ALGEBRA I

EDWARD O'CALLAGHAN

Contents

1. Prelude

TODO: Fix notation here...

2. Introduction

In this course we build up the rudiments of some important notions of algebraic structures. That is, a algebraic structure of an arbitrary set, or carrier set, coupled with various finitary operations defined on it. ..

3. Groups

Definition 3.1 (Binary operation)**.** A **binary operation** on a set $\mathcal{X}$ is a map $\circ : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$. **N.B.** that the binary operation is *closed*.

Definition 3.2 (Magma)**.** A **magma** is a set $\mathcal{M}$ equipped with a binary operation $\circ$. We denote the magma as the tuple pair $(\mathcal{M}, \circ)$.

Definition 3.3 (Semi-group)**.** A **semi-group** is a set $\mathcal{G}$ equipped with binary operation that is *associative*. Hence, a semi-group is a magma where the operation is *associative*; That is, given any $x, y, z \in \mathcal{G}$ then $x \circ (y \circ z) = (x \circ y) \circ z \in \mathcal{G}$. We denote the semi-group as the tuple pair $(\mathcal{G}, \circ)$, not to be confused with a magma from context.

Definition 3.4 (Monoid)**.** A **semi-group with idenitity** or, **monoid** for short, is a semi-group $(\mathcal{G}, \circ)$ with a unique identity element $e \in \mathcal{G}$ such that $x \circ e = x = e \circ x \forall x \in \mathcal{G}$

Proof: unquieness of idenitity. Assume some other identity $e^{'}$ exists in $\mathcal{G}$ then, $e^{'} = e^{'} \circ e = e \circ e^{'} = e$. $\square$

Example 3.5. Given $\mathcal{G} = \mathbb{N}$ with the binary law of composition $\circ$ to be defined as arithmetic addition $+$. Then, $(\mathbb{N}, +)$ forms a semi-group with identity 0. Verify the axioms.

Definition 3.6 (Group)**.** A **group** is a monoid where every element has an inverse. An abelian group is a group that is commutative.

Example 3.7. Given $\mathcal{G} = \mathbb{Z}$ with the binary law of composition $\circ$ to be defined as arithmetic addition $+$. Then, $(\mathbb{Z}, +)$ forms a semi-group with identity 0. Verify the axioms.

Question 3.8. *Why does the set of naturals* $\mathbb{N}$ *not form a group under multiplication, however does form a monoid?*

Definition 3.9 (Subgroup)**.** A group $\mathcal{H}$ is a **subgroup** of a group $\mathcal{G}$ if the restriction of the binary operation $\circ : \mathcal{H} \times \mathcal{H} \to \mathcal{H}$ is a group operation on $\mathcal{H}$. In particular, A non-empty subset $\mathcal{H}$ of a group $\mathcal{G}$ is a subgroup of $\mathcal{G}$ if and only if $h_1 \circ h_2 \in \mathcal{H}$ whenever $h_1, h_2 \in \mathcal{H}$, and $h^{-1} \in \mathcal{H}$ whenever $h \in \mathcal{H}$. We denote the subgroup by $\mathcal{H} \leq \mathcal{G}$.

Theorem 3.10 (Smallest subgroup)**.** *If $\mathcal{A}$ is a subset of a group $\mathcal{G}$, there is a* smallest *subgroup $Gp(\mathcal{A})$ of $\mathcal{G}$ which contains $\mathcal{A}$, the subgroup* generated *by $\mathcal{A}$.*

Example 3.11. Suppose $\mathcal{A} = \{g\}$ then $\mathrm{Gp}(\mathcal{A}) = \mathrm{Gp}(g)$ and so $\mathrm{Gp}(g) = \{g^n : n \in \mathbb{Z}\}$, where $g^0 = e$, g^n is the product of n copies of g where $n > 0$, and g^n is the product of $|n|$ copies of g^{-1} when $n < 0$.

Definition 3.12 (Cyclic group)**.** A group $\mathcal{G}$ is *cyclic* if $\mathcal{G} = \mathrm{Gp}(g)$ for some $g \in \mathcal{G}$. Such a element is called a *generator* of the group.

Definition 3.13 (Group order)**.** If a group $\mathcal{G}$ has finitely many elements, then the *order* $o(\mathcal{G})$ is the number of elements of $\mathcal{G}$.

Definition 3.14 (Normal subgroup)**.** A subgroup $\mathcal{H}$ of a group $\mathcal{G}$ is a **normal**, or *self-conjugate*, if $g^{-1}hg \in \mathcal{H}\,\forall g \in \mathcal{G}$ and $h \in \mathcal{H}$. We denote the normal $\mathcal{H} \trianglelefteq \mathcal{G}$.

Definition 3.15 (Simple group)**.** A group $\mathcal{G}$ is **simple** if it has no normal subgroups other than $\{e\}$ and $\mathcal{G}$.

3.1. **Group Homomorphisms.** Homomorphisms are structure preserving mappings. In group homomorphisms we preserve the structure of the binary operation $\circ$ as follows;

Definition 3.16 (Homomorphism)**.** Let $\mathcal{G}$ and $\mathcal{H}$ be two groups. Then a mapping

$$\varphi : \mathcal{G} \to \mathcal{H}$$

is called a *homomorphism* if

$$\varphi(x \circ y) = \varphi(x) \circ \varphi(y) : x, y \in \mathcal{G}$$

It follows that, for some $g \in \mathcal{G}$ we have,

$$\begin{aligned} \varphi(e_g) &= \varphi(g \circ g^{-1}) \\ &= \varphi(g) \circ \varphi(g^{-1}) \\ &= \varphi(g) \circ (\varphi(g))^{-1} \\ &= e_h \in \mathcal{H}. \end{aligned}$$

That is the identity e has been preserved. Hence, it does not matter if we compose in $\mathcal{G}$ and map to $\mathcal{H}$ or take two elements in $\mathcal{G}$ then compose the mapped elements in $\mathcal{H}$, since the group structure has been preserved.

How much information about the elements inside the structure is, however, another quality to consider. Hence we fix some terminology here.

- A homomorphism that is injective is called monomorphic.
- A homomorphism that is surjective is called epimorphic.

- A homomorphism that is bijective is called isomorphic.

Thus we have the following definitions by considering a group homomorphism $\varphi : \mathcal{G} \to \mathcal{H}$.

Definition 3.17 (Monomorphic)**.** φ is **monomorphic** if for $\varphi(x) = \varphi(y) \implies x = y\, \forall x, y \in \mathcal{G}$.

Definition 3.18 (Epimorphic)**.** φ is **epimorphic** if $\forall h \in \mathcal{H} \exists g \in \mathcal{G}$ so that $\varphi(g) = h$.

Definition 3.19 (Isomorphic)**.** φ is **isomorphic** if φ is **both** mono- and epic- morphic.

Some special cases are sometimes of particular interest and we shall outline them now.

Definition 3.20 (Endomorphic)**.** A monomorphism $\mathcal{G} \to \mathcal{G}$ for a group $\mathcal{G}$ is called an *endomorphism* of $\mathcal{G}$.

Definition 3.21 (Automorphic)**.** A isomorphism $\mathcal{G} \to \mathcal{G}$ for a group $\mathcal{G}$ is called an *automorphism* of $\mathcal{G}$.

Remark. The set $Aut(\mathcal{G})$ of automorphisms of $\mathcal{G}$ forms a group, when composition of mappings is taken as the group law of composition.

3.2. **Properties of homomorphisms.**

Definition 3.22 (kernel)**.** If $\varphi : \mathcal{G} \to \mathcal{H}$ is a group homomorphism, then the *kernel* is the set $\ker(\varphi) = \{g \in \mathcal{G} : \varphi(g) = e_h \in \mathcal{H}\}$.

If $\varphi : \mathcal{G} \to \mathcal{H}$ is a group homomorphism, then observe that $\ker(\varphi)$ is a normal subgroup of of $\mathcal{G}$.

3.3. **Cosets.** Let $\mathcal{G}$ be a group and $\mathcal{H}$ be a subgroup of $\mathcal{G}$ with $g \in \mathcal{G} : g \notin H$, then

Definition 3.23 (Left Coset)**.** $gH = \{gh : h \in H\}$ is a **left coset of** $\mathcal{H}$ in $\mathcal{G}$.

Definition 3.24 (Right Coset)**.** $Hg = \{hg : h \in H\}$ is a **right coset of** $\mathcal{H}$ in $\mathcal{G}$.

Definition 3.25 (Normal Subgroup)**.** If $gH = Hg$ then $\mathcal{H}$ is a **normal** subgroup of $\mathcal{G}$, denoted by $\mathcal{H} \trianglelefteq \mathcal{G}$.

3.4. **Factor (or Quotient) groups.** Let $\mathcal{G}$ be a commutative group and consider a subgroup $\mathcal{H}$. Then $\mathcal{H}$ determines an equivalence relation in $\mathcal{G}$ given by

$$x \sim x' \text{ iff } x - x' \in \mathcal{H}.$$

..

3.5. **Non-commutative Groups.** A common class of non-commutative groups are transformation groups. Note:

Definition 3.26 (Transformation)**.** A bijective map $\varphi : X \to X$ is called a **transformation** of X.

Note. The most trivial case is the *idenitity map* id_X by $id_X(x) = x$, $\forall x \in X$.

Hence, there exists a inverse φ^{-1} of φ such that $\varphi^{-1} \circ \varphi = id_X = \varphi \circ \varphi^{-1}$. Now, take two transformations of X, φ and ψ, and let the product $\varphi \circ \psi$ be well defined. Then the set of all transformations of X form the group **Transf(X)**. Since, given $\varphi, \psi, \phi \in Transf(X)$ then we have associativity, $\varphi \circ (\psi \circ \phi) = (\varphi \circ \psi) \circ \phi$. We have identity $e = id_X \in Transf(X)$ and so, inverses $\forall \varphi \in Transf(X) \exists! \varphi^{-1} : \varphi \circ \varphi^{-1} = e$. Closure follows from the composition of two transformations φ and ψ, since $(\varphi \circ \psi)^{-1} = \psi^{-1} \circ \varphi^{-1}$.

A transformation group is a type of group action which describes symmetries of objects. More abstractly, since a group $\mathcal{G}$ is a category with a single object in which every morphism is bijective. Then, a group action is a *forgetful functor* $\mathcal{F}$ from the group $\mathcal{G}$ in the category **Grp** to the set category **Set** that is, $\mathcal{F} : \mathcal{G} \to$ **Set**.

3.6. **Group actions.** For any mathematical object (e.g. sets, groups, vector spaces) X an isomorphism of X is a symmetry of X. The set of all isomorphisms of X, or symmetries of X, form a group called the symmetry group of X, denoted $Sym(X)$. More formally;

Definition 3.27 (Group action)**.** An *action* of a group $\mathcal{G}$ on a mathematical object X is a mapping $\mathcal{G} \times X \to X$, defined by $(g, x) \mapsto g.x$ satisfying:

- $e.x = x \, \forall x \in X$ and
- $(gh).x = g.(h.x) \, \forall g, h \in \mathcal{G}, x \in X$.

That is, we have the (*left*) $\mathcal{G}$-action on X and denote this by $\mathcal{G} \curvearrowright X$.

