LECTURE 7. POLYNOMIALS. CALEY-HAMILTON THEOREM

Please consider proof of every proposition in this lecture except the Fundamental Theorem of Algebra to be an exercise. The notation $\mathbb{F}$ stands for a field.

1. Determinant

Definition 1.1. Assume that V is a finite-dimensional vector space, $A \in \mathrm{End}(V)$. Note that $\dim(\Lambda^{\dim V}(V)) = 1$. Therefore, $\Lambda^{\dim V}(A)$ is a multiple of the identity map. We will define the determinant $\det(A)$ of the linear operator A by the following equality

$$\boxed{\Lambda^{\dim V}(A) = \det(A) \cdot \mathrm{Id}.} \tag{1.1}$$

Proposition 1.2. *Suppose that* $\dim V < \infty$ *and* $A, B \in \mathrm{End}(V)$. *Then*

$$\boxed{\det(AB) = \det(A) \cdot \det(B).} \tag{1.2}$$

Proposition 1.3. *Suppose that* $\dim V = n < \infty$ *and* $A \in \mathrm{End}(V)$ *has a matrix* $(a_i^j)_{1 \le i,j \le n}$ *in some basis. Then*

$$\boxed{\det(A) = \sum_{\rho \in S_n} \mathrm{sgn}(\rho) \prod_{j=1}^{n} a^j_{\rho(j)}.} \tag{1.3}$$

2. Decomposition of polynomials

Definition 2.1. A polynomial in a variable x is a formal sum of the form

$$p = \sum_{j=0}^{k} a_j x^j, \quad k \ge 0,\ a_j \in \mathbb{F}. \tag{2.1}$$

We assume that two sums of that form are equal if they differ only by some monomials with zero coefficients. If $a_k \neq 0$ we say that the degree of the polynomial p $\deg(p) := k$. In that case, we will call a_k the leading coefficient of the polynomial p. If all coefficients of the polynomial equal zero, then we put $\deg(0) := -\infty$. The set of polynomials is a vector space with respect to standard addition and scalar multiplication. We denote this vector space by $\mathbb{F}[x]$. Also, we put

$$\mathbb{F}_n[x] := \{p \in \mathbb{F}[x] \mid \deg(p) \le n\}. \tag{2.2}$$

With every polynomial $p \in \mathbb{F}[x]$ we associate a function $p : \mathbb{F} \to \mathbb{F}$.

Note that $\forall a \in \mathbb{F}, a \neq 0$ we have that $\deg(a) = 0$. Next, we may define Euclidean division for polynomials.

Proposition 2.2. *For every two polynomials* $f, g \in \mathbb{F}[x], g \neq 0$ *there exist unique polynomials* $q, r \in \mathbb{F}[x]$ *such that the following conditions hold:*

(1) $f = q \cdot g + r$;
(2) $\deg(r) < \deg g$.

Date: March 22, 2024.

Definition 2.3. We say that a polynomial $g \in \mathbb{F}[x], g \neq 0$ divides a polynomial $f \in \mathbb{F}[x]$ if there exists a polynomial $q \in \mathbb{F}[x]$ such that $f = g \cdot q$. We write $g \mid f$ in this case. Note that $\deg(f) \geq \deg(g)$ in that case.

For arbitrary polynomials $f_1, \ldots, f_m \in \mathbb{F}[x]$ we are going to denote their greatest common divisor by $GCD(f_1, \ldots, f_m)$. The GCD is determined up to multiplication by a non-zero scalar, so we will assume that its leading coefficient equals one to ensure uniqueness.

Example 2.4. Suppose that

$$p_1(x) = x^3 - x, \quad p_2(x) = x^3 - x^2 + x - 1, \quad p_3(x) = x^2 - 1. \tag{2.3}$$

Then

$$p_3 \mid p_1, \quad p_3 \nmid p_2, \quad GCD(p_1, p_2, p_3) = x - 1. \tag{2.4}$$

Exercise 2.5. Show that

$$((x - a) \mid f) \quad \Longleftrightarrow \quad (f(a) = 0). \tag{2.5}$$

Now, we will state an amazing property. It has its analogs for integers and, more generally, for principal ideals domains.

Proposition 2.6 (Bézout's identity for polynomials)**.** *Suppose that $f_1, \ldots, f_m \in \mathbb{F}[x]$ and $g = GCD(f_1, \ldots, f_m)$. Then there exist $p_1, \ldots, p_m \in \mathbb{F}[x]$ such that*

$$g = \sum_{j=1}^{m} p_j \cdot f_j. \tag{2.6}$$

To prove that result, we may consider the set

$$I := \{\sum_{j=1}^{m} p_j \cdot f_j \mid p_j \in \mathbb{F}[x]\} \tag{2.7}$$

and put g to be a non-zero element of I of minimal degree. Again it is determined up to multiplication by a non-zero scalar.

Proposition 2.6 allows us to prove the following lemma.

Lemma 2.7. *Suppose that $g, p, q \in \mathbb{F}[x]$, $g \mid pq$ and $GCD(g, p) = 1$. Then $g \mid q$.*

Definition 2.8. A polynomial f is called irreducible if

(1) $\deg(f) \geq 1$;
(2) for every $g, q \in \mathbb{F}[x]$ such that $f = g \cdot q$ we have that either $\deg(g) = 0$ or $\deg(q) = 0$.

One may deduce the unique factorization property for polynomials in one variable from Lemma 2.7.

Proposition 2.9. *Every polynomial $f \in \mathbb{F}[x]$ can be factored into a product of a non-zero constant and a finite number of irreducible polynomials:*

$$f = a \prod_{j=1}^{m} q_j, \quad a \in \mathbb{F}, q_j \in \mathbb{F}[x]. \tag{2.8}$$

This decomposition is unique up to the permutation of the factors and the multiplication of the factors by non-zero constants whose product is 1. Again, we may assume that the leading coefficient of every q_j equals one to ensure uniqueness.

