Chem237: Lecture 11

Alan Robledo

4/10/18

Linear Algebra

Just as we've seen with a Fourier Series, some topics can be better understood when we think about it in terms of Linear Algebra. Therefore, we will dive right into the topic just as our Math Methods book does in Chapter 6.

These next few lectures should not be seen as the equivalent of taking your first Linear Algebra course because such a course handles a topic as abstract as this one very carefully by making sure every definition is given and then proven. In mathematical methods, we will find ourselves using ideas that have been proven rigorously by mathematicians and simply treating these ideas as if it is all well known information. With that said, we will run through all the well known information first and then see how all of this information can be used to obtain solutions to problems of interest.

Vectors and Vector Spaces

In every standard calculus course you are told about objects known as vectors. And you learn that a vector is a mathematical object that is defined by its direction and length. So in 2 dimensions, you can have a whole space of vectors which could look something like this,
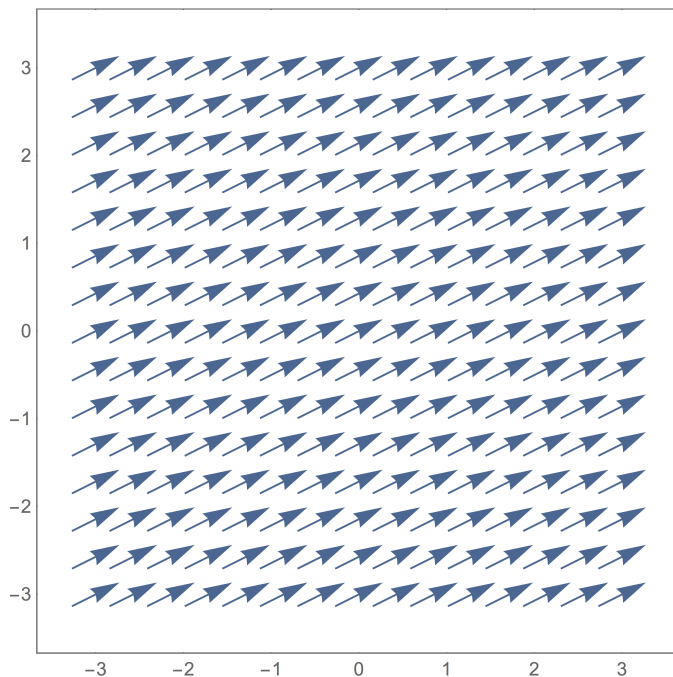

Figure 1: An example of a constant vector field $\vec{F}(x, y) = 2\hat{i} + \hat{j}$.

and we call this set of objects (vectors) a linear vector space.

Each vector space is said to be closed under addition and multiplication. In other words, for every vector $\vec{u}$ and $\vec{v}$ there exists another vector $\vec{u} + \vec{v}$, and for every scalar a and vector $\vec{u}$ there exists a vector $a\vec{u}$. This means that vectors

maintain properties such as the

$$
\begin{aligned}
\text{Commutative property:} \quad & \vec{u} + \vec{v} = \vec{v} + \vec{u} \\
\text{Associative property:} \quad & (\vec{u} + \vec{v}) + \vec{r} = \vec{u} + (\vec{v} + \vec{r}) \\
\text{Distributive scalar property:} \quad & a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v} \\
\text{Distributive vector property:} \quad & (a + b)\vec{u} = a\vec{u} + b\vec{u}
\end{aligned}
\tag{1}
$$

We also always assume that for any vector $\vec{u}$ there exists a null vector $\vec{0}$ such that,

$$
\vec{u} + \vec{0} = \vec{u}
\tag{2}
$$

and an additive inverse vector $-\vec{u}$ such that,

$$
\vec{u} + (-\vec{u}) = \vec{u} - \vec{u} = \vec{0}
\tag{3}
$$

These are some of the main properties of vector spaces that are important to know for solving problems that require the use of vectors. While figure 1 shows how we can represent vectors pictorally, we can represent vectors numerically as a set of numbers

$$
\vec{u} =< u_1, u_2, \ldots, u_n >
\tag{4}
$$

where n is the dimension of the vector space in question. These can be thought of as the coordinates of a vector. Some people use ( ) instead of $<>$ when writing vectors but either is acceptable, regardless of what anyone says. Just be consistent.