Notice that we may study properties of the symmetries of some mathematical object X without reference to the structure of X in particular.

3.7. **Permutations.** Take a finite set X with $|X| = n$, then the transformations of X are called **permutations** of the elements of X. In particular, the group of permutations of $X = \{1, 2, \cdots, n\}$ is a **symmetric group**, denoted S_n, with **order** $|S_n| = n!$. Thus, by taking any subgroup of S_n we have a **permutation group**. Also note that, for finite sets, *permutation* and *bijective maps* refer to the same operation, namely rearrangement of elements of X. Another way is to consider, a group $\mathcal{G}$ and set X. Then a group action is defined as a group homomorphism φ from $\mathcal{G}$ to the symmetric group of X. That is, the action $\varphi : \mathcal{G} \to S_n(X)$, assigns a permutation of X to each element of the group $\mathcal{G}$ in the following way:

- From the identity element $e \in \mathcal{G}$ to the identity transformation id_X of X, that is, $\varphi : e \to id_X$;

- A product of group homomorphisms $\varphi \circ \psi \in \mathcal{G}$ is then the composite of permutations given by φ and ψ in X.

Given that each element of $\mathcal{G}$ is represented as a permutation. Then a group action can also be consider as a permutation representation.

A permutation $\sigma \in S_n$ can be written,

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix} \text{ where } a_1 = \sigma(1), a_2 = \sigma(2), \cdots .$$

The idenitity permutation $id_n \in S_n$ is simply,

$$id_n = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}$$

.

Since $|S_n| = n!$ then the total number of ways n elements maybe permuted is $n!$.

Take any two permutations $\sigma, \pi \in S_n$ then composition is well defined as **functional composition** as follows.

Given,

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix} \; and \; \pi = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix}$$

then,

$$\begin{aligned} \pi \circ \sigma &= \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(a_1) & \pi(a_2) & \cdots & \pi(a_n) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & \cdots & n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix} \end{aligned}$$

.

A inverse of any permutation $\sigma \in S_n$ is given by,

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}^{-1} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 2 & \cdots & n \end{pmatrix}$$

3.8. **Permutation parity.** Consider the algebraic structure:

$$\triangle_n(x_1, \ldots, x_n) = \prod_{i<j} (x_i - x_j)$$

TODO..

3.9. **Fields.** We now may build higher order algebraic structures using the notion of a group.

Definition 3.28 (Field)**.** A **field** $\mathbb{F}$ is a set together with two binary operations, addition and multiplication, such that:

- addition forms an abelian group,
- multiplication forms a abelian quasi-group, i.e. a commutative multiplicative group on the set $\mathbb{F} - \{0\}$,

coupled together with a law of distribution between the two binary operations.

4. Exact sequence

An **exact sequence** may either be a finite or infinite sequence of objects and morphisms between them. Such a sequence is constructed so that the image of one morphism equals the kernel of the next.

In particular;

Definition 4.1 (Exact Sequence)**.** Consider the sequence of n group homomorphism between $n+1$ groups in the following way:

$$\mathcal{G}_0 \xrightarrow{\varphi_1} \mathcal{G}_1 \xrightarrow{\varphi_2} \mathcal{G}_2 \xrightarrow{\varphi_3} \dots \xrightarrow{\varphi_n} \mathcal{G}_n$$

Then the sequence is said to be *exact* if,

$$\ker(\varphi_{k+1}) = \operatorname{im}(\varphi_k)$$

for every $k \in \{1 \dots n\}$. For $n = 3$ the sequence is said to be a **short exact sequence**.

Example 4.2. Suppose we have $\mathcal{K} \trianglelefteq \mathcal{G}$ and that $q : \mathcal{G} \to \mathcal{G}/\mathcal{K}$ is the quotient mapping. Then,

$$1 \longrightarrow \mathcal{K} \xrightarrow{\subseteq} \mathcal{G} \xrightarrow{q} \mathcal{G}/\mathcal{K} \longrightarrow 1$$

is a short exact sequence.

5. First Isomorphism Theorem

Theorem 5.1. *Let $\mathcal{G}$ and $\mathcal{H}$, and let $\varphi : \mathcal{G} \to \mathcal{H}$ be a group homomorphism. Then:*

- *The kernel of φ is a normal subgroup of $\mathcal{G}$; $\ker(\varphi) \trianglelefteq \mathcal{G}$,*
- *The image of φ is a subgroup of $\mathcal{H}$; $\operatorname{im}(\varphi) \leq \mathcal{H}$, and*
- *The image of φ is also isomorphic to the factor group $\mathcal{G}/\ker(\varphi)$; $\operatorname{im}(\varphi) \cong \mathcal{G}/\ker(\varphi)$.*

In particular, if φ is epimorphic then $\mathcal{H} \cong \mathcal{G}/\ker(\varphi)$.

We may represent these fundamental relations in the following commutative diagram.

Notice the *exact sequence* that runs from the lower left to the upper right of the commutative diagram.

6. Determinants

Definition 6.1 (Determinant)**.**

$$\det(A) = \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

Example 6.2. Find the determinant of some matrix $A \in \mathcal{M}_{2,2}(\mathbb{F})$.

Suppose

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

and take the permutation group

$$S_2 = \{\sigma_1, \sigma_2\}$$

with $\sigma_1 = id$ and $\sigma_2 = [1\,2]$. Then, by definition, we have:

$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_2} sgn(\sigma) \prod_{i=1}^{2} a_{i,\sigma(i)} \\ &= sgn(\sigma_1)(a_{1,\sigma_1(1)} a_{2,\sigma_1(2)}) + sgn(\sigma_2)(a_{1,\sigma_2(1)} a_{2,\sigma_2(2)}) \\ &= +(a_{11}a_{22}) - (a_{12}a_{21}) \\ &= a_{11}a_{22} - a_{12}a_{21}. \end{aligned}$$

Example 6.3. Find the determinant of some matrix $A \in \mathcal{M}_{3,3}(\mathbb{F})$.

Suppose

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

and take the permutation group

$$S_6 = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6\}$$

with

$$\sigma_1 = [1\,2\,3]\,\sigma_2 = [1\,3\,2]$$
$$\sigma_3 = [2\,1\,3]\,\sigma_4 = [2\,3\,1]$$
$$\sigma_5 = [3\,2\,1]\,\sigma_6 = [3\,1\,2]$$

Then, by definition, we have:

$$\begin{aligned}
\det(A) &= \sum_{\sigma\in S_6} sgn(\sigma)\prod_{i=1}^{6} a_{i,\sigma(i)} \\
&= sgn(\sigma_1)(a_{1,\sigma_1(1)}a_{2,\sigma_1(2)}a_{3,\sigma_1(3)}a_{4,\sigma_1(4)}a_{5,\sigma_1(5)}a_{6,\sigma_1(6)}) \\
&+ sgn(\sigma_2)(a_{1,\sigma_2(1)}a_{2,\sigma_2(2)}a_{3,\sigma_2(3)}a_{4,\sigma_2(4)}a_{5,\sigma_2(5)}a_{6,\sigma_2(6)}) \\
&+ sgn(\sigma_3)(a_{1,\sigma_3(1)}a_{2,\sigma_3(2)}a_{3,\sigma_3(3)}a_{4,\sigma_3(4)}a_{5,\sigma_3(5)}a_{6,\sigma_3(6)}) \\
&+ sgn(\sigma_4)(a_{1,\sigma_4(1)}a_{2,\sigma_4(2)}a_{3,\sigma_4(3)}a_{4,\sigma_4(4)}a_{5,\sigma_4(5)}a_{6,\sigma_4(6)}) \\
&+ sgn(\sigma_5)(a_{1,\sigma_5(1)}a_{2,\sigma_5(2)}a_{3,\sigma_5(3)}a_{4,\sigma_5(4)}a_{5,\sigma_5(5)}a_{6,\sigma_5(6)}) \\
&+ sgn(\sigma_6)(a_{1,\sigma_6(1)}a_{2,\sigma_6(2)}a_{3,\sigma_6(3)}a_{4,\sigma_6(4)}a_{5,\sigma_6(5)}a_{6,\sigma_6(6)})
\end{aligned}$$

Lemma 6.4.

$$\det(A^T) = \det(A).$$

Proof. Consider some σ such that $\sigma_1, \ldots, \sigma_n$ is in order so that $\tau = \sigma^{-1}$ and take the matrix A^i_j so that $(A^i_j)^T = A^j_i$.

Then we see,

$$A^{\sigma(i)}_j = A^j_{\sigma^{-1}(j)} = A^j_{\tau(j)}$$

with $sgn(\tau) = sgn(\sigma)$. So,

$$\begin{aligned}
\det(A) &= \sum_{\sigma\in S_n} sgn(\sigma)\prod_{i=1}^{n} a_{i,\sigma(i)} \\
&= \sum_{\tau\in S_n} sgn(\tau)\prod_{i=1}^{n} a_{\tau(i),\tau\cdot\sigma(i)} \\
&= \sum_{\tau\in S_n} sgn(\tau)\prod_{i=1}^{n} a_{\tau(i),i} \\
&= \det(A^T).
\end{aligned}$$

□

Lemma 6.5.

$$\det(AB) = \det(A)\det(B).$$

Proof. Let $A = [a^i_j]$ and $B = [b^i_j]$ with $A, B \in \mathcal{M}_{nn}(\mathbb{F})$ so that $AB = \left[\sum_{k=1}^n a^i_k b^k_j\right]$.

$$\begin{aligned}\det(AB) &= \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^n \left[\sum_{k=1}^n a^i_k b^k_{\sigma(i)}\right] \\ &= ..\end{aligned}$$

□

7. Adjugate (Classical Adjoint)

Definition 7.1 (Cofactor)**.** For some square matrix $A \in \mathcal{M}_{n,n}(\mathbb{F})$ then the minor of the a_{ij} entry, denoted by $M_{i,j}$, is defined to be the determinant of the *submatrix* obtained by removing the $(i,j)^{th}$ row and column from A. That is,

$$C_{i,j} = (-1)^{i+j} M_{i,j}$$

where $C_{i,j}$ is called the *cofactor* of $a_{i,j}$.

Example 7.2. Consider the matrix $A \in \mathcal{M}_{3,3}$ and suppose

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

We may find the $C_{2,3}$ cofactor in the following way;

Observe that the minor $M_{2,3}$ may be found as follows,

$$\begin{aligned}M_{2,3} &= \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}a_{32} - a_{12}a_{31}\end{aligned}$$

and so the cofactor $C_{2,3}$ is by definition,

$$\begin{aligned}C_{2,3} &= (-1)^{2+3} M_{2,3} \\ &= (-1)^5 (a_{11}a_{32} - a_{12}a_{31}) \\ &= a_{12}a_{31} - a_{11}a_{31}.\end{aligned}$$

Notice that we may now find the determinant of some $n \times n$ square matrix in terms of its cofactors. The process of this cofactor expansion is called the Laplace expansion.