Example 2.10. We have the following decompositions

$$2x^3 + x \in \mathbb{R}[x], \quad 2x^3 + x = 2x(x^2 + 1); \tag{2.9}$$

$$2x^3 + x \in \mathbb{C}[x], \quad 2x^3 + x = 2x(x + i)(x - i). \tag{2.10}$$

We are going to take for granted the next claim.

Proposition 2.11 (Fundamental Theorem of Algebra)**.** *Every irreducible complex polynomial equals* $(x - z)$ *for some* $z \in \mathbb{C}$.

We may deduce the analogous result for the field $\mathbb{R}$ from the complex case.

Proposition 2.12. *Every irreducible real polynomial equals either* $(x - a)$ *for some* $a \in \mathbb{R}$*, or* $x^2 + bx + c$*, where* $b^2 - 4c < 0$.

To prove this proposition, one can consider the conjugation in the field $\mathbb{C}$.

3. Characteristic polynomial

Definition 3.1. Let V be a finite-dimensional vector space. For $A \in \mathrm{End}(V)$ we define its characteristic polynomial by

$$\chi_A(\lambda) := \det(\lambda \cdot \mathrm{Id} - A). \tag{3.1}$$

Example 3.2. Suppose that

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}; \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}; \quad C = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}. \tag{3.2}$$

Then

$$\chi_A(\lambda) = \lambda^2 + 1; \quad \chi_B(\lambda) = \lambda^2; \quad \chi_C(\lambda) = (\lambda - \alpha)(\lambda - \beta). \tag{3.3}$$

Exercise 3.3. Suppose that $A \in \mathrm{End}(\mathbb{F}^n)$ is given by the matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & a_0 \\ 1 & 0 & 0 & \dots & 0 & a_1 \\ 0 & 1 & 0 & \dots & 0 & a_2 \\ 0 & 0 & 1 & \dots & 0 & a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & a_{n-1} \end{pmatrix}. \tag{3.4}$$

Find $\chi_A(\lambda)$.

Definition 3.4. Let $A \in \mathrm{End}(V)$. A subspace $U \subseteq V$ is called an invariant subspace of A, or A-invariant if $\forall u \in U$ we have that $Au \in U$. Note that if U is A-invariant, then the restriction of the map A to the subspace U is an endomorphism of U.

Proposition 3.5. *Suppose* U *is an invariant subspace of* $A \in \mathrm{End}(V)$*,* $(u_1, \dots, u_m)$ *is a basis of* U*, and* $\alpha = (u_1, \dots, u_m, w_1, \dots, w_k)$ *is a basis of* V*. Then the matrix of* A *in the basis* α *has the form*

$$\begin{pmatrix} B & C \\ 0 & D \end{pmatrix}, \quad \text{for some } B \in \mathrm{Mat}(m \times m, \mathbb{F}), C \in \mathrm{Mat}(m \times k, \mathbb{F}), D \in \mathrm{Mat}(k \times k, \mathbb{F}). \tag{3.5}$$

Exercise 3.6. Show that

$$\det \begin{pmatrix} B & C \\ 0 & D \end{pmatrix} = \det(B) \cdot \det(D), \tag{3.6}$$

where B and D are square matrices.

From Proposition 3.5 we may deduce the following result about characteristic polynomials.

Proposition 3.7. *Suppose U is invariant with respect to $A \in \mathrm{End}(V)$. Denote by B the restriction of A to the subspace U. Then the characteristic polynomial χ_B divides the characteristic polynomial χ_A.*

4. Caley-Hamilton Theorem

In this section, we begin to consider polynomials evaluated at linear operators.

Definition 4.1. For a polynomial $p(x) = \sum_{j=0}^{k} a_j x^j \in \mathbb{F}[x]$ and a linear operator $B \in \mathrm{End}(V)$ we define another linear operator $p(B) \in \mathrm{End}(V)$ by the following equality:

$$p(B) := a_0 \cdot \mathrm{Id} + \sum_{j=1}^{k} a_j B^j. \tag{4.1}$$

It is also convenient to adopt the notation $B^0 := \mathrm{Id}$.

Example 4.2. Suppose that

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}; \quad p(x) = x^2 - 4x + 4. \tag{4.2}$$

Then

$$p(A) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.3}$$

The following proposition will help us to prove the Caley-Hamilton Theorem.

Proposition 4.3. *Suppose that $A \in \mathrm{End}(V)$, $z \in V, z \neq 0$, $\{z, Az, A^2z, \ldots, A^{k-1}z\}$ is linearly independent, and*

$$A^k z = \sum_{j=0}^{k-1} b_j A^j z. \tag{4.4}$$

Then a) $U = \langle z, Az, A^2z, \ldots, A^{k-1}z \rangle$ is invariant subspace of A.

b) Denote by B the restriction of A to the subspace U. Then

$$\chi_B(\lambda) = \lambda^k - \sum_{j=0}^{k-1} b_j \lambda^j. \tag{4.5}$$

c) $(\chi_B(B))(z) = 0$.

Proposition 4.4 (Caley-Hamilton)**.** *Suppose that $A \in \mathrm{End}(V)$ then*

$$\chi_A(A) = 0. \tag{4.6}$$

We may pick an arbitrary $z \in V$ and then, using Proposition 4.3 show that $(\chi_A(A))(z) = 0$.

Chapters from the textbooks relevant to the lecture:

- Friedberg, Insel, Spence, Sections 5.4;
- Axler ch.4, 5.