LINEAR ALGEBRA -II

B V Rajarama Bhat

Indian Statistical Institute, Bangalore

Lecture 33: Cayley Hamilton theorem

- Recall: Let A be an $n \times n$ complex matrix. For any complex polynomial $f(x) = a_0 + a_1 x + \cdots + a_m x^m$, by definition,

$$f(A) = a_0 I + a_1 A + \ldots + a_m A^m.$$

Lecture 33: Cayley Hamilton theorem

- Recall: Let A be an $n \times n$ complex matrix. For any complex polynomial $f(x) = a_0 + a_1 x + \cdots + a_m x^m$, by definition,

$$f(A) = a_0 I + a_1 A + \ldots + a_m A^m.$$

- Consider

$$\mathcal{A} = \{f(A) : f \text{ is a polynomial}\}.$$

Lecture 33: Cayley Hamilton theorem

- Recall: Let A be an $n \times n$ complex matrix. For any complex polynomial $f(x) = a_0 + a_1 x + \cdots + a_m x^m$, by definition,

$$f(A) = a_0 I + a_1 A + \ldots + a_m A^m.$$

- Consider

$$\mathcal{A} = \{f(A) : f \ \text{ is a polynomial}\}.$$

- Clearly this is a subspace of $M_n(\mathbb{C})$. Actually, $\mathcal{A}$ is a 'sub-algebra' of $M_n(\mathbb{C})$, that is, it is also closed under taking products.

Lecture 33: Cayley Hamilton theorem

- Recall: Let A be an $n \times n$ complex matrix. For any complex polynomial $f(x) = a_0 + a_1 x + \cdots + a_m x^m$, by definition,

$$f(A) = a_0 I + a_1 A + \ldots + a_m A^m.$$

- Consider

$$\mathcal{A} = \{f(A) : f \text{ is a polynomial}\}.$$

- Clearly this is a subspace of $M_n(\mathbb{C})$. Actually, $\mathcal{A}$ is a 'sub-algebra' of $M_n(\mathbb{C})$, that is, it is also closed under taking products.
- Note that $M_n(\mathbb{C})$ is a vector space of dimension n^2. Therefore the dimension of $\mathcal{A}$ can't be more than n^2.

Lecture 33: Cayley Hamilton theorem

- Recall: Let A be an $n \times n$ complex matrix. For any complex polynomial $f(x) = a_0 + a_1x + \cdots + a_mx^m$, by definition,

$$f(A) = a_0I + a_1A + \ldots + a_mA^m.$$

- Consider

$$\mathcal{A} = \{f(A) : f \ \text{ is a polynomial}\}.$$

- Clearly this is a subspace of $M_n(\mathbb{C})$. Actually, $\mathcal{A}$ is a 'sub-algebra' of $M_n(\mathbb{C})$, that is, it is also closed under taking products.
- Note that $M_n(\mathbb{C})$ is a vector space of dimension n^2. Therefore the dimension of $\mathcal{A}$ can't be more than n^2.
- In particular, $I, A, A^2, \ldots, A^{n^2}$ are linearly dependent.

Continuation

- In other words, there exists a non-zero polynomial $q(x) = b_0 + b_1 x + \cdots + b_m x^m$ of degree at most n^2 such that

$$q(A) = b_0 I + b_1 A + b_2 A^2 + \cdots + b_m A^m = 0.$$

Continuation

- In other words, there exists a non-zero polynomial $q(x) = b_0 + b_1 x + \cdots + b_m x^m$ of degree at most n^2 such that

$$q(A) = b_0 I + b_1 A + b_2 A^2 + \cdots + b_m A^m = 0.$$

- Assume $b_m \neq 0$. Then $A^m = -\frac{1}{b_m}(b_0 I + b_1 A + \cdots + b_{m-1} A^{(m-1)})$.

Continuation

- In other words, there exists a non-zero polynomial $q(x) = b_0 + b_1 x + \cdots + b_m x^m$ of degree at most n^2 such that

$$q(A) = b_0 I + b_1 A + b_2 A^2 + \cdots + b_m A^m = 0.$$

- Assume $b_m \neq 0$. Then $A^m = -\frac{1}{b_m}(b_0 I + b_1 A + \cdots + b_{m-1} A^{(m-1)})$.
- This may help us to compute higher powers of A or to simplify higher degree polynomials in A.

