Higher Linear Algebra
Notes — Math2601 UNSW

Hussain Nawaz
hussain.nwz000@gmail.com

2022T2

Contents

1 Groups and Fields

1.1 Groups

Definition of Groups A group G is a non-empty set with a binary operation defined on it. It must satisfy the following four properties:

1. **Closure:** For all $a, b \in G$, a composition $a * b$ is defined and in G.
2. **Associativity:** (a*b)*c = a*(b*c).
3. **Identity:** There exists an $e \in G$ such that $a * e = e * a = a$ for all $a \in G$.
4. **Inverse:** For all $a \in G$, there exists an a' such that $a * a' = a' * a = e$.

Groups Order and Pair Groups are actually pairs of objects. The first is the set of elements in the group and the second, the operation defined on the group. Therefore, groups may be written as $(G, *)$.

If G is finite, then the order of G, that is $|G|$, is the number of element in G.

Abelian Groups A group is abelian if the operation is *commutative*. That is,

$$a * b = b * a \quad \forall a, b \in G.$$

Notes on the Composition Observe that the composition is actually a function $* : G \times G \to G$. $a * b$ is simply a more convenient notation than $*(a, b)$.

Though the operation $*$ is not restricted, it is often one of addition (only for abelian groups), multiplication ($\times$, often written as juxtaposition) or, composition of functions.

Notation for Repeated Composition We may often use power notation for repeated applications of a composition. That is, $a * a * \cdots * a$ (with n compositions) may be written as a^n.

Suppose that instead we are using $+$ as the group operation, then $a + a + \cdots + a$ (added n times), may be written as na. Do note that this is <u>not</u> multiplication.

Trivial Groups The trivial group consists of exactly one element, the identity. That is, $\{e\}$. Since the empty set cannot be a group, as there is required to be at least one element in a group, the trivial group is the smallest group that exists.

Examples of Groups $(\mathbb{Z}, +)$ is an abelian group under the usual addition operation. However, $(\mathbb{Z}, \times)$ is not a group, since the inverse property cannot be satisfied. Similarly, $(, \times)$ is also not a group as 0 has no multiplicative inverse. However $(\mathbb{R}\backslash 0)$ is a group.

For an integers in the set $\mathbb{Z}_m = \{0, 1, 2, \ldots, m-1\}$ is a group under addition, modulo m.

Function Composition and Groups For any S, the set F of bijective functions $f : S \to S$ is a group under composition but, it is not necessarily abelian.

Proof Composing two bijections gives a bijection so, the operation is closed. Associativity is of composition follows as

$$(f \circ (g \circ h))(x) = f(g(h(x))) = (f \circ g) \circ h(x).$$

The identity function is $e(x) = x$ is a clear bijection. The inverse exists by definition of the bijection.

More Properties of Groups

- There is only *one* inverse of each element. That is, the inverse is unique.
- For all $a \in G$, $(a^{-1})^{-1} = a$
- For all $a, b \in G$, $(a * b)^{-1} = b^{-1} * a^{-1}$.
- let $a, b, c \in G$. Then if $a*b = a*c$, then $b = c$. We may think of this as the cancellation property.

Permutation Groups Let $\Omega_n = \{1, 2, \ldots, n\}$. As a ordered set, Ω_n has $n!$ permutations.

We may think of these permutations as being functions $f : \Omega_n \to \Omega_n$. Clearly, these are bijections.

Observe that the set S_n of all permutations forms a group under the composition of order n as, the set of all bijections on a set is a group.

We may write these permutations f as a matrix where, the i, j entry represents how what element is mapped to the j-th index, by f_i.

Small Finite Groups We may visualise these with a multiplication table where, the row element is multiplied on the left of the column element.

In a multiplication table of a finite group, each row must be a permutation of the elements of the group. Otherwise, if there was a repetition in a row then $xa = xb$ implies $a = b$ by the cancellation property. Thus, each element occurs no more than once is a row.

If $a^2 = a$ then, by cancellation property, $a = e$. So, the identity must be the only element that is fixed.

1.2 Fields

Definition of Fields A field is a set $\mathbb{F}$ with two binary operations on it, addition $(+)$ and, multiplication, $(\times)$ such that, the following hold

1. $(\mathbb{F}, +)$ is an abelian group.
2. $(\mathbb{F}^* \backslash \{0\}, \times)$ is an abelian group, where 0 is the additive identity.
3. The distributive laws $a \times (b + c) = (a \times b) + (a \times c)$ and $(a + b) \times c = a \times c + b \times c$ hold.

Fields and Notation

- Under the obvious operations, typically refer to the field as just $\mathbb{F}$.
- We use juxtaposition for multiplication under fields and, 1 as the multiplicative identity and often 0 as the additive identity.
- By our definition of fields as groups, it is equivalent to say that if $\mathbb{F}$ is a field then, it satisfies the $12 = 5 + 5. + 2$ number laws.
- The smallest possible fields only has two elements, the multiplicative and additive identity. That is, $\{0, 1\}$.
- We let $-b$ be the inverse of b under addition and may write $a + (-b)$ as $a - b$ as a shorthand. Similarly, we may write $\frac{a}{b}$ rather than ab^{-1} where b^{-1} is the multiplication inverse and $b \neq 0$.

Finite Fields The only finite fields that exists are those of the size p^k for some positive integer k and prime p (also known as, the characteristic of the field).

These may be called *Galois fields* of size p^k. That is, $GF(p^k)$. Note that $GF(p^k) \neq \mathbb{Z}_{p^k}$ unless $k = 1$.

Properties of Fields If $\mathbb{F}$ is a field and $a, b, c \in \mathbb{F}$ then,

- $a0 = 0$
- $a(-b) = -(ab)$
- $a(b - c) = ab - ac$
- If $ab = 0$ then either $a = 0$ or, $b = 0$.

1.3 Subgroups and Subfields

Defining Subgroups Let $(G, *)$ be a group and H be a non-empty subset of G. Suppose that $(H, *)$ satisfies the requirements of a group, then it is a subgroup of G.

We may write $H \leq G$ such that H inherits the group structure from G.

The Subgroup Lemma Let $(G, *)$ be a group and H a non-empty subset of G. H is a subgroup if and only if

1. $a * b \in H \quad \forall a, b \in H$
2. $a^{-1} \in H \quad \forall a \in H$

That is, H is closed under $*$ and $^{-1}$.

