Mathematics for Artificial Intelligence
Compilation: Yogesh Kulkarni

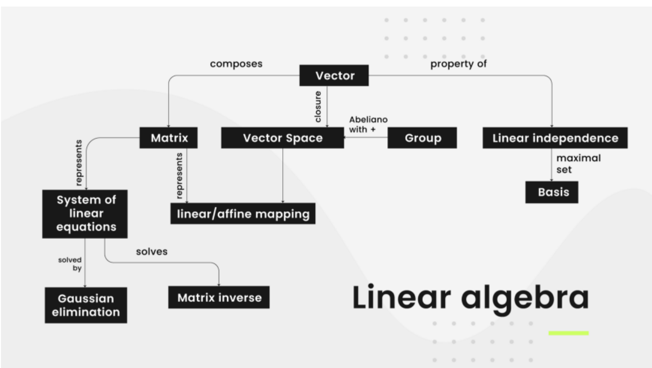
Linear Algebra

Linear Algebra

Basic Entities

- Scalars?
- Vectors?
- Matrices?
- Next? (or Whats this called collectively?)
- Point in n-dimensional space is represented by?

Landscape: Linear Algebra



(Ref: The NOT definitive guide to learning math for machine learning - Favio Vazquez)

Vectors

Vectors

- At its simplest, a vector is an entity that has both magnitude and direction.
- The magnitude represents a distance (for example, “2 miles”) and the direction indicates which way the vector is headed (for example,“East”).
- One more way is $\bar{v} = 2\hat{i} + 3\hat{j}$; meaning?
- Is Magnitude-Direction form equivalent to i-j form?
- Inter-convertible? How?
- Can it have just two components?

Vectors

Two-dimensional example:

- A vector that is defined by a point in a two-dimensional plane
- A two dimensional coordinate consists of an x and a y value, and in this case we’ll use 2 for x and 1 for y
- Its is written in matrix form as : $\vec{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$
- Describes the movements required to get to the end point (of head) of the vector
- So, it is not a point in space. It gives Direction, like a movement recipe.
- When added to a point, results into a transformed point.
- In this case, we need to move 2 units in the x dimension, and 1 unit in the y dimension

Vectors

Two-dimensional example:

- Note that we don’t specify a starting point for the vector
- We’re simply describing a destination coordinate that encapsulate the magnitude and direction of the vector.
- Think about it as the directions you need to follow to get to **there** from **here**, without specifying where **here** actually is!
- Generally using the point 0,0 as the starting point (or origin). Also called as Position Vector.
- Our vector of (2,1) is shown as an arrow that starts at 0,0 and moves 2 units along the x axis (to the right) and 1 unit along the y axis (up).

Vectors

Calculating Magnitude

- $\|\vec{v}\| = \sqrt{{v_1}^2 + {v_2}^2}$
- Double-bars are often used to avoid confusion with absolute values.
- Note that the components of the vector are indicated by subscript indices $(v_1, v_2, \dots v_n)$
- In this case, the vector v has two components with values 2 and 1, so our magnitude calculation is:
- $\|\vec{v}\| = \sqrt{2^2 + 1^2} = \sqrt{5} \approx 2.24$

Vectors

Calculating Direction

- We can get the angle of the vector by calculating the inverse tangent; sometimes known as the arctan
- For our v vector (2,1):$tan(\theta) = \frac{1}{2}$
- $\theta = tan^{-1}(0.5) \approx 26.57^o$
- use the following rules:
  - Both x and y are positive: Use the tan-1 value.
  - x is negative, y is positive: Add 180 to the tan-1 value.
  - Both x and y are negative: Add 180 to the tan-1 value.
  - x is positive, y is negative: Add 360 to the tan-1 value.

