1. Vector Spaces

1.1 Characterization of Vector Spaces

Definition 1.1. A $\mathbb{F}$-***vector space*** or **vector space over a field** $\mathbb{F}$ is a set $\mathcal{V}$ paired with operations + (addition) and · (multiplication), which are functions defined as such:

$$+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}; \quad (x, y) \mapsto x + y,$$

and

$$\cdot : \mathbb{F} \times \mathcal{V} \to \mathcal{V}; \quad (c, x) \mapsto c \cdot x.$$

Whenever we say "let $\mathcal{V}$ be a vector space" we tacitly assume it is a vector space over an arbitrary field $\mathbb{F}$, unless otherwise specified.

Furthermore, the following axioms (labeled **A1-A10**) must be satisfied:

1. $x + y \in \mathcal{V}$ for any $x, y \in \mathcal{V}$ (Closure under Addition)
2. $x + y = y + x$ for all $x, y \in \mathcal{V}$ (Commutativity under Addition)
3. $(x + y) + z = x + (y + z)$ for all $x, y, z \in \mathcal{V}$ (Associativity under Addition)
4. There exists $0_{\mathcal{V}} \in \mathcal{V}$ such that $x + 0_{\mathcal{V}} = x$ for all $x \in \mathcal{V}$ (Existence of Additive Identity)
5. For all $x \in \mathcal{V}$, there exists $y \in \mathcal{V}$ such that $x + y = 0_{\mathcal{V}}$. (Existence of Additive Inverse)
6. $cx \in \mathcal{V}$ for any $c \in \mathbb{F}, x \in \mathcal{V}$ (Closure under Multiplication)
7. $c(x + y) = cx + cy$ for any $c \in \mathbb{F}, x, y \in \mathcal{V}$ (Distributivity of Scalars amongst Vectors)
8. $(a + b)x = ax + bx$ for any $a, b \in \mathbb{F}, x \in \mathcal{V}$ (Distributivity of Vectors amongst Scalars)
9. $a(bx) = (ab)x$ for any $a, b \in \mathbb{F}, x \in \mathcal{V}$. (Associativity under Multiplication)
10. $1x = x$ for any $x \in \mathcal{V}$. (Multiplicative Identity)

Theorem 1.1 (Cancellation Law for Vector Addition)**.** *Let $\mathcal{V}$ be a vector space over a field $\mathbb{F}$. If $x, y, z \in \mathcal{V}$ such that $x + z = y + z$, then $x = y$.*

Proof. By **A5**, we know that for $z \in \mathcal{V}$ there exists an additive inverse $u \in \mathcal{V}$ such that $z + u = 0_{\mathcal{V}}$. Thus,

$$\begin{aligned} x &= x + 0_{\mathcal{V}} = x + (z + u) = (x + z) + u \\ &= (y + z) + u = y + (z + u) = y + 0_{\mathcal{V}} = y \end{aligned}$$

follows from **A3** and **A4**. ■

Theorem 1.2. *Let $\mathcal{V}$ be a vector space over a field $\mathbb{F}$. Then following statements are true:*

1. *$\mathcal{V}$ has a unique additive identity, that is, the zero vector in $\mathcal{V}$ is unique.*
2. *Every vector $x \in \mathcal{V}$ has a unique additive inverse.*
3. *$0x = 0_{\mathcal{V}}$ for all $x \in \mathcal{V}$, where $0 \in \mathbb{F}$.*
4. *$(-c)x = -(cx) = c(-x)$ for all $x \in \mathcal{V}$ and $c \in \mathbb{F}$.*
5. *$c0_{\mathcal{V}} = 0_{\mathcal{V}}$ for all $c \in \mathbb{F}$.*

Proof. **(1)** Suppose $0_{\mathcal{V}}$ and $0'_{\mathcal{V}}$ are both additive identities in $\mathcal{V}$, then by **A4**, it follows that

$$0_{\mathcal{V}} + 0'_{\mathcal{V}} = 0_{\mathcal{V}} \quad \text{and} \quad 0'_{\mathcal{V}} + 0_{\mathcal{V}} = 0'_{\mathcal{V}}.$$

Now, by **A2**, $0_{\mathcal{V}} + 0'_{\mathcal{V}} = 0'_{\mathcal{V}} + 0_{\mathcal{V}}$; then, by **Theorem 1.1**, $0_{\mathcal{V}} = 0'_{\mathcal{V}}$, thereby establishing uniqueness of the additive identity in $\mathcal{V}$. **(2)** Let $x \in \mathcal{V}$, and suppose y and y' are additive inverses for x, such that $x + y = 0_{\mathcal{V}} = x + y'$. Then, by application of **A4** and**A3**

$$y = y + 0_{\mathcal{V}} = y + (x + y) = y + (x + y') = (y + x) + y' = 0_{\mathcal{V}} + y = y'.$$

Alternatively, we may see that from the hypothesis, $y = y'$ follows immediately from the lemma previously proven. Thus, uniqueness of the additive inverse in $\mathcal{V}$ is asserted. **(3)** By **A8** it follows that

$$0x = \underbrace{(0 + 0)\,x = 0x + 0x}_{\textbf{A8}}.$$

Now, since $0x \in \mathcal{V}$ (why?), we know by **A5** , it has an additive inverse $y \in \mathcal{V}$ such that $0x + y = 0_{\mathcal{V}}$. So, add y to both sides of the equation above, thereby obtaining

$$0_{\mathcal{V}} = 0x + y = \underbrace{(0x + 0x) + y = 0x + (0x + y)}_{\textbf{A3}} = \underbrace{0x + 0_{\mathcal{V}} = 0x}_{\textbf{A4}} \qquad (\star)$$

as desired. **(4)** Since $cx \in \mathcal{V}$, then we know it has unique additive inverse in $\mathcal{V}$, namely $-(cx)$, such that

$$cx + [-(cx)] = 0_{\mathcal{V}}.$$

Now, if $cx + (-c)\,x = 0_{\mathcal{V}}$ as well, uniqueness of the additive inverse suggests that $(-c)\,x = -(cx)$. Indeed, by **A8** and $(\star)$, it follows that

$$\underbrace{cx + (-c)\,x = [c + (-c)]\,x}_{\textbf{A8}} = \underbrace{0x = 0_{\mathcal{V}}}_{\star}.$$

Thus, since

$$cx + (-c)\,x = 0_{\mathcal{V}} = cx + [-(cx)]$$

it follows from **A2** (applied to both sides of this equation) and the lemma above that $(-c)\,x = -(cx)$. In particular $(-1)\,x = -x$, which in addition to **A9** implies that

$$c\,(-x) = c\,[(-1)\,x] = [c\,(-1)]\,x = (-c)\,x.$$

Consequently, estabilish the chain of equalities $(-c)\,x = -(cx) = c\,(-x)$, as desired. **(4)** By **A7**, it follows that

$$c0_{\mathcal{V}} = c\,(0_{\mathcal{V}} + 0_{\mathcal{V}}) = c0_{\mathcal{V}} + c0_{\mathcal{V}},$$

now, add the additive inverse of$c0_{\mathcal{V}}$(call it$y$) to both sides of the equation above, thereby obtaining

$$0_{\mathcal{V}} = c0_{\mathcal{V}} + y = \underbrace{(c0_{\mathcal{V}} + c0_{\mathcal{V}}) + y = c0_{\mathcal{V}} + (c0_{\mathcal{V}} + y)}_{\textbf{A3}} = \underbrace{c0_{\mathcal{V}} + 0_{\mathcal{V}} = c0_{\mathcal{V}}}_{\textbf{A4}}.$$

And the proof is complete. ∎

1.2 Examples of Vector Spaces

BlH

1.3 Subspaces

Definition 1.2. Let $\mathcal{V}$ be a vector space. We say that $\mathcal{W}$ is a subspace of $\mathcal{V}$ if $\mathcal{W} \subseteq \mathcal{V}$ and $\mathcal{W}$ is a vector space with the same operations of vector addition and scalar multiplication in $\mathcal{V}$.

Remark 1.1. The axioms **A2**, **A3**, and **A7-A10**, are automatically satisfied by any subset $\mathcal{W} \subseteq \mathcal{V}$ of a vector space $\mathcal{V}$. So, what needs be to satisfied in order for $\mathcal{W}$ to be a subspace of $\mathcal{V}$, are the following properties: A1,A4,A6, in addition to A5, the existence of the additive inverse. The following theorem shows that the zero vector in $\mathcal{W}$ must be that of $\mathcal{V}$, and the requirement for the existence of the additive inverse for every vector in$\mathcal{W}$ is redundant, because it can be met by satisftying A1 (**closure under vector addition**), A4 (**existence of neutral element or additive identity**), and A6 (**closure under scalar multiplication**).

Theorem 1.3. *Let $\mathcal{V}$ be a vector space and $\mathcal{W} \subseteq \mathcal{V}$. Then $\mathcal{W}$ is a subspace of $\mathcal{V}$ if and only if the following conditions hold for the operations defined in $\mathcal{V}$:*

1. $0_{\mathcal{V}} \in \mathcal{W}$
2. $x + y \in \mathcal{W}$, *whenever* $x, y \in \mathcal{W}$.
3. $cx \in \mathcal{W}$ *whenever* $c \in \mathbb{F}$ *and* $x \in \mathcal{W}$.

Proof. If $\mathcal{W}$ is a subspace of $\mathcal{V}$, then $\mathcal{W} \subseteq \mathcal{V}$ and $\mathcal{W}$ is a vector space endowed with the same operations of vector addition and scalar multiplication in $\mathcal{V}$. Thus, (2) and (3) follow immediately. Furthermore, there exists a vector $0_{\mathcal{W}} \in \mathcal{W}$ such that for each $w \in \mathcal{W}$, we have $w + 0_{\mathcal{W}} = w$. Now, since $\mathcal{W} \subseteq \mathcal{V}$, it follows that $w \in \mathcal{V}$ as well, therefore there exists $0_{\mathcal{V}} \in \mathcal{V}$ such that $w + 0_{\mathcal{V}} = w$. Hence, by the cancellation law of vector addition, we conclude that $0_{\mathcal{W}} = 0_{\mathcal{V}} \in \mathcal{W}$. Conversely, suppose conditions (1), (2), and (3) hold. It remains to be shown that additive inverse of any vector in $\mathcal{W}$, is also a member of $\mathcal{W}$. Indeed, by (3), we know that if $w \in \mathcal{W}$, then $(-1)\,w \in \mathcal{W}$. Furthermore, by **Theorem 1.2**, we know that $(-1)\,w = -w$, and so the additive inverse of w, namely $-w$, is in $\mathcal{W}$. ■

Theorem 1.4. *Let $\mathcal{V}$ be a vector space over a field $\mathbb{F}$. A subset $\mathcal{W}$ of $\mathcal{V}$ is a vector space iff for all $a, b \in \mathbb{F}$ and $x, y \in \mathcal{V}$, any of the following equivalent conditions hold:*

1. $\mathcal{W} \neq \emptyset$ *and* $ax + by \in \mathcal{W}$.
2. $\mathcal{W} \neq \emptyset$, $x + y \in \mathcal{W}$ *and* $ax \in \mathcal{W}$.
3. $\mathcal{W} \neq \emptyset$, $ax + y \in \mathcal{W}$.

Proof. Exercise. ■

Theorem 1.5. *Let $\mathcal{W}_1$ and $\mathcal{W}_2$ be subspaces of a vector space $\mathcal{V}$, then $\mathcal{W}_1 \cup \mathcal{W}_2$ is a subspace of $\mathcal{V}$ if and only if $\mathcal{W}_1 \subseteq \mathcal{W}_2$ or $\mathcal{W}_2 \subseteq \mathcal{W}_1$.*

Proof. Suppose $\mathcal{W}_1 \cup \mathcal{W}_2$ is a subspace, then for all $a \in \mathbb{F}$ and $x, y \in \mathcal{W}_1 \cup \mathcal{W}_2$, we have $ax + y \in \mathcal{W}_1 \cup \mathcal{W}_2$. Now, suppose for the sake of contradiction that neither $\mathcal{W}_1 \subseteq \mathcal{W}_2$ nor $\mathcal{W}_2 \subseteq \mathcal{W}_1$ hold. Then, consider $x \in \mathcal{W}_1 \setminus \mathcal{W}_2$ and $y \in \mathcal{W}_2 \setminus \mathcal{W}_1$. Then $x + y \notin \mathcal{W}_1$ and $x + y \notin \mathcal{W}_2$, because if, for example, $x + y \in \mathcal{W}_1$, then $y = (x + y) + (-x) \in \mathcal{W}_1$, a contradiction. A result $x + y \notin \mathcal{W}_1$ and $x + y \notin \mathcal{W}_2$ contradicts the hypothesis that $\mathcal{W}_1 \cup \mathcal{W}_2$ is a subspace.