Associativity for the subset follows from associativity of the group structure. The identity also follows from the closure under an inverse and multiplication since $a^{-1} * a = e \in H$.

Note that every non-trivial group G will have two subgroups. They are the $\{e\}$ and G.

General Linear Groups Let n be an integer such that $n \geq 1$. The set of invertible $n \times n$ matrices over $\mathbb{F}$ is a group under the operation of matrix multiplication. This is a special case of a bijection function $f : S \to S$ with $S = \mathbb{F}^n$. This group will be non-abelian if $n > 1$. This group is names the *General Linear Group*, denoted as $GL(n, \mathbb{F})$

Special Linear Group The special linear groups are a subset of the general linear groups denoted as $SL(n, \mathbb{F})$ with the requirement that the matrices all have a determinant of $\mathbb{R}$.

Orthogonal Matrix Group The set of $n \times n$ orthogonal matrices over $\mathbb{F}$ is a subgroup of $GL(n, \mathbb{F})$.

That is, $O(n) \leq GL(n, \mathbb{R})$.

There also exists $SO(n) = O(n) \cap SL(n, \mathbb{R})$ which is the intersection of the orthogonal matrices and special matrices which is also a group.

Subfields Let $(\mathbb{F}, +, \times)$ be a field and $\mathbb{E} \subseteq \mathbb{F}$ such that $\mathbb{E}$ is also a field under the same operations.

Then, $(\mathbb{E}, +, \times)$ is a subfield of $\mathbb{F}$. Equivalently, $\mathbb{E} \leq \mathbb{F}$.

Subfield Lemma Let $\mathbb{E} \neq \{0\}$ be a non-empty subset of a field $\mathbb{F}$. Then, $\mathbb{E}$ is a subfield of $\mathbb{F}$ if and only if, for all $a, b \in \mathbb{E}$,

1. $a + b \in \mathbb{E}$,
2. $-b \in \mathbb{E}$,
3. $a \times b \in \mathbb{E}$,
4. $b^{-1} \in \mathbb{E}$ given $b \neq 0$.

The distributive laws are inherited from $\mathbb{F}$ and need no checking. The rest of the proof may follow from applications of the subgroup lemma to each operation $\mathbb{E}, +$ and $\mathbb{E}, \times$.

Cool Rational + Irrational Alpha Field Let α be any non-rational real or complex number. We may define $\mathbb{Q}(\alpha)$ to be the smallest field containing both $\mathbb{Q}$ and α.

The smallest such field is of the form $\{a + b\alpha : a, b \in \mathbb{Q}\}$.

1.4 Morphisms

Morphisms are the *nice* maps between the members.

Homomorphism Definition Let $(G, *)$ and $(H, \circ)$ be two groups. A (group) homomorphism from G to H is a map $\phi : G \to H$ that respects the two operations.

That is,

$$\phi(a * b) = \phi(a) \circ \phi(b) \quad \forall a, b \in G$$

Isomorphism An isomorphism is a bijective homomorphism $\phi : G \to H$. The groups are then isomorphic. That is, $G \cong H$.

In terms of group theory, if two groups are isomorphic then, they are effectively the same group. Isomorphism is an equivalence relation on groups.

Isomorphism Lemmas let $(G, *)$ and $(H, \circ)$ be two groups and ϕ a homomorphism between them. Then,

- ϕ maps the identity of G to the G to the identity of H.
- ϕ maps the inverses to inverses. That is, $\phi(a^{-1}) = (\phi(a))^{-1}$, for all $a \in G$.
- if ϕ is a isomorphism from $G \to H$ then, ϕ^{-1} is an isomorphism from $H \to G$.

Images and Kernel Definition Let $\phi : G \to H$ be a group homomorphism with, e' the identity of H.

The kernel of ϕ is the set

$$\ker(\phi) = \{g \in G : \phi(g) = e'\}.$$

Observer that $\ker \phi \leq G$.

The image of ϕ is the set

$$\mathrm{im}(\phi) = \{h \in H : h = \phi(g),\ \text{some } g \in G\}.$$

Note that $\mathrm{im}\phi \leq G$.

One-to-One Homomorphisms A homomorphism is one-to-one if and only if $\ker \theta = \{e\}$, where e is the identity of G.

If ϕ is one-to-one then, $\mathrm{im}(\phi)$ is isomorphic to G.

Group Homomorphisms and General Linear Group It is common to seek a homomorphism $\phi : G \to GL(n, \mathbb{F})$ for some n and field $\mathbb{F}$.

If ϕ is one to one (each element maps to a unique matrix), then the representation is faithful.

Example of Group Homomorphisms on GL Consider $C_4 = \{e, a, a^2, a^3\}$ $(a^4 = e)$. Let $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Define $\phi : C_4 GL(2, \mathbb{R})$ by $\phi(a^k) = J^k$.

Then, ϕ is a faithful representation of C_4 on $\mathbb{R}^2$.

Often, we may represent to the image subgroup of matrices as the group itself.

Cayley's Theorem We may apply the lemma that $\mathrm{im}(\phi)$ is a subgroup, to the case where G is finite and H is S_n for some n.

In this case, we get a permutation representation of the group G as a subgroup of S_n.

Cayley's theorem states that every finite group G has a representation as a subgroup of S_n for $n \leq |G|$. That is, S_n and its subgroups are all that exists in the case of finite groups.

2 Vector Spaces

Motivation for Vector Spaces Vector spaces are a natural and important generalisation of $\mathbb{R}^n$. It is natural to consider them whenever it is possible to add objects and multiply them by scalars.

It may be convenient to consider a field $\mathbb{F}$ as a vector space over one of its subfields.

Definition of Vector Spaces Let $\mathbb{F}$ be a field. Then, a vector space over a field $\mathbb{F}$ consists of an abelian group $(V, +)$ and, a function from $\mathbb{F} \times V \to V$ called scalar multiplication and written as αV where the following properties hold.

