Linear Algebra II (S)

Deng Tianle

20 October 2024

Contents

1 Introduction

This is basically my notes based on the lectures by Dr. Arghya Sadhukhan for MA2101S at the National University of Singapore. All mistakes are almost certainly mine.

The main reference book is the standard textbook by Hoffman and Kunze.

2 Preliminaries

2.1 Cardinality

We give a rough collection of results For details, please refer to the materials in MA1100T.

Definition 2.1.1. *A set S is countable if S is finite, or there exists a bijection $\phi : \mathbb{Z}_{\geqslant 0} \to S$.*

Remark. Roughly, S is countable $\iff$ $\exists$(infinite) algorithm to list its elements.

Proposition 2.1.1. *Take $S \neq \emptyset$. The following are equivalent:*

1. *S is conutable*
2. *there exists an injection $\phi : S \to \mathbb{Z}_{\geqslant 0}$*
3. *there exists a countable set C and an injection $\psi : S \to C$.*

Corollary 2.1.2. *Every subset of a countable set is countable too.*

Proposition 2.1.3. *A finite product of countable sets is also countable.*

Proof. Let $S_1, \ldots, S_n$ be countable sets. Let $p_1, \ldots, p_n$ be distinct primes. We have injections $f_i : S_i \to \mathbb{Z}_{\geqslant 0}$ by 2. of Proposition 2.1.1. Define

$$\begin{aligned} \phi : S_1 \times \cdots \times S_n &\to \mathbb{Z}_{\geqslant 0} \\ (s_1, \ldots, s_n) &\mapsto p_1^{f_1(s_1)} \ldots p_n^{f_n(s_n)}. \end{aligned}$$

We claim that ϕ is injective. Indeed, we have

$$\begin{aligned} & p_1^{f_1(s_1)} \ldots p_n^{f_n(s_n)} = p_1^{f_1(s_1')} \ldots p_n^{f_n(s_n')} \\ \overset{\text{FTA}}{\implies} & f_i(s_i) = f_i(s_i') \, \forall i = 1, \ldots, n \\ \overset{\text{injectivity}}{\implies} & s_i = s_i' \, \forall i = 1, \ldots, n \\ \implies & (s_1, \ldots, s_n) = (s_1', \ldots, s_n'). \end{aligned}$$

□

Corollary 2.1.4. $\mathbb{Q}$ *is countable.*

Proposition 2.1.5. *The union of a countable collection of countable sets is countable.*

Lemma 2.1.6. *Let $S \neq \emptyset$, $f : S \to \mathcal{P}(S)$; then f cannot be surjective.*

Proof. Suppose f is surjective. Consider $X := \{s \in S : s \notin f(s)\}$. Then $X \in \mathcal{P}(S)$, so by assumption there exists $x \in S$ such that $f(x) = X$. Then we get $x \in X \iff x \notin X$, contradiction. □

Corollary 2.1.7. $\mathcal{P}(\mathbb{Z}_{\geqslant 0})$ *is countable.*

Corollary 2.1.8. $\mathbb{R}$ *is uncountable.*

2.2 Groups, Rings and Fields

Definition 2.2.1. *A binary operation on $S \neq \emptyset$ is a function $* : S \times S \to S$, $(a, b) \mapsto a * b$.*

Remark. Addition, subtraction and multiplcation on $\mathbb{R}$ are examples. Division on $\mathbb{R}$ is a non-example.

Definition 2.2.2. *We have the following terminologies:*

1. *$*$ is associative: $(a * b) * c = a * (b * c)$ for all $a, b, c \in S$*
2. *$*$ is commutative: $a * b = b * a$ for all $a, b \in S$*
3. *an identity $e \in S$ is an element such that $a * e = e * a = a$ for all $a \in S$*
4. *if $e \in S$ is the identity, then an inverse of $a \in S$ is an element $a^{-1} \in S$ such that $a * b = b * a = e$.*

Remark. The identity is unique if it exists. The inverse is also unique if it exists (this justifies the notation a^{-1}). We have $(a^{-1})^{-1} = a$.

Definition 2.2.3. *Let $G \neq \emptyset$ be a set and $*$ be a binary operation on G. If $*$ is associative with the existence of identity and existence of inverses for each element in G, we say that $(G, *)$ is a group. If $*$ is commutative, we say that G is abelian.*

Example.
1. An example is $(\mathbb{R}, +)$. A non-example is $(R, -)$.
2. Let p be a prime, $\mathbb{Z}_p := \{0, 1, \ldots, p-1\}$.
   a) $(\mathbb{Z}_p, +)$ is a group, where $+$ = addition modulo p; more generally $(\mathbb{Z}_n, +)$ is also a group for any $n \in \mathbb{Z}_{\geqslant 0}$
   b) $(\mathbb{Z}_p - \{0\}, \cdot)$ is a group
3. $(M_n(\mathbb{R}).+)$ is an abelian group
4. (set of invertible matrices in $M_n(\mathbb{R}), \cdot)$ is a non-abelian group denoted as $\mathrm{GL}_n(\mathbb{R})$.

Definition 2.2.4. *Let R be a ring. Let $+$ and $\cdot$ be two associative binary operations on R such that:*

1. *$(R, +)$ is an abelian group (we denote the additive identity as 0; for $a \in R$, we denote its additive inverse as $-a$)*
2. *for all $a, b, c \in R$, $a \cdot (b + c) = a \cdot b + a \cdot c$, $(b + c) \cdot a = b \cdot a + c \cdot a$*

Example. Both $(\mathbb{Z}_p, +, \cdot)$ and $(M_n(\mathbb{R}), +, \cdot)$ are rings. Let R be a ring and $R[x]$ be the set of polynomials with coefficients in R. Then $(R[x], +, \cdot)$ is a ring.

Lemma 2.2.1. *Let R be a ring with identity 1. Let $a \in R$.*

1. *$0 \cdot a = a \cdot 0 = 0$*

2. $(-1) \cdot a = -a = a \cdot (-1)$

Definition 2.2.5. *Let F be a ring with the additional property that $(F - \{0\}, \cdot)$ is an abelian group (with identity 1). Then we say that $(F, +, \cdot)$ is a field.*

Remark. Let F be a field. For all $a, b \in F$, we have $a, b \neq 0 \implies a \cdot b \neq 0$ (because $F - \{0\}$ is closed with respect to $\cdot$). There are at least two elements $0 \neq 1$ in F.

Example. Common examples of fields include $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$, $F_p = (\mathbb{Z}_p, +, \cdot)$.

Let F be a field. In the following discussion, to avoid confusion, we write 0_F and 1_F for the additive and multiplicative identities respectively. We would also abbreviate $\underbrace{1_F + 1_F + \cdots + 1_F}_{n \text{ copies}}$ as $n \cdot 1_F$ and omit the $\cdot$ for field multiplication.

Definition 2.2.6. *We define the characteristic of a field F,*

$$\operatorname{char} F := \begin{cases} \min\{n \in \mathbb{Z}_{>0} : n \cdot 1_F = 0_F\}, & \textit{if this set is non-empty} \\ 0, & \textit{otherwise.} \end{cases}$$

Example. We have $\operatorname{char} \mathbb{R} = 0$, $\operatorname{char} F_p = p$.

Lemma 2.2.2. *Let F be a field. Then* $\operatorname{char} F$ *is either* 0 *or a prime p.*

Proof. Assume $\operatorname{char} F \neq 0$. There exists a minimal $n \in \mathbb{Z}_{>0}$ such that $n \cdot 1_F = 0_F$. Write $n = ab$, we have $(ab) \cdot 1_F = (a \cdot 1_F)(b \cdot 1_F) = 0_F$, so $a \cdot 1_F = 0_F$ or $b \cdot 1_F = 0_F$. By minimality, n has to be prime. □

Remark. Let F be a finite field. It is not hard to show that if $\operatorname{char} F = p$, then $|F| = p^n$ for some n.

2.3 Polynomials

Let R be a ring. We define

$$R[x] := \left\{ \sum_{i \in I} a_i x^i : I \subset \mathbb{Z}_{\geqslant 0}, |I| < \infty, a_i \in R \, \forall i \in I \right\}.$$

It is clear that $(R[x], +, \cdot)$ is a ring, called the polynomial ring of R. We have a natural embedding of R into $R[x]$ by the rule $a \mapsto ax^0$. Let $f \in R[x]$. We define the degree of f, $\deg f = \max\{i \in I : a_i \neq 0\}$ for $f \neq 0$ and $\deg 0 = -\infty$. The elements of R are called constant polynomials.

Remark. We have $\deg(f \cdot g) = \deg f + \deg g$.

Remark. The lecture did not require R to be commutative. Note that when R is non-comutative, the evaluation map may not be a ring homomorphism anymore.

Assume now that $R = F$ is a field. We want to study $F[x]$. In this special case, we have the following proposition:

Proposition 2.3.1 (Division algorithm)**.** *Let $f(x), g(x) \in F[x]$, $g(x) \neq 0$. Then there exists unique polynomials $q(x), r(x) \in F[x]$ such that:*

1. $f(x) = q(x)g(x) + r(x)$
2. *(either $r(x) = 0$ or)* $\deg r < \deg g$.

In fact, one can generalise the concept of a division algorithm.

Definition 2.3.1. *An integral domain R is Euclidean if there exists a function $d : R - \{0\} \to \mathbb{Z}_{>0}$ such that for all $a, b \in R$, $b \neq 0$, there exists $q, r \in R$ such that*

$$a = bq + r,$$

with $r = 0$ or $d(r) < d(b)$.

Example. The following are all Euclidean domains:

1. $R = F[x]$, $d = \deg$
2. $R = \mathbb{Z}$, $d = |\,|$
3. $R = \mathbb{Z}[i] := \{a + bi : a, b \in \mathbb{Z}\}$, $d : a + bi \mapsto a^2 + b^2$.

