocks for each instructor for three consecutive days is as given in the following table

	Paul	Siddhartan	John
Day 1	\$1500	\$2600	\$950
Day 2	\$1600	\$2810	\$1020
Day 3	\$1400	\$2550	\$1000

You also know that the price of each stock on each of the three days was as follows:

	Apple	IBM	General Mills
Day 1	\$100	\$50	\$20
Day 2	\$110	\$50	\$22
Day 3	\$100	\$40	\$30

Now here's the second formulation question: how many stocks of each company does each professor own? Can you formulate this as a matrix equation? If so, what are the matrices/vectors? If not, why not?

Solution 9.4

1. This can be formulated as $\mathbf{Ax} = \mathbf{b}$ where

$$\mathbf{A} = \begin{bmatrix} 100 & 100 & 100 \\ 100 & 200 & 0 \\ 50 & 50 & 200 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} 100 \\ 50 \\ 20 \end{bmatrix}, \text{ and } \mathbf{b} = \begin{bmatrix} d_{jd} \\ d_{et} \\ d_{jg} \end{bmatrix}.$$

Doing the matrix multiplication shows that Paul has $d_{jd} = 17000$, Siddhartan has $d_{et} = 20000$, and John has $d_{jg} = 11500$.

2. There are several ways to do this. Perhaps the simplest is to compute each person's stock holding individually. To do this, we let $\mathbf{A}$ be a matrix with the stock prices

$$\mathbf{A} = \begin{bmatrix} 100 & 50 & 20 \\ 110 & 50 & 22 \\ 100 & 40 & 30 \end{bmatrix},$$

let $\mathbf{b}_{jd}$ be a vector representing the value of Jeff's stocks on each day,

$$\mathbf{b}_{jd} = \begin{bmatrix} 1500 \\ 1600 \\ 1400 \end{bmatrix},$$

and let $\mathbf{x}_{jd}$ be a vector representing Jeff's stock holdings (i.e., the first entry tells us how many stocks of Apple he has, the second entry is IBM, and the third is General Mills). This gives the equation $\mathbf{A}\mathbf{x}_{jd} = \mathbf{b}_{jd}$. By inverting A we can solve for $\mathbf{x}_{jd}$. Then we repeat this procedure for each of the other instructors.

But... we can do it quicker! Form a 3×3 matrix $\mathbf{X}$ whose columns are made the vectors $\mathbf{x}_{jd}$, $\mathbf{x}_{et}$, and $\mathbf{x}_{jg}$. Then form a 3×3 matrix $\mathbf{B}$ whose columns are made of the vectors $\mathbf{b}_{jd}$, $\mathbf{b}_{et}$, and $\mathbf{b}_{jg}$. This gives the equation $\mathbf{AX} = \mathbf{B}$. Inverting A, we can solve for $\mathbf{X}$:

	Jeff	Siddhartan	John
Apple	10	20	5
IBM	10	10	5
General Mills	0	5	10

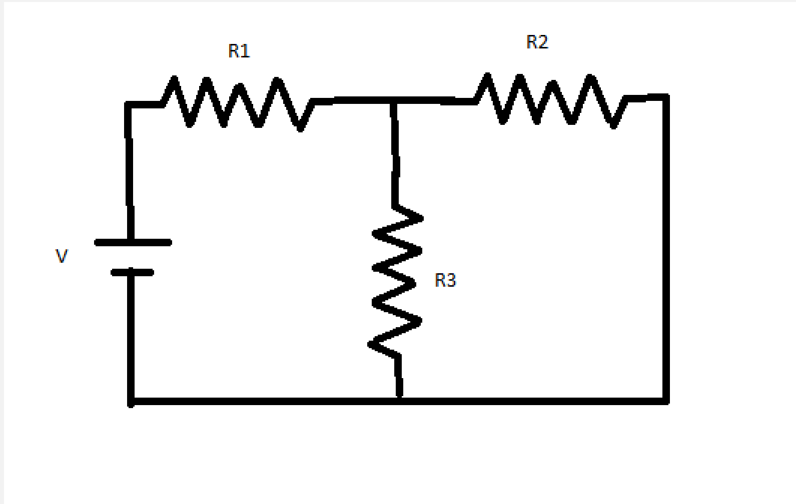
9.4.2 An Electrical Example

Remembering your circuit analysis back from ISIM, recall that Kirkhoff's laws:

- Kirkhoff's Voltage Law says that the sum of all the voltage drops around any loop of a circuit must sum to zero. (Batteries contribute a voltage increase of V, resistors contribute a voltage drop of IR.)
- Kirkhoff's Current Law says that the sum of all current going into and out of any junction of wires in the circuit must be zero.

