Eigenvectors and Diagonalization

(Determinant/Trace of Operators, Eigenspaces, Direct Sums, and Decomposition Criteria)

Contents

Chapter 1

Eigenvalues, the characteristic polynomial, and operator invariants

1.1 Eigenvectors revisited

Definition 1.1 (Eigenvector and eigenvalue)**.** Let V be a vector space over $\mathbb{F}$ and let $T \in \mathrm{End}_{\mathbb{F}}(V)$. A nonzero vector $u \in V \setminus \{0\}$ is an *eigenvector* of T if there exists $\lambda \in \mathbb{F}$ such that

$$T(u) = \lambda u.$$

In that case λ is called the *eigenvalue* associated to u.

Remark 1.2 (Kernel formulation)**.** The relation $T(u) = \lambda u$ is equivalent to

$$(T - \lambda \mathrm{id})(u) = 0, \qquad u \neq 0,$$

hence λ is an eigenvalue if and only if $\mathrm{Ker}(T - \lambda \mathrm{id}) \neq \{0\}$.

1.2 Determinant and trace of a linear operator

1.2.1 Motivation

In matrix theory, $\det(A)$ and $\mathrm{tr}(A)$ are invariants of similarity: if $A = P^{-1}BP$, then $\det(A) = \det(B)$ and $\mathrm{tr}(A) = \mathrm{tr}(B)$. This yields basis-independence of determinant/trace for linear operators via their matrix representations.

Definition 1.3 (Determinant and trace of an operator)**.** Let V be finite-dimensional over $\mathbb{F}$, and let $T \in \mathrm{End}_{\mathbb{F}}(V)$. Choose any ordered basis $\mathcal{B}$ of V and write $[T]_{\mathcal{B}} \in \mathrm{Mat}_n(\mathbb{F})$ for the matrix of T in $\mathcal{B}$. Define

$$\det(T) := \det([T]_{\mathcal{B}}), \qquad \mathrm{tr}(T) := \mathrm{tr}([T]_{\mathcal{B}}).$$

Proposition 1.4 (Well-definedness of $\det(T)$ and $\mathrm{tr}(T)$)**.** *The quantities* $\det(T)$ *and* $\mathrm{tr}(T)$ *do not depend on the chosen basis* $\mathcal{B}$*.*

Lemma 1.5 (Similarity invariance of determinant and trace). *Let $A, B \in \mathrm{Mat}_n(\mathbb{F})$ and assume A is similar to B, i.e. there exists $P \in \mathrm{GL}_n(\mathbb{F})$ such that*

$$A = P^{-1}BP.$$

Then

$$\det(A) = \det(B), \qquad \mathrm{tr}(A) = \mathrm{tr}(B).$$

Proof. Using multiplicativity of determinant,

$$\det(A) = \det(P^{-1}BP) = \det(P^{-1})\det(B)\det(P) = \det(P)^{-1}\det(B)\det(P) = \det(B).$$

For trace, use cyclicity $\mathrm{tr}(XY) = \mathrm{tr}(YX)$:

$$\mathrm{tr}(A) = \mathrm{tr}(P^{-1}BP) = \mathrm{tr}(BPP^{-1}) = \mathrm{tr}(BI) = \mathrm{tr}(B).$$

□

Proof of Proposition. Let $\mathcal{B}, \mathcal{C}$ be two ordered bases of V and let P be the change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$. Then the representing matrices satisfy

$$[T]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}P,$$

so $[T]_{\mathcal{B}}$ and $[T]_{\mathcal{C}}$ are similar. Lemma **??** implies $\det([T]_{\mathcal{B}}) = \det([T]_{\mathcal{C}})$ and $\mathrm{tr}([T]_{\mathcal{B}}) = \mathrm{tr}([T]_{\mathcal{C}})$. □

1.3 Eigenvalues and the determinant criterion

Proposition 1.6 (Eigenvalue criterion). *Let V be finite-dimensional and $T \in \mathrm{End}_{\mathbb{F}}(V)$. For $\lambda \in \mathbb{F}$,*

$$\lambda \text{ is an eigenvalue of } T \iff \det(T - \lambda \mathrm{id}) = 0.$$

Proof. λ is an eigenvalue $\iff \mathrm{Ker}(T - \lambda\mathrm{id}) \neq \{0\} \iff T - \lambda\mathrm{id}$ is not injective. In finite dimension, non-injectivity is equivalent to non-invertibility. For a linear operator, invertibility is equivalent to the representing matrix being invertible in any basis, i.e. having nonzero determinant. Thus $T - \lambda\mathrm{id}$ not invertible $\iff \det(T - \lambda\mathrm{id}) = 0$. □

