Teaching Linear Algebra with and to Computers

D.J.Jeffrey and Robert M. Corless
Ontario Research Centre for Computer Algebra
University of Western Ontario

Abstract

Three topics are discussed that relate to the teaching of linear algebra using computers (here the term *computers* includes calculators). The first topic is the variation in notation and

The major use of computers has been to assist students with

From the point of view of linear algebra, these operations are not important, because they are not matrix operations, but for teachers of introductory numerical analysis courses, the notation is the source of endless debugging problems for the students. For professionals, the compactness of this notation is a great convenience.

In Maple, the elementwise of the linear or

It can also be argued that we should look to the future. If the

3.3 Never learning the best method

A book on canoeing [5] recommends that the first paddling stroke to teach students is the *back*stroke,

> because we note a tendency to revert to the most familiar stroke when flustered.

Many readers will have heard stories of students graduating and going to work in industry, and then applying mathematics from their undergraduate textbooks. Perhaps they try solving 100 equations in 100 unknowns using Cramer's rule, or searching for the eigenvalue of a large matrix by trying to solve its characteristic polynomial. With today's arbitrary precision software, they

non-square matrix produced an error message; in MATLAB 6 (release 12), this is no longer the case, but Mathematica and the HP still insist on an invertible matrix.

In fact, any rectangular matrix A has the Turing factors

$$PA = LDUR\ .$$

Here, the R matrix is the unique reduced row-echelon form of A. The matrix P is a preconditioning matrix; it is usually a permutation matrix, but it can be a more general matrix. For example, the standard software package LAPACK uses row and column equilibration (routines xGESVX) and this can be described using preconditioning matrices.

4.2 What can we do with this?

We return to the eigenvalue problem above and use Turing factoring:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1-\lambda & 3 \\ -2 & 7-\lambda \end{pmatrix} =$$

$$\begin{pmatrix} 1 & 0 \\ \frac{1}{2}(\lambda-1) & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 \\ 0 & \frac{1}{2}(\lambda^2-8\lambda+13) \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{2}(\lambda-7) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Notice the rather useless row-echelon form at the end, and notice that the characteristic polynomial has appeared naturally in the D matrix. Students could be taught the simple rule "always check the cases $\det D = 0$ separately", but it is rewarding to understand where this rule comes from.

The importance of $\det D = 0$ comes because we are interested in special cases, and special cases are often points of discontinuity; indeed, this is why they are interesting and considered special. So we ask about the continuity properties of Turing factoring.

Consider the RREF of the matrix $A(k) = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$ as the parameter k passes through 0. The RREF is the identity I, except for a sudden, discontinuous change at $k = 0$. Linear algebra courses do not usually discuss the limit of a matrix, but under any reasonable definition, the limit of RREF(A) as $k \to 0$ must be the unit matrix. So we have $\lim_{k\to 0} RREF(A) = I \neq A(0)$. Under any definition of continuity, the RREF of A is discontinuous at $k = 0$.

Definition. A matrix $A(x)$ is continuous at $x = a$ if each of its elements is continuous at $x = a$.

Once we become accustomed to thinking of (interesting) special cases as discontinuities, we can frame the following theorem.

Theorem: Let $A(x)$ be a matrix depending upon one or more variables or parameters x, and let A be continuous at a point $x = a$. For any fixed x, let $A(x)$ have the Turing factoring given by $P(x)A(x) = L(x)D(x)U(x)R(x)$. If $\det D(x) \neq 0$ in some neighbourhood of $x = a$, then $R(x)$, $L(x)$, $D(x)$, $U(x)$ are all continuous at $x = a$ and moreover $P(x)$ may be taken constant in a neighbourhood of $x = a$.

This theorem is proved in [3] and means two things.

- A CAS can give an RREF which contains **visible failure** built in. Places where an RREF might fail are no longer invisible because of the definition.
- The discontinuity information is collected in a single place, namely, along the diagonal of D.

To return to the eigenvalue problem, the Turing factors of a matrix $A - \lambda I$, will always lead to a diagonal matrix D in which the diagonal entries are

$$p_1(\lambda), p_2(\lambda)/p_1(\lambda), \ldots, p_n(\lambda)/p_{n-1}(\lambda)$$

where $p_k(\lambda)$ is a polynomial of degree k, and in the last entry, $p_n(\lambda)$ is the characteristic polynomial. Thus, $\det D = p_n(\lambda)$ and only the roots of the characteristic polynomial are special cases. For some matrices, a fraction p_k/p_{k-1} might simplify; this would simply mean that a preliminary splitting of the characteristic polynomial had been found during the computation.

4.3 The benefits

The immediate benefit to the teacher of Turing factoring is the combining together of row reduction and LU factoring. If LU factoring was not previously in the course material, then it comes along at no extra cost to the student. A common objection to Turing factoring is that it is "a bit rich" for beginning students. Its computation also threatens a great deal of computation. However, the point of computers is exactly to take over the burden of computation. Provided students know what Turing factors are, computer algebra systems can easily obtain them for the student.

The immediate benefit to the system designer is that a mechanism becomes available for returning special case information back to the user. This obviates the need to develop new user interfaces that allow the passing back to the user of proviso information. The benefit to the student is a gentle introduction to one of the most powerful ideas of modern linear algebra: factoring.

References

[1] Robert M. Corless. Six, lies, and calculators. *The American Mathematical Monthly*, 100(4):344–350, 1993.

[2] Robert M. Corless and David J. Jeffrey. Well, it isn't quite that simple. *SIGSAM Bulletin*, 26(3):2–6, 1992.

[3] Robert M. Corless and David J. Jeffrey. The Turing factorization of a rectangular matrix. *SIGSAM Bulletin*, 31(3):20–28, 1997.

[4] Roger A. Horn. *The Hadamard*