1. **Associativity over scalar multiplication:** $\alpha(\beta \mathbf{v}) = (\alpha\beta)\mathbf{v}$ for all $v \in V$, $\alpha, \beta \in \mathbb{F}$
2. **Existence of 1:** $1\mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in V$
3. **Distributivity of scalar multiplication over addition:** $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$ for all $\mathbf{u}, \mathbf{v} \in V, \alpha \in \mathbb{F}$
4. **Distributivity of addition over scalar multiplication:** $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$

Properties and Notation for Vector Spaces dsdfa

1. Note that there are actually a total of ten axioms that exist. There is the four mentioned above, closure under scalar multiplication and, five that are inherited from the abelian group.
2. Addition in V is called *vector addition* to separate it from addition in $\mathbb{F}$.
3. V cannot be empty since it is a group.
4. Bold face letters $\mathbf{x}, \mathbf{y}, \mathbf{z}$ may be used instead of x, y, z, to denote vectors. More specifically, the identify of $(V, +)$ is denoted at $\mathbf{0}$ rather than the 0 that denotes a scalar in $\mathbb{F}$.
5. All the results from chapter 1 such as uniqueness of zero, negatives cancellation, ... all apply for vector addition.

Results on Combining Vectors Addition and Scalar Let V be a vector space over a field $\mathbb{F}$. Then for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $\lambda \in \mathbb{F}$,

1. $0v = \mathbf{0}$ and $\lambda\mathbf{0} = \mathbf{0}$,
2. $(-1)\mathbf{v} = -\mathbf{v}$,
3. $\lambda\mathbf{v} = \mathbf{0}$ implies either, $\lambda = 0$ or $\mathbf{v} = 0$
4. If $\lambda\mathbf{v} = \lambda\mathbf{w}$ where $\lambda \neq 0$, then, $\mathbf{v} = \mathbf{w}$.

2.1 Standard Examples of Vector Spaces

The space $\mathbb{F}^n$ over $\mathbb{F}$ The set $\mathbb{F}^n$ contains all n-tuples of elements of $\mathbb{F}$. That is,

$$\mathbb{F}^n = \left\{ \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \right\} : \alpha_i \in \mathbb{F}$$

Let $\mathbf{x} = (\alpha_i)_{1 \leq i \leq n}$ and $\mathbf{y} = (\beta_i)_{1 \leq i \leq n}$ be elements of $\mathbb{F}$. Then, vector addition is defined as

$$\mathbf{x} + \mathbf{y} = (\alpha_i + \beta_i)_{1 \leq i \leq n}.$$

Likewise, scalar multiplication on $\mathbb{F}^n$ is defined as

$$\lambda \mathbf{x} = (\lambda \mathbf{x}_i)_{1 \leq i \leq n}.$$

Geometric Vectors Geometric vectors are ordered pairs of points in $\mathbb{R}^n$ joined by label arrows. That is, they have direction and length. These may be added by placing head to tail where, scalar multiplication refers to increasing the length of the vector by a scalar value.

These vectors however, do not form a vector space. To do so, we define two geometric vectors to be equivalent if one is a translation of the other. Then, the set of equivalence classes of geometric vectors is a vector space. That is, we do not care about the position of the geometric vector, only its magnitude and direction.

Matrices For positive integers p, q the set $M_{p,q}(\mathbb{F})$ is the set of $p \times q$ matrices with elements from F. Then, $M_{p,q}$ is a vector space over $\mathbb{F}$ where vector addition and multiplication by a scalar are defined by adding each corresponding element or, multiplying each element by a the scalar.

Polynomials The set of all polynomials with coefficients in $\mathbb{F}$ denoted by $\mathcal{P}(\mathbb{F})$ is a vector space over $\mathbb{F}$ with

$$(f + g)(x) = f(x) + g(x) \qquad \text{for all } x \in \mathbb{F}, \tag{1}$$

$$(\lambda f)(x) = \lambda f(x) \qquad \text{for all } \lambda, x \in \mathbb{F}. \tag{2}$$

We may denote $\mathcal{P}(\mathbb{F})$ to be the set of all polynomials with degree n or less. This is also a vector space over $\mathbb{F}$.

Function Spaces Let X be a non-empty set and $\mathbb{F}$ be a field. Then,

$$\mathcal{F}[X] = \{f : X \to \mathbb{F}\}$$

where $\mathcal{F}[X]$is a vector space of F representing the set of all functions. We must define

1. The zero to be the zero function $x \to 0$ for all $x \in X$
2. $(f + g)(x) = f(x) + g(x)$ for all $x \in X$

3. $(\lambda f)(x) = \lambda(f(x))$ for all $x \in X$

Note that here, we use $\mathcal{F}$ to correspond with $\mathbb{F}$. If we were however using $\mathbb{R}$ as a field then, we may instead prefer to use $\mathcal{R}$ for the set of all functions instead. Similarly, we extend this for $\mathcal{Q}$ too.

2.2 Subspaces

Defining Vector Subspaces If V is a vector space over $\mathbb{F}$ and $U \subseteq V$ then, U is a subspace of V, denoted as $U \leq V$ if, it is a vector space over $\mathbb{F}$ with the same addition and scalar multiplication as V.

Observe that every vector space has $\{0\}$ (the trivial subspace) and itself as subspaces.

Subspace Lemma To check if U is a subspace of a vector space V, is is sufficient to just check for closure under addition and scalar multiplication. These conditions may be combined such that U is a subspace of V if and only if, for all $\alpha \in \mathbb{F}$, $\mathbf{u}, \mathbf{v} \in U$, $\alpha\mathbf{u} + \mathbf{v} \in \mathbb{F}$.

The other axioms may be inherited from V and it must be ensured that $\mathbf{0} \in U$.

All subspaces of $\mathbb{R}^3$ Trivially, every subspace must have $\mathbf{0}$ as an element so, it is clear that $\{\mathbf{0}\} \leq \mathbb{R}^3$. Then, the remaining subspaces must be of the form

$$\{\lambda\mathbf{a} : \lambda \in \mathbb{R}\} \qquad \text{or} \qquad \{\lambda\mathbf{a} + \mu\mathbf{b} : \lambda, \mu \in \mathbb{R}\}. \qquad \text{or} \qquad \{\lambda\mathbf{a} + \mu\mathbf{b} + v\mathbf{c} : \lambda, \mu, v \in \mathbb{R}\}.$$

That is, any line or plane through the origin or, all of $\mathbb{R}^3$.

Subspace Examples

- In $M_{p,p}$, the set of symmetric matrices is a subspace.
- Let X be any set and $Y \subseteq X$ The set $\{f \in \mathcal{F}[X] : f(y) = 0 \forall y \in Y\}$ is a subspace of $\mathcal{F}$.
- For any interval $I \subseteq \mathbb{R}$ the set $C(I)$ of continuous functions on I is a subspace of $\mathcal{R}(I)$. Similarly, if I is open, the set of differential functions, continuously differentiable, twice differentiable, ..., are a subspace of I, each one being a subspace of the previous.

