Eigenvectors and Diagonalization

Book-Style Lecture Notes

Contents

i

Chapter 1

Eigenvectors, invariant lines, and diagonalization

1.1 Linear operators and the "choose a good basis" principle

Remark 1.1 (Structural motivation). A linear operator $T : V \to V$ may appear complicated in an arbitrary basis, but can become extremely transparent after a judicious change of basis. The guiding principle is to find T-invariant subspaces that decompose V into simple pieces (ideally one-dimensional). This is the conceptual origin of diagonalization.

Notation 1.2. Throughout this chapter:

- $\mathbb{F}$ denotes a field (typical cases: $\mathbb{R}$ or $\mathbb{C}$).

- V denotes a finite-dimensional $\mathbb{F}$-vector space.

- $T \in \operatorname{End}_{\mathbb{F}}(V)$ denotes a linear operator.

1.2 Eigenvectors and eigenvalues

Definition 1.3 (Eigenvector and eigenvalue). Let $T : V \to V$ be $\mathbb{F}$-linear. A nonzero vector $u \in V \setminus \{0\}$ is an *eigenvector* of T if there exists $\lambda \in \mathbb{F}$ such that

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

The scalar λ is called the *eigenvalue* associated to u.

Remark 1.4 (Invariant line viewpoint). If $u \neq 0$ and $T(u) = \lambda u$, then the one-dimensional subspace $L = \mathbb{F}u$ is T-stable:

$$T(L) \subseteq L, \qquad T(cu) = cT(u) = c\lambda u = \lambda(cu) \quad (\forall c \in \mathbb{F}).$$

Equivalently, the restriction $T|_{\mathbb{F}u} \colon \mathbb{F}u \to \mathbb{F}u$ acts as scalar multiplication by λ.

Definition 1.5 (Eigenspace)**.** For $\lambda \in \mathbb{F}$, define the *eigenspace*

$$E_\lambda(T) := \ker(T - \lambda \mathrm{id}_V) = \{v \in V : T(v) = \lambda v\}.$$

If $E_\lambda(T) \neq \{0\}$, then λ is an eigenvalue of T. The set of eigenvalues is the *spectrum* of T, denoted $\sigma(T) \subseteq \mathbb{F}$.

Proposition 1.6 (Basic properties)**.** *Let* $T \in \mathrm{End}_\mathbb{F}(V)$*.*

(i) *Each $E_\lambda(T)$ is a linear subspace of V.*

(ii) *If $\lambda \neq \mu$, then $E_\lambda(T) \cap E_\mu(T) = \{0\}$.*

(iii) *Eigenvectors corresponding to distinct eigenvalues are linearly independent.*

Proof. (i) is immediate since $E_\lambda(T) = \ker(T - \lambda \mathrm{id})$.

(ii) If $v \in E_\lambda(T) \cap E_\mu(T)$, then $T(v) = \lambda v$ and $T(v) = \mu v$, hence $(\lambda - \mu)v = 0$. If $\lambda \neq \mu$, then $v = 0$.

(iii) Suppose $v_1, \ldots, v_k$ are eigenvectors with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$. If $\sum_{i=1}^{k} a_i v_i = 0$, apply T and subtract λ_k times the original relation to eliminate v_k; induct on k. $\square$

1.3 Diagonalization: operator-theoretic and matrix-theoretic formulations

Definition 1.7 (Diagonalizable operator)**.** A linear operator $T : V \to V$ is *diagonalizable over $\mathbb{F}$* if there exists a basis $\mathcal{B}$ of V such that the matrix $[T]_\mathcal{B}$ is diagonal. Equivalently, T is diagonalizable if V admits a basis consisting of eigenvectors of T.

Theorem 1.8 (Diagonalization criterion)**.** *Let* $T \in \mathrm{End}_\mathbb{F}(V)$ *with* $\dim_\mathbb{F} V = n$*. The following are equivalent:*

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

(2) *V decomposes as a direct sum of eigenspaces:*

$$V = \bigoplus_{\lambda \in \sigma(T)} E_\lambda(T).$$

(3) *There exist eigenvalues $\lambda_1, \ldots, \lambda_r$ such that*

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

Proof. $(1) \Rightarrow (2)$: If T is diagonal in some basis, then V is spanned by eigenvectors and each basis vector lies in an eigenspace. Distinct diagonal entries correspond to distinct eigenspaces, and the direct sum is forced by the trivial intersection property.

$(2) \Rightarrow (3)$: Taking dimensions yields $n = \sum_\lambda \dim E_\lambda(T)$, and only finitely many summands are nonzero.