Remark. For Euclidean domains, we do not require q and r to be unique. However, observe that the uniqueness holds for $F[x]$. It turns out that this gives a characterisation of $F[x]$: an Euclidean domain with given d is either a field or a polynomial domain over a field, if the q and r referred to above are unique.

We now end the diversion and come back to the basic concepts for polynomials.

Definition 2.3.2. *Let $(R, +, \cdot)$ be a ring. The invertible elements in $(R - \{0\}, \cdot)$ are called units.*

Definition 2.3.3. *Let F be a field and let $f \in F[x]$ be a polynomial. We say that f is irreducible if $\deg f \geqslant 1$ and f cannot be written as $f = g \cdot h$ where both g and h are nonconstant polynomials. We say that f is monic if its leading coefficient (that is, the coefficient for $x^{\deg f}$) is 1_F.*

It is helpful to consider the follwing analogies:

$\mathbb{Z}$	$F[x]$
± 1	non-zero constants (ring units)
prime	irreducible
$\lvert n \rvert$	$\deg f$ (for division algorithm)
positive	monic polynomials

Also, recall the usual chain of reasoning on $\mathbb{Z}$:

$$\begin{aligned} \text{division algorithm} &\to \text{Bezout identity} \\ &\to \text{Euclid lemma} \\ &\to \text{uniqueness of prime factorisation.} \end{aligned}$$

We now adapt these considerations to $F[x]$.

Definition 2.3.4. *Let $f, g \in F[x]$. We define* $\gcd(f, g)$ *to be the unique polynomial such that:*

1. *it is a common divisor of f and g*
2. *it has the maximum degree among all such common divisors*
3. *it is monic.*

We say that f, g are coprime if $\gcd(f, g) = 1$.

It is clear that the Bezout identity holds with this definition of gcd. In fact, the usual chain of arguments holds with little modification. Similarly, we can define the gcd of multiple polynomials. We now record a lemma which (perhaps surprisingly) will be extremely useful later:

Lemma 2.3.2 (Generalised Bezout)**.** *Let $g_1, \ldots, g_n \in F[x]$ be such that* $\gcd(g_1, \ldots, g_n) = 1$. *Then there exists $f_1, \ldots, f_n \in F[x]$ such that $\sum_{i=1}^{n} f_i(x)g_i(x) = 1$.*

Proof. Consider the ideal I generated by $g_1, \ldots, g_n$. Because $F[x]$ is a PID, I is generated by a single polynomial. This generator is a common divisor of all g_i, so in this case, we see that 1 generates I, so in particular $1 \in I$. □

Example. To calculate $\gcd(x^2 + 1, x + 1)$ in $\mathbb{R}[x]$, we have

$$x^2 + 1 = (x - 1)(x + 1) + 2$$

$$x + 1 = \left(\frac{1}{2}x + \frac{1}{2}\right)(2) + 0$$

So $\gcd(x^2 + 1, x + 1) = 1$ (monic polynomial).

Theorem 2.3.3. *Let $f \in F[x]$ be non-zero. Then we can write*

$$f = af_1 \cdot \cdots \cdot f_r$$

where $a \in F - \{0\}$, $f_i \in F[x]$ monic irreducible and $r \in \mathbb{Z}_{\geqslant 0}$. Moreover, this is unique up to reordering.

Definition 2.3.5. *Let F be a field. We say that F is algebraically closed if every non-constant polynomial in $F[x]$ has at least one root. (We say that ξ is a root of $f \in F[x]$ if $f(\xi) = 0$.)*

Example.
1. $\mathbb{C}$ is algebraically closed; in fact, $f \in \mathbb{C}[x]$ has exactly $\deg f$ many roots
2. $\mathbb{R}$ is not algebraically closed, e.g. $f(x) = x^2 + 1$
3. In $\mathbb{Z}/8\mathbb{Z}[x]$, consider $f(x) = x^2 - 1$, then $x = 1, 3, 5, 7$ are roots even though $degf = 2$.

Lemma 2.3.4. *F is algebraically closed if and only if every irreducible polynomial has* deg 1.

Example. 1. In $\mathbb{C}[x]$, irreducibles are linear polynomials like $x - \xi$

2. In $\mathbb{R}[x]$, $x^2 + 1$ is also irreducible; in fact, one can show that irreducibles are linear or quadratic in this case.

Proposition 2.3.5. *Let F be a field, let $m \in F^\times - (F^\times)^2$ (where $F^\times$ denotes $F - \{0\}$ and the whole notation denotes elements m such that $x^2 - m$ has no root in F). Then $K = F(\sqrt{m}) := \{a + b\sqrt{m} : a, b \in F\}$ is a field containing $F = \{a + 0\sqrt{m} : a \in F\}$ as a subfield. Here addition and multiplcation are:*

$$(a + b\sqrt{m}) + (c + d\sqrt{m}) = (a + c) + (b + d)\sqrt{m},$$

$$(a + b\sqrt{m})(c + d\sqrt{m}) := (ac + bdm) + (ad + bc)\sqrt{m}.$$

Proof. One can check the field axioms one by one. We only provide the check for multiplicative inverse: if $a + b\sqrt{m} \neq 0 + 0\sqrt{m} = 0$, then

$$\frac{1}{a + b\sqrt{m}} = \frac{a - b\sqrt{m}}{(a + b\sqrt{m})(a - b\sqrt{m})} = \frac{a}{a^2 - mb^2} - \frac{b}{a^2 - mb^2}\sqrt{m},$$

where $a^2 - mb^2 \neq 0$ due to assumptions about a, b and m. □

Remark. In particular, for $F = \mathbb{Z}_p$, K is a field with p^2 elements.

Remark. $\mathbb{Z}_p(\sqrt{m}) \cong \mathbb{Z}_p/(x^2 - m)$.

Theorem 2.3.6. *Let f be an irreducible polynomial of degree n in $\mathbb{Z}_p[x]$. Then $\mathbb{Z}_p[x]/(f)$ is a field of order p^n.*

3 Basic concepts

3.1 Vector spaces

Definition 3.1.1. *Let F be a field. A vector space V over F is a non-empty set with*

1. *vector addition: $V \times V \to V$, $(v_1, v_2) \mapsto v_1 + v_2$*
2. *scalar multiplication: $F \times V \to V$, $(\lambda, v) \mapsto \lambda v$*

such that:

1. *$(V, +)$ is an abelian group; call $0 (= 0_V)$ the identity element; $-v :=$ inverse of $v \in V$.*
2. *associative law for scalar multiplication: for all $\lambda, \mu \in F$, $v \in V$*

$$\lambda(\mu v) = (\lambda\mu)v$$

3. *distributive laws: for all* $\lambda, \mu \in F$, $u, v \in V$

$$(\lambda + \mu)v = \lambda v + \mu v$$

$$\lambda(u + v) = \lambda u + \lambda v$$

4. $1_F v = v$ *for all* $v \in V$.

Lemma 3.1.1. 1. $0_F v = 0_V$

2. $\lambda 0_V = 0_V$

3. $(-1_F)v = -v$

Example. 1. $V = F$; more generally, $V = F^n$

2. $V = \mathrm{Hom}(F, F)$; more generally, for any set X, V = $\mathrm{Hom}(X, F)$, where:

$$(f + g)(x) := f(x) + g(x)$$

$$(\lambda f)(x) := \lambda f(x).$$

3. $V = F[x]$

4. $V = M_{m \times n}(F)$

5.

$$\begin{aligned} V &= \text{set of sequences with entries in } F \\ &= \{(a_1, \dots, a_i, \dots) : a_i \in F \, \forall i \in \mathbb{Z}_{\geqslant 0}\} \end{aligned}$$

6.

$$V' = \{(a_1, \dots, a_i, \dots) : a_i \in F \, \forall i \in \mathbb{Z}_{\geqslant 0}, \exists N \in \mathbb{Z}_{\geqslant 0} \text{ such that } a_i = 0 \, \forall i \geqslant N\}$$

We have $V' \subset V$. Note that this is isomorphic to $V = F[x]$.

7. If $F = \mathbb{R}, \mathbb{C}$ (fields with 'analytic/topological flavour'). We can talk about functions with extra properties, e.g. $V = C^\infty([0, 1])$, $V = C^\infty(\mathbb{R})$.

Definition 3.1.2. *Let* V *be an* F*-vector space,* $W \subset F$*. We say that* W *is a subspace of* V *if* W *is a vector space itself with the vector addition and scalar multiplication on* V*.*

Proposition 3.1.2. W *is a subspace of* V *if and only if:*

1. $W \neq \emptyset$

2. W *is closed with respect to vector addition and scalar multiplication i.e. whenever* $w_1, w_2 \in W$, $\lambda \in F$, *we must have* $w_1 + w_2 \in W$ *and* $\lambda w_i \in W$.

Remark. For the closure above, it suffices to check that $w_1 + \lambda w_2 \in W$.

3.2 The lattice of subspaces

Proposition 3.2.1. *Let V be an F-vector space. Let $\{S_\alpha : \alpha \in I\}$ be a collection of subspaces of V. Then $X := \bigcap_{\alpha \in I} S_\alpha$ is also a subspace.*

Proof. Note that $0_V \in S_\alpha$ for all $\alpha \in I$, i.e. $0_V \in X$. Pick $v, w \in X$, $\lambda, \mu \in F$. Then $v, w \in S_\alpha$ for all $\alpha \in I$. By the definition of subspace, $\lambda v + \mu w \in X$. □

Let $\mathcal{S}(V)$ be collection of subspaces of V. We define a partial order on $\mathcal{S}$: $W_1 \leqslant W_2 \iff W_1 \subset W_2$.

