In lieu of introduction

Linear algebra is a vast subject. At this early stage I cannot tell you what it is. However, I can tell you what it is about: it is about finite-dimensional vector spaces and linear transformations acting on those spaces. At least that is the subject's core.

We will formally define vector spaces and linear transformations in Sections 1 and 2. Upon first reading, you will, most likely, find these definitions to be mind-boggling abstractions. That is because they *are* mind-boggling abstractions. And, as the course syllabus explains, that is precisely what you are here to learn. There is a popular misconception that a good mathematics instructor can somehow eliminate the abstract and make a subject like linear algebra easy and accessible. That is like saying that a good basketball coach can make the game easy by eliminating running and jumping. Linear algebra, both pure and applied, derives its power from abstraction. Remove the abstract and the subject loses its meaning.

If you find the following sections difficult to read, embrace the challenge. That *is* the game. Accept that the first few homework assignments may seem almost impossible. There may be many reasons for that and we can discuss them during office hours, if you wish. Remember that failing, repeatedly, is normal in a class like Math 145. In fact, that is how it *should be.* You are a student. So, by definition, you are here to make mistakes and learn from them, with some help from me. Be patient with your progress and show some resilience in the face of adversity. As in basketball, or any sport for that matter, one must play the linear algebra game hard for the results to be worthwhile. So do that, and you may even have some fun in the process.

1 Fundamental objects of linear algebra

Every mathematical subject has its main objects of study. In Euclidean geometry the main objects are points, lines, and planes. In Calculus, they are real-valued functions of one or more variables. The main objects in linear algebra are *vector spaces*. These are not the only objects, however. In order to define vector spaces, we first need to define *scalars* which we proceed to do.

1.1 Scalars

In principle, the word 'scalar' is a synonym for the word 'number'. So, scalars are numbers. At this point you may wonder why we suddenly need a synonym for 'numbers'. There are two reasons for that. Firstly, the word 'scalar' is derived from the verb 'to scale' which suggests an operation. Indeed, scalars are numbers *that scale vectors*. More on that later. Secondly, the use of the term 'scalar' allows us to abstract from the particulars of various numbers. We shall have occasion to work with real and complex numbers and, perhaps, with more exotic number systems. Using 'scalars' instead of, say, 'reals' will allow us to formulate results of greater scope and utility.

We now arrive at the key question:

> Just what are numbers?

Think about this for a few moments before turning the page.

When I ask this question in class—What are numbers?—the usual response is that numbers are symbols used to represent quantities. It is certainly true that numbers can be used to represent quantities. Yet, that is not the question! The question is: What *are* numbers? What numbers can represent is a wholly different story.

To properly answer this question and similar questions about vectors, linear transformations, matrices, and so on, think about the *algebraic operations* that you can perform on numbers. Some reflection should convince you that, at their core, numbers are algebraic objects that can be added and multiplied; subtraction and division are special cases of addition and multiplication. Think about it! From the purely mathematical point of view, it does not matter how numbers look or what they may represent. What matters is arithmetics. Remove addition and multiplication and numbers become mere symbols, like letters of the alphabet. Trust me, you do not want a number system that is just an alphabet.

And that is not all! We cannot conclude that scalars, being numbers, are simply things that can be added and multiplied. Scalar addition and multiplication must conform to a rigid set of rules which, so far, you have been taking for granted.

Henceforth I will use lower case Greek letters α, β, γ, and so on, to denote scalars. The following axioms of scalar addition must hold.

(A1) addition is commutative, $\alpha + \beta = \beta + \alpha$,

(A2) addition is associative, $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$,

(A3) there must be a unique scalar 0, called *zero*, such that $\alpha + 0 = \alpha$ for every α,

(A4) to every scalar α corresponds a unique scalar $-\alpha$ such that $\alpha + (-\alpha) = 0$.

The axioms of multiplication are similar and what you would expect.

