Linear Algebra

Hasan Baig

Michaelmas 2021

Contents

1 Vector Spaces, Subspaces

Notation. Let F be an arbitrary field (e.g. $F = \mathbb{R}$ or $\mathbb{C}$)

Definition (F vector space)**.** A **F-vector space** (a vector space over F) is an abelian group $(V, +)$ equipped with a function:

$$F \times V \to V, \quad (\lambda, v) \mapsto \lambda v$$

(λ is a scalar, v is a vector, λv is a vector)
Such that:

- $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$
- $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$
- $\lambda(\mu v) = (\lambda\mu)v$
- $1 \cdot v = v$

We know how to:

- sum two vectors
- multiply a vector $v \in V$ by a scalar $\lambda \in F$.

Examples. (i) $n \in \mathbb{N}, F^n$: column vectors of length n with entries in F

$$v \in F,\ v = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},\ x_i \in F,\ 1 \leq i \leq n$$

$$v + w = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{bmatrix}, \quad \lambda v = \begin{bmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{bmatrix}$$

F^n is an F-vector space.

(ii) $\mathbb{R}^X = \{f : X \to \mathbb{R}\}$ (set of real valued functions on X). We have that $\mathbb{R}^X$ is a $\mathbb{R}$ vector space:
- $(f_1 + f_2)(x) = f_1(x) + f_2(x)$
- $(\lambda f)(x) = \lambda f(x), \lambda \in \mathbb{R}$

(iii) $M_{n,m}(F) \equiv n \times m$ matrices with entries in F

Remark. The axiom of scalar multiplication implies that: $\forall v \in V, 0 \cdot v = 0$

Definition (Subspace)**.** Let V be a vector space over F. The subset U of V is a **vector subspace** of V (noted $U \leq V$) if:

- $0 \in U$
- $(u_1, u_2) \in U \times U \implies u_1 + u_2 \in U$
- $(\lambda, u) \in F \times U \implies \lambda u \in U$

Equivalently:

- $0 \in U$
- $\forall (\lambda_1, \lambda_2) \in F \times F, \forall (u_1, u_2) \in U \times U, \lambda_1 v_1 + \lambda_2 v_2 \in U$

This property means that U is stable by:

- scalar multiplication
- vector addition

Examples. (i) $V = \mathbb{R}^{\mathbb{R}}$ space of functions $\mathbb{R} \to \mathbb{R}$; $\mathcal{C}(\mathbb{R}) \leq V$ space of continuous functions $\mathbb{R} \to \mathbb{R}$; $\mathbb{P}(\mathbb{R}) \leq \mathcal{C}(\mathbb{R})$ space of polynomials

(ii) $\left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3, x_1 + x_2 + x_3 = t \right\}$ (can check that this is a subspace for $t = 0$ only).

Prop 1.1 (Intersection of two subspaces is a subspace)**.** Let V be an F vector space. Let $U, W \leq V$. Then:

$$U \cap W \leq V$$

Proof.

$$0 \in U,\, 0 \in W \implies 0 \in U \cap W$$

Stability: given $(\lambda_1, \lambda_2) \in F^2$ and $(v_1, v_2) \in (U \cap W)^2$, we have that

$$\lambda_1 v_1 + \lambda_2 v_2 \in U \text{ and } \lambda_1 v_1 + \lambda_2 v_2 \in W$$

And so

$$\lambda_1 v_1 + \lambda_2 v_2 \in U \cap W$$

Warning. The union of two subspaces is generally NOT a subspace.
(Typically not stable by addition)

Definition (Sum of subspaces)**.** Let V be an F vector space, let $U \leq V, W \leq V$. The **sum** of U and W is the set:

$$U + W = \{u + w : (u, w) \in U \times W\}$$

Prop 1.2 (Sum of two spaces is a subspace)**.** For V a F vector space, $(U \leq V, W \leq V) \implies U + W \leq V$

Proof.

$$0 = 0_{\in U} + 0_{\in W} \in U + W$$

Given $\lambda_1, \lambda_2 \in F$ and $f, g \in U + W$, we have

$$f = f_1 + f_2$$
$$g = g_1 + g_2$$

with $f_1, g_1 \in U$ and $f_2, g_2 \in W$. Hence

$$\lambda_1 f + \lambda_2 g = \lambda_1(f_1 + f_2) + \lambda_2(g_1 + g_2) = (\lambda_1 f_1 + \lambda_2 g_1) + (\lambda_1 f_2 + \lambda_2 g_2) \in U + W$$

(first bracket in U, second bracket in W)

Exercise: Show that $U + W$ is the smallest subspace of V which contains U and W.

1.1 Subspaces and quotient

Definition (Quotient)**.** Let V be an F vector space. Let $U \leq V$. The **quotient space** V/U is the abelian group V/U equipped with the scalar multiplication:

$$F \times V/U \to V/U, \quad (\lambda, v + U) \mapsto \lambda v + U$$

Note. We must check that the multiplication operator is well-defined. Indeed,

$$\begin{aligned} v_1 + U = v_2 + U &\implies v_1 - v_2 \in U \\ &\implies \lambda(v_1 - v_2) \in U \\ &\implies \lambda v_1 + U = \lambda v_2 + U \in V/U \end{aligned}$$

Prop 1.3 (Quotient spaces are vector spaces)**.** V/U is an F vector space.

Proof. Exercise.

1.2 Spans, Linear Independence and Steinitz Exchange Lemma

Definition (Span of a family of vectors)**.** Let V be an F vector space. Let $S \subset V$ be a subset (so S is a set of vectors). We define:

$$\begin{aligned} \langle S \rangle &= \{\text{finite linear combinations of elements of } S\} \\ &= \left\{ \sum_{s \in S} \lambda_s v_s : v_s \in s, \text{ only finitely many } \lambda_s \text{ are non-zero} \right\} \end{aligned}$$

Write $\langle S \rangle$ for **span** S. By convention, $\langle \varnothing \rangle = \{0\}$

Remark. $\langle S \rangle$ = smallest vector subspace of V which contains S.

Examples. (i) $V = \mathbb{R}^3$:

$$S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \\ -4 \end{bmatrix} \right\} \implies \langle S \rangle = \left\{ \begin{bmatrix} \alpha \\ \beta \\ 2\beta \end{bmatrix}, (\alpha, \beta) \in \mathbb{R}^2 \right\}$$

(ii)

$$V = \mathbb{R}^n = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, x_i \in \mathbb{R}, 1 \leq i \leq n \right\}, \quad e_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

(1 on i-th position) so $V = \text{span }(e_i)_{1 \leq i \leq n}$

(iii) X set, $V = \mathbb{R}^X = \{f : X \to \mathbb{R}\}$

$$S_x : X \to \mathbb{R}, \quad y \mapsto \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$

$\text{Span}(S_x)_{x \in X} \equiv \{f \in \mathbb{R}^X : f \text{ has finite support }\}$
$(\text{Supp } f = \{x : f(x) \neq 0\})$

Definition. Let V be an F-vector space. Let S be a subset of V. We say that S **spans** V if $\langle S \rangle = V$

Definition (Finite dimension)**.** Let V be an F-vector space. We say that V is **finite dimensional** if it is spanned by a finite set. We say V is **infinite dimensional** if there is no family S with finitely many elements which span V.

Example. Let $V = \mathbb{P}[x]$, the set of polynomials in $\mathbb{R}$. Let $V_n = \mathbb{P}_n[x]$, the set of polynomials in $\mathbb{R}$ with degree $\leq n$, $n \in \mathbb{N}$. $V_n = \langle \{1, x, \dots, x^n\} \rangle$ so V_n is finite dimensional

Claim. $V = \mathbb{P}[x]$ is infinite dimensional

Proof. Exercise.

If V is finite dimensional, is there a minimal number of vectors in the family required so that the family spans V?

Definition (Independence)**.** We say that $(v_1, \dots, v_n)$ elements of V are **linearly independent** if

$$\sum_{i=1}^{n} \lambda_i v_i = 0,\ \lambda_i \in F \implies \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$$

Equivalently, $(v_1, \dots, v_n)$ are not linearly independent if one of them is a linear combination of the $(n-1)$ reaining ones.
Indeed, $\exists(\lambda_1, \dots, \lambda_n)$, $j \in [1, n]$ s.t. $\sum_{i=1}^{n} \lambda_i v_i = 0$ and $\lambda_j \neq 0$.
Which implies

$$v_j = -\frac{1}{\lambda_j} \sum_{i \neq j}^{n} \lambda_i v_i$$

Remark. $(v_i)_{1 \leq i \leq n}$ linearly independent $\implies \forall i \in [1, n]$, $v_i \neq 0$

Definition (Basis)**.** A subset S of V is a **basis** of V if:
(i) $\langle S \rangle = V$
(ii) S linearly independent

Remark. When S spans V, we say that S is a generating family. So a basis is a linearly independent generating family.

Examples.
(i) Let $V = \mathbb{R}^n$ and e_i as before. Then $(e_i)_{1 \leq i \leq n}$ is a basis for V (exercise)
(ii) $V = \mathbb{C}$. $\mathbb{C} \equiv \mathbb{C}(= F)$ vector space, $\{1\}$ a basis but also $\mathbb{C} \equiv \mathbb{R}(= F)$ vector space, $\{1, i\}$ a basis
(iii) For $V = \mathbb{P}[x] = \{P(x)$ polynomials on $\mathbb{R}\}$, $S = \{x^n,\ n \geq 0\}$ is a basis of V

Lemma 1.4 (Unique decomposition for everything equivalent to being a basis)**.** Let V be a F vector space. Then $(v_1, \dots, v_n)$ is a basis of V if and only if any vector $v \in V$ has a unique decomposition:

$$v = \sum_{i=1}^{n} \lambda_i v_i,$$

Proof.

$$\langle v_1, \dots, v_n \rangle = V \implies \forall v \in V, \exists(\lambda_1, \dots, \lambda_n) \in F^n \text{ s.t. } v = \sum_{i=1}^{n} \lambda_i v_i$$

If

$$v = \sum_{i=1}^{n} \lambda_i v_i = \sum_{i=1}^{n} \lambda'_i v_i$$

then

$$\sum_{i=1}^{n} (\lambda_i - \lambda') v_i = 0$$

so we must have $\lambda_i = \lambda'_i$, $\forall 1 \leq i \leq n$ since $(v_i)_{1 \leq i \leq n}$ linearly independent

Lemma 1.5 (Some subset of a spanning set is a basis)**.** If $(v_1, \dots, v_n)$ spans V, then some subset of this family is a basis of V.

Proof. If $(v_1, \dots, v_n)$ are linearly independent, done. If they are not, then up to changing indices,

$$\begin{aligned} v_n \in \text{span}(v_1, \dots, v_{n-1}) &\implies \langle v_1, \dots, v_n \rangle = \langle v_1, \dots, v_{n-1} \rangle \\ &\implies \langle v_1, \dots, v_{n-1} \rangle = V \end{aligned}$$

Iterate this process

Theorem 1.6 (Steinitz Exchange Lemma)**.** Let V be a finite dimensional vector space over F. Take $(v_1, \dots, v_m)$ linearly independent, and $(w_1, \dots, w_n)$ which spans V. Then:

(i) $m \leq n$

(ii) Up to reordering,
$(v_1, \dots, v_m, w_{m+1}, \dots, w_n)$ spans V

Proof (Induction)**.** Suppose that we have replced $l (\geq 0)$ of the w_i. Reordering if necessary, $\langle v_1, \dots, v_l, w_{l+1}, \dots, w_n \rangle = V$. If $m = l$, done. Assume $l < m$. Then: $v_{l+1} \in V$.

$$v_{l+1} = \sum_{i \leq l} \alpha_i v_i + \sum_{i > l} \beta_i w_i$$

Since the $(v_i)_{1 \leq i \leq m}$ $(l + 1 \leq m)$ are linearly independent, one of the β_i, is non-zero. So, up to reordering:

$$w_{l+1} = \frac{1}{\beta_{l+1}}(v_{l+1} - \sum_{i \leq l} \alpha_i v_i - \sum_{i > l} \beta_i w_i)$$

$\implies V$ is spanned by $(v_1, \dots, v_{l+1}, w_{l+2}, \dots, w_n)$. And so we are done after m steps thus we must have replaced m of the w_i so $m \leq n$

1.3 Bases, dimension, direct sums

Corollary 1.7 (Dimension fixed)**.** V be a finite dimensional vector space over F, then: any two basis of V have the same number of vectors called the dimension of V, $\dim_F(V)$.

Proof. $(v_1, \dots, v_n), (w_1, \dots, w_m)$ basis of V over F. Then :

- $(v_i)_{1 \leq i \leq n}$ free, $(w_i)_{1 \leq i \leq m}$ generating $\implies n \leq m$
- $(w_i)_{1 \leq i \leq m}$ free, $(v_i)_{1 \leq i \leq n}$ generating $\implies m \leq n$

Corollary 1.8 ($|$Independent$| \leq |$basis$| \leq |$spanning$|$)**.** Let V be an F vector space with finite dimension n. Then:

(i) any independent set of vectors has at most n elements, with equality iff it is a basis

(ii) any spanning set of vectors has at least n elements, with equality iff it is a basis.

Proof. Trivial.

Prop 1.9 (Dimension of sum of subspaces)**.** Let U, W be subspaces of V. If U and W are finite dimensional, then so $U + W$ and:

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$$

Proof. Pick a basis $v_1, \ldots, v_p$ of $U \cap W$. Extend to a basis: $v_1, \ldots, v_p, u_1, \ldots, u_m$ of U and $v_1, \ldots, v_p, w_1, \ldots, w_n$ of W

Claim. $(v_1, \ldots, v_p, u_1, \ldots, u_m, w_1, \ldots, w_n)$ is a basis of $U + W$.

Generating family of $U + W$: obvious.
Free family (linearly independent):

$$\sum_{i=1}^{p} \alpha_i v_i + \sum_{i=1}^{m} \beta_i u_i + \sum_{i=1}^{n} \gamma_i w_i = 0$$

$$\implies \sum_{i=1}^{p} \alpha_i v_i + \sum_{i=1}^{m} \beta_i u_i = -\sum_{i=1}^{n} \gamma_i w_i$$

LHS in U, RHS in W

$$\implies \sum_{i=1}^{n} \gamma_i w_i \in U \cap W \implies \sum_{i=1}^{p} S_i v_i = \sum_{i=1}^{n} \gamma_i w_i$$

As $v_1, \ldots, v_p$ basis of $U \cap W$

$$\begin{aligned} &\implies \sum_{i=1}^{p} (\alpha_i + S_i) v_i + \sum_{i=1}^{m} \beta_i u_i = 0 \\ &\implies \alpha_i = -S_i,\ \beta_i = 0 \\ &\implies \sum_{i=1}^{p} \alpha_i v_i + \sum_{i=1}^{n} \gamma_i w_i = 0 \\ &\implies \alpha_i = \gamma_i = 0 \end{aligned}$$

As $(v_1, \ldots, v_p, w_1, \ldots, w_n)$ free

$$\implies \alpha_i = \beta_i = \gamma_i = 0$$

Prop 1.10 (Dimension of quotient space)**.** If V is a finite dimensional vector space over F and $U \leq V$ (subspace), then U and V/U are also finite dimensional and:

$$\dim V = \dim U + \dim V/U$$

Proof. Let $(u_1, \ldots, u_l)$ be a basis of U and extend it to a basis $(u_1, \ldots, u_l, w_{l+1}, \ldots, w_n)$ of V. We can show that $(w_{l+1} + U, \ldots, w_n + U)$ is a basis of V/U

Remark. For V a vector space over F and $U \leq V$, we say that U is a proper subspace if $U \neq V$. Then U proper $\implies \dim U < \dim V$. ($V/U \neq \{0\} \implies \dim V/U > 0 \implies \dim U < \dim V$)

Definition (Direct Sum)**.** V vector space over F and $U, W \leq V$ (subspaces)
We say that: $V = U \bigoplus W$ ("V is the **direct sum** of U and W")
iff every element $v \in V$ can be written:

$$v = u + w \text{ with } (u, w) \in U \times W \text{ and this decomposition is unique.}$$

Equivalently: $V = U \bigoplus W$
$\iff \forall v \in V, \exists!(v, w) \in U \times W$ s.t. $v = u + w$ (uniqueness is important)

Warning. We say that W is a direct complement of U in V. There is no uniqueness of such a complement.

Lemma 1.11 (Direct sum $\iff$ sum with trivial intersection $\iff$ union of bases gives basis)**.** Let $U, W \leq V$, then:
The following are equivalent:

(i) $V = U \bigoplus W$

(ii) $V = U + W$ and $U \cap W = \{0\}$

(iii) For any basis $\mathcal{B}_1$ of U, B_2 of W, the union $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ is a basis of V

Proof. (ii) $\implies$ (i): $V = U + W \implies \forall v \in V, \exists (u, w) \in U \times W$ s.t. $v = u + w$.
Uniqueness: $u_1 + w_1 = u_2 + w_2 = v$
$\implies u_1 - u_2 = w_2 - w_1$ (LHS $\in U$, RHS $\in W$
$\implies u_1 = u_2$ and $w_1 = w_2$, as $U \cap W = \{0\}$

(i) $\implies$ (iii): $\mathcal{B}_1$ basis of U, $\mathcal{B}_2$ basis of W.
Let $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$

- generating family of $U + W$ obvious
- $\mathcal{B}$ free family: $\sum_{v_i \in \mathcal{B}} \lambda_i v_i = 0 = 0_U + 0_W$

$$\sum_{u \in \mathcal{B}_1} \lambda_u u + \sum_{w \in \mathcal{B}_2} \lambda_w w = 0$$

Thus by uniqueness:

$$\sum_{u \in \mathcal{B}_1} \lambda_u u = \sum_{w \in \mathcal{B}_2} \lambda_w w = 0$$

$$\implies \lambda_u = 0,\ \lambda_w = 0$$

As $\mathcal{B}_1$ basis, $\mathcal{B}_2$ basis
$\implies \mathcal{B}$ free family

(iii) $\implies$ (ii) We need to show $V = U + W$ and $U \cap W = \{0\}$.
$\mathcal{B}_1$ basis of U and $\mathcal{B}_2$ basis of $W \implies \mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ basis of V. (from (iii))

$$\forall v \in V,\ v = \sum_{u \in \mathcal{B}_1} \lambda_u u + \sum_{w \in \mathcal{B}_2} \lambda_w w$$

$$\implies V = U + W$$

Let $v \in U \cap W$, then:

$$v = \sum_{u \in \mathcal{B}_1} \lambda_u u = \sum_{w \in \mathcal{B}_2} \lambda_w w$$

$$v = \sum_{u \in \mathcal{B}_1} \lambda_u u - \sum_{w \in \mathcal{B}_2} \lambda_w w = 0$$

$$\implies \lambda_u = \lambda_w = 0$$

As $\mathcal{B}_1 \cup \mathcal{B}_2$ free

Definition. For$V$ a vector space over F
$v_1, \ldots, v_p \leq V$ (subspaces)

(i)
$$\sum_{i=1}^{p} V_i = \{v_1 + \cdots + v_p,\, v_j \in V_j,\, 1 \leq j \leq p\}$$

(ii) The sum is direct:
$$\sum_{i=1}^{p} V_i = \bigoplus_{i=1}^{p} V_i$$

iff: $v_1 + \cdots + v_p = v_1' + \cdots + v_p'$
$\implies v_1 = v_1', \ldots, v_p = v_p'$

Equivalently:
$$V = \bigoplus_{i=1}^{p} V_i$$
$$\iff \forall v \in V, \exists!(v_1, \ldots, v_k) \in \prod_{i=1}^{p} : v = \sum_{i=1}^{p} v_i$$

Claim (Generalisation of previous lemma)**.** TFAE:

(i) $\sum_{i=1}^{p} V_i = \bigoplus_{i=1}^{p} V_i$ (sum is direct)

(ii)
$$\forall i,\, V_i \cap \left(\sum_{j \neq i} V_i\right) = \{0\}$$

(iii) For any basis $\mathcal{B}_i$ of $V_i, \mathcal{B} = \bigcup_{i=1}^{p} \mathcal{B}_i$ is a basis of $\sum_{i=1}^{p} V_i$

Proof. Exercise.

1.4 Linear maps, isomorphisms and the rank-nullity Theorem

Definition (Linear Map)**.** V, W are F-vector spaces. A map $\alpha : V \to W$ is **linear** iff:
$$\forall(\lambda_1, \lambda_2) \in F^2,\, \forall(v_1, v_2) \in V \times V,$$
$$\alpha(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \alpha(v_1) + \lambda_2 \alpha(v_2)$$

Examples. (i) Matrices $\mathbb{R}^n \to \mathbb{R}^m$
(ii)

$$\alpha : \mathcal{C}([j,1]) \to \mathcal{C}([j,1]), \quad f \mapsto \alpha(f)(x) = \int_{\infty}^{x} f(t)\,\mathrm{d}t$$

is a linear map

Remark. For U, V, W F vector spaces
(i) $\mathrm{Id}_V : V \to V$ i a linear map
(ii) $U \to V \to W$, composition of 2 linear maps is linear.