Remark. Given $W_1, W_2 \in \mathcal{S}(V)$, there is always a 'bigger' element in $\mathcal{S}(V)$ that is smaller than both W_1, W_2 (analogous to gcd) which is $W_1 \cap W_2$.

We now explore the analogy to lcm in $\mathcal{S}(V)$.

Proposition 3.2.2. *1. Let $W_1, W_2 \in \mathcal{S}(V)$. Then $W_1 \cup W_2 \in \mathcal{S}(V) \iff W_1 \subset W_2$ or $W_2 \subset W_1$.*

2. If F is infinite, V can never be the union of a finite number of proper subspaces.

Proof.

1. The $\Leftarrow$ direction is obvious. Now suppose W_1 and W_2 are not subsets of each other. Then we can pick $u_1 \in W_2 - W_1$, $u_2 \in W_1 - W_2$. Let $v = u_1 + u_2 \in W_1 \cup W_2$. We get contradiction by considering $v - u_1$ and $v - u_2$.

2. We induct on the number of proper subspaces in the union. The base case is obvious. Consider $U_1 \cup U_2 \cup \cdots \cup U_n$ to be union of proper subspaces of V. We can pick $u_1 \in U_1$ but $u_1 \notin U_i$ for all $i > 1$ (otherwise, we are reduced to the induction hypothesis), and pick $v \in V - U_1$. Then consider vectors of the form $v + \lambda u_1$ which cannot be in U_1. Furthermore, for $\lambda_1 \neq \lambda_2$, $v + \lambda_1 u_1 - (v + \lambda_2 u_1) = (\lambda_1 - \lambda_2) u_1 \notin U_i$ for all $i > 1$. Because F is infinite, there exists λ_0 such that $v + \lambda_0 u_1 \notin U_1 \cup U_2 \cup \cdots \cup U_n$.

□

Definition 3.2.1. *Let $U, W \in \mathcal{S}(V)$;*

$$U + W := \{u + w : u \in U, w \in W\}.$$

We define $U_1 + U_2 + \cdots + U_n$ similarly.

Proposition 3.2.3. *$U + W$ is a vector subspace, and it is the smallest subspace containing both U and W. Similarly, $U_1 + U_2 + \cdots + U_n$ is the smallest subspace containing U_i for all $1 \leqslant i \leqslant n$.*

Proof. Note that

1. $0 = 0_V = 0_V + 0_V \in U + W$

2. let $v_1, v_2 \in U + W$; $\lambda, \mu \in F$. Then

$$\begin{aligned}\lambda v_1 + \mu v_2 &= \lambda(u_1 + w_1) + \mu(u_2 + w_2) \\ &= (\lambda u_1 + \mu u_2) + (\lambda w_1 + \mu w_2) \\ &\in U + W\end{aligned}$$

3. $U + 0$ and $0 + W$ and contained in $U + W$.

4. Suppose V' is the smallest subspace containing U and W. Then clearly $U+W \subset V'$, so $V' = U + W$ by definition of V'.

□

Proposition 3.2.4. *$U_1, \ldots, U_n \in \mathcal{S}(V)$. The following are equivalent:*

1. *whenever $\sum_{i=1}^n u_i = \sum_{i=1}^n u_i'$ for $u_i, u_i' \in U_i$, we must have $u_i = u_i'$ for all i.*
2. *whenever $\sum_{i=1}^n u_i = 0$ for $u_i \in U_i$, we must have $u_i = 0$ for all i.*
3. *$(\sum_{i \neq j} U_i) \cap U_j = \{0\}$ for all j.*

Definition 3.2.2. *We say that the sum $\sum_{i=1}^n U_i$ is a direct sum and write $\bigoplus_{i=1}^n U_i$ whenever the conditions in the previous proposition are met.*

Example. $F = \mathbb{R}$, $V = \mathbb{R}^3$; let

$$\begin{aligned}U_1 &= \{(x, 0, 0) : x \in \mathbb{R}\} \\ U_2 &= \{(0, y, 0) : y \in \mathbb{R}\} \\ U_3 &= \{(0, 0, z) : z \in \mathbb{R}\} \\ U_4 &= \{(a, a, 0) : a \in \mathbb{R}\}.\end{aligned}$$

$U_1 + U_2 + U_3$ is a direct sum; $U_1 + U_2 + U_4$ is not.

Definition 3.2.3. *For an arbitrary set I, suppose $\{U_i : i \in I\}$ is a collection of subspaces of V.*

$$\sum_{i \in I} U_i = \bigcup_{\substack{A \subset I \\ |A| < \infty}} \sum_{\alpha \in A} U_\alpha$$

Example. Any matrix $A \in M_n R$, where R is a ring in which 2 is a unit. Then we can write

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$$

i.e. A is a sum of a symmetric and a skew-symmetric matrix. One can check that Sym_n and Skew-Sym_n are subpaces, and furthermore $M_n(R) = \text{Sym}_n \oplus \text{Skew-Sym}_n$.

3.3 Span and linear independence

Definition 3.3.1. *Let V be an $F-$vector space. Let $S \subset V$. The subspace spanned by S*

$$\text{span}(S) = \langle S \rangle := \{r_1 v_1 + \cdots + r_n v_n : r_i \in F, v_i \in S\}.$$

We say that $r_1 v_1 + \cdots + r_n v_n$ is a linear combination of $v_1, dots, v_n$. If $V = \text{span}(S)$, then we say that S is a spanning set of V.

Remark. Any superset of a spanning set is also a spanning set.

Definition 3.3.2. *A nonempty set $S \subset V$ is linearly independent if for any $v_1, \ldots, v_n \in S$, $r_1, \ldots, r_n \in F$, we have*

$$\sum_{i=1}^{n} r_i v_i = 0 \implies r_1 = r_2 = \cdots = r_n = 0.$$

Remark.
1. If S is linearly independent, then $0 \notin S$.
2. Any nonempty subset of a linearly independent set is linearly independent.

Lemma 3.3.1. *Let $S \subset V$; the following are equivalent:*

1. *S is linearly independent*
2. *every vector in $\text{span}(S)$ is a unique linear combination of vectors in S*
3. *no vector in S is a linear combination of other vectors in S, i.e. for all $v \in S$, $v \notin \text{span}(S - \{v\})$.*

Theorem 3.3.2. *Let $S \subset V$; the following are equivalent:*

1. *S linearly independent and spans V*
2. *For each vector $v \in V$, there exists unique $v_1, \ldots, v_n \in S$ and $r_1, \ldots, r_n \in F$ such that*
$$v = \sum_{i=1}^{n} r_i v_i.$$
3. *S spans V but any proper subset of S does not span V*
4. *S is linearly independent but any proper superset of S is not linearly independent.*

Definition 3.3.3. *Let $S \subset V$; We say that S is a basis of V if it satisfies any of the equivalent criteria above.*

Corollary 3.3.3. *A finite set $S = \{v_1, \ldots, v_n\}$ is a basis of V if and only if*

$$V = \text{span}(v_1) \oplus \text{span}(v_2) \oplus \cdots \oplus \text{span}(v_n).$$

Example. Let $V = F^n = \{(a_1, \dots a_n) : a_i \in F\}$. Then $S = \{e_1, \dots, e_n\}$ is a basis of V.

Remark. The zero vector space $\{0\}$ does not have a basis.

Theorem 3.3.4. *Let V be a non-zero F-vector space. Let I be a linearly independent set in V, and S is a spanning set of V containing I. Then there exists a basis $\mathcal{B}$ such that such that $I \subset B \subset S$. In particular:*

1. *every non-zero vector space has a basis*
2. *every linearly independent set can be extended to be a basis*
3. *every spanning set can be shrunk to be a basis.*

Lemma 3.3.5 (Zorn's lemma)**.** *Let S be a poset. If every totally ordered subset (chain) of S has an upper bound, then S has a maximal element.*

Proof of theorem. Consider

$$\mathcal{P} := \{\Lambda \subset V : I \subset \Lambda \subset S, \Lambda \text{ is linearly independent}\}.$$

Observe that $\mathcal{P}$ is nonempty ($I \in \mathcal{P}$). We claim that if $\mathcal{C} = \{I_k : k \in K\}$ is chain in $\mathcal{P}$, then $U = \bigcup_{k\in K} I_k$ is an upper bound in $\mathcal{P}$ (in particular, it is in $\mathcal{P}$). Indeed, we check that firstly, since $I \subset I_k \subset S$, we have $I \subset U \subset S$. Secondly, U is linearly independent, because given $v_1, \dots, v_l \in U$ such that $\sum_{i=1}^{l} c_i v_i = 0$, then since $\mathcal{C}$ is a chain, we can find some I_m which contains all $v_1, \dots, v_l$. Then we get $c_1 = \cdots = c_l = 0$ by linear independence of I_m.

By Zorn's lemma, there exists $\mathcal{B} \in \mathcal{P}$ which is maximal. We now claim that $\mathcal{B}$ is a basis. Indeed, $\mathcal{B}$ is linearly independent. Also, suppose $v \in V$ is such that $v \notin \operatorname{span} \mathcal{B}$. Then $\mathcal{B} \cup \{v\}$ is linearly independent. This contradicts the maximality of $\mathcal{B}$. So $\mathcal{B}$ is also spanning. □

Remark. One might try to argue with

$$\mathcal{P}' := \{\Lambda \subset V : I \subset \Lambda \subset S, \Lambda \text{ is spanning}\}.$$

One then need to check that if $\{I_k : k \in K\}$ are all spanning, then $\bigcap I_k$ is also spanning. Alas, this is not true. For example, $V = \mathbb{Q}$, $F = \mathbb{Q}$, $S = \mathbb{Q} = \{r_1, r_2, \dots\}$. Let $I_k := S - \{r_1, \dots, r_k\}$. Then I_k are all spanning, but $\bigcap I_k = \emptyset$.