Vectors

- Vectors are defined by an n-dimensional coordinate that describe a point in space that can be connected by a line from an arbitrary origin.
- Are n-dimensional Points and Vectors equivalent? How?
- $\|\vec{v}\| = \sqrt{{v_1}^2 + {v_2}^2 ... + {v_n}^2}$

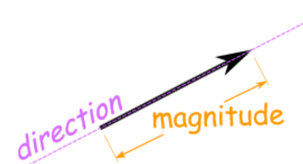
Definition A *vector* is a matrix with one column.
Example

$$\begin{bmatrix} 1 \\ 2 \\ -5 \\ 9 \end{bmatrix}$$

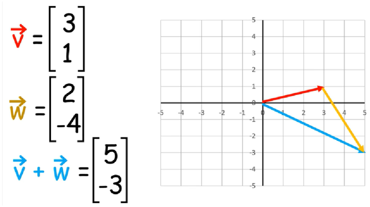
Note Two vectors are equal precisely when they have the same number of rows and all their corresponding entries are equal.

Vectors (Recap)

- A vector has magnitude (how long it is) and direction
- A point can be a vector (position vector, from Origin)
- A data row is a point in n-dimensions, thus a vector as well.



Vector Addition



- To add these vectors: We just add the individual components, so 3 plus 2 is 5; and 1 plus -4 is -3.
- It is simply adding another leg to the journey; so if we follow vector V along 3 and up 1, and then follow vector W along 2 and down 4, we end up at the head of the vector we calculated by adding V and W together.

Vector Addition

Definition We define the sum and of two vectors by

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$$

and the product of a scalar and a vector by

$$\alpha \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} \alpha u_1 \\ \alpha u_2 \\ \vdots \\ \alpha u_n \end{bmatrix}$$

Example

$$\begin{bmatrix} 1 \\ 3 \\ -5 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \\ 7 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix} \qquad \text{and} \qquad 3 \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 15 \\ 6 \\ 3 \end{bmatrix}$$

Exercise

Let $\vec{u}$ and $\vec{v}$ be given by

$$\vec{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \vec{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Plot $\vec{u}$, $\vec{v}$, $2\vec{u}$ and $\vec{u} + \vec{v}$.

Parallelogram rule for vector addition Suppose $\vec{u}$ and $\vec{v} \in \mathbb{R}^2$. Then $\vec{u} + \vec{v}$ corresponds to the fourth vertex of the parallelogram whose opposite vertex is $\vec{0}$ and whose other two vertices are $\vec{u}$ and $\vec{v}$.

Exercise

Let $\vec{u} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$. Display $\vec{u}$, $-2/3\vec{u}$, $\vec{v}$ and $-2/3\vec{u} + \vec{v}$ on a graph.

$\mathbb{R}^n$

In general we will consider vectors in $\mathbb{R}^n$, that is, having n real entries.

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^n$$

The zero vector is $\vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ having n entries, each equal to 0.

Properties of R^n

Theorem Suppose that $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ and $c, d \in \mathbb{R}$. Then,

- $\vec{u} + \vec{v} = \vec{v} + \vec{u}$.
- $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
- $\vec{u} + \vec{0} = \vec{0} + \vec{u} = \vec{u}$
- $\vec{u} + -\vec{u} = -\vec{u} + \vec{u} = \vec{0}$ $\qquad (-\vec{u} = (-1)\vec{u})$
- $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$
- $(c + d)\vec{u} = c\vec{u} + d\vec{u}$
- $c(d\vec{u}) = (cd)\vec{u}$
- $1 \cdot \vec{u} = \vec{u}$

Vector Multiplication

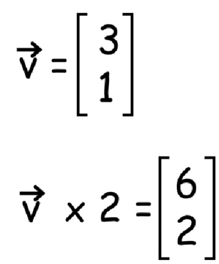
Vector Multiplication

Vector Multiplication is slightly complicated that plain Vector Addition. There are a few types of it.

- Scalar into Vector resulting in a vector: e.g. You have a list (a vector) of people's income. Tax rate is 15%. How do you get a list of Tax amounts?
- Vector into Vector resulting in a scalar: e.g. You have different amounts of foreign currencies. You know each ones conversion-to-INR rate. How do you compute total INRs you have?
- Vector into Vector resulting in a vector: e.g. Area of a parallelogram with a right hand rule direction.

