Comments on linear algebra

October 4, 2005

A first encounter with ideas of linear algebra could occur if one tries to find solutions to a system of linear equations. The manipulations involved in finding the solutions are algebraic in the sense that contain additions and multiplications, but not limits as in calculus, and pictures are not so important as in geometry.

After working on some simple systems one reaches two main conclusions. First, the algebraic computations needed to find the solution get very complicated when the systems under study get bigger, although the logic behind these calculations remains simple. The logic is this: substitute a variable of one equation into other equation, recursively. The difficulty in large systems comes from the last word, recursively. Systems with more than two variables could fall into the complicated category. The second conclusion is that some systems have only one solution, but others can have infinitely many solutions, and there are systems with no solutions at all.

Here is an example, find the solutions x_1, x_2 of the system

$$a_{11}x_1 + a_{12}x_2 = b_1, \tag{1}$$
$$a_{21}x_1 + a_{22}x_2 = b_2, \tag{2}$$

where the a's and b's are given constants. The x_1, x_2 are called the variables of the system. Let us use the substitution method mentioned above to find the solutions. At least one of the a's coefficients should be nonzero, so suppose that $a_{22} \neq 0$. Then

$$x_2 = \frac{1}{a_{22}}(b_2 - a_{21}x_1), \tag{3}$$

and substitute this expression in Eq. (1) which gives

$$(a_{11}a_{22} - a_{12}a_{21})\, x_1 = a_{22}b_1 - a_{12}b_2. \tag{4}$$

At this point one meets several possibilities.

- If $a_{11}a_{22} - a_{12}a_{21} \neq 0$, then the system has a unique solution given by

$$\begin{aligned} x_1 &= \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}, \\ x_2 &= \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}, \end{aligned}$$

where the second equations comes from the substitution of the first one into Eq. (3).

- If $a_{11}a_{22} - a_{12}a_{21} = 0$, then there are two more possibilities:
  - If b_1 and b_2 satisfy that $a_{22}b_1 - a_{12}b_2 = 0$, then there are infinitely many solutions, of the form

    $$x_2 = \frac{1}{a_{22}}(b_2 - a_{21}x_1),$$

    with x_1 not determined by Eq. (4), because this equation is trivially satisfied for all x_1. This is the reason to call such a variable a free variable. It is an equation of the form $0 = 0$ for all x_1. Any value of x_1 provides a different solution x_1, $(b_2 - a_{21}x_1)/a_{22}$;
  - If b_1 and b_2 satisfy that $a_{22}b_1 - a_{12}b_2 \neq 0$, then there is no solution at all, because Eq. (4) says that $0 = 1$ which is false, so the original hope of existence of solutions is just a mirage.

The first problem, the complications to find the solutions of linear systems through the substitution method, can be tackled with the technique known as Elementary Row Operations. The idea is to transform the original linear system of equations into an equivalent system, where equivalent means that both systems have the same solutions. The transformation is performed by three operations, namely: (a) to add to one equation a multiple of the other, (b) to multiply an equation by a nonzero constant, and (c) to switch the order the equations are written down. The main claim is this: *these operations can always be used to find a system of equations equivalent to the starting system, but having a particular form, known as echelon form.* Having the system this form, the solution can be read out using the backward substitution. This case of substitution is simply carried out in echelon form systems.

While carrying out the elementary row operations, one idea pops up immediately: All these operations can be carried out only with the coefficients of the linear system, the a's, and with the right hand side of the equations, the b's, without writing the variables, the x's. One even gives a name to these coefficients, a matrix of coefficients, and the augmented matrix.

The second problem, to characterize the solutions of a linear system, is more involved. It can be seen in the example above that even the very simple case of a 2×2 system is nontrivial. In analyzing this issue is that several ideas start to pop up. One idea is to think a linear system as a linear combination of column vectors in some appropriate $\mathbb{R}^n$. Back to our example, this idea is to think Eqs. (1), (2) as follows,

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 = \mathbf{b},$$

where

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

We introduced the addition law of vectors as follows,

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}.$$

Therefore, one converts the problem to find x_1, x_2, into a problem involving vectors. The solution exists if and only if the vector $\mathbf{b}$ is a linear combination of the vectors $\mathbf{a}_1$, $\mathbf{a}_2$. If this combination is unique, the solution x_1, x_2 is unique. The vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ define a plane, so the question of existence of solutions is translated into a question of whether $\mathbf{b}$ is tangent to that plane.

A second idea pops up, *different from the previous one,* but deeply related, though. What about if one thinks the variable itself as a vector, and the coefficients in the linear system as a function acting on that vector? Back to our example, this idea has the following form. Introduce the function $A : \mathbb{R}^2 \to \mathbb{R}^2$ given by the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

This function acts on two vectors, say $\mathbf{u}$, and the result is again a two vector, say $A\mathbf{u} = \mathbf{w}$, by the rule

$$A\mathbf{u} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} a_{11}u_1 + a_{12}u_2 \\ a_{21}u_1 + a_{22}u_2 \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \mathbf{w}.$$

We introduced the notation

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}.$$

Then, the original linear system given by Eqs. (1), (2), can be seen as follows: Given a vector $\mathbf{b}$ in the range of A, find the vector $\mathbf{x}$ in the domain of A such that

$$A\mathbf{x} = \mathbf{b},$$

where we have introduced the vector

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Existence of solutions to the linear system are now translated into the question: Is $\mathbf{b}$ in the range of A? Uniqueness of the solutions are now translated into the question: Is the function A injective?

Maybe it is easier to think in this association of a function to a linear system after one performs some elementary row operations to find solutions of linear systems. Recall that the association of a matrix to a linear system appears in some natural way after applying elementary row operations to the linear system. This is half the way to the association of a function to a linear system.

Both ideas mentioned above, although different, converge into a single issue: *The characterization of solutions of linear systems, like Eqs. (1), (2), is deeply* *related to linear combinations of vectors in* $\mathbb{R}^n$ *for some appropriate* n*, and is also deeply related to functions in these spaces given by matrices, which are called linear transformations.*

This is the reason why the study of solutions of linear equations starts to evolve into the study of a richer structure. This structure is known as a *Vector Space*, from which $\mathbb{R}^n$ with the operation of addition of vectors and multiplication of a vector by a number, is an important example. One finds out that matrices are functions on these spaces, called linear transformations. Linear Algebra is the study of Vector Spaces. Linear systems are slowly taken out of the main stage, once the vector space structure is unveiled. We will dedicate a central part of our course to understand this structure.