LINEAR ALGEBRA -II

B V Rajarama Bhat

Indian Statistical Institute, Bangalore

Lecture 32: Polynomials of a matrix

- 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 32: Polynomials of a matrix

- 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 32: Polynomials of a matrix

- 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 32: Polynomials of a matrix

- 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 32: Polynomials of a matrix

- 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.
- 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. ■

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 $x = A$,

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

A wrong proof

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

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

- Taking $x = A$,

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

- 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.

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 $x = A$,
$$p(A) = \det(A.I - A) = \det(A - A) = \det(0) = 0. \qquad (1)$$
- 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.
- We can't blindly substitute $x = A$ and do determinant computations.
- We will see a correct proof in the next lecture.

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.
- We can't blindly substitute $x = A$ and do determinant computations.
- We will see a correct proof in the next lecture.
- END OF LECTURE 32.