Linear Algebra II

Felix Chen

Contents

§1 Introduction

Teacher: An Jinpeng
Homepage: https://www.math.pku.edu.cn/teachers/anjp/algebra

§1.1 recap

Direct sums of vector spaces Given a field F, let $V_1, \ldots, V_k$ be vector spaces over F. The set

$$V_1 \times \cdots \times V_k = \{(v_1, \ldots, v_k) \mid v_i \in V_i\}$$

forms a vector space by the operations

$$(v_1, \ldots, v_k) + (w_1, \ldots, w_k) = (v_1 + w_1, \ldots, v_k + w_k)$$

and

$$c \cdot (v_1, \ldots, v_k) = (cv_1, \ldots, cv_k).$$

We call this vector space the **external direct sum** of $V_1, \ldots, V_k$, denoted by $\bigoplus_{i=1}^k V_i$.

Obviously $(U \oplus V) \oplus W \simeq U \oplus (V \oplus W)$.

For every i, we have an injective linear map:

$$\begin{aligned} \tau_i : V_i &\to \bigoplus_{j=1}^k V_j \\ v_i &\mapsto (0, \ldots, v_i, \ldots, 0) \end{aligned}$$

Lemma 1.1
If $\mathcal{B}_i$ are the bases of V_i, then $\bigcup_{i=1}^k \tau_i(\mathcal{B}_i)$ is a basis for $\bigoplus_{i=1}^k V_i$.
In paricular,

$$\dim \bigoplus_{i=1}^k V_i = \sum_{i=1}^k \dim V_i.$$

Proof. Spanning part:
For any $(v_1, \dots, v_k) \in \bigoplus_{i=1}^k V_i$,

$$v_i \in V_i = \operatorname{span}(\mathcal{B}_i) \implies \tau_i(v_i) \in \operatorname{span}\left(\tau_i(\mathcal{B})_i\right) \implies (v_1, \dots, v_k) \in \operatorname{span}\left(\bigcup_{i=1}^k \tau_i(\mathcal{B}_i)\right)$$

Linearly independent part:
If $\bigcup_{i=1}^k \tau_i(\mathcal{B}_i)$ is linearly dependent, i.e. exists $e_{ij} \in \mathcal{B}_i$ satisfying $\exists c_{ij} \in F$,

$$\sum_{i,j} c_{ij}\tau_i(e_{ij}) = 0.$$

This expands to

$$\left(\sum_{j=1}^{m_1} c_{1j}e_{1j}, \dots, \sum_{j=1}^{m_k} c_{kj}e_{kj}\right) = 0.$$

but e_{1j} are linear independent, which implies $c_{1j} = 0$. □

Remark 1.2 — Let V be a vector space over F, and $V_1, \dots, V_k$ are subspaces of V.
Consider a linear map $\Phi : V_1 \oplus \cdots \oplus V_k \to V$ by $(v_1, \dots, v_k) \mapsto v_1 + \cdots + v_k$.
Then $\operatorname{Im}(\Phi) = V_1 + \cdots + V_k$. If Φ is injective, i.e. $V_1, \dots, V_k$ are independent, we say $V_1 + \cdots + V_k$ the **internal direct sum** of $V_1, \dots, V_k$.
In this case Φ gives an isomorphism of external and internal sums:

$$\Phi : \bigoplus_{i=1}^k V_i \xrightarrow{\sim} \sum_{i=1}^k V_i.$$

Lemma 1.3

The following statements are equivalent:

1. $V_1, \dots, V_k$ are independent;
2. For $v_i \in V_i, (i = 1, \dots, k)$, if $\sum_{i=1}^k v_i = 0$, then $v_i = 0$.
3. For any $1 \le i \le k$, $V_i \cap (V_1 + \cdots + V_{i-1}) = \{0\}$.
4. Given arbitary bases $\mathcal{B}_i$ of V_i, they are disjoint and their union is a basis of $\bigoplus_{i=1}^k V_i$.
5. If $\dim V < +\infty$, they are also equivalent to:

$$\dim \sum_{i=1}^k V_i = \sum_{i=1}^k \dim V_i.$$

Proof. It's easy but verbose so I leave it out. □

Example 1.4

Let $\operatorname{char} F \ne 2$, $V = F^{n\times n}$, $V_1 = \{A \in V \mid A^t = A\}$, $V_2 = \{A \in V \mid A^t = -A\}$.