Lemma 1.12 (Linear maps can be idenfied by where they send basis)**.** For V, W F vector spaces with $\mathcal{B}$ basis for V, if $\alpha_0 : \mathcal{B} \to V$ is any map, then there is a unique linear map $\alpha : V \to W$ extending α_0 (i.e. $\forall v \in \mathcal{B}, \alpha_0(v) = \alpha(v)$

Proof. $v \in V$, $v = \sum_{i=1}^{n} \lambda_i v_i$ and $\mathcal{B} = (v_1, \dots, v_n)$.
Necessarily by linearity:

$$\alpha(v) = \alpha\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i \alpha(v_i)$$

Remark. (i) True for ∞ dimensional spaces as well. Often, to define a linear map, we define its value on a basis and "extend by linearity."
(ii) $\alpha_1, \alpha_2 : V \to W$ linear. If they agree on a basis $\mathcal{B}$ of V, they are equal.

Definition (Isomorphism)**.** For V, W vector spaces over F. A map $\alpha : V \to W$ is called an **isomorphism** iff:
(i) α linear
(ii) α bijection
If such an α exists, we note: $V \cong W$ (V is isomorphic to W)

Remark. $\alpha : V \to W$ linear isomorphism $\implies$ $\alpha^{-1} : W \to V$ is linear

Lemma 1.13 ('is isomorphic to' is an equivalence relation)**.** $\cong$ is an equivalence relation on the class of all vector spaces over F.
(i) $i_V : V \to V$ isomorphism
(ii) $\alpha : V \to W$ isomorphism $\implies$ $\alpha^{-1} : W \to V$ isomorphism
(iii) If $U \to V \to W$ (maps β then α isomorphisms)
$\implies$ $U \to W$ given by $\alpha \circ \beta$ is an isomorphism.

Proof. Exercise.

Theorem 1.14 (Dimension n implies isomorphic to F^n)**.** If V is a vector space over F of dimension n, then:

$$V \cong F^n$$

Proof. Let $\mathcal{B} = (v_1, \dots, v_n)$ be a basis of V.
Then $\alpha : V \to F^n$
$v = \sum_{i=1}^{n} \lambda_i v_i \mapsto \begin{bmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{bmatrix}$ is an isomorphism. (Exercise)

Remark. Choosing a basis of V is like choosing an isomorphism from V to F^n

Theorem 1.15 (Isomorphic iff same dimension (for finite dimensions))**.** Let V, W be F vector spaces with finite dimension. Then $V \cong W$ iff they have the same dimension.

Proof. $\Longleftarrow$: $\dim V = \dim W = n$
$\Longrightarrow$ $V \cong F^n, W \cong F^n$ so $V \cong W$
$\Longrightarrow$: $\alpha : V \to W$ isomorphism, $\mathcal{B}$ is a basis for V, then:

Claim. $\alpha(V)$ basis for W.

- $\alpha(\mathcal{B})$ spans V follows from surjectivity of α
- $\alpha(\mathcal{B})$ free family follows from injectivity of α

Proof. Exercise.

Definition (Kernel and image of a linear map)**.** Let V, W vector spaces over F.
Let $\alpha : V \to W$ linear map. We define:
$\ker \alpha = \{v \in V : \alpha(v) = 0\}$ (**kernel** of α) $\operatorname{Im} \alpha = \{w \in W : \exists v \in V,\, w = \alpha(v)\}$ (**image** of α)

Lemma 1.16 (kernel and image are vector spaces)**.** $\ker \alpha$ and $\operatorname{Im} \alpha$ are subspaces respectively V and W

Proof. $(\lambda_1, \lambda_2) \in F^2$, $(v_1, v_2) \in \ker \alpha \times \ker \alpha$,
$\alpha(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \alpha(v_1) + \lambda_2 \alpha(v_2) = 0 + 0 = 0$
$\Longrightarrow$ $\lambda_1 v_1 + \lambda_2 v_2 \in \ker \alpha$
$(\lambda_1 v_1 + \lambda_2 v_2) \in F^2$, $(w_1, w_2) \in (\operatorname{Im} \alpha)^2$,
$w_1 = \alpha(v_1)$, $w_2 = \alpha(v_2)$
$\Longrightarrow$ $\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 \alpha(w_1) + \lambda_2 \alpha(v_2) = \alpha(\lambda_1 v_1 + \lambda_2 v_2) \in \operatorname{Im} \alpha$

Theorem 1.17 (Quotient by kernel isomorphic to image)**.** Let V, W be F vector spaces. Let $\alpha : V \to W$ linear map. then:

$$\overline{\alpha} : V/\ker\alpha \mapsto \operatorname{Im}\alpha$$

$$\overline{\alpha}(v + \ker\alpha) \mapsto \alpha(v)$$

is an isomorphism.

Proof. $\overline{\alpha}$ well defined: trivial check
$\overline{\alpha}$ linear: follows immediately from α linear.
$\overline{\alpha}$ bijection:

- injectivity: $\overline{\alpha}(v + \ker\alpha) = 0$
  $\implies \alpha(v) = 0 \implies v \in \ker\alpha$
  $v + \ker\alpha = 0 + \ker\alpha$
- surjectivity: follows from the definition fo $\operatorname{Im}\alpha$:

$$w \in \operatorname{Im}\alpha, \exists v \in V : w = \alpha(v) = \overline{\alpha}(v)$$

Definition (Rank and nullity)**.** $r(\alpha) = \dim \operatorname{Im}(\alpha)$ (**rank**)
$n(\alpha) = \dim\ker(\alpha)$ (**nullity**)

Theorem 1.18 (Rank-nullity Theorem)**.** Let U, V be vector spaces over F, $\dim_F U < +\infty$
Let $\alpha : U \to V$ be a linear map, then:

$$\dim U = r(\alpha) + n(\alpha)$$

Proof. We have proved that:
$U/\ker\alpha \cong \operatorname{Im}(\alpha)$
$\implies \dim(U/\ker\alpha) = \dim\operatorname{Im}\alpha$
$\implies \dim(U) - \dim\ker\alpha = \dim\operatorname{Im}\alpha$
$\implies \dim U = r(\alpha) + n(\alpha)\square$

Lemma 1.19 (Characterization of isomorphism)**.** V, W vector spaces over F of equal finite dimension.
Let $\alpha : V \to W$ linear map, then TFAE:
(i) α injective
(ii) α surjective
(iii) α isomorphism

Proof. Exercise. Follows directly from the rank-nullity theorem.

1.5 Linear maps from V to W and matrices

Definition. The **space of linear maps** from V to W over F is:
$L(V, W) = \{\alpha : V \to W \text{ linear }\}$

Prop 1.20 (Set of linear maps between V and W is a vector space)**.** $L(V,W)$ is a vector space over F under the operations:

$$(\alpha_1+\alpha_2)(v) = \alpha_1(v) + \alpha_2(v)$$

$$(\lambda\alpha)(v) = \lambda\alpha(v)$$

Moreover if V and W are finite dimensional, then so is $L(V,W)$ and $\dim_F L(V,W) = (\dim_F V)(\dim_F W)$

Proof. $L(V,W)$ vector space is an exercise
Dimension statement proved later.

1.5.1 Matrices and linear maps

Definition. An $m\times n$ **matrix** over F is an array with m rows and n columns with entries in F:

$$(a_{ij})_{\substack{1\le i\le m\\1\le j\le n}} = \begin{bmatrix} & \vdots & \\ \dots & a_{ij} & \dots \\ & \vdots & \end{bmatrix}$$

$a_{ij}\in F$ with i row, j column.
$M_{m,n}(F) = \{\text{set of } m\times n \text{ matrices over } F\}$

Prop 1.21 (Set of $m\times n$ matrices over a field is a vector space)**.** $M_{m,n}(F)$ is an F vector space under operations:

$$(a_{ij}) + (b_{ij}) = (a_{ij}+b_{ij})$$

$$\lambda(a_{ij}) = (\lambda a_{ij})\ \lambda\in F$$

Proof. Exercise.

Prop 1.22 (Dimension of $M_{m,n}$)**.** $\dim_F M_{m,n}(F) = m\times n$

Proof. $1\le i\le m$, $1\le j\le n$. Define elementary matrix:

$$\begin{bmatrix} 0 & \vdots & 0 \\ \dots & 1_{ij} & \dots \\ 0 & \vdots & 0 \end{bmatrix}$$

Then $(E_{ij})_{\substack{1\le i\le m\\1\le j\le n}}$ basis of $M_{m,n}(F)$

Spans obvious: $M = (a_{ij}))_{\substack{1\le i\le m\\1\le j\le n}} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}E_{ij}$

Free family obvious □

1.5.2 Representation of linear maps by matrices

- V, W vector spaces over F and $\alpha : V \to W$ linear.
- Basis: $\mathcal{B} = (v_1, \dots, v_n)$ of V and $\mathcal{C} = (w_1, \dots, w_m)$ of W
- If $v \in V, v = \sum_{j=1}^n \lambda_j v_j = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} \in F^n$ (coordinates of v in the basis $\mathcal{B}$)

  $\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = [v]_{\mathcal{B}}$
- Similarly, for $w \in W$, we note:
  $[w]_{\mathcal{C}} =$ vector of coordinates of w in the basis $\mathcal{C}$.

Definition (Matrix of α in $\mathcal{B}$,$\mathcal{C}$ basis)**.** $[\alpha]_{\mathcal{B},\mathcal{C}} \equiv$ matrix of α wrt $\mathcal{B}$,$\mathcal{C}$
$\equiv \big([\alpha(v_1)]_{\mathcal{C}}, [\alpha(v_2)]_{\mathcal{C}}, \dots, \dots, [\alpha(v_n)]_{\mathcal{C}}\big) \in M_{m \times n}(F)$

Observation: If we let

$$[\alpha]_{\mathcal{B},\mathcal{C}} = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$$

Then, by definition: $1 \le j \le n$

$$\alpha(v_j) = \sum_{i=1}^m a_{ij} w_i$$

Lemma 1.23 (Writing in a vector in a different basis)**.** $\forall v \in V$,

$$[\alpha(v)]_{\mathcal{C}} = [\alpha]_{\mathcal{B},\mathcal{C}} [v]_{\mathcal{B}}$$

Proof. Given $v \in V$, $v = \sum_{j=1}^n \lambda_j v_j$

$$\alpha(v) = \alpha\left(\sum_{j=1}^n \lambda_j v_j\right)$$
$$= \sum_{j=1}^n \lambda_j \alpha(v_j) = \sum_{j=1}^n \lambda_j \sum_{i=1}^m a_{ij} w_i$$

Lemma 1.24 (Matrix that is the composition of linear maps)**.** $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$ linear and $U \xrightarrow{\alpha\circ\beta} W$
With: $\mathcal{A}$ basis of U
$\mathcal{B}$ basis of V
$\mathcal{C}$ basis of W
$\implies [\alpha \circ \beta]_{\mathcal{A},\mathcal{C}} = [\alpha]_{\mathcal{B},\mathcal{C}} \cdot [\beta]_{\mathcal{A},\mathcal{B}}$

Proof. $u_l \in \mathcal{A}$

$$\begin{aligned}(\alpha \circ \beta)(u_l) &= \alpha(\beta(u_l)) \\ &= \alpha\left(\sum_j b_{jl} v_j\right) \\ &= \sum_j b_{jl}\alpha(v_j) = \sum_j b_{jl} \sum_i a_{ij} w_i \\ &= \sum_i \left(\sum_j a_{ij} b_{jl}\right) w_i\end{aligned}$$

With sum in brackets is the (i, l) entry of product of the 2 matrices.

Prop 1.25 (space of linear maps isomorphic to space of matrices from V to W)**.** Given V and W vector spaces over F with $\dim_F V = n$ and $\dim_F W = m$

$$L(V, W) \cong M_{m,n}(F)$$

Proof. Fix $\mathcal{B}, \mathcal{C}$ basis of V, W.

Claim. $\theta : L(V, W) \to M_{m,n}(F)$
$\alpha \mapsto [\alpha]_{\mathcal{B},\mathcal{C}}$
is an isomorphism.

Proof.

- θ linear: $[\lambda_1\alpha_1 + \lambda_2\alpha_2]_{\mathcal{B},\mathcal{C}} = \lambda[\alpha_1]_{\mathcal{B},\mathcal{C}} + \lambda_2[\alpha_2]_{\mathcal{B},\mathcal{C}}$
- θ surjective: Indeed, pick $A = (a_{ij})_{\substack{1\le i\le m\\1\le j\le n}}$

  Consider the map:
  $\alpha : v_j \mapsto \sum_{i=1}^{m} a_{ij} w_i$, $1 \le j \le n$
  so α is a map defined on $(v_1, \dots, v_m) \equiv$ basis of V.
  ("extend by linearity")
  $\implies \alpha$ linear map, which by definition:

  $$[\alpha]_{\mathcal{B},\mathcal{C}} = (a_{ij})_{\substack{1\le i\le m\\1\le j\le n}} = A$$

- θ injective: $[\alpha]_{\mathcal{B},\mathcal{C}} = 0 \implies \alpha = 0$

Remark. $\mathcal{B}$ basis of V
$\mathcal{C}$ basis of V
$\varepsilon_{\mathcal{B}} : V \to F^n \quad \varepsilon_{\mathcal{C}} : W \to F^m$
$v \mapsto [v]_{\mathcal{B}} \qquad w \mapsto [w]_{\mathcal{C}}$

then the following diagram commutes:

$$\begin{array}{ccc} V & \overset{\alpha}{\longleftrightarrow} & W \\ {\scriptstyle e_{\mathcal{B}}}\updownarrow & & \updownarrow{\scriptstyle e_{\mathcal{C}}} \\ F^n & \underset{}{\overset{[\alpha]_{\mathcal{B},\mathcal{C}}}{\longleftrightarrow}} & F^m \end{array}$$

Claim (linear map between subspaces induces quotient map)**.** $Y \leq V$, $\alpha(Y) = Z \leq W$. α induces:

$$\overline{\alpha} : V/Y \to W/Z$$
$$v + Y \mapsto \alpha(v) + Z$$

Proof.

- Well-defined: $v_1 + Y \mapsto v_2 + Y$

$$\implies v_1 - v_2 \in Y$$
$$\alpha(v_1 - v_2) \in Z$$
$$\implies \alpha(v_1) + Z = \alpha(v_2) + Z$$

- $\overline{\alpha}$ linear obvious (α linear)

1.6 Change of basis and equivalent matrices.

- $\alpha : V \to W$
  $\mathcal{B}$ basis of V, $\mathcal{C}$ basis of W
  $[\alpha(v)]_{\mathcal{C}} = [\alpha]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}$
  $[\alpha]_{\mathcal{B},\mathcal{C}} = (\alpha(v_1)|\dots|\alpha(v_n))$ wrt basis $\mathcal{C}$
- $U \xrightarrow{\beta} V \xrightarrow{\alpha} W$
  $\mathcal{A}, \mathcal{B}, \mathcal{C}$ basis U, V, W
  $\implies [\alpha \circ \beta]_{\mathcal{A},\mathcal{C}} = [\alpha]_{\mathcal{B},\mathcal{C}}[\beta]_{\mathcal{A},\mathcal{B}}$

1.6.1 Change basis

$$\begin{array}{ccc} V & \xrightarrow{\alpha} & W \\ \mathcal{B} = \{v_1, \dots, v_n\} & & \mathcal{C} = \{w_1 \dots, w_m\} \\ \mathcal{B}' = \{v_1' \dots, v_n'\} & & \mathcal{C}' = \{w_1' \dots, w_m'\} \end{array}$$

Aim: Find equation to relate $[\alpha]_{\mathcal{B},\mathcal{C}}, [\alpha]_{\mathcal{B}',\mathcal{C}'}$

Definition. The **change of basis matrix** from $\mathcal{B}'$ to $\mathcal{B}$ is $P = (p_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq n}}$ is given by:

$$P = \big([v'_1]_{\mathcal{B}}\big|[v'_2]_{\mathcal{B}}\big|\dots\big|[v'_n]_{\mathcal{B}}\big)$$
$$(\equiv [\mathrm{Id}]_{\mathcal{B}',\mathcal{B}})$$

Lemma 1.26 (writing vector in different basis)**.**

$$[v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}$$

Proof.

- $[\alpha(v)]_{\mathcal{C}} = [a]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}$
- $P = [\mathrm{Id}]_{\mathcal{B},\mathcal{B}'}$

$[\mathrm{Id}(v)]_{\mathcal{B}} = [\mathrm{Id}]_{\mathcal{B}',\mathcal{B}}[v]_{\mathcal{B}'}$
using $(\mathcal{B} = \mathcal{C}, \mathcal{B}' = \mathcal{B})$
$\implies [v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}\square$

Remark. P is an $n \times n$ invertible matrix, and: $P^{-1} \equiv$ change of basis matrix from $\mathcal{B}$ to $\mathcal{B}'$.

Indeed: $[\alpha \circ \beta]_{\mathcal{A},\mathcal{C}} = [\alpha]_{\mathcal{B},\mathcal{C}}[\beta]_{\mathcal{A},\mathcal{B}}$

$$[\mathrm{Id}]_{\mathcal{B},\mathcal{B}'}[\mathrm{Id}]_{\mathcal{B}',\mathcal{B}} = [\mathrm{Id}]_{\mathcal{B}',\mathcal{B}'} = I_n = \begin{bmatrix}1 & & \\ & \ddots & \\ & & 1\end{bmatrix}$$

$$[\mathrm{Id}]_{\mathcal{B}',\mathcal{B}}[\mathrm{Id}]_{\mathcal{B},\mathcal{B}'} = [\mathrm{Id}]_{\mathcal{B},\mathcal{B}} = I_n = \begin{bmatrix}1 & & \\ & \ddots & \\ & & 1\end{bmatrix}$$

$\implies [\mathrm{Id}]_{\mathcal{B},\mathcal{B}'}P = P[\mathrm{Id}]_{\mathcal{B},\mathcal{B}'} = I_n$ $\square$

Warning. $[v]_{\mathcal{B}}$, $P([v'_1]_{\mathcal{B}}, \dots, [v'_n]_{\mathcal{B}})$
$[v]_{\mathcal{B}'} = I^{-1}[v]_{\mathcal{B}}$
$\implies$ need to invert P!

- We changed $\mathcal{B}$ to $\mathcal{B}'$ in V.
- We can also change basis $\mathcal{C}$ to $\mathcal{C}'$ in W
  $(\alpha : V \to W)$

V — W
$\mathcal{B}, \mathcal{B}'$ — $\mathcal{C}, \mathcal{C}'$
$P = [\mathrm{Id}]_{\mathcal{B}',\mathcal{B}}$ — $P = [\mathrm{Id}]_{\mathcal{C}',\mathcal{C}}$
$\alpha : V \to W$

How do $[\alpha]_{\mathcal{B},\mathcal{C}}$ and $[\alpha]_{\mathcal{B}',\mathcal{C}'}$ relate

Prop 1.27 (Writing linear map in different basis)**.** $A = [\alpha]_{\mathcal{B},\mathcal{C}}$, $A' = \alpha_{\mathcal{B}',\mathcal{C}'}$ and $P = [\text{Id}]_{\mathcal{B}',\mathcal{B}}$, $Q = [\text{Id}]_{\mathcal{C}',\mathcal{C}}$
$\implies A' = Q^{-1}AP$

Proof. Have:

$$\begin{aligned}[\alpha(v)]_{\mathcal{C}} &= [\alpha]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}} \\ [\alpha \circ \beta]_{\mathcal{A},\mathcal{C}} &= [\alpha]_{\mathcal{B},\mathcal{C}}[\beta]_{\mathcal{A},\mathcal{B}} \\ [v]_{\mathcal{B}} &= P[v]_{\mathcal{B}'}\end{aligned}$$

So:

$$\begin{aligned}[\alpha(v)]_{\mathcal{C}} &= Q[\alpha(v)]_{\mathcal{C}'} \\ &= Q[\alpha]_{\mathcal{B}',\mathcal{C}'}[v]_{\mathcal{B}'}\end{aligned}$$

$$\begin{aligned}[\alpha(v)]_{\mathcal{C}} &= [\alpha]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}} \\ &= AP[v]_{\mathcal{B}'}\end{aligned}$$

$\implies \forall v \in V$, $QA'[v]_{\mathcal{B}'} = AP[v]\mathcal{B}'$
$\implies QA' = AP$
$\implies A' = Q^{-1}AP$ □

Definition (Equivalent matrices)**.** Two matrices $A, A' \in M_{m,n}(F)$ are **equivalent** if:

$$A' = Q^{-1}AP$$

Where $Q \in M_{m\times m}$ invertible
$P \in M_{n\times n}$ invertible

Remark. This defines an equivalence relation on $M_{m,n}(F)$.

