Contents

Last update: 2024 November 1

1

Introduction

"It is my experience that proofs involving matrices can be shortened by 50% if one throws the matrices out"

-Emil Artin

1.1 Basic Properties of Vector Spaces

Definition 1.1.1. Let F be any field. Let V be a nonempty set with binary operations:

$$V \times V \to B$$
$$(v, w) \mapsto v + w$$

called *vector addition* and

$$F \times V \to V$$
$$(c, v) \mapsto cv$$

called *scalar multiplication*. Then V is an *F-vector space* if the following properties are satisfied:

(1) V is an abelian group, that is:

  (i) there exists a $0_v \in V$ such that $0_v + v = v = v + 0v$,

  (ii) for every $v \in V$ there exists a $-v \in V$ such that $v + (-v) = 0_v = (-v) + v$,

  (iii) for every $u, v, w \in V$, $(u + v) + w = u + (w + v)$, and

  (iv) $v + w = w + v$ for all $v, w \in V$.

(2) $c(v + w) = cv + cw$ for all $c \in F$, $v, w \in V$,

(3) $(c + d)v = cv + dv$ for all $c, d \in F$, $v \in V$,

(4) $(cd)v = c(dv)$ for all $c, d \in F$, $v \in V$,

(5) there exists a $1_F \in F$ such that $1_F v = v$.

Example 1.1.1.

(1) Let F be any field. Define $F^n = \{(a_1, ..., a_n) \mid a_i \in F\}$ as *affine n-space*. Then F^n is an F-vector space.

(2) Let $n \in \mathbf{Z}_{\geqslant 0}$. Define $P_n(F) = \{a_0 + a_1 x + ... + a_n x^n \mid a_i \in F\}$. This is an F-vector space with polynomial addition and scalar multiplication. Define $F[x] = \bigcup_{n \geqslant 0} P_n(F)$. This is also an F-vector space, but either via polynomial addition *or* polynomial multiplication.

(3) Let $m, n \in \mathbf{Z}_{\geqslant 0}$. Set $V = \mathrm{Mat}_{n,m}(F) = \{\text{all } m \times n \text{ matrices with entries in } F\}$. This is an F-vector space with matrix addition and scalar mutliplication. If $m = n$ then write $\mathrm{Mat}_n(F)$ for $\mathrm{Mat}_{n,n}(F)$.

Lemma 1.1.1. *Let V be an F-vector space.*

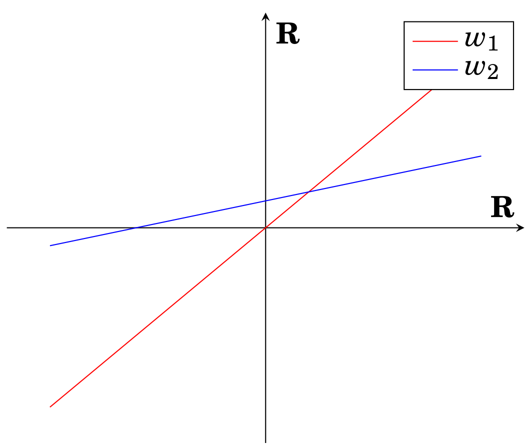
1. *The element $0_v \in V$ is unique,*
2. *$0v = 0_v$ for all $v \in V$,*
3. *$(-1_F)v = -v$ for all $v \in V$.*

Proof. (1) Let $0, 0'$ satisfy the following properties: $0 + v = v$ and $0' + v = v$ for all $v \in V$. Observe that $0 = 0' + 0 = 0 + 0' = 0'$. (2) Note that $0_F v = (0_F + 0_F)v = 0_F v + 0_F v$. Subtracting both sides by $0_F v$ yields $0 = 0_F v$. (3) Observe that $(-1_F)v + v = (-1_F)v + 1_F v = (-1_F + 1_F)v = 0_F v = 0$. Hence $(-1_F)v = -v$. □

Definition 1.1.2. Let V be an F-vector space. We say $W \subseteq V$ is an *F-subspace* (or just *subspace* if F is obvious by context) if W is an F-vector space under the same addition and scalar multiplication.

Example 1.1.2.

(1) Consider the plane $V = \mathbf{R}^2$. Let w_1, w_2 be subsets of $\mathbf{R}^2$ as follows:



Note that w_2 is not a subspace, as it does not contain $0_{\mathbf{R}^2}$. On the other hand w_1 is a subspace; note that every element of w_1 is of the form (x, ax), hence $(x_1, ax_1) + (x_2, ax_2) = (x_1 + x_2, a(x_1 + x_2))$. The other axioms follow similarly.

(2) Let $V = \mathbf{C}$ and $W = \{a + 0i \mid a \in \mathbf{R}\}$. If $F = \mathbf{R}$, then clearly W is an $\mathbf{R}$-subspace. If $F = \mathbf{C}$, then W is not a $\mathbf{C}$-subspace; given $2 \in W$ and $i \in \mathbf{C}$, $2i \notin W$.

(3) $\mathrm{Mat}_2(\mathbf{R})$ is not a subspace of $\mathrm{Mat}_4(\mathbf{R})$, as $\mathrm{Mat}_2(\mathbf{R}) \not\subseteq \mathrm{Mat}_4(\mathbf{R})$.

(4) Let $m, n \in \mathbf{Z}_{\geqslant 0}$. If $m \leqslant n$, then $P_m(F)$ is a subspace of $P_n(F)$.

Lemma 1.1.2. *Let V be an F-vector space and $W \subseteq V$. Then W is an F-subspace of V if:*

(1) W is nonempty,

(2) W is closed under addition, and

(3) W is closed under scalar multiplication.

Proof. Let $x, y \in W$ and $\alpha \in F$, then by assumption $x + \alpha y \in W$. Take $\alpha = -1$, then $x - y \in W$ which implies W is an abelian subgroup of V. Then by (3) it must be the case that W is an F-subspace of V. □

Definition 1.1.3. Let V, W be F-vector spaces. Let $T : V \to W$. We say T is a *linear transformation* (or *linear map*) if for every $v_1, v_2 \in V$ and $c \in F$ we have

$$T(v_1 + cv_2) = T(v_1) + cT(v_2).$$

The collection of all linear maps from V to W is denoted $\mathrm{Hom}_F(V, W)$ (some textbooks write this as $\mathcal{L}(V, W)$).

Example 1.1.3.

(1) Let V be an F-vector space. Define $\mathrm{id}_v : V \to V$ by $\mathrm{id}_v(v) = v$. This is a linear map; i.e., $\mathrm{id}_v \in \mathrm{Hom}_F(V, V)$ because $\mathrm{id}_v(v_1 + cv_2) = v_1 + cv_2 = \mathrm{id}_v(v_1) + c\,\mathrm{id}_v(v_2)$.

(2) Let $V = \mathbf{C}$. Define $T : V \to V$ by $z \mapsto \overline{z}$. Observe that:

$$\begin{aligned} T(z_1 + cz_2) &= \overline{z_1 + cz_2} = \overline{z_1} + \overline{c}\,\overline{z_2} \\ T(z_1) + cT(z_2) &= \overline{z_1} + c\,\overline{z_2}. \end{aligned}$$

Note that these two are only equal if $c = \overline{c}$. Hence $T \in \mathrm{Hom}_F(\mathbf{C}, \mathbf{C})$ if $F = \mathbf{R}$ but not if $F = \mathbf{C}$.

(3) Let $A \in \mathrm{Mat}_{m,n}(F)$. Define $T_A : F^n \to F^m$ by $x \mapsto Ax$. Then $T_A \in \mathrm{Hom}_F(F^n, F^m)$.

(4) Recall that $C^\infty(\mathbf{R})$ is the set of all smooth functions $f : \mathbf{R} \to \mathbf{R}$ (another way of saying "smooth" is "infinitely differentiable"). Let $V = C^\infty(\mathbf{R})$. This is an $\mathbf{R}$-vector space under pointwise addition and scalar multiplication. If $a \in \mathbf{R}$ then:

- $E_a : V \to \mathbf{R}$ defined by $f \mapsto f(a)$ is an element of $\mathrm{Hom}_{\mathbf{R}}(V, \mathbf{R})$,
- $D : V \to V$ defined by $f \mapsto f'$ is an element of $\mathrm{Hom}_{\mathbf{R}}(V, V)$,
- $I_a : V \to V$ defined by $f \mapsto \int_a^x f(t)dt$ is an element of $\mathrm{Hom}_{\mathbf{R}}(V, V)$, and
- $\tilde{E}_a : V \to V$ defined by $f \mapsto f(a)$ (where $f(a)$ is the constant function) is an element of $\mathrm{Hom}_{\mathbf{R}}(V, V)$.

From this, we can express the fundamental theorem of calculus as follows:

$$\begin{aligned} D \circ I_a &= \mathrm{id}_v \\ I_a \circ D &= \mathrm{id}_v - \tilde{E}_a. \end{aligned}$$

Proposition 1.1.3. $\mathrm{Hom}_F(V, W)$ *is an F-vector space.*

Proof. do this □

Lemma 1.1.4. *Let* $T \in \mathrm{Hom}_F(V, W)$. *Then* $T(0_v) = 0_w$.

Proof. do this □

Definition 1.1.4. Let $T \in \mathrm{Hom}_F(V, W)$ be invertible; i.e., there exists a linear transformation $T^{-1} : W \to V$ such that $T \circ T^{-1} = \mathrm{id}_w$ and $T^{-1} \circ T = \mathrm{id}_v$. If this is the case we say T is an *isomorphism* and say V and W are *isomorphic*, written as $V \cong W$.

Proposition 1.1.5. *Let* $T \in \mathrm{Hom}_F(V, W)$ *be an isomorphism. Then* $T^{-1} \in \mathrm{Hom}_F(W, V)$.

Proof. do this □

Example 1.1.4.

(1) Let $V = \mathbf{R}^2$ and $W = \mathbf{C}$. Define $T : \mathbf{R}^2 \to \mathbf{C}$ by $(x, y) \mapsto x + iy$. This is an isomorphism: note that $T \in \mathrm{Hom}_{\mathbf{R}}(\mathbf{R}^2, \mathbf{C})$ because

$$\begin{aligned} T((x_1, y_1) + r(x_2, y_2)) &= \ldots\text{fill this out} \\ &= T((x_1, y_1)) + rT((x_2, y_2)). \end{aligned}$$

Defining $T^{-1} : \mathbf{C} \to \mathbf{R}^2$ by $x + iy \mapsto (x, y)$ (and showing it's linear) clearly satisfies $(T \circ T^{-1})(x + iy) = x + iy$ and $(T^{-1} \circ T)((x, y)) = (x, y)$, hence $\mathbf{R}^2 \cong \mathbf{C}$ as $\mathbf{R}$-vector spaces.

(2) Set $V = P_n(F)$ and $W = F^{n+1}$. Define $T : P_n(F) \to F^{n+1}$ by

$$a_0 + a_1 x + \ldots + a_n x^n \mapsto (a_0, a_1, \ldots, a_n).$$

This is an isomorphism; $P_n(F) \cong F^{n+1}$.

Definition 1.1.5. Let $T \in \mathrm{Hom}_F(V, W)$. Define the *kernel* of T as:

(1) The *kernel of T* is defined as $\ker(T) = \{v \in V \mid T(v) = 0_w\}$.

(2) The *image of T* is defined as $\mathrm{im}(T) = \{w \in W \mid T(v) = w \text{ for some } v \in V\}$.

Lemma 1.1.6. *Let* $T \in \mathrm{Hom}_F(V, W)$. *Then:*

(1) $\ker(T)$ *is a subspace of* V,

(2) $\mathrm{im}(T)$ *is a subspace of* W.

Proof. Let $v_1, v_2 \in \ker(T)$ and $\alpha \in F$. Observe that $T(v_1 + \alpha v_2) = T(v_1) + \alpha T(v_2) = 0_w + \alpha 0_w = 0_w$, hence $v_1 + \alpha v_2 \in \ker(T)$ establishing $\ker(T)$ as a subspace of V.

Let $w_1, w_2 \in \mathrm{im}(T)$ and $\alpha \in F$. Then there exists $v_1, v_2 \in V$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$. Observe that $w_1 + \alpha w_2 = T(v_1) + \alpha T(v_2) = T(v_1 + \alpha v_2)$, hence $w_1 + \alpha w_2 \in \mathrm{im}(T)$ establishing $\mathrm{im}(T)$ as a subspace of W. □

Lemma 1.1.7. *Let* $T \in \mathrm{Hom}_F(V, W)$*.* T *is injective if and only if* $\ker(T) = \{0_v\}$

Proof. Let T be injective. Let $v \in \ker(T)$. Then $T(v) = 0_w = T(0_v)$, and since T is injective $v = 0_v$.

Conversely, assume $\ker(T) = 0_v$. Let $v_1, v_2 \in V$ with $T(v_1) = T(v_2)$. Subtracting both sides by $T(v_2)$ gives $T(v_1) - T(v_2) = 0_w$, and since T is a linear transformation yields $T(v_1 - v_2) = 0_w$. Since the kernel is trivial, it must be the case that $v_1 = v_2$. □

Example 1.1.5. Let $m > n$. Define $T : F^m \to F^n$ by

$$(a_0, a_1, \ldots, a_{n-1}, a_n, a_{n+1}, \ldots, a_m) \mapsto (a_0, a_1, \ldots, a_n)$$

Then $\mathrm{im}(T) = F^n$ and $\ker(T) = \{(0, \ldots, 0, a_{n+1}, a_{n+2}, \ldots, a_m) \in F^m\} \cong F^{m-n}$.

2

Bases and Dimension

2.1 Basic Definitions

Unless otherwise stated assume V to be an F-vector space.

Definition 2.1.1. Let $\mathcal{B} = \{v_i\}_{i \in I}$ be a subset of V where I is an indexing set (possibly infinite). We say $v \in V$ is an *F-linear combination of $\mathcal{B}$* (or just *linear combination*) if there is a set $\{a_i\}_{i \in I}$ with $a_i = 0$ for all but finitely many i such that $v = \sum_{i \in I} a_i v_i$. The collection of F-linear combinations is denoted $\operatorname{span}_F(\mathcal{B})$.

Example 2.1.1. Let $V = P_2(F)$.

(1) Set $\mathcal{B} = \{1, x, x^2\}$. We have $\operatorname{span}_F(\mathcal{B}) = P_2(F)$.

(2) Set $C = \{1, (x-1), (x-1)^2\}$. We have $\operatorname{span}_F(C) = P_2(F)$.

Definition 2.1.2. Let $\mathcal{B} = \{v_i\}_{i \in I}$ be a subset of V. We say $\mathcal{B}$ is *F-linearly independent* (or just *linearly independent*) if whenever $\sum_{i \in I} a_i v_i = 0$ then $a_i = 0$ for all $i \in I$.

Definition 2.1.3. Let $\mathcal{B} = \{v_i\}_{i \in I}$ be a subset of V. We say $\mathcal{B}$ is an *F-basis* (or just *basis*) of V if:

- $\operatorname{span}_F(\mathcal{B}) = V$, and
- $\mathcal{B}$ is linearly independent.

Example 2.1.2. Let $V = F^n$. Let $\mathcal{E}_n = \{e_1, \ldots, e_n\}$ with

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0) \\ e_2 &= (0, 1, 0, \ldots, 0) \\ &\vdots \\ e_n &= (0, 0, 0, \ldots, 1). \end{aligned}$$

We have that $\mathcal{E}_n$ is a basis of F^n and is referred to as the *standard basis*.

2.2 Every Vector Space Admits a Basis

Definition 2.2.1. A *relation* from A to B is a subset $R \subseteq A \times B$. Typically, when one says "a relation on A" that means a relation from A to A; i.e., $R \subseteq A \times A$.

Definition 2.2.2. Let A be a set. An *ordering* of A is a relation R on A that is

(1) *reflexive*: $(a, a) \in R$ for all $a \in A$,

(2) *transitive*: $(a, b), (b, c) \in R$ implies $(a, c) \in R$, and

(3) *antisymmetric*: $(a, b), (b, a) \in R$ implies $a = b$.

If this is the case, we write $(a, b) \in R$ as $a \leqslant_R b$. If A is an ordered set we write it as the ordered pair $(A, \leqslant_R)$ (or just A if the ordering is obvious by context).

Definition 2.2.3. An ordered set $(X, \leqslant_R)$ is *total* if for all $a, b \in X$ we have that $a \leqslant_R b$ or $b \leqslant_R a$.

Definition 2.2.4. Let $(X, \leqslant)$ be an ordered set and $A \subseteq X$ nonempty.

(1) A is called a *chain* if $(A, \leqslant)$ is a total ordering.

(2) A is called *bounded above* if there exists an element $u \in X$ with $a \leqslant u$ for all $a \in A$. Such a u is called an *upperbound* for A.

(3) A *maximal element of A* is an element $m \in A$ such that if $a \geqslant m$, then $a = m$.

Lemma 2.2.1 (Zorn's Lemma)**.** *Let X be an ordered set with the property that every chain has an upperbound. Then X contains a maximal element.*

Theorem 2.2.2. *Let $\mathcal{A}$ and C be subsets of V with $\mathcal{A} \subseteq C$. Assume $\mathcal{A}$ is linearly independent and* $\operatorname{span}_F(C) = V$. *Then there exists a basis $\mathcal{B}$ of V with $\mathcal{A} \subseteq \mathcal{B} \subseteq C$*[1].

Proof. Let $X = \{\mathcal{B}' \subseteq V \mid \mathcal{A} \subseteq \mathcal{B}' \subseteq C,\ \mathcal{B}' \text{ is linearly independent}\}$. We have $\mathcal{A} \in X$, so $X \neq \emptyset$. X is ordered with respect to inclusion, and has an upperbound of C. By Zorn's Lemma we have a maximal element in X, call it $\mathcal{B}$.

Claim: $\operatorname{span}_F(\mathcal{B}) = V$. Suppose towards contradiction it's not, then there exists a $v \in C$ with $v \notin \operatorname{span}_F(\mathcal{B})$. But then $\mathcal{B} \cup \{v\}$ is still linearly independent, and $\mathcal{B} \cup \{v\} \subseteq C$. This gives $\mathcal{B} \subseteq \mathcal{B} \cup \{v\}$, which is a contradiction because $\mathcal{B}$ is maximal in X. Thus $\operatorname{span}_F(\mathcal{B}) = V$. □

2.3 Cardinality and Dimension

Lemma 2.3.1. *A homogenous system of m linear equations in n unknowns with $m < n$ has a nonzero solution.*

Proof. do this □

Corollary 2.3.2. *Let $\mathcal{B} \subseteq V$ with* $\operatorname{span}_F(\mathcal{B}) = V$ *and $|\mathcal{B}| = m$. Any set with more than m elements cannot be linearly independent.*

[1]Given any linearly-independent set $\mathcal{A}$, we can constructing a basis $\mathcal{B}$ by adding elements. Given any spanning set C, we can construct a basis $\mathcal{B}$ by removing elements.

Proof. Let $C = \{w_1, \dots, w_n\}$ with $n > m$. We will show C cannot be linearly independent. Write $\mathcal{B} = \{v_1, \dots, v_m\}$ with $\operatorname{span}_F(\mathcal{B}) = V$. For each i, write

$$w_i = \sum_{j=1}^{m} a_{ji} v_j \text{ for some } a_{ji} \in F.$$

Consider the equations

$$\sum_{i=1}^{n} a_{ji} x_i = 0.$$

By Lemma 2.3.1 there exists nonzero solutions $(x_1, \dots, x_n) = (c_1, \dots, c_n) \neq (0, \dots, 0)$. We have

$$\begin{aligned} 0 &= \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ji} c_i \right) v_j \\ &= \sum_{i=1}^{n} c_i \left(\sum_{j=1}^{m} a_{ji} v_j \right) \\ &= \sum_{i=1}^{n} c_i w_i. \end{aligned}$$

Thus $C = \{w_1, \dots, w_n\}$ is not linearly independent. □

Corollary 2.3.3. *If $\mathcal{B}$ and C are both finite bases of V, then $|\mathcal{B}| = |C|$.*

Proof. Let $|\mathcal{B}| = m$ and $|C| = n$. Because $\operatorname{span}_F(\mathcal{B}) = V$ and C is linearly independent, it must be the case that $n \leqslant m$. But since $\operatorname{span}_F(C) = V$ and $\mathcal{B}$ is also linearly independent, it must be the case that $m \leqslant n$. By antisymmetry, $n = m$. □

Definition 2.3.1. Let $\mathcal{B}$ be a basis of V. The *dimension* of V, written $\dim_F(V)$, is the cardinality of $\mathcal{B}$; i.e., $\dim_F(V) = |\mathcal{B}|$.

Theorem 2.3.4. *Let V be a finite dimensional vector space with $\dim_F(V) = n$. Let $C \subseteq V$ with $|C| = m$.*

(1) If $m > n$, then C is not linearly independent.

(2) If $m < n$, then $\operatorname{span}_F(C) \neq V$.

(3) If $m = n$, then the following are equivalent:

- *C is a basis;*
- *C is linearly independent;*
- $\operatorname{span}_F(C) = V$.

Corollary 2.3.5. *Let $W \subseteq V$ be a subspace. We have $\dim_F(W) \leqslant \dim_F(V)$. If $\dim_F(V) < \infty$, then $V = W$ if and only if $\dim_F(V) = \dim_F(W)$.*

Example 2.3.1. Let $V = \mathbf{C}$.

(1) If $F = \mathbf{C}$, then $\mathcal{B} = \{1\}$ is a basis and $\dim_{\mathbf{C}}(\mathbf{C}) = 1$.

(2) If $F = \mathbf{R}$, then $\mathcal{B} = \{1, i\}$ is a basis and $\dim_{\mathbf{R}}(\mathbf{C}) = 2$.

(3) If $F = \mathbf{Q}$, then $|\mathcal{B}| = \mathfrak{c}$ and $\dim_{\mathbf{Q}}(\mathbf{C}) = \mathfrak{c}$ (the *continuum*).

Example 2.3.2. Let $V = F[x]$ and let $f(x) \in F[x]$. We can use this polynomial to split $F[x]$ into equivalence classes analogous to how one creates the field $\mathbf{F}_p$. Define $g(x)\ h(x)$ if $f(x) \mid (g(x)-h(x))$. This is an equivalence relation. We let $[g(x)]$ denote the equivalence class containing $g(x) \in F[x]$. Let $F[x]/(f(x)) = \{[g(x)] \mid g(x) \in F[x]\}$ denote the collection of equivalence classes. Define $[g(x)]+[h(x)] = [g(x)+h(x)]$ and $\alpha[g(x)] = [\alpha g(x)]$, this makes $F[x]/(f(x))$ into a vector space.

Set $n = \deg(f(x))$. Let $\mathcal{B} = \{[1], [x], \ldots, [x^{n-1}]\}$. We will show this is a basis for $F[x]/(f(x))$. Suppose there exists $a_0, \ldots, a_{n-1} \in F$ with $a_0[1] + a_1[x] + \ldots + a_{n-1}[x^{n-1}] = [0]$. So $[a_0 + a_1 x + \ldots + a_{n-1}x^{n-1}] = [0]$, hence $f(x) \mid (a_0 + a_1 x + \ldots + a_{n-1}x^{n-1})$. But $\deg(f(x)) = n$, so we must have $a_0 = a_1 = \ldots = 0$ (linear independence).

