Lecture 21

Diagonalizing a Real Symmetric Matrix

This lecture will be devoted to proving the following theorem:

Theorem Suppose A is a real symmetric n by n matrix. Then A is diagonalizable over $\mathbb{R}$. This means there is a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ for $\mathbb{R}^n$ such that for each j, $1 \leq j \leq n$, v_j is an eigenvector for A — so there exists λ_j with $Av_j = \lambda_j v_j$. Furthermore, we may assume $\mathcal{B}$ is an orthonormal basis. Hence, if we let P be the matrix with columns which are the coordinates of the eigenvectors v_j, $1 \leq j \leq n$, so

$$P = \begin{array}{c} \begin{matrix} v_1 & v_2 & & v_n \end{matrix} \\ \begin{pmatrix} \big\downarrow & \big\downarrow & \cdots & \big\downarrow \end{pmatrix} \end{array}.$$

Then

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

So we have diagonalized A.

Finally P is an orthogonal matrix

We will now prove the theorem. A much more general theorem is proved in the text, Theorem (32.16), pg 285. The key result is the following Proposition (existence of a real eigenvalue). It will require nine pages to prove (up to the end of pg. 11).

Proposition 1

Every real symmetric n by n matrix has at least one real eigenvalue.

Remark In fact all the eigenvalues of a real symmetric matrix are real — that will come later.

In order to prove Proposition 1 and the Theorem we will need to expand our point of view in two ways. First, we will have to consider the linear transformation T on $\mathbb{R}^n$ associated to A and the standard basis. And second, we will need to go from $\mathbb{R}^n$ to $\mathbb{C}^n$.

Symmetric Linear Transformations

Let V be a real vector space and let $(\ ,\ )$ be a scalar (inner) product on V (see Lecture 9). Let $u_1, u_2, \dots, u_n$ be an orthonormal basis for V so $(u_i, u_j) = \delta_{ij}$ (so 1 if $i=j$ and 0 if $i \neq j$).

<u>Definition</u> Let $T \in L(V,V)$. Then T is symmetric if

$$(Tv_1, v_2) = (v_1, Tv_2), \text{ all } v_1, v_2 \in V.$$

Recall that the transpose ${}^tT$ of T is defined by

$$(Tv_1, v_2) = (v_1, {}^tTv_2), \text{ all } v_1, v_2 \in V$$

So we see that T is a symmetric <u>linear transformation</u> if and only if

$$(v_1, Tv_2) = (v_1, {}^tTv_2), \text{ all } v_1, v_2 \text{ in } V.$$

This last equation holds if and only if

$$T = {}^tT \quad (*)$$

Note that $(*)$ is an equality of linear transformations not matrices.

Lemma 1

T is a symmetric linear transformation on a vector space V if and only if the matrix of T relative to any orthonormal basis for V is a symmetric matrix.

Proof Let $A = (a_{ij})$ be the matrix of T relative to an orthonormal basis $u_1, u_2, \ldots, u_n$ for V. Hence

$$Tu_j = \sum_{k=1}^{n} a_{kj} u_k . \quad (*)$$

Take the scalar product of each side of $(*)$ with u_i and use ~~that~~ $(u_k, u_i) = 0$, $k \neq i$ and $(u_i, u_i) = 1$ to obtain

$$(Tu_j, u_i) = a_{ij} .$$

Interchanging i and j we obtain

$$(Tu_i, u_j) = a_{ji}$$

Now T is a symmetric linear transformation if and only if for all i, j we have

$$a_{ij} = (T(u_j), u_i) = (u_j, T(u_i)) = (T(u_i), u_j) = a_{ji}$$

symmetry of T

But the equation $a_{ij} = a_{ji}$, all i, j, is the condition that A is a symmetric matrix.

$\square$

Lemma 2

Suppose T is a symmetric linear transformation and $U \subset V$ is an invariant subspace under T. Let $U^\perp$ be the orthogonal complement of U under $(,)$. Then $U^\perp$ is also T-invariant.

Proof

By definition, $v \in U^\perp$ if and only if $(v, u) = 0$, all $u \in U$.

So we want to prove ~~that~~ if $(v, u) = 0$, all $u \in U$, then $(Tv, u) = 0$, all $u \in U$.

So let $u \in U$. Because T is symmetric $(Tv, u) = (v, Tu)$. But because U is invariant under T we have $Tu \in U$. so $(v, Tu) = 0$. Hence $(Tv, u) = 0$.

Since this holds for all $u \in U$, $Tv \in U^\perp$. $\square$

Remark This lemma will play a critical role in the proof of the Theorem (see pg 13). Any result that provides an invariant complement to an invariant subspace is always important.