Scalar Vector Multiplication

Scalar Vector Multiplication



Multiply each element of the vector by the scalar

Scalar Vector Multiplication

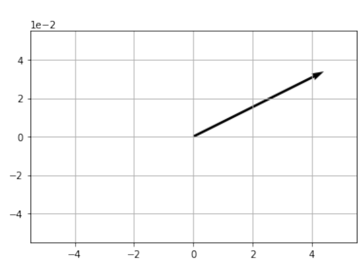
```
import numpy as np
import matplotlib.pyplot as plt
import math

v = np.array([2,1])

w = 2 * v
print(w)

# Plot w
origin = [0], [0]
plt.grid()
plt.ticklabel_format(style='sci', axis='both',
    scilimits=(0,0))
plt.quiver(*origin, *w, scale=10)
plt.show()
```

Scalar Vector Multiplication



Scalar Vector Multiplication

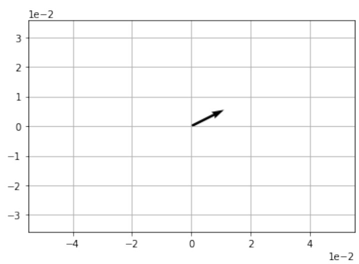
$\vec{b} = \frac{\vec{v}}{2}$

```
b = v / 2
print(b)

# Plot b
origin = [0], [0]
plt.axis('equal')
plt.grid()
plt.ticklabel_format(style='sci', axis='both',
    scilimits=(0,0))
plt.quiver(*origin, *b, scale=10)
plt.show()
```

[1. 0.5]

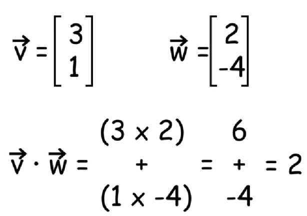
Scalar Vector Multiplication



Dot Product

Vector Vector Multiplication

Dot Product



Multiply the corresponding elements of the vectors and add the results In this case, 3 times 2 is 6, and 1 times -4 is -4; and adding these together gives us our scalar result of 2.

Vector Vector Multiplication

$\vec{v} \cdot \vec{s} = (v_1 \cdot s_1) + (v_2 \cdot s_2) ... + \ (v_n \cdot s_n)$

```
import numpy as np

v = np.array([2,1])
s = np.array([-3,2])
d = np.dot(v,s)
print (d)
```

-4

Vector Vector Multiplication

- Another form: $\vec{v} \cdot \vec{s} = \|\vec{v}\| \|\vec{s}\| \cos(\theta)$
- So for our vectors v (2,1) and s (-3,2), our calculation looks like this:
- $\cos(\theta) = \frac{(2 \cdot -3) + (-3 \cdot 2)}{\sqrt{2^2+1^2} \times \sqrt{-3^2+2^2}}$
- So $\cos(\theta) = -0.496138938357$
- $\theta \approx 119.74$

Angle Between Two Vectors

- Suppose we have two vectors $\vec{v} = (v, 0)$ lying on X axis and $\vec{w} = (w_1, w_2)$
- $w_1 = ||\vec{w}||cos\theta$, so $\theta = cos^{-1}(\frac{w_1}{||\vec{w}||})$
- Now, dot product is given as $\vec{v} \cdot \vec{w} = v_1.w_1 + 0.w_2 = v_1.w_1$
- Putting value of w_1, eqn becomes $= v_1.||\vec{w}||cos\theta = ||\vec{v}||||\vec{w}||cos\theta$
- Therefore: $cos\theta = \frac{\vec{v} \cdot \vec{w}}{||\vec{v}||||\vec{w}||}$
- Applicable to Higher Dimensions also!!

