Linear Algebra

June 3, 2023

Contents

Lectures

Start of lecture 1

1 Vector spaces and subspaces

Let F be an arbitrary field (eg $\mathbb{R}$ or $\mathbb{C}$).

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

$$F \times V \to V$$

$$(\lambda, v) \mapsto \lambda v$$

such that:

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

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:

$$v \in F^n, v = \begin{vmatrix} x_1 \\ \vdots \\ x_n \end{vmatrix}, \qquad x_i \in F, 1 \le i \le n$$

$$v + w = \begin{vmatrix} v_1 \\ \vdots \\ v_n \end{vmatrix} + \begin{vmatrix} w_1 \\ \vdots \\ w_n \end{vmatrix} = \begin{vmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{vmatrix}$$

$$\lambda v = \begin{vmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{vmatrix} \qquad \lambda \in F$$

check: F^n is an F vector space.

(ii) Any set X,

$$\mathbb{R}^X = \{f : X \to \mathbb{R}\}$$

(set of real valued functions on X) Then $\mathbb{R}^X$ is an $\mathbb{R}$ vector space

$$(f_1 + f_2)(x) = f_1(x) + f_2(x)$$

$$(\lambda f)(x) = \lambda f(x), \qquad \lambda \in \mathbb{R}$$

(iii) $\mathcal{M}_{n,m}(F) \equiv n \times m$ F valued matrices. Sum is sum of entries, $\lambda M = (\lambda m_{ij})$.

Remark. The axiom of scalar multiplication imply that:

$$\forall\ v \in V, \quad 0_F v = 0_V$$

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

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

The last two properties can be combined into a single property:

- $\forall (\lambda_1, \lambda_2, u_1, u_2) \in F \times F \times U \times U, \quad \lambda_1 u_1 + \lambda_2 u_2 \in U\ (*)$

Property $(*)$ means that U is stable by

- scalar multiplication
- vector addition

Example. V is an F vector space, and $U \leq V$. Then U is an F vector space.

Examples

(1) $V = \mathbb{R}^{\mathbb{R}}$ space of functions $f : \mathbb{R} \to \mathbb{R}$.

- Let $\mathcal{C}(\mathbb{R})$ be the space of continuous functions $f : \mathbb{R} \to \mathbb{R}$. Then $\mathcal{C}(\mathbb{R}) \leq V$.
- Let $\mathbb{P}(\mathbb{R})$ be the space of polynomials of one variable. Then $\mathbb{P}(\mathbb{R}) \leq V$.

(2) Let

$$V = \left\{ \begin{vmatrix} x_1 \\ x_2 \\ x_3 \end{vmatrix} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = t \right\}$$

check: that this is a subspace of $\mathbb{R}^3$ for $t = 0$ only.

Warning. The union of two subspaces is generally *not* a subspace. (It is typically not stable by addition).

Example. $V = \mathbb{R}^2$, with $U_1 = \{(x, 0) : x \in \mathbb{R}\}$, $U_2 = \{(0, y) : y \in \mathbb{R}\}$. Both subspaces, but the union isn't since

$$\underbrace{(1,0)}_{\in U_1} + \underbrace{(0,1)}_{\in U_2} = (1,1) \notin U \cup V$$

Proposition. 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: let $(\lambda_1, \lambda_2, v_1, v_2) \in F \times F \times (U \cap W) \times (U \cap W)$. Then

$$\underbrace{\lambda_1 v_1}_{\in U} + \underbrace{\lambda_2 v_2}_{\in U} \in U$$

and similarly for W, hence

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

□

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

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

Example. Use $V = \mathbb{R}^2$ and U_1, U_2 from the previous example. Then $U_1 + U_2 = V$.

Proposition. Let V be an F vector space, with $U, W \leq V$. Then

$$U + W \leq V$$

Proof.

- $0 = \underbrace{0}_{\in U} + \underbrace{0}_{\in W} \in U + W$
- Consider $\lambda_1 f + \lambda_2 g$ for $\lambda_1, \lambda_2 \in F$ and $f, g \in U + W$. Then let:

$$f = \underbrace{f_1}_{\in U} + \underbrace{f_2}_{\in W}$$

$$g = \underbrace{g_1}_{\in U} + \underbrace{g_2}_{\in W}$$

so

$$\begin{aligned}\lambda_1 f + \lambda_2 g &= \lambda_1(f_1 + f_2) + \lambda_2(g_1 + g_2) \\ &= \underbrace{(\lambda_1 \underbrace{f_1}_{\in U} + \lambda_2 \underbrace{g_1}_{\in U})}_{\in U} + \underbrace{(\lambda_1 \underbrace{f_2}_{\in W} + \lambda_2 \underbrace{g_2}_{\in W})}_{\in W} \\ &\in U + W\end{aligned}$$

□

Exercise: Show that $U + W$ is the *smallest* subspace of V which contains both 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 product multiplication:

$$F \times V/U \to V/U$$

$$(\lambda, v + U) \mapsto \lambda v + U \tag{$*$}$$

Proposition. V/U is an F vector space.

Remark. The multiplication is well defined:

$$v_1 + U = v_2 + U$$
$$\implies v_1 - v_2 \in U$$
$$\implies \lambda(v_1 - v_2) \in U$$
$$\implies \lambda v_1 + U = \lambda v_2 + U \in V/U$$

Exercise: Prove that V/U is an F vector space.

Start of lecture 2

1.2 Spans, linear independence and the Steinitz exchange lemma

Definition (Spand of a family of vectors)**.** Let V be an F vector space. Let $S \subset V$ be a subset (S = collection of vectors). We define:

$$\underbrace{\langle S \rangle}_{\text{"span of } S\text{"}} = \{\text{finite linear combination of elements of } S\}$$
$$= \left\{ \sum_{s \in J} \lambda_s v_s, v_s \in S, \lambda_s \in F, J \text{ is finite} \right\}$$

Convention: $\langle \emptyset \rangle = \{0\}$.

Remark. $\langle S \rangle$ is the smallest vector subspace which contains S.

Examples

(1) $V = \mathbb{R}^3$

$$S = \left\{ \begin{vmatrix} 1 \\ 0 \\ 0 \end{vmatrix}, \begin{vmatrix} 0 \\ 1 \\ 2 \end{vmatrix}, \begin{vmatrix} 2 \\ -2 \\ -4 \end{vmatrix} \right\}$$
$$\implies \langle S \rangle = \left\{ \begin{vmatrix} a \\ b \\ 2b \end{vmatrix}, (a,b) \in \mathbb{R}^2 \right\}$$

(2)

$$V = \mathbb{R}^n = \left\{ \begin{vmatrix} x_1 \\ \vdots \\ x_n \end{vmatrix}, x_i \in \mathbb{R}, 1 \le i \le n \right\}$$

Define:

$$e_i = \left| \begin{matrix} 0 \\ \vdots \\ 1 \quad \text{(in position } i\text{)} \\ \vdots \\ 0 \end{matrix} \right.$$

$$\implies V = \langle e_1, \dots, e_n \rangle$$

(3) X is a set, $V = \mathbb{R}^X = \{f : X \to \mathbb{R}\}$.

$$S_x : X \to \mathbb{R}$$

$$y \mapsto \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$

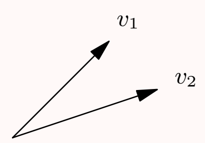
$$\begin{aligned} \langle (S_x)_{x \in X} \rangle &= \operatorname{span}((S_x)_{x \in X}) \\ &= \{f \in \mathbb{R}^X : f \text{ has finite support}\} \end{aligned}$$

(Support of f is $\{x \in X : f(x) \neq 0\}$)

Definition. Let V be an F vector space. Let S be a subset of V. We say that S *spans* V if:

$$\langle S \rangle = V.$$

Example. $V = \mathbb{R}^2$



$\{v_1, v_2\}$ spans 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.

Example. Let $V_1 = P[x]$ be the set of polynomials over $\mathbb{R}$, and let $V = P_n[x]$ be the set of polynomials over $\mathbb{R}$ with degree $\leq n$. Then $\{1, x, \ldots, x^n\}$ spans $P_n[x]$, so $P_n[x] = \langle 1, x, \ldots, x^n \rangle$. So $P_n[x]$ is finite dimensional.
On the other hand, $P[x]$ is *not* finite dimensional: it is infinite dimensional, because there is no family of V with finitely many elements which spans V.

Question: If V is finite dimensional, is there a *minimal* number of vectors in the family so that they span V.

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

$$\sum_{i=1}^{n} \lambda_i v_i = 0,\ \lambda_i \in F \implies \lambda_i = 0\ \forall i$$

Remark. (1) We also say that the family $(v_1, \ldots, v_n)$ is *free*.

(2) Equivalently, $(v_1, \ldots, v_n)$ are *not* linearly independent if one of these vectors is a linear combination of the remaining $(n-1)$ ones. Indeed, $\exists(\lambda_1, \ldots, \lambda_n)$ *not all zero* (that is, there exists j such that $\lambda_j \neq 0$), such that

$$\sum_{i=1}^{n} \lambda_i v_i = 0 \implies v_j = -\frac{1}{\lambda_j} \sum_{\substack{i=1 \\ i \neq j}}^{n} \lambda_i v_i$$

Example. $V = \mathbb{R}^3$. If (v_1, v_2) free, and v_3 is coplanar with both, then $(v_1, v_2, v_3$ is not free.

Remark. $(v_i)_{1 \leq i \leq n}$ free family (linearly independent) then $\forall 1 \leq i \leq n$, $v_i \neq 0$.

Definition (Basis)**.** A sub set S of V is a *basis* of V if and only if:

- $\langle S \rangle = V$ (generating family)
- S linearly independent / free

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

Examples

(1)

$$V = \mathbb{R}^n = \left\{ \left| \begin{matrix} x_1 \\ \vdots \\ x_n \end{matrix} \right., x_i \in \mathbb{R}, 1 \le i \le n \right\}$$

$$e_i = \left| \begin{matrix} & 0 \\ & \vdots \\ 1 & \text{(in position } i\text{)} \\ & \vdots \\ & 0 \end{matrix} \right.$$

Then $(e_i)_{1 \le i \le n}$ is a basis of V.

(2) $V = \mathbb{C}$. If $F = \mathbb{C}$ then $\{1\}$ is a basis of V. If $F = \mathbb{R}$ then $\{1, i\}$ is a basis of V.

(3) $V = P[x] = \{\text{polynomials over } \mathbb{R}\}$, $S = \{x^n : n \ge 0\}$. Then S is a basis for V.

Lemma. Let V be an 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:

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

Notation. $(\lambda_1, \dots, \lambda_n)$ are the *coordinates* of v in the basis $(v_1, \dots, v_n)$.

Proof. By assumption, $\langle v_1, \dots, v_n \rangle = V$ so

$$\forall v \in V, \exists (\lambda_1, \dots, \lambda_n) \in F^n \qquad v = \sum_{i=1}^{n} \lambda_i v_i$$

Uniqueness: let

$$\begin{aligned} v = \sum_{i=1}^{n} \lambda_i v_i &= \sum_{i=1}^{n} \lambda_i' v_i \\ \implies \sum_{i=1}^{n} (\lambda_i - \lambda_i') v_i &= 0 \\ \implies \forall 1 \le i \le n,\ \lambda_i &= \lambda_i' \end{aligned}$$

□

Lemma. 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 then done. Let's assume they are not independent. Then by possible reordering the vectors,

$$v_n \in \langle v_1, \dots, v_{n-1} \rangle$$

(v_n is a linear combination of $v_1, \dots, v_{n-1}$) so

$$V = \langle v_1, \dots, v_n \rangle = \langle v_1, \dots, v_{n-1} \rangle$$

Now we can iterate until the resulting set is a basis of V. (We only have to iterate finitely many times since n is finite). □

Theorem (Steinitz exchange lemma)**.** Let V be a *finite dimensional* vector space over F. Take:

(i) $(v_1, \dots, v_m)$ free

(ii) $(w_1, \dots, w_n)$ generating ($\langle w_1, \dots, w_n \rangle = V$).

Then $m \le n$, and up to reordering,

$$(v_1, \dots, v_m, w_{m+1}, \dots, w_n)$$

spans V.

Start of lecture 3

Proof. Induction. Suppose that we have replaced l (≥ 0) of the w_i. Reordering if necessary:

$$\langle v_1, \dots, v_l, w_{l+1}, \dots, w_n \rangle = V$$

If $m = l$ then we are done. So assume $l < m$. Then take $v_{l+1} \in V$, and we must have

$$v_{l+1} = \sum_{i \le l} a_i v_i + \sum_{i > l} \beta_i w_i$$

Since the family $(v_1, \dots, v_{l+1})$ is free, we must have that one of the β_i is non zero. So up to reordering, $\beta_{l+1} \ne 0$

$$\implies w_{l+1} = \frac{1}{\beta_{l+1}} \left[v_{l+1} - \sum_{i \le l} \alpha_i v_i - \sum_{i > l+1} \beta_i w_i \right]$$

So

$$w_{l+1} \in \langle v_1, \dots, v_{l+1}, w_{l+2}, \dots, w_n \rangle$$

hence we have that

$$\begin{aligned} V &= \langle v_1, \dots, v_l, w_{l+1}, \dots, w_n \rangle \\ &= \langle v_1, \dots, v_{l+1}, w_{l+2}, \dots, w_n \rangle \end{aligned}$$

so we can induct up on l. The base case $l = 0$ is trivial, so we deduce the last part of the lemma (which also trivially proves that $m \leq n$). □

1.3 Basis, dimension, direct sums

Corollary (of Steinitz)**.** Let 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, denoted $\dim_F V$ $(\in \mathbb{N})$.

Proof. $(v_1, \dots, v_n)$, $(w_1, \dots, w_m)$ basis of V over F. Then since $(v_i)_{1 \leq i \leq n}$ free, $(w_i)_{1 \leq i \leq m}$ generating, by Steinitz exchange lemma, $n \leq m$. Similarly $m \leq n$, so $n = m$. □

Corollary. Let V be a vector space over F with dimension $n \in \mathbb{N}$.

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

(ii) any spanning (generating) set of vectors has at *least* n elements with equality if and only if it is a basis.

Proof. Exercise. □

Proposition. Let U, W be subspaces of V. If U and W are finite dimensional, then so is $(U + W)$ and:

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

Proof. Pick $(v_1, \dots, v_n)$ basis of $U \cap W$. Extend to bases:

$$\langle v_1, \dots, v_l, u_1, \dots, u_m \rangle = U$$

$$\langle v_1, \dots, v_l, w_1, \dots, w_n \rangle = W$$

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

Proving it is a generating family is an exercise. Proving it is a free family:

$$\underbrace{\sum_{i=1}^{l} \alpha_i v_i + \sum_{i=1}^{m} \beta_i u_i}_{\in U} + \underbrace{\sum_{i=1}^{n} \gamma_i w_i}_{\in W} = 0 \tag{$*$}$$

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

$$\implies \sum_{i=1}^{l} S_i v_i = \sum_{i=1}^{n} \gamma_i w_i$$

$$\overset{(*)}{\implies} \sum_{i=1}^{l} (\alpha_i - S_i) v_i + \sum_{i=1}^{m} \beta_i u_i = 0$$

$$\overset{U \text{ basis}}{\implies} \beta_i = 0, \quad \alpha_i = S_i$$

$$\overset{(*)}{\implies} \sum_{i=1}^{l} \alpha_i v_i + \sum_{i=1}^{n} \gamma_i w_i = 0$$

$$\overset{W \text{ basis}}{\implies} \alpha_i = \gamma_i = 0$$

so the set is free, so it's a basis. □

Proposition. Let V be a finite dimensional vector space over F. Let $U \leq V$. Then U and V/U are both finite dimensional and:

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

Proof. Let $(u_1, \dots, u_l)$ be a basis of U. Complete it to a basis $(u_1, \dots, u_l, w_{l+1}, \dots, w_n)$ of V.

Claim. $(w_{l+1} + U, \dots, w_n + U)$ is a basis of V/U.

Exercise. □

Remark. V vector space over F with $U \leq V$. We say that U is *proper* if $U \neq V$. U proper implies $\dim U < \dim V$, since $V/U \neq \{\emptyset\}$.

Definition (Direct sum)**.** Let V be a vector space, and $U, W \leq V$. We say

$$V = U \oplus W$$

We say "V is the *direct* sum of U and W" if and only if any element $v \in V$ can be *uniquely* decomposed:

$$v = u + w, \quad u \in U, \ w \in W$$

Equivalently,

$$V = U \oplus W \iff \forall v \in V, \exists!(u, w) \in U \times W \ \ v = u + w$$

Warning 1. If $V = U \oplus W$, we say that W is *a* complement of U in V. There is no uniqueness of such a complement.

Example. $V = \mathbb{R}^2 = \langle(1, 0)\rangle \oplus \langle(0, 1)\rangle = \langle(1, 0)\rangle \oplus \langle(1, 1)\rangle$.

Notation. We will in the sequel systematically use the following notation. Let two collections of vectors:

$$\mathcal{B}_1 = \{v_1, \dots, v_l\}$$

$$\mathcal{B}_2 = \{w_1, \dots, w_m\}$$

then

$$\mathcal{B}_1 \cup \mathcal{B}_2 = \{u_1, \dots, u_l, w_l, \dots, w_m\}$$

not a set, because we care about the order. (it is more like a list) With this notation:

$$\{u_1\} \cup \{u_1\} = \{u_1, u_1\}$$

so the collection $\{u_1\} \cup \{u_1\}$ is never a free family.

Lemma. $U, W \leq V$. Then the following are equivalent (TFAE):

(i) $V = U \oplus W$

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

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

Proof. (ii) $\implies$ (i) $V = U + W$ implies that $\forall v \in V$, there exists $(u, w) \in U \times W$ such that

$v = u + w$. So it is generating. To show it is free, let $u_1 + w_1 = u_2 + w_2 = v$. Then

$$\underbrace{u_1 - u_2}_{\in U} = \underbrace{w_2 - w_1}_{\in W}$$

$$\implies u_1 - u_2, w_1 - w_2 \in U \cap W = \{0\}$$

$$\implies u_1 = u_2,\ w_1 = w_2$$

(i) $\implies$ (iii) $\mathcal{B}_1$ basis of U, $\mathcal{B}_2$ basis of W. Let $\mathcal{B} = \mathcal{B}_1 + \mathcal{B}_2$. It is clearly a generating family of $U + W = V$ It is a free family because

$$\sum \lambda_i v_i = 0$$

must be decomposed as $0_U + 0_W$ since $V = U \oplus W$. So

$$\sum_{u_1 \in \mathcal{B}_1} \lambda_i u_i = 0$$

$$\sum_{w_1 \in \mathcal{B}_2} \lambda_i w_i = 0$$

so $\lambda_i = 0$ for all i.

(iii) $\implies$ (ii) We need to show

$$V = U + W, \quad U \cap W = \{0\}$$

This is obvious.

□

Start of lecture 4

Definition. Let V be a vector space over F. Let $V_1, \dots, V_l \leq V$ (subspaces).

(i) Notation: $\sum_{i=1}^{l} V_i = \{x_1 + \dots + v_l, v_j \in V_j, 1 \leq j \leq l\}$

(ii) The sum is *direct*, denoted by:

$$\sum_{i=1}^{l} V_i = \bigoplus_{i=1}^{l} V_i$$

if and only if

$$v_1 + \dots + v_l = v_1' + \dots + v_l' \implies v_1 = v_1', \dots, v_l = v_l'$$

Equivalently:

$$V = \bigoplus_{i=1}^{l} V_i \iff \forall v \in V\ \exists! v_i\ v = \sum_i v_i$$

Exercise: the following are equivalent:

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

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

(iii) For *any* basis of V_i,

$$\mathcal{B} = \bigcup_{i=1}^{l} \mathcal{B}_i$$

is a basis of $\sum_{i=1}^{l} V_i$.

