MTH 100 : Lecture 32

Polynomials Applied to Matrices

- **Matrix Powers:** If A is an $n \times n$ matrix (we may take the entries either real or complex), then the product matrix $A.A$ is well-defined and can be written as A^2, $A.A.A = A^3$, ...
  In general, $A^m = A.A \ldots A$ (m times) for any positive integer m.

  For convenience, we define $A^0 = I_n$, the identity matrix.

- If A is invertible and A^{-1} is its inverse, then for any positive integer m, $\boxed{(A^m)^{-1} = (A^{-1})^m}$

- **Remark:**
  $$A^i . A^j = A^{i+j} \text{ and } (A^i)^j = A^{ij}$$
  where i, j can be arbitrary integers if A is invertible and non-negative integers if A is not invertible.

Definition: If p is a polynomial given by
$$p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_m t^m$$
and A is an $n \times n$ matrix, then $p(A)$ is the matrix given by
$$p(A) = a_0 I + a_1 A + a_2 A^2 + \cdots + a_m A^m$$

- Note that this is a new use of the symbol p because we are applying it to matrices and not just scalars.
- If p and q are two polynomials, then

$$pq(A) = p(A)q(A) = q(A)p(A) = (qp)(A)$$

(where pq is the polynomial defined by
$pq(t) = p(t)q(t)$ (usual multiplication of polynomials))

The Minimal Polynomial of a Matrix

Definition: Given an $n \times n$ matrix A, the minimal polynomial of A is the (non zero) monic polynomial of minimal degree such that $p(A) = 0$ (i.e. the zero matrix)

(Monic polynomial means the nonzero coefficient of highest power of the variable is equal to 1.
The monic condition is inserted so as to make the minimal polynomial unique.)

Note: Every square matrix must have a minimal polynomial:

Suppose A is an $n \times n$ matrix with entries from a field F, then the set $\{I, A, A^2, \dots, A^{n^2}\}$ cannot be linearly independant because this set has $\boxed{n^2+1}$ matrices and $\dim(F^{n\times n}) = \boxed{n^2}$

Let m be the smallest positive integer such that $\{I, A, A^2, \dots, A^m\}$ is linearly dependent. Then A^m is a linear combination of the preceding matrices.

Thus there exist scalars $a_0, a_1, \dots, a_{m-1}$ such that

$$\boxed{a_0 I + a_1 A + a_2 A^2 + \dots + a_{m-1} A^{m-1} + A^m = 0}$$

Note: Reference: Problem 5 of Worksheet 6.

A Famous Result

Theorem (Caley - Hamilton Theorem):

Let q denote the characteristic polynomial of an $n\times n$ matrix A.

Then $\boxed{q(A) = 0}$

Corollary: The degree of minimal polynomial of any $n \times n$ matrix is atmost n.

(Recall that the degree of the characteristic polynomial of A is n.)

Note: We will omit the proof of Caley-Hamilton Theorem. You may refer to advanced textbooks.

Remark:

Using Remainder Theorem for Polynomial Division, we can see that if $p(x)$ is the minimal polynomial of A and if $q(x)$ is any other polynomial satisfied by A, then $p(x)$ divides $q(x)$.

Review of Polynomials:

- We use the notation $F[t]$ to indicate the vector space of polynomials with coefficients from the field F (F could be either $\mathbb{R}$ or $\mathbb{C}$)
- $\lambda \in F$ is called a root of a polynomial $p(t)$ if $p(\lambda) = 0$

- **Lemma:** Suppose $p \in F[t]$ is a polynomial of degree $m \geq 1$.
  Then λ is a root of p if and only if there exists a polynomial $q \in F[t]$ with degree $(m-1)$ such that
  $$\boxed{p(t) = (t-\lambda)\, q(t)}$$

- **Lemma:** Suppose $p \in F[t]$ is a polynomial of degree $m \geq 0$, then p has atmost m distinct roots in F.

- **Lemma (Division Algorithm or Remainder Theorem):**
  Suppose $p, q \in F[t]$ with $p \neq 0$.
  Then there exist polynomials $r, s \in F[t]$ with $q(t) = p(t)\,s(t) + r(t)$ and either $r = 0$ or $\deg r < \deg p$.

- **Fundamental Theorem of Algebra:**
  Suppose $p \in \mathbb{C}[t]$ is a polynomial of degree $m \geq 1$. Then p has a root.
  Furthermore, p has a factorization of the form $\boxed{p(t) = c(t-\lambda_1)(t-\lambda_2)\cdots\cdots(t-\lambda_m)}$

Lemma: Suppose $q \in F[t]$.

Then $q(A)=0$ if and only if the minimal polynomial of A divides q.

Proof: $\Rightarrow$: Suppose $q(A)=0$

Let $p(t)$ be the minimal polynomial of A.

Then $p \neq 0$.

Using Remainder Theorem, we can write

$$q(t) = p(t)s(t) + r(t) \quad \ldots\ldots (1)$$

where either $r=0$ or $\deg r < \deg p$

Evaluating (1) at A,

$$q(A) = p(A)s(A) + r(A)$$

Since $\underline{q(A)=0}$, $\quad p(A)s(A) + r(A) = 0 \quad \ldots\ldots (2)$

Now, since p is the minimal polynomial of A, $p(A)=0$ and hence from (2), $r(A)=0$

But this is not possible unless $r=0$ because $\deg r < \deg p$ and p is the minimal polynomial of A

Therefore from (1), $q(t) = p(t)s(t)$ and so p divides q.

$\Leftarrow$ If $p(t)$ is the minimal polynomial of A, then $p(A) = 0$.

If $p(t)$ divides $q(t)$, there exists a polynomial $s(t)$ such that

$$q(t) = p(t)s(t)$$

Evaluating at A,

$$q(A) = p(A)s(A) = 0 \cdot s(A) = 0$$

$$\Rightarrow \boxed{q(A) = 0}$$

(QED).

Diagonalization of Matrices:

- If A is a diagonal matrix, then its diagonal elements are its eigen values and the standard basis vectors are its eigen vectors.
- i.e. if $A = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$ then $Ae_i = \lambda_i e_i$ where $e_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \to$ ith entry

for $i = 1, 2, \ldots, n$

Definition:

An $n \times n$ matrix A is said to be diagonalizable if A is similar to a diagonal matrix D. i.e. if there exists an invertible matrix P and a diagonal matrix D such that

$$\boxed{A = PDP^{-1}}$$

Note: If A is diagonalizable then its powers are easy to compute.

Note: If A is diagonalizable, then its eigen values can be found by inspection of D. However, in practice, we have to do things the other way round. First, we find the eigenvalues from the characteristic equation, then we find P and the diagonal matrix D.