# *Night 1: Introduction to Matrices*

## *Overview and Orientation*

In this night assignment, we will learn some of the foundational 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 for 2D and 3D matrices

- 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 track down an instructor, 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

## *2.1    Linear Algebra, Vectors, and Matrices*

In your concept-maps for eigenfaces, most, if not all of you would have included something about linear algebra, matrices, and vectors. These topics are used very heavily in many different areas, including in data analysis. For the next couple of weeks, you will spend a good deal of time learning about these things and how to apply them.

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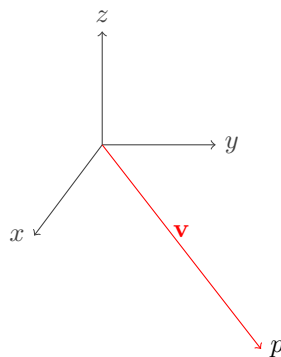
Consider the point $p = (1, 2, -1)$ in 3-dimensional space. We can associate a position vector $\mathbf{v}$ with this point, which is the vector from the origin to this point,

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

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

$$\mathbf{u}^T = \begin{bmatrix} p \\ q \\ r \end{bmatrix} \tag{2.1}$$

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

$$\mathbf{v}^T = \begin{bmatrix} x & y & z \end{bmatrix}. \tag{2.2}$$

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

$$\mathbf{u}\mathbf{v} = \begin{bmatrix} p & q & r \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = px + qy + zr \tag{2.3}$$

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 *dot* 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 $\theta$ is the angle between $\mathbf{v}$ and $\mathbf{w}$ and

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}^{1/2}$$

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

---

### Exercise 2.1

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

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

---

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

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^T \mathbf{w}. \tag{2.4}$$

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 2.2

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

---

$$\mathbf{f} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \tag{2.5}$$

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.

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 functions.*

### Introduction to 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

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}. \tag{2.6}$$

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 $3 \times 2$ 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 2.3

Recall the matrix $\mathbf{G}$ defined in equation (2.6) and the vector $\mathbf{f}$ defined in Exercise 2.2, which kept track of the number of fruit of different types. What does the vector $\mathbf{Gf}$ represent?

---

### Exercise 2.4

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

$$\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{2.7}$$

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

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{2.8}$$

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

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

$$\mathbf{B} = \begin{bmatrix} \cos(30°) & -\sin(30°) \\ \sin(30°) & \cos(30°) \end{bmatrix} \tag{2.9}$$

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

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

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{2.10}$$

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

8. Please draw a new spatial vector

$$\mathbf{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{2.11}$$

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

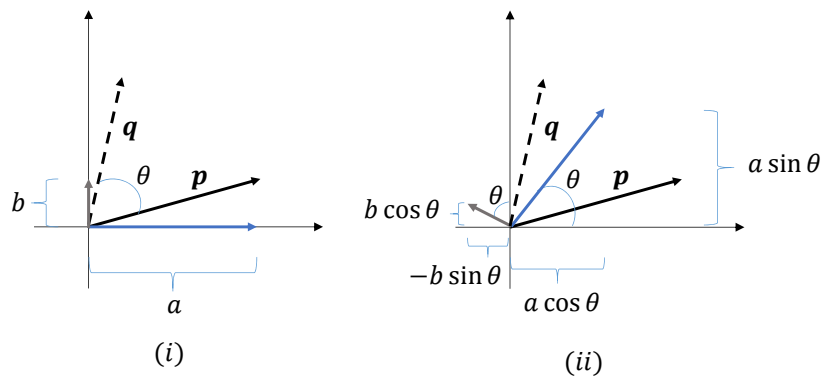10. What does multiplying *any* vector by $\mathbf{R}$ do?



(i)                    (ii)

Figure 2.1: Rotation of vectors

You may have guessed that $\mathbf{R}$ defined above, rotates a vector counter-clockwise by $\theta$. 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 41.4 (i). Suppose that we wish to rotate the vector $\mathbf{p}$ counter-clockwise by $\theta$, which will result in the vector $\mathbf{q}$. From the figure, we see that

$$\mathbf{p} = \begin{bmatrix} a \\ b \end{bmatrix}, \tag{2.12}$$

and $\mathbf{p}$ is the sum of the gray and blue vectors. If we now rotate the blue and gray vectors counter-clockwise by $\theta$, we see that $\mathbf{q}$ is the sum of the rotated versions of the blue and gray vectors, as shown in Figure 41.4 (ii). By using trigonometry, we see that the blue vector in Figure 41.4 (ii) is