Note that $V_1 \cap V_2 = \{0\}$, and $V_1 + V_2 = V$, hence $V_1 \oplus V_2 = V$ is an internal direct sum.

§2 Diagonization

Example: google page rank?

Given a linear map T, it can be represented as different matrices under different bases. Thus a question arises: What's the simpliest matrix representation of a linear map?

Definition 2.1 (Diagonizable maps)**.** Let V be a vector space over F, $T \in L(V)$ is a linear map from V to itself. If the matrix of T under a certain basis is diagonal, we say T is **diagonizable**.

In this case the linear map T can be simply described as a diagonal matrix, thus we'll study under what condition is T diagonizable.

§2.1 Eigen-things

Definition 2.2 (Eigenvalue)**.** Let $T : V \to V$ be a linear map, for $c \in F$, let

$$V_c = \{v \in V \mid Tv = cv\} = \ker(T - c \cdot \mathrm{id}_V).$$

If $V_c \ne \{0\}$, we call c an **eigenvalue** of T, and V_c the **eigenspace** of T with respect to c. the vectors in V_c are called **eigenvectors**.

All the eigenvalues of T are called the **spectrum** of T, denoted by $\sigma(T)$.

Proposition 2.3

Let $\mathcal{B}$ be a basis of V, then $[T]_{\mathcal{B}}$ is diagonizable $\iff$ vectors in $\mathcal{B}$ are all eigenvectors.

Proof. Let $\mathcal{B} = \{e_1, \dots, e_k\}$, $A = [T]_{\mathcal{B}}$.

$$Te_j = \sum_{i=1}^{k} A_{ij} e_i.$$

So A is diagonal $\iff A_{ij} = 0$ when $i \neq j$,
$\iff \exists c_j \in F, Te_j = c_j e_j$,
$\iff$ all the vectors e_j are eigenvectors. □

Example 2.4

Let $V = F^{n \times n}$, then V_{sym} is the eigenspace of 1, and $V_{antisym}$ is the eigenspace of -1.

Lemma 2.5

Let T be a linear operator, then

$$\sigma(T) = \{c \in F \mid \det(c \cdot id_V - T) = 0\}.$$

Proof. $V_c = \ker(c \cdot \mathrm{id}_V - T)$,

$$c \in \sigma(T) \iff V_c \neq \{0\} \iff \det(c \cdot \mathrm{id}_V - T) = 0.$$

□

§2.2 Characteristic polynomial

To define the characteristic polynomial rigorously, we need to introduce one more concept:

Definition 2.6 (Rational function field)**.** Let F be a field, and $F[x]$ be its polynoiaml ring. Define the **rational function field**:

$$H := \{(f, g) \mid f, g \in F[x], g \neq 0\} = F[x] \times (F[x] \backslash \{0\}).$$

This process is similar to the extension from $\mathbb{Z}$ to $\mathbb{Q}$: We define a equivalent relation on H:

$$(f_1, g_1) \sim (f_2, g_2) \iff f_1 g_2 = f_2 g_1.$$

Let $F(x)$ be the set of all the equivalence classes.

Next we define the addition and multiplication as the usual way, and check they are well-defined (here it is left out).

Remark 2.7 — This process can be adapted to any integral domain R, which gives its fraction field $\mathrm{Frac}(R)$.

In general, we can define $F(x_1, \dots, x_n) = \mathrm{Frac}(F[x_1, \dots, x_n])$.

Let F be a field, and V a finite dimensional vector space over F, T is a linear operator on V.

We want to find the eigenvalues of T, by Lemma 2.5, we need to solve the equation

$$\det(c \cdot \mathrm{id}_V - T) = 0.$$

Definition 2.8 (Characteristic polynomial)**.** Let $A \in F^{n\times n}$, consider

$$xI - A \in F[x]^{n\times n} \subset F(x)^{n\times n}.$$

So

$$\det(xI - A) =: f_A(x) \in F(x).$$

The polynomial $f_A(x)$ is called the **characteristic polynomial** of A. Observe that its roots are all the eigenvalues of A.

