# Linear algebra - version 1.0

10th Nov. 2011

## 1 Fields

Given a non-empty set $F$ and two functions $+_f$ and $._f$ as:
$+_f : F \times F \to F$ that is $(a, b) \mapsto a +_f b$ and $._f : F \times F \to F$ that is $(a, b) \mapsto a._f b$
$F$ is said to be a *field* iff the following hold[1]:

1. $F$ is closed w.r.t $+$, ..
   $\forall a, b \in F \ a + b \in F, \ a.b \in F.$
2. $+$, . are associative.
   $\forall a, b, c \in F \ a + (b + c) = (a + b) + c, \ a.(b.c) = (a.b).c$
3. $+$, . are commutative.
   $\forall a, b \in F \ a + b = b + a, \ a.b = b.a$
4. Two distinct identity elements $0_f$ and $1_f$ exist for $+$ and . respectively.
   $\forall a \in F, \ a + 0_f = a$, $\forall a \in F \ a.1_f = a$
5. An inverse exists for every element w.r.t $+$. Except for $0_f$ an inverse exists for every element w.r.t ..
   $\forall a \in F \ \exists - a \in F \ni a + (-a) = 0_f$ and $\forall a \in F, a \neq 0_f \ \exists a^{-1} \in F \ni a.a^{-1} = 1_f$
6. . distributes over $+$.
   $\forall a, b, c \in F \ a.(b + c) = a.b + a.c$

*Example* 1. $\mathcal{R}$ - set of reals, $\mathcal{Q}$ - set of rationals (a subfield of $\mathcal{R}$).
$0_f$, $1_f$ are the expected elements 0 and 1 respectively.
The inverse of an element is the negative of that element.

*Example* 2. $\mathcal{C} = (a, b), \ a, b \in \mathcal{R}$
$(a, b) + (c, d) = (a + c, b + d)$
$(a, b).(c, d) = (a.c - b.d, a.d + b.c)$
$0_f = (0, 0)$, $1_f = (1, 0)$
$-(a, b) = (-a, -b)$, $(a, b)^{-1} = (\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2})$ where $(a, b) \neq (0, 0)$

*Example* 3. $(a, b) \in \mathcal{Q} \times \mathcal{Q}$, let $p$ be a prime:
$(a, b) + (c, d) = (a + b, c + d)$
$(a, b).(c, d) = (a.c + b.d.p, a.d + b.c)$
$0_f = (0, 0)$, $1_f = (1, 0)$
Each element of the field can be written as: $(a + b\sqrt{p})$. We get the complex numbers if allow $p = -1$, though $-1$ is not a prime.

[1] We use $+$ for $+_f$ and . for $._f$ to avoid clutter. Later when we need to distinguish between two different '+' or '.' we will use the subscripted versions

*Example* 4. Let $n > 1$ and $F = \{0, 1, \ldots, n-1\}$. For $k \geq 0$ $k \; mod \; n \in F$. Let us denote it by $[k]$. Define $+$ and $.$ as follows:
$[k] + [h] = [k + h]$ and
$[k].[h] = [k.h]$
$+_I = 0$ and $._I = 1$.
If $n$ is a prime then $F$ is a finite field, usually written $\mathcal{Z}/(n)$ and called a Galois field. If $n$ is not a prime then $F$ is not a field.

*Exercise: Show that $F$ is not a field if $n$ is not a prime.*

For a Galois field it is possible that $n.1 = 0$, $n > 0$. For example, in $\mathcal{Z}/(2)$, $2.1 = 0$. If such an $n$ exists it is called the characteristic of the field, if not the field has characteristic 0. $\mathcal{Q}$, $\mathcal{R}$, $\mathcal{C}$ are characteristic 0 fields. If the characteristic is non-zero then it must be a prime.

$\mathcal{Z}$ is not field since it does not have a multiplicative inverse. Instead, it is true that:

$$a.b = 0 \Rightarrow a = 0 \text{ or } b = 0$$

If a set $F$ satisfies all properties of a field except having a multiplicative inverse but satisfies the property above then it is called an *integral domain*. $\mathcal{Z}$ is an integral domain.

**Lemma 5.** *Let $F$ be a field and $a, b \in F$ then the following holds:*
*1) $\exists c$, $c$ unique, satisfiying $a + c = b$*
*2) If $a \neq 0_f$ then $\exists d \in F$, $d$ unique, satisfying $a.d = b$.*

*Proof.* We prove *1). 2)* is left as an exercise.
Let $c = b - a$ then

$$\begin{aligned} a + c &= a + (b - a) \\ &= a + (b + (-a)) \\ &= a + ((-a) + b) \\ &= (a + (-a)) + b \\ &= 0_f + b \\ &= b \end{aligned}$$

It now remains to argue that $c$ is unique. Let $c' \in F$ satisfy $a + c' = b$ then

$$\begin{aligned} c' &= 0_f + c' \\ &= (-a + a) + c' \\ &= -a + (a + c') \\ &= -a + b \\ &= b - a \\ &= c \end{aligned}$$

□

*Exercise: Prove 2) above.*

Some proporties of fields are summarized in the lemma below.

**Lemma 6.** *If $a, b, c \in F$ then*

1. $0_f.a = 0_f$

2. $(-1_f).a = -a$
3. $a.(-b) = -a.b = (-a).b$
4. $-(-a) = a$
5. $(-a).(-b) = a.b$
6. $-(a+b) = -a - b$
7. $a.(b-c) = a.b - a.c$
8. $a \neq 0_f \rightarrow (a^{-1})^{-1} = a$
9. $a, b \neq 0_f \rightarrow (a.b)^{-1} = a^{-1}.b^{-1}$
10. $a + c = b + c \rightarrow a = b$
11. $c \neq 0_f, a.c = b.c \rightarrow a = b$
12. $a.b = 0_f \rightarrow a = 0_f \text{ or } b = 0_f$

*Proof.* We prove item 1 of lemma 6. The rest are left as exercises.

$0_f.a + 0_f.a = (0_f + 0_f).a = 0_f.a$
Add $-(0_f.a)$ to both sides
$0_f.a + 0_f.a - (0_f.a) = 0_f.a - (0_f.a)$
$0_f.a = 0_f$ □

# 2 Vector Spaces

**Definition 7** (Vector space)**.** A set of vectors $V$ is a vector space over a field $F$ if there exist functions $+_v : V \times V \rightarrow V$ and $._v : F \times V \rightarrow V$ satisfying:

1. Associativity of $+_v$. $\forall v_1, v_2, v_3 \in V, (v_1 +_v v_2) +_v v_3 = v_1 +_v (v_2 +_v v_3)$.
2. Commutativity of $+_v$. $\forall v_1, v_2 \in V, v_1 +_v v_2 = v_2 +_v v_1$
3. There exists an identity element $0_v$ with respect to $+_v$ such that $\forall v \in V, v + 0_v = v$.
4. $\forall v \in V$ there exists an inverse $-v$ such that $v + (-v) = 0_v$.
5. $._v$ distributes over $+_v$ $\forall v_1, v_2 \in V, a \in F,\ a._v(v_1 +_v v_2) = a._v v_1 +_v a._v v_2$.
6. $._v$ distributes over $+_f$. $\forall a_1, a_2 \in F, v \in V (a_1 +_f a_2)._v v = a_1._v v +_v a_2._v v$.
7. Associativity of scalar multiplication. $\forall a_1, a_2 \in F, v \in V, a_1._v(a_2._v v) = (a_1._f a_2)._v v$.
8. There exists an identity $1_f \in F$ with respect to scalar multiplication. $\forall v \in V,\ 1_f._v v = v$.

◁

*Example* 8. $F$ over $F$ is a vector space.

*Example* 9. Let $V$ be a vector space over $F$, $k, n > 0 \in \mathcal{Z}$, $\Omega = \{(i,j) | 1 \leq i \leq k, 1 \leq j \leq n\}$. Define $f : \Omega \rightarrow V$ then $f$ corresponds to a $k \times n$ matrix over $V$. The set of all $k \times n$ matrices over $V$, $M_{k \times n}$ is a vector space over $F$. Addition and multiplication by a scalar are defined as for matrices. The identity element is $0_{k \times n}$.

*Example* 10. $V = \mathcal{R}^{\mathcal{R}}$: set of all functions from $\mathcal{R} \to \mathcal{R}$ is a vector space over $\mathcal{R}$. The operations are defined as: $\forall f, g \in V, a \in \mathcal{R},\ (f+g)(a) = f(a) + g(a)$, $\forall a, c \in \mathcal{R}(c._v f)(a) = c +_f (f(a))$.

Similarly, $\mathcal{R}^{\mathcal{N}}$ is also a vector space.

Similar to fields the following properties hold for vector spaces and are very useful in proofs.

**Lemma 11.** *Let $V$ be a vector space over field $F$ and $v, w \in V$ then $\exists y \in V$, $y$ is unique such that $v + y = w$.*

**Lemma 12.** *Let $V$ be a vector space over field $F$. For any $v, w \in V$ and any $a \in F$:*

1. $a \cdot 0_v = 0_v$.
2. $0_f \cdot v = 0_v$.
3. $(-1_f) \cdot v = -v$.
4. $(-a) \cdot v = -(a \cdot v) = a \cdot (-v)$.
5. $-(-v) = v$.
6. $a \cdot v = (-a) \cdot (-v)$.
7. $a(v + w) = -v - w$.
8. $a \cdot (v - w) = a \cdot v - a \cdot w$.
9. $a \cdot v = 0 \to a = 0_f$ *or* $v = 0_v$

*Exercise: Prove lemmas 11, 12.*

**Definition 13** (Subspace)**.** Let $V$ be a vector space over field $F$. $\phi \neq W \subseteq V$ is a subspace of $V$ if $W$ is a vector space over $F$ with respect to $+_v, ._v$.
◁

To show that $W$ is a subspace the following lemma is very useful.

**Lemma 14.** *Let $V$ be a vector space. $W$ is a subspace of $V$ iff $W$ is closed with respect to $+_v$ and $._v$.*

*Proof.* One way is easy. By definition if $W$is a subspace it is closed with respect to $+_v$ and $._v$. To go the other way assume that $W$ is closed with respect to $+_v$ and $._v$. It is clear that the associativity, commutativity and distributivity properties will hold since $W \subseteq V$ and $V$is a vector space. It remains to show that $W$ contains the identity element and inverse elements. Consider $w \in W$, $0_f \cdot w = 0_v \to 0_v \in W$. The same identity element works for $W$. Similarly, $-1_f \cdot w = -w \in W$ and $w \in V$ so $-w$ is the additive inverse of $w$ and clearly this holds for all $w \in W$. □

We will see numerous examples of subspaces later. We just note two special subspaces.

*Example* 15. $\{0_v\}$ is a subspace - called the *trivial subspace*. Similarly, $V$ is also a subspace and is called the *improper subspace*.