Definition 3.3.4. *Let V be a vector space. If there is a basis of V consisting of finitely many elements, we say that V is finite-dimensional. Otherwise, we say that V is infinite dimensional.*

Theorem 3.3.6. *Let V be an $F-$vector space; assume $\{v_1, \dots, v_n\}$ is a linearly independent set and $\{s_1, \dots, s_m\}$ is a spanning set for V, then $n \leqslant m$.*

Proof. Consider listing the vectors in the following way, where the left side is always spanning:

$$s_1, s_2, \ldots, s_m; v_1, v_2, \ldots, v_n$$

$$v_1, s_1, s_2, \ldots, s_m; v_2, \ldots, v_n.$$

Since v_1 is a linearly combinations of s_i's, one of the s_i's, say s_1, can be written using $\{v_1, s_2, s_3, \ldots, s_m\}$.

$$v_1, s_2, \ldots, s_m; v_2, \ldots, v_n.$$

We can repeat the argument and get (wlog)

$$v_1, v_2, \ldots, s_m; v_3, \ldots, v_n.$$

Suppose $n > m$, then eventually we get

$$v_1, v_2, \ldots, v_m; v_{m+1}, \ldots, v_n.$$

Now the left side is still spanning, contradicting the assumption that $\{v_1, \ldots, v_n\}$ is linearly independent. □

Corollary 3.3.7. *If V is finite-dimensional, and $\mathcal{B}_1, \mathcal{B}_2$ are bases of V, then $|\mathcal{B}_1| = |\mathcal{B}_2|$.*

Definition 3.3.5. *The dimension of a finite-dimensional vector space V is the number of elements in its basis.*

Proposition 3.3.8. *Let S, T be two subspaces of a finite-dimensional vector space V. Then*

$$\dim(S + T) = \dim(S) + \dim(T) - \dim(S \cap T).$$

In particular, if S is a complement of T (i.e. $S + T = V$ and $S \cap T = \{0\}$), then $V = S \oplus T$ and

$$\dim V = \dim(S) + \dim(T).$$

Proof. Pick a basis $\mathcal{B} = \{b_i\}$ of $S \cap T$. Then we can extend $\mathcal{B}$ to a basis $\mathcal{B} \cup \mathcal{A}$ of S and extend $\mathcal{B}$ to a basis $\mathcal{B} \cup \mathcal{C}$ of T. It is not hard to check that $\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ is a basis of $S + T$. Then

$$\begin{aligned} \dim S + \dim T &= |\mathcal{B} \cup \mathcal{A}| + |\mathcal{B} \cup \mathcal{C}| \\ &= |\mathcal{A}| + |\mathcal{C}| + 2|B| \\ &= |\mathcal{A}| + |\mathcal{B}| + |\mathcal{C}| + |\mathcal{B}| \\ &= \dim(S + T) + \dim(S \cap T). \end{aligned}$$

□

4 Linear maps

4.1 Linear maps and matrices

Let U, V be $F-$vector spaces. We want to understand linear maps $U \to V$.

Definition 4.1.1. *A map $\alpha : U \to V$ is linear if for all $u, u' \in U$, $\lambda \in F$,*

1. $\alpha(u + u') = \alpha(u) + \alpha(u')$
2. $\alpha(\lambda u) = \lambda\alpha(u)$.

Example. 1. $\mathrm{Id} = \mathrm{Id}_V : V \to V$; $c\,\mathrm{Id}_V$; zero map $V \to V$

2. For $A \in M_{m\times n}(F)$, we have

$$\begin{aligned} \mathcal{L}_A : F^n &\to F^m \\ x &\mapsto Ax \end{aligned}$$

3. $D^\infty(\mathbb{R}) :=$ infinitely many times differentiable functions

$$\begin{aligned} F : D^\infty(\mathbb{R}) &\to D^\infty(\mathbb{R}) \\ f &\mapsto f' \end{aligned}$$

4. $C^\infty(\mathbb{R}) :=$ continuous functions $\mathbb{R} \to \mathbb{R}$

$$f \mapsto \left(x \mapsto \int_0^x f(t)\,dt \right)$$

5.

$$\begin{aligned} f : F &\to F \\ x &\mapsto x^2 \end{aligned}$$

where $\mathrm{char}\, F = 2$.

We use the notation $\mathcal{L}(V, W)$ to denote the set of all linear transformations $V \to W$.

Lemma 4.1.1. *Let $\alpha \in \mathcal{L}(F^n, F^m)$. Then there exsits a unique $A \in M_{m\times n}(F)$ such that $\alpha = \mathcal{L}_\alpha$ (refer to the previous example for the notation).*

Proof. The i-th column of A is $\alpha(e_i^{(n)})$. □

Remark. Let U, V be vector spaces, let $\mathcal{B}$ be a basis of U. Pick $\{v_b \in V : b \in \mathcal{B}\}$. Then the map

$$\begin{aligned} \alpha : U &\to V \\ \sum_{b\in\mathcal{B}} \lambda_b b &\mapsto \sum_{b\in\mathcal{B}} \lambda_b v_b \end{aligned}$$

is the unique linear transformation such that $\alpha(b) = v_b$ for all $b \in \mathcal{B}$.

Proposition 4.1.2. *Let S be a nonempty set and V, W be vector spaces. Then Fun(S, W) is a vector space. In particular $\mathcal{L}(V, W)$ is a vector space.*

Lemma 4.1.3. *Composition of linear transformations is a linear transformation.*

Definition 4.1.2. *Let $\alpha \in \mathcal{L}(U, V)$.*

1. $\ker \alpha := \{u \in U : \alpha(u) = 0_V\} = \alpha^{-1}(0_V)$
2. $\operatorname{im} \alpha := \{\alpha(u) : u \in U\} \subset V$.

Lemma 4.1.4. 1. $\ker \alpha$ *is a subspace of U.*

2. $\operatorname{im} \alpha$ *is a subspace of V.*

Theorem 4.1.5 (rank-nullity theorem)**.** *Let $\alpha \in \mathcal{L}(U, V)$, $\dim U < \infty$. Then*

$$\dim(\operatorname{im} \alpha) + \dim(\ker \alpha) = \dim U.$$

Proof. Pick a basis $\{u_1, \ldots, u_k\}$ of $\ker(\alpha) \subset U$. We can extend this basis to a basis $\{u_1, \ldots, u_k, u_{k+1}, \ldots, u_n\} \subset U$ of U. We claim that $\{\alpha(u_{k+1}), \ldots, \alpha(u_n)\}$ is a basis of $\operatorname{im}(\alpha)$. Indeed,

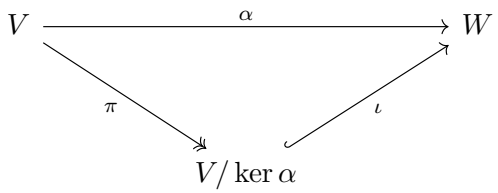
$$\begin{aligned} &\sum_{i=k+1}^{n} c_i \alpha(u_i) = \alpha(\sum_{i=k+1}^{n} c_i u_i) = 0_V \\ \implies &\sum_{i=k+1}^{n} c_i u_i \in \ker \alpha \\ \implies &\sum_{i=k+1}^{n} c_i u_i = \sum_{i=1}^{k} d_i u_i \\ \implies &c_i = 0 \, \forall i, \end{aligned}$$

showing linear independence. Moreover, let $v \in \operatorname{im} \alpha$, then

$$v = \alpha\left(\sum_{i=1}^{n} c_i u_i\right) = \alpha\left(\sum_{i=1}^{k} c_i u_i\right) + \sum_{i=k+1}^{n} c_i \alpha(u_i) = \sum_{i=k+1}^{n} c_i \alpha(u_i).$$

Hence $\{\alpha(u_{k+1}), \ldots, \alpha(u_n)\}$ spans $\operatorname{im} \alpha$. □

Remark. This is a consequence of the first isomorphism theorem for vector spaces:



where $\pi : V \to V/\ker\alpha$ is the quotient map $\pi(v) = v + \ker\alpha$. We define $\iota : V/\ker\alpha \to W$ by $\iota(v + \ker\alpha) = \alpha(v)$ (check that this is well-defined!). Hence, the diagram is commutative. The first isomorphism theorem says that $V/\ker\alpha \cong \operatorname{im}\alpha$. Now assuming everything is finite-dimensional, we get

$$\dim V - \dim\ker\alpha = \dim V/\ker\alpha = \dim\operatorname{im}\alpha$$

which is the rank-nullity theorem. The advantage of this approach is that it makes sense even in the infinite-dimensional case.

Corollary 4.1.6. *Let $\alpha \in \mathcal{L}(U, V)$. The following are equivalent:*

1. *α injective*
2. *$\ker\alpha = \{0_U\}$*
3. *$\dim\ker\alpha = 0$.*

Proposition 4.1.7. *Let $\alpha \in \mathcal{L}(U, V)$, $\dim U = \dim V < \infty$. Then α is injective if and only if it is surjective.*

Proof.

$$\alpha \text{ injective} \iff \dim\ker\alpha = 0 \iff \dim\operatorname{im}\alpha = \dim U = \dim V \iff \alpha \text{ surjective.}$$

□

Let V be an $F-$vector space such that $\dim V = n < \infty$. We want to understand $\mathcal{L}(V, V)$, i.e. $M_{n\times n}(F)$. In particular, we can consider the trace:

$$\begin{aligned} M_{n\times n}(F) &\to F \\ (a_{ij}) &\mapsto \sum_{i=1}^{n} a_{ii}. \end{aligned}$$

We would like to prove that things like this are indepedent of the choice of basis. In fact, we would see later that $\mathcal{L}(V, V) \cong V^* \otimes V$ and the trace can be defined as

$$\begin{aligned} V^* \otimes V &\to F \\ f \otimes v &\mapsto f(v). \end{aligned}$$

This definition is basis-free, and even generalises to infinite-dimensional case with some restriction. This is one motivation for our subsequent discussions.