Theorem 7.3. *Suppose $A \in \mathcal{M}_{n,n}$ is some $n \times n$ square matrix then*

in terms of the i^{th} row we have,

$$\det(A) = \sum_{i=1}^{n} a_{ij} C_{ij}$$

or in terms of the j^{th} column

$$= \sum_{j=1}^{n} a_{ij} C_{ij}.$$

Proof. TODO: Prove Laplace expansion. □

Example 7.4. Consider the matrix

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

Then,

$$\begin{aligned} \det(A) &= 1 \cdot \begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} - 3 \cdot \begin{vmatrix} 2 & 3 \\ 5 & 1 \end{vmatrix} + 5 \cdot \begin{vmatrix} 2 & 1 \\ 5 & 2 \end{vmatrix} \\ &= 1 \cdot (-5) - 3 \cdot (-13) + 5 \cdot (-1) = 29. \end{aligned}$$

Recall that the determinant is a measure of overall scaling of a matrix. Observe that each cofactor are essentially sub determinants or sub scalings. By this we are motivated to form a matrix of these cofactors in the following way:

Definition 7.5 (Matrix of cofactors)**.** Suppose $A \in \mathcal{M}_{n,n}(\mathbb{R})$ is some $n \times n$ real square matrix given by,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

with cofactors C_{ij} of a_{ij}. Then the *cofactor matrix* C is defined as,

$$C = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{pmatrix}.$$

Now since a matrix of a linear transformation is really just a linear dilation (Invariant Subspaces) of each dimension of the coordinate system (Invariant Subspaces). Then for a real matrix we may find the inverse by means of dilating each respective dimension back. This is essentially what the adjugate matrix is. In fact, even when the matrix is *singular* (non-invertible) the adjugate is still well defined. That is, the adjugate is in actual fact, the pre-image of a real linear transformation.

More formally, we have the following definition;

Definition 7.6 (Adjugate (Classical adjoint))**.** Suppose $A \in \mathcal{M}_{n,n}(\mathbb{R})$ is some real $n \times n$ square matrix with cofactor matrix C. Then the *adjugate* or classical adjoint of A is defined as,

$$adj(A) = C^T.$$

Thus we have the following result:

Lemma 7.7. *For some non-singular $n \times n$ real square matrix $A \in \mathcal{M}_{n,n}(\mathbb{R})$ we have,*

$$adj(A) = A^{-1}\det(A).$$

Proof. TODO: prove this lemma.. □

Example 7.8.

Suppose $A \in \mathcal{M}_{2,2}(\mathbb{R})$ is given by,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

then the adjugate of A is given by,

$$\begin{aligned} adj(A) &= \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix}^T \\ &= \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}. \end{aligned}$$

Some useful properties follow from the adjugate:

- $adj(I) = I$,

- $adj(AB) = adj(B)adj(A)$,
- $adj(A^T) = adj(A)^T$.

Proof. TODO prove the above properties. □

8. Vector Spaces

Definition 8.1 (Vector space)**.** A **vector space** $\mathcal{V}$ over a field $\mathbb{F}$ is an abelian additive group $(\mathcal{V}, +)$, coupled with a mapping (scalar multiplication) $(\lambda, x) \to \lambda x$ of $\mathbb{F} \times \mathcal{V}$ into $\mathcal{V}$ which satisfies

- $1.x = x$,
- $(\lambda + \mu)(x) = \lambda x + \mu x$,
- $\lambda(\mu x) = (\lambda\mu)x$,
- $\lambda(x + y) = \lambda x + \lambda y$,

for $\lambda, \mu \in \mathbb{F}$ and $x, y \in \mathcal{V}$. We note here that the elements of $\mathcal{V}$ are called *vectors* and that the elements of $\mathbb{F}$ are called *scalars*.

Remark. Note that any field $\mathbb{F}$ forms a degenerate vector space over itself.

Example 8.2. For a set X, the set of functions from $X \to \mathbb{F}$, denoted $Fun(X, \mathbb{F})$, is a vector space with pointwise addition and scalar multiplication. That is, for some $\varphi, \psi \in Fun(X, \mathbb{F})$, $x \in X$ and $\lambda \in \mathbb{F}$ we have,

$$(\varphi + \psi)(x) = \varphi(x) + \psi(x)$$

and

$$(\lambda\varphi)(x) = \lambda(\varphi(x)).$$

8.1. Subspaces.

Definition 8.3. A non-empty subset $\mathcal{U}$ of a vector space $\mathcal{V}$ over $\mathbb{F}$ is a *linear subspace* if it is a subgroup of $\mathcal{V}$ and if $\lambda x \in \mathcal{U}$ whenever $\lambda \in \mathbb{F}$ and $x \in \mathcal{U}$ written $\mathcal{U} \leq \mathcal{V}$.

Corollary. *A non-empty subset $\mathcal{U}$ of a vector space $\mathcal{V}$ over a field $\mathbb{F}$ is a linear subspace if and only if, for any $u, u' \in \mathcal{U}$ and $\lambda \in \mathbb{F}$ we have, $u + \lambda u' \in \mathcal{U}$.*

Example 8.4. Let the set of all real-valued functions on $\mathbb{R}^n$ be denoted by $\mathcal{F}(\mathbb{R}^n)$. Notice that, for any $f, g \in \mathcal{F}(\mathbb{R}^n)$, $\mathbf{x} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ we have, $(f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$ and $(\lambda f)(\mathbf{x}) = \lambda f(\mathbf{x})$. That is we claim that, $\mathcal{F}(\mathbb{R}^n)$ is a vector space.

Now, observe that the set of all continuous real-valued functions, denoted $\mathcal{C}(\mathbb{R}^n)$ is a subspace of $\mathcal{F}(\mathbb{R}^n)$.

Example 8.5. Given two subspaces $\mathcal{U}, \mathcal{V} \leq \mathcal{W}$ of vector space $\mathcal{W}$, then $\mathcal{U} \cap \mathcal{V} \leq \mathcal{W}$.

Proof.

Suppose $x, \lambda x' \in \mathcal{U} \cap \mathcal{V}$ with any $\lambda \in \mathbb{F}$. Then,

$$x, \lambda x' \in \mathcal{U} \text{ and } x, \lambda x' \in \mathcal{V}.$$

Now, since both $\mathcal{U}$ and $\mathcal{V}$ are subspaces then,

$$\begin{aligned} & x + \lambda x' \in \mathcal{U} \text{ and } x + \lambda x' \in \mathcal{V} \\ \implies & x + \lambda x' \in \mathcal{U} \cap \mathcal{V}. \end{aligned}$$

□

Lemma 8.6. *In general, for a vector space $\mathcal{V}$ and any family of subspaces $\mathcal{U}_i \leq \mathcal{V}$ for every $i \in I$, we have $\bigcap_{i \in I} \mathcal{U}_i \leq \mathcal{V}$.*

Example 8.7. Consider the subspaces $\mathcal{V}', \mathcal{V}'', \mathcal{W} \leq \mathcal{V}$. Prove that

$$(\mathcal{V}' \cap \mathcal{W}) + (\mathcal{V}'' \cap \mathcal{W}) \subseteq (\mathcal{V}' + \mathcal{V}'') \cap \mathcal{W}.$$

Proof.

Pick some

$$x \in (\mathcal{V}' \cap \mathcal{W}) + (\mathcal{V}'' \cap \mathcal{W}) : x = u + v$$

where $u \in \mathcal{V}' \cap \mathcal{W}$ and $v \in \mathcal{V}'' \cap \mathcal{W}$.

Then,

$$u \in \mathcal{V}', u \in \mathcal{W}$$

and

$$v \in \mathcal{V}'', v \in \mathcal{W}.$$

Now, since $u, v \in \mathcal{W}$ and that $\mathcal{W}$ is a subspace,

$$x = u + v \in \mathcal{W}.$$

Also, since

$$u \in \mathcal{V}' \text{ and } v \in \mathcal{V}''$$

then

$$u + v \in \mathcal{V}' + \mathcal{V}''.$$

Hence,

$$x \in W \cap (\mathcal{V}' + \mathcal{V}''). \qquad \square$$

Does the reverse inclusion always hold?

Proof.

Pick some $y \in (\mathcal{V}' + \mathcal{V}'') \cap W$ so that

$$y \in (\mathcal{V}' + \mathcal{V}'') \textbf{ or } y \in \mathcal{W}.$$

Now, write $y = y' + y''$ with $y' \in \mathcal{V}'$ and $y'' \in \mathcal{V}''$. Observe that,

$$\begin{aligned} &\text{for } y' \in \mathcal{W} \text{ so that } y' \in \mathcal{V}' \cap \mathcal{W} \\ &\not\Rightarrow y'' \in \mathcal{W} \text{ so that } y'' \in \mathcal{V}'' \cap \mathcal{W} \end{aligned}$$

and so,

$$y = y' + y'' \notin (\mathcal{V}' \cap \mathcal{W}) + (\mathcal{V}'' \cap \mathcal{W}).$$

Hence no. $\square$

Example 8.8. Consider the subspaces $\mathcal{V}', \mathcal{V}'' \leq \mathcal{V}$ and suppose that $\mathcal{V}' \cup \mathcal{V}''$ is also a subspace of $\mathcal{V}$. Then show that either $\mathcal{V}' \subseteq \mathcal{V}''$ or $\mathcal{V}'' \subseteq \mathcal{V}'$.

Proof.

Let $x \in \mathcal{V}'$ and $y \in \mathcal{V}''$ so that $x \in \mathcal{V}' \cup \mathcal{V}''$ and $y \in \mathcal{V}' \cup \mathcal{V}''$.

Now, since $\mathcal{V}' \cup \mathcal{V}''$ is a subspace then,

$$x + y \in \mathcal{V}' \cup \mathcal{V}''$$

so,

$$x + y \in \mathcal{V}' \textbf{ or } x + y \in \mathcal{V}''.$$

Case i.) If $x + y \in \mathcal{V}'$ while $x \in \mathcal{V}'$ and that, since $\mathcal{V}'$ is a subspace then, $y \in \mathcal{V}'$. So,

$$y \in \mathcal{V}'' \implies y \in \mathcal{V}'$$
$$\mathcal{V}' \subseteq \mathcal{V}''.$$

Case ii.) If $x + y \in \mathcal{V}''$ while $y \in \mathcal{V}''$ and that, since $\mathcal{V}''$ is a subspace then, $x \in \mathcal{V}''$. So,

$$x \in \mathcal{V}' \implies x \in \mathcal{V}''$$
$$\mathcal{V}'' \subseteq \mathcal{V}'.$$

□

Remark. Observe that, in general, unions of subspaces are not necessarily subspaces.

Example 8.9. Consider the subspaces $\mathcal{W}, \mathcal{W}' \leq \mathcal{V}$ of some vector space $\mathcal{V}$. Show that, if $\mathcal{W}''$ is any subspace of $\mathcal{V}$ containing $\mathcal{W}$ and $\mathcal{W}'$ then $\mathcal{W} + \mathcal{W}' \subseteq \mathcal{W}''$.

Proof.