**Theorem 16.** *Let $\Omega$ be an index set (e.g. $\mathcal{N}$) $V$ a vector space and $\{V_i | i \in \Omega\}$ a family of subspaces of $V$. Then $\bigcup_{i \in \Omega} V_i$ is also a subspace.*

*Proof.* The basic idea is to use closure. Let $W = \bigcup_{i \in \Omega}$ and consider $w, w' \in W$. Since $W$is an intersection $w, w' \in V_i$ for each $i \in \Omega$. Since $V_i$ is a subspace $w + w' \in V_i$ for each $i \in \Omega$ which implies that $w + w' \in W$. Similarly, if $a \in F$ and $w \in W$ $w \in V_i$ for each $i \in \Omega$ and $a \cdot w \in V_i$ for each $i \in \Omega$ implying $a \cdot w \in W$. Using the closure lemma 14 it follows that $W$ is a subspace. □

**Definition 17** (Linear combination or l.c.)**.** Let $\phi \neq D \subset V$, $V$ a vector space over field $F$. Then $v \in V$ is a linear combination of $D$ iff $\exists v_1, \ldots v_n \in D$ and $a_1, \ldots, a_n \in F$ such that $v = a_1 \cdot v_1 + \ldots + a_n \cdot v_n$. $\triangleleft$

Let $FD$ be the set of all possible linear combinations of elements in $D$. For example, if $D = \{v\}$ then $FD = \{a \cdot v | a \in F\}$. $F\phi = \{0_v\}$ by definition. The set $FD$ is called the *span* of $D$. $D$ is called the *generating set*(gs) of $FD$. $D \subset V$ is a subspace iff $D = FD$.

**Lemma 18.** *Let $V$ be a vector space over field $F$ and let $D \subset V$ then:*

1. *$FD$ is a subspace of $V$ containing $D$.*
2. *Any subspace of $V$ containing $D$ also contains $FD$.*
3. *$FD$ is the intersetion of all subspaces of $V$ containing $D$.*

*Proof.* If $D = \phi$ then by definition $FD$ is the trivial subspace. Let $D \neq \phi$. Clearly, $D \subset FD$, also $FD$ is closed under $+_v$, $._v$ and so by the closure lemma 14 $FD$ is a subspace. This proves 1.

To establish 2 let $W$ be a subspace of $V$and $D \subset W$. Then for any $v_1, \ldots, v_n \in D$ and $a_1, \ldots, a_n \in F$, $a_1 \cdot v_1 + \ldots, a_n \cdot v_n \in W$ since $W$is a subspace and therefore $FD \subset W$.

Every subspace that contains $D$ also contains $FD$. But by 1 $FD$ itself is a subspace containing $D$. Therefore, the intersection of all such subspaces will be $FD$. So, $FD$ is the smallest such subspace. $\square$

*Example* 19 (Generating set). If $F$is a field and $V = F^3$ then $D_1 = \{[1,0,0],[0,1,0],[0,0,1]\}$ is a gs for $V$. $D_2 = \{[0,1,1],[1,0,1],[1,1,0]\}$ is also a gs for $V$if $F$does not have characteristic 2 (i.e. $F \neq \mathcal{Z}/(2)$). If $F$has characteristic 2 then $[1,1,1]$ cannot be a l.c. of $D_2$ so $FD_2$ is a subspace of $V$. $D_3 = [1,0,0],[0,1,0],[1,1,0]$ is not a gs for $V$for any $F$since $[0,0,1]$ is not a l.c. of elements in $D_3$.

**Lemma 20.** *Let $V$be a vector space over field $F$and $D \subseteq D' \subseteq FD$ be subsets of $V$. Then $FD' = FD$.*

*Proof.* Since $D \subseteq D'$, $FD \subseteq FD'$. To show the converse observe that $FD$ is a subspace of $V$containing $D'$ consequently from lemma 18 $FD' \subseteq FD$ so $FD = FD'$. $\square$

Lemma 20 implies that adding l.c.s to $D$ does not change the generated subspace. And this is intuitively clear since adding l.c.s only adds linear combinations which are present in $FD$ by definition.

A vector space $V$over field $F$is *finitely generated* iff its generating set $D$ is finite.

*Example* 21. $\mathcal{R}$ is a finitely generated over $\mathcal{R}$ with $D = \{1\}$. It is not finitely generated over $\mathcal{Q}$ - set of rationals.

*Example* 22. If $n > 0$ then $F^n$ is a vector space over $F$ finitely generated by $D = \{[1,0\ldots,0],\ldots,[0,\ldots,0]\}$ containing $n$ vectors. Generally, if $V$is a vector space over field $F$finitely generated by $v_1, \ldots, v_k$ then $V^n$ is finitely generated by the $k.n$ vectors $\{[v_1,0,\ldots,0],\ldots,[v_k,0,\ldots,0],\ldots,[0,\ldots,v_1],\ldots,[0,\ldots,v_k]\}$.

*Example* 23. If $V$is a finitely generated vector space over field $F$and $k, n > 0$ then $M_{k\times n}(V)$ is finitely generated over $F$. This is the set of all $k \times n$ matrics with entries from $V$.

# 3 Linear independence, dimension, basis

**Definition 24** (Linear Dependence and Independence)**.** Let $V$be a vector space over field $F$. A set of vectors $v_1, \ldots, v_n \in V$ is *linearly dependent* if $\exists a_1, \ldots, a_n \in F$ not all $a_i = 0_f$ such that $a_1 \cdot v_1 + \ldots + a_n \cdot v_n = 0_v$.

A set of vectors $v_1, \ldots, v_n$ that is not linearly dependent is called *linearly independent*. In this case all $a_i, 1 \leq i \leq n$ are $0_f$.

A set $\phi \neq D \subset V$ is *linearly dependent* iff $\exists v_1, \ldots, v_n \in D$ and $a_1, \ldots, a_n \in F$ not all $a_i = 0_f$ and $a_1 \cdot v_1 + \ldots + a_n \cdot v_n = 0_v$. If no such elements exist then $D$ is *linearly independent*. In this case all $a_i = 0_f$. $\triangleleft$

It immediately follows from the definition that any subset of a linearly independent set is linearly independent and any set having a linearly dependent subset is itself linearly dependent.

Further, the set $\{0_v\}$ is trivially linearly dependent and so any set containing $0_v$ is linearly dependent. The empty set $\phi$ is always linearly independent. We now prove an important lemma.

**Lemma 25.** *Let $V$ be a vector space over field $F$ and let $V = F\{v_1, \ldots, v_n\}$, that is $V$ is a span of $\{v_1, \ldots, v_n\}$, and let $y_1, \ldots, y_m$ be a set of linearly independent vectors in $V$ then $m \leq n$.*

*Proof.* Since $\{v_1, \ldots, v_n\}$ span $V$ we have $0_v \neq y_1 = a_1 \cdot v_1 + \ldots, a_n \cdot v_n$ where not $a_i = 0_v$. Let $a_i \neq 0_v$ then $v_i$ can be expressed as a l.c. of $y_1 \cup (\{v_1, \ldots, v_n\} - \{v_i\})$, that is $y_1$ and the other $v_j, j \neq i$. If $m \geq n$ repeat the above step $n-1$ more times to yield $y_1, \ldots, y_n$ that span $V$. If $m > n$ then $y_{n+1} \in V$ will be a l.c. of $y_1, \ldots, y_n$ and $y_1, \ldots, y_m$ cannot be linearly independent. So, $m \leq n$. $\square$

**Definition 26** (Basis)**.** A linearly independent set of vectors $\{v_1, \ldots, v_n\}$ such that $F\{v_1, \ldots, v_n\} = V$ is called a *basis* of $V$. $\triangleleft$

Note that we are concerned largely with vector spaces with finite bases. So, generally when we say basis we mean finite basis. The size of the basis set is called the *dimension* of the vector space spanned by the basis and we shall write $dim(V)$ for dimension of $V$.

**Lemma 27.** *A vector space $V$ over field $F$ spanned by a finite set of vectors $\{v_1, \ldots, v_n\}$ has a finite basis.*

*Proof.* If the set $\{v_1, \ldots, v_n\}$ is linearly dependent then some $v_i$ whose coefficient $a_i \neq 0_f$ can be written as a linear combination of the others so the set $\{v_1, \ldots, v_n\} - \{v_i\}$ can still span $V$. If this set is still linearly dependent then repeat the process until the set is linearly independent. It still spans $V$ and is, therefore, a basis. $\square$

The *canonical basis* of an $n-$dimensional vector space is usually defined as the set of $n$ vectors $\{[1, 0, \ldots, 0], [0, 1, 0, \ldots, 0], [0, 0,$

*Example* 28. Consider the vector space $\mathcal{R}^3$ over $\mathcal{R}$. The canonical basis for this vector space is: {[1,0,0],[0,1,0],[0,0,1]}. Consider $B = \{v_1 = [1, 0, 0], v_2 = [1, 1, 0], v_3 = [1, 1, 2]$. To see that $B$ is also a basis we have to show it is linearly independent i.e.

$$a_1.v_1 + a_2.v_2 + a_3.v_3 = 0_v = [0, 0, 0]$$

This gives the following equations:

$$\begin{aligned} a + b + c &= 0 \quad (1) \\ b + c &= 0 \quad (2) \\ 2.c &= 0 \quad (3) \end{aligned}$$

It is easy to see that the only solution to the above set of linear equations is $a = 0, b = 0, c = 0$. This proves that $B$ is linearly independent and is therefore a basis.

**Theorem 29.** *If $V$ is a vector space over field $F$ then $\phi \neq D \subset V$ is a basis iff every $v \in V$ can be written as a l.c. of elements of $D$ in exactly one way.*

*Proof.* If $D$ is a basis then trivially every $v \in V$ is a l.c. of elements in $V$ since $FD = V$. It remains to argue that this l.c is unique. Consider $v \in V$ with two possible l.c. expansions $a_1.d_1, \ldots, a_n.d_n$ and $b_1.d_1, \ldots, b_n.d_n$ with $d_1, \ldots, d_n \in D$. Now $v - v = 0_v = \sum_j = 1^n (a_j - b_j).d_j$. Since $D$ is a basis $a_j - b_j = 0$ for each $j \in 1..n$ so there both expansions are identical.