Let $[g(x)] \in F[x]/(f(x))$. The Euclidean algorithm of polynomials gives $g(x) = f(x)q(x)+r(x)$ for some $q(x), r(x) \in F[x]/(f(x))$ with $r(x) = 0$ or $\deg(r(x)) \leqslant \deg(g(x))$. Observe that $[g(x)] = [f(x)q(x)+r(x)] = [f(x)q(x)]+[r(x)] = [r(x)]$. Since $[r(x)]$ can be written as a linear combination of basis elements from $\mathcal{B}$, we have $[g(x)] \in \operatorname{span}_F(\mathcal{B})$. Note that any element of $\operatorname{span}_F(\mathcal{B})$ is clearly contained in $F[x]/(f(x))$, establishing $\operatorname{span}_F(\mathcal{B}) = F[x]/(f(x))$.

Lemma 2.3.6. *Let V be an F-vector space and $C = \{v_i\}_{i\in I}$ be a subset of V. Then C is a basis if and only if each $v \in V$ can be written uniquely as a linear combination of elements of C.*

Proof. Suppose C is a basis. Let $v \in V$ and suppose

$$v = \sum_{i\in I} a_i v_i = \sum_{i\in I} b_i v_i,$$

for some $a_i, b_i \in F$. Observe that:

$$0_v = \sum_{i\in I}(a_i - b_i)v_i.$$

Since C is a basis, it is linearly independent, so $a_i - b_i = 0$ for all i. Thus $a_i = b_i$ for all i establishing that the expansion is unique.

Conversely, suppose every vector $v \in V$ is a unique linear combination of C. Certainly we have $\operatorname{span}_F(C) = V$. Suppose $0_v = \sum_{i\in I} a_i v_i$ for some $a_i \in F$. We also have that $0_v = \sum_{i\in I} 0 \cdot v_i$. Uniqueness gives $a_i = 0$ for all $i \in I$; i.e., C is linearly independent. □

Proposition 2.3.7. *Let V, W be F-vector spaces.*

(1) Let $T \in \operatorname{Hom}_F(V, W)$. We have that T is determined by what it does to a basis (where it maps it).

(2) Let $\mathcal{B} = \{v_i\}_{i\in I}$ be a basis of V and $C = \{w_i\}_{i\in I}$ be a subset of V. If $|\mathcal{B}| = |C|$, there is a $T \in \operatorname{Hom}_F(V, W)$ such that $T(v_i) = w_i$ for all $i \in I$.

Proof. (1) Let $v \in V$. Let $\mathcal{B} = \{v_i\}_{i \in I}$ be a basis of V and write $v = \sum_{i \in I} a_i v_i$. We have $T(v) = T(\sum_{i \in I} a_i v_i) = \sum_{i \in I} a_i T(v_i)$.

(2) Define $T : V \to W$ by $v \mapsto \sum_{i \in I} a_i w_i$. If $v = \sum_{i \in I} a_i v_i$ this map is linear (show this). □

Corollary 2.3.8. *Let $T \in \mathrm{Hom}_F(V, W)$ with $\mathcal{B} = \{v_i\}_{i \in I}$ a basis of V and $C = \{w_i = T(v_i)\}_{i \in I}$ a subset of W. We have C is a basis of W if and only if T is an isomorphism.*

Proof. Suppose C is a basis of W. Using the result from Proposition 2.3.7, define $S \in \mathrm{Hom}_F(W, V)$ with $S(w_i) = v_i$. Check $T \circ S = \mathrm{id}_W$ and $S \circ T = \mathrm{id}_V$. Thus T is an isomorphism.

Conversely, let T be an isomorphism. Let $w \in W$. As T is surjective, there exists a $v \in V$ such that $T(v) = w$. Using $\mathcal{B}$ as a basis of V, write $v = \sum_{i \in I} a_i v_i$. So observe that:

$$w = T(v) = T\left(\sum_{i \in I} a_i v_i\right) = \sum_{i \in I} a_i T(v_i) \in \mathrm{span}_F(C),$$

hence $W = \mathrm{span}_F(C)$ (note the other direction is trivial —you never need to show that). Now suppose there exists a collection of elements $a_i \in F$ with $\sum_{i \in I} a_i T(v_i) = 0_W$. Since T is linear, this is equivalent to $T(\sum_{i \in I} a_i v_i) = 0_W$, and since T is injective it must be the case that $\sum_{i \in I} a_i v_i = 0_V$. Since $\mathcal{B}$ is a basis we get $a_i = 0$ for all $i \in I$, establishing that C is linearly independent. □

Theorem 2.3.9 (Rank-Nullity Theorem)**.** *Let V be an F-vector space with $\dim_F(V) < \infty$. Then:*

$$\dim_F(V) = \dim_F(\ker(T)) + \dim_F(\mathrm{im}(T)).$$

Proof. Let $\dim_F(\ker(T)) = k$ and $\dim_F(V) = n$. Let $\mathcal{A} = \{v_1, \ldots, v_k\}$ be a basis of $\ker(T)$. Extend this to a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ of V. We'd like to show that $C = \{T(v_{k+1}), \ldots, T(v_n)\}$ is a basis of $\mathrm{im}(T)$.

Let $w \in \mathrm{im}(T)$. So there exists a $v \in V$ with $T(v) = w$. Write $v = \sum_{i=1}^{n} a_i v_i$. We have:

$$\begin{aligned} w &= T(v) \\ &= T\left(\sum_{i=1}^{n} a_i v_i\right) \\ &= \sum_{i=1}^{n} a_i T(v_i) \\ &= \sum_{i=k+1}^{n} a_i T(v_i) \in \mathrm{span}_F(C). && \text{b/c } v_1, \ldots, v_k \in \ker(T) \end{aligned}$$

Thus $\mathrm{span}_F(C) = \mathrm{im}(T)$. Now suppose we have $\sum_{i=k+1}^{n} a_i T(v_i) = 0_W$. Since T is linear we have $T(\sum_{i=1}^{n} a_i v_i) = 0_w$, which gives $\sum_{i=1}^{n} a_i v_i \in \ker(T)$. Thus we can write it in terms of the basis $\mathcal{A}$ of $\ker(T)$: there exists $a_1, \ldots, a_k$ such that

$$\sum_{i=k+1}^{n} a_i v_i = \sum_{i=1}^{k} a_i v_i,$$

which is equivalent to $\sum_{i=1}^{k} a_i v_i + \sum_{i=k+1}^{n} a_i v_i = 0_V$. However, $\mathcal{B}$ is a basis of V so $a_1 = \ldots = a_n = 0$. □

Corollary 2.3.10. *Let V, W be F-vector spaces with $\dim_F(V) = n$. Let $V_1 \subseteq V$ be a subspace with $\dim_F(V_1) = k$ and $W_1 \subseteq W$ a subspace with $\dim_F(W_1) = n-k$. Then there exists a $T \in \mathrm{Hom}_F(V, W)$ such that $\ker(T) = V_1$ and $\mathrm{im}(T) = W_1$.*

Proof. do it □

Corollary 2.3.11. *Let $T \in \mathrm{Hom}_F(V, W)$ with $\dim_F(V) = \dim_F(W) < \infty$. The following are equivalent:*

(1) T is an isomorphism.

(2) T is injective.

(3) T is surjective.

Proof. do it □

Corollary 2.3.12. *Let $A = \mathrm{Mat}_n(F)$. The following are equivalent:*

(1) A is invertible.

(2) There exists an element $B \in \mathrm{Mat}_n(F)$ such that $BA = 1_n$.

(3) There exists an element $B \in \mathrm{Mat}_n(F)$ such that $AB = 1_n$.

Proof. do it □

Corollary 2.3.13. *Let $\dim_F(V) = m$ and $\dim_F(W) = n$.*

(1) If $m < n$ and $T \in \mathrm{Hom}_F(V, W)$, then T is not surjective.

(2) If $m > n$ and $T \in \mathrm{Hom}_F(V, W)$, then T is not injective.

(3) If $m = n$ then $V \cong W$.

Example 2.3.3. This follows shortly after corollary 2.2.30 (write it down later)

2.4 Direct Sums and Quotient Spaces

Definition 2.4.1. Let V be an F-vector space and $V_1, \ldots, V_k$ be subspaces. The <u>*sum*</u> of $V_1, \ldots, V_k$ is

$$V_1 + \ldots + V_k = \{v_1 + \ldots + v_k \mid v_i \in V_i\}.$$

Proposition 2.4.1. *Let V be an F-vector space and $V_1, \ldots, V_k$ be subspaces. Then $V_1 + \ldots + V_k$ is also a subspace of V.*

Proof. do this □

Definition 2.4.2. Let $V_1, \dots, V_k$ be subspaces of V. We say $V_1, \dots, V_k$ are *independent* if whenever $v_1 + \dots + v_k = 0_V$ then $v_i = 0_V$.

Definition 2.4.3. Let $V_1, \dots, V_k$ be subspaces of V. We say V is the *direct sum* of $V_1, \dots, V_k$ and write $V = V_1 \oplus \dots \oplus V_k$ if:

(1) $V = V_1 + \dots + V_k$, and

(2) $V_1, \dots, V_k$ are independent.

Example 2.4.1.

(1) Let $V = F^2$ with $V_1 = \{(x, 0) \mid x \in F\}$ and $V_2 = \{(0, y) \mid y \in F\}$. Then

$$\begin{aligned} V_1 + V_2 &= \{(x,0) + (0,y) \mid x, y \in F\} \\ &= \{(x,y) \mid x, y \in F\} \\ &= V \end{aligned}$$

If $(x, 0) + (y, 0) = (0, 0)$, then $x = y = 0$ which means V_1 and V_2 are independent. Hence $F^2 = V_1 \oplus V_2$.

(2) Let $V = F[x]$ and $V_1 = F$, $V_2 = Fx = \{\alpha x \mid \alpha \in F\}$, and $V_3 = P_1(F)$. Note that $P_1(F) = V_1 \oplus V_2$. But V_1, V_3 are not independent because $1_F \in V_1$ and $-1_F \in V_3$ and $(-1_F) + 1_F = 0$.

(3) Let $\mathcal{B} = \{v_1, \dots, v_n\}$ be a basis of V and $\operatorname{span}_F(v_i) = V_i$. Then $V = V_1 \oplus \dots \oplus V_n$.

Lemma 2.4.2. *Let V be an F-vector space with $V_1, \dots, V_k$ as subspaces. We have $V = V_1 \oplus \dots \oplus V_k$ if and only if every $v \in V$ can be written uniquely in the form $v = v_1 + \dots + v_k$ for all $v_i \in V_i$.*

Proof. Suppose $V = V_1 \oplus \dots \oplus V_k$. Let $v \in V$. Suppose $v = v_1 + \dots + v_k = \tilde{v_1} + \dots + \tilde{v_k}$ for $v_i, \tilde{v_i} \in V_i$. Then $0_V = (v_1 - \tilde{v_1}) + \dots + (v_k - \tilde{v_k})$. Since $V_1, \dots, V_k$ are independent and $v_i - \tilde{v_i} \in V$, this gives $v_i - \tilde{v_i} = 0_V$ for all i. Thus the expansion for V is unique.

Conversely, suppose every $v \in V$ can be written uniquely in the form $v = v_1 + \dots + v_k$ with $v_i \in V_i$. Then $V = V_1 + \dots + V_k$ by definition of sums of subspaces. If $0_V = v_1 + \dots + v_k$ for some $v_i \in V_i$, and $0_v = 0_v + \dots + 0_v$, then (by uniqueness) it must be the case that $v_i = 0_V$ for all i. □

Note 1. It suffices to show that $\dim_F(V) = \dim_F(V_1) + \dots + \dim_F(V_k)$ and $V_1 \cap \dots \cap V_k = \{0_V\}$.

Exercise 2.4.1. *Let $V_1, \dots, V_k$ be subspaces of V. For each $1 \leqslant i \leqslant k$, let $\mathcal{B}_i$ be a basis of V_i. Let $\mathcal{B} = \bigcup_{i=1}^{k} \mathcal{B}_i$. Show that:*

(1) $\mathcal{B}$ spans V if and only if $V = V_1 + \dots + V_k$.

(2) $\mathcal{B}$ is linearly independent if and only if $V_1, \dots, V_k$ are independent.

(3) $\mathcal{B}$ is a basis if and only if $V = V_1 \oplus \dots \oplus V_k$.

Proof. do this shit □

Lemma 2.4.3. *Let $U \subseteq V$ be a subspace. Then U has a complement.*

Proof. do this shit □

Definition 2.4.4. Let $W \subseteq V$ be a subsapce. Define $v_1 \sim v_2$ if $v_1 - v_2 \in W$ for some $v_1, v_2 \in V$. This forms an equivalence relation. Denote the equivalence class containing v as $[v]_W = v + W = \{\tilde{v} \in V \mid v\ \tilde{v}\} = \{v + w \mid w \in W\}$. The set containing all equivalence classes over W is denoted $V/W = \{v + W \mid v \in V\}$.

Proposition 2.4.4. *Let $v_1 + W, v_2 + W \in V/W$ and $\alpha \in F$. With addition and scalar multiplication defined as follows:*

$$\begin{aligned}(v_1 + W) + (v_2 + W) &= (v_1 + v_2) + W \\ \alpha(v_1 + W) &= \alpha v_1 + W,\end{aligned}$$

it's operations are well-defined and V/W forms an F-vector space.

Proof. Let $v_1 + W = \tilde{v_1} + W$ and $v_2 + W = \tilde{v_2} + W$. Then $v_1 = \tilde{v_1} + w_1$ and $v_2 = \tilde{v_2} + w_2$ for some $w_1, w_2 \in W$. Observe that:

$$\begin{aligned}(v_1 + W) + (v_2 + W) &= (v_1 + v_2 + W) \\ &= (\tilde{v_1} + w_2 + \tilde{v_2} + w_2) + W \\ &= (\tilde{v_1} + \tilde{v_2}) + W \\ &= (\tilde{v_1} + W) + (\tilde{v_2} + W).\end{aligned}$$

$$\begin{aligned}c(v_1 + W) &= cv_1 + W \\ &= c(\tilde{v_1} + w) + W \\ &= c\tilde{v_1} + W \\ &= c(\tilde{v_1} + W).\end{aligned}$$

Hence addition and scalar multiplication are well-defined. show the vector space axioms here. □

Example 2.4.2. Let $V = \mathbf{R}^2$ and $W = \{(x, 0) \mid x \in \mathbf{R}\}$. Let $(x_0, y_0) \in V$. We have that $(x_0, y_0) \sim (x, y)$ if $(x_0, y_0) - (x, y) = (x_0 - x, y_0 - y) \in W$. So $(x_0, y_0) + W = \{(x, y_0) \mid x \in \mathbf{R}\}$. Then V/W is a vector space only when $y = 0$.

Define $\tau : \mathbf{R} \to V/W$ by $y \mapsto (0, y) + W$. This is an isomorphism. Let $y_1, y_2, c \in \mathbf{R}$. Observe that:

$$\begin{aligned}\tau(y_1 + cy_2) &= (0, y_1 + cy_2) + W \\ &= ((0, y_1) + (0, cy_2)) + W \\ &= ((0, y_1) + c(0, y_2)) + W \\ &= ((0, y_1) + W) + c((0, y_2) + W) \\ &= \tau(y_1) + c\tau(y_2).\end{aligned}$$

Hence $\tau \in \mathrm{Hom}_F(\mathbf{R}, V/W)$. Let $(x, y) + W \in V/W$. Then $(x, y) + W = (0, y) + W$. So τ is surjective because $\tau(y) = (0, y) + W$. Now let $y \in \ker(\tau)$. Then $\tau(y) = (0, y) + W = (0, 0) + W$. This implies $y = 0$, meaning the kernel is trivial and so τ is injective.

Alternatively, it is routine to show that $\tau^{-1} \in \mathrm{Hom}_F(V/W, \mathbf{R})$ with $\tau^{-1} \circ \tau = \mathrm{id}_{\mathbf{R}}$ and $\tau \circ \tau^{-1} = \mathrm{id}_{V/W}$.

Definition 2.4.5. Let $W \subseteq V$ be a subspace. The <u>*canonical projection map*</u> $\pi_W : V \to V/W$ is defined by $v \mapsto v + W$. Note that $\pi_W \in \mathrm{Hom}_F(V, V/W)$.

Note 2. To define a map $T : V/W \to V'$, you always have to check it is well-defined.

Theorem 2.4.5 (First Isomorphism Theorem)**.** *Let $T \in \mathrm{Hom}_F(V, W)$. Define $\overline{T} : V/\ker(T) \to W$ by $v + \ker(T) \mapsto T(v)$. Then $\overline{T}$ is a linear map. Moreover, $V/\ker(T) \cong \mathrm{im}(T)$.*

Proof. finish this □

2.5 Dual Spaces

Note that when one refers to something as *"canonical"*, it means the object in question does not depend on a basis.

Definition 2.5.1. Let V be an F-vector space. The <u>*dual space*</u>, denoted $V^\vee$, is defined to be $V^\vee = \mathrm{Hom}_F(V, F)$.

Theorem 2.5.1. *We have V is isomorphic to a subspace of $V^\vee$. If $\dim_F(V) < \infty$, then $V \cong V^\vee$.*

Proof. Let $\mathcal{B} = \{v_i\}_{i \in I}$ be a basis (hence this theorem is not canonical). For each $i \in I$, define:

$$v_i^\vee(v_j) = \begin{cases} 1, & i = j \\ 0, & \text{otherwise.} \end{cases}$$

We get $\{v_i^\vee\}_{i\in}$ are elements of $V^\vee$. We obtain $T \in \mathrm{Hom}_F(V, V^\vee)$ by $T(v_i) = v_i^\vee$. To show that V is isomorphic to a subspace of $V^\vee$, it is enough to show T is injective, then by the first isomorphism theorem $V \cong \mathrm{im}(T)$ (a subspace of $V^\vee$).

Let $v \in \ker(T)$, then $T(v) = 0_{V^\vee}$. Write $v = \sum_{i \in I} a_i v_i$. So:

$$\begin{aligned} 0_{V^\vee} &= T(v) \\ &= T\left(\sum_{i \in I} a_i v_i\right) \\ &= \sum_{i \in I} a_i T(v_i) \\ &= \sum_{i \in I} a_i v_i^\vee. \end{aligned}$$

Towards contradiction, pick some j with $a_j \neq 0$. Note that $0_{V^\vee} = \sum_{i \in I} a_i v_i^\vee(v_j) = a_j$ (every term except for $a_j v_j^\vee(v_j)$ equals 0). This is a contradiction, hence T is injective.

Now assume $\dim_F(V) = n$ and write $\mathcal{B} = \{v_1, \ldots, v_n\}$. Let $v^\vee \in V^\vee$. Define $a_i = v^\vee(v_i)$. Set $v = \sum_{i=1}^n a_i v_i$ and define $S : V^\vee \to V$ by $S(v^\vee) = v = \sum_{i=1}^n v^\vee(v_i) v_i$. We'd like to show that $S \in \mathrm{Hom}_F(V^\vee, V)$ and is the inverse of T. Let $v^\vee, w^\vee \in V^\vee$ and $c \in F$. Set $a_i = v^\vee(v_i)$ and $b_i = w^\vee(v_i)$. Then:

$$\begin{aligned}
S(v^\vee + cw^\vee) &= \sum_{i=1}^n \left[(v^\vee + cw^\vee)(v_i)\right] v_i \\
&= \sum_{i=1}^n \left[v^\vee(v_i) + cw^\vee(v_i)\right] v_i \\
&= \sum_{i=1}^n v^\vee(v_i) v_i + c \sum_{i=1}^n w^\vee(v_i) v_i \\
&= S(v^\vee) + cS(w^\vee).
\end{aligned}$$

Hence S is linear. Now observe that:

$$\begin{aligned}
(S \circ T)(v_j) &= S(T(v_j)) \\
&= S(v_j^\vee) \\
&= \sum_{i=1}^n v_j^\vee(v_i) v_i \\
&= v_j
\end{aligned}$$

Let $v^\vee \in V^\vee$. Note that $(T \circ S)(v^\vee)$ is a function, so it will require an input. Observe that

$$\begin{aligned}
(T \circ S)(v^\vee)(v_j) &= T(S(v^\vee))(v_j) \\
&= T(\sum_{i=1}^n v^\vee(v_i) v_i)(v_j) \\
&= \left[\sum_{i=1}^n v^\vee(v_i) T(v_i)\right](v_j) \\
&= \sum_{i=1}^n v^\vee(v_i)(v_i^\vee(v_j)) \\
&= v^\vee(v_j).
\end{aligned}$$

□

Definition 2.5.2. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of V. The <u>*dual basis*</u> for $V^\vee$ is $\mathcal{B}^\vee = \{v_1^\vee, \ldots, v_n^\vee\}$.

Proposition 2.5.2. *There is a canonical injective linear map from V to $(V^\vee)^\vee$. If $\dim_F(V) < \infty$, this is an isomorphism.*

Proof. Let $v \in V$. Define $\hat{v} : V^\vee \to F$ by $\varphi \mapsto \varphi(v)$[2]. We can easily verify that $\hat{v}$ is linear. Therefore, we have $\hat{v} \in \mathrm{Hom}_F\left(V^\vee, F\right) = (V^\vee)^\vee$. We have a map:

$$\Phi : V \to (V^\vee)^\vee \ \text{ defined by } \ v \mapsto \hat{v}.$$

We want to verify that Φ is an injective linear map. Let $v_1, v_2 \in V$ and $c \in F$. Let $\varphi \in V^\vee$, then:

$$\begin{aligned}
\Phi(v_1 + cv_2)(\varphi) &= \widehat{v_1 + cv_2}(\varphi) \\
&= \varphi(v_1 + cv_2) \\
&= \varphi(v_1) + c\varphi(v_2) \\
&= \hat{v_1}(\varphi) + c\hat{v_2}(\varphi) \\
&= \Phi(v_1)(\varphi) + c\Phi(v_2)(\varphi).
\end{aligned}$$

We will now show that Φ is injective. Let $v \in V$ and assume $v \neq 0_V$. We will form a basis $\mathcal{B}$ of V that contains v (why is this still canonical?). Let $v^\vee \in V^\vee$, then $v^\vee(v) = 1$ and $v^\vee(w) = 0$ for all $w \in \mathcal{B}$, $w \neq v$. Now assume $v \in \ker(\Phi)$. Then $\Phi(v)(\varphi) = \varphi(v) = 0$ for all $\varphi \in V^\vee$. But picking $\varphi = v^\vee$ gives:

$$\begin{aligned}
0 &= \Phi(v)(v^\vee) \\
&= v^\vee(v) \\
&= 1.
\end{aligned}$$

This is a contradiction, hence Φ is injective. $\square$

Definition 2.5.3. Let $T \in \mathrm{Hom}_F\left(V, W\right)$. We get an induced map $T^\vee : W^\vee \to V^\vee$ with $T^\vee(\varphi) = \varphi \circ T$. The following diagram commutes:

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[dr, dashed, "T^\vee(\varphi)"'] & W \arrow[d, "\varphi"] \\
& F.
\end{tikzcd}$$

[2] This can be notated as eval_v, but $\hat{v}$ appears more often in literature

3

Linear Transformations and Matrices

3.1 Choosing Coordinates

Example 3.1.1 (Choosing Coordinates)**.** Let V be an F-vector space with $\dim_F(V) < \infty$. Let $\mathcal{B} = \{v_1, ..., v_n\}$ be a basis for V. This basis fixes an isomorphism $V \cong F^n$. Let $v \in V$, write $v = \sum_{i=1}^n a_i v_i$.

$$\text{Define } T_{\mathcal{B}}(v) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in F^n.$$

This is an isomorphism. Given $v \in V$, we write $[v]_{\mathcal{B}} = T_{\mathcal{B}}(v) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$. We refer to this as *choosing coordinates* on V. asdf

Example 3.1.2.

