Remark: Up to this point, everything we have proved relies only on the Zermelo - Fraenkel Set theory (ZF). This is a list of ~8 axioms written in the language of first-order logic from which other theorems are proved. For example:

Thm of Finite Choice The cartesian product of a finite family of non-empty sets is non-empty.

is a theorem in ZF. It is easily proved using the Principal of Mathematical Induction, which is also a theorem in ZF. Finite choice asserts that we can choose 1 element from each set and form a new set. But the statement cannot be proved or disproved if the family is arbitrary.

Most mathematicians assume the statement as an axiom in addition to the ZF axioms:

Axiom of Choice If $\{X_i\}_{i \in I}$ is a set of non-empty sets, then the Cartesian product

$$\prod_{i \in I} X_i := \left\{ f : I \to \bigcup_{i \in I} X_i \;\middle|\; f(i) \in X_i,\ \forall i \in I \right\}$$

is nonempty. $\square$

The elements of $\prod X_i$ are functions $f : I \to \bigcup X_i$

having the property that $f(i) \in X_i$ for all $i \in I$. So the Axiom of Choice asserts the existence of a "choice function". Sometimes, choice functions can be constructed w/ AoC:

$$\prod_{X \in P(\mathbb{N}) \setminus \{\phi\}} X$$

is nonempty because

$$f : P(\mathbb{N}) \setminus \{\phi\} \longrightarrow \mathbb{N}$$
$$X \longmapsto \min(X)$$

is a choice function. How about

$$\prod_{X \in P(\mathbb{R}) \setminus \{\phi\}} X \quad ?$$

Its not possible to construct a choice function here. But AoC guarantees one exists.

So in Zermelo - Fraenkel set theory w/ choice (ZFC) one can prove the following theorem:

1.4.17 Theorem (Zorn's Lemma) Let $(P, \leq)$ be a non-empty partially ordered set. If every totally-ordered subset of P has an upper bound, then P has a maximal element. $\square$

Recall: • A partially ordered set is a set P w/ a relation $\leq$ that is

- reflexive $x \leq x \quad \forall x \in P$

- transitive $x \leq y \wedge y \leq z \Rightarrow x \leq z$
- antisymmetric $x \leq y \wedge y \leq x \Rightarrow x = y$

- **totally-ordered set** is a poset $(P, \leq)$ such that any 2 elements are comparable: $\forall x, y \in P$, $x \leq y$ or $y \leq x$.
- Let $Q \subseteq P$. An **upper bound** for Q is an element $x \in P$ s.t. $y \leq x$ $\forall y \in Q$.
- A **maximal element** of a poset P is an element $x \in P$ such that $x \leq y \Rightarrow x = y$. (E.g. a basis is a maximal element of the poset of all independent subsets of a v.s. V.)

In fact, Zorn's Lemma is actually equivalent to the Axiom of Choice! Let's use it!

1.4.18 Theorem **(existence of bases)** Let V be a vector space over F. Then V has a basis.

Proof. Here's the idea: by Prop 1.4.8 a basis for V is the same as a maximal element of the poset of all linearly independent subsets of V ordered by subset inclusion $\subseteq$. We can prove that this poset has a max'l element using Zorn's Lemma!

Let $L := \{ B \subseteq V \mid B \text{ is independent} \}$. Then $(L, \subseteq)$ is a partially ordered set. Moreover, L is nonempty since $\emptyset \in L$. Let $T \subseteq L$ be a totally ordered subset. We claim that $U := \bigcup_{B \in T} B$ is an upper bound for T. It suffices to show that $U \in L$, i.e., U is independent. Suppose

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$$

where $\alpha_i \in F$ and $v_i \in U$. Then there are $B_1, \ldots, B_n \in T$ such that $v_1 \in B_1, \ldots, v_n \in B_n$. Every finite subset of a totally ordered set has a maximum element, so there exists $j \in \{1, \ldots, n\}$ such that $B_i \subseteq B_j$ for all $i \in \{1, \ldots, n\}$. Thus, $v_1, \ldots, v_n \in B_j$. But B_j is independent, so $\alpha_1 = \cdots = \alpha_n = 0$. This proves that U is independent. Thus, by Zorn's lemma, L has a maximal element, i.e., a basis for V.

$\square$

Remark: One can actually prove that Thm 1.4.18 is actually equivalent to the Axiom of Choice. Most mathematicians work in ZFC because AOC seems reasonable. However, there are some bizarre consequences:

- Banach-Tarski paradox
- existence of non-measurable sets.