For the converse let every $v \in V$ be written uniquely as a l.c. of $D$. In particular $0_v = 0.d_1, \ldots, 0.d_n$ and that is the only way it can be written so $\{d_1, \ldots, d_n\}$ is an independent set of vectors and is therefore a basis. $\square$

**Definition 30** (Dimension)**.** If $\{v_1, \ldots, v_n\}$ is a basis of vector space $V$over field $F$then $n$ is said to be the dimension of $V$. We write this as: $dim(V)$. $\triangleleft$

**Theorem 31.** *Every finitely generated vector space $V$has a unique dimension or every basis of a finite dimensional vector space has the same number of vectors.*

*Proof.* Let $D_1,\ D_2$ be bases of $V$and $m = |D_1|$, $n = |D_2|$. Now $FD_2 = V$ and $D_1$ is a l.i. set so $m \leq n$ by lemma 25. Similarly, interchanging $D_1,\ D_2$ we get $n \leq m$ implying $m = n$. $\square$

**Proposition 32.** *If $V$is a finite dimensional vector space then any linearly independent subset of $V$is contained in a basis of $V$(equivalently every linearly independent subset of $V$can be extended to a basis).*

*Proof.* Let $D = \{v_1, \ldots, v_n\} \subset V$ be a l.i. set of vectors. If $FD = V$ then clearly $D$ is a basis and the proposition is true. If $FD \subset V$ then choose $v' \in V - FD$ and let $D' = D \cup \{v'\}$. Clearly, $D'$ is l.i. since $v' \notin FD$. If $FD' \neq V$ then repeat this process. Since $V$is finite dimensional this process must terminate after finitely many steps the corresponding $D'$ is clearly a basis and $D \subset D'$. $\square$

**Theorem 33.** *If $V$is a finite dimensional vector space over field $F$and $W$is a subspace of $V$then:*
*1) $W$is finite dimensional.*
*2) $dim(W) \leq dim(V)$.*
*3) Any basis of $W$is a subset of a basis of $V$and $dim(W) = dim(V)$ only when $W = V$.*

*Proof.* Let $D$ be a basis of $V$. Since $W$is a subspace of $V$any $w \in W$ is also in $V$and therefore $W \subset FD = V$ implying $dim(W) \leq dim(V)$ which also proves that $W$is finite dimensional. So 1) and 2) hold.

If $D'$ is a basis of $W$then from proposition 32 $D'$ is a subset of a basis of $V$. It also follows from theorem 31 that $dim(W) = dim(V)$ exactly when $W = V$. $\square$

**Proposition 34.** *If $V$is a vector space over field $F$then the following are equivalent:*
*1) $D$ is a minimal generating set of $V$.*
*2) $D$ is a maximal l.i. set of $V$.*
*3) $D$ is a l.i. generating set of $V$.*

*Proof.* To show 1) $\rightarrow$ 2). Assume 1) and let $D$ be l.d. then $d_i = a_i^{-1}.(\sum_{j \neq i}^{n} a_j.d_j)$ where $a_1, \ldots, a_n \in D$. Since $D$is a minimal g.s. of $V$for any $v \in V$ we have $v = c_1.d_1, \ldots, c_n.d_n$. But now $d_i$ can be expressed in terms of the other $d_k, k \neq i$ so $v = \sum_{k \neq i}^{n} c'_k.d_k$ that is any $v$ can be expressed in terms of the set $D - \{d_i\}$ and so $D$is not minimal - a contradiction.

To show 2) $\rightarrow$ 3). Assume 2) and let $v_0 \in V - D$. Then $D \cup \{v_0\}$ is l.d. and $a_0.v_0 + a_1.d_1, \ldots, a_n.d_n = 0_v$ and not all $a_i = 0_F, i \in 0..n$. In particular $a_0 \neq 0_F$ since otherwise $D$ will be l.d. So we get $v_0 = a_0^{-1}.(\sum_{j=1}^{n} a_j.d_j)$ implies that $D$is a g.s. for $V$since $v_0$ is any vector in $V - D$.

To show 3) $\rightarrow$ 1). Assume 3) and let $D' \subset D$ be a g.s. for $V$. Then $\exists v \in D - D'$ such that $v = a_1.d_1 +, \ldots + a_n.d_n$ and $d_i \in D', i \in 1..n$. This implies $D$is l.d. - a contradiction. $\square$

We now extend the notions of l.i. to subspaces. Let $W_1, W_2$ be subspaces of vector space $V$over field $F$. Let $W = W_1 + W_2 = \{w_1 + w_2 | w_1 \in W_1, w_2 \in W_2\}$ so each $w \in W$ can be written as $w = w_1 + w_2$ where $w_1 \in W_1$ and $w_2 \in W_2$. The breakup of $w$ into $w_1$ and $w_2$ is unique only if $W_1 \cap W_2 = \{0_v\}$. To see this assume $w_1 + w_2 = w = w'_1 + w'_2$ then $w_1 - w'_1 = w'_2 - w_2 = 0_v$ since $w_1 - w'_1 \in W_1$ and $w'_2 - w_2 \in W_2$ and the only common vector in $W_1$ and $W_2$ is $0_v$. This implies $w_1 = w'_1,\ w'_2 = w_2$. If $W_1 \cap W_2 = \{0_v\}$ the set $\{W_1, W_2\}$ is said to be *independent*. This idea of independence can be extended to a set of subspaces.

**Definition 35** (Independent set)**.** Let $W_1, \ldots, W_n$ be subspaces of vector space $V$. The set $\{W_1, \ldots, W_n\}$ is said to be independent if $\sum_{j=1}^{n} w_j = 0_v,\ w_j \in W_j$ implies $w_j = 0_v,\ \forall j \in 1..n$. $\triangleleft$

$D \subset V$, $D \neq \phi$ is l.i. if the set of subspaces $\{Fd | d \in D\}$ is independent. Whenever $V = W_1 + \ldots + W_n$ and the set $\{W_1, \ldots, W_n\}$ is independent we say that $V$is a *direct sum* of the set of subspaces $\{W_1, \ldots, W_n\}$ and write $V = W_1 \oplus \ldots \oplus W_n$.

*Example* 36. Let $V = \mathcal{R}^3$ and $W_1 = \{[a, 0, 0] | a \in \mathcal{R}\}$, $W_2 = \{[0, b, 0] | b \in \mathcal{R}\}$ and $W_3 = \{[0, 0, c] | c \in \mathcal{R}\}$. $\{W_1, W_2, W_3\}$ is independent and $V = W_1 \oplus W_2 \oplus W_3$.

**Proposition 37.** *Let $W_1, \ldots, W_n$ be subspaces of vector space $V$over field $F$. Then the following are equivalent:*
*1) $\{W_1, \ldots, W_n\}$ is independent.*
*2) Each $w \in \sum_{i=1}^n W_i$ has exactly one representation $w = w_1 + \ldots + w_n,\ w_i \in W_i, i \in 1..n$.*
*3) For each $1 \leq j \leq n$, $W_j \bigcap \sum_{i \neq j} W_i = \{0_v\}$.*

*Proof.* To show 1) $\rightarrow$ 2). The argument is similar to the one earlier. Let $w \in W = \sum_{j=1}^n W_j$ be written in two different ways: $w = w_1 + \ldots + w_n$ and $w = w'_1 + \ldots + w'_n$. Then $(w_1 - w'_1) + \ldots + (w_n - w'_n) = 0_v$. From 1) $w_i - w'_i = 0_v,\ \forall i \in 1..n$ implying $w_i = w'_i,\ \forall i \in 1..n$ and $w \in W$ has a unique representation as a sum.

To show 2) $\rightarrow$ 3). Each $W_i,\ i \in 1..n$ contains $0_v$ since they are subspaces. So, $\sum_{i \neq j} W_i$ contains $0_v$ and so does $W_j \bigcap \sum_{i \neq j} W_i$ for any $j \in 1..n$. Now let the intersection contain another vector $w' \neq 0_v$. So both $W_j$ and $\sum_{i \neq j} W_i$ contain the vectors $0_v$ and $w'$. Consider the vector $w = 0_v + w',\ 0_v \in W_j,\ w' \in \sum_{i \neq j} W_i$. This vector can also be written as $w = w' + 0_v,\ w' \in W_j,\ 0_v \in \sum_{i \neq j} W_i$ giving two different representations for $w$ thereby contradicting 2).

To show 3) $\rightarrow$ 1). Assume 3) and let $w_1 + \ldots + w_n = 0_v$ such that not all $w_i$ are $0_v$. Choose $j$ such that $w_j \neq 0_v$ then $w_j = \sum_{i \neq j} -w_i$ implying $w_j \in W_j \bigcap \sum_{i \neq j} W_i$, contradicting 3). □

Direct sums are often written $\bigoplus_{i=1}^n W_i$ instead of $\sum_{i=1}^n W_i$.

**Proposition 38.** *Let $V$be a vector space over field $F$and let $\{W_1, \ldots, W_n\}$ be an independent set of subspaces of $V$. Then the following are equivalent:*
*1) $V = \bigoplus_{i=1}^n W_i$.*
*2) If $D_i$ is a basis of $W_i$ then $\bigcup_{i=1}^n D_i$ is a basis of $V$.*

*Proof.* To show 1) $\rightarrow$ 2). For $v \in V$ we can write $v = w_1 + \ldots + w_n$, $w_i \in W_i,\ i \in 1..n$. Each $w_i$ can be written as a unique l.c. of vectors in $D_i$ since it is a basis of $W_i$. This means it is a unique l.c. of vectors in $\bigcup_{i=1}^n D_i$. So, the basis of $V$is $D = \bigcup_{i=1}^n D_i$.

To show 2) $\rightarrow$ 1). Since $D$ is a basis $v \in V$ can be written as a unique l.c. of vectors from $D$. This implies as a unique combination of vectors from the $D_i$s that is as a unique sum $w = w_1 + \ldots + w_n$ where each $w_i$ is a unique l.c. of vectors from $D_i$ this implies 1). □

An immediate consequence of the above proposition is the following corollary.

**Corollary 39.** *If $V = \bigoplus_{i=1}^n W_i$ where $W_i,\ i \in 1..n$ are subspaces of vector space $V$over field $F$then: $dim(V) = dim(W_1) + \ldots + dim(W_n)$.*

We now study some obvious direct sums, namely complements.

**Definition 40.** If $W$is a subspace of vector space $V$over field $F$then $W'$ is a complementary subspace or complement of $W$in $V$iff $V = W + W'$ ◁

**Proposition 41.** *Every subspace of vector space $V$over field $F$has a complement in $V$.*

*Proof.* Let $W \subseteq V$ be a subspace of $V$. If $W = V$ then $\{0_v\}$ is the complement and vice-versa. So, let $W \subset V$ and $D$ be a basis of $W$then $D \subset D'$ where $D'$ is a basis of $V$. Consider $W' = F(D' - D)$. Clearly, $D' - D$ is non-trivial and l.i. and $W'$ is a subspace of $V$and $V = W \oplus W'$ by construction and using proposition 38. □

Unlike sets a subspace can have more than one complement. The following example illustrates this.