Conversely, suppose without loss of generality, that $\mathcal{W}_1 \subseteq W_2$. Now, suppose $a \in \mathbb{F}$ and $x, y \in \mathcal{W}_1 \cup \mathcal{W}_2$. If $x, y \in \mathcal{W}_1$, then $ax + y \in \mathcal{W}_1 \subseteq \mathcal{W}_1 \cup \mathcal{W}_2$. Now, if $x \in \mathcal{W}_1$ and $y \in \mathcal{W}_2$, then $x \in \mathcal{W}_2$ as well, thus $ax + y \in \mathcal{W}_2 \subseteq \mathcal{W}_1 \cup \mathcal{W}_2$. In all cases, we have $ax + y \in \mathcal{W}_1 \cup \mathcal{W}_2$, thus $\mathcal{W}_1 \cup \mathcal{W}_2$ is a subspace. ∎

1.4 Sums and Products of Vector Spaces and Subspaces

Definition 1.3. The ***internal sum of a finite family of subspaces*** $\{\mathcal{W}_1, \mathcal{W}_2, \ldots, \mathcal{W}_n\} \subseteq \wp(\mathcal{V})$ is defined by

$$\sum_{i=1}^{n} \mathcal{W}_i := \left\{ \sum_{i=1}^{n} w_i \mid w_i \in \mathcal{W}_i \text{ for } 1 \leq i \leq n \right\}.$$

More generally, the ***internal sum of any family of subspaces*** $\{\mathcal{W}\}_{i \in I} \subseteq \wp(\mathcal{V})$ is the set of all finite sums of vectors from the union $\bigcup \mathcal{W}_i$, defined as:

$$\sum_{i \in I} \mathcal{W}_i := \left\{ \sum_{j=1}^{n} w_j \mid w_j \in \bigcup_{i \in I} \mathcal{W}_i \right\}.$$

Moreover, if $\{\mathcal{W}_i\}_{i \in I}$ is a family of subspaces of $\mathcal{V}$ and if:

1. $\mathcal{V} = \sum_{i \in I} \mathcal{W}_i$ and,
2. $\mathcal{W}_i \cap \sum_{i \neq j} \mathcal{W}_j = \{0_{\mathcal{V}}\}$ for all $i \in I$.

are satisfied, then $\mathcal{V}$ is the ***internal direct sum of a family of subspaces***, denoted by

$$\mathcal{V} = \bigoplus_{i \in I} \mathcal{W}_i := \left\{ \sum_{j=1}^{n} w_j \mid w_j \in \bigcup_{i \in I} \mathcal{W}_i, \text{ and } w_{j_i} \in \mathcal{W}_i \text{ if } w_{j_i} \neq 0_{\mathcal{V}} \right\}.$$

Now, if $\{\mathcal{W}_1, \ldots, \mathcal{W}_n\}$ is a finite family of subspaces of $\mathcal{V}$, and $\mathcal{V} = \sum_{i=1}^{n} \mathcal{W}_i$ and $\mathcal{W}_j \cap \sum_{i \neq j} \mathcal{W}_i = \{0_{\mathcal{V}}\}$, for all $j = 1, \ldots, n$, then the ***internal direct sum of a finite family of subspaces*** is

$$\mathcal{V} = \mathcal{W}_1 \oplus \mathcal{W}_2 \oplus + \cdots + \oplus \mathcal{W}_n = \bigoplus_{i=1}^{n} \mathcal{W}_i := \left\{ \sum_{i=1}^{n} w_i \mid w_i \in \mathcal{W}_i \right\}.$$

Theorem 1.6. *Let $\mathcal{V}$ be a vector space over a field $\mathbb{F}$. If $\mathscr{C} = \{\mathcal{W}_i\}_{i \in I}$ is a collection of subspaces of $\mathcal{V}$, then the following*

1. $\sum_{i \in I} \mathcal{W}_i$,
2. $\bigoplus_{i \in I} \mathcal{W}_i$,
3. $\bigcap_{i \in I} \mathcal{W}_i$

are subspaces.

Proof. (1) If $\mathscr{C} = \{\mathcal{W}_i\}_{i \in I}$ is a collection of subspaces, then they are nonempty by **Theorem 1.4**. Thus, if $a \in \mathbb{F}$ and $x, y \in \sum_{i \in I} \mathcal{W}_i$, then ∎

2. Spanning Sets and Linear Independence

2.1 Spanning Sets

Definition 2.1. Let $S = \{v_i\}$ be a subset of a vector space $\mathcal{V}$. A vector $x \in \mathcal{V}$ is a ***linear combination*** of the vectors in S if there is a set of scalars $\{c_i : c_i \in \mathbb{F}\}$, only finitely many of which are nonzero[1], such that $x = \sum_i c_i v_i$. If all $c_i = 0$, then we have $0_{\mathcal{V}} = \sum c_i v_i$ as the ***trivial linear combination*** of vectors in S, otherwise it is ***nontrivial***.

[1] Quoted from here: "Linear combinations of a (possibly infinite) family of vectors are obtained by multiplying a finite number among them by a nonzero scalar and adding up the results. The restriction to a finite number is essential and inevitable, because in algebra it makes no sense to add up infinitely many nonzero vectors (in analysis such infinite sums may be considered, but this requires additional notions of convergence and limits). One can formally admit infinite sums of vectors provided that only finitely many among them are nonzero (and the others do not affect the value of the sum); in this sense one can say a linear combination of an infinite family of vectors is obtained by associating scalars to every vector, but only for finitely many of them the scalar is nonzero, and adding everything up."

Remark 2.1. If $S = \{\}$, the empty set, then the only linear combination we have is the ***empty linear combination*** (which is trivial), whose value is $0_{\mathcal{V}} \in \mathcal{V}$. An infinite set

Definition 2.2. Let $S = \{v_i\}_{i \in I}$ be a nonempty subset of the vector space $\mathcal{V}$. Then the ***span*** of S is the subspace of $\mathcal{V}$ comprised of all linear combinations of finitely many vectors in S, that is,

$$\operatorname{span}(S) := \left\{ \sum_{i \in I} c_i v_i : c_i \in \mathbb{F} \right\}.$$

Furthermore, if $\operatorname{span}(S) = \mathcal{V}$, then S is a ***spanning set*** or ***generating set*** of $\mathcal{V}$. We also say S ***spans*** or ***generates*** $\mathcal{V}$. By definition, if $S = \varnothing$, then $\operatorname{span}\varnothing := \{0\}$.

Lemma 2.1. *If $\mathcal{V}$ is a vector space, then for all $S \subseteq \mathcal{V}$, it follows that* $\operatorname{span}(S)$ *is a subspace of $\mathcal{V}$. Moreover, if $\mathcal{W}$ is a subspace of $\mathcal{V}$ and $S \subseteq \mathcal{W}$, then* $\operatorname{span}(S) \subseteq \mathcal{W}$.

Proof. If $S = \varnothing$, then by definition $\operatorname{span}\varnothing = \{0_{\mathcal{V}}\}$, which is a subspace of $\mathcal{V}$ (prove it!!!). Moreover, if $\mathcal{W}$ is a subspace of $\mathcal{V}$, then $\varnothing \subseteq \mathcal{W}$ and so $\operatorname{span}\varnothing = \{0_{\mathcal{V}}\} \subseteq \mathcal{W}$. Now, if $S \neq \varnothing$, then for $x, y \in S$, we have $x = \sum_{i=1}^n a_i v_i$ and $y = \sum_{j=1}^m b_j w_j$ for $a_i, b_j \in \mathbb{F}$ and $v_i, w_i \in S$ with $(1 \le i \le n, 1 \le j \le m)$. Thus, if $a \in \mathbb{F}$,

$$ax + y = a\sum_{i=1}^n a_i v_i + \sum_{j=1}^m b_j w_j = \sum_{i=1}^n (a a_i) v_i + \sum_{j=1}^m b_j w_j \in \operatorname{span}(S)$$

so, by **Theorem 1.4**, we see that $\operatorname{span}(S)$ is a subspace of $\mathcal{V}$. Now, suppose $\mathcal{W}$ is a subspace of $\mathcal{V}$ and $S \subseteq \mathcal{W}$, then for $x \in \operatorname{span}(S)$, we have $x = \sum_{i=1}^n a_i v_i$ for some $a_i \in \mathbb{F}$ and $v_i \in S$; $(1 \le i \le n)$. Thus, since $S \subseteq \mathcal{W}$, we must have $v_i \in \mathcal{W}$ as well, whence it follows that $x = \sum_{i=1}^n a_i v_i \in \mathcal{W}$, as desired. Hence, $\operatorname{span}(S) \subseteq \mathcal{W}$. ∎

Theorem 2.1. *If $\mathcal{V}$ is a vector space, then $\mathcal{W} \subseteq \mathcal{V}$ is a subspace of $\mathcal{V}$ if and only if* $\operatorname{span}(\mathcal{W}) = \mathcal{W}$.

Proof. Suppose $\mathcal{W} \subseteq \mathcal{V}$ is a subspace of $\mathcal{V}$, then by the second statment of the previous lemma it follows that $\operatorname{span}(\mathcal{W}) \subseteq \mathcal{W}$. Now, if $w \in \mathcal{W}$, then $w = 1 \cdot w$ for $1 \in \mathbb{F}$, thus $w \in \operatorname{span}(\mathcal{W})$, and so $\mathcal{W} \subseteq \operatorname{span}(\mathcal{W})$. Consequently, $\operatorname{span}(\mathcal{W}) = \mathcal{W}$. Conversely, if $\operatorname{span}(\mathcal{W}) = \mathcal{W} \subseteq \mathcal{V}$, then $\mathcal{W}$ is a subspace of $\mathcal{V}$ by the first statement of the previous lemma. ∎

Theorem 2.2. *Let $\mathcal{V}$ be a vector space over a field $\mathbb{F}$ and $S_1, S_2 \subseteq \mathcal{V}$, then the following hold:*

1. *If $S_1 \subseteq S_2$, then* $\operatorname{span}(S_1) \subseteq \operatorname{span}(S_2)$.
2. *If $S_1 \subseteq S_2$ and* $\operatorname{span}(S_1) = \mathcal{V}$*, then* $\operatorname{span}(S_2) = \mathcal{V}$.
3. $\operatorname{span}(S_1 \cup S_2) = \operatorname{span}(S_1) + \operatorname{span}(S_2)$
4. $\operatorname{span}(S_1 \cap S_2) \subseteq \operatorname{span}(S_1) \cap \operatorname{span}(S_2)$.

Proof. **(1)** Suppose $S_1 \subseteq S_2$, then for $x \in S_1$, we have $x = 1 \cdot x \in \operatorname{span}(S_1)$. Now, since $S_1 \subseteq S_2$, it follows that $x = 1 \cdot x \in \operatorname{span}(S_2)$. Consequently, $\operatorname{span}(S_1) \subseteq \operatorname{span}(S_2)$.

(2) If $S_1 \subseteq S_2$ and $\operatorname{span}(S_1) = \mathcal{V}$, then by the preceeding result, we know that $\operatorname{span}(S_1) = \mathcal{V} \subseteq \operatorname{span}(S_2)$. Now, to show that $\operatorname{span}(S_2) \subseteq \operatorname{span}(S_1)$, observe that any vector in $v \in S_2$, may be expressed may be expressed $v = 1 \cdot v \in \operatorname{span}(S_2)$. Furthermore, since since $S_2 \subseteq \mathcal{V} = \operatorname{span}(S_1)$, it follows that $v = 1 \cdot v \in \operatorname{span}(S_1)$. Hence, $\operatorname{span}(S_2) \subseteq \operatorname{span}(S_1)$. Consequently, $\operatorname{span}(S_1) = \mathcal{V} = \operatorname{span}(S_2)$, as desired.

(3) Suppose $x \in \operatorname{span}(S_1 \cup S_2)$, then there exists $v_i \in S_1$, $w_j \in S_2$ and $a_i, b_j \in \mathbb{F}$ such that

$$x = \sum_{i=1}^{n} a_i v_i + \sum_{j=1}^{m} b_j w_j, \quad (1 \le i \le n, 1 \le j \le m)$$

where $\sum_i a_i v_i \in \operatorname{span}(S_1)$ and $\sum_j b_j w_j \in \operatorname{span}(S_2)$, whence $x \in \operatorname{span}(S_1) + \operatorname{span}(S_2)$. Thus, $\operatorname{span}(S_1 \cup S_2) \subseteq \operatorname{span}(S_1) + \operatorname{span}(S_2)$. Now, suppose $v = x + y \in \operatorname{span}(S_1) + \operatorname{span}(S_2)$, where $x \in \operatorname{span}(S_1)$ and $y \in \operatorname{span}(S_2)$. Then by (1), it follows that since $S_1 \subseteq S_1 \cup S_2$ and $S_2 \subseteq S_1 \cup S_2$, that $\operatorname{span}(S_1) \subseteq \operatorname{span}(S_1 \cup S_2)$ and $\operatorname{span}(S_2) \subseteq \operatorname{span}(S_1 \cup S_2)$. Hence, it follows that $x, y \in \operatorname{span}(S_1 \cup S_2)$, and since $\operatorname{span}(S_1 \cup S_2)$ is a subspace (by **Lemma 2.1**), we have $v = x + y \in \operatorname{span}(S_1 \cup S_2)$. Thus, $\operatorname{span}(S_1) + \operatorname{span}(S_2) \subseteq \operatorname{span}(S_1 \cup S_2)$. We therefore conclude $\operatorname{span}(S_1 \cup S_2) = \operatorname{span}(S_1) + \operatorname{span}(S_2)$.

(4) Suppose $x \in \operatorname{span}(S_1 \cap S_2)$, then by ■

2.2 *Linear Independence*

Definition 2.3. Let $S = \{v_i\}$ be subset of a vector space $\mathcal{V}$. Then we say that S is ***linearly independent*** if and only if the linear combination of elements of $\mathcal{V}$ that is equal to zero, is the trivial linear combination, that is, $0_{\mathcal{V}} = \sum_i c_i v_i$ implies $c_i = 0$ for all i.

Definition 2.4. A *nonempty* subset S of a vector space $\mathcal{V}$ is ***linearly dependent*** if and only if there *exists a finite number of distinct vectors* $v_1, v_2, \ldots, v_k$ in S and scalars (**not all zero**) $c_1, c_2, \ldots, c_k \in \mathbb{F}$, such that $\sum_{i=1}^{k} c_i v_i = 0$. Alternatively, we may say S is linearly dependent if and only if there exists some collection T of distinct vectors in S, such that we can obtain a nontrivial linear combination of the vectors in T.

Remark 2.2. The empty set $\varnothing = \{\}$ is linearly independent since linearly dependent sets must, by definition, be nonempty.

