Matrices

Last week we worked with **vectors** where numbers were ordered in any one dimension eg. rainfall in Sarjapura on each day of October 2022. This vector would have 31 elements, each representing the amount of rainfall in mm that fell in Sarjapura on those days.

Now, there would be such a vector for every year - 2021, 2020, 2019, 2018 etc. So, if we wanted to represent the daily rainfall in October over the last five year, one way could be to make one long vector of length 5x31=155. But there are actually two dimensions to this set of numbers (or this data). One is days, which ranges from 1 to 31, and the other is years, which ranges from 2018-2022. Hence, we can take five vectors for each of the five years and arrange them vertically to give a 2D structure to the numbers.

What we would get is a rectangle of width 31 and height 5. Each row (i.e. a horizontal line) would be a vector of length 31 giving the daily October rainfall for a particular year. Each column (i.e. a vertical line) would also be vector, of length 5, giving the rainfall for a particular day in October in each of the five years.

This 2 dimensional structure of numbers is called a **matrix**. One can think of it as a collection of row-vectors stacked one-below-the-other, and simultaneously as a collection of column-vectors stacked from left to right. Just as the number of elements in a vector is called its length, the number of elements in a matrix is signified by its **order** and is specified by mentioning the number of rows and columns. For example, the matrix mentioned above would be of the order 5x31.

The elements of a vector **x** were represented as x_i, where the subscript i would denote the i^{th} element i.e, x_1 would be the first element, x_2 the second element and so on. For a matrix **A** (matrices are typically denoted by capital letters), its element is represented by a_{ij} . The two subscripts i and j represent the row and the column of the element. For instance, in a 2x2 matrix **A** given by

$$\mathbf{A} = \begin{pmatrix} 3 & -2 \\ 5 & 8 \end{pmatrix}$$

$a_{11} = 3$, $a_{12} = -2$, $a_{21} = 5$, and $a_{22} = 8$,

Note that a row vector of length n is a matrix of order 1xn, and a column vector of length n is a matrix of the order nx1.

Application: Social networks

Matrices are used in a wide variety of applications. The study of the properties of matrices and its used is called **linear algebra.** We will see one of my favorite examples of the use of matrices, which is to study networks, and specifically social networks.

Think of 3 people whom we index as person 1, person 2 and person 3. We can define a 3x3 matrix that shows who follows whom on a particular social network (or who is friends with whom in real life). We can define the elements of this matrix as follows
a_{ij} = 1 if i is followed by j, and 0 otherwise

Let person 1 follow person 2 (a_{21}=1), and person 2 also follow person 1 (a_{12}=1). Let person 2 follow person 3 but is not followed back (a_{32}=1, a_{23}=0). Let person 3 follow person 1 but is not followed back (a_{13}=1, a_{31}=0). Also, a_{11}, a_{22} and a_{33} would be 0 as no one is following themselves.

So, the matrix would look like this:

0	1	1
1	0	0
0	1	0

We can depict this as an actual network using some basic code (try https://graphonline.ru/en/)

Any such matrix that depicts the connections in a network is called an adjacency matrix.

Q. Can this be an adjacency matrix?

0	1	1
1	0	1

[A: No, the number of rows and columns have to be equal as it is reflecting the connections between the same people]

A matrix where the number of rows and columns are equal is called a **square matrix.**

Q. What can we say about a social network if a_{ij}=a_{ji} for all i and j. For example:

0	1	1
1	0	0
1	0	0

[A. This would mean that all relationships are both ways i.e. if i follows j, then j also follows i]

A square matrix where a_{ij}=a_{ji} for all i and j is called a **symmetric matrix.**

Now, suppose we have two adjacency matrices **A** and **B** , which give data on people's connections in two different online social networking platforms, can we add them to get the combined information from both platforms i.e. create a matrix **C**=**A**+**B ?**

[A. Yes, if they are of the same order and represent the same people. Then we can do element-by-element addition to get combined information]

So, if **A=**

0	1	1
1	0	0
1	0	0

And **B=**

0	1	1
1	0	0
0	1	0

Then **A**+**B=**

0	2	2
2	0	0
1	1	0

We will come back to this example later to see how we can use adjacency matrices to get information about networks that is not obvious otherwise

Application: Linear equations

Recall the example we did in vectors of a catering firm with different kinds of meals with prices given by a vector **p** and quantities given by a vector **q** and hence the revenue was given by **p.q**, the inner product.