Definition 4.1.3. *An isomorphism is a bijective linear map.*

Lemma 4.1.8. *$V \cong W \iff |\mathcal{B}_V| = |\mathcal{B}_W|$ where $\mathcal{B}_V, \mathcal{B}_W$ are bases of V and W respectively.*

Proof. ($\Rightarrow$) Pick $T : V \cong W$. Then $\{T(v) : v \in \mathcal{B}_V\}$ is a basis of W.

($\Leftarrow$) Given a bijection $\alpha : \mathcal{B}_V \to \mathcal{B}_W$, extend α to a linear map $T : V \to W$. It is clear that T is bijective. □

Definition 4.1.4. *Let V be a finite-dimensional $F-$vector space, $\mathcal{B} = \{v_1, \ldots, v_n\}$ be an ordered basis. For $v = \sum_{i=1}^n \lambda_i v_i \in V$, define $[v]_{\mathcal{B}} := (\lambda_1, \lambda_2, \ldots, \lambda_n)^T$.*

Lemma 4.1.9. *Fix an ordered basis $\mathcal{B}$ of V. If $\dim V = n < \infty$, then*

$$\begin{aligned} \iota_{\mathcal{B}} : V &\to F^n \\ v &\mapsto [v]_{\mathcal{B}} \end{aligned}$$

is an isomorphism.

Definition 4.1.5. *Let V and W be finite-dimensional vector spaces. Let $\alpha \in \mathcal{L}(V, W)$. Fix $\mathcal{B}_V = \{v_1, \ldots, v_n\}, \mathcal{B}_W = \{w_1, \ldots, w_m\}$. Define*

$$[\alpha]_{\mathcal{B}_W, \mathcal{B}_V} := \begin{pmatrix} [\alpha(v_1)]_{\mathcal{B}_W} & [\alpha(v_2)]_{\mathcal{B}_W} & \ldots & [\alpha(v_n)]_{\mathcal{B}_W} \end{pmatrix} \in M_{m \times n}(F).$$

Lemma 4.1.10. *Let V and W be finite-dimensional vector spaces. Let $\alpha \in \mathcal{L}(V, W)$. Fix ordered bases $\mathcal{B}_V = \{v_1, \ldots, v_n\}, \mathcal{B}_W = \{w_1, \ldots, w_m\}$.*

1. *there is a unique $A \in M_{m \times n}(F)$ such that $[\alpha(v)]_{\mathcal{B}_W} = A[v]_{\mathcal{B}_V}$ for all $v \in V$. In fact, $A = [\alpha]_{\mathcal{B}_W, \mathcal{B}_V}$.*
2. $$\begin{aligned} \mathcal{L}(V, W) &\to M_{m \times n}(F) \\ \alpha &\mapsto [\alpha]_{\mathcal{B}_W, \mathcal{B}_V} \end{aligned}$$

   is an isomorphism.

Proof. 1. Our choices of basis induce isomorphisms $\iota_{\mathcal{B}_V} : V \to F^n$ and $\iota_{\mathcal{B}_W} : W \to F^m$. By Lemma 4.1.1, there is a unique $A \in M_{m \times n}(F)$ such that $\mathcal{L}_A = \iota_{\mathcal{B}_W} \circ \alpha \circ \iota_{\mathcal{B}_V}^{-1}$, namely the matrix whose $i-$th column is $\iota_{\mathcal{B}_W} \circ \alpha \circ \iota_{\mathcal{B}_V}^{-1}(e_i^{(n)}) = [\alpha(v_i)]_{\mathcal{B}_W}$. Hence A is precisely $[\alpha]_{\mathcal{B}_W, \mathcal{B}_V}$.

2. The map is clearly linear. Moreover, because $[\alpha]_{\mathcal{B}_W, \mathcal{B}_V}$ encodes $\alpha(v_i)$ (for all i), there is an inverse.

□

Corollary 4.1.11. *Let V, W, X be finite-dimensional. Let $\alpha \in \mathcal{L}(V, W), \beta \in \mathcal{L}(W, X)$. Pick ordered bases $\mathcal{B}_V, \mathcal{B}_W, \mathcal{B}_X$. We have*

$$[\beta \circ \alpha]_{\mathcal{B}_X, \mathcal{B}_V} = [\beta]_{\mathcal{B}_X, \mathcal{B}_W} [\alpha]_{\mathcal{B}_W, \mathcal{B}_V}.$$

Proof. For all $v \in V$

$$\begin{aligned}[\beta \circ \alpha]_{\mathcal{B}_X,\mathcal{B}_V}[v]_{\mathcal{B}_V} &= [(\beta \circ \alpha)(v)]_{\mathcal{B}_X} \\ &= [\beta(\alpha(v))]_{\mathcal{B}_X} \\ &= [\beta]_{\mathcal{B}_X,\mathcal{B}_W}[\alpha(v)]_{\mathcal{B}_W} \\ &= [\beta]_{\mathcal{B}_X,\mathcal{B}_W}[\alpha]_{\mathcal{B}_W,\mathcal{B}_V}[v]_{\mathcal{B}_V} \\ \implies [\beta \circ \alpha]_{\mathcal{B}_X,\mathcal{B}_V} &= [\beta]_{\mathcal{B}_X,\mathcal{B}_W}[\alpha]_{\mathcal{B}_W,\mathcal{B}_V}.\end{aligned}$$

□

Proposition 4.1.12. *Let V and W be finite-dimensional vector spaces. Let $\alpha \in \mathcal{L}(V,W)$. There exists ordered bases $\mathcal{B}_V = \{v_1,\dots,v_n\}, \mathcal{B}_W = \{w_1,\dots,w_m\}$ such that if $[\alpha]_{\mathcal{B}_W,\mathcal{B}_V} = (a_{ij})$, then*

$$a_{ij} = \begin{cases} 1,\ i = j = 1,2,\dots,\operatorname{rk}\alpha \\ 0,\ \textit{otherwise} \end{cases}$$

Proof. Pick a basis $\{v_{r+1},\dots,v_n\}$ of $\ker\alpha$. Extend to a basis $\{v_1,\dots,v_r,v_{r+1},\dots,v_n\}$ of V. Then $\{\alpha(v_1),\dots,\alpha(v_r)\}$ is a basis of $\operatorname{im}\alpha$. Then we extend this to a basis $\{w_1,\dots,w_r,w_{r+1},\dots,w_m\}$ of W. Then $\operatorname{rk}(\alpha) = r$ and $[\alpha]_{\mathcal{B}_W,\mathcal{B}_V}$ is of the desired form. □

4.2 Dual spaces

Definition 4.2.1. *Let V be an $F-$vector space. We define the dual space of V, $V^* := \mathcal{L}(V,F)$; elements of V^* are called linear functionals.*

Example. 1.

$$\begin{aligned}\operatorname{ev}_a : V = F[x] &\to F \\ f &\mapsto f(a)\end{aligned}$$

2.

$$\begin{aligned}\operatorname{Trace} : V = M_{n\times n}(F) &\to F \\ (a_{ij}) &\mapsto \sum_{i=1}^{n} a_{ii}.\end{aligned}$$

Remark. Let $\alpha : V \to F$. Note that $\operatorname{rk}\alpha$ is either 0 (α is identically zero) or 1. Applying rank-nullity, we have unless $\alpha = 0$, $\dim\ker\alpha = \dim V - 1$.

Proposition 4.2.1. *1. For all $v \in V - \{0\}$, there exists $f \in V^*$ such that $f(v) \neq 0$*

2. $v \in V$ is zero vector if and only if $f(v) = 0$ for all $f \in V^*$

3. If $f(v) \neq 0$, then $V = \operatorname{span}(v) \oplus \ker f$

4. *if $f, g \in V^*$, both non-zero, then $\ker f = \ker g$ if and only if $f = \lambda g$ for some $\lambda \in F$.*

Proof. 1. Extend $\{v\}$ to a basis, then define $f(v) \neq 0$.

2. Because f is a linear functional, $f(0) = 0$. The converse follows from above.

3. Pick $x \in \operatorname{span}(v) \cap \ker f$, then $x = av$ for some $a \in F$ and $f(x) = 0$. So $f(av) = af(v) = 0$, implying $x = 0$. Hence $\operatorname{span}(v) \cap \ker f = \{0\}$. Moreover, let $x \in V$. Then
$$x = x - \frac{f(x)}{f(v)}v + \frac{f(x)}{f(v)}v$$
and we can check that $f\left(x - \frac{f(x)}{f(v)}v\right) = f(x) - \frac{f(x)}{f(v)}f(v) = 0$.

4. If $f = \lambda g$, then obviously $\ker f = \ker g$. Conversely, suppose $\ker f = \ker g =: K$. Pick $x \notin K$, by above, $V = \operatorname{span}(x) \oplus K$. Of course, $f|_K = \lambda g|_K$ for any $\lambda \in F$. Then let $\lambda := \frac{f(x)}{g(x)}$ to ensure that
$$\lambda g(cx) = c\frac{f(x)}{g(x)}g(x) = cf(x),$$
so $f = \lambda g$ on the whole of V, by the previous part.

□

Definition 4.2.2 (Dual basis)**.** *Let $\mathcal{B} = \{v_i : i \in I\}$ be a basis of V. Define $v_i^* \in V^*$ by setting $v_i^* = \delta_{v_i}$.*

Theorem 4.2.2. *1. $\mathcal{B}^* = \{v_i^* : i \in I\}$ is linearly independent (in V^*)*

2. If V is finite-dimensional, then $\mathcal{B}^*$ is also spanning and hence is a basis (thus justifying the term 'dual basis')

Proof. 1. Suppose $\sum_{i=1}^n a_i v_i^* = 0$ in V^*. Then evaluating on v_i gives $a_i = 0$ for all i.