$(3) \Rightarrow (1)$: Choose a basis for each $E_{\lambda_i}(T)$ and concatenate them; by directness, this is a basis of V, and the matrix is diagonal with diagonal entries the corresponding eigenvalues. $\square$

1.3.1 Matrix version and similarity

Let $A \in \mathrm{Mat}_n(\mathbb{F})$. The associated linear map $L_A : \mathbb{F}^n \to \mathbb{F}^n$ is given by $L_A(x) = Ax$.

Definition 1.9 (Diagonalizable matrix). A matrix $A \in \mathrm{Mat}_n(\mathbb{F})$ is *diagonalizable over* $\mathbb{F}$ if there exists $P \in \mathrm{GL}_n(\mathbb{F})$ and a diagonal matrix D such that

$$P^{-1}AP = D.$$

Proposition 1.10 (Eigenbasis and similarity). *Let $A \in \mathrm{Mat}_n(\mathbb{F})$.*

(i) *If $v_1, \ldots, v_n$ is a basis of $\mathbb{F}^n$ consisting of eigenvectors of A, and $P = [v_1 \; \cdots \; v_n]$, then $P \in \mathrm{GL}_n(\mathbb{F})$ and $P^{-1}AP$ is diagonal.*

(ii) *Conversely, if $P^{-1}AP = D$ is diagonal, then the columns of P form an eigenbasis of A (with eigenvalues given by the diagonal entries of D).*

Proof. (i) Since the columns form a basis, $P \in \mathrm{GL}_n(\mathbb{F})$. Moreover $AP = [Av_1 \; \cdots \; Av_n] = [\lambda_1 v_1 \; \cdots \; \lambda_n v_n] = P \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, hence $P^{-1}AP$ is diagonal.

(ii) Rewrite $AP = PD$. Comparing columns yields $Av_i = d_{ii} v_i$ for each column v_i. $\qquad\square$

1.4 Characteristic polynomial and spectral invariants

Definition 1.11 (Characteristic polynomial). For $A \in \mathrm{Mat}_n(\mathbb{F})$, the *characteristic polynomial* of A is

$$\chi_A(t) := \det(A - tI_n) \in \mathbb{F}[t].$$

Proposition 1.12 (Eigenvalues are roots). *A scalar $\lambda \in \mathbb{F}$ is an eigenvalue of A if and only if $\chi_A(\lambda) = 0$.*

Proof. λ is an eigenvalue $\iff \exists v \neq 0$ with $(A - \lambda I)v = 0 \iff A - \lambda I$ is non-invertible $\iff \det(A - \lambda I) = 0$. $\qquad\square$

Theorem 1.13 (Trace and determinant as coefficients). *Let $A \in \mathrm{Mat}_n(\mathbb{F})$. Then $\chi_A(t)$ is monic of degree n, and*

$$\chi_A(t) = (-1)^n t^n + (-1)^{n-1}(\operatorname{tr} A)\, t^{n-1} + \cdots + \det(A).$$

Equivalently,

$$\chi_A(t) = t^n - (\operatorname{tr} A)\, t^{n-1} + \cdots + (-1)^n \det(A),$$

depending on the convention $\det(tI - A)$ vs. $\det(A - tI)$.

Proof sketch (suitable for a first reading). Expand $\det(A - tI)$ via permutations:

$$\det(A - tI) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n}(a_{i,\sigma(i)} - t\delta_{i,\sigma(i)}).$$

The coefficient of t^n arises only from the identity permutation and yields $(-1)^n t^n$. The coefficient of t^{n-1} arises by choosing $-t$ from exactly one diagonal factor and $a_{j,\sigma(j)}$ from all others; only $\sigma = \mathrm{id}$ contributes, giving $(-1)^{n-1} \sum_i a_{ii} = (-1)^{n-1} \operatorname{tr}(A)$. The constant term is $\det(A)$ (take no $-t$ factors). A clean induction on n can be organized using Laplace expansion. $\qquad\square$

Remark 1.14 (Diagonalizable case)**.** If A is diagonalizable over an algebraic closure $\overline{\mathbb{F}}$, with eigenvalues (counted with algebraic multiplicity) $\lambda_1, \ldots, \lambda_n \in \overline{\mathbb{F}}$, then

$$\chi_A(t) = \prod_{i=1}^{n}(\lambda_i - t), \qquad \mathrm{tr}(A) = \sum_{i=1}^{n}\lambda_i, \qquad \det(A) = \prod_{i=1}^{n}\lambda_i.$$

Chapter 2

A canonical example: rotation in the plane

2.1 The rotation matrix

Fix $\theta \in \mathbb{R}$. Consider the linear operator on $\mathbb{R}^2$ given in the standard basis by

