[11] N. McMillan, *The Irish Astronomical Network in the 19th century*, Irish Astronomical Journal (to appear).

[12] op.cit., F.P., p.xx.

[13] ibid, p.590.

[14] FitzGerald to Heaviside, February 4, 1889.

[15] A. M. Bork, *The 'FitzGerald' Contraction*, Isis, (1966), 57: pp. 192–207. The view that FitzGerald's claim to priority despite being of the most good natured and modest kind was presented here as being unfounded. In S. G. Brush, Notes on the History of the FitzGerald-Lorentz Contraction, Isis, (1967), 58, pp. 230–232, the published priority of FitzGerald is succinctly demonstrated in a devastating way to the Bork thesis.

[16] O. J. Lodge, Obituary Notice of the Royal Society, 1901.

[17] J. Larmor, obituary, F.P. p.xliv (emphasis added).

[18] ibid, p. xlliii.

School of Science
Regional Technical College
Carlow

NOTES

Diagonalising a Real Symmetric Matrix and the Interlacing Theorems

Donal P. O'Donovan

In most linear algebra textbooks the diagonalisation of a real symmetric matrix is accomplished by first proving that the eigenvalues are real and then proceeding to the orthogonal diagonalisation. Anton's book [1] notes in the preface that the first part of this can require an excursion into the theory of complex vector spaces. The purpose of this note is to show that a more direct route is possible if one proves the realness of the eigenvalues and the orthogonal diagonalisation simultaneously.

In itself this would be of very little interest, at least for mathematics students, who usually handle $\mathbf{C}^n$ as readily as $\mathbf{R}^n$. However what one is led to, is something much more, namely the Cauchy inequalities, between the eigenvalues of any finite dimensional self adjoint operator and its compressions [3], and also to the Courant-Fisher min-max formulae for the characteristic numbers [2[, and these are topics not usually found in Linear Algebra textbooks. So, many mathematicians must be unaware of them. Yet the interlacing that one finds is both elegant and useful. For example it gives in several lines, the proof, that a symmetric matrix is positive if the principal minors are all positive.

I find linear operators a better setting for diagonalisations than matrices, so we work with them. Recall that if U is any subspace of an inner product space V, and $T : V \to V$ is a linear operator then the compression of T to U is just the operator $P_U T : U \to U$, where P_U is the orthogonal projection onto U. For students who prefer matrices, if an orthonormal basis $u_1, ..., u_r$ for U is expanded to an orthonormal basis $u_1, ...u_r, u_{r+1} ..., u_r$ for V then the matrix for $P_U T$ is just that block of the matrix of T whose entries are in both the first r rows and first r columns. For those who like pictures.

$$[T] = \begin{pmatrix} t_{11} & \dots & t_{1r} & \dots & t_{1n} \\ \vdots & & \vdots & & \vdots \\ t_{r1} & \dots & t_{rr} & & \\ \vdots & & \vdots & \vdots & \\ t_{n1} & \dots & \dots & & t_{nn} \end{pmatrix} \Rightarrow [P_U T] = \begin{pmatrix} t_{11} & \dots & t_{1r} \\ \vdots & & \vdots \\ t_{r1} & \dots & t_{rr} \end{pmatrix}$$

Recall also that the characteristic numbers are just the eigenvalues of T, repeated according to multiplicity, arranged in descending order

$$\lambda_1(T) \geq \lambda_2(T) \dots \geq \lambda_n(T),$$

as we shall write them. We will use $C_T(\lambda)$ to denote the characteristic polynomial for T, that is det $(T - \lambda I)$.

Theorem *Let T be a symmetric linear transformation on a finite dimensional real inner product space V, then its eigenvalues are real and T can be diagonalised with respect to an orthonormal basis.*

Proof We proceed by induction on n, the dimension of V. If $n = 1$, the statements are trivial, so suppose both statements are proven for $n - 1$. Let V have dimension n and U be any $n - 1$ dimensional subspace. Let S be the compression of T to U. Then we may apply the induction hypothesis to S, obtaining real characteristic numbers, $\lambda_1(S), \dots, \lambda_{n-1}(S)$, and corresponding orthonormal eigenvectors $u_1, \dots, u_{n-1}$. Choose u_0 to complete the orthonormal basis $u_0, u_1, \dots u_{n-1}$, and consider the matrix representation for T.

$$T = \begin{bmatrix} t_0 & t_1 & . & . & . & t_{n-1} \\ t_1 & \lambda_1(S) & . & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ t_{n-1} & 0 & . & . & . & \lambda_{n-1}(S) \end{bmatrix} \qquad (*)$$

It is easy to see that

$$C_T(\lambda_1(S)) = -t_1^2 \prod_{j=2}^{n-1} (\lambda_j - \lambda_1)$$

So

$$C_T(\lambda_1(S)) = \begin{cases} \leq 0 & \text{if } n \text{ is even} \\ \geq 0 & \text{if } n \text{ is odd} \end{cases}$$

But $C_T(\lambda) \to +\infty$ or $-\infty$ when $\lambda \to +\infty$ according as n is even or odd. So in both cases the graph of $C_T(\lambda)$ crosses or at least touches the λ - axis. This gives a real root λ_0 (which must be $\geq \lambda_1(S)$). Now one solves for an eigenvector v_0 and applies the induction hypothesis to the compression of T to the orthogonal complement of v_0, giving the desired result.

We noted above the extra bit of information, namely that $\lambda_0 \geq \lambda_1(S)$. In fact we can deduce easily that if S is any compression of T to an $n - 1$ dimensional subspace then

$$\lambda_1(T) \geq \lambda_1(S) \geq \lambda_2(T) \geq \dots \geq \lambda_{n-1}(S) \geq \lambda_n(T)$$

which we will refer to as "interlacing".