1.4 The characteristic polynomial of an operator

Definition 1.7 (Characteristic polynomial). Let V be finite-dimensional over $\mathbb{F}$ and $T \in \mathrm{End}_{\mathbb{F}}(V)$. The *characteristic polynomial* of T is

$$\chi_T(t) := \det(T - t\mathrm{id}) \in \mathbb{F}[t].$$

Equivalently, for any basis $\mathcal{B}$,

$$\chi_T(t) = \det([T]_{\mathcal{B}} - tI).$$

Remark 1.8 (Basis-independence). Since $[T]_{\mathcal{B}}$ and $[T]_{\mathcal{C}}$ are similar for any two bases, $\det([T]_{\mathcal{B}} - tI)$ and $\det([T]_{\mathcal{C}} - tI)$ coincide as polynomials in t. Hence χ_T is well-defined.

Corollary 1.9. *A scalar $\lambda \in \mathbb{F}$ is an eigenvalue of T if and only if $\chi_T(\lambda) = 0$.*

Proof. By Proposition **??**,

$$\lambda \text{ eigenvalue} \iff \det(T - \lambda \mathrm{id}) = 0 \iff \chi_T(\lambda) = 0.$$

□

Chapter 2

Linear independence of eigenvectors and diagonalizability from eigenvalues

2.1 Eigenvectors for distinct eigenvalues are independent

Lemma 2.1 (Distinct-eigenvalue eigenvectors are independent)**.** *Let $T \in \mathrm{End}_{\mathbb{F}}(V)$. Suppose $u_1, \ldots, u_k \in V$ are eigenvectors with eigenvalues $\lambda_1, \ldots, \lambda_k \in \mathbb{F}$ satisfying $\lambda_i \neq \lambda_j$ for $i \neq j$. Then $\{u_1, \ldots, u_k\}$ is linearly independent.*

Proof. Assume for contradiction that the family is linearly dependent. Let $h \geq 1$ be the smallest integer such that $\{u_1, \ldots, u_h\}$ is linearly dependent. Then $h \geq 2$ and there exist scalars $a_1, \ldots, a_h \in \mathbb{F}$, not all zero, with

$$a_1 u_1 + \cdots + a_h u_h = 0. \tag{2.1}$$

By minimality of h, we must have $a_h \neq 0$ and $\{u_1, \ldots, u_{h-1}\}$ is linearly independent.

Apply T to (**??**):

$$a_1 T(u_1) + \cdots + a_h T(u_h) = 0 \quad \Longrightarrow \quad a_1 \lambda_1 u_1 + \cdots + a_h \lambda_h u_h = 0.$$

Multiply (**??**) by λ_h and subtract:

$$a_1(\lambda_1 - \lambda_h) u_1 + \cdots + a_{h-1}(\lambda_{h-1} - \lambda_h) u_{h-1} = 0.$$

Since $\lambda_i \neq \lambda_h$ for $1 \leq i \leq h-1$, each coefficient $(\lambda_i - \lambda_h)$ is nonzero. Because $\{u_1, \ldots, u_{h-1}\}$ is linearly independent, it follows that

$$a_i(\lambda_i - \lambda_h) = 0 \quad (1 \leq i \leq h-1) \implies a_i = 0 \quad (1 \leq i \leq h-1).$$

Then (**??**) reduces to $a_h u_h = 0$, forcing $a_h = 0$ (since $u_h \neq 0$), contradiction. □

2.2 A quick sufficient condition for diagonalizability

Corollary 2.2 (Distinct eigenvalues imply diagonalizable)**.** *Let V be an n-dimensional vector space over $\mathbb{F}$ and let $T \in \mathrm{End}_{\mathbb{F}}(V)$. If T has n distinct eigenvalues in $\mathbb{F}$, then T is diagonalizable over $\mathbb{F}$.*

Proof. Choose eigenvectors $u_1, \ldots, u_n$ corresponding to n distinct eigenvalues. By Lemma **??**, these vectors are linearly independent, hence form a basis of V. In that basis, T acts by $T(u_i) = \lambda_i u_i$, so $[T]$ is diagonal. $\square$

Remark 2.3 (Failure over $\mathbb{R}$ for rotations)**.** The rotation operator R_θ on $\mathbb{R}^2$ has complex eigenvalues $e^{\pm i\theta}$ when $\theta \not\equiv 0, \pi$, so it is not diagonalizable over $\mathbb{R}$ though it is diagonalizable over $\mathbb{C}$. This illustrates the field-dependence of diagonalization.

