Linear Algebra II

Felix Chen

Contents

Proof of **??**. First we prove a lemma:

Lemma 0.1

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

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

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

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

□

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

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

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

§0.1 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 imformation 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 **??**, T is diagonalizable $\iff$ $T_W, T_{V/W}$ are both diagonalizable.

Definition 0.2 (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 0.3

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 $\mathcal{B} = \bigcup \mathcal{B}_i$. $\square$

Definition 0.4 (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 0.5

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**. (Flag itself does not require T-invariant)

Remark 0.6 — In particular, when the base field is *algebraically closed*, the above statements always holds.

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

For (3) $\implies$ (4): We proceed by induction, for W_1 just take the space spanned by one of the eigenvectors of T.

Assume that we have constructed W_j for $0 \leq j \leq i$. Instead of finding an invariant subspace of dimension $i+1$, we'll find an invariant subspace of dimension 1 in V/W_i.

Let Q denote the quotient map $V \to V/W_i$. Consider the map $T_{V/W_i} : \alpha + W_i \mapsto T(\alpha) + W_i$.

We have

$$T_{V/W_i} \circ Q = Q \circ T.$$

Since $p_{T_{V/W_i}} \mid p_T \implies p_{T_{V/W_i}}$ is product of polynomials of degree 1, T_{V/W_i} must have an eigenvector. Let L denote the subspace spanned by this vector, and $W_{i+1} = Q^{-1}(L)$.

Clearly $\dim W_{i+1} = 1 + \dim W_i = i + 1$. It suffices to check that W_{i+1} is T-invariant:

$$T(W_{i+1}) = T(Q^{-1}(L)) = Q^{-1}(T_{V/W_i}(L)) \subset Q^{-1}(L) = W_{i+1}.$$

Now for the last part (3) $\implies$ (1):

Let $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$, such that $\operatorname{span}\{\alpha_1, \dots, \alpha_i\} = W_i$. The matrix of T under $\mathcal{B}$ is clearly an upper triangular matrix. $\square$

Proposition 0.7

Let F be an algebraically closed field. Suppose the elements of $\mathcal{F} \subset L(V)$ are pairwise commutative, then $\mathcal{F}$ is simultaneously triangulable.

Remark 0.8 — The inverse of this proposition is not true: Just let $\mathcal{F}$ be the set consisting of all the upper triangular matrices.

Lemma 0.9

There's a common eigenvector of $\mathcal{F}$.

Proof of lemma. WLOG $\mathcal{F}$ is finite. (In fact, $\operatorname{span} \mathcal{F} \subset L(V)$ is a finite dimensional vector space, so we can take a basis $\mathcal{F}_0$.)

Now by induction, if $T_1, \dots, T_{k-1}$ have common eigenvector α, let $T_i\alpha = c_i\alpha$. Then

$$W = \bigcap_{i=1}^{k-1} \ker(T_i - c_i \operatorname{id}_V) \neq \{0\}$$

is a T_k-invariant space.

So any eigenvector α' of $T_k\big|_W$ is the common eigenvector. $\square$

Proof of the proposition. It suffices to prove that there exists an $\mathcal{F}$-invariant flag. By the lemma, the proof is nearly identical as the proof of previous proposition. $\square$

§0.2 Decomposition of linear maps

In this section we mainly study how a linear map is decomposed into irreducible maps and the structure of irreducible maps.

Recall that every vector space V is an $F[x]$-module given a linear operator T. If a subspace $W \subset V$ is a T-invariant space, then W is a submodule of V.

Hence it leads to decompose V into direct sums of submodules.

Definition 0.10. Let V, W be isomorphic vector spaces. $T \in L(V)$, $T' \in L(W)$. If there exists an isomorphism $\Phi : V \to W$ s.t. $\Phi \circ T = T' \circ \Phi$, we say T and T' are **equivalent**.

Definition 0.11 (Primary maps)**.** Let $T \in L(V)$ be a linear map. We say T is **primary** if p_T is a power of prime polynomials.

