Linear algebra notes

Yoel Sanchez Araujo

September 20th 2020

1 Vectors

Before you read on, you should know that these notes will talk a lot about vectors and matrices, but linear algebra is much more than those objects.

To start, one way to think about a vector is as an ordered collection of things from some set. In particular, our set will be the set of real numbers. These are familiar to you, for example: 1.2, 0, -4.5, 6 , are all elements of the set of real numbers. Now we'll collect some of the real numbers and call that collection a vector.

Example:

$$(1, 2, -3, 2.3)$$

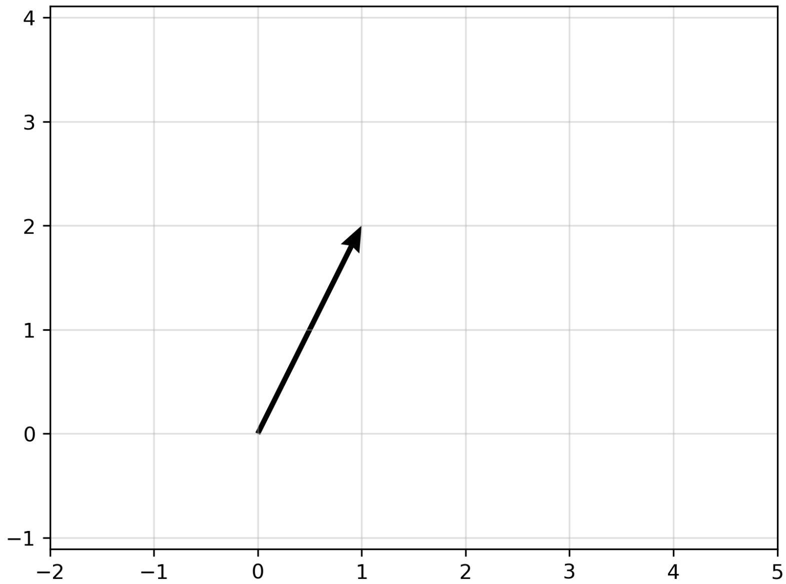
is a vector. We can also write that same vector as:

$$\begin{bmatrix} 1 \\ 2 \\ -3 \\ 2.3 \end{bmatrix}$$

For a more geometrically intuitive set of examples, we'll restrict ourselves for the moment, to only having at most 2 elements in our vectors. This will put us in the standard 2-D space, the Euclidean plane. First, think of the 2-D plane with the usual x and y coordinate axis. Now the vector:

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

represents one unit to the right, and two units up, in this plane starting at the point (0, 0):



1.1 Vector addition

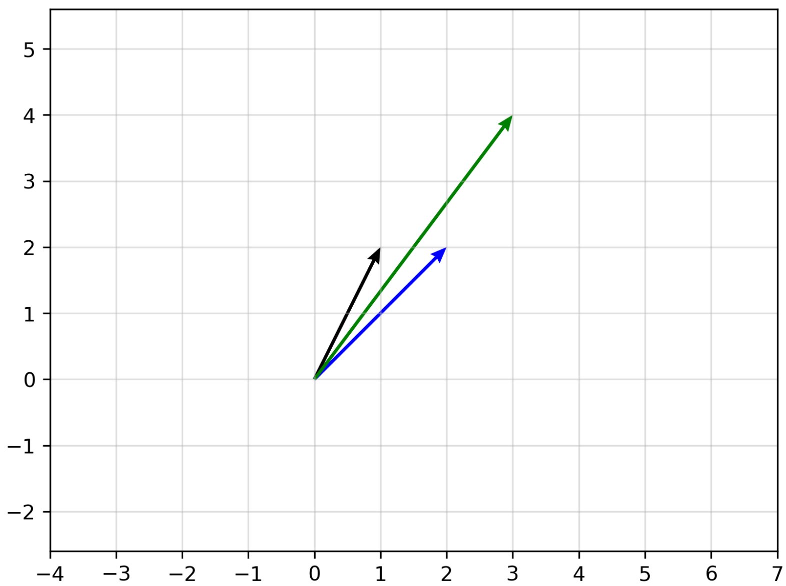
We are allowed to add vectors. To do this we add the individual elements at the same index between the two vectors. What does this imply about the lengths of both vectors? Example:

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

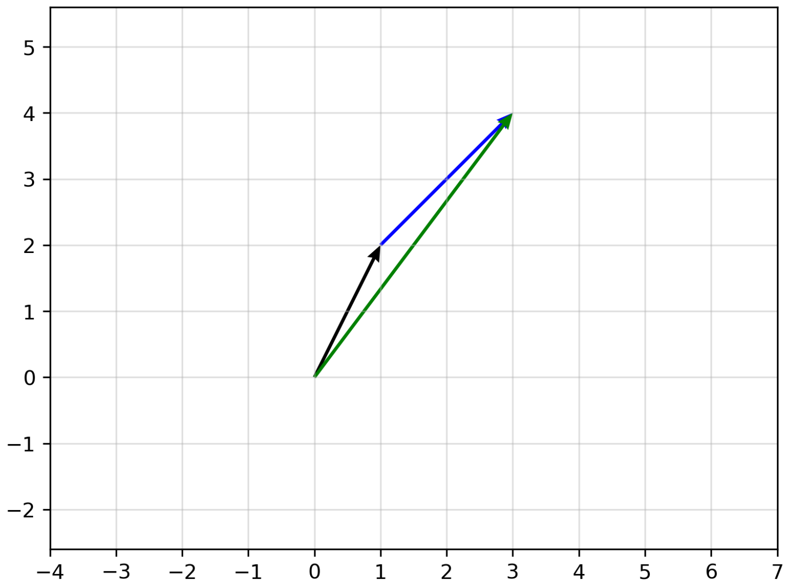
or more generally:

$$\begin{bmatrix}a\\b\end{bmatrix} + \begin{bmatrix}c\\d\end{bmatrix} = \begin{bmatrix}a+c\\b+d\end{bmatrix}$$

here's what the first numerical example of vector addition would look like on the 2D-plane:



The black arrow is again, the vector $[1, 2]$, the blue arrow is $[2, 2]$, and the green arrow is $[3, 4]$. Notice that if you start at zero, move one on the x-axis, then two up on the y-axis, you end up at the tip of the black arrow. Now from the tip of that black arrow move two to the right on the x-axis and two up on the y-axis and you end up at the tip of the green arrow. Said another way, if we put the blue and green arrows head-to-tail on the plot, we end up at the head of the green arrow. Here's the same plot but with the blue arrow's tail at the head of the black arrow:

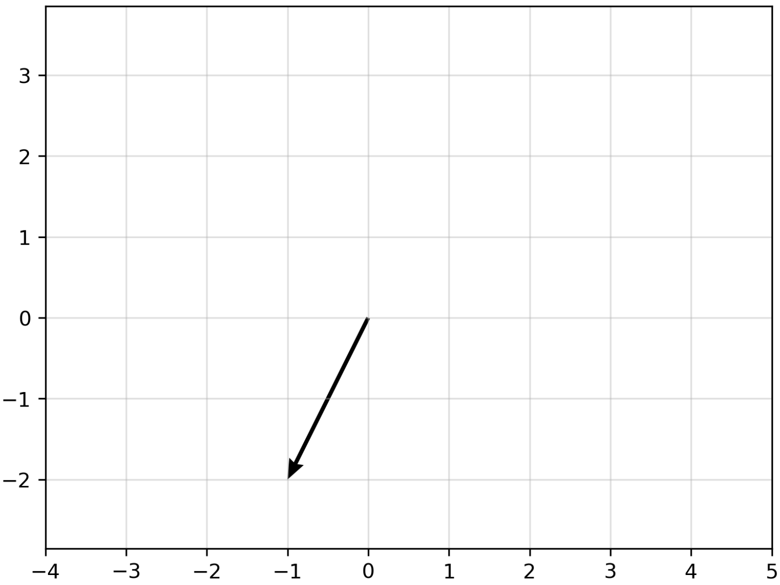


1.2 Vector - scalar multiplication

We are also allowed to multiply vectors by single numbers that we call scalars. For example:

$$\text{-1} \times \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

geometrically this flips the vector in the opposite direction:



we are also allowed to combine vector-scalar multiplication and vector addition. If we have two scalars, a and b, and two vectors of the same length $\mathbf{u}$ and $\mathbf{v}$ (we are using bold font to tell us that this variable is a vector) then we are allowed to do: $(a \times \mathbf{u}) + (b \times \mathbf{v})$. Often the multiplication symbol is omitted, so you'll see: $a\mathbf{u} + b\mathbf{v}$.