- $A = I_m^{-1}AI_n$
- $A' = Q^{-1}AP \implies A = (Q^{-1})^{-1}A'P^{-1}$
- $A' = Q^{-1}AP$ and $A'' = (Q')^{-1}A'P' \implies A'' = (QQ')^{-1}A(PP')$

Prop 1.28 (Can choose bases such that corresponding matrix diagonal)**.** Let V, W vector spaces over F and $\dim_F V = n$, $\dim_F W = m$
Let $\alpha : V \to W$ linear map. Then there exists $\mathcal{B}$ basis of V and $\mathcal{C}$ basis of W, s.t.:

$$[\alpha]_{\mathcal{B},\mathcal{C}} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

Proof. Choose $\mathcal{B}$ and $\mathcal{C}$ wisely.
Fix $r \in \mathbb{N}$ s.t. $\dim \ker \alpha = n - r$
$N(\alpha) = \ker(\alpha) = \{x \in V : \alpha(x) = 0\}$
Fix a basis of $N(\alpha) : v_{r+1}, \dots, v_n$
Extend it to a basis of $V \equiv \mathcal{B}$
$\mathcal{B} = (v_1, \dots, v_r, v_{r+1}, \dots, v_n)$

Claim. $(\alpha(v_1), \dots, \alpha(v_r))$ basis of Im α

Proof.
- Span: $v = \sum_{i=1}^{n} \lambda_i v_i$
  $\implies \alpha(v) = \sum_{i=1}^{n} \lambda_i \alpha(v_i) = \sum_{i=1}^{r} \lambda_i \alpha(v_i)$
  Let $y \in \text{Im}\,(\alpha)$, then: $\exists v \in V : \alpha(v) = y$
  $\implies y = \alpha(v) = \sum_{i=1}^{r} \lambda_i \alpha(v_i) \in \text{span}\,\{\alpha(v_1), \dots, \alpha(v_n)\}$
- Free:

$$\sum_{i=1}^{r} \lambda_i \alpha(v_i) = 0$$
$$\implies \alpha\left(\sum_{i=1}^{r} \lambda_i v_i\right) = 0$$
$$\implies \sum_{i=1}^{r} \lambda_i v_i \in \ker \alpha = \text{span}\,\{v_{r+1}, \dots, v_n\}$$
$$\implies \sum_{i=1}^{r} \lambda_i v_i = \sum_{r+1}^{n} \mu_i v_i$$
$$\implies \sum_{i=1}^{r} \lambda_i v_i - \sum_{r+1}^{n} \mu_i v_i = 0$$
$$\implies \lambda_i = \mu_i = 0 \implies \text{free}$$

We have proved that $(\alpha(v_1), \dots, \alpha(v_r))$ basis of Im α basis of Im α and $v_{r+1}, \dots, v_n$ basis of ker α
$\mathcal{B} = (v_1, \dots, v_r, v_{r+1}, \dots, v_n)$
$\mathcal{C} = (\alpha(v_1), \dots, \alpha(v_r), w_{r+1}, \dots, w_n)$ basis of W
$[\alpha]_{\mathcal{B},\mathcal{C}} = (\alpha(v_1)| \dots |\alpha(v_r)|\alpha(v_{r+1})| \dots |\alpha(v_n))$ wrt $\mathcal{C}$ is the desired matrix.

Remark. This provides another proof of the rank nullity Theorem:

$$r(\alpha) + N(\alpha) = n$$

Corollary 1.29 (Equivalence is determined by rank)**.** Any $m \times n$ matrix is equivalent to: $\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]$ where $r = r(\alpha)$

Definition. $A \in M_{m,n}(F)$

- The **column rank** of A, $r(A)$, is the dimension of the subspace of F^m spanned by the column vectors of A
  $A = (c_1| \dots |c_n)$, $c_i \in F^m$
  $r(A) = \dim_F \text{ span } \{c_1, \dots, c_n\}$
  Similarly, the **row rank** of A is the column rank of A^T

Remark. If α is a linear map represented by A with respect to some basis, then:

$$r(A) = r(\alpha)$$

(column rank = rank)

Prop 1.30 (Equivalence is determined by rank)**.** Two matrices are equivalent iff $r(A) = r(A')$

Proof. ($\implies$) If A, A' equivalent, they correspond to the same endomorphism α expressed in two different basis:

$$r(A) = r(\alpha) = r(A')$$

($\impliedby$)$r(A) = r(A') = r$, then A and A' are both equivalent to:
$\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]$
$\implies$ A and A' are equivalent.

Theorem 1.31 (Column rank = row rank)**.** $r(A) = r(A^T)$

Proof. $r = r(A)$

$$Q^{-1}AP = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]_{m\times n}$$

Take the transpose:

$$\begin{aligned} (Q^{-1}AP)^T &= P^T A^T (Q^{-1})^-1 \\ &= P^T A^T (Q^T)^{-1} \\ \implies P^T A^T (Q^T)^{-1} &= \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]^T \\ &= \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]_{n\times m} \\ \implies r(A^T) &= r(A)\square \end{aligned}$$

1.7 Elementary Operations and Elementary Matrices

Definition. A linear map $\alpha : V \to V$ is called an **endomorphism**

Equation. With P as the change of basis matrix from B' to B

$$[\alpha]_{\mathcal{B}',\mathcal{B}'} = P^{-1}[\alpha]_{\mathcal{B},\mathcal{B}}P$$

Definition. For A, A' $(n \times n)$ square matrices. We say that A and A' are **similar** (conjugate) iff

$$A' = P^{-1}AP$$

for P $(n \times n)$ invertible

Definition. An **elementary column operation** on a $m \times n$ matrix A is one of the following

(i) swap column i and j for $i \neq j$
(ii) replace column i by $\lambda\times$ (column i), $(\lambda \neq 0, \lambda \in F)$
(iii) add $\lambda\times$ (column i) to column j for $i \neq j$

and elementary row operations are defined analagously. We note that these operations are invertible and these operations can be relaligned through the action of elementary matrices:

(i) trivial to consider (swap rows in identity matrix). Let T_{ij} be the matrix that swaps row i and row j
(ii) $n_{i,\lambda}$ is the identity with ith row replaced by λ
(iii) $C_{i,j,\lambda} =$ Id $+ \lambda E_{ij}$ where E_{ij} just has 1 on the ith row and jth column

Remark. Link between elementary operations and elementary matrices: an elementary column (resp. row) operation can be performed by multiplying A by the corresponding elementary matrix from the right (resp. left)

Example.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}$$

Note. This gives a constructive proof that any $m \times n$ matrix is equivalent to

$$\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]$$

Method. (i) Start with A. If all entries are zero, done

(ii) Pick $a_{ij} = \lambda \neq 0$:
- swap rows i and 1
- swap columns j and 1

(iii) Multiply column 1 by $1/\lambda (\lambda \neq 0)$

(iv) Now clear out row 1 and column 1 using elmenetary operations of type (iii)

(v) continue with remaining $(n-1) \times (n-1)$ matrix until we end the process

$$Q^{-1}AP = E'_l \dots E'_1 A E_1 \dots E_k$$

rows on left, columns on right

Remark. Variations

(i) Gauss' pivot algorithm. If you use only row operations, you can reach the so called "row-echelon" form of the matrix by the following
- Assume that $a_{j1} \neq 0$ for some i
- Swap rows i and 1
- Divide first row by $\lambda = a_{i1}$ to get 1 in (1,1)
- Use 1 to clean the rest of the 1st column and similar for 2nd column etc.

Note. This is exactly how we solve systems of linear equations

(ii) Representation of square invertible matrices

Lemma 1.32 (Only need to operate by rows/columns to get I_n if invertible)**.** If A is a $(n \times n)$ square invertible matrix, then we can obtain I_n using row elementary operations only (resp. column operations only)

Proof. We do the proof for column operations. We argue by induction on the number of rows. Suppose we can reach a form with I_k in the top left, zeros to the left and 'stuff' below. We want to obtain the form with $k+1$ instead.
Easy to prove $\exists j > k$ s.t. $a_{k+1,j} = \lambda > 0$ by considering spans. Then, we can swap column $k+1$ and j then divide $k+1$ by $\lambda = a_{k+1,j} \neq 0$ and, as expected, use this to clear the rest of the $k+1$th row using type (iii) elementary operations

Our outcome is:

$$AE_1 \dots E_N = I$$

$$\implies A^{-1} = E_1 \dots E_N$$

Prop 1.33 (Can decompose invertible matrix into elementary matrices)**.** Any invertible square matrix is a product of elementary matrices

Proof. Proved above.

2 Dual Spaces and Dual Maps

Definition. Let V be a vector space over F. We say V^* is the **dual** of V which is

$$V^* = L(V, F) = \{\alpha : V \to F \text{ linear}\}$$

Notation. We say $\alpha : V \to F$ linear is a linear form

Examples. (i)

$$\text{Tr } : N_{n,n} \to F_n$$

$$A = (a_{ij}) \mapsto \sum_{i=1}^{n} a_{ii}$$

$$\text{Tr } \in N_{n,n}^*$$

(ii)

$$f : [0,1] \to \mathbb{R}$$

$$x \mapsto f(x)$$

$$T_f : \mathcal{C}([0,1], \mathbb{R}) \to \mathbb{R}$$

$$f \mapsto \int_0^1 f(x)g(x) \, \mathrm{d}x$$

$$T_f = \text{ linear form on } \mathcal{C}^\infty([0,1], \mathbb{R})$$

So you can construct f knowing T_f

Lemma 2.1 (We have a basis for $\mathcal{B}^*$ by the 'row vectors')**.** Let V be a vector space over F with a finite basis

$$\mathcal{B} = \{e_1, \ldots, e_n\}$$

Then there exists a basis for V^* given by

$$\mathcal{B}^* = \{\varepsilon_1, \ldots, \varepsilon_n\}$$

where

$$\varepsilon_j(\sum_{i=1}^{n} a_i e_i) = a_j,\ 1 \leq j \leq n$$

$$\mathcal{B}^* \equiv \text{ dual basis of } \mathcal{B}$$

Proof.

- $(\varepsilon_1, \ldots, \varepsilon_n)$ free family

$$\sum_{j=1}^{n} \lambda_j \varepsilon_j = 0$$

$$(\sum_{j=1}^{n} \lambda_j \varepsilon_j)(e_i) = 0 = \sum_{k=1}^{n} \lambda_j S_{ji} = \lambda_i,\ forall 1 \leq i \leq n$$

- Span: $\alpha \in V^*,\ x \in V$

$$\alpha(x) = \alpha(\sum_{j=1}^{n} \lambda_j e_j) = \sum_{j=1}^{n} \lambda_j \alpha(e_j)$$

On the other hand,

$$\sum_{j=1}^{n} \alpha(e_j) \varepsilon_j \in V$$

Then

$$\begin{aligned}(\sum_{j=1}^{n} \alpha(e_j)\varepsilon_j)(x) &= \sum_{j=1}^{n} \alpha(e_j) \sum_{k=1}^{n} \lambda_k \varepsilon_j(e_k) \\ &= \sum_{j=1}^{n} \alpha(e_j)\lambda_j = \alpha(x)\end{aligned}$$

We have shown

$$\alpha = \sum_{j=1}^{n} \alpha(e_j)\varepsilon_j$$

Notation. Kronecker symbol:

$$S_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

With this notation, we have:

$$\varepsilon_j(\sum_{i=1}^{n} a_i e_i) = a_j \iff \varepsilon_j(e_i) = S_{ij}$$

Corollary 2.2 (Dual same dimension)**.** V finite dimensional

$$\implies \dim V^* = \dim V$$

Remark. It is sometimes convenient to think of V^* as the space of row vectors of length m over F.
$(e_1, \dots, e_n)$ basis of V, $x = \sum_{i=1}^n x_i e_i \in V$.
$(\varepsilon_1, \dots, \varepsilon_n)$ dual basis of V^*, $\alpha = \sum_{i=1}^n x_i \varepsilon_i \in V^*$.

$$\begin{aligned}
\alpha(x) &= (\sum_{i=1}^n \alpha_i \varepsilon_i)(x) \\
&= \sum_{i=1}^n \alpha_i \varepsilon_i(x) \\
&= \sum_{i=1}^n \alpha_i \varepsilon_i(\sum_{j=1}^n x_j e_j) \\
&= \sum_{i,j} \alpha_i x_i \varepsilon_i(e_j) = \sum_{i=1}^n \alpha_i x_i
\end{aligned}$$

Definition. If $U \subset V$ (subset only) the **annihilator** of U is:

$$U^0 = \{\alpha \in V^* : \forall u \in U, \alpha(u) = 0\}$$

Lemma 2.3 (Annihilator is a subspace and finding its dimension)**.** (i) $U^0 \leq V^*$ (vector subspace)

(ii) If $U \leq V$ (vector subspace) and $\dim V < +\infty$ then

$$\dim V = \dim U + \dim U^0$$

Proof. (i) $0 \in U^0$ and if $\alpha, \alpha' \in U^0$ then

$$\forall u \in U, (\alpha + \alpha')(u) = \alpha(u) + \alpha'(i) = 0$$

$$\implies \alpha + \alpha' \in U^0$$

$$\forall u \in U, \forall \lambda \in F, (\lambda\alpha)(u) = \lambda\alpha(u) = 0$$

$$\implies \lambda\alpha \in U^0 \implies U^0 \leq V^*$$

(ii) Let $U \leq V$, $\dim V = n$ Let $(e_1, \dots, e_k)$ be a basis of U and complete it to a basis $\underbrace{(e_1, \dots, e_k, e_{k+1}, \dots, e_n)}_{\mathcal{B}}$ of V.

Let $(\varepsilon_1, \dots, \varepsilon_n) = \mathcal{B}^*$ be the dual basis of $\mathcal{B}$.

Claim.

$$U^0 = \langle \varepsilon_{k+1}, \dots, \varepsilon_n \rangle$$

If $i > k$, $\varepsilon_i(e_k) = S_{ik} = 0$, then

$$\varepsilon_i \in U^0$$

$$\implies \langle \varepsilon_{k+1}, \dots, \varepsilon_n \rangle \subset U_0$$

Conversely, let $\alpha \in U^0$. Then

$$\alpha = \sum_{i=1}^{n} \alpha_i \varepsilon_i$$

(B^* basis of V^*). For $i \leq k$

$$\alpha(e_i) = 0 \implies \alpha(e_i) = \sum_{j=1}^{n} \alpha_j \varepsilon_j(e_j) = \alpha_i$$

$$\implies \forall 1 \leq i \leq k,\ \alpha_i = 0$$

$$\implies \alpha = \sum_{k=1}^{n} \alpha_i e_i$$

so $\alpha \in \langle \varepsilon_{k+1}, \dots, \varepsilon_n \rangle$

$$\implies U^0 \subset \langle \varepsilon_{k+1}, \dots, \varepsilon_n \rangle$$

Definition. Let V, W be vector spaces over F and let $\alpha \in L(V, W)$. Then the map

$$\alpha^* : W^* \to V^*$$

$$\varepsilon \mapsto \varepsilon \circ \alpha$$

is an element of $L(W^*, V^*)$. It is called the **dual map** of α

Proof. $\varepsilon(\alpha) : V \to F$ linear by linearity of ε, α so $\varepsilon \circ \alpha \in V^*$.
α^* is linear as for $\theta_1, \theta_2 \in W^*$, then:

$$\alpha^*(\theta_1 + \theta_2) = (\theta_1 + \theta_2)(\alpha) = \theta_1 \circ \alpha + \theta_2 \circ \alpha = \alpha^*(\theta_1) + \alpha^*(\theta_1)$$

and similarly, $\forall \lambda \in F$:

$$\alpha^*(\lambda\theta) = \lambda\alpha^*(\theta)$$

thus

$$\alpha^* \in L(W^*, V^*)$$

Prop 2.4 (Writing dual map in dual basis)**.** Let V, W be finite dimensional vector spaces over F with basis repsectively $\mathcal{B}, \mathcal{C}$. Let $\mathcal{B}^*, \mathcal{C}^*$ be the dual basis of V^*, W^*. Then

$$[\alpha^*]_{\mathcal{C}^*, \mathcal{B}^*} = [\alpha]_{\mathcal{B}, \mathcal{C}}^T$$

Proof. It follows from the very definition

2.1 Properties of the Dual Map, Double Dual (Bidual)

Lemma 2.5 (Change of basis matrix for dual)**.** Change of basis matrix from $\mathcal{F}^* = (\eta_1, \ldots, \eta_n)$ to $\mathcal{E}^* = (\varepsilon_1, \ldots, \varepsilon_n)$ is $(P^{-1})^T$ where

$$P = [\mathrm{Id}]_{\mathcal{F}, \mathcal{E}}$$

$$\mathcal{E} = (e_1, \ldots, e_n) \text{ and } \mathcal{F} = (f_1, \ldots, f_n) \text{ bases of } V$$

Proof.

$$[\mathrm{Id}]_{\mathcal{F}^*, \mathcal{E}^*} = [\mathrm{Id}]_{\mathcal{F}, \mathcal{E}}^T = (P^{-1})^T$$

Lemma 2.6 (Nullity of dual is annihilator of image, image of dual is subspace of nullity of orginal map)**.** Let V, W be vector spaces over F. Let $\alpha \in L(V, W)$. Let $\alpha^* \in L(W^*, V^*)$ be the dual map. Then

(i)
$$N(\alpha^*) = (\text{Im } \alpha)^0$$
So α^* injective $\iff$ α surjective

(ii)
$$\text{Im } \alpha^* \leq (N(\alpha))^0$$
with equality iff V and W are finite dimensional.
(hence in this case, α^* surjective $\iff$ α injective)

Proof. (i) Let $\varepsilon \in W^*$. Then

$$\varepsilon \in N(\alpha^*) \iff \alpha^*(\varepsilon) = 0$$

$$\iff \alpha^*(\varepsilon) = \varepsilon \circ \alpha = 0$$
$$\iff \forall x \in V, \varepsilon(\alpha(x)) = 0$$
$$\iff \varepsilon \in (\text{Im } \alpha)^0$$

(ii) Let us first show that:
$$\text{Im } \alpha^* \leq (N(\alpha))^0$$
Indeed, let $\varepsilon \in \text{Im } \alpha^*$, then

$$\implies \varepsilon = \alpha^*(\varphi),\ \varphi \in W^*$$

$\implies \forall u \in N(\alpha)$

$$\begin{aligned}\varepsilon(u) &= (\alpha^*(\varphi))(u)\\ &= \varphi \circ \alpha(u)\\ &= \varphi(\alpha(u)) = 0\end{aligned}$$

$$\implies \varepsilon \in (N(\alpha))^0$$
$$\implies \text{Im } \alpha^* \leq (N(\alpha))^0$$

In finite dimension, we can compare the dimension of these two spaces:

$$\dim \text{Im } \alpha^* = r(\alpha^*) = r([\alpha^*]_{\mathcal{C}^*,\mathcal{B}^*}) = r([\alpha]^T_{\mathcal{B},\mathcal{C}}) = r([\alpha]_{\mathcal{B},\mathcal{C}})$$

$$\implies r(\alpha^*) = r(\alpha)$$

$$\begin{aligned}\dim \text{Im } \alpha^* &= r(\alpha^*)\\ &= \dim V - \dim N(\alpha)\\ &=\sim [(N(\alpha))^0]\end{aligned}$$

$$\implies \text{Im}\alpha^* \leq (N(\alpha))^0$$
$$\dim \text{Im } \alpha^* = \dim(N(\alpha))^0$$
$$\implies \text{Im } \alpha^* = (N(\alpha))^0$$

Note. In many applications, it is often simpler to understand α^* than α

2.1.1 Double Dual

Definition. Let V be a vector space over F

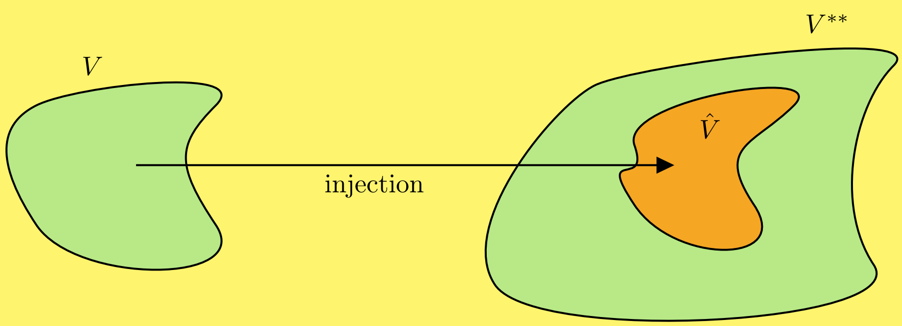
$$V^* = L(V, F) \text{ dual of } V$$

We define the **bidual** (double dual)

$$V^{**} = L(V^*, F) = (V^*)^*$$

Remark. This is a very important object. In general, there is no obvious relation between V and V^*. However, there are two fundamental facts:

(i) there is a CANONICAL embedding from V to V^{**}



Indeed, pick $v \in V$ and let

$$\hat{v} : V^* \to F$$
$$\varepsilon \mapsto \varepsilon(v)$$

Claim. $\hat{v} \in V^{**}$.