1.4 Linear maps, isomorphism and the rank-nullity Theorem

Definition (Linear map)**.** Let V, W be vector spaces of F. A map $\alpha : V \to W$ is *linear* if and only if:

$$\forall (\lambda_1, \lambda_2) \in F^2, \forall (v_1, v_2) \in V \times V$$

$$\alpha(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \alpha(v_1) + \lambda_2 \alpha(v_2)$$

Examples

(i) Matrices are linear maps.

(ii) $\alpha : \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$ defined by

$$f \mapsto \alpha(f)(x) = \int_0^x f(t)\mathrm{d}t$$

is a linear map

(iii) Fix $x \in [a,b]$. $\mathcal{C}([a,b]) \to \mathbb{R}$ defined by $f \mapsto f(\lambda)$ is a linear map.

Remark. Let U, V, W be F vector spaces.

(i) $\mathrm{id}_V : V \to V$ defined by $x \mapsto x$ is a linear map.

(ii) If $\beta : U \to V$ and $\alpha : V \to W$ are linear, then $\alpha \circ \beta : U \to W$ is linear. (linearity is *stable* by composition)

Lemma. Let V, W be F vector spaces, and $\mathcal{B}$ a basis of V. Let $\alpha_0 : \mathcal{B} \to W$ be any map, then there is a unique linear map $\alpha : V \to W$ extending α_0 (a map such that $\forall v \in \mathcal{B}$, $\alpha(v) = \alpha_0(v)$).

Proof. For all $v \in V$, $v = \sum_{i=1}^{n} \lambda_i v_i$. Denote $\mathcal{B} = (v_1, \dots, v_n)$. By linearity: $\alpha : V \to W$ linear, so

$$\begin{aligned} \alpha(v) &= \alpha\left(\sum_{i=1}^{n} \lambda_i v_i\right) \\ &= \sum_{i=1}^{n} \lambda_i \alpha(v_i) \\ &= \sum_{i=1}^{n} \lambda_i \alpha_0(v_i) \end{aligned}$$

□

Remark. This is true in the infinite dimensional case as well (and the proof is the same).

- Often, to define a linear map, we define its value on a basis and "extend by linearity".
- If $\alpha_1, \alpha_2 : V \to W$ are linear and agree on a basis of V, they are equal.

Definition (Isomorphism)**.** Let V, W be vector spaces over F. A map

$$\alpha : V \to W$$

is called an *isomorphism* if and only if:

(i) α is linear;

(ii) and α is bijective.

If such an α exists, we write $V \simeq W$ (V isomorphic to W).

Remark. If $\alpha : V \to W$ is an isomorphism then $\alpha^{-1} : W \to V$ is linear. Take $w_1 = \alpha(v_1)$, $w_2 = \alpha(v_2)$. Then

$$\begin{aligned}\alpha^{-1}(w_1 + w_2) &= \alpha^{-1}(\alpha(v_1) + \alpha(v_2)) \\ &= \alpha^{-1}(\alpha(v_1 + v_2)) \\ &= v_1 + v_2 \\ &= \alpha^{-1}(w_1) + \alpha^{-1}(w_2)\end{aligned}$$

Similarly, $\forall \lambda \in F$, $\forall v \in V$,

$$\alpha^{-1}(\lambda V) = \lambda \alpha^{-1}(v)$$

Lemma. $\simeq$ is an equivalence relation on the class of all vector spaces of F.

(i) $\mathrm{id}_V : V \to V$ is an isomorphism.

(ii) $\alpha : V \to W$ isomorphism then $\alpha^{-1} : W \to V$ is an isomorphism.

(iii) Let $\beta : U \to V$ and $\alpha : V \to W$ be isomorphisms. Then $\alpha \circ \beta$ is an isomorphism.

Theorem. If V is a vector space over F of dimension n, then:

$$V \simeq F^n$$

Proof. Let $\mathcal{B} = (v_1, \dots, v_n)$ be a basis of V. Then $\alpha : V \to f^n$ defined by

$$v = \sum_{i=1}^{n} \lambda_i v_i \mapsto \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

is an isomorphism (exercise). □

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

Theorem. Let V, W be vector spaces over F with finite dimension. Then:

$$V \simeq W \iff \dim_F V = \dim_F W$$

Proof. $\Leftarrow$ $\dim_F V = \dim_F W = n$ implies that $V \simeq F^n$, $W \simeq F^n$ so $V \simeq W$.

⇒ Let $\alpha : V \to W$ be an isomorphism. Let $\mathcal{B}$ be a basis of V. Then we claim that $\alpha(\mathcal{B})$ is a basis of W:

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

so V and W have the same size basis so $\dim_F V = \dim_F W$.

□

Definition (Kernel and Image of a linear map)**.** Let V, W be vector spaces over F. Let $\alpha : V \to W$ be a linear map. We define:

(i) $N(\alpha) = \ker \alpha = \{v \in V : \alpha(v) = 0\}$

(ii) $\operatorname{im}(\alpha) = \{w \in W : \exists v \in V, w = \alpha(v)\}$.

Lemma. $\ker \alpha$ is a vector subspace of V, and $\operatorname{im} \alpha$ is a vector subspace of W.

Proof.

- $\lambda_1, \lambda_2 \in F$, $v_1, v_2 \in \ker \alpha$ implies

$$\alpha(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \alpha(v_1) + \lambda_2 \alpha(v_2) = 0$$

hence $\lambda_1 v_1 + \lambda_2 v_2 \in \ker \alpha$.

- $\lambda_1, \lambda_2 \in F$, $w_1, w_2 \in \operatorname{im} \alpha$. Let $w_1 = \alpha(v_1)$, $w_2 = \alpha(v_2)$. Then

$$\lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 \alpha(v_1) + \lambda_2 \alpha(v_2) = \alpha(\lambda_1 v_1 + \lambda_2 v_2)$$

hence $\lambda_1 w_1 + \lambda_2 w_2 \in \operatorname{im} \alpha$

□

Example. $\alpha : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$, $f \mapsto \alpha(f) = f'' - f$. Then

- α is linear
- $\ker \alpha = \{f \in \mathcal{C}^\infty(\mathbb{R}) : f'' - f = 0\} = \operatorname{span}_{\mathbb{R}}\langle e^t, e^{-t}\rangle$
- $\operatorname{im} \alpha$? Exercise.

Remark. $\alpha : V \to W$ linear map. Then α injective is equivalent to $\ker \alpha = \{0\}$. $\alpha(v_1) = \alpha(v_2) \iff \alpha(v_1 - v_2) = 0$.

Theorem. Let V, W be vector spaces over F. Let $\alpha : V \to W$ be a linear map. Then

$$\overline{\alpha} : V/\ker\alpha \to \operatorname{im}\alpha$$
$$v + \ker\alpha \mapsto \alpha(v)$$

is an isomorphism.

Proof. This follows from linearity.

- $\overline{\alpha}$ is well defined:

$$\begin{aligned} v + \ker\alpha = v' + \ker\alpha \\ \implies v - v' \in \ker\alpha \\ \implies \alpha(v - v') = 0 \\ \implies \alpha(v) = \alpha(v') \end{aligned}$$

  so $\overline{\alpha}$ is well-defined.

- $\overline{\alpha}$ linear follows from the linearity of α.

- $\overline{\alpha}$ is a bijection:
  - injectivity $\overline{\alpha}(v + \ker\alpha) = 0$ implies that $\alpha(v) = 0$ hence $v \in \ker\alpha$. So $v + \ker\alpha = 0 + \ker\alpha$.
  - surjectivity: follows form the definition of the image: $w \in \operatorname{im}\alpha$, $\exists v \in V$ such that $w = \alpha(v) = \overline{\alpha}(v)$.

□

Start of lecture 5

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

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

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

Proof. We have proved that $U/\ker\alpha \simeq \operatorname{im}\alpha$. So $\dim(U/\ker\alpha) = \dim\operatorname{im}\alpha$. But $\dim(U/\ker\alpha) = \dim U - \dim\ker\alpha$. So $\dim U = \dim\ker\alpha + \dim\operatorname{im}\alpha = r(\alpha) + n(\alpha)$. □

Lemma. Let V, W be vector spaces over F of *equal* finite dimension. Let $\alpha : V \to W$ be a linear map. Then the following are equivalent:

- α is injective
- α is surjective
- α is an isomorphism

Proof. Follows immediately from the rank-nullity theorem. (Exercise) □

Example. Let $V = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$. Then consider $\alpha : \mathbb{R}^3 \to \mathbb{R}$ defined by $(x, y, z) \mapsto x + y + z$. Then $\ker \alpha = V$ and $\operatorname{im} \alpha = \mathbb{R}$, hence by rank nullity $3 = n(\alpha) + 1$ hence $\dim V = 2$.

1.5 Linear maps from V to W and matrices

The space of linear maps from V to W. Let V, W be vector spaces over F.

$$L(V, W) = \{\alpha : V \to W \text{ linear}\}$$

Proposition. $L(V, W)$ is a vector space over F with:

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

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

Moreover if V and W are finite dimensional over F, then so is $L(V, W)$ and:

$$\dim_F L(V, W) = (\dim_F V)(\dim_F W)$$

Proof. Proof that it is a vector space is an exercise.
We will prove the statement about dimensions soon. □

Matrices and linear maps

Definition (Matrix)**.** A $m \times n$ matrix over F is an array with m rows and n columns with entries in F.

Notation. $\mathcal{M}_{m,n}(F)$ is the set of $m \times n$ matrices over F.

Proposition. $\mathcal{M}_{m,n}(F)$ is an F vector space under operations:

- $(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$
- $\lambda(a_{ij}) = (\lambda a_{ij})$

Proof. Exercise. □

Proposition. $\dim_F \mathcal{M}_{m,n}(F) = m \times n$.

Proof. We exhibit a basis using *elementary* matrices. Pick $1 \leq i \leq m$, $1 \leq j \leq n$. Then we define E_{ij} to be the matrix which is 0 everywhere, except it is 1 in the entry that is in the i-th row and j-th column. Then (E_{ij}) is a basis of $\mathcal{M}_{m,n}(F)$. Clearly spans $\mathcal{M}_{m,n}(F)$. Family is free is an exercise. □

Representation of linear maps

- V, W vector spaces over F, $\alpha : V \to W$ linear map.
- Basis $\mathcal{B}(v_1, \ldots, v_n)$ basis of V, $\mathcal{C} = (w_1, \ldots, w_m)$ basis of W.
- Let $v \in V$, then we can write
$$v = \sum_{j=1}^{n} \lambda_j v_j$$
so we can consider the coordinates of v in the basis $\mathcal{B}$ $(\lambda_1, \ldots, \lambda_n \in F^n)$. We may write this as $[v]_{\mathcal{B}}$.
- Similarly for $w \in W$, we note $[w]_{\mathcal{C}}$ in a similar way.

Definition (Matrix of α in $\mathcal{B}, \mathcal{C}$ basis)**.**

$$[\alpha]_{\mathcal{B},\mathcal{C}} \equiv \text{matrix of } \alpha \text{ with respect to } \mathcal{B}, \mathcal{C}$$

We define it as:

$$[\alpha]_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} \vdots & \vdots & & \vdots \\ [\alpha(v_1)]_{\mathcal{C}} & [\alpha(v_2)]_{\mathcal{C}} & \cdots & [\alpha(v_n)]_{\mathcal{C}} \\ \vdots & \vdots & & \vdots \end{pmatrix}$$

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

Lemma. For any $v \in V$,

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

where

$$(Av)_i = \sum_{j=1}^{n} a_{ij}\lambda_j$$

Proof. Let $v \in V$, with

$$v = \sum_{j=1}^{n} \lambda_j v_j$$

Then

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

□

Lemma. Let $\beta : U \to V$, $\alpha : V \to W$ linear, and hence $\alpha \circ \beta : U \to W$ linear. Let $\mathcal{A}$ be a basis of U, $\mathcal{B}$ be a basis of V, and $\mathcal{C}$ a basis of W. Then

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

Proof. $A = [\alpha]_{\mathcal{B},\mathcal{C}}$, $B = [\beta]_{\mathcal{A},\mathcal{B}}$. Pick $u_l \in \mathcal{A}$. Then

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

□

Proposition. If V and W are vector spaces over F and $\dim_F V = n$ and $\dim_F W = m$. Then $L(V, W) \simeq \mathcal{M}_{m,n}(F)$, and in particular, $\dim L(V, W) = m \times n$.

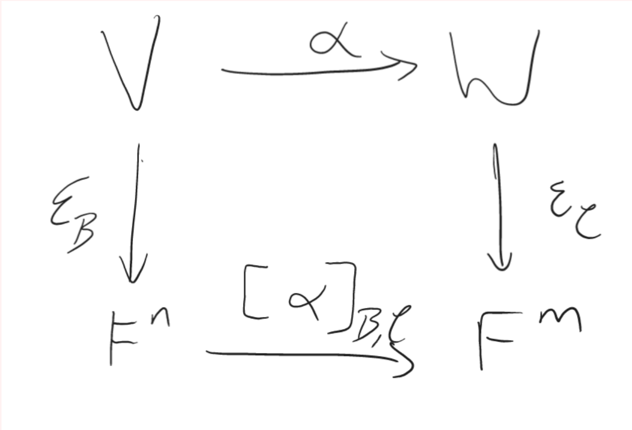
Proof. Fix $\mathcal{B}, \mathcal{C}$ basis of V, W.

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

- θ is linear:
$$[\lambda_1\alpha_1 + \lambda_2\alpha_2]_{\mathcal{B},\mathcal{C}} = \lambda_1[\alpha_1]_{\mathcal{B},\mathcal{C}} + \lambda_2[\alpha_2]_{\mathcal{B},\mathcal{C}}$$
- θ is surjective: let
$$A = (a_{ij})$$
  Consider the map:
$$\alpha : v_j \mapsto \sum_{i=1}^{m} a_{ij} w_i$$
  and extend by linearity. Then $[\alpha]_{\mathcal{B},\mathcal{C}} = A$.
- θ is injective because
$$[\alpha]_{\mathcal{B},\mathcal{C}} = 0 \implies \alpha \equiv 0$$

Hence, using θ, $L(V, W) \simeq \mathcal{M}_{m,n}(F)$. □

Remark. Let $\mathcal{B}, \mathcal{C}$ be bases of V, W. Let $\varepsilon_{\mathcal{B}} : V \to F^n$ be defined such that $v \mapsto [\alpha]_{\mathcal{B}}$, and similarly define $\varepsilon_{\mathcal{C}} : W \to F^m$ such that $w \mapsto [w]_{\mathcal{C}}$. Then the following diagram commutes:



Start of lecture 6

1.6 Change of basis and equivalent matrices

Let $\beta : U \to V$, $\alpha : V \to W$ and $\mathcal{A}, \mathcal{B}, \mathcal{C}$ bases of U, V, W.

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

Change basis

Let $\alpha : V \to W$ and let $\mathcal{B}, \mathcal{B}'$ and $\mathcal{C}, \mathcal{C}'$ be bases for V and W.

Definition. The "change of basis matrix" from $\mathcal{B}'$ to $\mathcal{B}$ is

$$P = (p_{ij})$$

given by

$$P = ([v'_1]_{\mathcal{B}} \cdots [v'_n]_{\mathcal{B}}) = [\mathrm{id}]_{\mathcal{B}',\mathcal{B}}$$

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

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

$$\implies [\mathrm{id}(v)]_{\mathcal{B}} = [\mathrm{id}]_{\mathcal{B}',\mathcal{B}} [v]_{\mathcal{B}'}$$
$$\implies [v]_{\mathcal{B}} = P[v]_{\mathcal{B}'}$$

□

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

Indeed

$$[\alpha \circ \beta]_{\mathcal{A},\mathcal{C}} = [\alpha]_{\mathcal{B},\mathcal{C}}[\beta]_{\mathcal{A},\mathcal{B}}$$
$$\implies [\text{id}]_{\mathcal{B},\mathcal{B}'}[\text{id}]_{\mathcal{B},\mathcal{B}'} = [\text{id}]_{\mathcal{B}',\mathcal{B}'} \equiv I_n$$
$$\implies [\text{id}]_{\mathcal{B}',\mathcal{B}}[\text{id}]_{\mathcal{B},\mathcal{B}'} = [\text{id}]_{\mathcal{B},\mathcal{B}} \equiv I_n$$

We changed $\mathcal{B}$ to $\mathcal{B}'$ in V. We can also change basis to $\mathcal{C}$ to $\mathcal{C}'$ in W.

Proposition. $A = [\alpha]_{\mathcal{B},\mathcal{C}}$, $A' = [\alpha]_{\mathcal{B}',\mathcal{C}'}$, $P = [\text{id}]_{\mathcal{B}',\mathcal{B}}$, $Q = [\text{id}]_{\mathcal{C}',\mathcal{C}}$. Then

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

Proof.

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

- $$[\alpha(v)]_{\mathcal{C}} = Q[\alpha(v)]_{\mathcal{C}} = Q[\alpha]_{\mathcal{B}',\mathcal{C}'} = QA'[v]_{\mathcal{B}'}$$
- $[\alpha(v)]_{\mathcal{C}} = [\alpha]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}} = AP[v]_{\mathcal{B}'}$.

So for all $v \in V$,

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

hence

$$QA' = AP \implies A' = Q^{-1}AP$$

□

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

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

with $Q \in \mathcal{M}_{m,m}$, $P \in \mathcal{M}_{n,n}$, with both invertible.

Remark. This defines an equivalence relation on $\mathcal{M}_{m,n}(F)$.

- $A = I_m^{-1} A I_n$
- $A' = Q^{-1}AP \implies A = (Q^{-1})^{-1}A'P^{-1}$
- $A' = Q^{-1}AP$, $A'' = (Q')^{-1}A'P'$. Then

$$A'' = (QQ')^{-1}A(PP')$$

Proposition. Let V, W be vector spaces over F, with $\dim_F V = n$, $\dim_F W = m$. Let $\alpha : V \to W$ be a linear map. Then there exists $\mathcal{B}$ basis of V and $\mathcal{C}$ basis of W such that

$$[\alpha]_{\mathcal{B},\mathcal{C}} = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

Proof. Choose $\mathcal{B}$ and $\mathcal{C}$ wisely.

- Fix $r \in \mathbb{N}$ such that $\dim \ker \alpha = n - r$.
- $N(\alpha) = \ker(\alpha) = \{x \in V, \alpha(x) = 0\}$
- Fix a basis of $N(\alpha)$: $v_{r+1}, \dots, v_n$. Extend it to a basis of V, so

$$\mathcal{B} = (v_1, \dots, v_r, \underbrace{v_{r+1}, \dots, v_n}_{\ker \alpha})$$

- Claim: $(\alpha(v_1), \dots, \alpha(v_r))$ is a basis of $\operatorname{im} \alpha$.
  - 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 $y \in \operatorname{im} \alpha$ then exists $v \in V$ such that $y = \alpha(v)$ then

$$y = \sum_{i=1}^{r} \lambda_i \alpha(v_i)$$

$$\implies y \in \langle \alpha(v_1), \dots, \alpha(v_r) \rangle$$

– Free:

$$\sum_{i=1}^r \lambda_i \alpha(v_i) = 0$$

$$\implies \alpha\left(\sum_{i=1}^r \lambda_i v_i\right) = 0$$

$$\implies \sum_{i=1}^r \lambda_i v_i \in \ker \alpha$$

$$\implies \sum_{i=1}^r \lambda_i v_i = \sum_{i=r+1}^n \mu_i v_i$$

$$\implies \sum_{i=1}^r \lambda_i v_i - \sum_{i=r+1}^n \mu_I v_i = 0$$

but since $\mathcal{B}$ is free, we must have $\lambda_i = 0, \mu_i = 0$ so it's free.