Chapter 3

Eigenspaces, rank–nullity, and direct sums

3.1 Eigenspaces and their basic structure

Definition 3.1 (Eigenspace)**.** Let $T \in \mathrm{End}_{\mathbb{F}}(V)$ and let $\lambda \in \mathbb{F}$. The *eigenspace* of T associated to λ is

$$E_\lambda(T) := \{u \in V : T(u) = \lambda u\} = \mathrm{Ker}(T - \lambda \mathrm{id}).$$

Proposition 3.2. *For every $\lambda \in \mathbb{F}$, $E_\lambda(T)$ is a linear subspace of V. Moreover,*

$$\dim E_\lambda(T) = \mathrm{Null}(T - \lambda \mathrm{id}),$$

the nullity of $T - \lambda \mathrm{id}$.

Proof. $E_\lambda(T) = \mathrm{Ker}(T - \lambda \mathrm{id})$ is a kernel, hence a subspace. Its dimension equals the nullity by definition. □

3.2 Rank–nullity and a useful dimension formula

Theorem 3.3 (Rank–nullity)**.** *Let $T : V \to W$ be linear with V finite-dimensional. Then*

$$\dim V = \dim \mathrm{Ker}(T) + \dim \mathrm{im}(T).$$

Equivalently,

$$\dim V = \mathrm{Null}(T) + \mathrm{rank}(T).$$

Proof. Define the quotient map $\pi : V \to V/\mathrm{Ker}(T)$ and the induced map $\widetilde{T} : V/\mathrm{Ker}(T) \to \mathrm{im}(T)$ by $\widetilde{T}(v + \mathrm{Ker}(T)) = T(v)$. This is well-defined and bijective. Hence $\dim(V/\mathrm{Ker}(T)) = \dim \mathrm{im}(T)$, i.e. $\dim V - \dim \mathrm{Ker}(T) = \dim \mathrm{im}(T)$. □

Corollary 3.4 (Eigenspace dimension via rank)**.** *For $T \in \mathrm{End}_{\mathbb{F}}(V)$ and $\lambda \in \mathbb{F}$,*

$$\dim E_\lambda(T) = \dim V - \mathrm{rank}(T - \lambda \mathrm{id}).$$

Proof. Apply Theorem **??** to $T - \lambda \mathrm{id} : V \to V$. □

3.3 Direct sums

Definition 3.5 (Sum of subspaces)**.** Let $U, W \leq V$. Define

$$U + W \coloneqq \{u + w : u \in U,\ w \in W\}.$$

Definition 3.6 (Direct sum of two subspaces)**.** Let $U, W \leq V$. We say V is the *direct sum* of U and W, written

$$V = U \oplus W,$$

if

$$V = U + W \quad \text{and} \quad U \cap W = \{0\}.$$

Proposition 3.7 (Uniqueness of decomposition)**.** *Let $U, W \leq V$. Then $V = U \oplus W$ if and only if every $v \in V$ can be written uniquely as*

$$v = u + w, \qquad u \in U,\ w \in W.$$

Proof. ($\Rightarrow$) If $v = u_1 + w_1 = u_2 + w_2$, then $u_1 - u_2 = -(w_1 - w_2) \in U \cap W = \{0\}$, so $u_1 = u_2$ and $w_1 = w_2$.

($\Leftarrow$) Existence of such a representation gives $V = U + W$. Uniqueness implies $U \cap W = \{0\}$ by applying it to $0 = u + w$. □

Definition 3.8 (Finite direct sums)**.** Let $U_1, \ldots, U_k \leq V$. We write

$$V = \bigoplus_{i=1}^{k} U_i$$

if $V = U_1 + \cdots + U_k$ and for each j,

$$U_j \cap \sum_{i \neq j} U_i = \{0\}.$$

Equivalently, each $v \in V$ has a unique representation $v = u_1 + \cdots + u_k$ with $u_i \in U_i$.

Chapter 4

Diagonalization via eigenspace decomposition

4.1 The eigenspace decomposition criterion

Theorem 4.1 (Diagonalization $\iff$ direct sum of eigenspaces)**.** *Let V be a finite-dimensional vector space over $\mathbb{F}$ with $\dim V = n$, and let $T \in \mathrm{End}_{\mathbb{F}}(V)$. Let $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ be the distinct eigenvalues of T (so $k \leq n$). Then the following are equivalent:*

(i) *T is diagonalizable over $\mathbb{F}$.*

(ii) *V decomposes as the direct sum of its eigenspaces:*

$$V = \bigoplus_{i=1}^{k} E_{\lambda_i}(T).$$

(iii) *The eigenspaces span V and their dimensions add to n:*

$$V = \sum_{i=1}^{k} E_{\lambda_i}(T) \quad \text{and} \quad \sum_{i=1}^{k} \dim E_{\lambda_i}(T) = n.$$