$$\begin{pmatrix} a \cos \theta \\ a \sin \theta \end{pmatrix} \tag{2.13}$$

and the gray vector in Figure 41.4 (ii) is

$$\begin{pmatrix} -b \sin \theta \\ b \cos \theta \end{pmatrix} \tag{2.14}$$

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

$$\mathbf{q} = \begin{pmatrix} a \cos \theta \\ a \sin \theta \end{pmatrix} + \begin{pmatrix} -b \sin \theta \\ b \cos \theta \end{pmatrix} = + \begin{pmatrix} a \cos \theta - b \sin \theta \\ a \sin \theta + b \cos \theta \end{pmatrix} = \mathbf{R}\mathbf{p}. \tag{2.15}$$

*General Notation*

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 **w**, is given by the following sum

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

### Other Matrix operations

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 in the next section. 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.

## 2.2    Matrix Operations in MATLAB

---

### Exercise 2.5

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!
```

---

To practice matrix operations, let's define a matrix and some vectors using MATLAB as follows:

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

Note that the semi-colon ends a row and begins a new row. You can also use returns between rows—try it! Square brackets enclose the matrix. To define the column vector **v** in MATLAB you can type the following command:

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

whilst to define the row vector **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 2.6

Using the definitions for **A**, **v**, and **u** from above, please predict the output of the following commands and then solve them using MATLAB.

```
A*v
u*A
A(1:2,:)*v
u*A(:,2)
```

---

For Night 1, you will also need to plot things. In MATLAB, you can use plot(xv,yv) to create a scatter plot.

```
yv=[1 7 4 5 3 9 2 4]
xv=[1 3 4 6 8 9 11 14]
plot(xv,yv)
```

If you want to know more about how to use a function like plot, use "help." Create the plot above, then type "help plot" into the command window and try to change something about your plot, such as using points instead of a line or adding axis labels.

Finally, you need to know how to use a for loop to repeat a set of commands a number of times. Here's an example for loop that makes a vector that's a sequence of squares: (Try it!)

```
for n=1:3% n is the index variable, which counts from 1 to 3 (call it whatever you want)
v(n)=n^2% assigns the nth component of v to the value n^2 and prints out v
% loop repeats, adding 1 to n each time, until i gets to 3
end% needed to end the loop!
```

---

### Exercise 2.7

Write a for loop that creates the following matrix:

```
M_squares=[1 1;2 4;3 9;4 16]
```

---

## 2.3    Elementary Matrix Operations, Properties, and Terminology

In this part of the assignment, you will learn a number of basic operations and properties of matrices which can then be used in applications. Admittedly, most of these exercises are a little dry, but they will be useful in the very near future, we promise!

*Matrix-Vector Multiply*

Here, you will work on examples of matrices multiplying vectors to get yourselves comfortable with matrix operations in MATLAB. First, let's define the matrix **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 **v** in MATLAB you can type the following command:

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

whilst to define the row vector **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 2.8

Using the definitions for **A**, **v**, and **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,:)`

---

*Addition, subtraction, scalar 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

$$\mathbf{A} = \begin{bmatrix} 3 & 4 & 1 \\ 3 & 1 & 1 \end{bmatrix} \tag{2.16}$$

$$\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 1 \end{bmatrix} \tag{2.17}$$

$$\mathbf{A} - \mathbf{B} = \begin{bmatrix} (3-1) & (4-2) & (1-3) \\ (3-2) & (2-1) & (1-1) \end{bmatrix} \tag{2.18}$$

$$= \begin{bmatrix} 2 & 2 & -2 \\ 1 & -1 & 0 \end{bmatrix} \tag{2.19}$$

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

$$3\mathbf{A} = \begin{bmatrix} 9 & 12 & 3 \\ 9 & 3 & 3 \end{bmatrix} \tag{2.20}$$

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,

$$\begin{bmatrix} 1 & 3 & 5 \\ 2 & 7 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 \\ 3 & 7 \\ 5 & 6 \end{bmatrix} \tag{2.21}$$

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

$$(\mathbf{A}\mathbf{v})^T = \mathbf{v}^T \mathbf{A}^T \tag{2.22}$$

---

### Exercise 2.9

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

---

### Exercise 2.10

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

---

**Exercise 2.11**

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

**Exercise 2.12**

How do you find the transpose of a vector or matrix in MATLAB?

## Matrix-Matrix Multiply

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 $\mathbf{C} = \mathbf{A}\mathbf{B}$ 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 $\mathbf{C} = \mathbf{A}\mathbf{B}$.

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

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

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:

$$\mathbf{A} = \left[ \begin{array}{rr} -2 & 4 \\ 0 & 3 \end{array} \right]$$

$$\mathbf{B} = \left[ \begin{array}{rr} 5 & -3 \\ -1 & -1 \end{array} \right]$$

---

**Exercise 2.13**

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

---

---

**Exercise 2.14**

Find the matrix product $\mathbf{BA}$

---

---

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.

$$\left[ \begin{array}{rr} 5 & -3 \\ -1 & -1 \end{array} \right] + \left[ \begin{array}{rr} 4 & -2 \\ -3 & -1 \end{array} \right] = \left[ \begin{array}{rr} 9 & -5 \\ -4 & -2 \end{array} \right]$$

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

$$\mathbf{C} = \left[ \begin{array}{rr} -5 & -1 \\ -3 & 2 \end{array} \right]$$

---

**Exercise 2.15**

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

---

---

---

**Exercise 2.16**

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

---

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 2.17**

1. $\mathbf{A}^2$

2. $\mathbf{B}^3$

---

*Special Types of Matrices*

---

**Exercise 2.18**

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

---

## 2.4   Matrices as transformation operators

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.

## Identity and Scaling Operations

The matrix

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.24}$$

when multiplying the vector

$$\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{2.25}$$

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 $4 \times 4$ matrix with 1s on the diagonal and zeros everywhere else, i.e.,

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.26}$$

---

### Exercise 2.19

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 $3 \times 3$ 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 $3 \times 3$ 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 $3 \times 3$ 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 $3 \times 3$ matrix which scales the $x$ component by $a$, the $y$ component by $b$, and the $z$ component by $c$.

## 2.5    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

$$\begin{bmatrix} 41 & 35 & 37 & 43 \\ 49 & 40 & 48 & 61 \end{bmatrix} \tag{2.27}$$

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 2.20

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. The data provided is given in Fahrenheit, and suppose you wish to convert it to Celsius using matrix operations (note that there are a number of ways of doing this, but we are focusing on using matrices here). The formula for converting a Fahrenheit temperature to Celsius is to first subtract 32 from the Fahrenheit temperature, and multiply the result by $\frac{5}{9}$.

   a) Define a matrix of the same shape as T with all its entries equalling 32, and call this matrix B. You will find MATLAB's `ones` function, which generates a matrix filled with 1's, useful here.

   b) Define a square, diagonal matrix of the appropriate dimensions which when multiplying another matrix scales all its entries by $\frac{5}{9}$. You should call this matrix A. You will find MATLAB's `eye` function, which generates an identity matrix, useful here.

c) Using your answers to the previous parts and appropriate matrix operations, please provide 1 line of MATLAB code which generates a new matrix Y which contains the temperature data in Celsius.

3. Lets go back to Fahrenheit for the rest of the assignment. 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.

4. Using MATLAB's mean function, which computes the average of the values in a vector, and guess, based on geography, which of the vectors corresponds to the temperature for which city.

5. What are the maximum and minimum temperatures for Boston in the 20 years for which you have data?

6. 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.

7. 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 $3 \times 365$ matrix of daily temperatures for the last year (for which you have data) in Boston, Providence and Washington DC.

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

$$T_n \approx 0.2235T_b + 0.4193T_p + 0.3856T_w. \tag{2.28}$$

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.

9. 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.)

## 2.6 Conceptual Quiz

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

1. **A** is a $3 \times 4$ matrix and **B** is a $4 \times 2$ matrix. What is the size of **AB**?

A. $2 \times 3$
B. $3 \times 1$
C. $3 \times 2$
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.

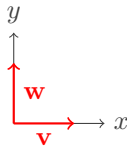$$\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}$$

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?
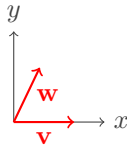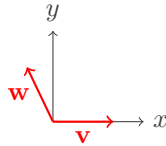
A.



B.



C.



D.
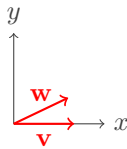
## Solution 2.1

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

## Solution 2.2

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.

## Solution 2.3

The vector $\mathbf{Gf}$ is a $2 \times 1$ vector whose first entry represents the total number of fruits and second entry represents the number of citrus fruits.
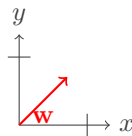
## Solution 2.4
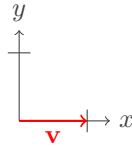
1. The vector $\mathbf{v}$ is



2. First, we compute
$$\mathbf{w} = \mathbf{Av} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix},$$

   which is visually represented as



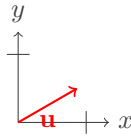3. Multiplying $\mathbf{v}$ by $\mathbf{A}$ rotated the vector counterclockwise by 45 degrees.

4. First we compute

$$\mathbf{u} = \mathbf{Bv} = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$$
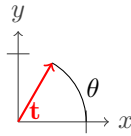
which is visually represented as



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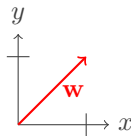
6. First we compute

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

which is visually represented as



7. Multiplying $\mathbf{v}$ by $\mathbf{R}$ rotated the vector counterclockwise by $\theta$ degrees.
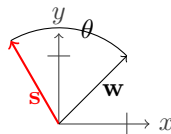
8. The vector $\mathbf{w}$ is



9. First we compute

$$\mathbf{s} = \mathbf{Rw} = \begin{bmatrix} \cos\theta - \sin\theta \\ \sin\theta + \cos\theta \end{bmatrix}$$

which is visually represented



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

## Solution 2.7