(M1) multiplication is commutative, $\alpha\,\beta = \beta\,\alpha$,

(M2) multiplication is associative, $(\alpha\,\beta)\,\gamma = \alpha\,(\beta\,\gamma)$,

(M3) there must be a unique scalar 1, called *one*, such that $\alpha\,1 = \alpha$ for every α,

(M4) to every nonzero scalar α corresponds a unique scalar α^{-1} such that $\alpha\,\alpha^{-1} = 1$.

Finally, the two scalar operations must be intertwined with a distributive law.

(D1) multiplication is distributive with respect to addition, that is: $\alpha\,(\beta + \gamma) = \alpha\,\beta + \alpha\,\gamma$.

We now have a formal definition of what is called a scalar or number *field*: it is a collection of objects that can be added and multiplied so that the above axioms hold. For extra emphasis, I would like to repeat that for the most part it will not matter to us how scalars look or what they represent. A scalar may look like this: 3.1416; or this: `400921fb54442d18`; or this: π. It may represent the ratio of circumference of a circle to its diameter, a one hundred and eighty degree angle, or nothing at all. What matters are the algebraic operations, not the objects themselves. This dictum applies to all objects in linear algebra, big and small.

1.2 Vectors

If you read the previous section on scalars carefully enough, then you already know how to define vectors which we will label using lower case Latin letters x, y, z, and so on. By the algebraic operations, of course! Yet, which ones? This is where things can get murky, especially if you learned about vectors in Calculus and/or Physics.

You most likely know that vectors can be added, scaled, dotted, and crossed. However only the first two of these operations are the defining ones. At the most elemental linear algebra level, vectors are abstract objects than can be added and scaled (by scalars); there is nothing else that can be done with vectors by default. The axioms of vector addition and scalar multiplications are listed below. You will hardly find them surprising.

To every pair, x and y, of vectors corresponds a vector $x + y$, called the *sum* of x and y, in such a way that

(A1) addition is commutative, $x + y = y + x$,

(A2) addition is associative, $(x + y) + z = x + (y + z)$,

(A3) there must be a unique vector 0, called *zero*, such that $x+0 = x$ for every x,

(A4) to every vector x corresponds a unique vector $-x$ such that $x + (-x) = 0$.

To every scalar α and vector x corresponds a vector $\alpha\,x$, called *scalar product*, in such a way that

(S1) multiplication by scalars is associative $\alpha\,(\beta\,x) = (\alpha\,\beta)\,x$,

(S2) $1\,x = x$ for every x.

Finally, the following two distributive laws must hold.

(D1) multiplication by scalars is distributive with respect to vector addition, $\alpha\,(x + y) = \alpha\,x + \alpha\,y$.

(D2) multiplication by vectors is distributive with respect to scalar addition, $(\alpha + \beta)\,x = \alpha\,x + \beta\,x$.

Notice that vector addition is indistinguishable from scalar addition. This may strike you as odd if you think of the former in terms of the seemingly very special parallelogram law. Bear in mind, however, that we are discussing vector *algebra* at the moment. From the algebraic point of view, there is no distinction between scalar and vector addition. In fact, scalars are vectors!

The converse to the last statement is not true. Vectors are not scalars because they cannot be multiplied! This must sound especially strange after Multivariate Calculus where not one but *two* vector products are introduced: the 'dot' and the 'cross'. Unfortunately, both are misnomers. The dot product is an operation that takes two vectors and outputs a *scalar*. It is used to endow vectors with geometric properties, such as length, and has nothing to do with vector algebra. The cross product is, arguably, closer to being a form of vector multiplication since it outputs vectors. However, it is neither commutative nor associative, and it only exists in dimension three. The cross product is very useful in Physics for describing rotational motion and in some geometric problems in three-space. Otherwise, its utility is very limited.

I would like to conclude this section by belaboring a point. At the moment it does not matter how vectors look, what they may represent, and what attributes they may have. What matters first and foremost is vector algebra which, by definition, has only *two* operations: vector addition and scalar multiplication. Later in the course we will dot vectors and maybe even cross them; we will manipulate vector *components*; we will draw vectors as arrows, just like you did in other classes. That will come *later*, after much thought and discussion. In the meantime, keep your focus on the algebraic nature of vectors. In order to truly understand constructs like the dot product, you need to first appreciate how much can be done with vector algebra alone.

