Real symmetric matrices

1 Eigenvalues and eigenvectors

We use the convention that vectors are row vectors and matrices act on the right.

Let A be a square matrix with entries in a field F; suppose that A is $n \times n$. An *eigenvector* of A is a non-zero vector $v \in F^n$ such that $vA = \lambda v$ for some $\lambda \in F$. The scalar λ is called an *eigenvalue* of A.

The *characteristic polynomial* of A is the polynomial f_A defined by

$$f_A(x) = \det(xI - A).$$

The *characteristic equation* is the equation $f_A(x) = 0$.

Theorem 1 *(a) The eigenvalues of A are the roots of the equation $f_A(\lambda) = 0$.*

(b) The matrix A satisfies its own characteristic equation; that is, $f_A(A) = O$, where O is the all-zero matrix.

Part (b) of this theorem is known as the *Cayley–Hamilton Theorem*. The *minimal polynomial* of A is the monic polynomial m_A of least degree such that $m_A(A) = O$. The Cayley–Hamilton Theorem is equivalent to the statement that $m_A(x)$ divides $f_A(x)$. It is also true that $f_A(x)$ and $m_A(x)$ have the same roots (possibly with different multiplicities).

A matrix D is *diagonal* if all its off-diagonal entries are zero. If D is diagonal, then its eigenvalues are the diagonal entries, and the characteristic polynomial of D is $f_D(x) = \prod_{i=1}^{n}(x - d_{ii})$, where d_{ii} is the (i,i) diagonal entry of D.

A matrix A is *diagonalisable* if there is an invertible matrix Q such that QAQ^{-1} is diagonal. Note that A and QAQ^{-1} always have the same eigenvalues and the same characteristic polynomial.

Theorem 2 *The matrix A is diagonalisable if and only if its minimal polynomial has no repeated roots.*

2 Symmetric and orthogonal matrices

For the next few sections, the underlying field is always the field $\mathbb{R}$ of real numbers.

We use $A^\top$ to denote the transpose of the matrix A: that is, it is the matrix whose (j,i) entry is the (i,j) entry of A. The matrix A is called *symmetric* if $A = A^\top$. The matrix Q is called *orthogonal* if it is invertible and $Q^{-1} = Q^\top$.

The most important fact about real symmetric matrices is the following theorem.

Theorem 3 *Any real symmetric matrix is diagonalisable. More precisely, if A is symmetric, then there is an orthogonal matrix Q such that $QAQ^{-1} = QAQ^\top$ is diagonal.*

The natural setting of this theorem is *real inner product spaces*, which we now describe.

3 Real inner product spaces

An *inner product* on the real vector space V is a function from $V \times V$ to $\mathbb{R}$ (we write the inner product of v and w as $v \cdot w$), satisfying the following three conditions:

- $v \cdot (a_1 w_1 + a_2 w_2) = a_1(v \cdot w_1) + a_2(v \cdot w_2)$;
- $w \cdot v = v \cdot w$;
- $v \cdot v \geq 0$, and $v \cdot v = 0$ if and only if $v = 0$.

The first two conditions imply

- $(a_1 v_1 + a_2 v_2) \cdot w = a_1(v_1 \cdot w) + a_2(v_2 \cdot w)$.

We summarise the first and fourth conditions by saying that the inner product is *bilinear*; the second condition says that it is *symmetric*, and the third that it is *positive definite*.

For any subspace W of V, we write $W^\perp$ for the subspace

$$\{v \in V : (v \cdot w) = 0 \text{ for all } w \in W\}.$$

This operation (read "perp") on subspaces has the properties

- $V = W \oplus W^\perp$ (this means that $V = W + W^\perp$ and $W \cap W^\perp = 0$);
- $\dim(W) + \dim(W^\perp) = \dim(V)$;

- if $W_1 \subseteq W_2$, then $W_1^\perp \supseteq W_2^\perp$;
- $(W^\perp)^\perp = W$.

Any real inner product space has an *orthonormal basis* $e_1, \ldots, e_n$. These vectors satisfy $e_i \cdot e_i = 1$ and $e_i \cdot e_j = 0$ for $i \neq j$. Such a basis is constructed by applying the *Gram–Schmidt process* to an arbitrary basis.

A linear transformation A of V is said to be *self-adjoint* if $v \cdot (wA) = (vA) \cdot w$ for all vectors v, w. A linear transformation Q which preserves the inner product (in the sense that $(vQ) \cdot (wQ) = v \cdot w$) is called *orthogonal*.

Now if we represent linear transformations with respect to an orthonormal basis, a transformation is self-adjoint if and only if its matrix is symmetric, and a transformation is orthogonal if and only if its matrix is orthogonal. Moreover, a transformation is orthogonal if and only if it maps an orthonormal basis to an orthonormal basis.

So, in order to prove the spectral theorem, it is enough to prove the following assertion:

Theorem 4 *Let A be a self-adjoint transformation of a real inner product space V. Then there is a decomposition $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$, and distinct scalars $\lambda_1, \ldots, \lambda_r$, with the properties*

(a) $W_i \subseteq W_j^\perp$ for $i \neq j$;

(b) $vA = \lambda_j v$ for all $v \in W_j$.