Conclusion: $(\alpha(v_1), \dots, \alpha(v_r))$ basis of $\operatorname{im} \alpha$, $(v_{r+1}, \dots, v_n)$ basis of $\ker \alpha$. Let

$$\mathcal{B} = (v_1, \dots, v_r, v_{r+1}, \dots, v_n)$$

$$\mathcal{C} = (\alpha(v_1, \dots, \alpha(v_r), w_{r+1}, \dots, w_m)$$

Then

$$[\alpha]_{\mathcal{B},\mathcal{C}} = (\alpha(v_1), \dots, \alpha(v_r), \alpha(v_{r+1}), \dots, \alpha(v_n))$$

□

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

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

Corollary. Any $m \times n$ matrix is equivalent to:

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

where $r = r(\alpha)$.

Start of lecture 7

Definition. $A \in \mathcal{M}_{m,n}(F)$

- The column rank of A, $r(A)$ is the dimension of the span of the column vectors of A in F^m, i.e. if $A = (c_1, \dots, c_n)$ then $r(A) = \dim_F \operatorname{span}\{c_1, \dots, c_n\}$.
- Similarly, the row rank is the column rank of $A^\top$.

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

$$r(A) = r(\alpha) = \dim \operatorname{im} \alpha$$

Proposition. Two matrices are equivalent if and only if $r(A) = r(A')$.

Proof. $\Rightarrow$ If A and A' are equivalent, then they correspond to the same linear map α written in two different bases

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

$\Leftarrow$ $r(A) = r(A') = r$, then both A and A' are equivalent to:

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

so A and A' are equivalent.

□

Theorem. $r(A) = r(A^\top)$ (column rank is the same as row rank)

Proof. Exercise. □

1.7 Elementary operations and elementary matrices

Special case of the change of basis formula.
Let $\alpha : V \to W$ be a linear map, $(\mathcal{B}, \mathcal{B}')$ bases of V, $(\mathcal{C}, \mathcal{C}')$ bases of W.

$$[\alpha]_{\mathcal{B},\mathcal{C}} \to [\alpha]_{\mathcal{B}',\mathcal{C}'}$$

If $V = W$, $\alpha : V \to V$ linear then we call it an endomorphism.

- $\mathcal{B} = \mathcal{C}$, $\mathcal{B}' = \mathcal{C}'$
- P is change of matrix from $\mathcal{B}'$ to $\mathcal{B}$.

then

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

> **Definition.** A, A' are $n \times n$ (square) matrices, we say that A and A' are *similar* (or *conjugate*) if and only if:
>
> - $A' = P^{-1}AP$
> - P is $n \times n$ square invertible.

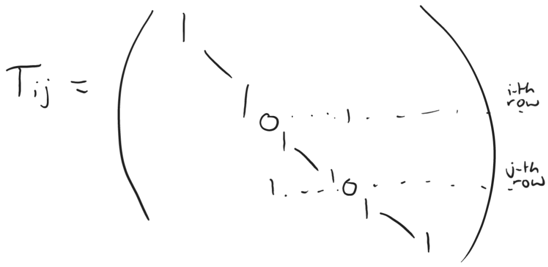
Central concept when we will study diagonalisation of matrices. (Spectral theory)

1.8 Elementary operations and elementary matrices

> **Definition.** Elementary *column* operation on an $m \times n$ matrix A:
>
> (i) swap columns i and j ($i \neq j$)
>
> (ii) replace column i by λ times column i ($\lambda \neq 0$, $\lambda \in F$)
>
> (iii) add λ times column i to column j (with $i \neq j$)

- Elementary row operations: analogous way
- Elementary operations are invertible
- These operations can be realised through the action of *elementary matrices.*

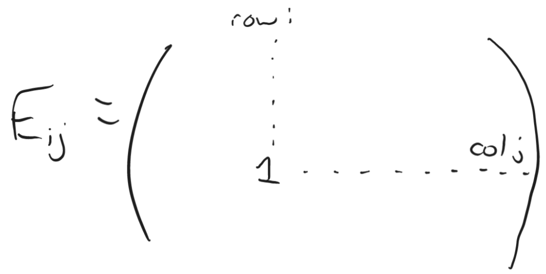
  (i) i, j, $i \neq j$.



  (ii) i

  (iii) i, j, λ, $i \neq j$

$$C_{i,j,\lambda} = \text{id} + E_{i,j}$$



Link between elementary operations / matrices:
an elementary column (row) operation can be performed by multiplying A by the corresponding elementary matrix from the right (left) $\rightarrow$ Exercise.

Now a constructive proof that any $m \times n$ matrix is equivalent to

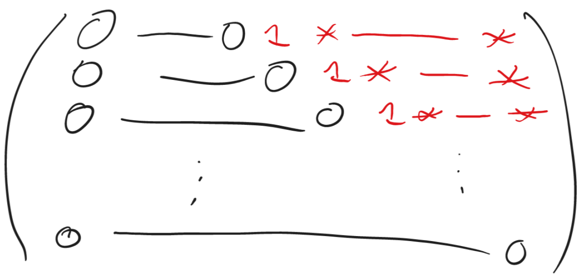
$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

- Start with A. If all entries are zero, done.
- Pick $a_{ij} = \lambda \neq 0$. Swap rows i and 1 and swap columns j and 1. Then λ is in position (1, 1)
- Multiply column 1 by $\frac{1}{\lambda}$ to get 1 in position (1, 1).
- Now clean out row 1 and column 1 using elementary operations of type (iii).
- Iterate with $\tilde{A}$ (the $(m-1) \times (n-1)$ sub matrix)
- Then at the end of the process we will have shown that

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) \equiv Q^{-1}AP = \underbrace{E'_p \cdots E'_1}_{\text{row operations}} \; A \; \underbrace{E_1 \cdots E_c}_{\text{column operations}}$$

Variation

- Gauss' pivot algorithm. If you use *only* row operations, we can reach the so called "row echelon form" of the matrix



- Assume that $a_{i1} \neq 0$ 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 first column
- Move to second column
- Iterate.

This procedure is exactly what you do when solving a linear system of equations: Gauss' pivot algorithm

Representation of square invertible matrices

Lemma. If A is $n \times n$ *square invertible* matrix, then we can obtain I_n using row elementary operations only (or column operations only).

Proof.

- We do the proof for column operations. We argue by induction on the number of rows
- Suppose that we could reach a form where the upper left corner is I_k. We want to obtain the same structure with $k \to k+1$.
- Claim: there exists $j > k$ such that $\lambda = a_{k+1,j} \neq 0$. Otherwise the vector $\delta_{i(k+1)}$ is *not* in the span of the column vectors of A (exercise) which contradicts the assumption that A is invertible.
- Swap column $k+1$ and j
- Divide column $k+1$ by $\lambda = a_{k+1,j} \neq 0$
- Use 1 to clear the rest of the $k+1$-th row using elementary operation of type (iii).
- This completes the inductive step.
- Continue until $k = n$.

□

Outcome:

$$AE_1 \cdots E_c = I_n$$
$$\implies A^{-1} = E_1 \cdots E_c$$

so this gives an algorithm for computing A^{-1}. (useful for solving $AX = F$, linear system of equations).

Proposition. Any invertible square matrix is a product of elementary matrices.

Start of lecture 8

1.9 Dual spaces and dual maps

Definition. Let V be a vector space. The we define

$$\begin{aligned} V^* &= \text{dual of } V \\ &= L(V, F) \\ &= \{\alpha : V \to F \text{ linear}\} \end{aligned}$$

Notation. $\alpha : V \to F$ linear. Then α is a linear form.

Examples

(i)

$$\begin{aligned} &\mathrm{Tr} : \mathcal{M}_{n,n}(F) \to F \\ &A = (a_{ij}) \mapsto \sum_{i=1}^n a_{ii} \\ &\implies \mathrm{Tr} \in \mathcal{M}_{n,n}^*(F) \end{aligned}$$

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

$$\begin{aligned} T_f &: \mathcal{C}^\infty([0,1], \mathbb{R}) \\ \varphi &\mapsto \int_0^1 f(x)\varphi(x)\mathrm{d}x \end{aligned}$$

then T_f is a linear form on $\mathcal{C}^\infty([0,1], \mathbb{R})$ ($\mathbb{R}$ vector space). Quantum mechanics. A function defines a linear form.

Lemma (Dual basis)**.** Let V be a vector space over F with a finite basis

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

Then there exists a basis for V^* given by

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

with

$$\varepsilon_j \left(\sum_{i=1}^n a_i e_i \right) = a_j, \quad 1 \le j \le n$$

We call $\mathcal{B}^*$ the dual basis of $\mathcal{B}$.

Remark. Kronecker symbol

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

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

Proof. Let $\{\varepsilon_1, \dots, \varepsilon_n\}$ be defined as above.

(1) Check that it is free: indeed, $\sum_{j=1}^{n} \lambda_j \varepsilon_j = 0$

$$\implies \sum_{j=1}^{n} \lambda_j \varepsilon_j(e_i) = 0$$

$$\implies \sum_{j=1}^{n} \lambda_i = 0 \quad \forall 1 \leq i \leq n$$

$$\implies \text{family is free}$$

(2) Check that it is generating: Pick $\alpha \in V^*$, then $x \in V$:

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

On the other hand, let the linear form:

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

Then:

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

□

Corollary. V finite dimensional,

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

Warning. These results about V^* are not relevant / very different when talking about infinite dimensional vector spaces instead.

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

$$\begin{aligned}
\alpha(x) &= \sum_{i=1}^n \alpha \varepsilon_i \left(\sum_{j=1}^n x_j e_j \right) \\
&= \sum_{i=1}^n a_i \sum_{j=1}^n x_j \varepsilon_i(e_j) \\
&= \sum_{i=1}^n a_i x_i \\
&= (\alpha_1, \dots, \alpha_n)
\end{aligned}$$

(scalar product structure)

Definition. If $U \le V$ (vector subspace), we define the *annihilator* of U by:

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

Lemma. (i) $U^0 \le V^*$ (vector subspace)

(ii) If $U \le V$ and $\dim V < \infty$ then $\dim V = \dim U + \dim U^0$.

Proof. (i) $0 \in U^0$. If $\alpha, \alpha' \in U^0$, then, for all $u \in U$,

$$(\alpha + \alpha')(u) = \alpha(u) + \alpha'(u) = 0$$

and for all $\lambda \in F$,

$$(\lambda \alpha)(u) = \lambda \alpha(u) = 0$$

so $U^0 \le V^*$.

(ii) Let $U \leq V$, $\dim V = n$. Let $(e_1, \dots, e_k)$ be a basis of U, complete it to a basis

$$\mathcal{B} = (e_1, \dots, e_k, e_{k+1}, \dots, e_n)$$

of V. Let $(\varepsilon_1, \dots, \varepsilon_n)$ be the dual basis of $\mathcal{B}$. We claim that $U^0 = \langle \varepsilon_{k+1}, \dots, \varepsilon_n \rangle$.

- Pick $i > k$, then:
$$\varepsilon_i(e_k) = \delta_{ik} = 0$$
so $\varepsilon_i \in U^0$, since $U = \langle e_1, \dots, e_k \rangle$. So
$$\langle \varepsilon_{k+1}, \dots, \varepsilon_n \rangle \leq U^0$$

- Let $\alpha \in U^0$, then let $\alpha \in V^*$, with
$$\alpha = \sum_{i=1}^n \alpha_i \varepsilon_i$$
Now for $i \leq k$:
$$\begin{aligned} \alpha \in U^0 &\implies \alpha(e_i) = 0 \ \forall 1 \leq i \leq k \\ &\implies \sum_{j=1}^n \alpha_j \varepsilon_j(e_i) = 0 \\ &\implies \alpha_i = 0 \quad \forall 1 \leq i \leq k \\ &\implies \alpha = \sum_{i=1}^n \alpha_i \varepsilon_i = \sum_{i=k+1}^n \alpha_i \varepsilon_i \\ &\implies \alpha \in \langle \varepsilon_{k+1}, \dots, \varepsilon_n \rangle \\ &\implies U^0 \leq \langle \varepsilon_{k+1}, \dots, \varepsilon_n \rangle \end{aligned}$$

□

Lemma. Let V, W be vector spaces over F. Let $\alpha \in \mathcal{L}(V, W)$. Then the map:

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

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

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

Proof.
- $\varepsilon \circ \alpha : V \to F$ linear follows by linearity of ε and α, so $\varepsilon \circ \alpha \in V^*$.

- α^* linear: let $\theta_1, \theta_2 \in W^*$, then
$$\begin{aligned}\alpha^*(\theta_1 + \theta_2) &= (\theta_1 + \theta_2)(\alpha) \\ &= \theta_1 \circ \alpha + \theta_2 \circ \alpha \\ &= \alpha^*(\theta_1) + \alpha^*(\theta_2)\end{aligned}$$
and similarly for all $\lambda \in F$,
$$\alpha^*(\lambda\theta) = \lambda\alpha^*(\theta)$$
so α^* is linear, i.e. $\alpha^* \in \mathcal{L}(V^*, W^*)$.

□

Proposition (Dual map matrix)**.** Let V, W be finite dimensional spaces over F with basis respectively $\mathcal{B}$ and $\mathcal{C}$. Let $\mathcal{B}^*$, $\mathcal{C}^*$ be the dual basis of $\mathcal{B}$ and $\mathcal{C}$. Then:
$$[\alpha^*]_{\mathcal{C}^*,\mathcal{B}^*} = [\alpha]_{\mathcal{B},\mathcal{C}}^\top$$

Proof. $\mathcal{B} = (b_1, \dots, b_n)$, $\mathcal{C} = (c_1, \dots, c_m)$, $\mathcal{B}^* = (\beta_1, \dots, \beta_n)$, $\mathcal{C}^* = (\gamma_1, \dots, \gamma_m)$. Say
$$[\alpha]_{\mathcal{B},\mathcal{C}} = A = (a_{ij})_{1\le i\le m, 1\le j\le n}$$
Recall: $\alpha^* : W^* \to V^*$. Let us compute:
$$\begin{aligned}\alpha^*(\gamma_r)(b_s) &= \gamma_r \circ \alpha(b_s) \\ &= \gamma_r(\alpha(b_s)) \\ &= \gamma_r\left(\sum_t a_{ts}c_t\right) \\ &= \sum_t a_{ts}\gamma_r(c_t) \\ &= a_{rs}\end{aligned}$$
Say
$$\begin{aligned}[\alpha^*]_{\mathcal{C}^*,\mathcal{B}^*} &= (\alpha^*(\gamma_1), \dots, \alpha^*(\gamma_m)) \begin{pmatrix}\beta_1 \\ \vdots \\ \beta_n\end{pmatrix} \\ &= (m_{ij})_{1\le i\le n, 1\le j\le m} \\ \implies \alpha^*(\gamma_r) &= \sum_{i=1}^n m_{ir}\beta_i \\ \implies \alpha^*(\gamma_r)(b_s) &= \sum_{i=1}^n m_{ir}\beta_i(b_s) \\ &= m_{sr}\end{aligned}$$
Conclusion $\alpha^*(\gamma_r)(b_s) = a_{rs} = m_{sr}$ so $[\alpha^*]_{\mathcal{C}^*,\mathcal{B}^*} = [\alpha]_{\mathcal{B},\mathcal{C}}^\top$. □

Start of lecture 9

1.10 Properties of the dual map, double dual

Let V, W be vector spaces over F, $\alpha \in L(V, W)$.

$$\mathcal{E} = (e_1, \dots, e_n)$$

basis of V

$$\mathcal{F} = (f_1, \dots, f_n)$$

another basis of V. Let

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

(change of basis matrix from $\mathcal{F}$ to $\mathcal{E}$)

$$\mathcal{E}^* = (\varepsilon_1, \dots, \varepsilon_n)$$

$$\mathcal{F}^* = (\eta_1, \dots, \eta_n)$$

Lemma. Let P be the change of basis matrix from $\mathcal{F}$ to $\mathcal{E}$. Then the change of basis matrix from $\mathcal{F}^*$to $\mathcal{E}^*$ is:

$$(P^{-1})^\top$$

Proof.

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

□

Properties of the dual map

Lemma. Let V, W be vector spaces over F. Let $\alpha \in \mathcal{L}(V, W)$ and $\alpha^* \in \mathcal{L}(W^*, V^*)$ be the dual map. Then:

(i) $N(\alpha^*) = (\mathrm{im}\,\alpha)^0$ (so α^* injective $\iff$ α surjective)

(ii) $\mathrm{im}\,\alpha^* \le (N(\alpha))^0$ with *equality* if V, W are finite dimensional (hence in this case, α^* surjective $\iff$ α injective).

Dual method: there are many problems (controllability) where the understanding of α^* is simpler than the understanding of α.

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

$$\begin{aligned} \varepsilon \in N(\alpha^*) &\iff \alpha^*(\varepsilon) = 0 \\ &\iff \alpha^*(\varepsilon) = \varepsilon(\alpha) = 0 \\ &\iff \forall x \in V,\ \varepsilon(\alpha)(x) = \varepsilon(\alpha(x)) = 0 \\ &\iff \varepsilon \in (\mathrm{im}\,\alpha) \end{aligned}$$

(ii) Let us first show that:

$$\operatorname{im}(\alpha^*) \leq (N(\alpha))^0$$

Indeed, let $\varepsilon \in \operatorname{im}(\alpha^*)$

$$\implies \varepsilon = \alpha^*(\varphi), \ \varphi \in W^*$$
$$\implies \forall u \in N(\alpha) \mid \varepsilon(u) = \alpha^*(\varphi)(u) = \varphi \circ \alpha(u)\varphi(\alpha(u)) = 0$$
$$\implies \varepsilon \in (N(\alpha))^0$$

In finite dimension, we can compute the dimensions of $\operatorname{im}(\alpha^*)$ and $(N(\alpha))^0$.

$$\dim(\operatorname{im}(\alpha^*)) = r(\alpha^*)$$

$$r(\alpha^*) = r([\alpha^*]_{\mathcal{C}^*,\mathcal{B}^*}) = r([\alpha]_{\mathcal{B},\mathcal{C}}^\top) = r([\alpha]_{\mathcal{B},\mathcal{C}}) = r(\alpha)$$
$$\implies r(\alpha) = r(\alpha^*)$$

so

$$\begin{aligned}\dim(\operatorname{im}\alpha^*) &= r(\alpha^*)\\ &+ r(\alpha)\\ &= \dim V - \dim N(\alpha)\\ &= \dim[(N(\alpha))^0]\end{aligned}$$

so $\operatorname{im}(\alpha^*) \leq (N(\alpha))^0$ and $\dim(\operatorname{im}(\alpha^*)) = \dim[(N(\alpha))^0]$ so $\operatorname{im}(\alpha^*) = [N(\alpha)]^0$. □

Double dual

- V vector space over F
- $V^* = \mathcal{L}(V, f)$ dual of V. We define the bidual:

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

Very important space in infinite dimension: in general, there is no obvious connection between V and V^* (unless Hilbertian structure). However, there is a large class of function spaces wuch that $V \simeq V^{**}$.

Example. $p > 2$,

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

Is a reflexive space.

In general, there is a canonical embedding of V into V^{**}. Indeed, pick $v \in V$, we define:

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

$$\varepsilon \mapsto \varepsilon(v)$$

linear:

- $\varepsilon \in V^*$ implies $\varepsilon(v) \in F$.
- linearity: $\lambda_1, \lambda_2 \in F, \varepsilon_1, \varepsilon_2 \in V^*$

$$\begin{aligned} \hat{v} = (\lambda_1\varepsilon_1 + \lambda_2\varepsilon_2) &= (\lambda_1\varepsilon_1 + \lambda_2\varepsilon_2)(v) \\ &= \lambda_1\varepsilon_1(v) + \lambda_2\varepsilon_2(v) \\ &= \lambda_1\hat{v}(\varepsilon_1) + \lambda_2\hat{v}(\varepsilon_2) \end{aligned}$$

  so $\hat{v} \in \mathcal{L}(V^*, F)$.

