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

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

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

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

We have

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

$$= w + v.$$

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

$$\begin{aligned}
(cd)v &= (cd)\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \\
&= \begin{pmatrix} (cd)v_1 \\ \vdots \\ (cd)v_n \end{pmatrix} \\
&= \begin{pmatrix} c\left(dv_1\right) \\ \vdots \\ c\left(dv_n\right) \end{pmatrix} \\
&= c\left(dv\right).
\end{aligned}$$

Finally,

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

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

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

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

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

$$f(x) + g(x) = \sum_{j=0}^{n} \left(a_j + b_j\right) x^j$$
$$cf(x) = \sum_{j=0}^{n} \left(ca_j\right) x^j.$$

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

We define

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

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

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

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

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 $\mathrm{Mat}_n(\mathbb{F})$ to denote $\mathrm{Mat}_{n,n}(\mathbb{F})$.

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.

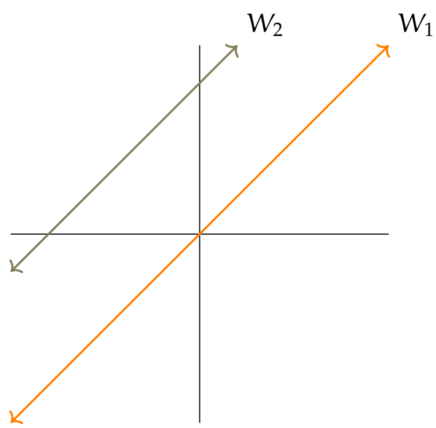
(3)

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

□

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 $\operatorname{Hom}_{\mathbb{F}}(V, W)$, or $\mathcal{L}(V, W)$.

Example (Identity Transformation)**.** Define

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

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

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

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

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

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

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 $T \in \operatorname{Hom}_{\mathbb{R}}(\mathbb{C}, \mathbb{C})$, but $T \notin \operatorname{Hom}_{\mathbb{C}}(\mathbb{C}, \mathbb{C})$.

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

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

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

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 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,

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

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 $\operatorname{span}_{\mathbb{F}}(C) = V$. 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 $\operatorname{span}_{\mathbb{F}}(\mathcal{B}) = V$. If not, there exists some $v \in C$ such that $v \notin \operatorname{span}_{\mathbb{F}}(\mathcal{B})$. However, if $v \notin \operatorname{span}_{\mathbb{F}}(\mathcal{B})$, 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, $\operatorname{span}_{\mathbb{F}}(\mathcal{B}) = V$. □

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 $\operatorname{span}_{\mathbb{F}}(\mathcal{B}) = V$ and $|\mathcal{B}| = m$.*

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

Proof. Let $C = \{w_1, \dots, w_n\}$ with $n > m$. We wish to show that 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,

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

Therefore,

$$f(x) | \left(a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} - 0\right),$$

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 $C = \{v_i\}_{i\in I}$ be a subset of V.*

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

Proof. Suppose 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. □

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) $V_1, \dots, V_k$ 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 $\mathcal{B} = \{v_1, \dots, v_n\}$ 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, $V_1, \dots, V_k$ 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 $V_1, \dots, V_k$ 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

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

□

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 $C = \mathcal{B} \setminus \mathcal{A}$, and $W = \operatorname{span}(C)$. □

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

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

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

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

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

Remark: To define a map $T : V/W \to U$, 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* $\overline{T} : V/\ker(T) \to W$ *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. □

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

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

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 $V \cong \operatorname{im}(T)$, 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 $v' \in V'$. Define a_i by

$$a_i = v'(v_i).$$

Set

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

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

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

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

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

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

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

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

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

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 $v' \in V'$. Then,

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

$\square$

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 $v' \in V'$, 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 $\varphi = v'$, 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 $\mathcal{B} = \{v_1, \dots, v_n\}$ 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

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

(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

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

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

Notice that

$$\begin{aligned} e_1 &= \frac{1}{2}v_1 + \frac{1}{2}v_2 \\ e_2 &= -\frac{1}{2}v_1 + \frac{1}{2}v_2. \end{aligned}$$

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

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

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:

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

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

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

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

Thus, $A = P\left[T_A\right]_{\mathcal{B}} P^{-1}$. 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 $A = \left(a_{ij}\right) \in \operatorname{Mat}_{m,n}(\mathbb{F})$. 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 $A \in \operatorname{Mat}_{m,n}(\mathbb{F})$. 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

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

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* $A \in \mathrm{Mat}_{m,n}(\mathbb{F})$*. 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,

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

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

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

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

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

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

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

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

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

If we look at the matrix, we then have

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

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

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