Introduction

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

Emil Artin

The goal of this course is to prove a lot of the essential results of linear algebra without basis dependence (as in, using the properties of the linear transformations themselves rather than matrices).

Contents

Vector Spaces

Vector Spaces and Linear Transformations

Remark: We let $\mathbb{F}$ be either $\mathbb{R}, \mathbb{Q}, \mathbb{C}, \mathbb{F}_p$ (where p is a prime). Primarily, we let $\mathbb{F} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$.

Example (Our First Vector Space)**.** The primary vector space we study in lower-division linear algebra is

$$\begin{aligned} V &= \mathbb{R}^n \\ &= \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \middle| a_1, \dots, a_n \in \mathbb{R} \right\} \end{aligned}$$

We know that for

$$v = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$
$$w = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

that

$$v + w = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}$$
$$cv = \begin{pmatrix} ca_1 \\ \vdots \\ ca_n \end{pmatrix},$$

where $c \in \mathbb{R}$ is some constant.

Definition (Vector Space)**.** Let V be a nonempty set with the following operations:

- $a : V \times V \to V$, $a(v, w) \mapsto v + w$ (vector addition);
- $m : F \times V \to V$, $m(c, v) \mapsto cv$ (scalar multiplication);

satisfying the following:

(1) there exists $0_v \in V$ such that $0_v + v = v = v + 0_v$ for all $v \in V$;

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

(3) for every $u, v, w \in V$, $(u + v) + w = u + (v + w)$;

(4) for every $v, w \in V$, $v + w = w + v$;

(5) for every $v, w \in V$ and $c \in \mathbb{F}$, $c(v + w) = cv + cw$;

(6) for every $c, d \in \mathbb{F}$, $v \in V$, $(c + d)v = cv + dv$;

(7) for every $c, d \in \mathbb{F}$, $v \in V$, $(cd)v = c(dv)$;

(8) for every $v \in V$, $(1_{\mathbb{F}})v = v$.

We say V is a $\mathbb{F}$-vector space.

Example ($\mathbb{F}^n$)**.** Let $\mathbb{F}$ be a field, $V = \mathbb{F}^n$.

$$V = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \middle| a_i \in \mathbb{F} \right\}.$$

Define:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}$$

$$c\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} ca_1 \\ \vdots \\ ca_n \end{pmatrix}.$$

We set

$$0_{\mathbb{F}^n} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Let

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$
$$w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$
$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix},$$

$c, d \in \mathbb{F}$. We observe that

$$\begin{aligned} 0_{\mathbb{F}^n} + v &= \begin{pmatrix} 0 + v_1 \\ \vdots \\ 0 + v_n \end{pmatrix} \\ &= \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}. \end{aligned}$$

Define

$$-v = \begin{pmatrix} -v_1 \\ \vdots \\ -v_n \end{pmatrix}.$$

Then,

$$\begin{aligned} v + (-v) &= \begin{pmatrix} v_1 + (-v_1) \\ \vdots \\ v_n + (-v_n) \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \\ &= 0_{\mathbb{F}^n}. \end{aligned}$$

Note that

$$(u + v) + w = \begin{pmatrix} (u_1 + v_1) + w_1 \\ \vdots \\ (u_n + v_n) + w_n \end{pmatrix}$$

$$\begin{aligned}
&= \begin{pmatrix} u_1 + (v_1 + w_1) \\ \vdots \\ u_n + (v_n + w_n) \end{pmatrix} \\
&= u + (v + w).
\end{aligned}$$

We have

$$\begin{aligned}
v + w &= \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix} \\
&= \begin{pmatrix} w_1 + v_1 \\ \vdots \\ w_n + v_n \end{pmatrix} \\
&= w + v.
\end{aligned}$$

Observe

$$\begin{aligned}
c\left(v + w\right) &= c\begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix} \\
&= \begin{pmatrix} c\left(v_1 + w_1\right) \\ \vdots \\ c\left(v_n + w_n\right) \end{pmatrix} \\
&= \begin{pmatrix} cv_1 + cw_1 \\ \vdots \\ cv_n + cw_n \end{pmatrix} \\
&= cv + cw, \\
(c + d)v &= (c + d)\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \\
&= \begin{pmatrix} (c + d)v_1 \\ \vdots \\ (c + d)v_n \end{pmatrix} \\
&= \begin{pmatrix} cv_1 + dv_1 \\ \vdots \\ cv_n + dv_n \end{pmatrix} \\
&= cv + dv,
\end{aligned}$$

and

$$\begin{aligned}
(cd)v &= (cd)\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \\
&= \begin{pmatrix} (cd)v_1 \\ \vdots \\ (cd)v_n \end{pmatrix}
\end{aligned}$$

$$
\begin{aligned}
&= \begin{pmatrix} c\left(dv_1\right) \\ \vdots \\ c\left(dv_n\right) \end{pmatrix} \\
&= c\left(dv\right).
\end{aligned}
$$

Finally,

$$
\begin{aligned}
1_{\mathbb{F}} &= 1_{\mathbb{F}} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \\
&= \begin{pmatrix} 1_{\mathbb{F}} v_1 \\ \vdots 1_{\mathbb{F}} \\ v_n \end{pmatrix} \\
&= \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \\
&= v.
\end{aligned}
$$

Example (Polynomials)**.** Let $n \in \mathbb{Z}_{\geqslant 0}$. We define

$$
P_n\left(\mathbb{F}\right) = \left\{a_0 + a_1 x + \cdots + a_n x^n \mid a_i \in \mathbb{F}\right\}.
$$

For $f(x) = \sum_{j=0}^{n} a_j x^j$ and $g(x) = \sum_{j=0}^{n} b_j x^j$ in $P_n(\mathbb{F})$, we have

$$
\begin{aligned}
f(x) + g(x) &= \sum_{j=0}^{n} \left(a_j + b_j\right) x^j \\
cf(x) &= \sum_{j=0}^{n} \left(c a_j\right) x^j.
\end{aligned}
$$

Note that these are not functions *per se*, we are only $f(x)$ and $g(x)$ to represent elements of $P_n\left(\mathbb{F}\right)$. We can verify that $P_n\left(\mathbb{F}\right)$ is a $\mathbb{F}$-vector space.

We define

$$
\mathbb{F}[x] = \bigcup_{n \geqslant 0} P_n\left(\mathbb{F}\right),
$$

which is also a $\mathbb{F}$-vector space.

Example (Matrices)**.** Let $m, n \in \mathbb{Z}_{>0}$. We set

$$
V = \mathrm{Mat}_{m,n}\left(\mathbb{F}\right),
$$

which is the set of $m \times n$ matrices with entries in $\mathbb{F}$. This is an $\mathbb{F}$-vector space with matrix addition and scalar multiplication.

In the case where $m = n$, we write $\mathrm{Mat}_n\left(\mathbb{F}\right)$ to denote $\mathrm{Mat}_{n,n}\left(\mathbb{F}\right)$.

Example (Complex Numbers)**.** Let $V = \mathbb{C}$. Then, V is a $\mathbb{C}$-vector space, an $\mathbb{R}$-vector space, and a $\mathbb{Q}$-vector space.

Note that the properties of a vector space change with the underlying scalar field.

Lemma (Basic Properties of Vector Spaces)**:** Let V be a $\mathbb{F}$-vector space.

(1) 0_V is unique.

(2) $0_{\mathbb{F}}v = 0_V$.

(3) $(-1_{\mathbb{F}})v = -v$.

Proof.

(1) Suppose toward contradiction that there exist $0, 0'$ both satisfy

$$0 + v = v \tag{$*$}$$
$$0' + v = v. \tag{$**$}$$

Then,

$$\begin{aligned} 0 + v &= v \\ 0 + 0' &= 0' && \text{by } (*) \text{ with } v = 0' \\ &= 0' + 0 \\ &= 0. && \text{by } (**) \text{ with } v = 0 \end{aligned}$$

(2) Note

$$\begin{aligned} 0_{\mathbb{F}}v &= (0_{\mathbb{F}} + 0_{\mathbb{F}})v \\ &= 0_{\mathbb{F}}v + 0_{\mathbb{F}}v. \end{aligned}$$

We subtract $0_{\mathbb{F}}v$ from both sides.

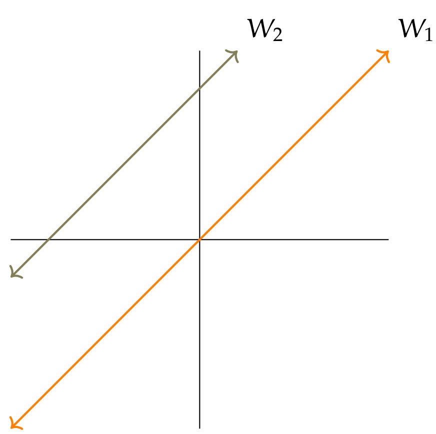
(3)

$$\begin{aligned} (-1_{\mathbb{F}})v + v &= (-1_{\mathbb{F}})v + 1_{\mathbb{F}}v \\ &= (-1_{\mathbb{F}} + 1_{\mathbb{F}})v \\ &= 0_{\mathbb{F}}v. \end{aligned}$$

□

Definition (Subspaces)**.** Let V be an $\mathbb{F}$-vector space. We say $W \subseteq V$ is an $\mathbb{F}$-subspace (henceforth subspace) if W is an $\mathbb{F}$-vector space under the same addition and scalar multiplication.

Example (Subspaces of $\mathbb{R}^2$)**.** Let $V = \mathbb{R}^2$.



Here, we see that W_1 is a subspace, and W_2 is not a subspace (as W_2 does not contain 0_V).

Example (Subspaces of $\mathbb{C}$). Let $V = \mathbb{C}$, $W = \{a + 0i \mid a \in \mathbb{R}\}$.

- If $\mathbb{F} = \mathbb{R}$, then W is a subspace of V.
- If $\mathbb{F} = \mathbb{C}$, then W is not a subspace; we can see that $2 \in W$, $i \in \mathbb{C}$, but $2i \notin W$.

Example (Matrices). It is not the case that $\text{Mat}_2(\mathbb{R})$ is a subspace of $\text{Mat}_4(\mathbb{R})$, since $\text{Mat}_2(\mathbb{R})$ is not a subset of $\text{Mat}_4(\mathbb{R})$.

Example (Polynomials). For the spaces $P_m(\mathbb{F})$ and $P_n(\mathbb{F})$, if $m \leqslant n$, then $P_m(\mathbb{F})$ is a subspace of $P_n(\mathbb{F})$.

Lemma (Proving Subspace Relation)**:** Let V be a $\mathbb{F}$-vector space, $W \subseteq V$. Then, W is a subspace of V if

(1) W is nonempty;

(2) W is closed under addition;

(3) W is closed under scalar multiplication.

Proof. The proof is an exercise. □

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

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

Note that on the left side, addition is in V, and on the right side, addition is in W.

The collection of all linear maps from V to W is denoted $\text{Hom}_{\mathbb{F}}(V, W)$, or $\mathcal{L}(V, W)$.

Example (Identity Transformation). Define

$$\text{id}_V : V \to V,$$

where $\text{id}_V(v) = v$. We can see that $\text{id}_V \in \text{Hom}_{\mathbb{F}}(V, V)$, since

$$\begin{aligned}\text{id}_V(v_1 + cv_2) &= v_1 + cv_2\\ &= \text{id}_V(v_1) + (c)(\text{id}_V(v_2))\end{aligned}$$

Example (Complex Conjugation). Let $V = \mathbb{C}$. Define $T : V \to V$ by $z \mapsto \overline{z}$.

We may ask whether $T \in \text{Hom}_{\mathbb{R}}(\mathbb{C}, \mathbb{C})$ or $T \in \text{Hom}_{\mathbb{C}}(\mathbb{C}, \mathbb{C})$.

$$\begin{aligned}T(z_1 + cz_1) &= \overline{z_1 + cz_2}\\ &= \overline{z_1} + (\overline{c})(\overline{z_2}).\end{aligned}$$

We can see that $T(z_1 + cz_2) = T(z_1)\, cT(z_2)$ if and only if $c = \overline{c}$, meaning c must be real. This means $T \in \text{Hom}_{\mathbb{R}}(\mathbb{C}, \mathbb{C})$, but $T \notin \text{Hom}_{\mathbb{C}}(\mathbb{C}, \mathbb{C})$.

Example (Matrices). Let $A \in \text{Mat}_{m,n}(\mathbb{F})$. We define

$$\begin{aligned}T_A : \mathbb{F}^n &\to \mathbb{F}^m\\ x &\mapsto Ax.\end{aligned}$$

Then, $T_A \in \text{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^m)$.

Example (Linear Maps on Smooth Functions). Let $V = C^\infty(\mathbb{R})$, which denotes the set of continuous functions with continuous derivatives at all orders. This is a vector space under pointwise addition and scalar multiplication.

$$\begin{aligned}(f + g)(x) &= f(x) + g(x)\\ (cf)(x) &= (c)(f(x)).\end{aligned}$$

Let $a \in \mathbb{R}$.

(1)
$$\begin{aligned} E_a : V &\to \mathbb{R} \\ f &\mapsto f(a). \end{aligned}$$
Then, $E_a \in \operatorname{Hom}_{\mathbb{R}}(V, \mathbb{R})$.

(2)
$$\begin{aligned} D : V &\to V \\ f &\mapsto f'. \end{aligned}$$
Then, $D \in \operatorname{Hom}_{\mathbb{R}}(V, V)$.

(3)
$$\begin{aligned} I_a : V &\to V \\ f &\mapsto \int_a^x f(t)\,dt. \end{aligned}$$
Then, $I_a \in \operatorname{Hom}_{\mathbb{R}}(V, V)$.

(4) Treating $f(a)$ as a (constant) function,
$$\begin{aligned} \tilde{E}_a : V &\to V \\ f &\mapsto f(a). \end{aligned}$$
Then, $\tilde{E}_a \in \operatorname{Hom}_{\mathbb{R}}(V, V)$.

Additionally,

- $D \circ I_a = \mathrm{id}_V$;
- $I_a \circ D = \mathrm{id}_V - \tilde{E}_a$ for some $a \in \mathbb{R}$.

Exercise: Show $\operatorname{Hom}_{\mathbb{F}}(V, W)$ is an $\mathbb{F}$-vector space.

Exercise: Let U, V, W be vector spaces. Let $S \in \operatorname{Hom}_{\mathbb{F}}(U, V)$ and $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$. Show $T \circ S \in \operatorname{Hom}_{\mathbb{F}}(U, W)$

Lemma (Image of Identity)**:** Let $T \in \operatorname{Hom}_{V,W}$. Then, $T(0_V) = 0_W$.

Definition (Isomorphism)**.** Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$ be invertible, meaning there exists $T^{-1}W \to V$ such that $T \circ T^{-1} = \mathrm{id}_W$ and $T^{-1} \circ T = \mathrm{id}_V$.

We say T is an isomorphism, and V, W are isomorphic.

Exercise: Show $T^{-1} \in \operatorname{Hom}_{\mathbb{F}}(W, V)$.

Example ($\mathbb{R}^2$ and $\mathbb{C}$)**.** Let $V = \mathbb{R}^2$, $W = \mathbb{C}$. Define $T : \mathbb{R}^2 \to \mathbb{C}$, $(x, y) \mapsto x + iy$.

We can verify that $T \in \operatorname{Hom}_{\mathbb{R}}\left(\mathbb{R}^2, \mathbb{C}\right)$. Let $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ and $r \in \mathbb{R}$. Then,
$$\begin{aligned} T\left((x_1, y_1) + r(x_2, y_2)\right) &= T\left((x_1 + rx_2, y_1 + ry_2)\right) \\ &= (x_1 + rx_2) + i(y_1 + ry_2) \\ &= x_1 + iy_1 + rx_2 + i(ry_2) \\ &= x_1 + iy_1 + r(x_2 + iy_2) \\ &= T\left((x_1, y_1)\right) + rT\left((x_2, y_2)\right). \end{aligned}$$

Define $T^{-1}\mathbb{C} \to \mathbb{R}^2$ by $x + iy \mapsto (x, y)$. We have $T \circ T^{-1}(x + iy) = x + iy$ is an inverse map and $T^{-1} \circ T((x, y)) = (x, y)$. Thus, $\mathbb{R}^2 \cong \mathbb{C}$ as $\mathbb{R}$-vector spaces.

Example ($P_n(\mathbb{F})$ and $\mathbb{F}^{n+1}$). Set $V = P_n(\mathbb{F})$ and $W = \mathbb{F}^{n+1}$.

Define $T : P_n(\mathbb{F}) \mapsto \mathbb{F}^{n+1}$,

$$a_0 + a_1 x + \cdots + a_n x^n \mapsto \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

We can verify that T is linear, with inverse map $T^{-1} : \mathbb{F}^{n+1} \to P_n(\mathbb{F})$

$$\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto a_0 + a_1 x + \cdots + a_n x^n.$$

Thus, $P_n(\mathbb{F}) \cong \mathbb{F}^{n+1}$.

Definition (Kernel)**.** Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$. Define

$$\ker(T) = \{v \in V \mid T(v) = 0_W\}.$$

We call this the kernel of T.

Definition (Image)**.** Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$. Define

$$\begin{aligned} \operatorname{im}(T) &= T(V) \\ &= \{w \in W \mid \exists v \in V \text{ such that } T(v) = w\} \end{aligned}$$

Lemma (Kernel and Image are Subspaces)**:** The kernel, $\ker(T)$, is a subspace of V, and the image, $\operatorname{im}(T)$, is a subspace of W.

Proof. Since $T(0_V) = 0_W$, we know that both $\ker(T)$ and $\operatorname{im}(T)$ are nonempty.

Let $c \in \mathbb{F}$ and $v_1, v_2 \in \ker(T)$. Then,

$$\begin{aligned} T(v_1 + cv_2) &= T(v_1) + cT(v_2) \\ &= 0. \end{aligned}$$

Thus, $v_1 + cv_2 \in \ker(T)$.

Let $w_1, w_2 \in \operatorname{im}(T)$. Then, there exist $u_1, u_2 \in V$ such that $T(u_1) = w_1$ and $T(u_2) = w_2$. We have

$$\begin{aligned} T(u_1 + cu_2) &= T(u_1) + cT(u_2) \\ &= w_1 + cw_2, \end{aligned}$$

meaning $w_1 + cw_2 \in \operatorname{im}(T)$, meaning $\operatorname{im}(T)$ is a subspace of W. □

Lemma (Injectivity of a Linear Transformation)**:** T is injective and only if $\ker(T) = \{0_V\}$.

Proof. Suppose T is injective. Let $v \in V$ be such that $T(v) = 0_W$. We also know that $T(0_V) = 0_W$. Since T is injective, this means $v = 0_V$.

Let $\ker(T) = \{0_V\}$. Suppose $T(v_1) = T(v_2)$. Then,

$$\begin{aligned} T(v_1) - T(v_2) &= 0_W \\ T(v_1 - v_2) &= 0_W, \end{aligned}$$

meaning $v_1 - v_2 \in \ker(T)$, meaning $v_1 - v_2 = 0_V$. Thus, $v_1 = v_2$. □

Example (Projection Map)**.** Let $m > n$. Define $T : \mathbb{F}^m \to \mathbb{F}^n$ by

$$\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \mapsto \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

We can see that $\operatorname{im}(T) = \mathbb{F}^n$.

To examine the kernel, let

$$\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \in \ker(T).$$

Then,

$$\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

with n entries. Thus,

$$\begin{aligned} \ker(T) &= \left\{ \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ a_{n+1} \\ \vdots \\ a_m \end{pmatrix} \middle| a_i \in \mathbb{F}^m \right\} \\ &\cong \mathbb{F}^{m-n}. \end{aligned}$$

Bases and Dimension

For this section, we let V be a $\mathbb{F}$-vector space.

Definition (Linear Combination)**.** Let $\mathcal{B} = \{v_i\}_{i\in I}$ be a subset of V. We say $v \in V$ is an $\mathbb{F}$-linear combination of $\mathcal{B}$ if there is a set $\{a_i\}_{i\in I}$ with $a_i = 0$ for all but finitely many i such that

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

We write $v \in \operatorname{span}_{\mathbb{F}}(\mathcal{B})$.

Example. Let $V = P_2(\mathbb{F})$. Set $\mathcal{B} = \left\{1, x, x^2\right\}$. We have $\operatorname{span}_{\mathbb{F}}(\mathcal{B}) = P_2(\mathbb{F})$.

Definition (Linear Independence)**.** Let $\mathcal{B} = \{v_i\}_{i\in I}$ be a subset of V. We say $\mathcal{B}$ is $\mathbb{F}$-linearly independent if whenever

$$\sum_{i\in I} a_i v_i = 0_V,$$

we have $a_i = 0$ for all $i \in I$. Note that these are finite sums.

Definition (Hamel Basis)**.** Let $\mathcal{B} = \{v_i\}_{i\in I}$ be a subset of V. We say $\mathcal{B}$ is a $\mathbb{F}$-basis for V if

(1) $\operatorname{span}(\mathcal{B}) = V$

(2) $\mathcal{B}$ is linearly independent.

Example (Standard Basis for $\mathbb{F}^n$)**.** Let $V = \mathbb{F}^n$. We let

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

where

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$
$$e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$
$$\vdots$$
$$e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

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

We wish to show that every vector space has a basis. In order to do so, we require Zorn's lemma.

Theorem (Zorn's Lemma)**:** Let X be a nonempty partially ordered set. If every totally ordered subset of X has an upper bound, then there exists at least one maximal element in X.

Theorem: Let $\mathcal{A}$ and C be subsets of V with $\mathcal{A} \subseteq C$. Assume $\mathcal{A}$ is linearly independent and $\text{span}_{\mathbb{F}}(C) = V$. Then, there exists a basis $\mathcal{B}$ of V with $\mathcal{A} \subseteq \mathcal{B} \subseteq C$.

Proof. Take

$$X = \left\{\mathcal{B}' \subseteq V \mid \mathcal{A} \subseteq \mathcal{B}' \subseteq C, \mathcal{B} \text{ linearly independent}\right\}.$$

We have $\mathcal{A} \in X$, meaning X is nonempty. We know that X is partially ordered with respect to inclusion, and has an upper bound of C.

Thus, by Zorn's lemma, we have a maximal element in X. We call this maximal element $\mathcal{B}$. By the definition of X, $\mathcal{B}$ is linearly independent.

We claim that $\text{span}_{\mathbb{F}}(\mathcal{B}) = V$. If not, there exists some $v \in C$ such that $v \notin \text{span}_{\mathbb{F}}(\mathcal{B})$. However, if $v \notin \text{span}_{\mathbb{F}}(\mathcal{B})$, then $\mathcal{B} \cup \{v\} \subseteq C$ is linearly independent. However, since $\mathcal{B} \subsetneq \mathcal{B} \cup \{v\}$, this implies that $\mathcal{B}$ is not maximal, which is a contradiction. Thus, $\text{span}_{\mathbb{F}}(\mathcal{B}) = V$. □

Remark: This proof applies to all vector spaces, not just those with finite dimensions.

Lemma: A homogeneous system of m linear equations in n unknowns with $m < n$ has a nonzero solution.

Corollary: Let $\mathcal{B} \subseteq V$ with $\text{span}_{\mathbb{F}}(\mathcal{B}) = V$ and $|\mathcal{B}| = m$.

Then, any set with more than m elements cannot be linearly independent.

Proof. Let $\mathcal{C} = \{w_1, \dots, w_n\}$ with $n > m$. We wish to show that $\mathcal{C}$ cannot be linearly independent.

Write $\mathcal{B} = \{v_1, \dots, v_m\}$ with $\operatorname{span}_{\mathbb{F}}(\mathcal{B}) = V$. For each i, write $w_i = \sum_{j=1}^{m} a_{ji} v_j$ for some $a_{ji} \in \mathbb{F}$.

Consider the equations

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

We have a solution to this $(c_1, \dots, c_n) \neq (0, \dots, 0)$.

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, $\mathcal{C}$ is not linearly independent. □

Corollary: If $\mathcal{B}$ and $\mathcal{C}$ are bases over V, with $\mathcal{B}$ and $\mathcal{C}$ finite, then $\operatorname{card} \mathcal{B} = \operatorname{card} \mathcal{C}$.

Proof. Let $|\mathcal{B}| = m$, $|\mathcal{C}| = n$. Since $\mathcal{C}$ is linearly independent, we know that $n \leqslant m$. We reverse the roles to see that $m \leqslant n$. □

Definition (Dimension)**.** Let V be a $\mathbb{F}$-vector space with Hamel basis $\mathcal{B}$. Then, we define $\dim_{\mathbb{F}} V = \operatorname{card} \mathcal{B}$.

Theorem: Let V be finite-dimensional with $\dim_{\mathbb{F}} V = n$. Let $\mathcal{C} \subseteq V$ with $\operatorname{card} \mathcal{C} = m$.

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

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

(3) If $m = n$, then the following are equal:
   - $\mathcal{C}$ is a basis;
   - $\mathcal{C}$ is linearly independent;
   - $\operatorname{span}_{\mathbb{F}}(\mathcal{C}) = V$.

Corollary: Let $W \subseteq V$ be a subspace. We have $\dim_{\mathbb{F}} W \leqslant \dim_{\mathbb{F}} V$.

If $\dim_{\mathbb{F}} V < \infty$, then $V = W$ if and only if $\dim_{\mathbb{F}} W = \dim_{\mathbb{F}} V$.

Example. Let $V = \mathbb{C}$.

If $\mathbb{F} = \mathbb{C}$, then $\mathcal{B} = \{1\}$, and $\dim_{\mathbb{C}} \mathbb{C} = 1$.

If $\mathbb{F} = \mathbb{R}$, then $\mathcal{B} = \{1, i\}$, and $\dim_{\mathbb{R}} \mathbb{C} = 2$.

Example. Let $V = \mathbb{F}[x]$, and let $f(x) \in \mathbb{F}[x]$ be fixed.

Define an equivalence relation $g(x) \equiv h(x)$ if $f(x) \,|\, (g(x) - h(x))$.

Given $g(x) \in \mathbb{F}[x]$, write $[g(x)]$ for the equivalence class containing $g(x)$.

Define $W = \mathbb{F}[x]/(f(x)) = \{[g(x)] \mid g(x) \in \mathbb{F}[x]\}$.

Define

$$\begin{aligned}[g(x)] + [h(x)] &= [g(x) + h(x)]\\ c[g(x)] &= [cg(x)].\end{aligned}$$

This makes W into a vector space. Set $n = \deg f(x)$.

Then, we claim

$$\mathcal{B} = \left\{[1], [x], \dots, \left[x^{n-1}\right]\right\}.$$

Suppose there exist $a_0, \dots, a_{n-1} \in \mathbb{F}$ with

$$a_0[1] + a_1[x] + \cdots + a_{n-1}\left[x^{n-1}\right] = [0].$$

Then,

$$\left[a_0 + a_1x + \cdots + a_{n-1}x^{n-1}\right] = [0].$$

Therefore,

$$f(x) | \left(a_0 + a_1x + \cdots + a_{n-1}x^{n-1} - 0\right),$$

which means we must have $a_0 = a_1 = \cdots = a_{n-1}$.

Let $[g(x)] \in W$. By the Euclidean algorithm,

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

for some $q(x), r(x) \in \mathbb{F}[x]$ with $r(x) = 0$ or $\deg r(x) < n$. Thus, we have

$$\begin{aligned}[g(x)] &= [f(x)q(x)] + [r(x)]\\ &= [r(x)].\end{aligned}$$

Since $r(x) = 0$ or $\deg r(x) < n$, we must have $[g(x)] = [r(x)] \in \operatorname{span}_{\mathbb{F}}(\mathcal{B})$.

Lemma: Let V be an $\mathbb{F}$-vector space, with $\mathcal{C} = \{v_i\}_{i \in I}$ be a subset of V.

Then, $\mathcal{C}$ is a basis if and only if each $v \in V$ can be uniquely written as a linear combination of elements of $\mathcal{C}$.

Proof. Suppose $\mathcal{C}$ is a basis. Let $v \in V$, and suppose

$$\begin{aligned}v &= \sum_{i \in I} a_i v_i\\ &= \sum_{i \in I} b_i v_i\end{aligned}$$

for some $a_i, b_i \in \mathbb{F}$. Then,

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

Since C is a basis, $a_i - b_i = 0$ for all i, meaning $a_i = b_i$, so the expression is unique.

Suppose every v can be written as a unique linear combination of C. Certainly, this means $\operatorname{span}_{\mathbb{F}}(C) = V$. Suppose

$$0_V = \sum_{i \in I} a_i v_i$$

for some $a_i \in \mathbb{F}$. It is also true that $0_V = \sum_{i \in I} 0 v_i$, meaning $a_i = 0$ for all i by uniqueness; thus, C is linearly independent. □

Proposition: Let V, W be $\mathbb{F}$-vector spaces.

(1) Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$. We have T is uniquely determined by the image of the basis of V.

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

Proof.

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

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

(2) Define T by setting

$$T(v) = \sum_{i \in I} a_i w_i,$$

for $v = \sum_{i \in I} a_i v_i$. We can verify that T is linear.

□

Corollary: Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$, with $\mathcal{B} = \{v_i\}$ a basis of V and $C = \{w_i\} \subseteq W$, with $w_i = T(v_i)$. Then, we have C is a basis of W if and only if T is an isomorphism.

Proof. Let C be a basis for W. Since C is a basis of W, we use the proposition to define $S \in \operatorname{Hom}_{\mathbb{F}}(W, V)$ with $S(w_i) = v_i$. We can verify that $T \circ S = \mathrm{id}_W$ and $S \circ T = \mathrm{id}_V$, meaning $S = T^{-1}$ and T is an isomorphism.

Suppose T is an isomorphism. Let $w \in W$. Since T is an isomorphism, T is surjective, meaning there exists $v \in V$ such that $T(v) = w$. Since $\mathcal{B}$ is a basis of V, we expand v to have

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

Combining these two facts, we have

$$\begin{aligned} w &= T(v) \\ &= T\left(\sum_{i \in I} a_i v_i\right) \\ &= \sum_{i \in I} a_i T(v_i) \\ &\in \operatorname{span}_{\mathbb{F}}(C). \end{aligned}$$

Thus, $W = \operatorname{span}_{\mathbb{F}}(C)$.

