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).

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

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

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);
- $m : \mathbb{F} \times V \to V$, $m(c, v) \mapsto cv$ (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

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

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

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

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

Observe

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

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

and

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

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

$$\mathcal{P}_n\left(\mathbb{F}\right) = \left\{a_0 + a_1 x + \cdots + a_n x^n \,|a_i \in \mathbb{F}\right\}.$$

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

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

$$cf(x) = \sum_{j=0}^{n} (ca_j)\, x^j.$$

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

We define

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

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

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

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.

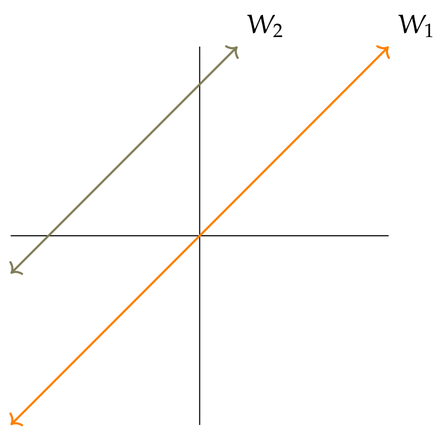
(3)

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

□

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 $\mathcal{P}_m(\mathbb{F})$ and $\mathcal{P}_n(\mathbb{F})$, if $m \leqslant n$, then $\mathcal{P}_m(\mathbb{F})$ is a subspace of $\mathcal{P}_n(\mathbb{F})$.

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

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

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

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

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

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

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

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

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, $E_a \in \mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R})$.

(2)

$$\begin{aligned}D : V &\to V\\ f &\mapsto f'.\end{aligned}$$

Then, $D \in \mathrm{Hom}_{\mathbb{R}}(V, V)$.

(3)

$$\begin{aligned}I_a : V &\to V\\ f &\mapsto \int_a^x f(t)\, dt.\end{aligned}$$

Then, $I_a \in \mathrm{Hom}_{\mathbb{R}}(V, V)$.

(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* $T \in Hom_{V,W}$. *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 ($\mathcal{P}_n(\mathbb{F})$ and $\mathbb{F}^{n+1}$)**.** Set $V = \mathcal{P}_n(\mathbb{F})$ and $W = \mathbb{F}^{n+1}$.

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

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

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

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

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

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

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

We call this the kernel of T.

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

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

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

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

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

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

Thus, $v_1 + cv_2 \in \ker \mathsf{T}$.

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

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

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

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

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

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

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

meaning $v_1 - v_2 \in \ker \mathsf{T}$, meaning $v_1 - v_2 = 0_V$. Thus, $v_1 = v_2$. □

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

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

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

To examine the kernel, let

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

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 $V = \mathcal{P}_2(\mathbb{F})$. Set $\mathcal{B} = \left\{1, x, x^2\right\}$. We have $\operatorname{span}_{\mathbb{F}}(\mathcal{B}) = \mathcal{P}_2(\mathbb{F})$.

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

$$0 = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ji} c_i \right) v_j$$

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

Thus, C is not linearly independent. □

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

Proof. Let $|\mathcal{B}| = m$, $|C| = n$. Since 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 $C \subseteq V$ with* $\operatorname{card} C = m$.

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

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

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

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

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

$$[g(x)] + [h(x)] = [g(x) + h(x)]$$
$$c[g(x)] = [cg(x)].$$

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

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

Thus,

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

Thus,

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

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 $[g(x)] = [r(x)] \in \operatorname{span}_{\mathbb{F}}(\mathcal{B})$.