Exercise 9.5

In the following circuit, consider that there is a current I_1 going through resistor R_1, a current I_2 going through resistor R_2 and a current I_3 going through resistor R_3. Find a linear algebra expression for the vector of our three unknown currents.



Solution 9.5

$$\begin{bmatrix} R_1 & 0 & R_3 \\ 0 & -R_2 & R_3 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} v \\ 0 \\ 0 \end{bmatrix}$$

9.5 Types of Linear Systems and Types of Solutions

Consider the linear system of algebraic equations expressed in matrix-vector form as,

$$\mathbf{Ax} = \mathbf{b}.$$

If $\mathbf{b} = \mathbf{0}$ the system of linear algebraic equations is *homogeneous* and if $\mathbf{b} \neq \mathbf{0}$ the system is *non-homogeneous*. As mentioned before, we've already dealt with systems like this before when we were transforming geometrical objects, but in that case we already knew $\mathbf{x}$ and we were simply multiplying by $\mathbf{A}$ in order to get $\mathbf{b}$. Here, we are considering the so-called *inverse* problem, and trying to find $\mathbf{x}$ given $\mathbf{A}$ and $\mathbf{b}$. However, let's back up and consider some small examples to explore the solution possibilities a little.

9.5.1 Elimination of Variables

In high school you probably learned some basic techniques for solving small linear systems of algebraic equations. Consider the following linear system of algebraic equations,

$$\begin{aligned} 2x_1 + 3x_2 &= 6 && (9.13) \\ 4x_1 + 9x_2 &= 15 && (9.14) \end{aligned}$$

The basic technique, called *elimination of Variables*, proceeds as follows: First, solve equation (2) for x_1

$$x_1 = 3 - \frac{3}{2}x_2 \tag{9.15}$$

Now substitute this expression for x_1 into equation (3)

$$4(3 - \frac{3}{2}x_2) + 9x_2 = 15$$

Now we simplify this equation

$$\begin{aligned} 12 - 6x_2 + 9x_2 &= 15 \\ \Rightarrow 3x_2 &= 3 \end{aligned}$$

and solve for x_2 to give $x_2 = 1$. Now we substitute this solution back into equation (2) or (4) to determine $x_1 = \frac{3}{2}$. The original linear system of algebraic equations therefore has a unique solution, $\mathbf{x} = \begin{bmatrix} 1 \\ 3/2 \end{bmatrix}$.

However, not all linear systems of algebraic equations have a unique solution. For example, the system

$$x_1 + 2x_2 = 1 \tag{9.16}$$
$$2x_1 + 4x_2 = 2 \tag{9.17}$$

has an infinite number of solutions because equation (6) is just a multiple of equation (5). Solving equation (5) for x_1 gives

$$x_1 = 1 - 2x_2$$

and choosing an arbitrary value of $x_2 = \alpha$ gives

$$\begin{aligned} x_1 &= 1 - 2\alpha \\ x_2 &= \alpha \end{aligned}$$

or in vector form

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

This defines an infinite number of solutions since α is any real number. What do you notice about each part of this vector?

It's also possible that a linear system of algebraic equations has no solution. For example, the system

$$x_1 + 2x_2 = 1 \tag{9.18}$$
$$2x_1 + 4x_2 = 1 \tag{9.19}$$

has no solution. Solving equation (8) for x_2 gives

$$x_2 = \frac{1}{4} - \frac{1}{2}x_1$$

and replacing into equation (7) gives

$$x_1 + 2(\frac{1}{4} - \frac{1}{2}x_1) = 1$$

which on simplification gives

$$\frac{1}{2} = 1$$

which hopefully we all agree is incorrect. We assumed that there was a solution, performed elimination and substitution and found a statement that contradicts our assumption: no solution therefore exists.

Exercise 9.6

1. Using the technique of elimination of variables described above, determine which values of h and k result in the following system of linear algebraic equations having (a) no solution, (b) a unique solution, and (c) infinitely many solutions?

$$\begin{aligned} x_1 + hx_2 &= 1 \\ 2x_1 + 3x_2 &= k \end{aligned}$$

2. Using the technique of elimination of variables described above, determine whether the following linear systems of algebraic equations have zero, one, or infinitely many solutions. If solution(s) exist, determine the actual solution(s).