Suppose there exists $a_i \in \mathbb{F}$ with $\sum_{i\in I} a_i T(v_i) = 0_W$. Since T is linear, we have

$$\sum_{i\in I} a_i T(v_i) = T\left(\sum_{i\in I} a_i v_i\right).$$

Since T is injective, we have

$$\sum_{i\in I} a_i v_i = 0_V.$$

Since $\mathcal{B}$ is a basis, we have $a_i = 0$. □

Theorem (Rank–Nullity)**:** Let V be finite-dimensional vector space over $\mathbb{F}$. Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$. Then,

$$\dim_{\mathbb{F}}(V) = \dim_{\mathbb{F}}(\ker(T)) + \dim_{\mathbb{F}}(\operatorname{im}(T))$$

Proof. Let $\dim_{\mathbb{F}}(\ker(T)) = k$ and $\dim_{\mathbb{F}}(V) = n$. Let $\mathcal{A} = \{v_1, \dots, v_k\}$ be a basis of $\ker(T)$. We extend $\mathcal{A}$ to a basis $\mathcal{B} = \{v_1, \dots, v_n\}$ of V.

We want to show that $C = \{T(v_{k+1}), \dots, T(v_n)\}$ is a basis of $\operatorname{im}(T)$.

Let $w \in \operatorname{im}(T)$. Then, there is $v \in V$ such that $T(v) = w$. We write

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

meaning

$$\begin{aligned} w &= T(v) \\ &= T\left(\sum_{i=1}^{n} a_i v_i\right) \\ &= \sum_{i=1}^{n} a_i T(v_i) \\ &= \sum_{i=k+1}^{n} a_i T(v_i) \\ &\in \operatorname{span}_{\mathbb{F}}(C), \end{aligned}$$

since $\{v_1, \dots, v_k\} \subseteq \ker(T)$, meaning $\operatorname{span}_{\mathbb{F}}(C) = \operatorname{Im}(T)$.

Suppose we have

$$\sum_{i=k+1}^{n} a_i T(v_i) = 0_W.$$

Then, we have

$$T\left(\sum_{i=k+1}^{n} a_i v_i\right) = 0_W,$$

meaning $\sum_{i=k+1}^{n} a_i v_i \in \ker(T)$. This means there exist $a_1, \dots, a_k$ such that

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

meaning

$$\sum_{i=1}^{k} a_i v_i + \sum_{i=k+1}^{n} (-a_i) v_i = 0_V.$$

Since $\{v_i\}$ are a basis, this means $a_i = 0$ for all i. $\square$

Corollary: Let V, W be $\mathbb{F}$-vector spaces with $\dim_{\mathbb{F}}(V) = n$. Let $V_1 \subseteq V$ be a subspace with $\dim_{\mathbb{F}}(V_1) = k$, and $W_1 \subseteq W$ a subspace with $\dim_{\mathbb{F}}(W_1) = n - k$. Then, there exists $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$ such that $\ker(T) = V_1$ and $\operatorname{im}(T) = W_1$.

Corollary: Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$ with $\dim_{\mathbb{F}}(V) = \dim_{\mathbb{F}}(W) < \infty$. Then, the following are equivalent:

(1) T is an isomorphism;

(2) T is injective;

(3) T is surjective.

Corollary: Let $A \in \operatorname{Mat}_n(\mathbb{F})$. The following are equivalent:

(1) A is invertible;

(2) There exists $B \in \operatorname{Mat}_n(\mathbb{F})$ such that $BA = I_n$;

(3) There exists $B \in \operatorname{Mat}_n(\mathbb{F})$ such that $AB = I_n$.

Corollary: Let $\dim_{\mathbb{F}}(V) = m$ and $\dim_{\mathbb{F}}(W) = n$.

(1) If $m < n$ and $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$, then T is not surjective.

(2) If $m > n$ and $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$, then T is not injective.

(3) We have $m = n$ if and only if $V \cong W$.

Direct Sums and Quotient Spaces

Definition (Sum of Subspaces)**.** Let V be a vector space, and $V_1, \dots, V_k$ be subspaces. Then, the sum of $V_1, \dots, V_k$ is

$$V_1 + \cdots + V_k = \left\{ \sum_{i=1}^{k} v_i \mid v_i \in V_i \right\}.$$

This is a subspace of V.

Definition (Independence of Subspaces)**.** Let $V_1, \dots, V_k$ be subspaces of V. We say $V_1, \dots, V_k$ are independent if whenever $v_1 + \cdots v_k = 0_V$, we have $v_i = 0_V$.

Definition (Direct Sum of Subspaces)**.** 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

$$V = V_1 \oplus \cdots \oplus V_k,$$

if the following conditions hold.

(1) $V = V_1 + \cdots V_k$;

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

Example (A Very Simple Direct Sum). Let $V = \mathbb{F}^2$, with $V_1 = \{(x, 0) \mid x \in \mathbb{F}\}$ and $V_2 = \{(0, y) \mid y \in \mathbb{F}\}$, we can see that

$$\begin{aligned} V_1 + V_2 &= \{(x,0) + (0,y) \mid x, y \in \mathbb{F}\} \\ &= \{(x,y) \mid x, y \in \mathbb{F}\} \\ &= \mathbb{F}^2. \end{aligned}$$

If $(x, 0) + (0, y) = 0$, then $x = 0$ and $y = 0$, meaning $\mathbb{F}^2 = V_1 \oplus V_2$.

Example (Direct Sum Constructions). Let $V = \mathbb{F}[x]$.

Define $V_1 = \mathbb{F}$, $V_2 = \mathbb{F}x = \{\alpha x \mid \alpha \in \mathbb{F}\}$, $V_3 = P_1(\mathbb{F})$.

We can see that

$$P_1 = V_1 \oplus V_2.$$

However, V_1 and V_3 are not independent, since $1_{\mathbb{F}} \in V_1$ and $-1_{\mathbb{F}} \in V_3$ with $1_{\mathbb{F}} + (-1_{\mathbb{F}}) = 0_{\mathbb{F}}$.

Example. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of V, with $V_i = \operatorname{span}(v_i)$. Then,

$$V = V_1 \oplus \cdots \oplus V_n.$$

Lemma: Let V be a vector space, $V_1, \ldots, V_k$ subspaces. We have $V = V_1 \oplus \cdots \oplus V_k$ if and only if every $v \in V$ can be written uniquely in the form

$$v = v_1 + \cdots + v_k$$

for $v_i \in V_i$.

Proof. Suppose $V = V_1 \oplus \cdots \oplus V_k$. Let $v \in V$. Then, $v = v_1 + \cdots + v_k$ for some $v_i \in V_i$ since $V = V_1 + \cdots + V_k$. Suppose

$$\begin{aligned} v &= v_1 + \cdots v_k \\ &= \tilde{v}_1 + \cdots + \tilde{v}_k \end{aligned}$$

for $v_i, \tilde{v}_i \in V_i$. Then,

$$0_V = (v_1 - \tilde{v}_1) + \cdots + (v_k - \tilde{v}_k).$$

Since $V_1, \ldots, V_k$ are linearly independent, $v_i - \tilde{v}_i \in V_i$, we have $v_i - \tilde{v}_i = 0_V$, meaning the expression for v is unique.

Suppose that every $v \in V$ can be written uniquely in the form $v = v_1 + \cdots + v_k$ with $v_i \in V_i$. Then,

$$V = V_1 + \cdots V_k$$

by the definition of $V_1 + \cdots + V_k$. If

$$0_V = v_1 + \cdots v_k$$

for $v_i \in V_i$, and it is also the case that

$$0_V = 0_V + \cdots + 0_V,$$

with $0_V \in V_i$, then it must be the case that $v_i = 0_V$ for all i by uniqueness. Thus, the V_i are independent, so

$$V = V_1 \oplus \cdots \oplus V_k.$$

$\square$

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

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

(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 $V = V_1 \oplus \cdots \oplus V_k$.

Lemma (Existence of Complement)**:** Let V be a vector space, and $U \subseteq V$ be a subspace. Then, U has a complement W such that $U \oplus W = V$.

Proof. Let $\mathcal{A}$ be a basis for U. Extend $\mathcal{A}$ to a basis $\mathcal{B}$ of V. Let $\mathcal{C} = \mathcal{B} \setminus \mathcal{A}$, and $W = \operatorname{span}(\mathcal{C})$. □

Example (Constructing a Quotient Group)**.** To introduce quotient spaces, consider the construction of the quotient group.

Let $n \in \mathbb{Z}_{>1}$. We say $a \equiv b$ modulo n if and only if $n|(a-b)$. This is an equivalence relation; we form $\mathbb{Z}/n\mathbb{Z} = \{[a]_n \mid a \in \mathbb{Z}\} = \{[0]_n, \dots, [n-1]_n\}$.

However, we also do this by defining $n\mathbb{Z} = \{nk \mid k \in \mathbb{Z}\}$, and taking $a \equiv b \bmod n$ if and only if $a - b \in n\mathbb{Z}$. Our equivalence classes are now

$$\begin{aligned}[a]_n &= \{a + nk \mid k \in \mathbb{Z}\} \\ &= a + n\mathbb{Z}.\end{aligned}$$

Definition (Quotient Space)**.** Let $W \subseteq V$ be a subspace. We say $v_1 \sim v_2$ if $v_1 - v_2 \in W$. Note that if $w \in W$, then $w \sim 0_V$ since $w - 0_V \in W$.

This is an equivalence relation.

- Reflexivity: since W is a subspace, $0_V \in W$, meaning $v - v \in W$ for all $v \in V$.
- Symmetry: if $v_1 \sim v_2$, then $v_1 - v_2 \in W$, meaning $-(v_1 - v_2) \in W$, so $v_2 - v_1 \in W$, or $v_2 \sim v_1$.
- Transitivity: Let $v_1 \sim v_2$ and $v_2 \sim v_3$. Then, $v_1 - v_2 \in W$ and $v_2 - v_3 \in W$. Since W is a subspace, $(v_1 - v_2) + (v_2 - v_3) \in W$, meaning $v_1 - v_3 \in W$, so $v_1 \sim v_3$.

We denote the equivalence classes by

$$\begin{aligned}[v] &= [v]_W \\ &= v + W \\ &= \{\tilde{v} \in V \mid v \sim \tilde{v}\} \\ &= \{v + w \mid w \in W\}.\end{aligned}$$

We set

$$V/W := \{v + W \mid v \in V\}.$$

We need to define vector addition and scalar multiplication on V/W. Let $v_1 + W, v_2 + W \in V/W$ and $c \in \mathbb{F}$. Define

$$\begin{aligned}(v_1 + W) + (v_2 + W) &= (v_1 + v_2) + W \\ c(v_1 + W) &= cv_1 + W.\end{aligned}$$

We will show that addition and scalar-multiplication are well-defined.

Addition: Let $v_1 + W = \tilde{v}_1 + W$, $v_2 + W = \tilde{v}_2 + W$, meaning $v_1 = \tilde{v}_1 + w_1$ and $v_2 = \tilde{v}_2 + w_2$ for some $w_1, w_2 \in W$. We have

$$\begin{aligned}(v_1 + W) + (v_2 + W) &= (v_1 + v_2) + W \\ &= (\tilde{v}_1 + w_1 + \tilde{v}_2 + w_2) + W \\ &= (\tilde{v}_1 + \tilde{v}_2) + W\end{aligned}$$

Scalar Multiplication: Let $v + W = \tilde{v} + W$. Then, we have $v = \tilde{v} + w$ for some $w \in W$. For $c \in \mathbb{F}$, we have

$$\begin{aligned} c\left(v + W\right) &= cv + W \\ &= c\left(\tilde{v} + w\right) + W \\ &= c\tilde{v} + W \\ &= c\left(\tilde{v} + W\right). \end{aligned}$$

We say V/W is the quotient space of V by W.

Example (Quotient Space of $\mathbb{R}^2$)**.** Let $V = \mathbb{R}^2$, and $W = \{(x, 0) \mid x \in \mathbb{R}\}$.

Let $(x_0, y_0) \in V$. We have

$$(x_0, y_0) \sim (x, y)$$

if

$$(x_0 - x, y_0 - y) \in W.$$

The only condition is thus that the y-coordinates in $\mathbb{R}^2$ must be equal. Therefore,

$$(x_0, y_0) + W = \{(x, y_0) \mid x \in \mathbb{R}\}.$$

Define $\tau : \mathbb{R} \to V/W$, $y \mapsto (0, y) + W$. We claim that τ is an isomorphism.

Let $y_1, y_2, c \in \mathbb{R}$. We have

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

Thus, we see that τ is a linear map.

To show surjectivity, let $(x, y) + W \in V/W$. We have $(x, y) + W = (0, y) + W$. Thus, τ is surjective, since

$$\begin{aligned} \tau\left(y\right) &= (0, y) + W \\ &= (x, y) + W. \end{aligned}$$

Finally, to show injectivity, we let $y \in \ker\left(\tau\right)$. We have

$$\begin{aligned} \tau\left(y\right) &= (0, y) + W \\ &= (0, 0) + W, \end{aligned}$$

implying that $y = 0$. Thus, τ is injective.

Example (Quotient Space of Polynomials)**.** Let $V = \mathbb{F}[x]$, $f(x) \in V$, and

$$W = \{g(x) \in \mathbb{F}[x] \mid f(x) \mid g(x)\}.$$

We can see that W is a subspace, which we refer to as $\langle f(x)\rangle$.

We defined an equivalence class $g(x) \sim h(x)$ if $f(x) | \left(g(x) - h(x)\right)$, where we then constructed a vector space from this set.

In particular, this construction is realized as V/W.[1]

[1]The ramifications of this construction are covered in depth in Algebra II.

Definition (Canonical Projection)**.** Let $W \subseteq V$ be a subspace. The canonical projection map π_W is defined by

$$\begin{aligned} \pi_W : V &\rightarrow V/W \\ v &\mapsto v + W. \end{aligned}$$

Note that $\pi_W \in \operatorname{Hom}_{\mathbb{F}}(V, V/W)$.

Remark: To define a map $T : V/W \rightarrow U$, one must always verify that T is well-defined.

Theorem (First Isomorphism Theorem for Vector Spaces)**:** Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$. Define $\overline{T} : V/\ker(T) \rightarrow W$ by taking $v + \ker(T) \mapsto T(v)$. Then, $\overline{T} \in \operatorname{Hom}_{\mathbb{F}}(V/\ker(T), W)$. Moreover, $V/\ker(T) \cong \operatorname{im}(T)$.

Proof. We will first show that $\overline{T}$ is well-defined. Let $v_1 + \ker(T) = v_2 + \ker(T)$. Then, for some $\tilde{v} \in \ker(T)$, we have $v_1 = v_2 + \tilde{v}$. Then,

$$\begin{aligned} \overline{T}(v_1 + \ker(T)) &= T(v_1) \\ &= T(v_2 + \tilde{v}) \\ &= T(v_2) + T(\tilde{v}) \\ &= T(v_2) \\ &= \overline{T}(v_2 + \ker(T)). \end{aligned}$$

Let $v_1 + \ker(T), v_2 + \ker(T) \in V/\ker(T)$, and $c \in \mathbb{F}$. Then, we have

$$\begin{aligned} \overline{T}((v_1 + \ker(T)) + c(v_2 + \ker(T))) &= \overline{T}((v_1 + cv_2) + \ker(T)) \\ &= T(v_1 + cv_2) \\ &= T(v_1) + cT(v_2) \\ &= \overline{T}(v_1 + \ker(T)) + c\overline{T}(v_2 + \ker(T)). \end{aligned}$$

Let $w \in \operatorname{im}(T)$. Then, $w = T(v)$ for some $v \in V$, meaning

$$\begin{aligned} w &= T(v) \\ &= \overline{T}(v + \ker(T)). \end{aligned}$$

Thus, $\overline{T}$ is surjective onto $\operatorname{im}(T)$.

Let $v + \ker(T) \in \ker\left(\overline{T}\right)$. Then,

$$\overline{T}(v + \ker(T)) = 0_W.$$

This gives

$$T(v) = 0_W,$$

meaning $v \in \ker(T)$, meaning $v + \ker(T) = 0_V + \ker(T)$. Thus, $\overline{T}$ is injective. $\square$

Dual Spaces

Definition (Dual Space)**.** Let V be an $\mathbb{F}$-vector space. The dual space, V',[II] is defined to be

$$V' := \operatorname{Hom}_{\mathbb{F}}(V, \mathbb{F}).$$

[II]My professor denotes this as $V^{\vee}$, but it's too hard to type that out in real time, so I will use the $'$ to denote the algebraic dual, just as V^* denotes the continuous dual of V.

Theorem: We have V is isomorphic to a subspace of V'. If $\dim_{\mathbb{F}}(V) < \infty$, then $V \cong V'$.

Remark: The isomorphism between V and V' in the finite-dimensional case is not canonical — that is, it depends on a basis.

Proof. Let $\mathcal{B} = \{v_i\}_{i\in I}$ be a basis for V.

For each $i \in I$, let $v_i'(v_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta. We get $\left\{v_i'\right\}_{i\in I}$ are elements of V'. We obtain

$$T \in \mathrm{Hom}_{\mathbb{F}}(V, V')$$

by $T(v_i) = v_i'$.

To show V is isomorphic to a subspace of V', it suffices to show that T is injective, since $V \cong \mathrm{im}(T)$, which is a subspace of V'.

Let $v \in V$ with $T(v) = 0_{V'}$. We write

$$\begin{aligned} v &= \sum_{i\in I} a_i v_i \\ 0_{V'} &= T(v) \\ &= \sum_{i\in I} a_i T(v_i) \\ &= \sum_{i\in I} a_i v_i'. \end{aligned}$$

Pick j with $a_j \neq 0$. Note that

$$\begin{aligned} \sum_{i\in I} a_i v_i'(v_j) &= 0 \\ &= a_j, \end{aligned}$$

which contradicts $a_j \neq 0$. Thus, $v = 0_V$, and T is injective.

Suppose $\dim_{\mathbb{F}}(V) = n$, with $\mathcal{B} = \{v_1, \dots, v_n\}$. Let $\in V'$. Define a_i by

$$a_i = (v_i).$$

Set

$$v = \sum_{i=1}^{n} a_i v_i.$$

Define the map $S : V' \to V$ by taking

$$S() = \sum_{i=1}^{n} (v'(v_i))\, v_i.$$

We want to show that $S \in \mathrm{Hom}_{\mathbb{F}}(V', V)$, and S is the inverse to T.

Let $, w' \in V'$, $c \in \mathbb{F}$. Set $a_i = v'(v_i)$ and $b_i = w'(v_i)$. Then,

$$S(+cw') = \sum_{i=1}^{n} (v'cw')(v_i)\, v_i$$

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

We compute $S \circ T(v_i)$.

$$
\begin{aligned}
S \circ T\left(v_j\right) &= S\left(T\left(v_j\right)\right) \\
&= S\left(v_j'\right) \\
&= \sum_{i=1}^{n} v_j'(v_i) v_i \\
&= \sum_{i=1}^{n} \delta_{ij} v_i \\
&= v_j.
\end{aligned}
$$

Note that for $T \circ S$, we have $T \circ S$ maps V' to V', meaning we need to check that $T \circ S$ is the identity map on V'. Let $\in V'$. Then,

$$
\begin{aligned}
(T \circ S)()\left(v_j\right) &= T\left(S\left(v'\right)\right)\left(v_j\right) \\
&= T\left(\sum_{i=1}^{n} (v_i) v_i\right)\left(v_j\right) \\
&= \left(\sum_{i=1}^{n} (v_i) T(v_i)\right)\left(v_j\right) \\
&= \sum_{i=1}^{n} (v_i)\left(v_i'\left(v_j\right)\right) \\
&= \sum_{i=1}^{n} (v_i) \delta_{ij} \\
&= \left(v_j\right).
\end{aligned}
$$

□

Definition (Dual Basis)**.** Let $\mathcal{B} = \{v_1, \dots, v_n\}$ be a basis of V. The dual basis for V' is

$$
\mathcal{B}' = \left\{v_i', \dots, v_n'\right\}.
$$

Remark: It is possible to continue taking duals; in the case of finite-dimensional V, we have

$$
\begin{aligned}
V &\cong V' \\
V' &\cong V''.
\end{aligned}
$$

Despite the isomorphism between V and V' not being canonical, it is the case that the isomorphism between V and V'' *is* canonical (i.e., not dependent on a basis).

Proposition: There is a canonical injective linear map from V to V''. If $\dim_{\mathbb{F}}(V) < \infty$, this is an isomorphism.

Proof. Let $v \in V$. Define $\hat{v} : V' \to \mathbb{F}$, $\varphi \mapsto \varphi(v)$.[III] We can easily verify that $\hat{v}$ is a linear map.

[III] This can be notated as eval_v, but $\hat{v}$ is faster to type (and it's used in functional analysis).

Therefore, we have $\hat{v} \in \operatorname{Hom}_{\mathbb{F}}(V', \mathbb{F}) = V''$. We have a map

$$\begin{aligned} \Phi : V &\rightarrow V'' \\ v &\mapsto \hat{v}. \end{aligned}$$

We want to verify that Φ is a linear and injective map. Let $v_1, v_2 \in V$, $c \in \mathbb{F}$. Let $\varphi \in V'$.

$$\begin{aligned} \Phi(v_1 + cv_2)(\varphi) &= (\hat{v}_1 + c\hat{v}_2)(\varphi) \\ &= \varphi(v_1 + cv_2) \\ &= \varphi(v_1) + c\varphi(v_2) \\ &= \hat{v}_1(\varphi) + c\hat{v}_2(\varphi) \\ &= \Phi(v_1)(\varphi) + c\Phi(v_2)(\varphi). \end{aligned}$$

We will show that Φ is injective. Let $v \in V$; suppose $v \neq 0_V$. We form a basis $\mathcal{B}$ of V that contains v. Note that $\in V'$, with $v'(v) = 1$ and $v'(w) = 0$ for $w \in \mathcal{B}$ and $w \neq v$.

Assume $v \in \ker(\Phi)$. Then, for any $\varphi \in V'$,

$$\begin{aligned} \Phi(v)(\varphi) &= 0 \\ \varphi(v) &= 0. \end{aligned}$$

However, this is a contradiction, as we can take $\varphi =$, where $\varphi(v) = 1$. Thus, it must be the case that Φ is injective. □

Definition (Dual Operator)**.** Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$. We get an induced map $T' : W' \rightarrow V'$. We define $T'(\varphi) = \varphi \circ T$.

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

Choosing Coordinates

Linear Transformations and Matrices

Let V be a finite-dimensional $\mathbf{F}$-vector space. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis. This vector space fixes an isomorphism $V \cong \mathbf{F}^n$.

Let $v \in V$. We can write $v = \sum_{i=1}^{n} a_i v_i$ for some $a_i \in \mathbb{F}$. We take the map

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

It is easy to see that T is an isomorphism. Given $v \in V$, we write $[v]_{\mathcal{B}} = T_{\mathcal{B}}(v)$. We refer to this process as choosing coordinates.

Example. Let $V = \mathbb{Q}^2$, and $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$. We can check that $\mathcal{B}$ is a basis of V.

Let $v \in V$, $v = \begin{pmatrix} a \\ b \end{pmatrix}$. We have

$$v = \frac{a+b}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{a-b}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

To represent v in terms of this basis, we have

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

If we chose a different basis, such as the standard basis $\mathcal{E}_2 = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$. In that case, we have

$$[v]_{\mathcal{E}_2} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Example. Let $V = P_2(\mathbb{R})$. Let $\mathcal{C} = \left\{1, (x-1), (x-1)^2\right\}$. We know that $\mathcal{C}$ is a basis of V.

Let $f(x) = a + bx + cx^2 \in P_2(\mathbb{R})$. We can write f in terms of this basis by taking

$$f(x) = (a+b+c) + (b+2c)(x-1) + c(x-1)^2.$$

In this case, we then have

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

Recall that given $A \in \mathrm{Mat}_{m,n}(\mathbb{F})$, we obtain a linear map $T_A \in \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^m)$ by $T_A(v) = Av$. The converse is true as well. Given any map $T \in \mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^m)$, there is a matrix A such that $T = T_A$.

Let $\mathcal{E}_n = \{e_1, \dots, e_n\}$ be the standard basis of $\mathbb{F}^n$ and $\mathcal{F}_m = \{f_1, \dots, f_m\}$ be the standard basis of $\mathbb{F}^m$.

We have $T(e_j) \in \mathbb{F}^m$ for each j, meaning we have $a_{ij} \in \mathbb{F}$ with $T(e_j) = \sum_{i=1}^m a_{ij} f_j$.

Define $A = (a_{ij})_{ij} \in \mathrm{Mat}_{m,n}(\mathbb{F})$. We want to show that $T_A(e_j) = T(e_j)$ for every j.

Then, we have

$$\begin{aligned} T_A(e_j) &= Ae_j \\ &= \sum_{a_{ij}} f_i \\ &= T(e_j). \end{aligned}$$

Let $T \in \mathrm{Hom}_{\mathbb{F}}(V, W)$. Let $\mathcal{B} = \{v_1, \dots, v_n\}$ be a basis for V and $\mathcal{C} = \{w_1, \dots, w_m\}$ be a basis for W.

Define $P = T_{\mathcal{B}} : V \to \mathbb{F}^n$, $v \mapsto [v]_{\mathcal{B}}$, $Q = T_{\mathcal{C}} : W \to \mathbb{F}^m$, $w \mapsto [w]_{\mathcal{C}}$. This yields the following diagram:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "T_{\mathcal{B}}"'] & W \arrow[d, "T_{\mathcal{C}}"] \\ \mathbb{F}^n \arrow[r, "T_{\mathcal{C}} \circ T \circ T_{\mathcal{B}}^{-1}"'] & \mathbb{F}^m \end{tikzcd}$$

In particular, this means T is given by a matrix $A \in \mathrm{Mat}_{m,n}(\mathbb{F})$, which we write as $[T]_{\mathcal{B}}^{\mathcal{C}} = A$.

In particular, $[T]_{\mathcal{B}}^{\mathcal{C}}$ is the unique matrix that satisfies

$$[T]_{\mathcal{B}}^{\mathcal{C}}\left([v]_{\mathcal{B}}\right) = [T(v)]_{\mathcal{C}}.$$

To compute $[T]_{\mathcal{B}}^{\mathcal{C}}$, we have

$$\begin{aligned}
T\left(v_j\right) &= \sum_{i=1}^{m} a_{ij} w_i && a_{ij} \in \mathbb{F} \\
\left[T\left(v_j\right)\right]_{\mathcal{C}} &= \left[\sum_{i=1}^{m} a_{ij} w_j\right]_{\mathcal{C}} \\
&= \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.
\end{aligned}$$

Similarly, since $[v]_{\mathcal{B}} = e_j$, we have

$$\begin{aligned}
[T]_{\mathcal{B}}^{\mathcal{C}}\left(e_j\right) &= \left[T\left(v_j\right)\right]_{\mathcal{C}} \\
&= \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix},
\end{aligned}$$

which is exactly the jth column of $[T]_{\mathcal{B}}^{\mathcal{C}}$.

We thus get a matrix of the form

$$[T]_{\mathcal{B}}^{\mathcal{C}} = \begin{pmatrix} [T(v_1)]_{\mathcal{C}} & \cdots & [T(v_n)]_{\mathcal{C}} \end{pmatrix},$$

where $\left[T\left(v_j\right)\right]_{\mathcal{C}}$ are column vectors.

Example. Let $V = P_3(\mathbb{R})$. Define $T \in \operatorname{Hom}_{\mathbb{R}}(V, V)$ by $T(f(x)) = f'(x)$.

We take $\mathcal{B} = \left\{1, x, x^2, x^3\right\}$ as our basis. Then, we have

$$\begin{aligned}
T(1) &= 0 \\
T(x) &= 1 \\
T\left(x^2\right) &= 2x \\
T\left(x^3\right) &= 3x^2.
\end{aligned}$$

As we fill in our matrix, we have

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

We can view each column as a basis vector of $\mathcal{B}$ and each row as the corresponding representation in $\mathcal{C}$ (where, in this case, $\mathcal{C} = \mathcal{B}$).

Example. Let $V = P_3(\mathbb{R})$, $T(f(x)) = f'(x)$. Let $\mathcal{B} = \left\{1, x, x^2, x^3\right\}$ and $\mathcal{C} = \left\{1, (x-1), (x-1)^2, (x-1)^3\right\}$.

$$\begin{aligned}
T(1) &= 0 \\
T(x) &= 1 \\
T\left(x^2\right) &= 2x = 2 + 2(x-1)
\end{aligned}$$

$$T\left(x^3\right) = 3x^2 = -9 - 6(x-1) + 3(x-1)^2.$$

Thus, our matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$ is

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

Exercise:

(1) Let $\mathcal{A}$ be a basis of U, $\mathcal{B}$ a basis of V, and $\mathcal{C}$ a basis of W. Let $S \in \operatorname{Hom}_{\mathbb{F}}(U, V)$ and $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$.

Show that

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

(2) We know that given $A \in \operatorname{Mat}_{m,k}(\mathbb{F})$ and $B \in \operatorname{Mat}_{n,m}(\mathbb{F})$, we have corresponding T_A and T_B linear maps.

Show that you recover the definition of matrix multiplication by using Part 1 to define matrix multiplication.

Note: To refer to $[T]_{\mathcal{B}}^{\mathcal{B}}$, we will write $[T]_{\mathcal{B}}$.

Let V be a vector space, with $\mathcal{B}$ and $\mathcal{B}'$ bases of V. We want to be able to transfer information about V in terms of $\mathcal{B}$ to information about V in terms of $\mathcal{B}'$ (i.e., change the basis).[IV]

Let $\mathcal{B} = \{v_1, \dots, v_n\}$ and $\mathcal{B}' = \left\{v_1', \dots, v_n'\right\}$. Define

$$\begin{aligned} T : V &\to \mathbb{F}^n \\ v &\mapsto [v]_{\mathcal{B}} \\ S : V &\to \mathbb{F}^n \\ v &\mapsto [v]_{\mathcal{B}'}. \end{aligned}$$

In terms of a diagram, we have

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

In particular, the change of basis matrix is

$$[\mathrm{id}_V]_{\mathcal{B}}^{\mathcal{B}'}.$$

Exercise: Let $\mathcal{B} = \{v_1, \dots, v_n\}$. Show that

$$[\mathrm{id}_V]_{\mathcal{B}}^{\mathcal{B}'} = \begin{pmatrix} [v_1]_{\mathcal{B}'} & \cdots & [v_n]_{\mathcal{B}'} \end{pmatrix}.$$

Example. Let $V = \mathbb{Q}^2$, $\mathcal{B} = \mathcal{E}_2 = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$. Let

$$\mathcal{B}' = \left\{ v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}.$$

Notice that

$$e_1 = \frac{1}{2}v_1 + \frac{1}{2}v_2$$

[IV]Note that $\mathcal{B}'$ does not refer to the algebraic dual.

$$e_2 = -\frac{1}{2}v_1 + \frac{1}{2}v_2.$$

In particular, we have

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

Thus,

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

Let

$$v = \left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right).$$

We have

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

Example. Let $V = P_2(\mathbb{R})$, $\mathcal{B} = \left\{1, x, x^2\right\}$, $\mathcal{B}' = \left\{1, (x-2), (x-2)^2\right\}$.