Continuation

- In other words, there exists a non-zero polynomial $q(x) = b_0 + b_1 x + \cdots + b_m x^m$ of degree at most n^2 such that

$$q(A) = b_0 I + b_1 A + b_2 A^2 + \cdots + b_m A^m = 0.$$

- Assume $b_m \neq 0$. Then $A^m = -\frac{1}{b_m}(b_0 I + b_1 A + \cdots + b_{m-1} A^{(m-1)})$.
- This may help us to compute higher powers of A or to simplify higher degree polynomials in A.
- So we would look for a non-zero polynomial q of lowest degree satisfying $q(A) = 0$.

Continuation

- In other words, there exists a non-zero polynomial $q(x) = b_0 + b_1 x + \cdots + b_m x^m$ of degree at most n^2 such that

$$q(A) = b_0 I + b_1 A + b_2 A^2 + \cdots + b_m A^m = 0.$$

- Assume $b_m \neq 0$. Then $A^m = -\frac{1}{b_m}(b_0 I + b_1 A + \cdots + b_{m-1} A^{(m-1)})$.
- This may help us to compute higher powers of A or to simplify higher degree polynomials in A.
- So we would look for a non-zero polynomial q of lowest degree satisfying $q(A) = 0$.
- We may scale such a polynomial to make the leading coefficient one, i. e. we may take it to be monic.

Annihilating polynomials and division algorithm

- Definition 32.1: A polynomial f is said to be annihilating for a matrix A if $f(A) = 0$.

Annihilating polynomials and division algorithm

- Definition 32.1: A polynomial f is said to be annihilating for a matrix A if $f(A) = 0$.
- Theorem 32.2: Let f, g be non-zero annihilating polynomials of a matrix A and suppose degree $(g) \leq$ degree (f). Then

$$f(x) = g(x)s(x) + r(x)$$

for some polynomials s, r, where either $r = 0$ or degree $(r) <$ degree (g) and $r(A) = 0$.

Annihilating polynomials and division algorithm

- Definition 32.1: A polynomial f is said to be annihilating for a matrix A if $f(A) = 0$.
- Theorem 32.2: Let f, g be non-zero annihilating polynomials of a matrix A and suppose degree $(g) \leq$ degree (f). Then

$$f(x) = g(x)s(x) + r(x)$$

  for some polynomials s, r, where either $r = 0$ or degree $(r) <$ degree (g) and $r(A) = 0$.
- Proof: This is clear from the division algorithm for polynomials. As $f(A) = 0 = g(A).s(A)$, we get $r(A) = 0$.

Minimal polynomial

- Theorem 32.3: Let A be an $n \times n$ complex matrix. Then there exists a unique monic polynomial q of lowest degree such that $q(A) = 0$.

Minimal polynomial

- Theorem 32.3: Let A be an $n \times n$ complex matrix. Then there exists a unique monic polynomial q of lowest degree such that $q(A) = 0$.
- Proof: Suppose q_1, q_2 are two distinct non-zero monic polynomials of lowest degree such that $q_1(A) = q_2(A) = 0$.

Minimal polynomial

- Theorem 32.3: Let A be an $n \times n$ complex matrix. Then there exists a unique monic polynomial q of lowest degree such that $q(A) = 0$.
- Proof: Suppose q_1, q_2 are two distinct non-zero monic polynomials of lowest degree such that $q_1(A) = q_2(A) = 0$.
- Then clearly $q_1 - q_2$ is a lower degree polynomial with $(q_1 - q_2)(A) = 0$.

Minimal polynomial

- Theorem 32.3: Let A be an $n \times n$ complex matrix. Then there exists a unique monic polynomial q of lowest degree such that $q(A) = 0$.
- Proof: Suppose q_1, q_2 are two distinct non-zero monic polynomials of lowest degree such that $q_1(A) = q_2(A) = 0$.
- Then clearly $q_1 - q_2$ is a lower degree polynomial with $(q_1 - q_2)(A) = 0$.
- We may scale it suitably to make it monic. This contradicts minimality of q_1, q_2. ■

Factorization

- Definition 32.4: Given a matrix A, the unique monic polynomial of lowest degree q, satisfying $q(A) = 0$ is defined as the minimal polynomial of A.