Lemma 2.2. *Let $\mathcal{V}$ be a vector space, the following statements are true:*

1. *The set $S = \{0_{\mathcal{V}}\}$ is linearly dependent.*
2. *The singleton set $S = \{v\}$such that $v \neq 0_{\mathcal{V}}$ is linearly independent.*
3. *Suppose $S_1 \subseteq S_2 \subseteq \mathcal{V}$. If S_1 is linearly dependent, then so is S_2.*
4. *Suppose $S_1 \subseteq S_2 \subseteq \mathcal{V}$. If S_2 is linearly independent, then so is S_1.*
5. *Any set containing the zero vector is linearly dependent.*

Proof. (1) Since $0_{\mathcal{V}} \in S$ and $0_{\mathcal{V}} \in \mathcal{V}$, then there exists a multiplicative indentity in $\mathbb{F}$, such that

$$0_{\mathcal{V}} = 1 \cdot 0_{\mathcal{V}},$$

hence the representation of $0_{\mathcal{V}}$ as a linear combination of itself is non-trivial with $c = 1$.(2)Indeed, suppose instead that the set $\{u\} \subseteq \mathcal{V}$ where $u \neq 0_{\mathcal{V}}$ is linearly dependent, then $cu = 0_{\mathcal{V}}$ for some non-zero scalar $c \in \mathbb{F}$. Hence, there exists $c^{-1} \in \mathbb{F}$ such that

$$u = c^{-1}(cu) = c^{-1}0_{\mathcal{V}}$$

whereby (5) in **Theorem 1.2**, it follows that $c^{-1}0_{\mathcal{V}} = 0_{\mathcal{V}}$, thereby implying that $u = 0_{\mathcal{V}}$, a contradiction. (3) If S_1 is linearly dependent, then we have a finite number of distinct vectors $v_1, v_2, \dots, v_n \in S_1$ and scalars (not all zero) $c_1, c_2, \dots, c_n \in \mathbb{F}$, such that $\sum_{i=1}^{n} c_i v_i = 0_{\mathcal{V}}$is nontrivial. Furthemore, this nontrivial representaion of $0_{\mathcal{V}}$ as linear combination of vectors in S_1serves as a nontrivial linear combination of vectors in S_2 since $S_1 \subseteq S_2$. (4) This is just the contrapositive of the previous statement. (5) Let $S \subseteq \mathcal{V}$ be the set containing the zero vector. Then $\{0_{\mathcal{V}}\} \subseteq S$, and by (1), we know that $\{0_{\mathcal{V}}\}$ is linearly dependent, thus, it follows from (4), that S is linearly dependent. ∎

Definition 2.5. Let S be a nonempty set of vectors in a vector space $\mathcal{V}$, that is, $S \subseteq \mathcal{V} \setminus \varnothing$. A nonzero vector x in $\mathcal{V}$ is an ***essentially unique linear combination*** of the vectors in S if, up to order of terms, there is one and only one way to express a vector $x \in \mathcal{V}$ as a linear combination

$$x = \sum_{i=1}^{n} c_i v_i$$

where $v_i \in S$ are distinct and $c_i \in \mathbb{F}$ are nonzero. More precisely, if there are $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{F} \setminus \{0\}$, and distinct $v_1, \dots, v_n \in S$ and distinct $w_1, \dots, w_m \in S$, then (re-indexing b_is if necessary), we have

$$\begin{cases} x = \sum_{i=1}^{n} a_i v_i \\ x = \sum_{i=1}^{m} b_i w_i \end{cases} \Rightarrow m = n \text{ and } \begin{cases} a_i = b_i \\ u_i = v_i \end{cases} \quad \text{for all } i.$$

Theorem 2.3. *Let $S \neq \{0_{\mathcal{V}}\}$ be a nonempty set of vectors in $\mathcal{V}$. The follow statements are equivalent:*

1. *S is linearly independent.*
2. *Every nonzero vector in the span of S is an essentially unique linear combination of the vectors in S.*
3. *No vector in S can be expressed as a linear combination of other vectors in S.*

Proof. ($\mathbf{1} \Rightarrow \mathbf{2}$). Suppose S is linearly indepedent and that for $x \in \mathcal{V} \setminus \{0_{\mathcal{V}}\}$, we have

$$x = \sum_{i=1}^{n} c_i v_i \quad \text{and} \quad x = \sum_{i=1}^{m} c_i' w_i$$

where $c_i, c_i' \in \mathbb{F} \setminus \{0\}$, $v_i \in S$ are distinct (amongst themselves) and $w_i \in S$ are distinct (amongst themselves). Then subtracting both sums and grouping k terms where the vectors v_i and w_i are equal, we obtain

$$0_{\mathcal{V}} = \sum_{r=1}^{k} \left(c_{i_r} - c_{i_r}'\right) v_{i_r} + \sum_{r=k+1}^{n} c_{i_r} v_{i_r} - \sum_{r=k+1}^{m} c_{i_r}' w_{i_r}.$$

Now, since S is linearly independent, the only representation we must have of $0_{\mathcal{V}} \in \mathcal{V}$ as a linear combination of vectors in S is trivial[2]. Thus, it follows that $k = m = n$, $c_{i_r} = c'_{i_r}$, and $v_{i_r} = w_{i_r}$for all i. $(\mathbf{2} \Rightarrow \mathbf{3})$ Now, suppose (2) holds. Also, suppose, for the sake of contradiction, that $v \in S$ is a linear combination of other vectors in S, that is, suppose that for distinct $v_i \in S$, that $v = \sum_i c_i v_i$ where v_i are distinct from v. Now, for those $v_i = kv_j$, we collect like terms, and remove all terms with zero as their coefficient. We are then left with an expression that contradictsthe definition of an essentially unique linear combination. $(\mathbf{3} \Rightarrow \mathbf{1})$ Assume no v in S can be written as a linear combination of other vectors in S, and for the sake of contradiction, that S is linearly dependent, that is, there exist scalars $c_1, c_2, \ldots, c_n$, not all zero, such that

$$0_{\mathcal{V}} = c_1 v_1 + \cdots + c_{i-1} v_{i-1} + c_i v_i + c_{i+1} v_{i+1} + \cdots + c_n v_n.$$

Now, without loss of generality, let i be the largest index such that $c_i \neq 0$. If $i = 1$, then we have $0_{\mathcal{V}} = c_1 v_1$, where $c_1 \neq 0$ and $v_1 = 0$, but this is a contradiction since $v_1 \in S$ and $S \neq \{0_{\mathcal{V}}\}$. Now, if $i \geq 2$, then we have

$$\begin{aligned} v_i &= -\frac{1}{c_i}\left(c_1 v_1 + \cdots + c_{i-1} v_{i-1} + c_{i+1} v_{i+1} + \cdots + c_n v_n\right), \\ &= c_i^{-1} c_1 v_1 + c_i^{-1} c_{i-1} v_{i-1} + c_i^{-1} c_{i+1} v_{i+1} + \cdots + c_i^{-1} c_n v_n \end{aligned}$$

in which case, v_i is linear combination of other vectors in S, a contradiction. ∎

[2] Thus $c_{i_r} - c'_{i_r} = 0$, and so $c_{i_r} = c'_{i_r}$.

Here's an example to see why the outcome in $(\mathbf{2} \Rightarrow \mathbf{3})$ contradicts (2). Suppose we have distinct $v_1, v_2, v_3 \in S$, and $v_1 = av_2 \Leftrightarrow \frac{1}{a} v_1 = v_2$. Now, suppose $v \in S$ is distinct from these other vectors and can be expressed as:

$$\begin{aligned} v &= v_1 + v_2 + c_3 v_3 \\ &= (a+1) v_2 + c_3 v_3 \\ &= \left(1 + \frac{1}{a}\right) v_1 + c_3 v_3. \end{aligned}$$

Then, by the definition of an essentially unique linear combination, we deduce that $v_1 = v_2$, which is a contradiction.

Corollary 2.1. *Let $\mathcal{V}$ be a vector space over a field $\mathbb{F}$ and $S \subseteq \mathcal{V}$ be a nonempty subset of $\mathcal{V}$. S is linearly dependent iff $S = \{0\}$ or there exist distinct vectors in $v, u_1, \ldots, u_n \in S$ such that v is a linear combination of $u_1, \ldots, u_n$.*

Proof. From the previous theorem we know that statements(1) and (3) are equivalent. Therefore, taking the contrapositive, this result immediately follows. ∎

2.3 Bases and Dimension

Theorem 2.4. *Let β be a subset of a vector space $\mathcal{V}$. The following statements are equivalent:*

1. *β is linearly independent and spans $\mathcal{V}$*
2. *Every nonzero vector $v \in \mathcal{V}$ is an essentially unique linear combination of vectors in β.*
3. *β is a minimal spanning set for $\mathcal{V}$, that is, β spans $\mathcal{V}$ but any proper subset of β does not span $\mathcal{V}$.*
4. *β is a maximal linearly independent subset of $\mathcal{V}$, that is, β is linearly independent, however, any proper superset of β is not linearly independent.*

Proof. $(1 \Leftrightarrow 2)$ Follows from **Theorem 1.2**. $(1 \Rightarrow 3)$ Suppose β is linearly independent and spans $\mathcal{V}$. Now, suppose for the sake of contradiction that any proper subset of β spans $\mathcal{V}$, that is, $\beta' \subset \beta$ spans $\mathcal{V}$. Then any vector in $\beta \setminus \beta'$ would be a linear combination of the vectors in β', there by contradicting the linear independence of β. $(3 \Rightarrow 1)$ Conversely, suppose β is a minimal spanning set for $\mathcal{V}$, yet not linearly dependent. Them some vector $v \in \beta$ would be a linear combination of other vectors in β, and so $\beta \setminus \{v\}$ is a proper spanning subset of $\mathcal{V}$, a contradiction. $(1 \Rightarrow 4)$ Suppose β is linearly independent and spans $\mathcal{V}$, however β isn't a maximal linearly independent set, thus there exists $x \in \mathcal{V} \setminus \beta$ such that the superset $\beta \cup \{x\}$ is linearly independent; thus x cannot be expressed as a linear combination of the vectors in β, thereby contradicting the hypothesis that β spans $\mathcal{V}$. $(4 \Rightarrow 1)$ Suppose β is a maximal linearly independent subet of $\mathcal{V}$; however it doesn't span $\mathcal{V}$. Then there exists $x \in \mathcal{V} \setminus \{\beta\}$ such that $x \notin \operatorname{span} \beta$. Now, consider the superset $\{x\} \cup \beta$. This superset is linearly independent, thereby contradicting the maximality of β as a linearly independent subset of $\mathcal{V}$. ∎

Definition 2.6. A subset β of a vector space $\mathcal{V}$ satisfying the conditions in the theorem above is called a ***basis*** of $\mathcal{V}$.

Theorem 2.5. *Let $\mathcal{V}$ be a vector space and let $\beta = \{v_i\}_{i \in I}$ be a set of vectors in $\mathcal{V}$. Then β is a basis of $\mathcal{V}$ if and only if every $x \in \mathcal{V}$ can be expressed as a unique linear combination of the vectors in β, that is x, can be uniquely written as $x = \sum_{i \in I} c_i v_i$ where $c_i \in \mathbb{F}$, all but finitely many zero.*

Proof. Suppose β is a for $\mathcal{V}$, then β spans $\mathcal{V}$, that is,

$$\operatorname{span} \beta = \left\{ \sum_{i \in I} c_i v_i \,:\, c_i \in \mathbb{F} \right\} = \mathcal{V}. \tag{$\star$}$$

Furthermore, β is linearly independent, that is, $0 = \sum_i c_i v_i$ implies that $c_i = 0$ for all i. Now, suppose $x \in \mathcal{V}$, and by virtue of $(\star)$, let $x = \sum_i c_i v_i$ and $x = \sum_i c_i' v_i$ where $c_i, c_i' \in \mathbb{F}$. Then

$$x - x = 0_{\mathcal{V}} = \sum_{i \in I} c_i v_i - \sum_{i \in I} c_i' v_i = \sum_{i \in I} (c_i - c_i') v_i.$$

Now, from this, we see that $(c_i - c_i') = 0$ for all i because the set $\beta = \{v_i\}$ is linearly independent. Hence, $c_i = c_i'$, and so the linear combination $x = \sum_i c_i v_i$ is unique[3]. Conversely, let $\beta = \{v_i\}$ be a set of vectors in $\mathcal{V}$, and suppose every $x \in \mathcal{V}$ can be expressed uniquely as $x = \sum_i c_i v_i$. Consequently, β spans $\mathcal{V}$. Now, consider $0_{\mathcal{V}} \in \mathcal{V}$, it therefore follows by our hypothesis that $0_{\mathcal{V}}$ can be uniquely expressed as

$$0_{\mathcal{V}} = \sum_i c_i v_i \tag{$\star$}$$

Clearly, $0_{\mathcal{V}} = \sum_i 0 v_i$, so then by uniqueness of $(\star)$, it must follow that $c_i = 0$ for all i. Thus, β is linearly independent and it spans $\mathcal{V}$, making it a basis of $\mathcal{V}$. ∎

[3] Excluding order of summation!

Definition 2.7. Let $\mathcal{V}$ be a vector space. The ***dimension*** of $\mathcal{V}$, denoted by $\dim \mathcal{V}$, is the number of vectors in any basis of $\mathcal{V}$.

2.4 Sums and Products of Vector Spaces and Subspaces

- Consider a family of subspaces of V, $\{W_i\}_{i \in I}$, where each $w \in \mathcal{V}$, may be expressed uniquely as

  $$w = \sum_{i \in I} w_i$$

  where each $w_i \in W_i$. Furthemore, $W_i = \langle x_i \rangle$, that is, all multiples of x_i. Then

  $$w = \sum_{i \in I} w_i = \sum_{i \in I} c_i x_i$$

  for some scalars $c_i \in \mathbb{F}$.

Definition 2.8. Let $\mathcal{V}$ be a vector space and let $\mathscr{C} = \{\mathcal{W}_1, \ldots, \mathcal{W}_k\}$ be a collection of subspaces of $\mathcal{V}$, where each $\mathcal{W}_i$ consists of multiples of a single nonzero vector $v_i \in \mathcal{V}$.

Notation 1. If one wants to describe a k-tuple of vectors chosen from the n-tuple $(x_1, \ldots, x_n)$, we must used indexed indices. That is, each k-tuple can be written as $(x_{i_1}, \ldots, x_{i_k})$, where i_α are integers with $1 \leq i_a \leq n$ and $\alpha = 1, \ldots k$. If in addition, we require that no w_i be used more than one, one must suppose that the i_α be pairwise distinct, that is, $i_\alpha \neq i_\beta$ for $\alpha \neq \beta$.

