Lecture 23

Putting a Matrix with Complex Entries into Triangular Form and the Cayley-Hamilton Theorem.

Upper Triangular Form

Definition Let A be an n by n complex matrix. We say A is upper triangular if all the entries of A strictly below the diagonal are equal to zero

$$\begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix} \quad (3 \text{ by } 3 \text{ case})$$

Theorem (Existence of Upper Triangular Form for Complex Matrices)

Let A be an n by n complex matrix. Then there is a basis $\mathcal{B}$ for $\mathbb{C}^n$ such that the matrix of A relative to $\mathcal{B}$ is upper triangular.

Equivalently there is an invertible matrix P such that $P^{-1}AP$ is upper triangular.

Proof We will prove the theorem by induction on n. Clearly it is true for the case $n = 1$ (all one-by-one matrices are upper-triangular, in fact diagonal, by definition)

Assume we have proved the theorem for $\mathbb{C}^{n-1}$ and let A be an n by n complex matrix. Since we are working over $\mathbb{C}$, A has an eigenvalue λ_1 with a corresponding eigenvector $v \neq 0$.

We are going to construct a basis $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ for $\mathbb{C}^n$ such that the matrix of A relative to $\mathcal{B}$ is upper triangular.

The first step — take the first basis vector v_1 to be the eigenvector v so the first column of the matrix will be

$$\begin{pmatrix} \lambda_1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Now comes the induction step.
Choose a complement U in V
to the line $\mathbb{C}v_1$ through v_1. Hence
we have

$$V = \mathbb{C}v_1 + U \quad (*)$$

We define the projection p_U
from V onto U (using the decomposition
$(*)$) by

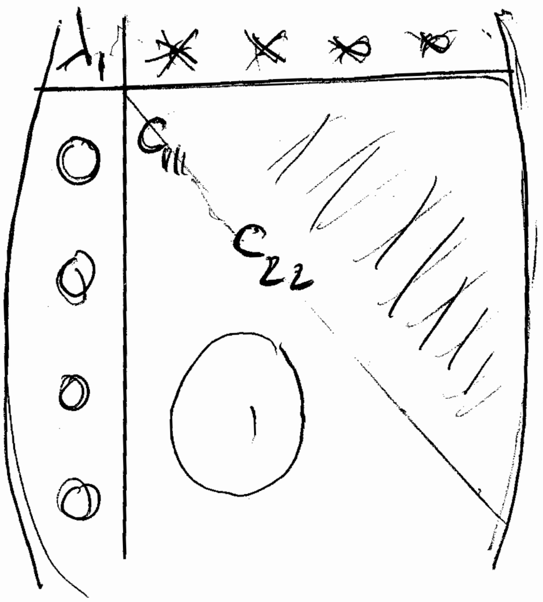
$$p_U(zv_1 + u) = u, \quad z \in \mathbb{C},\ u \in U.$$

Define $S \in L(U, U)$ by

$$S(u) = p_U(T(u)).$$

By induction there is a
basis $\{u_2, \ldots, u_n\}$ for U
so that S is an upper triangular

have

$B=$ 

From the picture we see that C upper triangular $n-1$ by $n-1$ forces B to be upper triangular n by n

With this the theorem is proved $\square$

The Characteristic Polynomial of an Upper Triangular Matrix

Theorem Suppose A is an upper triangular n by n matrix

so

$$A = \begin{pmatrix} a_{11} & a_{12} & & \\ & a_{22} & & \\ & & \ddots & \\ & 0 & & a_{nn} \end{pmatrix}$$

Then the characteristic polynomial $h_A(x)$ of A is given by

$$h_A(x) = (x - a_{11})(x - a_{22}) \cdots (x - a_{nn}).$$

Proof

$$\begin{aligned} h_A(x) &= \det(xI - A) \\ &= \det(-(A - xI)) \\ &= (-1)^n \det(A - xI) \end{aligned}$$

$$= (-1)^n \det \begin{pmatrix} a_{11}-x & & & \\ & a_{22}-x & & \\ & & \ddots & \\ 0 & & & a_{nn}-x \end{pmatrix}$$

$$= (-1)^n \left(a_{11}-x\right)\left(a_{22}-x\right)\cdots\left(a_{nn}-x\right)$$

$$= (x-a_{11})(x-a_{22})\cdots(x-a_{nn}).$$

$\square$

The Cayley-Hamilton Theorem
(text pg 205)

Theorem Let $T \in L(V,V)$ and $h(x)$ be the characteristic polynomial of T. Then

$$h(T) = 0$$

Proof Choose a basis $\mathcal{B} = \{v_1, \dots, v_n\}$ for V so that T is triangular relative to $\mathcal{B}$. Let A be the matrix of T relative to $\mathcal{B}$ so

$$A = \begin{pmatrix} a_{11} & a_{12} & & \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}$$

Hence $Av_1 = a_{11}v_1$, $Av_2 = a_{12}v_1 + a_{22}v_2$

and more generally

$$A\,\mathrm{span}(v_1, \dots, v_j) \subset \mathrm{span}(v_1, \dots, v_j), \; 1 \le j \le n$$

We have seen the characteristic polynomial $h_A(x)$ is given by

$$h_A(x) = (x-a_{11})(x-a_{22})\cdots(x-a_{nn})$$

and hence

$$h_A(A) = (A-a_{11}I)(A-a_{22}I)\cdots(A-a_{nn}I) \quad (*)$$

We now prove by induction on j that

$$h(A)(\mathrm{span}(v_1,\dots,v_j)) = 0, \quad 1 \leq j \leq n. \quad (**)$$

The key point is that the factors $(A-a_{jj}I)$ in the product on the right-hand side of $(*)$ commute with each other so we can permute them as needed.