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Proposition 2.1 (Characteristic polynomial of R_θ). *Over any field containing $\cos\theta$ and $\sin\theta$, the characteristic polynomial of R_θ is*

$$\chi_{R_\theta}(t) = \det(R_\theta - tI_2) = t^2 - 2(\cos\theta)\,t + 1.$$

Proof. Compute:

$$\det\begin{pmatrix} \cos\theta - t & -\sin\theta \\ \sin\theta & \cos\theta - t \end{pmatrix} = (\cos\theta - t)^2 + \sin^2\theta = t^2 - 2(\cos\theta)t + (\cos^2\theta + \sin^2\theta),$$

and $\cos^2\theta + \sin^2\theta = 1$. $\qquad\square$

Corollary 2.2 (Eigenvalues over $\mathbb{C}$). *Over $\mathbb{C}$,*

$$\chi_{R_\theta}(t) = t^2 - 2(\cos\theta)t + 1 = (t - e^{i\theta})(t - e^{-i\theta}),$$

so the eigenvalues are $e^{\pm i\theta}$.

Proposition 2.3 (Non-diagonalizability over $\mathbb{R}$ for genuine rotations). *Assume $\theta \not\equiv 0, \pi \pmod{2\pi}$. Then R_θ has no real eigenvectors, hence is not diagonalizable over $\mathbb{R}$.*

Proof. A real eigenvalue $\lambda \in \mathbb{R}$ would be a real root of $\chi_{R_\theta}(t) = t^2 - 2(\cos\theta)t + 1$. But the discriminant is $\Delta = 4(\cos^2\theta - 1) = -4\sin^2\theta < 0$ when $\sin\theta \neq 0$. Hence there are no real roots and therefore no real eigenvectors. $\qquad\square$

Proposition 2.4 (Diagonalization over $\mathbb{C}$)**.** *Assume $\theta \not\equiv 0, \pi \pmod{2\pi}$. Then R_θ is diagonalizable over $\mathbb{C}$, with diagonal form*

$$P^{-1} R_\theta P = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \quad \text{for a suitable } P \in \mathrm{GL}_2(\mathbb{C}).$$

Proof. Over $\mathbb{C}$, χ_{R_θ} splits with two distinct roots $e^{\pm i\theta}$ (since $\theta \not\equiv 0, \pi$). Distinct eigenvalues imply the existence of a basis of eigenvectors, hence diagonalizability.

For concreteness, solve $(R_\theta - e^{i\theta} I)v = 0$. One convenient eigenvector for $e^{i\theta}$ is

$$v_+ = \begin{pmatrix} 1 \\ -i \end{pmatrix}, \qquad \text{and for } e^{-i\theta} \text{ one may take } v_- = \begin{pmatrix} 1 \\ i \end{pmatrix},$$

which can be verified by direct multiplication using $e^{\pm i\theta} = \cos\theta \pm i\sin\theta$. Let $P = [v_+ \; v_-]$. Then $P \in \mathrm{GL}_2(\mathbb{C})$ and $P^{-1} R_\theta P = \mathrm{diag}(e^{i\theta}, e^{-i\theta})$. $\qquad \square$

Remark 2.5 (Field-dependence of diagonalization)**.** Diagonalizability is not an intrinsic property of a matrix alone; it is a property of the pair $(A, \mathbb{F})$. The same A can fail to diagonalize over $\mathbb{R}$ but diagonalize over $\mathbb{C}$ (as with R_θ).

Chapter 3

Exercises (aligned with the handwritten sheet)

Exercise 3.1 (Trace/determinant from characteristic polynomial). Let $A \in \mathrm{Mat}_n(\mathbb{F})$ and $\chi_A(t) = \det(A - tI_n)$. Show that:

(a) the constant term of $\chi_A(t)$ equals $\det(A)$;

(b) the coefficient of t^{n-1} equals $(-1)^{n-1} \mathrm{tr}(A)$;

(c) more generally, the coefficients are (up to sign) the elementary symmetric polynomials in the eigenvalues (over an algebraic closure).

Exercise 3.2 (Rotation: explicit eigenvectors over $\mathbb{C}$). Let R_θ be as above with $\theta \not\equiv 0, \pi$ (mod 2π).

(a) Compute $\chi_{R_\theta}(t)$ and factor it over $\mathbb{C}$.

(b) Find eigenvectors $v_\pm \in \mathbb{C}^2$ for eigenvalues $e^{\pm i\theta}$.

(c) Write down an explicit $P \in \mathrm{GL}_2(\mathbb{C})$ such that $P^{-1} R_\theta P = \mathrm{diag}(e^{i\theta}, e^{-i\theta})$.

