Linear Algebra II

Felix Chen

Contents

§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}.

§1.1 Minimal polynomials and Cayley-Hamilton

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, $p_T \mid f_T$.

Proof of Proposition 1.3. First we prove a lemma:

Lemma 1.10

Let $T_1, \ldots, T_k \in L(V)$, $\dim V < \infty$. Then

$$\dim\ker(T_1T_2\ldots T_n) \leq \sum_{i=1}^{k} \dim\ker(T_i).$$

Proof of the lemma. By induction we only need to prove the case $k = 2$.

Note that $\ker(T_1T_2) = \ker(T_2) + \ker(T_1|_{\operatorname{im} T_2})$. So

$$\dim\ker(T_1T_2) = \dim\ker(T_2) + \dim\ker(T_1|_{\operatorname{im} T_2}) \leq \dim\ker(T_2) + \dim\ker(T_1).$$

□

If T is diagonalizable, suppose the matrix of T is $diag\{c_1, \ldots, c_r\}$, then $g = \prod_{i=1}^{r}(x - c_i)$ is an annihilating polynomial of T.

Conversely, if $\prod_{i=1}^{r}(T - c_iI) = 0$, by lemma

$$n = \ker\left(\prod_{i=1}^{r}(T - c_iI)\right) \leq \sum_{i=1}^{r} \ker(T - c_iI) = \sum_{i=1}^{r} \dim V_{c_i}.$$

This forces $V = \bigoplus_{i=1}^{r} V_{c_i}$, which completes the proof. □

§1.2 Invariant subspaces

There may not exist a subspace W' s.t. $W \oplus W' = V$, so we can instead study the quotient space.

Let $W \subset V$ be a T-invariant subspace. Define $T_W = T|_W \in L(W)$, $T_{V/W} \in L(V/W)$: $T_{V/W}(\alpha + W) = T(\alpha) + W$. It's clear that $T_{V/W}$ is well-defined.

However, this decomposition loses some information about T, i.e. they can't determine T completely. For example when $T = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$, the matrix B will not be carried to T_W and $T_{V/W}$ as their matrices are A, C respectively.

Since $\det T = \det T_W \det T_{V/W}$, $f_T = f_{T_W} \cdot f_{T_{V/W}}$. The minimal polynomials satisfy

$$lcm(p_{T_W}, p_{T_{V/W}}) \mid p_T, \quad p_T \mid p_{T_W} p_{T_{V/W}}.$$

This follows by the definition of $T_W, T_{V/W}$, readers can check it manually. Hint: The image of $p_{T_{V/W}}(T)$ is in W. So by Proposition 1.3, T is diagonalizable $\iff$ $T_W, T_{V/W}$ are both diagonalizable.

Definition 1.11 (Simultaneous diagonalization)**.** Let $\mathcal{F} \subset L(V)$, if there exists $\mathcal{B}$ s.t. $\forall T \in \mathcal{F}$, $[T]_{\mathcal{B}}$ is diagonal matrix, then we say $\mathcal{F}$ can be simultaneously diagonalized.

Proposition 1.12

Let $\mathcal{F} \subset L(V)$, TFAE:

- $\mathcal{F}$ can be simultaneously diagonalized;
- Any element in $\mathcal{F}$ is diagonalizable, and any two elements commute with each other.

Proof. It's obvious the fisrt statement implies the second.

On the other hand, we proceed by induction on the dimension of the space V.

Assume $\dim V = n \geq 2$, WLOG $T \in \mathcal{F}$ is not a scalar matrix.

Let $\sigma(T) = \{c_1, \ldots, c_r\}, V = \bigoplus_{i=1}^{r} V_{c_i}$, where $r \geq 2$, $V_{c_i} \neq V$. Since T commutes with other elements in $\mathcal{F}$, so $V_{c_i} = \ker(T - c_i \operatorname{id}_V)$ is invariant under all the maps in $\mathcal{F}$.

Hence we can restrict $\mathcal{F}$ to V_{c_i} and apply induction hypothesis, i.e. for any $U \in \mathcal{F}$, $U|_{V_{c_i}}$ can be simultaneously diagonalized.

Therefore $\exists \mathcal{B}_i$ s.t. $[U|_{V_{c_i}}]_{\mathcal{B}_i}$ is diagonal $\implies$ $[U]_{\mathcal{B}}$ is diagonal, where $\mathcal{B} = \bigcup \mathcal{B}_i$. □

Definition 1.13 (Triangulable matrix)**.** Let $T \in L(V)$. If $[T]_{\mathcal{B}}$ is an upper triangular matrix for some basis $\mathcal{B}$, we say T is **triangulable**.

Proposition 1.14

Let $\dim V = n$, for $T \in L(V)$, TFAE:

1. T is triangulable;
2. f_T(or p_T) can be decomposed to product of polynomials of degree 1.
3. There exists a sequence of T-invariant subspaces $\{0\} = W_0 \subset W_1 \subset \cdots \subset W_n = V$.

   This kind of sequence is called a **flag**.(not require T-invariant)

Remark 1.15 — In particular, when the base field is *algebraically closed*, the above statements always holds.

Proof. It's obvious that (1) $\implies$ (2). □