(1) Let $V = \mathbf{Q}^2$ and $\mathcal{B} = \{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\}$. This forms a basis of V. Let $v \in V$ with $v = \begin{pmatrix} a \\ b \end{pmatrix}$. We have:

$$v = \frac{a+b}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{a-b}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \text{ hence } [v]_{\mathcal{B}} = \begin{pmatrix} \frac{a+b}{2} \\ \frac{a-b}{2} \end{pmatrix}.$$

Had we considered the standard basis $\mathcal{E}_2 = \{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$, then $[v]_{\mathcal{E}_2} = \begin{pmatrix} a \\ b \end{pmatrix}$.

(2) Let $V = P_2(\mathbf{R})$. Let $C = \{1, (x-1), (x-1)^2\}$. This forms a basis of V. Let $f(x) = a + bx + cx^2 \in P_2(\mathbf{R})$. Written in terms of C, we have $f(x) = (a+b+c) + (b+2c)(x-1) + c(x-1)^2$.

$$\text{Thus } [f(x)]_C = \begin{pmatrix} a+b+c \\ b+2c \\ c \end{pmatrix}$$

Example 3.1.3 (Linear Transformations as Matrices)**.** Recall that given a matrix $A \in \mathrm{Mat}_{m,n}(F)$, we obtain a linear map $T_A \in \mathrm{Hom}_F(F^n, F^m)$ by $T_A(v) = Av$. This process "works in reverse" —given a linear transformation $T \in \mathrm{Hom}_F(F^n, F^m)$, there is a matrix A so that $T = T_A$.

Let $\mathcal{E}_n = \{e_1, ..., e_n\}$ be the standard basis of F^n and $\mathcal{F}_m = \{f_1, ..., f_m\}$ be the standard basis of F^m. We have that $T(e_j) \in F^m$ for each j, meaning we have elements $a_{ij} \in F$ with $T(e_j) = \sum_{i=1}^m a_{ij} f_i$. Define $A = (a_{ij}) \in Mat_{m,n}(F)$. Observe that:

$$T_A(e_j) = Ae_j = \sum_{i=1}^m a_{ij} f_i = a_{1j} f_1 + ... + a_{mj} f_m.$$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ \vdots & \ddots & & & \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ a_{m1} & a_{m2} & \dots & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1_j \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

Working "in reverse", let $T \in \mathrm{Hom}_F(V, W)$ with $\mathcal{B} = \{v_1, \dots, v_n\}$ a basis for V and $C = \{w_1, \dots, w_m\}$ a basis for W. Define:

$$P = T_{\mathcal{B}} : V \to F^n \text{ by } v \mapsto [v]_{\mathcal{B}}$$
$$Q = T_C : W \to F^m \text{ by } w \mapsto [w]_C$$

From the following diagram:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "P"'] & W \arrow[d, "Q"] \\ F^n \arrow[r, dashed, "Q\circ T\circ P^{-1}"'] & F^m \end{tikzcd}$$

we have that $Q \circ T \circ P^{-1}$ corresponds to a matrix $A \in \mathrm{Mat}_{m,n}(F)$. Write $[T]_{\mathcal{B}}^{C} = A$, this is the unique matrix that satisfies $[T]_{\mathcal{B}}^{C}[v]_{\mathcal{B}} = [T(v)]_C$. Given $T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$, observe that:

$$[T]_{\mathcal{B}}^{C} v_j = [T(v_j)]_C = \left[\sum_{i=1}^{m} a_{ij} w_i\right]_C = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

So $[T]_{\mathcal{B}}^{C} v_j$ corresponds to the j^{th} column of the matrix $[T]_{\mathcal{B}}^{C}$ Thus we have:

$$[T]_{\mathcal{B}}^{C} = \Big([T(v_1)]_C \quad | \quad \dots \quad | \quad [T(v_n)]_C \Big)$$

Example 3.1.4.

(1) Let $V = P_3(\mathbf{R})$ with $\mathcal{B} = \{1, x, x^2, x^3\}$. Define $T \in \mathrm{Hom}_{\mathbf{R}}(V, V)$ by $T(f(x)) = f'(x)$. Following Example 3.1.3 gives:

$$T(1) = 0 = 0 \cdot 1 + 0 \cdots x + 0 \cdot x^2 + 0 \cdot x^3$$
$$T(x) = 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3$$
$$T(x^2) = 2x = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 + 0 \cdot x^3$$
$$T(x^3) = 3x^2 = 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2 + 0 \cdot x^3$$

$$[T(1)]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$
$$[T(x)]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$
$$\left[T(x^2)\right]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$$
$$\left[T(x^3)\right]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 0 \\ 3 \\ 0 \end{pmatrix}$$

$$[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(2) Let $V = P_3(\mathbf{R})$ with $\mathcal{B} = \left\{1, x, x^2, x^3\right\}$ with $C = \left\{1, (1-x), (1-x)^2, (1-x^3)\right\}$. Then

$$\begin{aligned} T(1) &= 0 \\ T(x) &= 1 \\ T(x^2) &= 2 + 2(x-1) \\ T(x^3) &= -9 - 6(x-1) + 3(x-1)^2 \end{aligned}$$

$$[T(1)]_C = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$
$$[T(x)]_C = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$
$$\left[T(x^2)\right]_C = \begin{pmatrix} 2 \\ 2 \\ 0 \\ 0 \end{pmatrix}$$
$$\left[T(x^3)\right]_C = \begin{pmatrix} -9 \\ -6 \\ 3 \\ 0 \end{pmatrix}$$

$$[T]_{\mathcal{B}}^{C} = \begin{pmatrix} 0 & 1 & 2 & -9 \\ 0 & 0 & 2 & -6 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Exercise 3.1.1.

(1) *Let $\mathcal{A}$ be a basis of U, $\mathcal{B}$ a basis of V and C a basis of W. Let $S \in \mathrm{Hom}_F(U,V)$ and $T \in \mathrm{Hom}_F(V,W)$. Show*
$$[T \circ S]_{\mathcal{A}}^{C} = [T]_{\mathcal{B}}^{C}\,[S]_{\mathcal{A}}^{\mathcal{B}}.$$

(2) *Given $A \in \mathrm{Mat}_{m,k}(F)$ and $B \in \mathrm{Mat}_{n,m}(F)$, we have corresponding linear maps T_A and T_B. Show that you can recover the definition of matrix multiplication by using part (1).*

Note 3. Instead of $[T]_{\mathcal{B}}^{\mathcal{B}}$ we will write $[T]_{\mathcal{B}}$.

Example 3.1.5 (Change of Basis)**.** Let V be an F-vector space and $\mathcal{B}, \mathcal{B}'$ bases of V. Given V expressed in terms of $\mathcal{B}$, we'd like to express it in terms of $\mathcal{B}'$ (or vice versa).

Let $\mathcal{B} = \{v_1, \dots, v_n\}$ and $\mathcal{B}' = \{v_1', \dots, v_n'\}$. Define:
$$T : V \to F^n \text{ by } v \mapsto [v]_{\mathcal{B}}$$
$$S : V \to F^n \text{ by } v \mapsto [v]_{\mathcal{B}'}.$$

We obtain a diagram similar to Example 3.1.3:

$$\begin{tikzcd} V \arrow[r, "\mathrm{id}_V"] \arrow[d, "T"'] & V \arrow[d, "S"] \\ F^n \arrow[r, "S\circ \mathrm{id}_V \circ T^{-1}"'] & F^n \end{tikzcd}$$

Hence the change of basis matrix is $[\mathrm{id}_V]_{\mathcal{B}}^{\mathcal{B}'}$

Exercise 3.1.2. *Let $\mathcal{B} = \{v_1, \dots, v_n\}$. Show that* $[\mathrm{id}_V]_{\mathcal{B}}^{\mathcal{B}'} = \left([v_1]_{\mathcal{B}'} \mid \dots \mid [v_n]_{\mathcal{B}'}\right)$.

Example 3.1.6.

(1) Let $V = \mathbf{Q}^2$ with $\mathcal{B} = \{e_1 = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), e_2 = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\}$ and $\mathcal{B}' = \{v_1 = \left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right), v_2 = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)\}$. Observe that:
$$e_1 = \frac{1}{2}v_1 + \frac{1}{2}v_2$$
$$e_2 = -\frac{1}{2}v_1 + \frac{1}{2}v_2$$

$$[e_1]_{\mathcal{B}} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$
$$[e_2]_{\mathcal{B}} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

$$[\mathrm{id}_V]_{\mathcal{E}_2}^{\mathcal{B}'} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Consider $v = \left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right) \in \mathbf{Q}^2$. We can express v in terms of $\mathcal{B}'$ by doing the following calculation:

$$\begin{aligned}
[\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{E}_2} [v_2]_{\mathcal{E}_2} &= \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} \\
&= \begin{pmatrix} -\frac{1}{2} \\ \frac{5}{2} \end{pmatrix} \\
&= [v]_{\mathcal{B}'}.
\end{aligned}$$

(2) Let $V = P_2(\mathbf{R})$ with $\mathcal{B} = \{1, x, x^2\}$ and $\mathcal{B}' = \{1, (x-2), (x-2)^2\}$. Then:

$$\begin{aligned}
1 &= 1 \cdot 1 + 0 \cdot (x-2) + 0 \cdot (x-2)^2 \\
x &= 2 \cdot 1 + 1 \cdot (x-2) + 0 \cdot (x-2)^2 \\
x^2 &= 4 \cdot 1 + 4 \cdot (x-2) + 1 \cdot (x-2)^2
\end{aligned}$$

$$\begin{aligned}
[1]_{\mathcal{B}'} &= \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) \\
[x]_{\mathcal{B}'} &= \left(\begin{smallmatrix} 2 \\ 1 \\ 0 \end{smallmatrix}\right) \\
[x^2]_{\mathcal{B}'} &= \left(\begin{smallmatrix} 4 \\ 4 \\ 1 \end{smallmatrix}\right)
\end{aligned}$$

$$[\mathrm{id}_V]^{\mathcal{B}'}_{\mathcal{B}} = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}.$$

Example 3.1.7 (Similar Matrices)**.** Let $A, B \in \mathrm{Mat}_n(F)$. Let $\mathcal{E}_n$ be the standard basis for F^n and $T_A \in \mathrm{Hom}_F(F^n, F^n)$ such that $A = [T_A]_{\mathcal{E}_n}$. We can relate A in terms of an arbitrary basis $\mathcal{B}$ as follows:

$$\begin{tikzcd}
F^n \arrow[r, "T_A"] \arrow[d, "T_{\mathcal{B}}"'] & F^n \arrow[d, "T_{\mathcal{B}}"] \\
F^n \arrow[r, "{[T_A]_{\mathcal{B}}}"'] & F^n.
\end{tikzcd}$$

But by extending our diagram using our change of basis algorithm, we obtain the following:

$$\begin{tikzcd}
F^n \arrow[r, "\mathrm{id}_{F^n}"] \arrow[d, "T_{\mathcal{B}}"'] & F^n \arrow[r, "T_A"] \arrow[d, "T_{\mathcal{E}_n}"'] & F^n \arrow[r, "\mathrm{id}_{F^n}"] \arrow[d, "T_{\mathcal{E}_n}"] & F^n \arrow[d, "T_{\mathcal{B}}"] \\
F^n \arrow[r, "{[\mathrm{id}_{F^n}]^{\mathcal{E}_n}_{\mathcal{B}}}"'] & F^n \arrow[r, "{[T_A]_{\mathcal{E}_n}}"'] & F^n \arrow[r, "{[\mathrm{id}_{F^n}]^{\mathcal{B}}_{\mathcal{E}_n}}"'] & F^n
\end{tikzcd}$$

So $[T_A]_{\mathcal{B}} = [\mathrm{id}_{F^n}]_{\mathcal{B}}^{\mathcal{E}_n} [T_A]_{\mathcal{E}_n} [\mathrm{id}_{F^n}]_{\mathcal{E}_n}^{\mathcal{B}}$. Assigning $P^{-1} = [\mathrm{id}_{F^n}]_{\mathcal{B}}^{\mathcal{E}_n}$ and $P = [\mathrm{id}_{F^n}]_{\mathcal{E}_n}^{\mathcal{B}}$ yields the familiar equation $[T_A]_{\mathcal{B}} = P^{-1}AP$; i.e., $A = P[T_A]_{\mathcal{B}}P^{-1}$. In particular, the matrix $A = [T_A]_{\mathcal{E}_n}$ is similar to $[T_A]_{\mathcal{B}}$ for any basis $\mathcal{B}$.

Example 3.1.8. Let $A = \begin{pmatrix} 1 & 3 & -5 \\ -2 & -1 & 6 \\ 3 & 2 & 1 \end{pmatrix}$. Let $\mathcal{E}_3 = \{e_1, e_2, e_3\}$ be the standard basis of F^3. We have:

$$\begin{aligned} T_A(e_1) &= e_1 - 2e_2 + 3e_3 \\ T_A(e_2) &= 3e_1 - e_2 + 2e_3 \\ T_A(e_3) &= 3e_1 + 2e_2 + e_3. \end{aligned}$$

Now consider $\mathcal{B} = \{v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}\}$. One can check this is indeed a basis. Observe that:

$$\begin{aligned} e_1 &= -2v_1 + -3v_2 + v_3 \\ e_2 &= 3v_1 + 3v_2 - v_3 \\ e_3 &= -2v_1 - 2v_2 + v_3. \end{aligned}$$

So the change of basis matrix from $\mathcal{E}_3$ to $\mathcal{B}$ is given by $P = [\mathrm{id}_{F^3}]_{\mathcal{E}_3}^{\mathcal{B}} = \begin{pmatrix} -2 & 3 & -2 \\ -3 & 3 & -2 \\ 1 & -1 & 1 \end{pmatrix}$. We have $P^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix}$. Thus A is similar to the matrix $B = P^{-1}AP = \begin{pmatrix} -29 & 32 & -25 \\ -38 & 45 & -31 \\ -20 & 27 & -15 \end{pmatrix}$.

3.2 Row Operations

Definition 3.2.1. Let $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(F)$. We say a_{kl} is a *pivot* of A if $a_{kl} \neq 0$ and $a_{ij} = 0$ if $i > k$ or $j < l$.

Example 3.2.1. Let $A = \begin{pmatrix} 2 & 1 & 4 & 5 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Then 2, 1, and 5 are pivots.

Definition 3.2.2. Let $A \in \mathrm{Mat}_{m,n}(F)$. We say A is in *row echelon form* if all its nonzero rows have a pivot and all its zero rows are located below nonzero rows. We say it is *reduced row echelon form* if it is in row echelon form and all of its pivots are 1 and the only nonzero elements in the columns containing pivots.

Example 3.2.2. From the previous example, expressing $A = \begin{pmatrix} 2 & 1 & 4 & 5 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ in reduced row echelon form yields $A' = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Example 3.2.3. Let $A = \begin{pmatrix} 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 \\ 1 & 1 & 2 & 3 \end{pmatrix}$. Then $T_A : F^4 \to F^4$. Let $\mathcal{B}_4 = \{e_1, e_2, e_3, e_4\}$ and $\mathcal{F}_3 =$

$\{f_1, f_2, f_3\}$. So $A = [T_A]_{\mathcal{B}_3}^{\mathcal{F}_3}$. We have the following set of equations:

$$\begin{aligned} T_A(e_1) &= 3f_1 + f_2 + f_3 \\ T_A(e_2) &= 4f_1 + 2f_2 + f_3 \\ T_A(e_3) &= 5f_1 + 3f_2 + 2f_3 \\ T_A(e_4) &= 6f_1 + 4f_2 + 3f_3. \end{aligned}$$

We are going to perform row operations of A by making substitutions to its basis elements. Consider the operation $R_1 \leftrightarrow R_3$.

$$\mathcal{F}_3^{(2)} = \{f_1^{(2)} = f_3, f_2^{(2)} = f_2, f_3^{(2)} = f_1\}.$$

$$\begin{aligned} T_A(e_1) &= f_1^{(2)} + f_2^{(2)} + 3f_3^{(2)} \\ T_A(e_2) &= f_1^{(2)} + 2f_2^{(2)} + 4f_3^{(2)} \\ T_A(e_3) &= 2f_1^{(2)} + 3f_2^{(2)} + 5f_3^{(2)} \\ T_A(e_4) &= 3f_1^{(2)} + 4f_2^{(2)} + 6f_3^{(2)}. \end{aligned}$$

So $[T_A]_{\mathcal{B}_3}^{\mathcal{F}_3^{(2)}} = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 \\ 3 & 4 & 5 & 6 \end{pmatrix}$. Now consider the row operation $-R_1 + R_2 \leftrightarrow R_2$.

$$\mathcal{F}_3^{(3)} = \{f_1^{(3)} = f_1^{(2)} + f_2^{(2)}, f_2^{(3)} = f_2^{(2)}, f_3^{(3)} = f_3^{(2)}\}.$$

$$\begin{aligned} T_A(e_1) &= f_1^{(2)} + f_2^{(2)} + 3f_3^{(2)} \\ &= f_1^{(3)} + 3f_3^{(3)}. \end{aligned}$$

$$\begin{aligned} T_A(e_2) &= f_1^{(2)} + 2f_2^{(2)} + 4f_3^{(2)} \\ &= f_1^{(2)} + f_2^{(2)} + f_2^{(2)} + 4f_3^{(2)} \\ &= f_1^{(3)} + f_2^{(3)} + 4f_3^{(3)}. \end{aligned}$$

$$\begin{aligned} T_A(e_3) &= \ldots \\ T_A(e_4) &= \ldots \end{aligned}$$

So $[T_A]_{\mathcal{B}_3}^{\mathcal{F}_3^{(3)}} = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 1 \\ 3 & 4 & 5 & 6 \end{pmatrix}$. Now consider the row operation $-3R_1 + R_3 \leftrightarrow R_3$.

$$\mathcal{F}_3^{(4)} = \{f_1^{(4)} = f_1^{(3)} + 3f_3^{(3)}, f_2^{(4)} = f_2^{(3)}, f_3^{(4)} = f_3^{(3)}\}.$$

$$\begin{aligned} T_A(e_1) &= f_1^{(3)} + 3f_3^{(3)} \\ &= f_1^{(4)} \end{aligned}$$

$$\begin{aligned} T_A(e_2) &= \dots \\ T_A(e_3) &= \dots \\ T_A(e_4) &= \dots \end{aligned}$$

The rest of the steps to convert A to reduced row echelon form follow similarly.

Theorem 3.2.1. *Let $A \in \mathrm{Mat}_{m,n}(F)$. The matrix A can be put in row echelon form through a series of row operations of the form:*

(1) $R_i \leftrightarrow R_j$

(2) $R_i \leftrightarrow cR_i$

(3) $cR_i + R_J \leftrightarrow R_j$.

Example 3.2.4. Instead of directly changing the basis of a matrix, we can use linear maps to perform row operations. Let $C = \{w_1, \dots, w_n\}$ be a basis of W.

(1) Define $T_{i,j} : W \longrightarrow W$ by

$$\begin{aligned} T_{i,j}(w_k) &= w_k \text{ if } k \neq i, j, \\ T_{i,j}(w_i) &= w_j, \\ T_{i,j}(w_j) &= w_i. \end{aligned}$$

Then $E_{i,j} = \left[T_{i,j}\right]_C^C$ corresponds to the identity matrix except the i^{th} and j^{th} rows are switched.

(2) Let $c \in F$, $c \neq 0$. Define $T_i^{(c)} : W \longrightarrow W$ by:

$$\begin{aligned} T_i^{(c)}(w_j) &= w_j \text{ if } j \neq i, \\ T_i^{(c)}(w_i) &= cw_i \end{aligned}$$

Then $E_i^{(c)} = \left[T_i^{(c)}\right]_C^C$ corresponds to the identity matrix with the i^{th} row multiplied by c.

(3) Define $T_{i,j}^{(c)} : W \longrightarrow W$ by:

$$\begin{aligned} T_{i,j}^{(c)}(w_k) &= w_k \text{ if } k \neq j, \\ T_{i,j}^{(c)}(w_j) &= w_j + cw_i \end{aligned}$$

Then $E_{i,j}^{(c)} = \left[T_{i,j}^{(c)}\right]_C^C$ corresponds to the identity matrix with the what does this mean?

Now let $T_A : F^4 \to F^3$ with $A = \left(\begin{smallmatrix} 3&4&5&6\\ 1&2&3&4\\ 1&1&2&3 \end{smallmatrix}\right)$ and $\mathcal{E}_4$ and $\mathcal{F}_3$ their respective standard bases. Performing the row operation $R_1 \leftrightarrow R_3$ using the above method yields:

$$\begin{aligned}(T_{1,3} \circ T_A)(e_1) &= T_{1,3}(3f_1 + f_2 + f_3)\\ &= 3T_{1,3}(f_1) + T_{1,3}(f_2) + T_{1,3}(f_3)\\ &= 3f_3 + f_2 + f1\end{aligned}$$

$$\begin{aligned}\left[T_{1,3} \circ T_{A}{}_{\mathcal{E}_4}^{\mathcal{F}_3}\right] &= \left[T_{1,3}\right]_{\mathcal{F}_3}^{\mathcal{F}_3} \left[T_A\right]_{\mathcal{E}_4}^{\mathcal{F}_3}\\ &= E_{1,3}A\\ &= \begin{pmatrix} 1 & 1 & 2 & 3\\ 1 & 2 & 3 & 4\\ 3 & 4 & 5 & 6 \end{pmatrix}.\end{aligned}$$

The rest of the row operations follow similarly. The reduced-row echelon form of A can then be expressed as:

$$\left[T_{1,3}^{(-1)} \circ T_{2,3}^{(-1)} \circ T_{(3)}^{(\frac{1}{2})} \circ T_{3,2}^{(-1)} \circ T_{3,1}^{(-3)} \circ T_{1,2}^{(-1)} \circ T_{1,3} \circ T_A\right]_{\mathcal{E}_4}^{\mathcal{F}_3}.$$

3.3 Column-space and Null-space

Definition 3.3.1. Let $A \in \mathrm{Mat}_{m,n}(F)$.

(1) The <u>*column-space*</u> of A is the F-span of the column vectors, denoted as $CS(A)$.

(2) The <u>*null-space*</u> of A is the F-span of vectors $v \in F^n$ such that $Av = 0_V$, denoted as $NS(A)$.

(3) The <u>*rank*</u> of A is $\operatorname{rank} A = \dim_F CS(A)$.

Example 3.3.1. Let $T_A \in \mathrm{Hom}_F(F^n, F^m)$ where $\mathcal{E}_n = \{e_1, \ldots, e_n\}$ is the standard basis of F^n and $\mathcal{F}_n = \{f_1, \ldots, f_m\}$ is the standard basis of F^m. Since

$$[T_A]_{\mathcal{E}_n}^{\mathcal{F}_m} = A = \Big(T_A(e_1) \mid \ \ldots \ \mid T_A(e_n)\Big),$$

we have that $CS(A) = \mathrm{im}\,(T_A)$, so $\operatorname{rank} A = \dim_F \mathrm{im}\,(T_A)$. Recall from an introductory linear algebra course that the column space is calculated by:

(a) Put A into row echelon form,

(b) Look at which columns have pivots,

(c) The same columns in A are then a basis of $CS(A)$.