Definition

Suppose that $\vec{u}, \vec{v} \in \mathbb{R}^n$. We define the **inner product** or **dot procuct** or $\vec{u}$ and $\vec{v}$ as

$$u \cdot v = u^t v = \sum_{i=1}^{n} u_i v_i.$$

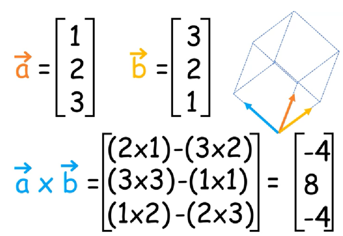
Example

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} = (1)(-1) + (2)(-2) + (3)(1) = -2.$$

Cross Product

Vector Vector Multiplication

Cross Product (for 3D vectors)



Skipping the current row and column, calculate determinant value of remaining sub matrix for that position.

Vector Vector Multiplication

Cross Product

- $\vec{p} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \quad \vec{q} = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$
-

$$r_1 = p_2 q_3 - p_3 q_2 \tag{1}$$
$$r_2 = p_3 q_1 - p_1 q_3 \tag{2}$$
$$r_3 = p_1 q_2 - p_2 q_1 \tag{3}$$

- $\vec{r} = \vec{p} \times \vec{q} = \begin{bmatrix} (3 \cdot -2) - (1 \cdot 2) \\ (1 \cdot 1) - (2 \cdot -2) \\ (2 \cdot 2) - (3 \cdot 1) \end{bmatrix} = \begin{bmatrix} -6 - 2 \\ 1 - -4 \\ 4 - 3 \end{bmatrix} = \begin{bmatrix} -8 \\ 5 \\ 1 \end{bmatrix}$

Vector Vector Multiplication

Cross Product

```
import numpy as np

p = np.array([2,3,1])
q = np.array([1,2,-2])
r = np.cross(p,q)
print (r)
```

[-8 5 1]

Matrix

Meaning of a Matrix

- Matrix is organization of data into rows and columns
- Example: columns can be various aspects of a person, such as height,weight, salary, etc, where as rows can represent different persons
- This Excel sheet like data can be thought of as a Matrix (especially in Data Science, Machine Learning)

Matrix

A matrix is an array of numbers that can be arranged into rows and columns. We generally name matrices with a capital letter.

$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$

```
 import numpy as np

A = np.array([[1,2,3],
              [4,5,6]])
print (A)

[[1 2 3]
 [4 5 6]]
```

Matrix

Definition A matrix with m rows and n columns is referred to as an $m \times n$ matrix. The number of rows always comes before the number of columns.

$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{bmatrix}$

```
import numpy as np

M = np.matrix([[1,2,3],
               [4,5,6]])
print (M)

[[1 2 3]
 [4 5 6]]
```

Matrix Addition

You can add or subtract matrices of the same size by simply adding or subtracting the corresponding elements in the two matrices.

$$A = \begin{bmatrix} 3 & 5 & 1 \\ 1 & 4 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 2 & -2 & 4 \\ -1 & 3 & 1 \end{bmatrix}$$

$$A + B = \begin{bmatrix} 5 & 3 & 5 \\ 0 & 7 & 4 \end{bmatrix}$$

$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 7 & 7 \\ 7 & 7 & 7 \end{bmatrix}$

Matrix Addition

```
import numpy as np

A = np.array([[1,2,3],
              [4,5,6]])
B = np.array([[6,5,4],
              [3,2,1]])
print(A + B)

[[7 7 7]
 [7 7 7]]
```

Matrix Subtraction

$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} - \begin{bmatrix} 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} -5 & -3 & -1 \\ 1 & 3 & 5 \end{bmatrix}$

```
import numpy as np

A = np.array([[1,2,3],
              [4,5,6]])
B = np.array([[6,5,4],
              [3,2,1]])
print(A - B)

[[-5 -3 -1]
 [ 1  3  5]]
```

The Transpose of a Matrix

Definition The transpose of a $m \times n$ matrix A is the matrix A^T having (i, j)-entry a_{ji}. That is,

$$(A^T)_{ij} = a_{ji}.$$

Example For example, $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ has transpose $A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$.