Factorization

- Definition 32.4: Given a matrix A, the unique monic polynomial of lowest degree q, satisfying $q(A) = 0$ is defined as the minimal polynomial of A.
- Theorem 32.5: Let A be an $n \times n$ complex matrix. Let q be the minimal polynomial of A. Suppose f is an annihilating polynomial of A, then there exists a polynomial s such that $f(x) = q(x)s(x)$. In other words, the minimal polynomial is a factor of every annihilating polynomial.

Factorization

- Definition 32.4: Given a matrix A, the unique monic polynomial of lowest degree q, satisfying $q(A) = 0$ is defined as the minimal polynomial of A.
- Theorem 32.5: Let A be an $n \times n$ complex matrix. Let q be the minimal polynomial of A. Suppose f is an annihilating polynomial of A, then there exists a polynomial s such that $f(x) = q(x)s(x)$. In other words, the minimal polynomial is a factor of every annihilating polynomial.
- Proof: This is clear from the minimality of q and the division algorithm on dividing f by q. $\blacksquare$

Example-I

- Example 32.6: Consider

$$C = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Example-I

- Example 32.6: Consider

$$C = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

- Then for any polynomial f,

$$f(C) = \begin{bmatrix} f(2) & 0 & 0 \\ 0 & f(2) & 0 \\ 0 & 0 & f(3) \end{bmatrix}.$$

Example-I

- Example 32.6: Consider

$$C = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

- Then for any polynomial f,

$$f(C) = \begin{bmatrix} f(2) & 0 & 0 \\ 0 & f(2) & 0 \\ 0 & 0 & f(3) \end{bmatrix}.$$

- Therefore, f is an annihilating polynomial for C if and only if $f(2) = f(3) = 0$.

Example-I

- Example 32.6: Consider

$$C = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

- Then for any polynomial f,

$$f(C) = \begin{bmatrix} f(2) & 0 & 0 \\ 0 & f(2) & 0 \\ 0 & 0 & f(3) \end{bmatrix}.$$

- Therefore, f is an annihilating polynomial for C if and only if $f(2) = f(3) = 0$.
- In particular, the unique minimal polynomial of C is given by $q(x) = (x-2)(x-3) = x^2 - 5x + 6$.

Example -II

- Example 32.7: Consider

$$D = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Example -II

- Example 32.7: Consider

$$D = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

- Now the unique minimal polynomial of D is given by $q(x) = (x-2)^2(x-3)$.

Eigenvalues and annihilating polynomials

- Theorem 32.8: Suppose A is a complex matrix and a is an eigenvalue of A. If f is an annihilating polynomial of A then $f(a) = 0$. In particular, every eigenvalue is a root of the minimal polynomial.

Eigenvalues and annihilating polynomials

- Theorem 32.8: Suppose A is a complex matrix and a is an eigenvalue of A. If f is an annihilating polynomial of A then $f(a) = 0$. In particular, every eigenvalue is a root of the minimal polynomial.
- Proof: Suppose v is an eigenvector of A with eigenvalue a.

Eigenvalues and annihilating polynomials

- Theorem 32.8: Suppose A is a complex matrix and a is an eigenvalue of A. If f is an annihilating polynomial of A then $f(a) = 0$. In particular, every eigenvalue is a root of the minimal polynomial.
- Proof: Suppose v is an eigenvector of A with eigenvalue a.
- Clearly, $A^k v = a^k v$ for every k.

