Mathematical Apology Number 16: Factorising and Linear Equations

John Butcher, The University of Auckland

Mathematics education has now become an academic subject in its own right. There was a time when any mathematician or mathematics teacher, who cared about his or her subject, had something to say on pedagogical issues with as much of the confidence of the expert as anybody else, but those days are long past. Normally I don't take part in seminars, lectures or workshops in the mathematics education specialty, because I have become an outsider. But I had to make an exception during a conference I took part in, during September in the lovely town of Kastoria in the Northern part of Greece. Although the conference was concerned mainly with other topics, a single lecture in the area of secondary school mathematics education was allocated to a session that I had been asked to chair. Suen Che Yin from Singapore spoke of the use of graphics calculators as educational aids, especially in Asian countries, where computers cannot be easily afforded. As an example, he showed how a calculator might be used to find the factors of a quadratic polynomial, such as x^2+x-2, by plotting the function on a calculator, and noting where it crossed the x–axis. The speaker went further than I was able to agree with, when he proposed that the old tedious drills of factorising quadratic expressions should now be removed completely from teaching practice, because machines can do this, or at least help with it. A well-presented argument, even if I don't entirely agree with its conclusions, always forces me to think about the questions it raises, as carefully as I can.

I don't believe students are taught to factorise quadratic expressions because this is a needed skill, with practical applications, but because it is one of the many early steps in the development of a mathematical mind. In the long run, a mathematician needs to be able to factorise all sorts of things. Along the way, a mathematician in the making needs to understand various functions and how they behave; a quadratic function of one variable is a crucial first example after straight lines.

Factorising as a basic mathematical idea breaks a complicated problem up into smaller problems, which might be solved in sequence, to yield the solution to the larger problem, or in parallel to yield alternative answers to the same problem. I will give some examples from solving linear equation systems where factorisation plays a crucial role.

The problem:

$$\begin{pmatrix} 1 & 3 \\ -2 & -5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

becomes easier if it is known that

$$\begin{pmatrix} 1 & 3 \\ -2 & -5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$$

because it then amounts to solving two simple problems in sequence

$$\text{first} \quad \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix} \quad \text{and then} \quad \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Thus the problem of finding "triangular factors" for a matrix, becomes a crucial step in linear algebra computations, because it replaces a single difficult problem by a sequence of two simple problems. I will say a little about how such factorisations are arrived at in a future Apology.

Now two further linear equation systems; in each of these there are now more equations than unknowns.

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -5 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -4 \\ 2 \\ 4 \end{pmatrix}.$$

The second of these two systems cannot be solved, and the best we can do seems to be to choose the unknowns to satisfy the first three of the four equations, and to accept the fact that nothing can be done about the fourth equation: it just cannot be satisfied. The solution, when we do this, is $x_1=4$, $x_2=-3$, $x_3=1$. A justification for this course of action might be that we have really found a least squares solution; that is we have found a vector x such that the length of the vector $Ax-b$ is minimized and this is being offered as a "solution" to an insolvable system $Ax=b$. If we were to attempt the same approach to the first of these two overdetermined systems, the details become much more complicated. However, if we were to somehow find out that:

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix},$$

and that

$$\begin{pmatrix} 1 \\ -5 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 \\ -4 \\ 2 \\ 4 \end{pmatrix},$$

then the least squares solution is the same as for the second problem, because the matrix:

$$Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

does not change the length of a vector it multiplies. The reason is that the columns of Q each have length 1 and any two different columns are orthogonal. Hence, the columns of Q could be used as orthogonal axes. I will show how to find a suitable Q for this sort of problem next time.

Many mathematicians are most comfortable with problems that are well-posed. It is very satisfying to know that a solution exists, and that this solution is the only solution. However, mathematics is the servant of sciences which often have to work with information which is sometimes inconclusive, and sometimes contradictory. To conclude this discussion, the following linear equation problem is proposed

$$\begin{pmatrix} -1 & -1 & -4 \\ -4 & 2 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}.$$

There is a family of solutions, but is there any good reason for preferring any one of them over the others?

Author

Emeritus Professor John Butcher, Department of Mathematics, The University of Auckland, Private Bag 92-019, Auckland. Phone (09) 373 7599, Ext. 8747, Fax (09) 373 7457, E-mail: butcher@math.auckland.ac.nz.