Contents

Contents i

1 Introduction 1
- 1.1 Basic Properties of Vector Spaces . . . 1

2 Bases and Dimension 6
- 2.1 Basic Definitions . . . 6
- 2.2 Every Vector Space Admits a Basis . . . 6
- 2.3 Cardinality and Dimension . . . 7
- 2.4 Direct Sums and Quotient Spaces . . . 11
- 2.5 Dual Spaces . . . 14

3 Linear Transformations and Matrices 17
- 3.1 Choosing Coordinates . . . 17
- 3.2 Row Operations . . . 22
- 3.3 Column-space and Null-space . . . 25
- 3.4 The Transpose of a Matrix . . . 26

4 Generalized Eigenvectors and Jordan Canonical Form 29
- 4.1 Diagonalization . . . 29
- 4.2 Eigenvalues and Eigenvectors . . . 31
- 4.3 Characteristic Polynomials . . . 39

Last update: 2024 October 5

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:

$$\begin{aligned} V \times V &\to B \\ (v, w) &\mapsto v + w \end{aligned}$$

called *vector addition* and

$$\begin{aligned} F \times V &\to V \\ (c, v) &\mapsto cv \end{aligned}$$

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.*

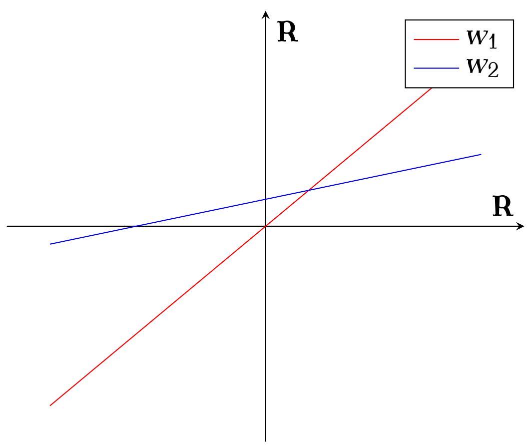
1. *The element $0_v \in V$ is unique,*
2. *$0v = 0_v$ for all $v \in V$,*
3. *$(-1_F)v = -v$ for all $v \in V$.*

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 $(-1_F)v = -v$. $\square$

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

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

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* $T \in \operatorname{Hom}_F(V, W)$*.* 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 $v_1 = v_2$. $\square$

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

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

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

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

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 $\mathcal{E}_n = \{e_1, \dots, e_n\}$ with

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

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

Proof. Let $X = \{\mathcal{B}' \subseteq V \mid \mathcal{A} \subseteq \mathcal{B}' \subseteq \mathcal{C},\ \mathcal{B}' \text{ is linearly independent}\}$. We have $\mathcal{A} \in X$, so $X \neq \emptyset$. X is ordered with respect to inclusion, and has an upperbound of $\mathcal{C}$. 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 $v \in \mathcal{C}$ with $v \notin \operatorname{span}_F(\mathcal{B})$. But then $\mathcal{B} \cup \{v\}$ is still linearly independent, and $\mathcal{B} \cup \{v\} \subseteq \mathcal{C}$. This gives $\mathcal{B} \subsetneq \mathcal{B} \cup \{v\}$, 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 $\mathcal{C}$, we can construct a basis $\mathcal{B}$ by removing elements.

Proof. Let $\mathcal{C} = \{w_1, \ldots, w_n\}$ with $n > m$. We will show $\mathcal{C}$ cannot be linearly independent. Write $\mathcal{B} = \{v_1, \ldots, v_m\}$ 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 $(x_1, \ldots, x_n) = (c_1, \ldots, c_n) \neq (0, \ldots, 0)$. 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 $\mathcal{C} = \{w_1, \ldots, w_n\}$ is not linearly independent. □

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

Proof. Let $|\mathcal{B}| = m$ and $|\mathcal{C}| = n$. Because $\operatorname{span}_F(\mathcal{B}) = V$ and $\mathcal{C}$ is linearly independent, it must be the case that $n \leqslant m$. But since $\operatorname{span}_F(\mathcal{C}) = V$ 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 <u>*dimension*</u> 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 $\mathcal{C} \subseteq V$ with $|\mathcal{C}| = m$.*

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

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

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

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

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 $\mathscr{B} = \{1\}$ is a basis and $\dim_{\mathbf{C}}(\mathbf{C}) = 1$.