We have

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

Thus, we have

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

Therefore,

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

For example, if we let $f(x) = -7 + 3x + 4x^2$, we have

$$\begin{aligned}
[f(x)]_{\mathcal{B}} &= \begin{pmatrix} -7 \\ 3 \\ 4 \end{pmatrix}\\
[f(x)]_{\mathcal{B}'} &= [\mathrm{id}_V]_{\mathcal{B}}^{\mathcal{B}'}[f(x)]_{\mathcal{B}}\\
&= \begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -7 \\ 3 \\ 4 \end{pmatrix}\\
&= \begin{pmatrix} 15 \\ 19 \\ 4 \end{pmatrix}
\end{aligned}$$

meaning

$$f(x) = 15 + 19\left(x-2\right) + 4\left(x-2\right)^2.$$

Exercise (Group Work)**:** Let $V = P_2(\mathbb{R})$, $\mathcal{B} = \left\{1, (x-1), (x-1)^2\right\}$ and $\mathcal{B}' = \left\{1, (x+1), (x+1)^2\right\}$. Find the change of basis matrix, and find $\left[2 - 6(x-1) + 2(x-1)^2\right]_{\mathcal{B}'}$.

Solution: We have

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

Thus, the change of basis matrix is

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

Thus, we have

$$\begin{aligned}
\left[2 - 6(x-1) + 2(x-1)^2\right]_{\mathcal{B}'} &= \begin{pmatrix} 1 & -2 & 4 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ -6 \\ 2 \end{pmatrix}\\
&= \begin{pmatrix} 22 \\ -14 \\ 2 \end{pmatrix}
\end{aligned}$$

Definition (Similar Matrices)**.** Given $A, B \in \mathrm{Mat}_n(\mathbb{F})$, we say A and B are similar if there exists $P \in \mathrm{GL}_n(\mathbb{F})$[v] such that $A = PBP^{-1}$.

We wish to rephrase this definition in terms of matrices. Given $A \in \mathrm{Mat}_n(\mathbb{F})$, there exists $T_A \in \mathrm{Hom}_{\mathbb{F}}(F^n, H^n)$ with $T_A(v) = Av$. Given a basis $\mathcal{B}$, we have the following diagram:

$$\begin{array}{ccc}
\mathbb{F}^n & \xrightarrow{T_A} & \mathbb{F}^n \\
{\scriptstyle T_{\mathcal{B}}}\downarrow & & \downarrow{\scriptstyle T_{\mathcal{B}}} \\
\mathbb{F}^n & \xrightarrow[{[T_A]_{\mathcal{B}}}]{} & \mathbb{F}^n
\end{array}$$

If $\mathcal{E}_n$ is the standard basis, then $A = [T_A]_{\mathcal{E}_n}$, meaning we have the following diagram:

[v] $\mathrm{GL}_n(\mathbb{F}) = \left\{C \in \mathrm{Mat}_n(\mathbb{F}) \mid C^{-1} \text{ exists}\right\}$

$$\begin{tikzcd}
\mathbb{F}^n \arrow[r, "\mathrm{id}_{\mathbb{F}^n}"] \arrow[d, "T_{\mathcal{B}}"'] & \mathbb{F}^n \arrow[r, "T_A"] \arrow[d, "T_{\mathcal{E}_n}"'] & \mathbb{F}^n \arrow[r, "\mathrm{id}_{\mathbb{F}^n}"] \arrow[d, "T_{\mathcal{E}_n}"] & \mathbb{F}^n \arrow[d, "T_{\mathcal{B}}"] \\
\mathbb{F}^n \arrow[r, "{P^{-1}=[\mathrm{id}_{\mathbb{F}^n}]_{\mathcal{B}}^{\mathcal{E}_n}}"'] & \mathbb{F}^n \arrow[r, "A"'] & \mathbb{F}^n \arrow[r, "{P^{-1}=[\mathrm{id}_{\mathbb{F}^n}]_{\mathcal{E}_n}^{\mathcal{B}}}"'] & \mathbb{F}^n
\end{tikzcd}$$

Thus, $A = P\,[T_A]_{\mathcal{B}}\,P^{-1}$. In other words, $A \sim B$ if and only if $A = [T_A]_{\mathcal{B}}$ for some basis $\mathcal{B}$ and $B = [T_A]_{\mathcal{C}}$.

Row Operations, Column Space, and Null Space

Definition (Pivot)**.** Let $A = \left(a_{ij}\right) \in \mathrm{Mat}_{m,n}(\mathbb{F})$. We say $a_{k\ell}$ is a pivot of A if and only if $a_{k\ell} \neq 0$ and $a_{ij} = 0$ if $i \geqslant k$ or $j \leqslant \ell$, with $(i,j) \neq (k,\ell)$.

Example. For the matrix

$$A = \begin{pmatrix} \boxed{2} & 1 & 4 & 5 \\ 0 & 0 & \boxed{1} & 7 \\ 0 & 0 & 0 & \boxed{5} \end{pmatrix},$$

the boxed entries are pivots.

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

Example. We have

$$A = \begin{pmatrix} 2 & 1 & 4 & 5 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

is in row echelon form, and

$$B = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Example. Let

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

We are going to put this matrix into reduced row echelon form. We have $T_A : \mathbb{F}^4 \to \mathbb{F}^3$. Let $\mathcal{E}_4 = \{e_1, e_2, e_3, e_4\}$ and $\mathcal{F}_3 = \{f_1, f_2, f_3\}$. Then, $A = [T_A]_{\mathcal{E}_4}^{\mathcal{F}_3}$. We have

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

Step 1: We switch $R_1 \leftrightarrow R_3$, yielding

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

yielding

$$[T_A]_{\mathcal{E}_4}^{\mathcal{F}_3^{(2)}} = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 \\ 3 & 4 & 5 & 6 \end{pmatrix}$$

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

Step 2: Our next step is $-R_1 + R_2 \to R_2$, yielding

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

Our new matrix is

$$[T_A]_{\mathcal{E}_4}^{\mathcal{F}_3^{(3)}} = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 1 \\ 3 & 4 & 5 & 6 \end{pmatrix}$$

$$\begin{aligned} T_A(e_1) &= \left(f_1^{(2)} + f_2^{(2)}\right) + 3f_3^{(2)} \\ &= f_1^{(3)} + 3f_3^{(3)} \\ T_A(e_2) &= \left(f_1^{(2)} + f_2^{(2)}\right) + f_2^{(2)} + 4f_3^{(2)} \\ &= f_1^{(3)} + f_2^{(2)} + 4f_3^{(3)} \\ &\vdots \end{aligned}$$

Step 3: Next, we have $-3R_1 + R_3 \to R_3$, which yields

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

Our matrix is now

$$[T_A]_{\mathcal{E}_4}^{\mathcal{F}_3^{(4)}} = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & -1 & -3 \end{pmatrix}$$

Step 4: Next, we have $-R_2 + R_3 \to R_3$, which yields

$$\mathcal{F}_3^{(5)} = \left\{ f_1^{(5)} = f_1^{(4)}, f_2^{(5)} = f_2^{(4)} + f_3^{(4)}, f_3^{(5)} = f_3^{(4)} \right\},$$

and a matrix of

$$[T_A]_{\mathcal{E}_4}^{\mathcal{F}_3^{(5)}} = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -2 & -4 \end{pmatrix}.$$

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

- switching two rows: $R_i \leftrightarrow R_j$;
- multiplying a row by a scalar: $\mathbb{R}_i \to cR_i$;
- replacing a row by adding a scalar multiple of another row: $aR_i + R_j \to R_j$.

Sketch of a Proof. For any matrix, we switch rows such that the value of a_{11} is nonzero. Then, we take

$$\begin{aligned} f_1^{(2)} &= \sum_{j=1}^{m} a_{ji} f_j \\ f_k^{(2)} &= f_k. \end{aligned}$$

□

Instead of directly changing the bases, we can use linear maps to change the bases.

We define $T_{i,j} : W \to W$ to be

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

Thus,

$$E_{i,j} = \left[T_{i,j}\right]_{\mathcal{C}}^{\mathcal{C}}$$

is the identity matrix except for switching the i and j rows.

Let $c \in \mathbb{F}$, define $T_i^{(c)} : W \to W$ by

$$\begin{aligned} T_i^{(c)}(w_k) &= w_k && k \neq i \\ T_i^{(c)}(w_i) &= cw_i, \end{aligned}$$

with

$$E_i^{(c)} = \left[T_i^{(c)}\right]_{\mathcal{C}}^{\mathcal{C}}$$

being the identity matrix except for row i multiplied by c.

Finally, we define $T_{i,j}^{(c)} : W \to W$ by

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

with

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

as the identity map with c in the ijth entry.

Example. Let

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

Define $T_A : \mathbb{F}^4 \to \mathbb{F}^3$, $\mathcal{E}_4 = \{e_1, e_2, e_3, e_4\}$, and $\mathcal{F}_3 = \{f_1, f_2, f_3\}$. We have

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

First, we interchange the rows by $T_{1,3} : \mathbb{F}^3 \to \mathbb{F}^3$, Then,

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

If we look at the matrix, we then have

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

For the full reduced row echelon form, we would have the following series of transformations:

$$\left[T_{1,3}^{(-1)} \circ T_{2,3}^{(-1)} \circ T_3^{(-2)} \circ T_{3,1}^{(-3)} \circ T_{1,2}^{-1} \circ T_{1,3} \circ T_A\right]_{\mathcal{E}_4}^{\mathcal{F}_3} = \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & -1\\ 0 & 0 & 1 & 2\end{pmatrix}.$$

Definition (Column Space, Null Space, and Rank)**.** Let $A \in \text{Mat}_{m,n}(\mathbb{F})$. The column space of A is the $\mathbb{F}$-span of the column vectors. This is denoted $\text{CS}(A)$.

The null space, $\text{NS}(A)$, is the $\mathbb{F}$-span of the vectors $v \in \mathbb{F}^n$ such that $Av = 0_{\mathbb{F}^m}$.

The rank of A, denoted $\text{rank}(A)$, is $\text{rank}(A) = \dim_{\mathbb{F}}(\text{CS}(A))$.

Let $\mathcal{E}_n = \{e_1, \dots, e_n\}$ be the standard basis for $\mathbb{F}^n$, with $T_A \in \text{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^n)$, and $\mathcal{F}_m = \{f_1, \dots, f_m\}$ the standard basis of $\mathbb{F}^m$.

We have $[T_A]_{\mathcal{E}_n}^{\mathcal{F}_m} = A$. We know that

$$A = \begin{pmatrix} T_A(e_1) & \cdots & T_A(e_n)\end{pmatrix}.$$

Thus, $\text{CS}(A) = \text{im}(T_A)$, meaning $\text{rank}(A) = \dim_{\mathbb{F}}(\text{im}(T_A))$.

In order to calculate $\text{CS}(A)$, we put the matrix A into row echelon form, look at the columns that have pivots, and those columns form the basis for $\text{CS}(A)$.

We have an isomorphism $E : \mathbb{F}^m \to \mathbb{F}^m$ such that

$$[E \circ T_A]_{\mathcal{E}_n}^{\mathcal{F}_m} = [E]_{\mathcal{F}_m}^{\mathcal{F}_m}$$

is in row echelon form. In particular, the column space of $[E \circ T_A]_{\mathcal{E}_n}^{\mathcal{F}_m}$ has as its basis the columns containing pivots:

$$\underbrace{\overbrace{[E \circ T_A(e_{i_1})]_{\mathcal{F}_m}}^{w_1}, \dots, \overbrace{[E \circ T_A(e_{i_k})]_{\mathcal{F}_m}}^{w_k}}_{\text{basis of } \text{CS}\left([E \circ T_A]_{\mathcal{E}_n}^{\mathcal{F}_m}\right)}$$

We have an inverse $E^{-1} : \mathbb{F}^m \to \mathbb{F}^m$. In particular,

$$\underbrace{E^{-1}(w_1), \dots, E^{-1}(w_k)}_{=[T_A(e_{i_1})]_{\mathcal{F}_m}, \dots, [T_A(e_{i_k})]_{\mathcal{F}_m}}$$

are linearly independent since E^{-1} is an isomorphism.

If there is a vector $v \in \mathrm{CS}(A)$ that is not in the span of $[T_A(e_{i_1})]_{\mathcal{F}_m}, \dots, [T_A(e_{i_k})]_{\mathcal{F}_m}$, then $E(v)$ cannot be in the span of $w_1, \dots, w_k$.

Thus, the columns $[T_A(e_{i_1})]_{\mathcal{F}_m}, \dots [T_A(e_{i_k})]_{\mathcal{F}_m}$ give a basis for $\mathrm{CS}(A)$.

Example. Consider the matrix

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

We put A into row echelon form as

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

Examining the pivots, we have the column space as

$$\mathrm{CS}(B) = \mathrm{span}_{\mathbb{F}}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \right),$$

implying the basis of the column space for A is

$$\mathrm{CS}(A) = \mathrm{span}_{\mathbb{F}}\left(\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ 3 \\ 2 \end{pmatrix} \right).$$

We have $v \in \mathrm{NS}(A)$ if and only if $Av = 0_{\mathbb{F}^m}$. Since $Av = T_A(v)$, we have $\mathrm{NS}(A) = \ker(T_A)$.

Example. Let

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

The reduced row echelon form of A is

$$B = \begin{pmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Thus,

$$\mathrm{CS}(A) = \mathrm{span}_{\mathbb{F}}\left(\begin{pmatrix} 4 \\ -4 \\ 2 \end{pmatrix}, \begin{pmatrix} -4 \\ 4 \\ -1 \end{pmatrix} \right).$$

We know that $(A) = \ker(T_A) \subseteq \mathbb{F}^3$-domain of T_A. When we put a matrix into reduced row echelon form, we do not impact the basis vectors of the domain of T_A, implying that NS(A) = NS(B).

In particular, we want

$$\begin{aligned}\begin{pmatrix}1 & 0 & 1/2\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix} &= \begin{pmatrix}x_1 + (1/2)\,x_3\\ x_2\\ 0\end{pmatrix}\\ &= \begin{pmatrix}0\\0\\0\end{pmatrix}.\end{aligned}$$

Therefore, we have $x_2 = 0$, $x_1 = -1/2x_3$, meaning

$$\mathrm{NS}(A) = \operatorname{span}_{\mathbb{F}}\left(\begin{pmatrix}-1/2\\ 0\\ 1\end{pmatrix}\right).$$

Transpose of a Matrix

Recall that, given a linear map $T \in \operatorname{Hom}_{\mathbb{F}}(V, W)$, there is an induced map $T' \in \operatorname{Hom}_{\mathbb{F}}(W', V')$ on the dual space given by $T'(\varphi) = \varphi \circ T$.

Let $A \in \operatorname{Mat}_{m,n}(\mathbb{F})$, $\mathcal{E}_n = \{e_1, \dots, e_n\}$ and $\mathcal{F}_m = \{f_1, \dots, f_m\}$ be standard bases for $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively. Let $T_A \in \operatorname{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^m)$, meaning $A = [T_A]_{\mathcal{E}_n}^{\mathcal{F}_m}$.

We have $\mathcal{E}_n' = \left\{e_1', \dots, e_n'\right\}$ and $\mathcal{F}_m' = \left\{f_1', \dots, f_m'\right\}$. The dual map $T_A' \in \operatorname{Hom}_{\mathbb{F}}(\mathbb{F}^m, \mathbb{F}^n)$, and the transpose of A is defined by

$$A^T = \left[T_A'\right]_{\mathcal{F}_m'}^{\mathcal{E}_n'}.$$

Lemma: Let $A = \left(a_{ij}\right) \in \operatorname{Mat}_{m,n}(\mathbb{F})$. Then,

$$A^T = \left(b_{ij}\right) \in \operatorname{Mat}_{n,m}(\mathbb{F})$$

with $b_{ij} = a_{ji}$.

Proof. Let $A \in \operatorname{Mat}_{m,n}(\mathbb{F})$, $\mathcal{E}_n = \{e_1, \dots, e_n\}$ and $\mathcal{F}_m = \{f_1, \dots, f_m\}$ be standard bases for $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively. Let $\mathcal{E}_n'$ and $\mathcal{F}_m'$ denote the dual bases.

Let $T_A \in \operatorname{Hom}_{\mathbb{F}}(\mathbb{F}^n, \mathbb{F}^m)$, meaning $A = [T_A]_{\mathcal{E}_n}^{\mathcal{F}_m}$. In particular, we have

$$T_A(e_i) = \sum_{k=1}^{m} a_{ki} f_k. \tag{$*$}$$

We have

$$\begin{aligned}A^t &= \left[T_A'\right]_{\mathcal{F}_m'}^{\mathcal{E}_n'}\\ &= \left(b_{ij}\right)\end{aligned} \tag{$**$}$$

Now, we have

$$T_A'\left(f_j'\right) = \sum_{j=1}^{n} b_{kj} e_k'.$$

Apply f'_j to $(*)$. Then,

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

Apply $(**)$ to e_i. Then,

$$\begin{aligned}T'_A\left(f'_j\right)(e_i) &= \sum_{k=1}^{n} b_{kj} e'_k(e_i)\\ &= b_{ij}.\end{aligned}$$

We have

$$\left(f'_j \circ T_A\right)(e_i) = \left(T'_A\left(f'_j\right)\right)(e_i)$$

by the definition of T'_A, meaning $b_{ij} = a_{ji}$. □

Exercise: Let $A_1, A_2 \in \mathrm{Mat}_{m,n}(\mathbb{F})$, $c \in \mathbb{F}$. Use the definition of the transpose to show

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

Lemma: Let $A \in \mathrm{Mat}_{m,n}(\mathbb{F})$, $B \in \mathrm{Mat}_{p,m}(\mathbb{F})$. Then,

$$(BA)^T = A^T B^T.$$

Proof. Let $\mathcal{E}_m$, $\mathcal{E}_n$, and $\mathcal{E}_p$ be standard bases.

We have

$$\begin{aligned}[T_A]_{\mathcal{E}_n}^{\mathcal{E}_m} &= A\\ [T_B]_{\mathcal{E}_m}^{\mathcal{E}_p} &= B.\end{aligned}$$

So,

$$BA = [T_B \circ T_A]_{\mathcal{E}_n}^{\mathcal{E}_p}.$$

Thus,

$$\begin{aligned}(BA)^T &= \left[(T_B \circ T_A)'\right]_{\mathcal{E}'_p}^{\mathcal{E}'_n}\\ &= \left[T'_A \circ T'_B\right]_{\mathcal{E}'_p}^{\mathcal{E}'_n}\\ &= \left[T'_A\right]_{\mathcal{E}'_m}^{\mathcal{E}'_n}\left[T'_B\right]_{\mathcal{E}'_p}^{\mathcal{E}'_m}\\ &= A^T B^T.\end{aligned}$$

□

Lemma: Let $A \in \mathrm{GL}_n(\mathbb{F})$. Then,

$$\left(A^{-1}\right)^T = \left(A^T\right)^{-1}.$$

Proof. We will show that $A^{T}\left(A^{-1}\right)^{T} = I_n = \left(A^{-1}\right)^{T} A^{T}$, and use the fact that inverses are unique.

We have

$$\begin{aligned} A &= [T_A]_{\mathcal{E}_n}^{\mathcal{E}_n} \\ A^{-1} &= \left[T_A^{-1}\right]_{\mathcal{E}_n}^{\mathcal{E}_n} \end{aligned}$$

We have

$$\begin{aligned} I_n &= \left[\mathrm{id}'_{\mathbb{F}^n}\right]_{\mathcal{E}'_n}^{\mathcal{E}'_n} \\ &= \left[\left(T_A^{-1} \circ T_A\right)'\right]_{\mathcal{E}'_n}^{\mathcal{E}'_n} \\ &= \left[T'_A \circ \left(T_A^{-1}\right)'\right]_{\mathcal{E}'_n}^{\mathcal{E}'_n} \\ &= \left[T'_A\right]_{\mathcal{E}'_n}^{\mathcal{E}'_n}\left[\left(T_A^{-1}\right)'\right]_{\mathcal{E}'_n}^{\mathcal{E}'_n} \\ &= A^{T}\left(A^{-1}\right)^{T}. \\ I_n &= \left[\left(T_A \circ T_A^{-1}\right)'\right]_{\mathcal{E}'_n}^{\mathcal{E}'_n} \\ &= \left[\left(T_A^{-1}\right)' \circ T'_A\right]_{\mathcal{E}'_n}^{\mathcal{E}'_n} \\ &= \left[\left(T_A^{-1}\right)'\right]_{\mathcal{E}'_n}^{\mathcal{E}'_n}\left[T'_A\right]_{\mathcal{E}'_n}^{\mathcal{E}'_n} \\ &= \left(A^{-1}\right)^{T} A^{T}. \end{aligned}$$

□

Generalized Eigenvectors and Jordan Canonical Form

Eigenvalues and Eigenvectors

Recall that we say $A \sim B$ if $A = PBP^{-1}$ for some $P \in GL_n(\mathbb{F})$. In particular, this means that $A = [T]_{\mathcal{A}}$ and $B = [T]_{\mathcal{B}}$ for some bases $\mathcal{A}$ and $\mathcal{B}$.

Definition (Diagonalizable)**.** We say A is diagonalizable if $A \sim D$ for some D a diagonal matrix.

If $A = [T]_{\mathcal{A}}$, A is diagonalizable if there is a basis $\mathcal{B}$ if $[T]_{\mathcal{B}} = D$ for D a diagonal matrix.

If $A \sim B$, 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. Let $V = \mathbb{F}^2$, $T \in \mathrm{Hom}_{\mathbb{F}}(V, V)$. We take $T(e_1) = 3e_1$ and $T(e_2) = -2e_2$.

In particular, we can see that

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

When we look at $V = V_1 \oplus V_2$, with $V_1 = \mathrm{span}_{\mathbb{F}}(e_1)$ and $V_2 = \mathrm{span}_{\mathbb{F}}(e_2)$.

In this case, we have $T(V_1) \subseteq V_1$ and $T(V_2) \subseteq V_2$, which allows us to write T as a diagonal matrix.

Example. Let $V = \mathbb{F}^2$, $T \in \mathrm{Hom}_{\mathbb{F}}(V, V)$. We take $T(e_1) = 3e_1$ and $T(e_2) = e_1 + 3e_2$.

In particular, we can see that

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

We still have $V = V_1 \oplus V_2$ with $V_1 = \mathrm{span}_{\mathbb{F}}(e_1)$ and $V_2 = \mathrm{span}_{\mathbb{F}}(e_2)$.

While we have $T(V_1) \subseteq V_1$, we do not have $T(V_2) \subseteq V_2$. We will find a diagonalization (or lack thereof) of T.

Suppose we have $W_1, W_2 \neq \{0\}$ with $V = W_1 \oplus W_2$ with $T(W_1) \subseteq W_1$ and $T(W_2) \subseteq W_2$.

Write $W_i = \mathrm{span}_{\mathbb{F}}(w_i)$. 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 would be able to write

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

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

$$\begin{aligned} \alpha w_1 &= T(w_1) \\ &= aT(e_1) + bT(e_2) \\ &= a(3e_1) + b(e_1 + 3e_2) \\ &= (3a + b)e_1 + 3be_2 \end{aligned}$$

Thus, $\alpha(ae_1 + be_2) = (3a + b)e_1 + 3be_2$, meaning $\alpha a = 3a + b$, $\alpha b = 3b$. Either $b = 0$ or $\alpha = 3$, but we still end with $\alpha = 3$. Thus, $T(w_1) = 3w_1$.

Applying to w_2, we have

$$\beta w_2 = (3c + d)e_1 + (3d)e_2,$$

implying $\beta c = ec + d$ and $\beta d = 3d$, meaning either $\beta = 3$ (which contradicts the first equation)or $w_2 = ce_1$, which contradicts w_1, w_2 being a basis.

Example. Let

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

Let $\mathbb{F} = \mathbb{Q}$. Can we find $P \in GL_2(\mathbb{Q})$ such that $P^{-1}AP = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$.

If we write $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we have

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

By the definition of diagonal matrix, we must have

$$3a^2 + 3ac - 2c^2 = 0.$$

If $c = 0$, then $a = 0$, which is a contradiction since P is invertible. We have $c \neq 0$, meaning we can divide by c^2 and set $x = a/c$

$$3x^2 + 3x - 2 = 0$$

$$\begin{aligned} x &= \frac{-3 \pm \sqrt{33}}{6} \\ a &= \frac{-3 \pm \sqrt{33}}{6} c. \end{aligned}$$

Since $c \neq 0$, $\frac{-3\pm\sqrt{33}}{6} c \notin \mathbb{Q}$. Thus, we cannot diagonalize A over $\mathbb{Q}$.

If we take $\mathbb{F} = \mathbb{Q}\left(\sqrt{33}\right)$, then we take

$$\begin{aligned} \mathcal{B} &= \left\{ v_1 = \begin{pmatrix} 1 \\ \frac{3+\sqrt{33}}{4} \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ \frac{3-\sqrt{33}}{4} \end{pmatrix} \right\}, \\ [T]_{\mathcal{B}} &= \begin{pmatrix} \frac{5+\sqrt{33}}{2} & 0 \\ 0 & \frac{5-\sqrt{33}}{2} \end{pmatrix}. \end{aligned}$$

Recall: The fundamental question we are investigating is whether given a $A \in \operatorname{Mat}_n(\mathbb{F})$, can we choose $P \in GL_n(\mathbb{F})$ such that PAP^{-1} is diagonal.

We saw that if $\mathbb{F}^2 = V_1 \oplus V_2$ with $A(V_1) \subseteq V_1$, $A(V_2) \subseteq V_2$, then it is possible to diagonalize A.

Definition. Let V be an $\mathbb{F}$-vector space with $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$. We say a subspace $W \subseteq V$ is T-invariant or T-stable if $T(W) \subseteq W$.

Theorem: Let $\dim_{\mathbb{F}}(V) = n$, $W \subseteq V$ a k-dimensional subspace.

Let $\mathcal{B}_W = \{v_1, \dots, v_k\}$ be a basis for W, and extend to a basis $\mathcal{B} = \{v_1, \dots, v_n\}$ of V.

Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$.

Then, W is T-stable if and only if $[T]_{\mathcal{B}}$ is block-upper triangular of the form

$$[T]_{\mathcal{B}} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

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

Example. Let $V = \mathbb{Q}^4$, $\mathcal{E}_4 = \{e_1, e_2, e_3, e_4\}$ the standard basis. Define T by

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

Notice that if we set $W = \operatorname{span}_{\mathbb{Q}}(e_1, e_3)$, then W is T-stable. We set $\mathcal{B}_W = \{e_1, e_3\}$, $\mathcal{B} = \{e_1, e_2, e_3, e_4\}$.

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

A special case is when $\dim_{\mathbb{F}}(W) = 1$. If $W = \operatorname{span}_{\mathbb{F}}(w_1)$, and W is T-stable, then $T(w_1) \in W$, meaning $T(w_1) = \lambda w_1$ for some $\lambda \in \mathbb{F}$.

We can rewrite this as $T(w_1) - \lambda(w_1) = 0_V$, meaning $(T - \lambda \operatorname{id}_V)(w_1) = 0_V$, meaning $w_1 \in \ker(T - \lambda \operatorname{id}_V)$.

Definition. Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$, 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 an eigenvector.

The set $E_\lambda^1 = \ker(T - \lambda \operatorname{id}_V)$ is called the eigenspace associated with λ.

Exercise: Show E_λ^1 is a subspace of V.

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

Example. Let

$$A = \begin{pmatrix} -12 & 35 \\ -6 & 17 \end{pmatrix} \in \operatorname{Mat}_2(\mathbb{Q}),$$

with $T_A \in \operatorname{Hom}_{\mathbb{Q}}\left(\mathbb{Q}^2, \mathbb{Q}^2\right)$ the associated linear map.

We have

$$\begin{pmatrix} -12 & 35 \\ -6 & 17 \end{pmatrix}\begin{pmatrix} 1 \\ 2/5 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 2/5 \end{pmatrix}$$
$$\begin{pmatrix} -12 & 35 \\ -6 & 17 \end{pmatrix}\begin{pmatrix} 1 \\ 3/7 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 3/7 \end{pmatrix}.$$

Therefore, T_A has eigenvalues of 2 and 3, with

$$E_2 = \operatorname{span}_{\mathbb{Q}}\left(\begin{pmatrix} 1 \\ 2/5 \end{pmatrix}\right) = \operatorname{span}_{\mathbb{Q}}(v_1)$$
$$E_3 = \operatorname{span}_{\mathbb{Q}}\left(\begin{pmatrix} 1 \\ 3/7 \end{pmatrix}\right) = \operatorname{span}_{\mathbb{Q}}(v_2),$$

meaning

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

Notation: Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$. We write $T^m = \underbrace{T \circ \cdots \circ T}_{m \text{ times}}$.

If $f(x) \in \mathbb{F}[x]$, $f(x) = a_m x^m + \cdots + a_1 x + a_0$, then

$$\begin{aligned} f(T) &= a_m T^m + \cdots + a_1 T + a_0 \operatorname{id}_V \\ &\in \operatorname{Hom}_{\mathbb{F}}(V, V). \end{aligned}$$

If $f(x) = g(x)h(x)$, then

$$f(T) = g(T) \circ h(T)$$

Example. If $g(x) = 2x^2 + 3$, then

$$g(T) = 2T^2 + 3\operatorname{id}_V$$
$$g(T)(v) = 2T(T(v)) + 3v.$$

Let $\dim_{\mathbb{F}}(V) = n$. Recall that $\operatorname{Hom}_{\mathbb{F}}(V, V)$ is an $\mathbb{F}$-vector space, meaning $\operatorname{Hom}_{\mathbb{F}}(V, V) \cong \operatorname{Mat}_n(\mathbb{F})$. Thus, $\dim_{\mathbb{F}}(\operatorname{Hom}_{\mathbb{F}}(V, V)) = n^2$.