3. Linear Transformations

Definition 3.1. A ***linear transformation*** (***general linear operato***r or ***homomorphism***) of a vector space $\mathcal{V}$ onto another vector space $\mathcal{W}$ is a mapping $\mathsf{T} : \mathcal{V} \to \mathcal{W}$ that assigns to each vector $x \in \mathcal{V}$, some vector $\mathsf{T}(x) \in \mathcal{W}$, and it satisfies the following properties:

In these notes I will frequently interchange the use of general linear operator and linear transformation.

1. $\mathsf{T}(x + y) = \mathsf{T}(x) + \mathsf{T}(y)$ (Additivity)

2. $\mathsf{T}(cx) = c\mathsf{T}(x)$ (Homogeneity)

for all $x, y \in \mathcal{V}$ and $c \in \mathbb{F}$.

Remark. To show that some operator T is linear it suffices to show that

$$\mathsf{T}(cx + y) = c\mathsf{T}(x) + \mathsf{T}(y)$$

where $x, y \in \mathcal{V}$ and $c \in \mathbb{F}$.

Definition 3.2. Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over a scalar field $\mathcal{F}$. We denote the set $\mathcal{L}(\mathcal{V}, \mathcal{W})$ as **the set of all linear transformations from $\mathcal{V}$ into $\mathcal{W}$ over a field $\mathbb{F}$**. This is also denoted as $\mathrm{Hom}_{\mathbb{F}}(\mathcal{V}, \mathcal{W})$ for **the space of $\mathbb{F}$-homomorphisms from $\mathcal{V}$ to $\mathcal{W}$**. Thus, if $\mathsf{T} : \mathcal{V} \to \mathcal{W}$, is linear, we say that T belongs to $\mathcal{L}(\mathcal{V}, \mathcal{W})$, that is, $\mathsf{T} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$.

Notation 2. If the argument of some linear transformation $\mathsf{T} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ isn't too complicated, we shall often write $\mathsf{T}x$ instead of $\mathsf{T}(x)$, where the context will make it clear that x is a vector in $\mathcal{V}$.

Definition 3.3. Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces and let $\mathsf{T} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$. The ***kernel*** or ***nullspace*** of T is the set of all vectors x in $\mathcal{V}$ such that $\mathsf{T}x = 0_{\mathcal{W}}$, that is,

$$\ker(\mathsf{T}) := \{x \in \mathcal{V} : \mathsf{T}x = 0_{\mathcal{W}}\}.$$

The ***range*** or ***image*** of T, often denoted by $\mathcal{R}(\mathsf{T})$ or $\mathrm{range}(\mathsf{T})$, is the subset of $\mathcal{W}$ comprised of all images of vectors in $\mathcal{V}$ under T, that is,

$$\mathrm{range}(\mathsf{T}) := \{\mathsf{T}x : x \in \mathcal{V}\}.$$

The ***rank*** of T is the dimension of its range space, that is,

$$\mathrm{rank}(\mathsf{T}) := \dim(\mathrm{range}(\mathsf{T}))$$

and the ***nullity*** of T is dimension of its kernel space, that is,

$$\mathrm{nullity}(\mathsf{T}) := \dim(\ker(\mathsf{T})).$$

Definition 3.4. Let $\mathcal{V}$ be a vector space over a field $\mathbb{F}$, and $\mathsf{T} \in \mathcal{L}(\mathcal{V})$, be a linear operator. A subspace $\mathcal{W}$ of $\mathcal{V}$ is said to be a T-***invariant subspace*** if $\mathsf{T}(x) \in \mathcal{W}$ for all $x \in \mathcal{W}$, that is, $\mathsf{T}(\mathcal{W}) \subseteq \mathcal{W}$. The ***restriction of*** T ***to a*** T-***invariant subspace*** $\mathcal{W}$ is denoted by $\mathsf{T}_{\mathcal{W}} : \mathcal{W} \to \mathcal{W}$ defined by $\mathsf{T}_{\mathcal{W}}(x) = \mathsf{T}(x)$ for all $x \in \mathcal{W}$.

3.1 Basic Properties of Linear Transformations

Theorem 3.1. *Let* $\mathsf{T} : \mathcal{V} \to \mathcal{W}$ *be a linear transformation, and* $u, v, w, 0_{\mathcal{V}} \in \mathcal{V}$, *and* $0_{\mathcal{W}} \in \mathcal{W}$. *Then:*

1. $\mathsf{T}(0_{\mathcal{V}}) = 0_{\mathcal{W}}$
2. $\mathsf{T}(-x) = -\mathsf{T}(x)$
3. $\mathsf{T}(x - y) = \mathsf{T}(x) - \mathsf{T}(y)$
4. $\mathsf{T}\left(\sum_{i=1}^{n} x\right) = n\mathsf{T}(x)$.
5. $\mathsf{T}(-cx) = -c\mathsf{T}(x)$, where $c \in \mathbb{F}$.

Proof. (1) For $0_{\mathcal{V}} \in \mathcal{V}$, we know from vector space properties, that $0_{\mathcal{V}} + 0_{\mathcal{V}} = 0_{\mathcal{V}}$, thus

$$\mathsf{T}(0_{\mathcal{V}}) = \mathsf{T}(0_{\mathcal{V}} + 0_{\mathcal{V}}) = \mathsf{T}(0_{\mathcal{V}}) + \mathsf{T}(0_{\mathcal{V}}),$$

whereby subtracting $\mathsf{T}(0_{\mathcal{V}})$ from both sides, we obtain $0_{\mathcal{W}} = \mathsf{T}(0_{\mathcal{V}})$, as desired. Now, for (2), note that for $x \in \mathcal{V}$, we have

$$\mathsf{T}(0_{\mathcal{V}}) = \mathsf{T}(x + (-x)) = \mathsf{T}(x) + \mathsf{T}(-x).$$

Thus, in view of our previous result in (1), it follows that

$$\mathsf{T}(-x) = \mathsf{T}(0_{\mathcal{V}}) - \mathsf{T}(x) = 0_{\mathcal{W}} - \mathsf{T}(x) = -\mathsf{T}(x),$$

as desired. Now, for (3), it follows that

$$\mathsf{T}(x - y) = \mathsf{T}(x + (-y)) = \mathsf{T}(x) + \mathsf{T}(-y) = \mathsf{T}(x) - \mathsf{T}(y),$$

where the last equality follows from (2). Now, for (4), we should know that $\sum_{i=1}^{n} x = nx$, where $n \in \mathbb{N}$. Thus, $\mathsf{T}\left(\sum_{i=1}^{n} x\right) = \mathsf{T}(nx) = n\mathsf{T}(x)$, by homogeneity of T. Lastly, for (5), we have $\mathsf{T}(-cx) = \mathsf{T}(-(cx)) = -\mathsf{T}(cx) = -c\mathsf{T}(x)$. ∎

Theorem 3.2. *If* $\mathcal{V}$ *and* $\mathcal{W}$ *are subspaces over the same field* $\mathbb{F}$ *and* $\mathsf{T} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, *then* $\ker(\mathsf{T})$ *and* $\operatorname{range}(\mathsf{T})$ *are subspaces of* $\mathcal{V}$ *and* $\mathcal{W}$, *respectively. Moreover, if* $\mathsf{T} \in \mathcal{L}(\mathcal{V})$, *then* $\{0_{\mathcal{V}}\}$, $\ker(\mathsf{T})$, $\operatorname{range}(\mathsf{T})$, *and* $\mathcal{V}$ *are* T*-invariant.*

Proof. Since $\mathsf{T}(0_{\mathcal{V}}) = 0_{\mathcal{W}}$, we know that $0_{\mathcal{V}} \in \ker(\mathsf{T})$. Now, suppose $x, y \in \ker(\mathsf{T})$, then by definition $\mathsf{T}x = 0_{\mathcal{W}}$ and $\mathsf{T}y = 0_{\mathcal{W}}$. Furthermore,

$$\mathsf{T}(x + y) = \mathsf{T}x + \mathsf{T}y = 0_{\mathcal{W}} + 0_{\mathcal{W}} = 0_{\mathcal{W}}.$$

Hence, $x + y \in \ker(\mathsf{T})$. Now, suppose $c \in \mathbb{F}$ and $x \in \ker(\mathsf{T})$, then $\mathsf{T}x = 0$; furthermore,

$$\mathsf{T}cx = c\mathsf{T}x = c \cdot 0_{\mathcal{W}} = 0_{\mathcal{W}},$$

thus $cx \in \ker(\mathsf{T})$, and we conclude that $\ker(\mathsf{T})$ is a subspace of $\mathcal{V}$. Now, to show $\operatorname{range}(\mathsf{T})$ is a subspace of $\mathcal{W}$, observe that because $\mathsf{T}(0_{\mathcal{V}}) = 0_{\mathcal{W}}$, it follows by definition of range that the image of $0_{\mathcal{V}}$ under T is contained in $\mathcal{W}$, thus, $0_{\mathcal{W}} \in \operatorname{range}(\mathsf{T})$. Now, suppose $x, y \in \operatorname{range}(\mathsf{T})$ and $c \in \mathbb{F}$, then there exists $u, v \in \mathcal{V}$, such that $\mathsf{T}u = x$ and $\mathsf{T}v = y$, hence

$$\mathsf{T}(u + v) = \mathsf{T}u + \mathsf{T}v = x + y$$

implies that $x + y \in \operatorname{range}(\mathsf{T})$ for $u + v \in \mathcal{V}$. Furthermore, $\mathsf{T}(cu) = c\mathsf{T}u = c \cdot x$, thus $c \cdot x \in \operatorname{range}(\mathsf{T})$ for $cu \in \mathcal{V}$. Hence, we conclude that $\operatorname{range}(\mathsf{T})$ is a subspace of $\mathcal{W}$. Let $\mathsf{T} : \mathcal{V} \to \mathcal{V}$, then we have $\mathsf{T}(\{0_{\mathcal{V}}\}) \subseteq \mathcal{V}$, since $\mathsf{T}(0_{\mathcal{V}}) = 0_{\mathcal{V}}$; hence $\{0_{\mathcal{V}}\}$ is T-invariant. Furthermore, $\mathcal{V}$ is T-invariant, since the co-domain of $\mathsf{T} : \mathcal{V} \to \mathcal{V}$ is $\mathcal{V}$, hence, $\mathsf{T}(\mathcal{V}) \subseteq \mathcal{V}$. Now, for $\ker(\mathsf{T})$, we see that for all $x \in \ker(\mathsf{T})$, that $\mathsf{T}x = 0_{\mathcal{V}}$, then since $\ker(\mathsf{T})$ is a subspace of $\mathcal{V}$, we know that $0_{\mathcal{V}} \in \ker(\mathsf{T})$; thus $\mathsf{T}(\ker(\mathsf{T})) \subseteq \mathcal{V}$. Lastly, if $x \in \operatorname{range}(\mathsf{T})$, then there exists $y \in \mathcal{V}$, such that $\mathsf{T}x = y \in \mathcal{V}$; hence, $\mathsf{T}(\operatorname{range}(\mathsf{T})) \subseteq \mathcal{V}$. ∎

3.2 *Types of Linear Transformations*

Definition 3.5. Let $\mathsf{T} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, then T is:

1. an ***endomorphism*** or ***linear operator*** if $\mathcal{V} = \mathcal{W}$. This is not to be confused with a general linear operator because the former requires that $\mathcal{V} = \mathcal{W}$. The **set of all linear operators** is denoted by $\mathcal{L}(\mathcal{V})$ or $\operatorname{End}_{\mathbb{F}}(\mathcal{V})$.

2. a ***monomorphism*** if it is injective, that is if $x_1, x_2 \in \mathcal{V}$, then $\mathsf{T}(x_1) = \mathsf{T}(x_2)$ implies $x_1 = x_2$.

3. an ***epimorphism*** if it is surjective, that is, for all $y \in \mathcal{W}$, there exists $x \in \mathcal{V}$ such that $\mathsf{T}(x) = y$.

4. an ***isomorphism*** if it is surjective and injective, that is, bijective. The **set of all linear isomporphisms from $\mathcal{V}$ to $\mathcal{W}$** is denoted by $\operatorname{GL}(\mathcal{V}, \mathcal{W})$, analoguous to the general linear group.

5. an ***automorphism*** if it is bijective and $\mathcal{V} = \mathcal{W}$, essentially a bijective linear operator.

Remark 3.1. If an isomporphism from $\mathcal{V}$ to $\mathcal{W}$ exists, then we say that $\mathcal{V}$ is **isomorphic to** $\mathcal{W}$, usually denoted as $\mathcal{V} \cong \mathcal{W}$.

Definition 3.6. Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces.

1. The linear transformation $\mathsf{I}_{\mathcal{V}} : \mathcal{V} \to \mathcal{V}$ defined by $\mathsf{I}_{\mathcal{V}}(x) = x$ for all $x \in \mathcal{V}$ is the ***identity linear transformation***.

2. The linear transformation $\mathsf{T}_0 : \mathcal{V} \to \mathcal{W}$ defined by $\mathsf{T}_0(x) = 0$ for all $x \in \mathcal{V}$ is the ***zero transformation***.

3. The linear operator $\mathsf{T} \in \mathcal{L}(\mathcal{V})$ such that $\mathsf{T}^2 = \mathsf{T}$ is the **idempotent operator**.

Definition 3.7. Let $\mathsf{T}, \mathsf{U} : \mathcal{V} \to \mathcal{W}$ be arbitrary functions, where $\mathcal{V}$ and $\mathcal{W}$ are vector spaces over a field $\mathbb{F}$. Then, we define:

1. $\mathsf{T} + \mathsf{U} : \mathcal{V} \to \mathcal{W}$ by $(\mathsf{T} + \mathsf{U})(x) = \mathsf{T}(x) + \mathsf{U}(x)$ for all $x \in \mathcal{V}$.

2. $c\mathsf{T} : \mathcal{V} \to \mathcal{W}$ by $(c\mathsf{T})(x) = c\mathsf{T}(x)$ for all $x \in \mathcal{V}$, $c \in \mathbb{F}$.

