Linear Algebra II

Felix Chen

Contents

§0.1 Characteristic polynomial

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

Definition 0.1 (Rational function field)**.** Let F be a field, and $F[x]$ be its polynomial 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 0.2 — 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 **??**, we need to solve the equation

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

Definition 0.3 (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 0.4 — 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 0.5

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 0.6 (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 0.7

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 0.8

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 0.9

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: $\{e_1, \ldots, e_n\}$, 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. □

Proposition 0.10

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

Proof. Let $\sigma(T) = \{c_1, \ldots, c_r\}$, 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 0.11

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 $\{e_1, \ldots, e_d\}$ of V_c and extend it to a basis of V: $\{e_1, \ldots, e_n\}$.

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 0.12

The followings are equivalent:

1. T is diagonalizable;
2. $V = \bigoplus\limits_{c\in\sigma(T)} V_c$;
3. $\dim V = \sum\limits_{c\in\sigma(T)} \dim V_c$;
4. $f_T(x) = \prod\limits_{c\in\sigma(T)} (x-c)^{\dim V_c}$.

Proof. This follows immediately by previous propositions. □

§1 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 1.1 (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 1.2 (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.

Proposition 1.3

T is diagonalizable $\iff \exists f \in M_T$ s.t. f is the product of different polynomials of degree 1.

Before we prove this proposition, let us take a look at the properties of annihilating polynomials.

Since $F[x]$ is a PID, M_T must be generated by one element, namely p_T, the minmimal polynomial of T. Thus we can WLOG assmue $f = p_T$ in the above proposition.

Speaking of polynomials and linear maps, one thing that pops into our mind is the characteristic polynomial f_T. In fact there is strong relations between p_T and f_T:

Theorem 1.4 (Cayley-Hamilton)

The characteristic polynomial of a linear operator T is its annihilating polynomial, i.e. $f_T(T) = 0$.

This theorem is also true when T is a matrix on a module. To prove it more generally, we introduce the concept of modules.

Definition 1.5 (Modules over commutative rings)**.** Let R be a commutative ring. A set M is called a **module** over R or an **R-module** if:

- There is a binary operation (addition) $M \times M \to M : (\alpha, \beta) \mapsto \alpha + \beta$ such that M becomes a commutative group under this operation.
- There is an operation (scaling) $R \times M \to M : (r, \alpha) \mapsto r\alpha$ with assosiativity and distribution over addition (both left and right). We also require $1_R\alpha = \alpha$ for all $\alpha \in M$.

Example 1.6

A commutative group automatically has a structure of $\mathbb{Z}$-module. (view the group operation as addition in definition of modules)

Example 1.7

Let $R = F[x]$, T a linear operator on V. Define $R \times V \to V : (f, \alpha) \mapsto f\alpha := f(T)\alpha$. We can check V becomes a module over R.

We can also define matrices on a commutative ring R, with addition and multiplication identical to the usual matrices. So the determinant and characteristic polynomial make sense as well.

Note that each $m \times n$ matrix represents a homomorphism $R^m \to R^n$.

Proof of Theorem 1.4. Take a basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ of V. Let $A = [T]_{\mathcal{B}}$. If we view V as a R-module ($R = F[x]$),

$$(\alpha_1, \ldots, \alpha_n)A = (T\alpha_1, \ldots, T\alpha_n) = (x\alpha_1, \ldots, x\alpha_n) = (\alpha_1, \ldots, \alpha_n) \cdot xI_n.$$

This implies $(\alpha_1, \ldots, \alpha_n)(xI_n - A) = (0, \ldots, 0)$.

Claim 1.8. If $f \in F[x]$ s.t. $\exists B \in R^{n\times n}$ s.t. $(xI_n - A)B = fI_n$, then $f(T) = 0$.

Proof of the claim.

$$0 = (\alpha_1, \ldots, \alpha_n)(xI_n - A)B = (\alpha_1, \ldots, \alpha_n) \cdot fI_n = (f(T)\alpha_1, \ldots, f(T)\alpha_n).$$

Since $\alpha_1, \ldots, \alpha_n$ is a basis, $f(T)$ must equal to 0. □

Now it's sufficient to prove f_T satisfies the condition in the claim. This follows from letting $B = A^{\mathrm{adj}}$, the adjoint matrix of A. □

Remark 1.9 — In fact this proof is derived from the proof of Nakayama's lemma, which is an important result in commutative algebra.

As a corollary, $p_T \mid f_T$.