If,

$$\mathcal{W} \cup \mathcal{W}' \subseteq \mathcal{W}'' \text{ with } \mathcal{W}'' \leq \mathcal{V}$$

then,

$$\mathcal{W} \subseteq \mathcal{W}'' \text{ and } \mathcal{W}' \subseteq \mathcal{W}''.$$

Now, suppose $x = x' + x''$ with $x' \in \mathcal{W}$ and $x'' \in \mathcal{W}'$. Then we have,

$$x \in \mathcal{W} \cup \mathcal{W}'$$

and so

$$x \in \mathcal{W}''.$$

Hence,

$$\mathcal{W} + \mathcal{W}' \subseteq \mathcal{W}''.$$

□

9. Vector space Homomorphisms

Definition 9.1 (Vector space Homomorphism)**.** A vector space homomorphism is the structure preserving mapping between two vector spaces. That is, for some vector spaces $\mathcal{U}$ and $\mathcal{V}$ over common field $\mathbb{F}$ with morphism $\varphi : \mathcal{U} \to \mathcal{V}$ we have,

$$\varphi(u_1 + \lambda u_2) = \varphi(u_1) + \lambda\varphi(u_2) : u_1, u_2 \in \mathcal{U} \text{ and } \lambda \in \mathbb{F}.$$

Remark. A vector space homomorphism preserves linearity and so we typically call the homomorphism a linear map.

Definition 9.2 (Kernel)**.** Consider the linear map $\varphi : \mathcal{U} \to \mathcal{V}$, the *kernel* of φ, denoted $\ker(\varphi)$ is given by:

$$\ker(\varphi) = \{u \in \mathcal{U} : \varphi(u) = \mathbf{0} \in \mathcal{V}\}.$$

The set or space of elements inside the kernel is thus called the *null space.*

Notice that the null space is a subspace of the linear map where any elements inside this subspace are taken to the trivial vector space of $\mathcal{V} = \mathbf{0}$ under the linear map.

Example 9.3. Consider the linear map $T : \mathbb{R}^3 \to \mathbb{R}^2$ where,

$$T = \begin{pmatrix} 1 & 3 & 2 \\ 1 & 2 & 5 \end{pmatrix}.$$

To find the kernel, $\ker(T)$, we must use the definition and hence solve, $T(\mathbf{x}) = \mathbf{0}$ for some $\mathbf{x} \in \mathbb{R}^3$. Hence, by Gaussian elimination augmenting $(T|\mathbf{0})$, we have;

$$\ker(T) = \{\lambda \begin{pmatrix} -11 \\ 3 \\ 1 \end{pmatrix} : \lambda \in \mathbb{F}\}$$

Definition 9.4 (Nullity)**.** The *nullality* is the dimension of the null space or kernel.

Remark. Note that the nullity gives a measure of the injectivity of the linear map φ.

Definition 9.5 (Image)**.** For some morphism $\varphi : \mathcal{U} \to \mathcal{V}$ the *image* is defined as,

$$img(\varphi) = \varphi(\mathcal{U}) = \{v \in \mathcal{V} : v = \varphi(u) \text{for some} u \in \mathcal{U}\}.$$

Definition 9.6 (Preimage)**.** For some morphism $\varphi : \mathcal{U} \to \mathcal{V}$ the *preimage* is defined as,

$$\overleftarrow{\varphi}(\mathcal{V}) = \{u \in \mathcal{U} : \varphi(u) \in \mathcal{V}\}$$

Remark. The preimage can be defined even when no inverse morphism exists.

Theorem 9.7. *Let $\mathcal{U}$ and $\mathcal{V}$ be vector spaces over some common field $\mathbb{F}$ and some linear map $\varphi : \mathcal{U} \to \mathcal{V}$ then*

$$\dim(img(\varphi)) + \dim(\ker(\varphi)) = \dim(\mathcal{U})$$

Definition 9.8 (Injective)**.** If, for the linear map $\varphi : \mathcal{U} \to \mathcal{V}$, we have

$$\varphi(u_1) = \varphi(u_2) \implies u_1 = u_2 \, \forall u_1, u_2 \in \mathcal{U}$$

then φ is said to be *injective.*

Definition 9.9 (Surjective)**.** If, for the linear map $\varphi : \mathcal{U} \to \mathcal{V}$, we have

$$\forall v \in \mathcal{V} \exists u \in \mathcal{U} : \varphi(u) = v$$

then φ is said to be *surjective*.

Definition 9.10 (Bijective)**.** If the linear map $\varphi : \mathcal{U} \to \mathcal{V}$ is both injective and subjective then φ is said to be *bijective*, or, isomorphic in the context vector space homomorphisms.

9.1. **Properties of linear maps.** One may note that a linear map can always be represented in the form of a matrix and as such the determinant can be taken given that linear maps are from vector spaces to vector spaces over a common field. Hence, the conceptual question arises, geometrically, what exactly is the determinant? It turns out the determinant is a *measure* of dilation much like how scaler multiplies of the unit scalar dilate it. That is, $3 * 1 = 3$ is the dilation of the unit scalar 1 by a measure of 3. Now, recall that a field is a vector space over itself and so we may use this to generalise our intuition from above. Thus, we may view a linear map as a linear dilation of some vector in one vector space to another vector space and the determinant as the measure of dilation of this mapping.

TODO.. motivate to prove the determinant definition here.. and motiave to the charteristic polynomial and eigen spaces..

10. Direct Sum

Consider two subspaces $\mathcal{U}, \mathcal{V} \leq \mathcal{W}$ of $\mathcal{W}$. Then $\mathcal{U}$ and $\mathcal{V}$ are said to be **complementry** if every vector $w \in \mathcal{W}$ has a *unique* decomposition $w = u + v : u \in \mathcal{U}$ and $v \in \mathcal{V}$. The vector space $\mathcal{W}$ is then said to be the **internal direct sum** of subspaces $\mathcal{U}$ and $\mathcal{V}$, written $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$.

Definition 10.1 (Direct sum)**.** Consider two vector spaces $\mathcal{U}$ and $\mathcal{V}$. Then the sum $\mathcal{U} + \mathcal{V}$ is called *direct* if and only if $\mathcal{U} \cap \mathcal{V} = \{\mathbf{0}\}$, written $\mathcal{U} \oplus \mathcal{V}$.

Definition 10.2 (Complementary subspaces)**.** Both $\mathcal{U}$ and $\mathcal{V}$ are complementary subspaces of $\mathcal{W}$ if and only if,

- $\mathcal{W} = \mathcal{U} + \mathcal{V}$ and,
- $\mathcal{U} \cap \mathcal{V} = \{\mathbf{0}\}$.

That is, if $\mathcal{U}, \mathcal{V} \leq \mathcal{W}$ we have that $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$.

Definition 10.3 (External direct sum)**.** Consider two arbitrary vector spaces $\mathcal{U}$ and $\mathcal{V}$ over some field $\mathbb{F}$. We may define their direct sum constructively by setting $\mathcal{U} \oplus \mathcal{V} = \mathcal{U} \times \mathcal{V} = \{(u, v) : u \in \mathcal{U}, v \in \mathcal{V}\}$ with vector addition and scalar multiplication defined by:

- $(u, v) + (u', v') = (u + u', v + v')$ and
- $\lambda(u, v) = (\lambda u, \lambda v) : \lambda \in \mathbb{F}$.

Consider two vector spaces $\mathcal{U}$ and $\mathcal{V}$ with the linear map $\varphi : \mathcal{U} \to \mathcal{V}$ and the short exact sequence

$$\mathcal{U} \xrightarrow{\alpha} \mathcal{U}/\ker(\varphi) \xrightarrow{\gamma} \operatorname{im}(\varphi) \xrightarrow{\beta} \mathcal{V}.$$

In particular, observe that the map $\alpha : u \mapsto (u, 0)$ and $\beta : v \mapsto (0, v)$ are injective. While the maps $\alpha^{'} : (u, 0) \mapsto u$ and $\beta^{'} : (0, v) \mapsto v$ are subjective. That is γ is isomorphic. Hence we have the following commutative diagram:

Observe that the internal and external direct sums are isomorphic in this case.

We can now write higher dimensional vector spaces are direct sums of one dimensional subspaces by the natural isomorphism outlined above.

Example 10.4. Let $\mathbb{R}_1 = \{(\mathbf{x}_1, \mathbf{0}) : \mathbf{x}_1 \in \mathbb{R}\}$ and $\mathbb{R}_2 = \{(\mathbf{0}, \mathbf{x}_2) : \mathbf{x}_2 \in \mathbb{R}\}$. Then we see that $\mathbb{R}^2 = \mathbb{R}_1 \oplus \mathbb{R}_2 \cong \mathbb{R} \oplus \mathbb{R}$.

More generally, consider the subspace $\mathbb{R}_i$ with vectors of the form

$$\mathbb{R}_i = \{(\mathbf{0}_1, \ldots, \mathbf{0}_{i-1}, \mathbf{x}_i, \mathbf{0}_{i+1}, \ldots, \mathbf{0}_n) : \mathbf{x}_i \in \mathbb{R}\}.$$

Then we have the following:

$$\begin{aligned}\mathbb{R}^n &= \bigoplus_{i=1}^{n} \mathbb{R}_i \\ &= \mathbb{R}_1 \oplus \mathbb{R}_2 \oplus \cdots \oplus \mathbb{R}_n \\ &\cong \underbrace{\mathbb{R} \oplus \cdots \oplus \mathbb{R}}_{\text{n times}} \\ &= \bigoplus_{i=1}^{n} \mathbb{R}.\end{aligned}$$

A subspace of $\mathbb{R}^n$ that is isomorphic to $\mathbb{R}_i$ is called the *canonical copy* of $\mathbb{R}_i$ in $\mathbb{R}$.

Example 10.5.

Let,

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$B = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \text{ and } C \qquad = (1).$$

Then,

$$A \oplus B \oplus C = \begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{pmatrix}.$$

Remark. Note that direct sums are *associative*, that is, $(A \oplus B) \oplus C = A \oplus (B \oplus C)$.

11. Invariant Subspaces

We begin our treatment of invariance of a linear map by considering some linear map $T : \mathcal{V} \to \mathcal{V}$ and asking which non-zero vectors are dilated **only**? That is, for which $\mathbf{x} \in \mathcal{V}$ does $T(\mathbf{x}) = \lambda \mathbf{x}$ for some scalar $\lambda \in \mathbb{F}$. We define this as the eigen-invariance of the linear map T where we call, $\lambda \in \mathbb{F}$ the *eigenvalue* and $\mathbf{x} \in \mathcal{V}$ the corresponding eigenvector. The word *eigen* is German for *self* or intrinsic and so the eigenvectors span the intrinsic bases of the linear map. The set of eigenvalues of the linear map are called the *spectrum* of the linear map. It is these eigenvalues the fully characterise the map, and for any two linear maps that share the same set of eigenvalues are therefore equivalent under eigen-invariance. You will see why this is the case as we progress.

Definition 11.1 (Characteristic polynomial)**.** Consider some matrix $A \in \mathcal{M}_{n,n}(\mathbb{F})$ with the eigenvector $\mathbf{x}$ and eigenvalue $\lambda \in \mathbb{F}$ relation,

$$A\mathbf{x} = \lambda \mathbf{x} : \mathbf{x} \neq \mathbf{0}$$

then we have,

$$\begin{aligned} A\mathbf{x} - \lambda \mathbf{x} &= \mathbf{0} \\ \Rightarrow (A - \lambda I_n)\mathbf{x} &= \mathbf{0}. \end{aligned}$$

Hence, the characteristic polynomial is given by

$$\chi_A(\lambda) = \left| A - \lambda I_n \right|.$$

Theorem 11.2 (Clayley-Hamilton)**.** *For any matrix $A \in \mathcal{M}_{n,n}(\mathbb{F})$ with characteristic polynomial $\chi_A(\lambda) : \lambda \in \mathbb{F}$. Then A satisfies $\chi_A(\lambda)$, that is, $\chi_A(A) = 0$.*

Example 11.3. Suppose some matrix $A \in \mathcal{M}_{n,n}(\mathbb{F})$ is *idempotent*, that is $A^2 = A$. Then show that each eigenvalue of A is either zero or one.