It turns out that AoC is independent of ZF:

it can not be proved from the other axioms. It is also consistent w/ ZF, in the sense that no statement and its negation are both true in ZFC. However, this is not provable in ZFC, one must work in a "larger" axiomatic system.

1.4.19 Example Let F be a field and X a set.

$$F^X := \mathrm{Maps}(X,F) = \{ f: X \to F : f \text{ a function} \}$$

is an F-vector space w/ "coordinate-wise" addition & scalar mult.

- $X = \{1,\dots,n\}$, a function $f:\{1,\dots,n\} \to F$ is just an ordered tuple $(f(1),\dots,f(n)) \in F^n$

  So $F^{\{1,\dots,n\}} = F^n$

- $X = \mathbb{N}$, a function $f: \mathbb{N} \to F$ is by definition a sequence $(f(1), f(2), \dots)$

  So $F^{\mathbb{N}}$ is the space of of infinite seq. in F.

There is an important subspace of F^X:

$$F^{(X)} := \left\{ f: X \to F \;\middle|\; \begin{array}{l} f(x)=0 \text{ for all but finitely} \\ \text{many elmts of } X \end{array} \right\}$$

The vector space $F^{(X)}$ has an obvious basis: for each $x \in X$, define $e_x : X \to F$ via

$$e_x(y) = \begin{cases} 1, & y=x \\ 0, & y \neq x. \end{cases}$$

Then clearly $\{e_x : x \in X\}$ is a basis for $F^{(X)}$. However, there is no general way to write down a basis for F^X (if X is infinite). Even though Thm 1.4.18 guarantees the existence of basis, the proof is nonconstructive - it doesnt tell us how to build a basis. In order understand infinite-dimensional vector spaces, one needs new ideas. $\square$

1.4.20 Definition **(free vector space)** Let F be a field and let X be a set. The vector space $F^{(X)}$ described in Example 1.4.19 is called the **free F-vector space with basis X**. $\square$

Note: one reason why we care: for any set X, there is a vector space w/ basis (in bijection w/) X.

Chapter II: Linear Maps

2.1 Linear Maps

2.1.1 Definition Let V and W be vector spaces over the same field F. A **linear map** from V to W is a function $L: V \to W$ satisfying:

(a) $L(u+v) = L(u) + L(v)$ for all $u, v \in V$

(b) $L(\alpha u) = \alpha L(u)$ for all $u \in V$, $\alpha \in F$.

If $V = W$, then L is called a **linear operator.**

Linear maps are also called linear transformation.

$\square$

Note:
- $L(0) = 0$ $\quad (L(0) = L(0) + L(0) \Rightarrow L(0) = 0)$
- $L(-v) = -L(v)$ $\quad (L(-v) = L((-1)v) = (-1)L(v) = -L(v))$
- $L(\sum_{i=1}^{n} \alpha_i v_i) = \sum_{i=1}^{n} \alpha_i L(v_i)$ (Proof by induction).

2.1.2 Example

(a) The zero map
$$\begin{aligned} 0: V &\longrightarrow W \\ v &\longmapsto 0 \end{aligned}$$

(b) The identity map
$$\begin{aligned} \mathrm{id}_V: V &\longrightarrow V \\ v &\longmapsto v \end{aligned}$$

(c) Transpose of a matrix:
$$\begin{aligned} (-)^T: F^{m\times n} &\longrightarrow F^{n\times m} \\ A &\longmapsto A^T \end{aligned}$$

Note: this is a bijection since $(A^T)^T = A$

(d)
$$\begin{aligned} F^{n\times 1} &\longrightarrow F^n \\ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} &\longmapsto (a_1, \ldots, a_n) \end{aligned} \qquad \begin{aligned} F^{1\times n} &\longrightarrow F^n \\ (a_1, \ldots, a_n) &\longmapsto (a_1, \ldots, a_n) \end{aligned}$$

These are both linear bijection. From now on, we identify the spaces of row vectors, column vectors, and tuples.