Now we start the induction.

$\underline{j=1}$ (so we want to prove $h_A(A)v_1 = 0$)

$$h_A(A)v_1 = (A-a_{11}I)(A-a_{22}I)\cdots(A-a_{nn}I)v_1$$

$$= (A-a_{22}I)(A-a_{33}I)\cdots(A-a_{nn}I)(A-a_{11}I)v_1$$

But $Av_1 = a_{11}v_1$ so $(A-a_{11}I)v_1 = 0$.

So $(A-a_{22}I)(A-a_{33}I)\cdots(A-a_{nn}I)\underbrace{(A-a_{11}I)v_1}_{0}$

$$= (A-a_{22}I)(A-a_{33}I)\cdots(A-a_{nn}I)\cdot 0$$

$= 0$ (any linear transformation applied to 0 gives 0)

<u>Induction step $j \longrightarrow j+1$</u>

Assume we have proved (~~**~~) for j.

Then $h_A(A)v_1 = 0,\ h(A)v_2 = 0, \dots, h(A)v_j = 0$

We want to prove that $h_A(A)v_{j+1} = 0$.

But

$$h_A(A)v_{j+1} = (A-a_{11}I)\cdots(A-a_{jj}I)(A-a_{j+1,j+1}I)(A-a_{j+2,j+2}I)\cdots(A-a_{nn}I)v_{j+1}$$

$$= (A-a_{j+2,j+2}I)\cdots(A-a_{nn}I)\cdot(A-a_{11}I)(A-a_{22}I)\cdots(A-a_{jj}I)(A-a_{j+1,j+1}I)v_{j+1}$$

But $u = Av_{j+1} - a_{j+1,j+1}v_{j+1} \in \mathrm{Span}(v_1, \dots, v_j)$

and so $(A-a_{11}I)(A-a_{22}I)\cdots(A-a_{jj}I)(A-a_{j+1,j+1})v_{j+1}$

$= (A-a_{11}I)(A-a_{22}I)\cdots(A-a_{jj}I)u$.

But since $u \in \mathrm{span}(v_1, \dots, v_j)$ we have by induction

$$(A-a_{11}I)(A-a_{22}I)\cdots(A-a_{jj}I)u = 0$$

So,
$$(A-a_{j+2,j+2}I)\cdots(A-a_{nn}I)\underbrace{(A-a_{11}I)(A-a_{22}I)\cdots(A-a_{j+1,j+1}I)v_{j+1}}_{0}$$

$= 0.$

Hence $h_A(A)v_{j+1} = 0.$
and we have completed the induction steep.

Hence $h_A(A)$ annihilates $\mathrm{span}(v_1, \ldots, v_n)$
or equivalently
$h_A(A)v_1 = 0, \ldots, h_A(A)v_n = 0.$

But $v_1, v_2, \ldots, v_n$ span V so $h_A(A)v = 0$, all $v \in V$.

Hence $h_A(A) = 0.$

$\square$

Extending the Cayley-Hamilton Theorem to any Subfield F of $\mathbb{C}$ (e.g. $\mathbb{R}$ or $\mathbb{Q}$)

Theorem Let F be a subfield of $\mathbb{C}$ and V be a vector space over F. Let $T \in L(V,V)$. Let $h_T(x)$ be the characteristic polynomial of T so $h_T(x) \in F[x]$. Then

$$h_T(T) = 0.$$

Proof Choose a basis $\mathcal{B} = \{v_1, \dots, v_n\}$ for V over F and let A be the matrix of T relative to $\mathcal{B}$ so $A \in M_n(F)$.

But since $F \subset \mathbb{C}$ we have

$$M_n(F) \subseteq M_n(\mathbb{C})$$

and also $F[x] \subset \mathbb{C}[x]$.

Let $A \in M_n(F)$, then $\det(A) \in F$ is the same whether we compute it considering A as an element of $M_n(F)$ or $M_n(\mathbb{C})$. The same statement holds for the characteristic polynomial

$h_A(x) = \det(xI - A) \in F[x]$.

But if we consider $A \in M_n(\mathbb{C})$ we just proved $h_A(A) = 0$.

But the matrix $h_A(A)$ is the same if we consider $A \in M_n(F)$ so

$h_A(A) = 0$. $\square$

<u>Remark</u> The Cayley-Hamilton Theorem holds for matrices over <u>any</u> field F.

Block Upper-Triangular Form for Real Matrices

The previous theorem does not hold for real matrices. Here is the best one can do for a general real matrix.

Definition

Let B be an n by n real matrix. Then we say B is in block upper triangular form if B has the form

$$B = \begin{pmatrix} B_1 & * & \cdots & * \\ 0 & B_2 & & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_n \end{pmatrix}$$

where each B_j is either one by one, or two by two.

Example

$$B = \left(\begin{array}{cc|cc} 0 & -1 & b_{13} & b_{14} \\ 1 & 0 & b_{23} & b_{24} \\ \hline 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

$\leftarrow$ this is the best you can do, you can't make it upper triangular real

Theorem

Let A be a real n by n matrix. Then there is a basis $\mathcal{B}$ for $\mathbb{R}^n$ so that if we rewrite A in terms of $\mathcal{B}$ then the resulting matrix is in block upper triangular form.

Equivalently, there exists an invertible n by n matrix P so that $P^{-1}AP = B$ is block upper triangular.

Example The matrix $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ (rotation by 90°) cannot be put in real upper triangular form (it has eigenvalues i and $-i$).