Theorem. If V is a *finite dimensional* vector space over F, then:

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

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

is an isomorphism.

(in infinite dimension we can show under canonical assumptions (Banach space) that: $\hat{}$ is injective)

Proof.

- V finite dimensional. Given $v \in V$, $\hat{v} \in V^{**} \in \mathcal{L}(V^*, F)$.
- $\hat{}$ linear: let $v_1, v_2 \in V$, $\lambda_1, \lambda_2 F$, $\varepsilon \in V^*$:

$$\begin{aligned} (\widehat{\lambda_1 v_1 + \lambda_2 v_2})(\varepsilon) &= \varepsilon(\lambda_1 v_1 + \lambda_2 v_2) \\ &= \lambda_1\varepsilon(v_1) + \lambda_2\varepsilon(v_2) \\ &= \lambda_1\hat{v}_1(\varepsilon) + \lambda_2\hat{v}_2(\varepsilon) \\ \implies (\widehat{\lambda_1 v_1 + \lambda_2 v_2}) &= \lambda_1\hat{v}_1 + \lambda_2\hat{v}_2 \end{aligned}$$

- $\hat{}$ injective: indeed, let $e \in V \setminus \{0\}$. I extend $(e, e_2, \ldots, e_n)$ basis of V. Let $(\varepsilon, \varepsilon_2, \ldots, \varepsilon_n)$ the dual basis of (of V^*), then

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

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

$$\implies N(\hat{}) = \{0\}$$

  so $\hat{}$ is injective.

- $\hat{}$ isomorphism. We can compute dimensions:

$$\dim V = \dim V^* = \dim[(V^*)^*] = \dim(V^{**})$$

As a conclusion: $\hat{}: V \to V^{**}$ is injective, $\dim V = \dim V^{**}$, so $\hat{}$ is surjective, so $\hat{}$ is an isomorphism.

□

Lemma. Let V be a finite dimensional vector space over F, let $U \leq V$. Then

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

so after identification of V and V^{**}, we have

$$U \simeq U^{00}$$

Proof. Let us show that: $U \leq U^{00}$.

- Indeed, let $u \in U$:

$$\begin{aligned} &\forall \varepsilon \in U^0, \varepsilon(u) = 0 \\ \implies &\forall \varepsilon \in U^0, \varepsilon(u) = \hat{u}(\varepsilon) = 0 \\ \implies &\hat{u} \in U^{00} \\ \implies &\hat{U} \subset U^{00} \end{aligned}$$

- Commute dimensions

$$\dim U^{00} = \dim V - \dim U^0 = \dim U$$

□

Remark. $T \leq V^*$

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

Start of lecture 10

Remark. $T \leq V^+$, we can define

$$T^0 = \{v \in V \mid \theta(v) = 0, \theta \in T\}$$

Lemma. 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. (i) Exercise.

(ii) Take ^{0} of (i) and use $U^{00} = U$.

□

1.11 Bilinear Forms

$\implies$ Quadratic algebra.

Definition. U, V vector spaces over F. Then:

$$\varphi : U \times V \to F$$

is a *bilinear form* if it "linear in both components":

- $\varphi(u, \bullet) : V \to F$ is linear for all $u \in U$ $(v \mapsto \varphi(u, v))$.
- $\varphi(\bullet, v) : U \to F$ is linear for all $v \in V$ $(u \mapsto \varphi(u, v))$

Examples

(i) $V \times V^* \to F$

$$(v, \theta) \mapsto \theta(v)$$

(ii) Scalar product / dot product on $U = V = \mathbb{R}^n$

(iii) $U = V = \mathcal{C}([0,1], \mathbb{R})$

$$\varphi(f, g) = \int_0^1 f(t)g(t)\mathrm{d}t$$

("infinite dimensional scalar product)

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

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

Lemma. $\varphi(u, v) = [u]_{\mathcal{B}}^{\top}[\varphi]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{C}}$. $(*)$

Link between the bilinear form and its matrix in given basis.

Proof. $u = \sum_{i=1}^{m} \lambda_i e_i$, $v = \sum_{j=1}^{n} \mu_j f_j$. Then by linearity:

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

□

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

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

$$\varphi_L : U \to V^*$$

$$\varphi_L(u) : V \to F$$

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

define φ_R similarly.

Lemma. $\mathcal{B} = (e_1, \ldots, e_m)$ basis of U, $B^* = (\varepsilon_1, \ldots, \varepsilon_m)$ dual basis of U^*. $\mathcal{C} = (f_1, \ldots, f_n)$ basis of V, $\mathcal{C}^* = (\eta_1, \ldots, \eta_n)$ dual basis of V^*. Let $A = [\varphi]_{\mathcal{B},\mathcal{C}}$ then:

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

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

Proof.

$$\begin{aligned}
\varphi_L(e_i)(f_j) = \varphi(e_i, f_j) = A_{ij} \\
\implies \varphi_L(e_i) = \sum A_{ij}\eta_j
\end{aligned}$$

similarly for φ_R. □

Definition (Degenerate / non degenerate bilinear form)**.** $\ker \varphi_L$: "left kernel of φ", $\ker \varphi_R$: "right kernel of φ". We say that φ is non-degenerate if

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

Otherwise, we say that φ is degenerate.

Lemma. U, V finite dimensional. $\mathcal{B}$ basis of U, $\mathcal{C}$ basis of V. $\varphi : U \times V \to F$ bilinear form, $A = [\varphi]_{\mathcal{B},\mathcal{C}}$. Then φ non degenerate $\iff$ A invertible.

Corollary. φ non degenerate

$$\implies \dim U = \dim V$$

Proof.

$$\begin{aligned} \varphi \text{ non degenerate } &\iff \ker \varphi_L = \{0\} \text{ and } \ker \varphi_R = \{0\} \\ &\iff n(A^\top) = 0 \text{ and } n(A) = 0 \\ &\iff r(A^\top) = \dim U \text{ and } r(A) = \dim V \\ &\iff A \text{ invertible and then: } \dim U = \dim V \end{aligned}$$

□

Remark. $\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ scalar product, then φ is non degenerate, and if we take the usual bases, then $[\varphi]_{\mathcal{B},\mathcal{B}} = I_n$.

Corollary. When U and V are finite dimensional, then choosing a *non degenerate* bilinear form $\varphi : U \to V \to F$ is equivalent to choosing an isomorphism $\varphi_L : U \to V^*$.

Definition. $T \subset U$, we define:

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

Similarly define for $S \subset V$

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

Change basis for bilinear forms

Proposition.
- $\mathcal{B}$, $\mathcal{B}'$ basis of U, $P = [\mathrm{id}]_{\mathcal{B}',\mathcal{B}}$
- $\mathcal{C}, \mathcal{C}'$ basis of V, $Q = [\mathrm{id}]_{\mathcal{C}',\mathcal{C}}$.

Let $\varphi : U \times V \to F$ bilinear form, then

$$[\varphi]_{\mathcal{B}',\mathcal{C}'} = P^\top [\varphi]_{\mathcal{B},\mathcal{C}} Q$$

change of basis formula for bilinear forms.

Proof.

$$\begin{aligned}
\varphi(u, v) &= [u]_{\mathcal{B}}^\top [\varphi]_{\mathcal{B},\mathcal{C}} [v]_{\mathcal{C}} \\
&= (P[u]_{\mathcal{B}'})^\top [\varphi]_{\mathcal{B},\mathcal{C}} (Q[v]_{\mathcal{C}'}) \\
&= [u]_{\mathcal{B}'}^\top (P^\top [\varphi]_{\mathcal{B},\mathcal{C}} Q)[v]_{\mathcal{C}'}
\end{aligned}$$

□

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

Indeed, $r(P^\top AQ) = r(A)$ for any invertible P, Q.

Remark. $r(\varphi) = r(\varphi_R) = r(\varphi_L)$. (we computed matrices in a basis and $r(A) = r(A^\top)$)

More applications later: scalar product.

1.12 Determinant and Traces

Definition. $A \in \mathcal{M}_n(F) = \mathcal{M}_{n\times n}(F)$ We define the trace of A

$$\mathrm{Tr}\, A = \sum_{i=1}^{n} A_{ii}$$

$$A = (A_{ij})_{1\le i,j\le n}$$

Remark. $\mathcal{M}_n(F) \to F$ linear form ($A \mapsto \mathrm{Tr}\, A$).

Lemma. $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

Proof.

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

□

Start of lecture 11

Corollary. Similar matrices have the same trace

Proof.

$$\begin{aligned}\mathrm{Tr}(P^{-1}AP) &= \mathrm{Tr}(APP^{-1}) \\ &= \mathrm{Tr}(A)\end{aligned}$$

□

Definition. If $\alpha : V \to V$ linear (endomorphism) we can define:

$$\mathrm{Tr}\, \alpha = \mathrm{Tr}([\alpha]_{\mathcal{B}})$$

in any basis $\mathcal{B}$ (does not depend on the choice $\mathcal{B}$).

Lemma. $\alpha : V \to V$, $\alpha^* : V^* \to V^*$ dual map, then

$$\mathrm{Tr}\, \alpha = \mathrm{Tr}\, \alpha^*$$

Proof.

$$\begin{aligned}\mathrm{Tr}\, \alpha &= \mathrm{Tr}([\alpha]_{\mathcal{B}}) \\ &= \mathrm{Tr}([\alpha]_{\mathcal{B}}^\top) \\ &= \mathrm{Tr}([\alpha^*]_{\mathcal{B}^*}) && (\text{because } [\alpha]_{\mathcal{B}}^\top = [\alpha^*]_{\mathcal{B}^*})\end{aligned}$$

□

1.13 Determinants

Permutations and transpositions

- permutation: S_n = group of permutations of $\{1, \ldots, n\}$

$$\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$$

  is a bijection. Then σ is a permutation.

- Transposition $k \neq l$, $\tau_{kl} \in S_n$ just swaps k and l.

- Decomposition: any permutation of σ can be decomposed as a product of transpositions

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

  τ_i transposition.

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

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

  $\varepsilon(\sigma)$ = signature of σ. *and* ε is a group homomorphism.

Definition (Determinant)**.** $A \in \mathcal{M}_n(F)$ (square matrix),

$$A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$$

We define the determinant of A as:

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

Example.

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} a_{11}a_{22} - a_{12}a_{21}$$

Lemma. If $A = (a_{ij})$ is an upper (lower) triangular matrix with 0 on the diagonal:

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

then $\det A = 0$.

Proof. For $a_{\sigma(1)1} \cdots a_{\sigma(n)n}$ not to be zero, I need $\sigma(j) < j$ for all $j \in \{1, \ldots, n\}$ which is *impossible* for $\sigma \in S_n$. So all the terms are 0, so $\det A = 0$. $\square$

Exercise: Show similarly that if instead we allow the diagonal elements to be nonzero, then the determinant is the product of the diagonal elements.

Lemma. $\det A = \det(A^\top)$

Proof.

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \\
&= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i} \\
&= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^{n} a_{j\sigma^{-1}(j)}
\end{aligned}$$

Now remember $\varepsilon(\sigma\sigma^{-1}) = \varepsilon(\sigma)\varepsilon(\sigma^{-1})$ so since $\varepsilon(\sigma) \in \{-1, 1\}$,

$$\implies \varepsilon(\sigma^{-1}) = \varepsilon(\sigma)$$

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i\sigma^{-1}(i)} \\
&= \sum_{\sigma \in S_n} \varepsilon(\sigma^{-1}) \prod_{i=1}^{n} a_{i\sigma^{-1}}(i) \\
&= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} \\
&= \det(A^\top)
\end{aligned}$$

$\square$

Why this formula for $\det A$?

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

$$\underbrace{F^n \times \cdots \times F^n}_{n \text{ times}} \to F$$

such that

(i) d is multilinear: for any $1 \leq i \leq n$, for all $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n \in F^n$,

$$v \mapsto d(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n)$$

is linear (i.e. an element of $(F^n)^*$) (linear with respect to all coordinate)

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

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

We want to show that there is in fact only *one* (up to a multiplicative constant) volume form on $F^n \times \cdots \times F^n$ which is given by the determinant:

$$A = (a_{ij}) = (A^{(1)} \mid \cdots \mid A^{(n)})$$

(column vectors)

$$\det A = \det(A^{(1)}, \ldots, A^{(n)})$$

Lemma. $F^n \times \cdots \times F^n \to F$

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

is a volume form.

Proof. (i) multilinear $\sigma \in S_n$, then $\prod_{i=1}^{n} a_{\sigma(i)i}$ is multilinear: there is only one term from each column appearing in the expression. The sum of multilinear maps is multilinear, so det is multilinear.

(ii) Alternate: Assume $k \neq l$, $A^{(k)} = A^{(l)}$. I want to show $\det A = 0$. Indeed: let τ be the transposition which swaps k and l. Then since $A^{(k)} = A^{(l)}$ then $a_{ij} = a_{i\tau j}$ for all i, j. We can decompose:

$$S_n = A_n \sqcup \tau A_n$$

then

$$\begin{aligned}\det A &= \sum_{\sigma\in S_n} \varepsilon(\sigma)\prod_{i=1}^n a_{i\sigma(i)} \\ &= \sum_{\sigma\in A_n}\prod_{i=1}^n a_{i\sigma(i)} + \sum_{\sigma\in\tau A_n}\varepsilon(\sigma)\prod_{i=1}^n a_{i\tau\sigma(i)} \\ &= \sum_{\sigma\in A_n}\prod_{i=1}^n a_{i\sigma(i)} - \sum_{\sigma\in A_n} a_{i\tau\sigma(i)} \\ &= \sum_{\sigma\in A_n}\prod_{i=1}^n a_{i\sigma(i)} - \sum_{\sigma\in A_n}\prod_{i=1}^n a_{i\sigma(i)} \\ &= 0\end{aligned}$$

□

Lemma. Let d be a volume form. Then swapping two entries changes the sign.

Proof. Equivalent definition of "alternate".

$$\begin{aligned}0 &= d(v_1,\dots,v_i+v_j,\dots,v_i+v_j,\dots,v_n) \\ &= d(v_1,\dots,v_i,\dots,v_i,\dots,v_n) + d(v_1,\dots,v_i,\dots,v_j,\dots,v_n) \\ &\quad + d(v_1,\dots v_j,\dots,v_i,\dots,v_n) + d(v_1,\dots,v_j,\dots,v_j,\dots,v_n) \\ &= 0 + d(v_1,\dots,v_i,\dots,v_j,\dots,v_n) + d(v_1,\dots,v_j,\dots,v_i,\dots,v_n)\end{aligned}$$

□

Start of lecture 12

Corollary. $\sigma\in S_n$, d volume form, then:

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

Proof. $\sigma = \prod_{i=1}^{n\sigma}\tau_i$. □

Theorem. Let d be a volume form on F^n. Let $A = (A^{(1)}|\cdots|A^{(n)})$. Then

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

Up to a constant, det is the *only* volume form on F^n.

Proof.

$$\begin{aligned}
d(A^{(1)}, \ldots, A^{(n)}) &= d\left(\sum_{i=1}^{n} a_{i1} e_i, \ldots, A^{(n)}\right) \\
&= \sum_{i=1}^{n} a_{i1} d(e_i, A^{(2)}, \ldots, A^{(n)}) \\
&= \sum_{i=1}^{n} a_{i1} d\left(e_i, \sum_{j=1}^{n} a_{j2} e_j, \ldots, A^{(n)}\right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i1} a_{j2} d(e_i, e_j, \ldots, A^{(n)}) \\
&= \sum_{\substack{1 \le i_1 \le n \\ 1 \le i_2 \le n \\ \vdots \\ 1 \le i_n \le n}} \left(\prod_{k=1}^{n} a_{i_k k} d(e_{i_1}, e_{i_2}, \ldots, e_{i_n})\right)
\end{aligned}$$

The last d term is nonzero only if all the i_k are different, so we can write the i_k as a permutation. This means we can continue and get

$$\begin{aligned}
d(A^{(1)}, \ldots, A^{(n)}) &= \sum_{\sigma \in S_n} \prod_{k=1}^{n} a_{\sigma(k)k} d(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) \\
&= \sum_{\sigma \in S_n} \left[\prod_{k=1}^{n} a_{\sigma(k)k}\right] \varepsilon(\sigma) d(e_1, \ldots, e_n) \\
&= d(e_1, \ldots, e_n) \left[\sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{k=1}^{n} a_{\sigma(k)k}\right] \\
&= d(e_1, \ldots, e_n) \det A
\end{aligned}$$

□

Corollary. det is the *only* volume form such that

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

1.14 Some properties of determinants

Lemma. $A, B \in \mathcal{M}_n(F)$, then:

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

Proof. Indeed, pick A. Consider the map:

$$d_A : \underbrace{F^n \times \cdots \times F^n}_{n} \to F$$

defined by

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

Then:

- d_A is multilinear: $v_i \mapsto Av_i$ is linear.
- d_A is alternate: if $v_i = v_j$ then $Av_i = Av_j$.

so d_A is a volume form. In particular,

$$d_A(v_1, \ldots, v_n) = C \det(v_1, \ldots, v_n)$$

Now we compute C. $Ae_i = (A)$ so

$$d_A(e_1, \ldots, e_n) = \det(Ae_1, \ldots, Ae_n) = \det(a_1, \ldots, A_n) = \det A$$

So

$$C = \det A$$

We have proved:

$$\begin{aligned} d_A(v_1, \ldots, v_n) &= d(Av_1, \ldots, Av_n) \\ &= (\det A) \det(v_1, \ldots, v_n) \end{aligned}$$

Now observe:

$$AB = ((AB)_1, \ldots, (AB)_n)$$
$$(AB)_i = AB_i$$

so

$$\begin{aligned} \det(AB) &= \det(AB_1, \ldots, AB_n) \\ &= \det(A) \det(B_1, \ldots, B_n) \\ &= \det(A) \det(B) \end{aligned}$$

□

Definition. $A \in \mathcal{M}_n(F)$, we say that:

(i) A is *singular* if $\det A = 0$

(ii) A is *non singular* if $\det A \neq 0$.

Lemma. A is invertible implies A is non singular.

Proof. A is invertible.

$$\begin{aligned}
&\implies \exists A^{-1}, AA^{-1} = A^{-1}A = I_n \\
&\implies \det(AA^{-1}) = \det(A^{-1}A) = \det I_n = 1 \\
&\implies (\det A)(\det A^{-1}) = 1 \\
&\implies \det A \neq 0
\end{aligned}$$

□

Remark. We have proved that

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

Theorem. Let $A \in \mathcal{M}_n(F)$. Then the following are equivalent:

(i) A is invertible

(ii) A is non singular

(iii) $r(A) = n$

Proof. (i) $\iff$ (iii) done (rank nullity Theorem). (i) $\implies$ (iii) is lemma above. We need to show (ii) $\implies$ (iii). Indeed, assume $r(A) < n$. Then

$$\begin{aligned}
&\iff \dim \operatorname{span}\{c_1, \dots, c_n\} < n \\
&\implies \exists(\lambda_1, \dots, \lambda_n) \neq (0, \dots, 0)
\end{aligned}$$

$$\sum_{i=1}^{n} \lambda_i c_i = 0$$

I pick j such that $\lambda_j \neq 0$

$$\implies c_j = -\frac{1}{\lambda_j} \sum_{i \neq j} \lambda_i c_i$$

$$\begin{aligned}
\implies \det A &= \det(c_1, \dots, c_j, \dots, c_n) \\
&= \det\left(c_1, \dots, -\frac{1}{\lambda_j} \sum_{i \neq j} \lambda_i c_i, \dots, c_n\right) \\
&= \sum_{i \neq j} -\frac{1}{\lambda_j} \det(c_1, \dots, c_i, \dots, c_n) \\
&= 0
\end{aligned}$$

□

Remark. This gives us the sharp criterion for invertibility of a linear system of n equations with n unknowns:

$$Y \in F^n$$

$$A \in \mathcal{M}_n(F)$$

$$AX = Y, X \in F^n$$

exists a unique solution if and only if A is invertible, which happens if and only if $\det A \neq 0$.