Let us take the same case as the last example with three types of meals. Hence, both vectors are of length 3. Let us say we observe the number of meals and the revenue but not the prices. Suppose the first day they sold 20 meals of the first type, 40 of the second type, and 50 of the third type and earned a revenue of 4000 rupees. On the second day they sold 35 meals of the first type, 30 of the second type and 45 of the third type and earned revenue 3900 rupees.

Q. Can you express this as simultaneous linear equations?

20p1 + 40p2 +50p3 = 4000
35p1 + 30p2 +45p3 = 1950

Now we will try and express it in terms of vectors. We know that (p1,p2,p3) is the unknown vector. The first day's quantities, let us call it $\mathbf{q}_{day1}$ are given by (20,40,50) and the second day's $\mathbf{q}_{day2}$= (35,30,45). Also, we know that $\mathbf{p}.\mathbf{q}_{day1}$=2200, and $\mathbf{p}.\mathbf{q}_{day2}$=1950

Since there is a two dimensional structure there, we can arrange all the data into a matrix.

Q=			
	20	40	50
	35	30	45

We know that **p** is a vector of length 3 that can be written either as a column or row vector and hence a matric.

Now, revenue **r** is also a vector of length 2!

What we want to be able to say is that **Q** multiplied by **p** gives **r**. To do this let us look at how matrix multiplication is defined.

Matrix multiplication: A matrix **A** of order m x n, i.e. with m rows and n columns can only be multiplied with a matrix that has n rows. Let **B** be a matrix of order n x k i.e with n rows and k columns. The product **C=AB** will be a matrix of order m x k i..e with m rows and k columns, with the element c_{ij} given by the inner product of the ith row of matrix A and jth column of matrix B.

For example, let **A** by a 3x3 matrix

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

And **B** be a 3x2 matrix

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

Then the matrix multiplication is performed as follows

$$\mathbf{AB} = \begin{pmatrix} 0 & 1 & 2 \\ \boxed{2 \quad 3 \quad 1} \\ 4 & -1 & 6 \end{pmatrix} \begin{pmatrix} \boxed{\begin{matrix} 3 \\ 1 \\ -1 \end{matrix}} & \begin{matrix} 2 \\ 0 \\ 1 \end{matrix} \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ \boxed{8} & 5 \\ 5 & 14 \end{pmatrix}$$

Note that **AB** is not equal to **BA**. In fact, in this case **BA** is not defined as **B** has only two columns but **A** has 3 rows.

Some examples to get you used to it:

Let A =		
	0	2
	1	-1
	3	0

B =	1	2	1

C=	0
	1

Which one of the following are defined? Perform the multiplication where it is defined

AB, BA, BC, CB, AC, CA

Some properties of matrix multiplication:

- **AB ≠ BA**
- **A(B+C)=AB + AC - (**can you show this?)
- **(AB)C = A(BC)**

Note that multiplying a number with a matrix works exactly like it does for vectors

- **B**=k**A,** where k is a number (scalar) then b_{ij}=k a_{ij}

Identity matrix: A square matrix of order nxn with all its diagonal elements as 1 and all its other elements as 0 is called an identity matrix and is denoted by **I**.
Q. Show that **AI**=**A,** where **A** is any matrix of the order mxn and **I** is an identity matrix of order nxn. Also **IA=A,** where **A** is any matrix of the order mxn and **I** is an identity matrix of order mxm.

Coming back to our example: **Q** is a matrix of the order 2 x 3

Now, if we define **p** as a column vector of length 3 i.e. a matrix of the order 3 x 1, then **Qp** is a column vector of length 2, or a matrix of order 2 x 1.