Hermitian Matrices and Linear Transformations

In order to prove Proposition 1 we will need to pass from $\mathbb{R}^n$ to $\mathbb{C}^n$ and prove results generalizing those of the previous pages. To start with we need to define a complex-valued scalar product on $\mathbb{C}^n$. (the analogue for a general complex vector space V is called a Hermitian scalar product on V, see the text, Def (32.2), pg. 279)

Definition

Let $u = (z_1, z_2, \dots, z_n)$ and $v = (w_1, w_2, \dots, w_n) \in \mathbb{C}^n$. Then we define the Hermitian scalar product (u, v) of u and v by

$$(u, v) = \sum_{i=1}^{n} z_i \overline{w_i}$$

Lemma 3

$$(v, u) = \overline{(u, v)}$$

Proof By definition $(v, u) = \sum_{i=1}^{n} w_i \overline{z_i}$

But $$\overline{(u,v)} = \overline{\sum_{i=1}^{n} z_i \overline{w_i}} = \sum_{i=1}^{n} \overline{z_i \overline{w_i}}$$

$$= \sum_{i=1}^{n} \overline{z_i}\, \overline{\overline{w_i}} = \sum_{i=1}^{n} \overline{z_i}\, w_i = (v, u)$$

$\square$

Lemma 4 If $e_1, e_2, \ldots, e_n$ is the standard basis of $\mathbb{C}^n$ then

$$(e_i, e_j) = \delta_{ij}, \quad 1 \le i, j \le n$$

Proof clear

Lemma 5 If $u \in \mathbb{C}^n$ and $u \neq 0$ then $(u,u) > 0$ (so it is real and positive)

Proof Suppose $u = (z_1, \ldots, z_n)$ and $z_j = x_j + iy_j$, $1 \le j \le n$. Then

$$(u,u) = \sum_{j=1}^{n} z_j \overline{z_j} = \sum_{j=1}^{n} |z_j|^2 = \sum_{j=1}^{n} (x_j^2 + y_j^2) > 0$$

(Recall if $z \in \mathbb{C}$, $z = x + iy$ then $\bar{z} = x - iy$ and $z\bar{z} = |z|^2 = x^2 + y^2$) $\square$

Definition (adjoint of a matrix) Let $A = (a_{ij})$ be a complex n by n matrix. Then A^*, the adjoint of A, is the matrix obtained by taking the complex conjugates of the entries of A then taking the transpose of the resulting matrix so

$$A^* = {}^t\bar{A} \quad \text{where } \bar{A} = (\overline{a_{ij}})$$

Definition (adjoint of a linear transformation). Let V be a complex vector space and $T: V \to V$ be a linear transformation.

Then the adjoint T^* of the linear transformation T is the unique linear transformation satisfying

$$(Tv_1, v_2) = (v_1, T^* v_2), \text{ all } v_1, v_2.$$

Lemma 5

Let A be the matrix of T relative to the standard basis. Then the matrix of T^* relative to the standard basis is A^*.

Proof Let $A = (a_{ij})$. Then by definition $Te_j = \sum_{i=1}^{n} a_{ij} e_i$, $1 \le j \le n$.

Hence $a_{ij} = \underline{(Te_j, e_i)} = (e_j, T^* e_i)$

$$= \overline{(T^* e_i, e_j)} \qquad (*)$$

Suppose $B=(b_{ij})$ is the matrix of T^* relative to the standard basis.

Then $T^*e_i = \sum_{k=1}^{n} b_{ki} e_k$.

Hence $(T^*e_i, e_j) = b_{ji}$.

and $\overline{(T^*e_i, e_j)} = \overline{b_{ji}}$

Substituting $\overline{b_{ji}}$ for $\overline{(T^*e_i, e_j)}$ in $(*)$ we obtain

$$a_{ij} = \overline{b_{ji}}$$

and so

$$b_{ij} = \overline{a_{ji}}.$$

Hence $B = A^*$

$\square$

Definition

An n by n complex matrix A is Hermitian if $A^* = A$.

Remark Note that if A is real (so $\overline{A} = A$) and symmetric so ${}^tA = A$ then it is Hermitian. Indeed

$$A^* = {}^t(\overline{A}) = {}^tA = A.$$

Hence Proposition 1 will follow from

Proposition 2 (more general)

Let A be an n by n complex Hermitian matrix. Then all eigenvalues of A are real.

Proof Let λ be an eigenvalue (so $\lambda \in \mathbb{C}$) of A and $v \in \mathbb{C}^n$ be a corresponding eigenvector so $v \neq 0$ and

$$Av = \lambda v \qquad (*)$$

Then λ is real if and only if $\overline{\lambda} = \lambda$.

Take the scalar product of both sides of (*) with v to obtain

$$(Av,v) = \lambda(v,v)$$

so

$$\lambda = \frac{(Av,v)}{(v,v)}.$$

But we have seen that (v,v) is real and strictly positive so λ is real if and only if (Av,v) is real if and only if

$$\overline{(Av,v)} = (Av,v) \qquad (**)$$

We now prove (**)

By Lemma 3, $\overline{(Av,v)} = (v,Av)$

and by definition $(v,Av) = (A^*v,v)$.