Sub-Vector space but not Subspace It is entirely possible to have a subset of a vector space be a vector space but, not a subspace. A usual example is $\mathbb{R}^+$ where multiplication is used as addition. This set is a subset of $\mathbb{R}$ and a vector space but it is not a subspace.

Also consider, $\mathbb{C}^2$ as a vector space and let and let $V = \{\mathbf{v} \in \mathbb{C}^2 : v_1 \in \mathbb{R}\}$.

Here, V is not a subspace of $\mathbb{C}^2$ over $\mathbb{C}$. This stems from the fact that over $\mathbb{C}$, complex scalars are allowed. However, we could consider $\mathbb{C}^2$ as a vector space over $\mathbb{R}$ such that we are only allowed to multiply by scalars in $\mathbb{R}$ but, still allowed complex elements in the vector.

2.3 Linear Combinations, Spans and Independence

Linear Combinations Let V be a vector space over $\mathbb{F}$. A (finite) linear combination of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ in V is any vector that for scalars a_k, may be expressed as

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n;$$

Span If S is a subset of V then, the space of S is

$$\text{span}(S) = \{\text{all finite linear combinatons of vectors in } S\}.$$

If $\text{span}(S) = V$ then, S is a spanning set of V. That is, S spans V. It may also be natural to say that V generates S.

Spans as Subspaces If S is a non-empty subset of a vector space V then, $\text{span}(S)$ is a subspace of V.

- The zero element trivially exists and, the span is non-empty as S is non-empty.
- For $\mathbf{v}, \mathbf{w} \in S$, $\lambda \mathbf{v} + \mathbf{w}$ is a linear combination of two linear combinations of S which makes it a linear combination of element in V and thus, is in $\text{span}(S)$.

Linear Independence A subset S of a vector space V is linearly independent if for all vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in S$, (where $n \geq 1$), with $\alpha_i in \mathbb{F}$,

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n = 0,$$

implies that $a_i = 0$ for all $i = 1 \dots n$. That is, the solution to the above equation is unique.

Linear Dependence A set $S = \{\mathbf{v}_1, \dots \mathbf{v}_n\}$ is linearly dependent in V then, there exists an $i, 2 \leq i \leq n$ such that

$$\mathbf{v}_i = \sum_{j=1}^{i-1} \beta_j \mathbf{v}_j.$$

Properties and Theorems for Linear Dependence

- Any subset of a linearly independent set is linearly independent,
- If $\mathbf{v} \in \text{span}(S)$ then $S \cap \{\mathbf{v}\}$ is linearly dependent,
- If S is linearly independent and $S \cup \{\mathbf{v}\}$ is linearly dependent then, $v \in \text{span}(S)$,
- If $S_1 \subseteq S_2$ then $\text{span}(S_1) \subseteq span(S_2)$,
- If $S_1 \subseteq \text{span}(S_2)$ then $\text{span}(S_1) \subseteq span(S_2)$,
- $\text{span}(S \cup \{\mathbf{v}\}) = span(S)$ if and only if $\mathbf{v} \in \text{span}(S)$.
- If S is linearly dependent then, there exists $\mathbf{v} \in S$ such that $S \backslash \{\mathbf{v}\} = \text{span} S$.
- In $\mathbb{F}^n$, if $P \in GL(p, \mathbb{F})$ is also an invertible matrix and $\{\mathbf{v}_i\}$ is a linearly independent set then, $\{P\mathbf{v}_i\}$ is also linearly independent. This may be proved by contradiction.

2.4 Bases

Let $S \subseteq V$. The set S is a *basis* for V over $\mathbb{F}$ if and only if $V = \text{span}(S)$ and S is linearly independent.

That is, all vectors of V can be generated by a linear combination of vectors from S.

Standard Basis - Matrices We define the standard basis as

$$E_{ij} = (e_{hl}) = \begin{cases} 1 & \text{if } i = h \text{ and } j = l, \\ 0 & \text{otherwise.} \end{cases}$$

The set $\mathcal{B} = \{E_{ij} : 1 \leq i \leq p, 1 \leq j \leq q\}$, is the standard basis of $M_{p,q}(\mathbb{F})$ as a vector space over $\mathbb{F}$.

Function Spaces The space $\mathcal{F}(X)$ has no obvious basis unless X is finite.

Let $X = \{a_i, \ldots, a_n\}$ and for each $i = 1, \ldots, n$ define $f_i : X \to \mathbb{F}$ by

$$f_i(a_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

The set $\mathcal{B} = \{f_1, f_2, \ldots, f_n\}$ is a basis for $\mathcal{F}(X)$. We may call δ_{ij} the *Kronecker delta symbol.*

Fields The set $\{1, i\}$ is a basis for $\mathbb{C}$ as as vector space over $\mathbb{F}$. Similarly, $\mathbb{Q}(\sqrt(2))$ as a vector space over $\mathbb{Q}$ has a basis $\{1, \sqrt(2)\}$.

Zorn's Lemma All vector spaces have bases. However, it is not true that all vector spaces have a finite bases.

Standard set theory allows us to choose whether they have bases or not, as an axiom independent of the others. Assuming they do, is equivalent to the axiom of choice; or equivalently, Zorn's lemma.

2.5 Dimension

Definition Dimension Suppose that V is a vector space with a finite spanning set. Then, it has a finite basis and, all bases contain the same number of elements. The size of the basis is the dimension of V, denoted as $\dim(V)$. If we need to denote the field that the vector space exists over, we may use $\dim_{\mathbb{F}}(V)$.

If S is a finite spanning set for V then, S contains a finite basis for V.

The Exchange Lemma Suppose that S is a finite spanning set for V and that T is a (finite) linearly independent subset of V with $|T| \leq |S|$. Then, there exists a spanning set S' of V such that

$$T \subseteq S' \quad \text{and} \quad |S'| = |S|.$$

Corollary for Exchange Lemma If S is a finite spanning set for vector space V and T is a linearly independent subset of V then, T is finite and $|T| \leq |S|$.

That is, independent sets are no larger than spanning sets - if the spanning sets are finite.

Extending Linear Independence to Basis Let V be a vector space over $\mathbb{F}$ with a finite spanning set and let T be a linearly independent subset of V. Then, there is a basis B of V which contains T. That is, all linearly independent sets can be extended to a basis by adding elements - given that a finite spanning set exists.