   (a)
   $$\begin{aligned} x_1 + x_2 + x_3 &= 6 \\ x_2 + x_3 &= 2 \\ x_1 - 2x_3 &= 4 \end{aligned}$$

   (b)
   $$\begin{aligned} x_1 + x_2 + x_3 &= -6 \\ 2x_1 + x_2 - x_3 &= 18 \\ x_1 - 2x_3 &= 4 \end{aligned}$$

   (c)
   $$\begin{aligned} x_1 + x_2 + x_3 &= 6 \\ 2x_1 + x_2 - x_3 &= 10 \\ x_1 - 2x_3 &= 4 \end{aligned}$$

Solution 9.6

1. Rearrange the equations to linear form $y = mx + b$. If the lines are identical, there are infinitely many solutions; if the lines are parallel, but don't overlap, there are zero solutions; if the lines are not parallel, there is one solution.
   (a) h=3/2, k$\neq$2,
   (b) h$\neq$3/2,
   (c) h=3/2, k=2
2. (a) $x = \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix}$
   (b) No Solution
   (c) Infinite Solutions

9.5.2 Solving a linear system of algebraic equations in MATLAB

Exercise 9.7

In the last class, you worked with an example of fruits in your refrigerator, and we asked you questions like how to calculate the total weight of the fruits, how many fruits there are, etc. We can use matrix operations to calculate *inverse problems* as well, as this question illustrates. Suppose that you know that you have apples and oranges in the fridge and that in the genetically engineered future, the weights of all apples are 3oz and all oranges are 4oz. Because of inflation in this genetically engineered future, the price of each apple is \$1 and the price of each orange is \$2. Suppose that you also know that you paid \$13 total for your fruit and the total weight of the fruit is 33 oz. We can use this information and tools we have developed to figure out how many apples and oranges we have. Let n_o and n_a be the numbers of oranges and apples in your fridge respectively, and that you don't know what these numbers are. Define the following vectors

$$\mathbf{n} = \begin{bmatrix} n_o \\ n_a \end{bmatrix} \tag{9.20}$$

$$\mathbf{d} = \begin{bmatrix} 13 \\ 33 \end{bmatrix} \tag{9.21}$$

1. Write an equation relating $\mathbf{n}$ and $\mathbf{d}$, using a matrix-vector product.
2. Calculate how many oranges and apples you have.
3. Why this kind of problem is often called an inverse problem?

Solution 9.7

1. $\begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} n_o \\ n_a \end{bmatrix} = \begin{bmatrix} 13 \\ 33 \end{bmatrix}$
2. $\begin{bmatrix} n_o \\ n_a \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$
3. In this case we know the result $\mathbf{b}$, and are working backwards to find the number of apples and oranges. We also use a matrix inverse to find the result.

Exercise 9.8

1. Consider the example with the fruits that you worked out earlier. Now, in addition to apples and oranges, suppose you also had an unknown number of pears which each weigh 3 oz, and cost \$3. Additionally, suppose that the total weight of the fruits is 45 oz, and you paid a total of \$21 for the fruit.

   (a) If possible find the numbers of oranges, apples and pears. If not, please explain why.

   (b) Suppose that you additionally know that you have a total of 14 fruits. Can you formulate and solve a matrix-vector equation to find out the numbers of oranges, apples and pears you have?

   (c) What is the determinant of the matrix you have set up to solve this?

2. The fruit vendors bought the pricing algorithm from Uber. Oranges are still \$2, pears are now only \$1.50, and (due to an influx of teachers) apples are now surging at \$1.50 each. Their weights stay the same. You return to the market, and again purchase 14 fruits, which have the same total weight and total cost.

   (a) Can you formulate and solve a matrix-vector equation to find out the numbers of oranges, apples and pears you have?

   (b) What is the determinant of the matrix you have set up to solve this?

3. Recall the example with fruits from class: Suppose that you have a total number of 14 apples, oranges and pears in your fridge. Suppose that each apple costs \$1, each orange costs \$ 2 and each pear costs \$3. Assume also that the weights of every apple is 3 oz, every orange is 4 oz and every pear is 3 oz. Additionally, suppose that the total weight of the fruits is 45 oz, and you paid a total of \$21 for the fruit.

