01. Linear systems and their solutions

Consider the linear system below.

$$\begin{cases} \frac{1}{2}x + \frac{1}{2}y & = 5 \\ x - y & = 4 \end{cases} . \tag{1}$$

How might *Sherlock Holmes* solve such a system? We certainly claim no expertise in Sherlockiana; we are guided by general folklore. It is said that Sherlock would examine a situation with great clarity and notice everything that is present. (Later we will adjoin to this an important addendum; watch for it.) Thus the first equation in the system (1) may be interpreted to say that the average of x and y is 5 and the second equation may be viewed to state that their difference is 4. `[Reader: watch out for ambiguity here.]` So the two numbers are 'centered' around 5 and are 4 lengths apart, so they are $3, 7$, respectively. `[Reader: ambiguity clearing is needed here, to make precise which is` x `and which is` y`.]`

This method works well for Sherlock Holmes, who typically faces a given problem type just once. But what if we had to consider another linear system, say,

$$\begin{cases} 2x + y & = 5 \\ x + 2y & = 4 \end{cases} . \tag{2}$$

What heuristic interpretation might we use now? Let's try a not so Sherlockian approach to (2). We will label the two equations of this system by the Roman numerals I and II. As the reader knows, Roman numerals are rather cumbersome to use in computations. On the other hand, when we discuss linear systems in class we will either look at a fairly small system, rarely larger than 4×4, or a large system, sometimes $n \times n$, or even $p \times q$ $(p \neq q)$, with liberal use of '...' `[Reader:What do English professors call 'dot-dot-dot'?]` Thus, it is not so cumbersome to use just the roman numerals $I, \ldots, IV$ to refer to linear equations in a system. By convention, I refers to the first or *leftmost* or *top* equation, etc.

Take equation I in (2) and add (-2) times equation II to it. We denote this as follows:

$$I \longrightarrow I + (-2)II.$$

The reader should reflect on this notation, on its meaning and on the validity of the implied assertions. The new system we obtain is

$$\begin{cases} 0x + (-3)y & = (-3) \\ x + 2y & = 4 \end{cases} . \tag{3}$$

At this point Sherlock would surely observe that I has the solution $y = 1$. (Note: the symbol I has been recycled. It now refers to the first equation of (3). The reader should reflect on this notational point.) With this insight one sees that equation II now says $x + 2 \cdot 1 = 4$, so that $x = 2$.

A high-altitude overview of our treatment of (2) shows that we are starting by assuming that a solution exists, or equivalently, "if (2) has a solution (x, y) then ...". Thus we have only shown that if (2) has a solution, then that solution must be $x = 2, y = 1$. To complete the problem we must check to see if this is indeed a solution, by substitution, or "plugging in". Alternatively, one may argue that the steps taken above are reversible, and hence that we know *a priori* that the putative solution $(2, 1)$ is in fact an actual solution. `[Reader: what's` *a-priori* `and how is it used in mathematical discussions?]` Needless to say, the reader should reflect on these points.

The two systems (1) and (2) have some qualities in common. In particular, they both have unique solutions. This is not the case with every linear system, e.g.,

$$\begin{cases} x + y & = a \\ 2x + 2y & = b \end{cases} . \tag{4}$$

Here we treat a, b as constants (which are not specified at the moment). The symbols x and y, on the other hand, get to be designated as *variables*. Sherlock Holmes would call this system *redundant* if $b = 2a$ and *inconsistent* if $b \neq 2a$. `[Reader: reflect.]` System (4) is qualitatively different from systems (1) and (2) . `[Reader: elaborate.]` Now consider the 3×3 *walking around* system:

$$\begin{cases} 1x + 2y + 3z = \alpha \\ 4x + 5y + 6z = \beta \\ 7x + 8y + 9z = \gamma \end{cases} . \tag{5}$$

