# Basic Linear Algebra

Antonio Prgomet
https://www.linkedin.com/in/antonioprgomet

## Background

In this short document, I will introduce basic Linear Algebra that is required in order to understand technical aspects within Artificial Intelligence (AI).

# 1 Vectors

A vector is either a row vector or a column vector. If nothing is said it is implicitly assumed that it is a column vector. By convention vectors are written with small bold letters (for instance $\boldsymbol{v}$) or small letters with an arrow above them (for instance $\vec{v}$).

**Example**
The column vector $\boldsymbol{v}$ is a $3 \times 1$ vector since it contains 3 rows and 1 column.

$$\boldsymbol{v} = \begin{bmatrix} 5 \\ 4 \\ 8 \end{bmatrix}$$

**Example**
The row vector $\boldsymbol{x}$ is a $1 \times 7$ vector since it contains 1 row and 7 column.

$$\boldsymbol{x} = \begin{bmatrix} 5 & 2 & 12 & 4 & 5 & 10 & 4 \end{bmatrix}$$

There are several applications of vectors. In Mathematics, a point in a a 3 dimensional space can be specified as a vector. For example the vector $\boldsymbol{v} = [5, 4, 8]$ could represent a point as seen from figure 1. In general we can also view higher dimensional vectors as a point in a n-dimensional space, but if $n > 3$ we cannot visualize it.

In applications within Statistics, Data Analysis and Artificial Intelligence vectors are often used to store data. The vector $\boldsymbol{x} = [5, 2, 12, 4, 5, 10, 4]$ could for example represent the number of ice creams a company sold each day during a week.

## 1.1 Operations on Vectors

The first example we are going to look at is the transpose operator. It simply flips a row vector to a column vector and vice versa.

**Example - Transpose.** A row vector $\boldsymbol{r} = [2, 1, 3, 5]$ transposed would be written in this way:

$$\mathbf{r} = [2, 1, 3, 5]^\top = \begin{bmatrix} 2 \\ 1 \\ 3 \\ 5 \end{bmatrix}$$

We will now show two examples where we multiply a vector with a scalar and where we do vector addition. Note that a scalar is simply an ordinary number.

**Example - Multiplying a vector with a scalar.**
Define the vector $\boldsymbol{a} = [4, 2, 12, 6]$

Then we can multiply a vector by a scalar.

$$\mathbf{3a} = \begin{bmatrix} 3 \times 4 \\ 3 \times 2 \\ 3 \times 12 \\ 3 \times 6 \end{bmatrix} = \begin{bmatrix} 12 \\ 6 \\ 36 \\ 18 \end{bmatrix}$$

**Example - Vector addition.**
Define the two vectors $\boldsymbol{a} = [4, 2, 12, 6]$ and $\boldsymbol{b} = [1, 1, 3, 5]$.
Then we can add the two vectors.

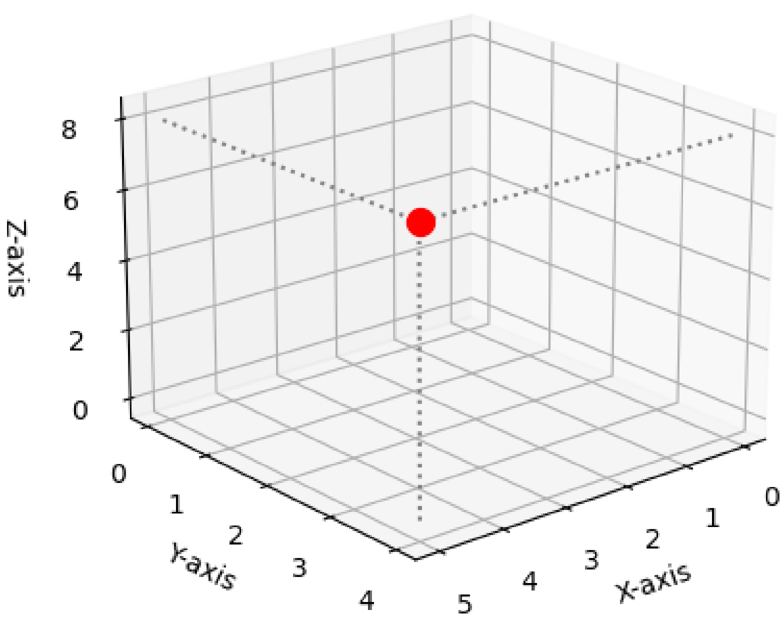


Figure 1: A vector representing the point $(5, 4, 8)$ in a 3 dimensional plane.

$$\mathbf{a}+\mathbf{b} = \begin{bmatrix} 4+1 \\ 2+1 \\ 12+3 \\ 6+5 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 15 \\ 11 \end{bmatrix}$$

**Example - Vector addition not defined.**
If we have a vector called $\boldsymbol{x_1} = [5, 1, 6]$ and $\boldsymbol{x_2} = [3, 3]$ then $\boldsymbol{x_1} + \boldsymbol{x_2}$ is not defined since $\boldsymbol{x_1}$ has the dimension $1 \times 3$ and $\boldsymbol{x_2}$ has the dimension $1 \times 2$. The dimensions are not the same hence vector addition is not defined in this case.