Given $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$, consider

$$\left\{\operatorname{id}_V, T, T^2, \dots, T^{n^2}\right\} \subseteq \operatorname{Hom}_{\mathbb{F}}(V, V).$$

Since this set contains $n^2 + 1$ elements, it must be linearly dependent. Let m be the smallest integer such that $a_m T^m + \cdots + a_1 T + a_0 \operatorname{id}_V = 0_{\operatorname{Hom}_{\mathbb{F}}(V,V)}$. Since m is minimal, $a_m \neq 0$.

Define $f(x) = x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0 \in \mathbb{F}[x]$, where $b_i = \frac{a_i}{a_m}$.

Observe that $f(T) = 0_{\operatorname{Hom}_{\mathbb{F}}(V,V)}$. In other words, $f(T)(v) = 0_V$ for all $v \in V$.

Theorem: Let $\dim_{\mathbb{F}}(V) = n$. There is a unique monic polynomial $m_T(x) \in \mathbb{F}[x]$ of lowest degree such that

$$m_T(T)(v) = 0_V$$

for every $v \in V$. Moreover, $\deg(m_T(x)) \leqslant n^2$

Proof of Uniqueness. Suppose $f(x) \in \mathbb{F}[x]$ satisfies $f(T)(v) = 0$ for all $v \in V$.

We write

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

for some $q(x), r(x) \in \mathbb{F}[x]$, with $r(x) = 0$ or $\deg r(x) < \deg m_T(x)$.

Plugging in T, we have for all $v \in V$,

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

Thus, $r(T)(v) = 0$ for all $v \in V$; thus, it must be the case that $r(T) = 0$.

Thus, $m_T(x) | f(x)$. However, if $m_T(x)$ and $f(x)$ are monic and of minimal degree, with $m_T(x)|f(x)$, then $m_T(x) = f(x)$. □

Definition. The unique monic polynomial $m_T(x)$ is called the minimal polynomial.

Corollary: If $f(x) \in \mathbb{F}[x]$ satisfies $f(T)(v) = 0$ for all $v \in V$, then $m_T(x)|f(x)$.

Example. Let $F = \mathbb{Q}$,

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

We can see that for any $a_0 \in \mathbb{Q}$,

$$A - a_0 I_2 \neq 0_{\operatorname{Mat}_2(\mathbb{Q})}.$$

However, for

$$A^2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix},$$

we have

$$A^2 - 5A - 2I_2 = 0_{\operatorname{Mat}_2(\mathbb{Q})},$$

yielding $m_A(x) = x^2 - 5x - 2$.

The roots of $m_A(x)$ are $\frac{5 \pm \sqrt{33}}{2}$.

Example. Let $V = \mathbb{Q}^3$, $\mathcal{E}_3 = \{e_1, e_2, e_3\}$, with T_A given by

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

We can find

$$A^2 = \begin{pmatrix} 1 & 4 & 8 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$A^3 = \begin{pmatrix} 1 & 6 & 11 \\ 0 & 1 & 4 \\ 0 & 0 & -1 \end{pmatrix}.$$

Thus, we find

$$A^3 - A^2 - A + I = 0,$$
$$(x-1)^2 (x+1) = m_{T_A}(x)$$

Theorem: Let V be an $\mathbb{F}$-vector space, and let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$. We have λ is an eigenvalue if and only if λ is a root of $m_T(x)$.

In particular, if $(x - \lambda) \,|\, m_T(x)$, then $E_\lambda^1 \neq \{0_V\}$.

Proof. Let λ be an eigenvalue with eigenvector v, and write $m_T(x) = x^m + \cdots + a_1 x + a_0$. Notice that $T^k(v) = \lambda^k(v)$.

We have

$$\begin{aligned}
0_V &= m_T(T)(v) \\
&= \left(T^m + a_{m-1}T^{m-1} + \cdots + a_1 T + a_0 \operatorname{id}_V\right)(v) \\
&= T^m(v) + a_{m-1}T^{m-1}(v) + \cdots + a_1 T(v) + a_0 v \\
&= \lambda^m v + a_{m-1}\lambda^{m-1} v + \cdots + a_1 \lambda v + a_0 v \\
&= \left(\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0\right) v \\
&= m_T(\lambda)\, v,
\end{aligned}$$

meaning $m_T(\lambda) v = 0_V$. Since $m_T(\lambda) \in \mathbb{F}$ and $v \neq 0_V$, it is the case that $m_T(\lambda) = 0$, meaning λ is a root of $m_T(x)$.

Suppose $m_T(\lambda) = 0$. This gives

$$m_T(x) = (x - \lambda) f(x)$$

for some $f(x) \in \mathbb{F}[x]$. Therefore, $\deg(f(x)) < \deg(m_T(x))$. There must exist a nonzero vector $v \in V$ such that $f(T)(v) \neq 0_V$. Set $w = f(T)(v)$. Observe that $m_T(T)(v) = 0_V$, so $(T - \lambda \operatorname{id}_V) f(T)(v) = 0_V$, meaning $(T - \lambda \operatorname{id}_V)(w) = 0_V$, so $T(w) = \lambda w$. Thus, λ is an eigenvalue. □

Corollary: Let $\lambda_1, \dots, \lambda_m \in \mathbb{F}$ be distinct eigenvalues of T. For each i, let v_i be an eigenvector with eigenvalue λ_i. Then, $\{v_1, \dots, v_m\}$ is linearly independent

Proof. We can write

$$m_T(x) = (x - \lambda_1) \cdots (x - \lambda_m) f(x).$$

Suppose $a_1v_1 + \cdots + a_mv_m = 0_V$ for some $a_i \in \mathbb{F}$.

Define $g_1(x) = (x - \lambda_2)\cdots(x - \lambda_m) f(x)$. Note that $g_1(T)(v_i) = 0_V$ for all $2 \leqslant i \leqslant m$. Then,

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

Since $g_1(\lambda_1) \neq 0$, and $v_1 \neq 0$, it must be the case that $a_1 = 0$. Symmetry provides the case for $2, \dots, m$. □

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

Example. Let

$$\begin{aligned}
A &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \\
B &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.
\end{aligned}$$

These matrices are not similar. However, $m_A(x) = m_B(x) = (x-1)(x-2)$.

Therefore, the minimal polynomial does not provide enough information about a matrix's similarity class.

Example. Let

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

We found that the minimal polynomial for A was $m_A(x) = (x-1)^2(x+1)$.

We can see that $Ae_1 = e_1$, meaning $\operatorname{span}_{\mathbb{F}}(e_1) = E_1^1$. Note that

$$Ae_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix},$$

meaning $e_2 \notin E_1^1$.

We can see that

$$(A - I_3)^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & -8 \\ 0 & 0 & 4 \end{pmatrix}.$$

However,

$$(A - I_3)^2(e_2) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

meaning $e_1, e_2 \in \ker\left((T_A - \mathrm{id}_{\mathbb{F}^3})^2\right)$.

Though we do not have distinct eigenvectors, we *kinda* have them.

Definition (Generalized Eigenvector)**.** Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$. For $k \geqslant 1$, the kth generalized eigenspace of T with eigenvalue λ is

$$\begin{aligned} E_{\lambda}^{k} &= \ker\left((T_A - \lambda\,\mathrm{id}_V)^k\right) \\ &= \left\{v \in V \mid (T - \lambda\,\mathrm{id}_V)^k v = 0_V\right\}. \end{aligned}$$

Elements in E_{λ}^{k} are called generalized λ-eigenvectors.

We set

$$E_{\lambda}^{\infty} = \bigcup_{k \geqslant 1} E_{\lambda}^{k}.$$

Example. In the previous example, we saw that $\operatorname{span}_{\mathbb{F}}(e_1, e_2) \subseteq E_1^2$, and we have -1 is an eigenvalue of A with eigenvector

$$v_3 = \begin{pmatrix} 1/2 \\ -1/2 \\ 1 \end{pmatrix}.$$

We can verify that $v_3 \notin E_1^2$.

Thus, $\dim_F E_1^2 \leqslant 2$, meaning $E_1^2 = \operatorname{span}_{\mathbb{F}}(e_1, e_2)$.

Example. Let $\mathcal{B} = \{v_1, \dots, v_n\}$ be a basis for V, and $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$, $\lambda \in \mathbb{F}$ such that

$$A = [T]_{\mathcal{B}} = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix},$$

which is a matrix of λ along the diagonal and 1 along the superdiagonal. In particular, we can see that $A - \lambda I_n$ is the matrix with 1 along the superdiagonal and 0 everywhere else.

Notice that $(A - \lambda I_n)(v_1) = 0$, $(A - \lambda I_n)(v_2) = v_1$, etc.

Thus, we get that $E_{\lambda}^{1} = \operatorname{span}_{\mathbb{F}}(v_1)$, $E_{\lambda}^{2} = \operatorname{span}_{\mathbb{F}}(v_1, v_2)$, etc.

In general, $E_{\lambda}^{k} = \operatorname{span}_{\mathbb{F}}(v_1, \dots, v_k)$ for $1 \leqslant k \leqslant n$.

Thus, $E_{\lambda}^{\infty} = E_{\lambda}^{n} = V$.

Exercise: Describe the generalized eigenspaces of

$$\begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix}$$

We can see that we used $E_{\lambda}^{i} \subseteq E_{\lambda}^{i+1}$; this is true more generally.

More generally, let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$. We claim that if $i \geqslant j$, then $\ker\left(T^j\right) \subseteq \ker\left(T^i\right)$.

Write $i = j + k$. Let $v \in \ker\left(T^j\right)$. Then,

$$T^i(v) = T^{j+k}(v)$$

$$\begin{aligned} &= T^k\left(T^j(v)\right)\\ &= T^k(0_V)\\ &= 0_V. \end{aligned}$$

This gives $E_\lambda^1 \subseteq E_\lambda^2 \subseteq \cdots \subseteq E_\lambda^\infty$.

Lemma: Let V be a finite dimensional vector space with $\dim_{\mathbb{F}}(V) = n$, and $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$. Then, there exists m with $1 \leqslant m \leqslant n$ such that

$$\ker(T^m) = \ker\left(T^{m+1}\right).$$

Moreover, for such an m, $\ker(T^m) = \ker\left(T^{m+j}\right)$ for all $j \geqslant 0$.

Proof. We have

$$\ker\left(T^1\right) \subseteq \ker\left(T^2\right) \subseteq \cdots \subseteq \ker(T^\infty).$$

If these containments are strict, then the dimension goes up indefinitely, contradicting $\dim_{\mathbb{F}}(V) = n$.

Thus, we have $1 \leqslant m \leqslant n$ with

$$\ker(T^m) = \ker\left(T^{m+1}\right).$$

Let m be the smallest value such that $\ker(T^m) = \ker\left(T^{m+1}\right)$.

We use induction on j. The base case of $j = 1$ is what defines m. Assume $\ker(T^m) = \ker\left(T^{m+j}\right)$ for all $1 \leqslant j \leqslant N$.

Let $v \in \ker\left(T^{m+N+1}\right)$. This gives

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

meaning $T^N(v) \in \ker\left(T^{m+1}\right)$. However, $\ker\left(T^{m+1}\right) = \ker(T^m)$, meaning $T^N(v) \in \ker(T^m)$, hence

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

meaning $v \in \ker\left(T^{m+N}\right)$. The inductive hypothesis gives $\ker\left(T^{m+N}\right) = \ker(T^m)$, meaning $v \in \ker(T^m)$. Thus, $\ker\left(T^{m+N+1}\right) \subseteq \ker\left(T^{m+N}\right)$, meaning $\ker\left(T^{m+N+1}\right) = \ker\left(T^{m+N}\right)$. ◻

Corollary: If $\dim_{\mathbb{F}}(V) = n$, and $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$, there exists m with $1 \leqslant m \leqslant n$ such that for any $\lambda \in \mathbb{F}$,

$$E_\lambda^\infty = E_\lambda^m.$$

Theorem: Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$, $\lambda \in \mathbb{F}$, with $(x-\lambda)^j \mid m_T(x)$. We have

$$\dim_{\mathbb{F}}\left(E_\lambda^j\right) \geqslant j.$$

Proof. Write $m_T(x) = (x-\lambda)^k f(x)$, $f(x) \in \mathbb{F}[x]$, $f(x) \neq 0$.

Define $g_j(x) = (x-\lambda)^j$. We have $g_{k-1}f(x)$ is not the minimal polynomial, meaning there is $v \in V$ such that

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

Set $v_k = f(T)v$. Note that $v_k \neq 0_V$.

Observe that

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

Thus, $v \in E_\lambda^k$.

Moreover, by construction,

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

Thus, $v_k \notin E_\lambda^{k-1}$.

Define

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

Note that

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

meaning $v_{k-1} \in E_\lambda^{k-1}$.

Additionally,

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

meaning $v_{k-1} \in E_\lambda^{k-1} \setminus E_\lambda^{k-2}$.

Continuing the process, we construct $\{v_1, \dots, v_k\}$ linearly independent. □

Example. Let $T_A \in \operatorname{Hom}_{\mathbb{F}}(\mathbb{F}^3, \mathbb{F}^3)$ given by

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

We can verify that $m_T(x) = (x-2)^3$.

Observe that

$$(A - 2I_3)^2 = \begin{pmatrix} 0 & 0 & 4\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

Notice that $(A - 2I_3)^3(e_3) = 4e_3 \neq 0$, meaning we set $v_3 = e_3$.

Note that $(T - 2\,\mathrm{id}_V)^3(e_3) = 0$, meaning $e_3 \in E_2^3$.

We find $v_2 = (A - 2I_3)(v_3)$, meaning

$$\begin{aligned} v_2 &= \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}. \end{aligned}$$

Finally,

$$\begin{aligned} v_1 &= (A - 2I_3)(v_2) \\ &= \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix}. \end{aligned}$$

Thus, our generalized eigenvectors are

$$E_2^3 = \operatorname{span}\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix}\right).$$

If we say $\mathcal{B} = \{v_1, v_2, v_3\}$, then our matrix $[T_A]_{\mathcal{B}}$ is

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

Remark: This matrix is in what is known as Jordan canonical form.

Characteristic Polynomials and the Cayley–Hamilton Theorem

Definition. Let $A \in \mathrm{Mat}_n(\mathbb{F})$. The characteristic polynomial is $c_A(x) = \det(xI_n - A)$.

Remark: The Cayley–Hamilton theorem states that

$$c_A(A) = 0_n.$$

Definition. Let $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in \mathbb{F}[x]$. The companion matrix of $f(x)$ is given by $C(f(x))$, which consists of $-a_{n-1}$ through $-a_0$ along the first column, 0 on the rest of the diagonal, and 1 along the superdiagonal.

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

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

Proof. Let $A = PBP^{-1}$ for some $P \in GL_n(\mathbb{F})$. Then, we have

$$\begin{aligned} c_A(x) &= \det(xI_n - A) \\ &= \det\left(xI_n - PBP^{-1}\right) \\ &= \det\left(P(xI_n)P^{-1} - PBP^{-1}\right) \\ &= \det\left(P(xI_n - B)P^{-1}\right) \end{aligned}$$

$$\begin{aligned}
&= \det(P)\det(xI_n - B)\det\left(P^{-1}\right)\\
&= \det(xI_n - B)\\
&= c_B(x).
\end{aligned}$$

□

Definition (Characteristic Polynomial of Linear Transformation)**.** For $T \in \mathrm{Hom}_{\mathbb{F}}(V, V)$, let $\mathcal{B}$ be a basis of V and set

$$c_T(x) = c_{[T]_{\mathcal{B}}}(x).$$

Theorem: Let $v \in V$, $v \neq 0$. Let $\dim_{\mathbb{F}}(V) < \infty$. Then, there is a unique monic polynomial $m_{T,v}(x) \in \mathbb{F}[x]$ of minimal degree such that $m_{T,v}(T)(v) = 0_V$.

Moreover, if $f(x) \in \mathbb{F}[x]$ with $f(T)(v) = 0$, then $m_{T,v}(x) | f(x)$.

Proof. Consider the set $\{v, T(v), \dots, T^n(v)\}$. This collection consists of $n + 1$ elements of V, meaning it is linearly dependent. Let

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

for some $m \leqslant n$ of minimal degree with not all $a_i = 0$. Set

$$p(x) = x^m + \frac{a_{m-1}}{a_m}x^{m-1} + \cdots + \frac{a_1}{a_m}x + \frac{a_0}{a_m}.$$

Thus, $p(t)(v) = 0_V$ by construction.

Set

$$I_v = \{g(x) \in \mathbb{F}[x] \mid g(T)(v) = 0_V\}.$$

We know $p(x) \in I_v$, and $p(x) \neq 0$. We have $p(x)$ is a nonzero monic polynomial in I_v of minimal degree.

Set $m_{T,v}(x) = p(x)$.

Let $f(x) \in I_v$. We want to show that $m_{T,v}(x) | f(x)$.

Write $f(x) = q(x)m_{T,v}(x) + r(x)$ for some $q(x), r(x) \in \mathbb{F}[x]$, with $r(x) = 0$ or $\deg(r(x)) < \deg m_{T,v}(x)$. We have $r(x) = f(x) - q(x)m_{T,v}(x)$, implying

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

implying $r(x) \in I_v$. Since $m_{T,v}(x)$ was defined to have minimal degree, it has to be the case that $r(x) = 0$.

If $h(x) \in I_v$ with $\deg(h(x)) = \deg(m_{T,v}(x))$ with $h(x)$ monic, then $m_{T,v}(x) | h(x)$ implies $h(x) = m_{T,v}(x)$. □

We will refer to $m_{T,v}(x)$ as the T-annihilator of v.

Example. Let $V = \mathbb{F}^n$, $\mathcal{B} = \{e_1, \dots, e_n\}$. Define $T \in \mathrm{Hom}_{\mathbb{F}}(V, V)$ by

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

Let $f(x) = x$. Then, $f(T)(e_1) = T(e_1) = 0_V$, implying that $m_{T,e_1}(x) | x$; thus, $m_{T,e_1}(x) = 1$ or $m_{T,e_1}(x) = x$, but $\mathrm{id}(e_1) = e_1 \neq 0_V$, meaning $m_{T,e_1}(x) = x$.

Let $g(x) = x^2$. Then,

$$\begin{aligned} g(T)(e_2) &= T^2(e_2) \\ &= T(T(e_1)) \\ &= T(0_V) \\ &= 0_V. \end{aligned}$$

This gives $m_{T,e_2}(x) | x^2$, so $m_{T,e_2}(x) = 1, x, x^2$. If $m_{T,e_2}(x) = 1$, then $\mathrm{id}_V(e_2) = e_2 = 0_V$, which is not the case. Similarly, if $m_{T,e_2}(x) = x$, then $T(e_2) = e_1 = 0_V$, so $m_{T,e_2}(x) = x^2$.

For each $1 \leqslant j \leqslant n$, $m_{T,e_j}(x) = x^j$.

Example. Let $V = \mathbb{Q}^2$, $T \in \mathrm{Hom}_{\mathbb{Q}}\left(\mathbb{Q}^2, \mathbb{Q}^2\right)$, with

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

We wish to find the annihilating polynomial for e_1.

We know that $m_{T,e_1}(x)$ has degree 1 or 2. Additionally, $m_{T,e_1}(x)$ cannot have degree 1, as if $m_{T,e_1}(x) = x + a$, then

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

Thus, m_{T,e_1} is of degree 2.

$$\begin{aligned} T^2(e_1) &= T(e_1 + 3e_2) \\ &= T(e_1) + 3T(e_2) \\ &= e_1 + 3e_2 + 3(2e_1 + 4e_2) \\ &= 7e_1 + 15e_2. \end{aligned}$$

We want to find $b, c \in \mathbb{Q}$ such that

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

Solving the resulting system of linear equation yields $b = -5$ and $c = -2$. The annihilating polynomial is, thus,

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

Exercise:

(1) Show that $m_{T,e_2}(x) = x^2 - 5x - 2$.

(2) Calculate $m_{T,e_1}(x)$ and $m_{T,e_2}(x)$ for $\mathbb{F} = \mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$.

Theorem: Let $\dim_{\mathbb{F}}(V) = n$, and $\mathcal{B} = \{v_1, \dots, v_n\}$ be a basis of V. Let $T \in \mathrm{Hom}_{\mathbb{F}}(V, V)$. We have

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

Proof. Let $f(x) = \operatorname{lcm}_{1\leqslant i\leqslant n} m_{T,v_i}(x)$. Then,

$$m_T(T)(v_i) = 0$$

meaning $m_{T,v_i}|m_T(x)$ for each i, so $f(x)|m_T(x)$.

Let $v \in V$; write $v = \sum_{i=1}^{n} a_iv_i$. Then,

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

since $m_{T,v_i}(x)|f(x)$ for all i. Thus, $m_T(x)|f(x)$. □

Lemma: Let $T \in \operatorname{Hom}_{\mathbb{F}}(V,V)$. Let $v_1,\dots,v_k \in V$, and set $p_i(x) = m_{T,v_i}(x)$. Suppose $p_i(x)$ are pairwise relatively prime. Set

$$v = v_1 + \cdots v_k.$$

Then,

$$m_{T,v}(x) = \prod_{j=1}^{k} p_j(x).$$

Proof. We will prove this for $k = 2$.

Since $p_1(x)$ and $p_2(x)$ are relatively prime, we can write

$$1 = p_1(x)q_1(x) + p_2(x)q_2(x).$$

Particularly,

$$\mathrm{id}_V = p_1(T)q_1(T) + p_2(T)q_2(T).$$

Set $v = v_1 + v_2$. Then,

$$\begin{aligned} v &= \mathrm{id}_V(v)\\ &= (p_1(T)q_1(T) + p_2(T)q_2(T))(v)\\ &= p_1(T)q_2(T)(v) + p_2(T)q_2(T)(v)\\ &= p_1(T)q_2(T)(v_1+v_2) + p_2(T)q_2(T)(v_1+v_2)\\ &= \underbrace{p_1(T)q_2(T)(v_2)}_{w_2} + \underbrace{p_2(T)q_2(T)(v_2)}_{w_1} \end{aligned}$$

meaning

$$v = w_1 + w_2.$$

Note that

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

meaning $w_1 \in \ker(p_1(T))$, and similarly, $w_2 \in \ker(p_2(T))$.

Let $r(x) \in \mathbb{F}[x]$ with $r(T)(v) = 0$. We have $v = w_1 + w_2$ and $w_2 \in \ker(p_2(T))$, meaning

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

Thus,

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

Similarly, $r(T)q_1(T)p_1(T)(w_1) = 0_V$ since $w_1 \in \ker(p_1(T))$. 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

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

Thus, $p_1(x) | r(x)p_2(x)q_2(x)$. Additionally,

$$\begin{aligned} 1 &= p_1(x)q_1(x) + p_2(x)q_2(x) \\ \gcd(p_1(x), p_2(x)q_2(x)) &= 1, \end{aligned}$$

implying $p_1(x) | r(x)$, and similarly for $p_2(x) | r(x)$.

Since $\gcd(p_1(x), p_2(x)) = 1$, we have

$$\operatorname{lcm}(p_1(x), p_2(x)) = p_1(x)p_2(x),$$

so $p_1(x)p_2(x) | r(x)$. If we take $r(x) = m_{T,v}(x)$, implying $p_1(x)p_2(x) | m_{T,v}(x)$. Thus, $m_{T,v}(x) = p_1(x)p_2(x)$. □

Exercise: Prove for $k > 2$.

Theorem: Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$. There exists $v \in V$ such that $m_{T,v}(x) = m_T(x)$. In particular, $\deg m_T(x) \leqslant n$.

Proof. Let $\mathcal{B} = \{v_1, \dots, v_n\}$ be a basis for V.

We know that

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

Factor

$$m_T(x) = p_1(x)^{e_1} \cdots p_k(x)e^k,$$

with p_i relatively prime, $e_i \geqslant 1$.

For $1 \leqslant j \leqslant k$, there exists $i_j \in \{1, \dots, n\}$ and $q_{i_j}(x) \in \mathbb{F}[x]$ with

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

Define $w_j = q_{i_j}(T)\left(v_{i_j}\right)$. This gives

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

Set $w = w_1 + \cdots + w_k$. The previous result gives

$$\begin{aligned} m_{T,w}(x) &= \prod_{j=1}^{k} p_j(x)^{e_j} \\ &= m_T(x). \end{aligned}$$

□

Recall: We defined $m_{T,v}(x)$, and that $m_T(x)$ is $m_{T,v}(x)$ for some $v \in V$, meaning $\deg(m_T(x)) < n$.

Lemma: Let W be a T-invariant subspace. We get a map $\overline{T} \in \operatorname{Hom}_{\mathbb{F}}(V/W, V/W)$ defined by

$$\overline{T}(v + W) = T(v) + W.$$

Let $v \in V$. Then,

$$m_{\overline{T},v+W}(x) | m_{T,v}(x)$$

and similarly,

$$m_{\overline{T}}(x) | m_T(x).$$

Proof. We have

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

Thus, $m_{\overline{T},v+W} | m_{T,v}(x)$. □

Definition. Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$, $\mathcal{A} = \{v_1, \dots, v_k\}$ a set of vectors. The T-span of $\mathcal{A}$ is

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

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

Example. Let $V = \mathbb{Q}^4$. Take $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$ by

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

Let $\mathcal{A} = \{e_1\}$. We want the T-span of $\mathcal{A}$. Set $p(x) = 1$, meaning $p(T)(e_1) = \mathrm{id}(e_1) = e_1$.

Set $q(x) = \frac{1}{3}(x-2)$. If we take $q(T)(e_1)$, we have

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

Thus, $\operatorname{span}_{\mathbb{F}}(e_1, e_3) \subseteq$ T-span of $\mathcal{A}$.

However, we also know that $\operatorname{span}_{\mathbb{F}}(e_1, e_3)$ is T-invariant and contains $\mathcal{A}$.

Thus, the T-span of $\mathcal{A}$ is $\operatorname{span}_{\mathbb{F}}(e_1, e_3)$.

If we set $f(x) = x^2 - 5x - 1$, then $f(T)(e_1) = 0_V$, meaning $m_{T,e_1}(x) | f(x)$. However, f is irreducible over $\mathbb{Q}$, so $m_{T,e_1}(x) = x^2 - 5x - 1$. Note that $\deg(m_{T,e_1}(x)) = \dim_{\mathbb{F}}(\text{T-span } \{e_1\})$.

Lemma: Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$, $w \in V$, and W the subspace of V that is generated by T on $\{w\}$.

Then, $\dim_{\mathbb{F}}(W) = \deg(m_{T,w}(x))$.

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

Theorem: Let $\dim_{\mathbb{F}}(V) = n$.

(1) We have $m_T(x) | c_T(x)$.

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

Proof. 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 $\{v\}$, with $\dim_{\mathbb{F}}(W_1) = K$

Set $v_k = v$, $v_{k-i} = T^i(v)$. We have

$$\mathcal{B} = \{v_1, dots, v_k\}$$

is a basis of W_1, and

$$\left[T\big|_{W_1}\right]_{\mathcal{B}_1} = c(m_T(x)).$$

If $k = n$, then $W_1 = V$, so $[T]_{\mathcal{B}_1} = c(m_T(x))$ which has characteristic polynomial $m_T(x)$, meaning $m_T(x) = c_T(x)$.

Suppose $k < n$. Expand $\mathcal{B}_1$ to a full basis of V, $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$, with $\mathcal{B}_2 = \{v_{k+1}, \dots, v_n\}$. In the upper triangular matrix

$$[T]_{\mathcal{B}} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

we have $A = c(m_T(x))$, so

$$\begin{aligned} c_T(x) &= \det\left(xI_n - [T]_{\mathcal{B}}\right) \\ &= \det\begin{pmatrix} xI_k - A & B \\ 0 & xI_{n-k} - D \end{pmatrix} \end{aligned}$$

$$\begin{aligned} &= \det(xI_k - A)\det(xI_{n-k} - D) \\ &= c_A(x)\det(xI_{n-k} - D) \\ &= m_T(x)\det(xI_{n-k} - D), \end{aligned}$$

meaning $m_T(x)|c_T(x)$.

To prove (2), we induct on $\dim_{\mathbb{F}}(V) = n$. If $n = 1$, then both characteristic polynomials are monic of degree 1, so they are equal.

If $\deg(m_T(x)) = n$, then $m_T(x)|c_T(x)$, and both have degree n and are monic, so $c_T(x) = m_T(x)$.

Suppose $\deg(m_T(x)) = k < n$. Pick v such that $m_T(x) = m_{T,v}(x)$. Define W_1 to be the T-span of $\{v\}$, with $\mathcal{B}_1 = \{v_1, \dots, v_k\}$ defined as above. Extend $\mathcal{B}_1$ to $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ as above.

Consider $\overline{T} : V/W_1 \to V/W_1$, and $\overline{\mathcal{B}} = \{v_{k_1} + W_1, \dots, v_n + W_1\} = \pi_{W_1}(\mathcal{B})$.

In the upper triangular matrix

$$[T]_{\mathcal{B}} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

the matrix $\left[\overline{T}\right]_{\overline{\mathcal{B}}} = D$.

Since $\dim_{\mathbb{F}}(V/W_1) < \dim_{\mathbb{F}}(V)$, the induction hypothesis holds that $m_{\overline{T}}(x)$ and $c_{\overline{T}}(x)$ have the same irreducible factors.

Earlier, we had

$$c_T(x) = m_T \det(xI_{n-k} - D),$$

yielding

$$c_T(x) = m_T(x)c_{\overline{T}}(x).$$

Let $p(x)$ be an irreducible factor of $c_T(x)$. If $p(x)|m_T(x)$, we are done. Else, $p(x)|c_{\overline{T}}(x)$. However, $c_{\overline{T}}(x)$ and $m_{\overline{T}}(x)$ have the same irreducible factors, so $p|m_{\overline{T}}(x)$. However, $m_{\overline{T}}(x)|m_T(x)$, so $p(x)|m_T(x)$. □

Example. Let

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 7 & 0 & 0 \\ 0 & 0 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & -5 & 6 \\ 0 & 0 & 0 & 0 & 2 & -3 \end{pmatrix} \in \mathrm{Mat}_6(\mathbb{Q}).$$

We can verify that

$$c_A(x) - \left(x^2 - 5x - 2\right)\left(x^2 - x - 1\right)\left(x^2 + 8x + 3\right),$$

implying that

$$m_A(x) = \left(x^2 - 5x - 2\right)\left(x^2 - x - 1\right)\left(x^2 + 8x + 3\right).$$

Theorem (Cayley–Hamilton)**:**

(1) Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$, $\dim_{\mathbb{F}}(V) < \infty$. Then,

$$c_T(T) = 0_{\operatorname{Hom}_{\mathbb{F}}(V,V)}$$

(2) Let $A \in \operatorname{Mat}_n(\mathbb{F})$. Then,

$$c_A(A) = 0_n.$$

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

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

□

Jordan Canonical Form

For the purposes of this section, V is always finite dimensional, and all polynomials split into linear factors over their respective fields.