*Example* 42. Let $V = \mathcal{R}^2$ then each $W_i$ is a complement of the others:
1) $W_1 = \{[a, 0]|a \in \mathcal{R}\}$ 2) $W_2 = \{[0, b]|b \in \mathcal{R}\}$
3) $W_3 = \{[c, c]|c \in \mathcal{R}\}$ 4) $W_4 = \{[d, 2.d]|d \in \mathcal{R}\}$

*Example* 43. Let $V = \mathcal{M}_{n\times n}(\mathcal{Q})$ - the set of all square matrices with rational entries. Define $W_1 = \{A = [a_{ij}]|a_{ij} = a_{ji},\ i, j \in 1..n\}$, $W_2 = \{A = [a_{ij}]|a_{ij} = -a_{ji},\ i, j \in 1..n\}$. Clearly, $W_1,\ W_2$ are subspaces of $V$, namely the set of all symmetric and skew symmetric matrices respectively. Also, $W_1 \bigcap W_2 = \{0_v\}$ so $\{W_1, W_2\}$ is an independent set. For any $A \in V$, we can write $A = B + C$ where $b_{ij} = \frac{a_{ij}+a_{ji}}{2}$ and $c_{ij} = \frac{a_{ij}-a_{ji}}{2}$. Clearly, $B \in W_1$ and $C \in W_2$ and $V = W_1 \oplus W_2$.

From example 42 it is clear that a subspace can have infinitely many complements. The following proposition says exactly when that happens.

**Proposition 44.** *If $F$ is an infinite field and $V$ a vector space over $F$ and $dim(V) > 1$ then any non-trivial subspace $W$ of $V$ has infinitely many complements.*

*Proof.* The proof is left as an exercise.

Hint: $W$ has at least one complement $W'$ in $V$. Let $D$ be a basis of $W'$. Choose $w \in W,\ w \neq 0_v$ then $Fw$ is an infinite subset of $W$. Construct $Y_w = F\{u + w|u \in D\}$ for each $w \in W$. The claim is: $Y_w$ is a complement of $W$ for each $w$. Now complete the proof. □

We know that if $\{W_1, W_2\}$ is an independent set of subspaces of vector space $V$ over field $F$ such that $V = W_1 + W_2$ then $dim(V) = dim(W_1) + dim(W_2)$. This can be generalized.

**Theorem 45** (Grassman's theorem)**.** *If $W_1,\ W_2$ are two subspaces of a vector space $V$ over field $F$ then $dim(W_1+W_2) = dim(W_1) + dim(W_2) - dim(W_1 \bigcap W_2)$.*

*Proof.* Let $U = W_1 \bigcap W_2$ then $U$ is a subspace of $W_1$ and $W_2$. Let $U_1'$ be the complement of $U$ in $W_1$ and $U_2'$ the complement of $U$ in $W_2$. We have $W_1 + W_2 = U + U_1' + U_2'$. We now argue that $\{U, U_1', U_2'\}$ is an independent set giving $W_1 + W_2 = U \oplus U_1' \oplus U_2'$.

Let $u + u_1' + u_2' = 0_v$ with $u \in U,\ u_1' \in U_1',\ u_2' \in U_2'$. Then $u_1' = -u - u_2'$. The l.h.s is in $W_1$ the r.h.s is in $W_1 \bigcap W_2$ so $u_1' \in W_1 \bigcap W_2$. This only possible if $u_1' = 0_v$. A similar argument shows $u_2' = 0_v$ implying $u = 0_v$. Consequently, $\{U, U_1', U_2'\}$ is an independent set and $W_1 + W_2 = U \oplus U_1' \oplus U_2'$. From corollary 39 we have:

$$\begin{aligned} dim(W_1 + W_2) &= dim(U) + dim(U_1') + dim(U_2') \\ &= dim(W_1) + dim(U_2') + dim(U) - dim(U) \\ &= dim(W_1) + dim(W_2) - dim(U) \\ &= dim(W_1) + dim(W_2) - dim(W_1 \cap W_2) \end{aligned}$$

□

# 4 A Short Detour on Functions/Mappings

A function $f : V \to W$ is *1-1* or *injective* or *monic* or a *monomorphism* iff distinct $v \in V$ map to distinct $w \in W$. That is $f(v) = f(v') \Rightarrow v = v'$ or equivalently $f(v) \neq f(v')$ iff $v \neq v'$.

A function $f : V \to W$ is *onto* or *surjective* or *epic* or an *epimorphism* iff every $w \in W$ has some $v \in V$ associated with it. That is for each $w \in W$, $\exists v \in V$ such that $f(v) = w$.

A function $f : V \to W$ is *1-1, onto* or *bijective* or *isomorphic* or an *isomorphism* if it is both 1-1 and onto or injective and surjective or monic and epic or a monomorphism and epimorphism.

A bijective function $f$ always has an inverse $f^{-1}$ and $f^{-1} \circ f$ is the identity function on $V$. Similarly, $f \circ f^{-1}$ is an identity on $W$.

# 5 Linear Transformations

Let $V$, $W$ be vector spaces over $F$. A function $f : V \to W$ is called a *linear transformation* or *homomorphism* if it satsifies:

1. $f(v_1 + v_2) = f(v_1) + f(v_2)$, $v_1, v_2 \in V$.
2. $f(a \cdot v) = a \cdot f(v)$, $v \in V$, $a \in F$.

From 2 above we can see: $f(0_v) = f(0_f \cdot 0_v) = 0_f \cdot f(0_v)$. So, the additive identity of $V$ maps to the additive identity of $W$. Henceforth, we write l.t. for linear transformation.

*Example* 46. Simple scaling, $v \mapsto av$ is a l.t.
Similarly, $v \mapsto 0_v$ and $v \mapsto v$ the nullary and identity transformations are also l.t.s.

*Example* 47. Every matrix $A = [a_{ij}] \in \mathcal{M}_{m\times n}(F)$ is a l.t. from $F^n \to F^m$ using the standard multiplication of a matrix and a vector. $Av = w$, where $w_i = \sum_{j=1}^{n} a_{ij} \cdot v_j$ for $i \in 1 \dots m$. $v$ is a $n \times 1$ vector and $w$ is a $m \times 1$ vector.

*Example* 48. Let $F$ be a field of characteristic 0 (e.g. $\mathcal{R}$, $\mathcal{C}$) and $P_n(x)$ the set of all polynomials of degree $n$ with coefficients in $F$. Consider the functions:

$f : \sum_{i=0}^{n} a_i \cdot x_i \mapsto \sum_{i=1}^{n} (i \cdot 1_f) a_i \cdot x^{i-1}$

$g : \sum_{i=0}^{n} a_i \cdot x_i \mapsto \sum_{i=0}^{n} ((i+1) \cdot 1_f)^{-1} a_i \cdot x^{i+1}$

$f$, $g$ are l.t.s. namely differentiation and integration.

*Example* 49. Let $V$, $W$ be vector spaces over $F$ and $m$, $n > 0 \in \mathcal{N}$. Let $f_{ij} : V \to W$ be a set of l.t.s then $f : \mathcal{M}_{m\times n}(V) \to \mathcal{M}_{m\times n}(W)$ defined by $f : [v_{ij}] = [f_{ij}(v_{ij})]$ is a l.t.

*Example* 50. Let $V_1$, $V_2$ be vector spaces over $F$. The graph of any function $f : V_1 \to V_2$ is defined as $graph(f) = \{(v, f(v)) | v \in V_1\} \subset V_1 \times V_2$. $f$ is a l.t. iff $graph(f)$ is a subspace of $V_1 \times V_2$.
*Exercise*: Show this.

Let $V, W$ be vector spaces over $F$. Since l.t.s are functions from $V \to W$ they are in $W^V -$ set of all functions from $V$to $W$. Let $f$, $g$ be l.t.s from $V \to W$ and define:

$$(f + g)(v) = f(v) + g(v)$$
$$(a \cdot f)(v) = a \cdot f(v),\ a \in F$$

The following is easy to see:

$$\begin{aligned}(f + g)(v_1 + v_2) &= f(v_1 + v_2) + g(v_1 + v_2)\\ &= f(v_1) + f(v_2) + g(v_1) + g(v_2)\\ &= f(v_1) + g(v_1) + f(v_2) + g(v_2)\\ &= (f + g)(v_1) + (f + g)(v_2)\end{aligned}$$

and

$$\begin{aligned}(f + g)(a \cdot v) &= f(a \cdot v) + g(a \cdot v)\\ &= a \cdot f(v) + a \cdot g(v)\\ &= a \cdot (f + g)(v)\end{aligned}$$

Similarly, we can see that $[a \cdot f](v_1 + v_2) = [a \cdot f](v_1) + [a \cdot f](v_2)$, $v_1, v_2 \in V$, $a \in F$ and $[a \cdot f](c\dot{v}) = c \cdot [a \cdot f](v)$, $v_1, v_2 \in V$, $a, c \in F$.

So the set of all l.t.s from $V \to W \subset W^V$ is a subspace of $w^V$ over $F$. This subspace is often written as $Hom(V, W)$ - the set of all homomorphisms from $V$ to $W$.

**Proposition 51.** *Let $V$, $W$ be vector spaces over $F$. Let $D$ be a basis of $V$ and $f$ any function $f : D \to W$. Then there exists a unique l.t. $g : V \to W$ satisfying $g(d) = f(d)$, $d \in D$.*

This says that for any function there is a unique l.t. that agrees with it on the basis.

*Proof.* For any $v \in V$ we have $v = \sum_{i=1}^{n} a_i \cdot d_i$, $n = |D|$, $\{d_1, \ldots, d_n\}$ and the expansion is unique. Define $g : v \mapsto \sum_{i=1}^{n} a_i \cdot f(d_i)$. This is a l.t. and is well defined since the expansion of $v$ is unique. Clearly, $g(d_i) = f(d_i)$ for $d_i \in D$. To show it is unique assume there is a $h$ such that $h(d_i) = f(d_i)$ then

$$\begin{aligned} h(v) &= \sum_{i=1}^{n} a_i \cdot h(d_i) \\ &= \sum_{i=1}^{n} a_i . f(d_i) \\ &= \sum_{i=1}^{n} a_i . g(d_i) \\ &= g(\sum_{i=1}^{n} a_i \cdot d_i) \\ &= g(v) \end{aligned}$$

This holds for all $v \in V$ implying $h = g$. □

**Proposition 52.** *Let $V$, $W$, $X$ be vector spaces over $F$. Let $f : V \to W$, $g : W \to X$ be l.t.s. then the composition $g \circ f : V \to X$ is also a l.t.*

*Proof.* Verify that $g \circ f$ satisfies the linearity constraints. □

**Definition 53** (Inv, Ker)**.** Let $V$, $W$ be vector spaces over $F$, and $f : V \to W$ a l.t. Then:

$Inv(f, w) = \{v | f(v) = w, w \in W\}$ that is all $v$ that map $w \in W$ via $f$.