Eigenvalues and annihilating polynomials

- Theorem 32.8: Suppose A is a complex matrix and a is an eigenvalue of A. If f is an annihilating polynomial of A then $f(a) = 0$. In particular, every eigenvalue is a root of the minimal polynomial.
- Proof: Suppose v is an eigenvector of A with eigenvalue a.
- Clearly, $A^k v = a^k v$ for every k.
- Hence for any polynomial f,

$$f(A)v = f(a)v.$$

Eigenvalues and annihilating polynomials

- Theorem 32.8: Suppose A is a complex matrix and a is an eigenvalue of A. If f is an annihilating polynomial of A then $f(a) = 0$. In particular, every eigenvalue is a root of the minimal polynomial.
- Proof: Suppose v is an eigenvector of A with eigenvalue a.
- Clearly, $A^k v = a^k v$ for every k.
- Hence for any polynomial f,

$$f(A)v = f(a)v.$$

- Since $v \neq 0$, if $f(A)v = 0$ then $f(a) = 0$. Now the result is immediate. ■

Eigenvalues and annihilating polynomials

- Theorem 32.8: Suppose A is a complex matrix and a is an eigenvalue of A. If f is an annihilating polynomial of A then $f(a) = 0$. In particular, every eigenvalue is a root of the minimal polynomial.
- Proof: Suppose v is an eigenvector of A with eigenvalue a.
- Clearly, $A^k v = a^k v$ for every k.
- Hence for any polynomial f,

$$f(A)v = f(a)v.$$

- Since $v \neq 0$, if $f(A)v = 0$ then $f(a) = 0$. Now the result is immediate. ■
- Now we may guess the following result.

Cayley Hamilton theorem

- Theorem 32.9 (Cayley Hamilton theorem): Let A be a complex $n \times n$ matrix and let p be the characteristic polynomial of A. Then

$$p(A) = 0.$$

In other words, the characteristic polynomial is an annihilating polynomial for A.

Cayley Hamilton theorem

- Theorem 32.9 (Cayley Hamilton theorem): Let A be a complex $n \times n$ matrix and let p be the characteristic polynomial of A. Then

$$p(A) = 0.$$

  In other words, the characteristic polynomial is an annihilating polynomial for A.
- Corollary 32.9: For any square matrix, the minimal polynomial is a factor of the characteristic polynomial.

A wrong proof

- Wrong proof: By the definition of the characteristic polynomial:
$$p(x) = \det(xI - A).$$

A wrong proof

- Wrong proof: By the definition of the characteristic polynomial:
$$p(x) = \det(xI - A).$$
- Taking $x = A$,
$$p(A) = \det(A.I - A) = \det(A - A) = \det(0) = 0. \qquad (1)$$

A wrong proof

- Wrong proof: By the definition of the characteristic polynomial:
$$p(x) = \det(xI - A).$$
- Taking $x = A$,
$$p(A) = \det(A.I - A) = \det(A - A) = \det(0) = 0. \qquad (1)$$
- This is a wrong proof, as in the equation above, on the left we have a matrix, where as, on the right we have a scalar.

A wrong proof

- Wrong proof: By the definition of the characteristic polynomial:

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

- Taking $x = A$,

$$p(A) = \det(A.I - A) = \det(A - A) = \det(0) = 0. \quad (1)$$

- This is a wrong proof, as in the equation above, on the left we have a matrix, where as, on the right we have a scalar.
- We can't blindly substitute $x = A$ and do determinant computations.

A wrong proof

- Wrong proof: By the definition of the characteristic polynomial:
$$p(x) = \det(xI - A).$$
- Taking $x = A$,
$$p(A) = \det(A.I - A) = \det(A - A) = \det(0) = 0. \quad (1)$$
- This is a wrong proof, as in the equation above, on the left we have a matrix, where as, on the right we have a scalar.
- We can't blindly substitute $x = A$ and do determinant computations.
- END OF REVIEW.

Similarity instead of unitary equivalence

- Example 33.1: Consider

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$$

Similarity instead of unitary equivalence

- Example 33.1: Consider

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$$

- Then clearly A is not normal. Hence A can not be diagonalized using unitary equivalence.

