GENERAL ASSEMBLY

INTRO TO DATA SCIENCE
LESSON 2: LINEAR ALGEBRA

WHAT IS DATA SCIENCE
DATA EXPLORATION AND WORKFLOW
PYTHON DATA STRUCTURES

ANY QUESTIONS?

I. LINEAR ALGEBRA REVIEW
II. THE PYTHON CONTROL FLOW

LAB:
III. MATRIX MULTIPLICATION IN PYTHON
IV. ADDING CONTROL FLOW INTO CLICKS AGGREGATION

I. LINEAR ALGEBRA REVIEW

In order to best understand most machine learning algorithms, we need some basis of linear algebra.

In order to best understand most machine learning algorithms, we need some basis of linear algebra.

Linear algebra is best defined as mathematics in the multidimensional space and the mapping between said spaces.

FOR EXAMPLE...

$$y = mx + b$$

$$y = m_1x_1 + m_2x_2 + b$$

$$y = m_1x_1 + m_2x_2 + m_3x_3 + m_4x_4 + b$$

$$y = m_1x_1 + m_2x_2 + m_3x_3 + m_4x_4 + m_5x_5 + m_6x_6 + m_7x_7 + m_8x_8 + m_9x_9 + m_{10}x_{10} + b$$

Matrices are an array of real numbers with m rows and n columns

Each value in a matrix is called an entry.

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

Matrices are an array of real numbers with m rows and n columns

Each value in a matrix is called an entry.

$$A_{2\,1} \rightarrow \begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{matrix}$$

Vectors are a special kind of matrix, as they only consist of one dimension of real numbers.

These look most like a numeric array (or **list**) in Python.

$$[1 \quad 3 \quad 9 \quad 2]$$

Likewise, you can refer to each index or value similarly
(a[0] in Python is the same entity as 0 in vector a)

Rule 1!
Matrices can be added together only when they are the same size.
If they are not the same size, their sum is **undefined**.

$$\begin{bmatrix} 1 & 3 & 9 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 5 & 9 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 8 & 18 & 6 \end{bmatrix}$$

Rule 1!
Matrices can be added together only when they are the same size.
If they are not the same size, their sum is **undefined**.

$$\begin{bmatrix} 1 & 3 & 9 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 5 & 9 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 8 & 18 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 72 & 3 & 1 \end{bmatrix} + \begin{bmatrix} 17 & 55 & 3 & 10 \end{bmatrix} = ?$$

Rule 2!
Matrices can be multiplied by a scalar (single entity) value.
Each value in the matrix is multiplied by the scalar value.

$$[\,1 \quad 3 \quad 9 \quad 2\,] * 3 = [\,3 \quad 9 \quad 27 \quad 6\,]$$

$$[\,8 \quad 72 \quad 3 \quad 1\,] * 2 = ?$$

Rule 3!

Matrices and vectors can be multiplied together given that the matrix columns are as wide as the vector is long.

The result will always be a vector.

$$\begin{matrix} 1 & 3 & 9 & 2 \\ 2 & 4 & 6 & 8 \end{matrix} * \begin{matrix} 2 \\ 3 \\ 6 \\ 5 \end{matrix} = \begin{matrix} 2+6+54+10 \\ 4+8+36+40 \end{matrix} = \begin{matrix} 72 \\ 88 \end{matrix}$$

Matrices represent the multiple dimensions in our data! If we had a a vector that suggested how important each dimension of our data was, we could use that to find our best **linear model**!

Matrices represent the multiple dimensions in our data! If we had a a vector that suggested how important each dimension of our data was, we could use that to find our best **linear model**!

We will see matrices quite often in **all** of our data, so pay careful attention to how data is structured and how different algorithms interact with them