To show this, let λ be any eigenvalue of A, and W the corresponding eigenspace $\{w \in V : wA = \lambda w\}$. We claim that $W^\perp$ is invariant under A. For take $v \in W^\perp$; we must show that $vA \in W^\perp$. But, for any $w \in W$, we have

$$(vA) \cdot w = v \cdot (wA) = v \cdot (\lambda w) = \lambda(v \cdot w) = 0;$$

so the assertion is proved.

But then the restriction of A to $W^\perp$ is a self-adjoint transformation. By induction on the dimension, there is a decomposition of $W^\perp$ into subspaces with the properties of the theorem. Adding W to this decomposition, we obtain the result.

4 The spectral theorem

Let W be a subspace of the inner product space V. Then $V = W \oplus W^\perp$. Define a map P on V by the rule that $(w+x)P = w$, where $w \in W$ and $x \in W^\perp$. (Any vector has a unique representation of this form.) We call P the *projection* of V onto W. Note that the image of P is W, and $wP = w$ for all $w \in W$; hence $P^2 = P$. Moreover, it is easy to check that P is self-adjoint.

Conversely, if P is self-adjoint and satisfies $P^2 = P$, then P is the projection of V onto some subspace W (namely, the image of P). So these conditions define projections in an intrinsic way. This means that we can recognise the matrix of a projection; it is a matrix satisfying $P = P^\top = P^2$.

The main theorem about real symmetric matrices can be re-phrased in terms of projections. In this form it is often referred to as the *spectral theorem*.

Theorem 5 *Let A be a real symmetric matrix with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$. Then there are projection matrices $P_1, \ldots, P_r$ satisfying*

(a) $P_1 + P_2 + \cdots + P_r = I$;

(b) $P_i P_j = O$ for $i \neq j$;

(c) $A = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_r P_r$.

We take P_i to be the projection onto the eigenspace V_i associated with λ_i (the set of all vectors v satisfying $vA = \lambda_i v$. Since these spaces are pairwise orthogonal and satisfy $V_1 \oplus V_2 \oplus \cdots \oplus V_r$, conditions (a) and (b) hold. Part (c) is proved by noting that the two sides agree on any vector in V_i, for any i, and so agree everywhere.

5 Commuting symmetric matrices

There is a remarkable extension of the spectral theorem to sets of symmetric matrices. The essential condition is that the matrices should commute.

Theorem 6 *Let $A_1, \ldots, A_s$ be a set of symmetric matrices satisfying $A_i A_j = A_j A_i$ for all i, j. Then there is an orthogonal matrix P such that PAP^{-1} is diagonal for*

$i = 1, \ldots, s$. Equivalently, there are projections $P_1, \ldots, P_r$ such that conditions (a) and (b) of Theorem 5 hold, and for $j = 1, \ldots$ we have

$$A_j = \sum_{i=1^r} \lambda_{ij} P_i$$

for some scalars λ_{ij}.

In terms of the inner product space, the theorem reads as follows. Given a set of s pairwise commuting self-adjoint transformations; then $\mathbb{R}^n$ is the orthogonal direct sum of subspaces V_i, on each of which all of the transformations are scalars.

The proof is by induction on s, the spectral theorem itself being the case $s = 1$. So suppose that the theorem is true for $s-1$. Now decompose $\mathbb{R}^n$ into eigenspaces of A_s, using the spectral theorem; let V_j correspond to the eigenvalue λ_j of A_s.

We claim that V_j is invariant under $A_1, \ldots, A_{s-1}$. For let $v \in V_j$ and $i \leq s-1$. Then

$$(vA_i)A_s = (vA_s)A_i = (\lambda_j v)A_i = \lambda_j(vA_i),$$

where the first equality holds because A_i and A_s commute. Now this equation shows that vA_i is an eigenvector of A_s with eigenvalue λ_j; so $vA_i \in V_j$, as claimed.

Now V_j is an inner product space, and the restrictions of $A_1, \ldots, A_{s-1}$ to it are pairwise commuting self-adjoint transformations of it. So we can write V_j as an orthogonal direct sum of subspaces, each of which consists of eigenvectors for $A_1, \ldots, A_{s-1}$ (and also of course for A_s, since all vectors in V_j are eigenvectors for A_s). The proof is complete.

This result is crucial in the theory of association schemes.

6 Hermitian, normal and unitary matrices

Although our main interest lies in real symmetric matrices, there is a parallel theory over the complex numbers, which is of great importance. The definition of an inner product on a complex vector space is slightly different: the definition is

- $v \cdot (a_1 w_1 + a_2 w_2) = a_1(v \cdot w_1) + a_2(v \cdot w_2)$;
- $w \cdot v = \overline{v \cdot w}$, where $\overline{\ }$ denotes complex conjugation;
- $v \cdot v$ is a non-negative real number, and is zero if and only if $v = 0$.