Proof.

Suppose that,

$$A\mathbf{x} = \lambda \mathbf{x} : \mathbf{x} \neq \mathbf{0}$$

we see that,

$$\begin{aligned} A^2\mathbf{x} &= \lambda^2\mathbf{x} \\ \Rightarrow A\mathbf{x} &= \lambda^2\mathbf{x} \\ &= \lambda\mathbf{x} \\ \Rightarrow \lambda &= \lambda^2 \text{ iff } \lambda \in \{0, 1\}. \end{aligned}$$ □

Example 11.4. Suppose some matrix $A \in \mathcal{M}_{n,n}(\mathbb{F})$ is *nilpotent*, that is $A^k = \mathbf{0}$ for some positive integer k. Then show that all the eigenvalues of A are zero.

Proof.

Suppose that,

$$A\mathbf{x} = \lambda\mathbf{x} : \mathbf{x} \neq \mathbf{0}$$

by induction we see that,

$$\begin{aligned} A^k\mathbf{x} &= \lambda^k\mathbf{x} \\ \Rightarrow \mathbf{0} &= \lambda^k\mathbf{x} \\ \Rightarrow \lambda^k &= 0 \\ \text{and so, } \lambda &= 0 \text{ for every } \lambda \in \mathbb{F}. \end{aligned}$$ □

Definition 11.5 (Eigen Subspace)**.** For some matrix $A \in \mathcal{M}_{n,n}$. Define a eigen-subspace $\mathcal{E}_\lambda$ for every eigenvalue λ as the space that contains all the eigenvectors for that particular eigenvalue. That is,

$$\mathcal{E}_\lambda(k) = \ker(A - \lambda I_n)^k.$$

Definition 11.6 (Spectrum)**.** The *spectrum* of a linear operator $A \in \mathcal{M}_{n,n}(\mathbb{C})$ is the set of all its eigenvalues, denoted by $\sigma(A)$. That is,

$$\sigma(A) = \{\lambda \in \mathbb{C} : A\mathbf{x} = \lambda\mathbf{x} : \mathbf{x} \neq \mathbf{0}\}.$$

Definition 11.7 (Spectral radius)**.** The *spectral radius* of a linear operator $A \in \mathcal{M}_{n,n}(\mathbb{C})$ is the radius of the smallest disc centered at the origin in the complex plane that includes all the eigenvalues of A, denoted by $\rho(A)$. That is,

$$\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}.$$

Definition 11.8 (T-invariant subspace)**.** For a vector space $\mathcal{V}$ consider the linear endomorphism $T : \mathcal{V} \to \mathcal{V}$. Then, for some subspace $\mathcal{W} \leq \mathcal{V}$, $\mathcal{W}$ is said to be T-invariant if $T(\mathcal{W}) \subseteq \mathcal{W}$.

Example 11.9. Show that,

$$W = span\left\{\begin{pmatrix}1\\1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\0\\2\\0\end{pmatrix}\right\} \text{ and } W' = span\left\{\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\-1\\0\end{pmatrix}\right\}$$

are T-invariant subspaces of the linear map:

$$T = \begin{pmatrix}4 & -2 & -1 & -1\\3 & -1 & -1 & -1\\-2 & 2 & 2 & 0\\1 & -1 & 0 & 1\end{pmatrix}.$$

Observe that,

$$T(W) = span\left\{\begin{pmatrix}1\\1\\0\\1\end{pmatrix}, \begin{pmatrix}2\\1\\2\\1\end{pmatrix}\right\} \subseteq W$$

and that

$$T(W') = span\left\{\begin{pmatrix}2\\2\\0\\0\end{pmatrix}, \begin{pmatrix}3\\3\\-2\\0\end{pmatrix}\right\} \subseteq W'$$

and so, both W and W' are T-invariant subspaces.

12. Primary decomposition

Theorem 12.1 (Bezout's theorem)**.** *Let $f(x), g(x) \in \mathbb{F}[x]$ be polynomials such that* $\gcd(f(x), g(x)) = \alpha \in \mathbb{F}$*. Then there exists some $a(x), b(x) \in \mathbb{F}[x]$ such that*

$$a(x)f(x) + b(x)g(x) = 1.$$

Proof. TODO.. □

Theorem 12.2 (Primary decomposition)**.** *Let $f(x), g(x) \in \mathbb{F}[x]$ such that* $\gcd(f(x), g(x)) = \alpha \in \mathbb{F}$ *and let $T : \mathcal{V} \to \mathcal{V}$ be a linear map such that $f(T).g(T) = 0$. Then,*

$$\mathcal{V} = \ker f(T) \oplus \ker g(T).$$

Proof.

Pick some $a(x), b(x) \in \mathbb{F}[x]$. Then by Bezout's theorem we have

$$a(x)f(x) + b(x)g(x) = 1$$

and so we prove in both directions as follows.

First we show that $\ker f(T) + \ker g(T)$ is direct by showing that $\ker f(T) \cap \ker g(T) = \{\mathbf{0}\}$.

Suppose we fix some $\mathbf{v} \in \ker f(T) \cap \ker g(T)$. Then

$$\begin{aligned} \mathbf{v} &= (a(T)f(T) + b(T)g(T))\,\mathbf{v} \\ &= a(T)f(T)\mathbf{v} + b(T)g(T)\mathbf{v} \\ &= 0 \end{aligned}$$

and so we have

$$\ker f(T) \cap \ker g(T) = \{\mathbf{0}\}$$

hence $\ker f(T) + \ker g(T)$ is direct.

Conversely we now wish to show that $\ker f(T) \oplus \ker g(T) = \mathcal{V}$.

Fix some $\mathbf{v} \in \mathcal{V}$ so that

$$a(T)f(T)\mathbf{v} + b(T)g(T)\mathbf{v} = \mathbf{v}$$

and so

$$a(T)f(T)\mathbf{v} + b(T)g(T)\mathbf{v} = \mathbf{0}$$

if and only if

$$a(T)f(T)\mathbf{v} \in \ker g(T)$$

and

$$b(T)g(T)\mathbf{v} \in \ker f(T).$$

So we show that

$$a(T)f(T)\mathbf{v} \in \ker g(T)$$

by

$$\begin{aligned} g(T)\left(a(T)f(T)\mathbf{v}\right) &= \left(f(T).g(T)\right)a(T)\mathbf{v} \\ &= (0)a(T)\mathbf{v} \\ &= \mathbf{0} \end{aligned}$$

and that

$$b(T)g(T)\mathbf{v} \in \ker f(T)$$

by

$$\begin{aligned} f(T)\left(b(T)g(T)\mathbf{v}\right) &= \left(f(T).g(T)\right)b(T)\mathbf{v} \\ &= (0)b(T)\mathbf{v} \\ &= \mathbf{0}. \end{aligned}$$

□

Theorem 12.3 (Generalised eigensubspace decomposition)**.** *Let $T : \mathcal{V} \to \mathcal{V}$ be linear with finite dimension* $\dim \mathcal{V} < \infty$ *and consider the characteristic polynomial*

$$\chi_T(\lambda) = \prod_{i=1}^{r} (\lambda - \lambda_i)^{a_i}$$

where each eigenvalue λ_i is distinct with corresponding algebraic multiplicity a_i. Then we have the T-invariant direct sum

$$\mathcal{V} = \bigoplus_{i=1}^{r} \mathcal{E}_{\lambda_i}(a_i).$$

Proof. TODO.. □

Proposition 12.4 (Linear independence modulo a subspace)**.** *Consider the vector space $\mathcal{V}$ and let $\mathcal{W} \leq \mathcal{V}$. Observe that the set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_m : \mathbf{v}_i \in \mathcal{V}\}$ is linearly independent* modulo $\mathcal{W}$ *if and only if the set S is linear independent and spans $\mathcal{W}^{\perp}$ in $S + \mathcal{W}$. That is, the sum $S + \mathcal{W}$ is direct.*

Example 12.5. Suppose that we are given some subspace $\mathcal{W} \leq \mathcal{V}$ and some linearly independent set $S = \{\mathbf{v}, \ldots, \mathbf{v}_m : \mathbf{v}_i \in \mathcal{V}\}$ modulo $\mathcal{W}$. We wish to prove that $m \leq \dim \mathcal{V} - \dim \mathcal{W}$.

Proof.

Notice that,

$$\mathcal{W} \cap S = \{\mathbf{0}\}$$
$$\implies \mathcal{W} \oplus S \leq \mathcal{V}$$

and so,

$$\dim \mathcal{W} + \dim S \leq \dim \mathcal{V}$$
$$\implies \dim \mathcal{W} + m \leq \dim \mathcal{V}$$

hence,

$$m \leq \dim \mathcal{V} - \dim \mathcal{W}.$$ □

13. Similarity

For some matrix $A \in \mathcal{M}_{n,n}(\mathbb{C})$ with $A : \mathbb{C}^n \to \mathbb{C}^n$. The similarity transformation of the linear map A is a representation into another basis on $\mathbb{C}^n$.

Hence we study similarity in order to study the properties which are intrinsic to a linear transformation and so common to all the various basis representations.

Definition 13.1 (Similarity)**.** A matrix $B \in \mathcal{M}_{n,n}(\mathbb{F})$ is said to be *similar* to a matrix $A \in \mathcal{M}_{n,n}(\mathbb{F})$ if there exists a nonsingular matrix $S \in \mathcal{M}_{n,n}$ such that

$$B = S^{-1}AS.$$

The transformation $A \mapsto S^{-1}AS$ is called a *similarity transformation* by the similarity matrix S.

Proposition 13.2. *Similarity is an* equivalence relation *on $\mathcal{M}_{n,n}$; that is; similarity is*

- *reflexive: $A \sim A$,*
- *symmetric: $B \sim A \implies A \sim B$; and*
- *transitive: $C \sim B$ and $B \sim A \implies C \sim A$.*

Proof. For a matrix $B \in \mathcal{M}_{n,n}(\mathbb{F})$ and nonsingular change of basis matrix S. We wish to prove that the similarity transformation $B = S^{-1}AS$ is a equivalence relation, $B \sim A$.