2. For any $f \in V^*$, check that $f = \sum_{i \in I} f(v_i)v_i^*$ (this is a finite sum!) by evaluating both sides at v_i's.

□

Remark.

Corollary 4.2.3. *When V is finite-dimensional, $\dim V^* = \dim V$, so $V^* \cong V$.*

Note that Corollary 4.2.3 is false when $\dim V = \infty$:

Example. Let V be any infinite-dimensional vector space over $\mathbb{Z}_2$ with a countable basis $\mathcal{B}$. Any vector $v \in V$ is an indication of a finite subset of $\mathcal{B}$, i.e. there is an injection from V to the set of finite subsets of $\mathcal{B}$, which is countable. So V is (at most) countable. However, there is a surjection $V^* \to \mathcal{P}(\mathcal{B})$. So V^* has strictly greater cardinality than V.

We can take dual of the dual space V^*. This gives rise to the double dual $V^{**} = (V^*)^* = \mathcal{L}(V^*, F)$. Unlike for V^*, there is a canonical/natural isomorphism $V \cong V^{**}$ when $\dim V < \infty$: define

$$\begin{aligned} \mathrm{ev} : V &\to V^{**} \\ v &\mapsto (T \mapsto T(v)). \end{aligned}$$

One can check that this is a linear map.

Proposition 4.2.4. *The above $ev : V \to V^{**}$ is an isomorphism when $\dim V < \infty$.*

Proof. It suffices to show that ev is injective. Let $v \in \ker \mathrm{ev}$, then $\mathrm{ev}(v) \in \mathcal{L}(V^*, F)$ is the zero map, i.e. $\mathrm{ev}(f) = 0$ for all $f \in V^*$. By Proposition 4.2.1, $v = 0$. □

4.3 Invariant subspaces

We write $\mathcal{L}(V) := \mathcal{L}(V, V)$. For $\alpha \in \mathcal{L}(V)$, we abbreviate $[\alpha]_{B_V, B_V}$ as $[\alpha]_{B_V}$.

Definition 4.3.1. *Let $\alpha \in \mathcal{L}(V)$, $W \subset V$ a subspace. We say that W is $\alpha-$invariant if $\alpha(W) \subset W$, i.e. $\alpha|_W$ is in $\mathcal{L}(W)$.*

Example.
1. $\ker \alpha, \operatorname{im} \alpha$ are always α-invariant.
2. Let $v \in V$. Then $\operatorname{span}\{\alpha^i(v)\}_{i \in \mathbb{Z}_{>0}}$ is the smallest $\alpha-$invariant subspace containing v.

Lemma 4.3.1. *Let $\alpha \in \mathcal{L}(v)$, $\dim V < \infty$; suppose $V = U \oplus W$ and U, W are $\alpha-$invariant. Pick ordered basis $\mathcal{B}_V = \mathcal{B}_U \cup \mathcal{B}_W$. Then*

$$[\alpha]_{\mathcal{B}_V} = \begin{pmatrix} [\alpha]_{\mathcal{B}_U} & \\ & [\alpha]_{\mathcal{B}_W} \end{pmatrix}.$$

4.4 The minimal polynomial

We observe that $\mathcal{L}(V)$ can be considered as a non-commutative ring (with composition as the ring multiplication). Let $p = \sum_{i=0}^{\deg p} a_i x^i \in F[x]$ and $\alpha \in \mathcal{L}(V)$. Then $p(\alpha) = \sum a_i \alpha^i \in \mathcal{L}(V)$.

Remark.
1. $\alpha\beta = \beta\alpha \implies p(\alpha)q(\beta) = q(\beta)p(\alpha)$
2. $\operatorname{span}\{\alpha^i(v)\}_{i \in \mathbb{Z}_{>0}} = \{p(\alpha)v : p \in F[x]\}$

Definition 4.4.1. *Let $\alpha \in \mathcal{L}(V)$, $p \in F[x] - \{0\}$. Then we say that α satisfies $p(x)$ if $p(\alpha) = 0$ in $\mathcal{L}(V)$.*

Lemma 4.4.1. *If $\dim V < \infty$, then there exists $p(x) \in F[x]$ with $\deg p \leqslant n^2$ such that α satisfies $p(x)$.*

Proof. The collection $\{1, \alpha, \alpha^2, \ldots, \alpha^{n^2}\} \subset \mathcal{L}(V)$ has $n^2 + 1 > \dim L(V)$ elements, so it is not linearly indepedent. □

Remark. The above fails if $\dim V = \infty$. Consider 'shift' in $F[x]$. This shows that there is no minimal polynomial in this case.

Definition 4.4.2. *Let $\alpha \in \mathcal{L}(V)$. The minimal polynomial of α, denoted $m_\alpha(x) \in F[x]$, is some polynomial (if exists) such that:*

1. $m_\alpha(\alpha) = 0$
2. *if $p(x) \in F[x] - \{0\}$ is such that $p(\alpha) = 0$, then $\deg p \geqslant \deg m_\alpha$*
3. *m_α is monic.*

Lemma 4.4.2. *Let $\alpha \in \mathcal{L}(V)$; suppose α satistifies some non-zero polynomial. Then m_α exsists and is unique. Moreover, α satistifes $p(x)$ if and only if $m_\alpha \mid p(x)$.*

Proof. Define $X_\alpha := \{p(x) \in F[x] : p(\alpha)\}$. By assumption $X_\alpha \neq \emptyset$. Then we can define $m_\alpha =$ a monic polynomial in X_α having minimal degree. Moreover, we can write $p(x) = q(x)m_\alpha(x) + r(x)$ where $r(x) = 0$ or $\deg r < \deg m_\alpha$. Now if $p(\alpha) = 0$, then $r(\alpha) = 0$, so $r(x) = 0$ by minimality of m_α. Conversely, if $m_\alpha \mid p(x)$, then clearly α satistifes $p(x)$. Finally, it is then clear that m_α is unique. □

Remark. Notice that the above argument does not use the linearity of α. In fact, observe that $X_\alpha \subset F[x]$ is an ideal. So this argument is essentially showing that any Euclidean domain is a principal ideal domain. In this case, m_α is the generator of the ideal X_α.

Corollary 4.4.3. *Let $\alpha \in \mathcal{L}(V)$ with a minimal polynomial m_α. Suppose $U \subset V$ is α-invariant. Then $m_{\alpha|_U} \mid m_\alpha$ and $m_{\tilde{\alpha}} \mid m_\alpha$, where $\tilde{\alpha} : V/W \to V/W$, $\tilde{\alpha}(v+U) = \alpha(v)+U$.*

Proof. Note that $m_\alpha(\alpha|_U) = 0$, so $m_{\alpha|_U} \mid m_\alpha$ by minimality of $m_{\alpha|_U}$. Similarly, $m_\alpha(\tilde{\alpha})(v + U) = m_\alpha(\alpha)v + U = 0 + U$. □

Corollary 4.4.4. *Let $\alpha \in \mathcal{L}(V)$; $W_1, \ldots, W_k \subset V$ are α-invariant. Set $m_i(x) = m_{A|_{W_i}}(x)$; if $\sum W_i = V$, then $m_\alpha(x) = \operatorname{lcm}(m_1, \ldots, m_k)$.*

Proof. We know that $m_i \mid m_\alpha$ for all i. Moreover, suppose $m \in F[x]$ is a common multiple of all m_i. Then $m(\alpha)v = m(\alpha)(w_1+\cdots+w_k) = m(\alpha)w_1+\cdots+m(\alpha)w_k = 0$. □

Remark. We would see later that $\chi_\alpha(x) = \Pi\chi_{A|_{W_i}}(x)$.

Theorem 4.4.5. *Let $\alpha \in \mathcal{L}(V)$, $\dim V = n < \infty$. Then $\deg m_\alpha \leqslant n$.*

Proof. To find $p(x) \in F[x]$ such that $p(\alpha) = 0$ and $\deg p \leqslant n$. If $\dim V = 1$, then $\alpha(v) = kv$, so $p(x) := x - k$; α satisfies $p(x)$.

Now we proceed by induction; assume the claim is true for all finite-dimensional having $\dim < n$.

Do induction on n; pick $v \neq 0 \in V$. Consider $\{v, \alpha(v), \ldots, \alpha^n(v)\} \subset V$. Then $\sum c_i a^i(v) = 0$, so $p(\alpha)v = 0$. Now let $W = \ker p(\alpha) \subset V$. Observe that if $\alpha \circ \beta = \alpha \circ \beta$, then $\ker(\alpha)$ and $\operatorname{im}(\alpha)$ are $\beta-$invariant. Now apply this to α and $p(\alpha)$, we get that

W is α-invariant. Then the map $\tilde{\alpha} : V/W \to V/W$, $\tilde{\alpha}(v+W) = \alpha(v) + W$ is well-defined. Let $q = m_{\tilde{\alpha}}$ and $r = m_{\alpha|_W}$. By hypothesis $\deg q \leqslant \dim V/W = \dim V - \dim W$, $\deg r \leqslant \dim W$. Now we claim that $q(\alpha)r(\alpha)v = 0$ for all $v \in V$. Indeed, because $q(\alpha)r(\alpha)v = r(\alpha)q(\alpha)v$, it sufficies to show that $q(\alpha)v \in W$, i.e. $q(\alpha)v + W = 0 + W$. By definition of q, $q(\alpha)v + W = q(\tilde{\alpha})(v+W) = 0 + W$. Finally, $\deg q(\alpha)r(\alpha) \leqslant \dim V - \dim W + \dim W = n$, completing the proof. □

Remark. This also follows from Theorem 4.5.11 later.