Properties of Dimension Let V be a vector field of finite dimension and suppose $n = \dim(V)$.

1. The number of elements in any spanning set is *at least* n.
2. The number of elements in any independent set is no more than n.
3. If $\text{span}(S) = V$ and $|S| = n$ then S is a basis.
4. If S is a linearly independent set and $|S| = n$ then S is a basis.

Spans and Independence Let V be a finite vector space over $\mathbb{F}$. Then, $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for V if and only if all $\mathbf{x} \in V$ can be written *uniquely* as

$$\mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i \quad , \text{for } \alpha_i \in \mathbb{F}.$$

2.6 Coordinates

Suppose that V is a vector space of dimension n over $\mathbb{F}$ and $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an ordered basis for V over $\mathbb{F}$.

Then, if $\mathbf{v} \in \mathbb{F}$ then for unique α_i,

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

Then, we call $\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$ the coordinate vector of v with respect to $\mathcal{B}$ and refer to α_i as the coordinates of $\mathbf{v}$.

A useful notation is

$$\alpha = [\mathbf{v}]_{\mathcal{B}} \quad \text{if} \quad \mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i.$$

Finding Coordinates in Non-standard Basis The standard way of solving for a coordinate in any basis is to solve linear equations. Suppose that we want to find the coordinates of $\mathbf{v}$ over the basis $\mathcal{B}$ with elements $\{\mathbf{b}_1, \ldots, vb_n\}$ then, we may solve the set of linear equations for each component of $\mathbf{v}$.

2.7 Sums and Direct Sums

Sums of Subspaces The sum $S+T$ of two subspaces is vector space

$$S + T \coloneqq \{\mathbf{a} + \mathbf{b} : \mathbf{a} \in S, \mathbf{b} \in T\}.$$

Direct Sums If $S \cap T = \{\mathbf{0}\}$ then, $S+T$ is a direct sum and denoted as $S \oplus T$.

Implications of Sums and Spans If $X = Y = \mathbb{R}$ then obviously $X \oplus Y = \mathbb{R}$. Also, $\operatorname{span}\{\mathbf{v}\} \oplus \operatorname{span}\{\mathbf{w}\} = \operatorname{span}\{\mathbf{v}, \mathbf{w}\}$.

Uniqueness of Direct Sums The sum of subspace S, T is direct if and only if any vector $\mathbf{x} \in S + T$ can be written uniquely as $\mathbf{x} = \mathbf{v} + \mathbf{t}$ where $\mathbf{v} \in S, \mathbf{t} \in T$.

Dimensions of Sums Suppose that S and T are of finite dimension. Then,

$$\dim(S) + \dim(T) = \dim(S+T) + \dim(S \cap T).$$

Observe that for direct sums, $\dim(S \cap T) = 0$.

Complementary Space If V is a finite dimensional vector space and $X \leq V$ then, there is a subspace Y such that $X \oplus Y = V$.

We call Y a complementary space, which is not unique because we can pick Y to be any basis of $V - X$.

2.8 External Direct Sums

External Direct Sums Let X, Y be two vector spaces over the same field $\mathbb{F}$. Then, the cartesian product $X \times Y$ can be made into a vector space over $\mathbb{F}$ with the obvious definitions

$$(\mathbf{x}_1, \mathbf{x}_2) + (\mathbf{y}_1, \mathbf{y}_2) = (\mathbf{x}_1 + \mathbf{y}_1, \mathbf{x}_2 + \mathbf{y}_2)$$

and

$$\lambda(\mathbf{x}_1, \mathbf{x}_2) = (\lambda \mathbf{x}_1, \lambda \mathbf{x}_2).$$

This cartesian product is the <u>external</u> direct sum of X, Y. That is, $X \oplus Y$. This is opposed to the *internal* direct sum covered earlier.

External Direct Sums and Dimension Obviously, for X, Y of finite dimension,

$$\dim(X \oplus Y) = \dim(X) + \dim(Y).$$

If X, Y are subset of some vector space V then the external direct sum can still be identified, even if $X \cup Y$ is non-trivial.

The internal and external sums of subspaces can be identified if and only if, the subspaces have trivial intersection.

3 Linear Transformations

3.1 Linear Transformations

Linear Transformations The morphisms (nice maps) respect both operations on a vector space. Suppose that V, W are vector spaces over a field $\mathbb{F}$. Then, a function $T : V \to W$ is a linear transformation only if

$$T(\mathbf{v} + \mathbf{u}) = T(\mathbf{v}) + T(\mathbf{u}), \text{ and } T(\lambda\mathbf{v}) = \lambda T(\mathbf{v}),$$

for all $\mathbf{v}, \mathbf{u} \in V$ and $\lambda \in \mathbb{R}$.

Linearity Test Lemma A function $T : V \to W$ between vector spaces is linear if and only if

$$T(\lambda\mathbf{u} + \mathbf{v}) = \lambda T(\mathbf{u}) + T(\mathbf{v}).$$

Linear Transformations are Vector Spaces Let V, W be vector spaces over $\mathbb{F}$. Then, the set $L(V, W)$ of linear transformations from V to W is a vector space under the operations

$$(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v}), \text{ and } (\lambda S)(\mathbf{v}) = \lambda S(\mathbf{v}).$$

Composition Of Linear Transformations are Vector Spaces Let $T : V \to W$ and $S : W \to X$ be linear maps between vector spaces. Then the following composition is also linear $S \circ T : V \to X$.

Linearity of Inverse Let $T : V \to W$ be an invertible linear map between vector spaces over $\mathbb{F}$. Then, $T^{-1} : V \to W$ is linear.

Invertible Linear are Groups The invertible linear maps $L(V, V)$ form a group under composition. Note that composition of maps is always associative so and the inverse exists by definition of $L(V, V)$, only closure and the identity need to be proved.

Closure exists since composition of linear transformations are vector spaces. The identity map is linear and clearly invertible and so, also exists in the group.

Taking Coordinates is Linear Let V be a finite-dimensional vector space over $\mathbb{F}$ with a basis $\mathcal{B} = \{\mathbf{v}_1, \ldots \mathbf{v}_n\}$. Define $S : V \to \mathbb{F}^n$ as $S(\mathbf{x}) = [\mathbf{x}]_b$; that is, S reprsents a change of coordinated to the basis $\mathcal{B}$. Then, S is linear.

