COURSE 13

Eigenvectors and Eigenvalues

Let K be a field and let V be a K-vector space.

Definition 1. Let $f : V \to V$ be a K-linear map, i.e. $f \in End_K(V)$. A non-zero vector $x \in V$ is an **eigenvector of** f if there exists $\lambda \in K$ such that $f(x) = \lambda x$. The above scalar λ is an **eigenvalue of** f corresponding to x. The set of all the eigenvalues of f is **the spectrum of** f.

Remarks 2. a) An eigenvector has a unique corresponding eigenvalue.

Indeed, if $x \in V$, $x \neq 0$, is an eigenvector of f and λ, λ' are eigenvalues of f corresponding to x then

$$f(x) = \lambda x \text{ şi } f(x) = \lambda x' \Rightarrow \lambda x = \lambda x' \Rightarrow (\lambda - \lambda')x = 0 \stackrel{x\neq 0}{\Longrightarrow} \lambda - \lambda' = 0 \Rightarrow \lambda = \lambda'.$$

b) If $\lambda \in K$ is an eigenvalue of f and $V(\lambda)$ is the subset of V consisting of the zero vector and the eigenvectors of f corresponding to the eigenvalue λ, i.e.

$$V(\lambda) = \{x \in V \mid f(x) = \lambda x\},$$

then $V(\lambda)$ is a subspace of V called **the eigenspace** (or **the characteristic space**) **of** f associated with λ.

Indeed,

$$x \in V(\lambda) \Leftrightarrow f(x) = \lambda x \Leftrightarrow (f - \lambda 1_V)(x) = 0 \Leftrightarrow x \in \operatorname{Ker}(f - \lambda 1_V)$$

hence $V(\lambda) = \operatorname{Ker}(f - \lambda 1_V)$. Since the kernel of a linear map is a subspace, $V(\lambda) \leq_K V$.

c) If $\lambda \in K$ is an eigenvalue of $f \in End_K(V)$ then $\dim V(\lambda) \geq 1$.

Indeed, since $V(\lambda) \leq_K V$ is not the zero subspace, $\dim V(\lambda) > 0$, hence $\dim V(\lambda) \geq 1$.

d) If $\lambda \in K$ is an eigenvalue of $f \in End_K(V)$ then $f(V(\lambda)) \subseteq V(\lambda)$.

Indeed,

$$x \in V(\lambda) \Rightarrow f(x) = \lambda x \Rightarrow f(f(x)) = \lambda f(x) \Rightarrow f(x) \in V(\lambda).$$

For the next part of the course, we consider that $\dim V = n (\in \mathbb{N}^*)$**.**

Theorem 3. Let $f \in End_K(V)$, $B = (v_1, \ldots, v_n)$ a basis of V and let $A = (a_{ij}) \in M_n(K)$ be the matrix of f in the basis B, i.e. $A = [f]_B$. The eigenvalues λ of f are the solutions from K of the equation $\det(A - \lambda I_n) = 0$, i.e. the equation

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} = 0 \qquad (1)$$

called **the characteristic equation of the matrix** A. If $\lambda \in K$ is a solution of the equation (1), then the coordinates $x_1, \ldots, x_n$ in the basis B of the vectors from $V(\lambda)$ result by solving the homogeneous linear system

$$\begin{cases} (a_{11} - \lambda)x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x + (a_{22} - \lambda)x_2 + \cdots + a_{2n}x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + (a_{nn} - \lambda)x_n = 0 \end{cases} \qquad (2)$$

Proof. A scalar $\lambda \in K$ is an eigenvalue of f if and only if there exists a non-zero vector $x \in V$ such that $f(x) = \lambda x$. But

$$f(x) = \lambda x \Leftrightarrow (f - \lambda 1_V)(x) = 0.$$

If $x = x_1v_1 + \cdots + x_nv_n$ is the representation of x in the basis B, the coordinates of $(f - \lambda 1_V)(x)$ are linear combinations of the coordinates of x having as coefficients the entries of the rows of $[f - \lambda 1_V]_B$. Therefore,

$$(f - \lambda 1_V)(x) = 0 \;\Leftrightarrow\; [f - \lambda 1_V]_B \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

But $[f - \lambda 1_V]_B = [f]_B - \lambda[1_V]_B$ and $[1_V]_B = I_n$. Hence the above equality can be rewritten

$$\begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{3}$$

The matrix equation (3) is equivalent to the homogeneous linear system (2), and (2) has non-trivial solutions if and only if the determinant of the system's matrix is zero, i.e. λ is a solution of the equation (1). □

Definition 4. The determinant $\det(A - \lambda I_n)$ from the left side of (1) is a polynomial expression $p_A(\lambda)$ of degree n in λ called **the characteristic polynomial of the linear map f in the basis B** or **the characteristic polynomial of the matrix $A = [f]_B$**. More precisely, the characteristic polynomial results by replacing the scalar λ in $\det(A - \lambda I_n)$ with the indeterminate X.