Does this system behave like (1),(2) or more like (4)? A casual glance by the untrained observer reveals no redundancy nor inconsistency. But let's apply some *row operations*, starting with $II \longrightarrow II - 4 \cdot I$, and followed by $III \longrightarrow III - 7 \cdot I$:

$$\begin{cases} 1x + 2y + 3z = \alpha \\ 0x - 3y - 6z = \beta - 4 \cdot \alpha \\ 0x - 6y - 12z = \gamma - 7 \cdot \alpha \end{cases} . \tag{6}$$

Now Sherlock may observe that (6) is redundant if $\gamma - 7 \cdot \alpha = 2 \cdot (\beta - 4 \cdot \alpha)$ and is inconsistent otherwise. `[Reader:reflect on the difference in the phrases` *'redundant equation'* `versus` *'redundant system'*`.]`

This discussion suggests the concept of *archetypes* of linear systems. As often happens in mathematics, we have some idea, possibly a vague idea of what we want *archetypes* to mean, and we need to make it more precise. Whatever this term means, (1) and (2) seem to be of the same archetype and, to some extent, (4) and (5) share a different archetype.

<u>Exercise:</u> work out all possible archetypes of 2×2 systems of linear equations.

(Note that, in our context, this exercise is not precisely defined.)

To save some typesetting effort as well as raw materials used in writing, we often represent linear systems in a simplified notation, e.g., system (5) is represented as:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}. \tag{7}$$

Reading (7) we have a 3×3 *matrix* times a 3×1 matrix equals a 3×1 matrix. This last sentence suggests definitions of the concept of *matrix* and of *multiplication* of matrices, which the reader should look up. An even more concise representation is possible, using what is called an *augmented* matrix:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & \alpha \\ 4 & 5 & 6 & \beta \\ 7 & 8 & 9 & \gamma \end{array} \right). \tag{8}$$

The reader will notice that the variables x, y, z do not appear at all in (8). This is a good point for reflection and for some looking up, though the intuitive picture will probably suffice.

Reflecting on this entire section we recall an important quality of top detectives such as Sherlock Holmes. In examining a scene they notice not only everything that is there, but also *everything that is not there.* All the linear systems that we have displayed have coefficients that are whole numbers, i.e., integers. (Okay, there's one system that has some half-integers.) The set of all integers is denoted by $\mathbb{Z}$, probably because of the German word for numbers, *Zhal.* But this is only for the purpose of convenient exposition and computation. We intend to do linear algebra "over the reals $\mathbb{R}$" (or the complex numbers $\mathbb{C}$; for a while it will not make much difference which of these two we use). Thus, we could just as well have worked with the following linear system:

$$\begin{cases} \pi x + \sqrt{2} y & = e \\ \\ 4x - \sqrt{5} y & = \sqrt{\pi} \end{cases}. \tag{9}$$

The coefficient 4 is the only one in system (9) that is an integer. The others are not even *rational.* (Reader: review the definition of rational numbers, $\mathbb{Q}$.)

Here is an ancient tablet depicting methods for solving linear equations by a method which, many years later, became known as Gaussian Elimination, after Carl F. Gauss, born millennia after the carving of the tablet:

Figure 1: Captions go above figures

It began in cuneiform tablets like VAT 8389. Vorderasiatisches Museum in Berlin, 12.1 by 7.9 cm.

Here are some general references for linear algebra, including the books [2, 3] and the papers [1, 4, 5]:

References

[1] Beezer, R.A. (2014). Extended Echelon Form and Four Subspaces, *American Math Monthly*, 121:7, 644-647.

[2] Beezer, R.A. (2016). *A First Course In Linear Algebra*, Open Source book, `http://linear.ups.edu`, accessed May, 2020.

[3] Hoffman, K., Kunze, R. (1971). *Linear Algebra*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall.

[4] Strang, G. (2014) The Core Ideas in Our Teaching *Notices Am. Math.* Soc. 61, no. 10, 1243-1245, doi:10.1090/noti1174.

[5] Trotter, H. F. (1961). A Canonical Basis for Nilpotent Transformations. *Amer. Math. Monthly* 68:8, 779-780.