3.2 Kernel and Image

Definiton: Kernel and Image Let $T : V \to W$ be a linear transformation. The *kernel* is the set such that

$$\ker T = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\}.$$

If $U \leq V$ then, the *image* of U is the set

$$T(U) = \{T(\mathbf{u}) : \mathbf{u} \in U\}.$$

Also, the *image* of T (or range) is defined as the image of all V so $\text{im}(T) = T(V)$.

Kernel and Image of Linear Transformations Let $T : V \to W$ be a linear transformation between vector spaces over $\mathbb{F}$ where $U \leq V$. Then,

1. $\ker T$ is a subspace of V
2. $T(U)$ is a subspace of W and thus, $\text{im}(T) \leq W$.
3. If U is finite-dimensional, so is $T(U)$ and thus if V is finite dimensional, so is $\text{im}(T)$.

Rank and Nullity Let T be a linear transformation. The nullity is the dimension of the kernel of T. The rank is the dimension of the image of T.

Rank-Nullity Theorem If V is a finite dimensional vector space over $\mathbb{F}$ and $T : V \to W$ is linear then,

$$\text{rank}(T) + \text{nullity}(T) = \dim(V).$$

Rank and Nullity - One to One A linear map T is one-to-one if and only if $\text{nullity}(T) = 0$.

Bijetions, Injections and Surjections Let V, W be vector spaces over $\mathbb{F}$ and suppose $\dim(V) = \dim(W)$ which are both finite. Let $T : V \to W$ be linear. The following are equivalent

1. **Bijective:** T is invertible,
2. **Injective:** T is one-to-one,
3. **Surjective:** T is onto. That is, $\text{rank}(T) = \dim(V)$.

Isomorphic Maps An invertible linear map $T : V \to W$ is an isomorphism of vector spaces V and W.

Isomorphism is an equivalence relation on vector spaces.

1. **Reflexive:** The identity map $\mathrm{id} : V \to V$ is an isomorphism from V to itself.
2. **Symmetric:** If $T : V \to W$ is an isomorphism then, T^{-1} is an isomorphism from $W \to V$.
3. **Transitive:** If $T : V \to W$ is an isomorphism and $S : W \to X$ is an isomorphism then $S \circ T : V \to X$ is an isomorphism.

Easy Test for Isomorphism Finite dimension spaces V, W are isomorphism over $\mathbb{F}$ if and only if,

$$\dim(V) = \dim(W).$$

Isomorphism: Coordinates and Matrices Taking coordinates in some chosen basis is an isomorphism. All isomorphisms to $\mathbb{F}^p$ can be described in this way.

Isomorphism of Row and Column Matrices Note that all the vector spaces $\mathbb{F}^p, M_{p,1}(\mathbb{F}), M_{1,p}(\mathbb{F})$ are all isomorphic.

3.3 Spaces Associated with Matrices

Kernel, Image Rank - Matrix as Map Let A be a $p \times q$ matrix over a field $\mathbb{F}$. Define a map $T : \mathbb{F}^q \to \mathbb{F}^p$ as $T(\mathbf{x}) = A\mathbf{x}$.

Then, the kernel, rank, image and nullity of A are by definition, the same as those defined under the map T.

Image Suppose A has columns $c_1, \ldots, c_q \in \mathbb{F}$. Then,

$$\begin{aligned} \mathrm{im}(A) &= \{A\mathbf{x} : x \in \mathbb{F}^q\} \\ &= \{x_1, \mathbf{c}_1 + \ldots, x_q\mathbf{c}_q : x_i \in \mathbb{F}\} && = \mathrm{span}(\{\mathbf{c}_1, \ldots, \mathbf{c}_{nq}\}). \end{aligned}$$

That is, $\mathrm{im}(A)$ is the space spanned by the columsn of A, also knows as the column space of A. This may be denoted as $\mathrm{col}(A) \leq \mathbb{F}$.

The rank of A is the dimension of the column space of A.

Rank-Nullity Theorem: Matrices For $A \in M_{p,q}$, let q be the number of columns of A. Then,

$$\mathrm{rank}(A) + \mathrm{nullity}(A) = q.$$

Row Space The row space of A, $\mathrm{row}(A)$ is the space spanned by the rows of A. It follows that

$$\mathrm{row}(A) = \mathrm{col}(A^T) = \mathrm{im}(A^T)..$$

Rank of Matrix The rank of the matrix is identical to the column space of the matrix. The row-rank (dimension of the row space) is equivalent to the rank.

3.4 Matrix of a Linear Map

Matrices of Linear Maps Let V, W be two finite dimensional vector spaces over $\mathbb{F}$. Suppose that $\dim(V) = q$ and V and basis $\mathcal{B}$. Also suppose that W has basis $\mathbb{C}$ and $\dim(W) = p$.

If $T : V \to W$ is a linear map then, there exists a unique $A \in M_{p,q}\mathbb{F}$ with

$$[T(\mathbf{v})]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}}.$$

Conversely, for any $A \in M_{p,q}(\mathbb{F})$, the equation above defines a unique linear map from $V \to W$.

Notation We call the matrix A above, the matrix of T with respect to $\mathcal{B}$ and $\mathcal{C}$ A useful notation for this the matrix $[T]^{\mathcal{B}}_{\mathcal{C}}$. Then, the equation from the theorem takes the following form:

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

You may think of the $\mathcal{B}$ as cancelling out. This is not very rigourous.

Composition of Linear Maps as Matrices Let $T : V \to W$ and $S : W \to X$ be linear maps betweeen vector spaces. Suppose V, W, X have bases $\mathcal{A}, \mathcal{B}, \mathcal{C}$ respectively. Then, the matrix $S \circ T : V \to X$ is the product of teh matrices of T and S with the appropriate bases as follows:

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

Inverting Matrices as Transformations If $T : V \to W$ is linear and invertible then, the matrix of T^{-1} is the inverse of the matrix of T. Alternatively

$$[T^{-1}]^{\mathcal{C}}_{\mathcal{B}} = \left([T]^{\mathcal{C}}_{\mathcal{B}}\right)^{-1}.$$

Change of Basis Matrix Suppose that a vector space has two bases $\mathcal{B}$ and $\mathcal{C}$. Then the matrix $[id]^{\mathcal{B}}_{\mathcal{C}}$ is called the change of basis matrix (from $\mathbb{B} \to \mathbb{C}$).