$Ker(f) = Inv(f, 0_w)$, all $v$ that map to $0_w$. ◁

**Proposition 54.** *Let $f : V \to W$ be a l.t. over $F$ for vector spaces $V$, $W$. Then*

1. *$Ker(f)$ is a subspace of $V$.*
2. *$f$ is 1-1 (injective) only if $Ker(f) = 0_v$.*

*Proof.* Left as an exercise. □

Similar to $Ker(f)$ define $Im(f) = \{f(v) | v \in V\} \subset W$. Since $0_v \mapsto 0_w$ $Im(f)$ is always non-empty. $Im(f)$ is the image of l.t. $f$ in $W$.

**Proposition 55.** *Let $V$, $W$ be vector spaces over $F$, and $f : V \to W$ a l.t. Then*

1. *$Im(f)$ is a subspace of $W$.*
2. *The function is onto (surjective) only iff $Im(f) = W$.*

*Proof.* Left as an exercise. □

While $Ker(f)$, $Im(f)$ are subspaces of $V$ and $W$ respectively. In general $Inv(f, w),\ w \neq 0_w$ is not a subspace of $V$. But we have the following:

**Proposition 56.** *Let $f : V \to W$ be an l.t., $w \in Im(f)$. If $v_0 \in V$ satisfies $f(v_0) = w$ then $Inv(f, w) = \{v + v_0 | v \in Ker(f)\}$*

*Proof.* Let $v \in Ker(f)$, $f(v_0 + v) = f(v_0) + f(v) = w + 0_w = w$, so $v_0 + v \in Inv(f, w)$. Consider $v' \in Inv(f, w)$, define $v = v' - v_0$ then $fv) = f(v') - f(v_0) = w - w = 0_w$ so $v \in Ker(f)$. □

*Example* 57. Consider the set $\mathcal{R}^{[0,1]}$ of all everywhere differentiable functions on $[0, 1]$. Define $\delta : W \to \mathcal{R}^{[0,1]}$ be a l.t which maps $f \in W$ to its derivative $f'$ in $\mathcal{R}^{[0,1]}$. Then $Ker(\delta)$ is the set of all constant functions in $\mathcal{R}^{[0,1]}$. Let $g \in Im(\delta),\ g = \delta(f)$ then $f(x) = \int_0^x g(t)dt\ +\ c$, $c$ is a constant function.

**Proposition 58.** *A l.t. $f : V \to W$ is an isomorphism between $V$, $W$ iff there exists $g : W \to V$ such that $g(f(v)) = v,\ v \in V$ and $f(g(w)) = w,\ w \in W$*

Isomorphism between $V$ and $W$will be written $V \sim W$.

*Proof.* Exercise. □

**Corollary 59.** *$\sim$ is an equivalence relation. That is:*

1. *$V \sim V$.*
2. *$V \sim W \to W \sim V$.*
3. *$V \sim W$ and $W \sim X$ implies $V \sim X$.*

Also, it is clear that for $f$ an isomorphism if $D$ is a basis of $V$ then $f(D) = D'$ is a basis for $W$.

**Proposition 60.** *If $V$, $W$ are vectors spaces over $F$ with dimension $n$ then $V \sim W$.*

*Proof.* Hint: Show there is a 1-1, onto mapping by using bases of $V$, $W$and a 1-1 mapping between the bases $D_v$ and $D_w$. □

It follows that any finite dimensional vector space with dimension $n$ is isomorphic to $F^n$. The following generalization of proposition 60 follows.

**Proposition 61.** *Let $V$, $W$ be finitely generated vectors spaces over $F$. Then,*

1. *There exists a 1-1 l.t. of $V$to $W$ iff $dim(V) \leq dim(W)$.*
2. *There exists an onto l.t. of $V$to $W$ iff $dim(V) \geq dim(W)$.*

From the above we also get the converse of proposition 60, namely if $V \sim W$ then $dim(V) = dim(W)$ where $V$, $W$ are finitely generated.

*Proof.* To show 1 let $D_v$ be a basis of $V$, $f$ the 1-1 map and $D_w = f(D_v) = \{f(d_i)|d_i \in D_v\}$. Then $FD_w \subseteq W$ and therefore $dim(V) \leq dim(W)$.

To see the converse let $dim(V) \leq dim(W)$ and let $D_v$ and $D_w$ be bases of $V$ and $W$ respectively. Map the first $n$ vectors of $D_w$ to the $n$ basis vectors in $D_v$ where $n = |D_v|$. Let $f(d_i) = w_i \in D_w$. Then $f(v) = f(a_1 \cdot d_1 + \ldots + a_n \cdot d_n) = a_1 \cdot f(d_1) + \ldots + a_n \cdot f(d_n) = a_1 \cdot w_1 + \ldots + a_n \cdot w_n = w \in W$. Now $v$ has a unique expansion in terms of the $d_i$s (basis vectors) and similarly $w$ also is unique for the same reason so $f$ is 1-1.

The proof of 2 is similar. Let $|D_v| = n$, $|D_w| = m$ and $n \geq m$ map $\{d_1, \ldots, d_m\}, d_i \in D_v$ to $w_1, \ldots w_m \in D_w$. This map is onto since every $w \in W$ is uniquely expressed in terms of vectors in $D_w$. The $v \in V$ that maps to this $w$ is $w = f(v) = a_1 \cdot f(d_1) + \ldots + a_n \cdot f(d_m)$ and $F\{d_1, \ldots, d_m\} \subseteq V$.

In the reverse direction assume $f$ is an onto l.t. from $V$ to $W$. Consider $Inv(f, W) = \cup_{w \in W} Inv(f, w)$. Clearly, $FInv(f, W) \subseteq V$ therefore $dim(W) \leq dim(V)$. □

There is an important relation between the the dimensions of $V$, $Im(f)$ and $Ker(f)$.

**Proposition 62.** *Let $f : V \to W$ be a l.t. where $V$, $W$ are finitely generated vector spaces over $F$. Then $dim(V) = dim(Im(f)) + dim(Ker(f))$.*

*Proof.* Let $V'$ be the complement of $Ker(f)$ then $dim(V) = dim(Ker(f)) \oplus dim(V')$. We claim $V' \sim Im(f)$. Then by proposition 61 it follows that $dim(V') = dim(Im(f))$.

Let $f'$ be a restriction of $f$ to $V'$. Then $f'$ is a l.t. from $f' : V' \to Im(f)$. If $v' \in Ker(f')$ then $v' = 0_v$ since $V'$ is a complement of $Ker(f)$, so $v' \in V' \cap Ker(f') = 0_v$. This implies that $f'$ is 1-1 (refer to proposition 54).

To show that $f'$ is onto let $w \in Im(f)$ then there is a $v \in V$ such that $f(v) = w$. Now, $v = v' + v''$ with $v' \in Ker(f)$ and $v'' \in V'$. This implies: $w = f(v) = f(v') + f(v'') = 0_w + f(v'') = f'(v'')$. So, $w \in Im(f')$ and $Im(f') \sim V'$, that is $f'$ is an isomorphism. □

$dim(Im(f))$ is called the *rank* of $f$ and $dim(Ker(f))$ is called the *nullity* of $f$.

**Theorem 63** (Sylvester's Theorem)**.** *Let $V$, $W$, $X$ be finitely generated vector spaces and $f : V \to W$ and $g : W \to X$ be l.t.s then:*

1. $dim(Im(gf)) \leq min\{dim(Im(f)), dim(Im(g)))\}$.
2. $dim(Ker(gf)) \leq dim(Ker(f)) + dim(Ker(g))$.
3. $dim(Im(f)) + dim(Im(g)) - dim(W) \leq dim(Im(gf))$.

*Proof.* To show 1 observe that $Im(gf) = g(Im(f)) \subseteq Im(g)$. This implies $dim(Im(gf) \leq dim(Im(g))$.

Also, $Im(f) \subseteq W$ and $Im(gf)$ is a restriction of $g$ to $Im(f)$ so $dim(Im(gf)) \leq dim(Im(f))$ and $dim(Im(gf)) \leq min(dim(Im(f)), dim(Im(g)))$.

To prove 2 let $g' : Im(g) \to X$ then $Ker(g') \subseteq Ker(g)$. Also, using proposition 62

$$\begin{aligned} dim(Ker(gf)) &= dim(V) - dim(Im(gf)) \\ &= dim(V) - dim(Im(f)) + dim(Im(f)) - dim(Im(gf)) \\ &= dim(Ker(f)) + dim(Ker(g') \\ &\leq dim(Ker(f)) + dim(Ker(g)) \end{aligned}$$

To show 3 from proposition 62 we have:

$$dim(Im(f)) = dim(V) - dim(Ker(f)) \quad (4)$$
$$dim(Im(g)) = dim(W) - dim(Ker(g)) \quad (5)$$
$$dim(Im(gf)) = dim(V) - dim(Ker(gf)) \quad (6)$$

(4) + (5) - (6) gives:

$$\begin{aligned} dim(Im(f)) + dim(Im(g)) - dim(W) &= dim(Im(gf)) - [dim(Ker(f)) + dim(Ker(g)) - dim(Ker(gf))] \\ &= dim(Im(gf)) + \alpha, \alpha > 0 \text{ due to 2} \\ &\leq dim(Im(gf)) \end{aligned}$$

□

# 6 Endomorphisms, Automorphisms

A special case of a l.t. is $f : V \to V$, that is the range vector space is the same as the domain vector space. Such functions are called *endomorphisms* and $End(V)$ is the set of all endomorphisms. Some obvious endomorphisms are: $v \mapsto 0_v$ (let us symbolize this by $0_{0_v}$), $v \mapsto v$ (identity, let us symbolize this by $I$), $v \mapsto av,\ a \in F$ (scaling, symbolized by $S_a$).

Observe that if $f,\ g$ are endomorphisms so is $f + g$ and $f.g$, usually written $fg$. So, we have two operations $+,\ .$ defined on $End(V)$. So, it is useful to ask what is the algebraic structure of $End(V)$. Note that $+$ is closed, associative, commutative, has an identity $(0_{0_v})$ and an inverse. The operation . is associative, not commutative, has an identity $(I)$ but does not always have an inverse. Also, . distributes over $+$. So, we see that $End(V)$ is actually a *ring*.

*Example* 64. Consider endomorphisms $f,\ g : F^3 \to F^3$, $F$ a field defined by: $f : [a, b, c] \mapsto [a, 0_f, 0_f]$ and $g : [a, b, c] \mapsto [b, a, c]$ then $fg : [a, b, c] \mapsto [b, 0_f, 0_f]$ while $gf : [a, b, c] \mapsto [0_f, a, 0_f]$. So, $fg \neq gf$.

Also, note that $Ker(f) \neq [0_f, 0_f, 0_f]$ implying that $f$ is not 1-1 and therefore not a bijection. Therefore, an inverse $f^{-1}$ does not exist.