Definition. Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$. A Jordan basis for V with regard to T is a basis $\mathcal{B}$ such that $[T]_{\mathcal{B}}$ has some $\lambda \in \mathbb{F}$ along the diagonal and 1 along the superdiagonal.

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

Definition. A matrix with λ along the diagonal and 1 along the superdiagonal is called a Jordan block associated with eigenvalue λ.

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

Definition. We say a matrix is in Jordan canonical form if it is block diagonal with Jordan blocks.

Theorem:

(1) Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$. Suppose $c_T(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_k)^{e_k}$ over $\mathbb{F}$. Then, V has a Jordan basis $\mathcal{B}$. Moreover, $J = [T]_{\mathcal{B}}$ is unique up to the order of the blocks.

(2) Let $A \in \operatorname{Mat}_n(\mathbb{F})$ with $c_A = (x - \lambda_1)^{e_1} \cdots (x - \lambda_k)^{e_k}$ over $\mathbb{F}$. Then A is similar to a matrix in Jordan canonical form that is unique up to the order of the blocks.

Lemma: Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$. We have $\ker\left((T - \lambda \operatorname{id}_V)^j\right)$ and $\operatorname{im}\left((T - \lambda \operatorname{id}_V)^j\right)$ are T-invariant subspaces for all $j \geqslant 0$.

Proof. Note that

$$T \circ (T - \lambda \operatorname{id}_V)^j = (T - \lambda \operatorname{id}_V)^j \circ T.$$

Let $v \in \ker\left((T - \lambda \operatorname{id}_V)^j\right)$. We have

$$(T - \lambda \operatorname{id}_V)^j (T(v)) = T\left((T - \lambda \operatorname{id}_V)^j (v)\right)$$

$$= T(0_V)$$
$$= 0_V.$$

Thus, $T(v) \in \ker\left((T - \lambda\,\mathrm{id}_V)^j\right)$.

Let $w \in \mathrm{im}\left((T - \lambda\,\mathrm{id}_V)^j\right)$. We can write

$$w = (T - \lambda\,\mathrm{id}_V)^j\,(v)$$

for some $v \in V$. Applying T to both sides, we have

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

meaning $T(w) \in \mathrm{im}\left((T - \lambda\,\mathrm{id}_V)^j\right)$. ◻

We know there exists m such that $E_\lambda^\infty = E_\lambda^m$. We also know that if $(x - \lambda)^k \,|\, m_T(x)$, then $\dim_{\mathbb{F}}\left(E_\lambda^k\right) \geqslant k$.

Lemma: Suppose $m_T(x) = (x - \lambda)^m\, p(x)$ with $p(\lambda) \neq 0$. Then,

$$E_\lambda^\infty = E_\lambda^m.$$

Proof. Let $v \in E_\lambda^\infty$ and e be the least positive integer such that

$$(T - \lambda\,\mathrm{id}_V)^e\,(v) = 0_V.$$

Suppose toward contradiction that $e > m$. We have $m_{T,v}(x) |\, (x - \lambda)^e$, but $m_{T,v}(x) \nmid (x - \lambda)^{e-1}$. In particular, $m_{T,v}(x) = (x - \lambda)^e$. However, $m_{T,v}(x) | m_T(x)$. ◻

Lemma: Let $\dim_{\mathbb{F}}(V) = n$. Let $m_T(x) = (x - \lambda)^m\, p(x)$ with $p(\lambda) \neq 0$. Then, we have

$$V = E_\lambda^m \oplus \mathrm{im}\left((T - \lambda\,\mathrm{id}_V)^m\right).$$

Proof. Recall that

$$E_\lambda^m = \ker\left((T - \lambda\,\mathrm{id}_V)^m\right).$$

Therefore, the dimensions line up. All we need show is that $E_\lambda^m \cap \mathrm{im}\left((T - \lambda\,\mathrm{id}_V)^m\right) = \{0_V\}$.

Let $v \in E_\lambda^m \cap \mathrm{im}\left((T - \lambda\,\mathrm{id}_V)^m\right)$. We have

$$v = (T - \lambda\,\mathrm{id}_V)^m\,(w)$$

for some $w \in V$. Applying $(T - \lambda\,\mathrm{id}_V)^m$ to both sides, we have

$$\begin{aligned} (T - \lambda\,\mathrm{id}_V)^m\,(v) &= (T - \lambda\,\mathrm{id}_V)^{2m}\,(w)\\ &= 0_V, \end{aligned}$$

since $v \in \ker\left((T - \lambda\,\mathrm{id}_V)^m\right)$. Thus,

$$(T - \lambda)^{2m}\,(w) = 0_V.$$

Thus, $w \in E_\lambda^{2m}$. However, since $E_\lambda^\infty = E_\lambda^m$, so too is E_λ^{2m}, so $w \in E_\lambda^m$, meaning

$$(T - \lambda)^m\,(w) = 0_V,$$

so $v = 0_V$. ◻

Theorem: Assume $m_T(x) = (x-\lambda_1)^{m_1}\cdots(x-\lambda_k)^{m_k}$ with $\lambda_j \in \mathbb{F}$, λ_j distinct, $m_j \geqslant 1$.

Then,

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

Proof. We will 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

$$V = E_{\lambda_1}^{m_1}.$$

Assume the result is true for any vector space W and $S \in \operatorname{Hom}_{\mathbb{F}}(W, W)$, where $m_S(x)$ splits completely over $\mathbb{F}$ and has fewer than k distinct roots.

We can break our vector space V to be

$$V = E_{\lambda_1}^{m_1} \oplus \operatorname{im}\left((T - \lambda_1 \operatorname{id}_V)^{m_1}\right).$$

Set $W = \operatorname{im}\left((T-\lambda_1)^{m_1}\right)$. We have W is T-invariant. Thus, $T_W := T|_W \in \operatorname{Hom}_{\mathbb{F}}(W, W)$.

We claim that $m_{T_W}(x) = (x-\lambda_2)^{m_2}\cdots(x-\lambda_k)^{m_k}$.

Set $p(x) = (x-\lambda_2)^{m_2}\cdots(x-\lambda_k)^{m_k}$. Suppose $w \in W$ satisfies $p(T_W)(w) \neq 0_V$. At the same time, we have $m_T(T)(w) = 0_V$. Thus,

$$(T - \lambda_1 \operatorname{id}_V)^{m_1}(p(T)(w)) = 0_V,$$

meaning $p(T)(w) \in E_{\lambda_1}^{m_1}$. This is a contradiction, since $p(T)(w) = p(T_W)(w) \in W$.

Thus, $m_{T_W}|p(x)$.

Suppose m_{T_W} is a proper divisor of $p(x)$. If we set $f(x) = m_{T_W}(x)(x-\lambda_1)^{m_1}$. For $v \in V$, write

$$v = v_1 + w$$

with $v_1 \in E_{\lambda_1}^{m_1}$ and $w \in W$. Notice that

$$\begin{aligned}
f(T)(v) &= f(T)(v_1) + f(T)(w)\\
&= m_{T_W}\left((T-\lambda_1 \operatorname{id}_V)^{m_1}\right)(v) + (T - \lambda_1 \operatorname{id}_V)^{m_1} M_{T_W}(w)\\
&= 0_V + 0_V\\
&= 0_V.
\end{aligned}$$

Thus, $m_T(x)|f(x)$, which is a contradiction if m_{T_W} is a proper divisor of $p(x)$.

Thus, $m_{T_W}(x) = p(x)$ as claimed.

We have that

$$V = E_{\lambda_1}^{m_1} \oplus W,$$

and we apply the induction hypothesis to W to yield

$$V = E_{\lambda_1}^{m_1} \oplus \left(E_{\lambda_2}^{m_2} \oplus \cdots \oplus E_{\lambda_k}^{m_k}\right).$$

□

If T has minimal polynomial of the form $m_T(x) = (x-\lambda)^m p(x)$ with $p(\lambda) \neq 0$, then we get at least one Jordan block with size m.

Lemma: Let $m_T(x) = c_T(x) = (x-\lambda)^n$, with $\dim_{\mathbb{F}}(V) = n$. Then, a Jordan basis for V exists.

Proof. Let $w_1 \in V$ with $m_{T,w_1}(x) = m_T(x) = c_T(x)$. Let W_1 be the space generated by T on $\{w_1\}$. We claim $W_1 = V$.

Set $v_n = w_1$ and

$$v_i = (T - \lambda\, \mathrm{id}_V)^{n-i}(v_n).$$

Note that

$$\begin{aligned} v_i &= (T - \lambda\, \mathrm{id}_V)^{n-i}(v_n) \\ &= (T - \lambda\, \mathrm{id}_V)(T - \lambda\, \mathrm{id}_V)^{n-i-1}(v_n) \\ &= (T - \lambda\, \mathrm{id}_V)(v_{i+1}), \end{aligned}$$

meaning $T(v_{i+1}) = v_i + \lambda v_{i+1}$.

We claim that $\{v_1, \dots, v_n\}$ is a basis of V.

Suppose

$$c_1 v_1 + \cdots + c_n v_n = 0_V$$

for some $c_i \in \mathbb{F}$> This gives

$$c_1 (T - \lambda\, \mathrm{id}_V)^{n-1} + \cdots + c_n v_n = 0_V.$$

Set $p(x) = c_1(x-\lambda)^{n-1} + \cdots + c_{n-1}(x-\lambda) + c_n$.

Then,

$$\begin{aligned} p(T)(v_n) &= 0_V \\ &= p(T)(w_1), \end{aligned}$$

meaning

$$m_{T,w_1}(x) \,|\, p(x),$$

but $\deg(m_{T,w_1}(x)) = n$, meaning $p(x) = 0$, so $c_i = 0$ for all i.

Thus, $\{v_1, \dots, v_n\}$ is a Jordan basis. □

Proposition: Let $\dim_{\mathbb{F}}(V) = n$ and $m_T(x) = (x-\lambda)^k$ for some $1 \leqslant k \leqslant n$. Then, a Jordan basis for V exists.

Proof. We have $V = E_\lambda^\infty = E_\lambda^k$. We know the result if $k = n$. Assume $k < n$.

We claim that given any subspace W_1 of V with $W_1 \cap \ker\left((T - \lambda\, \mathrm{id}_V)^{k-1}\right) = \{0_V\}$, there is a T-stable subspace U of V with

$$V = \underbrace{\left(W_1 + (T - \lambda\, \mathrm{id}_V)(W_1) + \cdots + (T - \lambda\, \mathrm{id}_V)^{k-1}(W_1)\right)}_{k \times k \text{ Jordan block}} \oplus U.$$

We know there exists $v_k \in V$ with $(T - \lambda\,\mathrm{id}_V)^{k-1}(v_k) \neq 0_V$. Set $W_1 = \mathrm{span}_{\mathbb{F}}(v_k)$. We have

$$W_1 \cap \ker\left((T-\lambda\,\mathrm{id}_V)^{k01}\right) = \{0_V\}.$$

Write

$$V = W_1 \oplus \ker\left((T-\lambda\,\mathrm{id}_V)^{k-1}\right) \oplus W_2.$$

Note that W_2 consists of other $k \times k$ Jordan block generators, though it can also be 0_V.

Set $W = W_1 \oplus W_2$. We have

$$(T-\lambda\,\mathrm{id}_V)(W) \subseteq \ker\left((T-\lambda\,\mathrm{id}_V)^{k-1}\right).$$

We also have

$$(T-\lambda\,\mathrm{id}_V)(W) \cap \ker\left((T-\lambda\,\mathrm{id}_V)^{k-2}\right) = \{0_V\}.$$

If $w \in (T-\lambda\,\mathrm{id}_V)(W) \cap \ker\left((T-\lambda\,\mathrm{id}_V)^{k-2}\right)$, then

$$w = (T-\lambda\,\mathrm{id}_V)(w_1 + w_2)$$

for $w_i \in W_i$, and

$$(T-\lambda\,\mathrm{id}_V)^{k-2}(w) = 0_V,$$

meaning

$$\begin{aligned}(T-\lambda\,\mathrm{id}_V)^{k-2}(T-\lambda\,\mathrm{id}_V)(w_1+w_2) &= 0_V\\ (T-\lambda\,\mathrm{id}_V)^{k-1}(w_1) + (T-\lambda\,\mathrm{id}_V)^{k-1} w_2 &= 0_V,\end{aligned}$$

implying $w_1 = w_2 = 0_V$, since

$$V = W_1 \oplus W_2 \oplus \underbrace{\ker\left((T-\lambda\,\mathrm{id}_V)^{k-1}\right)}_{\tilde{V}}.$$

Note that $\dim_{\mathbb{F}}\left(\tilde{V}\right) < n$. We also know that $\tilde{V}$ is T-stable.

Let $\tilde{W} = (T-\lambda\,\mathrm{id}_V)(W)$. We have

$$\tilde{W} \cap \ker\left((T-\lambda\,\mathrm{id}_V)^{k-2}\right) = \{0_V\}.$$

We apply the induction hypothesis to $\tilde{V}$ and $\tilde{W}$ to get a T-stable subspace $\tilde{U}$ such that

$$\tilde{V} = \left(\tilde{W} + (T-\lambda\,\mathrm{id}_V)\left(\tilde{W}\right) + \cdots + (T-\lambda\,\mathrm{id}_V)^{k-2}\left(\tilde{W}\right)\right) \oplus \tilde{U}.$$

Define

$$U = \left(W_2 + (T-\lambda\,\mathrm{id}_V)(W_2) + \cdots + (T-\lambda\,\mathrm{id}_V)^{k-1}(W_2)\right) + \tilde{U}.$$

We have U is T-stable. We need to show that

$$V = \left(W_1 + (T-\lambda\,\mathrm{id}_V)(W_1) + \cdots + (T-\lambda\,\mathrm{id}_V)^{k-1}(W_1)\right) \oplus U.$$

We have

$$\begin{aligned}
V &= W_1 + W_2 + \ker\left((T - \lambda\,\mathrm{id}_V)^{k-1}\right)\\
&= W_1 + W_2 + \tilde{V}\\
&= W_1 + W_2 + \left(\tilde{W} + (T - \lambda\,\mathrm{id}_V)\left(\tilde{W}\right) + \cdots + (T - \lambda\,\mathrm{id}_V)^{k-2}\left(\tilde{W}\right)\right) + \tilde{U}\\
&= W_1 + W_2 + \left((W_1 + W_2) + (T - \lambda\,\mathrm{id}_V)(W_1 + W_2) + \cdots + (T - \lambda\,\mathrm{id}_V)^{k-2}(W_1 + W_2)\right) + \tilde{U}\\
&= W_1 + (T - \lambda\,\mathrm{id}_V)(W_1) + \cdots + (T - \lambda\,\mathrm{id}_V)^{k-1}(W_1) + U.
\end{aligned}$$

Let $v \in \left(W_1 + (T - \lambda\,\mathrm{id}_V)(W_1) + \cdots + (T - \lambda\,\mathrm{id}_V)^{k-1}(W_1)\right) \cap U$. Then,

$$v = \sum_{j=0}^{k-1} (T - \lambda\,\mathrm{id}_V)^j \left(w_j\right)$$

for $w_0, \dots, w_{k-1} \in W_1$. Additionally,

$$v = \sum_{j=0}^{k-1} (T - \lambda\,\mathrm{id}_V)^j \left(w_j'\right) + \tilde{u}$$

for $w_0', \dots, w_{k-1}' \in W_2$ and $\tilde{u} \in \tilde{U}$.

Applying $(T - \lambda\,\mathrm{id}_V)^{k-1}$ to both expressions for v, yielding

$$(T - \lambda\,\mathrm{id}_V)^{k-1}(w_0) = (T - \lambda\,\mathrm{id}_V)\left(w_0'\right)$$

since $\tilde{u} \in \ker (T - \lambda\,\mathrm{id}_V)^{k-1}$. Thus,

$$(T - \lambda\,\mathrm{id}_V)^{k-1}\left(w_0 - w_0'\right) = 0_V,$$

meaning $w_0 - w_0' \in \ker\left((T - \lambda\,\mathrm{id}_V)^{k-1}\right)$, and $w_0 - w_0' \in W$, so $w_0 = w_0' \in W_1 \cap W_2 = \{0_V\}$.

To extract the basis, let $W_1 = \operatorname{span}(v_k)$, $v_j = (T - \lambda\,\mathrm{id}_V)^{k-j}(v_k)$, and

$$\mathcal{B}_{\mathcal{W}} = \{v_1, \dots, v_k\}$$

is a Jordan basis for $\mathcal{W} := W_1 + (T - \lambda\,\mathrm{id}_V) W_1 + \cdots + (T - \lambda\,\mathrm{id}_V)^{k-1}(W_1)$. Thus, we have

$$V = \mathcal{W} \oplus U,$$

with U having Jordan basis $\mathcal{B}_U$ by induction. Thus,

$$\mathcal{B} = \mathcal{B}_{\mathcal{W}} \cup \mathcal{B}_U$$

is a Jordan basis for V. □

Thus, we have that for

$$m_T(x) = (x - \lambda)^k,$$

then V has has a Jordan basis with respect to T.

Theorem:

(1) Let $T \in \operatorname{Hom}_{\mathbb{F}}(V, V)$. Suppose $m_T(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k}$ over $\mathbb{F}$. Then, V has a basis $\mathcal{B}$ such that $J = [T]_{\mathcal{B}}$ is in Jordan canonical form. Moreover, J is unique up to order of the Jordan blocks.

(2) Let $A \in \mathrm{Mat}_n(\mathbb{F})$. Suppose $m_A(x) = (x-\lambda_1)^{m_1}\cdots(x-\lambda_k)^{m_k}$ over $\mathbb{F}$. Then, A is similar to a matrix J in Jordan canonical form. The matrix J is unique up to the order of the Jordan blocks.

Proof. We can write

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

We know that $m_{T|_{E_{\lambda_j}^{m_j}}} = \left(x-\lambda_j\right)^{m_j}$, meaning we have a Jordan basis for $T|E_{\lambda_j}^{m_j}$, $\mathcal{B}_j$.

Set

$$\mathcal{B} = \bigcup_{j=1}^{k} \mathcal{B}_j.$$

To show uniqueness, we know that the generators of the $j \times j$ Jordan blocks are

$$\ker\left((T-\lambda_i\,\mathrm{id}_V)^{j-1}\right) \setminus \ker\left((T-\lambda_i\,\mathrm{id}_V)^{j-2}\right).$$

$\square$

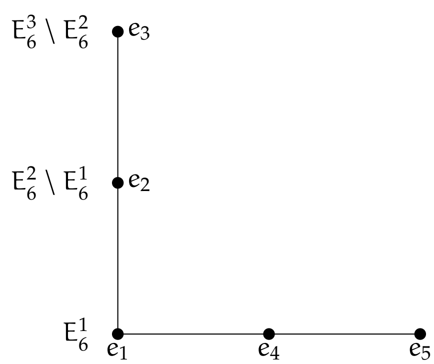
Example. Let

$$A = \begin{pmatrix} 6 & 1 & & & & & & \\ & 6 & 1 & & & & & \\ & & 6 & & & & & \\ & & & 6 & & & & \\ & & & & 6 & & & \\ & & & & & 7 & 1 & \\ & & & & & & 7 & \\ & & & & & & & 7 \end{pmatrix}$$

We have $c_A(x) = (x-6)^5(x-7)^3$.

We have

$$\begin{aligned} E_6^1 &= \mathrm{span}_{\mathbb{F}}(e_1, e_4, e_5) \\ E_6^2 &= \mathrm{span}_{\mathbb{F}}(e_1, e_2, e_3, e_5) \\ E_6^3 &= \mathrm{span}_{\mathbb{F}}(e_1, e_2, e_3, e_4, e_5). \end{aligned}$$



Reading this diagram, the first vertical line denotes the Jordan block of size 3 over e_1, e_2, e_3, then the other two points at e_4 and e_5 denote the Jordan blocks of size 1.

Example. Let

$$A = \begin{pmatrix} 3 & 3 & 0 & 0 & 0 & -1 & 0 & 2 \\ -3 & 4 & 1 & -1 & -1 & 0 & 1 & -1 \\ 0 & 6 & 3 & 0 & 0 & -2 & 0 & 4 \\ -2 & 4 & 0 & 1 & -1 & 0 & 2 & -5 \\ -3 & 2 & 1 & -1 & 2 & 0 & 1 & -2 \\ -1 & 1 & 0 & -1 & -1 & 3 & 1 & -1 \\ -5 & 10 & 1 & -3 & -2 & -1 & 6 & -10 \\ -3 & 2 & 1 & -1 & -1 & 0 & 1 & 1 \end{pmatrix}$$

We have $c_A(x) = (x-2)(x-3)^5\left(x^2 - 6x + 21\right)$. Notice that this does not split over $\mathbb{Q}$, so ith as no Jordan canonical form over $\mathbb{Q}$. However, if we take c_A into the extension field $\mathbb{Q}\left(\sqrt{-3}\right)$, we have

$$c_A(x) = (x-2)(x-3)^5\left(x - \left(3 + 2\sqrt{-3}\right)\right)\left(x - \left(3 - 2\sqrt{-3}\right)\right).$$

From this characteristic polynomial, we then obtain

$$\begin{aligned} E_2^{\infty} &= E_2^1 \\ E_{3\pm 2\sqrt{-3}}^{\infty} &= E_{3\pm 2\sqrt{-3}}^2. \end{aligned}$$

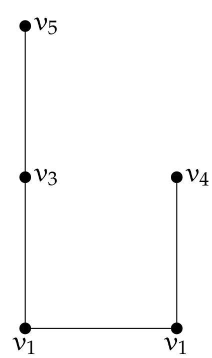
Calculating E_3^{∞}, we get

$$\begin{aligned} \dim_{\mathbb{C}}\left(E_3^1\right) &= \dim_{\mathbb{C}}\left(\ker\left(A - 3I_8\right)\right) \\ &= 2, \end{aligned}$$

we have

$$\dim_{\mathbb{C}}\left(E_3^2\right) = 4,$$

meaning our diagram is



This means there is one Jordan block of size 3 and one Jordan block of size 2 for the generalized eigenvalue 3.

Diagonalization

Theorem: If $c_T(x)$ does not split into a product of linear factors over $\mathbb{F}$, 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 λ, $E_\lambda^\infty = E_\lambda^1$;

(3) $m_T(x)$ 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 $e_\lambda = \dim_{\mathbb{F}}\left(E_\lambda^1\right)$;

(5) if we set set $d_\lambda = \dim_{\mathbb{F}}\left(E_\lambda^1\right)$, then $\sum_\lambda d_\lambda = \dim_{\mathbb{F}}(V)$;

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

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

Tensor Products and Determinants

Extension of Scalars

If V is a $\mathbb{C}$-vector space, then V is also an $\mathbb{R}$-vector space, as we can restrict the scalars of $\mathbb{C}$.

However, we may be interested in the opposite direction. If V is an $\mathbb{R}$-vector space, can we "extend" V to be a $\mathbb{C}$-vector space?

Example (Our First Complexification). Let's start with $V = \mathbb{R}$. We cannot make $\mathbb{R}$ into a $\mathbb{C}$-vector space. However, we do have $V \hookrightarrow \mathbb{C}$ by $x \mapsto x + 0i$, and $\mathbb{C}$ is a $\mathbb{C}$-vector space.

Turning our attention to $\mathbb{C}$, we have $z \in \mathbb{C}$ can be written as $z = x + yi$. We can see that $\mathbb{C}$ is isomorphic to $\mathbb{R}^2 = \mathbb{R} \oplus \mathbb{R}$ as $\mathbb{R}$-vector spaces by

$$x + yi \mapsto \begin{pmatrix} x \\ y \end{pmatrix}.$$

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

$$(a + ib)(x + iy) = (ax - by) + (ay + bx)i,$$

meaning in $\mathbb{R}^2$, we define

$$(a + bi)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax - by \\ ay + bx \end{pmatrix}.$$

With the scalar multiplication as defined above, we have $\mathbb{R} \oplus \mathbb{R}$ is a $\mathbb{C}$-vector space, and $\mathbb{R} \oplus \mathbb{R} \cong \mathbb{C}$ as a $\mathbb{C}$-vector space. We denote this version of $\mathbb{R} \oplus \mathbb{R}$ as $\mathbb{R}_{\mathbb{C}}$. This is known as the complexification of $\mathbb{R}$.

Example (Complexification of a Real Vector Space). Given a real vector space V, we define $V_{\mathbb{C}} = V \oplus V$, defining the complex scalar product by taking

$$(a + bi)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} av_1 - bv_2 \\ av_2 + bv_1 \end{pmatrix}.$$

In particular, via the complexification, this yields $V_{\mathbb{C}}$ as a $\mathbb{C}$-vector space. Notice that

$$i\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -v_2 \\ v_1 \end{pmatrix},$$

meaning that

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ 0_V \end{pmatrix} + \begin{pmatrix} 0_V \\ v_2 \end{pmatrix}$$

$$= \begin{pmatrix} v_1 \\ 0_V \end{pmatrix} + i \begin{pmatrix} v_2 \\ 0_V \end{pmatrix},$$

which looks like $v_1 + iv_2$.

Let $z_1 = x_1 + y_1 i$, $z_2 = x_2 + y_2 i$, $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in V_{\mathbb{C}}$. We want to show that

$$(z_1 z_2) v = z_1 (z_2 v).$$

This yields

$$\begin{aligned} (z_1 z_2) v &= ((x_1 + y_1 i)(x_2 + y_2 i)) v \\ &= ((x_1 x_2 - y_1 y_2) + (x_1 y_2 + x_2 y + 1) i) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\ &= \begin{pmatrix} (x_1 x_2 - y_1 y_2) v_1 - (x_1 y_2 + x_2 y_1) v_2 \\ (x_1 x_2 - y_1 y_2) v_2 + (x_1 y_2 + x_2 y_1) v_1 \end{pmatrix} \\ z_1 (z_2 v) &= z_1 (x_1 + y_1 i) \begin{pmatrix} v_2 \\ v_2 \end{pmatrix} \\ &= z_1 \begin{pmatrix} x_2 v_1 - y_2 v_2 \\ x_2 v_1 + y_2 v_1 \end{pmatrix} \\ &= (x_1 + y_1 i) \begin{pmatrix} x_2 v_1 - y_2 v_2 \\ x_2 v_1 + y_2 v_1 \end{pmatrix} \\ &= \begin{pmatrix} x_1 (x_2 v_1 - y_2 v_2) - y_1 (x_2 v_2 + y_2 v_1) \\ x_1 (x_2 v_2 + y_2 v_1) + y_1 (x_2 v_1 - y_2 v_2) \end{pmatrix}. \end{aligned}$$

Upon simplification, we see that these two expressions are equal.

Exercise: Verify that $V_{\mathbb{C}}$ is a $\mathbb{C}$-vector space.[VI]

We have an embedding $V \hookrightarrow V_{\mathbb{C}}$ by taking $v \mapsto \begin{pmatrix} v \\ 0_V \end{pmatrix}$. The set

$$\left\{ \begin{pmatrix} v \\ 0_V \end{pmatrix} \mid v \in V \right\}$$

is a real subspace of $V_{\mathbb{C}}$.

This method works great for the particular case of $\mathbb{R}$ and $\mathbb{C}$, but we need a different method for more arbitrary vector spaces. Eventually, we will show that for any linear map T, there is a unique map $T_{\mathbb{C}}$.

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \downarrow & & \downarrow \\ V_{\mathbb{C}} & \xrightarrow{T_{\mathbb{C}}} & W_{\mathbb{C}} \end{array}$$

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

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

$$\begin{aligned} w_1 &= \sum_{j \in I} a_j v_j \\ w_1 &= \sum_{j \in I} b_j v_j \end{aligned}$$

[VI]Don't actually do this exercise.

for some $a_i, b_i \in \mathbb{R}$. We have

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

Suppose we have

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

Then,

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

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 $\left\{\left(v_j, 0_V\right)\right\}_{j\in I}$ are linearly independent. □

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

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, hook, "\iota_V"] & W \arrow[d, hook, "\iota_W"] \\
V_{\mathbb{C}} \arrow[r, "T_{\mathbb{C}}"] & W_{\mathbb{C}}
\end{tikzcd}$$

Proof. We define

$$T_{\mathbb{C}}(v_1, v_2) = \left(T(v_1), T(v_2)\right).$$

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

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

$$= (T(v), 0_W),$$

and

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

We claim that $T_{\mathbb{C}}$ is $\mathbb{C}$-linear. Let $x + iy \in \mathbb{C}$, $(v_1, v_2), (v_1', v_2') \in V_{\mathbb{C}}$. Then,

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

Suppose we have $S \in \operatorname{Hom}_{\mathbb{C}}(V_{\mathbb{C}}, W_{\mathbb{C}})$ such that the following diagram commutes.

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, hook, "\iota_V"] & W \arrow[d, hook, "\iota_W"] \\
V_{\mathbb{C}} \arrow[r, "S"] & W_{\mathbb{C}}
\end{tikzcd}$$

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

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

Thus, $S = T_{\mathbb{C}}$, so $T_{\mathbb{C}}$ is unique. $\square$

We are aware that every vector space has a basis. However, we may ask if, given a set Γ, can we build a vector space that has Γ as a basis element?

The answer is yes.

Theorem (Existence of a Free Vector Space)**:** Let $\mathbb{F}$ be a field, and Γ a set. There is an $\mathbb{F}$-vector space, $\mathbb{F}(\Gamma)$, that has X as a basis.

Moreover, $\mathbb{F}(\Gamma)$ has the following universal property: if W is any $\mathbb{F}$-vector space, and $t : \Gamma \to W$ is a map of sets, there is a unique $T \in \operatorname{Hom}_{\mathbb{F}}(\mathbb{F}(\Gamma), W)$ 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, we take $\mathbb{F}(\Gamma) = \{0\}$.