First notice that,

$$A = I^{-1}AI$$

and so,

$$A \sim A.$$

Next observe that,

$$\begin{aligned} B \sim A &\implies B = S^{-1}AS \\ &\implies SB = AS \\ &\implies A = SBS^{-1} \\ &\implies A = S'^{-1}BS' \text{ with } S' = S^{-1} \\ &\implies A \sim B \end{aligned}$$

and so,

$$B \sim A \implies A \sim B.$$

Finally, consider

$$C = S^{-1}BS \text{ and } B = S^{-1}AS$$

and then observe that,

$$\begin{aligned} C \sim B &\implies C = S^{-1}BS \\ &\implies C = S^{-1}\left(S^{-1}AS\right)S \\ &= (S^{-1})^2A(S)^2 \end{aligned}$$

now note $S^{-1}S^{-1}SS = S^{-1}S = I \Leftrightarrow (S^{-1})^2 = (S^2)^{-1}$ hence,

$$C = S'^{-1}AS' \text{ with } S' = S^2$$

and so,

$$C \sim B \text{ and } B \sim A \implies C \sim A. \qquad \square$$

Theorem 13.3. *Let $A, B \in \mathcal{M}_{n,n}(\mathbb{F})$. If B is similar to A, then the characteristic polynomial of B is the same as that of A.*

Proof.

For any λ we have

$$\begin{aligned}
\chi_B(\lambda) &= \det(B - \lambda I) \\
&= \det\left(S^{-1}AS - \lambda S^{-1}IS\right) \\
&= \det\left(S^{-1}(A - \lambda I)S\right) \\
&= \det(S^{-1})\det(A - \lambda I)\det(S) \\
&= \det(S)^{-1}\det(S)\det(A - \lambda I) \\
&= \det(A - \lambda I) = \chi_A(\lambda).
\end{aligned}$$

□

Corollary. *If $A, B \in \mathcal{M}_{n,n}(\mathbb{F})$ and if A and B are similar, then they have the same eigenvalues, counting multiplicity.*

14. Jordan Chains

Definition 14.1 (Jordan chain)**.** Consider the linear map $T : \mathcal{V} \to \mathcal{V}$ and some nilpotent matrix $N = T - \lambda I$ for some eigenvalue in the spectrum of T $\lambda \in \sigma(T)$. Then let $\mathbf{v} \in (\mathcal{E}_\lambda(n) - \mathcal{E}_\lambda(n-1))$ so that $N^n\mathbf{v} = \mathbf{0}$ and that $N^{n-1}\mathbf{v} \neq \mathbf{0}$.

A *Jordan chain* of length n is a row matrix $C \in \mathcal{V}$ of the form

$$C = \begin{pmatrix} N^{n-1}\mathbf{v} & N^{n-2}\mathbf{v} & \dots & \mathbf{v} \end{pmatrix}.$$

The *Jordan chain space* associated to $C : \mathbb{F}^n \to \mathcal{V}$ is $\operatorname{im} C \leq \mathcal{V}$.

Proposition 14.2. *Suppose*

$$C = \begin{pmatrix} \mathbf{v}_{n-1} & \mathbf{v}_{n-2} & \dots & \mathbf{v}_0 \end{pmatrix}$$

is some Jordan chain. Then the Jordan chain space, $\operatorname{im} C$*, is T-invariant.*

15. Jordan Canonical Form

The Jordan (normal) canonical form is a set of "almost diagonal" matrices, called the *Jordan blocks*. The Jordan blocks contain the diagonal components composed of eigenvalues. We call this form, the *block diagonal form.* The set of Jordan blocks form a *equivalence class* under similarity invariance of the eigen-invariant subspaces.

Definition 15.1 (Jordan block)**.** A Jordan block $\mathcal{J}_k(\lambda)$ is a $k \times k$ *upper triangular* square matrix of the form:

$$\mathcal{J}_k(\lambda) = \begin{pmatrix} \lambda & 1 & & & \\ 0 & \lambda & 1 & & \Large 0 \\ & & \ddots & \ddots & \\ \Large 0 & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}$$

where there are $k-1$ $(+1)$'s in the *superdiagonal* and k of a particular scalar eigenvalue λ in the main diagonal while all other entries are zero. The trivial 1×1 Jordan block of λ is $\mathcal{J}_1(\lambda) = (\lambda)$.

Definition 15.2 (Jordan matrix)**.** A Jordan matrix $\mathcal{J} \in \mathcal{M}_{n,n}(\mathbb{F})$ is a direct sum (10) of Jordan blocks. That is,

$$\mathcal{J} = \bigoplus_{i=1}^{N} \mathcal{J}_{\lambda_i, m_i} : 1 \leq i \leq N.$$

where each $\lambda_i \in \sigma(A)$ for some matrix $A \in \mathcal{M}(\mathbb{F})$ need not be distinct.

Example 15.3.

Suppose,

$$A = \begin{pmatrix} 3 & 1 & 1 \\ -4 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We wish to find the Jordan canonical form so that we have the usual change of basis given by, $C\mathcal{J} = AC$. First we find the characteristic polynomial,

$$\begin{aligned} \chi_A(\lambda) &= (3-\lambda)(\lambda+1)^2 - 4(\lambda+1) \\ &= -(\lambda+1)\left\{(\lambda+1)(\lambda-3)+4\right\} \\ &= -(\lambda+1)(\lambda^2 - 2\lambda + 1) \\ \Rightarrow \chi_A(\lambda) &= (\lambda+1)(\lambda-1)^2 \end{aligned}$$

and so, $\lambda_1 = -1$ and $\lambda_2 = 1$ with algebraic multiplicity two. Next recall that,

$$\mathcal{E}_\lambda(n) = \ker(A - \lambda I)^n$$

and, in particular, that

$$\text{geometric multiplicity} = \dim \mathcal{E}_\lambda(1).$$

For λ_1 we recall that the geometric multiplicity is always less than or equal to the algebraic multiplicity and since λ_1 has algebraic multiplicity of one it has geometric multiplicity of one, hence has the trivial Jordan block:

$$\mathcal{J}_1(\lambda_1) = \mathcal{J}_1(-1).$$

To find the corresponding eigenvector to λ_1 we look in the span of the kernel as follows

$$\begin{aligned}
\mathbf{v}_1 \in \mathcal{E}_{\lambda_1}(1) = \mathcal{E}_{-1}(1) \\
&= \ker(A+I)^1 \\
&= \ker \begin{pmatrix} 4 & 1 & 1 \\ -4 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
&= \ker \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} && \text{(after a row reduction)} \\
\Rightarrow \mathbf{v}_1 \in \mathit{span} \left\{ \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}.
\end{aligned}$$

Note that it is important we find $\mathbf{v}_1$ first as we shall see. Now, see that

$$\begin{aligned}
\mathcal{E}_{\lambda_2}(1) = \mathcal{E}_1(1) &= \ker(A-I)^1 \\
&= \ker \begin{pmatrix} 2 & 1 & 1 \\ -4 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \\
&= \ker \begin{pmatrix} 2 & 1 & 1 \\ -2 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\
&= \ker \begin{pmatrix} 0 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} && \text{(after row reductions.)}
\end{aligned}$$

Hence, $\dim \mathcal{E}_{\lambda_2}(1) = 2$ and so we have the corresponding Jordan block,

$$\mathcal{J}_2(\lambda_2) = \mathcal{J}_2(1).$$

Now, to find the corresponding eigenvectors to λ_2 we need only pick a vector **not** in the span of $\mathcal{E}_{\lambda_2}(1)$ and is linearly independent from $\mathbf{v}_1$. Consider the vector $\mathbf{v}_2 \notin \mathcal{E}_{\lambda_2}(1)$ with,

$$\mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Then by the Jordan chain we see that,

$$\begin{aligned}\mathbf{v}_3 &= \mathcal{E}_{\lambda_2}(1)\mathbf{v}_2 \\ &= \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \in span\mathcal{E}_{\lambda_2}(1)\end{aligned}$$

thus, $\mathcal{E}_{\lambda_2}(1)\mathbf{v}_3 = \mathbf{0}$ and since the geometric multiplicity of λ_2 was two we now have the spanning set of corresponding eigenvectors $\mathbf{v}_2, \mathbf{v}_3$. So,

$$\begin{aligned} C\mathcal{J} &= AC \\ \Rightarrow A &= C\mathcal{J}C^{-1} \\ \Rightarrow A &= C\mathcal{J}_1(\lambda_1) \oplus \mathcal{J}_2(\lambda_2)C^{-1} \\ A &= \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & -2 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & -2 \\ -1 & 0 & 0 \end{pmatrix}^{-1} . \end{aligned}$$

16. Algebras

Definition 16.1 (Algebras)**.** A finite-dimensional (associative) *algebra* $\mathcal{A}$ over a field $\mathbb{F}$ is a finite-dimensional vector space over $\mathbb{F}$ equipped with a law of composition, that is, a mapping (multiplication) $(a, b) \to ab$ from $\mathcal{A} \times \mathcal{A}$ into $\mathcal{A}$ which satisfies

- $(ab)c = a(bc)$ (associativity),
- $a(b + c) = ab + ac$,
- $(a + b)c = ac + bc$,
- $\lambda(ab) = (\lambda a)b = a(\lambda b)$,

for $\lambda \in \mathbb{F}$ and $a, b, c \in \mathcal{A}$.

Remark. A algebra $\mathcal{A}$ is called *unital* if there exists $1 \in \mathcal{A}$, the *idenity element*, such that $1a = a1 = a \, \forall a \in \mathcal{A}$.

Remark. A algebra $\mathcal{A}$ is *commutative* if $ab = ba \, \forall a, b \in \mathcal{A}$.

17. Dual Spaces

Recall that; If $\mathcal{U}$ and $\mathcal{V}$ are two vector spaces over a field $\mathbb{F}$, then the set of all linear maps from $\mathcal{U} \to \mathcal{V}$, denoted $L(\mathcal{U}, \mathcal{V})$, is a vector space over $\mathbb{F}$.

Definition 17.1. Let $\mathcal{V}$ be a vector space over $\mathbb{F}$. The dual space of $\mathcal{V}$, denoted $\mathcal{V}^*$, is the set of all linear maps from $\mathcal{V}$ to $\mathbb{F}$. That is, $\mathcal{V}^* = L(\mathcal{V}, \mathbb{F})$. The elements of $\mathcal{V}^*$ are called *linear functionals* on $\mathcal{V}$. Frequently, linear functionals on $\mathcal{V}$ are called *covectors* or *1-forms.*

Remark. Since $L(\mathcal{V}, \mathbb{F})$ is a vector space over $\mathbb{F}$ then so too is the dual $\mathcal{V}^*$, with point-wise addition and scalar multiplication. Moreover, if $n = \dim \mathcal{V} < \infty$, then $\dim \mathcal{V}^* = \dim L(\mathcal{V}, \mathbb{F}) = n$.

Example 17.2. Fix $x \in [0,1]$ and define the *evaluation map* $ev_x : \mathcal{C}([0,1]) \to \mathbb{R}$ by $ev_x(f) = f(x)$ for any $f \in \mathcal{C}([0,1])$. Then $ev_x \in (\mathcal{C}([0,1]))^*$.

18. Normed vector space

A norm on a complex vector space $\mathcal{V}$ is a real-valued function $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ that, to any vector $x \in \mathcal{V}$ associates a real number $\|x\|$, satisfies the following:

- $\|x\| \geq 0$, $\forall x \in \mathcal{V}$ and in particular $\|x\| = 0 \Leftrightarrow x = 0$ (positive definite),
- $\|\alpha x\| = |\alpha|\|x\| \, \forall x \in \mathcal{V}$, and $\alpha \in \mathbb{C}$ (uniform scaling),
- $\|x + y\| \leq \|x\| + \|y\| \, \forall x, y \in \mathcal{V}$ (triangle inequality).