Determinant of linear maps

Lemma. Conjugate matrices have the same determinant.

Proof.

$$\begin{aligned}\det(P^{-1}AP) &= \det(P^{-1})\det A \det P \\ &= \frac{1}{\det P}\det A \det P \\ &= \det A\end{aligned}$$

(P invertible implies $\det P \neq 0$). □

Definition. $\alpha : V \to V$ linear (endomorphism). We define

$$\det \alpha = \det([\alpha]_{\mathcal{B}})$$

$\mathcal{B}$ is any basis of V. This number does not depend on the choice of the basis!

Theorem. $\det : L(V, V) \to F$ satisfies:

(i) $\det \mathrm{id} = 1$

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

(iii) $\det(\alpha) \neq 0$ if and only if α is invertible and then

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

Proof. Pick a basis and express in terms of $[\alpha]_{\mathcal{B}}$ and $[\beta]_{\mathcal{B}}$. □

Determinant of block matrices

Lemma. $A \in \mathcal{M}_k(F)$, $\mathcal{B} \in \mathcal{M}_l(F)$ and $C \in \mathcal{M}_{k,l}(F)$. Let

$$M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \in \mathcal{M}_n(F)$$

$(n = k + l)$ then

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

Proof.

$$\det M = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} m_{\sigma(i)i} \qquad (*)$$

Observation:

$$m_{\sigma(i)i} = 0$$

if $i \le k$, $\sigma(i) > k$. So In $(*)$, we need only sum over $\sigma \in S_n$ such that:

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

(ii) and hence $\forall j \in [k + 1, n]$, $\sigma(j) \in [k + 1, n]$.

In other words, we restrict to σ of the form:

$$\sigma_1 : \{1, \dots, k\} \to \{1, \dots, k\}$$

$$\sigma_2 : \{k + 1, \dots, n\} \to \{k + 1, \dots, n\}$$

(i) $m_{\sigma(j)j}$ with $1 \le j \le k$, then $\sigma(j) \in \{1, \dots, k\}$, can be rewritten as

$$m_{\sigma(j)j} = a_{\sigma(j)j} = a_{\sigma_1(j)j}$$

(ii) Similarly, for $k + 1 \le j \le n$, $k + 1 \le \sigma(j) \le n$,

$$m_{\sigma(j)j} = b_{\sigma(j)j} = b_{\sigma_1(j)j}$$

Note that

$$\varepsilon(\sigma) = \varepsilon(\sigma_1)\varepsilon(\sigma_2)$$

so then

$$
\begin{aligned}
\det M &= \sum_{\sigma S_n} \varepsilon(\sigma) \prod_{i=1}^{n} m_{\sigma(i)i} \\
&= \sum_{\substack{\sigma_1 \in S_k \\ \sigma_2 \in S_l}} \varepsilon(\sigma_1 \circ \sigma_2) \prod_{i=1}^{k} a_{\sigma_1(i)i} \prod_{j=k+1}^{n} b_{\sigma_2(j)j} \\
&= \sum_{\substack{\sigma_1 \in S_k \\ \sigma_2 \in S_l}} \varepsilon(\sigma_1)\varepsilon(\sigma_2) \prod_{i=1}^{k} a_{\sigma_1(i)i} \prod_{j=k+1}^{n} b_{\sigma_2(j)j} \\
&= \left(\sum_{\sigma_1 \in S_k} \varepsilon(\sigma_1) \prod_{i=1}^{k} a_{\sigma_1(i)i} \right) \left(\sum_{\sigma_2 \in S_l} \varepsilon(\sigma_2) \prod_{j=k+1}^{n} a_{\sigma_2(j)j} \right) \\
&= (\det A)(\det B)
\end{aligned}
$$

□

Start of lecture 13

Corollary. $A_1, \dots, A_k$ are square matrices, then

$$
\det \begin{pmatrix} A_1 & * & * & \cdots & * \\ 0 & A_2 & * & \cdots & * \\ 0 & 0 & A_3 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & A_k \end{pmatrix} = (\det A_1) \cdots (\det A_k)
$$

Proof. By induction on k. □

In particular, if A is filled with zeros below the diagonal, then $\det A$ is the product of the entries on the diagonal. (But this is also quite easy to show directly from the definition of $\det A$).

Warning. In general:

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

Remark. In $\mathbb{R}^3$, we have that $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ is a volume form (and represents the volume of a parallelepiped), and in fact, $\det(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$.

1.15 Adjugate matrix

Observation: We know that swapping two column vectors flips the sign of the determinant, and we also know that $\det A = \det A^\top$. So we find that swapping two rows changes the determinant by a factor of -1.

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

Column (line) expansion and adjugate matrix

Column expansion is to reduce the computation of $n \times n$ determinants to $(n-1) \times (n-1)$ determinants. Very useful to compute determinants.

Definition. $A \in \mathcal{M}_n(F)$. Pick $i, j \leq n$. We define:

$$A_{\hat{i}\hat{j}} \in \mathcal{M}_{n-1}(F)$$

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

Example.

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

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

Lemma (Expansion of the determinant)**.** Let $A \in \mathcal{M}_n(F)$.

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

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

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

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

Powerful tool to compute determinants.

Example.

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 3 & -1 & 1 \\ 4 & 2 & -7 \end{pmatrix}$$

$$\det A = -(2)\begin{vmatrix} 3 & 1 \\ 4 & -7 \end{vmatrix} + (-1)\begin{vmatrix} 1 & -1 \\ 4 & -7 \end{vmatrix} - 2\begin{vmatrix} 1 & -1 \\ 3 & 1 \end{vmatrix}$$

Proof. Expansion with respect to the j-th column (row expansion formula follows by taking transpose). Pick $1 \leq j \leq n$.

- $A = (A^{(1)} \mid A^{(2)} \mid \cdots \mid A^{(j)} \mid \cdots \mid A^{(n)})$

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

- $$\begin{aligned} \det A &= \det\left(A^{(1)}, \ldots, \sum_{i=1}^{n} a_{ij} e_i, \ldots, A^{(n)}\right) \\ &= \sum_{i=1}^{n} a_{ij} \det(A^{(1)}, \ldots, e_i, \ldots, A^{(n)}) \\ \det(A^{(1)} \mid \cdots \mid e_i \mid \cdots \mid A^{(n)}) &= (-1)^{j-1} \det(e_i \mid A^{(1)} \mid A^{(j-1)} \mid A^{(j+1)} \mid \cdots \mid A^{(n)}) \\ &= (-1)^{i-1}(-1)^{j-1} \det(A_{\hat{i}\hat{j}}) \\ &= (-1)^{i+j} \det(A_{\hat{i}\hat{j}}) \end{aligned}$$

  so

$$\begin{aligned} \det A &= \sum_{i=1}^{n} a_{ij} \det(A^{(1)}, \ldots, a^{(j-1)}, e_i, \ldots, A^{(n)}) \\ &= \sum_{i=1}^{n} a_{ij} (-1)^{i+j} \det A_{\hat{i}\hat{j}} \end{aligned}$$

□

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

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

Theorem. Let $A \in \mathcal{M}_n(F)$, then

$$\operatorname{adj}(A)A = (\det A)I_n$$

In particular, when A is invertible,

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

Proof. We just proved: $(*)$

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

- For $j \neq k$ we have

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

  for $j \neq k$.

So done. □

Cramer rule

Proposition. Let $A \in \mathcal{M}_n(F)$ be invertible. Let $b \in F^n$. Then the unique solution to $Ax = b$ is given by:

$$x_o = \frac{1}{\det A}\det(A_{\hat{i}}b)$$

$1 \leq i \leq n$ where $A_{\hat{i}}b$ is obtained by replacing the i-th column of A by b.

Algorithmically, this avoids computing A^{-1}.

TODO: CHECK WHETHER NEEDS EDITING.

Start of lecture 14

1.16 Eigenvectors, eigenvalues and trigonal matrices

First step towards the diagonalisation of endomorphisms.

- V vector space over F, $\dim_F V = n < \infty$. $\alpha : V \to V$ linear (endomorphism of V). General problem: 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.

Reminder: $\mathcal{B}'$ another basis of V, P = change of basis matrix,

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

Equivalently: given a matrix $A \in \mathcal{M}_n(F)$, is it conjugated to a matrix with a "simple" form?

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

$$[\alpha]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

(ii) $\alpha \in \mathcal{L}(V)$ is *triangulable* if there exists $\mathcal{B}$ basis of V such that $[\alpha]_{\mathcal{B}}$ is triangular:

$$[\alpha]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Remark. A matrix is diagonalisable (respectively triangulable) if and only if it is conjugated to a diagonal (respectively triangular) matrix.

Definition (eigenvalue, eigenvector, eigenspace). (i) $\lambda \in F$ is an eigenvalue of $\alpha \in \mathcal{L}(V)$ if and only if there exists $v \in V \setminus \{0\}$ such that $\alpha(v) = \lambda v$.

(ii) $v \in V$ is an eigenvector of $\alpha \in \mathcal{L}(v)$ if and only if $v \neq 0$ and there exists $\lambda \in F$ such that $\alpha(v) = \lambda v$.

(iii) $V_\lambda = \{v \in V \mid \alpha(v) = \lambda v\} \leq V$ is the eigenspace associated to $\lambda \in F$.

Remark. Once can write evalue, evectors, espace.

Lemma. $\alpha \in L(v)$, $\lambda \in F$, then

$$\lambda \text{ eigenvalue} \iff \det(\alpha - \lambda \text{id}) = 0$$

Proof.

$$\begin{aligned}
\lambda \text{ eigenvalue} &\iff \exists v \in V \setminus \{0\} \mid \alpha(v) = \lambda v \\
&\iff \exists v \in V \setminus \{0\} \mid (\alpha - \lambda \text{id})(v) = 0 \\
&= \ker(\alpha - \lambda \text{id}) \neq \{0\} \\
&= \alpha - \lambda \text{id not injective} \\
&\iff \alpha - \lambda \text{id not surjective} \\
&\iff \alpha - \lambda \text{id not bijective} \\
&\iff \det(\alpha - \lambda \text{id}) = 0
\end{aligned}$$

□

Remark. If $\alpha(v_j) = \lambda v_j$, $v_j \neq 0$. I can complete it to a basis $(v_1, \dots, v_{j-1}, v_j, \dots, v_n)$ of V. Then

$$[\alpha]_{\mathcal{B}} = (| \cdots | (\lambda \text{ in } j\text{-th entry}) | \cdots)$$

Elementary facts about polynomials

We will study $P(\alpha)$, P polynomial. $\alpha \in \mathcal{L}(V)$.

- F field,

  $$f(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$$

  $a_i \in F$. $n \equiv$ 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 $f(t) \iff f(\lambda) = 0$.

Lemma. λ is a root of f, then $t - \lambda$ divides f:

$$f(t) = (t - \lambda)g(t), \quad g(t) \in F(t)$$

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

$$\begin{aligned} f(t) &= f(t) - f(\lambda) \\ &= a_n(t^n - \lambda^n) + \cdots + a_1(t - \lambda) \\ &= a_n(t - \lambda)(t^{n-1} + \lambda t^{n-2} + \cdots \lambda^{n-2} t + \lambda^{n-1}) + \cdots \end{aligned}$$

□

Corollary. A *nonzero* polynomial of degree n (≥ 0) has *at most* n roots (counted with multiplicity).

Proof. Induction on the degree. (Exercise) □

Corollary. f_1, f_2 polynomials of degree $< n$ such that $f_1(t_i) = f_2(t_i)$, $(t_i)_{1 \leq i \leq n}$ n distinct values. Then $f_1 \equiv f_2$.

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

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

So $f \in \mathbb{C}[t]$,

$$f(t) = c \prod_{i=1}^{r} (t - \lambda_i)_i^{\alpha} \quad c, \lambda_i \in \mathbb{C}, \alpha_i \in \mathbb{N}$$

$\to$ complex analysis.

Definition (characteristic polynomial)**.** Let $\alpha \in \mathcal{L}(V)$, the characteristic polynomial of α is

$$\chi_\alpha(t) = \det(A - t\mathrm{id})$$

Remark. The fact that $\det(A - \lambda \mathrm{id})$ is a polynomial in λ follows from the very definition of det.

Remark. Conjugate matrices have the same characteristic polynomial.

$$\begin{aligned}\det(P^{-1}AP - \lambda \mathrm{id}) &= \det(P^{-1}(A - \lambda \mathrm{id})P) \\ &= \det(A - \lambda \mathrm{id})\end{aligned}$$

So we can define

$$\chi_\alpha(t) = \det(A - \lambda \mathrm{id})$$

where $A = [\alpha]_{\mathcal{B}}$, and the polynomial does not depend on the choice of basis.

Theorem. $\alpha \in \mathcal{L}(V)$ is triangulable if and only χ_α can be written as a product of linear factors over F:

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

$\to$ If $F = \mathbb{C}$, any matrix is triangulable.

Proof. $\Rightarrow$ Suppose α triangulable, then

$$[\alpha]_{\mathcal{B}} = \begin{pmatrix} a_1 & * & \cdots & * \\ 0 & a_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}$$

so

$$\chi_\alpha(t) = \det \begin{pmatrix} a_1 & * & \cdots & * \\ 0 & a_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} = \prod_{i=1}^{n} (a_i - t)$$

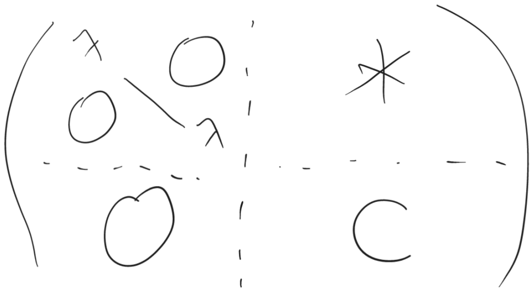
$\Leftarrow$ We argue by induction on $n = \dim V$.

- $n = 1$ easy.
- $n > 1$. By assumption, let $\chi_\alpha(t)$ which has a root λ. Then $\chi_\alpha(\lambda) = 0$ if and only if λ is an eigenvalue of α. Let $U = V_\lambda$ be associated eigenspace. Let $(v_1, \ldots, v_k)$ be a basis of U. We complete to $(v_{k+1}, \ldots, v_n)$ of V

$$\mathrm{span}(v_{k+1}, \ldots, v_n) = W$$

$$V = U \oplus W$$

- $[\alpha]_{\mathcal{B}} =$

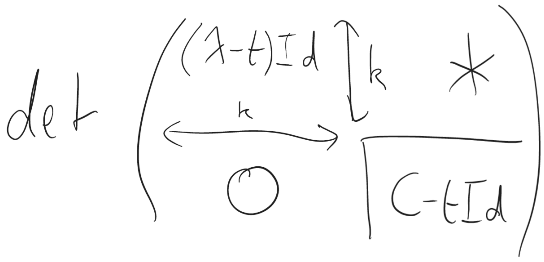


α induces an endormorphism $\overline{\alpha} : V/U \to V/U$,

$$C = [\overline{\alpha}]_{\overline{\mathcal{B}}}, \overline{\mathcal{B}} = (v_{k+1} + U, \ldots, v_n + U)$$

Then: (block product)

$$\det(\alpha - \mathrm{id}) =$$



$$= (\lambda - t)^k \det(C - t\mathrm{id}) = c \prod_{i=1}^{n} (t - a_i)$$

$$\implies \det(C - t\mathrm{id}) = \tilde{c} \prod_{i=k+1}^{n} (t - \tilde{a}_i)$$

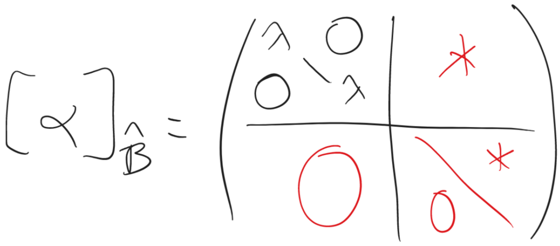
so use the induction because $\dim V/U = \dim V - \dim U < \dim V$. So $\tilde{\mathcal{B}} = (\tilde{v}_{k+1}, \ldots, \tilde{v}_n)$ basis of W where:

$$[C]_W = \begin{pmatrix} \tilde{a}_1 & * & \cdots & * \\ 0 & \tilde{a}_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{a}_n \end{pmatrix}$$

implies $V = U \oplus W$,

$$\hat{\mathcal{B}} = (v_1, \ldots, v_k, \tilde{v}_{k+1}, \ldots, \tilde{v}_n)$$

basis of V in which



$\rightarrow$ triangular form.

□

Start of lecture 15

Lemma. V n dimensional over $F = \mathbb{R}, \mathbb{C}$, $\alpha \in \mathcal{L}(V)$. Then $\chi_\alpha(t) = (-1)^n t^n + c_{n-1}t^{n-1} + \cdots + c_0$, $c_0 = \det A = \det \alpha$, $c_{n-1} = (-1)^{n-1} \operatorname{Tr} A$.

Proof.

- $\chi_\alpha(t) = \det(\alpha - t\mathrm{id})$

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

- Say that $F = \mathbb{R}$ or $\mathbb{C}$ (if $F = \mathbb{R}$, we can think of is as having complex entries as well). We know that α is triangulable over $\mathbb{C}$, so:

$$\begin{aligned} \chi_\alpha(t) &= \det \begin{pmatrix} a_1 - t & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_n - t \end{pmatrix} \\ &= \prod_{i=1}^n)a_i - 1) \\ &= (-1)^n t^n + c_{n-1} t^{n-1} + \cdots + c_0 \end{aligned}$$

$$c_{n-1} = (-1)^{n-1} \sum i = 1^n a_i = \operatorname{Tr} \alpha$$

□

1.17 Diagonalisation criterion and minimal polynomial

Notation (polynomial of an endomorphism)**.** Pick $p(t)$ polynomial over F

$$p(t) = a_n t^n + \cdots + a_1 t + a_0, \quad a_i \in F$$

$A \in \mathcal{M}_n(F)$, for all n, $A^n \in \mathcal{M}_n(F)$. We define:

$$p(A) = a_n A^n + \cdots + a_1 A + a_i \text{id} \in \mathcal{M}_n(F)$$

If $\alpha \in \mathcal{L}(V)$, we define

$$p(\alpha) = a_n \alpha^n + \cdots + a_1 \alpha + a_0 \text{id}$$

where $\alpha = \alpha \circ \cdots \circ \alpha \in \mathcal{L}(V)$.

$\to$ very useful.

Theorem (Sharp criterion of diagonalisability)**.**

- V vector space over F, $\dim_F V < \infty$
- $\alpha \in \mathcal{L}(V)$

Then α is diagonalisable if and only if there exists a polynomial p which is the product of *distinct linear factors* such that $p(\alpha) = 0$.

$$\alpha \text{ diagonalisable } \iff \exists(\lambda_1, \dots, \lambda_n) \text{ distinct }, \lambda_i \in F \text{ such that:}$$
$$p(t) = \prod_{i=1}^{k} (t - \lambda_i)$$
$$\text{and } p(\alpha) = 0$$

Proof. $\Rightarrow$ Suppose α is diagonalisable, with $\lambda_1, \dots, \lambda_k$ the distinct eigenvalues. Let $p(t) = \prod_{i=1}^k (t - \lambda_i)$. Let $\mathcal{B}$ be the basis of V made of eigenvectors of α (it is precisely the basis in which $[\alpha]_{\mathcal{B}}$ is diagonal). Then $v \in \mathcal{B}$, then $\alpha(v) = \lambda_i(v)$ for some $i \in \{1, \dots, k\}$, implies $(a - \lambda_i \text{id})(v) = 0$, implies

$$p(\alpha) = \left[\prod_{j=1}^{k} (\alpha - \lambda_j \text{id}) \right] (v) = 0$$

but the terms in the product commute, i.e.

$$(\alpha - \lambda_j \text{id})(\alpha - \lambda_k \text{id}) = (\alpha - \lambda_k \text{id})(\alpha - \lambda_j \text{id})$$

so for all $v \in \mathcal{B}$, $p(\alpha)(v) = 0$, so $p(\alpha)(v) = 0$ for all $v \in V$ (since $\mathcal{B}$ is a basis for V). So $p(\alpha) = 0$.