A vector can also be defined as either a column vector

$$
\vec{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}
$$

or a row vector

$$
\vec{u} = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}
$$

When we have a vector's coordinates, we can find out the length of the vector by finding its magnitude. Notation-wise, the length of a vector $\vec{u}$ is denoted as $|\vec{u}|$ and this length is often reffered to as just u, which is the name of the vector without the arrow over it. We can also define $|\vec{u}|$ as the norm of a vector $\vec{u}$, which is commonly given by the definition of the L2-norm

$$
|\vec{u}| = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}
\tag{5}
$$

In order to assign a value to a vector's length, we have to impose a **coordinate system/axes** and make use of the concept of **projection**. This concept was explained briefly in lecture 8 when we discussed Fourier Series/Transform.

Before we go into this, we should first talk about linear operators that operate on vectors, aka matrices.

Matrices

If we think of vectors as an array of numbers with size n, then matrices can be thought of as an array of arrays with size MxN, where M denotes the number of rows and N denotes the number of columns. So a matrix A of size MxN would be written as:

$$
\mathbf{A} = \begin{bmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & & \vdots \\ A_{M1} & \cdots & A_{MN} \end{bmatrix}
$$

Matrices aren't meant to be visualized the way vectors are. But matrices have properties that are important to know which are similar to vectors such as the

$$
\begin{aligned}
\text{Commutative property:} \quad & \mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \\
\text{Associative property:} \quad & (\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) \\
\text{Distributive scalar property:} \quad & a(\mathbf{A} + \mathbf{B}) = a\mathbf{A} + a\mathbf{B} \\
\text{Distributive matrix property:} \quad & (a + b)\mathbf{A} = a\mathbf{A} + b\mathbf{A}
\end{aligned}
\tag{6}
$$

If we had two matrices $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{A} \in \mathbb{C}^{MxN}$ and $\mathbf{B} \in \mathbb{C}^{NxK}$ we could define the multiplication of the two matrices as

$$\mathbf{AB} = \mathbf{C} \tag{7}$$

where $\mathbf{C}$ is a matrix such that $\mathbf{C} \in \mathbb{C}^{MxK}$. Therefore, the only way for the multiplication to work is if the number of columns of the left matrix is equal to the number of rows of the right matrix. A useful way to write the multiplication is with summation notation,

$$\sum_{n=1}^{N} A_{mn} B_{nk} = C_{mk} \tag{8}$$

where we sum over all the possible values of n. Notice how we used lowercase letters for m and k because we are saying that m can take on values from $1 \ldots M$ and k can take on values from $1 \ldots K$. Equation 8 is more preffered when you're working out a problem and need to keep track of your variables, especially when you are coding the multiplication of two matrices. Therefore, it might help to see what is actually going on in with the sum. Equation 8 says that

$$A_{m1} B_{1k} + A_{m2} B_{2k} + \ldots A_{m(N-1)} B_{(N-1)k} + A_{mN} B_{Nk} = C_{mk} \tag{9}$$

which means that for one element of the matrix C say the mk element, we need to single out the mth row of matrix A and the kth column of matrix B, multiply the corresponding entries together, and then add up the products.

$$\begin{bmatrix} \vdots & \vdots & & \vdots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{m(N-1)} & AmN \\ \vdots & \vdots & & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \cdots & B_{1k} & \cdots \\ \cdots & B_{2k} & \cdots \\ & \vdots & \\ \cdots & B_{(N-1)k} & \cdots \\ \cdots & B_{Nk} & \cdots \end{bmatrix} = \begin{bmatrix} & \cdots & \\ \vdots & C_{mk} & \vdots \\ & \cdots & \end{bmatrix}$$

We can also multiply matrices with vectors and the summation would look very similar. If we had the same matrix $\mathbf{A}$ with an M-dimensional vector $\vec{u}$, their multiplication would give a new vector $\vec{v}$

$$v_m = \sum_{n=1}^{N} A_{mn} u_n \tag{10}$$

Notice that the equation above tells us that we could find the mth coordinate of our new vector by taking the mth row of $\mathbf{A}$ and multiplying the corresponding values with the coordinates of $\vec{u}$.

A common mistake that people make is thinking of multiplying matrices the same way as multiplying scalars. If you have two matrices $\mathbf{A}$ and $\mathbf{B}$ such that,

$$\mathbf{AB} = \mathbf{C} \tag{11}$$

then you could solve for $\mathbf{B}$ by left multiplying both sides of the equation by the inverse of $\mathbf{A}$,

$$\mathbf{A}^{-1}\mathbf{AB} = \mathbf{A}^{-1}\mathbf{C} \quad \Rightarrow \quad \mathbf{B} = \mathbf{A}^{-1}\mathbf{C} \tag{12}$$

because of the property $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ and because the concept of dividing by matrices does not exist. Similarly, we could solve for $\mathbf{A}$ by right multiplying by the inverse of $\mathbf{B}$,

$$\mathbf{ABB}^{-1} = \mathbf{CB}^{-1} \quad \Rightarrow \quad \mathbf{A} = \mathbf{CB}^{-1} \tag{13}$$