Proof. (i)$\Rightarrow$(ii). Assume T is diagonalizable. Then there exists a basis $\mathcal{B}$ of V consisting of eigenvectors. Group the basis vectors by eigenvalue:

$$\mathcal{B} = \{u_{11}, \dots, u_{1m_1}, u_{21}, \dots, u_{2m_2}, \dots, u_{k1}, \dots, u_{km_k}\},$$

where $T(u_{ij}) = \lambda_i u_{ij}$. For each i, the set $\{u_{i1}, \dots, u_{im_i}\}$ spans a subspace contained in $E_{\lambda_i}(T)$. Conversely, since these vectors are eigenvectors with eigenvalue λ_i, their span is contained in $E_{\lambda_i}(T)$, and by maximality of a basis inside $E_{\lambda_i}(T)$ we may regard them as a basis of $E_{\lambda_i}(T)$ after refinement. Thus

$$V = \mathrm{span}(\mathcal{B}) = \sum_{i=1}^{k} \mathrm{span}(u_{i1}, \dots, u_{im_i}) = \sum_{i=1}^{k} E_{\lambda_i}(T).$$

To see the sum is direct, let $v_i \in E_{\lambda_i}(T)$ satisfy $v_1 + \cdots + v_k = 0$. Writing each v_i in the eigenbasis $\mathcal{B}$ shows all its coordinates are zero, hence each $v_i = 0$. Therefore $V = \bigoplus_{i=1}^{k} E_{\lambda_i}(T)$.

(ii)⇒(iii). If $V = \bigoplus_i E_{\lambda_i}(T)$, then $V = \sum_i E_{\lambda_i}(T)$ and

$$\dim V = \sum_{i=1}^{k} \dim E_{\lambda_i}(T) = n$$

by additivity of dimension on direct sums.

(iii)⇒(i). Assume $V = \sum_i E_{\lambda_i}(T)$ and $\sum_i \dim E_{\lambda_i}(T) = n$. For each i, choose a basis $\mathcal{B}_i$ of $E_{\lambda_i}(T)$. The union $\mathcal{B} := \bigcup_{i=1}^{k} \mathcal{B}_i$ spans $\sum_i E_{\lambda_i}(T) = V$. Moreover,

$$|\mathcal{B}| = \sum_{i=1}^{k} |\mathcal{B}_i| = \sum_{i=1}^{k} \dim E_{\lambda_i}(T) = n = \dim V.$$

Hence $\mathcal{B}$ is a spanning set of size $\dim V$, therefore a basis of V. By construction, every vector in $\mathcal{B}$ is an eigenvector, so T is diagonalizable. □

Remark 4.2 (Matrix form in an eigenbasis)**.** Under the decomposition $V = \bigoplus_{i=1}^{k} E_{\lambda_i}(T)$, if $\mathcal{B}_i$ is a basis of $E_{\lambda_i}(T)$ and $\mathcal{B}$ is the concatenation of the $\mathcal{B}_i$, then the matrix of T in $\mathcal{B}$ is diagonal:

$$[T]_{\mathcal{B}} = \operatorname{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{\dim E_{\lambda_1}}, \underbrace{\lambda_2, \dots, \lambda_2}_{\dim E_{\lambda_2}}, \dots, \underbrace{\lambda_k, \dots, \lambda_k}_{\dim E_{\lambda_k}}).$$

4.2 Worked example (as in the notes): a 3×3 matrix with two eigenvalues

Example 4.3 (A matrix may be diagonalizable without n distinct eigenvalues)**.** Consider the matrix

$$A = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{F}).$$

Its characteristic polynomial can be computed by expansion:

$$\chi_A(t) = \det(A - tI) = \det \begin{pmatrix} 4-t & 0 & 1 \\ 2 & 3-t & 2 \\ 1 & 0 & 4-t \end{pmatrix} = (4-t) \det \begin{pmatrix} 3-t & 2 \\ 0 & 4-t \end{pmatrix} + 1 \cdot \det \begin{pmatrix} 2 & 3-t \\ 1 & 0 \end{pmatrix}.$$

Hence

$$\chi_A(t) = (4-t)(3-t)(4-t) - (3-t) = (3-t)\big((4-t)^2 - 1\big) = -(t-3)^2(t-5).$$

Thus the only eigenvalues are 3 (algebraic multiplicity 2) and 5 (algebraic multiplicity 1). Nevertheless, one may check (e.g. via rank–nullity) that

$$\dim E_5(A) = 1, \qquad \dim E_3(A) = 2,$$

so $\dim E_5(A) + \dim E_3(A) = 3 = \dim \mathbb{F}^3$; by Theorem **??**(iii), A is diagonalizable.

Guide to conventions

- We use $\chi_T(t) = \det(T - t\,\mathrm{id})$ (equivalently $\chi_A(t) = \det(A - tI)$).
- Under this convention the leading term is $(-1)^n t^n$.
- If instead one defines $\det(tI - A)$, the polynomial is monic and the t^{n-1} coefficient becomes $-\operatorname{tr}(A)$.