Define $h : [a, b, c] \mapsto [0_f, 0_f, c]$. Then we see that $fh = hf = [0_f, 0_f, 0_f] = 0_{0_v}$ though neither $f$ nor $h$ is $0_{0_v}$. This is another property of a field that does not hold. Actually, it is not even an integral domain.

While every element of $End(V)$ does not have a multiplicative inverse some do. For example, all $S_a$ for $a \in F$ have multiplicative inverses and multiplication is commutative. Endomorphisms that have multiplicative inverses are called *automorphisms* - isomorphic endomorphisms. We represent the set of all automorphisms on vector space $V$ by $Aut(V)$.

We now look at some properties of endomorphisms.

**Proposition 65.** *Let $f$ be an endomorphism on $V$, a vector space of finite dimension over field $F$. Then the following are equivalent:*

1. *$f$ is an automorphism on $V$.*
2. *$f$ is 1-1.*
3. *$f$ is onto.*

*Proof.* Left as an exercise. □

Let $W$ be a subspace of $V$ and $f$ an endomorphism on $V$. Then $W$ is said to be *stable* under $f$ iff $\forall w \in W,\ f(w) \in W$.

**Proposition 66.** *Let $V$ be a vector space over field $F$, $W$ as subspace of $V$ and $f$ and endomorphism on $V$. Then the following are equivalent:*

1. *$W$ is stable under $f$.*
2. *if $p$ is a projection of $V$ onto $W$ then $pfp = fp$.*

*Proof.* Assume 1 holds. Use $p$ to breakup $V$as $V = W \oplus W'$ where $W'$ is the complement of $W$ in $V$. Then for any $v \in V$, $v = w + w'$, where $p(v) = w \in W$, $w' \in W'$. Now $pfp(v) = pf(w) = f(w)$ (since $f$ is stable) $= fp(v)$. Thus proving 2.

For the converse, let 2 hold. Since $p$ is a projection of $V$ to $W$ for $w \in W$, $p(w) = w$. So, $pf(w) = pfp(w) = fp(w) = f(w)$ therefore $f(w) \in W$ and $W$ is stable under $f$. □

# 7 Linear Transformations as Matrices

Let $V$ $W$ be finitely generated vector spaces over $F$ with bases $D_v\{v_1, \ldots, v_n\} =$ and $D_w = \{w_1, \ldots, w_m\}$ respectively. Let $f : V \to W$ be an l.t. For each $v_i \in D_v$ there exists a unique set of $a_{ij}$ such that:
$f(v_i) = w = \sum_{j=1}^{m} aijw_j$. This follows because every $w \in W$ can be uniquely written in terms of the basis vectors in $D_w$. The matrix $[a_{ij}] \in \mathcal{M}_{m\times n}(F)$ which completely determines the l.t. $f$ in the following sense: every $v \in V$ is uniquely written as $\sum_{k=1}^{n} b_k v_k$ so:

$$\begin{aligned} f(v) &= f(\sum_{k=1}^{n} b_k v_k) \\ &= \sum_{k=1}^{n} b_k f(v_k) \\ &= \sum_{k=1}^{n} b_k \sum_{j=1}^{m} a_{jk} w_j \\ &= \sum_{j=1}^{m} (\sum_{k=1}^{n} a_{jk} b_k) w_j \text{ (exchange sums since finite)} \end{aligned}$$

So given bases $V$, $W$ over $F$ with each l.t. $f : V \to W$ we can associate a matrix $[a_{jk}] \in \mathcal{M}_{m\times n}(F)$ which completely determines $f$. Conversely, every matrix $[a_{jk} \in \mathcal{M}_{m\times n}(F)$ uniquely defines a l.t. $f : V \to W$. $f : v = \sum_{k=1}^{n} b_k \mapsto \sum_{j=1}^{m} (\sum_{k=1}^{n} a_{jk} b_k) w_j$ This is summarized in the following theorem.

**Theorem 67.** *Let $V$, $W$ be vector spaces $F$ and of dimensions $n$, $m$ respectively. For each basis $D_v$ of $V$ and $D_w$ of $W$ there exists an isomorphism $\phi_{D_v D_w} : Hom(V, W) \to \mathcal{M}_{m\times n}(F)$ defined by $\phi_{D_v D_w}(f) = [a_{ij}$ where $f(v_j) = \sum_{i=1}^{m} a_{ij} w_i$ for $v_j \in D_v$, $j \in 1..n$.*

*Proof.* $\phi_{D_v D_w}$ is a bijection (discussion above) so we need to show it is an l.t. If $f, g \in Hom(V, W)$ and $\phi_{D_v D_w}(f) = [a_{ij}]$ and $\phi_{D_v D_w}(g) = [b_{ij}]$. Then:

$$\begin{aligned} (f + g)(v_j) &= \sum_{i=1}^{m} (a_{ij} + b_{ij}) w_i \\ &= \sum_{i=1}^{m} a_{ij} w_i + \sum_{i=1}^{m} b_{ij} w_i \\ &= f(v_j) + g(v_j) \end{aligned}$$

so $\phi_{D_vD_w}(f+g) = \phi_{D_vD_w}(f) + \phi_{D_vD_w}(g)$. For $c \in F$ we have

$$\begin{aligned}(cf)(v_j) &= \sum_{i=1}^{m} ca_{ij}w_i \\ &= c(\sum_{ij}^{m} w_i) \\ &= cf(v_j)\end{aligned}$$

So, $\phi_{D_vD_w}(cf) = c\phi_{D_vD_w}(f)$ is indeed linear. □

**Corollary 68.** *The vector space $Hom(V, W)$ has dimensions $m.n$.*

Note: A matrix is defined by the bases chosen and the order in which $v_i \in D_v$ are arranged (so order matters).

**Proposition 69.** *Let $V$, $W$, $X$ be vector spaces over $F$ and $D_v$, $D_w$, $D_x$ their bases respectively. Let l.t.s $f : V \to W$, $g : W \to X$. Then $\phi_{D_vD_x}(gf) = \phi_{D_wD_x}(g)\phi_{D_vD_w}(f)$.*

*Proof.* Left as an exercise. □

*Example* 70. Let $V = \mathcal{R}^3$, $W = \mathcal{R}^2$. Consider the following bases for $V$ and $W$,
$D_v = \{v_1 = [1/2, -1/2, 0], v_2 = [1/2, 0, -1/2], v_3 = [0, 1/2, -1/2]\}$ and $D_w = \{[1, 1], [1, 0]\}$. If $[x_1, x_2, x_3] \in \mathcal{R}^3$, there exists $a_1, a_2, a_3$ such that $[x_1, x_2, x_3] = a_1v_1 + a_2v_2 + a_3v_3 = [\frac{a_1+a_2}{2}, \frac{-a_1+a_3}{2}, \frac{-a_2+a_3}{2}]$.

This gives $x_1 = \frac{a_1+a_2}{2}$, $x_2 = \frac{-a_1+a_3}{2}$, $x_3 = \frac{-a_2+a_3}{2}$ and we get $a_1 = x_1 - x_2 + x_3,\ a_2 = x_1 + x_2 - x_3,\ a_3 = x_1 + x_2 + x_3$.

The matrix

$$\begin{bmatrix} 3 & 5 & 7 \\ 4 & 8 & 2 \end{bmatrix}$$

defines an l.t. $f : \mathcal{R}^3 \to \mathcal{R}^2$ where $a_1v_1 + a_2v_2 + a_3v_3 \mapsto ((3a_1 + 5a_2 + 7a_3)[1, 1], (4a_1 + 8a_2 + 2a_3)[1, 0])$ in $\mathcal{R}^2$. So, $f([x_1, x_2, x_3]) \mapsto (15x_1 + 9x_2 + 5x_3)[1, 1] + (14x_1 + 6x_2 - 2x_3)[1, 0] = [29x_1 + 15x_2 + 3x_3, 15x_1 + 9x_2 + 5x_3]$.

Let $f : V \to W$ be a l.t. from an $n$ dimensional vector space $V$ to an $m$ dimensional vector space $W$, both vector spaces over field $F$, with bases $D_v = \{v_1, \ldots, v_n\}$ and $D_w = \{w_1, \ldots, w_m\}$. Let $A = [a_{ij}]$ be the representation of the l.t. $f$ with respect to the bases $D_v,\ D_x$. Then we see that the row $R_j = [a_{j1}, \ldots, a_{jm}]$ stands for the vector $\sum_{i=1}^{m} a_{ji}w_i \in W$.

The matrix $A = [a_{ij}]$ can be seen as a sequence of$m$ column vectors: $[C_1, \ldots, C_m]$ where column $C_i = [a_{1i}, \ldots, a_{ni}]^T$ or as a sequence of row vectors $[R_1, \ldots, R_n]$ where each row $R_i = [a_{i1}, \ldots, a_{im}]$. Define the subspace spanned by the $R_i,\ i \in 1..n$ as the *row subspace* which will be a subspace of $\mathcal{R}^m$. The corresponding subspace spanned by the columns $C_j,\ j \in 1..m$ is the *column subspace* which will be a subspace of $\mathcal{R}^n$. We define *row rank* as the dimension of the row subspace and the *column rank* as the rank of the column subspace. The theorem below relates the row rank and the column rank.

**Theorem 71.** *The row rank and column rank of an $n \times m$ matrix $A_{n\times m} = [a_{ij}]$ are equal.*

*Proof.* Let $A_r$ be the row subspace with row rank $k$ and basis $D_r = \{v_1, \ldots, v_k\}$. Now a row vector $R_i,\ i \in 1..n$ is $R_i = [a_{i1}, \ldots, a_{im}]$. It can also a l.c. of the basis vectors $R_i = \sum_{j=1}^{k} a_{ij}v_j$ where $v_j \in D_r$. Let $v_j$ be the vector $v_j = [b_{j1}, \ldots, b_{jm}]$. Then we can write: $[a_{i1}, \ldots, a_{im}] = \sum_{j=1}^{k} a_{ij}v_j = \sum_{j=1}^{k} a_{ij}[b_{j1}, \ldots, b_{jm}]$. Equating coordinates gives $a_{il} = \sum_{j=1}^{k} a_{ij}b_{jl}$, $i \in 1..n$ and $l \in 1..m$. Writing this in matrix form gives:

$$\begin{bmatrix} a_{1l} \\ . \\ . \\ a_{nl} \end{bmatrix} = \begin{bmatrix} a_{11} \\ . \\ . \\ a_{n1} \end{bmatrix} b_{1l} + \ldots + \begin{bmatrix} a_{n1} \\ . \\ . \\ a_{nk} \end{bmatrix} b_{kl}$$

Observe that the left side of the matrix is just column $C_l$. So, a column is a l.c. of $k$ vectors and the column subspace can have a dimension at most $k$ so the column rank is at most equal to the row rank. A similar argument with columns $C_i$s instead of $R_i$s establishes that row rank is at most equal to the column rank. Together they imply that row rank and column rank are equal. □

Geometrically a l.t. scales, rotates or projects a vector.