Why does this work? There exists an isomorphism $E : F^n \to F^m$ so that $[E \circ T_A]_{\mathcal{E}_n}^{\mathcal{F}_m} = [E]_{\mathcal{E}_n}^{\mathcal{F}_m} A$ is in row echelon form. The column space of $[E \circ T_A]_{\mathcal{E}_n}^{\mathcal{F}_m}$ has as its basis the columns containing pivots (denoted $e_{i1}, \ldots, e_{ik}$):

$$\underbrace{[E \circ T_A(e_{i1})]_{\mathcal{F}_m}, \ \ldots \ , [E \circ T_A(e_{ik})]_{\mathcal{F}_m}}_{\text{this is a basis of } CS([E \circ T_A]_{\mathcal{E}_n}^{\mathcal{F}_m})}$$

Since E is an isomorphism, there is an inverse $E^{-1} : F^m \to F^m$ with:

$$\begin{aligned} E^{-1}(w_1) &= [E \circ T_A(e_{i1})]_{\mathcal{F}_m} \\ &\vdots \\ E^{-1}(w_k) &= [E \circ T_A(e_{ik})]_{\mathcal{F}_m} . \end{aligned}$$

These are linearly independent since E^{-1} is an isomorphism. If there is a vector $v \in CS(A)$ with $v \notin \operatorname{span}_F \left([E \circ T_A(e_{i1})]_{\mathcal{F}_m}, \ldots, [E \circ T_A(e_{ik})]_{\mathcal{F}_m}\right)$, then $E(v)$ cannot be in $\operatorname{span}_F (w_1, \ldots, w_k)$. So the columns
$[E \circ T_A(e_{i1})]_{\mathcal{F}_m}, \ldots, [E \circ T_A(e_{ik})]_{\mathcal{F}_m}$ give a basis for the column space of A.

Example 3.3.2. Let $A = \begin{pmatrix} 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 \\ 1 & 1 & 2 & 3 \end{pmatrix}$. Rewritten in row echelon form is $A' = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -2 & -4 \end{pmatrix}$. Thus:

$$\begin{aligned} CS(B) &= \operatorname{span}_F \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \right) \\ CS(A) &= \operatorname{span}_F \left(\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ 3 \\ 2 \end{pmatrix} \right) .e \end{aligned}$$

Example 3.3.3. We have $v \in NS(A)$ if and only if $Av = 0_{F^m} = T_A(v)$. Note that $T_A(v) = 0_{F^m}$ if and only if $v \in \ker(T_A)$, hence $NS(A) = \ker(T_A)$. In an introductory algebra class, the null space of a matrix A is calculated by:

(1) Putting A into reduced row echelon form,

(2) Solving the equation $A'x = 0_{F^n}$.

This works because given a map $T_A : F^n \to F^m$, row operations change the basis of the codomain, not the domain. So $NS(A) = NS(A')$.

Example 3.3.4. Let $A = \begin{pmatrix} 4 & -4 & 2 \\ -4 & 4 & -2 \\ 2 & -1 & 1 \end{pmatrix}$. The reduce row echelon form of A is $A' = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Solving the equation:

$$\begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

gives $x_2 = 0$ and $x_1 = -\frac{1}{2}x_3$. Hence $NS(A) = \operatorname{span}_F \begin{pmatrix} -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix}$.

3.4 The Transpose of a Matrix

Definition 3.4.1. Let $A \in \mathrm{Mat}_{m,n}(F)$ with $\mathcal{E}_n = \{e_1, ..., e_n\}$ and $\mathcal{F}_m = \{f_1, ..., f_m\}$ as standard bases. Then $A = [T_A]_{\mathcal{E}^n}^{\mathcal{F}_m}$, and furthermore $T_A \in \mathrm{Hom}_F(F^n, F^m)$ induces a dual map $T_A^\vee \in \mathrm{Hom}_F(F^{m\vee}, F^{n\vee})$. The <u>*transpose*</u> of A is defined as:

$$A^t = \left[T_A^\vee\right]_{\mathcal{F}_m^\vee}^{\mathcal{E}_n^\vee}.$$

Lemma 3.4.1. *Let $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(F)$. Then $A^t = (b_{ij}) \in \mathrm{Mat}_n, m(F)$ with $b_{ij} = a_{ji}$.*

Proof. We use the same setup as Definition 3.4.1. We have:

$$T_A(e_i) = \sum_{k=1}^{m} a_{ki} f_k$$
$$T_A^\vee(f_j^\vee) = \sum_{k=1}^{n} b_{kj} e_k^\vee.$$

Applying $f_j^\vee$ to $T_A(e_i)$ yields[1]:

$$\begin{aligned}(f_j^\vee \circ T_A)(e_i) &= f_j^\vee\left(\sum_{k=1}^{m} a_{ki} f_k\right) \\ &= \sum_{k=1}^{m} a_{ki} f_j^\vee(f_k) \\ &= a_{ji}.\end{aligned}$$

Evaluating the $T_A^\vee(f_j^\vee)$ at e_i gives:

$$\begin{aligned}T_A^\vee(f_j^\vee)(e_i) &= \sum_{k=1}^{n} b_{kj} e_k^\vee(e_i) \\ &= b_{ij}.\end{aligned}$$

By Definition 2.5.3, we have $(f_j^\vee \circ T_A)(e_i) = T_A^\vee(f_j^\vee)(e_i)$. Hence $a_{ji} = b_{ij}$ □

Exercise 3.4.1. *Let $A_1, A_2 \in \mathrm{Mat}_{m,n}(F)$ and $c \in F$. Show that:*

$$(A_1 + A_2)^t = A_1^t + A_2^t$$
$$(cA_1)^t = cA_1^t.$$

Lemma 3.4.2. *Let $A \in \mathrm{Mat}_{m,n}(F)$ and $B \in \mathrm{Mat}_{p,m}(F)$. Then $(BA)^t = A^t B^t$.*

[1] I was really confused about this. In short, given a $T \in \mathrm{Hom}_F(V, V)$ and basis $\mathcal{B}$ we have a matrix representation $[T]_{\mathcal{B}}$. It is natural to wonder what, $[T^\vee]_{\mathcal{B}^\vee}$ looks like, and it turns out to be the "transpose" we were familiar with from 214. Basically, applying $f_j^\vee$ to $T_A(e_i)$ gives us coefficients (by definition of dual basis elements) which correspond to a particular column vector of $[T_A]_{\mathcal{B}}$. Likewise, since we have that fancy property from Definition 2.5.3, naturally we should evaluate $T_A^\vee(f_j^\vee)$ at e_i, which gives us coefficients which correspond to column vectors of $[T_A^\vee]_{\mathcal{B}^\vee}$. The rest is self-explanatory.

Proof. Let $\mathcal{E}_m$, $\mathcal{E}_n$, and $\mathcal{E}_p$ be standard bases with $[T_A]_{\mathcal{E}_n}^{\mathcal{E}_m} = A$ and $[T_B]_{\mathcal{E}_m}^{\mathcal{E}_p} = B$. Then $BA = [T_B \circ T_A]_{\mathcal{E}_n}^{\mathcal{E}_p}$. Thus:

$$\begin{aligned}(BA)^t &= \left[(T_B \circ T_A)^\vee\right]_{\mathcal{E}_p^\vee}^{\mathcal{E}_n^\vee} \\ &= \left[T_A^\vee \circ T_B^\vee\right]_{\mathcal{E}_p^\vee}^{\mathcal{E}_n^\vee} \\ &= \left[T_A^\vee\right]_{\mathcal{E}_m^\vee}^{\mathcal{E}_n^\vee}\left[T_B^\vee\right]_{\mathcal{E}_p^\vee}^{\mathcal{E}_m^\vee} \\ &= A^t B^t.\end{aligned}$$

□

Lemma 3.4.3. *Let* $A \in \mathrm{GL}_n(F)$*. Then* $(A^{-1})^t = (A^t)^{-1}$.

Proof. Let $A = [T_A]_{\mathcal{E}_n}^{\mathcal{E}_n}$. Then $A^{-1} = \left[T_A^{-1}\right]_{\mathcal{E}_n}^{\mathcal{E}_n}$. We have:

$$\begin{aligned}1_n &= \left[\mathrm{id}_{F^n}^\vee\right]_{\mathcal{E}_n^\vee}^{\mathcal{E}_n^\vee} \\ &= \left[(T_A^{-1} \circ T_A)^\vee\right]_{\mathcal{E}_n^\vee}^{\mathcal{E}_n^\vee} \\ &= \left[T_A^\vee \circ (T_A^{-1})^\vee\right]_{\mathcal{E}_n^\vee}^{\mathcal{E}_n^\vee} \\ &= \left[T_A^\vee\right]_{\mathcal{E}_n^\vee}^{\mathcal{E}_n^\vee}\left[(T_A^{-1})^\vee\right]_{\mathcal{E}_n^\vee}^{\mathcal{E}_n^\vee} \\ &= A^t(A^{-1})^t.\end{aligned}$$

By the uniqueness of inverses, we must have that $(A^{-1})^t = (A^t)^{-1}$ Showing left invertibility follows identically. □

4

Generalized Eigenvectors and Jordan Canonical Form

4.1 Diagonalization

Recall. We say $A \sim B$ if and only if $A = PBP^{-1}$ for some $P \in \mathrm{GL}_n(F)$. In particular, this means $A = [T]_{\mathcal{A}}$ and $B = [T]_{\mathcal{B}}$ for some bases $\mathcal{A}$ and $\mathcal{B}$ (Example 3.1.7).

Definition 4.1.1. We say A is *diagonalizable* if $A \sim D$ for some diagonal matrix D. In terms of linear transformations, $A = [T]_{\mathcal{A}}$ is diagonalizable if there is a basis $\mathcal{B}$ such that $[T]_{\mathcal{B}} = D$.

Example 4.1.1. If $A \sim B$ then A is diagonalizable if and only if B is diagonalizable. If A and B are diagonalizable, they must be similar to the same diagonal matrix up to reordering the diagonals.

Example 4.1.2. Let $V = F^2$ and $T \in \mathrm{Hom}_F(V, V)$. Let $T(e_1) = 3e_1$ and $T(e_2) = -2e_2$. We have that:

$$[T]_{\mathcal{E}_2} = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}.$$

It follows that $V = V_1 \oplus V_2$, where $V_1 = \mathrm{span}_F(e_1)$ and $V_2 = \mathrm{span}_F(e_2)$. In this case, we have that $T(V_1) \subseteq V_1$ and $T(V_2) \subseteq V_2$, allowing us to write T as a diagonal matrix.

Example 4.1.3. Let $V = F^2$ and $T \in \mathrm{Hom}_F(V, V)$. Consider $T(e_1) = 3e_1$ and $T(e_2) = e_1 + 3e_2$. Then:

$$[T]_{\mathcal{E}_2} = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}.$$

Then $V = V_1 \oplus V_2$ with $V_1 = \mathrm{span}_F(e_1)$ and $V_2 = \mathrm{span}_F(e_2)$. But while we have $T(V_1) \subseteq V_1$, we do not have $T(V_2) \subseteq V_2$.

Suppose towards contradiction we have $W_1, W_2 \neq \{0\}$ with $T(W_1) \subseteq W_1$ and $T(W_2) \subseteq W_2$. Write $W_i = \mathrm{span}_F(w_i)$. In particular, this means we can write $T(w_1) = \alpha w_1$ and $T(w_2) = \beta w_2$. For $\mathcal{B} = \{w_1, w_2\}$, we have:

$$[T]_{\mathcal{B}} = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}.$$

Write $w_1 = ae_1 + be_2$ and $w_2 = ce_1 + de_2$. Then:

$$\begin{aligned} \alpha w_1 &= T(w_1) \\ &= aT(e_1) + bT(e_2) \\ &= a(3e_1) + b(e_1 + 3e_2) \\ &= (3a + b)e_1 + (3b)e_2. \end{aligned}$$

Thus, $\alpha(ae_1 + be_2) = (3a + b)e_1 + (3b)e_2$, meaning $\alpha a = 3b + b$ and $\alpha b = 3b$. Either $b = 0$ or $\alpha = 3$. It must be the case that $\alpha = 3$, hence $T(w_1) = 3w_1$. A similar argument for w_1 gives:

$$\begin{aligned} \beta w_2 &= T(w_2) \\ &= \dots \\ &= (3c + d)e_1 + (3d)e_2. \end{aligned}$$

This implies $\beta c = ec + d$ and $\beta d = 3d$. If $\beta = 3$, then this contradicts the first equation. If $w_2 = ce_1$, this contradicts w_1, w_2 being a basis.

Example 4.1.4. Let $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$. Let $F = \mathbf{Q}$. Let $P \in GL_2(\mathbf{Q})$, where $P = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. We have:

$$P^{-1}AP = \frac{1}{ad - bc}\begin{pmatrix} ad - 2ab + 2cd - 4bc & -3bd - 3b^2 + 2d^2 \\ 3ac + 3a^2 - 2c^2 & -bc + 3ab - 2cd + 4ad \end{pmatrix}.$$

We must have that $3a^2 + 4ac - 2c^2 = 0$. If $c = 0$, then $a = 0$, which contradicts P being invertible. So $c \neq 0$, meaning we can divide by c^2 and set $x = \frac{a}{c}$. Then the roots of $3x^2 + 3x - 2 = 0$ are:

$$x = \frac{-3 \pm \sqrt{33}}{6},$$

which gives:

$$a = \frac{-3 \pm \sqrt{33}}{6}c.$$

Since $c \neq 0$, $a \notin \mathbf{Q}$. Thus we cannot diagonalize A over $\mathbf{Q}$. But if we were to take $F = \mathbf{Q}(\sqrt{33})$, then we have that:

$$\mathcal{B} = \{v_1 = \begin{pmatrix} 1 \\ \frac{3+\sqrt{33}}{4} \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ \frac{3-\sqrt{33}}{4} \end{pmatrix}\},$$

$$[T]_{\mathcal{B}} = \begin{pmatrix} \frac{5+\sqrt{33}}{2} & 0 \\ 0 & \frac{5-\sqrt{33}}{2} \end{pmatrix}.$$

Definition 4.1.2. Let V be an F-vector space and $T \in \mathrm{Hom}_F(V, V)$. A subspace $W \subseteq V$ is said to be <u>*T-invariant*</u> or <u>*T-stable*</u> if $T(W) \subseteq W$.

Theorem 4.1.1. *Let* $\dim_F(V) = n$ *and* $W \subseteq V$ *a* k*-dimensional subspace. Let* $\mathcal{B}_W = \{v_1, \ldots, v_k\}$ *be a basis of* W *and extend to a basis* $\mathcal{B} = \{v_1, \ldots, v_n\}$ *of* V*. Let* $T \in \operatorname{Hom}_F(V, V)$*. We have* W *is* T*-stable if and only if* $[T]_{\mathcal{B}}$ *is block upper-triangular of the form*

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

where $A = [T|_W]_{\mathcal{B}_W}$.

Example 4.1.5. Let $V = \mathbf{Q}^4$ with basis $\mathcal{E}_4 = \{e_1, e_2, \ldots, e_4\}$ and define T by:

$$\begin{aligned} T(e_1) &= 2e_1 + 3e_3 \\ T(e_2) &= e_1 + e_4 \\ T(e_3) &= e_1 - e_3 \\ T(e_4) &= 2e_1 - 2e_2 + 5e_3 - 4e_4. \end{aligned}$$

Set $W = \operatorname{span}_{\mathbf{Q}}(e_1, e_3)$, then W is T-stable. Since $\mathcal{B}_W = \{e_1, e_3\}$ and $\mathcal{B} = \{e_1, e_2, e_3, e_4\}$, we have:

$$[T]_{\mathcal{B}} = \begin{pmatrix} 2 & 1 & 1 & 2 \\ 3 & -1 & 0 & 5 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 1 & -4 \end{pmatrix}$$

Example 4.1.6. A special case is when $\dim_F W = 1$. If $W = \operatorname{span}_F(w_1)$ and W is T-stable, then $T(w_1) \in W_1$; i.e., $T(w_1) = \lambda w_1$ for some $\lambda \in F$ Equivalently, this can be written as $(T - \lambda \operatorname{id}_V)(w_1) = 0_V$, meaning $w_1 \in \ker(T - \lambda \operatorname{id}_V)$.

4.2 Eigenvalues and Eigenvectors

Definition 4.2.1. Let $T \in \operatorname{Hom}_F(V, V)$ and $\lambda \in F$. If $\ker(T - \lambda \operatorname{id}_V) \neq \{0_V\}$, we say λ is an <u>*eigenvalue*</u> of T. Any nonzero vector in $\ker(T - \lambda \operatorname{id}_V)$ is called a <u>*λ-eigenvector*</u>. The set $E_\lambda^1 = \ker(T - \lambda \operatorname{id}_V)$ is called the <u>*eigenspace*</u> associated with λ.

Exercise 4.2.1. *Show that* E_λ^1 *is a subspace.*

Exercise 4.2.2. *Let* $T \in \operatorname{Hom}_F(V, V)$*. If* $\lambda_1, \lambda_2 \in F$ *with* $\lambda_1 \neq \lambda_2$*, then* $E_{\lambda_1}^1 \cap E_{\lambda_2}^1 = \{0_V\}$.

Example 4.2.1. Let $A = \left(\begin{smallmatrix} 12 & 35 \\ -6 & 17 \end{smallmatrix}\right) \in \mathrm{Mat}_2(\mathbf{Q})$ and $T_A \in \mathrm{Hom}_{\mathbf{Q}}(\mathbf{Q}^2, \mathbf{Q}^2)$. We have:

$$\begin{pmatrix} -12 & 35 \\ -6 & 17 \end{pmatrix}\begin{pmatrix} 1 \\ \frac{2}{5} \end{pmatrix} = 2\begin{pmatrix} 1 \\ \frac{2}{5} \end{pmatrix}$$

$$\begin{pmatrix} -12 & 35 \\ -6 & 17 \end{pmatrix}\begin{pmatrix} 1 \\ \frac{3}{7} \end{pmatrix} = 3\begin{pmatrix} 1 \\ \frac{3}{7} \end{pmatrix}$$

So T_A has eigenvalues of 2 and 3. Then

$$E_2^1 = \mathrm{span}_{\mathbf{Q}}\left(v_1 = \left(\begin{smallmatrix} 1 \\ 2/5 \end{smallmatrix}\right)\right)$$
$$E_3^1 = \mathrm{span}_{\mathbf{Q}}\left(v_2 = \left(\begin{smallmatrix} 1 \\ 3/7 \end{smallmatrix}\right)\right)$$

gives:

$$[T_A]_{\{v_1,v_2\}} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Example 4.2.2 ($F[x]$-Modules)**.** Let $T \in \mathrm{Hom}_F(V,V)$. Note that V is by definition an F-module, but we are able to view V as an $F[x]$-module given some linear transformation T. The action $F[x] \times V \to V$ is defined by $(f(x), v) \mapsto f(T)(v)$.

Write $T^m = \underbrace{T \circ T \circ \dots \circ T}_{m-\text{times}}$. Write $f(x) \in F[x]$ as $f(x) = a_m x^m + \dots + a_1 x + a_0$. Then

$$f(T) = a_m T^m + \dots + a_1 T + a_0 \,\mathrm{id}_V \in \mathrm{Hom}_F(V,V).$$

For example, let $g(x) = 2x^2 + 3 \in \mathbf{R}[x]$. Then $g(T) = 2T^2 + 3\,\mathrm{id}_V$ and $g(T)(v) = 2T(T(v)) + 3v$. If $f(x) = g(x)h(x)$ for some $g(x), h(x) \in F[x]$, then $f(T) = g(T) \circ h(T)$. Instead of writing $f(T)(v) = g(T)(h(T)(v))$, we will abuse notation and write $g(T)h(T)(v)$. Normally function composition does not commute, but these do for some reason.

Theorem 4.2.1. *Let* $\dim_F(V) = n$ *and* $T \in \mathrm{Hom}_F(V,V)$. *There is a unique monic polynomial* $m_T(x) \in F[x]$ *of lowest degree so that* $m_T(T)(v) = 0_V$ *for all* $v \in V$. *Moreover,* $\deg_{m_T}(T) \leqslant n^2$.

Proof. Recall that $\mathrm{Hom}_F(V,V)$ is an F-vector space. We have $\mathrm{Hom}_F(V,V) \cong \mathrm{Mat}_n(F)$, hence $\dim_F(\mathrm{Hom}_F(V,V)) = n^2$.

Given $T \in \mathrm{Hom}_F(V,V)$, consider the set $\{\mathrm{id}_V, T, T^2, \dots, T^{n^2}\} \subseteq \mathrm{Hom}_F(V,V)$. This has n^2+1 elements, so it must be linearly dependent (meaning a linear combination of some subset can equal 0). Let m be the smallest integer so that

$$a_m T^m + \dots + a_1 T + a_0 \,\mathrm{id}_V \text{ [1]} = 0_{\mathrm{Hom}_F(V,V)}.$$

We obtain a set $\{\mathrm{id}_V, T, T^2, \ldots, T^m\}$. Since m is minimal, $a_m \neq 0$. Define:

$$m_T(x) = x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0 \in F[x], \quad \text{where} \quad b_i = \frac{a_i}{a_m}.$$

This gives $m_T(T) = 0_{\mathrm{Hom}_F(V,V)}$; i.e., $m_T(T)(v) = 0_V$ for all $v \in V$. It remains to that $m_T(x)$ is unique. Suppose there exists an $f(x) \in F[x]$ which satisfies $f(T)(v) = 0_V$ for all $v \in V$. Write:

$$f(x) = m_T(x)q(x) + r(x)$$

for some $q(x), r(x) \in F[x]$ with $r(x) = 0$ or $\deg(r(x)) < \deg(m_T(x))$. We have for all $v \in V$:

$$\begin{aligned} 0_V &= f(T)(v) \\ &= q(T)m_T(T)(v) + r(T)(v) \\ &= q(T)(0_V) + r(T)(v) \\ &= r(T)(v) \end{aligned}$$

It must be the case that $r(x) = 0$, otherwise we have a polynomial of lower degree than $m_T(x)$ which kills all vectors. So $f(x) = m_T(x)q(x)$; i.e., $m_T(x) \mid f(x)$. But if $m_T(x)$ and $f(x)$ are both monic and of minimal degree, it must be the case that they are the same degree. This gives $m_T(x) = f(x)$. □

Definition 4.2.2. The unique monic polynomial $m_T(x)$ is called the *minimal polynomial* of T.

Corollary 4.2.2. *If $f(x) \in F[x]$ satisfies $f(T)(v) = 0_V$ for all $v \in V$, then $m_T(x) \mid f(x)$.*

Proof. wat □

Example 4.2.3. Let $F = \mathbf{Q}$ and $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. We can see that:

$$A - a_0 1_2 \neq 0_2 \quad \text{for any} \quad a_0 \in F.$$

But $A^2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}$ gives $A^2 - 5A - 2 \cdot 1_2 = 0_2$. Hence $m_A(x) = x^2 - 5x - 2$. Note the relationship between this example and Example 4.1.4.