Back to matrix properties, there exists a matrix $\mathbf{I}$ called a unit matrix or identity matrix such that for any matrix $\mathbf{A}$,

$$\mathbf{IA} = \mathbf{AI} = \mathbf{A} \tag{14}$$

which is often thought of as multiplying by one for matrices. Therefore, the identity matrix is a diagonal matrix of ones

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ldots \tag{15}$$

There also exists a matrix $\mathbf{0}$ called the null matrix such that each element of the matrix is zero. Therefore,

$$\mathbf{A0} = \mathbf{0A} = \mathbf{0} \quad \text{and} \quad \mathbf{A} + \mathbf{0} = \mathbf{A}. \tag{16}$$

Using these tools, we can discuss the concept of a matrix being associated with a linear transformation.

Linear Transformation

A linear transformation T is when we create a mapping from one vector space to another using a matrix, i.e. $T : \mathbb{C}^N \to \mathbb{C}^M$. Computing the transformation is as easy as multiplying any two matrices, because that's exactly what it is.

For example, consider two vector spaces in $\mathbb{R}^2$. We can take any vector and rotate it counterclockwise by some degree θ by using the rotation matrix $\mathbf{R}_\theta$

$$\mathbf{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{17}$$

If we wanted to visualize this, consider a vector $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. We can rotate the vector 45 degrees $= \frac{\pi}{4}$ radians by multiplying it with the rotation matrix $\mathbf{R}_{\frac{\pi}{4}}$

$$\mathbf{R}_{\frac{\pi}{4}} \vec{v}_1 = \begin{bmatrix} \cos\left(\frac{\pi}{4}\right) & -\sin\left(\frac{\pi}{4}\right) \\ \sin\left(\frac{\pi}{4}\right) & \cos\left(\frac{\pi}{4}\right) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \vec{v}_2 \tag{18}$$

and we can verify that the rotation was successful with a plot of $\vec{v}_1$ and $\vec{v}_2$.
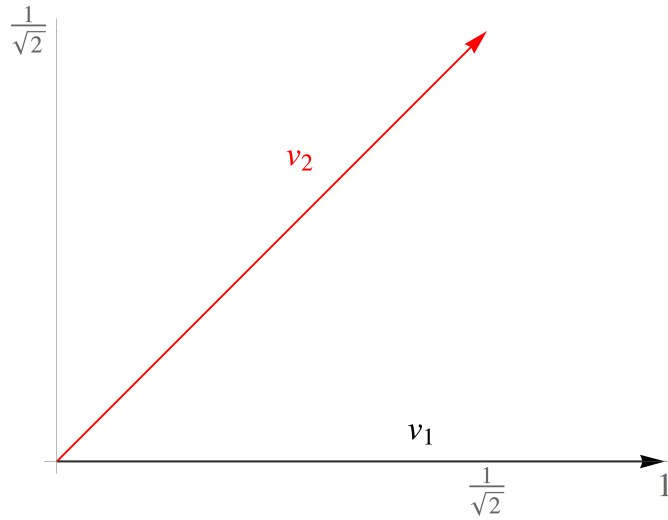

Figure 2: $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ being rotated 45 degrees counterclockwise to produce $\vec{v}_2$ using the rotation matrix.

The rotation matrix is a special matrix because it has the following properties:

1. $\mathbf{R}_\theta \mathbf{R}_{\theta'} = \mathbf{R}_{\theta'} \mathbf{R}_\theta$

2. $\mathbf{R}_\theta \mathbf{R}_{\theta'} = \mathbf{R}_{\theta+\theta'}$

3. $\mathbf{R}_\theta \mathbf{R}_{-\theta} = \mathbf{R}_0 = \mathbf{I}$

4. $\mathbf{R}_\theta^T = \mathbf{R}_\theta^{-1}$

and property number 4 is why we call the rotation matrix an **orthognal matrix**. The left hand side of property 4 is called the transpose matrix. Transposing a matrix means taking the columns of a matrix and turning them into rows. For example,

$$\mathbf{R}_\theta^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}. \tag{19}$$

So in general, for any matrix $\mathbf{A}$,

$$A_{mn}^T = A_{nm}. \tag{20}$$

Orthogonal matrices are convenient to work with because finding the inverse of the matrix is much more difficult than finding its transpose. In some cases, the inverse of a matrix may not exist. To determine the existence of an inverse, it would be easier to first compute the determinant of the matrix, which we will discuss next lecture.