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

(3) If $F = \mathbf{Q}$, then $|\mathscr{B}| = \mathfrak{c}$ 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 $\mathscr{B} = \{[1], [x], \dots, [x^{n-1}]\}$. We will show this is a basis for $F[x]/(f(x))$. Suppose there exists $a_0, \dots, a_{n-1} \in F$ with $a_0[1] + a_1[x] + \dots + a_{n-1}[x^{n-1}] = [0]$. So $[a_0 + a_1 x + \dots + a_{n-1}x^{n-1}] = [0]$, hence $f(x) \mid (a_0 + a_1 x + \dots + a_{n-1}x^{n-1})$. But $\deg(f(x)) = n$, so we must have $a_0 = a_1 = \dots = 0$ (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 $\mathscr{B}$, we have $[g(x)] \in \operatorname{span}_F(\mathscr{B})$. Note that any element of $\operatorname{span}_F(\mathscr{B})$ is clearly contained in $F[x]/(f(x))$, establishing $\operatorname{span}_F(\mathscr{B}) = F[x]/(f(x))$.

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

Proof. Suppose $\mathscr{C}$ 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 $\mathscr{C}$ 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 $\mathscr{C}$. Certainly we have $\operatorname{span}_F(\mathscr{C}) = V$. 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., $\mathscr{C}$ is linearly independent. $\square$

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 $\mathscr{B} = \{v_i\}_{i\in I}$ be a basis of V and $\mathscr{C} = \{w_i\}_{i\in I}$ be a subset of V. If $|\mathscr{B}| = |\mathscr{C}|$, 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). $\square$

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 $\mathcal{C} = \{w_i = T(v_i)\}_{i\in I}$ a subset of W. We have $\mathcal{C}$ is a basis of W if and only if T is an isomorphism.*

Proof. Suppose $\mathcal{C}$ 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:

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

hence $W = \mathrm{span}_F(\mathcal{C})$ (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 $\mathcal{C}$ is linearly independent. $\square$

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 $\mathcal{A} = \{v_1, \dots, v_k\}$ be a basis of $\ker(T)$. Extend this to a basis $\mathcal{B} = \{v_1, \dots, v_n\}$ of V. We'd like to show that $\mathcal{C} = \{T(v_{k+1}), \dots, T(v_n)\}$ 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:

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

Thus $\mathrm{span}_F(\mathcal{C}) = \mathrm{im}(T)$. 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 $a_1, \dots, a_k$ 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 $a_1 = \dots = a_n = 0$. $\square$

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 $\square$

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 $\square$

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 $\square$

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 $V_1, \dots, V_k$ be subspaces. The <u>sum</u> of $V_1, \dots, V_k$ is

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

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

Proof. do this $\square$

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

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

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

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

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

Proof. Suppose $V = V_1 \oplus \ldots \oplus V_k$. Let $v \in V$. Suppose $v = v_1 + \ldots + v_k = \tilde{v}_1 + \ldots + \tilde{v}_k$ for $v_i, \tilde{v}_i \in V_i$. Then $0_V = (v_1 - \tilde{v}_1) + \ldots + (v_k - \tilde{v}_k)$. Since $V_1, \ldots, V_k$ 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 $v = v_1 + \ldots + v_k$ with $v_i \in V_i$. Then $V = V_1 + \ldots + V_k$ by definition of sums of subspaces. If $0_V = v_1 + \ldots + v_k$ for some $v_i \in V_i$, and $0_v = 0_v + \ldots + 0_v$, then (by uniqueness) it must be the case that $v_i = 0_V$ for all i. □

Exercise 2.4.1. *Let $V_1, \ldots, V_k$ 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 $V = V_1 + \ldots + V_k$.