(e) Let $A \in F^{m\times n}$. Left multiplication by A

defines a linear map

$$\begin{aligned} A\cdot : F^{n\times k} &\longrightarrow F^{m\times k} \\ X &\longmapsto AX \end{aligned}$$

In particular, we have a linear map

$$\begin{aligned} A\cdot : F^{n} &\longrightarrow F^{m} \\ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} &\longmapsto A\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \end{aligned}$$

(f) Let $A \in F^{m\times n}$. Right mult. by A defines a linear map

$$\begin{aligned} \cdot A : F^{k\times n} &\longrightarrow F^{k\times n} \\ X &\longmapsto XA \end{aligned}$$

The linear maps $A\cdot$ and $\cdot A$ commute w/ $(-)^T$

$$\begin{tikzcd} F^{n\times k} \arrow[r, "A\cdot"] \arrow[d, "(-)^T"'] \arrow[dr, phantom, "\circlearrowright"] & F^{m\times k} \arrow[d, "(-)^T"] \\ F^{k\times n} \arrow[r, "\cdot A^T"'] & F^{k\times m} \end{tikzcd}$$

The diagram commutes in the sense that

$$(AX)^T = X^T A^T$$

for all $x \in F^{n\times k}$

When $k=1$

column vectors →

row vectors →

$$\begin{tikzcd} F^{n} \arrow[r, "A\cdot"] \arrow[r, bend left=50] \arrow[d, "\wr"'] \arrow[d, bend right=50] \arrow[dr, phantom, "\circlearrowright"] & F^{m} \arrow[d, "\wr"] \\ F^{n} \arrow[r, "\cdot A^T"'] \arrow[r, bend right=50] & F^{m} \arrow[u, bend right=50] \end{tikzcd}$$

So up to viewing elements of F^n as row or column vectors, the maps $A\cdot$ and $\cdot A^T$ are the same. For this reason, there is no loss in generality in considering only $A\cdot$. ☐

2.1.3 Proposition Let F be a field. There is a bijection

$$\mathrm{Hom}(F^n,F^m) := \left\{ \begin{matrix} \text{linear maps} \\ F^n \to F^m \end{matrix} \right\} \xleftrightarrow{\sim} \left\{ \begin{matrix} m \times n \text{ matrices} \\ \text{w/ coeff. in} \\ F \end{matrix} \right\} =: F^{m\times n}$$

Proof. Define $\Phi : \mathrm{Hom}(F^n, F^m) \longrightarrow F^{m \times n}$ as follows:
$\Phi(L)$ is the $m \times n$ matrix w/ i^{th} column equal to $L(e_i)$, that is

$$\Phi(L) = (L(e_1) \,|\, L(e_2) \,|\, \cdots \,|\, L(e_n)).$$

Define $\Psi : F^{m\times n} \longrightarrow \mathrm{Hom}(F^n, F^m)$ via $\Psi(A) := A\cdot$.
Then Φ and Ψ are mutually inverse bijections:

$$(\Psi \circ \Phi)(L)\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = (L(e_1) \,|\, L(e_2) \,|\, \cdots \,|\, L(e_n))\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

$$= a_1 L(e_1) + \cdots + a_n L(e_n)$$

$$= L(a_1 e_1 + \cdots + a_n e_n)$$

$$= L\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

the columns of A

$$(\Phi \circ \Psi)(A) = \Phi(A\cdot) = (Ae_1 \,|\, Ae_2 \,|\, \cdots \,|\, Ae_n) = A.$$

☐

The preceding prop. gives us a way to generate examples of linear maps.

2.1.4 Example Let $F = \mathbb{R}$

(a) $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, $\quad \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ Zero map

(b) $A = \begin{pmatrix}1&0\\0&1\end{pmatrix}$, $\quad \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ $\quad$ identity map

$$\begin{pmatrix}x\\y\end{pmatrix} \longmapsto \begin{pmatrix}x\\y\end{pmatrix}$$

(c) $A = \begin{pmatrix}0&1\\1&0\end{pmatrix}$, $\quad \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ $\quad$ reflection over the line $y = x$

$$\begin{pmatrix}x\\y\end{pmatrix} \longmapsto \begin{pmatrix}y\\x\end{pmatrix}$$

(d) $A = \begin{pmatrix}2&0\\0&2\end{pmatrix}$, $\quad \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ $\quad$ dilates by a factor of 2

$$\begin{pmatrix}x\\y\end{pmatrix} \longmapsto 2\begin{pmatrix}x\\y\end{pmatrix}$$

(e) $A = \begin{pmatrix}1&0\\0&0\end{pmatrix}$, $\quad \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ $\quad$ projection onto x axis

$$\begin{pmatrix}x\\y\end{pmatrix} \longmapsto \begin{pmatrix}x\\0\end{pmatrix}$$

(f) $A = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$, $\quad \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ $\quad$ rotation about origin by θ radians.