Similarity instead of unitary equivalence

- Example 33.1: Consider

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$$

- Then clearly A is not normal. Hence A can not be diagonalized using unitary equivalence.
- However, $\sigma(A) = \{1, 3\}$ and since the corresponding geometric multiplicities are at least 1, we can get a basis of eigenvectors of A. In other words, there exists an invertible matrix S such that

$$A = S \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} S^{-1}.$$

Similarity instead of unitary equivalence

- Example 33.1: Consider

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$$

- Then clearly A is not normal. Hence A can not be diagonalized using unitary equivalence.
- However, $\sigma(A) = \{1, 3\}$ and since the corresponding geometric multiplicities are at least 1, we can get a basis of eigenvectors of A. In other words, there exists an invertible matrix S such that

$$A = S \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} S^{-1}.$$

Similarity instead of unitary equivalence

- Example 33.1: Consider

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$$

- Then clearly A is not normal. Hence A can not be diagonalized using unitary equivalence.
- However, $\sigma(A) = \{1, 3\}$ and since the corresponding geometric multiplicities are at least 1, we can get a basis of eigenvectors of A. In other words, there exists an invertible matrix S such that

$$A = S \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} S^{-1}.$$

- This shows that some times it maybe more prudent not to insist on unitary equivalence. We may try to simplify A through similarity instead of unitary equivalence. This is done either when there is no underlying inner product or when we have a prescribed inner product but we choose to ignore it.

Upper triangular form

- Theorem 33.2: Let A be an $n \times n$ complex matrix. Then there exists an upper triangular matrix T and an invertible matrix S such that
$$A = STS^{-1}.$$

Upper triangular form

- Theorem 33.2: Let A be an $n \times n$ complex matrix. Then there exists an upper triangular matrix T and an invertible matrix S such that
$$A = STS^{-1}.$$
- Proof: We may consider the standard inner product on $\mathbb{C}^n$. Then by Schur's upper triangularization theorem, there exists a unitary U and an upper triangular matrix T such that
$$A = UTU^*.$$
Take $S = U$. Since $U^* = S^{-1}$, the proof is complete.

Upper triangular form

- Theorem 33.2: Let A be an $n \times n$ complex matrix. Then there exists an upper triangular matrix T and an invertible matrix S such that
$$A = STS^{-1}.$$
- Proof: We may consider the standard inner product on $\mathbb{C}^n$. Then by Schur's upper triangularization theorem, there exists a unitary U and an upper triangular matrix T such that
$$A = UTU^*.$$
Take $S = U$. Since $U^* = S^{-1}$, the proof is complete.
- Alternatively, we may imitate the proof of Schur's upper triangularization theorem. Choose an eigenvector v_1 corresponding to some eigenvalue a_1 of A, extend $\{v_1\}$ to a basis of $\mathbb{C}^n$.

Continuation

- In the new basis, the linear map A will have the form:

$$A = \begin{bmatrix} a_1 & y \\ 0 & B \end{bmatrix}$$

for some $1 \times (n-1)$ row vector y and $(n-1) \times (n-1)$ matrix B. Now use induction. $\blacksquare$.

A Lemma

- Lemma 33.3: Let T be an upper triangular matrix with diagonal entries $d_1, d_2, \ldots, d_n$. For $1 \leq k \leq n$, take

$$M_k = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} : x_1, x_2, \ldots, x_k \in \mathbb{C} \right\}.$$

Take $M_0 = \{0\}$. Then for every $1 \leq k \leq n$,

$$(T - d_k I)(M_k) \subseteq M_{k-1}.$$

A Lemma

- Lemma 33.3: Let T be an upper triangular matrix with diagonal entries $d_1, d_2, \ldots, d_n$. For $1 \leq k \leq n$, take

$$M_k = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} : x_1, x_2, \ldots, x_k \in \mathbb{C} \right\}.$$

Take $M_0 = \{0\}$. Then for every $1 \leq k \leq n$,

$$(T - d_k I)(M_k) \subseteq M_{k-1}.$$

- Proof: Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{C}^n$.

A Lemma

- Lemma 33.3: Let T be an upper triangular matrix with diagonal entries $d_1, d_2, \ldots, d_n$. For $1 \leq k \leq n$, take

$$M_k = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} : x_1, x_2, \ldots, x_k \in \mathbb{C} \right\}.$$

Take $M_0 = \{0\}$. Then for every $1 \leq k \leq n$,

$$(T - d_k I)(M_k) \subseteq M_{k-1}.$$

- Proof: Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{C}^n$.
- Then $M_k = \operatorname{span}\{e_1, e_2, \ldots, e_k\}$.