3.3 *Linear Transformations are Uniquely Determined by their Action on a Basis.*

Theorem 3.3. *Let $\mathcal{V}$ and $\mathcal{W}$ be vector spaces over the same field $\mathbb{F}$, and let $\beta = \{v_i\}_{i \in I}$ be a basis for $\mathcal{V}$ and $\{w_i\}_{i \in I}$ be an arbitrary set of vectors in $\mathcal{W}$. Then a unique well-defined* *linear transformation* T *can be obtained from* $\mathcal{L}(\mathcal{V},\mathcal{W})$ *by arbitrarily specifying the values* $w_i = \mathsf{T}(v_i) \in \mathcal{W}$ *for each* $i \in I$ *and by specifying that for all* $c_1, \ldots, c_n \in \mathbb{F}$, *we have* T *satisfy:*

$$\mathsf{T}\left(\sum_{i=1}^{n} c_i v_i\right) = \sum_{i=1}^{n} \mathsf{T}(c_i v_i) = \sum_{i=1}^{n} c_i \mathsf{T}(v_i) = \sum_{i=1}^{n} c_i w_i, \tag{$\star$}$$

thereby endowing T *with linearity.*

Proof. We know that for all $v \in \mathcal{V}$, we may express v as a unique linear combination of the vectors in the basis β, that is, v may be uniquely expressed as $v = \sum_{i\in I} c_i v_i$, all but finitely many zero. Now, defining the function $\mathsf{T} \in \mathcal{W}^{\mathcal{V}}$ (the set of all functions from $\mathcal{V}$ to $\mathcal{W}$)[4] as in $(\star)$, we proceed to show that T is linear, that is, $\mathsf{T} \in \mathcal{L}(\mathcal{V},\mathcal{W})$. Indeed, for $u, v \in \mathcal{V}$, we know by virtue of previous theorem (name it!), that there exist unique scalars $a_1, \ldots, a_n, b_1, \ldots, b_m \in \mathbb{F}$ and $u_1, \ldots u_n, v_1, \ldots v_m \in \beta$ such that

[4] Not necessarily linear, but we can say that $\mathcal{W}^{V} \supset \mathcal{L}(\mathcal{V},\mathcal{W})$

$$u = \sum_{i=1}^{n} a_i u_i, \quad v = \sum_{j=1}^{m} b_j v_j.$$

Thus, if $w_1, \ldots, w_n, z_1, \ldots, z_m \in \mathcal{W}$ are arbitrary vectors in $\mathcal{W}$, for which we've specified via $(\star)$,

$$\mathsf{T}(u) = \sum_{i=1}^{n} a_i \mathsf{T}(u_i) = \sum_{i=1}^{n} a_i w_i, \qquad \mathsf{T}(v) = \sum_{j=1}^{m} b_j \mathsf{T}(v_j) = \sum_{j=1}^{m} b_j z_j,$$

then for $a \in \mathbb{F}$, we have

$$\begin{aligned}\mathsf{T}(au+v) &= \mathsf{T}\left[a\left(\sum_{i=1}^{n} a_i u_i\right) + \left(\sum_{j=1}^{m} b_j v_j\right)\right] = \mathsf{T}\left(\sum_{i=1}^{n} a a_i u_i + \sum_{j=1}^{m} b_j v_j\right)\\ &= \sum_{i=1}^{n} \mathsf{T}(a a_i u_i) + \sum_{j=1}^{m} \mathsf{T}(b_j v_j) = \sum_{i=1}^{n} a a_i \mathsf{T}(u_i) + \sum_{j=1}^{m} b_j \mathsf{T}(v_j)\\ &= a\sum_{i=1}^{n} a_i \mathsf{T}(u_i) + \sum_{j=1}^{m} b_j \mathsf{T}(v_j) = a\mathsf{T}(u) + \mathsf{T}(v).\end{aligned}$$

Hence T is linear. Furthermore, since the expression of u and v as a linear combination of vectors in β is unique, we know T is well-defined. Now, to show that T is unique, suppose that we have $\mathsf{T}' \in \mathcal{L}(\mathcal{V},\mathcal{W})$, such that $\mathsf{T}' v_i = w_i = \mathsf{T} v_i$ for all $i \in I$. Then, for any $x \in \mathcal{V}$, we have $x = \sum_{i=1}^{n} c_i v_i$ for unique $c_i \in \mathbb{F}$, for which

$$\begin{aligned}\mathsf{T}'(x) &= \mathsf{T}'\left(\sum_{i=1}^{n} c_i v_i\right) = \sum_{i=1}^{n} \mathsf{T}' c_i v_i = \sum_{i=1}^{n} c_i \mathsf{T}' v_i\\ &= \sum_{i=1}^{n} c_i \mathsf{T} v_i = \sum_{i=1}^{n} \mathsf{T} c_i v_i = \mathsf{T}\left(\sum_{i=1}^{m} c_i v_i\right) = \mathsf{T}(x).\end{aligned}$$

Hence, $\mathsf{T}' = \mathsf{T}$, and so T is unique. ∎

3.4 Isomorphisms

Theorem 3.4. *If* $\mathcal{V}$ *and* $\mathcal{W}$ *are vector spaces over the same field* $\mathbb{F}$ *and* β *is a basis for* $\mathcal{V}$, *then*

---

4. Transformation Matrices

Remark 4.1. $\mathcal{L}(\mathcal{V},\mathcal{W})$ is actually a vector space. I leave it to you to determine that.

Definition 4.1. Let $\beta := \{\mathbf{e}_1, e_2, \ldots, e_n\}$ be a basis for a finite dimensional vector space $\mathcal{V}$. By definition, each $x \in \mathcal{V}$ has as unique linear decomposition of the form

$$x = \sum_{i=1}^{n} \alpha_i e_i.$$

The ***coordinates of*** x ***with respect to the basis*** β are $\alpha_1, \alpha_2, \ldots, \alpha_n$. The ***coordinate vector of*** x ***with respect to basis*** β is

$$[x]_\beta = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

The **coordinate map**, also known as the **standard representation of** $\mathcal{V}$ **with respect to** β is the function $\phi_\beta : \mathcal{V} \to \mathbb{F}^n$, given by $\phi_\beta(x) = [x]_\beta$.

Theorem 4.1. *If $\mathcal{V}$ is a finite dimensional vector space over a field $\mathbb{F}$ with ordered basis $\beta = (v_i)_{i=1}^n$, then the coordinate map$\phi_\beta : \mathcal{V} \to \mathbb{F}^n$ is an isomorphism, that is, $\phi_\beta \in GL(\mathcal{V}, \mathbb{F}^n)$.*

Proof. First, we show that ϕ_β is linear, that is, $\phi_\beta \in \mathcal{L}(\mathcal{V}, \mathcal{W})$. Indeed, for $x, y \in \mathcal{V}$ there exist unique scalars $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{F}$ such that

$$x = \sum_{i=1}^{n} a_i v_i \qquad y = \sum_{i=1}^{n} b_i v_i.$$

Now, by definition of the coordinate map $\phi_\beta : \mathcal{V} \to \mathbb{F}^n$, it follows that $\phi_\beta(x) = [x]_\beta$, thus

$$[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad [y]_\beta = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Now, for $c \in \mathbb{F}$, we have

$$cx + y = c\sum_{i=1}^{n} a_i v_i + \sum_{i=1}^{n} b_i v_i = \sum_{i=1}^{n} (ca_i + b_i) v_i.$$

Thus,

$$\begin{aligned} \phi_\beta(cx + y) = [cx + y]_\beta &= \begin{pmatrix} ca_1 + b_1 \\ ca_2 + b_2 \\ \vdots \\ ca_n + b_n \end{pmatrix} = \begin{pmatrix} ca_1 \\ ca_2 \\ \vdots \\ ca_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \\ &= c\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = c[x]_\beta + [y]_\beta = c\phi_\beta(x) + \phi_\beta(y). \end{aligned}$$

Hence, ϕ_β is linear. Now, to show ϕ_β is an isomorphism, we first observe that $\dim(\mathcal{V}) = \dim(\mathbb{F}^n)$. Thus, by (Theorem what?), it suffices to show that either $\ker(\phi_\beta) = \{0_\mathcal{V}\}$(injective) or that $\operatorname{range}(\phi_\beta) = \mathbb{F}^n$ (surjective). Thus, if we proceed to show that the kernel of ϕ_β is $\{0_\mathcal{V}\}$, then we can automatically establish that ϕ_β is surjective. Indeed, let $v \in \mathcal{V}$, be a vector such that $\phi_\beta(v) = 0_{\mathbb{F}^n}$. Thus, $[v]_\beta = 0_{\mathbb{F}^n}$ implies that for the unique expression $v = \sum_{i=1}^n a_i v_i$, we must have

$$[v]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \Leftrightarrow a_i = 0, \quad \text{for all } i.$$

Thus, $v = 0_{\mathcal{V}}$, and so $\ker(\phi_\beta) = \{0_{\mathcal{V}}\}$, implying that ϕ_β is injective. By (Theorem???), it follows that ϕ_β is surjective as well, and thus an isomorphism from $\mathcal{V}$ to $\mathbb{F}^n$. Consequently, we say $\mathcal{V}$ is isomorphic to $\mathbb{F}^n$, or $\mathcal{V} \cong \mathbb{F}^n$. ■

Finding the Transformation Matrix associated with T

Now we will show that a linear operator acting on one vector space into another can be uniquely associated with a matrix provided that these vector spaces are each assigned a proper basis. It is important to note that when a linear operator is applied to some $x \in \mathcal{V}$, that it is also acting on the the basis vectors assigned to $\mathcal{V}$. To see why, let us assume that $\mathcal{V}$ is a finite dimensional vector space, and we've assigned to it a suitable basis, namely $\beta := \{e_1, e_2, \ldots, e_n\}$. Also, let $\mathsf{T} : \mathcal{V} \to \mathcal{W}$ be a linear operator from $\mathcal{V}$ to the finite dimensional vector space $\mathcal{W}$. If $x \in \mathcal{V}$ is some arbitrary vector in $\mathcal{V}$, then it can be expressed as a unique linear combination of the basis vectors in β, that is,

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n,$$

where $\alpha_k \in \mathcal{F}$ for $1 \leq k \leq n$. Then, if we apply the linear operator T to $x \in \mathcal{V}$, we have

$$\begin{aligned} \mathsf{T}(x) &= \mathsf{T}(\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n) \\ &= \alpha_1 \mathsf{T}(e_1) + \alpha_2 \mathsf{T}(e_2) + \cdots + \alpha_n \mathsf{T}(e_n). \end{aligned} \tag{4.1}$$

Thus, we see that the operator T is indeed acting on the basis vectors of β. Furthermore, we see that the values of $\mathsf{T}(e_k)$ determine the value of T on some arbitrary vector x in$\mathcal{V}$ for $1 \leq k \leq n$. Now, by definition, we see that the operator T assigns the basis vectors e_k to some $\mathsf{T}(e_k) \in \mathcal{W}$ for $1 \leq k \leq n$. Letting $\gamma := \{w_1, w_2, \ldots, w_m\}$ be a basis for $\mathcal{W}$, we can then express each $\mathsf{T}(e_k)$ as unique linear combinations of these basis vectors for $\mathcal{W}$. Thus,

$$\begin{cases} \mathsf{T}(e_1) = a_{11} w_1 + a_{21} w_2 + \cdots + a_{m1} w_m \\ \mathsf{T}(e_2) = a_{12} w_1 + a_{22} w_2 + \cdots + a_{m2} w_m \\ \vdots \\ \mathsf{T}(e_n) = a_{1n} w_1 + a_{2n} w_2 + \cdots + a_{mn} w_m. \end{cases} \tag{4.2}$$

You may be wondering why we chose to use a_{ij} as coefficients in (1), however, bear in mind that our endgame is to associate T with a matrix $A \in M_{m \times n}(\mathcal{F})$. Therefore, we will later notice that these coefficients play some role as entries of the transformation matrix associated with T.

Essentially, what we have above is

$$\mathsf{T}(e_k) = \sum_{j=1}^m a_{ij} w_i$$

for all $1 \leq k \leq n$, where $a_{ij} \in \mathcal{F}$. The image of x under T, that is, $\mathsf{T}(x) \in \mathcal{W}$, has with respect to the basis γ, certain coordinates, $c_1, c_2, \ldots, c_m$, and it can be expressed as

$$\mathsf{T}(x) = c_1 w_1 + c_2 w_2 + \cdots + c_m w_m \tag{4.3}$$

which is a unique linear combination of the basis vectors in γ . Substituting each $\mathsf{T}(e_k)$ in (1.1) with the expressions obtained in (1.2) , we have

$$\begin{aligned}\mathsf{T}(x) &= \alpha_1(a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m) + \alpha_2(a_{12}w_1 + a_{22}w_2 + \cdots + a_{m2}w_m) + \cdots \\ &+ \alpha_n(a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m) \\ &= \sum_{j=1}^{n}\alpha_j\left(\sum_{i=1}^{m}a_{ij}w_i\right)\end{aligned}$$

whereby grouping the coefficients of each w_j,we have

$$\begin{aligned}&= \sum_{i=1}^{m}\left(\sum_{j=1}^{n}\alpha_j a_{ij}\right)w_i \\ &= (\alpha_1 a_{11} + \alpha_2 a_{12} + \cdots + \alpha_n a_{1n})w_1 + (\alpha_1 a_{21} + \alpha_2 a_{22} + \cdots + \alpha_n a_{2n})w_2 + \cdots \qquad (4.4) \\ &+ (\alpha_1 a_{m1} + \alpha_2 a_{m2} + \cdots + \alpha_n a_{mn})w_m.\end{aligned}$$

Now, since the linear combination in (1.3) is unique, we must have

$$c_i = \sum_{j=1}^{n}\alpha_j a_{ij}$$

for $1 \le i \le m$. Thus,

$$\begin{cases} c_1 = \alpha_1 a_{11} + \alpha_2 a_{12} + \cdots + \alpha_n a_{1n} \\ c_2 = \alpha_1 a_{21} + \alpha_2 a_{22} + \cdots + \alpha_n a_{2n} \\ \vdots \\ c_m = \alpha_1 a_{m1} + \alpha_2 a_{m2} + \cdots + \alpha_n a_{mn}. \end{cases} \tag{4.5}$$

Thus, we have expressions for the coordinates c_i in terms of α_j, which is an important observation since all of (1.5) describes the action a linear operator T has on the bases β and γ, assigned to $\mathcal{V}$ and $\mathcal{W}$ respectively. That being said, we associate this "action" with a matrix comprised of the coefficients a_{ij} (where $1 \le i \le m$, $1 \le j \le n$). This is essentially the transformation matrix associated with T w.r.t bases β and γ, denoted by

$$[\mathsf{T}]_\beta^\gamma = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}. \tag{4.6}$$

Definition 4.2. Let $\mathcal{V}$ and $\mathcal{W}$ be finite dimensional vector spaces with $\beta := \{e_1, e_2, \ldots, e_n\}$ and $\gamma := \{w_1, w_2, \ldots, w_m\}$ as their respective bases. The matrix $[\mathsf{T}]_\beta^\gamma$ in (1.6) is the ***transformation matrix of the linear operator*** $\mathsf{T} : \mathcal{V} \to \mathcal{W}$ ***with respect to bases*** β ***and*** γ.