# 2 Matrices

An $m \times n$ matrix is a rectangular array of numbers with m rows and n columns. For example, a two-by-three matrix A, with two rows and three columns can look like:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

where for instance $a_{12}$ represents the number in row 1 and column 2.

The convention is that matrices are denoted by capital letters such as $A$ or $B$.

## 2.1 Operations on Matrices

**Example - Transpose.** When transposing a matrix, the rows becomes columns and vice versa.

$$A = \begin{bmatrix} 4 & 10 & 5 \\ 1 & 1 & 2 \end{bmatrix}^\top = \begin{bmatrix} 4 & 1 \\ 10 & 1 \\ 5 & 2 \end{bmatrix}$$

**Example - Matrix Addition.** When doing matrix addition, the matrices must have the same size in order for the operation to be defined.

In this example we are first going to multiply each matrix with a scalar and then add them.

$$2\begin{bmatrix} 3 & 4 \\ 5 & 5 \\ 2 & 1 \end{bmatrix} + 5\begin{bmatrix} -1 & 1 \\ 10 & 0 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 10 \\ 4 & 2 \end{bmatrix} + \begin{bmatrix} -5 & 5 \\ 50 & 0 \\ -25 & 10 \end{bmatrix} = \begin{bmatrix} 1 & 13 \\ 60 & 10 \\ -21 & 12 \end{bmatrix}$$

Now we are going to define matrix multiplication. At first sight it might seem like it is a strange definition, but as you will see soon, it makes it really easy to write large systems of equations in a compact way. This is very practical in some AI and Statistics applications such as Linear Regression (we will not go into that in this document).

**Example - Matrix Multiplication.**
Matrix multiplication involves multiplying each element of a row from the first matrix with the corresponding element of a column from the second matrix and summing up these products. The resulting matrix has dimensions determined by the number of rows of the first matrix and the number of columns of the second matrix. In general, if you multiply a matrix A of size $m \times n$ with matrix B of size $n \times p$ then the multiplication is defined because A has as many columns (n) as B has rows (n), the dimension of the matrix that is resulting of the multiplication is $m \times p$.

Let

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

and

$$B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

The product $AB$ is calculated as follows:

$$AB = \begin{bmatrix} (2 \times 5 + 3 \times 7) & (2 \times 6 + 3 \times 8) \\ (4 \times 5 + 1 \times 7) & (4 \times 6 + 1 \times 8) \end{bmatrix} = \begin{bmatrix} 31 & 36 \\ 27 & 32 \end{bmatrix}$$

### 2.1.1 Why do we define matrix multiplication as we do?

Consider the system of linear equations:

$$\begin{aligned} 2x + 3y &= 8 \\ 4x + y &= 7 \end{aligned}$$

We can express this system using matrix multiplication as $Ax = B$, where:

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}, \quad x = \begin{bmatrix} x \\ y \end{bmatrix}, \quad B = \begin{bmatrix} 8 \\ 7 \end{bmatrix}.$$

The system can then be written as:

$$\begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \end{bmatrix}.$$

If we denote the matrix product $Ax$ as $C$, the system becomes:

$$C = \begin{bmatrix} 2x + 3y \\ 4x + y \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \end{bmatrix}.$$

Putting the elements equal to each other, we get the original system of equations back:

$$\begin{aligned} 2x + 3y &= 8 \\ 4x + y &= 7 \end{aligned}$$

This example was quite simple, but you can imagine how helpful this is when you have many equations. The reason is that instead of having to write each equation separately (there can be millions of them), we simply define the matrix $A$ and the vectors $x$ and $b$. See the following example.

**Example.** Consider the system of linear equations:

$$\begin{aligned} 2x + 3y - z &= 5 \\ 4x - y + 2z &= -1 \\ x + 2y - 3z &= 8 \\ 3x - y + z &= 3 \\ 2x - 4y + 5z &= -6 \end{aligned}$$

We can express this system using matrix multiplication as $Ax = b$, where:

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 4 & -1 & 2 \\ 1 & 2 & -3 \\ 3 & -1 & 1 \\ 2 & -4 & 5 \end{bmatrix}, \quad x = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad b = \begin{bmatrix} 5 \\ -1 \\ 8 \\ 3 \\ -6 \end{bmatrix}.$$

# 3 Norm

We can generalize the notion of "distance" in higher dimensions between two points/vectors by defining what is called the norm.

The norm of a vector $\boldsymbol{u} = [u_1, u_2, u_3]$, is defined by:

$$\|\mathbf{u}\| = (u^T u)^{0.5} = \left(u_1^2 + u_2^2 + u_3^2\right)^{0.5},$$

.
The vector $\boldsymbol{u}$ can have arbitrary many elements.