Definition 3.3 (Leibniz formula for the determinant). Let $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{F})$. The determinant of A is

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)},$$

where S_n is the symmetric group on $\{1, \dots, n\}$ and $\mathrm{sgn}(\sigma) \in \{\pm 1\}$ is the sign of σ.

Definition 3.4 (Sign via inversions). For $\sigma \in S_n$, an *inversion* is a pair (i, j) with $1 \le i < j \le n$ and $\sigma(i) > \sigma(j)$. Let $\mathrm{inv}(\sigma)$ be the number of inversions. Then

$$\mathrm{sgn}(\sigma) = (-1)^{\mathrm{inv}(\sigma)}.$$

Notation 3.5 (Principal submatrix). Let $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{F})$ and let $S \subseteq \{1, \dots, n\}$ with $S = \{i_1 < i_2 < \cdots < i_k\}$. We write $A[S, S] \in \mathrm{Mat}_k(\mathbb{F})$ for the *principal submatrix* obtained by restricting to rows and columns indexed by S, i.e.

$$\big(A[S, S]\big)_{pq} = a_{i_p, i_q} \qquad (1 \le p, q \le k).$$

The scalar $\det(A[S, S])$ is called a *principal minor* of A.

Theorem 3.6 (Trace and determinant as coefficients; principal-minor formula). *Let $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{F})$ and $\chi_A(t) := \det(A - tI_n) \in \mathbb{F}[t]$. Then*

$$\chi_A(t) = \sum_{k=0}^{n} (-1)^k \, e_k(A) \, t^k, \qquad e_k(A) = \sum_{\substack{S \subseteq \{1, \dots, n\} \\ |S| = n-k}} \det(A[S, S]),$$

where $A[S, S]$ denotes the principal submatrix of A indexed by S. In particular, the constant term is $\det(A)$ and the coefficient of t^{n-1} is $(-1)^{n-1} \mathrm{tr}(A)$.

Proof. By the Leibniz formula,

$$\det(A - tI) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} \big(a_{i,\sigma(i)} - t \, \delta_{i,\sigma(i)}\big).$$

Fix $\sigma \in S_n$ and let $F(\sigma) = \{i : \sigma(i) = i\}$ be its set of fixed points. A factor $(-t)$ can be chosen from the i-th term only if $i \in F(\sigma)$; otherwise the factor is $a_{i,\sigma(i)}$. Hence

$$\prod_{i=1}^{n} \big(a_{i,\sigma(i)} - t \, \delta_{i,\sigma(i)}\big) = \sum_{K \subseteq F(\sigma)} (-t)^{|K|} \prod_{i \notin K} a_{i,\sigma(i)}.$$

Therefore

$$\chi_A(t) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \sum_{K \subseteq F(\sigma)} (-t)^{|K|} \prod_{i \notin K} a_{i,\sigma(i)}.$$

Let $S = \{1, \dots, n\} \setminus K$. Then $|S| = n - |K|$, and the condition $K \subseteq F(\sigma)$ is equivalent to $\sigma(i) = i$ for all $i \notin S$, i.e. σ fixes S^c and permutes S. Reindexing yields

$$\chi_A(t) = \sum_{k=0}^{n} (-t)^k \sum_{\substack{S \subseteq \{1, \dots, n\} \\ |S| = n-k}} \sum_{\substack{\sigma \in S_n \\ \sigma(i) = i \ \forall i \notin S}} \mathrm{sgn}(\sigma) \prod_{i \in S} a_{i,\sigma(i)}.$$

If σ fixes S^c, then σ is determined by its restriction $\tau = \sigma|_S \in \mathrm{Sym}(S)$ and $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\tau)$. Thus the inner sum is precisely the Leibniz formula for $\det(A[S,S])$:

$$\sum_{\tau \in \mathrm{Sym}(S)} \mathrm{sgn}(\tau) \prod_{i \in S} a_{i,\tau(i)} = \det(A[S,S]).$$

Hence

$$\chi_A(t) = \sum_{k=0}^{n} (-t)^k \sum_{\substack{S \subseteq \{1,\dots,n\} \\ |S| = n-k}} \det(A[S,S]) = \sum_{k=0}^{n} (-1)^k e_k(A) t^k.$$

For $k = 0$, the only set is $S = \{1, \dots, n\}$, giving the constant term $\det(A)$. For $k = n-1$, the sets S have size 1, and $\det(A[\{i\},\{i\}]) = a_{ii}$, so $e_{n-1}(A) = \sum_i a_{ii} = \mathrm{tr}(A)$, and the coefficient of t^{n-1} is $(-1)^{n-1}\,\mathrm{tr}(A)$. $\qquad\square$