Theorem 5. If A and A' are matrices of $f \in End_K(V)$ in two bases then $p_A(\lambda) = p_{A'}(\lambda)$.

Proof. Let B, B' be two bases of V, S be the transition matrix from B to B', $A = [f]_B$ and $A' = [f]_{B'}$. Then $S \in GL_n(K)$ and $A' = S^{-1}AS$. Therefore,

$$p_{A'}(\lambda) = \det(A' - \lambda I_n) = \det(S^{-1}AS - \lambda S^{-1}I_nS) = \det(S^{-1}(A - \lambda I_n)S) = \det(S^{-1})\det(A - \lambda I_n)\det(S)$$

Since K is commutative and $\det S^{-1} = (\det S)^{-1}$,

$$\det(S^{-1})\det(A - \lambda I_n)\det(S) = \det(S^{-1})\det(S)\det(A - \lambda I_n) = \det(A - \lambda I_n) = p_A(\lambda).$$

Thus $p_{A'}(\lambda) = p_A(\lambda)$. □

Remarks 6. a) Theorem 5 shows that the characteristic polynomial of an endomorphism f in a certain basis does not depend on the basis of V, this is why we call it **the characteristic polynomial of f** and we denote it also by $p_f(\lambda)$. From (1) we get

$$p_f(\lambda) = (-1)^n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

where

$$a_{n-1} = (-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn}) \text{ and } a_0 = p_f(0) = \det A.$$

Since p_f does not depend on the basis of V, the sum $\text{Tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$ (called **the trace of** A) and $\det A$ are invariants of f. Therefore $a_{11} + a_{22} + \cdots + a_{nn}$ is also called **the trace of** f and $\det A$ is also called **the determinant of** f.
b) The characteristic polynomial of $f \in End_K(V)$ has the degree $n = \dim V$.
c) An endomorphism $f \in End_K(V)$ has at most $n = \dim V$ different eigenvalues.
d) If $K = \mathbb{C}$, $f \in End_K(V)$ and $n = \dim V$ then the characteristic polynomial of f has n roots in K (not necessarily different). This statement is no longer true for $K = \mathbb{R}$.

Definition 7. A matrix

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix} \in M_n(K) \tag{4}$$

is called **diagonal matrix**.

Definition 8. Let V be a K-vector space of dimension n. An **endomorphism** f of V is **diagonalizable** if there exists a basis $B = (v_1, \dots, v_n)$ of V such that the matrix $[f]_B$ is diagonal. A **matrix** $A \in M_n(K)$ is **diagonalizable** if there exists a diagonalizable endomorphism $f \in End_K(V)$ and a basis B of V such that $[f]_B = A$.

Remark 9. A matrix $A \in M_n(K)$ is diagonalizable if and only if there exists $S \in GL_n(K)$ such that $S^{-1}AS$ has the form (4).

Theorem 10. An endomorphism $f \in End_K(V)$ is diagonalizable if and only if the space V has a basis $B = (v_1, \dots, v_n)$ consisting only of eigenvectors of f.

Proof. $f \in End_K(V)$ is diagonalizable if and only if there exists a basis $B = (v_1, \dots, v_n)$ of V such that the matrix $[f]_B$ has the form (4). This means that

$$\begin{cases} f(v_1) = \lambda_1 v_1 + 0 \cdot v_2 + 0 \cdot v_3 + \cdots + 0 \cdot v_n \\ f(v_2) = 0 \cdot v_1 + \lambda_2 v_2 + 0 \cdot v_3 + \cdots + 0 \cdot v_n \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ f(v_n) = 0 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3 + \cdots + \lambda_n v_n \end{cases}$$

i.e. each v_i $(i = 1, \dots, n)$ is an eigenvector of f with the corresponding eigenvalue λ_i. □

Corollary 11. If $f \in End_K(V)$ is diagonalizable then all the roots of the characteristic polynomial of f are in K.