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

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

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 $v_1\ v_2$ 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 $\tau \in \operatorname{Hom}_F(\mathbf{R}, V/W)$. 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 $\tau^{-1} \in \operatorname{Hom}_F(V/W, \mathbf{R})$ 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 canonical projection map $\pi_W : V \to V/W$ is defined by $v \mapsto v + W$. Note that $\pi_W \in \operatorname{Hom}_F(V, V/W)$.

Note 1. 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 $T \in \operatorname{Hom}_F(V, W)$. Define $\overline{T} : V/\ker(T) \to W$ by $v + \ker(T) \mapsto T(v)$. Then $\overline{T}$ is a linear map. Moreover, $V/\ker(T) \cong \operatorname{im}(T)$.*

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 dual space, denoted $V^\vee$, is defined to be $V^\vee = \operatorname{Hom}_F(V, F)$.

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 $T \in \operatorname{Hom}_F(V, V^\vee)$ 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 $V \cong \operatorname{im}(T)$ (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 $\hat{v} \in \operatorname{Hom}_F(V^\vee, F) = (V^\vee)^\vee$. 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 $T \in \operatorname{Hom}_F(V, W)$. 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 $[v]_{\mathcal{B}} = T_{\mathcal{B}}(v) = \left(\begin{smallmatrix} a_1 \\ \vdots \\ a_n \end{smallmatrix}\right)$. We refer to this as *choosing coordinates* on V. asdf

Example 3.1.2.

(1) Let $V = \mathbf{Q}^2$ and $\mathcal{B} = \{\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)\}$. This forms a basis of V. Let $v \in V$ with $v = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$. 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 $\mathcal{E}_2 = \{\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)\}$, then $[v]_{\mathcal{E}_2} = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$.

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

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

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 $\mathcal{C} = \{w_1, \dots, w_m\}$ a basis for W. Define:

$$\begin{aligned} P &= T_{\mathcal{B}} : V \to F^n \text{ by } v \mapsto [v]_{\mathcal{B}} \\ Q &= T_{\mathcal{C}} : W \to F^m \text{ by } w \mapsto [w]_{\mathcal{C}} \end{aligned}$$

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 $[T]_{\mathcal{B}}^{\mathcal{C}} = A$, this is the unique matrix that satisfies $[T]_{\mathcal{B}}^{\mathcal{C}}[v]_{\mathcal{B}} = [T(v)]_{\mathcal{C}}$. Given $T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$, observe that:

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

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

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

Example 3.1.4.

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

$$\begin{aligned} T(1) &= 0 = 0 \cdot 1 + 0 \cdots x + 0 \cdot x^2 + 0 \cdot x^3 \\ T(x) &= 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 \\ T(x^2) &= 2x = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 \\ T(x^3) &= 3x^2 = 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2 + 0 \cdot x^3 \end{aligned}$$

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

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

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

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

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

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

Exercise 3.1.1.

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

$$[T \circ S]_{\mathcal{A}}^{\mathcal{C}} = [T]_{\mathcal{B}}^{\mathcal{C}} [S]_{\mathcal{A}}^{\mathcal{B}}.$$

(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 2. 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 $\mathcal{B} = \{v_1, \ldots, v_n\}$ and $\mathcal{B}' = \{v_1', \ldots, v_n'\}$. Define:

$$T : V \longrightarrow F^n \text{ by } v \longmapsto [v]_{\mathcal{B}}$$
$$S : V \longrightarrow F^n \text{ by } v \longmapsto [v]_{\mathcal{B}'}.$$

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 $\mathcal{B} = \{v_1, \ldots, v_n\}$. Show that $[\mathrm{id}_V]_{\mathcal{B}}^{\mathcal{B}'} = ([v_1]_{\mathcal{B}'} \mid \ldots \mid [v_n]_{\mathcal{B}'})$.*

Example 3.1.6.