In fact we can write f_A explicitly:

$$f_A(x) = \sum_{i=0}^{n} (-1)^i \sum \det B x^{n-i}$$

where $\sum \det B$ is over all $i \times i$ principal minors of A. In particular, $f_A(0) = (-1)^n \det A$.

> **Remark 2.9 —** In fact, the more intrinsic way to define the characteristic polynomial is to define it as $f_T(x) = (x - c_1)(x - c_2)\cdots(x - c_n)$, where c_i's are eigenvalues of a linear operator T. However, this definition requires the theory of Jordan forms, so it's hard to define it beforehand.

It's clear that similar matrices has the same characteristic polynomial since they represent the same linear operator.

Lemma 2.10

Let $A : F^r \to F^n$, $B : F^n \to F^r$ be linear maps. Then $f_{AB}(x) = x^{n-r} f_{BA}(x)$.

Proof 1. Note that

$$\begin{pmatrix} xI_n & A \\ B & I_r \end{pmatrix} \begin{pmatrix} I_n & 0 \\ -B & xI_r \end{pmatrix} = \begin{pmatrix} xI_n - AB & xA \\ 0 & xI_r \end{pmatrix}.$$

and

$$\begin{pmatrix} I_n & 0 \\ -B & xI_r \end{pmatrix} \begin{pmatrix} xI_n & A \\ B & I_r \end{pmatrix} = \begin{pmatrix} xI_n & A \\ 0 & xI_r - BA \end{pmatrix}.$$

By taking the determinant of both equations, we get:

$$x^r \det(xI_n - AB) = x^n \det(xI_r - BA).$$

□

Proof 2. By taking a suitable basis, we may assume $A = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}$. Suppose $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$, where B_{11} is an $m \times m$ matrix.

Compute

$$AB = \begin{pmatrix} B_{11} & B_{12} \\ 0 & 0 \end{pmatrix}, BA = \begin{pmatrix} B_{11} & 0 \\ B_{21} & 0 \end{pmatrix}.$$

we get $f_{AB}(x) = f_{B_{11}}(x) x^{n-m}$, $f_{BA}(x) = x^{r-m} f_{B_{11}}(x)$. □

If T is diagonalizable, then $f_T(x) = (x - c_1)\cdots(x - c_n)$, where $\{c_1, \ldots, c_n\} = \sigma(T)$.

Example 2.11 (How to diagonalize a matrix)

Let $A = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}$, we can compute $f_A(x) = (x-1)(x-2)^2$.

Next we compute the eigenspaces of each eigenvalue:

$$V_1 = (3, -1, 3), V_2 = \text{span}\{(2, 1, 0), (2, 0, 1)\}.$$

denote the eigenvectors by v_1, v_2, v_3.

At last we set $P = (v_1, v_2, v_3)$, we know $P^{-1}AP = \text{diag}\{1, 2, 2\}$.

Example 2.12

Let $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, $f_A(x) = x^2 - 2\cos\theta x + 1$, which has no real roots.

But if we regard it as a complex matrix, we can get $\sigma(A) = \{e^{i\theta}, e^{-i\theta}\}$, and $P = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$.

Example 2.13

Let $A = \begin{pmatrix} \lambda & a & b \\ 0 & \lambda & c \\ 0 & 0 & \lambda \end{pmatrix}$, where $\lambda, a, b, c \in \mathbb{R}$.

$f_A = (x-\lambda)^3$, but its eigenspace has dimension less than 3, so A is not diagonalizable.

From the examples we know not all the matrices can be diagonalized

- When f_A cannot decomposite to products of polynomials of degree 1;
- When the dimensions of eigenspaces can't reach $\dim V$.

The first case can be solved by putting it in a larger field; While the second case is intrinsic.

In what follows we'll take a closer look at the diagonalizable matrices, and find some equivalent statement of being diagonalizable.