A Lemma

- Lemma 33.3: Let T be an upper triangular matrix with diagonal entries $d_1, d_2, \ldots, d_n$. For $1 \leq k \leq n$, take

$$M_k = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} : x_1, x_2, \ldots, x_k \in \mathbb{C} \right\}.$$

Take $M_0 = \{0\}$. Then for every $1 \leq k \leq n$,

$$(T - d_k I)(M_k) \subseteq M_{k-1}.$$

- Proof: Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{C}^n$.
- Then $M_k = \text{span}\{e_1, e_2, \ldots, e_k\}$.
- Since T is upper triangular $(T - d_k I)$ is also upper triangular with k-th diagonal entry equal to 0.

Continuation

- In particular, the j-th column of $(T - d_k I)$ is in the span of $\{e_1, e_2, \ldots, e_{k-1}\}$ for $1 \leq j \leq k$.

Continuation

- In particular, the j-th column of $(T - d_k I)$ is in the span of $\{e_1, e_2, \ldots, e_{k-1}\}$ for $1 \leq j \leq k$.
- In other words $(T - d_k I)e_j \in M_{k-1}$ for $1 \leq k \leq n$. $\blacksquare$

Cayley-Hamilton Theorem

- Now we present a proof of this famous theorem.

Cayley-Hamilton Theorem

- Now we present a proof of this famous theorem.
- Theorem 32.9 (Cayley Hamilton theorem): Let A be a complex $n \times n$ matrix and let p be the characteristic polynomial of A. Then

$$p(A) = 0.$$

  In other words, the characteristic polynomial is an annihilating polynomial for A.

Cayley-Hamilton Theorem

- Now we present a proof of this famous theorem.
- Theorem 32.9 (Cayley Hamilton theorem): Let A be a complex $n \times n$ matrix and let p be the characteristic polynomial of A. Then
$$p(A) = 0.$$
In other words, the characteristic polynomial is an annihilating polynomial for A.
- Proof: By Theorem 33.2, there exists a non-singular matrix S and an upper triangular matrix T such that
$$A = STS^{-1}.$$

Cayley-Hamilton Theorem

- Now we present a proof of this famous theorem.
- Theorem 32.9 (Cayley Hamilton theorem): Let A be a complex $n \times n$ matrix and let p be the characteristic polynomial of A. Then

$$p(A) = 0.$$

  In other words, the characteristic polynomial is an annihilating polynomial for A.
- Proof: By Theorem 33.2, there exists a non-singular matrix S and an upper triangular matrix T such that

$$A = STS^{-1}.$$

- Note that for any polynomial f,

$$f(A) = Sf(T)S^{-1}.$$

Cayley-Hamilton Theorem

- Now we present a proof of this famous theorem.
- Theorem 32.9 (Cayley Hamilton theorem): Let A be a complex $n \times n$ matrix and let p be the characteristic polynomial of A. Then

  $$p(A) = 0.$$

  In other words, the characteristic polynomial is an annihilating polynomial for A.
- Proof: By Theorem 33.2, there exists a non-singular matrix S and an upper triangular matrix T such that

  $$A = STS^{-1}.$$

- Note that for any polynomial f,

  $$f(A) = Sf(T)S^{-1}.$$

- Let p be the characteristic polynomial of A and let $d_1, \dots, d_n$ be the diagonal entries of T.

Continuation

- Then p is also the characteristic polynomial of T and

$$p(x) = (x - d_1)(x - d_2) \cdots (x - d_n).$$

Continuation

- Then p is also the characteristic polynomial of T and

$$p(x) = (x - d_1)(x - d_2)\cdots(x - d_n).$$

- As $p(A) = Sp(T)S^{-1}$, it suffices to show that $p(T) = 0$.

Continuation

- Then p is also the characteristic polynomial of T and

$$p(x) = (x - d_1)(x - d_2)\cdots(x - d_n).$$

- As $p(A) = Sp(T)S^{-1}$, it suffices to show that $p(T) = 0$.
- We use the notation of previous lemma. Consider any $x \in \mathbb{C}^n = M_n$.