This may be used to change coordinates as follows:

$$[\mathbf{v}]_{\mathcal{C}} = [\mathrm{id}]^{\mathcal{B}}_{\mathcal{C}}[\mathbf{v}]_{\mathcal{B}}.$$

It may often be easier to find the basis change in the other direction and then take the inverse.

Finding a Transformation Matrix To find the matrix $T : V \to W$ with respect to bases $\mathcal{B}$ as the domain and $\mathcal{C}$ in the codomain, T is found through the following three mappings.

For any $\mathbf{v} \in V$:

1. Calculate $\mathrm{id}_V(\mathbf{v}) = \mathbf{v}$ where id_V is the identity map on V.
2. Calculate $T(\mathrm{id}_V(\mathbf{v})) = T(\mathbf{v})$
3. Calculate $\mathrm{id}_W\,(T(\mathrm{id}_V(\mathbf{v}))) = T(\mathbf{v})$ where id_W is the identity map on W.

TODO: commutative diagram

Rank and Nullity of Matrices Let $T : V \to W$ be a linear map between finite dimensional vector spaces over $\mathbb{F}$.Let A be the matrix for this transformation with respect to any two bases in V and W.

Then,

$$\text{nullity}(A) = \text{nullity}(T) \text{ and } \text{rank}(A) = \text{rank}(T).$$

Invariant Subspaces Let V be a vector space over $\mathbb{F}$ and $T : V \to V$ be a linear map. If $X \leq V$ such that $T(X) \leq X$ then, X is an invariant subspace of T.

Recall that we know there is always a complementary subspace Y such that $V = X \oplus Y$. Note, Y is not necessarily invariant.

Invertibility of Linear Map and its Matrix Let $T : V \to W$ be a linear map between finite dimensional vector spaces. Suppose that A is the matrix of T with respect to any two bases. Then, T is invertible if and only if A is invertible.

3.5 Similarity

Matrices $A, B \in M_{p,p}(\mathbb{F})$ are similar if there exists some invertible matrix P such that

$$B = P^{-1}AP.$$

Similarity is an equivalence relationship on square matrices of the same size. That is, it is reflexive, symmetric and transitive.

Similarity means that the matrices represent the same linear map with respect to the choice of their two bases. We can choose P to be a change of basis matrix.

Similarity Invariants A property of matrices is similar invariant if it is preserved under similarity. That is, it holds for all similar matrices.

Note that the determinant is a similarity invariant. Thus, if matrices have different determinants, then they are not similar. However, an equal determinant does not imply similarity.

Examples of Similarity Invariants The following are all similarity invariants:

- Trace
- Determinants
- Rank, Nullity

3.6 Multi-linear Maps

Defining Bi-Linearity A mapping $T : U \times V \to W$ is bilinear if, it is linear with respect to changes in any of its arguments. That is,

$$T(\lambda u + u', v) = \lambda T(u, v) + T(u', v)$$
$$T(u, \lambda v + v') = \lambda T(u, v) + T(u, v').$$

Extending to Multi-linear Maps Like bi-linear maps, multi-linear maps are linear with respect to changes in any of their arguments.

Symmetry and Alternation on Multi-linear Maps A multi-linear map is symmetric if swapping any of its arguments does not change its value. However, if swapping arguments changes the sign of the map, then it is alternating.

4 Inner Product Spaces

4.1 The Real Dot-Product

Definition For any two vectors $\mathbf{v}, vu$ in $\mathbb{R}^p$, recall that

$$\mathbf{v} \cdot \mathbf{u} = \sum_{i=1}^{p} v_i u_i.$$

Positive Definite A bilinear map T is a positive definite if and only if, for all $\mathbf{v}$,

$$T(\mathbf{v}, \mathbf{v}) \geq 0 \quad \text{and} \quad T(\mathbf{v}, \mathbf{v}) = 0 \iff \mathbf{v} = \mathbf{0}.$$

Cauchy-Schwarz Inequality For any $\mathbf{a}, \mathbf{b} \in \mathbb{R}^p$,

$$-||\mathbf{a}||||\mathbf{b}|| \leq \mathbf{a} \cdot \mathbf{b} \leq ||\mathbf{a}||||\mathbf{b}||.$$

It follows that

$$-1 \leq \frac{\mathbf{a} \cdot}{||\mathbf{a}||||\mathbf{b}||} \leq 1.$$

Triangle Inequality For any $\mathbf{a}, \mathbf{b}$,

$$||\mathbf{a} + \mathbf{b}|| \leq ||\mathbf{a}|| + ||\mathbf{b}||.$$

As a vector, the way to think about this is that $\mathbf{a} + \mathbf{b}$ is the hypotenuse between $\mathbf{a}$ and $\mathbf{b}$. It follows directly from the Cauchy-Schwarz Inequality.

Cosine Angle Between Vectors For non-zero $\mathbf{a}, \mathbf{b}$, the angle θ between $\mathbf{a}$ and $\mathbf{b}$ is defined with

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{||\mathbf{a}||||\mathbf{b}||}, \quad \theta \in [0, \pi].$$

Orthogonality Two vectors are orthogonal if and only if $\mathbf{a} \cdot \mathbf{b} = 0$.

Orthonormality A set is orthogonal if all elements in it are orthogonal to each other. A set is orthonormal if it is an orthogonal set and all elements have unit length.

Orthogonality and Linear Independence If a set is orthogonal, then it is linearly independent.

4.2 The Complex Dot Product

Defining the Complex Dot Product For any $\mathbf{a}, \mathbf{b} \in \mathbb{C}^p$,

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{p} \overline{a_i} b_i = \overline{\mathbf{a}}^T B.$$

That is, take the conjugate of the left term and then apply the real dot product.

For convenience, define the following shorthand:

$$\mathbf{a}^* = \overline{\mathbf{a}}^T.$$

Properties of the Dot-Product The complex dot-product has the following properties for all $\lambda \in \mathbb{C}$ and $a, b \in \mathbb{C}^p$:

- $a \cdot (\lambda b + c) = \lambda a \cdot b + \lambda a \cdot c$,
- **Conjugate Symmetry**: $a \cdot b = \overline{b \cdot a}$
- **Conjugate Linearity**: $(\lambda a + b) \cdot c = \overline{\lambda a} \cdot b + \overline{b} \cdot c$
- $||a||$ is a positive definite.