$\Leftarrow$ (Kernel lemma, Bezout's theorem for prime polynomials)

- Suppose $p(\alpha) = 0$ for:
$$p(t) = \prod_{i=1}^{k}(t - \lambda_i)$$
$\lambda_i \neq \lambda_j$, $i \neq j$.
- Let $V_{\lambda_i} = \ker(\alpha - \lambda_i \mathrm{id})$, we claim:
$$V = \bigoplus_{i=1}^{k} V_{\lambda_i}$$
Indeed let:
$$q_j(t) = \prod_{\substack{i=1\\ i\neq j}}^{k} \left(\frac{t - \lambda_i}{\lambda_j - \lambda_i}\right), \quad 1 \leq j \leq k$$
Then
$$q_j(\lambda_i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$
Hence let us consider:
$$q(t) = \sum_{j=1}^{k} q_j(t)$$
Then $\deg q_j \leq k - 1$, so $\deg q \leq k - 1$. Also $q(\lambda_j) = 1$ for all $1 \leq j \leq k$. So the polynomial $q(t) - 1$ has degree $\leq k - 1$ and at least k roots, so for all t, $q(t) = 1$. So for all t,
$$q_1(t) + \cdots + q_k(t) = 1$$
- Let us define the *projector*
$$\imath_j q_j(\alpha) \in \mathcal{L}(V)$$
Then
$$\begin{aligned} \sum_{j=1}^{k} \pi_j &= \sum_{j=1}^{k} q_j(\alpha) \\ &= \left(\sum_{j=1}^{k} q_j\right)(\alpha) \\ &= \mathrm{id} \end{aligned}$$
This means for all $v \in V$,
$$v = q(\alpha)(v) = \sum_{j=1}^{k} \pi_j(v) = \sum_{j=1}^{k} q_j(\alpha)(v)$$

Observe: pick $j \in \{1 \ldots, k\}$,

$$(\alpha - \lambda_j \mathrm{id}) q_j(\alpha)(v) = \frac{1}{\prod_{i \neq j}(\lambda_i - \lambda_j)} p(\alpha)(v) = 0$$

so

$$\begin{aligned} &\forall j \in \{1, \ldots, k\}, \quad (\alpha - \lambda_j \mathrm{id})\pi_j(v) = 0 \\ &\implies \forall j \in \{1, \ldots, k\} \pi_j(v) \in V_{\lambda_j} \end{aligned}$$

(π_j is a projector on V_{λ_j}) Now for all $v \in V$,

$$v = \sum_{j=1}^{k} \pi_j(v)$$

hence

$$V = \sum_{j=1}^{k} V_{\lambda_j}$$

We need to prove that the sum is *direct*. Indeed, let $v \in V_{\lambda_j} \cap \left(\sum_{i \neq j} V_{\lambda_i} \right)$.

- $v \in V_{\lambda_j}$. Then

$$\begin{aligned} \pi_j(v) &= q_j(\alpha)(v) \\ &= \prod_{i=1, i \neq j}^{k} \frac{(\alpha - \lambda_i \mathrm{id})}{\lambda_i - \lambda_j}(v) \\ &= \left[\prod_{i=1, i \neq j}^{k} \frac{(\lambda_i - \lambda_j)}{\lambda_i - \lambda_j} \right](v) \\ &= v \end{aligned}$$

  so $\pi_j \mid_{V_{\lambda_j}} = \mathrm{id}$.

- By assumption $v \in \sum_{i \neq j} V_{\lambda_i}$. Now, $i_0 \neq j$, $v \in V_{\lambda_{i_0}}$, $\alpha(v) = \lambda_{i_0} v$,

$$\begin{aligned} \implies \pi_j(v) &= q_j(\alpha)(v) \\ &= \left[\prod_{i \neq j} \frac{(\alpha - \lambda_i \mathrm{id})}{\lambda_i - \lambda_j} \right](v) \\ &= \left(\prod_{i \neq j} \frac{\lambda_{i_0} - \lambda_j}{\lambda_i - \lambda_j} \right) \\ &= 0 \end{aligned}$$

  so $\pi_j \mid_{V_{\lambda_i}} = 0$ for $i \neq j$.

As a conclusion: $v \in V_{\lambda_j} \cap \left(\sum_{i \neq j} V_{\lambda_i}\right)$

(1) $v \in V_{\lambda_i}$, implies $\pi_j(v) = v$

(2) $v \in \sum_{i \neq j} V_{\lambda_i}$ implies $\pi_j(v) = 0$

so $v = 0$. We have proved:

$$V = \bigoplus_{j=1}^{k} V_{\lambda_j}$$

$$\pi_j \mid_{V_{\lambda_j}} = \mathrm{id}$$

$$\pi_i \mid_{V_{\lambda_j}} = 0$$

for $i \neq j$.

Remark. We have proved the following: if $\lambda_1, \dots, \lambda_k$ are k distinct eigenvalues of α, then

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

(always true) (and we know the projectors)

This means that the only way diagonalisation fails is if:

$$\bigoplus_{i=1}^{k} V_{\lambda_i} = \sum_{j=1}^{k} V_{\lambda_j} \neq V$$

□

Example. $A \in \mathcal{M}_n(F)$, $F = \mathbb{C}$. A has finite order. (there exists $m \in \mathbb{N}$ such that $A^m = \mathrm{id}$). Then A is diagonalisable (over $\mathbb{C}$). TODO..

Start of lecture 16

Theorem (Simultaneous diagonalisation).

- $\dim_F V < \infty$
- $\alpha, \beta \in \mathcal{L}(V)$ diagonalisable

Then α, β are simultaneously diagonalisable ($\exists \mathcal{B}$ basis of V in which both $[\alpha]_{\mathcal{B}}$, $[\beta]_{\mathcal{B}}$ are diagonal) if and only if α and β commute.

Proof. $\Rightarrow$ Exists basis $\mathcal{B}$ of V in which

$$[\alpha]_{\mathcal{B}} = D_1$$

$$[\beta]_{\mathcal{B}} = D_2$$

D_1, D_2 both diagonal, then $D_1 D_2 = D_2 D_1$ so $\alpha\beta = \beta\alpha$.

$\Leftarrow$ Suppose α, β are both diagonalisable *and* $\alpha\beta = \beta\alpha$. Let $\lambda_1, \dots, \lambda_k$ be the k distinct eigenvalues of α. We have shown:

$$\alpha \text{ diagonalisable} \iff V = \bigoplus_{i=1}^{k} V_{\lambda_i}$$

V_{λ_i} is the eigenspace associated to λ_i.

Claim: V_{λ_I} stable by β: $\beta(V_{\lambda_i}) \le V_{\lambda_i}$.
Indeed, let $v \in V_{\lambda_i}$, then

$$\alpha\beta(v) = \eta\alpha(v) = \beta(\lambda_i v) = \lambda_i \beta(v)$$

so $\beta(v) \in V_{\lambda_i}$.

- We use criterion for diagonalisability: β is diagonalisable implies that there exists p with distinct linear factors such that $p(\beta) = 0$.

Now $B|_{V_{\lambda_j}}$ endomorphism ($\beta : V_{\lambda_j} \to V_{\lambda_j}$) and

$$p(\beta|_{V_{\lambda_j}}) = 0$$

p has distinct linear factors, so $\beta|V_{\lambda_j}$ is diagonalisable. So there exists $\mathcal{B}$ basis of V_{λ_j} in which $\beta|_{V_{\lambda_j}}$ is diagonal. Then

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

so $(\mathcal{B}_1, \dots, \mathcal{B}_k) = \mathcal{B}$ is a basis of V in which both α and β are in diagonal form.

$\square$

Minimal polynomial of an endormorphism

- Remainder: (Groups, Rings and Modules).
  Euclidean algorithm for polynomials: a, b polynomials over F, $b \neq 0$, then there exist polynomials q, r over F with:

$$\deg r < \deg b$$

$$a = qb + r$$

Definition (Minimal polynomial)**.** V vector space over F, $\dim_F V < \infty$. Let $\alpha \in \mathcal{L}(V)$. The *minimal polynomial* m_α of α is the (unique up to a constant) non zero polynomial with *smallest degree* such that

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

Existence and uniqueness follow from the following observations:

- $\dim_F V = n$, $\alpha \in \mathcal{L}(V)$. We know:

$$\dim_F \mathcal{L}(V) = n^2$$

$$\implies \text{id}, \alpha, \dots, \alpha^{n^2} \text{ cannot be free}$$

$$\implies a_{n^2}\alpha^{n^2} + \cdots + a_1\alpha + a_0 = 0$$

$$\implies \exists p \in F[t] \mid p(\alpha) = 0, p \neq 0$$

  That is, there does exist a polynomial p that kills α.

- Lemma: $\alpha \in \mathcal{L}(V)$, $p \in F[t]$. Then $p(\alpha) = 0$ if and only if m_α is a factor of p.
  Proof: $p \in F[t]$, $p(\alpha) = 0$, m_α is minimum polynomial of α. So $\deg m_\alpha \leq \deg p$. By Euclidean division:

$$p = m_\alpha q + r$$

$$\deg r < \deg m_\alpha$$

  Then

$$p(\alpha) = 0 = m_\alpha q(\alpha) r(\alpha)$$

  so $r(\alpha) = 0$. If $r \neq 0$, then this would contradict the definition of m_α. So $r \equiv 0$. So $p = m_\alpha q$, that is, m_α divides p.

- If m_1, m_2 are both polynomial with smallest degree which kill α then by the above lemma, $m_1 \mid m_2$, $m_2 \mid m_1$ so $m_2 = cm_1$, $c \in F$. That is, the minimal polynomial is unique up to a constant.

Example. $V = \mathbb{R}^2$

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

- Let $p(t) = (t-1)^2$, then $p(A) = p(B) = 0$. So minimal polynomial is either $t - 1$ or $(t-1)^2$.
- Check: $m_A = t - 1$, $m_B = (t-1)^2$. So A is diagonalisable but B is not.

1.18 Cayley Hamiton Theorem and multiplicity of eiganvalues

Theorem (Cayley Hamilton)**.** Let V be an F vector space, $\dim_F V < \infty$. Let $\alpha \in \mathcal{L}(V)$ with characteristic polynomial $\chi_\alpha(t) = \det(\alpha - t\text{id})$. Then $\chi_\alpha(\alpha) = 0$.

Corollary. $m_\alpha \mid \chi_\alpha$.

Proof. $F = \mathbb{C}$ (general proof is in the notes). $\alpha \in \mathcal{L}(V)$, $n = \dim_\mathbb{C} V$. Exists basis $\mathcal{B} = \{v_1, \dots, v_n\}$ such that

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

(triangulable). Let $U_j = \langle v_1, \dots, v_j \rangle$. Then because of the triangular form, $(\alpha - a_j \mathrm{id}) U_j \le U_{j-1}$.

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

$$\begin{aligned} &(\alpha - a_1 \mathrm{id}) \cdots (\alpha - a_{n-1}\mathrm{id})(\alpha - a_n \mathrm{id}) V \\ &\le (\alpha - a_1 \mathrm{id}) \cdots (\alpha - a_{n-1}\mathrm{id}) U_{n-1} \\ &\vdots \\ &\le 0 \end{aligned}$$

So $\chi_\alpha(\alpha) = 0$. For the general case, see the notes. □

Definition (algebraic / geometric multiplicity)**.** $\dim_F V < \infty$, $\alpha \in \mathcal{L}(V)$. Let λ eigenvalue of α. Then

$$\chi_\alpha(t) = (t - \lambda)^{a_\lambda} q(t)$$

$$q \in F[t], \quad q(\lambda) \neq 0$$

- a_λ is the algebraic multiplicity of λ.
- g_λ is the geometric multiplicity of λ, and $g_\lambda = \dim \ker(\alpha - \lambda \mathrm{id})$.

Remark. λ eigenvalue $\iff$ $\alpha - \lambda$ $\iff$ singular $\iff$ $\det(\alpha - \lambda \mathrm{id}) = \chi_\alpha(\lambda) = 0$

Lemma. λ eigenvalue of $\alpha \in \mathcal{L}(V)$, then $1 \le g_\lambda \le a_\alpha$.

Proof.
- $g_\lambda = \dim \ker(\alpha - \lambda \mathrm{id}) \ge 1$ since λ is an eigenvalue.
- Let us show that $g_\lambda \le a_\lambda$. Indeed, let $(v_1, \dots, v_{g_\lambda})$ basis of $V_\lambda = \ker(\alpha - \lambda \mathrm{id})$, and compute $\mathcal{B} = (v_1, \dots, v_{\lambda_g}, v_{g_\lambda + 1}, \dots, v_n)$ of V. Then

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

$$\implies \det[\alpha - t\mathrm{id}] = \det \begin{pmatrix} (\lambda - t)\mathrm{id}_{g_\lambda} & * \\ 0 & A_1 - t\mathrm{id} \end{pmatrix} = (\lambda - t)^{g_\lambda} \chi_{A_1}(t)$$
$$\implies g_\lambda \le a_\lambda$$

□

Lemma. λ eigenvalue of $\alpha \in \mathcal{L}(V)$. Let:

$$c_\lambda \equiv \text{multiplicity of } \lambda \text{ as a root of } m_\alpha \text{ (minimal polynomial)}$$

Then $1 \le c_\lambda \le a_\lambda$.

Proof.
- Cayley-Hamilton implies $m_\alpha \mid \chi_\alpha$. So $c_\lambda \le a_\lambda$.
- $c_\lambda \ge 1$. Indeed, there exists $b \neq 0$ such that $\alpha(v) = \lambda v$ so then for all $p \in F[t]$, $p(\alpha)(v) = (p(\lambda))v$ ($\alpha^n(v) = \lambda^n v$) so $m(\alpha)(v) = (m(\lambda))v$ so $m(\lambda) = 0$ so $c_\lambda \ge 1$.

□

Example.

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

m_A?

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

- So m_α is either $(t-1)^2(t-2)$ or $(t-1)(t-2)$. Check $(A-I)(A-2I) = 0$, so $m_\alpha = (t-1)(t-2)$ so A is diagonalisable.

Start of lecture 17

Example (Jordan block)**.**

$$J_\lambda = \begin{pmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & \lambda \end{pmatrix} \in \mathcal{M}_n(F)$$

Check $g_\lambda = 1$, $a_\lambda = n$, $c_\lambda = n$

Lemma (characterisation of diagonalisable endomorphisms over $F = \mathbb{C}$)**.** $F = \mathbb{C}$, $\dim_{\mathbb{C}} V = n < \infty$, $\alpha \in \mathcal{L}(V)$. The following are equivalent:

(i) α diagonalisable

(ii) $\forall \lambda$ eigenvalue of α, $a_\lambda = g_\lambda$

(iii) $\forall \lambda$ eigenvalue of α, $c_\lambda = 1$.

Proof. (i) $\iff$ (iii) done. We need (i) $\iff$ (ii). Indeed, let $(\lambda_1, \dots, \lambda_k)$ be the *distinct* eigenvalues of α. We showed:

$$\alpha \text{ diagonalisable} \iff V = \bigoplus_{i=1}^{k} V_{\lambda_i}$$

$$\dim V = n \qquad\qquad = \deg \chi_\alpha$$

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

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

so

$$\alpha \text{ diagonalisable} \iff \sum_{i=1}^{k} a_{\lambda_i} = \sum_{i=1}^{k} g_{\lambda_i} \tag{$*$}$$

We know: $\forall 1 \le i \le k$, $g_{\lambda_i} \le a_{\lambda_i}$. Hence $(*)$ holds $\iff$ $\forall 1 \le i \le k$, $a_{\lambda_i} = g_{\lambda_i}$. □

1.19 Jordan normal form

Note. In this subsection, $F = \mathbb{C}$.

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

$$A = \begin{pmatrix} J_{n_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{n_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{n_k}(\lambda_k) \end{pmatrix}$$

where:

- $k \geq 1$, k integer
- $n_1, \dots, n_k$ integers
- $\sum_{i=1}^k n_i = n$
- $\lambda_i \in \mathbb{C}$, $1 \leq i \leq k$: they need *not* be distinct
- $m \in \mathbb{N}$, $m \neq 0$, $\lambda \in \mathbb{C}$, $J_m(\lambda) = \lambda)$ if $m = 1$,

Example. $n = 3$,

$$A = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} J_1(\lambda & 0 & 0 \\ 0 & J_1(\lambda) & 0 \\ 0 & 0 & J_1(\lambda) \end{pmatrix}$$

so this is in Jordan Normal Form.

Theorem. Every matrix $A \in \mathcal{M}_n(\mathbb{C})$ is similar to a matrix in Jordan Normal Form, which is unique up to reordering the Jordan block.

Proof. Non examinable (in Groups, Rings and Modules class). (Proof is in lecturer's notes). □

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

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad m = (t-\lambda_1)(t-\lambda_2), \quad \lambda_1 \neq \lambda_2$$

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \quad m = (t-\lambda)$$

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \quad m = (t-\lambda)^2$$

Example. $n = 3$,

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \quad (t-\lambda_1)(t-\lambda_2)(t-\lambda_3) \quad \lambda_1, \lambda_2, \lambda_3 \text{ distinct}$$

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} \quad (t-\lambda_1)(t-\lambda_2)$$

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix} \quad (t-\lambda_1)(t-\lambda_2)^2$$

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \quad (t-\lambda)$$

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \quad (t-\lambda)^2$$

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \quad (t-\lambda)^3$$

Useful observation: which explains why JNF is unique. $\rightarrow$ we can directly compute in the JNF the quantities $a_\lambda, g_\lambda, c_\lambda$. Indeed, let $M \geq 2$ and let $J_m(\lambda)$. Then

$$J_m - \lambda Id = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$

$$(J_m - \lambda Id)^2 = \begin{pmatrix} 0 & 0 & 1 & & 0 \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & \ddots & 0 \\ 0 & & & & 0 \end{pmatrix}$$

By induction we can show:

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

for $k \leq m$, and is 0 for $k = m$. We say that the matrix $(J_m - \lambda \mathrm{id})$ is *nilpotent* of order m. ($u^m = 0$ and $u^{m-1} \neq 0$). So

$$a_\lambda \equiv \text{sum of sizes of blocks with eigenvalue } \lambda \equiv \text{number of } \lambda \text{ on the diagonal}$$

$$g_\lambda = \dim \ker(A - \lambda \mathrm{id}) = \text{number of blocks with eigenvalue } \lambda$$

$$c_\lambda J_m(\lambda) \to (t - \lambda)^m \text{kills it}$$

(because $(J_m - \lambda \mathrm{id})$ is nilpotent of order exactly m) so

$$c_\lambda \equiv \text{size of the largest block with eigenvalue } \lambda$$

Example.

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

Find a basis in which A is Jordan Normal Form?

$\chi_A(t) = (t-1)^2$ eigenvalue $\lambda = 1$. $A - \mathrm{id} \neq 0$ implies $m_A(t) = (t-1)^2$, and Jordan Normal Form

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

(ii) Eigenvectors:

$$A - \mathrm{id} = \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}$$

$\ker(A - \mathrm{id}) = \langle v_1 \rangle$, $v_1 = (1, -1)^\top$. I look for a (non-unique!) v_2 such that

$$(A - \mathrm{id})v_2 = v_1$$

$v_2 = (-1, 0)^\top$ works.

$$[A]_{\mathcal{B}} = J_1(1)$$

for $\mathcal{B} = (v_1, v_2)$.

$$P^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}$$

$$A = \underbrace{\begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}}_{P^{-1}} \underbrace{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}}_{J} \underbrace{\begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}^{-1}}_{P}$$

Exercise:

$$A = \begin{pmatrix} 3 & -2 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

Find a basis in which A is JNF.