Example 4.2.4. Let $V = \mathbf{Q}^3$, $\mathcal{E}_3 = \{e_1, e_2, e_3\}$, and

$$[T_A]_{\mathcal{E}_3} = A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & -1 \end{pmatrix}.$$

Let $W = \mathrm{span}_{\mathbf{Q}}(e_1)$ Then $T(W) = T(\alpha e_1) = \alpha e_1 \in W$. Hence $T(W) \subseteq W$, meaning W is T-stable. This gives 1 as an eigenvalue. On a completely unrelated note, $m_{T_A}(x) = (x-1)^2(x+1)$.

[1]This seems kind of out of nowhere, so think of it like this: Let $I_T = \{p \in F[x] \mid p(T)(v) = 0_V \text{ for all } v \in V\}$. $F[x]$ is a P.I.D., so every ideal is generated by a single element. The minimal polynomial $m_T(x)$ is the generator of this ideal.

Theorem 4.2.3. *Let V be an F-vector space and $T \in \operatorname{Hom}_F(V, V)$. We have λ is an eigenvalue if and only if λ is the root of $m_T(x)$. In particular, if $(x-\lambda) \mid m_T(x)$, then $E_\lambda^1 \neq \{0_V\}$ (i.e., there is a nonzero $v \in V$ such that $T(v) = \lambda v$).*

Proof. Let λ be an eigenvalue with eigenvector v and write $m_T(x) = x^m + \ldots + a_1 x + a_0$. We have:

$$\begin{aligned} 0_V &= m_T(T)(v) \\ &= (T^m + a_{m-1}T^{m-1} + \ldots + a_1 T + a_0 \operatorname{id}_V)(v) \\ &= T^m(v) + a_{m-1}T^{m-1}(v) + \ldots + a_1 T(v) + a_0 v \\ &= \lambda^m v + a_{m-1}\lambda^{m-1} v + \ldots + a_1 \lambda v + a_0 v \\ &= (\lambda^m + a_{m-1}\lambda^{m-1} + \ldots + a_1\lambda + a_0) v \\ &= m_T(\lambda) \cdot v. \end{aligned}$$

Since $v \neq 0$ and $m_T(\lambda) \in F$, it must be the case that $m_T(\lambda) = 0$. Hence λ is a root.

Now suppose $m_T(\lambda) = 0$. This gives $m_T(x) = (x-\lambda) f(x)$ for some $f(x) \in F[x]$. Since $\deg f(x) < \deg m_T(x)$, this gives a nonzero vector $v \in V$ so that $f(T)(v) \neq 0$ (since $m_T(x)$ is the smallest polynomial that satisfies $m_T(T)(v) = 0_V$, it must be the case that there is a nonzero $v \in V$ that satisfies $f(T)(v) \neq 0$). Set $w = f(T)(v)$, then:

$$\begin{aligned} 0_V &= (T - \lambda \operatorname{id}_V) f(T) \\ &= (T - \lambda \operatorname{id}_V) w, \end{aligned}$$

which simplifies to $T(w) = \lambda w$. Thus λ is an eigenvalue. □

Corollary 4.2.4. *Let $\lambda_1, \ldots, \lambda_n \in F$ be distinct eigenvalues of T. For each i, let v_i be an eigenvector with eigenvalue λ_i. The set $\{v_1, \ldots, v_m\}$ is linearly independent.*

Proof. We have $m_T(x) = (x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_m) f(x)$ for some $f(x) \in F[x]$. Suppose $a_1 v_1 + \ldots + a_m v_m = 0_V$ for $a_i \in F$. Define $g_1(x) = (x-\lambda_2)\ldots(x-\lambda_m) f(x)$. Note that $g_1(T)(v_i) = 0_V$ for $2 \leqslant i \leqslant m$. Then:

$$\begin{aligned} 0_V &= g_1(T)(0_V) \\ &= \sum_{j=1}^{m} a_j g_1(T)(v_j) \\ &= a_1 g_1(T)(v_1) \\ &= a_1 g_1(\lambda_1) v_1 \end{aligned}$$

But $g_1(\lambda_1) \neq 0$ and $v \neq 0$, so it must be that case that $a_1 = 0$. Inductively, it follows for $2, \ldots, m$. □

Corollary 4.2.5. *If $\deg(m_T(x)) = \dim_F(V)$ and $m_T(x)$ has distinct roots, all of which are in F, then we can find a basis $\mathcal{B}$ so that $[T]_\mathcal{B}$ is diagonal.*

Example 4.2.5. Let $A = \begin{pmatrix} 1&0&0 \\ 0&2&0 \\ 0&0&2 \end{pmatrix}$ and $B = \begin{pmatrix} 1&0&0 \\ 0&1&0 \\ 0&0&2 \end{pmatrix}$. These matrices are not similar, however $m_A(x) = m_B(x) = (x-1)(x-2)$. The minimal polynomial is not enough information on the similarity of matrices.

Example 4.2.6. Let:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & -1 \end{pmatrix}.$$

We have that $m_A(x) = (x-1)^2(x+1)$. Note that $Ae_1 = e_1$, so $E_1^1 \supseteq \operatorname{span}_F(e_1)$ (or, more simply, $e_1 \in E_1^1$). Note that $Ae_2 = \left(\begin{smallmatrix} 2 \\ 1 \\ 0 \end{smallmatrix}\right)$. So $e_2 \notin E_1^1$ (another way of saying this is $(A - 1_3)e_2 \neq \left(\begin{smallmatrix} 0 \\ 0 \\ 0 \end{smallmatrix}\right)$). But now consider:

$$(A - 1_3)^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & -8 \\ 0 & 0 & 4 \end{pmatrix}.$$

We have $(A - 1_3)^2 e_2 = \left(\begin{smallmatrix} 0 \\ 0 \\ 0 \end{smallmatrix}\right)$. Thus $e_1, e_2 \in \ker(A - \mathrm{id}_{F^3})^2$.

Definition 4.2.3. Let $T \in \operatorname{Hom}_F(V, V)$. For $k \geqslant 1$, the <u>k^{th} generalized eigenspace</u> of T associated to λ is $E_\lambda^k = \ker(T - \lambda\,\mathrm{id}_V)^k = \{v \in V \mid (T - \lambda\,\mathrm{id}_V)^k v = 0_V\}$. Elements of E_λ^k are called <u>*generalized eigenvectors*</u>. Set $E_\lambda^\infty = \bigcup_{k \geqslant 1} E_\lambda^k$.

Example 4.2.7. Continuing Example 4.2.6, let $\alpha e_1 + \beta e_2 \in \operatorname{span}_F(e_1, e_2)$. Then:

$$(A - 1_3)^2(\alpha e_1 + \beta e_2) = \alpha(A - 1_3)^2 e_1 + \beta(A - 1_3)^2 e_2 = \left(\begin{smallmatrix} 0 \\ 0 \\ 0 \end{smallmatrix}\right).$$

So $\operatorname{span}_F(e_1, e_2) \subseteq E_1^2$. We also have -1 as an eigenvalue with eigenvector $v_3 = \begin{pmatrix} \frac{1}{2} \\ -2 \\ 1 \end{pmatrix}$. Check that $v_3 \notin E_1^2$. So $\dim_F(E_1^2) \leqslant 2$; i.e., $E_1^2 = \operatorname{span}_F(e_1, e_2)$. why does $v_3 \notin E_1^2$ imply the dimension which implies containment in the other direction.

Lemma 4.2.6. *Let V be a finite dimensional F-vector space,* $\dim_F(V) = n$, *and* $T \in \operatorname{Hom}_F(V, V)$. *There exists m with $1 \leqslant m \leqslant n$ such that* $\ker(T^{,}) = \ker(T^{m+1})$. *Moreover, for such an m,* $\ker(T^m) = \ker(T^{m+j})$ *for all $j \geqslant 0$.*

Proof. We have $\ker(T^1) \subseteq \ker(T^2) \subseteq \ldots$ If these containments are always strict, then the dimension increases indefinitely, which contradicts $\dim_F(V) = n$. Hence we have an m with $1 \leqslant m \leqslant n$ and $\ker(T^m) = \ker(T^{m+1})$.

Let m be the smallest value where $\ker(T^m) = \ker(T^{m+1})$. We use induction on j. Base case of $j = 1$ is what defines m. Assume $\ker(T^m) = \ker(T^{m+j})$ for all $1 \leqslant j \leqslant N$. Let $v \in \ker(T^{m+N+1})$. This gives:

$$\begin{aligned} 0_V &= T^{m+N+1}(v) \\ &= T^{m+1}(T^N(v)). \end{aligned}$$

So $T^N(v) \in \ker(T^{m+1})$. However $\ker(T^{m+1}) = \ker(T^m)$, so $T^N(v) \in \ker(T^m)$. Hence:

$$\begin{aligned} 0_V &= T^m(T^n(v)) \\ &= T^{m+N}(v), \end{aligned}$$

so $v \in \ker(T^{m+N})$. Induction hypothesis gives $\ker(T^{m+N}) = \ker(T^m)$, giving $v \in \ker(T^m)$. Thus $\ker(T^{m+N+1}) \subseteq \ker(T^m)$. The other direction of containment is trivial. □

Example 4.2.8. Let $\mathcal{B} = \{v_1, \dots, v_n\}$ be a basis of V and $T \in \mathrm{Hom}_F(V, V)$, $\lambda \in F$ such that:

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{pmatrix}.$$

In other words, $[T]_{\mathcal{B}}$ contains λ along the diagonal and 1 along the super-diagonal. Let $A = [T]_{\mathcal{B}}$. Consider:

$$(A - \lambda 1_n) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We get:

$$\begin{aligned} (A - \lambda 1_n)v_1 &= 0_V \\ (A - \lambda 1_n)v_2 &= v_1 \\ &\vdots \\ (A - \lambda 1_n)v_n &= v_{n-1}. \end{aligned}$$

This gives $E_\lambda^1 = \operatorname{span}_F(v_1)$ (by the first equation). Now observe:

$$(A - \lambda 1_n)^2 v_1 = 0_V$$

$$\begin{aligned}(A - \lambda 1_n)^2 v_2 &= (A - \lambda 1_n)(A - \lambda 1_n) v_2 \\ &= (A - \lambda 1_n) v_1 \\ &= 0_V\end{aligned}$$

$$\begin{aligned}(A - \lambda 1_n)^2 v_3 &= v_1 \\ &\vdots \\ (A - \lambda 1_n)^2 v_n &= v_{n-2}.\end{aligned}$$

So $E_\lambda^2 = \operatorname{span}_F(v_1, v_2)$. In general, we have that $E_\lambda^k = \operatorname{span}_F(v_1, \ldots, v_k)$. Moreover, Lemma 4.2.6 gives $E_\lambda^1 \subseteq E_\lambda^2 \subseteq \ldots \subseteq E_\lambda^k$.

Corollary 4.2.7. *If $\dim_F(V) = n$ and $T \in \operatorname{Hom}_F(V, V)$, there exists an m with $1 \leqslant m \leqslant n$ so that for any $\lambda \in F$, $E_\lambda^\infty = E_\lambda^m$.*

Theorem 4.2.8. *Let $T \in \operatorname{Hom}_F(V, V)$, and $\lambda \in F$ with $(x - \lambda)^k \mid m_T(x)$. We have:*

$$\dim_F(E_\lambda^k) \geqslant k.$$

Proof. Write $m_T(x) = (x - \lambda)^k f(x)$ where $f(x) \in F[x]$, $f(\lambda) \neq 0$. Define $g_k(x) = (x - \lambda)^k$. We have that $(x - \lambda)^{k-1} f(x) = g_{k-1}(x) f(x)$ is *not* the minimal polynomial. So there is a $v \in V$ with $v \neq 0_V$ such that:

$$g_{k-1}(T) f(T)(v) \neq 0_V.$$

Set $v_k = f(T)(v)$. Observe that:

$$\begin{aligned}(T - \lambda \operatorname{id}_V)^k (v_k) &= (T - \lambda \operatorname{id}_V)^k f(T)(v) \\ &= m_T(T)(v) \\ &= 0_V.\end{aligned}$$

So $v_k \in E_\lambda^k$. Moreover, by our construction:

$$\begin{aligned}(T - \lambda \operatorname{id}_V)^{k-1}(v_k) &= g_{k-1}(T)(v_k) \\ &= g_{k-1}(T) f(T)(v) \\ &\neq 0_V.\end{aligned}$$

Hence $v_k \in E_\lambda^k \setminus E_\lambda^{k-1}$. Now set $v_{k-1} = (T - \lambda \operatorname{id}_V) v_k = (T - \lambda \operatorname{id}_V) f(T)(v)$. Note:

$$\begin{aligned}(T - \lambda \operatorname{id}_V)^{k-1}(v_{k-1}) &= (T - \lambda \operatorname{id}_V)^{k-1}(T - \lambda \operatorname{id}_V)(v_k) \\ &= (T - \lambda \operatorname{id}_V)^k (v_k) \\ &= (T - \lambda \operatorname{id}_V)^k f(T)(v) \\ &= m_T(T)(v) \\ &= 0_V.\end{aligned}$$

So $v_{k-1} \in E_\lambda^{k-1}$. Again, by our construction:

$$\begin{aligned}(T - \lambda\, \mathrm{id}_V)^{k-2}(v_{k-1}) &= (T - \lambda\, \mathrm{id}_V)^{k-2}(T - \lambda\, \mathrm{id}_V)(v_k) \\ &= (T - \lambda\, \mathrm{id}_V)^{k-1}(v_k) \\ &\neq 0_V.\end{aligned}$$

So $v_{k-1} \in E_\lambda^{k-1} \setminus E_\lambda^{k-2}$. Setting $v_{k-2} = (T - \lambda\, \mathrm{id}_V)^2 v_k$ gives a similar result. By this construction, we obtain a set $\{v_k, v_{k-1}, \ldots, v_2, v_1\}$. Claim: this set is linearly independent. Suppose towards contradiction it's not, that is, $a_1 v_1 + \ldots + a_k v_k = 0_V$ does not imply $a_1 = \ldots = a_k = 0$. This gives $v_k = \frac{-1}{a_k}(a_1 v_1 + \ldots + a_{k-1} v_{k-1}) \in E_\lambda^{k-1}$, which is a contradiction. It follows that $a_1 = \ldots = a_k = 0$, hence $\{v_k, v_{k-1}, \ldots, v_2, v_1\}$ is linearly independent (linear independent set $\subseteq$ a basis, so thats why the theorem is established). ◻

Example 4.2.9. Let $T_A \in \mathrm{Hom}_F(F^3, F^3)$ be defined by:

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 2 \end{pmatrix}.$$

We have that $m_T(x) = (x-2)^3$. Now observe:

$$(A - 2 \cdot 1_3)^2 = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Note $(A - 2 \cdot 1_3)^2 e_3 = 4e_3 \neq 0_F^3$, but $(A - 2 \cdot 1_3)^3 e_3 = 0_{F^3}$. Set $v_3 = e_3$, we have $v_3 \in E_2^3$. Now observe:

$$\begin{aligned} v_2 &= (A - 2 \cdot 1_3)(v_3) \\ &= \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}.\end{aligned}$$

Similarly:

$$\begin{aligned} v_1 &= (A - 2 \cdot 1_3)(v_2) \\ &= \dots \\ &= \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix}. \end{aligned}$$

Hence:

$$\begin{aligned} E_2^3 &= \mathrm{span}_F \left(\left(\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 3 \\ 4 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 4 \\ 0 \\ 0 \end{smallmatrix}\right) \right) \\ E_2^2 &= \mathrm{span}_F \left(\left(\begin{smallmatrix} 3 \\ 4 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 4 \\ 0 \\ 0 \end{smallmatrix}\right) \right) \\ E_2^1 &= \mathrm{span}_F \left(\left(\begin{smallmatrix} 4 \\ 0 \\ 0 \end{smallmatrix}\right) \right). \end{aligned}$$

Setting $\mathcal{B} = \{v_1, v_2, v_3\}$, we have:

$$[T_A]_{\mathcal{B}} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

4.3 Characteristic Polynomials

Definition 4.3.1. Let $A \in \mathrm{Mat}_n(F)$. The <u>characteristic polynomial</u> is $c_A(x) = \det(x1_n - A)$.

Definition 4.3.2. Let $f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0 \in F[x]$. The <u>companion matrix</u> of $f(x)$ is given by:

$$C(f(x)) = \begin{pmatrix} -a_0 & 0 & 0 & \dots & 0 \\ -a_1 & 1 & 0 & \dots & 0 \\ -a_2 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{n-1} & 0 & 0 & \dots & 1 \end{pmatrix}$$

The companion matrix shows that any polynomial $f(x) \in F[x]$ can be realized as the characteristic polynomial of a matrix.

Lemma 4.3.1. *If $A = C(f(x))$, then $c_A(x) = f(x)$.*

Lemma 4.3.2. *Let $A, B \in \mathrm{Mat}_n(F)$ be similar matrices. Then $c_A(x) = c_B(x)$.*

Proof. Let $A = PBP^{-1}$ for some $P \in GL_n(F)$. We have:

$$\begin{aligned} c_A(x) &= \det(x1_n - A) \\ &= \det(x1_n - PBP^{-1}) \\ &= \det(P(x1_n)P^{-1} - PBP^{-1}) \\ &= \det(P(x1_n - B)P^{-1}) \\ &= \det(P)\det(x1_n - B)\det(P^{-1}) \\ &= \det(x1_n - B) \\ &= c_B(x). \end{aligned}$$

□

Definition 4.3.3. For $T \in \mathrm{Hom}_F(V, V)$, let $\mathcal{B}$ be a basis of V and set $c_T(x) = c_{[T]_{\mathcal{B}}}(x)$.

Theorem 4.3.3. *Let $v \in V$, $v \neq 0_V$. Let $\dim_F(V) = n$. Then there is a unique monic polynomial $m_{T,v}(x) \in F[x]$ so that $m_{T,v}(T)(v) = 0_V$. Moreover, if $f(x) \in F[x]$ with $f(T)(v) = 0_V$, then $m_{T,v}(x) \mid f(x)$.*

Proof. Consider the set $\{v, T(v), T^2(v), \ldots, T^n(v)\}$. Since this set contains $n+1$ elements and the dimension of V is n, the set must be linearly dependent. Write:

$$a_m T^m(v) + \ldots + a_1 T(v) + a_0 = 0_V$$

for some $m \leqslant n$ of minimal order and $a_i \neq 0$ for all i. Set:

$$p(x) = x^m + \frac{a_{m-1}}{a_m}x^{m-1} + \ldots + \frac{a_1}{a_m}x + \frac{a_0}{a_m} \in F[x].$$

By construction $p(T)(v) = 0_V$. Set $I_v = \{g(x) \in F[x] \mid g(T)(v) = 0_V\}$. We have that $p(x)$ is a monic nonzero polynomial in I_v of minimal degree. Set $m_{T,v}(x) = p(x)$.

Let $f(x) \in I_v$. We'd like to show that $m_{T,v}(x) \mid f(x)$. Write:

$$f(x) = q(x)m_{T,v}(x) + r(x),$$

with $q(x), r(x) \in F[x]$ and $\deg(r(x)) = 0$ or $deg(r) < \deg(m_{T,v}(x))$. Observe that:

$$\begin{aligned} r(T)(v) &= f(T)(v) - q(T)m_{T,v}(T)(v) \\ &= 0_V - q(T)0_V \\ &= 0_V. \end{aligned}$$

So $r(x) \in I_v$. But $m_{T,v}(x)$ had minimal degree, so it must be the case that $r(x) = 0$. Thus $f(x) = q(x)m_{T,v}(x)$, implying $m_{T,v}(x) \mid f(x)$[2]. Now suppose $h(x) \in I_v$ with $\deg(h(x)) = \deg(m_{T,v}(x))$. Since both polynomials are monic and of equal degree, if $m_{T,v}(x) \mid h(x)$ then $m_{T,v}(x) = h(x)$. □

Definition 4.3.4. We refer to $m_{T,v}(x)$ as the <u>*T-annihilator*</u> of v.

[2]The proof of $F[x]$ being a P.I.D. follows identically. Instead of considering I_v we would consider an arbitrary polynomial in $F[x]$.

Example 4.3.1. Let $V = F^n$ and $\mathcal{E}_n = \{e_1, \ldots, e_n\}$. Define $T \in \mathrm{Hom}_F(V, V)$ by:

$$\begin{aligned} T(e_1) &= 0_v \\ T(e_j) &= e_{j-1} \text{ for } 2 \leqslant j \leqslant n. \end{aligned}$$

Consider $f(x) = x$. Then $f(T)(e_1) = T(e_1) = 0_V$. Hence $m_{T,e_1}(x) \mid x$. So either $m_{T,e_1}(x) = 1$ or $m_{T,e_1}(x) = x$. But $\mathrm{id}_V(e_1) = e_1 \neq 0_V$, hence it must be the case that $m_{T,e_1}(x) = x$.

Now consider $g(x) = x^2$. Then $g(T)(e_2) = T^2(e_2) = T(T(e_2)) = T(e_1) = 0_V$. Hence $m_{T,e_2}(x) \mid x^2$. So $m_{T,e_2}(x) = 1$ or x or x^2. If $m_{T,e_2}(x) = 1$, then $\mathrm{id}_V(e_2) = e_2 \neq 0_V$. If $m_{T,e_2}(x) = x$, then $T(e_2) = e_1 \neq 0$. So $m_{T,e_2}(x) = x^2$. It follows for $i \leqslant j \leqslant n$, $m_{T,e_j}(x) = x^j$.

Example 4.3.2. Let $V = \mathbf{Q}^2$. Define $T \in \mathrm{Hom}_{\mathbf{Q}}(\mathbf{Q}^2, \mathbf{Q}^2)$ by:

$$\begin{aligned} T(e_1) &= e_1 + 3e_2 \\ T(e_2) &= 2e_1 + 4e_2. \end{aligned}$$

We are trying to find $m_{T,e_1}(x)$. Since V is two-dimensional, $\deg(m_{T,e_1}(x)) = 1$ or 2. Write $m_{T,e_1}(x) = x + a$. Then:

$$\begin{aligned} m_{T,e_1}(T)(e_1) &= T(e_1) + ae_1 \\ &= e_1 + 3e_2 + ae_1 \\ &\neq 0_V. \end{aligned}$$

So it must be that $\deg(m_{T,e_1}(x)) = 2$. Note that:

$$\begin{aligned} T^2(e_1) &= T(e_1 + 3e_2) \\ &= T(e_1) + 3T(e_2) \\ &= 7e_1 + 15e_2. \end{aligned}$$

Now let:

$$T^2(e_1) + bT(e_1) + ce_1 = 0_V,$$

for some $b, c \in \mathbf{Q}$. This will yield a system of equations, and solving for it gives:

$$\begin{aligned} b &= -5 \\ c &= -2. \end{aligned}$$

Hence $m_{T,e_1}(x) = x^2 - 5x - 2$.

Exercise 4.3.1.

1. *Show* $m_{T,e_2}(x) = x^2 - 5x - 2$.
2. *Calculate* $m_{T,e_1}(x)$ *and* $m_{T,e_2}(x)$ *of* $F = \mathbf{F}_3$.

Theorem 4.3.4. *Let* $\dim_F(V) = n$ *and* $\mathcal{B} = \{v_1, \ldots, v_n\}$ *be a basis of* V. *Let* $T \in \mathrm{Hom}_F(V, V)$. *We have:*

$$m_T(x) = \operatorname*{lcm}_{1 \leqslant i \leqslant n} m_{T,v_i}(x).$$

Proof. Let $f(x) = \operatorname{lcm}_{1\leqslant i\leqslant n} m_{T,v_i}(x)$. Note that $m_T(T)(v_i) = 0_V$, so $m_{T,v_i}(x) \mid m_T(x)$ for each i. Hence $f(x) \mid m_T(x)$.