We accept the basic definition and properties of determinants for now. A detailed discussion is deferred to later. The goal is to define a map

$$\begin{aligned} F[x] &\to \mathcal{L}(V) \\ p(x) &\mapsto p(\alpha) \end{aligned}$$

to 'decompose' V, α.

Definition 4.4.3. *Let R be a commutative ring with 1_R. Let $A = (a_{ij}) \in M_n(R)$. Then*

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}$$

(We would not digress on the relevant facts in group theory.)

4.5 Eigenstuff

Definition 4.5.1. *Let $\alpha : V \to V$, $\lambda \in F$. Define $E_\lambda := \{v \in V : \alpha(v) = \lambda v\}$. If $E_\lambda \neq \{0\}$, then λ is an eigenvalue of α; any vector $v \in E_\lambda - \{0\}$ is an associated eigenvector.*

One can check that E_λ is a subspace. If W is a $1-$dimensional invariant subspace, then there exists λ such that $W = E_\lambda$.

Theorem 4.5.1. *Let V be finite-dimensional. The following are equivalent:*

1. *λ is an eigenvalue*
2. *$\alpha - \lambda I$ is not invertible*
3. *$\det(A - \lambda I) = 0$ for all $A = [\alpha]_{\mathcal{B}}$ where $\mathcal{B}$ is some basis of V.*

Remark. Note that if $\mathcal{B}, \mathcal{B}'$ are two bases of V, then $[\alpha]_{\mathcal{B}} = X^{-1}[\alpha]_{\mathcal{B}'}X$ for $X = [\mathrm{Id}]_{\mathcal{B}',\mathcal{B}}$. Then

$$\det([\alpha]_{\mathcal{B}} - \lambda I) = \det(X^{-1}[\alpha]_{\mathcal{B}'}X - \lambda I) = \det(X^{-1}([\alpha]_{\mathcal{B}'} - \lambda I)X) = \det([\alpha]_{\mathcal{B}'} - \lambda I).$$

Definition 4.5.2. *Let V be a finite dimensional vector space and $\alpha \in \mathcal{L}(V)$. The characteristic polynomial $\chi_\alpha(x) := \det(x \operatorname{Id} - \alpha)$.*

Remark. One can check that characteristic polynomial is a continuous function of the matrix. But the minimal polynomial is not (consider $\begin{pmatrix} 1 & \epsilon \\ 0 & 1 \end{pmatrix}$ and I).

Next, we see that the characteristic polynomial and minimal polynomial have the same roots:

Proposition 4.5.2. *Let $\alpha \in \mathcal{L}(V)$. The following are equivalent:*

1. *λ is an eigenvalue of α*
2. *λ is a root of χ_α*
3. *λ is a root of m_α.*

Proof. We know that 1. and 2. are equivalent, so it suffices to show that $m_\alpha(\lambda) = 0$ if and only if λ is an eigenvalue. Suppose that $m_\alpha(\lambda) = 0$, then we can write $m_\alpha(x) = (x - \lambda)q(x)$ where $\deg q < \deg m_\alpha$. Since $q(\alpha) \neq 0$, there exists $v \in V$ such that $q(\alpha)v \neq 0$. Then $(\alpha - \lambda \operatorname{Id})(q(\alpha)v) = m_\alpha(\alpha)v = 0$ shows that $q(\alpha)v$ is an eigenvcetor with eigenvalue λ. Conversely, suppose that λ is an eigenvalue with eigenvector v. Observe that for $p \in F[x]$, $p(\alpha)v = p(\lambda)v$. In particular, $m_\alpha(\alpha)v = m_\alpha(\lambda)v = 0$, so $m_\alpha(\lambda) = 0$. □

Example. Let $A = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}$. Then $\chi_A(x) = (x-2)^2(x-1)$. By above, $1, 2$ are roots of m_A. It turns out that $m_A(x) = (x-1)(x-2)$.

Definition 4.5.3. *Let V be a finite dimensional vector space and $\alpha \in \mathcal{L}(V)$. We say that α is diagonalisable if there exists a basis of V consisting of eigenvectors; equivalently, there exists P invertible such that $P^{-1}AP$ is a diagonal matrix.*

Proposition 4.5.3. *Let $\alpha \in \mathcal{L}(V)$ and $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues; then $\sum_{i=1}^k E_{\lambda_i}$ is in fact a direct sum $\bigoplus_{i=1}^k E_{\lambda_i}$.*

Proof. Let $v_i \in E_{\lambda_i}$ for $1 \leqslant i \leqslant k$ be such that $v_1 + \cdots + v_k = 0$. Observe that for any $p(x) \in F[x]$,

$$p(\alpha)(v_1 + \cdots + v_k) = p(\alpha)v_1 + \cdots + p(\alpha)v_k = p(\lambda_1)v_1 + \cdots + p(\lambda_k)v_k = 0.$$

But we can find polynomials p_i such that $p_i(\lambda_i) = 1$ and $p_i(\lambda_j) = 0$ for all $j \neq i$. Applying these p_i to above then gives $v_i = 0$ for all i. □

Corollary 4.5.4. *If $V = \sum E_{\lambda_i}$ for distinct λ_i, then $\mathcal{B} := \bigcup \mathcal{B}_{E_{\lambda_i}}$ (where $\mathcal{B}_{E_{\lambda_i}}$ is basis for E_{λ_i}) is a basis of V full of eigenvectors, i.e. α is diagonalisable.*

Corollary 4.5.5. *If α is diagonalisable, then $V = \bigoplus_{i=1}^k E_{\lambda_i}$ where E_{λ_i} are eigenspaces for distinct eigenvalues λ_i.*

Proof. This follows from observing that $\dim V \leqslant \sum \dim E_{\lambda_i}$. □

Lemma 4.5.6. *Suppose $m_\alpha(x) = (x-c_1)^{r_1}\dots(x-c_k)^{r_k}$ for $c_i \in F$, $r_k \in \mathbb{Z}_{>0}$. Let W be a proper α-invariant subspace. Then there exists $v \notin W$ such that $(\alpha - c_i \operatorname{Id})v \in W$ for some i.*

Proof. Since W is α-invariant, we have the usual map $\tilde{\alpha} : V/W \to V/W$, $\tilde{\alpha}(v+W) = \alpha(v)+W$. From Corollary 4.4.3, we have $m_{\tilde{\alpha}} \mid m_\alpha$, so there is some i such that c_i is a eigenvalue for $\tilde{\alpha}$. Now let $v+W$ be a corresponding eigenvector. We have

$$\alpha(v)+W = \tilde{\alpha}(v+W) = c_i(v+W) = c_iv+W \implies \alpha(v)-c_iv = (\alpha - c_i\operatorname{Id})v \in W$$

and $v \notin W$ because $v+W \neq 0+W$. □

Theorem 4.5.7. *Let V be a vector space of dimension $n < \infty$ and $\alpha \in \mathcal{L}(V)$. Then α is diagonalisable if and only if $m_\alpha(x) = (x-c_1)\dots(x-c_k)$ for distinct eigenvalues c_i and $k \leqslant n$.*

Proof. ($\Rightarrow$) There exists $\mathcal{B} = \{v_1, \dots v_n\}$ basis of V such that v_i are eigenvectors. Let $p(x) = (x-c_1)\dots(x-c_k)$. To show that $p(\alpha) = 0$, it suffices to check that $p(\alpha)v_i = 0$ for all i. But let c be the eigenvalue of v_i, we see that $(\alpha - c\operatorname{Id})v_i = 0$. Moreover, all eigenvalues are roots of m_α, p must be m_α itself.

($\Leftarrow$) Consider $W = \bigoplus E_\lambda \subset V$. Note that W is α-invariant. Suppose that $W \neq V$, then apply Lemma 4.5.6 to get $v \notin W$ such that $w := (\alpha - c\operatorname{Id})v \in W$. Write $w = w_1 + \dots + w_k$ where $\alpha(w_i) = c_iw_i$. Then for any $h(x) \in F[x]$, we have $h(\alpha)w = h(c_1)w_1 + \dots h(c_k)w_k$. Write $m_\alpha(x) = (x-c_i)q(x)$. We can write $q(x) - q(c_i) = (x-c_i)h(x)$. We have $q(\alpha)v - q(c_i)v = h(\alpha)(\alpha - c_i\operatorname{Id})v = h(\alpha)w \in W$. Note that $(\alpha - c_i\operatorname{Id})q(\alpha)v = m_\alpha(\alpha)v = 0$, so $q(\alpha)v \in W$, so $q(c_i)v \in W$. But $v \notin W$, so $q(c_i) = 0$, which is a contradiction. □

Remark. This also follows from Theorem 4.6.1 later.

This theorem provides an efficient way to check whether α is diagonalisable. Say $\chi_\alpha(x) = (x-1)^3(x-2)^2$. Then we just check whether $(\alpha-1)(\alpha-2) = 0$.

Example.

1. $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Then $\chi_A(x) = x^2+1$, so no real eigenvalues. One can check that $m_\alpha = x^2+1$ (e.g. by Cayley-Hamilton later).
2. $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ Then $\chi_B(x) = (x-1)^2$ and $m_\beta = (x-1)^2$. Indeed, B is also not diagonalisable.