Remark 4.2. Observe that the kth column of $[\mathsf{T}]_\beta^\gamma$ is simply the coordinate vector of $\mathsf{T}(e_k)$ with respect to γ . Thus,

$$[\mathsf{T}]_\beta^\gamma = \left([\mathsf{T}(e_1)]_\gamma \;\middle|\; [\mathsf{T}(e_2)]_\gamma \;\middle|\; \cdots \;\middle|\; [\mathsf{T}(e_n)]_\gamma\right) \tag{4.7}$$

☜ Simply apply **Definition 2.1** to (2.2) to verify this result.

4.1 Matrix Multiplication

5. Multilinear Forms

Supplementary Material on 2-forms

Definition 5.1. A ***bilinear form*** (2-form) $\varphi : \mathcal{V} \times \mathcal{V} \to F$ is a function $\varphi(x, y)$ such that for every fixed $\tilde{x} \in \mathcal{V}$, we have $\varphi(\tilde{x}, y)$ as a linear function of $y \in \mathcal{V}$, and for every fixed $\tilde{y} \in \mathcal{V}$, we have$\varphi(x, \tilde{y})$ as a linear function of $x \in \mathcal{V}$. Thus, it is linear in each of its arguments, and thereby satisfies both of the following:

$$\varphi(cx_1 + x_2, y) = c\varphi(x_1, y) + \varphi(x_2, y)$$

and

$$\varphi(x, cy_1 + y_2) = c\varphi(x, y_1) + \varphi(x, y_2)$$

for $x, y, x_1, x_2 y_1, y_2 \in \mathcal{V}$ and $c \in F$.

Definition 5.2. A bilinear form $\varphi : \mathcal{V} \times \mathcal{V} \to F$ on a vector space $\mathcal{V}$ is:

1. ***symmetric*** if $\varphi(x, y) = \varphi(y, x)$
2. ***antisymmetric*** if $\varphi(x, y) = -\varphi(y, x)$.
3. ***alternating*** if $\varphi(x, x) = 0$

for all $x, y \in \mathcal{V}$.

Proposition 5.1. *Every alternating bilinear form* $\delta : \mathcal{V} \times \mathcal{V} \to F$ *is antisymmetric, that is,* $\delta(x, y) = -\delta(y, x)$.

Proof. Suppose we have an alternating bilinear form $\delta : \mathcal{V} \times \mathcal{V} \to F$, then for $x, y \in \mathcal{V}$

$$\begin{aligned}\delta(x + y, x + y) &= \delta(x, x + y) + \delta(y, x + y) \\ &= \delta(x, x) + \delta(x, y) + \delta(y, x) + \delta(y, y) \\ &= \delta(x, y) + \delta(y, x)\end{aligned}$$

Now, since $\delta(x + y, x + y) = 0$ by virtue of being an alternating form, then the last equality above equals 0 if and only if $\delta(x, y) = -\delta(y, x)$. Therefore, we conclude that δ is indeed antisymmetric. ■

Example 5.1. Let $\mathcal{V} = \mathbb{R}^3$. The function $\delta_v : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ defined by $\delta_v(x, y) = v \bullet (x \times y)$ is an alternating bilinear form. Indeed, we first show that δ_v is bilinear and then proceed by demonstrating that $\delta_v(x, x) = 0$ for any $x \in \mathbb{R}^3$. Let input vectors below belong to $\mathcal{V}$, then

$$\begin{aligned}\delta_v(cx_1 + x_2, y) &= v \bullet ((cx_1 + x_2) \times y) \\ &= v \bullet ((cx_1 \times y) + (x_2 \times y)) \\ &= v \bullet (cx_1 \times y) + v \bullet (x_2 \times y) \\ &= cv \bullet (x_1 \times y) + v \bullet (x_2 \times y) = c\delta_v(x_1, y) + \delta_v(x_2, y)\end{aligned}$$

Similarly, we show that

$$\delta_v(x, cy_1 + y_2) = v \bullet (x \times (cy + y_2)) = c\delta_v(x, y_1) + \delta_v(x, y_2).$$

☜ The notation • represents the dot product (inner product), and × signifies the cross product (vector product). The dot product is defined as

$$x \bullet y = \sum_{i=1}^{n} x_i y_i$$

where $x, y \in \mathcal{V}$ and x_i,y_i are respective coordinates of these vectors. The cross product is defined as:

Thus we see that δ_v is bilinear. Now to see that it is alternating, observe that

$$\delta_v(x, x) = v \bullet (x \times x) = v \bullet 0 = 0.$$

Thus, we see that δ_v is an alternating bilinear form. ◇

Miscellany For Now

- Every $m \times n$ matrix A of rank r can be factored into

  $$A = U \Sigma V^T$$

  where the columns of U (m by m) are eigenvectors of AA^T, and the columns of V (n by n) are eigenvectors of $A^T A$. The r singular values on the diagonal of Σ (m by n) are the square roots of the nonzero eigenvalues of both AA^T and $A^T A$.

- If $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$ are the non-zero singular values of an $m \times n$ matrix A, then for each $k < r$, the distance from A to the closest matrix of rank k, that is A_k, is

  $$\sigma_{k+1} =$$

- The eigenvalues of a matrix A are the values of a vector x such that

  $$Ax = \lambda x.$$

  Such a vector x is known as an eigenvector and λ is the eigenvalue. Eigenvectors are important in many branches of science because they help reveal underlying symmetries induced by an operation (in partricular linear operators or general linear operators). A formal definition for a symmetry is as follows: we say that an object is **symmetric**, with respect to a mathematical operator, if it remains invariant under that operator's action.

- What do eigenvectors do? As quoted here: "the eigenvalues of a linear mapping is a measure of the distortion induced by the transformation and the eigenvectors tell you about how the distortion is oriented." Also, consider the eigenvector corresponding to the maximum eigenvalue of a matrix, then the action of the matrix is maximum. No other vector when acted by this matrix will get stretched as much as this eigenvector. In summary, eigenvectors are the directions that remain invariant under a linear transformation (general linear operator). The corresponding eigenvalues themselves serve as scaling factors that determine by how much these directions "axes" are stretched. Thus the relative change in length of each invariant direction (eigenvector) is determined by its corresponding eigenvalue. The **spectral radius**

  $$\rho(\mathsf{T}) = \max_{\lambda \in \sigma(\mathsf{T})} |\lambda|$$

  is the maximum eigenvalue.

- Two objects X and Y are isomorphic if there exists a correspondence between them that preserves (in both directions) all the structure currently of interest. So, if two objects are isomorphic, they differ only in their labeling, and are essentially two different representations of the same object. They can be used interchangeably provided that we are willing to "relabel" everything else that they interact with. So the "essence" of the objects is the part of them that does not depend on the particular choice of representation.

Consider the vector space $\mathcal{M}_{m\times n}(\mathbb{C})$, that is the vector space of m by n matrices with complex entries which has dimension $m \cdot n$. The magnitudes of matrices in this space can be obtained by employing any vector norm on this space. For example, consider the matrix

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

then by sifting through all entries row-wise into a four-component vector, we may then then induce the euclidean $\mathbb{R}^4$ norm, thereby obtaining

$$\|A\| = \left[1^2 + 2^2 + 3^2 + (-4)^2\right]^{1/2} \approx 5.477.$$

This type of matrix norm if formally addressed as the Frobenius norm, defined below:

Definition 5.3. Let $A \in \mathcal{M}_{m\times n}(\mathbb{C})$, then the ***Frobenius Norm*** is defined by:

$$\|A\|_F^2 = \sum_{i,j} \left|A_{ij}\right|^2 = \sum_i \|A_{i\bullet}\|_2^2 = \sum_j \|A_{\bullet j}\|_2^2 = \operatorname{trace}(A^\star A).$$

The dot in $A_{i\bullet}$ indicates an aggregation over the index j, while the dot in $A_{\bullet j}$ indicates an agregration over the index i.

- For every matrix $A \in \mathcal{M}_{m\times n}(\mathbb{C})$, there exists a permutation matrix P such that

  $$PA = LU$$

  where U is upper triangular (obtained by forward elimination) and L is lower triangular with 1s on the diagonal. If $m = n$, then $P = I$, thus $A = LU$. A permutation matrix P is a matrix where each row or column contains only one nonzero entry that is 1 while the remaining entries are zero. Essentially the rows and columns of P are those of I in some particular order. A permutation matrix is handy to have in advance because it encodes row exchanges. For example, consider the matrix

  $$A = \begin{pmatrix} 0 & 0 & a \\ 0 & b & 0 \\ c & 0 & 0 \end{pmatrix}, \text{ find a permutation matrix } P \text{ such that } PA = \begin{pmatrix} c & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & a \end{pmatrix}.$$

  Now, notice row 2 in A can be left alone, thus let $P_{13} = \left(\begin{smallmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{smallmatrix}\right) = P$, which encodes the exchange of row 1 with row 3. Consider this other example: let

  $$A = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ d & e & f \end{pmatrix}, \text{ we want } P \text{ such that } PA = \begin{pmatrix} d & e & f \\ 0 & a & b \\ 0 & 0 & c \end{pmatrix}.$$

  Suppose we want to exchange row 1 with row 3 first, then let $P_{13} = \left(\begin{smallmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{smallmatrix}\right)$. After that, let use exchange row 2 with row 3, so let $P_{23} = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{smallmatrix}\right)$. Now, let

  $$P_{23}P_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = P,$$

  takining note that when P acts on A the exchange between row 1 and 3 takes place first, followed by the exhange of row 2 with row 3 as intended, thus

  $$PA = P_{23}P_{13}A = \begin{pmatrix} d & e & f \\ 0 & a & b \\ 0 & 0 & c \end{pmatrix}.$$

- Characteristics of a matrix A in echelon form U and its reduced echelon form R:
  - The pivots are the first nonzero entries in their rows. If the matrix is in reduced echelon form, the pivots are 1s.
  - Below each pivot is a column of zeros, obtained by elimination. If the matrix is in reduced echelon form, then the entries above (with the exception of the first pivot) and below each pivot are zero.
  - Each pivot lies to right of the pivot in the row above.
- Among the variables in a system, those corresponding to the columns containing leading 1s (in reduced echelon form) are called the ***pivot variables***, while those corresponding to other columns (if any) are ***free variables***. For example, consider:

$$Rx = \begin{pmatrix} 1 & 3 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

  Then, the variables x_1, x_2, x_3, x_4 fall into two groups exclusively. Hence we see that the variables x_1, x_3 corresponding to the columns with pivots (or leading 1s if R is in reduced echelon form) are pivot variables. Meanwhile, x_2, x_4 correspond to columns without pivots and are therefore free variables. Pivot variables are sometimes referred to as free variables.

---

6. Change of Coordinates and Bases

Let $\mathcal{V}$ be a finite dimensional vector space over a field with two ordered bases β and γ. We know from previous results that the coordinate maps $\phi_\beta : \mathcal{V} \to \mathbb{F}^n$ and $\phi_\gamma : \mathcal{V} \to \mathbb{F}^n$ are isomorphisms. Futhermore, these coordinate maps may be used to define a change of coordinates transformation that enables us to translate the coordinate representation of vectors in $\mathcal{V}$ with respect to basis β to that of a coordinate representation with respect to another basis, namely γ. Such a transformation will be denoted by $\phi_{\beta,\gamma} : \mathbb{F}^n \to \mathbb{F}^n$, which changes β coordinates into γ coordinates. It is defined by

$$\phi_{\beta,\gamma} := \phi_\gamma \circ \phi_\beta^{-1}$$

such that

$$\phi_{\beta,\gamma}\left([v]_\beta\right) = [v]_\gamma .$$

Definition 6.1. Let $\mathcal{V}$ be a finite dimensional vector space over a field $\mathbb{F}$ and $\beta = (v_i)_{=1}^n$ and $\gamma = (w_i)_{i=1}^n$ be two ordered bases for $\mathcal{V}$. Since, ϕ_β and ϕ_γ establish isomorphisms from $\mathcal{V}$ to $\mathbb{F}^n$, the following diagram commutes,

$$\begin{tikzcd} \mathbb{F}^n \arrow[rr, "\phi_\gamma \circ \phi_\beta^{-1}"] & & \mathbb{F}^n \\ & \mathcal{V} \arrow[ul, "\phi_\beta"] \arrow[ur, "\phi_\gamma"'] & \end{tikzcd} ,$$

that is, $\phi_\gamma \circ \phi_\beta^{-1} := \phi_\gamma \circ \phi_\beta^{-1} \in \mathcal{L}(\mathbb{F}^n)$, which is the ***change of basis operator*** changing β-coordinates into γ-coordinates, an automorphism on $\mathbb{F}^n$. Its matrix representation, called ***change of coordinate matrix*** is given by:

$$Q_{\gamma\leftarrow\beta} = \left[\phi_{\beta,\gamma}\right]_\xi \in \mathcal{M}_{n\times n}(\mathbb{F})$$

where ξ is standard ordered basis for $\mathbb{F}^n$.

