Cayley-Hamilton for Eigenvalues

Robin Harte

The Cayley-Hamilton theorem says that a linear operator $T: X \to X$ on a finite dimensional space $X \cong \mathbf{C}^n$ satisfies its *characteristic equation*:

$$p_T(T) = 0 \tag{1}$$

where

$$p_T(T) = (z-\lambda_1)^{\nu_1}(z-\lambda_2)^{\nu_2}\dots(z-\lambda_k)^{\nu_k} \tag{2}$$

is the *characteristic polynomial* of T; thus $\lambda_1, \lambda_2, \dots, \lambda_k$ are the distinct *eigenvalues* of T and $\nu_1, \nu_2, \dots, \nu_k$ are their (algebraic) *multiplicities.* It is familiar that, if the inverse T^{-1} exists, then it can be expressed as a polynomial in T with the help of (2); dividing across by the non-vanishing constant term of p_T and bringing it across the equality sign gives

$$p'_T(T)T = I = Tp'_T(T) \tag{3}$$

This note arises from the problem of calculating the *eigenvectors* associated with the eigenvalues λ_j. In the process we rediscover a well-known theorem (which was obviously not well enough known to the author!).

Begin with the observation that (1) may well be valid for polynomials p_T other than the characteristic polynomial: it is possible for (1) to hold with integers ν_j in (2) smaller than the full algebraic multiplicities. If p_T is the polynomial given by (2) we shall write

$$\hat{p}_T(T) = (z-\lambda_1)(z-\lambda_2)\dots(z-\lambda_k), \tag{4}$$

and call this the *reduced* polynomial of T; then it may or may not happen that

$$\hat{p}_T(T) = 0. \tag{5}$$

If (5) holds we shall call the operator T *reduced.* The well-known theorem [4, Chapter IV Theorem 5] is simply stated:

Theorem *If $T: X \to X$ is a linear operator on a finite dimensional space then*

$$T \text{ reduced} \iff T \text{ diagonal.} \tag{6}$$

Proof *Diagonal* means of course that T is a direct sum of scalars:

$$X = \sum_{j=1}^{k}(T-\lambda_j I)^{-1}(0) = \bigoplus_{j=1}^{k}(T-\lambda_j I)^{-1}(0); \tag{7}$$

if this happens then (5) follows at once by considering $\hat{p}_T(T)$ separately on each eigenspace. For forward implication in (6) we need the notion of "exactness" [2, 3, Chapter 10]: the pair (R, S) of operators on X is called *exact* if

$$R^{-1}(0) = S(X). \tag{8}$$

Inclusion one way is just the condition

$$RS = 0; \tag{9}$$

for the opposite inclusion it is sufficient that there are operators S' and R' on X for which

$$R'R + SS' = I. \tag{10}$$

We have this, in particular, when R and S are polynomials in T without common divisor: if

$$S = q(T) \quad \text{and} \quad R = r(T) \quad \text{with} \quad \gcd(q, r) = 1 \tag{11}$$

then the Euclidean algorithm for polynomials gives polynomials $q'(z)$ and $r'(z)$ for which

$$q'(z)q(z) + r'(z)r(z) = 1,$$

giving (10) with $S' = q'(T)$ and $R' = r'(T)$. This happens many times over if T is reduced: if (5) holds then we get (11) with

$$S = q(T) = T - \lambda_j I \quad R = r(T) = \prod_{i \neq j}(T - \lambda_i I) \tag{12}$$

Further, in this case, everuthing in (9) and (10) commutes, so that also

$$S^{-1}(0) = R(X), \quad R^{-1}(0) \cap S^{-1}(0) = 0 \quad \text{and} \quad S(X) + R(X) = X; \tag{13}$$

thus

$$X = S^{-1}(0) \oplus R^{-1}(0).$$

Forward implication in (6) is now induction on the number k of distinct eigenvalues λ_j; for if $T: X \to X$ is diagonal on each of its invariant subspaces $S^{-1}(0)$ and $R^{-1}(0)$ then it is diagonal on their direct sum $S^{-1}(0) \oplus R^{-1}(0)$. On $S^{-1}(0)$ the operator T coincides with the scalar $\lambda_j I$; on $R^{-1}(0)$ T has only $k-1$ eigenvalues.

This theorem is not new, and can be found for example in Jacobson [4]. We believe our direct deduction from the Euclidean algorithm has some charm; the same argument gives, with no assumptions about T, the "primary decomposition"

$$X = \sum_{j=1}^{k} (T - \lambda_j I)^{-\nu_j}(0) = \bigoplus_{j=1}^{k} (T - \lambda_j I)^{-\nu_j}(0).$$

An alternative version of the argument, passing through the medium of "Taylor invertibility", is given by Gonzalez [1].

When an operator $T: X \to X$ is "reduced" in the sense of (5) then its eigenvectors can all be obtained without solving any more equations: with $S = T - \lambda_j I$ and R as in (12), the first part of (13) says that the eigenspace corresponding to λ_j is the range or "column space" of the matrix R built out of the remaining eigenvalues. Of course in practice it will usually be easier and pleasanter to solve the equations $Sx = 0$ than to compute the matrix R.

References

[1] M. Gonzalez, *Null spaces and ranges of polynomials of operators*, Pub. Math. U. Barcelona **32** (1988), 167–170.

[2] R.E. Harte, *Almost exactness in normed spaces*, Proc. Amer. Math. Soc. **100** (1987), 257–265.

[3] R.E. Harte, *Invertibility and singularity*, Marcel Dekker, New York, 1988.

[4] N. Jacobson, *Lectures in abstract algebra II*, Van Nostrand, New York, 1953.

Department of Mathematics
University College Cork

BOOK REVIEWS

Metric Spaces: iteration and application
by Victor Bryant, Cambridge University Press (1985), STG£5.95 (paperback).

Metric Spaces
by E.T. Copson, Cambridge Tracts in Mathematics number 57, Cambridge University Press (1968), STG£22.50 (hardback), STG£7.95 (paperback).

Introduction to Metric and Topological Spaces
by W.A. Sutherland, Oxford University Press (1981), STG£10.95 (paperback).

Of the three books, I like the one by Bryant the least. It claims, with some justification, to make the subject interesting. But the result is a book which might be more appropriately described as an introduction to iteration and fixed point theory that includes a little on metric spaces. To be somewhat objective, the book does touch on many of the basic concepts (limits of sequences; closed, complete, compact and connected sets). The applications include the existence and uniqueness of solutions for ordinary differential equations. But my basic objection is the second class treatment given to open sets and the less than enthusiastic treatment of continuity. On page 35, having introduced closed sets via limits of sequences, we are told that open sets are not really necessary because "all theorems about open sets can be stated in terms of closed sets". While this is undeniable, most textbooks do not take such an upside down view, and I do not consider that one can be said to have learned 'metric spaces' without being comfortable with the notion of open set. Who would like to volunteer to rewrite a standard text on multivariable analysis (never mind ones about complex analysis, functional analysis or elementary manifolds) mentioning only closed sets? The last chapter (marked optional) of Bryant's short book does make some amends by looking into continuity (even uniform continuity and the fact that the continuous image of a compact set is compact) and defining open sets.

This brings us to the question of what the book sets out to achieve. It claims to be intended for courses for engineering or 'combined honours' students, or really for those who have taken but not grasped a single variable