$\varepsilon \in V^*, \varepsilon(v) \in F$ and

$$\hat{v}(\lambda_1 \varepsilon_1 + \lambda_2 \varepsilon_2) = \lambda_1 \varepsilon_1(v) + \lambda_2 \varepsilon_2(v) = \lambda_1 \hat{v}(\varepsilon_1) + \lambda_2 \hat{v}(\varepsilon_2)$$

(ii) there are examples of infinite dimensional spaces where $V \simeq V^{**}$ (reflexive spaces, $L^P(\mathbb{R}^d)$)

$$L^P(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \mathbb{R}, \int_{\mathbb{R}^d} |f(x)|^p \,\mathrm{d}x < +\infty$$

Theorem 2.7 (Isomorphism between V and V^{**})**.** If V is finite dimensional, then

$$\hat{}: V \to V^{**}$$

$$v \mapsto \hat{v}$$

is an isomorphism

Proof. $\hat{}$ linear: trivial
$\hat{}$ injective: Indeed, let $e \in V \backslash \{0\}$. By extending e to a basis of V

$$(e, e_2, \ldots, e_n) \text{ basis of } V$$

Let $(\varepsilon, \varepsilon_2, \ldots, \varepsilon_n)$ be the dual basis of V^*, then:

$$\hat{e}(\varepsilon) = \varepsilon(e) = 1$$

$$\implies \hat{e} \neq \{0\}$$

$$\implies \hat{()} = \{0\}, \hat{} \text{ injective}$$

$\hat{}$ isomorphism: compute dimensions (trivial)

Moral. $\hat{}: V \to V^{**}$ isomorphism. This allows us to "identify" V and V^{**}

Lemma 2.8 (Annihilator of annihilator can be viewed as $\hat{U}$)**.** Let V be a finite dimensional vector space over F, and $U \leq V$. Then $\hat{U} = U^{00}$, so after identification of V and V^{**}, $U^{00} = U$

Proof. Let us show that $U \leq U^{00}$. Indeed, let $u \in U$ then

$$\forall \varepsilon \in U^0, \varepsilon(u) = 0$$

$$\implies \forall \varepsilon \in U^0, \varepsilon(u) = \hat{u}(\varepsilon) = 0$$

$$\implies \hat{u} = U^{00}$$

$$\implies \hat{U} \leq U^{00}$$

We compute dimensions

$$dim\hat{U} = \dim U = \dim U^{00}$$

$$\implies \hat{U} = U^{00}$$

Remark. Thanks to identification of V^{**} and V, we can define $T \leq V^*$

$$T^0 = \{v \in V : \theta(v) = 0 \forall \theta \in T\}$$

Lemma 2.9 (Annihilator of sums and intersections)**.** Let V be a finite dimensional vector space over F. Let $U_1, U_2 \leq V$, then

(i)
$$(U_1 + U_2)^0 = U_1^0 \cap U_2^0$$

(ii)
$$(U_1 \cap U_2)^0 = U_1^0 + U_2^0$$

Proof. trivial

Definition. Let U, V be vector spaces over F. Then $\varphi : U \times V \to F$ is a **bilinear form** if it is linear in both components

$$\varphi(u, \cdot),\ V \to F \in V^* \ (\forall u \in U)$$
$$\varphi(\cdot, v),\ U \to F \in U^* \ (\forall v \in V)$$

Example. (i)
$$V \times V^* \to F$$
$$(v, \theta) \mapsto \theta(v)$$

(ii) Canonical model: scalar product on $U = V = \mathbb{R}^n$

$$\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$
$$(x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}) \mapsto \sum_{i=1}^{n} x_i y_i$$

(iii) $U = V = \mathcal{C}([0,1], \mathbb{R})$
$$\varphi(f, g) = \int_0^1 f(t) g(t) \, \mathrm{d}t$$

infinite dimensional product (L^2)

Definition. $\mathcal{B} = (e_1, \ldots, e_m)$ basis of U
$\mathcal{C} = (f_1, \ldots, f_m)$ basis of V
$\varphi : U \times V \to F$ bilinear form
The matrix of φ wrt $\mathcal{B}$ and $\mathcal{C}$ is

$$[\varphi]_{\mathcal{B},\mathcal{C}} = (\varphi(e_i, f_j))_{1 \leq i \leq m, 1 \leq j \leq n}$$

Lemma 2.10 (Computing bilinear form)**.**

$$\varphi(u,v) = [u]_{\mathcal{B}}^T[\varphi]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{C}}$$

Proof.

$$u = \sum_{i=1}^{m} \lambda_i e_i$$

$$v_i = \sum_{j=1}^{n} \mu_j f_j$$

then by linearity

$$\begin{aligned}\varphi(u,v) &= \varphi(\sum_{i=1}^{m} \lambda_i e_i, \sum_{j=1}^{n} \mu_j f_j) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} \lambda_i \mu_j \varphi(e_i, e_j) \\ &= [u]_{\mathcal{B}}^T[\varphi]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{C}}\end{aligned}$$

Remark. $[\varphi]_{\mathcal{B},\mathcal{C}}$ is the only matrix such that

$$\varphi(u,v) = [u]_{\mathcal{B}}^T[\varphi]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{C}}$$

Notation. $\varphi : U \times V \to F$ bilinear form, then φ induces two linear maps:

$$\varphi_L : U \to V^*, \varphi_L(u) : V \to F$$

$$v \mapsto \varphi(u,v)$$

$$\varphi_R : V \to U^*, \varphi_R(v) : U \to F$$

$$u \mapsto \varphi(u,v)$$

In particular, by the very definition

$$\forall (u,v) \in U \times V$$

$$\varphi_L(u)(v) = \varphi(u,v) = \varphi_R(v)(u)$$

Lemma 2.11 (Writing left and right maps in terms of bases)**.** $\mathcal{B} = (e_1, \ldots, e_m)$ basis of U
$\mathcal{B}^* = (\varepsilon_1, \ldots, \varepsilon_m)$ dual basis
$\mathcal{C} = (f_1, \ldots, f_m)$ basis of V
$\mathcal{C}^* = (\eta_1, \ldots, \eta_m)$ dual basis
Let $A = [\varphi]_{\mathcal{B},\mathcal{C}}$, then

$$[\varphi_R]_{\mathcal{C},\mathcal{B}^*} = A,\ [\varphi_L]_{\mathcal{B},\mathcal{C}^*} = A^T$$

Proof.

$$\varphi_L(e_i)(f_j) = \varphi(e_i, f_j) = A_{ij}$$
$$\implies \varphi_L(e_i) = \sum A_{ij}\eta_j$$
$$\varphi_R(f_j)(e_i) = \varphi(e_i, f_j) = A_{ij}$$
$$\implies \varphi_R(f_j) = \sum_i A_{ij}\varepsilon_i$$

Definition.

$$\ker \varphi_L \equiv \textbf{ left kernel } \text{of } \varphi$$
$$\ker \varphi_R \equiv \textbf{ right kernel } \text{of } \varphi$$

Definition. We say that φ is **non-degenerate** if

$$\ker \varphi_L = \{0\} \text{ and } \ker \varphi_R = \{0\}$$

Otherwise, we say that φ is **degenerate**

Lemma 2.12 (non degenerate iff invertible)**.** $\mathcal{B}$ basis of U and $\mathcal{C}$ basis of V (U, V finite dimensional)

$$\varphi : U \times V \to F \text{ bilinear form}$$

$$A = [\varphi]_{\mathcal{B},\mathcal{C}}$$

Then φ non degenerate $\iff$ A is invertible

Proof. φ non degenerate iff both kernels $\{0\}$ iff

$$n(A^T) = 0 \text{ and } n(A) = 0$$
$$\iff r(A^T) = \dim U \text{ and } r(A) = \dim V$$
$$\iff A \text{ invertible}$$

and this forces $\dim U = \dim V$

Corollary 2.13 (non-degenerate forces same dimension)**.** If φ is non degenerate then

$$\dim U = \dim V$$

Remark. Cannonical example of non degenerate bilinear form

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

$$(x, y) \mapsto \sum_{i=1}^{n} x_i y_i$$

Corollary 2.14 (choosing non degenerate bilinear form same as choosing isomorphism)**.** When U and V are finite dimensional with the same dimension, then choosing a non degenerate bilinear form $\varphi : U \times U \to F$ is equivalent to choosing an isomorphism $\varphi_L : U \to V^*$

Definition. (i) $T \subset U$, we define

$$T^{\perp} = \{v \in V : \varphi(t, v) = 0, \forall t \in T\}$$

(ii) $S \subset V$

$$^{\perp}S = \{u \in U : \varphi(u, s) = 0 \forall s \in S\}$$

"**orthogonal**" of respectively T and S

Prop 2.15 (Change of basis formula for bilinear forms)**.** $\mathcal{B}, \mathcal{B}'$ basis of U, $P = [\mathrm{Id}]_{\mathcal{B}',\mathcal{B}}$
$\mathcal{C}, \mathcal{C}'$ basis of V, $Q = [\mathrm{Id}]_{\mathcal{C}',\mathcal{C}}$

$$\varphi : U \times V \to F \text{ bilinear form}$$

then:

$$[\varphi]_{\mathcal{B}',\mathcal{C}'} = P^T [\varphi]_{\mathcal{B},\mathcal{C}} Q$$

Proof.

$$\begin{aligned} \varphi(u, v) &= [u]_{\mathcal{B}}^T [\varphi]_{\mathcal{B},\mathcal{C}} [v]_{\mathcal{C}} \\ &= (P[u]_{\mathcal{B}'})^T [\varphi]_{\mathcal{B},\mathcal{C}} (Q[v]_{\mathcal{C}'}) \\ &= [u]_{\mathcal{B}'}^T (P^T [\varphi]_{\mathcal{B},\mathcal{C}} Q) [v]_{\mathcal{C}'} \\ &= [u]_{\mathcal{B}'}^T [\varphi]_{\mathcal{B}',\mathcal{C}'} [v]_{\mathcal{C}'} \end{aligned}$$

Definition. The rank of φ ($r(\varphi)$) is the rank of any matrix representing φ

Remark.

$$r(\varphi) = r(\varphi_R) = r(\varphi_L)$$

where we used $r(A) = r(A^T)$

3 Determinant and Traces

3.1 Trace

Definition. Let $A = M_n(F)$ (square matrix of size n). We definethe **trace of A** as

$$\text{tr } A = \sum_{i=1}^{n} A_{ii}$$

Remark. $M_n(F) \to F$, $A \mapsto \text{tr } A$ is a linear form

Lemma 3.1 (Can cycle around when working out trace)**.**

$$\text{Tr}(AB) = \text{Tr}(BA), \forall A, B \in M_n(F)$$

Proof.

$$\begin{aligned}\text{Tr}(AB) &= \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} b_{ji} \right) \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} b_{ji} a_{ij} \\ &= \text{Tr}(BA)\end{aligned}$$

Corollary 3.2. Similar matrices have the same trace

Proof. trivial

Definition. If $\alpha : V \to V$ linear, we can define $\text{Tr}(\alpha) = \text{Tr}([\alpha]_{\mathcal{B}})$ in any basis $\mathcal{B}$

Lemma 3.3 (Trace of map same as trace of dual)**.** $\alpha : V \to V$ linear. $\alpha^* : V^* \to V^*$ dual map. Then

$$\text{Tr } \alpha = \text{Tr } \alpha^*$$

Proof. Trivial by choosing a basis then trace of transpose same.

3.2 Determinants

3.2.1 Permutations and Transpositions

Definition. $S_n \equiv$ group of permutations of $\{1, 2, \dots, n\}$,

$$\sigma : \{1, \dots, n\} \to \{1, \dots, n\} \text{ bijection}$$

σ is a **permutation**

Definition. $k \neq l$, $\tau_{kl} \in S_n$ exchanges k and l, other elements are unchanged

Remark. Recall any permutation σ can be decomposed as a product of transpositions

$$\sigma = \prod_{i=1}^{n_\sigma} \tau_i$$

τ_i transposition

Definition. The **signature** of a permutation

$$\varepsilon : S_n \to \{-1, 1\}$$

$$\sigma \mapsto \begin{cases} 1 & \text{if } n_\sigma \text{ even} \\ 0 & \text{if } n_\sigma \text{ odd} \end{cases}$$

ε is a homomorphism:

$$\varepsilon(\sigma \circ \sigma') = \varepsilon(\sigma)\varepsilon(\sigma')$$

Definition (Leibniz Formula)**.** For

$$A \in M_n(F), A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$$

we define the **determinant** of A:

$$\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n}$$

has $n!$ summands and one term for each column and each row

Lemma 3.4 (Upper triangular has determinant zero)**.** If $A = (a_{ij})$ is an upper (lower) trianglar matrix:

$$a_{ij} = 0 \text{ for } i \geq j \text{ (resp } i < j)$$

then $\det A = 0$

Proof. Some term in the summand is zero (need $\sigma(j) \leq j$)

Lemma 3.5 (Determinant of transpose is the same)**.**

$$\det A = \det A^T$$

Proof. Same proof as in Vectors and Matrices, change sum by summing σ^{-1} instead

Definition. A **volume form** d on F^n is a function

$$\underbrace{F^n \times \dots \times F^n}_{n} \to F$$

such that

(i) d multilinear: for any $1 \le i \le n$, $\forall (v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n) \in F^n \times \dots \times F^n$

$$F^n \to F, v \mapsto d(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n)$$

is linear ($\in (F^n)^*$). I.e. d linear with respet to any entry

(ii) d alternate: if $v_1 = v_j$ for some $i \ne j$, then

$$d(v_1, \dots, v_n) = 0$$

We want to show that (up ot multiplication by a scalar), there is only one volume form on $F^n \times \dots \times F^n$ and it is given by the determinant.

Lemma 3.6 (Mapping columns to determinant is volume form)**.** Let

$$A = (a_{ij}) = \left[A^{(1)} \mid \quad \dots \quad \mid A^{(n)} \right]$$

Then

$$(A^{(1)}, \dots, A^{(n)}) \mapsto \det A$$

is a volume form

Proof. (i) True as product only contains one term in each column

(ii) Consider τ which exchanges k and l for $k \ne l$. Then $a_{ij} = a_{i\tau j}$ and since

$$S_n = A_n \cup \tau A_n$$

we can compute $\det A$ using the disjoin decomponsition and see that $\sigma \in A_n$ cancels with $\sigma \in \tau A_n$

Lemma 3.7 (Volume forms change sign on swapping two entries)**.** Let d be a volume form. Then swapping two entries changes the sign:

$$d(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = -d(v_1, \dots, v_j, \dots, v_i, \dots, v_n)$$

Proof. Indeed

$$d(v_1, \dots, v_i + v_j, \dots, v_i + v_j, \dots, v_n) = 0$$

Then we apply linearity

Corollary 3.8 (Calculating volume form of permutation of columns)**.** $\sigma \in S_n$, d volume form

$$d(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \varepsilon(\sigma) d(v_1, \dots, n_n)$$

Proof.

$$\sigma = \prod_{i=1}^{n_\sigma} \tau_i$$

where τ_i are transpositions

Theorem 3.9 (Computing volume form on columns of a matrix)**.** Let d be a volume form on F^n. Let $A = \left[A^{(1)} \mid \quad \dots \quad \mid A^{(n)}\right]$. Then

$$d(A^{(1)}, \dots, A^{(n)}) = (\det A) d(e_1, \dots, e_n)$$

$\det A$ is the only volume form such that

$$d(e_1, \dots, e_n) = 1$$

Proof. Write out coordinates and keep applying linearity and recognise d is alternate so require all the i_k to be different so rewrite as a permutation and use above corollary

Corollary 3.10 (Significance of det)**.** det is the unique volume form such that $d(e_1, \dots, e_n) = 1$

3.3 Some Properties of Determinants

Lemma 3.11 (Det is multiplicative)**.** Let $A, B \in M_n(F)$. Then

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

Proof.

$$d_A : F^n \times \cdots \times F^n \to F$$

$$(v_1, \ldots, v_n) \mapsto \det(Av_1, \ldots, Av_n)$$

d_A is multilinear ($v_i \mapsto Av_i$ is linear) as det multilinear.
d_A is alternate: (trivial check)
Thus d_A is a volume form so $\exists C_A$ s.t.

$$d_A(v_a, \ldots, v_n) = C_A \det(v_1, \ldots, v_n)$$

And letting $v_i = e_i$ gives us $C_A = \det A$.
Then consider $d_A(B_1, \ldots, B_n)$.

Definition. $A \in M_n(F)$, we say that:
(i) A is singular if $\det A = 0$
(ii) A is non singular if $\det A \neq 0$

Lemma 3.12 (Invertible implies non-singular)**.** A is invertible $\implies$ A is non singular

Proof. A invertible $\implies \exists A^{-1} \in M_n(F)$ s.t.

$$AA^{-1} = A^{-1}A = I_n$$

$$\implies \det\left(AA^{-1}\right) = (\det A)[\det(A)^{-1}]$$

$$\implies \det A \neq 0$$

Remark. We have proved that A ivertible $\implies \det A \neq 0$ and

$$\det\left(A^{-1}\right) = \frac{1}{\det A}$$

Theorem 3.13 (Invertible $\iff$ non-singular $\iff$ $r(A) = n$)**.** Let $A \in M_n(F)$. TFAE

(i) A is invertible

(ii) A is non singular

(iii) $r(A) = n$

Proof. (i) $\iff$ (iii) done (rank-nullity Theorem)
(i) $\implies$ (ii) is Lemma above. Need to show that (ii) $\implies$ (iii). Assume $r(A) < n$. Let us show that

$$\det A = 0$$

$$r(A) < n \implies \dim textSpan(A_1, \dots, A_n) < n$$

$$\implies \exists(\lambda_1, \dots, \lambda_n) \neq (0, 0, \dots, 0) \text{ s.t. } \sum_{i=1}^{n} \lambda_i A_i = 0$$

Let's say $\lambda_j \neq 0$, then:

$$A_j = -\frac{1}{\lambda_j} \sum_{i \neq j} \lambda_i A_i$$

$$\begin{aligned}\implies \det A &= \det(A_1, \dots, A_j, \dots, A_n) \\ &= \det\left(A_1, \dots, -\frac{1}{\lambda_j} \sum_{i \neq j} \lambda_i A_i, \dots, A_n\right) \\ &= -\frac{1}{\lambda_j} \sum_{i \neq j} \det(A_1, \dots, A_i, \dots, A_n) \\ &= 0\end{aligned}$$

Remark. Theorem gives the sharp critereon for invertibility of a set on n linear equations with n unknowns:

$$\mathbf{Y} \in F^n, A \in M_n(F)$$

$$AX = \mathbf{Y} \text{ with } X \in F^n \text{ has a unique solution } X \text{ for every } \mathbf{Y}$$

$$\iff \det A \neq 0$$

Lemma 3.14 (Determinant property of the linear map)**.** Conjugate matrices have the same determinant

Proof. trivial

Definition. $\alpha : V \to V$ linear. We define

$$\det \alpha = \det[\alpha_{\mathcal{B}}]$$

where B is any basis of V. This number does not depend on the choice of the basis.

Theorem 3.15 (Reformulation of previous facts in terms of linear maps)**.**

$$\det : L(V,V) \to F$$

satisfies:

(i)

$$\det\ \mathrm{Id} = 1$$

(ii)

$$\det(\alpha\beta) = (\det\alpha)(\det\beta)$$

(iii)

$$\det\alpha \neq 0 \iff \alpha \text{ is invertible}$$

and in this case:

$$\det\left(\alpha^{-1}\right) = \frac{1}{\det\alpha}$$

Proof. reformulation of above

Lemma 3.16 (Determinant of matrices with corner block of 0s)**.** $A \in M_k(F),\ B = M_l(F),\ C \in M_{k,l}(F)$. Let

$$N = \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \in M_n(F)$$

for $n = k + l$, then $\det N = (\det A)(\det B)$

Proof. I need to compute

$$\det N = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i} \tag{*}$$

Observe $m_{\sigma(i)i} = 0$ if $i \leq k,\ \sigma(i) > k$. Thus, in (*), we need only sum over $\sigma \in S_n$ such that

(i)

$$\forall j \in [1,k], \sigma(j) \in [1,k]$$

($[1,k] = \{1,\ldots,k\}$)

(ii)

$$\forall j[k+1,n], \sigma(j) \in [k+1,n]$$

(iii) In other words, in (*), we can consider σ decomposed into σ_1 permuting $\{1,\ldots,k\}$ and σ_2 permuting $\{k+1,\ldots,n\}$

$$\det N = \sum_{\sigma_1 \in S_k, \sigma_2 \in S_l} \varepsilon(\sigma_1)\varepsilon(\sigma_2) \prod_{i=1}^{k} a_{\sigma_1(i)i} \prod_{i=k+1}^{n} b_{\sigma(k)k} = (\det A)(\det B)$$

Corollary 3.17 (determinant of diagonal blocks with 0s below)**.** $A_1, \dots, A_N$ are square matricies, then:

$$\det \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_N \end{bmatrix} = (\det A_1) \dots (\det A_N)$$

Proof. Induct on number of blocks

Warning.

$$\det \begin{bmatrix} A & C \\ D & B \end{bmatrix} \neq \det A \det B - \det C \det D$$

for A, B, C, D square

Remark. (i) Reasoning behind name 'volume form'

$$\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$$

$$(a, b, c) \mapsto (a \times b) \cdot c$$

where

$$a \times b = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix}$$

Considering a parallelopiped with edges vectors a, b, c, we see

$$d(a, b, c) = \text{ signed volume of parallelopiped}$$

(ii)

$$\det(a, b, c) = (a \times b) \cdot c$$

3.4 Adjugate Matrix

Equation. For $A \in M_N(F)$, $A = (A^{(1)} | \dots | A^{(n)})$. We have that swapping two columns in determinant swaps the sign. Since $\det A = \det A^T$, we similarly see that swapping two rows chances the sign of the determinant

Remark. We could prove all properties of determinants using the decomposition of A into elementary matrices.

