

MATH 725 NOTES

EIGENVALUES & EIGENVECTORS

MARIO L. GUTIERREZ ABED

Definition. Let V be a vector space over \mathbb{F} .

- A scalar $\lambda \in \mathbb{F}$ is an **eigenvalue** of an operator $\tau \in \mathcal{L}(V)$ if there exists a nonzero vector $v \in V$ for which

$$\tau(v) = \lambda v.$$

In this case, v is an **eigenvector** of τ associated with λ .

- A scalar $\lambda \in \mathbb{F}$ is an **eigenvalue** for a matrix A if there exists a nonzero column vector x for which

$$Ax = \lambda x.$$

In this case, x is an **eigenvector** for A associated with λ .

- The set of all eigenvectors associated with a given eigenvalue λ , together with the zero vector, forms a subspace of V , called the **eigenspace** of λ , denoted by \mathcal{E}_λ . This applies to both linear operators and matrices.
- The set of all eigenvalues of an operator or matrix is called the **spectrum** of the operator or matrix. ★

Theorem 1. Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V . Then the eigenvalues of T are precisely the entries on the diagonal of the matrix.

Theorem 2. Suppose that $\lambda_1, \dots, \lambda_k$ are distinct eigenvalues of a linear operator $\tau \in \mathcal{L}(V)$. Then

- The corresponding eigenspaces meet only in the 0 vector, that is

$$\mathcal{E}_{\lambda_i} \cap \mathcal{E}_{\lambda_j} = \{0\}.$$

- Eigenvectors associated with distinct eigenvalues are linearly independent. That is, if $v_i \in \mathcal{E}_{\lambda_i}$, then the vectors $\{v_1, \dots, v_k\}$ are linearly independent.

Definition. Let τ be a linear operator on a finite dimensional vector space V . Then the **characteristic polynomial** $C_\tau(x)$ of τ is defined to be the product of the elementary divisors of τ . If M is any matrix that represents τ , then

$$C_\tau(x) = C_M(x) = \det(xI - M). \quad \star$$

Remark: Note that the characteristic polynomial is not a *complete* invariant under similarity. For example, the matrices

$$A = \begin{bmatrix} \delta & 0 \\ 0 & \delta \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} \delta & 0 \\ 1 & \delta \end{bmatrix}$$

have the same characteristic polynomial but are not similar.

Definition. The **companion matrix** of the monic polynomial

$$m(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + x^n$$

is the square matrix

$$C[m(x)] = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & 0 & \vdots \\ \vdots & \vdots & \ddots & 0 & -a_{n-2} \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}. \star$$

Lemma 1. If $C[p(x)]$ is the companion matrix of the polynomial $p(x)$, then

$$\det(xI - C[p(x)]) = p(x).$$

Proof. See proof on page 154, Roman's. □

Definition. Let V_τ be an $\mathbb{F}[x]$ -module defined by τ . The unique monic order of V_τ (i.e., the unique monic polynomial that generates $\text{ann}(V_\tau)$) is called the **minimal polynomial** for τ and we denote it by $m_\tau(x)$. Thus we have

$$\text{ann}(V_\tau) = \langle m_\tau(x) \rangle$$

and

$$p(x)V_\tau = \{0\} \iff p(\tau) = 0 \iff m_\tau(x) \mid p(x). \quad \star$$

Remark: In treatments of linear algebra that do not emphasize the role of the module V_τ , the **minimal polynomial** of a linear operator τ is simply defined as the unique monic polynomial $m_\tau(x)$ of *smallest degree* for which $m_\tau(\tau) = 0$. It follows that any other polynomial p that satisfies $p(\tau) = 0$ must be a multiple of m_τ . It is not hard to see that this definition is equivalent to the previous definition.

Also, the concept of a minimal polynomial is defined for matrices. If A is a square matrix over \mathbb{F} , the **minimal polynomial** $m_A(x)$ of A is defined as the unique monic polynomial $p(x) \in \mathbb{F}[x]$ of smallest degree for which $p(A) = 0$.