1.3 Properties and manipulations

Many of the algebraic properties that you're familiar with also apply to vectors. These are

- commutativity: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
- associativity: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
- additive identity: $\mathbf{0} + \mathbf{v} = \mathbf{v}$
- additive inverse: there is always a vector $\mathbf{w}$ such that $\mathbf{v} + \mathbf{w} = \mathbf{0}$
- multiplicative identity: $\mathbf{1v} = \mathbf{v}$
- distributive property: $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$ and $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$

1.4 Dot products

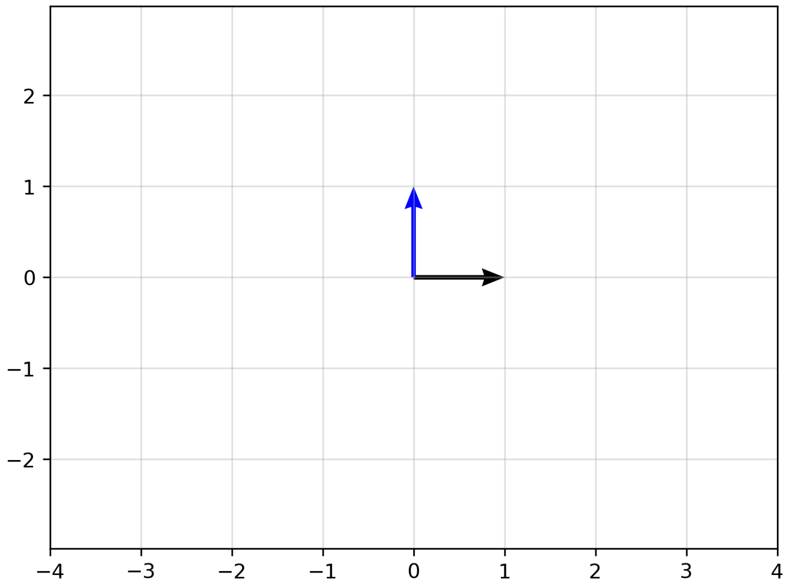
If we have two vectors $\mathbf{u} = [u_1, u_2, ..., u_n]$ and $\mathbf{v} = [v_1, v_2, ..., v_n]$, one operation defined on both of them is called the dot product, we will write it as $\mathbf{u^T v}$:

$$\mathbf{u^T v} = u_1 v_1 + u_2 v_2 + ... + u_n v_n$$

we will abbreviate this using the uppercase sigma symbol: $\sum$ and write it as:

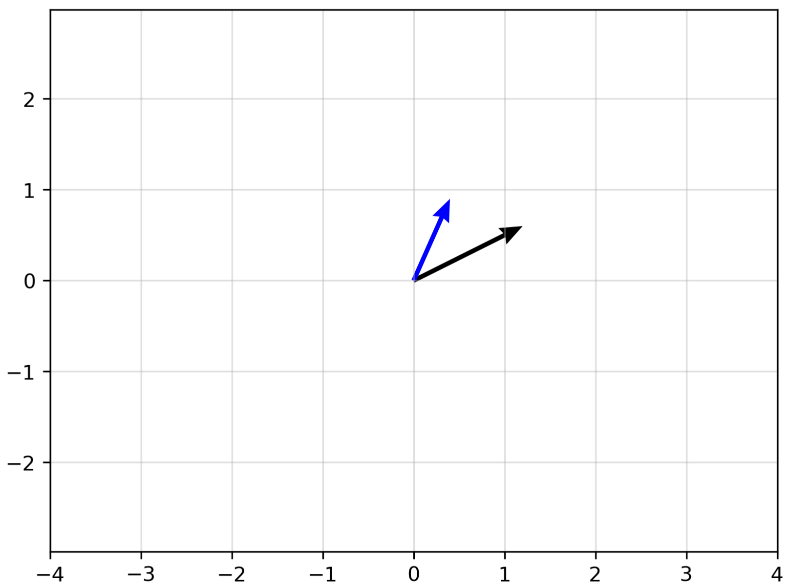
$$\mathbf{u^T v} = \sum_{k=1}^{n} u_k v_k$$