Let $\Gamma \neq \emptyset$. Define

$$\mathbb{F}(\Gamma) = \left\{f : \Gamma \to \mathbb{F} \mid f(x) \text{ finitely supported}\right\}.$$

Let $c \in \mathbb{F}$, $f, g \in \mathbb{F}(\Gamma)$. Then, $(f+g)(x) = f(x) + g(x)$. Since f, g are finitely supported, so too is $f + g$, and similarly $(cf)(x) = cf(x)$ is finitely supported. We can verify that $\mathbb{F}(\Gamma)$ is a vector space with the zero element $f(x) = 0_{\mathbb{F}(\Gamma)}$.

Given any $y \in \Gamma$, define f_y by $f_y(y) = 1$ and $f_y(x) = 0$ for $x \neq y$. Thus, $\Gamma \hookrightarrow \mathbb{F}(\Gamma)$ by $x \mapsto f_x$. We let $\mathcal{X} = \{f_x \mid x \in \Gamma\}$; we let $\iota = \Gamma \xrightarrow{\text{bijection}} \mathcal{X}$.

We claim that $\mathcal{X}$ is a basis for $\mathbb{F}(\Gamma)$. For any $f \in \mathbb{F}(\Gamma)$, we claim $f = \sum_{x \in \Gamma} f(x) f_x$. This gives $\operatorname{span}_{\mathbb{F}}(\mathcal{X}) = \mathbb{F}(\Gamma)$.

Note that

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

Suppose

$$\sum_{i=1}^{n} a_i f_{x_i} = 0_{\mathbb{F}(\Gamma)}.$$

In particular,

$$\sum_{i=1}^{n} a_i f_{x_i}(y) = 0$$

for all $y \in \Gamma$. Thus,

$$\begin{aligned} 0 &= \sum_{i=1}^{n} a_i f_{x_i}\left(x_j\right) \\ &= a_j, \end{aligned}$$

meaning $\{f_x\}_{x \in \Gamma}$ is a basis for $\mathbb{F}(\Gamma)$.

Suppose we have $t : X \to W$. Define

$$T : \mathbb{F}(\Gamma) \to W$$

by

$$\begin{aligned} T\left(\sum_{i=1}^{n} a_i f_{x_i}\right) &= \sum_{i=1}^{n} a_i t\left(f^{-1}(x_i)\right) \\ &= \sum_{i=1}^{n} a_i t(x_i). \end{aligned}$$

This is clearly linear, and makes the diagram commute, and is unique (since T is determined uniquely by the basis elements). $\square$

Example. If $\Gamma = \mathbb{R}$, we can form $\mathbb{F}_{\mathbb{R}}(\mathbb{R})$. Then, we can write any element of $\mathbb{F}_{\mathbb{R}}(\mathbb{R})$ by $2 \cdot \pi + 3 \cdot 2$, where π and 2 are basis elements and $2, 3$ are scalars.

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

Definition (Constructing $K \otimes V$)**.** Let K/F be an extension of fields. Let $X = \{(a, v) \mid a \in K, v \in V\} \in K \times V$, where V is an F-vector space.

Form $\mathbb{F}_K(K \times V)$. Elements of $\mathbb{F}_K(K \times V)$ look like finite sums

$$\sum_{i=1}^{n} c_i (a_i, v_i)$$

with $c_i \in K$, $(a_i, v_i) \in K \times V$. Our goal is, given V an F-vector space, construct a K-vector space that contains V as an F-subspace. We want to shrink $\mathbb{F}_K(K \times V)$ so that we still have the F-structure. We construct $\mathrm{Rel}_K(K \times V)$ such that

- $(a_1 + a_2) * v \sim a_1 * v + a_2 * v$ for $a_1, a_2 \in K, v \in V$;
- $a * (v_1 + v_2) \sim a * v_1 + a * v_2$ for $a \in K, v_1, v_2 \in V$;
- $(ac) * v \sim a * (cv)$ for $a \in K$, $c \in F$, and $v \in V$.

Thus, we define $\mathrm{Rel}_K(K \times V)$ to be the K-span of the following elements of $K \times V$

(1) $(a_1 + a_2, v) - (a_1, v) - (a_2, v)$ for $a_1, a_2 \in K, v \in V$;

(2) $(a, v_1 + v_2) - (a, v_1) - (a, v_2)$ for $a \in K, v_1, v_2 \in V$;

(3) $a_1(a_2, v) - (a_1 a_2, v)$ for $a_1, a_2 \in K, v \in V$;

(4) $(ac, v) - (a, cv)$ for $c \in F$, $a \in K$, and $v \in V$.

This allows $\mathrm{Rel}_K(K \times V)$ to contain all such elements that we expect to be equal under the tensor product. Since $\mathrm{Rel}_K(K \times V)$ is the K-span of a subset of $\mathbb{F}(K \times V)$, the set of relations is a K-subspace of $\mathbb{F}_K(K \times V)$. We define

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

Given $(a, v) \in \mathbb{F}_K(K \times V)$, we write $a \otimes v = (a, v) + \mathrm{Rel}_K(K \times V)$.

We convert the relation $\mathrm{Rel}_K(K \times V)$ into the language of the tensor product.

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

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

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

(4) $ca \otimes v = a \otimes cv$ for $a \in K$, $c \in F$, and $v \in V$.

Elements of $K \otimes_F V$ are of the form

$$\sum_{i \in I} c_i (a_i \otimes v_i) = \sum_{i \in I} b_i \otimes v_i,$$

where $b_i \in K$ and $v_i \in V$. A pure tensor is of the form $a \otimes v$.

The element $0_{K \otimes_F V}$ is of the form $0 \otimes 0_V$.

Finally, we need to verify that V is a F-subspace of $K \otimes V$.

Proposition: Let K/F be a field extension, and V an F-vector space. The K-vector space $K \otimes_F V$ contains a subspace isomorphic to V as an F-vector space.

Proof. Let $\mathcal{B} = \{v_i\}_{i \in I}$ be a basis of V. Define $T : V \to K \otimes_F V$ by taking $v_i \mapsto 1 \otimes v_i$. It is the case that T is a F-linear map.

Let $W = T(V)$. Then, $T : V \to W$ is a surjection.

Let $v \in \ker(T)$, meaning $1 \otimes v = 1 \otimes 0_V$. This means $(1, v) - (0, 0_V) \in \mathrm{Rel}_K(K \times V)$. Thus, $(1, v) - (1, 0_V) \in \mathrm{Rel}_K(K \times V)$. This is only true if $v = 0_V$. Thus, T is also injective. □

We refer to $K \otimes_F V$ as the extension of scalars of V from F to K.

Proposition: Let K/F be a field extension, and V an F-vector space with basis $\mathcal{B} = \{v_i\}_{i \in I}$.

Then, $\mathrm{span}_K \{1 \otimes v_i\}_{i \in I} = K \otimes_F V$.

Proof. Let $a \otimes v \in K \otimes_F V$. Write

$$v = \sum_{i \in I} c_i v_i$$

for $c_i \in F$. Note that we have

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

If we take

$$\sum_{j \in I} a_j \otimes v_j' \in K \otimes_F V,$$

then each $a_j \otimes v_j'$ can be written as a finite linear combination

$$a_j \otimes v_j' = \sum_{i \in I} b_{j_i} (1 \otimes v_i),$$

so

$$\sum_{j \in I} \left(\sum_{i \in I} b_{j_i} (1 \otimes v_i)\right) \in \mathrm{span}_K \{1 \otimes v_i\}_{i \in I}.$$

□

Theorem: Let $\iota_V : V \to K \otimes_F V$ be defined by $\iota_V(v) = 1 \otimes v$.

Let W be any K-vector space, and let $S \in \mathrm{Hom}_F(V, W)$. Then, there is a unique $T \in \mathrm{Hom}_K(K \otimes_F V, W)$ such that $S = T \circ \iota_V$. 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 $T \operatorname{Im} \operatorname{Hom}_K (K \otimes FV, W)$, then $T \circ \iota_V \in \operatorname{Hom}_F (V, W)$.

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

Define $t : K \times V \to W$, $(a, v) \mapsto aS(v)$.

The universal property for $\mathbb{F}(K \times V)$ gives a unique linear map $T : \mathbb{F}(K \times V) \to W$, given by $T(a, v) = t(a, v)$.

Since T is linear, we have

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

All we need to do now is show that T vanishes on $\operatorname{Rel}_K (K \times V)$. For instance, we need to show that T vanishes on

$$(a + b, v) - (a, v) - (b, v),$$

or

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

Thus, the diagram commutes.

We have shown that $\{1 \otimes v\}$ spans $K \otimes_F V$. To determine a K-linear map on $K \otimes_F V$, it is enough to see what T does to $1 \otimes v$.

Since $T(1 \otimes v) = S(v)$, if $\tilde{T}$ also made the diagram commute, then $S(v) = \tilde{(}1 \otimes v)$, meaning $T = \tilde{T}$. Thus, T is unique. ◻

Proposition: Let K/F be an extension of fields. Then,

$$K \otimes F \cong K$$

Proof. We have $i : F \hookrightarrow K$ by inclusion. Our diagram is

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

The universal property gives $T \in \operatorname{Hom}_K (K \otimes_F F, K)$ such that for $x \in F$, $T(1 \otimes x) = i(x) = x$.

For any

$$\sum a_i \otimes x_i \in K \otimes_F F,$$

since T is K-linear, we have

$$\begin{aligned} T\left(\sum a_i \otimes x_i\right) &= \sum T\left(a_i \otimes x_i\right) \\ &= \sum T\left(a_i\left(1 \otimes x_i\right)\right) \\ &= \sum a_i T\left(1 \otimes x_i\right) \\ &= \sum a_i x_i. \end{aligned}$$

Define $S : K \to K \otimes F$, $y \mapsto y \otimes 1$. For $a \in K$, $y_1, y_2 \in K$, we have

$$\begin{aligned} S\left(y_1 + a y_2\right) &= \left(y_1 + a y_2\right) \otimes 1 \\ &= y_1 \otimes 1 + a\left(y_2 \otimes 1\right) \\ &= S\left(y_1\right) + a S\left(y_2\right), \end{aligned}$$

meaning S is a K-linear map. Thus, $S \in \operatorname{Hom}_K\left(K \otimes_F F\right)$. We have

$$\begin{aligned} T \circ S(y) &= T\left(y \otimes 1\right) \\ &= y T\left(1 \otimes 1\right) \\ &= y, \end{aligned}$$

and

$$\begin{aligned} S \circ T\left(a \otimes x\right) &= S\left(T\left(a \otimes x\right)\right) \\ &= S\left(a T\left(1 \otimes x\right)\right) \\ &= S\left(a x\right) \\ &= a x \otimes 1 \\ &= a \otimes x. \end{aligned}$$

Thus, T is an isomorphism. □

Remark: This shows that $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R} \cong \mathbb{C}$.

Proposition: Let K/F be an extension of fields, and let V be 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 map $K \otimes_F V \to K^n$.

We will define a map $T : V \to K^n$ as follows. Set $\mathcal{B} = \{v_1, \dots, v_n\}$ to be a basis for V, and $\mathcal{E} = \{e_1, \dots, e_n\}$ as the standard basis for K^n. Define $t : \mathcal{B} \to \mathcal{E}$ by taking $v_i \mapsto e_i$. Since t is defined on the bases, t extends to a linear map $T : V \to K^n$. Thus, $T \in \operatorname{Hom}_F\left(V, K^n\right)$. The universal property for tensor products gives a K-linear map $\overline{T} \in \operatorname{Hom}_K\left(K \otimes_F V, K^n\right)$, where

$$\overline{T}\left(1 \otimes v_i\right) = e_i.$$

We can define $S \in \operatorname{Hom}_K\left(K^n, K \otimes_F V\right)$ by taking

$$S\left(e_i\right) = 1 \otimes v_i.$$

Since S and $\overline{T}$ are inverses of each other, $\overline{T}$ is an isomorphism, so $K \otimes_F V \cong K^n$.

Moreover, since S is an isomorphism, and $\{e_i\}_{i=1}^n$ is a basis for K^n, the collection $\{1 \otimes v_i\}_{i=1}^n$ forms a basis for $K \otimes_F V$. □

Proposition: Let K/F be an extension of fields. Let $\mathcal{B} = \{v_i\}_{i\in I}$ be a F-basis for V. Then, $\mathcal{B}_K = \{1 \otimes v_i\}_{i\in I}$ is a basis for $K \otimes_F V$.

Proof. We know that $\{1 \otimes v_i\}_{i\in I}$ is spanning for $K \otimes_F V$. Suppose

$$\sum_{i\in I} c_i (1 \otimes v_i) = 0_{K\otimes_F V}.$$

for some $c_i \in K$.

Fix $i_0 \in I$. Define

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

by $v \mapsto a_{i_0}$. This is a F-linear map.

$$\begin{aligned} t_{i_0}(v + c) &= t_{i_0}\left(\sum_{i\in I} \left(a_i c a_i'\right) v_i\right) \\ &= a_{i_0} + c a_{i_0}' \\ &= t_{i_0}(v) + c t_{i_0}(). \end{aligned}$$

Thus, $t \in \operatorname{Hom}_F(V, K)$. By the universal property of tensor products, we get a map $T_{i_0} \in \operatorname{Hom}_K(K \otimes_F V, K)$ such that

$$\begin{aligned} T_{i_0}(1 \otimes v) &= t_{i_0}(v) \\ &= a_{i_0}. \end{aligned}$$

Thus, we have

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

Thus, $\{1 \otimes v_i\}_{i\in I}$ is linearly independent. $\square$

Theorem: Let K/F be an extension of fields, and let V, W be F-vector spaces. Let $T \in \operatorname{Hom}_F(V, W)$. There is a map $T_K \in \operatorname{Hom}_K(K \otimes_F V, K \otimes_F W)$ such 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 a map $t : V \hookrightarrow K \otimes_F W$ by $v \mapsto 1 \otimes T(v)$. We have

$$\begin{aligned} t(v_1 + cv_2) &= 1 \otimes T(v_1 + cv_2) \\ &= 1 \otimes (T(v_1) + cT(v_2)) \\ &= 1 \otimes T(v_1) + c(1 \otimes T(v_2)) \end{aligned}$$

$$= t(v_1) + ct(v_2).$$

This gives $t \in \operatorname{Hom}_F(V, K \otimes_F W)$. The universal property for tensor products gives $T_K \in \operatorname{Hom}_K(K \otimes_F V, K \otimes_F W)$ satisfying

$$\begin{aligned} T_K(1 \otimes v) &= t(v) \\ &= 1 \otimes T(v). \end{aligned}$$

Let $v \in V$. Then,

$$\begin{aligned} T_K(\iota_V(v)) &= T_K(1 \otimes v) \\ &= 1 \otimes T(v) \\ &= \iota_W(T(v)). \end{aligned}$$

So, the diagram commutes. □

Remark: This shows that $\mathbb{C} \otimes_{\mathbb{R}} V \cong V_{\mathbb{C}}$.

Tensor Products of Vector Spaces and the Trace

Example. In multivariable calculus,

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R}^n &\xrightarrow{\cdot} \mathbb{R} \\ \left(\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}\right) &\mapsto \sum_{i=1}^{n} a_i b_i. \end{aligned}$$

Some of the properties we like about the dot product are as follows.

- $(v_1 + v_2) \cdot w = v_1 \cdot w + v_2 \cdot w$;
- $v \cdot (w_1 + w_2) = v \cdot w_1 + v \cdot w_2$;
- $c(v \cdot w) = cv \cdot w = v \cdot (cw)$.

This is an example of a bilinear form.

Example. In the case of

$$\begin{aligned} \mathbb{R}^3 \times \mathbb{R}^3 &\xrightarrow{\times} \mathbb{R}^3 \\ (v, w) &\mapsto v \times w, \end{aligned}$$

we have the following properties.

- $(v_1 + v_2) \times w = v_1 \times w + v_2 \times w$;
- $v \times (w_1 + w_2) = v \times w_1 + v \times w_2$;
- $c(v \times w) = (cv) \times w = v \times (cw)$.

This is an example of a bilinear form, this time not mapping to the scalar field.

Rephrasing the above two examples, if we let $t : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $t(v, w) = v \times w$, the above properties become the following.

- $t(v_1 + v_2, w) = t(v_1, w) + t(v_2, w)$;
- $t(v, w_1 + w_2) = t(v, w_1) + t(v, w_2)$;

- $ct(v,w) = t(cv,w) = t(v,cw)$.

Example. If we let V be an $\mathbb{F}$-vector space, and define $t : \mathbb{F}\times V \to V$, $(a,v) \mapsto av$, this is also a bilinear form.

Definition (Bilinear Map)**.** Let U, V, W be $\mathbb{F}$-vector spaces. Let $t : V \times W \to U$ be a map satisfying

- $t(v_1 + v_2, w) = t(v_1, w) + t(v_2, w)$;
- $t(v, w_1 + w_2) = t(v, w_1) + t(v, w_2)$;
- $ct(v,w) = t(cv,w) = t(v,cw)$.

We call such a map a bilinear map. The collection of bilinear maps is denoted $\operatorname{Hom}_{\mathbb{F}}(V, W; U)$.

We want to construct a vector space that contains $V \times W$, but treats V and W as separate vector spaces to "linearize" the bilinear map.

Let $X = V \times W$ as a set. We will form the vector space $\mathbb{F}(V \times W)$. We form $\operatorname{Rel}_{\mathbb{F}}(V \times W)$ to be the $\mathbb{F}$-span of $E = E_1 \cup E_2 \cup E_3 \cup E_4$, with

$$\begin{aligned}
E_1 &= \{(v_1 + v_2, w) - (v_1, w) - (v_2, w) \mid v_1, v_2 \in V, w \in W\}\\
E_2 &= \{(v, w_1 + w_2) - (v, w_1) - v(w_2) \mid v \in V, w_1, w_2 \in W\}\\
E_3 &= \{(cv, w) - (v - cw) \mid c \in \mathbb{F}, v \in V, w \in W\}\\
E_4 &= \{c(v, w) - (cv, w) \mid c \in \mathbb{F}, v \in V, w \in W\}.
\end{aligned}$$

We define $V \otimes W = \mathbb{F}(V \times W) / \operatorname{Rel}_{\mathbb{F}}(V \times W)$. We write $v \otimes w = (v, w) + \operatorname{Rel}_{\mathbb{F}}(V \times W)$. We have

- $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$;
- $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$;
- $(cv) \otimes w = v \otimes (cw)$;
- $c(v \otimes w) = (cv) \otimes w = v \otimes (cw)$.

Elements of $V \otimes W$ look like

$$\begin{aligned}
\sum_{i\in I} c_i (v_i \otimes w_i) &= \sum_{i\in I} (cv_i) \otimes w_i\\
&= \sum_{i\in I} v_i \otimes w_i.
\end{aligned}$$

We have

$$\begin{aligned}
\iota : V \times W &\hookrightarrow V \otimes W\\
(v, w) &\mapsto v \otimes w.
\end{aligned}$$

Exercise: Show $\iota \in \operatorname{Hom}_{\mathbb{F}}(V, W; V \otimes W)$.

Let U be another $\mathbb{F}$-vector space, and let $T \in \operatorname{Hom}_{\mathbb{F}}(V \otimes W, U)$. We have

$$\begin{aligned}
T((v_1 + v_2) \otimes w) &= T(v_1 \otimes w + v_2 \otimes w)\\
&= T(v_1 \otimes w) + T(v_2 \otimes w)\\
T(v \otimes (w_1 + w_2)) &= T(v \otimes w_1 + v \otimes w_2)\\
&= T(v \otimes w_1) + T(v \otimes w_2)\\
cT(v \otimes w) &= T(c(v \otimes w))\\
&= T((cv) \otimes w)\\
&= T(v \otimes (cw)).
\end{aligned}$$

Theorem (Universal Property for Tensor Products)**:** Let U, V, W be F-vector spaces.

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

(2) If $t \in \operatorname{Hom}_F(V, W; U)$, there is a unique $T \in \operatorname{Hom}_F(V \otimes W, U)$ such that $t = T \circ \iota$ — i.e., the following diagram commutes.

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

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

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

The same arguments works in the second variable. Thus, $t \in \operatorname{Hom}_F(V, W; U)$.

Let $t \in \operatorname{Hom}_F(V, W; U)$. This says $t : V \times W \to U$; the universal property for $\mathbb{F}(V \times W)$ gives 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 $\pi : \mathbb{F}(V \times W) \to \mathbb{F}(V \times W) / \operatorname{Rel}_F(V \times W)$, all we need to show is that T vanishes on $\operatorname{Rel}_F(V \times W)$.

Since T does vanish on $\operatorname{Rel}_F(V \times W)$, we have

$$\begin{tikzcd} & (v,w) \arrow[r, mapsto] & v \otimes w \arrow[dr, mapsto, bend left] & \\ (v,w) \arrow[ddrr, mapsto] & V \times W \arrow[r, hook, "\iota"] \arrow[dr, "t"'] & V \otimes W \arrow[d, "T"] & \\ & & U & T(v \otimes w) \\ & & t(v,w) & \end{tikzcd}$$

Since the diagram commutes, we must have $T(v \otimes w) = t(v, w)$. ◻

Corollary: Let U, V, W be F-vector spaces.

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

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

Proof. We will prove (2).

Fix $w \in W$. Define $t : U \times V \to U \otimes_F (V \otimes_F W)$, $(u, v) \mapsto u \otimes (v \otimes w)$. We claim that t is bilinear.

Let $u_1, u_2 \in U$, $v \in V$, $c \in F$. Then, we have

$$\begin{aligned}
t\left((u_1 + cu_2, v)\right) &= (u_1 + cu_2) \otimes (v \otimes w) \\
&= u_1 \otimes (v \otimes w) + c\left(u_2 \otimes (v \otimes w)\right) \\
&= t\left((u_1, v)\right) + ct\left((u_2, v)\right).
\end{aligned}$$

Linearity in the second variable follows similarly.

Since t is bilinear, the universal property gives a linear map $T : U \otimes_F V \to U \otimes_F (V \otimes_F W)$ by $u \otimes v \mapsto u \otimes (v \otimes w)$.

Define $s : (U \otimes_F V) \times W \to U \otimes_F (V \otimes_F W)$ by taking $(u \otimes v, w) \mapsto u \otimes (v \otimes w)$. Let $u_1 \otimes v_1 = u_2 \otimes v_2$. Then,

$$\begin{aligned}
s\left((u_1 \otimes v_1, w)\right) &= u_1 \otimes (v_2 \otimes w) \\
&= T(u_1 \otimes v_1) \\
&= T(u_2 \otimes v_2) \\
&= u_2 \otimes (v_2 \otimes w) \\
&= S\left((u_2 \otimes v_2, w)\right).
\end{aligned}$$

Additionally, s is a bilinear map, meaning that the universal property of tensor products gives a unique linear map $S : (U \otimes_F V) \otimes_F W \to U \otimes_F (V \otimes_F W)$, mapping $(u \otimes v) \otimes w \mapsto u \otimes (v \otimes w)$.

We do the same in the opposite direction. Fix $u \in U$.

$$\begin{aligned}
\tilde{t} : V \times W &\to (U \otimes_F V) \otimes_F W \\
(v, w) &\mapsto (u \otimes v) \otimes w.
\end{aligned}$$

Continuing the process, we get $\tilde{S} : U \otimes_F (V \otimes_F W) \to (U \otimes_F V) \otimes_F W$, mapping $u \otimes (v \otimes w) \mapsto (u \otimes v) \otimes w$. These are inverses of each other. $\square$

Theorem: Let U, V, W be F-vector spaces. There is an isomorphism

$$\begin{aligned}
(U \oplus V) \otimes_F W &\xrightarrow{\cong} (U \otimes_F W) \oplus (V \otimes_F W) \\
(u, v) \otimes w &\mapsto (u \otimes w, v \otimes w)
\end{aligned}$$

Proof. Define $t : (U \oplus V) \times W \to (U \otimes_F W) \oplus (V \otimes_F W)$ by taking

$$((u, v), w) \mapsto (u \otimes w, v \otimes w).$$

This is a bilinear map.

Thus, there is a unique linear map

$$T : (U \oplus V) \otimes_F W \to (U \otimes_F W) \oplus (V \otimes_F W)$$

mapping $(u, v) \otimes w \mapsto (u \otimes w, v \otimes w)$.

To find an inverse map, we define two maps

$$\begin{aligned}
s_1 : U \times W &\to (U \oplus V) \otimes_F W \\
(u, w) &\mapsto (u, 0_V) \otimes w
\end{aligned}$$

$$\begin{aligned}
s_2 : V \times W &\to (U \oplus V) \otimes_F W \\
(v, w) &\mapsto (0_U, v) \otimes w.
\end{aligned}$$

Let $u_1, u_2 \in U, w \in W, c \in F$. Then,

$$\begin{aligned}
s_1\left((u_1 + cu_2, w)\right) &= (u_1 + cu_2, 0_V) \otimes w\\
&= \left((u_1, 0_V) + c\,(u_2, 0_V)\right) \otimes w\\
&= (u_1, 0_V) \otimes w + c\,(u_2, 0_V) \otimes w\\
&= s_1\left((u_1, w)\right) + cs_1\left((u_2, w)\right)
\end{aligned}$$

Similarly, s_2 is bilinear, meaning we have well-defined linear maps

$$\begin{aligned}
S_1 : U \otimes_F W &\longrightarrow (U \oplus V) \otimes_F W\\
u \otimes w &\longmapsto (u, 0_V) \otimes w
\end{aligned}$$

$$\begin{aligned}
S_2 : V \otimes_F W &\longrightarrow (U \oplus V) \otimes_F W\\
v \otimes w &\longmapsto (0_U, v) \otimes_F W.
\end{aligned}$$

Define

$$\begin{aligned}
S : (U \otimes_F W) \oplus (V \otimes_F W) &\longrightarrow (U \oplus V) \otimes_F W\\
(u \otimes w_1, v \otimes w_2) &\longmapsto S_1\left(u \otimes w_1\right) + S_2\left(v \otimes w_2\right)
\end{aligned}$$

We have

$$\begin{aligned}
S \circ T\left((u, v) \otimes w\right) &= S\left((u \otimes w, v \otimes w)\right)\\
&= (u, 0_V) \otimes w + (0_U, v) \otimes w\\
&= \left((u, 0_V) + (0_U, v)\right) \otimes w\\
&= (u, v) \otimes w.
\end{aligned}$$

Similarly,

$$\begin{aligned}
T \circ S\left((u \otimes w_1, v \otimes w_2)\right) &= T\left((u, 0_V) \otimes w_1 + (0_U, v) \otimes w_2\right)\\
&= T\left((u, 0_V) \otimes w_1\right) + T\left((0_U, v) \otimes w_2\right)\\
&= \left(u \otimes w_1, 0_V \otimes w_1\right) + \left(0_U \otimes w_2 + v \otimes w_2\right)\\
&= \left(u \otimes w_1, v \otimes w_2\right).
\end{aligned}$$

□

Corollary (Bases of Tensor Products)**:** Let V, W be finite-dimensional F-vector spaces with bases $\mathcal{B} = \{v_1, \dots, v_m\}$ in V and $\mathcal{C} = \{w_1, \dots, w_n\}$ respectively.

Then, the collection

$$\mathcal{D} = \left\{v_i \otimes w_j\right\}_{\substack{1 \leqslant i \leqslant m\\ 1 \leqslant j \leqslant n}}$$

is a basis for $V \otimes_F W$. In particular, $\dim\left(V \otimes_F W\right) = \dim_F(V)\dim_F(W)$.

Proof. We define $t : V \times W \longrightarrow \operatorname{Mat}_{m,n}(F)$, mapping

$$\left(v_i, w_j\right) \longmapsto e_{ij},$$

where e_{ij} is the matrix with 1 in the ij position and 0 everywhere else.

Let $v \in V$ and $w \in W$. We write

$$v = \sum_{i=1}^{m} a_i v_i$$

$$w = \sum_{j=1}^{n} b_j w_j,$$

and define

$$t(v, w) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i b_j e_{ij}.$$

This is, by definition, bilinear. Thus, there is a unique linear map

$$T : V \otimes_F W \to \mathrm{Mat}_{m,n}(F)$$

such that $v_i \otimes w_j \mapsto e_{ij}$.

Define $S : \mathrm{Mat}_{m,n}(F) \to V \otimes_F W$ by $e_{ij} \mapsto v_i \otimes w_j$. Since $\{e_{ij}\}$ is a basis for $\mathrm{Mat}_{m,n}(F)$, it is the case that T is an isomorphism, so $\dim(T) = \dim(\mathrm{Mat}_{m,n}(F)) = mn$.

Since S is an isomoprhism, and $S(\{e_{ij}\}) = \{v_i \otimes w_j\}$, it is the case that $\{v_i \otimes w_j\}$ is a basis for $V \otimes_F W$. $\square$

Example. We have

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} \cong \mathbb{R}^4,$$

with basis $\{1 \otimes 1, 1 \otimes i, i \otimes 1, i \otimes i\}$. Meanwhile,

$$\mathbb{C} \otimes_{\mathbb{C}} \mathbb{C} \cong \mathbb{C},$$

with basis $\{1 \otimes 1\}$.

Theorem: Let V, W be F-vector spaces. Let $\mathcal{B}_V = \{v_i\}_{i \in I}$ and $\mathcal{B}_W = \{w_j\}_{j \in I}$ be bases. The set $\mathcal{B} = \{v_i \otimes w_j\}_{i,j \in I}$ is a basis for $V \otimes_F W$.

Proof. Let $v \in V$ and $w \in W$. We can write

$$v = \sum_{i \in I} a_i v_i$$
$$w = \sum_{j \in I} b_j w_j,$$

where the sums are finite.

Then,

$$v \otimes w = \sum_{i,j \in I} a_i b_j (v_i \otimes w_j),$$

so $\mathcal{B} = \{v_i \otimes w_j\}_{i,j \in I}$ is spanning for $V \otimes_F W$.

Suppose we can write

$$\sum_{i,j \in I} c_{i,j} (v_i \otimes w_j) = 0_{V \otimes_F W}$$

for some $c_{i,j} \in F$ as a finite sum.