# 8 Determinants

**Definition 72** (Determinant)**.** A mapping $D : M_{n\times n}(F) \to F$ is a determinantal mapping (detmap) if it satisfies: ($A \in M_{n\times n} = [a_1, \ldots, a_n]$ where $a_i$ is the $i^{th}$ column)

1. Multilinearity (linear function of each column) that is:
   D1: $D[\ldots, a_i + b_i, \ldots] = D[\ldots, b_i, \ldots] + D[\ldots, c_i, \ldots]$.
   D2: $D[\ldots, c.a_i, \ldots] = c.D[\ldots, a_i, \ldots]$.
2. Alternation that is:
   D3: $D[\ldots, a_i, \ldots, a_j, \ldots] = -D[\ldots, a_j, \ldots, a_i, \ldots]$.
3. Identity preserving, that is:
   D4: $D(I) = 1_f$

◁

*Exercise: Show that if D above satisfies D1 then it satisfies D3 iff it satisfies $D(A) = 0$.*

We now state several properties of determinants that are well known. We do not prove any of them. Look at any elementary linear algebra book for proofs. Since we are only concerned with square matrices in this section we will simply say 'matrix of size $n$' to mean $n \times n$.

**Proposition 73.** *If D satisfies D1 then it satisfies D3 iff it satisfies:*
*D3': $D(A) = 0$ whenever A has two identical columns.*

**Corollary 74.** *D is a detmap iff it satisfies D1, D2, D3', D4.*

A detmap of a matrix of size $n$ can be built up using detmaps of size $(n-1)$.

**Proposition 75.** *Let $n \geq 3$ and $D : \mathcal{M}_{n\times n}(f) \to F$ be a detmap. For $i \in 1..n$ define $f_i : \mathcal{M}_{n\times n}(F) \to F$ by:*

$$f_i(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} D(A_{ij})$$

*then each $f_i$ is a detmap. $A_{ij}$ is obtained from A by deleting the $i^{th}$ row and $j^{th}$ column*

Propostion 75 can be used to calculate $D(A)$, namely the *determinant* of $A$ - written $detA$, by the Laplace expansion.

$$detA = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} detA_{ij}$$

So, the determinant is the same along any row. Also if $A'$ is a transpose of $A$ then $detA = detA'$. So, the value of the determinant is the same along any column.

**Proposition 76.** *If $A, B \in \mathcal{M}_{n\times n}(F)$ then $detAB = detA.detB$*

**Proposition 77.** *If $A \in \mathcal{M}_{n\times n}(F)$ is invertible then $detA \neq 0$ and $detA_{-1} = \frac{1}{detA}$.*

**Definition 78** (Adjoint)**.** If $A \in \mathcal{M}_{n\times n}(F)$ then adjoint of $A$, written $adjA$ is the matrix $[adjA]_{ij} = (-1)^{i+j} detA_{ji}$. ◁

**Proposition 79.** *For every matrix $A$ of size $n$ $A.adjA = (detA)I_n = adjA.A$.*

A matrix $A$ is *singular* if $detA = 0$ else it is *non-singular*.

**Proposition 80.** *A matrix $A$ is invertible iff it is non-singular. The inverse matrix is: $A^{-1} = \frac{adjA}{detA}$*

**Proposition 81** (Cramer's rule)**.** *If $A$ is invertible then the system of linear equations $Ax = b$ has the solution $x = A^{-1}b$ and $x_i = \frac{detA_{i;b}}{detA}$, where $A_i; b$ means $A$ with its $i^{th}$ column replaced by $b$.*

# 9 Inner Product Spaces

To define a notion of distance we start by defining an *inner product* on a vector space. For simplicity let the vector space in this section be $\mathcal{R}^n$ (it can also be $\mathcal{C}^n$).

**Definition 82** (Inner Product)**.** An inner product is a function $\langle . \rangle : V \times V \to \mathcal{R}$ rm

satisfying:

1. $\langle x, x\rangle \geq 0$ and $\langle x, x\rangle = 0$ iff $x = 0_v$.
2. $\langle x, y\rangle = \langle y, x\rangle$.
3. $\langle ax, y\rangle = a\langle x, y\rangle$.
4. $\langle x + z, y\rangle = \langle x, y\rangle + \langle z, y\rangle$.

◁

**Definition 83** (Norm)**.** The norm of a vector $x$, written $\parallel x \parallel$, is $\parallel x \parallel = \sqrt{\langle x, x\rangle}$. ◁

A norm allows us to define distance between two vectors.

**Definition 84** (Distance)**.** The distance between two vectors $x,\ y$ is $d(x, y) = \parallel x - y \parallel = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$. ◁

An *inner product space* (ips henceforth) is a vector space with an inner product defined on it.

*Example* 85. Let $V = \mathcal{C}$ and $\langle z, w\rangle = Re(z\overline{w})$ defines an ips.

*Example* 86. Let $V$ be the set of all continuous functions on $[0, 1]$ and define the inner product by $ppfg = \int_0^1 f(t)g(t)dt$.

*Exercise:* For both examples confirm that the inner product defined satisfies the properties of an inner product.

Further any vector $v$ can be written as $v = \parallel v \parallel u$ where $u$ is a unit vector parallel to $v$. For a unit vector, $\parallel u \parallel = 1$. We now look at some important properties of inner products and norms.

**Proposition 87** (Cauchy-Schwarz-Bunyakovsky Inequality)**.** *Let $u, v \in V$ be vectors in an ips then $|\langle u, v\rangle|^2 \leq \langle u, u\rangle\langle v, v\rangle$ or $|\langle u, v\rangle| \leq \parallel u \parallel\parallel v \parallel$. Equality holds only if $u,\ v$ are linearly dependent.*

*Proof.* If either $u$ or $v$ is $0_v$ then it holds trivially. So, assume $u, v \neq 0_v$. Write $a = -\langle v, u\rangle$ and $b = \langle u, u\rangle$.

$$\begin{aligned} 0 &\leq \langle au + bv, au + bv\rangle \\ &= a^2\langle u, u\rangle + b^2\langle v, v\rangle + ab\langle u, v\rangle + ba\langle v, u\rangle \\ &= a^2 b + b^2\langle v, v\rangle - 2a^2 b \\ &= b^2\langle v, v\rangle - a^2 b \\ a^2 &\leq b\langle v, v\rangle \\ |\langle u, v\rangle|^2 &\leq \langle u, u\rangle\langle v, v\rangle \end{aligned}$$

It is clear that equality holds above if $u = cv$, that is $u, v$ are linearly dependent. □

**Proposition 88.** *Let $V$ be an ips and $u, v \in V$ and $a \in \mathcal{R}$ then:*

1. $\| av \| = |a| \, \| v \|$.
2. $\| v \| \geq 0$ *and* $\| v \| = 0$ *iff* $v = 0_v$.
3. $\| u + v \| \leq \| u \| + \| v \|$ - *Minkowski inequality.*
4. $\| u + v \|^2 + \| u - v \|^2 = 2(\| u \|^2 + \| v \|^2)$ - *parallelogram law.*

*Proof.*

1. $\| av \| = \langle av, av\rangle^{1/2} = (a^2\langle v, v\rangle)^{1/2} = |a| \, \| v \|$.
2. Follows from definition of inner product and norm.
3. $$\begin{aligned} \| u + v \|^2 &= \langle u + v, u + v\rangle \\ &= \langle u, u\rangle + \langle v, v\rangle + 2\langle u, v\rangle \\ &= \| u \|^2 + \| v \|^2 + 2\langle u, v\rangle \\ &\leq \| u \|^2 + \| v \|^2 + 2 \| u \| \| v \| \quad \text{using CSB} \\ &\leq (\| u \| + \| v \|)^2 \\ \| u + v \| &\leq (\| u \| + \| v \|) \end{aligned}$$
4. This follows by a straighforward expansion and simplification of the left hand side.

□

While we defined the norm using the inner product a more general definition of norm is: $\| \, . \, \| : v \mapsto \| v \| \in F$, $F$ a scalar field and it satisfies conditions 1. to 3. in proposition 88 above. So, for example if $u, v \in \mathcal{R}^n$ then $|u| + |v|$ is a norm, but it is not induced by an inner product. Several other non inner product norms are possible and used.

**Definition 89** (Orthogonal)**.** Two vectors $u, v \in V$, $V$ an ips are orthogonal iff $\langle u, v\rangle = 0$. This is represented by $u \perp v$. ◁

**Proposition 90.** *Let $u, v, w \in V$ be vectors in ips $V$. Then*

1. $d(u, v) = d(v, u)$.
2. $d(u, v) \geq 0$. *Equality only if* $u = v$.

*3.* $d(u,v) + d(v,w) \geq d(u,w)$ *- the triangular inequality.*

*Proof.* Follows from the previous proposition. Left as an exercise. □

*Example* 91. Let $V$ be space of continuous functions over $[-1,1]$ to $\mathcal{R}$. Define $\langle f,g\rangle = \int_{-1}^{1} f(t)g(t)dt$. For $i \geq 0$ define: $p_0(x) = 1, p_1(x) = x, \ldots, p_{i+1}(x) = \frac{2h+1}{h+1}xp_h(x) - \frac{h}{h+1}p_{h-1}(x)$. These are the Legendre polynomials and $p_i \perp p_j,\ i \neq j$.

**Proposition 92.** *Let $V$ be an ips then 1) If $v \in V$ is such that $v \perp u$ for all $u \in V$ then $v = 0_v$. 2) Let $A \subseteq V$, $A \neq \phi$ and $v \perp u, \forall u \in A$. Then $v \perp w,\ \forall w \in W = FA$.*

*Proof.* For 1) observe that $v \perp v$ and this is only possible if $v = 0_v$.
For 2) $w \in FA$ implies $w = \sum_{i=1}^{|A|} a_i v_i,\ v_i \in A$. Let $v \in V,\ v \perp u,\ \forall u \in A$. We now calculate $\langle v, w\rangle$.
$\langle v, w\rangle = \langle v, \sum_{i=1}^{|A|} a_i v_i\rangle = \sum_{i=1}^{|A|}\langle v, v_i\rangle = 0$. □

We now argue that if vectors in a set are mutually orthogonal then they are linearly independent. This allows us to move towards the standard basis of $\mathcal{R}^n$ that is familiar to us.