Theorem 0.12 (Primary decomposition)

Let $T \in L(V)$, $p_T = \prod_{i=1}^k p_i^{r_i}$, where p_i are different monic prime polynomials of degree 1.

We have

$$V = \bigoplus_{i=1}^k W_i, \quad W_i = \ker\left(p_i^{r_i}(T)\right),$$

with $W_i \neq \{0\}$ and $T\big|_{W_i}$ primary.

Proof. Let $f_i = \prod_{j \neq i} p_j^{r_j}$, f_i and p_i are coprime.

Note that $f_i(T) \neq 0$ and $f_i(T)p_i^{r_i}(T) = p_T(T) = 0$, thus $p_i^{r_i}(T)$ is not inversible, which implies $W_i \neq \{0\}$.

W_i independent : If there exists $\alpha_j \in W_j$ s.t. $\sum_{j=1}^k \alpha_j = 0$, applying f_i we get $f_i(\alpha_i) = 0$. But $p_i^{r_i}(\alpha_i) = 0 \implies \alpha_i = 0, \forall i$.

To prove $V = \sum_{i=1}^k W_i$, observe that

$$\gcd(f_1, \dots, f_k) = 1 \implies \exists g_1, \dots, g_k \quad s.t. \quad 1 = \sum_{i=1}^k g_i f_i \implies \alpha = \sum_{i=1}^k g_i(f_i \alpha), \quad \forall \alpha \in V.$$

Since $f_i\alpha \in W_i$, W_i is T-invariant $\implies g_i f_i \alpha \in W_i$.

Lastly, we'll prove that the minimal polynomial q_i of $T\big|_{W_i}$ is $p_i^{r_i}$.

Clearly $p_i^{r_i}(T\big|_{W_i}) = 0$, so $q_i \mid p_i^{r_i}$.

On the other hand, $q_1 q_2 \dots q_k$ is an annihilating polynomial of T, hence

$$\prod_{i=1}^k p_i^{r_i} \mid \prod_{i=1}^k q_i \implies q_i = p_i^{r_i}, \quad \forall i.$$

□

§0.3 Cyclic decomposition

In the following contents we'll assmue $R = F[x]$ if it's not specified.

Definition 0.13 (Cyclic maps)**.** Let V be a finite dimensional vector space and $T \in L(V)$. For $\alpha \in V$, $R\alpha = \{f\alpha \mid f \in R\} = \text{span}\{\alpha, T\alpha, \dots\}$ is the smallest T-invariant subspace containing α.

We say T is **cyclic** if $\exists \alpha$ s.t. $V = R\alpha$. In this case α is called a **cyclic vector**.

Here $R\alpha$ is called the cyclic subspace spanned by α.

Remark 0.14 — The word "cyclic" comes from the theory of modules.

Note that $\dim R\alpha = 1 \iff \alpha$ is an eigenvector.

Example 0.15

Let $A = E_{21} \in F^{2\times 2}$. Then A is cyclic because $A\varepsilon_1 = \varepsilon_2$, $A\varepsilon_2 = 0$. This means ε_1 is a cyclic vector of A,

Now there's a natural question: When is T cyclic and how to find its cyclic vectors?

For a given vector α, let $M_\alpha = \{f \in R \mid f\alpha = 0\}$ is an ideal of R.

Note that $M_T \subset M_\alpha$ as $f \in M_T \implies f(T)\alpha = 0$, so M_α is nonempty, it has a generating element p_α, called the **annihilator** of α.

Proposition 0.16

Let $d = \deg p_\alpha$, then $\{\alpha, T\alpha, \dots, T^{d-1}\alpha\}$ is a basis of $R\alpha$. In particular, $\dim R\alpha = \deg p_\alpha$.

Proof. Linear independence:

If $\sum_{i=0}^{d-1} c_i T^i \alpha = 0$, let $g = \sum_{i=0}^{d-1} c_i x^1$.

$$g\alpha = 0 \implies g \in M_\alpha \implies p_\alpha \mid g.$$

But $\deg g \le d - 1 < d = \deg p_\alpha \implies g = 0$.