□

2.1.5 Example

More exotic examples.

(a) Formal differentiation of polynomials:

$$D : F[X] \longrightarrow F[X]$$

$$a_0 + a_1 x + \dots + a_n x^n = \sum_{i=0}^{n} a_i x^i \longmapsto \sum_{i=1}^{n} a_i \cdot i \, x^{i-1} = a_1 + 2a_2 X + \dots + n a_n x^{n-1}$$

(b) $C(\mathbb{R}, \mathbb{R})$ = continuous functions $\mathbb{R} \to \mathbb{R}$, a subspace of $\mathbb{R}^{\mathbb{R}}$. Integration over a compact interval $[a,b]$ defines a linear map:

$$C(\mathbb{R}, \mathbb{R}) \longrightarrow \mathbb{R}$$

$$f \longmapsto \int_a^b f(x)\,dx$$

(c) the set of all convergent sequences of real numbers is subspace of $\mathbb{R}^{\mathbb{N}}$. The limit defines a linear map

$$\begin{aligned} \{\text{convergent sequences}\} \subseteq \mathbb{R}^{\mathbb{N}} &\longrightarrow \mathbb{R} \\ (a_n) &\longmapsto \lim a_n. \end{aligned}$$

□

2.1.6 Definition Let $L: V \longrightarrow W$ be a linear map.

(a) The **image** of L is the set

$$\operatorname{im}(L) := \{ L(v) : v \in V \}$$

(b) The **kernel** of L is the set

$$\ker(L) := \{ v \in V : L(v) = 0 \}$$

□

Note: $L: V \longrightarrow W$ is injective $\Leftrightarrow$ $\ker L = \{0\}$

⌜($\Rightarrow$) Let $v \in \ker L$. Then $L(v) = 0 = L(0)$. Since L is inj, $v = 0$.

($\Leftarrow$) Assume $L(v) = L(w)$. Then $v - w \in \ker L$ since $L(v-w) = L(v) - L(w) = 0$. So $v - w = 0$. ⌟

2.1.7 Proposition Let $L: V \longrightarrow W$ be a linear map. Then

(a) $\operatorname{im}(L)$ is a subspace of W

(b) $\ker(L)$ is a subspace of V.

Proof. Exercise.

□

2.1.8 Definition Let $L: V \to W$ be a linear map. The **rank** of L is the dimension of $\mathrm{im}(L)$ and the **nullity** of L is the dimension of $\ker(L)$. We use the notation

$$\begin{aligned} \mathrm{rank}(L) &:= \dim_F(\mathrm{im}(L)) \\ \mathrm{nullity}(L) &:= \dim_F(\ker(L)). \end{aligned}$$

$\square$

2.1.9 Definition A bijective linear map is called an **isomorphism**. Two vector spaces V and W are **isomorphic** if there exists an isomorphism $L: V \xrightarrow{\sim} W$. We write $V \cong W$. $\square$

2.1.10 Proposition If $L: V \to W$ is an isomorphism, then so is $L^{-1}: W \to V$.

Proof. Let $w_1, w_2 \in W$. Then

$$\begin{aligned} L^{-1}(w_1 + w_2) &= L^{-1}\big(L(L^{-1}(w_1)) + L(L^{-1}(w_2))\big) \\ &= L^{-1}\big(L(L^{-1}(w_1) + L^{-1}(w_2))\big) \\ &= L^{-1}(w_1) + L^{-1}(w_2). \end{aligned}$$

Scalar multiplication is similar.

$\square$

2.1.11 Proposition If $K: U \to V$ and $L: V \to W$ are linear maps, then so is $L \circ K: U \to W$.

Proof. Exercise.

$\square$

2.1.12 Corollary The relation $\cong$ on the class of F-vector spaces is an equivalence relation.

Proof. Reflexive: id_V is an isomorphism.

Symmetric: Prop 2.1.10

Transitive: Prop 2.1.11

□

2.2 Coordinates

2.2.1 Definition Let V be a finite-dimensional vector space. An **ordered basis** for V is a tuple $B = (b_1, \ldots, b_n)$ consisting of the elements of a basis in a certain order.

□

2.2.2 Proposition Let $B = (b_1, \ldots, b_n)$ be an ordered basis for V. For every $v \in V$, there is a unique vector $(\alpha_1, \ldots, \alpha_n) \in F^n$ satisfying

$$v = \sum_{i=1}^{n} \alpha_i b_i .$$

Proof. This follows from Prop 1.4.5.