Theorem 3. We have the following two results concerning minimal polynomials and matrices:

- 1) If A and B are similar matrices, then $m_A(x) = m_B(x)$. Thus, the minimal polynomial is an invariant under similarity.
- 2) The minimal polynomial of $\tau \in \mathcal{L}(V)$ is the same as the minimal polynomial of any matrix that represents τ .

Remark: We also have that the following three statements are equivalent:

- λ is a root of $m_\tau(x)$
- λ is a root of the characteristic polynomial $C_M(x)$, where M is any matrix that represents τ .
- λ is an eigenvalue of the matrix M .

We express this formally on the following theorem:

Theorem 4. Let $\tau \in \mathcal{L}(V)$ have minimal polynomial $m_\tau(x)$ and characteristic polynomial $C_\tau(x)$.

- i) The polynomials $m_\tau(x)$ and $C_\tau(x)$ have the same prime factors and hence the same set of roots, called the spectrum of τ .
- ii) The eigenvalues of a matrix are invariants under similarity.
- iii) If λ is an eigenvalue of a matrix A , then the eigenspace \mathcal{E}_λ is the solution space to the homogeneous system of equations

$$(\lambda I - A)(x) = 0.$$

Theorem 5 (The Cayley-Hamilton Theorem). Let $\tau \in \mathcal{L}(V)$ have minimal polynomial $m_\tau(x)$ and characteristic polynomial $C_\tau(x)$. Then the minimal polynomial divides the characteristic polynomial. Another way to say this is that an operator τ satisfies its own characteristic polynomial, that is,

$$C_\tau(\tau) = 0.$$

Remark: One way to compute the eigenvalues of a linear operator τ is to first represent τ by a matrix M and then solve the characteristic equation

$$C_M(x) = 0.$$

Unfortunately, it is quite likely that this equation cannot be solved when $\dim(V) \geq 3$. As a result, the art of approximating the eigenvalues of a matrix is a very important area of applied linear algebra.

Note: Eigenvalues have two forms of multiplicity, as described in the next definition:

Definition. Let λ be an eigenvalue of a linear operator $\tau \in \mathcal{L}(V)$. Then

- The **algebraic multiplicity** of λ is the multiplicity of λ as a root of the characteristic polynomial $C_\tau(x)$.
- The **geometric multiplicity** of λ is the dimension of the eigenspace \mathcal{E}_λ . ★

Theorem 6. The geometric multiplicity of an eigenvalue λ of $\tau \in \mathcal{L}(V)$ is less than or equal to its algebraic multiplicity.

THE JORDAN CANONICAL FORM

Definition. The **Jordan block** associated with a scalar λ_i is given by the $e_{i,j} \times e_{i,j}$ matrix

$$\mathcal{J}(\lambda_i, e_{i,j}) = \begin{bmatrix} \lambda_i & 0 & \dots & \dots & 0 \\ 1 & \lambda_i & \ddots & & \vdots \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 1 & \lambda_i \end{bmatrix}. \quad \star$$

Remark: Note that a Jordan block has λ_i 's on the main diagonal, 1's on the subdiagonal and 0's elsewhere.

Theorem 7 (The Jordan Canonical Form). *Let $\dim(V) < \infty$ and suppose that the minimal polynomial of $\tau \in \mathcal{L}(V)$ splits over the base field \mathbb{F} , that is*

$$m_\tau(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_n)^{e_n}.$$

Then V has an ordered basis \mathcal{C} under which

$$[\tau]_{\mathcal{C}} = \begin{bmatrix} \mathcal{J}(\lambda_1, e_{1,1}) & & & & \\ & \ddots & & & \\ & & \mathcal{J}(\lambda_1, e_{1,k_1}) & & \\ & & & \ddots & \\ & & & & \mathcal{J}(\lambda_n, e_{n,k_n}) & \\ & & & & & \ddots & \\ & & & & & & \mathcal{J}(\lambda_n, e_{n,k_n}) \end{bmatrix}_{\text{block}},$$

where the polynomials $(x - \lambda_i)^{e_i}$ are the elementary divisors of τ . This block diagonal matrix is said to be in **Jordan canonical form** and is called the **Jordan canonical form of τ** .