note that the result of a dot product is one single number: $\alpha = \mathbf{u^T v} = \sum_{k=1}^{n} u_k v_k$. Continuing with our geometric examples, what does a dot product tell us geometrically about two vectors? Say $\mathbf{u} = [1, 0]$ and $\mathbf{v} = [0, 1]$, then $\mathbf{u^T v} = 0$ as you should verify. Notice that $\mathbf{u}$ and $\mathbf{v}$ are perpendicular to each other. This is an important result, a dot product between two vectors will be zero only if they are perpendicular (often people will say orthogonal instead of perpendicular).

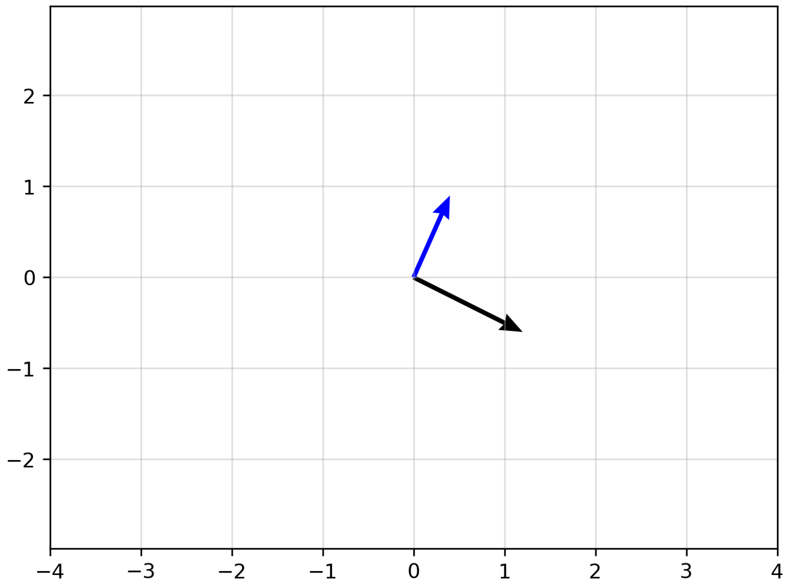


Imagine moving along the x-axis of the vector $\mathbf{u} = [1, 0]$, and now notice that no matter where you move along this vector, the vector $\mathbf{v} = [0, 1]$ is unchanged. In other words, the two vectors do not co-vary to any extent. Therefore, if two vectors are perpendicular, their dot product will be zero, and this will tell us, with a single number that they are neither similar nor dis-similar.

If we instead have the vectors $\mathbf{u} = [1.2, 0.4]$ and $\mathbf{v} = [0.6, 0.9]$, then $\mathbf{u}^{\mathbf{T}}\mathbf{v} = 1.08$. If we plot them:



now what about if $\mathbf{u} = [1.2, 0.4]$ and $\mathbf{v} = [-0.6, 0.9]$, then $\mathbf{u}^{\mathbf{T}}\mathbf{v} = -0.36$ and on the x-y plane we have:



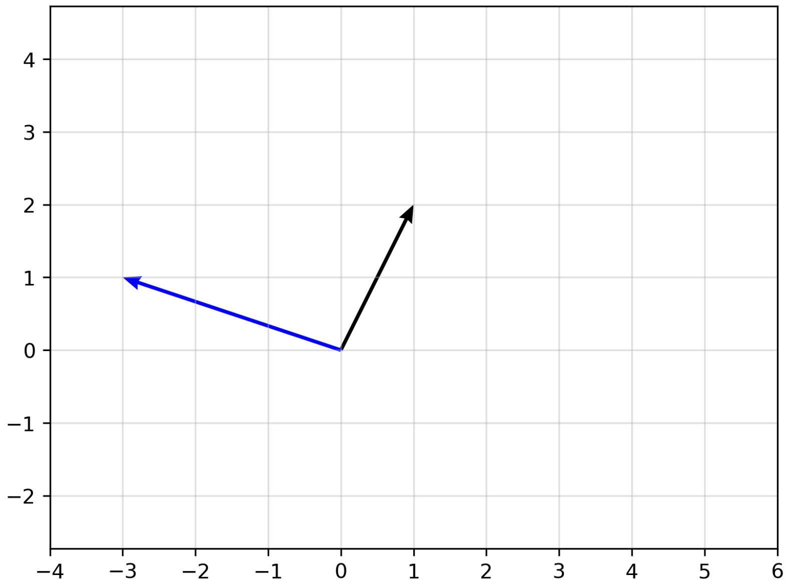
Notice that in the case of $\mathbf{u} = [1.2, 0.4]$ and $\mathbf{v} = [0.6, 0.9]$ as one vector increased in it's direction, so did the other and their dot product was positive. However for the case of $\mathbf{u} = [1.2, 0.4]$ and $\mathbf{v} = [-0.6, 0.9]$ as one increases, the other decreases, and their dot product is negative. Finally, to reiterate if they do not co-vary at all, their dot product is zero. Given this information, what can we say about what the dot product tells us between two vectors?

2 Matrices

So far we've only been working with one vector at a time. Now we'll have multiple vectors and have them form the columns of a matrix. Below we show the 2 by 2 matrix M

$$M = \begin{bmatrix} 1 & -3 \\ 2 & 1 \end{bmatrix}$$

notice that the first column of M can be seen as the vector $\mathbf{m_1} = [1, 2]$ and the second column can be seen as the vector $\mathbf{m_2} = [3, 1]$. The sub-script here references to the column number, instead of an element of the vector, as we previously saw in the vectors section. As we've been doing, we can plot the columns of this matrix:



This is important, if the two columns of the matrix M define different vectors, than by taking an infinite number of linear combinations of these vectors, we can get to any point on the 2-D plane. If the two columns are different vectors, then we say they are independent. By a combination of two vectors we mean using the rules of vector - scalar multiplication and vector addition that we defined in the sections above. Example:

$$0.2 \times \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1.4 \times \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

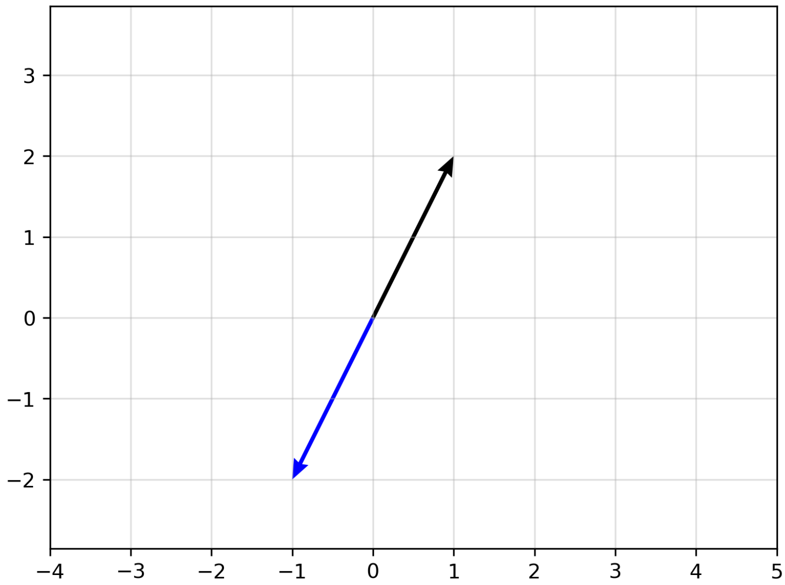
defines a linear combination of the two columns of the matrix M