Again we proceed by induction. For $n = 2$ we have

$$\begin{pmatrix} t_0 & t_1 \\ t_1 & \lambda_1(S) \end{pmatrix}$$

If $t_1 = 0$ the result is immediate. If $t_1 \neq 0$, then $C_T(\lambda_1(S)) = -t_1^2 < 0$, but $C_T(\lambda) \to +\infty$ as $\lambda \to \pm\infty$, so $\lambda_1(S)$ lies between the roots.

For arbitrary n, we have as before the matrix $(*)$ for T. First if any $t_i = 0$ then $\lambda_i(S)$ is an eigenvalue for T, and so are the eigenvalues of the matrix gotten by ignoring the i^{th} row and the i^{th} column. By the induction hypothesis $\{\lambda_j(S)\}_{j \neq i}$ interlace the eigenvalues of this second matrix. A moments thought shows that $\{\lambda_j(S)\}$ then interlace the set of all eigenvalues of T.

Next one sees that

$$\begin{aligned} C_T(\lambda_1(S)) &= -t_1^2 \prod_{j=2}^{n-1} (\lambda_j(S) - \lambda_1(S)) \\ &= -t_1^2 (-1)^{n-2} \prod_{j=2}^{n-1} |\lambda_j(S) - \lambda_1(S)| \end{aligned}$$

and

$$\begin{aligned} C_T(\lambda_2(S)) &= -t_2^2(\lambda_1(S)-\lambda_2(S))\prod_{j=3}^{n-1}(\lambda_j(S)-\lambda_2(S)) \\ &= -t_2^2|\lambda_1(S)-\lambda_2(S)|(-1)^{n-3}\prod_{j=3}^{n-1}|\lambda_j(S)-\lambda_2(S)| \end{aligned}$$

and so on.

Thus we see that if the $\lambda_i(S)$ are all distinct then the signs of $C_T(\lambda_i(S))$ alternate, and the roots of $C_T(\lambda)$ interlace the $\lambda_i(S)$. If some $\lambda_k(S)=\lambda_{k+1}(S)$, writing

$$C_T(\lambda)=t_0\prod_{j=1}^{n-1}(\lambda_j(S)-\lambda)-\sum_{i-1}^{n-1}t_i^2\prod_{j\neq i}(\lambda_j(S)-\lambda)$$

we have $\lambda_k(S)-\lambda$ as a factor of $C_T(\lambda)$. Hence $\lambda_k(S)$ is an eigenvalue for T, and it is immediate from the form of $C_T(\lambda)$ above, that

$$\frac{C_T(\lambda)}{\lambda_k(S)-\lambda}=C_R(\lambda)$$

where R is the $(n-1)\times(n-1)$ matrix

$$\begin{pmatrix} t_0 & t_1 & . & . & \sqrt{t_k^2+t_{k+1}^2} & t_{k+2} & . & t_{n-1} \\ t_1 & \lambda_1(S) & . & . & 0 & 0 & . & 0 \\ . & . & . & . & . & . & . & \\ . & . & . & . & . & . & . & \\ \sqrt{t_k^2+t_{k+1}^2} & 0 & . & . & \lambda_k(S) & . & . & 0 \\ t_{k+2} & 0 & . & . & . & \lambda_{k+2}(S) & . & 0 \\ . & . & . & . & . & . & . & . \\ t_{n-1} & 0 & . & . & 0 & 0 & . & \lambda_{n-1}(S) \end{pmatrix}$$

Applying the induction hypothesis to R gives the desired interlacing.

If $(t_{ij})_{i,j=1,n}$ is an $n\times n$ matrix then the principal minors $\Delta_k(T)$ are defined as

$$\Delta_k(T)=\det(t_{ij})_{i,j=1,k}$$

We want to show that if $\Delta_k(T)\geq 0$ for all $1\leq k\leq n$ for an $n\times n$ symmetric matrix T then T is positive. This follows easily if we show all the eigenvalues are positive. Again we proceed by induction. The case $n=1$ is clear. Now applying the induction hypothesis to $(t_{ij})_{i,j=1,n-1}$ we have all its eigenvalues positive. Then by the interlacing $n-1$ of the eigenvalues of T are positive. Then $\Delta_n(T)\geq 0$ shows that all of them are positive.

Finally, if $T^{(r)}$ is the compression of T to an $n-r$ dimensional subspace, we may successively invoke the interlacing r times to obtain the Cauchy inequalities

$$\lambda_i(T)\geq\lambda_i(T^{(r)})\geq\lambda_{i+r}(T)$$

It follows that

$$\lambda_i(T)\leq\lambda_1(T^{(i-1)})$$

But by diagonalisation, equality can be achieved, so

$$\lambda_i(T)=\min\lambda_1(T^{(i-1)})$$

where the min is over all $n-i+1$ dimensional compressions, and of course

$$\lambda_1(T^{(i-1)})=\max\{(T^{(i-1)}v,v):||v||=1\}$$

which gives the desired min-max characterisation.

We can also note that the interlacing result and its consequences are obviously also true for a self adjoint operator on a finite dimensional complex inner product space.

References

[1] H. Anton, *Elementary Linear Algebra*, Wiley, New York, 1984.

[2] R. Bellman, *Introduction to Matrix Analysis*, McGraw-Hill, New York, 1970.

[3] M. Marcus and H. Mink, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Boston, 1964.

School of Mathematics
Trinity College Dublin