Indeed, if there exists a bais B of V such that $[f]_B$ has the form (4) then

$$p_f(\lambda) = \det([f]_B - \lambda I_n) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \dots (\lambda_n - \lambda)$$

hence the n roots of p_f are $\lambda_i \in K$ $(i = 1, \dots, n)$.

Proposition 12. Let $f \in End_K(V)$ and let $\lambda_i \in K$ be a root of the polynomial $p_f(\lambda)$. If m_i is the multiplicity of λ_i in $p_f(\lambda)$ then $\dim V(\lambda_i) \leq m_i$.

Proof. (optional)

If $B = (v_1, \dots, v_{n_i})$ is a basis of $V(\lambda_i)$ and $B' = (v_1, \dots, v_{n_i}, v_{n_i+1}, \dots, v_n)$ is a completion of B to a basis of V then $f(v_1) = \lambda_i v_1, \dots,\ f(v_{n_i}) = \lambda_i v_{n_i}$. If we denote by B_1 the diagonal matrix from $M_{n_i}(K)$ which has λ_i on the main diagonal, i.e. $B_1 = \lambda_i I_{n_i}$ then

$$[f]_{v'} = \begin{pmatrix} B_1 & B_2 \\ O & B_3 \end{pmatrix} \tag{5}$$

where O is the zero matrix. From (5) we deduce

$$p_f(\lambda) = \det(B_1 - \lambda I_{n_i}) \cdot \det(B_3 - \lambda I_{n-n_i}) = (\lambda_i - \lambda)^{n_i} \cdot \det(B_3 - \lambda I_{n-n_i})$$

hence,

$$p_f(\lambda) = (\lambda_i - \lambda)^{n_i} \cdot p_{B_3}(\lambda). \tag{6}$$

From (6) we deduce $n_i \leq m_i$. □

Corollary 13. Let $f \in End_K(V)$ and let $\lambda_i \in K$ be a simple root of $p_f(\lambda)$. Then $\dim V(\lambda_i) = 1$.

Indeed, the multiplicity of λ_i in $p_f(\lambda)$ is $m_i = 1$ and

$$1 \leq \dim V(\lambda_i) \leq m_i = 1.$$

Thus $\dim V(\lambda_i) = m_i = 1$.

Next, we will see that mutually different eigenvalues determine linearly independent eigenvectors.

Theorem 14. If $f \in End_K(V)$ and $v_1, \dots, v_k \in V$ are eigenvectors of f corresponding to the mutually different eigenvalues $\lambda_1, \dots, \lambda_k$, respectively, then $v_1, \dots, v_k$ are linearly independent.

Proof. We prove the theorem by way of induction on $k \in \mathbb{N}^*$. For $k = 1$, since $v_1 \neq 0$, from $\alpha_1 v_1 = 0$ with $\alpha_1 \in K$, we deduce $\alpha_1 = 0$. Hence the statement is true for $k = 1$.

Assume the statement true for $k \geq 1$ and we prove it for $k+1$ mutually different eigenvalues. If $\alpha_1, \dots, \alpha_k, \alpha_{k+1} \in K$ and

$$\alpha_1 v_1 + \cdots + \alpha_k v_k + \alpha_{k+1} v_{k+1} = 0 \tag{7}$$

then, by applying f we get

$$\alpha_1 \lambda_1 v_1 + \cdots + \alpha_k \lambda_k v_k + \alpha_{k+1} \lambda_{k+1} v_{k+1} = 0. \tag{8}$$

Multiplying (7) by $-\lambda_{k+1}$ and adding (8) it follows that

$$\alpha_1(\lambda_1 - \lambda_{k+1}) v_1 + \cdots + \alpha_k(\lambda_k - \lambda_{k+1}) v_k = 0.$$

By our assumption, $v_1, \dots, v_k$ are linearly independent, therefore

$$\alpha_1(\lambda_1 - \lambda_{k+1}) = \cdots = \alpha_k(\lambda_k - \lambda_{k+1}) = 0.$$

But $\lambda_1 \neq \lambda_{k+1}, \dots,\ \lambda_k \neq \lambda_{k+1}$. Hence $\alpha_1 = \cdots = \alpha_k = 0$. Now (7) implies $\alpha_{k+1} = 0$. Thus $v_1, \dots, v_k, v_{k+1}$ are linearly independent. □

Corollary 15. If $f \in End_K(V)$, $n = \dim V$ and f has n mutually different eigenvalues, then V has a basis which consists only of eigenvectors, hence f is diagonalizable.