If the base field \mathbb{F} is algebraically closed then, except for the order of the blocks in the matrix, the Jordan canonical form is a canonical form for similarity; that is, up to order of the blocks, each similarity class contains exactly one matrix in Jordan canonical form.

Proof. As to the uniqueness, suppose that \mathcal{J} is a matrix in Jordan canonical form that represents the operator τ with respect to some ordered basis \mathcal{B} , and that \mathcal{J} has Jordan blocks $\mathcal{J}_1(\lambda_1, f_1), \dots, \mathcal{J}_m(\lambda_m, f_m)$, where the λ_i 's may not be distinct. Then V is the direct sum of τ -invariant subspaces, that is, submodules of V_τ , say

$$V = V_1 \oplus \cdots \oplus V_m.$$

Now consider a particular submodule V_k . It is easy to see from the matrix representation that $\tau|_{V_k}$ satisfies the polynomial $(x - \lambda_k)^{f_k}$ on V_k , but no polynomial of the form $(x - \lambda_k)^d$ for $d < f_k$, and so the order of V_k is $(x - \lambda_k)^{f_k}$. In particular, each V_k is a primary submodule of V_τ .

We claim that V_k is also a cyclic submodule of V_τ . To see this, let (v_1, \dots, v_{f_k}) be the ordered basis that gives the Jordan block $\mathcal{J}(\lambda_k, f_k)$. Then it is easy to see by induction that $\tau^j v_1$ is a linear combination of v_1, \dots, v_{j+1} , with coefficient of v_{j+1} equal to 1 or -1 . Hence, the set

$$\{v_1, \tau v_1, \tau^2 v_1, \dots, \tau^{f_k-1} v_1\}$$

is also a basis for V_k , from which it follows that V_k is a τ -cyclic subspace of V , that is, a cyclic submodule of V_τ .

Thus, the Jordan matrix \mathcal{J} corresponds to a primary cyclic decomposition of V_τ with elementary divisors $(x - \lambda_k)^{f_k}$. Since the multiset of elementary divisors is unique, so is the Jordan matrix representation of τ , up to order of the blocks. \square

Remark: Note that if τ has Jordan canonical form \mathcal{J} , then the diagonal elements of \mathcal{J} are precisely the eigenvalues of τ , each appearing a number of times equal to its algebraic multiplicity.

Lemma 2 (Schur's Lemma). *Let V be a finite-dimensional vector space over a field \mathbb{F} .*

- i) If $\tau \in \mathcal{L}(V)$ has the property that its characteristic polynomial $C_\tau(x)$ splits over \mathbb{F} , then τ is upper triangularizable.*
- ii) If \mathbb{F} is an algebraically closed field,¹ then all operators are upper triangularizable.*

Remark: Note that when $\mathbb{F} = \mathbb{R}$ not all operators are upper triangularizable, since \mathbb{R} is not algebraically closed (note, for instance, that the polynomial $(x^2 + 1) \in \mathbb{R}[x]$ has no zeroes in the reals).

Definition. *A linear operator $\tau \in \mathcal{L}(V)$ is **diagonalizable** if there is an ordered basis \mathcal{B} for which $[\tau]_{\mathcal{B}}$ is diagonal.* ★

Theorem 8. *A linear operator $\tau \in \mathcal{L}(V)$ is diagonalizable if and only if there is a basis for V that consists entirely of eigenvectors of τ . That is, if and only if*

$$V = \mathcal{E}_{\lambda_1} \oplus \cdots \oplus \mathcal{E}_{\lambda_k},$$

where $\lambda_1, \dots, \lambda_k$ are the distinct eigenvalues of τ .

Theorem 9. *A linear operator $\tau \in \mathcal{L}(V)$ on a finite-dimensional vector space is diagonalizable if and only if its minimal polynomial is the product of distinct linear factors.*

¹Recall that an algebraically closed field \mathbb{F} contains a root for every nonconstant polynomial in $\mathbb{F}[x]$, the ring of polynomials in the indeterminate x with coefficients in \mathbb{F} .