Fix $(i_0, j_0) \in I \times I$. Define

$$\begin{aligned} t_{i_0,j_0} &: V \times W \to F \\ (v_i, w_j) &\mapsto \delta_{(i,j)(i_0,j_0)}. \end{aligned}$$

Note that

$$\begin{aligned} t_{i_0,j_0}\left(\sum_{i\in I} a_i v_i, \sum_{j\in I} b_j w_j\right) &= \sum_{i\in I}\sum_{j\in I} a_i b_j t_{i_0,j_0}\left(v_i, w_j\right) \\ &= a_{i_0} b_{i_0}. \end{aligned}$$

Therefore, there is $T_{i_0,j_0} \in \mathrm{Hom}_F\left(V \otimes_F W, F\right)$ with

$$T_{i_0,j_0}\left(v_i \otimes w_j\right) = \delta_{\left(i_j\right)(i_0,j_0)}.$$

Therefore, we have

$$\begin{aligned} 0 &= T_{i_0,j_0}\left(0_{V\otimes_F W}\right) T \\ &=_{i_0,j_0}\left(\sum_{i,j\in I} c_{i,j} v_0 \otimes w_j\right) \\ &= c_{i_0,j_0} \\ &= 0, \end{aligned}$$

for each (i_0, j_0) in the sum, so $\left\{v_i \otimes w_j\right\}_{i,j\in I}$ is linearly independent. □

Definition (Trace). Let $A \in \mathrm{Mat}_n(F)$, $A = \left(a_{ij}\right)$. Then,

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Recall that

$$A = \left[T_A\right]_{\mathcal{B}}$$

for some $\mathcal{B}$. If trace is to mean anything, we should be able to define the trade to be basis-independent — that is,

$$\begin{aligned} \mathrm{tr}(T) &= \mathrm{tr}\left([T]_{\mathcal{B}_1}\right) \\ &= \mathrm{tr}\left([T]_{\mathcal{B}_2}\right) \end{aligned}$$

for different bases $\mathcal{B}_1$ and $\mathcal{B}_2$.

It may seem suspect that "summing the diagonal" is independent of choice of basis. Therefore, we want to define the trace to be basis-independent, then we will show that the "summing the diagonal" definition yields.

The trace should be defined

$$\mathrm{tr} : \mathrm{Hom}_F(V, V) \to F.$$

We need to bring the tensor product into this, defining a map on $\mathrm{Hom}_F(V, V)'$ is difficult.

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

Proof. Let

$$\begin{aligned} t : V \times V' &\to \operatorname{Hom}_F(V, V) \\ (v, \varphi) &\mapsto (w \mapsto \varphi(w) v). \end{aligned}$$

Let $v_1, v_2 \in V$, $c \in F$. Then,

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

meaning that the map $\mapsto \varphi(w) v$ is indeed a linear map.

We want to show that t is bilinear.

$$t(v_1 + cv_2, \varphi) = t(v_1, \varphi) + ct(v_2, \varphi),$$

and similarly,

$$t(v, \varphi_1 + c\varphi_2) = t(v, \varphi_1) + ct(v, \varphi_2).$$

We really need to show that

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

Computing, we have

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

Similarly,

$$t(v, \varphi_1 + c\varphi_2)(w) = t(v, \varphi_1)(w) + ct(v, \varphi_2)(w).$$

Thus, there is a well-defined map

$$\begin{aligned} \mathcal{T} : V \otimes_F V' &\to \operatorname{Hom}_F(V, V) \\ v \otimes \varphi &\mapsto (w \mapsto \varphi(w)v). \end{aligned}$$

In particular, we have

$$\mathcal{T}(v \otimes \varphi)(w) = \varphi(w) v.$$

Since both $V \otimes V'$ and $\operatorname{Hom}_F(V, V)$ have dimension n^2, it is enough to show that $\mathcal{T}$ is injective.

Let $\mathcal{B} = \{v_1, \dots, v_n\}$ be a basis for V, and $\mathcal{B}' = \left\{v_1', \dots, v_n'\right\}$ a basis for V'. Suppose we have

$$\mathcal{T}\left(\sum_{i,j} a_{i,j}\left(v_i \otimes v_j'\right)\right) = 0_{\operatorname{Hom}_F(V,V)}$$

for some $a_{i,j} \in \mathbb{F}$. Take $v_m \in \mathcal{B}$. Then, we have

$$0_V = \mathcal{T}\left(\sum_{i,j} a_{i,j}\left(v_i \otimes v_j'\right)\right)(v_m)$$

$$\begin{aligned}
&= \sum_{i,j} a_{i,j}\mathcal{T}\left(v_i \otimes v'_j\right)(v_m)\\
&= \sum_{i,j} a_{i,j} v'_j (v_m) v_i\\
&= \sum_{i} a_{i,m} v_i,\\
a_{i,m} &= 0.
\end{aligned}$$

However, since m was arbitrary, we have $a_{i,j} = 0$ for each i, j, so $\mathcal{T}$ is injective, hence $\mathcal{T}$ is an isomorphism. □

Recall that we have

$$\begin{aligned}
\operatorname{Hom}_F(V,V) \times \operatorname{Hom}_F(V,V) &\to \operatorname{Hom}_F(V,V)\\
(S,T) &\xmapsto{\text{comp}} S \circ T.
\end{aligned}$$

Therefore, we have the following diagram

$$\begin{tikzcd}
(V \otimes V') \times (V \otimes V') \arrow[r, dashed, "\Phi"] \arrow[d, "\mathcal{T}\times\mathcal{T}"'] & V \otimes V' \arrow[d, "\mathcal{T}"]\\
\operatorname{Hom}_F(V,V) \times \operatorname{Hom}_F(V,V) \arrow[r, "\text{comp}"] & \operatorname{Hom}_F(V,V)
\end{tikzcd}$$

We define

$$\begin{aligned}
\Phi : (V \otimes_F V') \times (V \otimes_F V') &\longrightarrow V \otimes_F V'\\
(v \otimes \varphi, w \otimes \psi) &\longmapsto \varphi(w) v \otimes \psi.
\end{aligned}$$

We need to verify that this map allows the diagram to commute. Let $x \in V$. Then, we have

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

In the other direction, we have

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

Indeed, the diagram does commute, and Φ is our map that corresponds to composition of functions.

Returning to the trace, let $T \in \operatorname{Hom}_F(V,V)$, with $\mathcal{B} = \{v_1, \dots, v_n\}$ a basis for V, and we write

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

Note that

$$a_{ij} = v'_i\left(T\left(v_j\right)\right).$$

In particular, we have

$$\operatorname{Tr}(A) = \sum_{i=1}^{n} a_{ii}$$

$$= \sum_{i=1}^{n} v_i' \left(T\left(v_i\right)\right)$$

Let

$$\begin{aligned} s : V \times V' &\to F \\ (v, \varphi) &\mapsto \varphi(v). \end{aligned}$$

This map is bilinear, meaning we have

$$\begin{aligned} S : V \otimes V' &\to F \\ v \otimes \varphi &\mapsto \varphi(v). \end{aligned}$$

Thus, we have the map

$$\begin{tikzcd} \operatorname{Hom}_F(V, V) \arrow[r, "\mathrm{tr}=S\circ T^{-1}"] & F \\ V \otimes V' \arrow[u, "T"] \arrow[ur, "S"'] & \end{tikzcd}$$

We know that $\mathcal{T}\left(v_i \otimes v_j'\right) = T_{ij} \in \operatorname{Hom}_F(V, V)$. Since $\mathcal{T}$ is an isomorphism, we know that $\left\{T_{ij} \mid T_{ij} = \mathcal{T}\left(v_i \otimes v_j\right)\right\}_{i,j}$ is a basis of $\operatorname{Hom}_F(V, V)$.

We take

$$\begin{aligned} \operatorname{Tr}\left(T_{k,\ell}\right) &= \operatorname{Tr}\left(\mathcal{T}\left(v_k \otimes v_\ell'\right)\right) \\ &= \sum_{i=1}^{n} v_i'\left(\mathcal{T}\left(v_k \otimes v_\ell\right)\left(v_i\right)\right) \\ &= \sum_{i=1}^{n} v_i'\left(v_\ell'\left(v_i\right) v_k\right) \\ &= \sum_{i=1}^{n} v_\ell'\left(v_i\right) v_i'\left(v_k\right) \\ &= v_\ell'\left(v_k\right) \\ &= \begin{cases} 1 & k = \ell \\ 0 & \text{else} \end{cases}. \end{aligned}$$

We also have

$$\begin{aligned} \operatorname{tr}\left(T_{k,\ell}\right) &= \left(S \circ \mathcal{T}^{-1}\right)\left(T_{k,\ell}\right) \\ &= \left(S \circ \mathcal{T}^{-1}\right)\left(\mathcal{T}\left(v_k \otimes v_\ell'\right)\right) \\ &= S\left(v_k \otimes v_\ell'\right) \\ &= v_\ell'\left(v_k\right) \\ &= \begin{cases} 1 & k = \ell \\ 0 & \text{else} \end{cases}. \end{aligned}$$

Thus, Tr = tr as they agree on the basis elements.

We immediately see that Tr is a linear map.

Corollary: Let $A, B \in \mathrm{Mat}_n(F)$. We have $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

Proof. Define

$$\begin{aligned} t_1 : \mathrm{Hom}_F(V,V) \times \mathrm{Hom}_F(V,V) &\to F \\ (S,T) &\mapsto \mathrm{tr}(S \circ T) \end{aligned}$$

$$\begin{aligned} t_2 : \mathrm{Hom}_F(V,V) \times \mathrm{Hom}_F(V,V) &\to F \\ (S,T) &\mapsto \mathrm{tr}(T \circ S). \end{aligned}$$

These are both bilinear maps, meaning we have maps

$$\begin{aligned} T_1 : \mathrm{Hom}_F(V,V) \otimes \mathrm{Hom}_F(V,V) &\to F \\ (S \otimes T) &\mapsto \mathrm{tr}(S \circ T) \end{aligned}$$

$$\begin{aligned} T_2 : \mathrm{Hom}_F(V,V) \otimes \mathrm{Hom}_F(V,V) &\to F \\ (S \otimes T) &\mapsto \mathrm{tr}(T \circ S). \end{aligned}$$

Let $v \otimes \varphi, w \otimes \psi \in V \otimes V'$. We need to show that

$$\mathrm{tr}(\mathcal{T}(v \otimes \varphi) \circ \mathcal{T}(w \otimes \psi)) = \mathrm{tr}(\mathcal{T}(w \otimes \psi) \circ \mathcal{T}(v \otimes \varphi)).$$

Recall that $\mathcal{T}(v \otimes \varphi) \circ \mathcal{T}(w \otimes \psi) \leftrightarrow \varphi(w)v \otimes \psi$, so we have

$$\begin{aligned} \mathrm{tr}(\mathcal{T}(v \otimes \varphi) \circ \mathcal{T}(w \otimes \psi)) &= \mathrm{tr}(\varphi(w)v \otimes \psi) \\ &= \varphi(w)\,\mathrm{tr}(v \otimes \psi) \\ &= \varphi(w)\psi(v) \\ \mathrm{tr}(\mathcal{T}(w \otimes \psi) \circ \mathcal{T}(v \otimes \varphi)) &= \mathrm{tr}(\psi(v)w \otimes \varphi) \\ &= \psi(v)\,\mathrm{tr}(w \otimes \varphi) \\ &= \varphi(w)\psi(v). \end{aligned}$$

□

Tensor Algebras, Exterior Algebras, and the Determinant

Now that we understand the trace, we want to build some more structure to understand the determinant.

Recall that we have

$$(U \otimes V) \otimes W \cong U \otimes (V \otimes W),$$

so we may write

$$U \otimes V \otimes W.$$

By induction, given $V_1, \dots, V_n$, we may write

$$V_1 \otimes \cdots \otimes V_n$$

as a unique vector space up to isomorphism.

Elements in $V_1 \otimes \cdots \otimes V_n$ look like

$$\sum a_{i_1,\dots,i_n}(v_{i_1} \otimes \cdots \otimes v_{i_n}),$$

where $v_{i_j} \in V_j$.

We write

$$\mathcal{T}^k(V) = \underbrace{V \otimes \cdots \otimes V}_{k \text{ copies}}$$

Definition. Let $V_1, \dots, V_n$ be F-vector spaces. A map

$$t : V_1 \times \cdots \times V_n \to W$$

is said to be multilinear if it is linear in each variable separately.

The collection of multilinear maps is denoted $\mathrm{Hom}_F(V_1, \dots, V_n; W)$.

Exercise: Show that $\mathrm{Hom}_F(V_1, \dots, V_n; W)$ is a vector space.

Theorem: Let $V_1, \dots, V_n, W$ be F-vector spaces, and let $\iota : V_1 \times \cdots \times V_n \to V_1 \otimes \cdots \otimes V_n$ by $\iota(v_1, \dots, v_n) = v_1 \otimes \cdots \otimes v_n$.

(1) Given $T \in \mathrm{Hom}_F(V_1 \otimes \cdots \otimes V_n, W)$, then $T \circ \iota \in \mathrm{Hom}_F(V_1, \dots, V_n; W)$.

(2) Given $t \in \mathrm{Hom}_F(V_1, \dots, V_n; W)$, there is a unique linear map $T \in \mathrm{Hom}_F(V_1 \otimes \cdots \otimes V_n, W)$ such that $t = T \circ \iota$.

Proof. Proof is an exercise. Adapt the proof from $V_1 \otimes V_2$. ∎

Corollary: Let $V_1, \dots, V_k$ be vector spaces of dimension $n_1, \dots, n_k$. Let

$$\mathcal{B}_i = \left\{e_1^i, \dots, e_{n_i}^i\right\}$$

be bases for V_i. Then,

$$\mathcal{B} = \left\{e_{i_1}^1 \otimes e_{i_2}^2 \otimes \cdots \otimes e_{i_k}^k\right\}$$

is a basis for $V_1 \otimes \cdots V_k$.

In particular,

$$\dim_F(V_1 \otimes \cdots \otimes V_k) = \prod_{j=1}^{k} \dim_F\left(V_j\right).$$

Example. Let $V = V_1 = V_2 = V_3 = \mathbb{C}$, $F = \mathbb{R}$.

We have $\mathcal{B}_1 = \mathcal{B}_2 = \mathcal{B}_3 = \{1, i\}$. For the basis of $V_1 \otimes V_2 \otimes V_3$, we then have

$$\mathcal{B} = \{1 \otimes 1 \otimes 1, i \otimes 1 \otimes 1, 1 \otimes i \otimes 1, 1 \otimes 1 \otimes i, \dots, i \otimes i \otimes i\}.$$

Definition (Exterior Product)**.** Let $k \geqslant 1$ We define the kth exterior product of V, denoted $\Lambda^k(V)$, by

$$\Lambda^k(V) = \mathcal{T}^k(V) / \mathcal{A}_k(V),$$

where

$$\mathcal{A}_k(V) = \mathrm{span}\left\{v_1 \otimes \cdots \otimes v_k \;\middle|\; v_i = v_j \text{ for some } i \neq j\right\}$$

We write

$$v_1 \otimes \cdots \otimes v_k + \mathcal{A}_k(V) = v_1 \wedge \cdots \wedge v_k.$$

We call this an elementary wedge product. Elements in $\Lambda^k(V)$ are finite sums of elementary wedge products.

We have

$$\begin{aligned}(v_1 + \tilde{v}_1) \wedge v_2 \wedge \cdots \wedge v_k &= v_1 \wedge v_2 \wedge \cdots \wedge v_k + \tilde{v}_1 \wedge v_2 \wedge \cdots \wedge v_k \\ v_1 \wedge \cdots \wedge v_{j-1} \wedge cv_j \wedge v_{j+1} \wedge \cdots v_k &= c\left(v_1 \wedge \cdots \wedge v_k\right).\end{aligned}$$

We also have

$$v_1 \cdots \wedge v_k = 0_{\Lambda^k(V)}$$

if $v_i = v_j$ for some $i \neq j$.

Let $v, w \in V$. Then, we have

$$\begin{aligned}0_{\Lambda^2(V)} &= (v + w) \wedge (v + w) \\ &= v \wedge v + w \wedge w + w \wedge v + v \wedge w \\ &= w \wedge v + v \wedge w,\end{aligned}$$

meaning

$$v \wedge w = -w \wedge v.$$

More generally, we have

$$v_1 \wedge \cdots \wedge v_i \wedge v_{i+1} \wedge \cdots \wedge v_k = -v_1 \wedge \cdots \wedge v_{i+1} \wedge v_i \wedge \cdots \wedge v_k$$

Definition (Alternating Maps)**.** Let V, W be F-vector spaces. Let $t \in \mathrm{Hom}_F(V, \dots, V; W)$. If

$$t(v_1, \dots, v_k) = 0_W$$

whenever $v_i = v_j$ for some $i \neq j$, then we say t is alternating.

We denote the set of alternating maps

$$\mathrm{Alt}^k(V; W).$$

We set $\mathrm{Alt}^0(V; W) = F$.

Example (Cross Product)**.** Let $V = W = \mathbb{R}^3$, and define $t : V \times V \to W$ by $t(v_1, v_2) = v_1 \times v_2$. We saw before that $t \in \mathrm{Hom}_F(V, V; W)$, and we remember from calculus that $v \times v = 0$. Thus, we also have $t \in \mathrm{Alt}^2(V, W)$.

Example (Determinant)**.** Let $\det : \mathrm{Mat}_n(F) \to F$ be the regular determinant map.

Given $A \in \mathrm{Mat}_n(F)$, we can write

$$\begin{aligned}A &= \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ a_{21} & \cdots & \cdots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{pmatrix} \\ &= \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix},\end{aligned}$$

where v_i is the ith column of A. We can identify $\mathrm{Mat}_n(F)$ with $F^n \times \cdots \times F^n$, meaning we can imagine $\det : F^n \times \cdots \times F^n \to F$.

We have det is multilinear; for instance,

$$\det\left(v_1, \dots, v_j + c\tilde{v}_j, \cdots v_n\right) = \det(v_1, \dots, v_n) + c\det\left(v_1, \dots, \tilde{v}_j, \dots, v_n\right).$$

Exercise: Prove using induction.

It is also the case that det is alternating — i.e., that $\det(v_1, \dots, v_n) = 0$ if $v_i = v_j$ for some $i \neq j$.

This shows that det is an alternating map.

Exercise:

(a) Show that $\operatorname{Alt}_F^k(V; W)$ is an F-subspace of $\operatorname{Hom}_F(V, \dots, V; W)$.

(b) Show that $\operatorname{Alt}^1(V; F) = \operatorname{Hom}_F(V, F) = V'$.

(c) If $\dim_F(V) = n$, show that $\operatorname{Alt}^k(V; F) = 0$ for all $k > n$.

Theorem (Universal Property for Alternating Linear Maps)**:** Let V, W be F-vector spaces, $k > 0$. Define

$$\begin{aligned}\iota : V \times \cdots \times V &\to \Lambda^k(V)\\ (v_1, \dots, v_k) &\mapsto v_1 \wedge \cdots \wedge v_k.\end{aligned}$$

(1) $\iota \in \operatorname{Alt}^k\left(V, \Lambda^k(V)\right)$

(2) If $T \in \operatorname{Hom}_F\left(\Lambda^k(V), W\right)$, then $T \circ \iota \in \operatorname{Alt}^k(V, W)$

(3) If $t \in \operatorname{Alt}^k(V, W)$, then there is a unique $T \in \operatorname{Hom}_F\left(\Lambda^k(V), W\right)$ such that $t = T \circ \iota$ making the following diagram commute.

$$\begin{tikzcd} V \times \cdots \times V \arrow[r, "\iota"] \arrow[dr, "t"'] & \Lambda^k(V) \arrow[d, "T"] \\ & W \end{tikzcd}$$

Proof.

(1) We have the composition

$$\begin{tikzcd} V \times \cdots \times V \arrow[r] & \mathcal{T}^k \arrow[r] & \Lambda^k(V) \\ (v_1, \dots, v_n) \arrow[r, mapsto] \arrow[rr, bend right, "\iota"'] & v_1 \otimes \cdots \otimes v_k \arrow[r, mapsto] & v_1 \wedge \cdots \wedge v_k \end{tikzcd}$$

implying that ι is multilinear.

We have

$$\begin{aligned} t(v_1 + c\tilde{v}_1, \dots, v_k) &= T \circ \iota(v_1 + c\tilde{v}_1, \dots, v_k)\\ &= T((v_1 + c\tilde{v}_1) \wedge v_2 \wedge \cdots \wedge v_k)\\ &= T(v_1 \wedge v_2 \wedge \cdots \wedge v_k) + cT(\tilde{v}_1 \wedge v_2 \wedge \cdots \wedge v_k)\\ &= T(\iota(v_1, \dots, v_k)) + cT(\iota(\tilde{v}_1, \dots, v_k))\\ &= t(v_1, \dots, v_k) + ct(\tilde{v}_1, \dots, v_k). \end{aligned}$$

Let $v_i = v_j$ with $i \neq j$. Then,

$$\begin{aligned} t(v_1, \dots, v_k) &= T(\iota(v_1, \dots, v_k))\\ &= T(0)\\ &= 0, \end{aligned}$$

as ι is alternating, meaning t is also alternating.

Let $t \in \operatorname{Alt}^k(V; W)$. We have

$$t : V \times \cdots \times V \to W$$

and $t \in \operatorname{Hom}_F(V, \dots, V; W)$. By the universal property of the tensor product, we have

$$T : \mathcal{T}^k(V) \to W,$$

with $T(v_1 \otimes \cdots \otimes v_k) = t(v_1, \dots, v_k)$.

We have

$$T|_{\mathcal{A}_k(V)} = 0$$

because T agrees with t and t is alternating. $\square$

Example. Let V be a F-vector space with $\dim_F(V) = 1$. Let $v \neq 0_V$, so $\mathcal{B} = \{v\}$ is a basis for V.

Consider $\Lambda^k(V)$. Elements in this set are finite sums

$$\omega = \sum_{i \in I} \omega_i,$$

where

$$\omega_i = a_{1i} v \wedge \cdots \wedge a_{ki} v$$

for some $a_{ji} \in F$, or

$$\omega_i = a_{1i} \cdots a_{ki} (v \wedge \cdots \wedge v).$$

If $k \geqslant 2$, then we have $v \wedge v \in \omega_i$, so $\omega_i = 0$. We have

$$\begin{aligned} \Lambda^k(V) &= 0 && k \geqslant 2 \\ \lambda^k(V) &= V && k = 1 \\ \lambda^k(V) &= F. && k = 0 \end{aligned}$$

Now, we let V be a 2-dimensional F-vector space, with basis $\mathcal{B} = \{v_1, v_2\}$.

Let $k = 2$. A typical element in $\Lambda^2(V)$ is a finite sum

$$\begin{aligned} \omega &= \sum_{i \in I} \omega_i \\ \omega_i &= (a_i v_1 + b_i v_2) \wedge (c_i v_1 + d_i v_2) \end{aligned}$$

for some $a_i, b_i, c_i, d_i \in F$. Then, we have

$$\begin{aligned} \omega_i &= a_i v_1 \wedge c_i v_1 + a_i v_1 \wedge d_i v_2 + b_i v_2 \wedge c_i v_1 + b_i v_2 \wedge d_i v_2 \\ &= (a_i d_i - b_i c_i)(v_1 \wedge v_2). \end{aligned}$$

Thus, we have

$$\Lambda^2(V) = \operatorname{span}\{v_1 \wedge v_2\}$$

is a 1-dimensional vector space.

Likewise, if $\dim_F(V) = 3$, then

$$\begin{aligned}\Lambda^0(V) &\cong F\\ \Lambda^1(V) &\cong V\\ \Lambda^2(V) \text{ has basis } &\{v_1\wedge v_2, v_1\wedge v_3, v_2\wedge v_3\}\\ \Lambda^3(V) \text{ has basis } &\{v_1\wedge v_2\wedge v_3\}\\ \lambda^k(V) &= 0 \text{ for all } k\geqslant 4\end{aligned}$$

Theorem: Let $\dim_F(V) = n$, with basis $\{v_1,\dots,v_n\}$. For $1\leqslant k\leqslant n$, then

$$\mathcal{B}_k = \{v_{i_1}\wedge\cdots\wedge v_{i_k} \mid 1\leqslant i_1<\cdots<i_k\leqslant n\}$$

forms a basis for $\Lambda^k(V)$, and for $k>n$, $\Lambda^k(V) = 0$.

In particular, for $1\leqslant k\leqslant n$,

$$\dim_F\left(\Lambda^k(V)\right) = \binom{n}{k}.$$

Proof. Recall that

$$\mathcal{C} = \{v_{i_1}\otimes\cdots\otimes v_{i_k}\}$$

is a basis for $\mathcal{T}^k(V)$. The projection of these vectors, $\{v_{i_1}\wedge\cdots\wedge v_{i_k}\}$ is a spanning set for $\Lambda^k(V)$.

However, we can order the indices using the property that $v\wedge w = -w\wedge v$. Thus, $\mathcal{B}_k$ is spanning.

We want to show that $\Lambda^k(V)\hookrightarrow\mathcal{T}^k(V)$, and elements of $\mathcal{B}_k$ map to basis elements.

In order to do this, we must find an alternating map

$$t: V\times\cdots\times V\to\mathcal{T}^k(V).$$

Suppose we have

$$0_{\Lambda^k(V)} = \sum_{i_1,\dots,e_k\in I} c_{i_1,\dots,i_k}\left(v_{i_1}\wedge\cdots\wedge v_{i_k}\right).$$

Define

$$\begin{aligned}t_k: V\times\cdots V&\to\mathcal{T}^k(V)\\ (\tilde{v}_1,\dots,\tilde{v}_k)&\mapsto\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)\,\tilde{v}_{\sigma(1)}\otimes\cdots\otimes\tilde{v}_{\sigma(k)}.\end{aligned}$$

Note that for

$$\begin{aligned}\Delta_k &= \prod_{1\leqslant i\leqslant j\leqslant k}\left(x_i - x_j\right)\\ \sigma(\Delta_k) &= \prod_{1i\leqslant j\leqslant k}\left(x_{\sigma(i)} - x_{\sigma(j)}\right)\\ &= \pm\Delta_k,\end{aligned}$$

where $\operatorname{sgn}(\sigma)$ denotes the plus or minus sign on Δ_k. Then, we have

$$t_k\left(\tilde{v}_1 + c\tilde{v}_1', \tilde{v}_2,\dots,\tilde{v}_k\right) = \sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)\left(\tilde{v}_{\sigma(1)} + c\tilde{v}'_{\sigma(1)}\right)\otimes\tilde{v}_{\sigma(2)}\otimes\cdots\otimes\tilde{v}_{\sigma(k)}$$

$$
\begin{aligned}
&= \sum_{\sigma\in S_k} \operatorname{sgn}(\sigma)\left(\left(\tilde{v}_{\sigma(1)}\right)\otimes\tilde{v}_{\sigma(2)}\otimes\cdots\otimes\tilde{v}_{\sigma(k)} + c\left(\tilde{v}'_{\sigma(1)}\otimes\tilde{v}_{\sigma(2)}\otimes\cdots\otimes\tilde{v}_{\sigma(k)}\right)\right)\\
&= \sum_{\sigma\in S_k} \operatorname{sgn}(\sigma)\left(\tilde{v}_{\sigma(1)}\right)\otimes\tilde{v}_{\sigma(2)}\otimes\cdots\otimes\tilde{v}_{\sigma(k)}\\
&+ c\sum_{\sigma\in S_k} \operatorname{sgn}(\sigma)\left(\tilde{v}'_{\sigma(1)}\otimes\tilde{v}_{\sigma(2)}\otimes\cdots\otimes\tilde{v}_{\sigma(k)}\right)\\
&= t(\tilde{v}_1,\dots,\tilde{v}_k) + ct\left(\tilde{v}'_1,\dots,\tilde{v}_k\right).
\end{aligned}
$$

Suppose $\tilde{v}_i = \tilde{v}_{i+1}$. We want to show that $t_k(\tilde{v}_1,\dots\tilde{v}_k) = 0$.

Set $N_i = \langle (i,i+1)\rangle \leqslant S_k$. We can write

$$
\begin{aligned}
S_k &= \bigsqcup_{\sigma\in S_k} N_i\sigma\\
&= \bigsqcup \left\{(i,i+1),(i,i+1)\sigma\right\}.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
\sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)\, t\left(\tilde{v}_{\sigma(1)},\dots,\tilde{v}_{\sigma(k)}\right) &= \sum_{\sigma\in S_k/N_i}\operatorname{sgn}(\sigma)\, t\left(\tilde{v}_{\sigma(1)},\dots,\tilde{v}_{\sigma(k)}\right)\\
&+ \operatorname{sgn}((i,i+1)\sigma)\, t\left(\tilde{v}_{(i,i+1)\sigma(1)},\dots,\tilde{v}_{(i,i+1)\sigma(k)}\right)\\
&= \sum_{\sigma\in S_k/N_i}\operatorname{sgn}(\sigma)\, t\left(\tilde{v}_{\sigma(1)},\dots,\tilde{v}_{\sigma(k)}\right)\\
&- \operatorname{sgn}(\sigma)\, t\left(\tilde{v}_{(i,i+1)\sigma(1)},\dots,\tilde{v}_{(i,i+1)\sigma(k)}\right)
\end{aligned}
$$

Note that

$$
(i,i+1)\,\sigma(j) = \begin{cases} \sigma(j) & \sigma(j)\neq i,i+1\\ i+1 & \sigma(j) = i\\ i & \sigma(j) = i+1\end{cases}.
$$

Suppose $\sigma(1) = i$ and $\sigma(2) = i+1$. Then,

$$
\begin{aligned}
t\left(\tilde{v}_{(i,i+1)\sigma(1)},\dots,\tilde{v}_{(i,i+1)\sigma(k)}\right) &= t\left(\tilde{v}_{i+1},\tilde{v}_i,\dots,\tilde{v}_{\sigma(k)}\right)\\
&= t\left(\tilde{v}_i,\tilde{v}_{i+1},\dots,\tilde{v}_k\right)\\
&= t\left(\tilde{v}_{\sigma(1)},\dots,\tilde{v}_{\sigma(k)}\right).
\end{aligned}
$$

In other words, we get the subtraction equal to zero for each σ.