Now let $v \in V$. Write $v = \sum_{i=1}^n a_i v_i$. We have:

$$\begin{aligned} f(T)(v) &= f(T)(\sum_{i=1}^n a_i v_i) \\ &= \sum_{i=1}^n a_i f(T)(v_i) \\ &= 0_V, \end{aligned}$$

because $m_{T,v_i}(x) \mid f(x)$ for all i. Hence $m_T(x) \mid f(x)$. i dont quite get this number theory stuff □

Lemma 4.3.5. *Let $T \in \operatorname{Hom}_F(V,V)$. Let $v_1, \ldots, v_k \in V$, and set $p_i(x) = m_{T,v_i}(x)$. Suppose $p_i(x)$ are pairwise relatively prime. Set $v = v_1 + \ldots + v_k$. Then:*

$$m_{T,v}(x) = p_1(x) \ldots p_k(x).$$

Proof. We prove this for $k \geqslant 2$; i.e., $m_{T,v_1+v_2}(x) = m_{T,v_1}(x)m_{T,v_2}(x)$. Since $p_1(x)$ and $p_2(x)$ are relatively prime, there exists $q_1(x), q_2(x) \in F[x]$ so that $1 = p_1(x)q_1(x)+p_2(x)q_2(x)$. In particular, $\mathrm{id}_V = p_1(T)q_1(T) + p_2(T)q_2(T)$. Set $v = v_1 + v_2$. We have:

$$\begin{aligned} v &= \mathrm{id}_V(v) \\ &= (p_1(T)q_1(T) + p_2(T)q_2(T))(v) \\ &= p_1(T)q_1(T)(v) + p_2(T)q_2(T)(v) \\ &= p_1(T)q_1(T)(v_1 + v_2) + p_2(T)q_2(T)(v_1 + v_2) \\ &= p_1(T)q_1(T)(v_2) + p_2(T)q_2(T)(v_2). \end{aligned}$$

Write $w_1 = p_1(T)q_1(T)(v_2)$ and $w_2 = p_2(T)q_2(T)(v_1)$. This means $v = w_1 + w_2$. Note:

$$\begin{aligned} p_1(T)(w_1) &= p_1(T)p_2(T)q_2(T)(v_1) \\ &= q_2(T)p_2(T)\underbrace{p_1(T)(v_1)} \\ &= 0_V. \end{aligned}$$

Hence $w_1 \in \ker(p_1(T))$. It follows similarly that $w_1 \in \ker(p_2(T))$. Let $r(x) \in F[x]$ with $r(T)(v) = 0_V$. We have $v = w_1 + w_2$ and $w_2 \in \ker(p_2(T))$, so:

$$\begin{aligned} p_2(T)(v) &= p_2(T)(w_1 + w_2) \\ &= p_2(T)(w_1). \end{aligned}$$

Thus:

$$\begin{aligned} 0_V &= p_2(T)q_2(T)(0_V) \\ &= p_2(T)q_2(T)r(T)(v) \\ &= r(T)p_2(T)q_2(T)(v) \\ &= r(T)p_2(T)q_2(T)(w_1). \end{aligned}$$

We also know $r(T)q_1(T)p_1(T)(w_1) = 0_V$ because $w_1 \in \ker(p_1(T))$. Hence:

$$\begin{aligned}0_V &= r(T)p_2(T)q_2(T)(w_1) + r(T)p_1(T)q_1(T)(w_1)\\ &= r(T)\underbrace{\left(p_2(T)q_2(T) + p_1(T)q_1(T)\right)}_{\mathrm{id}_V}(w_1)\\ &= r(T)(w_1).\end{aligned}$$

This gives:

$$\begin{aligned}0_V &= r(T)(w_1)\\ &= r(T)p_2(T)q_2(T)(v_1).\end{aligned}$$

So $r(T)p_2(T)q_2(T)(v_1) = 0_V$. Thus $p_1(x) \mid r(x)p_2(x)q_2(x)$. Now note that:

$$\gcd\left(p_1(x), p_2(x)q_2(x)\right) = 1,$$

which means $p_1(x) \mid r(x)$. A similar argument shows $p_2(x) \mid r(x)$. And since $\gcd(p_1(x), p_2(x)) = 1$, this gives $\mathrm{lcm}(p_1(x), p_2(x)) = p_1(x)p_2(x)$. So $p_1(x)p_2(x) \mid r(x)$. Since $r(x)$ was arbitrary, take $r(x) = m_{T,v}(x)$. Then $p_1(x)p_2(x) \mid m_{T,v}(x)$. Finally, since $p_1(x)p_2(x)(v) = 0_V$, $m_{T,v}(x) \mid p_1(x)p_2(x)$, establishing the lemma. □

Exercise 4.3.2. *Show inductively that $m_{T,v} = p_1(x)p_2(x)\ldots p_k(x)$*[3].

Theorem 4.3.6. *Let $T \in \mathrm{Hom}_F(V,V)$. There exists $v \in V$ such that $m_{T,v}(x) = m_T(x)$. In particular, $\deg(m_T(x)) \leqslant n$.*

Proof. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis. We know:

$$m_T(x) = \operatorname*{lcm}_{1\leqslant i\leqslant n} m_{T,v_i}(x).$$

Factor $m_T(x) = p_1(x)^{e_1}\ldots p_k(x)^{e_k}$, with each $p_i(x)$ relatively prime and $e_1 \geqslant 1$. For $1 \leqslant j \leqslant k$, there exists $i_j \in \{1, \ldots, n\}$ and $q_{i_j}(x) \in F[x]$ with:

$$m_{T,v_{i_j}}(x) = p_j(x)^{e_j} q_{i_j}(x).$$

Set $w_j = q_{i_j}(T)(v_{i_j})$. This gives:

$$m_{T,w_j}(x) = p_j(x)^{e_j}.$$

Now set $w = w_1 + \ldots + w_k$. The previous result gives $m_{T,w}(x) = p_1(x)^{e_1}\ldots p_k(x)^{e_k} = m_T(x)$????. □

Lemma 4.3.7. *Let $W \subseteq V$ be a T-invariant subspace. Then there is an induced map $\overline{T} \in \mathrm{Hom}_F(V/W, V/W)$ defined by $\overline{T}(v + W) = T(v) + W$.*

Lemma 4.3.8. *Let $v \in V$. Then $m_{\overline{T},[v]}(x) \mid m_{T,v}(x)$. Similarly, $m_{\overline{T}}(x) \mid m_T(x)$.*

[3] Pairwise coprime is a stronger statement than just coprime. It means that $\gcd(p_i, p_j) = 1$ for all $1 \leqslant i, j \leqslant k$

Proof. We have:

$$\begin{aligned} m_{T,v}(\overline{T})([v]) &= m_{T,v}(\overline{T})(v+W) \\ &= m_{T,v}(T)(v)+W \\ &= 0_V + W \\ &= 0_{V/W}. \end{aligned}$$

Then by definition of (in this case, $m_{\overline{T},[v]}(x)$) annihilator polynomials, $m_{\overline{T},[v]}(x) \mid m_{T,v}(x)$. □

Definition 4.3.5. Let $T \in \operatorname{Hom}_F(V,V)$ and $\mathcal{A} = \{v_1, \ldots, v_k\}$ a set of vectors in V. The *T-span* of $\mathcal{A}$ is the subspace:

$$W = \left\{ \sum_{i=1}^{k} p_i(T)(v_i) \mid v_i \in \mathcal{A},\ p_i(x) \in F[x] \right\}.$$

We say the subset W is *T-generated* by $\mathcal{A}$.

Exercise 4.3.3. *Show W is a T-invariant subspace of V. Moreover, show it is the smallest T-invariant subspace with respect to inclusion of V that contains $\mathcal{A}$.*

Example 4.3.3. Let $V = \mathbf{Q}^4$. Define $T \in \operatorname{Hom}_{\mathbf{Q}}(\mathbf{Q}^4, \mathbf{Q}^4)$ by:

$$\begin{aligned} T(e_1) &= 2e_1 + 3e_3 \\ T(e_2) &= e_1 + e_2 \\ T(e_3) &= e_1 - e_3 \\ T(e_4) &= 2e_1 - 2e_2 + 5e_3 - 4e_4. \end{aligned}$$

Let $\mathcal{A} = \{e_1\}$. Our goal is to find $T\text{-}\operatorname{span}_{\mathbf{Q}}(\mathcal{A})$. Set $p(x) = 1$, then $p(T)(e_1) = \operatorname{id}_V(e_1) = e_1$. Hence $e_1 \in T\text{-}\operatorname{span}_{\mathbf{Q}}(\mathcal{A})$. Now set $q(x) = \frac{1}{3}(x-2)$. Then:

$$\begin{aligned} q(T)(e_1) &= \frac{1}{3}(T - 2\operatorname{id}_V)(e_1) \\ &= \frac{1}{3}(T(e_1) - 2e_1) \\ &= e_3. \end{aligned}$$

Hence $e_3 \in T\text{-}\operatorname{span}_{\mathbf{Q}}(\mathcal{A})$. So $\operatorname{span}_{\mathbf{Q}}(e_1, e_3) \subseteq T\text{-}\operatorname{span}_{\mathbf{Q}}(\mathcal{A})$ (basically $\alpha p(x) + \beta q(x) \in \operatorname{span}_T, F(\mathcal{A})$, so plugging in a linear combination of e_1 and e_3 will give you back a linear combination of e_1 and e_3). Note that $\operatorname{span}_F(e_1, e_3)$ is T-invariant. By Exercise 4.3.3, since $T\text{-}\operatorname{span}_{\mathbf{Q}}(\mathcal{A})$ is the smallest T-invariant subspace by inclusion, it must be the case that $T\text{-}\operatorname{span}_{\mathbf{Q}}(\mathcal{A}) \subseteq \operatorname{span}_F(e_1, e_3)$. Hence $T\text{-}\operatorname{span}_{\mathbf{Q}}(\mathcal{A}) = \operatorname{span}_F(e_1, e_3)$.

Lemma 4.3.9. *Let $T \in \operatorname{Hom}_F(V,V)$, $w \in V$, and W the subspace of V that is T-generated by $\{w\}$. Then $\dim_F(W) = \deg(m_{T,w}(x))$.*

Proof. Let $\deg(m_{T,w}(x)) = k$. Consider the set $\{w, T(w), \ldots, T^{k-1}(w)\}$. This is a basis of the T-span of $\{w\}$. □

Theorem 4.3.10. *Let* $\dim_F(V) = n$.

(1) We have $m_T(x) \mid c_T(x)$.

(2) Every irreducible factor of $c_T(x)$ *is a factor of* $m_T(x)$.

Proof. (1) Let $\deg(m_T(x)) = k \leqslant n$. Let $v \in V$ with $m_T(x) = m_{T,v}(x)$. Let W_1 be the T-span of $\{v\}$. By Lemma 4.3.9, $\dim_F(W_1) = k$. Write:

$$\begin{aligned} v &= v_k \\ T(v) &= v_{k-1} \\ T^2(v) &= v_{k-2} \\ &\vdots \\ T^i(v) &= v_{k-i}. \end{aligned}$$

We have $\mathcal{B}_1 = \{v_1, \ldots, v_k\}$ is a basis of W_1 (see proof of previous lemma). Since:

$$\begin{aligned} 0_V &= m_T(T)(v) \\ &= T^k(v) + a_{k-1}T^{k-1}(v) + \ldots + a_1T(v) = a_0v, \end{aligned}$$

we have that $T^k(v) = -a_{k-1}T^{k-1}(v) - \ldots - a_1T(v) - a_0v$. Thus:

$$\begin{aligned} T(v_1) &= T(T^{k-1}(v)) = T^k(v) = -a_{k-1}T^{k-1}(v) - \ldots - a_1T(v) - a_0v. \\ T(v_2) &= T(T^{k-2}(v)) = T^{k-1}(v) \\ T(v_3) &= T(T^{k-3}(v)) = T^{k-2}(v) \\ &\vdots \end{aligned}$$

So $\left[T|_{W_1}\right]_{\mathcal{B}_1} = C(m_T(x))$. We proceed with cases:

Case 1: $k = n$. Then $W_1 = V$, and $[T]_{\mathcal{B}_1} = C(m_T(x))$, which has characteristic polynomial $m_T(x)$, meaning $m_T(x) = c_T(x)$.

Case 2: $k < n$. Expand $\mathcal{B}_1$ to a full basis of V as follows: Let $\mathcal{B}_2 = \{v_{k+1}, \ldots, v_n\}$ and write:

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2.$$

Since W_1 is T-invariant, by Theorem 4.1.1 $[T]_{\mathcal{B}}$ will be block diagonal. So we have:

$$[T]_{\mathcal{B}} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}, \ A = \left[T\big|_{W_1}\right]_{\mathcal{B}} = C(m_T(x)).$$

Hence:

$$\begin{aligned} c_T(x) &= \det(x1_n - [T]_{\mathcal{B}}) \\ &= \det\begin{pmatrix} x1_k - A & -B \\ 0 & x1_{n-k} - D \end{pmatrix} \\ &= \det(x1_k - A)\det(x1_{n-k}) - D \\ &= c_A(x)\det(x1_{n-k} - D) \\ &= m_T(x)\det(x1_{n-k} - D). \end{aligned}$$

Thus $m_T(x) \mid c_T(x)$.

For (2), we induct on $\dim_F(V) = n$. If $n = 1$, then both the characteristic polynomial and minimal polynomial are monic and of degree 1, hence they are equal. If $\deg(m_T(x)) = n$, then $m_T(x) \mid c_T(x)$. Since both are degree n and monic, they must be equal. Now suppose $\deg m_T(x) = k < n$ The rest of this proof is hard. □

Example 4.3.4. Consider:

$$A = \begin{pmatrix} 1 & 2 & & & & \\ 3 & 4 & & & & \\ & & 3 & 7 & & \\ & & -1 & 2 & & \\ & & & & -5 & 6 \\ & & & & 2 & -3 \end{pmatrix} \in \mathrm{Mat}_9(\mathbf{Q}).$$

We can verify that $c_A(x) = (x^2 - 5x - 2)(x^2 - x - 1)(x^2 + 8x + 3)$. Since every irreducible factor of $c_T(x)$ is a factor of $m_T(x)$, we have that $m_T(x) = (x^2 - 5x - 2)(x^2 - x - 1)(x^2 + 8x + 3)$.

Theorem 4.3.11 (Cayley-Hamilton)**.**

(1) Let $T \in \mathrm{Hom}_F(V, V)$ *and* $\dim_F(V) < \infty$. *Then* $c_T(T) = 0_{\mathrm{Hom}_F(V,V)}$.

(2) Let $A \in \mathrm{Mat}_n(F)$. *Then* $c_A(A) = 0_n$.

Proof. Write $c_T(x) = f(x)m_T(x)$. Then for any $v \in V$:

$$\begin{aligned} c_T(T)(v) &= f(T)m_T(T)(v) \\ &= f(T)(0_V) \\ &= 0_V. \end{aligned}$$

□

4.4 Jordan Canonical Form

For this section V is always finite-dimensional.

Definition 4.4.1. Let $T \in \mathrm{Hom}_F(V, V)$. A <u>*Jordan basis*</u> for V with respect to T is a basis $\mathcal{B}$ so that:

$$[T]_{\mathcal{B}} = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & 1 & \\ & \lambda_1 & 1 \\ & & \lambda_1 \end{matrix}} & & & & \\ & \boxed{\begin{matrix} \lambda_2 & 1 \\ & \lambda_2 \end{matrix}} & & & \\ & & \boxed{\lambda_3} & & \\ & & & \ddots & \\ & & & & \boxed{\begin{matrix} \lambda_n & 1 \\ & \lambda_n \end{matrix}} \end{pmatrix}.$$

for some $\lambda_1, \ldots, \lambda_n \in F$. More generally, if $V = V_1 \oplus \ldots \oplus V_k$ is a decomposition into T-invariant subspaces, and each V_i has a Jordan basis $\mathcal{B}_i$, we say $\mathcal{B} = \bigcup_{i=1}^{k} \mathcal{B}_i$ is a Jordan basis for V.

Definition 4.4.2. A matrix of the form:

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

is called a <u>*Jordan block*</u> associated to λ_i. We say a matrix J is in <u>*Jordan canonical form*</u> if it is a block diagonal matrix with each block representing a Jordan block.

$$J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_p \end{bmatrix}.$$

Lemma 4.4.1. *Let* $T \in \mathrm{Hom}_F(V, V)$*. We have that* $\ker(T - \lambda\,\mathrm{id}_V)^j$ *and* $\mathrm{im}(T - \lambda\,\mathrm{id}_V)^j$ *are* T*-invariant subspaces for all* $j \geqslant 0$*.*

Proof. Note that $T \circ (T - \lambda\,\mathrm{id}_V)^j = (T - \lambda\,\mathrm{id}_V)^j \circ T$ Let $v \in \ker(T - \lambda\,\mathrm{id}_V)^j$. We have:

$$\begin{aligned} (T - \lambda\,\mathrm{id}_V)^j(T(v)) &= T((T - \lambda\,\mathrm{id}_V)^j(v)) \\ &= T(0_V) \\ &= 0_V. \end{aligned}$$

So $T(v) \in \ker(T - \lambda \operatorname{id}_V)^j$. Now let $w \in \operatorname{im}(T - \lambda \operatorname{id}_V)^j$. We can write $w = (T - \lambda \operatorname{id}_V)^j(v)$ for some $v \in V$. Then:

$$\begin{aligned} T(w) &= T((T - \lambda \operatorname{id}_V)^j(v)) \\ &= (T - \lambda \operatorname{id}_V)(T(v)). \end{aligned}$$

Thus $T(w) \in \operatorname{im}(T - \lambda \operatorname{id}_V)^j$. □

Lemma 4.4.2. *Suppose $m_T(x) = (x - \lambda)^m p(x)$ with $p(\lambda) \neq 0$. Then $E_\lambda^\infty = E_\lambda^m$.*

Proof. Let $E_\lambda^\infty = E_\lambda^e$. Let $v \in E_\lambda^e \setminus E_\lambda^{e-1}$. Since $(T - \lambda \operatorname{id}_V)^e(v) = 0_V$, we know that $m_{T,v}(x) \mid (x - \lambda)^e$. Note that $m_{T,v}(x) \nmid (x - \lambda)^{e-1}$, otherwise $(T - \lambda \operatorname{id}_V)^{e-1}(v) = q(T)m_{T,v}(T)(v) = 0_V$, implying $v \in E_\lambda^{e-1}$ which is a contradiction. Now since $m_{T,v}$ is by definition the monic polynomial of minimal degree which kills v, it must be the case that $m_{T,v}(x) = (x - \lambda)^e$. So we have:

$$(x - \lambda)^e \mid (x - \lambda)^m p(x).$$

Note that $p(x)$ does not have irreducible factors of the form $(x - \lambda)$, hence $(x - \lambda)^e$ and $p(x)$ are coprime, implying:

$$(x - \lambda)^e \mid (x - \lambda)^m.$$

But this is a contradiction if $e > m$, so it must be the case that $e \leqslant m$. Now if we were to assume that $e < m$, then:

$$\begin{aligned} (T - \lambda)^m(v) &= g(T)(T - \lambda)^e(v) \\ &= g(T)(0_V) \\ &= 0_V. \end{aligned}$$

So $m_T(x) \mid (x - \lambda)^m$, a contradiction. Thus $e = m$. □

Lemma 4.4.3. *Let $\dim_F(V) = n$. Let $m_T(x) = (x - \lambda)^m p(x)$ with $p(\lambda) \neq 0$. We have $V = E_\lambda^m \oplus \operatorname{im}(T - \lambda)^m$.*

Proof. Recall that $E_\lambda^m = \ker(T - \lambda)^m$. The dimensions are correct by the rank-nullity theorem. It only remains to show that $E_\lambda^m \cap \operatorname{im}(T - \lambda)^m = \{0_V\}$.

Let $v \in E_\lambda^m \cap \operatorname{im}(T - \lambda)^m$. Since $v \in \operatorname{im}(T - \lambda)^m$, let $v = (T - \lambda)^m(w)$ for some $w \in V$. Applying $(T - \lambda)^m$ to both sides gives:

$$(T - \lambda)^m(v) = (T - \lambda)^{2m}(w).$$

Since $v \in E_\lambda^m$ by our assumption, we have that $0_V = (T - \lambda)^{2m}(w)$. But Lemma 4.4.2 gives that $E_\lambda^\infty = E_\lambda^m$, hence $(T - \lambda)^{2m}(w) = (T - \lambda)^m(w) = 0_V$. Thus $v = 0_V$. □

Theorem 4.4.4. *Assume $m_T(x) = (x - \lambda_1)^{m_1} \ldots (x - \lambda_k)^{m_k}$ with each $\lambda_i \in F$ distinct and $m_j \geqslant 1$. We have:*

$$V = E_{\lambda_1}^{m_1} \oplus \ldots \oplus E_{\lambda_k}^{m_k}.$$

Proof. We use induction on k. If $k = 1$, then $m_T(x) = (x-\lambda_1)^{m_1}$. Since $m_T(T)(v) = 0_V$ for all $v \in V$, we have $V = E_{\lambda_1}^{m_1}$. Now assume the result is true for any vector space W and $S \in \mathrm{Hom}_F(W, W)$ where $m_S(x)$ splits completely over F and has less than k distinct roots. Write:

$$V = E_{\lambda_1}^{m_1} \oplus \mathrm{im}(T - \lambda_1)^{m_1}.$$

Set $W = \mathrm{im}(T - \lambda_1)^{m_1}$. Lemma 4.4.1 gives that W is T-invariant, so $T|_W \in \mathrm{Hom}_F(W, W)$. We want to show that $m_{T|_W}(x) = (x - \lambda_2)^{m_2} \dots (x - \lambda_k)^{m_k}$. Set $p(x) = (x - \lambda_2)^{m_2} \dots (x - \lambda_k)^{m_k}$. Suppose $p(T)(w) \neq 0$. We have:

$$\begin{aligned} 0_V &= m_T(T)(w) \\ &= (T - \lambda_1)^{m_1} p(T)(w). \end{aligned}$$

So $p(T)(w) \in E_{\lambda_1}^{m_1}$. But also $p(T)(w) = p(T|_W)(w) \in W$. So Lemma 4.4.3. gives that $p(T)(w) = p(T|_W)(w) = 0_V$. Thus:

$$m_{T|_W}(x) \mid p(x).$$

Suppose $m_{T|_W}(x)$ is a proper divisor of $p(x)$, that is, $p(x) = T|_W(x)h(x)$ where $\deg(h(x)) > 1$. Consider $f(x) = (x - \lambda_1)^{m_1} m_{T|_W}(x)$. Let $v \in V$ and write $v = v_1 + w$ for some $v_1 \in E_{\lambda_1}^{m_1}$ and $w \in W$. Then:

$$\begin{aligned} f(T)(v) &= f(T)(v_1) + f(T)(w) \\ &= m_{T|_W}(T)(T - \lambda_1)^{m_1}(v_1) + (T - \lambda_1)^{m_1} m_{T|_W}(T)(w) \\ &= 0_V. \end{aligned}$$

Thus $m_T(x) \mid f(x)$. But note that:

$$\begin{aligned} m_T(x) &= (x - \lambda_1)^{m_1} p(x) \\ &= (x - \lambda_1)^{m_1} m_{T|_W}(x) h(x) \\ f(x) &= (x - \lambda_1)^{m_1} m_{T|_W}(x). \end{aligned}$$

This contradicts $m_T(x) \mid f(x)$, so our original assumption that $m_{T|_W}(x)$ is a proper divisor of $p(x)$ is false, it must be the case that $m_{T|_W}(x) = p(x)$. Since $V = E_{\lambda_1}^{m_1} \oplus W$, applying our induction hypothesis to W and $T|_W$ gives:

$$V = E_{\lambda_1}^{m_1} \oplus \left(E_{\lambda_1}^{m_2} \oplus \dots \oplus E_{\lambda_k}^{m_k} \right).$$

□

4.5 Diagonalization, II

The following theorem relates all that was discussed in this chapter.