1.3 Vector Spaces

We are now ready to define the fundamental objects of linear algebra.

Definition 1 (Vector Space)**.** *A vector space is a collection of vectors that is closed under vector addition and scalar multiplication.*

Say I have a collection of vectors V. It is said to be closed under vector operations if for any scalar α and any two vectors x and y in V the scalar product $\alpha\, x$ and the (vector) sum $x + y$ are also in V. In that case, and only in that case, V is called a vector space.

As a familiar example, take all vectors in the plane; the scalars are real numbers. This vector space is called $\mathbb{R}^2$ and we will write its elements as *column* vectors. In our linear algebra notation the vector operations in the plane will look like this:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}, \quad \alpha \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \alpha\, x_1 \\ \alpha\, x_2 \end{bmatrix}.$$

The reason for column notation will become clear once we get to matrix-vector multiplication. In the meantime, convince yourself that as long as $\mathbb{R}^2$ contains *all* of its vectors, it is closed under vector addition and scalar multiplication and is therefore a vector space. The same goes for $\mathbb{R}^n$ with any n.

As a counterexample, take V to be the collection of all vectors in $\mathbb{R}^2$ *except for the zero vector*. Let x be some vector in V: in the future, we will write that as $x \in V$. Then $x + (-x) = 0 \notin V$ (the sign $\notin$ means 'is not an element of'). So, V is *not* a vector space because it is not closed under vector addition. Convince yourself that V is not closed under scalar multiplication either.

Closure under vector operations has a number of consequences that are far from trivial. Say V is a vector space over the real numbers—a *real* vector space. How many vectors does V have? Owing to closure, there is a dichotomy: either one or infinitely many. In the first case V consists of just the zero vector and is therefore a *trivial* vector space. By definition, a nontrivial vector space must have at least one nonzero vector. Yet then it must have all scalar multiples of that vector of which there are infinitely many insofar as there are infinitely many scalars.

Another consequence of closure is *convexity* which is important in optimization. A convex combination of two vectors x and y has the form: $\alpha\, x + (1 - \alpha)\, y$ where $0 \le \alpha \le 1$. It is easy to see that a real vector space is always convex because it contains all convex combinations of its vectors.

2 Linearity

Most questions in applied linear algebra concern *linear transformations* acting on vector spaces rather than the vector spaces themselves. In order to define a linear transformation, we first need to introduce *linear combinations*.

2.1 Linear combinations

Given any collection of vectors $\{x_i\}_{i=1}^n$ in a vector space V and a sequence of scalars $\{\alpha_i\}_{i=1}^n$ in the underlying field, one can define a new vector $y \in V$ as a *linear combination*:

$$y = \alpha_1 \, x_1 + \ldots + \alpha_n \, x_n = \sum_{i=1}^{n} \alpha_i \, x_i.$$

Although we defined vectors using two operations—addition and scalar multiplication—from now on think of linear combinations as *the* fundamental vector operations. It should be clear that both scalar multiplication and vector addition are particular linear combinations.

2.2 Linear transformations

Suppose that we have a "function" T which takes as inputs vectors in some vector space V and produces as outputs vectors from another vector space W. Since functions are typically scalar-valued, the proper term for T is 'transformation' or 'mapping' and the symbolic notation is $T : V \mapsto W$.

Definition 2 (Linear transformation)**.** *We shall say that T is a* linear *transformation if it maps linear combinations of inputs into linear combinations of outputs. In symbols:*

$$T\left(\sum_{i=1}^{n} \alpha_i \, x_i\right) = \sum_{i=1}^{n} \alpha_i \, T(x_i).$$

A linear transformation that maps a vector space *into itself*, is often called a *linear operator*. Thus $T : V \mapsto V$ denotes a linear operator on V. By the way, in this course we will work almost exclusively with linear transformations and operators, so the adjective 'linear' will be often omitted.