Proposition 2.14

T can be diagonalize $\iff$ V can decomposite to direct sums of one-dimensional fixed subspaces.

Proof. Since there exists a basis consisting of eigenvectors: $\{e_1, \dots, e_n\}$, then $V = \bigoplus_{i=1}^n Fe_i$.

On the other hand, if $V = \bigoplus_{i=1}^n V_i$, where V_i's are 1-dimensional subspaces fixed under T, take $v_i \in V_i$, it's clear that v_i's form a basis of V, and they are all eigenvectors. This implies T is diagonalizable. $\square$

Proposition 2.15

The eigenspaces of different eigenvalues are independent. So their sum is acutually internal direct sums.

Proof. Let $\sigma(T) = \{c_1, \dots, c_r\}$, for any $v_i \in V_{c_i}$, if $v_1 + \cdots + v_r = 0$, let

$$S_1 = (T - c_2 \operatorname{id}_V) \cdots (T - c_r \operatorname{id}_V),$$

then $S_1(v_1 + \cdots + v_r) = Cv_1 = 0 \implies v_1 = 0$.
(As $S_1 v_i = (c_i - c_2) \cdots (c_i - c_r) v_i$ for $1 \le i \le r$.)
Similarly $v_i = 0$ for all i. □

Proposition 2.16

Suppose

$$f_T(x) = \prod_{c \in \sigma(T)} (x - c)^{m(c, f_T)}.$$

then $\forall c \in \sigma(T)$ we have $1 \le \dim V_c \le m(c, f_T)$.

Here $\dim V_c$ is called the **geometric multiplicy**, and $m(c, f_T)$ is the **algebraic multiplicy** of c.

Proof. Let $d = \dim V_c \ge 1$.
Take a basis $\{e_1, \dots, e_d\}$ of V_c and extend it to a basis of V: $\{e_1, \dots, e_n\}$.
Since $Te_i = ce_i, \forall i \le d$, so

$$[T]_{(e_i)} = \begin{pmatrix} cI_d & * \\ 0 & * \end{pmatrix}.$$

so $f_T(x) = (x - c)^d g(x)$, which means $m(c, f_T) \ge d$. □

Now we come to a conclusion:

Theorem 2.17

The followings are equivalent:

1. T is diagonalizable;
2. $V = \bigoplus_{c \in \sigma(T)} V_c$;
3. $\dim V = \sum_{c \in \sigma(T)} \dim V_c$;
4. $f_T(x) = \prod_{c \in \sigma(T)} (x - c)^{\dim V_c}$.

Proof. This follows immediately by previous propositions. □

§3 Jordan canonical form

It turns out that not all linear operators can be expressed as diagonal matrix. In this section we proceed in another direction: to find the "simpliest" matrix expression for a general operator.

Definition 3.1 (Irreducible maps)**.** Let T be a linear operator on V. We say T is **reducible** if V can be decompose to a direct sum of two T-invariant subspaces $W_1 \oplus W_2$. Otherwise we say T is **irreducible**.

In order to study T, we only need to study the "smaller" maps $T\big|_{W_1}$ and $T\big|_{W_2}$. In this case we denote $T = T\big|_{W_1} \oplus T\big|_{W_2}$. By decompose these smaller maps, we'll eventually get a decomposition of T consisting of irreducible maps:

$$T = \bigoplus_{i=1}^{r} T_{W_i}.$$

Then by taking a basis of each W_i, and they form a basis $\mathcal{B}$ of V. It's easy to observe that $[T]_{\mathcal{B}}$ is a block diagonal matrix.

In the special case when the W_i's are all 1-dimensional subspaces, the map T is diagonalizable. The eigenvectors are the elements in the W_i's and the eigenvalues are actually the map T_{W_i}.

Definition 3.2 (Annihilating polynomial)**.** Let $M_T = \{f \in F[x] \mid f(T) = 0\}$, we say the polynomial in M_T are the **annihilating polynomials** of T.

Note that M_T is an *nonzero* ideal of $F[x]$. This is because $\{\mathrm{id}, T, \ldots, T^{n^2}\} \subset \mathrm{Mat}_{n\times n}(F)$ must be linealy dependent.