Theorem 4.5.1. *If $c_T(x)$ does not split into a product of linear factors over F, T is not diagonalizable. If $c_T(x)$ does split into linear factors, the following are equivalent:*

(1) *T is diagonalizable;*

(2) *for every eigenvalue $\lambda, E_\lambda^\infty = E_\lambda^1$;*

(3) $\mathrm{m_T(x)}$ *splits into a product of (distinct) linear factors;*

(4) *for every eigenvalue λ, if $c_T(x) = (x-\lambda)^{e_\lambda} p(x)$ with $p(\lambda) \neq 0$, then $e_\lambda = \dim_F \left(E_\lambda^1\right)$;*

(5) *if we set set $d_\lambda = \dim_F \left(E_\lambda^1\right)$, then $\Sigma_\lambda d_\lambda = \dim_F(V)$;*

(6) *if $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of T, then*

$$V = E_{\lambda_1}^1 \oplus \cdots \oplus E_{\lambda_m}^1$$

5

Tensor Products and the Trace

5.1 Complexification

Recall that if V is a $\mathbf{C}$-vector space, then V is also an $\mathbf{R}$-vector space by restricting the scalars of $\mathbf{C}$. A natural question to ask is if V is an $\mathbf{R}$-vector space, can we "extend" V to be a $\mathbf{C}$-vector space?

Example 5.1.1 (Complexification of $\mathbf{R}$). Let $V = \mathbf{R}$. We cannot make $\mathbf{R}$ into a $\mathbf{C}$-vector space. However, we do have $\mathbf{R} \hookrightarrow \mathbf{C}$ by $x \mapsto x + 0i$, with $\mathbf{C}$ as a $\mathbf{C}$-vector space. But note that $z \in \mathbf{C}$ can we written as $z = x + yi$. There is an isomorphism $\mathbf{R} \oplus \mathbf{R} \cong \mathbf{C}$ as $\mathbf{R}$-vector spaces by:

$$x + yi \mapsto (x, y)$$

If we take $z = x + yi \in \mathbf{C}$ to be a vector, and $a + bi \in \mathbf{C}$ to be a scalar, we have:

$$(a + bi)(x + yi) = (ax - by) + (ay + bx)i,$$

meaning in $\mathbf{R} \oplus \mathbf{R}$ we define:

$$(a + bi)(x, y) = (ax - by, ay + bx)$$

With scalar multiplication defined as above, then $\mathbf{R} \oplus \mathbf{R}$ is a $\mathbf{C}$-vector space. Furthermore, we have $\mathbf{R} \oplus \mathbf{R} \cong \mathbf{C}$ as $\mathbf{C}$-*vector spaces*!

Definition 5.1.1. Let V be a real vector space. The *complexification* of V is denoted $V_{\mathbf{C}} = V \oplus V$, where complex scalar multiplication is defined by:

$$(a + bi)(v_1, v_2) = (av_1 - bv_2, av_2 + bv_1).$$

Upon investigation one can see:

$$i(v_1, v_2) = (-v_2, v_1).$$

Exercise 5.1.1. *Prove that $V_{\mathbf{C}}$ is a $\mathbf{C}$-vector space.*

Proposition 5.1.1. *Let $\mathcal{B} = \{v_i\}_{i \in I}$ be an $\mathbf{R}$-basis of V. The set $\mathcal{B}_{\mathbf{C}} = \{(v_j, 0_V)\}_{j \in I}$ is a $\mathbf{C}$-basis of $V_{\mathbf{C}}$.*

Proof. Let $(w_1, w_2) \in V_{\mathbf{C}}$. We can write:

$$w_1 = \sum_{j \in I} a_j v_j$$
$$w_2 = \sum_{j \in I} b_j v_j$$

for some $a_j, b_j \in \mathbf{R}$. We have:

$$\begin{aligned}
(w_1, w_2) &= \left(\sum_{j\in I} a_j v_j, \sum_{j\in I} b_j v_j\right) \\
&= \left(\sum_{j\in I} a_j v_j, 0_V\right) + \left(0_V, \sum_{j\in I} b_j v_j\right) \\
&= \sum_{j\in I} a_j (v_j, 0_V) + \sum_{j\in I} b_j (0_V, v_j) \\
&= \sum_{j\in I} a_j (v_j, 0_V) + \sum_{j\in I} i b_j (v_j, 0_V) \\
&\in \operatorname{span}_{\mathbf{C}} \left\{(v_j, 0_V)\right\}_{i\in I}.
\end{aligned}$$

Now suppose we have $(0_V, 0_V) = \sum_{j\in I} (a_j + i b_j)(v_j, 0_V)$. Then:

$$\begin{aligned}
(0_V, 0_V) &= \sum_{j\in I} (a_j + i b_j)(v_j, 0_V) \\
&= \sum_{j\in I} a_j (v_j, 0_V) + \sum_{j\in I} i b_j (v_j, 0_V) \\
&= \left(\sum_{j\in I} a_j v_j, 0_V\right) + i\left(\sum_{j\in I} b_j v_j, 0_V\right) \\
&= \left(\sum_{j\in I} a_j v_j, 0_V\right) + \left(\sum_{j\in I} 0_V, b_j v_j\right) \\
&= \left(\sum_{j\in I} a_j v_j, \sum_{j\in I} b_j v_j\right),
\end{aligned}$$

meaning:

$$\begin{aligned}
\sum_{j\in I} a_j v_j &= 0_V \\
\sum_{j\in I} b_j v_j &= 0_V.
\end{aligned}$$

So $a_j = 0$ for all j and $b_j = 0$ for all j. Thus $\left\{(v_j, 0_V)\right\}_{j\in I}$ are linearly independent. □

Proposition 5.1.2. *Let V, W be $\mathbf{R}$-vector spaces, and let $T \in \operatorname{Hom}_{\mathbf{R}}(V, W)$. There is a unique $T_{\mathbf{C}} \in \operatorname{Hom}_{\mathbf{C}}(V_{\mathbf{C}}, W_{\mathbf{C}})$ that makes the following diagram commute:*

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, hook, "\iota_V"'] & W \arrow[d, hook, "\iota_W"] \\
V_{\mathbf{C}} \arrow[r, "T_{\mathbf{C}}"'] & W_{\mathbf{C}}
\end{tikzcd}$$

Proof. Define

$$T_{\mathbf{C}}(v_1, v_2) = (T(v_1), T(v_2)).$$

Let $v \in V$. We have $\iota_V(v) = (v, 0_V)$, meaning:

$$\begin{aligned} T_{\mathbf{C}}(\iota_V(v)) &= T_{\mathbf{C}}((v, 0_V)) \\ &= (T(v), T(0_V)) \\ &= (T(v), 0_W), \end{aligned}$$

and:

$$\iota_W(T(v)) = (T(v), 0_W).$$

Hence the diagram commutes. We claim that $T_{\mathbf{C}}$ is $\mathbf{C}$-linear. Let $x + iy \in \mathbf{C}$ and $(v_1, v_2), (v_1', v_2) \in V_{\mathbf{C}}$. Then:

$$\begin{aligned} T_{\mathbf{C}}\left((v_1, v_2) + (x + \mathrm{iy})\left(v_1', v_2'\right)\right) &= T_{\mathbf{C}}\left((v_1, v_2) + \left(xv_1' - yv_2', xv_2' + yv_1'\right)\right) \\ &= T_{\mathbf{C}}\left(\left(v_1 + xv_1' - yv_2', v_2 + xv_2' + yv_1'\right)\right) \\ &= \left(T\left(v_1 + xv_1' - yv_2'\right), T\left(v_2 + xv_2' + yv_1'\right)\right) \\ &= \left(T\left(v_1\right), T\left(v_2\right)\right) + x\left(\mathrm{T}\left(v_1'\right), T\left(v_2'\right)\right) + y\left(-T\left(v_2'\right), T\left(v_1'\right)\right) \\ &= \left(T\left(v_1\right), T\left(v_2\right)\right) + (x + iy)\left(T\left(v_1'\right), T\left(v_2'\right)\right) \\ &= T_{\mathbf{C}}\left(v_1, v_2\right) + (x + \mathrm{iy})T_{\mathbf{C}}\left(v_1', v_2'\right). \end{aligned}$$

Hence $T_{\mathbf{C}}$ is linear. Now suppose there is an $S \in \mathrm{Hom}_{\mathbf{C}}(V_{\mathbf{C}}, W_{\mathbf{C}})$ making the following diagram commute:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, hook, "\iota_V"'] & W \arrow[d, hook, "\iota_W"] \\ V_{\mathbf{C}} \arrow[r, "S"'] & W_{\mathbf{C}} \end{tikzcd}$$

Let $v_1, v_2 \in V_{\mathbf{C}}$. Then:

$$\begin{aligned} S((v_1, v_2)) &= S((v_1, 0_V) + (0_V, v_2)) \\ &= S((v_1, 0_V) + i(v_2, 0_V)) \\ &= S((v_1, 0_V)) + iS((v_2, 0_V)) \\ &= S(\iota_V(v_1)) + iS(\iota_V(v_2)) \\ &= \iota_W(T(v_1)) + i\iota_W(T(v_2)) \\ &= (T(v_1), 0_W) + i(T(v_2), 0_W) \\ &= (T(v_1, 0_W)) + (0_W, T(v_2)) \\ &= (T(v_1), T(v_2)) \\ &= T_{\mathbf{C}}((v_1, v_2)). \end{aligned}$$

Thus $T_{\mathbf{C}}$ is unique. ◻

5.2 Free Vector Spaces

We showed in Section 2.2 that every vector space has a basis. In this section we show that given a set X, we can build a vector space that "has" X as a basis. We will give a few basic definitions before investigating the properties of these spaces.

Definition 5.2.1. Let $f : \Omega \to F$ be a map whose domain is an arbitrary set Ω. The _support_ if f, denoted $\mathrm{supp}(f)$ is the set of points in Ω where f is nonzero:

$$\mathrm{supp}(f) = \{x \in \Omega \mid f(x) \neq 0\}.$$

Definition 5.2.2. Let F be a field. The set of all functions from Ω to F is denoted:

$$\mathcal{F}(\Omega, F) = \{f \mid f : \Omega \to F\}.$$

Exercise 5.2.1. *Show that $\mathcal{F}(\Omega, F)$ is an F-vector space.*

Example 5.2.1. Fix $t \in \Gamma$. Recall that $\delta_t : \Gamma \to F$ is defined by:

$$\delta_t(s) = \begin{cases} 1, & s = t \\ 0, & s \neq t \end{cases}.$$

We have that $\delta_t \in \mathcal{F}(\Gamma, F)$, and furthermore $\mathrm{supp}(\delta_t) = \{t\}$. If $f \in \mathcal{F}(\Gamma, F)$ has finite support, then $\mathrm{supp}(f) = \{t_1, ..., t_n\}$ for some $t_i \in F$. If:

$$\begin{aligned} f(t_1) &= \alpha_1 \neq 0 \\ f(t_2) &= \alpha_2 \neq 0 \\ &\vdots \\ f(t_n) &= \alpha_n \neq 0, \end{aligned}$$

then we can write $f = \sum_{j=1}^{n} \alpha_j \delta_{t_j}$.

Theorem 5.2.1 (Existence of Free Vector Spaces)**.** *Let F be a field and Γ a set. There is an F-vector space $\mathbb{F}(\Gamma)$ that has a basis isomorphic to Γ as sets. Moreover, $\mathbb{F}(\Gamma)$ has the following universal property: if W is any F-vector space and $t : \Gamma \to W$ is a map of sets, there is a unique $T \in \mathrm{Hom}_F(\mathbb{F}(\Gamma), W)$ such that $T(x) = t(x)$ for every $x \in \Gamma$; i.e., the following diagram commutes:*

$$\begin{tikzcd} \Gamma \arrow[r, hook, "\iota"] \arrow[dr, "t"'] & \mathbb{F}(\Gamma) \arrow[d, "T"] \\ & W \end{tikzcd}$$

Proof. If Γ is the empty set, take $\mathbb{F}(\Gamma) = \{0\}$. Let $\Gamma \neq \emptyset$. Define:

$$\mathbb{F}(\Gamma) = \{f : \Gamma \to F \mid \mathrm{supp}(f) < \infty\}.$$

Since $\mathbb{F}(\Gamma) \subseteq \mathcal{F}(\Gamma, F)$, this space inherits a natural vector space structure. In particular, if f, g are finitely supported functions and $c \in F$, then $(f+g)(x) = f(x) + f(x)$ and $(cf)(x) = cf(x)$ will be finitely supported. Moreover, the zero element of this set if $f(x) = 0_{\mathbb{F}(\Gamma)}$. The rest of the vector space axioms are left as an exercise.

We obtain an inclusion $\iota : \Gamma \hookrightarrow \mathbb{F}(\Gamma)$ by $x \mapsto \delta_x$. Let $\mathcal{X} = \{\delta_x \mid x \in \Gamma\}$. This a subset of $\mathbb{F}(\Gamma)$ and furthermore we have a bijection $\Gamma \hookrightarrow\!\!\!\!\rightarrow \mathcal{X}$.

Let $f \in \mathbb{F}(\Gamma)$. We can write $f = \sum_{x\in\Gamma} f(x)\delta_x \in \operatorname{span}_F(\mathcal{X})$. Hence $\operatorname{span}_F(\mathcal{X}) = \mathbb{F}(\Gamma)$. Note that:

$$\begin{aligned} f(y) &= f(y)\delta_y(y) \\ &= f(y)\delta(y)(y) + \sum_{x\neq y} f(x)\delta_x(y) \\ &= \sum_{x\in\Gamma} f(y)\delta_x(y). \end{aligned}$$

Note that $f(y)$ is just a scalar in F, hence an arbitrary element of $\mathbb{F}(\Gamma)$ looks like $\sum_{i=1}^n a_i\delta_{x_i}$. Suppose then that $\sum_{i=1}^n a_i\delta_{x_i} = 0_{\mathbb{F}(\Gamma)}$. We have that $\sum_{i=1}^n a_i\delta_{x_i}(y) = 0_F$ for all $y \in \Gamma$. Thus:

$$\begin{aligned} 0_F &= \sum_{i=1}^n a_i\delta_{x_i}(x_j) \\ &= a_j. \end{aligned}$$

This establishes $\mathcal{X}$ as a basis for $\mathbb{F}(\Gamma)$.

Now suppose we have $t : \Gamma \to W$. Define $T : \mathbb{F}(\Gamma) \to W$ by:

$$T\left(\sum_{i=1}^n a_i\delta_{x_i}\right) = \sum_{i=1}^n a_i t(\iota^{-1}(\delta_{x_i})).$$

Because $\mathcal{X}$ is a basis, this gives a well-defined linear map. It is unique because any linear map that agrees with t on $\mathcal{X}$ must agree with T on $\mathbb{F}(\Gamma)$, establishing the proof. □

Example 5.2.2. If $\Gamma = \mathbf{R}$, we can form $\mathbb{F}_{\mathbf{R}}(\mathbf{R})$. An example of an element of $\mathbb{F}_{\mathbf{R}}(\mathbf{R})$ is $2\cdot\pi + 3\cdot 2$, where $\pi, 2$ are basis elements and $2, 3$ are scalars. Note that, from this construction, we cannot simplify this expression.

Exercise 5.2.2. *Show that if* $\Gamma = \{x_1, \ldots, x_n\}$, *then* $\mathbb{F}(\Gamma) \cong F^n$.

5.3 Extension of Scalars

Let V be an F-vector space and K/F an extension of fields. We can naturally consider K as an F-vector space. As we did with complexification, we want to define a way to "multiply" vectors in V by scalars in K. The way we define "multiplication" should be obvious: Let $a, a_1, a_2 \in K$, $c \in F$, and $v, v_1, v_2 \in V$. We want multiplication to satisfy:

(1) $(a_1 + a_2) \star v$;

(2) $a \star (v_1 + v_2) = a \star v_1 + a \star v_2$;

(3) $(ac) \star v = a \star (cv)$.

We will construct a vector space that satisfies exactly this by constructing the *tensor product* of V with K.

Definition 5.3.1. Let V be an F-vector space and K/F be an extension of fields. Let $K \times V$ be the Cartesian product of K and V and define:

$$\begin{aligned}
\mathcal{A}_1 &= \{(a_1 + a_2, v) - (a_1, v) - (a_2, v) \mid a_1, a_2 \in K, v \in V\}, \\
\mathcal{A}_2 &= \{(a, v_1 + v_2) - (a, v_1) - (a, v_2) \mid a \in K, v_1, v_2 \in V\}, \\
\mathcal{A}_3 &= \{(ca, v) - (a, cv) \mid c \in F, a \in K, v \in V\}, \\
\mathcal{A}_4 &= \{a_1(a_2, v) - (a_1 a_2, v) \mid a_1, a_2 \in K, v \in V\}.
\end{aligned}$$

Define $\mathrm{Rel}_K(K \times V) = \mathrm{span}_F(\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4)$. The <u>*tensor product*</u> of K and V over F is:

$$K \otimes_F V = \mathbb{F}(K \times V)/\mathrm{Rel}_K(K \times V).$$

For any arbitrary element $\sum_{\text{finite}} c_i \delta_{a_i, v_i} \in \mathbb{F}(K \times V)$, we denote the equivalence class $\sum_{\text{finite}} c_i \delta_{a_i, v_i} + \mathrm{Rel}_K(K \times V)$ as:

$$\begin{aligned}
\sum_{\text{finite}} c_i (a_i \otimes v_i) &= \sum_{\text{finite}} c_i a_i (1 \otimes v_i) \\
&= \sum_{\text{finite}} b_i \otimes v_i
\end{aligned}$$

for some $b_i \in K$. An element of the form $a \otimes v$ is referred to as a <u>*pure tensor*</u>. Both arbitrary elements of $K \otimes_F V$ and pure tensors admit the following properties:

(1) $(a_1 + a_2) \otimes v = a_1 \otimes v + a_2 \otimes v$ for all $a_1, a_2 \in K$, $v \in V$;

(2) $a \otimes (v_1 + v_2) = a \otimes v_1 + a \otimes v_2$ for all $a \in K$, $v_1, v_2 \in V$;

(3) $ca \otimes v = a \otimes cv$ for all $c \in F$, $a \in K$, and $v \in V$;

(4) $a_1(a_2 \otimes v) = (a_1 a_2) \otimes v$ for all $a_1, a_2 \in K$, $v \in V$.

Note 4.

(1) An **arbitrary element of $K \otimes_F V$ is a finite sum.** It is a common mistake when working with tensor products to check things for pure tensors and not with arbitrary elements.

(2) Since $K \otimes_F V$ is a quotient space, there must be care in checking things are well-defined when working with tensor products.

Exercise 5.3.1. *Show that $K \otimes_F V$ is a K-vector space. (Hint: $0 \otimes 0_V$ is the additive identity in $K \otimes_F V$).*

Proposition 5.3.1. *Let K/F be a field extension and V an F-vector space with basis $\mathcal{B} = \{v_i\}_{i \in I}$. We have* $\mathrm{span}_K \{1 \otimes v_i\}_{i \in I} = K \otimes_F V$.

Proof. Let $a \otimes v \in K \otimes_F V$. Write $v = \sum_{i=1}^n c_i v_i$ for some $c_i \in F$. We have:

$$\begin{aligned} a \otimes v &= a \otimes \left(\sum_{i=1}^n c_i v_i \right) \\ &= \sum_{i=1}^n a \otimes c_i v_i \\ &= \sum_{i=1}^n a c_i \otimes v_i \\ &= \sum_{i=1}^n a c_i (1 \otimes v_i). \end{aligned}$$

From this, it follows that every pure tensor $a \otimes v$ is also in the span of $\{1 \otimes v_i\}_{i \in I}$. This gives that all finite sums of the form $\sum_{j \in I} a_j \otimes \tilde{v}_j$ are also in the span of $\{1 \otimes v_i\}_{i \in I}$. Hence $\operatorname{span}_F \{1 \otimes v_i\}_{i \in I} = K \otimes_F V$. □

Theorem 5.3.2. *Let K/F be an extension of fields, V an F-vectorspace, and $\iota_V : V \to K \otimes_F V$ defined by $\iota_V(v) = 1 \otimes v$. Let W be any K-vector space and $S \in \mathrm{Hom}_F(V, W)$. There is a unique $T \in \mathrm{Hom}_K(K \otimes_F V, W)$ so that $S = T \circ \iota_V$; i.e., the following diagram commutes:*

$$\begin{tikzcd} V \arrow[r, "\iota_V"] \arrow[dr, "S"'] & K \otimes_F V \arrow[d, "T"] \\ & W \end{tikzcd}$$

Conversely, if $T \in \mathrm{Hom}_K(K \otimes_F V, W)$, then $T \circ \iota_V \in \mathrm{Hom}_F(V, W)$.

Proof. Let $S \in \mathrm{Hom}_F(V, W)$. Recall we constructed $K \otimes_F V$ as a quotient of $\mathbb{F}(K \times V)$. Define:

$$t : K \times V \to W \text{ by } (a, v) \mapsto aS(v)$$

as a map of sets. Theorem 5.2.1 tells us that t extends to a map $T \in \mathrm{Hom}_K(\mathbb{F}(K \times V), W)$ such that $T((a, v)) = t((a, v))$. Since T is linear:

$$\begin{aligned} T\left(\sum_{i \in I} c_i (a_i, v_i) \right) &= \sum_{i \in I} T(c_i (a_i, v_i)) \\ &= \sum_{i \in I} c_i T((a_i, v_i)) \\ &= \sum_{i \in I} c_i t((a_i, v_i)) \\ &= \sum_{i \in I} c_i a_i S(v_i). \end{aligned}$$

We must check that T is the zero map when restricted to $\mathrm{Rel}_K(K \times V)$. We have:

$$\begin{aligned} T((a + b, v) - (a, v) - (b, v)) &= T((a + b, v)) - T((a, v)) - T((b, v)) \\ &= (a + b)S(v) - aS(v) - bS(v) \\ &= 0_W. \end{aligned}$$

The rest of the relations are left as an exercise. Thus we have $T \in \mathrm{Hom}_K(K \otimes_F V, W)$ defined by $T\left(\sum_{i \in I} c_i(a_i \otimes v_i)\right) = \sum_{i \in I} c_i a_i S(v_i)$. To see that the diagram commutes, observe that:

$$T(\iota_V(v)) = T(1 \otimes v) = 1 \cdot S(v) = S(v).$$

From Proposition 5.3.1, we saw that $K \otimes_F V$ is spanned by elements of the form $1 \otimes v$. Hence any linear map on $K \otimes_F V$ is determined by the image of these elements. Since $T(1 \otimes v) = S(v)$, we get T is uniquely determined by S.