Theorem 6.1. *The change of coordinate matrix $Q_{\gamma\leftarrow\beta}$ satisfies the following conditions:*

1. $Q_{\gamma\leftarrow\beta} = \left[\phi_{\beta,\gamma}\right]_\xi = \left[\phi_\gamma\right]_\beta^\xi = \Big([v_1]_\gamma \quad | \quad [v_2]_\gamma \quad | \quad \cdots \quad | \quad [v_n]_\gamma \Big)$
2. *For any $v \in \mathcal{V}$, $[v]_\gamma = Q_{\gamma\leftarrow\beta}[v]_\beta$.*
3. *$Q_{\gamma\leftarrow\beta}$ is invertible, and $Q_{\gamma\leftarrow\beta}^{-1} = Q_{\beta\leftarrow\gamma} = \left[\phi_\beta\right]_\gamma^\xi$.*

Proof. (1) By Theorem and Theorem, we have:

$$Q_{\gamma\leftarrow\beta} = \left[\phi_{\beta,\gamma}\right]_\xi = \left[\phi_\gamma \circ \phi_\beta^{-1}\right]_\xi = \left[\phi_\gamma\right]_\beta^\xi \left[\phi_\beta^{-1}\right]_\xi^\beta = \left[\phi_\gamma\right]_\gamma^\xi \left(\left[\phi_\beta\right]_\beta^\xi\right)^{-1} = \left[\phi_\gamma\right]_\beta^\xi I_n^{-1} = \left[\phi_\gamma\right]_\beta^\xi I_n = \left[\phi_\gamma\right]_\beta^\xi.$$

(2) Now, let $v \in \mathcal{V}$ be arbitrary, it follows from Theorem ???, that

$$\phi_\gamma(v) = (\phi_\gamma \circ \mathsf{I})(v) = \left(\phi_\gamma \circ \left(\phi_\beta^{-1} \circ \phi_\beta\right)\right)(v) = \left(\left(\phi_\gamma \circ \phi_\beta^{-1}\right) \circ \phi_\beta\right)(v),$$

Thus,

$$[v]_\gamma = \left[\phi_\gamma(v)\right]_\xi = \left[\left(\left(\phi_\gamma \circ \phi_\beta^{-1}\right) \circ \phi_\beta\right)(v)\right]_\xi = \left[\phi_\gamma \circ \phi_\beta^{-1}\right]_\xi \left[\phi_\beta(v)\right]_\xi = \left[\phi_{\beta,\gamma}\right]_\xi [v]_\beta = Q_{\gamma\leftarrow\beta}[v]_\beta.$$

(3) Since ϕ_β and ϕ_γ are in $\mathrm{GL}(\mathcal{V}, \mathbb{F}^n)$, it follows that $\phi_{\beta,\gamma} := \phi_\gamma \circ \phi_\beta^{-1}$ is an isomorphism. Furthermore, $\phi_{\beta,\gamma}^{-1} \in \mathcal{L}(\mathbb{F}^n)$ is an isomorphism, because the diagram in Definition ??? commutes, we must have

$$\phi_{\beta,\gamma}^{-1} = \left(\phi_\gamma \circ \phi_\beta^{-1}\right)^{-1} = \phi_\beta \circ \phi_\gamma^{-1} = \phi_{\beta,\gamma},$$

thus by (1), we have

$$Q_{\gamma\leftarrow\beta}^{-1}\left(\left[\phi_{\beta,\gamma}\right]_\xi\right)^{-1} = \left[\phi_{\beta,\gamma}^{-1}\right]_\xi = \left[\phi_{\gamma,\beta}\right]_\xi = \left[\phi_\beta\right]_\gamma^\xi = Q_{\beta\leftarrow\gamma},$$

as desired. Alternatively, by Theorem ???,

$$Q_{\gamma\leftarrow\beta}^{-1} = \left(\left[\phi_{\beta,\gamma}\right]_\xi\right)^{-1} = \left[\phi_{\beta,\gamma}^{-1}\right]_\xi = \left[\phi_{\gamma,\beta}\right]_\xi = \left[\phi_\beta\right]_\gamma^\xi = Q_{\beta\leftarrow\gamma}.$$

■

Definition 6.2. Let $\mathcal{V}$ be an vector space with ordered bases $\beta = (u_i)_{i=1}^n$ and $\gamma = (w_i)_{i=1}^n$, and let $\mathsf{I} = \mathsf{I}_\mathcal{V}$ be the indentity transformation on $\mathcal{V}$, then the ***change of basis matrix*** $Q = [\mathsf{I}]_\beta^\gamma$, is the unique matrix such that

$$Q[v]_\beta = [\mathsf{I}_\mathcal{V}]_\beta^\gamma [v]_\beta = [v]_\gamma$$

for every $v \in \mathcal{V}$. Sometimes, we'll write Q as $Q_{\gamma\leftarrow\beta}$, to indicate that this matrix changes β-coordinates in to γ-coordinatees.

Remark 6.1. If Q changes β-coordinates into γ-coordinates, then Q^{-1} changes γ-coordinates into $\beta-$coordinates. We shall prove this later.

Lemma 6.1. *The matrix $Q = [\mathsf{I}_{\mathcal{V}}]_{\beta}^{\gamma}$ is given bywhich is the matrix whose j-th column is $\left[v_j\right]_{\gamma}$. We see here that $\left[v_j\right]_{\gamma}$ is the coordinate vector of v_j with respect to the basis γ.*

Theorem 6.2. *Let $\mathcal{V}$ be a finite dimensional vector space, then:*

1. *For any basis β of $\mathcal{V}$, $Q_{\beta\leftarrow\beta} = I$, the identity matrix.*
2. *For any two bases β and γ of $\mathcal{V}$, $Q_{\gamma\leftarrow\beta}$ is invertible and $\left(Q_{\gamma\leftarrow\beta}\right)^{-1} = Q_{\beta\leftarrow\gamma}$.*
3. *For any three bases β, γ, φ, of $\mathcal{V}$, $Q_{\varphi\leftarrow\beta} = Q_{\varphi\leftarrow\gamma}Q_{\gamma\leftarrow\beta}$.*

Remark 6.2. The matrix of the identity operator $\mathsf{I}_{\mathcal{V}}$ need not be the identity matrix. Here's why: suppose that $\mathcal{V}$ is a n-dimensional vector space and β_1, β_2 are two distinct ordered bases for $\mathcal{V}$, then, $[\mathsf{I}_{\mathcal{V}}]_{\beta_1}^{\beta_2} = Q_{\beta_2\leftarrow\beta_1}$. Indeed, let $\beta_1 = \{v_i\}_{i=1}^{n}$ and $\beta_2 = \{w_i\}_{i=1}^{n}$ be ordered bases for $\mathcal{V}$, then

$$[\mathsf{I}_{\mathcal{V}}] = ()$$

---

7. Rank

- The rank of a matrix is the dimension of its column space or row space.

Let $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ be an $m \times n$ matrix with entries over a field $\mathbb{F}$ and $\mathsf{L}_A : \mathbb{F}^n \to \mathbb{F}^m$ be a general linear operator defined by $\mathsf{L}_A(x) = Ax$ for each $x \in \mathbb{F}^n$. The following properties apply:

Table 1. Useful Properties

Property of L_A	**Meaning in Terms of L_A**	**Meaning in terms of A.**	**Test on A**
$b \in \text{Range}(\mathsf{L}_A)$	There exists $x \in \mathbb{F}^n$ such that $\mathsf{L}_A(x) = b$.	There exists $x \in \mathbb{F}^n$ such that $Ax = b$.	Augmented column of $[A : b]$ has no pivot.
L_A is surjective (onto)	For every $b \in \mathbb{F}^m$, there exists $x \in \mathbb{F}^n$ such that $\mathsf{L}_A(x) = b$.	$Ax = b$ has a solution for every $b \in \mathbb{F}^m$	Every row of A has a pivot position.
L_A is injective	Given $x_1, x_2 \in \mathbb{F}^n$, $\mathsf{L}_A(x_1) = \mathsf{L}_A(x_2)$ implies $x_1 = x_2$ or $\ker \mathsf{T} = \{0\}$	0 is the the only solution of $Ax = 0$ or $\ker A = \{0\}$.	Every column of A has a pivot positon.
b has unique pre-image under L_A	There exists a unique $x \in \mathbb{F}^n$ such that $\mathsf{L}_A(x) = b$.	$Ax = b$ has a unique solution.	For $[A : b]$, every column of A has a pivot position but augmented column has no pivot position.

Remark: the augmented column is the last column in $[U : c]$, namely c, where U is in upper triangular form. We often use R to denote a matrix that is in reduced echelon form, while U denotes a matrix in echelon (upper triangular) form.

Now, let us the consider the possible scenarios in vector spaces and determine how to address them accordingly. Let $S = \{v_1, v_2, \ldots, v_k\}$ be a linearly independent set of a vector space $\mathcal{V}$. Let $\beta = \{w_1, w_2, \ldots, w_n\}$ be an ordered basis for $\mathcal{V}$. Also, let $B = (v_1|\ldots|v_k|w_1|\ldots|w_n)$ be the matrix whose columns are the amalgamation of the ordered vectors in both S and β. Now, let $A \in \mathcal{M}_{m\times n}(\mathbb{F})$ and assume $|S| \leq \dim(\mathcal{V}) = n$, then we address the following scenarios:

Table 2. Useful Properties

Problem in Vector Space	**Solution**
Extend S to a basis of $\mathcal{V}$.	Obtain the pivot columns of B and use those as a basis.
Obtain a basis for the column space (or range) of A.	Obtain the pivot columns of A and use those as basis for the column space. It is helpful to write A in echelon form U to find the corresponding pivot columns.
Find dimension of $\mathrm{Col}(A)$ or $\mathrm{Rank}(A)$.	Determine the number of pivot columns for A. This can be easily done by expressing A in echelon form.
Obtain a basis for the null space of A or $\ker A$.	Express the general solution of $Ax = 0$ as a linear combination of vectors with the free variables as coefficients. These vectors constitute a basis for $\ker A$.
Find dimension of $\ker A$ or $\mathrm{Nullity}(A)$.	Obtain the number of non-pivot columns in A.
Find basis for the row space of A.	Obtain the non-zero rows of the echelon form of A and use those to form a basis.
Find dimension of the row space of A.	Determine the number of non-zero rows in the echelon form of A or simply find the rank of A, because row rank equals column rank.

Remark: the row space and the column space of A have the same dimension, that is, the rank of A. For a square matrix, if the columns are linearly independent, then so are the columns.

8. Miscellaneous Exercises to Consider

Exercise 8.1. Consider the real vector space $\mathbb{R}^4$. Let

$$v_1 = (1, -3, 0, 2), \quad v_2 = (-2, 1, 1, 1), \quad v_3 = (-1, -2, 1, 3)$$

Determine whether v_1, v_2, and v_3 are linearly independent. Find the dimension and a basis for the subspace span v_1, v_2, v_3.

Exercise 8.2. Let $\mathcal{V}$ be the subspace of $\mathbb{R}^4$ spanned by the quadruples:

$$v_1 = (1, 2, 3, 4), \quad v_2 = (2, 3, 4, 5), \quad v_3 = (3, 4, 5, 6), \quad v_4 = (4, 5, 6, 7).$$

Find a basis of $\mathcal{V}$ and $\dim \mathcal{V}$. **Solution:** Let $\mathcal{W} \subseteq \mathbb{R}^4$ be the subspace spanned by $S = \{v_i\}_{i=1}^4$, that is,

$$\operatorname{span} S = \left\{ \sum_{i=1}^{4} c_i v_i \,:\, c_i \in \mathbb{R} \right\} = \mathcal{W}$$

Let us represent v_i as column vectors, and those vectors in $\mathcal{W}$ as column vectors as well. Then, every vector $w \in \mathcal{W}$ can be expressed as linear combination of the column vectors in S. So, let us concatenate these vectors into a 4×4 matrix A, where

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{pmatrix}.$$

To find a basis for $\mathcal{W}$ it now suffices to find a basis for the column space of A, and find the rank of A to obtain the dimension of the column space, which in turn gives us the dimension of $\mathcal{W}$. The reason for doing this is that the columns of A span $\mathcal{W}$, and so finding a basis for the column space enables us to find a minimal spanning set $\mathcal{W}$. So, let us proceed by writing A in echelon form to find its pivot columns. The pivot columns of A will provide the basis for $\mathcal{W}$ we seek. So, by elementary row operations we obtain:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & -3 \\ 0 & -2 & -4 & -6 \\ 0 & -3 & -6 & -9 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & -3 & -6 & -9 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = U$$

The pivot columns in U are columns 1 and 2, so the corresponding pivot columns in A are its first and second columns. Thus we use these first two columns in A as a basis for the column space of A. Futhermore, these columns form a basis for $\mathcal{W}$, namely $\beta = \{v_1, v_2\}$. Consequently the dimension of $\mathcal{W}$ is $\operatorname{card}(\beta) = 2$.

Exercise 8.3. Let $\mathcal{W} := \left\{ \left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix} \right) : a, b, c \in \mathbb{R} \right\}$. Show that $\mathcal{W}$ is a subspace of $\mathcal{M}_{2\times 2}(\mathbb{R})$ over $\mathbb{R}$ and that the following matrices form a basis for $\mathcal{W}$:

$$v_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Find the coordinates of the matrix $\left(\begin{smallmatrix} 1 & -2 \\ -2 & 3 \end{smallmatrix} \right)$ under the basis. **Answer:** check your papers.

- Consider the following vector spaces: $\mathcal{P}(\mathbb{F}) := \{\text{polynomials } p(x) \text{ with coefficients in } \mathbb{F}\}$ and $\mathcal{P}_n(\mathbb{F}) = \{\text{polynomials } p(x) \text{ of degree at most } n \text{ and coefficients in } \mathbb{F}\}$, where n is a non-negative integer. The degree of the 0-polynomial is undefined, furthermore the zero polynomial, denoted 0_p is a member of $\mathcal{P}_n(\mathbb{F})$ for every n. We say that $\mathcal{P}_m(\mathbb{F})$ is a subspace of $\mathcal{P}_n(\mathbb{F})$, whenever $m \leq n$.