3.5 Column (row) Expansion and the Adjugate Matrix

Column expansion aims to reduce the computation of $n \times n$ determinants to $(n-1) \times n-1)$ determinants to reduce dimension

Definition. For $A \in M_n(F)$, pick $i, j \in \{1, \ldots, n\}$. We define

$$A_{\hat{i}\hat{j}} \in M_{n-1}(F)$$

obtained by removing the i-th row and the j-th column from A.

Example.

$$A = \begin{bmatrix} 1 & 2 & 7 \\ 2 & 1 & 0 \\ -3 & 6 & 1 \end{bmatrix}$$

$$A_{\hat{3}\hat{2}} = \begin{bmatrix} 1 & -7 \\ 2 & 0 \end{bmatrix}$$

Lemma 3.18 (Expansion of the Determinant)**.** Let $A \in M_n(F)$

(i) Expansion with respect to the j-th column: pick $1 \leq j \leq n$, then:

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{\hat{i}\hat{j}} \tag{*}$$

(ii) Expansion with respect to the i-th row: pick $1 \leq i \leq n$, then

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

Proof. Expansion with respect to the j-th column (row expansion formula then follows by taking transpose). We have result as a direct consequence of the volume form property: pick $1 \leq j \leq n$

$$A^{(j)} = \sum_{i=1}^{n} a_{ij} e_j,\ (e_i)_{1 \leq i \leq n}$$

Canonical basis

$$\begin{aligned}\det A &= \det\left(A^{(1)}, \cdots \sum_{i=1}^{n} a_{ij} e_j, \ldots, A^{(n)}\right)\\ &= \sum_{i=1}^{n} a_{ij} \det\left(A^{(1)} | \ldots | e_i | \ldots, A^{(n)}\right)\end{aligned}$$

$$\begin{aligned}\det\left(A^{(1)} | \ldots | e_i | \ldots, A^{(n)}\right) &= (-1)^{j-1} \det\left(e_i | A^{(2)} | \ldots | A^{(n)}\right)\\ &= (-1)^{j-1}(-1)^{i-1} \det \begin{bmatrix} 1 & a_{i1} & a_{i2} & \ldots & a_{in} \\ 0 & & & & \\ \vdots & & A_{\hat{i}\hat{j}} & & \\ 0 & & & & \end{bmatrix}\\ &= (-1)^{i+j} \det A_{\hat{i}\hat{j}}\end{aligned}$$

We have proved

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

Remark. We have proved that

$$\det\left(A^{(1)}, \ldots, A^{(j-1)}, e_i, A^{(j+1)}, \ldots, A^{(n)}\right) = (-1)^{i+j} \det\left(A_{\hat{i}\hat{j}}\right)$$

Definition (Adjugate matrix)**.** Let $A \in M_n(F)$. The adjugate matrix $\operatorname{adj}(A)$ is the $n \times n$ matrix with (i, j) entry given by:

$$(-1)^{i+j} \det\left(A_{\hat{j}\hat{i}}\right)$$

$$\det\left(A^{(1)}, \dots, A^{(j-1)}, e_i, A^{(j+1)}, \dots, A^{(n)}\right) = (\operatorname{adj}(A))_{ji}$$

Theorem 3.19 (Adjugate key property)**.** Let $A \in M_n(F)$, then

$$\operatorname{adj}(A)A = (\det A)I_d = \begin{bmatrix} \det A & & 0 \\ 0 & \ddots & \\ 0 & & \det A \end{bmatrix}$$

In particular, when A is invertible:

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

Proof. We just proved:

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} \det A_{\hat{i}\hat{j}} a_{ij} = \sum_{i=1}^{n} (\operatorname{adj}(A))_{ji} a_{ij} = (\operatorname{adj}(A)A)_{jj}$$

For $j \neq k$, we have

$$\begin{aligned} 0 &= \det\left(A^{(1)}, \dots, A^{(k)}, \dots, A^{(k)}, \dots, A^{(n)}\right) \\ &= \det\left(A^{(1)}, \dots, \sum_{i=1}^{n} a_{ik} e_i, \dots, A^{(k)}, \dots, A^{(n)}\right) \\ &= \sum_{i=1}^{n} (\operatorname{adj}(A))_{ji} a_{ik} \\ &= (\operatorname{adj}(A)A)_{jk} \end{aligned}$$

3.6 Cramer Rule

Prop 3.20 (Solving linear equations)**.** Let $A \in M_n(f)$ invertible and let $b \in F^n$. Then the unique solution to $Ax = b$ is given by

$$x_i = \frac{1}{\det A} \det(A_{\hat{i}b}),\ 1 \leq i \leq n$$

where $A_{\hat{i}b}$ is the matrix obtained by replacing the ith column of A by b

Proof. A invertible implies $\exists! x \in F^n : Ax = b$. Let x be this solution, then:

$$\begin{aligned}
\det(A_{\hat{i}b}) &= \det\Big(A^{(1)}, \ldots, A^{(i-1)}, b, A^{(i+1)} \ldots, A^{(n)}\Big) \\
&= \det\Big(A^{(1)}, \ldots, A^{(i-1)}, Ax, A^{(i+1)}, \ldots, A^{(n)}\Big) \\
&= x_i \det\Big(A^{(1)}, \ldots, A^{(i)}, \ldots, A^{(n)}\Big) \\
&= x_i \det A
\end{aligned}$$

4 Eigenvectors, Eigenvalues and Trigonal Matrices

Moral. This is the first step twoards diagonalisation of endomorphisms

Let V be a vector space over F with $\dim V = n < +\infty$. Let $\alpha : V \to V$ be a linear map (endomorphism of V). Can we find a basis $\mathcal{B}$ of V such that in this basis,

$$[\alpha]_{\mathcal{B}} = \alpha_{\mathcal{B},\mathcal{B}}$$

has a "nice" form?

Equation. Reminder: If $\mathcal{B}'$ is another basis and P is the change of basis matrix,

$$[\alpha]_{\mathcal{B}'} = P^{-1}[\alpha]_{\mathcal{B}}P$$

Definition. (i) $\alpha \in L(V)$ ($\alpha : V \to V$ linear) is **diagonalisable** if $\exists \mathcal{B}$ basis of V such that

$$[\alpha]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \text{ (diagonal)}$$

(ii) α is **triangulable** if $\exists \mathcal{B}$ basis of V such that $[\alpha]_{\mathcal{B}}$ is triangulat.

$$[\alpha]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & * & * \\ & \ddots & * \\ & & \lambda_n \end{bmatrix}$$

Remark. This can be expressed equivalently in terms of conjugation of matrices

Definition. (i) $\lambda \in F$ is an **eigenvalue** of $\alpha \in L(V)$ iff:

$$\exists v \in V \backslash \{0\} : \alpha(v) = \lambda v$$

(ii) $v \in V$ is an **eigenvector** of α iff

$$v \neq 0 \text{ and } : \exists \lambda \in F : \alpha(v) = \lambda v$$

(iii)

$$V_\lambda = \{v \in V : \alpha(v) = \lambda v\} \leq V$$

is the **eigenspace** associated to λ

Lemma 4.1 (Eigenvalue in terms of determinant)**.** $\alpha \in L(v), \lambda \in F$. λ an eigenvalue $\iff$ $\det(a - \lambda \text{ Id}) = 0$

Proof.

$$\begin{aligned}\lambda \text{ eigenvalue} &\iff \exists v \in V\backslash\{0\} : \alpha(v) = \lambda v \\ &\iff \exists v \in V\backslash\{0\} : (\alpha - \lambda \text{ Id})(v) = 0 \\ &\iff \ker(\alpha - \lambda \text{ Id}) \neq \{0\} \\ &\iff \alpha - \lambda \text{ Id nor injective} \\ &\iff \alpha - \lambda \text{ Id not surjective} \\ &\iff \alpha - \lambda \text{ Id not invertible} \\ &\iff \det(\alpha - \lambda \text{ Id}) = 0\end{aligned}$$

Remark. If $\alpha(v_j) = \lambda v_j j$, v_j eigenvector, $v_j \neq 0$. I can complete $(v_1, \dots, v_j, \dots, v_n) = \mathcal{B}$ basis of V

4.1 Elementary Facts About Polynomials

For F a field,

- For
$$f(t) = a_n t^n + \cdots + a_n t + a_0,\ a_i \in F$$
n is the largest exponent such that $a_n \neq 0$, $n = \deg f$
- $\deg\{f + g\} \leq \max\{\deg f, \deg g\}$
- $\deg fg = \deg f + \deg g$
- $F[t] = \{\text{polynomials with coefficients in } F\}$
- λ root of $f \iff f(\lambda) = 0$

Lemma 4.2 (root of polynomial gives factor)**.** λ root of $f \implies (t - \lambda)$ divides f

Proof.
$$f(t) = a_n t^n + \cdots + a_1 t + a_0$$
$$f(\lambda) = a_n \lambda^n + \cdots + a_1 \lambda + a_0 = 0$$

$$\begin{aligned}\implies f(t) &= f(t) - f(\lambda) \\ &= a_n(t^n - \lambda^n) + \cdots + a_1(t - \lambda)\end{aligned}$$

Remark. We say that λ is a root of multiplicity k if $(t - \lambda)^k$ divides f, but $(t - \lambda)^{k+1}$ does not.

$$f(t) = (t - \lambda)^k g(t),\ g(\lambda) \neq 0$$

Corollary 4.3 (Bound on number of roots)**.** A non-zero polynomial of degree n (≥ 0) has at most n roots (counted with multiplicity)

Proof. Induction on the degree

Corollary 4.4 (Agreeing on too many points implies equivalent)**.** For $f_1,\ f_2$ polynomials of degree $< n$ s.t.

$$f_t(t_i) = f_2(t_i),\ (t_i)_{1\leq i\leq n}\ n \text{ distinct values}$$

we have $f_1 \equiv f_2$

Proof. $f_1 - f_2$ has degree $< n$ and n distinct roots so $f_1 - f_2 \equiv 0$

Theorem 4.5 (FTA)**.** Any $f \in \mathbb{C}[t]$ of positive degree has a (complex) root, hence exacty $\deg f$ roots when counted with multiplicity

Note. $f \in \mathbb{C}[t] = c\prod_{i=1}^{r}(t-\lambda_i)^{\alpha_i},\ c \in \mathbb{C}, \lambda_i \in \mathbb{C},\ \alpha_i \in \mathbb{N}^+$

Definition. For $\alpha \in L(v)$, the **characteristic polynomial** of α is

$$\chi_\alpha(\lambda) = \det(\alpha - \lambda \text{ Id})$$

Remark. The fact that χ_α is a polynomial in λ follows from the very definition of the determinant

Remark. Conjugate matrices have the same characteristic polynomial: (trivial to show)

Theorem 4.6 (Triangulable same as being able to write char poly as linear factors)**.** $\alpha \in L(V)$ is triangulable $\iff \chi_\alpha(t)$ can be written as a product of linear factors over F:

$$\chi_\alpha(t) = c\prod_{i=1}^{n}(t-\lambda_i)$$

If $F = \mathbb{C}$, every matrix is triangulable

Proof. $\implies$: suppose α triangulable

$$[\alpha]_{\mathcal{B}} = \begin{bmatrix} a_1 - t & & 0 \\ & \ddots & \\ 0 & & a_n - t \end{bmatrix} = \prod_{i=1}^{n}(a_i - t)$$

$\impliedby$: we argue by induction on $n = \dim V$:

- $n = 1$ trivial
- $n > 1$ by assumption, let $\chi_\alpha(t)$ have a root λ. Then:

$$\chi_\alpha(\lambda) = 0 \iff \lambda \text{ eigenvalue of } \alpha$$

Let $U = V_\lambda \equiv$ eigenspace associated to λ. Let $(v_1, \ldots, v_k)$ be a basis of U. We complete it to $(v_{k+1}, \ldots, v_n)$ basis of V. $\text{Span}(v_{k+1}, \ldots, v_n) = W$, $V = U \oplus W$. $\mathcal{B} = (v_1, \ldots, v_n)$

$$[\alpha]_{\mathcal{B}} = \begin{bmatrix} \lambda \text{ Id} & * \\ 0 & C \end{bmatrix}$$

α induces an endomorphism: $\overline{\alpha} : V/U \to V/U$, C represents $\overline{\alpha}$ wrt $(v_{k+1}+U, \ldots, v_n+U)$. By induction (since $k \geq 1$), we know that we can find a basis $(\tilde{v}_{k+1}, \ldots, \tilde{v}_n)$ in which C has a a triangular form T.

$$[\alpha]_{\tilde{\mathcal{B}}} = \begin{bmatrix} \lambda \text{ Id} & * \\ 0 & T \end{bmatrix}$$

$$\implies \alpha \text{ has a triangular form}$$

Lemma 4.7 (Char poly coefficients)**.** If V is n dimensional over $F = \mathbb{R}$ or $\mathbb{C}$ and $\alpha \in L(V)$. Then:

$$\chi_\alpha(t) = (-1)^n t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0$$

with $c_{n-1} =$ Trα and $c_0 = \det\alpha$

Proof.

$$\chi_\alpha(t) = \det(\alpha - t \text{ Id})$$

$$\chi_\alpha(0) = \det\alpha = c_0$$

$F = \mathbb{R}, \mathbb{C}$, we know that α is triangulable over $\mathbb{C}$:

$$\chi_\alpha(t) = \begin{bmatrix} a_1 - t & & \\ & \ddots & \\ & & a_n - t \end{bmatrix} = \prod_{i=1}^{n}(a_i - t)$$

which gives us the form as desired once we expand

5 Diagonalisation Critereon and Minimal Polynomial

Notation. For $p(t)$ polynomial over F and $p(t) = a_n t^n + \cdots + a_1 t + a_0,\ a_i \in F$. For $A \in M_n(F)$, we define

$$p(A) = a_n A^n + \cdots + a_1 A + a_0 \text{ Id} \in M_n(F)$$

$\alpha \in L(V)$, ($\alpha : V \to V$ linear)

$$p(\alpha) = a_n \alpha^n + \cdots + a_1 \alpha + a_0 \text{ Id} \in L(V)$$

$$\alpha^n = \underbrace{\alpha \circ \cdots \circ \alpha}_{n}$$

Theorem 5.1 (Sharp Criterion of Diagonalisability)**.** If V is a vector space over F with $\dim V = n < +\infty$, $\alpha \in L(V)$, then α is diagonalisable $\iff \exists$ a polynomial p which is a product of distinct linear factors such $p(\alpha) = 0$

Proof. α diagonalisable $\iff \exists(\lambda_1, \dots, \lambda_k)$ distinct such that

$$p(t) = \prod_{i=1}^{k}(t - \lambda_i) \implies p(\alpha) = 0$$

$\implies$: Suppose that α is diagonalisable, with $(\lambda_1, \dots, \lambda_k)$ the k distinct eigenvalues. Let

$$p(t) = \prod_{i=1}^{k}(t - \lambda_i)$$

Let $\mathcal{B}$ be a basis of V formed of eigenvectors. Let $v \in \mathcal{B}$ s.t. $\alpha(v) = \lambda_i v$ for some i. Then

$$(\alpha - \lambda_i \text{ Id})(v) = 0 \implies p(\alpha)(v) = \prod_{i=1}^{k} \underbrace{(\alpha - \lambda_i \text{ Id})}_{\text{commute}}(v) = 0$$

$$\implies \forall v \in \mathcal{B}, [p(\alpha)](v) = 0$$

$$\implies p(\alpha) = 0$$

$\impliedby$: Suppose $p(\alpha) = 0$ for some $p(t) = \prod_{i=1}^{k}(t - \lambda_i)$ with $\lambda_i \neq \lambda_j$ for $i \neq j$.
Let $V_{\lambda_i} = \ker(\alpha - \lambda_i \text{ Id})$. We claim:

$$V = \bigoplus_{i=1}^{k} V_{\lambda_i} \tag{*}$$

Indeed, let us consider the polynomials:

$$q_j(t) = \prod_{i=1, i\neq j}^{k} \frac{(t - \lambda_i)}{(\lambda_j - \lambda_i)}, 1 \leq j \leq k$$

By definition:

$$q_j(\lambda_i) = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases} = S_{ij}$$

Let us define the polynomial

$$q(t) = q_1(t) + \dots + q_k(t)$$

this has degree at most $k - 1$. On the other hand: for $1 \leq i \leq k$

$$q(\lambda_i) = 1 \implies q(t) = 1 \ \forall t$$

$$\implies \forall t, q_1(t) + \dots + \dots q_k(t) = 1 \tag{**}$$

Proof (continued)**.** Let us consider (projection factor),

$$\Pi_j = q_j(\alpha) \in L(V)$$

for $1 \leq j \leq k$. Then by construction, for (**)

$$\sum_{j=1}^{k} \Pi_j = \sum_{j=1}^{k} q_j(\alpha) = \text{ Id}$$

$$\sum_{j=1}^{k} q_j(t) = 1 \implies \sum_{j=1}^{k} q_j(\alpha) = \text{ Id}$$

$$\implies \forall v \in V, \quad \text{Id}(v) = v = \sum_{j=1}^{k} \Pi_j(v) \iff v = \sum_{j=1}^{k} q_j(\alpha)(v)$$

Observe

$$\begin{aligned}(\alpha - \lambda_j \text{ Id})q_j(\alpha)(v) &= \frac{1}{\prod_{i\neq j}(\lambda_j - \lambda_i)}(\alpha - \lambda_j \text{ Id})[\prod_{i=1, i\neq j}^{k}(t - \lambda_i)](\alpha)(v) \\ &= \frac{1}{\prod_{i\neq j}(\lambda_j - \lambda_i)}\prod_{i=1}^{k}(\alpha - \lambda_i \text{ Id})(v) = 0 \ \forall v \in V\end{aligned}$$

This means

$$\begin{aligned}(\alpha - \lambda_j \text{ Id})q_j(\alpha)(v) = 0 \ \forall v \in V \\ \implies (\alpha - \lambda_j \text{ Id})\Pi_j(v) = 0 \\ \implies \Pi_j(v) \in \ker(\alpha - \lambda_j \text{ Id}) = V_j \ \forall v \in V\end{aligned}$$

Proof (continued)**.** We have proved $\forall v \in V$

$$v = \sum_{j=1}^{k} \underbrace{\Pi_j(v)}_{\in V_j}$$

$$V = +_{j=1}^{k} V_j$$

It remains to show that the sum is direct: indeed for $v \in V_{\lambda_j} \cap (\sum_{i \neq j} V_{\lambda_i})$, we need to show that $v = 0$. Let us apply Π_j to $v \in V_{\lambda_j} \cap (\sum_{i \neq j} V_{\lambda_i})$. For $v \in V_{\lambda_j}$:

$$\Pi_j(v) = q_j(\alpha)(v) = \prod_{i=1, i \neq j}^{k} \frac{(\alpha - \lambda_i \text{ Id})(v)}{(\lambda_j - \lambda_i)} = \prod_{i=1, i \neq j}^{k} \frac{(\lambda_i - \lambda_i)}{(\lambda_j - \lambda_i)} v = v$$

(Π_j is the projector onto V_{λ_j}). For $v \in \sum_{i \neq j} V_{\lambda_i} = \sum_{i \neq j} \omega_i,\ \omega_i \in V_{\lambda_i}$:

$$\Pi_j(\omega_i) = \Pi_{m=1, m \neq j}^{k} \frac{(\alpha - \lambda_j \text{ Id})(\omega_i)}{(\lambda_m - \lambda_j)} = 0$$

$$\implies \Pi_j(v) = \sum_{i \neq j} \Pi_j(\omega_i) = 0$$

$$\implies \Pi_j(v) = 0$$

But $v = \Pi_j(v) \implies v = 0$ so

$$+_{i=1}^{k} V_{\lambda_i} = \bigoplus_{i=1}^{k} V_{\lambda_i}$$

We have proved

$$V = \bigoplus_{i=1}^{k} V_{\lambda_i}$$

$\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_k)$ with $\mathcal{B}_i$ basis of V_{λ_i} has $\mathcal{B}$ basis of V and so $[\alpha]_{\mathcal{B}}$ is diagonal

Remark. We have shown something more general: if $\lambda_1, \dots, \alpha_j$ are k distinct eigenvalues of α, then the sum

$$\sum_{i=1}^{k} V_{\lambda_i} = \bigoplus_{i=1}^{k} V_{\lambda_i}$$

The only way diagonalisation fails is if

$$\sum_{i=1}^{k} V_{\lambda_i} < V$$

Example. Many applications of the diagonalisation critereon, $F = \mathbb{C}, A \in M_n(F)$ such: A has finite order $\iff \exists m \in M$ s.t.$A^m =$ Id $\implies A$ is diagonalisable

$$t^m - 1 = p(t) = \prod_{j=0}^{m-1} (t - \xi_m^j)$$

where $\xi_m = e^{\alpha_i \pi / m}$ has $p(A) = 0$

Theorem 5.2 (Simultaneous Diagonalisation)**.** Let $\alpha, \beta \in L(V)$ diagonalisable. Then α, β are simultaneously diagonalisable (i.e. there exists a basis in which α, β have a diagonal matrix) iff α and β commute