Spanning:

Clearly $T^i\alpha \in R\alpha$. $\forall f \in R$, let $f = qp_\alpha + r$ with $\deg r < \deg p_\alpha$. Hence $f\alpha = r\alpha \in \operatorname{span}\{\alpha, T\alpha, \dots, T^{d-1}\alpha\}$. ☐

Since α is a cyclic vector $\iff \dim R\alpha = \dim V$, and $\deg p_\alpha \le \deg p_T \le \deg f_T = \dim V$, so we care whether these two inequalities can attain the equality.

Proposition 0.17

There exists $\alpha \in V$ s.t. $p_\alpha = p_T$.

Proof. Let $p_T = \prod_{i=1}^k p_i^{r_i}$.

$$W_i = \ker(p_i^{r_i}(T)) \implies V = \bigoplus_{i=1}^k W_i.$$

We claim that $\ker(p_i^{r_i-1}) \subsetneq W_i$ as $p_{T_{W_i}} = p_i^{r_i}$.

Take a vector $\alpha_i \in W_i \backslash \ker(p_i^{r_i-1}(T))$. By definition $p_{\alpha_i} \mid p_i^{r_i}, p_{\alpha_i} \nmid p_i^{r_i-1} \implies p_\alpha = p_i^{r_i}$.

Let $\alpha = \sum_{i=1}^k \alpha_i$. If $f\alpha = 0$, then $f\alpha_i = 0$ for $i = 1, \dots, k$ as $f\alpha_i \in W_i$.

$$f\alpha_i = 0 \implies p_{\alpha_i} \mid f \implies p_T \mid f.$$

This means we must have $p_\alpha = p_T$. ☐

Now we come to a conclusion:

Corollary 0.18

T is cyclic $\iff \deg p_T = \dim V \iff p_T = f_T$.

In this case, α is a cyclic vector $\iff p_\alpha = p_T$.

Let $n = \dim V$, T be a cyclic map, α be a cyclic vector. By previous proposition, $\{\alpha, T\alpha, \dots, T^{n-1}\alpha\}$ is a basis of V. Denote the basis by $\mathcal{B}$.

Observe that $[T]_\mathcal{B}$ is equal to

$$A = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & 0 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -c_{n-2} \\ 0 & 0 & 0 & \cdots & 1 & -c_{n-1} \end{pmatrix}$$

where c_i are the coefficients of $p_\alpha = p_T = f_T = \sum_{i=0}^n c_i x^i$. For a monic polynomial f, define C_f to be the matrix as above, called the **companion matrix** of f.

Proposition 0.19

If exists a basis $\mathcal{B}$ s.t. $[T]_{\mathcal{B}} = C_f$ for some monic polynomial f, then T is cyclic and $p_T = f$.

Proof. Let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$, we have $T^i\alpha_1 = \alpha_{i+1} \implies R\alpha_1 = V$ and $p_{\alpha_1} = f$. $\square$

Remark 0.20 — In fact we can check directly that f is the characteristic polynomial of C_f.

This gives another proof of Cayley-Hamilton theorem:

Proof. For any $\alpha \in V$, consider $T_{R\alpha} \mid f_T$.

$$f_{T_{R\alpha}} = f_{C_{p_\alpha}} = p_\alpha \mid f_T$$

This implies that f_T is an annihilating polynomial of α, which means $f_T(\alpha) = 0, \forall \alpha \in V$, i.e. $f_T(T) = 0$. $\square$

Theorem 0.21 (Cyclic decomposition)

Let $T \in L(V)$, $\dim V = n$. There exists $\alpha_1, \ldots, \alpha_r \in V$ s.t. $V = \bigoplus_{i=1}^r R\alpha_i$.

Furthermore, $p_{\alpha_r} \mid \cdots \mid p_{\alpha_1} = p_T$, $f_T = \prod_{i=1}^r p_{\alpha_i}$.

Here p_{α_i}'s are called the **invariant factors** of T. The invariant factors are *totally determined* by T.