Definition 4.5.4. *We say that $\alpha \in \mathcal{L}(V)$ is traingularisable if there exists some basis $\mathcal{B}$ of V such that $[\alpha]_\mathcal{B}$ is traingular.*

Theorem 4.5.8. *α is traingularisable if and only if $m_\alpha(x) = (x-c_1)^{r_1}\dots(x-c_k)^{r_k}$ for $c_i \in F$, $r_k \in \mathbb{Z}_{>0}$.*

Proof. ($\Rightarrow$) Let $[\alpha]_{\mathcal{B}} = (a_{ij})$, then $\chi_\alpha(x) = (x-a_{11})\dots(x-a_{nn}) = (x-c_1)^{t_1}\dots(x-c_k)^{t_k}$. By Theorem 4.5.11, $m_\alpha \mid \chi_\alpha$, so $m_\alpha(x) = (x-c_1)^{r_1}\dots(x-c_k)^{r_k}$ where $r_i \leqslant t_i$.

($\Leftarrow$) α being traingularisable with repsect to $\mathcal{B} = \{v_1, \dots v_n\}$ means that $\alpha(v_i) \in \operatorname{span}\{v_1, \dots v_i\}$. Apply Lemma 4.5.6 to $W = \{0\}$ to obtain v_1. Then apply lemma again to $W = \operatorname{span}\{v_1\}$ (which is α-invariant) to obtain v_2. We see that doing this repeatedly yields the desired v_i. □

Remark. Actually, this is equivalent to Theorem 4.5.11. However, ($\Leftarrow$) does not use Theorem 4.5.11 and we actually use this direction to prove it below.

Corollary 4.5.9. *If F is algebraically closed, then every linear operator on a finite-dimensional F-vector space V is triangularisable.*

We take without proof the following theorem:

Theorem 4.5.10. *Given any field F, there exists a field $\bar{F}$ such that $F \subset \bar{F}$ and $\bar{F}$ is algebraically closed.*

Theorem 4.5.11 (Cayley-Hamilton)**.** $\chi_\alpha(\alpha) = 0$.

Proof. Passing to $\bar{F}$, split $m_\alpha(x) = \Pi(x-\lambda_i)$ where $\lambda_i \in \bar{F}$. Then α is traingularisable in $M_n(\bar{F})$. Now we observe that $\chi_A(A) = 0$ for upper triangular A ('killing' row by row from the bottom). Since $\chi_A \in F[x]$, the theorem follows. □

Corollary 4.5.12. $m_\alpha \mid \chi_\alpha$.

4.6 Canonical forms

Theorem 4.6.1 (primary decomposition)**.** *$\alpha \in \mathcal{L}(V)$; let $m_\alpha = p_1^{r_1}\dots p_k^{r_k}$ where p_i are distinct irreducible monic polynomials in $F[x]$. Set $W_i = \ker p_i^{r_i}(\alpha)$. Then*

1. *$V = \bigoplus W_i$*
2. *W_i is α-invariant*
3. *$m_{\alpha|_{W_i}} = p_i^{r_i}$ for all i.*

Moreover, our W_i is the unique collection of subspaces satisfying the above three conditions.

Proof. 1. Set $f_i(x) = \frac{m_\alpha(x)}{p_i(x)^{r_i}}$, then $\gcd(f_1, \dots, f_k) = 1$, so by Lemma 2.3.2 we can write

$$1 = \sum_{i=1}^{k} g_i(x) f_i(x)$$

for some $g_i \in F[x]$. Evaluating at α gives

$$\mathrm{Id} = \sum_{i=1}^{k} h_i(\alpha) \implies v = \sum_{i=1}^{k} h_i(\alpha) v$$

where we write $h_i(\alpha) := f_i(\alpha)g_i(\alpha)$. Now observe that $m_\alpha \mid p_i(x)^{r_i}h_i(x)$, showing that $h_i(\alpha)v \in W_i$. Thus $V = \sum W_i$. Now suppose for contradiction that it is not a direct sum. Then there exists a shortest list of length $l \geqslant 2$ such that

$$w_1 + \cdots + w_l = 0$$

where w_j are all non-zero and from distinct W_i. Now say $w_1 \in W_i$. Then

$$p_i^{r_i}(\alpha)(w_1 + \cdots + w_l) = p_i^{r_i}(\alpha)w_2 + \ldots p_i^{r_i}(\alpha)w_l = 0$$

where crucially, $p_i^{r_i}(\alpha)w_j \neq 0$ for $2 \leqslant j \leqslant l$: this is because otherwise, say there exists $w_n \in W_j$, $2 \leqslant n \leqslant l$ such that $p_i^{r_i}(\alpha)w_n = 0$, then it contradicts Lemma 2.3.2 applied to $p_i^{r_i}$ and $p_j^{r_j}$. This contradicts the minimality of l.

2. This is obvious.

3. It is clear that $m_{\alpha|_{W_i}} \mid p_i^{r_i}$. Moreover, if $m_{\alpha|_{W_i}} = p_i^{r_i'}$ for some $r_i' < r_i$, then it contradicts the minimality of m_α.

For uniqueness, note that from 3. we get that $W_i \subset \ker p_i^{r_i}(\alpha)$. □

Corollary 4.6.2 (Jordan-Chevalley decomposition)**.** *Let F be algebraically closed; $\alpha \in \mathcal{L}(V)$ where V is finite-dimensional; then there exists a decomposition $\alpha = D + N$ where D is diagonal, N is nilpotent and $DN = ND$.*

Proof. $m_\alpha(x) = (x - \lambda_1)^{r_1} \ldots (x - \lambda_k)^{r_k}$. We have from above,

$$V = \bigoplus_{i=1}^{k} \ker(\alpha - \lambda_i \operatorname{Id})^{r_i}$$

So we collect the corresponding bases $\mathcal{B} = \bigcup \mathcal{B}_i$. Then

$$[\alpha_\mathcal{B}] = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_k \end{pmatrix} = \begin{pmatrix} \lambda_1 I & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_k I \end{pmatrix} + N$$

where $(A_i - \lambda_i I)^{r_i} = 0$, so $N_i := A_i - \lambda_i I$ is nilpotent, implying that $A_i = \lambda_i I + N_i$. We note finally that $ND = DN$. □

Todo: tidy up below.

Random: consider the projection maps for $V = \bigoplus W_i$. Then the decomposition is α-invariant if and only if $\pi_i \alpha = \alpha \pi_i$.

Now we focus on the simplest α-invariant subspaces i.e. the cyclic subspaces $\langle v \rangle_\alpha := \{p(\alpha) : p \in F[x]\} = \operatorname{span}\{\alpha^i v : i \in \mathbb{Z}_{>0}\}$. We have the following algorithm: form $\{v, Av, \ldots, A^i v\}$ until it first becomes linearly dependent. Then $\sum_{i=0}^{k} c_i A^i v = 0$, and we get the minimal polynomial of $\langle v \rangle_\alpha$.

Proposition 4.6.3. *1. $\langle v\rangle_\alpha$ is finite dimensional if and only if there exists $p(x) \in F[x] - \{0\}$ such that $p(\alpha)v = 0$. We define $m_{\alpha,v}(x)$ to be the unique such monic polynomial of least degree.*

2. Let $\dim \langle v\rangle_\alpha = n$. Then $\deg m_{\alpha,v}(x) = n$.

3. $\chi_{\alpha|_{\langle v\rangle_\alpha}}(x) = m_{\alpha|_{\langle v\rangle_\alpha}}(x) = m_{\alpha,v}(x)$.

4. $\mathcal{B} = \{v, \alpha(v), \dots, \alpha^{n-1}(v)\}$ is a basis of $\langle v\rangle_\alpha$.

5. $[\alpha|_{\langle v\rangle_\alpha}]_\mathcal{B} = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$*, where $m_{\alpha,v}(x) = x^n + \sum_{i=0}^{n-1} a_i x^i$. This is the companion matrix $C(m_{\alpha,v}(x))$.*

Proof. We only prove 3.: □

We introduce the notation $\{v\}_\alpha^n := \{v, \alpha v, \dots \alpha^{n-1}v\}$, so that $\langle v\rangle_\alpha := \operatorname{span}\{v\}_\alpha^n$.

Lemma 4.6.4. *Let $\alpha \in \mathcal{L}(V)$ be such that $m_\alpha = f(x)^k$ for $f(x)$ monic irreducible of degree n. Let $v \in \ker f(\alpha)^k - \ker f(\alpha)^{k-1}$ (which is nonempty by minimality of m_α).*

1. $m_{\alpha,v}(x) = f(x)^k$

2. $\bigcup_{i=0}^{k-1}\{f(\alpha)^i v\}_\alpha^n$ is a basis of $\langle v\rangle_\alpha$.

3. $\bigcup_{i=k-j}^{k-1}\{f(\alpha)^i v\}_\alpha^n$ is a basis of $\langle v\rangle_\alpha \cap \ker f(\alpha)^j$, for all $j = 1, \dots, k$.

Remark. The setting is that V having a basis $\mathcal{B} = \{v_{11}, \dots v_{nk}\}$ and we want to decompose into the flag

$$\ker f(\alpha) \subset \ker f(\alpha)^2 \subset \cdots \subset \ker f(\alpha)^k = V.$$

Proof. 1. Since $m_{\alpha,v}(x) \mid m_\alpha(x)$ for any $v \in V$. So $m_{\alpha,v}(x) = f(x)^j$ for some $j \leqslant k$. But note that $j \geqslant k$, otherwise

2. Note $P_{nk-1} := \{p(x) \in F[x] : \deg p \leqslant nk - 1\}$ has two bases

$$\{1, x, \dots, x^{nk-1}\}$$

and (note that $\deg f = n$)

$$\{x^j f(x)^i : 0 \leqslant j \leqslant n-1, 0 \leqslant i \leqslant k-1\}.$$

We can write x^l in terms of $x^j f(x)^i$

3. tutorial?

□

Lemma 4.6.5. *Let* $v_1, \ldots, v_r \in \ker f(\alpha)^l$ *such that*

$$v_i \notin \ker f(\alpha)^{l-1} = \sum_{j=1}^{i-1} \langle v_j \rangle.$$

Then