Proof. $\implies$: $\exists \mathcal{B}$ s.t $[\alpha]_{\mathcal{B}} = D_1$ and $[\beta]_{\mathcal{B}} = D_2$ with $D-1, D_2$ diagonal, then

$$D_1 D_2 = D_2 D_1 \implies \alpha\beta = \beta\alpha$$

$\impliedby$: Suppose α, β diagonalisable

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$$

$\lambda_1, \dots, \lambda_k$ are the k distinct eigenvalues of α

Claim.

$$\beta(V_{\lambda_j}) \leq V_{\lambda_j}$$

(V_{λ_j} is stable by β)

Indeed: for $v \in V_{\lambda_j}$,

$$\alpha\beta(v) = \beta\alpha(v) = \beta(\lambda_j v) = \lambda_j \beta(v) \implies \alpha(\beta(v)) = \lambda_j \beta(v)$$

$$\implies \beta(v) \in V_{\lambda_j}$$

By assumption, β is diagonalisable so $\exists p$ with distinct linear factors such that $p(\beta) = 0$. Now

$$\beta(V_{\lambda_j}) \leq V_{\lambda_j} \implies B|_{V_{\lambda_j}} \in L(V_{\lambda_j})$$

I can compute $p(\beta|_{V_{\lambda_j}}) = 0$ so $\beta|_{V_{\lambda_j}}$ is diagonalisable. Now I take the $\mathcal{B}_i$ basis of V_{λ_i} in which $\beta|_{V_{\lambda_i}}$ is diagonal

Reminder: Euclidean algorithm for polynomials: given a, b polynomials over F with $b \neq 0$, there exist polynomials q, r over F with

$$\deg r < \deg b$$

$$a = qb + r$$

Definition. For V vector space over F, $\alpha \in L(V)$, $\dim(V) < +\infty$, the **minimal polynomial** m_α of α is the non-zero polynomial with smallest degree such that

$$m_\alpha(\alpha) = 0$$

Remark. For $\dim_F V = n < +\infty$ $\alpha \in L(V)$, have $\dim_F L(V) = n^2$ hence $\{\mathrm{Id}\ , \alpha, \dots, \alpha^{n^2}\}$ cannot be free

$$\implies \exists (a_0, \dots, a_n) \neq (0, \dots, 0)$$

s.t.

$$\underbrace{a_0 \,\mathrm{Id} + a_1\alpha + \cdots + a_{n^2}\alpha^{n^2}}_{p(\alpha)} = 0$$

Lemma 5.3 (minimal polynomial indeed minimal)**.** For $\alpha \in L(V)$, $p \in F[t]$, we have $p(\alpha) = 0 \iff m_\alpha$ is a factor of p (m_α divides p). (in particular, m_α is well defined)

Proof. $\deg m_\alpha < \deg p$ by minimality so Euclidean algorithm gives $p = m_\alpha q + r$ with $\deg r < \deg m_\alpha$ so

$$\begin{aligned} p(\alpha) = 0 &= m_\alpha(\alpha)q(\alpha) + r(\alpha) \\ &\implies r(\alpha) = 0 \end{aligned}$$

so $r \equiv 0 \implies m_\alpha q$
In particular, if m_1, m_2 are both minimal polynomials that kill α, then m_2 divides m_1 and vice versa so $m_2 = cm_1,\ c \in F$

Examples. $V = F^2$

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$p(t) = (t-1)^2 \implies p(A) = p(B) = 0$
minimal polynomial (of A or B) has to be either $t-1$ or $(t-1)^2$. We can check $m_A = t-1,\ m_B = (t-1)^2$. Thus A is diagonalisable but B is not diagonalisable

5.1 Cayley-Hamilton Theorem and Multiplicity of Eigenvectors

Theorem 5.4 (Cayley Hamilton)**.** Let V be a F vector space, $\dim_F V < +\infty$. Let $\alpha \in L(V)$ with characteristic polynomial

$$\chi_\alpha = \det(\alpha - t \text{ Id})$$

then

$$\chi_\alpha(\alpha) = 0$$

Proof. For $F = \mathbb{C}, \mathcal{B} = \{v_1, \dots, v_n\}$ and $n = \dim_F V$

$$[\alpha]_\mathcal{B} = \begin{bmatrix} a_1 & * & \\ \ddots & & \\ 0 & & a_n \end{bmatrix}$$

Let $U_j = \text{ span}\{v_1, \dots, v_j\}$

$$\chi)\alpha(t) = \prod_{i=1}^n (a_i - t)$$

$$\chi_\alpha(\alpha) = (\alpha - a_1 \text{ Id}) \dots (\alpha - a_{n-1} \text{ Id})(\alpha - a_n \text{ Id})$$

for $v \in V = U_n$

$$\begin{aligned} \chi_\alpha(\alpha)(v) &= (\alpha - a_1 \text{ Id}) \dots (\alpha - a_{n-1} \text{ Id}) \underbrace{(\alpha - a_n \text{ Id})(v)}_{\in U_{n-1}} \\ &= (\alpha - a_1 \text{ Id})(v) \\ &= 0 \end{aligned}$$

Proof (alternative)**.** For any field F. $A \in M_n(F)$

$$\begin{aligned}\det(t\ \mathrm{Id} - A) &= (-1)^n \chi_A(t)\\ &= t^n + a_{n-1}t^{n-1} + \cdots + a_0\end{aligned}$$

For any matrix B, we have proved

$$B \cdot\ \mathrm{adj}(B) = (\det B)\ \mathrm{Id} \tag{*}$$

Applying (*) to $B = t\ \mathrm{Id} - A$. Let

$$\mathrm{adj}(B) = B_{n-1}t^{n-1} + \cdots + B_1 t + B_0$$

(*) gives us

$$(t\ \mathrm{Id} - A)[B_{n-1}t^{n-1} + \cdots + B_1 t + B_0] = (t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0)I$$

Equating coefficients of t^k, we get:

$$\begin{aligned}\mathrm{Id} &= B_{n-1}\\ a_{n-1}\ \mathrm{Id} &= B_{n-2} - AB_{n-1}\\ &\vdots\\ a_0\ \mathrm{Id} &= -AB_n\end{aligned}$$

Multiplying the top by A^n, second by A^{n-1} etc. and summing both sides yields

$$A^n + a_{n-1}A^{n-1} + \cdots + a_0\ \mathrm{Id} = 0 \tag{**}$$

Definition. For $\dim_F V = n < +\infty$, $\alpha \in L(V)$ where λ an eigenvalue of α, $\chi_\alpha(t) = (t-\lambda)^{a_\lambda} q(t)$ for $q \in F[t], (t-\lambda) \nmid q$ has a_λ is the **algebraic multiplicity** of λ. We define $g_\lambda = \dim \ker(\alpha - \lambda\ \mathrm{Id})$ is the **geometric multiplicity** of λ (it is the dimension of the eigenspace associated to λ)

Remark. λ eigenvalue $\iff$ $\chi_\alpha(\lambda) = 0$

Lemma 5.5 (A.M $\geq$ G.M.)**.** λ eigenvalue of $\alpha(V)$ implies $1 \leq g_\lambda \leq a_\lambda$

Proof.

$$g_\lambda = \dim \ker(\alpha - \lambda\ \mathrm{Id})$$

λ eigenvalue $\implies \exists v \neq 0 : v \in \ker(\alpha - \lambda\ \mathrm{Id})$. So $g_\lambda = \dim\ker(\alpha - \lambda\ \mathrm{Id}) \geq 1$. Let $(v_1, \ldots, v_{g_\lambda})$ be a basis of $V_\lambda = \ker(\alpha - \lambda\ \mathrm{Id})$. Complete it to a basis $\mathcal{B} = (v_1, \ldots, v_{g_\lambda}, v_{g_\lambda+1}, \ldots, v_n)$ then

$$[\alpha]_{\mathcal{B}} = \begin{bmatrix} \lambda\ \mathrm{Id}_{g_\lambda} & * \\ 0 & A_1 \end{bmatrix}$$

for some A_1 and so

$$\det(a - t\ \mathrm{Id}) = (\lambda - t)^{g_\lambda} \underbrace{\det(A_1 - t\ \mathrm{Id})}_{\text{polynomial}} \implies g_\lambda \leq a_\lambda$$

Lemma 5.6 (Minimal multiplicity $\leq$ algebraic multiplicity)**.** Let λ be an eigenvlaue of α. Let c_λ be the multiplicity of λ as a root of the minimal polynomial m_α. Then $1 \leq c_\lambda \leq a_\lambda$

Proof. Caycley-Hamilton:

$$\chi_\alpha(\alpha) = 0 \implies m_\alpha \mid \chi_\alpha$$
$$\implies c_\lambda \leq a_\lambda$$

$c_\lambda \geq 1$ as if λ an eigenvalue then $\exists v \neq 0 : \alpha(v) = \lambda v$ and so $\alpha^p v = \lambda^p v$

Example.

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix}$$

$$\chi_A(t) = (t-1)^2(t-2)$$

m_A is either

- $(t-1)^2(t-2)$
- $(t-1)(t-2)$

we have (ii) hods so A is diagonalisable

Lemma 5.7 (Characterization of diagonalisable endomorphisms for $F = \mathbb{C}$)**.** Have $F = \mathbb{C}$. V an F vector space with $\dim F < \infty$, $\alpha \in L(V)$. TFAE:

(i) α is diagonalisable
(ii) $\forall \lambda$ eigenvalue of $\alpha, a_\lambda = g_\lambda$
(iii) $\forall \lambda$ eigenvalue of α, $c_\lambda = 1$

Proof. (i) $\iff$ (iii): already done.
(i) $\iff$ (ii): let $(\lambda_1, \dots, \lambda_k)$ be the distinct eigenvalues of α. We showed: α diagonalisable $\iff V = \bigoplus_{i=1}^k V_{\lambda_i}$

$$\dim V = n = \sum_{i=1}^k a_{\lambda_i}$$

$$\dim \bigoplus_{i=1}^k V_{\lambda_i} = \sum_{i=1}^k \dim V_{\lambda_i} = g_{\lambda_i}$$

and since $\forall 1 \leq i \leq k$, $g_{\lambda_i} \leq a_{\lambda_i}$, we have equality iff $\forall 1 \leq i \leq k$, $g_{\lambda_i} = a_{\lambda_i}$

6 Jordan Normal Form

Remark. In this chapter, $F = \mathbb{C}$

Definition. Let $A \in M_n(\mathbb{C})$, we say that A is in **Jordan Normal Form** (JNF) if it is a block diagonal matrix:

$$A = \begin{bmatrix} J_{n_1}(\lambda_1) & & & \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{n_k}(\lambda_k) \end{bmatrix}$$

where $k \geq 1$ nad $n_1, \ldots, n_k$ integers satisfying

$$\sum_{i=1}^{k} n_i = n$$

(need not be distinct). For $m \geq 1,\ \lambda \in \mathbb{C}$, define $J_m(\lambda)$

$$J_1(\lambda) = \begin{bmatrix} \lambda \end{bmatrix}$$

$$J_m(\lambda) = \begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

($J_m(\lambda)$ is a **Jordan Block**)

Remark. for $n = 3$

$$A = \begin{bmatrix} \lambda & & 0 \\ 0 & \lambda & 0 \\ & & \lambda \end{bmatrix}$$

is in jordan normal form as we have $J_1(\lambda)$ on diagonal

Theorem 6.1 (Can write in JNF in $\mathbb{C}$)**.** Every matrix $A \in M_n(\mathbb{C})$ is similar to a matrix in JNF, which is unique up to reordering of the Jordan block

Proof. Non examinable

Example. for $n = 2$, the possible JNF in this case

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}, \quad \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

Characterised by minimal polynomials

Theorem 6.2 (Generalised eigenspace decomposition)**.** For V a $\mathbb{C}$ vector space, $\dim_{\mathbb{C}} V = n < +\infty$, $\alpha \in L(V)$, and $m_\alpha(t) = (t-\lambda)^{c_1} \dots (t-\lambda_2)^{c_k}$ where (λ_i) are the distinct eigenvalues of α then:

$$V = \bigoplus_{j=1}^{k} V_j$$

where

$$V_j = \ker[(\alpha - \lambda_j \text{ Id})^{c_j}]$$

Proof. The key is that projectors onto V_j are "explicit".
Indeed

$$m_\alpha(t) = \prod_{j=1}^{k} (t - \lambda_j^{c_j})$$

We introduce

$$p_j(t) = \prod_{i \neq j} (t - \lambda_i)^{c_i}$$

Then the p_j polynomials have no common factor, so by Euclid's algorithm, we can find polynomials $q_1, \dots, q_k$

$$\sum_{i=1}^{k} q_i p_i = 1$$

define the projectors

$$\Pi_j = q_j p_j(\alpha)$$

(i) by construction

$$\begin{aligned} \sum_{j=1}^{k} \Pi_j(v) &= (\sum_{j=1}^{k} q_j p_j)(\alpha(v)) \\ &= \text{Id } (v) \\ &= v \end{aligned}$$

$$\implies \forall v \in V,\ v = \sum_{j=1}^{k} \Pi_j(v)$$

(ii) $\Pi_j(v) \in V_j$ (trivial check) we have shown

$$V = \sum_{j=1}^{k} \Pi_j(v) = +_{j=1}^{k} V_j$$

(iii) We need to show that the sum is direct. We have $\Pi_i \Pi_j = 0$ if $i \neq j$

$$\Pi_i = \Pi_i(\sum_{j=1}^{k} \Pi_j) = \Pi_i^2$$

$$\implies \Pi_i|_{V_j} = \begin{cases} \text{Id} & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

This immediately implies the direct sum property (trivial)

Remark. V_j is stable by α: $\alpha(V_j) \leq V_j$.
Let $(\alpha - \lambda_j \text{ Id})|_{V_j} = u_j$. Then u_j is a nilpotent endomorphism i.e.:

$$u_j^{c_j} = 0$$

thus the JNF decomposition is now a statement about nilpotent matrices

Notation. $V_j = \ker[(\alpha - \lambda_j \text{ Id})] \equiv$ generalized eigenspace associated to λ_j

Remark. When α is diagonalisable, $c_j = 1$ and hence theorem holds

Remark. We can compute on the JNF the quantities $a_\lambda, g_\lambda, c_\lambda$.

- Indeed, let $m \geq 2$, Considering $(J_m - \lambda \text{ Id})^2$, we get

  $$(J_m - \lambda \text{ Id})^k = \begin{bmatrix} 0 & I_{m-k} \\ 0 & 0 \end{bmatrix}$$

  for $k < m$ and 0 for $k = m$. Thus $(J_m - \lambda \text{ Id})$ is nilpotent of order exactly m.
- a_λ = sum of sizes of blocks with eigenvalue λ
- g_λ = number of blocks with eigenvlaue λ
- c_λ = size of the largest block with eigenvalue λ

Example.

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix}$$

To find a basis where A is JNF:

(i)

$$\chi_A(t) = (t-a)^2$$

so have one eigneigenvalue $\lambda = 1$ so our JNF is

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

(ii) Eigenvectors:

$$\ker(A - \text{ Id}) = \langle v_1 \rangle, \quad v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Look for v_2 s.t.

$$(A - \text{ Id})v_2 = v_1$$

and $v_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ works.

$$P^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix}$$

7 Bilinear Forms

Have $\varphi : V \times V \to F$ a bilinear form in this section, $n = \dim_F V < \infty$, $\mathcal{B}$ basis of V, $\mathcal{B} = (e_1, \dots, e_n)$.
$[\varphi]_{\mathcal{B}} = [\varphi]_{\mathcal{B},\mathcal{B}} = (\varphi(e_i, e_j))_{1 \le i,j \le n}$

Lemma 7.1 (Change of basis for bilinear forms)**.** For $\varphi : V \times V \to F$ a bilinear form with $\mathcal{B}, \mathcal{B}'$ bases for V and with $P = [\,\mathrm{Id}]_{\mathcal{B}',\mathcal{B}}$. We have

$$[\varphi]_{\mathcal{B}'} = P^T [\varphi]_{\mathcal{B}} P$$

Proof. Special case of general formula

Definition. We say $A, B \in M_n(F)$ are **congruent** if $\exists P \in M_n(F)$ invertible s.t.:

$$A = P^T B P^T$$

Remark. This defines an equivalence relation

Definition. A bilinear form φ on V is **symmetric** if:

$$\varphi(u, v) = \varphi(v, u), \ \forall u, v \in V$$

Remark. For $A \in M_n(F)$, we say that A is **symmetric** if A^T if

$$A^T = A$$

or equivalently

$$A = (a_{ij})_{1 \le i,j \le n}, \quad a_{ij} = a_{ji}$$

φ is symmetric $\iff$ $[\varphi]_{\mathcal{B}}$ is symmetric in any basis $\mathcal{B}$. To be able to represent φ by a diagonal matrix in some basis $\mathcal{B}$, it is necessary that φ is symmetric:

$$P^T A P = D \implies D^T = P^T A^T P$$

which implies $A^T = A$, so φ is symmetric

Definition. A map $Q : V \to F$ is a **quadratic form** if: there exists a bilinear form $\varphi : V \times V \to F$ such that

$$\forall u \in V, \quad Q(u) = \varphi(u, u)$$

Remark. With $\mathcal{B}$ and A defined as above, let $u = \sum_{i=1}^{n} \lambda_i e_i$, then

$$\begin{aligned} Q(u) &= \varphi(u, u) \\ &= \varphi(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} x_j e_j) \\ &= \sum_{i,j=1}^{n} x_i x_j \varphi(e_i, e_j) \\ &= \sum_{i,j=1}^{n} a_{ij} x_i x_j \\ &= x^T A x \end{aligned}$$

where $x = [u]_{\mathcal{B}}$ and $A = [\varphi]_{\mathcal{B}}$

Note.

$$\begin{aligned} x^T A x &= \sum_{i,j=1}^{n} a_{ij} x_i x_j \\ &= \sum i, j = 1^n \left(\frac{a_{ij} + a_{ji}}{2} \right) x_i x_j \\ &= x^T \left(\frac{A + A^T}{2} \right) x \end{aligned}$$

Prop 7.2 (Quadratic form $\leftrightarrow$ symmetric bilinear form)**.** If $Q : V \to F$ is a quadratic form, then there exists a unique symmetric bilinear form $\varphi : V \times V \to F$ such that

$$Q(u) = \varphi(u, u) \quad \forall u \in V$$

Proof. Let ψ be a bilinear form on V s.t. $\forall u \in V$, $Q(u) = \psi(u, u)$. Let

$$\varphi(u, v) = \frac{1}{2}(\psi(u, v) + \psi(v, u))$$

Then ψ is a symmetric bilinear form

$$\varphi(u, u) = \psi(u, u) = Q(u)$$

Thus $\exists \varphi$ bilinear symmetric such that

$$\varphi(u, u) = Q(u)$$

$$A \to \frac{1}{2}(A^T + A)$$

Uniqueness: let φ be a symmetric bilinear form such that

$$\forall u \in V, \quad \varphi(u, u) = Q(u)$$

Then

$$\begin{aligned} Q(u + v) &= \varphi(u + v, u + v) \\ &= \varphi(u, u) + 2\varphi(u, v) + \varphi(v, v) \\ &= Q(u) + 2\varphi(u, v) + Q(v) \end{aligned}$$

$$\implies \varphi(u, v) = \frac{1}{2}[Q(u + v) - Q(u) - Q(v)]$$

Theorem 7.3 (Diagonalisation of symmetric bilinear forms)**.** Let $\varphi : V \times V \to F$ be a symmetric bilinear form with $\dim_F V = n < +\infty$. Then there exists a basis $\mathcal{B}$ of V such that $[\varphi]_{\mathcal{B}}$ is diagonal

Proof. We induct on dimension. $n = 1$ trivial. Suppose Theorem holds for all dimensions $< n$. If $\varphi(u,u) = 0\ \forall u \in V$, then $\varphi \equiv 0$, done.
If $\varphi \not\equiv 0$, then $\exists u \in V \backslash \{0\}$ s.t. $\varphi(u,u) \neq 0$. Let us call $u = e_1$

$$U = (\langle e_1 \rangle)^{\perp} = \{v \in V : \varphi(e_r, v) = 0\} = \ker\{\varphi(e_1, \cdot) : V \to F\}$$

Rank nullity on $\varphi(e_1, \cdot)$ gives

$$\dim V = n = \dim U + 1$$

$$\implies \dim U = n - 1$$

We claim $U + \langle e_1 \rangle = U \oplus \langle e_1 \rangle$. Indeed, for $v = \langle e_1 \rangle \cap U$ so $v = \lambda e_1$ for $\lambda \in F$

$$\implies 0 = \varphi(e_1, v) = \varphi(e_1, \lambda e_1) = \lambda \varphi(e_1, e_1)$$

$$\implies \lambda = 0 \implies v = 0$$

$$\implies V = U \oplus \langle e_1 \rangle$$

Pick $(e_2, \dots, e_n)$ basis of U, $\mathcal{B} = (e_1, e_2, \dots, e_n)$ basis of V, then

$$[\varphi]_{\mathcal{B}} = (\varphi(e_i, e_j))_{1 \leq i,j \leq n} = \begin{bmatrix} \varphi(e_1, e_1) & 0 \\ 0 & A' \end{bmatrix}$$

as for $j \geq 2$, $\varphi(e_1, e_j) = \varphi(e_j, e_1) = 0$, with $A' = [\varphi|_U]_{\mathcal{B}'}$ defined as expected and define induction hypothersis