Continuation

- Then p is also the characteristic polynomial of T and

$$p(x) = (x - d_1)(x - d_2) \cdots (x - d_n).$$

- As $p(A) = Sp(T)S^{-1}$, it suffices to show that $p(T) = 0$.
- We use the notation of previous lemma. Consider any $x \in \mathbb{C}^n = M_n$.
- By the lemma

$$(T - d_n I)x \in M_{n-1}.$$

Continuation

- Then p is also the characteristic polynomial of T and

$$p(x) = (x - d_1)(x - d_2) \cdots (x - d_n).$$

- As $p(A) = Sp(T)S^{-1}$, it suffices to show that $p(T) = 0$.
- We use the notation of previous lemma. Consider any $x \in \mathbb{C}^n = M_n$.
- By the lemma

$$(T - d_n I)x \in M_{n-1}.$$

- As $(T - d_{n-1} I)M_{n-1} \subseteq M_{n-2}$ we get

$$(T - d_{n-1} I)(T - d_n I)x \in M_{n-2}.$$

Continuation

- Continuing this way (i.e., by induction) :

$$(T - d_1 I)(T - d_2 I)\cdots(T - d_n I)x \in M_0 = \{0\}.$$

Continuation

- Continuing this way (i.e., by induction) :

$$(T - d_1 I)(T - d_2 I) \cdots (T - d_n I)x \in M_0 = \{0\}.$$

- In other words, $p(T)x = 0$ for every $x \in \mathbb{C}^n$.

Continuation

- Continuing this way (i.e., by induction) :

$$(T - d_1 I)(T - d_2 I) \cdots (T - d_n I)x \in M_0 = \{0\}.$$

- In other words, $p(T)x = 0$ for every $x \in \mathbb{C}^n$.
- This proves the claim. ■

Example

- Example 33.4 Suppose D is a diagonal matrix with diagonal entries $d_1, d_2, \ldots, d_n$. Then the characteristic polynomial of D is given by

$$p(x) = (x - d_1)(x - d_2) \cdots (x - d_n).$$

Example

- Example 33.4 Suppose D is a diagonal matrix with diagonal entries $d_1, d_2, \ldots, d_n$. Then the characteristic polynomial of D is given by

$$p(x) = (x - d_1)(x - d_2)\cdots(x - d_n).$$

- It is clear that $p(D) = 0$.

Example

- Example 33.4 Suppose D is a diagonal matrix with diagonal entries $d_1, d_2, \ldots, d_n$. Then the characteristic polynomial of D is given by

$$p(x) = (x - d_1)(x - d_2)\cdots(x - d_n).$$

- It is clear that $p(D) = 0$.
- Corollary 33.5: Suppose A is an $n \times n$ matrix. Then the dimension of

$$\mathcal{A} = \{f(A) : f \text{ is a polynomial}\}$$

Example

- Example 33.4 Suppose D is a diagonal matrix with diagonal entries $d_1, d_2, \ldots, d_n$. Then the characteristic polynomial of D is given by

$$p(x) = (x - d_1)(x - d_2) \cdots (x - d_n).$$

- It is clear that $p(D) = 0$.
- Corollary 33.5: Suppose A is an $n \times n$ matrix. Then the dimension of

$$\mathcal{A} = \{f(A) : f \text{ is a polynomial}\}$$

- is at most n.

Example

- Example 33.4 Suppose D is a diagonal matrix with diagonal entries $d_1, d_2, \ldots, d_n$. Then the characteristic polynomial of D is given by

$$p(x) = (x - d_1)(x - d_2)\cdots(x - d_n).$$

- It is clear that $p(D) = 0$.
- Corollary 33.5: Suppose A is an $n \times n$ matrix. Then the dimension of

$$\mathcal{A} = \{f(A) : f \text{ is a polynomial}\}$$

- is at most n.
- Proof: This is now clear, as the Cayley Hamilton theorem tells us that A^n is a linear combination of $\{I, A, \ldots, A^{n-1}\}$.