2.1 Linear dependence

Say that now the matrix M is instead defined as:

$$M = \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix}$$

if we again plot the columns, we'll see something interesting:



if we imagine a line that goes through zero and stretches in both directions forever, we see that both columns of the new M matrix are on that line. In particular, one column is just negative one times the other column. As opposed to our previous example, we cannot fill the entire 2-D plane by taking linear combinations of these two columns, because all combinations will still be on that same line. This is what it means for the columns of a matrix to be linearly dependent. The concepts on matrices outlined so far are crucial to having a solid conceptual foundation for linear algebra. If you'd like to see a better explanation with animations see: *this video*

2.2 Matrix - vector multiplication

We can multiply vectors and matrices under certain conditions. In particular, the product is defined between a matrix of dimension N by P and a vector of length P. Where an N by P matrix is a matrix with N rows and P columns, and a vector of length P is a vector with P elements in it. Example, the matrix-vector product:

$$\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

works because the matrix has 2 columns, and the vector has 2 elements. The way in which the operations work is by taking a times the first column and adding it to b times the second column:

$$\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \times 1 + b \times 3 \\ a \times 2 + b \times 1 \end{bmatrix}$$

as you can see, the result of the matrix - vector product is a vector. Notice that we could have also written the above as:

$$\begin{bmatrix} a \times \begin{bmatrix} 1 \\ 2 \end{bmatrix} + b \times \begin{bmatrix} 3 \\ 1 \end{bmatrix} \end{bmatrix}$$

2.3 Matrix - matrix multiplication

Matrix - matrix multiplication has similar mechanics as matrix-vector multiplication. The only difference now, is that we have more than one vector that we are multiplying and each vector creates a new column in the result (which is now a matrix).

For two matrices M and C, the number of columns in M must match the number of rows in C. Why is this? Recall that matrix-vector multiplication ended up being a series of dot-products. For dot products to work, each vector must have the same number of elements. So, for matrix-matrix multiplication of MC we end up having a series of dot products between the columns of C and the rows of M. Example:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \left[\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e \\ g \end{bmatrix}, \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} f \\ h \end{bmatrix} \right]$$

We can see that the resulting matrix will be a 2 by 2 matrix.

2.4 Identity and permutation matrices

2.4.1 Identity matrices

There is a special matrix called the identity matrix that has the same role as the number 1 that we are use to in the algebra we learned in grade-school. This is called an identity matrix. It will always be a square matrix (same number of columsn and rows), and have ones along the main diagonal. Example:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

above, the sub-script 2 is telling us that it is the 2 by 2 identity matrix. Perform the following operation and convince yourself that any matrix times an appropriate identity matrix will result in the original matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

by using the identity matrix as teaching tool, we can see more clearly what matrix multiplication is doing. First, the operation:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

is saying "give me 1 times the first column plus 0 times the second column". This gives back just the first column. The following operation would give back just teh second column, and so the result is that you get back the original matrix when multiplying by an appropriate identity matrix.

2.4.2 Permutation matrices

A permutation matrix looks like an identity matrix in that it also only has the values 1 and 0 in it. However, the order of the columns or rows is permuted. The effect that this has is that it will swap either the rows or columns of the matrix it is multiplying, depending on if it on the right or left hand side. Example:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} b & a \\ d & c \end{bmatrix}$$

again, please note that we are using permutation matrices here to hammer home at the conceptual understanding of what matrix multiplication does, that will hopefully augment the mechanical operations of matrix multiplication.

2.5 Non-square matrix multiplication

So far, we've only used square 2 - by - 2 matrices because they are both easy to work with, and easy to plot for geometric examples. Recall however, that the condition for matrix multiplication to work is simply that the number of columns for the matrix on the left must match the number of rows for the matrix on the right. We'll now provide an example of non-square matrix multiplication:

$$\begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 4 & 1 & 2 & 0 \\ 1 & 5 & 3 & 7 \end{bmatrix}$$

$$= \left[\begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 7 \end{bmatrix} \right]$$

$$= \begin{bmatrix} 6 & 11 & 8 & 14 \\ 13 & 8 & 9 & 7 \\ 2 & 10 & 6 & 14 \end{bmatrix}$$

References

[1] S. Axler. Linear Algebra Done Right. Third Edition.