**Proposition 93.** *Let $V$ be an ips, $\phi \neq A \subset V$. For any $v_i, v_j \in A,\ v_i, v_j \neq 0_v$ and $i \neq j$ let $v_i \perp v_j$ then $A$ is linearly independent.*

*Proof.* Let $\sum_{i=1}^{|A|} a_i v_i = 0_v,\ v_i \in A$. Then we must show that $a_i = 0,\ i \in 1..|A|$, where no $v_i$ is $0_v$. Let $|A| = n$.

$$\begin{aligned} a_i\langle v_j, v_j\rangle &= \sum_{i=1}^{n} a_i\langle v_i, v_j\rangle,\ \langle v_i, v_j\rangle = 0,\ i \neq j \\ &= \langle \sum_{i=1}^{n} a_i v_i, v_j\rangle \\ &= \langle 0_v, v_j\rangle \\ &= 0 \end{aligned}$$

This is possible only if $a_i = 0_f$ for all $i \in 1..n$ since $\langle v_j, v_j\rangle > 0$. So, $A$ is l.i. □

We close this section by arguing that we can find an orthogonal basis for every finite dimensional vector space.

**Proposition 94** (Gram-Schmidt algorithm)**.** *A finitely generated ips $V$ over $F$ has a mutually orthogonal basis.*

*Proof.* The proof is by induction on the dimension of $V$. It is trivially true if dimension is 1. Let it be true for dimension $m$ and let $W \subset V$ be an $m$-dimensional subspace of $V$ with basis $\{v_1, \ldots, v_m\}$. Consider the vector $v \in V - W$ and define $a_i = \frac{\langle v, v_i\rangle}{\langle v_i, v_i\rangle}$ for each $i \in 1..m$. Let $v_{m+1} = v - \sum_{i=1}^{m} a_i v_i$. $v_{m+1} \notin W$ since $v \notin W$. For each $j \in 1..m$ we have,

$$\begin{aligned} \langle v_{m+1}, v_j\rangle &= \langle v - \sum_{i=1}^{m} a_i v_i,\ v_j\rangle \\ &= \langle v, v_j\rangle - \langle \sum_{i=1}^{m} a_i v_i,\ v_j\rangle \\ &= \langle v, v_j\rangle - a_j\langle v_j, v_j\rangle,\ \text{since } v_i \perp v_j,\ i \neq j \\ &= \langle v, v_j\rangle - \frac{\langle v, v_j\rangle}{\langle v_j, v_j\rangle}\langle v_j, v_j\rangle \\ &= 0 \end{aligned}$$

So, $\{v_1, \ldots, v_{m+1}\}$ is mutually orthogonal and is therefore linearly independent by propositon 93. □

The Gram-Schmidt process is an algorithm to obtain an orthogonal (actually we can easily get an orthonormal) basis starting from some basis.

**Input**: Basis $V = \{v_1, \ldots, v_n\}$
**Output**: Orthonormal basis $\{u_1, \ldots, u_n\}$
**for** $1 \leq k \leq n$ **do**
  $u'_k = v_k - \sum_{j=1}^{k-1} \langle v_k, u_j \rangle u_j$;
  $u_k = \| u'_k \|^{-1} u'_k$;
**end**

**Algorithm 1:** Gram-Schmidt algorithm

**Definition 95** (Orthogonal Complement)**.** Let $W$ be a subspace of $V$. Let $W^\perp$ be the subspace of all $v \in V$ such that $v \perp w$ for all $w \in W$. $W^\perp$ is a subspace of $V$. It is called the orthogonal complement of $W$ in $V$. $\triangleleft$

**Proposition 96.** *If $V$ is a finitely generated ips over $F$ and $W$ is a subspace of $V$ then $V = W \oplus W^\perp$ and $(W^\perp)^\perp = W$.*

*Proof.* Left as an exercise. $\square$

# 10 Eigenvalues and Eigenvectors

Let us consider an endomorphism $f : V \to V$ on a finitely generated vector space $V$over $F$. We want to find a matrix representation $\phi_{DD}(f)$ that will be as 'nice and convenient' as possible. Let us agree to call a subspace $W$of $V$*stable* under an endomorphism $f : V \to V$ iff $\forall w \in W, f(w) \in W$ i.e. $f$ restricted to subspace $W$is an endomorphism on $W$.

**Definition 97** (Eigenvalue, eigenvector)**.** Let $f : V \to V$ be an endomorphism. A scalar $a \in F$ is an eignevalue of $f$ iff $\exists v \in V,\ v \neq 0_v,$ such that $fv = av$. The vector $v$ is called an eigenvector of $f$. $\triangleleft$

Observe that:

1. $v \in V,\ v \neq 0_v$ is an eigenvector of $f$ iff $Fv$ is stable under $f$.
2. Every eigenvector is associated with exactly one eigenvalue.
3. An eigenvalue may be associated with multiple eigenvectors.
4. An endomorphism may not have any eigenvalues. For example, $F = \mathcal{R},\ f : \mathcal{R}^2 \to \mathcal{R}^2, [x, y] \mapsto [-y, x]$ has no eigenvalues.

The set of all eigenvalues of $f$ is called the *spectrum* of $f$ - written *spec(f)*.

*Example* 98. $F = \mathcal{R},\ f : \mathcal{R}^2 \to \mathcal{R}^2, [x, y] \mapsto [y, x]$. We have $spec(f) = \{1, -1\}$ and eigenvectors associated with 1 are $\{[a, a] | 0_f \neq a \in F\}$ and with $-1$ $\{[a, -a] | 0_f \neq a \in F\}$. So, we see that each eigenvalue has infinitely many eigenvectors associated with it.

**Proposition 99.** *Let $f : V \to V$ be an endomorphism on $V$a vector space over field $F$with an eigenvalue $a \in F$. The subset $W = \{0_v\} \cup \{$All eigenvectors of $a\}$ is a subspace of $V$.*

*Proof.* The proof is easy. We see that $f(0_v) = 0_v$ and $f(w_1 + w_2) = f(w_1) + f(w_2) = a.w_1 + a.w_2 = a.(w_1 + .w_2) \to w_1 + w_2 \in W$. Similarly, $f(cw) = c.f(w) = a.(c.w) = c.(a.w) \to a.w \in W$. $\square$

*Example* 100. Let $V = \mathcal{R}^3$ over $\mathcal{R}$ and let $f$ be an endomorphism defined by $f : [a, b, c] \to [a, 0, c]$. Then 1 is an eigenvalue of $f$ and the corresponding eigenspace is the subspace $F\{[1, 0, 0], [0, 0, 1]\}$

**Proposition 101.** *If $V$ is a finitely generated vector space and $f : V \to V$ is an endomorphism then following are equivalent for $c \in F$:*

1. *$c$ is an eigenvalue of $f$.*
2. *$f - c\sigma_1$ of $V$ is not an automorphism.*
3. *If $A$ is the matrix for $f$ with respect to some basis of $V$ then $|cI - A| = 0_f$.*

*Proof.* To be completed. □

**Proposition 102.** *If $V$ is a finitely generated vector space over $F$ and $f$, $g$ are endomorphisms on $V$ then $spec(fg) = spec(gf)$.*

**Proposition 103.** *If $V$ is a finitely generated vector space over $F$ and $f$ an endomorphism on $V$ then the following two conditions for a basis $D = \{d_1, \ldots, d_n\}$ of $V$ are equivalent:*

1. *Each $d_i$ is an eigenvector of $f$.*
2. *$\phi_{DD}(f)$ is a diagonal matrix.*

**Proposition 104.** *Let $V$ be a finitely generated vector space over $F$ and $f$ an endomorphism on $V$ with distinct eigenvalues $c_1, \ldots, c_k$. For each $i \in 1..k$ let $v_i$ be the eigenvector associated with $c_i$. Then $\{v_1, \ldots, v_k\}$ is linearly independent.*

The following are simple consequences of the propositions above.

**Corollary 105.** *Let the conditions of proposition 104 hold. Then for each $i \in 1..k$ the set of eigenspaces $\{W_1, \ldots, W_k\}$ is independent where $W_i$ is the eigenspace associated with eigenvalue $c_i$ of $f$.*

**Corollary 106.** *Let the conditions be as in proposition 104 then the number of distinct eigenvalues of $f$ is at most $dim(V)$.*

**Corollary 107.** *Let $V$ be a finitely generated vector space over $F$ and $f$ an endomorphism on $V$ with $n$ distinct eigenvalues then there is a basis $D$ of $V$ such that $\phi_{DD}(f)$ is a diagnoal matrix.*

Let $A \in M_{n\times n}$ then $|XI - A|$, that is the determinant of the matrix $XI - A$, is called the *characteristic polynomial* of $A$. The polynomial is an injection and has degree $n$. The eigenvalues of $A$ are the zeroes of $A$. For each non-singular matrix $P \in M_{n\times n}(F)$ we can define the following function $f_p : M_{n\times n}(F) \to M_{n\times n}(F)$ defined by $f_p : A \mapsto P^{-1}AP$.

*Exercise: Argue that $f_p$ is an automorphism.*

The set of all automorphisms of the above kind defines an equivalence relation. $A \sim B$ iff $B = P^{-1}AP$ and matrices $A$, $B$ are *similar*.

**Proposition 108.** *If matrices $A$, $B$ are similar then they have the same characteristic polynomial. However, the converse is not true.*

Note that given an endomorphism $f : V \to V$ where $V$is a finite dimensional vector space over field $F$we can identify the *characteristic polynomial* of $f$ as the characteristic polynomial of the matrix representing $f$ with respect to any basis of $V$.

Let $V$be a vector space, $f$ an endomorphism as above. Let $c$ be an eigenvalue of $f$ then the *algebraic multiplicity* of $c$ is $k$ if it is the largest integer such that $(X - c)^k$ is a factor of the characteristic polynomial of $f$. The geometric multiplicity of $c$ is the dimension of the eigenspace associated with $c$. The geometric multiplicity is less than equal to the algebric multiplicity.

**Proposition 109.** *Let $V$ be a finitely generated vector space over $F$ and $f$ an endomorphism on $V$ with distinct eigenvalues $c_1, \ldots, c_k$. Then the following two conditions are equivalent:*

1. *There is a basis $D$ of $V$ such that $\phi_{DD}(f)$ is a diagnonal matrix.*
2. *The algebraic multiplicity of $c_j$ is equal to its geometric multiplicity for $j \in 1..k$.*

If $p(X)$ is a polynomial of finite degree and $A \in M_{n\times n}(F)$ then we can define a function $p(A) = \sum_{i=0}^{n} a_i A^i$, $A^0 = I$. if $p(A) = O$ then $A$ is said to *annihilate* the polynomial $p(X)$. The set of all polynomials annihilated by a fixed matrix $A$ is a subspace of $F[X]$ - set of finite degree polynomials.

*Exercise: Show the above and then argue that similar matrices annihilate the same set of polynomials.*

**Proposition 110.** *Let $F$ be a field and $n > 0$. For $A \in M_{n\times n}$ there is a polynomial $p(X) \in F[X]$ of positive degree satisfying $p(A) = O$.*

**Theorem 111** (Cayley-Hamilton)**.** *Let $F$ be a field and $n > 0$. Then for $A \in M_{n\times n}(F)$ annihilates its characteristic polynomial.*