Note that for the real-case conjugate symmetry and conjugate linearity are just symmetry and linearity. The mix of linearity in the second argument and conjugate linear in the first argument is called sesquilinearity.

Further, the triangle and Cauchy-Schwarz inequalities hold for the complex case too.

These properties exist to satisfy that the dot-product is an inner-product.

4.3 Inner Product Space

Let V be a vector space over $\mathbb{F}$ (typically let $\mathbb{F} = \mathbb{C}$ unless stated otherwise).

Then an inner product is a mapping $\langle , \rangle$ from $V \times V$ to $\mathbb{F}$ that satisfies the following properties for all $\lambda \in \mathbb{F}$

1. $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$,
2. $\langle u, \lambda v\rangle = \lambda\langle u, v\rangle$,
3. $\langle v, u\rangle = \overline{\langle u, v\rangle}$,
4. $\langle v, v\rangle$ is a real, positive definite.

Observe that this is just a sesquilinear, conjugate symmetric, real positive definite. The dot-products are examples of inner products.

Standard Inner Product for Continuous Functions The inner product over $\mathcal{C}[a, b]$ is defined as

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx.$$

4.4 Orthogonality

For an inner-space V, two non-zero vectors $u, v \in V$ are orthogonal if and only if $\langle u, v\rangle = 0$. Notate this as $u \perp v$. The usual definitions for othogonal and orthonormal sets holds.

Projections For $\mathbf{u}, \mathbf{v} \in V$, we define the projection of $\mathbf{u}$ onto $\mathbf{v}$ as

$$\text{proj}_{\mathbf{v}}(v) = \frac{\langle \mathbf{v}, \mathbf{u}\rangle}{\langle v, v\rangle}\mathbf{v}$$

Note that projection is linear.

Also note that $\mathbf{u} - \alpha\mathbf{v}$ if and only if

$$\mathbf{u} - \alpha\mathbf{v} = \mathbf{u} - \text{proj}_{\mathbf{v}}(\mathbf{u})$$

Orthogonal Sets as Spans Suppose that $V = \{v_1, \ldots, v_k\}$ is an orthogonal set. Then it is also a linearly independent set and so for any $\mathbf{v} \in V$ we can write

$$\mathbf{v} = \sum_{i=1}^{k} \alpha_i v_i.$$

for unique α_i.

For our orthogonal sets,

$$\alpha_i = \frac{\langle \mathbf{v}_i, v\rangle}{||v_i||^2}.$$

If the set is orthonormal, then we can simplify this to

$$\alpha_i = \langle \mathbf{v}_i, v \rangle.$$

Alternatively, for the orthogonal case,

$$\mathbf{v} = \sum_{i=1}^{k} \text{proj}_{\mathbf{v}_i}(\mathbf{v}).$$

4.5 Gram-Schmidt Process

Existence of Orthonormal Bases Any finite dimensional, inner-product space V has an orthonormal basis. We can construct this basis using the Gram-Schmidt process.

The Process Suppose that $S = \{v_1, \dots v_k\}$ is a basis for some V over $\mathbb{F}$. Then, we can construct an orthonormal basis $S' = \{w_1, \dots, w_k\}$ that independently spans V with the following formula:

$$w_i = v_i - \sum_{j=1}^{i-1} \text{proj}_{w_j}(v_i).$$

Expansion over $\mathbb{R}^3$ Let $\{v_1, v_2, v_3\}$ be an orthogonal set spanning $\mathbb{R}^3$. Then the corresponding orthonormal set is $\{w_1, w_2, w_3\}$ where

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1, \\
\mathbf{w}_2 &= \mathbf{v}_2 - \text{proj}_{\mathbf{w}_1}(\mathbf{v}_2), \\
\mathbf{w}_3 &= \mathbf{v}_3 - \text{proj}_{\mathbf{w}_1}(\mathbf{v}_3) - \text{proj}_{\mathbf{w}_2}(\mathbf{v}_3).
\end{aligned}$$

4.6 Orthogonal Complement

Orthogonal Complement Let X be a subspace of some vector space V. Then, the orthogonal complement $X^\perp$ is the set where every element of $X^\perp$ is orthogonal to X. That is

$$X^\perp = \mathbf{y} : \mathbf{y} \cdot \mathbf{x} = 0 \quad \forall \mathbf{x} \in X.$$

Orthogonal Complements as Direct Sum For any subspace V,

$$V = X \oplus X^\perp, \quad \text{and} \quad \left(X^\perp\right)^\perp = X.$$

Then, for any $\mathbf{v} \in V$, we can write the following for unique $\mathbf{x}, \mathbf{y}$:

$$\mathbf{v} = \mathbf{x} + \mathbf{y}, \quad \mathbf{x} \in X, \mathbf{y} \in X^\perp.$$

Annotate this as $\mathbf{x} = \text{proj}_X(\mathbf{v}), \mathbf{y} = \text{proj}_{X^\perp}(\mathbf{v})$.

Projection on Subspace From the previous result, for any subspace W with orthogonal basis, we calculate the projection onto that as follows:

$$\text{proj}_W(\mathbf{v}) = \sum_{i=1}^{k} \text{proj}_{\mathbf{w}_i}(\mathbf{v}).$$

This follows directly from linearity of projection and the result of orthogonal sets as spans.

Properties of the Projection Function Suppose that W is a subspace of V with $\mathbf{v} \in V, \mathbf{w} \in W$. Then,

- $\mathbf{v} - \text{proj}_W(\mathbf{v})$ is in $W^\perp$,
- $\text{proj}_W(\mathbf{w}) = \mathbf{w}$.
- The projection mapping is idempotent. That is $\text{proj}_W \circ \text{proj}_W = \text{proj}_W$.
- Projection does not extend lengths. That is, $||\text{proj}_W(\mathbf{v})|| \leq ||\mathbf{v}||$.
- $\text{proj}_W + \text{proj}_{W^\perp}$ is an identity mapping.

Projection Minimises Recall that projection does not extend lengths. In-fact we extend this to note that projection will actually minimise the length of the difference between the original vector and the projection. That is, $\mathbf{v} - \text{proj}_W(\mathbf{v})$ is the shortest distance between $\mathbf{v}$ and W. Consequently,

$$||\mathbf{v} - \mathbf{w}|| \geq ||\mathbf{v} - \text{proj}_W(\mathbf{v})||$$