Hence, we can write the system of equations as

Qp=**r**

In fact, we can write any system of linear equations in this fashion:

$3x_1+2x_2+5x_3+... = 54$
$2x_1-3x_2+10x_3+..=30$
...

If we denote the j^{th} coefficient of the i^{th} equation as a_{ij} , and the right hand side constant for the i^{th} equation as b_i , then the system of equations becomes

$a_{11}x_1 + a_{12}x_2 + a_{13}x_3+... = b_1$
$a_{21}x_1 + a_{22}x_2 + a_{23}x_3+..= b_2$

…

In matrix form we can write these system of equations as just one equation:

Ax=**b**

Remember that here **x** is a vector of unknowns, and **A** and **b**, are both known.

What we eventually want to do is something like
x=**b**/**A** !!

But for that we will have to define division by a matrix - we will do something to that effect in a little while. But for now let us look at what use we can put matrix multiplication to when applied to adjacency matrices.

Consider an adjacency matrix **A**. We can work with the one we constructed earlier as an example

0	1	1
1	0	0
0	1	0

Q. If **k** is a column vector of all 1s of the length same as one dimension of the adjacency matrix (3 in this case), then what is the interpretation of **A k** ?
[Number of followers]

Q. Let us denote **AA** as $\mathbf{A}^2$ (note that only square matrices can be squared - why?). What do the elements of $\mathbf{A}^2$ signify?
[How many paths of length 2 are there between j and i]

And therefore $\mathbf{A}^2$ **k** gives the total number of followers of the followers for every person.

Exercise:
Find the squares and the cube of the following matrix:

A=

2	1
-1	0

Ans: $\mathbf{A}^2$ =

3	2
-2	-1

$\mathbf{A}^3$ =

4	3
-3	-2

Application: Influence vector

Q. How did Google start? What is a search engine? How does it work?
[Small discussion on how they think a search engine may work]

When Google first came out, its advantage over existing search engines was its ability to give important results first. The way it did that was using an algorithm called PageRank. Here we can see a simplified version of the algorithm.

Let us first do it in the context of a social network and then we can draw an analogy to the worldwide web.

Q. Let **A** be the adjacency matrix of all users to an online platform (FB, Insta, Twitter etc.). We want a way to rank the users by their 'influence'. How do we do that?

A first guess would be to look at the number of followers. **Ak** will give us that. But it is not just the number but the quality of followers which also matters. Being followed by a person with 1000 followers themselves is not the same as being followed by a bot with 0 followers. So, we need to look at followers of followers. $\mathbf{A}^2\ \mathbf{k}$ will give us that. But similarly, the quality of the followers of followers also matters - so we need to look at $\mathbf{A}^3\ \mathbf{k}$, and so on.

To find the overall influence vector, we will add all of these up but we should give less weightage to followers of followers, than to followers, and so on. Let this less weightage be signified by multiplying by a fraction, say 0.5.

$\mathbf{Ak} + 0.5\mathbf{A}^2\ \mathbf{k} + 0.5\text{x}0.5\ \ \mathbf{A}^3\ \mathbf{k}\ \ldots$

$=(\mathbf{I} + 0.5\mathbf{A} + (0.5\mathbf{A})^2 + \ldots)\ \mathbf{Ak}$

(Here **I** is the identity matrix of the appropriate size)

You should be able to see that the terms inside the brackets look similar to a geometric progression $1 + x + x^2 + \ldots$, which is equal to 1/1-x if x<1

Similarly in this case, we can get something like **Ak** / (**I**-0.5**A**). Again, we need to define division by a matrix or something similar to that

Before we do that, let us look at how this is applicable to PageRank. Suppose **A** denotes the entire web with each node (i.e. each row/column of the matrix) being a webpage, and the element a_{ij}=1 if webpage *j* links to webpage *i*, and is 0 otherwise.

So, the vector **A k** will give the number of webpage that directly link to each webpage on the net. But a link from an unknown blog is not the same as a link from, say Wikipedia. So we want to know how linked are the 'follower' webpage. Hence the same logic as social network influence applies and we get a similar expression for the 'influence' of each webpage which is then used to display the higher ranked webpages first.