Exercise 8.4. Consider $\mathcal{P}_n(\mathbb{R})$

1. Show that $\mathcal{P}_n(\mathbb{R})$ is a vector space of $\mathbb{R}$ under the ordinary addition and scalar multiplication for polynomials.
2. Show that $\alpha = \{1, x, x^2, \ldots, x^{n-1}\}$ is a basis for $\mathcal{P}_n(\mathbb{R})$, and so is

   $$\beta = \left\{1, (x-a), (x-a)^2, \ldots, (x-a)^{n-1}\right\},$$

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

3. Find the coordinate of $p(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1} \in \mathcal{P}_n(\mathbb{R})$ with respect to the basis β.

4. Let $a_1, a_2, \ldots, a_n \in \mathbb{R}$ be distinct. For $i = 1, 2, \ldots, n$, let
$$f_i(x) = (x - a_1) \cdots (x - a_{i-1})(x - a_{i+1}) \cdots (x - a_n)$$
. Show that $\{f_1(x), \ldots, f_n(x)\}$is also a basis for $\mathcal{P}_n(\mathbb{R})$.

5. Show that $\mathcal{W} = \{p(x) \in \mathcal{P}_n(\mathbb{R}) : p(1) = 0\}$ is a subspace of $\mathcal{P}_n(\mathbb{R})$. Find its dimension and a basis.

Exercise 8.5. Let $\alpha = \{v_i\}_{i=1}^n$ be an ordered basis for a vector space $\mathcal{V}$, where $n \geq 2$. Show that $\beta = \left\{\sum_{i=1}^j v_i\right\}_{j=1}^n$ is also a basis for $\mathcal{V}$. Is the set
$$\gamma = \{v_1 + v_2, v_2 + v_3, \ldots, v_{n-1} + v_n, v_n + v_1\}$$
a basis for $\mathcal{V}$ too? How about the converse?

Exercise 8.6. Show that
$$\alpha = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\} \quad \beta = \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix} \right\}$$
are bases for $\mathbb{R}^3$. Find the matrix $A = [A]_\alpha^\beta$, that is, the matrix from basis α to β such that
$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 1 & 4 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix} A.$$

If a vector $x \in \mathbb{R}^3$ has coordinate $(2, 0, -1)$ under the basis α, what is the coordinate of x under β.

Exercise 8.7. Let
$$\mathcal{W} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{C}^4 : x_3 = x_1 + x_2 \text{ and } x_4 = x_1 - x_2 \right\}$$

1. Prove that $\mathcal{W}$ is a subspace of $\mathbb{C}^4$.
2. Find a basis for $\mathcal{W}$. What is the dimension of $\mathcal{W}$?
3. Prove that $\left\{k(1, 0, 1, 1)^t : k \in \mathbb{C}\right\}$ is a subspace of $\mathcal{W}$.

Exercise 8.8. Find the dimension and a basis for the solution space of the system
$$\begin{aligned} x_1 - x_2 + 5x_3 - x_4 &= 0 \\ x_1 + x_2 - 2x_3 + 3x_4 &= 0 \\ 3x_1 - x_2 + 8x_3 + x_4 &= 0 \\ x_1 + 3x_2 - 9x_3 + 7x_4 &= 0. \end{aligned}$$

Exercise 8.9. Find a basis for the solution space of the following system of $n + 1$ linear equations of $2n$ unknowns:
$$\begin{aligned} x_1 + x_2 + \cdots + x_n &= 0 \\ x_2 + x_3 + \cdots + x_{n+1} &= 0 \\ &\vdots \\ x_{n+1} + x_{n+2} + \cdots + x_{2n} &= 0. \end{aligned}$$

Exercise 8.10. Let $A \in \mathcal{M}_{n\times n}(\mathbb{F})$ and write $A = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix}$, where each $\alpha_i \in \mathbb{F}^n$:

1. Show that
$$\dim(\operatorname{span}\{\alpha_1, \ldots, \alpha_n\}) = \operatorname{rank}(A).$$

2. Let P be an $n \times n$ invertible matrix, where
$$PA = \begin{pmatrix} P\alpha_1 & P\alpha_2 & \cdots & P\alpha_n \end{pmatrix} = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix}.$$
Show that $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_r}$ are linearly independent if and only if $\beta_{i_1}, \beta_{i_2}, \ldots, \beta_{i_r}$ are linearly independent over $\mathbb{F}$.

3. Find the dimension and a basis of the space spanned by
$$\gamma_1 = \begin{pmatrix} 2 \\ 1 \\ 3 \\ 1 \end{pmatrix}, \quad \gamma_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix}, \quad \gamma_3 = \begin{pmatrix} 0 \\ 2 \\ -2 \\ 1 \end{pmatrix}, \quad \gamma_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Exercise 8.11. Let $\mathcal{W}_1$ and $\mathcal{W}_2$ be the vector spaces over $\mathbb{R}$ spanned, respectively, by
$$\gamma_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ -2 \end{pmatrix}, \quad \gamma_2 = \begin{pmatrix} 3 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \gamma_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ -1 \end{pmatrix}$$
and
$$\beta_1 = \begin{pmatrix} 2 \\ 5 \\ -6 \\ -5 \end{pmatrix}, \quad \beta_2 = \begin{pmatrix} -1 \\ 2 \\ -7 \\ 3 \end{pmatrix}.$$
Find the dimensions and bases for $\mathcal{W}_1 \cap \mathcal{W}_2$ and $\mathcal{W}_1 + \mathcal{W}_2$. Hint: use dimension theorem.
Hint: $\dim(V + W) = \dim(V) + \dim(W) - \dim(V \cap W)$.

9. Appendix

Rudimentary Set Theory

- A ***class*** is an arbitrary collection of objects. A ***set*** is a class that occurs as a member of some class. A class which is not a set is a ***proper class***.

- Suppose a non-empty set I known as the index set is given such that for each $i \in I$ an object $x_i$is associated, then the ***family*** $(x_i)_{i \in I}$ is the rule which assigns to each element $i \in I$ its object x_i.

  - Consider the following example: families with index set $I = \{2, 4, 6, 8\}$ are tuples of the form (x_2, x_4, x_6, x_8), and families with index set $I = \mathbb{N}$ are sequences of the form $(x_1, x_2, x_3, x_4, \ldots)$. Also, if $I = \{1, 2\}$, then families with this index set are simply ordered pairs of the form (x_1, x_2).

- A set A is called a ***subset*** of a set B if every element of A is an element of B, in which case we write $A \subseteq B$. If A is not a subset of B we write $A \nsubseteq B$. Now, $A \neq B$, but $A \subseteq B$, then we say that A is a ***proper subset*** of B, thereby expressing this as $A \subset B$.

Definition 9.1. Let A and B be sets, then:

1. the ***union*** $A \cup B$ of A and B consists of all elements belonging to A or B, or both. We must take note that the "or" statement in mathematics si non-exclusive. Now, if $(A_i)_{i \in I}$ is a family of indexed sets (where I is non-empty index set), then the union $\bigcup_{i \in I} A_i$ consists of all elements belonging to atleast of one of the sets A_i, that is, $x \in \bigcup_{i \in I} A_i$ if and only if there exists an index $i_0 \in I$ such that $x \in A_{i_0}$.

2. the ***intersection*** $A \cap B$ of A and B consists of all elements common to both A and B. Now, if $(A_i)_{i \in I}$ is a family of indexed sets (where I is non-empty index set), then the intersection $\bigcap_{i \in I} A_i$ consists of all elements belonging to all of the sets A_i, that is, $x \in \bigcap_{i \in I} A_i$ if and only if $x \in A_i$ for all $i \in I$.

3. The ***complement*** $A \setminus B$ of B in A consists of all elements in A not belonging to B, hence $A \setminus B$ is obtained from A by removing those elements belonging to B as well, if any.

Definition 9.2. Let X be a set, then:

1. Two subsets $A, B \subseteq X$ are disjoint if and only if $A \cap B = \emptyset$. If $(A_i)_{i \in I}$ is a family of indexed subsets of X, then the sets A_i are ***pairwise disjoint*** if and only if $A_i \cap A_j = \emptyset$ whenever $i \neq j$.

2. If $X = \bigcup_{i \in I} A_i$, where A_i are pairwise disjoint, then we say X is the ***disjoint union*** of the sets A_i, and that the sets A_i form a ***partition*** of X, in which case we write $X = \biguplus_{i \in I} A_i$.

Definition 9.3. Let A and B be sets, then the ***cartestian product***, denoted by $A \times B$ is set of all pairs of (a, b) for which $a \in A$ and $b \in B$. More generally if $A_1, A_2, \ldots, A_n$ are sets, then the cartesian product $A_1 \times A_2 \times \cdots \times A_n$ of these sets if the set of all n-tuples $(a_1, a_2, \ldots, a_n)$ where $a_i \in A_i$ for $1 \leq i \leq n$. Even more generally if $(A_i)_{i \in I}$ is any family of indexed sets, then the cartesian product $\prod_{i \in I} A_i$ of the sets A_i is the set of all tuples $(a_i)_{i \in I}$, that is, all families of elements a_i where $a_i \in A_i$ for each $i \in I$.

Definition 9.4. Let X and Y be sets, then a ***function***, ***map***, or ***mapping***, of X into Y is a rule which assigns to each $x \in X$ a unique[5] element $y = f(x) \in Y$. We write $f : X \to Y$ to say that f is a function of X into Y. X is called the ***domain*** and Y is called the ***codomain*** of X. If $f(x) = y$, we say that y is the ***value of the argument*** x under the function f. For any set $A \subseteq X$, we call $f(A) := \{f(a) : a \in A\}$ the ***image*** or ***range*** of A under f. For any set $B \subseteq Y$, we call $f^{-1}(B) := \{x \in X : f(x) \in B\}$ the ***pre-image*** of B under f. The set $\Gamma(f) := \{(x, f(x)) : x \in X\}$ is the ***graph*** of f. The set of all functions $f : X \to Y$ is denoted by Y^X. Two functions $f : X \to Y$ and $g : X \to Y$ are ***equal*** if and only if $f(x) = g(x)$ for all $x \in X$.

[5] The assignment of a unique element in the codomain Y to each element in the domain X via the function f is what makes the function f **well-defined**. More, precisely $f : X \to Y$ is well-defined if for each $x \in X$ there is a unique $y \in Y$ such that $f(x) = y$.

Definition 9.5. If $A \subseteq B$, then the mapping $\iota : A \to B$ defined by $\iota(a) = a$ for all $a \in A$ is called the ***inclusion map*** of A into B. Let X, Y be sets. If $f : X \to Y$ is a function and if $A \subseteq X$, then the mapping $f\big|_A(a) = f(a)$ for all $a \in A$ is called the ***restriction*** of f to A in$X$. Conversely, if $g : A \to Y$ is a mapping then every function $f : X \to Y$ with $g = f\big|_A$, such that $g(a) = f(a)$ for all $a \in A$, is an ***extension*** of g to X. The ***identity map*** id_X of X is the mapping $f : X \to X$ defined by $f(x) = x$ for all $x \in X$.

Definition 9.6. Let X, Y, Z be sets and $f : X \to Y$, $g : Y \to Z$ be functions. Then the ***composition*** $g \circ f : X \to Z$ is the function from X into Z which is defined by $(g \circ f)(x) := g(f(x))$.

Definition 9.7. Let $f : X \to Y$ be a function. We say that f is ***injective*** or ***one-to-one*** if for $x_1, x_2 \in X$, $f(x_1) = f(x_2)$ implies $x_1 = x_2$, that is, each element of Y occurs as the image of at most one element in X. We say that f is ***surjective*** or ***onto*** if for each element $y \in Y$ there is an element $x \in X$ such that $f(x) = y$. We say f is ***bijective*** if and only if f is both injective and surjective.

Definition 9.8. Let $(A_i)_{i \in I}$ be an indexed family of sets (I is a non-empty index set). The ***product*** of the sets $\{A_i\}_{i \in I}$, written as $\prod_{i \in I} A_i$ consists of all functions a with domain, the indexing set I, having the property that for each $i \in I$, $a(i) \in A_i$. We may express this compactly as:

$$\prod_{i \in I} A_i := \left\{ a : I \to \bigcup_{i \in I} A_i \; : \; a(i) \in A_i \; \forall i \in I \right\}.$$

☞ This is rather confusing but let's take a closer look at it. What the set builder notation here is telling us is that if $a \in \prod_{i \in I} A_i$, then $a(i) \in A_i$, so suppose we consider the cartesian product of two sets, namely, $A_1 \times A_2$, which can be expressed as

$$A_1 \times A_2 = \prod_{i \in \{1,2\}} A_i,$$

then an element of this product can be a function $a : \{1, 2\} \to A_1 \cup A_2$, where a is a function of $I = \{1, 2\}$ into the union $A_1 \cup A_2$, for which $a(1) \in A_1$ and $a(2) \in A_2$. Thus a represents a function from the set $\{1, 2\}$ to the set $A_1 \cup A_2$, and the arguments 1 and 2, take on the values A_1 and A_2, respectively under the function a.

Axiom 1 (Axiom of Choice)**.** *If we have an indexed family of non-empty sets, that is, $(A_i)_{i \in I}$, then we can choose an element from each A_i at a time. More precisely, there exists a **choice function** $f : I \to \bigcup_{i \in I} A_i$ such that $f(i) \in A_i$ for all i, again meaning that f picks one element from each of the sets A_i.*

Definition 9.9. Suppose that a diagram shows arrows between points which represent mappings between sets. Such a diagram is called a commutative diagram if all compositions of mappings leading from one point in the diagram to another coincide.

Example 9.1. Consider the following diagrams:

$$\begin{tikzcd} A \arrow[rr, "g"] \arrow[dr, "f"'] & & B \\ & C \arrow[ur, "h"'] & \end{tikzcd} \qquad \begin{tikzcd} \mathcal{V} \arrow[r, "\mathsf{T}"] \arrow[d, "\phi_\beta"'] & \mathcal{W} \arrow[d, "\phi_\gamma"] \\ \mathbb{F}^n \arrow[r, "\mathsf{L}_A"'] & \mathbb{F}^m \end{tikzcd}$$

These commute since for the first diagram, we have $f = h \circ g$ and for the second, we have $\mathsf{L}_A \phi_\beta = \phi_\gamma \mathsf{T}$.