Theorem (Generalised eigenspace decomposition).

- $\dim_{\mathbb{C}} V = n < \infty$
- $\alpha \in \mathcal{L}(V)$.
- $m_\alpha(t) = (t-\lambda_1)^{c_1} \cdots (t-\lambda_k)^{c_k}$
- $\lambda_1, \ldots, \lambda_k$ distinct eigenvalues of α.

Then

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

$$V_j = \ker[(\alpha - \lambda \mathrm{id})^{c_j}]$$

(V_j is generalised eigenspace)

Remark. α diagonalisable, $c_j = 1$. THen V_j eigenspace associated to λ_j.

Proof. projectors onto V_j are *explicit*. Indeed, let

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

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

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

We define

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

(i) By $(*)$,

$$\mathrm{id} = \sum_{j=1}^{k} q_j p_j(\alpha) = \sum_{j=1}^{k} \pi_j$$

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

(ii) $m_\alpha(\alpha) = 0$, $m_\alpha = \prod_{j=1}^{k} (t - \lambda_j)^{c_j}$

$$\implies (\alpha - \lambda_j \mathrm{id})^{c_j} \pi_j = (\alpha - \lambda_j \mathrm{id})^{c_j} q_j p_j(\alpha) = 0$$

$$\implies \forall v \in V, \pi(v) \in V_j$$

$$V_j = \ker(\alpha - \lambda_j \mathrm{id})^{c_j}$$

Hence $\forall v \in V$

$$v = \sum_{j=1}^{k} \pi_j(v)$$

$$\implies V = \sum_{j=1}^{k} V_{\lambda_j}$$

(iii) Show that:

$$\sum_{j=1}^{k} V_{\lambda_j} = \bigoplus_{j=1}^{k} V_{\lambda_j}$$

Indeed, $\pi_i \pi_j = 0$ if $i \neq j$ and so $\pi_i = \pi_i \left(\sum_{j=1}^{k} \pi_j\right) = \pi_i^2$.

$$\implies \pi_i|_{V_{\lambda_i}} = \text{id}$$

$\implies$ direct sum projection follows:

$$v = V_{\lambda_i} \cap \left(\sum_{i \neq j} V_{\lambda_i}\right)$$

$$v = \sum_{i \neq j} v_j, \quad v_j \in V_{\lambda_j}$$

If apply π_i and use:

$$\pi_i|_{V_{\lambda_i}} = \text{id}$$
$$\pi_i|_{V_{\lambda_j}} = 0 \text{ for } j \neq i$$

so $v = 0$.

□

$$V = \bigoplus_{i=1}^{k} V_{\lambda_i}$$
$$V_{\lambda_i} = \ker(\alpha - \lambda_i \text{id})^{c_{\lambda_i}}$$

By definition

$$(\alpha - \lambda_i \text{id})|_{V_{\lambda_i}}$$

is nilpotent, since

$$(\alpha - \lambda_i \text{id})^{c_{\lambda_i}}|_{V_{\lambda_i}} = 0$$

$\implies$ all I need to do is to find JNF nilpotent endomorphism.

- $\alpha \in \mathcal{L}(V)$, $\dim_{\mathbb{C}} V = n$, $a^k = 0$, $a^{k-1} \neq 0$. $\overset{?}{\implies}$ JNF with blocks $J_m(0)$. $\to$ By induction on the dimension.

$$\alpha^k = 0, \alpha^{k-1} \neq 0$$

$$\implies \exists x \in V, (x, \alpha(x), \dots, \alpha^{k-1}(x))$$

free.

Question: $F = \operatorname{span}\langle x, \alpha(x), \dots, \alpha^{k-1}(x)\rangle$. Can I find G such that:

$$V = F \oplus G$$

G stable by α?
$\to$ done.

Start of lecture 18

1.20 Bilinear Forms

Bilinear form: $\varphi : V \times V \to F$.

- $\dim_F V < \infty$, $\mathcal{B}$ basis of V.
- $[\varphi]_{\mathcal{B}} = [\varphi]_{\mathcal{B},\mathcal{B}} = (\varphi(e_i, e_j))_{1 \leq i,j \leq n}$. $\mathcal{B} = (e_i)_{1 \leq i \leq n}$.

Lemma. $\varphi : V \times V \to F$ bilinear, $\mathcal{B}, \mathcal{B}'$ two basis of V, $P = [\mathrm{id}]_{\mathcal{B}',\mathcal{B}}$ then

$$[\varphi]_{\mathcal{B}'} = P^\top [\varphi]_{\mathcal{B}} P$$

Proof. Special case of the general formula $\to$ Lecture 10. □

Definition (Congruent matrices)**.** $A, B \in \mathcal{M}_n(F)$, we say that A and B are *congruent* if and only if there exists P invertible such that

$$A = P^\top B P$$

Remark. This defines an equivalence relation.

Definition (Symmetric)**.** A bilinear form φ on V is *symmetric* if:

$$\varphi(u, v) = \varphi(v, u) \quad \forall (u, v) \subset V \times V$$

Remark.

- $A \in \mathcal{M}_n(F)$, we say that A is symmetric if and only if $A = A^\top$

$$\iff A = (a_{ij})_{1 \le i,j \le n}, a_{ij} = a_{ji}$$

- φ symmetric $\iff$ $[\varphi]_{\mathcal{B}}$ is symmetric in *any* basis $\mathcal{B}$ of V.

Remark. To be able to represent φ by a diagonal matrix, then φ must be symmetric

$$P^\top AP = D \implies D^\top P^\top A^\top P$$

$$\implies A = A^\top$$

Definition (Quadratic form)**.** A map $Q : V \to F$ is a *quadratic form* if and only if there exists a bilinear form $\varphi : V \times V \to F$ such that $\forall u \in V$,

$$Q(u) = \varphi(u, u)$$

Remark. $\mathcal{B} = (e_i)_{1 \le i \le n}$, $A = [\varphi]_{\mathcal{B}} = (\underbrace{\varphi(e_i, e_j)}_{a_{ij}})_{1 \le i,j \le n}$. Then

$$u = \sum_{i=1}^{n} x_i e_i, x = (x_1, \dots, x_n)^\top$$

Then

$$\begin{aligned} Q(u) &= \varphi(u, u) \\ &= \varphi\left(\sum_{i=1}^{n} x_i e_i, \sum_{i=1}^{n} x_i e_i\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \underbrace{\varphi(e_i, e_j)}_{a_{ij}} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j a_{ij} \\ &= x^\top A x \end{aligned}$$

so

$$\boxed{Q(u) = x^\top A x}$$

Observation:

$$\begin{aligned} x^\top A x &= \sum_{i,j=1}^n a_{ij} x_i x_j \\ &= \sum_{i,j=1}^n a_{ji} x_i x_j \\ &= \frac{1}{2} \sum_{i,j=1}^n (a_{ij} + a_{ji}) x_i x_j \\ &= \frac{1}{2} x^\top (A + A^\top) x \end{aligned}$$

and $\frac{1}{2}(A + A^\top)$ is symmetric.

Proposition. 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:

$$\forall u \in V, Q(u) = \varphi(u, u)$$

Proof. Let ψ bilinear form on V such that

$$\forall u, Q(u) = \psi(u, u)$$

Let

$$\varphi(u, v) = \frac{1}{2}(\psi(u, v) + \psi(v, u))$$

- φ symmetric
- $\varphi(u, u) = \psi(u, u) = Q(u)$.

$\to$ existence of φ symmetric. Now uniqueness. Let φ be a symmetric bilinear form such that $\varphi(u, u) = Q(u) \forall u \in V$. Then

$$\begin{aligned} Q(u + v) &= \varphi(u + v, u + v) \\ &= \varphi(u, u) + \varphi(v, u) + \varphi(v, u) + \varphi(v, v) \\ &= Q(u) + 2\varphi(u, v) + Q(v) \end{aligned}$$

so

$$\boxed{\varphi(u, v) = \frac{1}{2}[Q(u + v) - Q(u) - Q(v)]}$$

$\equiv$ POLARIZATION IDENTITY. $\square$

Theorem (Diagonalisation of symmetric bilinear forms)**.** Let $\varphi : V \times V \to F$ be a symmetric bilinear form ($\dim_F V < \infty$). Then there exists a basis $\mathcal{B}$ of V such that:

$$[\varphi]_{\mathcal{B}} \text{ is diagonal}$$

$\implies$ extensions to infinite dimensional cases.

Proof. • $\dim_F V < \infty$.

- Induction on the dimension n.
- $n = 0, 1$ are clear.
- Suppose that the theorem holds for all dimensions $< n$.
  (1) Let $\varphi : V \times V \to F$ be a symmetric bilinear form. If $\varphi(u,u) = 0$, $\forall u \in V$, φ is identically zero. (polarization identity).

  $$\implies \exists u \in V \setminus \{0\} \mid \varphi(u,u) \neq 0$$

  (because $\varphi \neq 0$).

  (2) Let us call $u = e_1$. ($e_1 \neq 0$, $\varpi(e_1, e_1) \neq 0$). Let us define

  $$\begin{aligned} U &= (\langle e_1 \rangle)^{\perp} \\ &= \{v \in V \mid \varphi(e_1, v) = 0\} \\ &= \ker\{\varphi(e_1, \bullet) : V \to F, v \mapsto \varphi(e_1, v)\} \end{aligned}$$

  (linear because φ is bilinear). Now rank nullity:

  $$\dim V = n = 1 + \dim U$$

  ($r(\varphi(e_1, \bullet)) = 1$ $\varphi(e_1, e_1) \neq 0$) So $\dim U = n - 1$.

  (3) Claim $U + \langle e_1 \rangle = U \oplus \langle e_1 \rangle$. Indeed, $v \in \langle e_1 \rangle \cap U$ then $v = \lambda e_1$, $\lambda \in F$.

  $$\varphi(e_1, v) = 0 \quad (v \in U)$$

  so $0 = \varphi(e_1, \lambda e_1) = \lambda \varphi(e_1, e_1)$ so $\lambda = 0$, so $v = 0$.

  (4) Conclusion $V = \langle e_1 \rangle \oplus U$, by counting dimensions.

  (5) Complete $(e_2, \ldots, e_n)$ basis of U. So $\mathcal{B} = (e_1, e_2, \ldots, e_n)$ basis of V. And:

  $$[\varphi]_{\mathcal{B}} = \left(\begin{array}{c|c} \varphi(e_1, e_1) & 0 \longleftrightarrow 0 \\ \hline 0 & \\ \updownarrow & A' \\ 0 & \end{array}\right)$$

$(\varphi(e_j, e_1) = \varphi(e_1, e_j) = 0$ for $2 \leq j \leq n)$.

$$A' = (\varphi(e_i, e_k))_{2 \leq i,j \leq n}$$

Then $(A')^\top = A'$ since φ symmetric.

$$\implies \varphi|_U : U \times U \to F$$

bilinear symmetric with matrix A'. By the induction hypotheses, I can find $\mathcal{B}' = (e_2', \dots, e_n')$ basis of U in which $[\varphi|_U]_{\mathcal{B}'}$ is diagonal. Then

$$[\varphi]_{(e_1, e_2', \dots, e_n')}$$

diagonal form. □

Remark. $\varphi(e_1, e_1) \neq 0$

$$\implies V = \langle e_1 \rangle \oplus U$$

$$U = \langle e_1 \rangle^\perp$$

Example. $V = \mathbb{R}^3$

- $Q(x_1, x_2, x_3) = x_1^2 + x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 - 2x_2x_3 = x^\top Ax$ where

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

- Diagonalise: Two ways.
  (1) Follow the proof of diagonalisation $\to$ algorithm.
  (2) "Complete the square"

$$\begin{aligned} Q(x_1, x_2, x_3) &= x_1^2 + x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 - 2x_2x_3 \\ &= (x_1 + x_2 + x_3)^2 + x_3^2 - 4x_2x_3 \\ &= \underbrace{(x_1 + x_2 + x_3)}_{x_1'}{}^2 + \underbrace{(x_3 - 2x_2)}_{x_2'}{}^2 - \underbrace{(2x_2)}_{x_3'}{}^2 \end{aligned}$$

  – P,

$$P^\top AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

  – To find P, remember:

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 2 & 0 \end{pmatrix} = P^{-1}$$

Start of lecture 19

1.21 Sylvester's law / Sesquilinear forms

Recall:

Theorem. $\dim_F V < \infty$, $\varphi : V \times V \to F$ is a *symmetric* bilinear form $\implies$ there exists $\mathcal{B}$ basis of V in which $[\varphi]_\mathcal{B}$ is diagonal.

Remark. We take $F = \mathbb{R}$ or $\mathbb{C}$ in this subsection.

Corollary. $F = \mathbb{C}$, $\dim_{\mathbb{C}} V < \infty$, φ symmetric bilinear form on V. Then there exists basis of V such that

$$[\varphi]_{\mathcal{B}} = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right), \qquad r = \operatorname{rank}(\varphi)$$

Proof. Pick a basis $\mathcal{E} = (e_1, \dots, e_n)$ such that

$$[\varphi]_{\mathcal{E}} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}$$

Reorder e_i such that $a_i \neq 0$ for $1 \leq i \leq r$, $a_i = 0$ for $i > r$. For $i \leq r$, I let $\sqrt{a_i}$ be a choice of complex root of a_i, we define:

$$v_i = \begin{cases} \frac{e_i}{\sqrt{a_i}} & \text{for } 1 \leq i \leq r \\ e_i & \text{for } i < r \end{cases}$$

$\mathcal{B} = (v_1, \dots, v_r, e_{r+1}, \dots, e_n)$, $\mathcal{B}$ basis of V

$$\implies [\varphi]_{\mathcal{B}} = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

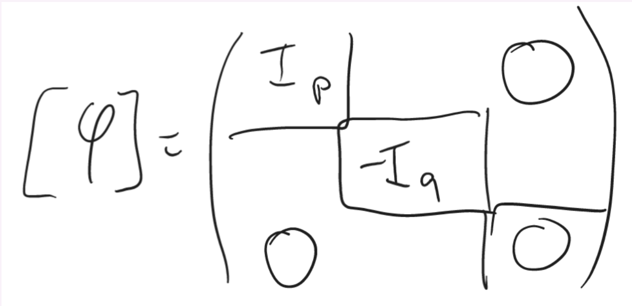
□

Corollary. Every *symmetric* matrix of $\mathcal{M}_n(\mathbb{C})$ is *congruent* to a *UNIQUE* matrix of the form:

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline & \\ 0 & 0 \end{array}\right)$$

We want to address the same problem with $F = \mathbb{R}$. → we cannot take complex roots this time.

Corollary. $F = \mathbb{R}$, $\dim_{\mathbb{R}} V < \infty$, φ *symmetric* bilinear form on V. Then there exists $\mathcal{B} = (v_1, \dots, v_n)$ basis of V such that



$p, q \geq 0$, $p + q = r(\varphi)$.

Proof. $\mathcal{E} = (e_1, \dots, e_n)$ basis of V such that

$$[\varphi]_{\mathcal{E}} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} \qquad a_i \in \mathbb{R}$$

Reorder a_i so that:

- $a_i > 0$ for $1 \leq i \leq p$
- $a_i < 0$ for $p + 1 \leq i \leq q$
- $a_i = 0$ for $i \geq q + 1$

We define:

$$v_i = \begin{cases} \frac{e_i}{\sqrt{a_i}} & 1 \leq i \leq p \\ \frac{e_i}{\sqrt{|a_i|}} & p + 1 \leq i \leq q \\ e_i & i \geq q + 1 \end{cases}$$

Then let $\mathcal{B} = (v_1, \dots, v_n)$ and then $[\varphi]_{\mathcal{B}}$ has the announced form. □

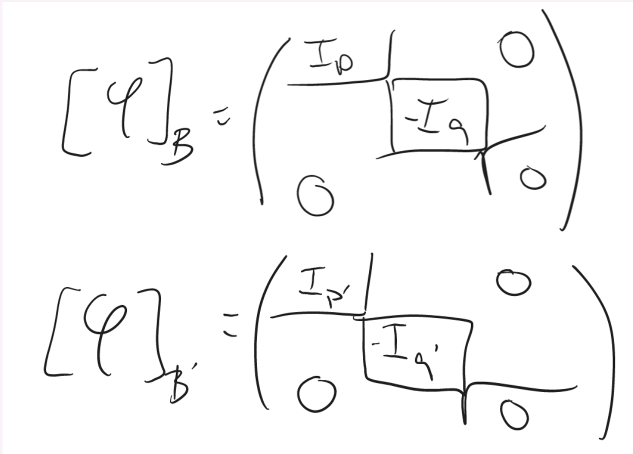
Definition (signature)**.** We define (under the assumptions above)

$$s(\varphi) = p - q \equiv \text{signature of } \varphi$$

(we also speak of the signature of the associated quadratic form $Q(u) = \varphi(u, u)$)

This definition makes sense:

Theorem (Sylvester's law of inertia)**.** $F = \mathbb{R}$, $\dim_{\mathbb{R}} V < \infty$, φ symmetric bilinear form on V. If φ is represented by:



with $\mathcal{B}, \mathcal{B}'$ bases of V. Then $p = p'$ and $q = q'$.

Definition. φ be a symmetric bilinear form on a real valued vector space ($F = \mathbb{R}$). We say that:

(i) φ is *positive definite*

$$\iff \forall u \in V \setminus \{0\}, \quad \varphi(u, u) > 0$$

(ii) φ is *positive semi definite*

$$\iff \forall u \in V, \quad \varphi(u, u) \geq 0$$

(iii) φ is *negative definite*

$$\iff \forall u \in V \setminus \{0\}, \quad \varphi(u, u) < 0$$

(iv) φ is negative semi definite

$$\iff \forall u \in V, \quad \varphi(u, u) \leq 0$$

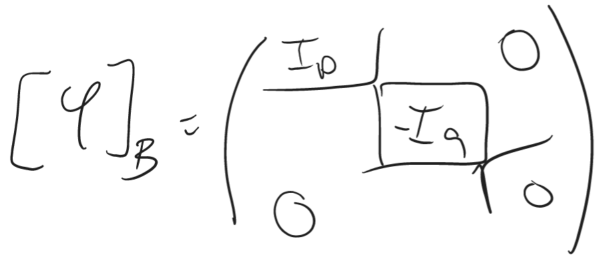
Example.

$$\left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & 0 \end{array}\right)$$

positive definite for $p = n$

- positive semi definite for $p \leq n$.

Proof. (Of Sylvester's law of inertia)
In order to prove that p is independent of the choice of the basis, we show that p has a geometric interpretation:
Claim: p is the largest dimension of subspace on which φ is positive definite.
Proof:
Say $\mathcal{B} = (v_1, \ldots, v_n)$ in which:

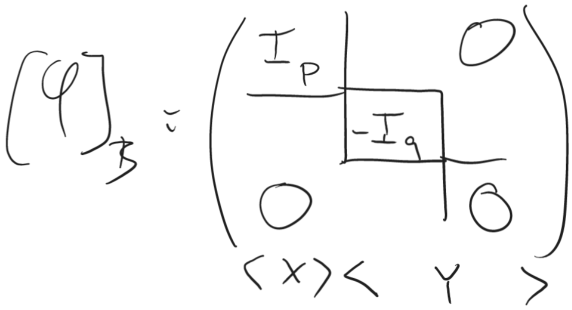


(1) Let $X = \langle v_1, \ldots, v_p \rangle$. Then φ is positive definite on X. Indeed, $u = \sum_{i=1}^{p} \lambda_i v_i$,

$$\begin{aligned} Q(u) &= \varphi(u, u) \\ &= \varphi\left(\sum_{i=1}^{p} \lambda_i v_i, \sum_{i=1}^{p} \lambda_i v_i\right) \\ &= \sum_{i,j=1}^{n} \lambda_i \lambda_j \varphi(v_i, v_j) \\ &= \sum_{i=1}^{p} \lambda_i^2 && > 0 && \text{for } u \neq 0 \end{aligned}$$

$\dim X = p$, $\varphi|_{X \times X}$ is positive definite.