Example. $V = \mathbb{R}^3, (e_1, e_2, e_3)$ a basis

$$Q(x_1, x_2, x_3) = x_1^2 + x_2^2 + 2x^3 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3 = x^T Ax$$

where

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

(i) Diagonalise using the proof algortithm

(ii) Complete the square

$$\begin{aligned} Q(x_1, x_2, x_3) &= x_1^2 + x_2^2 + 2x^3 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3 \\ &= (x_1 + x_2 + x_3)^2 + x_3^3 - 4x_2x_3 \\ &= (\underbrace{x_1 + x_2 + x_3}_{x_1'})^2 + (\underbrace{x_3 - 2x_2}_{x_2'})^2 - (\underbrace{2x_2}_{x_3'})^2 \end{aligned}$$

$$\implies P^T AP = \begin{bmatrix} 1 & & \\ & 1 & \\ & & -1 \end{bmatrix}$$

To fine P note that

$$\begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & -2 & 0 \end{bmatrix}}_{P^{-1}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

7.1 Sylvester's Law and Sesquelinear Forms

Theorem 7.4 (Can diagonalise symmetric bilinear forms)**.** For $\dim_F V < \infty$, $\varphi : V \times V \to F$ a symmetric bilinear form, $\exists \mathcal{B}$ basis of V w.r.t. $[\varphi]_{\mathcal{B}}$ is diagonal

Corollary 7.5 (Can choose 'nice' basis for symmetric bilinear forms on $\mathbb{C}$)**.** For $F = \mathbb{C}$, $\dim_{\mathbb{C}} V = n < +\infty$, φ symmetric bilinear form on $V \times V$, $\exists \mathcal{B}$ basis of V s.t.:

$$[\varphi]_{\mathcal{B}} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \quad r = r(\varphi)$$

Proof. Pick $\mathcal{E} = (e_1, \ldots, e_n)$ such that

$$[\varphi]_{\mathcal{E}} = \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix}$$

Order a_i such that $a_i \neq 0$ for $1 \leq i \leq r$ and $a_i = 0$ for $i > r$. Then, for $i \leq r$, let $\sqrt{a_i}$ be a choice of complex root for a_i. Let

$$v_i = \frac{e_i}{\sqrt{a_i}} \text{ for } 1 \leq i \leq r$$

$$v_i = e_i \text{ for } i > r$$

Then $\mathcal{B} = (v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ basis of V and we can check

$$[\varphi]_{\mathcal{B}} = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

Corollary 7.6 (Congruence of symmetric matrices in $\mathbb{C}$ determined by rank)**.** Every symmetric matrix of $M_n(\mathbb{C})$ is congruent to a unique matrix of the form

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

Corollary 7.7 (Can choose 'nice' basis for symmetric bilinear forms on $\mathbb{R}$)**.** For $F = \mathbb{R}$, $\dim_{\mathbb{R}} V = n < \infty$ and φ symmetric bilinear form of $V \times V$, we have $\exists \mathcal{B} = (v_1, \dots, v_n)$ basis of V such that

$$[\varphi]_{\mathcal{B}} = \begin{bmatrix} I_p & & \\ & I_q & \\ & & 0 \end{bmatrix}$$

for some $p - q \geq 0$ and $p + q = r(\varphi)$

Proof. $\mathcal{E} = (a_1, \dots, e_n)$ s.t.

$$[\varphi]_{\mathcal{E}} = \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix}$$

Reorder indices such that $a_i > 0,\ 1 \leq i \leq p$, $a_i < 0, p+1 \leq i \leq q$ and $a_i = 0,\ i \geq p+q+1$. Define

$$v_i = \begin{cases} \frac{e_i}{\sqrt{a_i}} & \text{for } 1 \leq i \leq p \\ \frac{e_i}{\sqrt{-a_i}} & \text{for } p+1 \leq i \leq p+q \\ e_i & \text{for } i \geq p+q+1 \end{cases}$$

$$\implies \mathcal{B} = (c_1, \dots, v_n)$$

works

Definition. For $F = \mathbb{R}$, $s(\varphi) = p - q \equiv$ signature of φ (or the signature of the associated quadratic form Q)

Remark. Need to show that $s(\varphi)$ is intrinsic to φ: does not change if the basis $\mathcal{B}$ changes

Definition. For φ symmetric bilinear form on a real vector space V. We say that

(i) φ is **positive definite** $\iff$ $\varphi(u,u) > 0\ \forall u \in V \backslash \{0\}$
(ii) φ is **positive semi definite** $\iff$ $\varphi(u,u) \geq 0\ \forall u \in V \backslash \{0\}$
(iii) φ is **negative definite** $\iff$ $\varphi(u,u) < 0\ \forall u \in V \backslash \{0\}$
(iv) φ is **negative semi definite** $\iff$ $\varphi(u,u) \leq 0\ \forall u \in V \backslash \{0\}$

Example.

$$\begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix}$$

is positive definite for $p = n$, positive semi definite for $1 \leq p < n$

Theorem 7.8 (Sylvester's law of inertia). $F = \mathbb{R}$, $\dim_F V = n < \infty$. If a real symmetric bilinear form is represented by

$$\begin{bmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{bmatrix}$$

in $\mathcal{B}$ basis of V and

$$\begin{bmatrix} I_{p'} & & \\ & -I_{q'} & \\ & & 0 \end{bmatrix}$$

in $\mathcal{B}'$ basis of V

$$\implies p = p', \quad q = q'$$

Proof. In order to prove uniqueness of p, it is enough to show that p is the largest dimension of a subspace of V on which φ is definite positive. Say $\mathcal{B} = (v_1, \dots, v_n)$

$$[\varphi]_{\mathcal{B}} = \begin{bmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{bmatrix}$$

Let $X = \langle v_1, \dots, v_p \rangle$. Then φ is positive definite on X:

$$u \in X, \; u = \sum_{i=1}^{p} \lambda_i v_i$$

$$\begin{aligned} Q(u) &= \varphi(u, u) \\ &= \varphi(\sum_{i=1}^{p} \lambda_i v_i, \sum_{i=1}^{p} \lambda_i v_i) \\ &= \sum_{i=1}^{p} \lambda_i^2 \\ &> 0 \text{ as long as } v \neq 0 \end{aligned}$$

Suppose that φ is definite positive on another subspace X'. Let

$$X = \langle v_1, \dots, v_p \rangle$$

$$Y = \langle v_{p+1}, \dots, v_n \rangle$$

Then arguing verbating as above, we know φ is negative semidefinite on Y. This implies that $Y \cap X' = \{0\}$. Indeed if $y \in \varphi \cap X'$, then

$$Q(y) \leq 0 \leq Q(y) \implies y = 0$$

$$\implies Y + X' = Y \oplus X'$$

$$n = \dim_{\mathbb{R}} V \geq \dim(Y + X') = \dim Y + \dim X'$$

$$\implies n \geq n - p + \dim X'$$

$$\implies \dim X' \leq p$$

Similarly, we show that q is the largest subspace on which φ is definite negative and so we have a geometric characterisation of p, q

Definition.

$$K = v \in V : \forall u \in V, \varphi(u, v) = 0$$

is the **kernel** of the bilinear form

Remark.

$$\dim K + r(\varphi) = n$$

One can show using the above notation that there si a subspace T of dimension:

$$n - (p + q) + \min p, q$$

such that $\varphi|_T = 0$ (just consider 'cancellations' in matrix)

$$T = \langle v_1 + v_{p+1}, \ldots, v_q + v_{p+q}, v_{p+q+1}, \ldots, v_n \rangle$$

Moreover, one can show that the dimension of T is the largest possible dimension of a subspace T such that $\varphi|_T = 0$

7.1.1 Sesquilinear Forms

We have standard inner product on $\mathbb{C}^n$ given by

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \bar{y}_i$$

Warning.

$$(x, y) \mapsto \langle x, y \rangle = \sum_{i=1}^{n} x_i \bar{y}_i$$

is NOT a bilinear form on $\mathbb{C}$.

Definition. If V, W are vector spaces over $\mathbb{C}$. A **sesquilinear form** on $V \times W$ is a function

$$\varphi : V \times W \to \mathbb{C}$$

such that:

(i)

$$\varphi(\lambda_1 v_1 + \lambda_2 v_2, w) = \lambda_1 \varphi(v_1, w) + \lambda_2 \varphi(v_2, w)$$

$(\forall \lambda_1, \lambda_2 \in \mathbb{C},\ \forall v_1, v_2 \in V, \forall w \in W)$

(ii)

$$\varphi(v, \lambda_1 w_1 + \lambda_2 w_2) = \overline{\lambda}_1 \varphi(v, w_1) + \overline{\lambda}_2 \varphi(v, w_2)$$

(antilinear with respect to the second coordinate)

Lemma 7.9 (Evaluating sesquilinear form w.r.t. bases)**.** If $\mathcal{B} = (v_1, \ldots, v_m)$ is a basis of V and $\mathcal{C} = (w_1, \ldots, w_n)$ basis of W and $[\varphi]_{\mathcal{V},\mathcal{C}} = (\varphi(v_i, w_j))_{1 \le i \le m, 1 \le j \le n}$ then

$$\varphi(v, w) = [v]_{\mathcal{B}}^T [\varphi]_{\mathcal{B},\mathcal{C}} \overline{[w]_{\mathcal{C}}}$$

Lemma 7.10 (Writing matrix for sesquilinear form)**.** If $\mathcal{B}, \mathcal{B}'$ bases for V with $P = [\text{Id}]_{\mathcal{B},\mathcal{B}'}$ and $\mathcal{C}, \mathcal{C}'$ bases for W with $Q = [\ \text{Id}]_{\mathcal{C},\mathcal{C}'}$

$$[\varphi]_{\mathcal{B}',\mathcal{C}'} = P^T [\varphi]_{\mathcal{B},\mathcal{C}} \overline{Q}$$

7.2 Hermitian Forms and Skew Symmetric Forms

Definition. A sesquilinear form $\varphi : V \times V \to \mathbb{C}$ is **Hermitian** if

$$\forall (u, v) \in V \times V \quad \varphi(u, v) = \overline{\varphi(v, u)}$$

Remark. φ Hermitian $\implies \varphi(u, u) \in \mathbb{R}$
Moreover,

$$\varphi(\lambda u, \lambda u) = |\lambda|^2 \varphi(u, u)$$

This allows us to talk about positive or negative definite Hermitian forms

Lemma 7.11 (Hermitian iff matrix same as conjugate transpose for any basis)**.** A sesquilinear form $\varphi : V \times V \to \mathbb{C}$ is Hermitian iff: for any basis $\mathcal{B}$ of V

$$[\varphi]_{\mathcal{B}} = \overline{[\varphi]_{\mathcal{B}}^T}$$

Proof. Let $A = [\varphi]_{\mathcal{B}} = (a_{ij})_{1 \leq i,j \leq n}$ and it is trivial. Conversely, write u and v in terms of $\mathcal{B}$ and apply linearity to show equality

Claim (Polarization Identity)**.** A Hermitian form φ on a complex vector space V is entirely determined by

$$Q : V \to \mathbb{R} \quad v \mapsto \varphi(v, v)$$

via the formula

$$\varphi(u, v) = \frac{1}{4}[Q(u + v) - Q(u - v) + iQ(u + iv) - iQ(u - iv)]$$

Proof. Trivial check, similar to symmetric bilinear forms

Theorem 7.12 (Hermitian formulation of Sylvester's Law)**.** Let $n = \dim_{\mathbb{C}} V < +\infty$. Let $\varphi : V \times V \to \mathbb{C}$ be a Hermitian form on V. Then $\exists \mathcal{B} = (v_1, \dots, v_n)$ of V s.t.

$$[\varphi]_{\mathcal{B}} = \begin{bmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{bmatrix}$$

where p, q depend only on φ

Proof. Mainly identical to the case of real symmetric bilinear forms
Existence $\varphi = 0$ done. Otherwise, using the polarization identity, there exists $e_1 \neq 0$ s.t. $\varphi(e_1, e_1) \neq 0$. Then rescale e_1 to get $\varphi(v_1, v_1) = \pm 1$ Consider the orthogonal:

$$W = \{w \in V : \varphi(v_1, w) = 0\}$$

then can check

$$V = \langle v_1 \rangle \oplus W$$

and we argue by induction on dimension to diagonalise.
Uniqueness of p: p is the maximal dimension of a subspace on which φ is definite positive.

Definition (Skew symmetric bilinear forms)**.** A bilinear form $\varphi : V \times V \to \mathbb{R}$ is **skew symmetric** if

$$\forall (u, v) \in V \times V \quad \varphi(u, v) = -\varphi(v, u)$$

Remark. (i)

$$\varphi(u, u) = -\varphi(u, u) \implies \varphi(u, u) = 0$$

(ii) $\forall \mathcal{B}$ basis of V,

$$[\varphi]_{\mathcal{B}} = -[\varphi]_{\mathcal{B}}^T$$

(iii) $\forall A \in M_n(\mathbb{R})$

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$$

symmetric + skew symmetric

Theorem 7.13 (Sylvester form of skew symmetric matrices)**.** Let φ be a skew symmetric bilinear form over V (vector space over $\mathbb{R}$), then $\exists$ a basis $\mathcal{B}$ of V

$$\mathcal{B} = (v_1, w_1, v_2, w_2, \ldots, v_m, w_m, v_{2m+1}, v_{2m+2}, \ldots, v_n)$$

s.t.

$$[\varphi]_{\mathcal{B}} = \begin{bmatrix} J & & & & \\ & J & & & \\ & & \ddots & & \\ & & & J & \\ & & & & 0 \end{bmatrix}$$

$$J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Proof. Induction on dimension of V. $\varphi \equiv 0$ done. $\varphi \neq 0 \implies \exists (v_1, w_1) : \varphi(v_1, w_1) \neq 0$. After scaling say w_1, we can assume

$$\varphi(v_1, w_1) = 1 \implies \varphi(w_1, v_1) = -1$$

Observe (v_1, w_1) are linearly independent.
Let $U = \langle v_1, w_1 \rangle$

$$W = \{v \in V : \varphi(v_1, v) = \varphi(w_1, v) = 0\}$$

then we can show $V = U \oplus W$ by induction

Corollary 7.14. Skew symmetric matrices have an even rank

8 Inner Product Spaces

Have for definite positive bilinear forms a scalar product and a norm. We have an infinite dimensional counterpart – Hilbert Spaces.

Definition (Inner product, scalar product)**.** Let V be a vector space over $\mathbb{R}$ (resp $\mathbb{C}$). An **inner product** on V is a positive definite symmetric (resp Hermitian) symmetric form φ on V

Notation. $\varphi(u,v) = \langle u, v\rangle$. V is called a real (resp complex) inner product space

Examples. (i)

$$\mathbb{R}^n, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$$

(ii) $\mathbb{C}^n$

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y}_i$$

(iii) $V = \mathcal{C}^0([0,1], \mathbb{C})$

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g}(t)\,\mathrm{d}t$$

(iv) We can fix a wright $w : [0,1] \to \mathbb{R}_+^*$ and define on $V = \mathcal{C}^0([0,1], \mathbb{C})$:

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g}(t)w(t)\,\mathrm{d}t$$

Note. One can checkthat all the examples are inner product

Remark. The study of L^2 spaces is the heart of the definitian of a new integral: Lebesgue Integral

Definition (norm/ length)**.**

$$\|v\| = (\langle v, v\rangle)^{1/2}$$

Remark. $\langle v, v\rangle \in \mathbb{R}_+$ and

$$\|v\| = 0 \iff v = 0$$

Lemma 8.1 (Cauchy-Schwarz).

$$|\langle u, v\rangle| \le \|u\|\|v\|$$

Proof. With $F = \mathbb{R}$ or $\mathbb{C}$. Let $t \in F$, then

$$\begin{aligned} 0 \le \|tu - v\|^2 &= \langle tu - v, tu - v\rangle \\ &= t\bar{t}\langle v, u\rangle - t\langle v, u\rangle - \bar{t}\langle v, u\rangle + \|v\|^2 \\ &= |t|^2\|u\|^2 - 2\,\mathrm{Re}(t\langle v, u\rangle) + \|v\|^2 \end{aligned}$$

Choose explicitly

$$t = \frac{\overline{\langle v, u\rangle}}{\|u\|^2}$$

which gives result and we can also show that if there is equality in Cauchy-Schwarz, then the two vectors are colinear

Corollary 8.2 (Triangle inequality).

$$\|u + v\| \le \|u\| + \|v\|$$

Proof. trivial

Remark. $\|\cdot\|$ is a norm

Definition (Orthogonal/ orthonormal families). A set $(e_1, \ldots, e_k)$ of vectors of V is

(i) Orthogonal if

$$\langle e_i, e_j\rangle = 0 \text{ if } i \neq j$$

(ii) Orthonormal if

$$\langle e_i, e_j\rangle = S_{ij} = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases}$$

Lemma 8.3 (Orthogonal non-zero set is linearly independent). If $(e_1, \ldots, e_k)$ are orthogonal (all non zero) vectors, then they are linearly independent

$$v = \sum_{j=1}^{k} \lambda_j e_j$$

$$\lambda_j = \frac{\langle v, e_j\rangle}{\|e_j\|^2}$$

Proof. Just take inner products

Lemma 8.4 (Parseval's identity)**.** If V is a finite dimensional inner product space and $(e_1, \ldots, e_n)$ is an orthonormal basis, then

$$\langle u, v\rangle = \sum_{i=1}^{n} \langle u, e_i\rangle \overline{\langle v, e_i\rangle}$$

$$\|u\| = \sum_{i=1}^{n} |\langle u, e_i\rangle|^2$$

Proof. Just write u and v using the previous lemma and take scalar products

Theorem 8.5 (Gram-Schmidt orthogonalization process)**.** Given V an inner product space. Let $(v_i)_{i\in I}$ be such that I countable (or finite) and $v_i \in V$, $(v_i)_{i\in I}$ are linearly independent. Then there exists a family $(e_i)_{i\in I}$ of orthonormal vectors such that

$$\forall k \geq 1, \quad \text{span}\langle v_1, \ldots, v_k\rangle = \text{span}\langle e_1, \ldots, e_k\rangle$$

Proof. We give an explicity algorithm to compute the family $(e_i)_{i\in\mathbb{N}}$. Induction on k:

- $k = 1$, $e_1 = v_1/\|v_1\|$ since $v_1 \neq 0$
- Say we have found $(e_1, \ldots, e_k)$ orthonormal with

  $$\text{span}\{v_1, \ldots, v_k\} = \text{span}\{e_1, \ldots, e_k\}$$

- Let us compute e_{k+1}. We define:

  $$e'_{k+1} = v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, e_i\rangle e_i$$

  (notice we can interpret this as projection)
- $e'_{k+1} \neq 0$: otherwise,

  $$v_{k+1} \in \text{span}\{e_1, \ldots, e_k\} = \text{span}\{v_1, \ldots, v_k\}$$

  by induction on k
- For $j \in \{1, \ldots, k\}$:

  $$\begin{aligned}\langle e'_{k+1}, e_j\rangle &= \langle v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, e_i\rangle e_i, e_j\rangle \\ &= \langle v_{k+1}, e_j\rangle - \sum_{i=1}^{k} \langle v_{k+1}, e_j\rangle \langle e_i, e_j\rangle \\ &= \langle v_{k+1}, e_j\rangle - \langle v_{k+1}, e_j\rangle = 0\end{aligned}$$

  $$\implies \forall 1 \leq j \leq k \quad e'_{j+1} \perp e_j$$

- $\text{span}\{v_1, \ldots, v_k\} = \text{span}\{e_1, \ldots, e_k, e'_{k+1}\}$ (follows from formula for e'_{k+1})
- $e'_{k+1} \neq 0$ so $e_{k+1} = e'_{k+1}/\|e_{k+1}\|$ does the job

Corollary 8.6 (Can extend any orthogonal set to orthonormal basis)**.** If V is a finite dimensional inner product space, then any orthogonal set of vectors can be extended to an orthonormal basis of V

Proof. Pick $(e_1, \dots, e_k)$ orthonormal. Then they are linearly independent and we can extend to $(e_1, \dots, e_k, v_{k+1}, \dots, v_n)$ basis of V. We apply the Gram-Schmidt algorithm to this set to get $(e_1, \dots, e_n)$ orthonormal with

$$\text{span}\{e_1, \dots, e_n\} = \text{span}\{e_1, \dots, e_k, v_{k+1}, \dots, v_n\} = V$$

$\implies (e_1, \dots, e_n)$ is an orthonormal basis

Remark. For $A \in M_n(\mathbb{R})$ or $(M_n(\mathbb{C}))$, then the column vectors of A are orthogonal iff $A^T = A = \text{Id}$ in $\mathbb{R}$ case or $A^T\overline{A} = \text{Id}$ in $\mathbb{C}$ case