18.1. **Linear Operators.** In functional analysis one considers linear maps between normed vector spaces. In this context such linear mappings are termed *linear operators* or just operators. In this light one is motivated to talk about notions of *continuity* of such operators.

For any normed vector space $(\mathcal{U}, \|\cdot\|_{\mathcal{U}})$ consider the closed unit ball in $\mathcal{U}$ given by $B(\mathcal{U}) = \{x \in \mathcal{U} : \|x\|_{\mathcal{U}} \leq 1\}$. Hence, we may now define boundedness of an operator formally as follows:

Definition 18.1 (Bounded linear operator)**.** A linear operator $T : \mathcal{U} \to \mathcal{V}$ between normed vector spaces $(\mathcal{U}, \|\cdot\|_{\mathcal{U}})$ and $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ is called a *bounded* linear operator if it maps the closed unit ball $B(\mathcal{U})$ in $\mathcal{U}$ into a bounded set (a ball) in $\mathcal{V}$.

Corollary. *A linear operator $T : \mathcal{U} \to \mathcal{V}$ between normed vector spaces $(\mathcal{U}, \|\cdot\|_{\mathcal{U}})$ and $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ is continuous if and only if it is bounded.*

19. Inner Product Spaces

Let V be a complex vector space.

Definition 19.1 (Inner product)**.** An *inner product* in V is a mapping $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ satisfying

- Hermitian symmetry: $\langle x, y \rangle = \overline{\langle y, x \rangle} \, \forall x, y \in V$;
- Linearity: $\langle \alpha x_1 + \beta x_2, y \rangle = \alpha \langle x_1, y \rangle + \beta \langle x_2, y \rangle \ \forall x_1, x_2, y \in V$ and $\alpha, \beta \in \mathbb{C}$;
- Positive definite: $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ iff $x = 0$, $\forall x \in V$;

Remark. Note that condition (i) implies that $\langle x, x \rangle \in \mathbb{R}$ and so condition (ii) is consistent.

Remark. Note that condition (ii) defines that the inner product is linear in the first argument. However, the inner product is *conjugate linear* in the second argument as shown by:

Proof.

$$\begin{aligned}
\langle x, \alpha y_1 + \beta y_2 \rangle &= \overline{\langle \alpha y_1 + \beta y_2, x \rangle} \\
&= \overline{\langle \alpha y_1, x \rangle + \langle \beta y_2, x \rangle} \\
&= \bar{\alpha}\overline{\langle y_1, x \rangle} + \bar{\beta}\overline{\langle y_2, x \rangle} \\
&= \bar{\alpha}\langle x, y_1 \rangle + \bar{\beta}\langle x, y_2 \rangle.
\end{aligned}$$

□

A vector space V together with an inner product $\langle \cdot, \cdot \rangle$ is termed an *inner product space.*

Example 19.2. Consider the vector space $V = M_{mn}(\mathbb{F})$. We wish to show that, for some $X, Y \in V$, that $tr(X^*Y)$ defines an inner product in V. Hence we have, $\langle X, Y \rangle \doteq tr(X^*Y)$ for all $X, Y \in V$.

Proof. First suppose that $\langle X, Y \rangle = tr(X^*Y)$ and so;

We check that $tr(X^*Y)$ is Hermitian symmetric:

$$\begin{aligned}
\langle X, Y \rangle &= tr(X^*Y) \\
&= \overline{tr(\overline{X^*Y})} \\
&= \overline{tr(\overline{YX^*})} \\
&= \overline{tr(\overline{(X^*Y)}^T)} \\
&= \overline{tr(Y^*X)} \\
&= \overline{\langle Y, X \rangle}.
\end{aligned}$$

Check that $tr(X^*Y)$ is linear in the first parameters by considering some $\alpha, \beta \in \mathbb{F}$ so that

$$\begin{aligned}
\langle \alpha X_1 + \beta X_2, Y \rangle &= tr((\alpha X_1 + \beta X_2)^*Y) \\
&= tr((\alpha X_1)^*Y + (\beta X_2)^*Y) \\
&= tr((\alpha X_1)^*Y) + tr((\beta X_2)^*Y) \\
&= \bar{\alpha} tr(X_1^*Y) + \bar{\beta} tr(X_2^*Y) \\
&= \bar{\alpha}\langle X_1, Y \rangle + \bar{\beta}\langle X_2, Y \rangle.
\end{aligned}$$

and check that $tr(X^*Y)$ is positive definite for $X = Y$,

$$\langle X, X \rangle = tr(X^*X) \geq 0$$

and when

$$\begin{aligned} X = 0 &\implies X^* = 0 \\ &\implies tr(X^*X) = 0. \end{aligned}$$ □

TODO..

Example 19.3. Let $\mathcal{V}$ be the vector space of continuous real-valued functions $[0,1]$ endowed with the standard inner product. If $f(x) = x$ and $g(x) = e^x$ then find $\langle f, g \rangle$ and $\|f\|$.

For $\langle f, g \rangle$ we have,

$$\begin{aligned} \langle f, g \rangle &= \int_0^1 f(x)\overline{g(x)}\mathrm{d}x \\ &= \int_0^1 xe^x \mathrm{d}x && \text{(by parts we have,)} \\ &= xe^x\|_0^1 - \int_0^1 e^x \mathrm{d}x \\ &= e - (e-1) = 1. \end{aligned}$$

Now, for $\|f\|$ we have,

$$\begin{aligned} \|f\| &= \left(\int_0^1 x^2 \mathrm{d}x\right)^{\frac{1}{2}} \\ &= \sqrt{\frac{1}{x^3}}|_0^1 \\ &= \frac{1}{\sqrt{3}}. \end{aligned}$$

Theorem 19.4. *Let $\mathcal{V}$ be a complex vector space with inner product $\langle \cdot, \cdot \rangle$. Then the inner product induces a norm $\| \cdot \|$ in $\mathcal{V}$ by the definition*

$$\|x\| = \sqrt{\langle x, x \rangle} \text{ for } x \in \mathcal{V}.$$

Proof.

Notice that $\| \cdot \|$ is trivially positive definite as follows:

$$\|x\| = \sqrt{\langle x, x \rangle} \geq 0 \Leftrightarrow x = 0, \forall x \in \mathcal{V}.$$

By linearity we see that,

$$\begin{aligned}
\|\alpha x\| &= \sqrt{\langle \alpha x, \alpha x\rangle} \\
&= \sqrt{\alpha\bar{\alpha}\langle x, x\rangle} \\
&= \sqrt{|\alpha|^2\langle x, x\rangle} = |\alpha|\|x\| \; \forall \alpha \in \mathbb{C} \text{ and } \forall x \in \mathcal{V}.
\end{aligned}$$

Finally the triangle inequality follows as shown,

$$\begin{aligned}
\|x+y\|^2 &= \langle x+y, x+y\rangle \\
&= \|x\|^2 + \|y\|^2 + \langle x, y\rangle + \langle y, x\rangle \\
&= \|x\|^2 + \|y\|^2 + \langle x, y\rangle + \overline{\langle x, y\rangle} \\
&= \|x\|^2 + \|y\|^2 + 2\mathbb{R}[\langle x, y\rangle] \\
&\leq \|x\|^2 + \|y\|^2 + 2|\langle x, y\rangle| \\
&\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \\
&= (\|x\| + \|y\|)^2 \; \forall x, y \in \mathcal{V}.
\end{aligned}$$

□

Definition 19.5 (Isometric)**.** Consider two inner product spaces $\mathcal{U}$ and $\mathcal{V}$ with the map $T : \mathcal{U} \to \mathcal{V}$. Then the map T preserves inner products (*isometric*) if, for any $\mathbf{u}, \mathbf{u'} \in \mathcal{U}$, we have $\langle T(\mathbf{u}), T(\mathbf{u'})\rangle = \langle \mathbf{u}, \mathbf{u'}\rangle$.

Remark. A map need not be linear for it to be isometric.

Theorem 19.6. *Isometric maps are monomorphic.*

Proof. Suppose some inner product preserving maps $T : \mathcal{U} \to \mathcal{V}$. Then it is sufficient to show that the dimension of the kernel of T is zero for injectivity.

Pick some $\mathbf{u} \in \ker T$ so that,

$$\begin{aligned}
\langle T(\mathbf{u}), T(\mathbf{u})\rangle &= \langle 0, 0\rangle \\
&= 0.
\end{aligned}$$

However, since T preserves inner products then

$$\langle \mathbf{u}, \mathbf{u}\rangle = \langle T(\mathbf{u}), T(\mathbf{u})\rangle$$

and so,

$$\langle \mathbf{u}, \mathbf{u} \rangle = 0$$
$$\implies \mathbf{u} = \mathbf{0}$$

hence T is monomorphic.

$\square$

TODO.. MOVE ME..

Example 19.7. Find a basis for the orthogonal complement of $W = span \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix} \right\} \in \mathbb{R}^4$.

Therefore we do the following:

$$\begin{aligned} W^{\perp} &= \ker \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 2 & 1 & 1 \end{pmatrix} \\ &= \ker \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} && \text{(after a row reduction)} \\ &= span \left\{ \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}. \end{aligned}$$

20. QR decomposition

The QR factorisation of a matrix $A \in \mathcal{M}_{n,m}(\mathbb{F})$ is the decomposition into a product of an orthogonal matrix Q and upper triangular matrix R, written $A = QR$. However, first we must look at the Gram-Schmidt algorithm in order to obtain an *orthonomalised babsis* of A.

Definition 20.1 (Projection operator)**.** Consider two $n \times 1$ vectors $\mathbf{u}, \mathbf{v} \in \mathbb{F}^n$. Then the **orthogonal projection** *of* $\mathbf{v}$ *onto* the line spanned by $\mathbf{u}$ is given by,

$$proj_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}.$$

Recall that when the field $\mathbb{F}$ is taken to be real for two $n \times 1$ vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ then the inner product is then $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T \mathbf{v}$. When the field $\mathbb{F}$ is taken to be complex for two $n \times 1$ vectors $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ then the inner product is then $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^* \mathbf{v}$. In fact, since the conjugate of a real number is itself the usual complex version can simply be used without issue.

20.1. **Gram-Schmidt.** The Gram-Schmidt algorithm is a method of orthonormalising a finite set of linearly independent vectors that span an inner product space into a orthonormal basis.

Consider a finite sequence of vectors given by the spanning set,

$$S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}.$$

Then the sequence of Gram-Schmidt orthogonalised vectors, $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is given by,

$$\mathbf{u}_k = \mathbf{v}_k - \sum_{j=1}^{k-1} proj_{\mathbf{u}_k}(\mathbf{v}_k).$$

Hence, we may normalise the sequence by,

$$\mathbf{e}_k = \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}.$$

Example 20.2.

Consider the spanning set,

$$S = \left\{ \mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} \right\}.$$

Let $\mathbf{u}_1 = \mathbf{v}_1$ and so,

21. Quadratic Form

A quadratic or more generally, a *quadric*, form is the linear combination of quadratic terms with scalar coefficients with possible cross terms. More precisely, quadratic forms are in fact homogeneous quadratic polynomials in n variables.