Note The rows of A become the columns of A^T and vice versa.

Meaning of a Matrix Multiplication

- Matrix is organization of data into rows and columns
- Example: columns can be various aspects of a person, such as height,weight, salary, etc, where as rows can represent different persons
- This Excel sheet like data can be thought of as a Matrix (especially in Data Science, Machine Learning)
- If you have another matrix like this, what is the meaning of their multiplication?
- Geometrically: say first matrix represents points of a shape, a polygon, where each row is a point, and each column represents X, Y, Z coordinates.
- Second matrix is typically a Homogeneous transformation matrix, such as rotation, when multiplied gets rotated shape.

Matrix Multiplication Rules

Theorem Let A and B be matrices whose sizes are appropriate for the following sums and products to be defined

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$.
- For any scalar r, $(rA)^T = rA^T$.
- $(AB)^T = B^T A^T$

Example

$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, and $B = \begin{bmatrix} 5 & 1 & -1 \\ 1 & 2 & 2 \end{bmatrix}$ then

$$AB = \begin{bmatrix} 7 & 5 & 3 \\ 9 & 11 & 5 \end{bmatrix} \qquad (AB)^T = \begin{bmatrix} 7 & 9 \\ 5 & 11 \\ 3 & 5 \end{bmatrix} = B^T A^T$$

but A^T is 2×2 and B^T is 3×2, so $A^T B^T$ isn't even defined.

Matrix Transpose

Exchange rows and columns

$$A = \begin{bmatrix} 3 & 5 & 1 \\ 1 & 4 & 3 \end{bmatrix}$$

$$A^T = \begin{bmatrix} 3 & 1 \\ 5 & 4 \\ 1 & 3 \end{bmatrix}$$

Matrix Transpose

$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$

```
import numpy as np

A = np.array([[1,2,3],
              [4,5,6]])
print(A.T)

[[1 4]
 [2 5]
 [3 6]]
```

Matrix Multiplication

Here are the cases to consider:

- Scalar multiplication, which is multiplying a matrix by a single number
- Element wise matrix multiplication (rarely used, called Hadamard multiplication, shown with circle instead of dot)
- Dot product matrix multiplication, or multiplying a matrix by another matrix.

Matrix Scalar Multiplication

To multiply a matrix by a scalar value, you just multiply each element by the scalar to produce a new matrix:

$2 \times \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{bmatrix}$

```
import numpy as np
import numpy as np

A = np.array([[1,2,3],
              [4,5,6]])
print(2 * A)

[[ 2  4  6]
 [ 8 10 12]]
```

Matrix Multiplication Defined

Definition If A is an $m \times n$ matrix, and if $B = [\vec{b}_1, \vec{b}_2 \ldots, \vec{b}_p]$ is a $n \times p$ matrix, then the matrix product AB is the following $m \times p$ matrix.

$$AB = [A\vec{b}_1 \quad A\vec{b}_2 \quad \ldots \quad A\vec{b}_p]$$

Example Let $A = \begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$ and let $B = \begin{bmatrix} 3 & -1 & 6 \\ 7 & 5 & 3 \end{bmatrix}$. Compute AB.

Multiplying Matrices

Row-Column Rule If A is $m \times n$ and if B is $n \times p$ the (i, j)-entry of AB is given by

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Note $\text{Row}_i(AB) = \text{Row}_i(A) \cdot B$.

Matrix Operations

Additions

- Commutative: $A + B = B + A$
- Associative: $A + (B + C) = (A + B) + C$

Multiplication

- Scalar : sA: multiplying all elements by s
- Commutative: $AB \neq BA$
- Associative: $A(BC) = (AB)C$
- Distributive: $A(B + C) = AB + AC$
- Identity: $I_m A_{mn} = A_{mn} I_n = A$