The first two conditions imply

- $(a_1v_1 + a_2v_2) \cdot w = \overline{a_1}(v_1 \cdot w) + \overline{a_2}(v_2 \cdot w)$;
- $v \cdot v$ is real for any vector v.

As before, a linear transformation is *self-adjoint* if $v \cdot (wA) = (vA) \cdot w$ for any v and w. We say that Q is *unitary* if it preserves the inner product, that is, $(vQ) \cdot (wQ) = v \cdot w$ for all v, w.

If we take an orthonormal basis for the space, then the matrix representing a self-adjoint transformation is *Hermitian*, that is, $\overline{A}^\top = A$; and the matrix representing a unitary transformation satisfies $\overline{Q}^\top = Q^{-1}$ (we also call such a matrix *unitary*).

Now the complex spectral theorem asserts that, if A is Hermitian, then there is a unitary matrix Q such that QAQ^{-1} is diagonal. The reformulation in terms of projections and the extension to a set of commuting Hermitian matrices are as in the real case.

Over the real numbers, a matrix which can be diagonalised by an orthogonal matrix must necessarily be symmetric. In the complex case, however, we can go a little further. The matrix Z is said to be *normal* if it commutes with its conjugate transpose (that is, $Z\overline{Z}^\top = \overline{Z}^\top Z$).

Theorem 7 *Let Z be a square matrix over $\mathbb{C}$. Then there exists a unitary matrix Q such that QZQ^{-1} is diagonal if and only if Z is normal.*

For any complex matrix Z can be written uniquely in the form $Z = X + iY$, where X and Y are self-adjoint. (This is the analogue of the real and imaginary parts of a complex number.) Now Z is normal if and only if $XY = YX$. If this holds, then there is a unitary matrix which diagonalises both X and Y, and hence Z. The converse is trivial.

7 An application: Moore graphs

Here is an application to show how the theory of real symmetric matrices can be used in combinatorics. This analysis is due to Hoffman and Singleton [1].

Let Γ be a connected regular graph with valency k, and v a vertex of Γ. Each neighbour of v is joined to at most $k-1$ non-neighbours of v. So, if Γ has diameter 2, then it has at most $1+k+k(k-1) = k^2+1$ vertices altogether, with equality if and only if it has *girth* 5 (no circuits shorter than 5). Dually, if Γ has girth 5, then it has at least k^2+1 vertices, with equality if and only if it has diameter 2.

A *Moore graph* of diameter 2 is a graph attaining these bounds, that is, a connected regular graph with diameter 2 and girth 5. It has $n = k^2 + 1$ vertices, where k is the valency. Our question is: *which Moore graphs exist?*

The adjacency matrix A of such a graph satisfies

$$A^2 = kI + (J - I - A),$$

that is, $A^2 + A - (k-1)I = J$, where J is the all-1 matrix. Thus J commutes with A, and so A and J are simultaneously diagonalisable. Now J has the all-1 vector j as an eigenvector with multiplicity n; all its other eigenvalues are zero. Clearly j is an eigenvalue of A with multiplicity 1. Its other eigenvalues thus satisfy

$$\lambda^2 + \lambda - (k-1) = 0,$$

so $\lambda = \frac{1}{2}(-1 \pm \sqrt{4k-3})$. If their multiplicities are f and g, we have

$$\begin{aligned} f + g &= n - 1 = k^2, \\ \tfrac{1}{2}(-1 + \sqrt{4k-3})f + \tfrac{1}{2}(-1 - \sqrt{4k-3})g &= -k, \end{aligned}$$

the second equation coming from the fact that $\mathrm{Trace}(A) = 0$ (since all the diagonal entries are zero). From these equations we find that

$$(f - g)\sqrt{4k-3} = \tfrac{1}{2}k(k-2).$$

If $k = 2$, then $n = 5$ and the unique graph is the pentagon. Suppose not. Then $f - g$ is an integer, so $4k - 3$ must be a perfect square, indeed the square of an odd number; say

$$4k - 3 = (2s+1)^2,$$

so $k = s^2 + s + 1$. In addition, $2s+1$ must divide $k(k-2) = (s^2+s+1)(s^2+s-1)$. It follows that $2s+1$ divides $3 \cdot 5 = 15$, so (since $s > 0$) we have

$$\begin{aligned} 2s+1 &= 3, \quad 5, \quad 15 \\ k &= 3, \quad 7, \quad 57 \\ n &= 10, \quad 50, \quad 3250 \end{aligned}$$

It is known that there is a unique 3-valent Moore graph on 10 vertices (the Petersen graph) and a unique 7-valent Moore graph on 50 vertices (the Hoffman–Singleton graph). The existence of a 57-valent Moore graph on 3250 vertices is unknown.

References

[1] A. J. Hoffman and R. R. Singleton, On finite Moore graphs, *IBM J. Research Develop.* **4** (1960), 497–504.

Peter J. Cameron
September 9, 2002