The converse statement that for any $T \in \mathrm{Hom}_K(K \otimes_F V)$ one has $T \circ \iota_V \in \mathrm{Hom}_F(V, W)$ is left as an exercise. □

Exercise 5.3.2. *Complete the proof that T vanishes on all the relations.*

Exercise 5.3.3. *Given $T \in \mathrm{Hom}_K(K \otimes_F V, W)$, show $T \circ \iota_V \in \mathrm{Hom}_F(V, W)$.*

Proposition 5.3.3. *Let K/F be an extension of fields. Then $K \otimes_F F \cong K$ as K-vector spaces.*

Proof. There is a natural inclusion map $i : F \to K$. Let $\iota : F \to K \otimes_F F$. By the universal property we obtain a unique K-linear map $T : K \otimes_F F \to K$ so that the following diagram commutes:

$$\begin{tikzcd} F \arrow[r, "\iota"] \arrow[dr, hook, "i"'] & K \otimes_F F \arrow[d, "T"] \\ & K \end{tikzcd}$$

We see that $T(1 \otimes x) = i(x) = x$. Since T is K-linear this completely determines T because, for $\sum_{i \in I} a_i \otimes x_i \in K \otimes_F F$, we have:

$$\begin{aligned} T\left(\sum_{i \in I} a_i \otimes x_i\right) &= \sum_{i \in I} T(a_i \otimes x_i) \\ &= \sum_{i \in I} T(a_i(1 \otimes x_i)) \\ &= \sum_{i \in I} a_i T(1 \otimes x_i) \\ &= \sum_{i \in I} a_i x_i. \end{aligned}$$

If we show T has an inverse map, then we obtain an isomorphism. Let $S : K \to K \otimes_F F$ defined by $y \mapsto y \otimes 1$. Let $a \in K$, and $y_1, y_2 \in K$. Then:

$$S(y_1 + a y_2) = \ldots$$

Hence $S \in \mathrm{Hom}_K(K, K \otimes_F F)$. Since S, T are linear, it is enough to check that they are inverses with pure tensors. Observe that:

$$\begin{aligned} T(S(y)) &= T(y \otimes 1) = yT(1 \otimes 1) = y \\ S(T(a \otimes x)) &= S(aT(1 \otimes x)) = S(ax) = ax \otimes 1 = a \otimes x. \end{aligned}$$

Thus $T^{-1} = S$, and so $K \otimes_F F \cong K$ as K-vector spaces. □

Example 5.3.1. From the previous section, we can now see that $\mathbf{R_C} = \mathbf{C} \otimes_{\mathbf{R}} \mathbf{R} \cong \mathbf{C}$.

Example 5.3.2. It is not always obvious that an element of $K \otimes_F F$ is nonzero. Take for example $\mathbf{Z}/2\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}$. We have that $1 \otimes 2 = 2 \otimes 1 = 0_{\mathbf{Z}/2\mathbf{Z}} \otimes 1 = 0_{\mathbf{Z}/2\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}}$.

Exercise 5.3.4. *Show that* $V_{\mathbf{C}} \cong \mathbf{C} \otimes_{\mathbf{R}} V$.

Proposition 5.3.4. *Let* K/F *be an extension of fields and* V *an* F*-vector space with* $\dim_F V = n$. *Then* $K \otimes_F V \cong K^n$ *as* K*-vector spaces.*

Proof. We want a K-linear map $K \otimes_F V \to K^n$. Take $\mathcal{B} = \{v_1, \ldots, v_n\}$ to be a basis for V and $\mathcal{E}_n = \{e_1, \ldots, e_n\}$ the standard basis for K^n. Define a map $t : V \to K^n$ by $t(v_i) = e_i$. Since this map is defined on a basis, it extends to an F-linear map. So $t \in \operatorname{Hom}_F(V, K^n)$. The universal property gives $T \in \operatorname{Hom}_K(K \otimes_F V, K^n)$ so that $T(1 \otimes v_i) = e_i$. We will show that T has an inverse. Define $S \in \operatorname{Hom}_K(K^n, K \otimes_F V)$ by $S(e_i) = 1 \otimes v_i$. These are clearly inverse maps, so $K \otimes_F V \cong K^n$. Moreover, since S is invertible and the e_i are a basis, $\{S(e_i)\}_{i=1}^n$ gives a basis of $K \otimes_F V$; i.e., $\{1 \otimes v_i\}_{i=1}^n$ is a basis. □

Proposition 5.3.5. *Let* K/F *be an extension of fields and* V *an* F*-vector space. Let* $\mathcal{B} = \{v_i\}_{i \in I}$ *be an* F*-basis of* V*. We have* $\mathcal{B}_K = \{1 \otimes v_i\}_{i \in I}$ *is a basis of* $K \otimes_F V$.

Proof. We saw in Proposition 5.3.1 that $\mathcal{B}_K$ spans $K \otimes_F V$. Suppose there exists a linear dependence $\sum_{i \in I} c_i(1 \otimes v_i) = 0_{K \otimes_F V}$. Given $(b, v) \in K \times V$, write $(b, \sum_{i \in I} a_i v_i)$ for some $a_i \in F$. Fix $i_0 \in I$ and define:

$$t_{i_0} : V \to K$$

by $t_{i_0}(v) = t_{i_0}(\sum_{i \in I} a_i v_i) = a_{i_0}$. One can check that $t_{i_0} \in \operatorname{Hom}_F(V, K)$. The universal property extends this to $T_{i_0} \in \operatorname{Hom}_K(K \otimes_F V, K)$ so that $T_{i_0}(1 \otimes v) = t_{i_0}(v) = a_{i_0}$. Recall that $\sum_{i \in I} c_i(1 \otimes v_i) = 0_{K \otimes_F V}$. Observe that:

$$\begin{aligned} 0_K &= T_{i_0}(0_{K \otimes_F V}) \\ &= T_{i_0}\left(\sum_{i \in I} c_i(1 \otimes v_i)\right) \\ &= \sum_{i \in I} c_i T_{i_0}(1 \otimes v_i) \\ &= \sum_{i \in I} c_i t_{i_0}(v_i) \\ &= c_{i_0}. \end{aligned}$$

Since $i_0 \in I$ was arbitrary, we have that $c_i = 0$ for all $i \in I$; i.e., $\mathcal{B}_K$ is linearly independent. Thus $\mathcal{B}_K$ is a basis of $K \otimes_F V$. □

Theorem 5.3.6. *Let* K/F *be an extension of fields and* V, W *both* F*-vector spaces. Given* $T \in \operatorname{Hom}_F(V, W)$*, there is a unique map* $T_K \in \operatorname{Hom}_K(K \otimes_F V, K \otimes_F W)$ *so that the following diagram commutes:*

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, hook, "\iota_V"'] & W \arrow[d, hook, "\iota_W"] \\ K \otimes_F V \arrow[r, "T_K"'] & K \otimes_F W \end{tikzcd}$$

Proof. Define $t : V \to K \otimes_F W$ by $t(v) = 1 \otimes T(v)$. It can be shown that $t \in \mathrm{Hom}_F(V, K \otimes_F W)$. The universal property extends this to a unique map $T_K \in \mathrm{Hom}_K(K \otimes_F V, K \otimes_F W)$ so that $t = T_K \circ \iota_V$. Let $v \in V$. We have that $\iota_W(T(v)) = 1 \otimes T(v) = t(v) = (T_K \circ \iota_V)(v)$, meaning the diagram commutes. □

Exercise 5.3.5. *Let V be an $\mathbf{R}$-vector space. We have $\mathbf{C} \otimes_{\mathbf{R}} V \cong V_{\mathbf{C}}$.*

5.4 Tensor Products of Vector Spaces

Definition 5.4.1. Let U, V, W be F-vector spaces. If $t : V \times W \to U$ satisfies:

(1) $t(v_1 + v_2, w) = t(v_1, w) + t(v_2, w)$;

(2) $t(v, w_1 + w_2) = t(v, w_1) + t(v, w_2)$;

(3) $ct(v, w) = t(cv, w) = t(v, cw)$;

we call t a _bilinear map_. The collection of bilinear maps is denoted $\mathrm{Hom}_F(V, W; U)$. If $t \in \mathrm{Hom}_F(V, V; F)$, then we say t is a _bilinear form_.

Example 5.4.1.

1. The standard dot product $\mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ is a bilinear form.
2. The standard cross-product in $\mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}^3$ is a bilinear map.

Definition 5.4.2. Let V, W be F-vector spaces and consider the Cartesian product $V \times W$. Let:

$$asdf$$

and define $\mathrm{Rel}_F(V \times W) = \mathrm{span}_F\{\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_4\}$. The _tensor product_ of V and W over F is:

$$V \otimes_F W = \mathbb{F}(V \times W)/\mathrm{Rel}_F(V \times W).$$

For any arbitary element $\sum_{\text{finite}} c_i \delta_{v_i, w_i} \in \mathbb{F}(V \times W)$, we denote the equivalent class $\sum_{\text{finite}} c_i \delta_{v_i, w_i} + \mathrm{Rel}_F(V \times W)$ as:

$$\begin{aligned} \sum_{\text{finite}} c_i(v_i \times w_i) &= \sum_{\text{finite}} c_i v_i (1 \otimes w_i) \\ &= \sum_{\text{finite}} v_i \otimes w_i, \end{aligned}$$

where v_i is reassigned as $c_i v_i$. An element of the form $v \otimes w$ is referred to as a _pure tensor_. Both arbitrary elements of $V \otimes_F W$ and pure tensors admit the following properties:

(1) 1

(2) 2

(3) 3

(4) 4

Exercise 5.4.1. *Define* $\iota : V \times W \to V \otimes_F W$ *by* $(v, w) \mapsto v \otimes w$. *Show that* $\iota \in \mathrm{Hom}_F(V, W; V \otimes_F W)$.

Theorem 5.4.1. *Let* U, V, W *be* F*-vector spaces.*

(1) If $T \in \mathrm{Hom}_F(V \otimes_F W, U)$*, then* $T \circ \iota \in \mathrm{Hom}_F(V, W; U)$.

(2) If $t \in \mathrm{Hom}_F(V, W; U)$*, then there exists a unique* $T \in \mathrm{Hom}_F(V \otimes_F W, U)$ *so that* $t = T \otimes \iota$*; i.e., the following diagram commutes:*

$$\begin{tikzcd} V \times W \arrow[r, "\iota"] \arrow[dr, "t"'] & V \otimes_F W \arrow[d, "T"] \\ & U \end{tikzcd}$$

Proof. Let $T \in \mathrm{Hom}_F(V \otimes_F W, U)$ and set $t = T \otimes \iota$. Let $v_1, v_2 \in V$, $w \in W$, and $c \in F$. We have:

$$\begin{aligned} t(v_1 + cv_2, w) &= T(\iota(v_1 + cv_2, w)) \\ &= T((v_1 + cv_2) \otimes w) \\ &= T(v_1 \otimes w + c(v_2 \otimes w)) \\ &= T(v_1 \otimes w) + cT(v_2 \otimes w) \\ &= t(v_1, w) + ct(v_2, w). \end{aligned}$$

The same argument is true for the second variable. Thus $t \in \mathrm{Hom}_F(V, W; U)$.

Now let $t \in \mathrm{Hom}_F(V, W; U)$. This says that $t : V \times W \to U$. By the universal property of free vector spaces, t extends to a unique F-linear map $T : \mathbb{F}(V \times W) \to U$ that satisfies $t(v, w) = T(v, w)$. Taking the canonical projection $\pi : \mathbb{F}(V \times W) \to \mathbb{F}(V \times W)/\mathrm{Rel}_F(V \times W)$, it remains to show that T vanishes on $\mathrm{Rel}_F(V \times W)$. T does indeed vanish, and by diagram chasing we can show that $T(v, w) = t(v, w)$, establishing the proof. □

Exercise 5.4.2. *Show that* $\mathrm{Hom}_F(V, W; U) \cong \mathrm{Hom}_F(V \otimes_F W, U)$ *as* F*-vector spaces.*

Corollary 5.4.2. *Let* U, V, W *be* F*-vector spaces. We have:*

(1) $V \otimes_F W \cong W \otimes_F V$*;*

(2) $(U \otimes_F V) \otimes_F W \cong U \otimes_F (V \otimes_F W)$.

Proof. (1) is left an exercise. (2) Fix $w \in W$. Define: $t : U \times V \to U \otimes_F (V \otimes)$ □

buncha stuff to rewrite

5.5 The Trace

Given $A \in \mathrm{Mat}_n(F)$ with $A = (a_{ij})$, then $\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$. But in Chapter 3 we showed that $A = [T]_{\mathcal{B}}$ for some basis $\mathcal{B}$. We should be able to define the trace on linear maps so that:

$$\mathrm{tr}(T) = \mathrm{tr}([T]_{\mathcal{B}_1}) = \mathrm{tr}([T]_{\mathcal{B}_2})$$

for different bases $\mathcal{B}_1$ and $\mathcal{B}_2$. Normally trace is defined as $\mathrm{tr}(\cdot) : \mathrm{Mat}_n(F) \to F$. We want to "convert this" somehow to $\mathrm{tr}(\cdot) : \mathrm{Hom}_F(V,V) \to F$. Working with $\mathrm{Hom}_F(V,V)$ is rather tedious, so this will be our first obstacle to overcome.

Lemma 5.5.1. *Let V be a finite dimensional F-vector space. Then $V \otimes_F V^\vee \cong \mathrm{Hom}_F(V,V)$.*

Proof. Define:

$$t : V \times V^\vee \to \mathrm{Hom}_F(V,V)$$

by:

$$(v, \varphi) \mapsto (v' \mapsto \varphi(v') \cdot v).$$

Let $v_1, v_2 \in V$ and $c \in F$. We have:

$$\begin{aligned} t(v,\varphi)(v_1 + cv_2) &= \varphi(v_1 + cv_2)v \\ &= (\varphi(v_1) + c\varphi(v_2))v \\ &= \varphi(v_1)v + c\varphi(v_2)v \\ &= t(v,\varphi)(v_1) + ct(v,\varphi)(v_2). \end{aligned}$$

So for each $(v,\varphi) \in V \times V^\vee$, we have that $t(v,\varphi) \in \mathrm{Hom}_F(V,V)$. We'd like to now show that:

$$\begin{aligned} t(v_1 + cv_2, \varphi) &= t(v_1,\varphi) + ct(v_2,\varphi) \\ t(v, \varphi_1 + c\varphi_2) &= t(v,\varphi_1) + ct(v,\varphi_2). \end{aligned}$$

Since these are both functions, we need to show that they are equal for all $v' \in V$; i.e.:

$$t(v_1 + cv_2, \varphi)(v') = t(v_1,\varphi)(v') + ct(v_2,\varphi)(v').$$

Observe that:

$$\begin{aligned} t(v_1 + cv_2, \varphi)(v') &= \varphi(v')(v_1 + cv_2) \\ &= t(v_1,\varphi)(v') + ct(v_2,\varphi)(v'). \end{aligned}$$

Showing t is linear in the second variable is left as an exercise. This gives a well-defined linear map $\mathcal{T} : V \otimes_F V^\vee \to \mathrm{Hom}_F(V,V)$ defined by $v \otimes \varphi \mapsto (v' \mapsto \varphi(v')v)$. Since $\dim_F(V \otimes_F V^\vee) = \dim_F(\mathrm{Hom}_F(V,V)) = n^2$, it is enough to show that $\mathcal{T}$ is injective. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of V and $\mathcal{B}^\vee = \{v_1^\vee, \ldots, v_n^\vee\}$ its dual basis. Suppose $\mathcal{T}(\sum_{i,j} a_{ij}(v_i \otimes v_j^\vee)) = 0_{\mathrm{Hom}_F(V,V)}$. Take $v_m \in \mathcal{B}$. We have:

$$\begin{aligned} 0_V &= \mathcal{T}(\sum_{i,j} a_{ij}(v_i \otimes v_j^\vee))(v_m) \\ &= \sum_{i,j} a_{ij}(\mathcal{T}(v_i \otimes v_j^\vee)(v_m)) \\ &= \sum_{i,j} a_{ij} v_j^\vee(v_m) v_i \\ &= \sum_i a_{im} v_i. \end{aligned}$$

Thus $a_{im} = 0$ for all i. Since m was arbitrary, this gives that $a_{ij} = 0$ for all i, j. Thus $\mathcal{T}$ is injective, establishing $V \otimes_F V^\vee \cong \mathrm{Hom}_F(V,V)$. □

Proposition 5.5.2. *Let $C : \mathrm{Hom}_F(V,V) \times \mathrm{Hom}_F(V,V) \to \mathrm{Hom}_F(V,V)$ be defined by the usual function composition. There exists a unique $\Psi : (V \otimes_F V^\vee) \times (V \otimes_F V^\vee) \to V \otimes_F V^\vee$ making the following diagram commute:*

$$\begin{tikzcd} (V \otimes_F V^\vee) \times (V \otimes_F V^\vee) \arrow[r, dotted, "\Psi"] \arrow[d, "\mathcal{T}\times\mathcal{T}"] & V \otimes_F V^\vee \arrow[d, "\mathcal{T}"] \\ \mathrm{Hom}_F(V,V) \times \mathrm{Hom}_F(V,V) \arrow[r, "C"'] & \mathrm{Hom}_F(V,V) \end{tikzcd}$$

Proof. Define $\Psi : (V \otimes_F V^\vee) \times (V \otimes_F V^\vee) \to V \otimes_F V^\vee$ by $(v \otimes \varphi, w \otimes \psi) \mapsto \varphi(w) v \otimes \psi$. Given $x \in V$, from the following diagram chase:

$$\begin{tikzcd} (v \otimes \varphi, w \otimes \psi) \arrow[d, "\mathcal{T}\times\mathcal{T}"] & \\ (\mathcal{T}(v \otimes \varphi), \mathcal{T}(w \otimes \psi)) \arrow[r, "C"] & \mathcal{T}(v \otimes \varphi) \circ \mathcal{T}(w \otimes \psi), \end{tikzcd}$$

we can see that:

$$\begin{aligned} \mathcal{T}(v \otimes \varphi)(\mathcal{T}(w \otimes \psi)(x)) &= \mathcal{T}(v \otimes \varphi)(\psi(x)(w)) \\ &= \psi(x)\mathcal{T}(v \otimes \varphi)(w) \\ &= \psi(x)\varphi(w)v. \end{aligned}$$

Another diagram chase:

$$\begin{tikzcd} (v \otimes \varphi, w \otimes \psi) \arrow[r, "\Psi"] & \varphi(w)v \otimes \psi \arrow[d, "\mathcal{T}"] \\ & \mathcal{T}(\varphi(w)v \otimes \psi) \end{tikzcd}$$

gives:

$$\begin{aligned}\mathcal{T}(\varphi(w)v \otimes \psi)(x) &= \varphi(w)\mathcal{T}(v \otimes \psi)(x)\\ &= \varphi(w)\psi(x)v.\end{aligned}$$

Hence our diagram commutes. □

Remark. Note that $\Psi = \mathcal{T}^{-1} \circ \mathcal{C} \circ (\mathcal{T} \times \mathcal{T})$ might not necessarily be definable! When we constructed $\mathcal{T} : V \otimes_F V^\vee \to \mathrm{Hom}_F(V, V)$ we had to pick a basis for V. This is not an issue for vector-spaces, but had we been working over modules it may be impossible to express what $\mathcal{T}^{-1}$ is.

Remark. We also did this to show that $V \otimes_F V^\vee$ admits the same properties as Hom; i.e., that there is some semblance of "composition."

Definition 5.5.1. Let $T \in \mathrm{Hom}_F(V, V)$, $\mathcal{B} = \{v_1, .., v_n\}$ a basis of V, and $A = [T]_{\mathcal{B}}$. The *trace* of A is:

$$\begin{aligned}\mathrm{Tr}(A) &= \sum_{i=1}^{n} v_i^\vee(T(v_i))\\ &= \sum_{i=1}^{n} v_i^\vee\left(\sum_{j=1}^{n} a_{ij} v_j\right)\\ &= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij} v_i^\vee(v_j)\right)\\ &= \sum_{i=1}^{n} a_{ii}.\end{aligned}$$

Definition 5.5.2. Let $\mathcal{T} : V \otimes_F V^\vee \to \mathrm{Hom}_F(V, V)$ be defined by $v \otimes \varphi \mapsto \varphi(v')v$. Given $s : V \times V^\vee \to F$ defined by $(v, \varphi) \mapsto \varphi(v)$ and the induced map $S : V \otimes_F V^\vee \to F$ defined by $v \otimes \varphi \mapsto \varphi(v)$, the *trace* is $\mathrm{tr} = \mathcal{T}^{-1} \circ S$.

$$\begin{tikzcd}
\mathrm{Hom}_F(V, V) \arrow[r, "\mathrm{tr}"] & F \\
V \otimes_F V \arrow[u, "\mathcal{T}"] \arrow[ur, "S"'] &
\end{tikzcd}$$

Proposition 5.5.3. $\mathrm{tr} = \mathrm{Tr}$.

Proof. Define $\mathcal{T}(v_i \otimes v_j^\vee) = \mathcal{T}_{ij} \in \mathrm{Hom}_F(V, V)$. Since $\mathcal{T}$ is an isomorphism, $\{\mathcal{T}_{ij}\}_{i,j}$ is a basis of

$\mathrm{Hom}_F(V, V)$. Observe that:

$$\begin{aligned}
\mathrm{Tr}(\mathcal{T}_{kl}) &= \mathrm{Tr}(\mathcal{T}(v_k \otimes v_l^\vee)) \\
&= \sum_{i=1}^{n} v_i^\vee(\mathcal{T}(v_k \otimes v_l^\vee)(v_i)) \\
&= \sum_{i=1}^{n} v_i^\vee(v_l^\vee(v_i)v_k) \\
&= \sum_{i=1} v_l^\vee(v_i)v_i^\vee(v_k) \\
&= v_l^\vee v_k
\end{aligned}$$

$$\begin{aligned}
\mathrm{tr}(\mathcal{T}_{kl}) &= (S \circ \mathcal{T}^{-1})(\mathcal{T}_{kl}) \\
&= S(v_k \otimes v_l^\vee) \\
&= v_l^\vee(v_k)
\end{aligned}$$

Thus both of our definitions of trace are consistent. □

Corollary 5.5.4. *Let* $A, B \in \mathrm{Mat}_n(F)$. *We have* $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

Proof. Define:

$$\begin{aligned}
t_1 &: \mathrm{Hom}_F(V, V) \times \mathrm{Hom}_F(V, V) \to F \quad \text{by} \quad (S, T) \mapsto \mathrm{Tr}(S \circ T) \\
t_2 &: \mathrm{Hom}_F(V, V) \times \mathrm{Hom}_F(V, V) \to F \quad \text{by} \quad (S, T) \mapsto \mathrm{Tr}(T \circ S).
\end{aligned}$$

One can check that these are both bilinear. This induces:

$$\begin{aligned}
\tilde{t}_1 &: \mathrm{Hom}_F(V, V) \otimes_F \mathrm{Hom}_F(V, V) \to F \quad \text{by} \quad S \otimes T \mapsto \mathrm{Tr}(S \circ T) \\
\tilde{t}_2 &: \mathrm{Hom}_F(V, V) \otimes_F \mathrm{Hom}_F(V, V) \to F \quad \text{by} \quad S \otimes T \mapsto \mathrm{Tr}(T \circ S).
\end{aligned}$$

Let □

6
Tensor Algebras, Exterior Algebras, and the Determinant

6.1 Tensor Algebras