The special case from the quadratic equation,

$$ax^2 + bx + c = 0$$

where $a \in \mathbb{R} : a \neq 0$, $x \in \mathbb{F}$ and fix $b = c = 0$ so to be homogeneous, then it remains that we have the unary quadratic function,

$$q(x) = ax^2$$

This is a primary example of a simple one dimensional quadratic form. It however, turns out more generally that a *quadric* has a more generalised linear algebraic form known as the *quadratic form* which is defined herein.

Definition 21.1 (Quadratic Form)**.** For an $n \times 1$ vector $\mathbf{x}$ and $n \times n$ symmetric matrix A the *quadratic form* is defined as:

$$Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}.$$

Notice that in the special case of the unary quadratic function the x term is a scalar in the field $\mathbb{F}$. However in the more generalised form $\mathbf{x}$ is now a vector with components in some field $\mathbb{F}$ where the vector space is defined on, typically $\mathbb{F} = \mathbb{R}$. Observe that, if the quadratic form $Q(\mathbf{x})$ is defined over some vector space $\mathcal{V}(\mathbb{F})$ then the quadratic form defines a mapping in following way, $Q : \mathcal{V} \to \mathbb{F}$. Such a quadratic mapping motivates the following definition:

Definition 21.2 (Quadratic Space)**.** A *quadratic space* is the pairing $(\mathcal{V}, q)$ of some vector space $\mathcal{V}(\mathbb{F})$ over some field $\mathbb{F}$ and the quadratic map $q : \mathcal{V} \to \mathbb{F}$ defined by the quadratic form $q = Q(\mathbf{v}) = \mathbf{v}^T A\mathbf{v}$ for some symmetric matrix A that is well defined for vectors from the vector space $\mathcal{V}$.

Example 21.3. Consider the vector space $\mathbb{R}^2$ and the quadratic map, $q : \mathbb{R}^2 \to \mathbb{R}$ defined by the square of the Euclidean norm. That is, $q(\mathbf{v}) = \|\mathbf{v}\|^2 = x^2 + y^2$. Hence a Euclidean normed vector space is a *quadratic space*.

In any case, digressing back to the quadratic form consider for example, some of the following motivating cases given,

Example 21.4.

$$q(x) = \alpha x^2 \tag{unary}$$

$$q(x, y) = \alpha x^2 + \beta xy + \gamma y^2 \tag{binary}$$

$$q(x, y, z) = \alpha x^2 + \beta y^2 + \gamma z^2 + axy + bxz + cyz \tag{ternary}$$

By re-representing these forms in terms of the more compact matrix representations we may make use of various linear algebraic tools such as eigen-decomposition.

Remark. The quadratic form is analogous to the notion of *completing the square*. By way of this, we can remember the quadratic form in the geometric context as completing the hypersquares.

Quadratic forms allow us to characterise various geometric curves and surfaces. In particular,

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = c^2 \tag{eclipse}$$

$$\left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = c^2 \tag{hyperbola}$$

Consider then an examples of a quadric form,

Example 21.5. $Q(x,y) = x^2 + 2\sqrt{2}xy + 4y^2$.

It is not yet apparent as to what this surface looks like geometrically. However, by a change of bases we may derive a more familiar representation to characterise the surface. The change of bases we pick will be the normalised orthogonal eigenbasis. The rational for this choose is that the eigenvectors are *stretched only* by a scalar amount (the corresponding eigenvalues), and so are invariant to rotation or reflection under the curve. We are guarantied that the eigenvectors are orthogonal by the spectral theorem when A is symmetric and that all the eigenvalues are real. Moreover, by normalising the eigenvectors we remain with a representation in a new coordinate system that does not 'interfere' with the 'look', if you like, of the curve. This orthonormal basis is called the *principle axes* given in the following theorem.

First recall that,

Lemma 21.6. *If A is a real symmetric matrix, then by the spectral theorem A has all real eigenvalues and the eigenspaces of A are mutually orthogonal and so A is orthonormally diagonalisable.*

Theorem 21.7 (Principal Axis Theorem)**.** *Given the quadratic form, $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ where A is a real symmetric matrix. Then by the above lemma, we have that, in particular by Gram-Schmidt, A is orthogonally diagonalisable into an orthonormal eigenbasis.*

Suppose then that $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ are the eigenvalues of A and $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ the eigenvectors of A. Then let $\{\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \dots, \hat{\mathbf{u}}_n\}$ denote the set of normalised eigenvectors of A and let S be the matrix of the orthogonal eigenvectors normalised,

$$\begin{aligned} S &= \begin{pmatrix} \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} & \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} & \dots & \frac{\mathbf{u}_n}{\|\mathbf{u}_n\|} \end{pmatrix} \\ \Rightarrow S &= \begin{pmatrix} \hat{\mathbf{u}}_1 & \hat{\mathbf{u}}_2 & \dots & \hat{\mathbf{u}}_n \end{pmatrix}. \end{aligned}$$

Then with the substitution $\mathbf{x} = S\mathbf{y}$ we have the principle axes form*:*

$$Q(\mathbf{x}) = \mathbf{y}^T \Lambda \mathbf{y} = \sum_{i=1}^{n} \lambda_i y_i^2.$$

The set of normalised eigenvectors, $\{\hat{\mathbf{u}}_1, \hat{\mathbf{u}}_2, \dots, \hat{\mathbf{u}}_n\}$ is said to be the principle axes *of the quadratic form $Q(\mathbf{x})$.*

Proof. Consider the quadratic form, $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ and suppose that the matrix A is real and symmetric, i.e. $A = A^T$. Suppose then that the set $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ are the eigenvalues of A and

$\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ are the eigenvectors of A. Then let,

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

be the matrix with eigenvalues down the diagonal as so that we may diagonalise A by the change of basis $A = \mathcal{B}\Lambda\mathcal{B}^{-1}$ where $\mathcal{B}$ is the matrix,

$$\mathcal{B} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_n \end{pmatrix}$$

of eigenvectors. Let us now write a matrix S of the orthogonal eigenvectors normalised,

$$S = \begin{pmatrix} \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} & \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} & \ldots & \frac{\mathbf{u}_n}{\|\mathbf{u}_n\|} \end{pmatrix}.$$

Hence we have that,

$$\begin{aligned} Q &= \mathbf{x}^T A\mathbf{x} \\ &= \mathbf{x}^T (S\Lambda S^{-1})\mathbf{x} \\ &= \mathbf{x}^T (S\Lambda S^T)\mathbf{x} \\ &= (S^T\mathbf{x})^T \Lambda(S^T\mathbf{x}). \end{aligned}$$

It remains that,

$$Q = (S^T\mathbf{x})^T \Lambda(S^T\mathbf{x}).$$

Consider now the substitution $\mathbf{x} = S\mathbf{y}$

$$\begin{aligned} \Rightarrow Q &= (S^T(S\mathbf{y}))^T \Lambda(S^T(S\mathbf{y})) \\ &= \mathbf{y}^T(S^TS)\Lambda(S^TS)\mathbf{y} \\ &= \mathbf{y}^T\Lambda\mathbf{y} \\ &= \sum_{i=1}^{n} \lambda_i y_i^2. \end{aligned}$$

□

Lemma 21.8. *For the quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, if all the eigenvalues of A are strictly positive then the quadric $Q(\mathbf{x})$ is an* ellipsoid*. If all the eigenvalues are strictly negative then the quadric is an* imaginary ellipsoid*. If some eigenvalues are negative and some positive then the quadric is said to be an* hyperboloid*. Finally, if there exists at least one eigenvalues of scale zero then the quadric is some kind of* paraboloid *to be determined.*

Example 21.9. Consider again the quadric form, $Q(x, y) = x^2 + 2\sqrt{2}xy + 4y^2$ and rewrite this in the matrix quadratic form in the following way,

$$\begin{aligned} Q(x,y) = Q(\mathbf{x}) &= \mathbf{x}^T A \mathbf{x} \\ &= \mathbf{x}^T \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 4 \end{pmatrix} \mathbf{x} \qquad \text{(division by two much like completing the square.)} \end{aligned}$$

Now we wish to diagonalise the matrix A and so we find the eigenvectors and corresponding eigenvalues of A in the usual way. By direct computation we see that A has eigenvalues $\lambda_1 = 5$ and $\lambda_2 = 2$. Hence, by the principal axes theorem we have that,

$$Q(\mathbf{y}) = 5y_1^2 + 2y_2^2.$$

We may wish to sketch the level curve when $Q = 1$. Consider then the curve $5y_1^2 + 2y_2^2 = 1$ and observe that this is an *ellipse* that intersects with the y_1-axis at $-1/\sqrt{5}$ and $1/\sqrt{5}$ and intersects with the y_2 axis at $1/\sqrt{2}$ and $-1/\sqrt{2}$.

By direct computation we find the eigenvectors are,

$$\mathbf{v}_1 = \begin{pmatrix} \sqrt{2} \\ -1 \end{pmatrix},$$

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}.$$

Recall now the substitution $\mathbf{x} = S\mathbf{y}$ and set the matrix S as the orthogonal eigenvectors normalised in the following way,

$$S = \frac{1}{\sqrt{3}} \begin{pmatrix} \sqrt{2} & 1 \\ -1 & \sqrt{2} \end{pmatrix}.$$

Where $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$. Hence, if we then plot the normalised eigenbasis on the XY-plane and then plot the ellipse on the normalised eigenbasis we may observe that the original quadratic form of $Q(x, y) = x^2 + 2\sqrt{2}xy + 4y^2$ is merely a rotated ellipse in the usual XY-plane. In particular, the matrix S is a rotation matrix of the ellipse on the principal axes (normalised eigenbasis) into the XY-plane.

Finally we can find the closes points to the origin by observing that in the eigen coordinate system, that is the Y_1Y_2-plane, we have the closes points as $\pm\frac{1}{\sqrt{5}}$.

Example 21.10. FIXME: Fix this example!! Given $3x^2 - 10xy + 3y^2 = 24$ find the shortest distance to the origin.

Consider the quadratic form, $Q = \mathbf{x}^T A\mathbf{x}$, and let

$$A = \begin{pmatrix} 3 & -5 \\ -5 & 3 \end{pmatrix}$$

which has eigenvalues $\lambda_1 = 8$ and $\lambda_2 = -2$ with eigen vectors,

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Let us know write a matrix S of the othogonal eigenvectors normalised,

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix}$$

and matrix Λ with eigenvalues down the diagonal as so,

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Hence we have that,

$$\begin{aligned} Q &= \mathbf{x}^T A\mathbf{x} \\ &= \mathbf{x}^T (S\Lambda S^{-1})\mathbf{x} \\ &= (S^T\mathbf{x})\Lambda(S^T\mathbf{x}). \end{aligned}$$

It remains that,

$$\begin{aligned} Q &= (S^T\mathbf{x})\Lambda(S^T\mathbf{x}) \\ &= \frac{1}{2}\begin{pmatrix} (x-y) & (x+y) \end{pmatrix} \begin{pmatrix} 8 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} (x-y) \\ (x+y) \end{pmatrix} \\ &= 4(x-y)^2 - 1(x+y)^2. \end{aligned}$$

Thus we have the hyperboloid $4(x-y)^2 - 1(x+y)^2 = 24$.