As an example of a linear transformation, consider $D : \mathbb{R}^3 \mapsto \mathbb{R}^2$ defined by:

$$D : \begin{bmatrix} a \\ b \\ c \end{bmatrix} \mapsto \begin{bmatrix} 2\,a \\ b \end{bmatrix}$$

As a simple exercise, verify that D is indeed linear. If we interpret vectors in $\mathbb{R}^n$ as coefficients of polynomials listed in descending order, e.g.,

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \equiv a\,x^2 + b\,x + c$$

then D is nothing but the first derivative acting on the space of quadratics. We could also define the derivative as an operator on $\mathbb{R}^3$:

$$D' : \begin{bmatrix} a \\ b \\ c \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 2\,a \\ b \end{bmatrix}$$

Although both D and D' can be interpreted as derivatives of quadratics, they are, technically, different linear transformations because the target spaces are different.

As a counterexample, consider $f : \mathbb{R}^1 \mapsto \mathbb{R}^1$ defined by $f(x) = m\,x + b$. In Calculus, f is called 'linear function' because its graph is a straight line. However, considered as an operator on $\mathbb{R}^1$, f is nonlinear! Indeed, $f(x + y) \neq f(x) + f(y)$. Confusing? The homework should help you sort some of this out.

Homework

Some of the exercises in this homework provide practice with the new language. Others probe your prerequisite knowledge, as well as your writing and programming skills. In all cases, work with awareness after having studied the handout. Read the problems carefully, follow the directions and provide not just answers but complete *solutions*. Remember that I am not asking you these questions because I cannot answer them myself. Rather I need to make sure that you can think logically, communicate abstract ideas, and make sense of challenging mathematics.

1. Let S consist of all ordered pairs of real numbers or, equivalently, points in the plane. Define addition and multiplication on S componentwise. That is:

   $$(\alpha_1, \beta_1) + (\alpha_2, \beta_2) = (\alpha_1 + \alpha_2, \beta_1 + \beta_2).$$

   Similarly,

   $$(\alpha_1, \beta_1)\,(\alpha_2, \beta_2) = (\alpha_1\,\alpha_2, \beta_1\,\beta_2).$$

   Can we call elements of S 'numbers'? Support your answer with reasoning.

2. Let $C[a, b]$ consist of all real-valued functions of one variable that are continuous on the interval $[a, b]$. Can functions be considered vectors? If "Yes", can $C[a, b]$ be an example of a vector space? Explain your answers.

3. In this problem you will first need to (verbally) introduce some notation. Let l be a line passing through the origin and let P be an arbitrary point; you can consider this situation either in two or three dimensions.

   (a) Find the point Q on l that is closest to P.

   (b) Regarding l as a mirror, find the reflection R of the point P. *Hint:* R is the same distance away from Q as P and is located in the plane of l and P on the "other" side from P.

   Explain your solution in both cases and confirm it with numerical experiments.

4. Let P have coordinates (x_0, y_0) and let Q be the point obtained by rotating P about the origin of $\mathbb{R}^2$ by angle θ. Find the coordinates of Q. Explain and validate your answer.

5. Consider the triangle with vertices $\mathrm{O}(0, 0)$, $\mathrm{A}(x_1, y_1)$, and $\mathrm{B}(x_2, y_2)$. Derive formulas for the following quantities.

   (a) Area of the triangle.

   (b) Cosine of the angle $\angle AOB$.

   Use whatever means necessary. Remember that I am not interested in answers alone, so explain and validate your work.

6. Consider the following snippet of `Matlab` code:

```
s = linspace(0,2*pi);
t = 2*pi*rand(3,1);

figure
plot(cos(s),sin(s))
hold on
patch(cos(t),sin(t),'r')
axis equal
title('Random triangle')
```

Save this code as a script file where `Matlab` can find it. Then run the code and save the figure it produces. Include the figure in your homework.

7. The snippet given in the previous problem generates a random triangle whose vertices lie on the circumference of the unit circle. Suppose one generates a thousand of such random triangles and computes their areas. What is going to be the average area? Modify the code snippet to answer this question. If you are skilled at `Matlab` and know enough statistics, produce a histogram of areas and compute both the mean and the standard deviation.