Lecture 34

Andrei Antonenko

May 5, 2003

1 Functions of the operators

Let $\mathcal{A}$ be a linear operator. If $f(t)$ is a polynomial,

$$f(t) = a_m t^m + a_{m-1}t^{m-1} + \cdots + a_1 t + a_0,$$

then we can define $f(\mathcal{A})$ as

$$f(\mathcal{A}) = a_m\mathcal{A}^m + a_{m-1}\mathcal{A}^{m-1} + \cdots + a_1\mathcal{A} + a_0\mathcal{I}.$$

Since the space of linear operators is finite-dimensional (we are considering only finite-dimensional vector spaces here), then there is only a finite number of linearly independent powers of $\mathcal{A}$. Thus, there exists polynomial $f(t)$, such that $f(\mathcal{A}) = 0$. Such polynomials are called **annihilating polynomials**. The annihilating polynomials of minimal degree are called **minimal polynomials** of the operator $\mathcal{A}$. We will denote minimal polynomials as $m_{\mathcal{A}}(t)$.

Example 1.1. *The minimal polynomial of the zero operator is*

$$m_0(t) = t.$$

Example 1.2. *The minimal polynomial of the identity operator $\mathcal{I}$ is*

$$m_{\mathcal{I}}(t) = t - 1.$$

Lemma 1.3. *The minimal polynomial of the Jordan block of the size m with the eigenvalue λ is equal to $(t-\lambda)^m$.*

Proof. Let $\mathcal{A}$ be a linear operator, given by this Jordan block. Then $\mathcal{N} = \mathcal{A} - \lambda\mathcal{I}$ is a nilpotent operator of the height m, i.e.

$$(\mathcal{A} - \lambda\mathcal{I})^m = 0, \qquad (\mathcal{A} - \lambda\mathcal{I})^{m-1} \neq 0.$$

Thus $(t-\lambda)^m$ is an annihilating polynomial, but none of the smaller powers of $(t-\lambda)$ is not an annihilating polynomial. Thus $(t-\lambda)^m$ is a minimal polynomial. $\square$

Now, if the space V is decomposed into a direct sum of invariant subspaces of the operator $\mathcal{A}$, then the minimal polynomial of $\mathcal{A}$ is equal to the least common multiple of the minimal polynomials of the restrictions of $\mathcal{A}$ to these subspaces.

Now, let's assume that $\mathcal{A}$ is a linear operator, and it's characteristic polynomial can be factored into linear terms. Let $\lambda_1, \lambda_2, \ldots, \lambda_s$ be all the roots of the characteristic polynomial. Then the following theorem is an immediate corollary of the previous remark and lemma:

Theorem 1.4. *The minimal polynomial of the operator $\mathcal{A}$ is equal to*

$$m_{\mathcal{A}}(t) = \prod_{i=1}^{s} (t - \lambda_i)^{m_i}, \tag{1}$$

where m_i is a maximal size of Jordan blocks with the eigenvalue λ_i in the Jordan canonical form of $\mathcal{A}$.

Example 1.5. *The minimal polynomial of the operator $\mathcal{A}$ with Jordan canonical form*

$$J = \begin{pmatrix} 3 & 1 & 0 & & & & \\ 0 & 3 & 1 & & & & \\ 0 & 0 & 3 & & & & \\ & & & 3 & 1 & & \\ & & & 0 & 3 & & \\ & & & & & 2 & 1 \\ & & & & & 0 & 2 \end{pmatrix}$$

is $(t-3)^3(t-2)^2$.

Corollary 1.6. *The Jordan canonical form of the operator is diagonal only if its minimal polynomial does not have roots of multiplicity more than 1.*

This theory allows us to solve some interesting problems.

Example 1.7. *Let us describe all the linear operators $\mathcal{A}$ such that $\mathcal{A}^3 = \mathcal{A}^2$. It means that $t^3 - t^2$ is an annihilating polynomial of $\mathcal{A}$, and thus the minimal polynomial of $\mathcal{A}$ divides $t^3 - t^2 = t^2(t-1)$. By the previous theorem it happens only if the Jordan canonical form of the operator has only following Jordan blocks:*

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \end{pmatrix}.$$

The Jordan canonical form can have any number of such blocks. Moreover all the matrices, satisfying the given condition can be described as $C^{-1}JC$, where C is an arbitrary invertible matrix, and J is a Jordan canonical form described above.

The immediate corollary of this theorem is the following fact:

Theorem 1.8 (Cayley-Hamilton theorem)**.** $p_{\mathcal{A}}(\mathcal{A}) = 0$*, i.e. the characteristic polynomial of the operator* $\mathcal{A}$ *annihilates operator* $\mathcal{A}$*.*

Example 1.9. *For any* 2×2 *matrix* $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *we have*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 - (a+d)\begin{pmatrix} a & b \\ c & d \end{pmatrix} + (ad-bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$