So (**) becomes (for any A)

$$\overline{(Av,v)} = (A^*v,v) \qquad (***)$$

But A is Hermitian so $A^* = A$

and we obtain

$$\overline{(Av,v)} = (Av,v)$$

$\square$

Proof of the Theorem

Let A be a real symmetric n by n matrix. We will prove it is diagonalizable by induction on n.

<u>$n=1$</u> Then A is a 1 by 1 matrix so it is already diagonal.

<u>Induction step $n-1 \to n$</u>

Assume any $n-1$ by $n-1$ real symmetric matrix B is diagonalizable and there is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of B.

Let A be an n by n real symmetric matrix. By Proposition 1, A has a real eigenvalue λ corresponding to a nonzero real eigenvector v.

Let T be the symmetric linear transformation on $\mathbb{R}^n$ with matrix relative to the standard basis equal to A. Then

$$Tv = \lambda v.$$

Let U be the line through v.
Then U is invarient under T.
Hence by Lemma 2, $U^{\perp}$ is
also invarient under T (this is
where we use that A and hence T is symmetric)).

Lemma 6

The restriction of T to $U^{\perp}$
is symmetric.

Proof

Let $w_1, w_2 \in U^{\perp}$. Since $U^{\perp} \subset V = \mathbb{R}^n$
and T is symmetric

$$(Tw_1, w_2) = (w_1, Tw_2)$$

But this proves T restricted to
U is symmetric.

$\square$

But $\dim U^{\perp} = n-1$, hence, by induction, there is an orthonormal basis $w_1, w_2, \ldots, w_{n-1}$ for $U^{\perp}$ consisting of eigenvectors for T. Since $v \in U$ and $w_i \in U^{\perp}$, $1 \leq i \leq n-1$, we have

$$(v, w_i) = 0, \quad 1 \leq i \leq n-1,$$

and hence $v, w_1, \ldots, w_{n-1}$ is a set of n orthogonal vectors in $\mathbb{R}^n$. Hence, they are independent and since $\dim \mathbb{R}^n = n$ they are an orthogonal basis for $\mathbb{R}^n$. To get an orthonormal basis we replace v by $u = \frac{v}{\|v\|}$. Then $\|u\| = 1$,

$(u, w_i) = 0$, $1 \leq i \leq n-1$ and u is still an eigenvector of T corresponding to ~~the~~ eigenvalue λ, that is

$$Tu = \lambda u.$$

(Any scalar multiple of an eigenvector is still an eigenvector).

Hence $\mathcal{B} = \{u, w_1, \ldots, w_{n-1}\}$ is a basis for $\mathbb{R}^n$ consisting of eigenvectors for A.

$\square$

How to Diagonalize a Real Symmetric Matrix Using the Change of Basis Formula (Lecture 8)

Let A be a real symmetric n by n matrix corresponding to a linear transformation $T: V \to V$ where $V = \mathbb{R}^n$ using the standard basis $\mathcal{S} = \{e_1, e_2, \ldots, e_n\}$ for $V = \mathbb{R}^n$. Hence, in the notation of Lecture 6

$$A = {}_{\mathcal{S}}[T]_{\mathcal{S}} .$$

Now let $\mathcal{E} = \{v_1, v_2, \ldots, v_n\}$ be a basis for V consisting of eigenvectors of T. Hence $Tv_j = \lambda_j v_j$, $1 \leq j \leq n$, and B, the matrix of T relative to $\mathcal{E}$ is

$$B = {}_{\mathcal{E}}[T]_{\mathcal{E}} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

<u>Theorem</u> Let $P = P_{\mathcal{S} \leftarrow \mathcal{E}}$ be the matrix which is the change of basis matrix from $\mathcal{E}$ to $\mathcal{S}$. So the columns of P are the coordinates of the eigenvectors $v_1, v_2, \ldots, v_n$.

Then

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} = B$$

<u>Proof</u> By the Change of Basis formula for Linear Transformations, Lecture 8,

$$B = {}_{\mathcal{E}}[T]_{\mathcal{E}} = P_{\mathcal{E} \leftarrow \mathcal{S}} \; {}_{\mathcal{S}}[T]_{\mathcal{S}} \; P_{\mathcal{S} \leftarrow \mathcal{E}}$$

But ${}_{\mathcal{S}}[T]_{\mathcal{S}} = A$ and $P_{\mathcal{S} \leftarrow \mathcal{E}} = P$ $\square$

<u>How to Remember This</u>

"When you diagonalize a matrix A the matrix of eigenvectors P goes on the right"

For an example see pages 5–7 of Lecture 21.