(1) Let $V = \mathbf{Q}^2$ with $\mathcal{B} = \{e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$ and $\mathcal{B}' = \{v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}\}$. 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:

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

(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:

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

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:

$$T_A(e_1) = e_1 - 2e_2 + 3e_3$$
$$T_A(e_2) = 3e_1 - e_2 + 2e_3$$
$$T_A(e_3) = 3e_1 + 2e_2 + e_3.$$

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:

$$e_1 = -2v_1 + -3v_2 + v_3$$
$$e_2 = 3v_1 + 3v_2 - v_3$$
$$e_3 = -2v_1 - 2v_2 + v_3.$$

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:

$$T_A(e_1) = 3f_1 + f_2 + f_3$$
$$T_A(e_2) = 4f_1 + 2f_2 + f_3$$
$$T_A(e_3) = 5f_1 + 3f_2 + 2f_3$$
$$T_A(e_4) = 6f_1 + 4f_2 + 3f_3.$$

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 $-3R_1 + R_3 \leftrightarrow R_3$.

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

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

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

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 $\mathcal{C} = \{w_1, ..., w_n\}$ be a basis of W.

(1) Define $T_{i,j} : W \to W$ 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 $E_{i,j} = \left[T_{i,j}\right]_{\mathcal{C}}^{\mathcal{C}}$ corresponds to the identity matrix except the i^{th} and j^{th} rows are switched.

(2) Let $c \in F$, $c \neq 0$. Define $T_i^{(c)} : W \to W$ 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 $E_i^{(c)} = \left[T_i^{(c)}\right]_{\mathcal{C}}^{\mathcal{C}}$ corresponds to the identity matrix with the i^{th} row multiplied by c.

(3) Define $T_{i,j}^{(c)} : W \to W$ 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 $E_{i,j}^{(c)} = \left[T_{i,j}^{(c)}\right]_{\mathcal{C}}^{\mathcal{C}}$ 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:

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

$$\left[T_{1,3} \circ T_{A}{}^{\mathcal{F}_3}_{\mathcal{E}_4}\right] = \left[T_{1,3}\right]^{\mathcal{F}_3}_{\mathcal{F}_3} \left[T_A\right]^{\mathcal{F}_3}_{\mathcal{E}_4}$$

$$= E_{1,3}A$$

$$= \begin{pmatrix} 1 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 \\ 3 & 4 & 5 & 6 \end{pmatrix}.$$

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

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

3.3 Column-space and Null-space

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

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

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

(3) The *rank* 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 $\mathcal{E}_n = \{e_1, ..., e_n\}$ is the standard basis of F^n and $\mathcal{F}_n = \{f_1, ..., f_m\}$ is the standard basis of F^m. Since

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

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 $[E \circ T_A]^{\mathcal{F}_m}_{\mathcal{E}_n} = [E]^{\mathcal{F}_m}_{\mathcal{E}_n} A$ is in row echelon form. The column space of $[E \circ T_A]^{\mathcal{F}_m}_{\mathcal{E}_n}$ has as its basis the columns containing pivots (denoted $e_{i1}, ..., e_{ik}$):

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

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 $\mathcal{E}_n = \{e_1, \ldots, e_n\}$ and $\mathcal{F}_m = \{f_1, \ldots, f_m\}$ as standard bases. Then $A = [T_A]_{\mathcal{E}^n}^{\mathcal{F}_m}$, and furthermore $T_A \in \operatorname{Hom}_F(F^n, F^m)$ induces a dual map $T_A^\vee \in \operatorname{Hom}_F(F^{m\vee}, F^{n\vee})$. 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:

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

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:*

$$\begin{aligned} (A_1 + A_2)^t &= A_1^t + A_2^t \\ (cA_1)^t &= cA_1^t. \end{aligned}$$

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

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:

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

□

[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.

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:

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

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

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:

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

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 *T-invariant* or *T-stable* 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 $A = \left[T|_W\right]_{\mathcal{B}_W}$.

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 *eigenvalue* of T. Any nonzero vector in $\ker(T - \lambda \operatorname{id}_V)$ is called a *λ-eigenvector*. The set $E^1_\lambda = \ker(T - \lambda \operatorname{id}_V)$ is called the *eigenspace* associated with λ.