```
for n=1:4
M_squares(n,1) = n; %assigns the nth row, 1st column the value n
M_squares(n,2) = n^2; %assigns the nth row, 2nd column the value n^2
end
```

## Solution 2.8

1. `[-3; -7; 4]`

2. `[-5 9]`

3. Does not work because the inner matrix dimensions must agree and here we have a $3 \times 2$ matrix multiplied by a $1 \times 3$ matrix

4. Does not work because the inner matrix dimensions must agree and here we have a $2 \times 1$ matrix multiplied by a $3 \times 2$ 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 $1 \times 3$ matrix multiplied by a $1 \times 2$ matrix.

## Solution 2.9

$$4\mathbf{A} - 5\mathbf{B} = \begin{bmatrix} 7 & 6 & -11 \\ 2 & -6 & -1 \end{bmatrix} \tag{2.23}$$

## Solution 2.10

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

## Solution 2.11

$\mathbf{Av}$ is $4 \times 1$ and $(\mathbf{Av})^T$ is $1 \times 4$.

## Solution 2.12

You use the apostrophe: $(\mathbf{A})^T$ is A' in Matlab.

### Solution 2.13

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

### Solution 2.14

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

### Solution 2.15

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

### Solution 2.16

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

### Solution 2.17

1.

$$\mathbf{A}^2 = \begin{bmatrix} 4 & 4 \\ 0 & 9 \end{bmatrix}$$

2.

$$B^3 = \begin{bmatrix} 152 & -72 \\ -24 & 8 \end{bmatrix}$$

### Solution 2.18

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.,

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

## Solution 2.19

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:

$$\mathbf{M} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

## Solution 2.20

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.  a) We can define a matrix B of the same size as T but filled with 1s by typing » `B = ones(4,7670)` and then we can multiply it by 32 by typing » `B = 32*B`. Alternatively we could include the 32 from the start by typing » `B = 32*ones(4,7670)`

    b) There are two ways to do this: we can multiply either on the left or right.

      - Option 1: To multiply a $4 \times 7670$ matrix on the left, we need a $4 \times 4$ matrix, which we would create by typing » `A = eyes(4)*5/9`.
      - Option 2: To multiply a $4 \times 7670$ matrix on the left, we need a $7670 \times 7670$ matrix, which we would create by typing » `A = eyes(7670)*5/9`.

c) Now that we have the pieces in place we could simply type » `Y = A*(T-B)` (or `Y = (T-B)*A` if you chose the right multiply option). We can actually make this a lot simpler because MATLAB will create matrices of the correct size on the fly, so the following will work » `Y = (T-32)*5/9`. Normally we would expect `T-32` to be a problem because we are subtracting a scalar from a matrix, but MATLAB simply assumes that we wish to subtract 32 from every element in the matrix.

3. 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.

4. 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.

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

6. 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.

7. 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.

8. 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.

9. 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%.