**Example** Let $\boldsymbol{v} = [3, 4]$, then

$$\|\mathbf{v}\| = \left(3^2 + 4^2\right)^{0.5} = (9 + 16)^{0.5} = (25)^{0.5} = 5$$

.

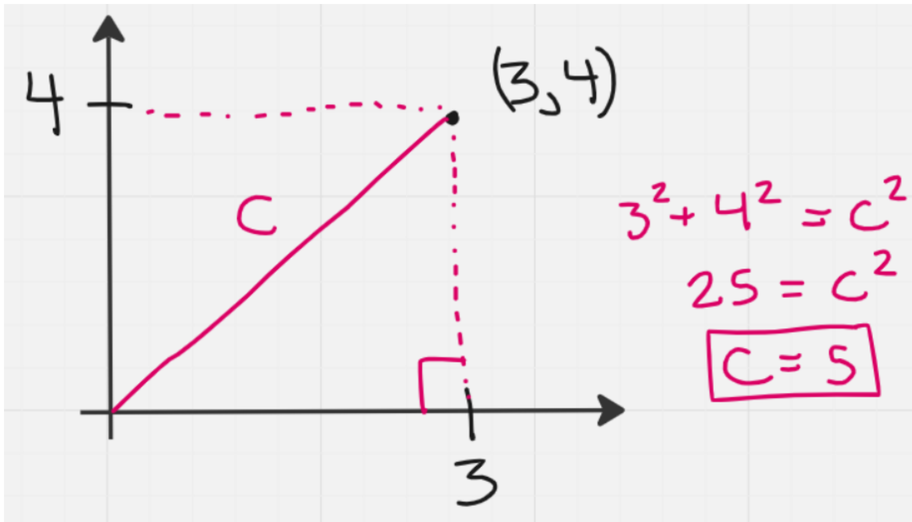


Figure 2: Calculating the distance from the point/vector $(3, 4)$ to origo $(0, 0)$.

So the norm is 5. We can interpret this as $\|\mathbf{v}\| = \|\mathbf{v} - \mathbf{0}\|$ which is saying that the distance from the point $\boldsymbol{v} = (3, 4)$ to origo $\mathbf{0} = (0, 0)$ is 5. This can easily be shown with the Pythagorean Theorem. Recall that the Pythagorean theorem says that for a right triangle with sides $a, b$ and hypotenuse $c$, the relationship $a^2 + b^2 = c^2$ holds. See figure 2, do you understand what the figure shows?

So in general the norm of the difference between two arbitrary, same sized vectors, $\mathbf{a}$ and $\mathbf{b}$, $\|\mathbf{a} - \mathbf{b}\|$ allows us to measure distance in higher dimensions. Why is this interesting? As we will se later (not in this document), when appplying models within Machine Learning which is an import subject within Artificial Intelligence, we often want to minimize the distance between numbers that we know are true and our model predictions. Probably this does not tell you much right now, the main message is that it is conceptually important and you will se that later on.

## 4 Exercises

**Question 1**
Define the matrices:

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

$$D = \begin{bmatrix} 3 & 4 \\ 4 & 3 \end{bmatrix}, \quad E = \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \quad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

If defined, do the computations.
a) $2A$
b) $B - 2A$
c) $3C - 2E$
d) AC
e) CD
f) CB
g) CI

**Question 2** Define the matrix

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

Calculate $AA^T$.

**Question 3**
Let

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 4 & 3 \\ 0 & 2 \end{bmatrix}.$$

Verify that $AB = AC$ and yet $B \neq C$. What I want to show with this example is that we have defined new objects (matrices in this case) and the usual rules for arithmetic does not necessarily hold. We will not go into further details about this.

**Question 4**
Define the vectors $\boldsymbol{v_1} = [4, 3, 1, 5]$ and $\boldsymbol{v_2} = [2, 3, 1, 1]$.
a) Calculate $\|\mathbf{v_1}\|$.
b) Calculate $\|\mathbf{v_1} - \mathbf{v_2}\|$.

**Question 5**
Write the system of linear equations below on the form $Ax = b$.

$$\begin{aligned} 3x_1 + 2x_2 + 4x_3 &= 7 \\ 2x_1 + 3x_2 + 8x_3 &= 4 \\ 4x_1 + x_2 + 3x_3 &= 11 \\ 7x_1 + x_2 + 5x_3 &= 9 \end{aligned}$$

Verify that your solution is correct by writing out the matrix multiplication of $Ax = b$ and check that you get the original system of equations back.

## Videos

You can watch the following videos that explains the concepts covered in this document:
1. Definition of a matrix: `https://www.youtube.com/watch?v=JhikgDtwpLM&feature=youtu.be`

2. Matrix addition and multiplication:
`https://www.youtube.com/watch?v=nSNebx6C5Vg`
`https://www.youtube.com/watch?v=MG7t6SWBnwA`

3. Transpose matrix: `https://www.youtube.com/watch?v=wwXCDY9-bAA`

3. Inner and outer products: `https://www.youtube.com/watch?v=wwXCDY9-bAA`