□

2.2.3 Definition Let $B = (b_1, \ldots, b_n)$ be an ordered basis for V. The **coordinates** of $v \in V$

w/ respect to B are the unique scalars $\alpha_1, \dots, \alpha_n \in F$ satisfying $v = \sum_{i=1}^{n} \alpha_i b_i$. The coordinate vector of v w/ respect to B is the vector

$$[v]_B := (\alpha_1, \dots, \alpha_n) \in F^n.$$

$\square$

Remark: Prop 2.2.2 asserts that the assignment $v \longmapsto [v]_B$ defines a function

$$V \longrightarrow F^n$$
$$v \longmapsto [v]_B$$

Moreover, the function is a bijection w/ invers

$$F^n \longrightarrow V$$
$$(\alpha_1, \dots, \alpha_n) \longmapsto \sum_{i=1}^{n} \alpha_i b_i \text{ .}$$

We will prove that this is an isomorphism.

2.2.4 Proposition Let $B = (b_1, \dots, b_n)$ be an ordered basis for V. Then for all $u, v \in V$ and $\lambda \in F$

(a) $[u+v]_B = [u]_B + [v]_B$

(b) $[\lambda u]_B = \lambda [u]_B$.

Proof. Suppose $[u]_B = (\alpha_1, \dots, \alpha_n)$ and $[v]_B = (\beta_1, \dots, \beta_n)$. Then

$$u + v = \sum_{i=1}^{n} \alpha_i b_i + \sum_{i=1}^{n} \beta_i b_i$$
$$= \sum_{i=1}^{n} (\alpha_i + \beta_i) b_i$$

This shows that

$$[u+v]_B = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)$$
$$= (\alpha_1, \dots, \alpha_n) + (\beta_1, \dots, \beta_n)$$
$$= [u]_B + [v]_B$$

This proves (a). Part (b) is similar. $\square$

2.2.5 Corollary Let $\mathcal{B}$ be an ordered basis for V. Then the function

$$\begin{array}{ccc} V & \longrightarrow & F^n \\ v & \longmapsto & [v]_{\mathcal{B}} \end{array}$$

is an isomorphism.

Proof. Follows immediately from Prop 2.2.4. $\square$

In particular, we now know that every finite-dimensional vector space V over F is isomorphic to $F^{\dim V}$.

2.3 Linear Maps and Bases

Linear maps are uniquely determined by their values on a basis.

2.3.1 Theorem Let $\mathcal{B} = (b_1, \dots, b_n)$ be an ordered basis for a vector space V and let W be any other vector space. For every sequence $c_1, \dots, c_n$ of vectors in W, there is a unique linear map

$$L : V \longrightarrow W$$

such that $L(b_i) = c_i$ for all $i = 1, \dots, n$.

Proof. Define $L : V \to W$ as follows. If $v \in V$ w/ $[v]_{\mathcal{B}} = (\alpha_1, \dots, \alpha_n)$, then $L(v) := \sum_{i=1}^{n} \alpha_i c_i$. This is a linear map: suppose $w \in V$ w/ $[w]_{\mathcal{B}} = (\beta_1, \dots, \beta_n)$.

Then $[v+u]_{\mathcal{B}} = (\beta_1+\alpha_1, \dots, \beta_n+\alpha_n)$ by Prop 2.2.4.

Then
$$\begin{aligned} L(v+u) &= \sum_{i=1}^{n} (\beta_i+\alpha_i) c_i \\ &= \sum_{i=1}^{n} (\beta_i c_i + \alpha_i c_i) \\ &= \sum_{i=1}^{n} \beta_i c_i + \sum_{i=1}^{n} \alpha_i c_i \\ &= L(v) + L(u). \end{aligned}$$

Similarly, $L(\lambda v) = \lambda L(v)$ for all $\lambda \in F$. So L is linear. Moreover, $L(b_i) = c_i$ since $[b_i]_{\mathcal{B}} = e_i$. This proves existence. Suppose $K : V \longrightarrow W$ is another linear map satisfying $K(b_i) = c_i$. Then
$$K(v) = K\left(\sum_{i=1}^{n} \alpha_i b_i\right) = \sum_{i=1}^{n} \alpha_i K(b_i) = \sum_{i=1}^{n} \alpha_i c_i = L(v).$$

This proves uniqueness.

□