Example

- Example 33.4 Suppose D is a diagonal matrix with diagonal entries $d_1, d_2, \ldots, d_n$. Then the characteristic polynomial of D is given by

$$p(x) = (x - d_1)(x - d_2) \cdots (x - d_n).$$

- It is clear that $p(D) = 0$.
- Corollary 33.5: Suppose A is an $n \times n$ matrix. Then the dimension of

$$\mathcal{A} = \{f(A) : f \text{ is a polynomial}\}$$

- is at most n.
- Proof: This is now clear, as the Cayley Hamilton theorem tells us that A^n is a linear combination of $\{I, A, \ldots, A^{n-1}\}$.
- It is then easy to see that A^m for $m \geq n$ are also in the span of $\{I, A, \ldots, A^{n-1}\}$.

Multivariable case

- Remark 33.6: Suppose A, B are commuting $n \times n$ matrices. What is the maximum possible dimension of $\{f(A, B) : f \text{ is a two variable polynomial}\}$?

Multivariable case

- Remark 33.6: Suppose A, B are commuting $n \times n$ matrices. What is the maximum possible dimension of $\{f(A, B) : f \text{ is a two variable polynomial}\}$?
- Surprisingly the answer is still n. This is known as Gerstenhaber's theorem.

Multivariable case

- Remark 33.6: Suppose A, B are commuting $n \times n$ matrices. What is the maximum possible dimension of $\{f(A, B) : f \text{ is a two variable polynomial}\}$?
- Surprisingly the answer is still n. This is known as Gerstenhaber's theorem.
- Question: Suppose A, B, C are commuting $n \times n$ matrices. Is the maximum possible dimension of

$$\{f(A, B, C) : f \text{ is a three variable polynomial}\}$$

Multivariable case

- Remark 33.6: Suppose A, B are commuting $n \times n$ matrices. What is the maximum possible dimension of $\{f(A, B) : f \text{ is a two variable polynomial}\}$?
- Surprisingly the answer is still n. This is known as Gerstenhaber's theorem.
- Question: Suppose A, B, C are commuting $n \times n$ matrices. Is the maximum possible dimension of

$$\{f(A, B, C) : f \text{ is a three variable polynomial}\}$$

- equal to n?

Multivariable case

- Remark 33.6: Suppose A, B are commuting $n \times n$ matrices. What is the maximum possible dimension of $\{f(A, B) : f \text{ is a two variable polynomial}\}$?
- Surprisingly the answer is still n. This is known as Gerstenhaber's theorem.
- Question: Suppose A, B, C are commuting $n \times n$ matrices. Is the maximum possible dimension of

$$\{f(A, B, C) : f \text{ is a three variable polynomial}\}$$

- equal to n?
- Ans: Not known.

References (Optional)

- (1) B. Sethuraman, The algebra generated by three commuting matrices, Ramanujan Math. Soc. Newsletter, 21 number 2, September 2011, 26-31.

References (Optional)

- (1) B. Sethuraman, The algebra generated by three commuting matrices, Ramanujan Math. Soc. Newsletter, 21 number 2, September 2011, 26-31.
- (2) Holbrook and C.O'Meara, Some thoughts on Gerstenhaber's theorem, Linear Algebra and its Applications Volume 466, 1 February 2015, Pages 267-295.

References (Optional)

- (1) B. Sethuraman, The algebra generated by three commuting matrices, Ramanujan Math. Soc. Newsletter, 21 number 2, September 2011, 26-31.
- (2) Holbrook and C.O'Meara, Some thoughts on Gerstenhaber's theorem, Linear Algebra and its Applications Volume 466, 1 February 2015, Pages 267-295.

- It is known that the dimension of the algebra generated can be more than n, when there are four or more commuting matrices.

References (Optional)

- (1) B. Sethuraman, The algebra generated by three commuting matrices, Ramanujan Math. Soc. Newsletter, 21 number 2, September 2011, 26-31.
- (2) Holbrook and C.O'Meara, Some thoughts on Gerstenhaber's theorem, Linear Algebra and its Applications Volume 466, 1 February 2015, Pages 267-295.

- It is known that the dimension of the algebra generated can be more than n, when there are four or more commuting matrices.
- END OF LECTURE 33.