   (a) Formulate (or look up your formulation from class) and write down (but don't solve it yet) a matrix-vector equation to find out the numbers of oranges, apples and pears you have.

   (b) Solve this equation to find the numbers of apples, oranges and pears using the following approaches (they will of course give you the same results, but we want you to get familiar with using the different operations here).

      i. Using MATLAB, compute the inverse of the matrix in part a and use it to find the numbers of apples, oranges and pears.

      ii. Use MATLAB's `linsolve` function to find the numbers of apples, oranges and pears.

      iii. Use MATLAB's \ operator to find the numbers of apples, oranges and pears.

Solution 9.8

1. (a) No, you have three unknowns and only two equations.

   (b) Yes, you now have three equations and three unknowns.

$$\begin{bmatrix} 2 & 1 & 3 \\ 4 & 3 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} n_o \\ n_a \\ n_p \end{bmatrix} = \begin{bmatrix} 21 \\ 45 \\ 14 \end{bmatrix}$$

$$\begin{bmatrix} n_o \\ n_a \\ n_p \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \\ 2 \end{bmatrix}$$

   (c) $det(\mathbf{A}) = 2$

2. (a)

$$\begin{bmatrix} 2 & 1.50 & 1.50 \\ 4 & 3 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} n_o \\ n_a \\ n_p \end{bmatrix} = \begin{bmatrix} 21 \\ 45 \\ 14 \end{bmatrix}$$

   This forumation cannot be solved because A is not invertible.

   (b) $det(\mathbf{A}) = 0$

3. (a)

$$\begin{bmatrix} 2 & 1 & 3 \\ 4 & 3 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} n_o \\ n_a \\ n_p \end{bmatrix} = \begin{bmatrix} 21 \\ 45 \\ 14 \end{bmatrix}$$

   (b) i.

$$\begin{bmatrix} n_o \\ n_a \\ n_p \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \\ 2 \end{bmatrix}$$

      ii.

$$\begin{bmatrix} n_o \\ n_a \\ n_p \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \\ 2 \end{bmatrix}$$

iii.

$$\begin{bmatrix} n_o \\ n_a \\ n_p \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \\ 2 \end{bmatrix}$$

9.6 Conceptual Quiz

1. Select the matrices which are invertible.

   (a) $\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$

   (b) $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$

   (c) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

   (d) $\begin{bmatrix} 1 & 2 \\ 4 & 8 \end{bmatrix}$

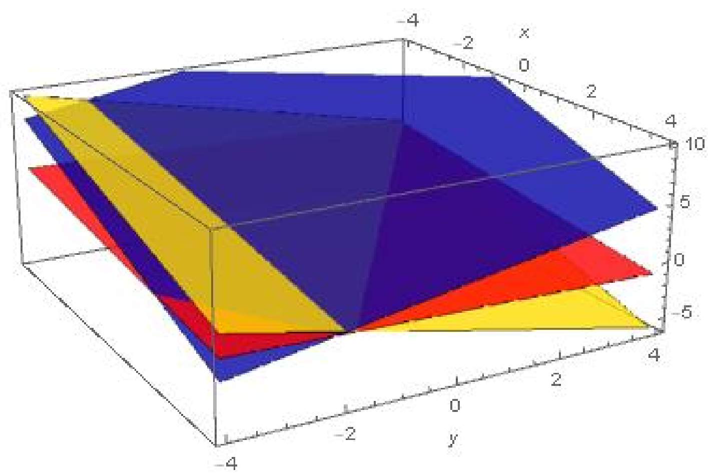
   (e) $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

2. Let (a, b, c) be the point of interesection for the following three planes, pictured below:

$$z = 2 - x - y$$

$$z = (31 - 6x + 4y)/5$$

$$z = (13 - 5x - 2y)/2$$



   What is a?

3. How many solutions does the following system of equations have?

$$x + y = 9$$

$$x - z = 2$$

$$y + z = 7$$

A. Zero
B. One
C. Two
D. Infinitely many

4. What is the area of a parallelogram whose vertices are $(0, 0)$, $(2, 4)$, $(5, 1)$ and $(7, 5)$?

5. Solve the following system of linear equations

$$x - y = 2$$

$$3x + z = 11$$

$$y - 2z = -3$$

What is the value of y?