Definition. $A \in M_n(\mathbb{R})$ $(M_n(\mathbb{C}))$ is:

- $\mathbb{R}$ orthogonal if:
$$A^T A = \text{Id}(\iff A^{-1} = A^T)$$
- $\mathbb{C}$ unitary if:
$$A^T\overline{A} = \text{Id}(\iff A^{-1} = \overline{A}^T)$$

Prop 8.7 (Decomposing into upper triangular and orthogonal)**.** If $A \in M_n(\mathbb{R})$ is non-singular, then A can be written as

$$A = RT$$

where T is upper triangular and R is orthogonal (unitary)

Proof. Exercise (apply Gram Schmidt to the column vectors of A)

8.1 Orthogonal Complement and Projection

Definition. Let V be an inner product space with $V_1, V_2 \leq V$. We say that V is the **orthogonal direct sum** of V_1 and V_2 if

(i) $V = V_1 \oplus V_2$

(ii) $\forall (v_1, v_2) \in V_1 \times V_2$

$$\langle v_1, v_2 \rangle = 0$$

Notation. $V = V_1 \oplus^{\perp} V_2$

Remark. $\forall (v_1, v_2) \in V_1 \times V_2,\ \langle v_1, v_2 \rangle = 0$ so $V_n \cap V_2 = \{0\}$

Definition. For V an inner product space with $W \leq V$, we define

$$W^{\perp} = \{v \in V : \forall w \in W, \langle v, w \rangle = 0\}$$

Lemma 8.8 (Subspace and orthogonal complement form direct sum)**.** For V an inner product space with $\dim V < +\infty$, and $W \leq V$, we have

$$V = W \oplus^{\perp} W^{\perp}$$

Proof. For $\omega \in W, \omega \in W^{\perp}$

$$\|\omega^2\| = \langle \omega, \omega \rangle = 0$$

$$\implies \omega = 0$$

Need to show $V = W + W^{\perp}$. Let $(e_1, \dots, e_k)$ be an orthonormal basis of W. Extend it to $(e_1, \dots, e_k, e_{k+1}, \dots, e_n)$ orthonormal basis of W. Observe that $(e_{k+1}, \dots, e_n) \in W^{\perp}$

$$\implies V = W + W^{\perp}$$

Remark.

$$V = W \oplus W^{\perp}$$

Definition (Projection map)**.** Suppose $V = U \oplus W$ (U is a complement of W in V). We define $\Pi : V \to W \;\; v = u + w \mapsto w$

- Π is well defined
- Π is linear
- $\Pi^2 = \Pi$

We say that Π is the **projection operator** onto W

Remark. Id $-\Pi \equiv$ projection onto U. We can make the projection map very explicit when $U = W^{\perp}$ (U is the orthogonal complement of W in V)

Lemma 8.9 (Evaluating projection map in terms of inner products)**.** Let V be an inner product space. Let $W \leq V$, with W finite dimensional. Let $(e_1, \dots, e_k)$ be an orthonormal basis of W. Then

(i)
$$\Pi(v) = \sum_{i=1}^{k} \langle v, e_i \rangle e_i, \quad \forall v \in V$$

(ii) $\forall v \in V, \forall w \in W$
$$\|v - \Pi(v)\| \leq \|v - w\|$$
with equality iff $w = \Pi(v)$

Proof. (i) We define: for $v \in V$
$$\Pi(v) = \sum_{i=1}^{k} \langle v, e_i \rangle e_i$$
$W =$ span$\{e_1, \dots, e_k\}$ so $\Pi(v) \in W$. We write
$$v = v - \Pi(v) + \Pi(v)$$
And we claim $v - \Pi(v) \in W^{\perp}$. Indeed, we need to show $\forall w \in W$, $\langle v - \Pi(v), w \rangle = 0$. We compute
$$\begin{aligned}\langle v - \Pi(v), e_j \rangle &= \langle v, e_j \rangle - \langle \sum_{i=1}^{k} \langle v, e_i \rangle e_i, e_j \rangle \\ &= \langle v, e_j \rangle - \langle v, e_j \rangle \\ &= 0\end{aligned}$$
$$v - \Pi(v) \in W^{\perp}$$
thus
$$V = W \oplus^{\perp} W^{\perp}$$

(ii) Let $v \in V$, $w \in W$, let us compute:
$$\begin{aligned}\|v - w\|^2 &= \|\underbrace{v - \Pi(v)}_{\in W^{\perp}} + \underbrace{\Pi(v) - w}_{\in W}\|^2 \\ &= \|v - \Pi(v)\|^2 + \|\Pi(v) - w\|^2 \\ &\geq \|v - \Pi(v)\|^2\end{aligned}$$
with equality iff $w = \Pi(v)$. We have shows: $\forall w \in W$,
$$\|v - w\|^2 \geq \|v - \Pi(v)\|^2$$

8.2 Adjoint Maps

This is a fundamental object with deep infinite dimensional generalisations

Definition. Let V, W be finite dimensional inner product spaces, $\alpha \in L(V, W)$. Then there is a unique linear map $\alpha^* : W \to V$ such that $\forall (v, w) \in V \times W$,

$$\langle \alpha(v), w \rangle = \langle v, \alpha^*(w) \rangle$$

Claim (Writing adjoint map)**.** If $\mathcal{B}$ orthonormal basis of V and $\mathcal{C}$ orthonormal basis of W then $[\alpha^*]_{\mathcal{C},\mathcal{B}} = (\overline{[\alpha]_{\mathcal{B},\mathcal{C}}})^T$

Proof. Brute force computation

$$\mathcal{B} = \{v_1, \ldots, v_n\}$$
$$\mathcal{C} = \{w_1, \ldots, w_n\}$$

$$A = [\alpha]_{\mathcal{B},\mathcal{C}} = (a_{ij})$$

Existence: let $[\alpha^*]_{\mathcal{C},\mathcal{B}} = (c_{ij})$ we can compute

$$\begin{aligned} \langle \alpha(\sum \lambda_i v_i), \sum \mu_j w_j \rangle &= \langle \sum_{i,k} \lambda_i a_{ki} w_k \sum \mu_j w_j \rangle \\ &= \sum_{i,j} \lambda_i a_{ji} \overline{\mu_j} \qquad (*) \end{aligned}$$

$$\begin{aligned} \langle \sum_i \lambda_i v_i, \alpha^*(\sum_j \mu_j w_j) \rangle &= \langle \sum_i \lambda_i v_i, , \sum_{j.k} \mu_j c_{kj} v_k \rangle \\ &= \sum_{i,j} \lambda_i \overline{c_{ij} \mu_j} \qquad (**) \end{aligned}$$

and so

$$\overline{c_{ij}} = a_{ji}$$

and uniquely defined as $(*) = (**)$ for any vector iff $\overline{c_{ij}} = a_{ji}$

Notation.

$$\overline{A}^T = A^\dagger$$

Remark. We are using the same notation α^* for the adjoint (as just defined) and the dual map. If V, W are real product inner spaces, and $\alpha \in L(V, W)$

$$\psi_{R,V} : V \to V^* \quad v \mapsto \langle \cdot, v \rangle$$

$$\psi_{R,W} : W \to W^* \quad w \mapsto \langle \cdot, w \rangle$$

then the adjoint of α is given by:

$$W \to W^* \to V^* \to V$$

by $\psi_{R,W}$, dual of α and $\psi_{R,V}^{-1}$

8.3 Self Adjoint Maps and Isometries

Definition. For V an inner product space, $\alpha \in L(V)$ and $\alpha^* \in L(V)$ the adjoint map, we have the following:

Condition	Equivalent	Name
$\langle \alpha v, w\rangle = \langle v, \alpha w\rangle$	$\alpha = \alpha^*$	Self Adjoint: $\mathbb{R}$ **Symmetric** $\mathbb{C}$ **Hermitian**
$\langle \alpha v, \alpha w\rangle = \langle v, w\rangle$	$\alpha^* = \alpha^{-1}$	Isometry: $\mathbb{R}$ **Orthogonal**, $\mathbb{C}$ **Unitary**

Proof. We check equivalence for isometries Have $\langle \alpha(v), \alpha(v)\rangle = \langle v, w\rangle$ so $\|\alpha(v)\|^2 = \|v\|^2$ so the kernel is trivial and thus α is a bijection so α^{-1} well defined

$$\langle v, \alpha^*(w)\rangle = \langle \alpha v, w\rangle = \langle \alpha v, \alpha(\alpha^{-1}w)\rangle = \langle v, \alpha^{-1}w\rangle$$

So we have shown $\forall v \forall w$

$$\langle v, (\alpha^* - \alpha^{-1})w\rangle = 0$$

Choose $v = (\alpha^* - \alpha^{-1})(w)$ to get $\forall w$

$$\|(\alpha^* - \alpha^{-1})(w)\|^2 = 0$$

$$\implies \forall w, (\alpha^* - \alpha^{-1})(w) = 0$$

$$\implies \alpha^* = \alpha^{-1}$$

And for the reverse

$$\langle \alpha v, \alpha w\rangle = \langle v, \alpha^* \alpha w\rangle = \langle v, w\rangle$$

from the definition of α^* and that $\alpha^* = \alpha^{-1}$

Remark. Using the polarization identity, one can show α isometry $\iff \forall v \in V, \|\alpha(v)\| = \|v\| \iff \forall (v, w) \in V \times W, \langle \alpha(v), \alpha(w)\rangle = \langle v, w\rangle$

Lemma 8.10 (Classifying self adjoint maps and isometries)**.** For V a finite dimensional real (complex) inner product space, we have $\alpha \in L(V)$ is:
- (i) self adjoint iff for any orthonormal basis $\mathcal{B}$ of V, $[\alpha]_{\mathcal{B}}$ is symmetric (Hermitian)
- (ii) an isometry iff for any orthonormal basis of V, $[\alpha]_{\mathcal{B}}$ is orthogonal (unitary)

Proof. Let $\mathcal{B}$ be an orthnormal basis:

$$[\alpha^*]_{\mathcal{B}} = \overline{[\alpha]_{\mathcal{B}}^T}$$

- (i) Self adjoint: $\overline{[\alpha]_{\mathcal{B}}^T} = [\alpha]_{\mathcal{B}}$
- (ii) Isometry: $\overline{[\alpha]_{\mathcal{B}}^T} = [\alpha]_{\mathcal{B}}^{-1}$

Definition. For V a finite dimensional inner product space with:
- $F = \mathbb{R}$, $O(V) = \{\alpha \in L(V) : \alpha \text{ is an isometry}\} \equiv$ **orthogonal** group of V
- $F = \mathbb{C}$, $U(V) = \{\alpha \in L(V) : \alpha \text{ is an isometry}\} \equiv$ **unitary** group of V

9 Spectral Theory for Self Adjoint Maps

Spectral theory is the study of the spectrum of operators

Lemma 9.1 (Self adjoint operators have real eigenvalues and an orthogonal set of eigenvectors)**.** Let V be a finite dimensional inner product space. Let $\alpha \in L(V)$ be self adjoint: $\alpha = \alpha^*$. Then

(i) α has real eigenvalues

(ii) Eigenvectors of α with respect to different eigenvalues are orthogonal

Proof. (i) Take $\lambda \in \mathbb{C}, v \in V\backslash\{0\}$ s.t. $\alpha(v) = \lambda v$. Then

$$\begin{aligned}\langle v, v\rangle &= \lambda\|v\|^2 \\ \langle \alpha v, v\rangle &= \langle v, \alpha v\rangle = \langle v, \lambda v\rangle \\ &= \overline{\lambda}\|v\|^2\end{aligned}$$

$$\implies (\lambda - \overline{\lambda})\|v\|^2 = 0$$

$$\implies \lambda = \overline{\lambda},\ \lambda \in \mathbb{R}$$

(ii) Let us consider two eigenvectors for different eigenvalues

$$\alpha v = \lambda v \quad \alpha w = \mu w$$

with $\lambda, \mu \in \mathbb{R}$ non-zero. Then

$$\begin{aligned}\lambda\langle v, w\rangle &= \langle \lambda v, w\rangle \\ &= \langle \alpha(v), w)\rangle \\ &= \langle v, \alpha(w))\rangle \\ &= \langle v, \mu w\rangle \\ &= \overline{\mu}\langle v, w\rangle = \mu\langle v, w\rangle\end{aligned}$$

$$\implies (\lambda - \mu)\langle v, w\rangle = 0$$

$$\implies \langle v, w\rangle = 0$$

Theorem 9.2 (Adjoint operators are diagonalisable)**.** Let V be a finite dimensional inner product space and let $\alpha \in L(V)$ be self adjoint. Then V has an orthonormal basis of eigenvectors of α (so α is diagonalisable)

Proof. $F = \mathbb{R}$ or $\mathbb{C}$. We argue by induction on the dimension of V

- $n = 1 : V$
- Say $A = [\alpha]_{\mathcal{B}}$ wrt fundamentabl basis $\mathcal{B}$. By the fundamental theorem of algebra, we know that $\chi_A(\lambda)$ has a root. This root is a n eigenvalue of α so the root is real. Let us call this real eigenvalue $\lambda \in \mathbb{R}$. Pick $v_1 \in V \backslash \{0\}$ s.t.

$$\alpha(v_1) = \lambda v_1 \quad \|v_1\| = 1$$

Let $U = \langle v_1 \rangle^\perp \leq V$. Then U is stable by α. Indeed, let $u \in U$, then

$$\begin{aligned}\langle \alpha(u), v \rangle &= \langle u, \alpha^*(v_1) \rangle \\ &= \langle u, \alpha(v_1) \rangle \\ &= \langle u, \lambda v_1 \rangle \\ &= \lambda \langle u, v_1 \rangle = 0\end{aligned}$$

$$\implies \alpha(u) \perp v_1 \implies \alpha(u) \in U$$

- Hence we may consider $\alpha|_U \in L(U)$ which is self adjoint, and

$$n = \dim V = \dim U + 1$$

$$\implies \dim U = n - 1$$

$\implies \exists (v_1, \ldots, v_{n-1})$ orthonormal basis of U of eigenvectors of $\alpha|_U$ so $(v_1, \ldots, v_n)$ is an orthonormal basis of V of eigenvectors of α

$$V = \langle v_1 \rangle \oplus^\perp U$$

Corollary 9.3 (Decompose V into orthogonal direct sum of eigenspaces)**.** V finite dimensional inner product space. If $\alpha \in L(V)$ is self adjoint, then V is the orthogonal direct sum of all the eigenspaces of α

9.1 Spectral Theory for Unitary Maps

Lemma 9.4 (Unitary maps have unit modulus eigenvectors which are orthogonal)**.** Let V be a complex inner product space (Hermitian sesquilinear structure). Let $\alpha \in L(V)$ be unitary ($\alpha^* = \alpha^{-1}$) then

(i) All eigenvalues of α lie on the unit circle

(ii) Eigenvectors corresponding to different eigenvalues are orthogonal

Proof. (i) Let $\lambda \in \mathbb{C}$, $v \in L \backslash \{0\}$ s.t. $\alpha(v) = \lambda v$

- $\lambda \neq 0$: α isometry $\implies$ α invertible so $\ker \alpha = \{0\}$
- We compute

$$\lambda \langle v, v \rangle = \langle v, \alpha^{-1} v \rangle$$

and $\alpha(v) = \lambda v \implies v = \lambda \alpha^{-1} v$ and so

$$\lambda \langle v, v \rangle = \frac{1}{\bar{\lambda}} \langle v, v \rangle$$

$$\implies (|\lambda|^2 - 1)\|v\|^2 = 0$$

$$\implies |\lambda| = 1$$

(ii) Let v, w be two eigenvectors for two distnct eigenvalues

$$\alpha(v) = \lambda v \quad \alpha(w) = \mu w$$

Then

$$\lambda \langle v, w \rangle = \mu \langle v, w \rangle$$

$$\implies \langle v, w \rangle = 0$$

Theorem 9.5 (Spectral Theorem for unitary maps)**.** Let V be a finite dimensional complex inner product space. Let $\alpha \in L(V)$ be unitary. Then V has an orthonormal basis consisting of eigenvectors of α

Note. Equivalently, α is diagonalisable in an orthonormal basis of V

Proof. $A = [\alpha]_{\mathcal{B}}$, $\mathcal{B}$ orthonormal basis. Fix $v_1 \in V \backslash \{0\}$ s.t.

$$\alpha(v_1) = \lambda v_1 \quad \|v_1\| = 1$$

Let $U = \langle v_1 \rangle^{\perp}$, we claim: $\alpha(U) \leq U$. Indeed, for $u \in U$

$$\langle \alpha(u), v_1 \rangle = \frac{1}{\overline{\lambda}} \langle u, v_1 \rangle = 0$$

$$\implies \alpha(u) \in U$$

Hence $\alpha|_U \in L(U)$ which is unitary and $\dim U = n - 1$, $n = \dim_{\mathbb{C}} V$. By induction, get $(v_2, \ldots, v_n)$ orthonormal basis of U made up of eigenvectors of $\alpha|_U$.

$$V = \langle v_1 \rangle \oplus^{\perp} U$$

So $(v_1, \ldots, v_n)$ orthonormal basis of V made of eigenvectors of α.

Warning. We used the complex structure. In general a real orthonormal matrix A s.t. $AA^T = \mathrm{Id}$ CANNOT be diagonalised over $\mathbb{R}$ e.g. rotation in $\mathbb{R}^2$

9.2 Application to Bilinear Forms

Corollary 9.6 (Can diagonalise symmetric matrices with $P^{-1} = P^T$)**.** Let $A \in M_n(\mathbb{R})$ (resp $M_n(\mathbb{C})$) be a symmetric (resp Hermitian) matrix. Then there is an orthonormal (resp unitary) matrix P such that $P^T AP$ (resp $P^{\dagger} AP$) is diagonal with real valued entries

Proof. $F = \mathbb{R}$ ($\mathbb{C}$). Let $\langle\ \rangle$ be the standard inner product over $\mathbb{R}^n$ (resp $\mathbb{C}^n$). Then $A \in L(F^n)$ is self adjoint hence we can fnd an orthonormal basis F^n such that A is diagonal in this basis, say $\mathcal{B} = (v_1, \ldots, v_n)$. Let $P = (v_1 | \ldots | v_n)$ with $(v_1 | \ldots | v_n)$ orthonormal basis. Have this iff P unitary $\iff P^T P =$ Id. I know $P^{-1} AP = D$ diagonal with real diagonal.
Then, as $P^{-1} = P^T$, $P^T AP = D$

Corollary 9.7 (Can diagonalise symmetric forms)**.** Let V be a finite dimensional real (complex) inner product space. Let

$$\varphi : V \times V \to F$$

by a symmetric (resp Hermitian) form. Then there exists an orthnormal basis of V such that φ in this basis is represented by a diagonal matrix

Proof. $\mathcal{B} = (v_1, \ldots, v_n)$ orthonormal basis of V. Let:

$$A = [\varphi]_{\mathcal{B}}$$

$$\implies A^T = A$$

and hence there is an orthogonal (unitary) matrix P such that: $P^T AP$ $(P^\dagger AP)$ is diagonal, say D.
Let (v_i) be the ith row of P^T $(P^\dagger)$, then $(v_1, \ldots, v_n)$ is an orthonormal basis $\mathcal{B}'$ of V and

$$[\varphi]_{\mathcal{B}'} = D$$

Remark. The diagonal entries of $P^T AP$ are the eigenvalues of A. Moreover

$$s(\varphi) = \text{ number positive eigenvalues of } A \text{ - number negative eigenvalues of } A$$

Corollary 9.8 (Simultaneous diagonalization)**.** Let V be a finite dimensional real (complex) vector space. Let

$$\varphi, \psi : V \times V \to F$$

be symmetric (Hermitian) bilinear forms. Assume φ is definite positive. Then $\exists (v_1, \ldots, v_n)$ basis of V with respect to which both forms are respected by a diagonal matrix

Proof. Key point: φ is definite positive so V equipped with φ is a finite dimensional inner product space

$$\langle u, v \rangle = \varphi(u, v)$$

Hence there exists an orthonormal (for the φ induced scalar product) basis of V in whihc ψ is respresented by a diagonal matrix. Observe that φ in this basis is represented by the identity matrix. $(\varphi(v_i, v_j) = \langle v_i, v_j \rangle = S_{ij}\ )$

Corollary 9.9 (Simultaneous diagonalization for matrices)**.** $A, B \in M_n(\mathbb{R})$ (resp $M_n(\mathbb{C})$) symmetric (Hermitian). Assume

$$\forall x \neq 0,\ \overline{x}^T Ax > 0 \qquad (*)$$

. Then there exists $Q \in M_n(\mathbb{R})$ invertible such that: both matrices $Q^T AQ,\ Q^T BQ$ are diagonal

Proof. The condition (*) just expresses the fact that A induces φ definite positive..
Similarly $\tilde{Q}(x) = \overline{x}^T Bx$, $\overline{Q}(x) = \psi(x, x)$ symmetric so we just apply the previous simultaneous diagonalization Theorem to φ, ψ. We use change of basis formula for quadratic forms.