Since we have an alternating map, we can use the universal property for exterior products to get our linear map

$$
\begin{aligned}
T_k &: \Lambda^k(V)\to\mathcal{T}^k(V)\\
\tilde{v}_1\wedge\cdots\wedge\tilde{v}_k &\mapsto \sum_{\sigma\in S_k}\operatorname{sgn}(\sigma)\left(\tilde{v}_{\sigma(1)}\otimes\cdots\otimes\tilde{v}_{\sigma(k)}\right).
\end{aligned}
$$

Thus,

$$
0_{\Lambda^k(V)} = \sum_{i_1,\dots,i_k\in I} c_{i_1,\dots,i_k}\left(v_{i_1}\wedge\cdots\wedge v_{i_k}\right)
$$

$$\begin{aligned} 0_{\mathcal{T}^k(V)} &= \sum_{i_1,\dots,i_k \in I} c_{i_1,\dots,i_k} T_k\left(v_{i_1} \wedge \cdots \wedge v_{i_k}\right) \\ &= \sum_{i_1,\dots,i_k \in I} c_{i_1,\dots,i_k} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)\, v_{\sigma(i_1)} \otimes \cdots \otimes v_{\sigma(i_k)} \end{aligned}$$

is a sum of a subset of a basis for $\mathcal{T}^k(V)$, so $c_{i_1,\dots,i_k} = 0$ for each k. □

We want to make use of the fact that $\dim_F(\Lambda^n(V)) = 1$ for $\dim_F(V) = 1$, with basis $v_1 \wedge \cdots \wedge v_n$ for $\mathcal{B} = \{v_1, \dots, v_n\}$ a basis for V.

Proposition: Let $T \in \operatorname{Hom}_F(V, W)$. There is a unique linear map $\Lambda^k(T) \in \operatorname{Hom}_F\left(\Lambda^k(V), \Lambda^k(W)\right)$ such that

$$\Lambda^k(T)(v_1 \wedge \cdots \wedge v_k) = T(v_1) \wedge \cdots \wedge T(v_k)$$

for $v_1, \dots, v_k \in V$. Moreover,

$$\Lambda^k(\mathrm{id}_V) = \mathrm{id}_{\Lambda^k(V)},$$

and if $S \in \operatorname{Hom}_F(U, V)$, then

$$\Lambda^k(T \circ S) = \Lambda^K(T) \circ \Lambda^k(S).$$

Proof. Define $t : V \times \cdots \times V \to \Lambda^k(W)$ by $(v_1, \dots, v_k) \mapsto T(v_1) \wedge \cdots T(v_k)$. Since T is linear, t is multilinear, and is alternating by the definition of the wedge product.

By the universal property, we have a linear map

$$\begin{aligned} \Lambda^k(T) : \Lambda^k(V) &\to \Lambda^k(W) \\ v_1 \wedge \cdots \wedge v_k &\mapsto T(v_1) \wedge \cdots \wedge T(v_k). \end{aligned}$$

□

Example. Let $V = F^3$, $\mathcal{E}_3 = \{e_1, e_2, e_3\}$. Let T be the linear map such that

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

Consider the map

$$\Lambda^2(T) : \Lambda^2\left(F^3\right) \to \Lambda^2\left(F^3\right).$$

We know a basis for $\Lambda^2\left(F^3\right)$ is

$$\mathcal{B} = \{e_1 \wedge e_2, e_1 \wedge e_3, e_2 \wedge e_3\}.$$

We consider the matrix

$$\left[\Lambda^2(T)\right]_{\mathcal{B}}.$$

Consider

$$\begin{aligned} \Lambda^2(T)(e_1 \wedge e_2) &= T(e_1) \wedge T(e_2) \\ &= (e_1 \wedge 3e_3) \wedge (2e_2 - e_3) \\ &= 2(e_1 \wedge e_2) - (e_1 \wedge e_3) - 6(e_2 \wedge e_3). \\ \Lambda^2(T)(e_1 \wedge e_3) &= T(e_1) \wedge T(e_3) \end{aligned}$$

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

Thus, we have

$$\left[\Lambda^2 (T)\right]_{\mathcal{B}} = \begin{pmatrix} 2 & 1 & -4 \\ -1 & -5 & 2 \\ 6 & -3 & 3 \end{pmatrix}$$

Given $T \in \operatorname{Hom}_F (V, V)$, we have

$$
\begin{aligned}
\Lambda^n (T) : \Lambda^n (V) &\to \Lambda^n (V) \\
v_1 \wedge \cdots \wedge v_n &\mapsto \underbrace{\alpha (v_1 \wedge \cdots \wedge v_n)}_{T(v_1)\wedge \cdots T(v_n)}
\end{aligned}
$$

Definition. For $T \in \operatorname{Hom}_F (V, V)$, we define $\det(T)$ to be such that

$$\Lambda^n (T) (\omega) = \det(T)\omega$$

for any $\omega \in \Lambda^n (V)$.

Example. Let $V = F^2$, $\mathcal{E}_2 = \{e_1, e_2\}$. Let $T \in \operatorname{Hom}_F (V, V)$, we have

$$[T]_{\mathcal{E}_2} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then,

$$
\begin{aligned}
\Lambda^2 (T) (e_1 \wedge e_2) &= (ae_1 + ce_2) \wedge (be_1 + de_2) \\
&= \underbrace{(ad - bc)}_{\det(T)} (e_1 \wedge e_2).
\end{aligned}
$$

Exercise: Verify that we recover the cofactor expansion on a 3×3 matrix.

Lemma: Let $S, T \in \operatorname{Hom}_F (V, V)$. We have $\det (T \circ S) = \det(T) \det(S)$.

Proof. We have

$$
\begin{aligned}
\det(T \circ S) (v_1 \wedge \cdots \wedge v_n) &= \Lambda^n (T \circ S) (v_1 \wedge \cdots \wedge v_n) \\
&= \Lambda^n (T) \circ \Lambda^n (S) (v_1 \wedge \cdots \wedge v_n) \\
&= \Lambda^n (T) (\det(S) (v_1 \wedge \cdots \wedge v_n)) \\
&= \det(S)\Lambda^n (T) (v_1 \wedge \cdots \wedge v_n) \\
&= \det(S) \det(T) (v_1 \wedge \cdots \wedge v_n),
\end{aligned}
$$

meaning $\det (T \circ S) = \det(T) \det(S)$. □

Our ultimate goal is to prove that, for a given $A \in \operatorname{Mat}_n (F)$, that

$$\det (T_A) = \det (A),$$

where the determinant of the matrix defined by the cofactor expansion.

We view

$$\begin{aligned}A &= \begin{pmatrix} a_{11} & \cdots & a_{nn} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}\\ &\leftrightarrow \left(\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix}\right)\end{aligned}$$

as an element of $F^n \times \cdots \times F^n$, where F^n are the column vectors of A.

Theorem: We have $\det \in \text{Alt}^n (F^n, F)$, with $\det(I_n) = 1$.

Proof. Let $\mathcal{E}_n = \{e_1, \dots, e_n\}$ be the standard basis of F^n, and write

$$w_i = a_{1i}e_1 + \cdots + a_{ni}e_n$$

for $i = 1, \dots, n$. We let

$$w = b_{11}e_1 + \cdots + b_{n1}e_n.$$

Let $c \in F$.

We want to show that

$$\det(w_1 + cw, w_2, \dots, w_n) = \det(w_1, w_2, \dots, w_n) + c\det(w, w_1, \dots, w_n).$$

Define

$$\begin{aligned} T_1 : F^n &\to F^n \\ e_i &\mapsto w_i, \end{aligned}$$

with

$$[T_1]_{\mathcal{E}_n} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

We define

$$\begin{aligned} T_2 : F^n &\to F^n \\ e_1 &\mapsto w \\ e_i &\mapsto w_i, \end{aligned}$$

with

$$[T_2]_{\mathcal{E}_n} = \begin{pmatrix} b_{11} & a_{12} & \cdots & a_{1n} \\ b_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Finally, we define

$$T_3 : F^n \to F^n$$

$$\begin{aligned} e_1 &\mapsto w_1 + cw_2 \\ e_i &\mapsto w_i, \end{aligned}$$

with

$$[T_3]_{\mathcal{E}_n} = \begin{pmatrix} a_{11} + cb_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} + cb_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} + cb_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

In particular,

$$\begin{aligned} \det(T_1) &= \det(w_1, \dots, w_n) \\ \det(T_2) &= \det(w, w_2, \dots, w_n) \\ \det(T_3) &= \det(w_1 + cw, w_2, \dots, w_n). \end{aligned}$$

We want to show that

$$\det(T_3) = \det(T_1) + c\det(T_2).$$

Note that

$$\begin{aligned} \det(T_3)(e_1 \wedge \cdots \wedge e_n) &= \Lambda^n(T_3)(e_1 \wedge \cdots \wedge e_n) \\ &= T_3(e_1) \wedge T_3(e_2) \cdots \wedge T_3(e_n) \\ &= (w_1 + cw_2) \wedge w_2 \wedge \cdots \wedge w_n \\ &= w_1 \wedge w_2 \wedge \cdots \wedge w_n + c(w \wedge w_2 \wedge \cdots \wedge w_n) \\ &= T_1(e_1) \wedge T_1(e_2) \wedge \cdots T_1(e_n) + c(T_2(e_1) \wedge T_2(e_2) \wedge \cdots T_2(e_n)) \\ &= \Lambda^n(T_1)(e_1 \wedge \cdots \wedge e_n) + c\Lambda^n(T_2)(e_1 \wedge \cdots \wedge e_n) \\ &= \det(T_1)(e_1 \wedge \cdots \wedge e_n) + c\det(T_2)(e_1 \wedge \cdots \wedge e_n) \\ &= (\det(T_1) + c\det(T_2))(e_1 \wedge \cdots \wedge e_n). \end{aligned}$$

Thus, $\det(T_3) = \det(T_1) + c\det(T_2)$.

Let $w_i = w_j$ for some $i \neq j$.

$$\begin{aligned} \det(w_1, \dots, w_n)(e_1 \wedge \cdots \wedge e_n) &= \det(T_1)(e_1 \wedge \cdots \wedge e_n) \\ &= \Lambda^n(T_1)(e_1 \wedge \cdots \wedge e_n) \\ &= T_1(e_1) \wedge \cdots \wedge T_1(e_n) \\ &= w_1 \wedge \cdots \wedge w_n \\ &= 0_{\Lambda^n(V)}. \end{aligned}$$

Thus, det is an alternating map.

We have

$$\begin{aligned} \det(I_n)(e_1 \wedge \cdots \wedge e_n) &= \Lambda^n(I_n)(e_1 \wedge \cdots \wedge e_n) \\ &= e_1 \wedge \cdots \wedge e_n, \end{aligned}$$

so $\det(I_n) = 1$. □

Lemma: Let $t \in \text{Alt}^k(V, F)$. Then,

$$t(v_1, \dots, v_k) = -t(v_1, \dots, v_{i+1}, v_i, \dots, v_k).$$

Proof. Define

$$\psi(x, y) = T(v_1, \dots, v_{i-1}, x, y, v_{i+2}, \dots, v_k).$$

It is enough to show that $\psi(x, y) = -\psi(y, x)$. Since t is alternating,

$$\begin{aligned} 0 &= \psi(x + y, x + y) \\ &= \psi(x, x) + \psi(x, y) + \psi(y, x) + \psi(y, y) \\ &= \psi(x, y) + \psi(y, x) \\ \psi(x, y) &= -\psi(y, x). \end{aligned}$$

□

Lemma:

(1) Let $t \in \mathrm{Alt}^k(V, F)$. Then,

$$t\left(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(n)}\right) = \mathrm{sgn}(\sigma)\, t(v_1, \dots, v_k).$$

(2) If v_i is replaced by $v_i + cv_j$ for any $i \neq j$ and $c \in F$, then the value of t is unchanged.

Proposition: Let $t \in \mathrm{Alt}^k(V, F)$. Suppose for some $v_1, \dots, v_n \in V$ and $w_1, \dots, w_n \in V$, $a_{ij} \in F$, we have

$$w_i = a_{1i}v_i + \cdots + a_{ni}v_n.$$

Then,

$$t(w_1, \dots, w_n) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)\, a_{\sigma(1)1} \cdots a_{\sigma(n)n}\, t\left(v_{\sigma(1)}, \dots, v_{\sigma(n)}\right).$$

Proof. Expanding,

$$\begin{aligned} t(w_1, \dots, w_n) &= t\left(\sum_{j=1}^{n} a_{j1}e_j, \dots, a_{jn}e_j\right) \\ &= a_{i_1 1} \cdots a_{i_n n}\, t(v_{i_1}, \dots, v_{i_n}), \end{aligned}$$

where all the i_j are distinct. We have a bijection between the possible tuples $(i_1, \dots, i_n)$ and all possible tuples $(\sigma(1), \dots, \sigma(n))$ with $\sigma \in S_n$. From this, we get

$$\begin{aligned} t(w_1, \dots, w_n) &= \sum_{\sigma \in S_n} a_{\sigma(1)1} \cdots a_{\sigma(n)n}\, t\left(v_{\sigma(1)}, \dots, v_{\sigma(n)}\right) \\ &= \sum_{\sigma \in S_n} a_{\sigma(1)1} \cdots a_{\sigma(n)n}\, \mathrm{sgn}(\sigma)\, t(v_1, \dots, v_n). \end{aligned}$$

□

Corollary: The determinant is the unique function in $\mathrm{Alt}^n(F^n, F)$ with $\det(I_n) = 1$.

. Let $\mathcal{E}_n = \{e_1, \dots, e_n\}$ be the standard basis, $t \in \mathrm{Alt}^n(F^n, F)$ with $t(I_n) = 1$. This is the same as saying $t(e_1, \dots, e_n) = 1$.

Let $v_1, \dots, v_n \in V$, and write

$$v_i = a_{1i}e_1 + \cdots a_{ni}e_n.$$

We have

$$t(v_1, \dots, v_n) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma)\, a_{\sigma(1)1} \cdots a_{\sigma(n)n}\, t(e_1, \dots, e_n)$$

$$\begin{aligned}
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\, a_{\sigma(1)1}\cdots a_{\sigma(n)n}\\
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\, a_{\sigma(1)1}\cdots a_{\sigma(n)n} \det(e_1,\dots,e_n)\\
&= \det(v_1,\dots,v_n),
\end{aligned}$$

meaning $t = \det$. □

We can now determine the trace through the exterior product.

Let $T \in \operatorname{Hom}_F(V, V)$. Define

$$\begin{aligned}
t &: V\times\cdots\times V \to \Lambda^n(V)\\
(v_1,\dots,v_n) &\mapsto \sum_{j=1}^{n} \left(v_1\wedge\cdots\wedge v_{j-1}\wedge T\left(v_j\right)\wedge v_{j+1}\wedge\cdots\wedge v_n\right).
\end{aligned}$$

Exercise: Show this map is multilinear.

We want to show this map is alternating. Suppose $v_1 = v_2$.

$$\begin{aligned}
t(v_1, v_2,\dots,v_n) &= T(v_1)\wedge v_1\wedge\cdots\wedge v_n + v_1\wedge T(v_2)\wedge\cdots\wedge v_n + \sum_{j=3}^{n} v_1\wedge v_2\wedge\cdots\wedge T\left(v_j\right)\wedge\cdots\wedge v_n\\
&= T(v_1)\wedge v_2\wedge\cdots\wedge v_n + v_1\wedge T(v_2)\wedge\cdots\wedge v_n\\
&= T(v_1)\wedge v_2\wedge\cdots\wedge v_n - T(v_2)\wedge v_1\wedge\cdots\wedge v_n\\
&= 0.
\end{aligned}$$

Thus, $t \in \operatorname{Alt}^n(V, \Lambda^n(V))$, meaning we have a map $\varphi_T \in \operatorname{Hom}_f(\Lambda^n(V), \Lambda^n(F))$ such that

$$\varphi_T(v_1,\dots,v_n) = \sum_{j=1}^{n} v_1\wedge\cdots\wedge T\left(v_j\right)\wedge\cdots\wedge v_n.$$

Since $\dim_F(\Lambda^n(V)) = 1$, we know that φ_T is multiplication by a scalar.

We claim this scalar is $\operatorname{tr}(T)$. Let $A \in \operatorname{Mat}_n(F)$, with T_A the corresponding linear transformation. Then,

$$\begin{aligned}
\varphi_{T_A}(e_1\wedge\cdots\wedge e_n) &= \sum_{j=1}^{n} e_1\wedge\cdots\wedge T_A\left(e_j\right)\wedge\cdots\wedge e_n\\
&= \sum_{j=1}^{n} e_1\wedge\cdots\wedge\left(\sum_{i=1}^{n} a_{ij}e_i\right)\wedge\cdots\wedge e_n\\
&= \sum_{j=1}^{n} e_1\wedge\cdots\wedge\left(a_{jj}e_j\right)\wedge\cdots\wedge e_n\\
&= \sum_{j=1}^{n} a_{jj}(e_1\wedge\cdots\wedge e_n)\\
&= \left(\sum_{j=1}^{n} a_{jj}\right)(e_1\wedge\cdots\wedge e_n)\\
&= \operatorname{tr}(T_A)(e_1\wedge\cdots\wedge e_n).
\end{aligned}$$

Bilinear and Sesquilinear Forms

Consider $V = \mathbb{R}^3$. We know that

$$\begin{aligned} \varphi : V \times V &\to \mathbb{R} \\ (v, w) &\mapsto v \cdot w \end{aligned}$$

is such that $v \cdot v = \|v\|^2$ and

$$\varphi\left(v_1 + cv_2, w\right) = \varphi\left(v_1, w\right) + c\varphi\left(v_2, w\right),$$

meaning $\varphi \in \operatorname{Hom}_{\mathbb{R}}\left(\mathbb{R}^3, \mathbb{R}^3; \mathbb{R}\right)$.

Turning our attention to $V = \mathbb{C}$ with $F = \mathbb{C}$, we have for $z = x + iy$,

$$\begin{aligned} \|z\|^2 &= x^2 + y^2 \\ &= z\overline{z}. \end{aligned}$$

We define

$$\begin{aligned} \varphi : \mathbb{C} \times \mathbb{C} &\to \mathbb{C} \\ (z, w) &\mapsto z\overline{w}. \end{aligned}$$

This allows us to have $\varphi\left(z, z\right) = z\overline{z} = \|z\|^2$. We have

$$\varphi\left(z_1 + cz_2, w\right) = \varphi\left(z_1, w\right) + c\varphi\left(z_2, w\right),$$

but

$$\varphi\left(z, w_1 + cw_2\right) = \varphi\left(z, w_1\right) + \overline{c}\varphi\left(z, w_2\right).$$

This is not bilinear per se, but it's close.

In this case, we say φ is an example of a sesquilinear form.

Basic Definitions and Facts

Definition (Bilinear Form)**.** A bilinear form is an element of the space $\operatorname{Hom}_F\left(V, V; F\right)$.

Example (Bilinear Forms)**.**

(1) The dot product on F^n.

(2) Let $A \in \operatorname{Mat}_n\left(F\right)$. Define

$$\varphi_A\left(v, w\right) = v^T A w$$

for $v, w \in F^n$. We have $\varphi_A \in \operatorname{Hom}_F\left(F^n, F^n; F\right)$.

Exercise: Let $B \in \operatorname{Mat}_n\left(F\right)$. Define φ on $V = F^n$ by taking

$$\varphi\left(v, w\right) = \left(Bv\right) \cdot w.$$

Show $\varphi \in \operatorname{Hom}_F\left(F^n, F^n; F\right)$. What is the relationship between φ and φ_B.

Example. Let $V = \mathbb{R}^3$. Then, for $x, y, Z \in V$, recall that

$$|x \cdot (y \times z)|$$

is the volume of the parallelepiped defined by x, y, z.

Fixing x, the map

$$\varphi_x (y, z) = x \cdot (y \times z)$$

is bilinear.

Example. Let $p, q \in \mathbb{Z}_{\geqslant 0}$ with $p + q = n$. Set $V = F^n$.

Let $x, y \in V$. Define

$$\varphi_{p,q} (x, y) = \sum_{j=1}^{p} x_j y_j - \sum_{j=p+1}^{n} x_j y_j.$$

This is a bilinear form.

We denote the vector space with this bilinear form as $F^{p,q}$.

For instance, $\mathbb{R}^{3,1}$ is known as Minkowski space in the theory of relativity.

Example. Let $V = F^{2n}$. Let $x, y \in V$.

We define

$$\varphi (x, y) = \sum_{j=1}^{n} \left(x_{2j-1} y_{2j} - x_{2j} y_{2j-1}\right).$$

This is a bilinear form.

Definition. Let $\varphi \in \operatorname{Hom}_F (V, V; F)$. We say φ is right non-degenerate if, given $w_0 \in V$ such that

$$\varphi (v, w_0) = 0$$

for every $v \in V$, then $w_0 = 0$.

Similarly, φ is left non-degenerate if, given $v_0 \in V$ such that

$$\varphi (v_0, w) = 0$$

for every $w \in V$, then $v_0 = 0$.

If φ is both left non-degenerate and right non-degenerate, we say φ is non-degenerate.

Example. Let $\varphi \in \operatorname{Hom}_F (F^n, F^n; F)$ be the usual dot product. Suppose φ is right degenerate — i.e., there exists $w_0 \in F^n$ such that $\varphi (v, w_0) = 0$ for all $v \in F^n$. In particular, we have

$$\begin{aligned} \varphi (e_i, w_0) &= w_{0,i} \\ &= 0, \end{aligned}$$

so $w_{0,i} = 0$ for each i. Thus, we must have $w_0 = 0$, so φ is right non-degenerate.

Similarly, φ is left non-degenerate.

Exercise: Show that $\varphi_{p,q}$ is left and right non-degenerate.

Example. Let $V = F^3$ and define $\varphi(x, y) = x_1 y_1 + x_2 y_2$ for $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$.

It is the case that φ is a bilinear form, but φ is both left and right degenerate, as selecting either x or y to be $v = (0, 0, 1)$, $\varphi(x, v) = 0$ for all x and $\varphi(v, y) = 0$ for all y.

If V is finite-dimensional, we will see that left and right non-degenerate are equivalent. However, if V is infinite-dimensional, they are not equivalent.[VII]

Recall that we had $V \cong V'$, but that this isomorphism is not canonical. We will see that the isomorphisms for V and V' are in bijection with non-degenerate forms. This doesn't work in the infinite-dimensional case since V is not necessarily isomorphic to V'.

Let $\varphi \in \operatorname{Hom}_F(V, V; F)$. Fix some element $v_0 \in V$. The map

$$\varphi(\cdot, v_0) \in V',$$

as φ is bilinear, so $\varphi(\cdot, v_0)$ is linear. This gives a map

$$\begin{aligned} R_\varphi &: V \to V' \\ &v_0 \mapsto \varphi(\cdot, v_0). \end{aligned}$$

We write this as $R_\varphi(v_0) = \varphi(\cdot, v_0)$.

Let $v_1, v_2 \in V$, and $a \in F$. Then,

$$\begin{aligned} R_\varphi(v_1 + a v_2)(w) &= \varphi(w, v_1 + a v_2) \\ &= \varphi(w, v_1) + a\varphi(w, v_2) \\ &= R_\varphi(v_1)(w) + a R_\varphi(v_2)(w). \end{aligned}$$

Thus, we see that $R_\varphi \in \operatorname{Hom}_F(V, V')$.

Similarly, we can have $L_\varphi(v) = \varphi(v, \cdot)$, so by a similar argument, we get

$$L_\varphi \in \operatorname{Hom}_F(V, V').$$

Lemma: A bilinear form φ is non-degenerate if and only if L_φ and R_φ are injections.

Proof. Suppose L_φ and R_φ are injections. Suppose we have w_0 such that $\varphi(v, w_0) = 0$ for all $v \in V$. Thus,

$$\begin{aligned} 0 &= \varphi(v, w_0) \\ &= R_\varphi(w_0)(v) \end{aligned}$$

for all $v \in V$, so $R_\varphi(w_0)$ is the zero map. However, we said that $R_\varphi(w_0)$ is injective, $w_0 = 0$. Thus, φ is right non-degenerate.

Similarly, if $\varphi(v_0, w) = 0$ for all $w \in V$, then $L_\varphi(v_0)$ is the zero map, so $v_0 = 0$. Thus, φ is left non-degenerate.

Assume L_φ or R_φ is not injective. Let R_φ be not injective. Then, there exists $w_0 \in V$ such that

$$R_\varphi(w_0) = 0_{V'},$$

so $R_\varphi(w_0)(v) = 0$, so $\varphi(v, w_0) = 0$ for all $v \in V$, so φ is right degenerate. □

[VII] In the sequence space ℓ_2, using the left and right shift operators allows us to find this.

Corollary: If V is finite-dimensional, then φ is non-degenerate if and only if L_φ and R_φ are isomorphisms.

Proof. If $\dim_F(V) < \infty$, then $\dim_F(V) = \dim_F(V')$, so injective is the same as bijective. □

Definition. We define the left and right kernels of φ to be

$$\begin{aligned}\ker_R(\varphi) &= \{w \in V \mid \varphi(v,w) = 0 \text{ for all } v \in V\}\\ \ker_L(\varphi) &= \{w \in V \mid \varphi(w,v) = 0 \text{ for all } v \in V\}\end{aligned}$$

Theorem: Let $\dim_F(V) < \infty$, $\varphi \in \operatorname{Hom}_F(V,V;F)$. The maps R_φ and L_φ are dual to each other. In other words, given

$$L_\varphi : V \to V',$$

we consider the dual map

$$L'_\varphi : V'' \to V'.$$

If we identify $V \cong V''$ (where each $\hat{v}$ is identified with the element v), then $L'_\varphi = R_\varphi$, and similarly, $R'_\varphi = L_\varphi$.

Proof. Recall that given $T \in \operatorname{Hom}_F(V,W)$, we have a dual map $T' : W' \to V'$ defined by

$$T'(\varphi)(v) = \varphi(T(v)).$$

Recall the canonical isomorphism given by $v \mapsto \hat{v}$, where $\hat{v}(\varphi) = \varphi(v)$.

Let $v, w \in V$. We have

$$\begin{aligned}L'_\varphi(\hat{v})(w) &= \hat{v}(L_\varphi(w))\\ &= \hat{v}(L_\varphi(w,\cdot))\\ &= L_\varphi(w,v)\\ &= \varphi(w,v)\\ &= R_\varphi(v)(w).\end{aligned}$$

Thus, $L'_\varphi = R_\varphi$.

Thus, if V is finite dimensional, then φ being non-degenerate is equivalent to φ being left non-degenerate or right non-degenerate. □

Lemma: Let $\dim_F(V) < \infty$. There is a bijection between isomorphisms $V \to V'$ and non-degenerate bilinear forms.

Proof. Given such a φ, we have $R_\varphi \in \operatorname{Hom}_F(V,V')$. Since φ is non-degenerate, R_φ is an injective, so R_φ is an isomorphism.

Suppose we have $T : V \to V'$ is an isomorphism. Define $\varphi(v,w) = T(w)(v)$. This is non-degenerate, as T is an isomorphism. □

Definition (Conjugation)**.** Let F be a field, $\operatorname{conj} : F \to F$ a map such that

(1) $\operatorname{conj}\left(\operatorname{conj}(x)\right) = x$

(2) $\operatorname{conj}(x+y) = \operatorname{conj}(x) + \operatorname{conj}(y)$

(3) $\operatorname{conj}(xy) = \operatorname{conj}(x)\operatorname{conj}(y)$.

We call conj a conjugation map. We say it is nontrivial if it is not the identity map.

Example. If $F = \mathbb{C}$, and we define $\operatorname{conj}(z) = \overline{z}$ is a conjugation map.

Example. Let $F + \mathbb{Q}\left(\sqrt{d}\right) = \left\{a + b\sqrt{d} \,\middle|\, a, b \in \mathbb{Q}\right\}$ with d not a perfect square. The map

$$a + b\sqrt{d} \xmapsto{\text{conj}} a - b\sqrt{d}$$

is a conjugation map.

Henceforth, we refer to conjugation maps by taking $x \mapsto \overline{x}$.

Lemma: Let F be a field with nontrivial conjugation. Assume $\operatorname{char}(F) \neq 2$.

(1) Let $F_0 = \{z \in F \mid z = \overline{z}\}$. Then, F_0 is a proper subfield of F.

(2) There is a nonzero element $j \in F$ such that $\overline{j} = -j$.

(3) Every element of F can be written as a $z = x + yj$ for some $x, y \in F_0$.

Proof. Exercise. □

Definition. Let V be an F-vector space, where F is a field with conjugation. A conjugation map on V is a map $\operatorname{conj} : V \to V$ such that

(1) $\operatorname{conj}\left(\operatorname{conj}(v)\right) = v$

(2) $\operatorname{conj}(v + w) = \operatorname{conj}(v) + \operatorname{conj}(w)$.

(3) $\operatorname{conj}(av) = \overline{a}\operatorname{conj}(v)$.

Example. If V is an F-vector space of dimension n and F has conjugation, we can define conjugation on V by taking

$$[v]_{\mathcal{B}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \xmapsto{\text{conj}} \begin{pmatrix} \overline{x_1} \\ \vdots \\ \overline{x_n} \end{pmatrix} = [w]_{\mathcal{B}}.$$

We set $\operatorname{conj}(v) = w$.

Definition. Let V, W be F-vector spaces, where F has conjugation. We say $T : V \to W$ is conjugate linear if

(1) $T(v_1 + v_2) = T(v_1) + T(v_2)$

(2) $T(av) = \overline{a}T(v)$.

We say T is a conjugate isomorphism if it is conjugate linear and bijective.

We can define a new vector space $\overline{V}$ by having $\overline{V} = V$ as a set, but with

$$\begin{aligned} m : F \times \overline{V} &\to \overline{V} \\ (a, v) &\mapsto a \cdot v = \overline{a}v. \end{aligned}$$

Exercise: Verify that $\overline{V}$ is a F-vector space.

We may ask what linear maps look like in $\operatorname{Hom}_F\left(\overline{V}, W\right)$.

$$\begin{aligned} T(v_1 + v_2) &= T(v_1) + T(v_2) \\ T(a \cdot v) &= aT(v) \\ T(a\overline{v}) &= \overline{a}T(v). \end{aligned}$$

Thus, we have T is conjugate linear on V if and only if it is linear on $\overline{V}$.

Definition (Sesquilinear Form)**.** Let $\varphi : V \times V \to F$ be a map that is linear in the first variable and conjugate linear in the second variable. Their collection is denoted $\operatorname{Hom}_F\left(V, \overline{V}; W\right)$.

Example. Let $V = \mathbb{C}^n$. Define

$$\varphi : \mathbb{C}^n \times \mathbb{C}^n$$

by

$$\begin{aligned} \varphi(v, w) &= v^T \overline{w} \\ &= \sum_{i=1}^{n} v_i \overline{w_i}. \end{aligned}$$

Example. Let $V \in F^n$, where F has conjugation. Let $A \in \operatorname{Mat}_n(F)$. Define

$$\varphi_A : V \times V \to F$$

by

$$\varphi_A(v, w) = v^T A \overline{w}.$$