Theorem 16. Let $n = \dim V$, $f \in End_K(V)$. The following statements are equivalent:
a) f is diagonalizable.
b) All the roots of the characteristic polynomial $p_f(\lambda)$ are in K, and if $\lambda_1, \dots, \lambda_k$ are these roots (mutually different) then, for any $i \in \{1, \dots, k\}$ the multiplicity m_i of λ_i is equal to $\dim V(\lambda_i)$.

(without proof)

As we saw in Corollary 13, the equality from b) always holds for the simple roots of p_f. In practice, for testing the diagonalizability of f we use the following corollary:

Corollary 17. With the notations of Theorem 16, f is diagonalizable if and only if all the roots of the characteristic polynomial p_f are in K and if $\lambda_1, \dots, \lambda_k$ are the (mutually different) roots of p_f,

$$m_i = n - \operatorname{rang}(f - \lambda_i 1_V), \ \forall\, i \in \{1, \dots, k\}. \tag{9}$$

Since $V(\lambda_i) = \operatorname{Ker}(f - \lambda_i 1_V)$, the equality from b) becomes (9) in the following way:

$$m_i = \dim V(\lambda_i) = \dim \operatorname{Ker}(f - \lambda_i 1_V) = \dim V - \dim(f - \lambda_i 1_V)(V) = n - \operatorname{rang}(f - \lambda_i 1_V).$$

Cayley-Hamilton Theorem

Let K be a field, $f = a_0 + a_1 X + \cdots + a_m X^m \in K[X]$ and $A \in M_n(K)$. Denote by $f(A)$ the matrix

$$f(A) = a_0 I_n + a_1 A + \cdots + a_m A^m.$$

If $f, g \in K[X]$ and $\alpha \in K$ then

$$(f+g)(A) = f(A) + g(A),\ (fg)(A) = f(A)g(A),$$

$$(\alpha f)(A) = \alpha f(A),\ f(A)g(A) = g(A)f(A).$$

Theorem 18. (Cayley-Hamilton Theorem). Any matrix $A \in M_n(K)$ satisfies its own characteristic equation, i.e.

$$p_A(A) = O$$

($O = O_n$ denotes the zero matrix from $M_n(K)$).

Proof. (*optional*)

We notice that any matrix $C \in M_n(K[X])$ can be uniquely written as

$$C = C_0 + C_1 X + \cdots + C_m X^m,\ \text{with } C_i \in M_n(K)\ (i = 0, 1, \dots, m).$$

If B is the adjugate matrix of $A - XI_n$, then

$$B \cdot (A - XI_n) = p_A(X) \cdot I_n \tag{10}$$

since $p_A(X) = \det(A - XI_n)$. The form of the characteristic polynomial p_A is

$$p_A(X) = a_0 + a_1 X + \cdots + a_n X^n. \tag{11}$$

The entries of B are the cofactors of the elements of $A - XI_n$. Therefore, these entries are polynomials from $K[X]$ with the degree at most $n-1$. Hence B can be written as

$$B = B_0 + B_1 X + \cdots + B_{n-1} X^{n-1} \tag{12}$$

where $B_i \in M_n(K)$ $(i = 0, 1, \ldots, n-1)$. From (10), (11) and (12) it follows that

$$(B_0 + B_1 X + \cdots + B_{n-1} X^{n-1})(A - XI_n) = (a_0 + a_1 X + \cdots + a_n X^n) I_n.$$

This leads us to the following equalities

$$\begin{cases} -B_{n-1} = a_n I_n \\ B_{n-1}A - B_{n-2} = a_{n-1} I_n \\ \vdots \\ B_1 A - B_0 = a_1 I_n \\ B_0 A = a_0 I_n \end{cases}$$

Multiplying the first equality with A^n on the right side, the second one with $A^{n-1}, \ldots$, the one before the last one with A and adding the resulting equalities we get

$$a_n A^n + a_1 A^{n-1} + \cdots + a_1 A + a_0 I_n = O\,, \tag{13}$$

i.e. $p_A(A) = O$. $\square$

Corollary 19. If the matrix $A \in M_n(K)$ has an inverse, then, from (13) one deduces that

$$A^{-1} = -\frac{1}{\det A}(a_1 I_n + a_2 A + \cdots + a_n A^{n-1}).$$