Exercise 4.2.1. *Show that E^1_λ 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 $E^1_{\lambda_1} \cap E^1_{\lambda_2} = \{0_V\}$.*

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

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

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 $\{\mathrm{id}_V, T, T^2, \ldots, T^{n^2}\} \subseteq \mathrm{Hom}_F(V, V)$. This has $n^2 + 1$ elements, so it must be linearly dependent. Let m be the smallest integer so that

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

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

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

[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.

This gives $m_T(T) = 0_{\text{Hom}_F(V,V)}$; 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 <u>*minimal polynomial*</u> 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 $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$. We can see that:

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

But $A^2 = \left(\begin{smallmatrix} 7 & 10 \\ 15 & 22 \end{smallmatrix}\right)$ 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 $W = \text{span}_{\mathbf{Q}}(e_1)$ 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, $m_{T_A}(x) = (x-1)^2(x+1)$.

Theorem 4.2.3. *Let V be an F-vector space and $T \in \text{Hom}_F(V,V)$. 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 $E^1_\lambda \neq \{0_V\}$ (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 $m_T(x) = x^m + ... + a_1x + a_0$. We have:

$$\begin{aligned} 0_V &= m_T(T)(v) \\ &= (T^m + a_{m-1}T^{m-1} + ... + a_1T + a_0 \operatorname{id}_V)(v) \\ &= T^m(v) + a_{m-1}T^{m-1}(v) + ... + a_1T(v) + a_0v \\ &= \lambda^m v + a_{m-1}\lambda^{m-1}v + ... + a_1\lambda v + a_0v \\ &= (\lambda^m + a_{m-1}\lambda^{m-1} + ... + a_1\lambda + a_0)v \\ &= m_T(\lambda) \cdot v. \end{aligned}$$

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

Proof. We have $m_T(x) = (x - \lambda_1)(x - \lambda_2)...(x - \lambda_m)f(x)$ for some $f(x) \in F[x]$. Suppose $a_1v_1 + ... + a_mv_m = 0_V$ for $a_i \in F$. Define $g_1(x) = (x - \lambda_2)...(x - \lambda_m)f(x)$. Note that $g_1(T)(v_i) = 0_V$ for $2 \leqslant i \leqslant m$. Then:

$$\begin{aligned} 0_V &= g_1(T)(0_V) \\ &= \sum_{j=1}^{m} a_j g_1(T)(v_j) \\ &= a_1g_1(T)(v_1) \\ &= a_1g_1(\lambda_1)v_1 \end{aligned}$$

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 $2, ..., m$. □

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 $v_3 = \left(\begin{smallmatrix} \frac{1}{2} \\ -2 \\ 1 \end{smallmatrix}\right)$. 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 $\ker(T^\cdot) = \ker(T^{m+1})$. 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 $T \in \operatorname{Hom}_F(V, V)$, $\lambda \in F$ 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 $E_\lambda^k = \operatorname{span}_F(v_1, \dots, v_k)$. Moreover, Lemma 4.2.6 gives $E_\lambda^1 \subseteq E_\lambda^2 \subseteq \dots \subseteq E_\lambda^k$.

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:

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

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

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:

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

Hence:

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

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 *characteristic polynomial* is $c_A(x) = \det(x1_n - A)$.

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

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

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 *T-annihilator* of v.

Example 4.3.1. Let $V = F^n$ and $\mathcal{E}_n = \{e_1, \ldots, e_n\}$. Define $T \in \operatorname{Hom}_F(V, V)$ 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 $T \in \operatorname{Hom}_{\mathbf{Q}}(\mathbf{Q}^2, \mathbf{Q}^2)$ 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:

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

[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 $F[x]$.

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* $T \in \mathrm{Hom}_F(V, V)$*. 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:

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

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* $m_{T,v} = p_1(x)p_2(x)...p_k(x)$.

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

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

Factor $m_T(x) = p_1(x)^{e_1}...p_k(x)^{e_k}$, with each $p_i(x)$ relatively prime and $e_1 \geqslant 1$. For $1 \leqslant j \leqslant k$, there exists $i_j \in \{1, ..., n\}$ 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 $w = w_1 + ... + w_k$. The previous result gives $m_{T,w}(x) = p_1(x)^{e_1}...p_k(x)^{e_k} = m_T(x)$. □