High level overview of the structure, connections and insights of linear algebra

Matias Frank Jensen

19 marts 2017

Introduction

This document is supposed to be a small supplement to the first year undergraduate course Linear Algebra at Aarhus University. The notes referenced here are the ones written by Jesper Funch used in the course.
My main purpose writing these is to give the reader an intuition for how all the elements presented in the linear algebra course fit together and what motivates the specific definitions. Therefore you will not find formal proving of statements here. For that the lecture notes for the course are excellent.
These notes are written with the assumption that the reader has seen and worked with the definitions, theorems, lemmas and propositions presented in the course notes and have made a serious attempt to understand the mathematical formalities. They can therefore not be used on its own to learn linear algebra, they are only a supplement to give the reader a sense of overview and intuition.
Final caution: I wrote this rather late over the course of a couple of hours without proofreading, so you will probably encounter misspellings, poor formulations, long sentences and a bit of rambling.

Why linear algebra

One can motivate the construction of math of linear algebra in multiple ways, I choose to do it like so: Solving real world problems and more abstract problems, one often finds that linear equations pop up and demand to be solved. Therefore it is of great interest to be able to effectively solve these linear equations and understand the underlying mathematical structure behind *linearity* and hopefully develop deeper insights into the inner workings of this particular part of mathematics.

First of all, it is natural to start by considering linear equations of the form

$$
\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + ...a_{1n}x_n = b_1 \\
&\vdots \\
&a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_n
\end{aligned}
$$

Where the x_i's and b_i's are real or complex numbers. This are the most familiar forms of equations and where the journey through the world of linear algebra will start.
It turns out that doing an arbitrary series of elementary operations on a system of linear equations preserves the solution set! This also explains why the *elementary operations* are defined as they are. These operations are exactly the fundamental operations that preserves the solution set.

So instead of solving the original problem, one can just transform it with the elementary operations into a new problem, with the same solution that is easier to solve. This is a very fundamental idea in mathematics: instead of solving the original problem transform it into an equivalent, but easier, one.
This naturally leads to the question: what equivalent form of a system of homogeneous equations is the easiest to solve? Answering this question one discovers the definition of a reduced linear equation and the first algorithm related to linear algebra, Gauss Elimination.
As a consequence of Gauss Elimination, it follows that a homogeneous system of linear equations

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + ...a_{1n}x_n = 0 \\ &\vdots \\ &a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = 0 \end{aligned}$$

have a non-zero solution if $m < n$. At this point it may not seem that important, but it turns out to be.

Matricies

After having done Gauss Elimination enough times, the notation becomes quite cumbersome and therefore matrices are introduced. At first, they are just an efficient way to represent systems of linear equations which makes them easier to solve. Later on one discovers they have far deeper properties.
Given a definition of matricies as representation of linear equations, it makes sense to define row operations equivalent to the elementary operations on the equations. After doing this, the addition, subtraction and scalar multiplication of matrices are also a quite straight forward addition, and these operations translate very nicely into the realm of linear equations.
Having made all these definitions the idea of a vector is a next logical step. At first they can be seen as just an efficient representation of our n variables $x_1, ..., x_n$ and n solutions $b_1, ..., b_n$. With these tools it is now very easy to represent the solution to an entire system of linear equations as just

$$Ax = b$$

Making sense of multiplying a vector x on a matrix A is done by observing that the point of the expression Ax is to represent

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + ...a_{1n}x_n \\ &\vdots \\ &a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n \end{aligned}$$

and if one views a matrix B as a sequence of vectors $(b_1, ..., b_n)$ matrixmultiplication can be seen as a sequence of matrix-vector multiplications, and the definition for matrixmultiplication follows naturally.

Given that we have defined matrices to represent systems of linear equatons, if one develops a deeper understanding of matrices this should hopefully translate into a deeper understanding of linear equations, which is the place from where all of this started.

First one can see that matrixmultiplication has some nice properties like the distributive and associate law and that $AI = A$ where I is the identity matrix. The fact that $AI = A$ is a really nice fact, and if one could, given a matrix A, find another matrix B so $AB = I$, one would effectively have discovered division for matrices. This search leads to the theory of invertible matrices, and it turns out invertible matrices have a whole slew of nice properties. Importantly, invertible matrices are precisely those matrices that are row equivalent to the identity matrix I, and the homogeneous system of linear equations I represents have only the zero solution. Therefore this facts translates to all invertible matrices as well.
All the nice properties of invertible matrices are a prime example of the close connection between linear equations and matrices.

Vectors and vector spaces

One of the deepest desires of mathematician is the process of *generalization.* First you understand a certain topic and later on realize these insights actually describe a far wider range of mathematical objects and concepts than first anticipated.
Linear algebra is no exception, and the concept of vectors is a perfect example. In the beginning we only considered linear equations where the variables could be real or complex numbers. But has we have seen, scalar multiplication and addition makes sense for other objects as well such as matrices and vectors. In fact, all the results developed so far has not really used the assumption at all that our variables are real or complex numbers. The only thing we have assumed is that addition and scalar multiplication makes sense, so for *any* object where this is the case, our linear algebra tools should works as well.
But what exactly are the properties objects must have so it makes sense to solve linear equations with them? It turns out there are 8 axioms that make it all work, and they are summarized in definition 5.1
Any set of objects that satisfies these axioms we call *vectors.* It turns out that our real and complex numbers are special case objects which we call *scalars.* The main difference being that one can multiply (commutatively) together two scalars and get another scalar as a result, but one cannot necessarily multiply vectors.
So any set of vectors and scalars can represent a system of linear equations, and all our tools automatically work. But what are vectors? Real and complex numbers can be vectors, so can matrices, vectors as you know them from $\mathbb{F}^d$, polynomials and other arbitrary functions. The point is, a vector is a very general concept that applies to many seemingly quite different objects.
One of the amazing things about this is that suddenly our matrices have gotten so much more powerful. They can now be used to solve problems with polynomials, other functions, other matrices even! A journey that started with the innocent goal of understanding linear equations of just real or complex numbers has turned out to be far more fruitful and insightful than one could initially have hoped.

But is any set of vectors interesting in a linear algebra sense? It turns out, no. Given a set of vectors V, if the addition of two vectors $v, u \in V$ is not contained in V, it doesn't really makes sense to solve linear equations with elements from V. Therefore we define a vector space to be a set of vectors that is closed under addition and scalar multiplication. One consequence of this is that the 0 vector must be an element of a vector space.

As we did with matrices, it can be very fruitful to study the properties of vector spaces since they are now a fundamental part of our understanding of linear algebra. This then leads to the definition of a sub space, linear combinations and span. A natural question could be *given vectors* $v_1, ..., v_n$, *what is the smallest vector space that contains* $v_1, ..., v_n$*?* Linear combinations and span answer this question.

From $\mathbb{F}^d$, we understand the concept behind *dimension* as the amount of numbers needed to describe a single point in $\mathbb{F}^d$. Since $\mathbb{F}^d$ is a vector space, maybe the idea of dimension will also make sense and be useful for arbitrary vector spaces? Again, the answer to this question turns out to be a resounding *yes.* Just as for $\mathbb{F}^d$ where every point (or vector in our linear algebra nomenclature) can be described by d numbers, for an arbitrary vector space V there exists a smallest integer d so every vector $v \in V$ is a linear combination of the same set of d vectors $v_1, ..., v_d$. Hence, we decide to define the *dimension* of a vector space to refer to this integer d, and a set of vectors $v_1, ..., v_d$ which span a vector space we call a *basis.* These are all very fundamental ideas and they are very important to truly understand.