(2) Suppose that φ is definite positive when restricted to another subspace X'. Let $X = \langle v_1, \ldots, v_p \rangle$, $Y = \langle v_{p+1}, \ldots, v_n \rangle$, $\mathcal{B} = (v_1, \ldots, v_n)$. Then



$\Longrightarrow$ We know that φ is negative semi definite on Y. So $Y \cap X' = \{0\}$. Indeed, if $u \in Y \cap X'$ and $u \neq 0$, then $u \in Y$ so $\varphi(u,u) \leq 0$, but $u \in X'$ so $\varphi(u,u) > 0$. So $Y \cap X' = \{0\}$. So $Y + X' = Y \oplus X'$, so $\dim Y + \dim X' \leq n$, and $\dim Y = n - p$ so $\dim X' \leq p$.

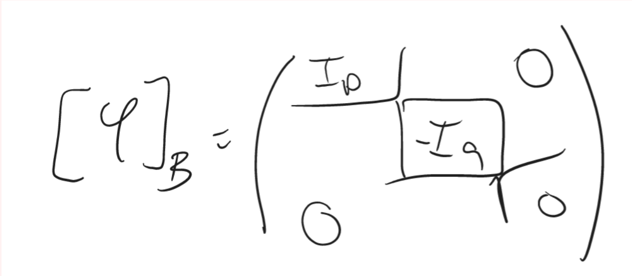
So now we know that p has a geometric interpretation / is unique. Then by considering $-\varphi$, we find that q is unique too. □

Remark. Similarly, q is the largest dimension of a subset on which φ is negative definite.

Definition. K = kernel of a bilinear form $\varphi = \{v \in V \mid \forall u \in V, \varphi(u,v) = 0\}$.

Remark. $\dim K + r(\varphi) = n$

Remark. $F = \mathbb{R}$. One notices that there is a subspace T of dimension $n - (p + q) + \min\{p, q\}$ such that $\varphi|_T = 0$. Indeed: $\mathcal{B} = (v_1, \dots, v_n)$,



$T = \langle \underbrace{v_1 + v_{p+1}, \dots, v_q + v_{p+q}}_{q}, \underbrace{v_{p+q+1}, \dots, v_n}_{n-(p+q)} \rangle$ (if $p \geq q$). Check $\varphi|_T = 0$ $(\forall (u, v) \in T \times T, \varphi(u, v) = 0)$. Moreover, one can show that this is the largest dimension of a subspace T' on which $\varphi|_{T' \times T'} = 0$

Sesquilinear Forms

- $F = \mathbb{C}$
- Standard *inner product* on $\mathbb{C}^n$ is $\langle x, y \rangle = \sum_{i=1}^n x_i \overline{y_i}$. In particular,

$$\|x\|^2 = \langle x, x \rangle = \underbrace{\sum_{i=1}^n |x_i|^2}_{\in \mathbb{R}^+}$$

Warning. $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$

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

is *not* a bilinear form: $\lambda \in \mathbb{C}$,

$$\langle \lambda x, y \rangle = \sum_{i=1}^n \lambda x_i \overline{y_i} = \lambda \langle x, y \rangle$$

$$\langle x, \lambda y \rangle = \sum_{i=1}^n x_i \overline{\lambda y_i} = \overline{\lambda} \langle x, y \rangle$$

$\to$ antilinear with respect to the second coordinate.

Definition. V, W $\mathbb{C}$ vector spaces. A sesquilinear form φ is a function $\varphi : V \times W \to \mathbb{C}$ such that:

(i) $\varphi(\lambda_1 v_1 + \lambda_2 v_2, w) = \lambda_1 \varphi(v_1, w) + \lambda_2 \varphi(v_2, w)$ (linear with respect to the first coordinate)

(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).

Start of lecture 20

$\dim_{\mathbb{C}} W < \infty$, $\dim_{\mathbb{C}} V < \infty$, φ sesquilinear, $V \times W \to \mathbb{C}$

- linear first variable: $\varphi(\lambda u, v) = \lambda \varphi(u, v)$
- multilinear second variable $\varphi(u, \lambda v) = \overline{\lambda} \varphi(u, v)$

Definition. $\mathcal{B} = (v_1, \dots, v_m)$ basis of V, $\mathcal{C} = (w_1, \dots, w_n)$ basis of W.

$$[\varphi]_{\mathcal{B},\mathcal{C}} = (\varphi(v_i, w_j))$$

$m \times n$ matrix.

Lemma. $\varphi(v, w) = [v]_{\mathcal{B}}^{\top} [\varphi]_{\mathcal{B},\mathcal{C}} \overline{[w]_{\mathcal{B}}}$

Proof. Exercise. □

Lemma. $\mathcal{B}$, $\mathcal{B}'$ basis for V, $P = [\mathrm{id}]_{\mathcal{B}',\mathcal{B}}$, $\mathcal{C}, \mathcal{C}'$ basis for W, $Q = [\mathrm{id}]_{\mathcal{C}',\mathcal{C}}$. Then

$$[\varphi]_{\mathcal{B}',\mathcal{C}'} = P^{\top} [\varphi]_{\mathcal{B},\mathcal{C}} \overline{Q}$$

Proof. Exercise. □

1.22 Hermitian Forms / $\mathbb{C}$, Skew Symmetric forms / $\mathbb{R}$

Hermitian form

$\dim_{\mathbb{C}} V < \infty$, $\varphi : V \times V \to \mathbb{C}$ sesquilinear ($W = V$).

Definition (Hermitian form)**.** A sesquilinear form $\varphi : V \times V \to \mathbb{C}$ is called *Hermitian* if

$$\forall (u, v) \in V \times V, \quad \varphi(u, v) = \overline{\varphi(v, u)}$$

Remark. φ Hermitian

$$\begin{aligned}&\implies \varphi(u,u) = \overline{\varphi(u,u)}\\&\implies \forall u \in V, \varphi(u,u) \in \mathbb{R}\end{aligned}$$

Allows us to speak of positive / negative (semi) definite Hermitian form.

Lemma. A sesquilinear form $\varphi : V \times V \to \mathbb{C}$ is Hermitian if and only if $\forall \mathcal{B}$ basis of V,

$$[\varphi]_{\mathcal{B}} = \overline{[\varphi]_{\mathcal{B}}^{\top}}$$

Proof. $A = [\varphi]_{\mathcal{B}} = (a_{ij})_{1\leq i,j\leq n}$, $a_{ij} = \varphi(e_i, e_j)$. Then $a_{ij} = \varphi(e_i,e_j)$, $a_{ji} = \varphi(e_j,e_i) = \overline{\varphi(e_i,e_j)} = \overline{a_{ij}}$.

$$\implies [\varphi]_{\mathcal{B}}^{\top} = \overline{[\varphi]_{\mathcal{B}}}$$

Conversely $[\varphi]_{\mathcal{B}} = A$, $A = \overline{A^{\top}}$

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

$$v = \sum_{i=1}^{n} \mu_i e_i$$

$\mathcal{B} = (e_1, \dots, e_n)$

$$\begin{aligned}\varphi(u,v) &= \varphi\left(\sum_{i=1}^{n} \lambda_i e_i, \sum_{i=1}^{n} \mu_i e_i\right)\\&= \sum_{i,j=1}^{n} \lambda_i \overline{\mu_j} \varphi(e_i, e_j)\\&= \sum_{i,j=1}^{n} \lambda_i \overline{\mu_j} a_{ij}\end{aligned}$$

Then

$$\begin{aligned}
\overline{\varphi(v,u)} &= \overline{\varphi\left(\sum_{i=1}^{n}\mu_i e_i, \sum_{i=1}^{n}\lambda_i e_i\right)} \\
&= \overline{\sum_{i=1}^{n}\mu_i\overline{\lambda_j}\varphi(e_i,e_j)} \\
&= \sum_{i,j=1}^{n}\overline{\mu_i}\lambda_j\overline{a_{ij}} \\
&= \sum_{i,j=1}^{n}\lambda_i\overline{\mu_j a_{ji}} \\
&= \sum_{i,j=1}^{n}\lambda_i\overline{\mu_j}a_{ij} \\
&= \varphi(u,v)
\end{aligned}$$

□

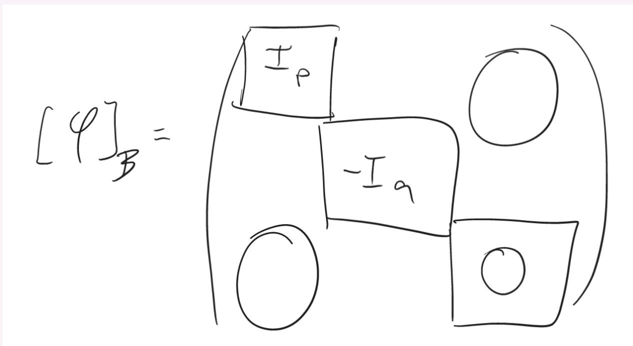
Polarization identity

A Hermitian form φ on a complex vector space V is *entirely determined* by: $Q : V \to \mathbb{R}$, $u \mapsto \varphi(u,u)$ via the formula:

$$\begin{aligned}
\varphi(u,v) &= \frac{1}{4}[Q(u+v) - Q(u-v) + iQ(u+iv) - iQ(u-iv)] \\
&= \text{polarization identity for Hermitian forms}
\end{aligned}$$

Proof. Exercise (just check). □

Theorem (Sylvester's law of inertia for Hermitian forms)**.** $\dim_{\mathbb{C}} V < \infty$, $\varphi : V \times V \to \mathbb{C}$ a Hermitian form on V. Then $\exists \mathcal{B} = (v_1, \dots, v_n)$ basis of V:



where P and q depend only on φ.

Proof. (Sketch: nearly identical to the real case of symmetric forms).

- Existence: $\varphi \equiv 0$, done. Assume $\varphi \neq 0$, then the polarization identity ensures that there exists $e_1 \neq 0$ such that
$$\varphi(e_1, e_1) \neq 0$$
Rescale:
$$v_1 = \frac{e_1}{\sqrt{|\varphi(e_1, e_1)|}}$$
$\implies \varphi(v_1, v_1) = \pm 1$. Then we consider the orthogonal:
$$W = \{w \in V \mid \varphi(v_1, w) = 0\}$$
and we check (verbatim like in the real case)
$$V = \langle v_1 \rangle \oplus W$$
($\dim W = n - 1$). Now argue by induction on the dimension on V by considering $\varphi\,|_W$ which is Hermitian on $W \times W$.

- Uniqueness of p: As in the real case,

  $p \equiv$ maximal dimension of a subspace on which φ is definite positive ($\varphi(u,u) \in \mathbb{R}$)

  Similarly for q.

□

Skew Symmetric Real Valued Forms

$F = \mathbb{R}$, V vector space over $\mathbb{R}$.

Definition (skew symmetric)**.** A bilinear form $\varphi : V \times V \to \mathbb{R}$ is *skew symmetric* if:
$$\varphi(u, v) = -\varphi(v, u) \quad \forall (u, v) \in V \times V$$

This is also often called antisymmetric.

Remark. (i) $\varphi(u,u) = -\varphi(u,u)$ so $\varphi(u,u) = 0$. $\forall u \in V$.

(ii) $\forall \mathcal{B}$ basis of V, $[\varphi]_{\mathcal{B}} = -[\varphi]_{\mathcal{B}}^{\top}$.

(iii) $A \in \mathcal{M}_n(\mathbb{R})$,
$$A = \frac{1}{2}(A + A^{\top}) + \frac{1}{2}(A - A^{\top})$$
i.e. decomposition into symmetric and antisymmetric / skew symmetric parts.

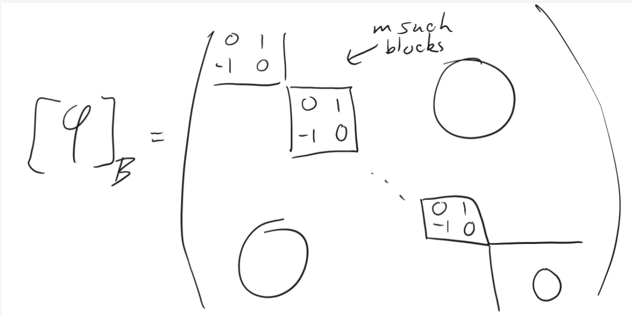
Theorem (Sylvester for skew symmetric form).

- V vector space over $\mathbb{R}$, $\dim_{\mathbb{R}} V < \infty$
- $\varphi : V \times V \to \mathbb{R}$ skew symmetric bilinear form.

Then there exists $\mathcal{B}$ basis of V,

$$\mathcal{B} = (v_1, w_1, v_2, w_2, \dots, v_m, w_m, v_{2m+1}, v_{2m+2}, \dots, v_n)$$

such that



Corollary. Skew symmetric matrices have an even rank.

Proof. (Sketch). Induction on the dimension of V.

- $\varphi \equiv 0$ then done.
- $\varphi \neq 0 \implies \exists (v_1, v_w) \in V \times V$ such that $\varphi(v_1, w_1) \neq 0$.
- $v_1 \neq 0$, $w_1 \neq 0$, after scaling:
$$\varphi(v_1, w_1) = 1$$
$$\implies \varphi(w_1, v_1) = -1$$
since skew symmetric.
- (v_1, w_1) linearly independent.
$$\varphi(v_1, \lambda v_1) = \lambda \varphi(v_1, v_1) = 0$$
since skew symmetric.
- Define $U = \langle v_1, w_1 \rangle$.
$$W = \{v \in V \mid \varphi(v_1, v) = \varphi(w_1, v) = 0\}$$
Exercise: show that $V = U \oplus W$.
- Now apply the induction hypothesis to $\varphi|_{W \times W}$ skew symmetric. □

Inner Product Spaces

- definite positive bilinear forms:
  - $\rightarrow$ Scalar product
  - $\rightarrow$ Norm (distance)

$\Longrightarrow$ SPECTACULAR generalisation / application to infinite dimensional spaces:

Hilbert Spaces

$\rightarrow$ part II (linear analysis, analysis of functions).

Definition (Inner product)**.** Let V be a vector space over $\mathbb{R}$ (respectively $\mathbb{C}$). An *inner product* is a positive definite symmetric (respectively Hermitian) bilinear form φ on V.

Notation. $\varphi(u,v) = \langle u, v\rangle$.

If such a bilinear form exists, V is called a real (respectively complex) inner product space.

Example. (i) $\mathbb{R}^n$, $x = (x_1, \dots, x_n)^\top$, $y = (y_1, \dots, y_n)^\top$,

$$\langle x, y\rangle = \sum_{i=1}^n x_i y_i$$

$\rightarrow$ inner product.

(ii) $\mathbb{C}^n$, $\langle x, y\rangle = \sum_{i=1}^n x_i \overline{y_i} \rightarrow$ inner product.

(iii) $V = \mathcal{C}([0,1], \mathbb{C})$

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\mathrm{d}t$$

"L^2 scalar product"

One can check that (i), (ii), (iii) are inner products.

$$\langle u, u\rangle = 0 \implies u = 0$$

$\rightarrow$ *definite* positive assumption.

Start of lecture 21

1.23 Gram Schmidt and orthogonal complement

- V vector space over $\mathbb{R}$ (or $\mathbb{C}$). An inner product is a *positive definite* symmetric (or Hermitian) bilinear form on V.

> **Notation.** $\varphi(u, v) = \langle u, v\rangle$.

- Norm: $\|v\| = \sqrt{\langle v, v\rangle}$ (the norm). Then $\|v\| \geq 0$ and $\|v\| = 0 \iff v = 0$.

$\to$ associated notion of length.

> **Lemma** (Cauchy-Schwartz).
>
> $$|\langle v, v\rangle| \leq \|u\|\|v\|$$
>
> More over, equality holds if and only if u and v are proportional.

Proof. $f = \mathbb{R}$ or $\mathbb{C}$. Let $t \in F$, then

$$\begin{aligned} 0 &\leq \|tu - v\|^2 \\ &= \langle tu - v, tu - v\rangle \\ &= t\bar{t}\langle u, u\rangle - t\langle v, v\rangle - \bar{t}\langle v, u\rangle + \|v\|^2 \\ &= |t|^2\|u\|^2 - 2\operatorname{Re}(t\langle v, u\rangle) + \|v\|^2 \end{aligned}$$

Explicitly: the minimum is taken at $t = \frac{\overline{\langle u, v\rangle}}{\|u\|^2}$

$$\implies 0 \leq \frac{|\langle u, v\rangle|^2}{\|u\|^2}\|u\|^2 - 2\operatorname{Re}\left(\frac{|\langle u, v\rangle|^2}{\|u\|^2}\right) + \|v\|^2$$

$$\implies 0 \leq \|v\|^2 - \frac{|\langle u, v\rangle|^2}{\|u\|^2}$$

$$\implies |\langle u, v\rangle|^2 \leq \|u\|^2\|v\|^2$$

Exercise: equality $\implies$ u and v are proportional. □

> **Corollary** (Triangle inequality).
>
> $$\|u + v\| \leq \|u\| + \|v\| \tag{$*$}$$

$\to$ key to show that $\|\bullet\|$ is a norm.

Proof.

$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle \\ &= \|u\|^2 + 2\operatorname{Re}(\langle u, v\rangle) + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2\end{aligned}$$

□

Definition. A set $(e_1, \ldots, e_k)$ of vectors of V is

(i) Orthogonal: if $\langle e_i, e_j\rangle = 0$ if $i \neq j$.

(ii) Orthonormal: if $\langle e_i, e_i\rangle = \delta_{ij}$ where

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Lemma. If $(e_1, \ldots, e_k)$ is orthogonal, then

(i) The family is free

(ii) $v = \sum_{j=1}^k \lambda_j e_j$, then

$$\lambda_j = \frac{\langle v, e_j\rangle}{\|e_j\|^2}$$

Proof. (i) $\sum_{j=1}^k \lambda_j e_j = 0$

$$\implies 0 = \left\langle \sum_{j=1}^k \lambda_j e_j, e_i \right\rangle = \sum_{j=1}^k \lambda_j \langle e_j, e_i\rangle = \lambda_i$$

so the family is free.

(ii) $v = \sum_{i=1}^k \lambda_i e_i$.

$$\begin{aligned}&\implies \langle v, e_j\rangle = \lambda_j \langle e_j, e_j\rangle = \lambda \|e_j\|^2 \\ &\implies \lambda_j = \frac{1}{\|e_j\|^2}\langle v, e_j\rangle\end{aligned}$$

□

Lemma (Parseval's Identity)**.** If V is a finite dimensional inner product space and $(e_1, \ldots, e_n)$ is an *orthonormal* basis, then

$$\langle u, v\rangle = \sum_{i=1}^{n} \langle u, e_i\rangle \overline{\langle v, e_i\rangle}$$

In particular, in an *orthonormal basis*,

$$\|v\|^2 = \langle v, v\rangle = \sum_{i=1}^{n} |\langle v, e_i\rangle|^2$$

$$v = \sum_{i=1}^{n} \langle v, e_i\rangle e_i$$

$(\|e_i\| = 1)$

Proof. $u = \sum_{i=1}^{n} \langle e, e_i\rangle e_i$, $\|e_1\| = 1$, $v = \sum_{i=1}^{n} \langle v, e_i\rangle e_i$

$$\implies \langle u, v\rangle \left\langle \sum_{i=1}^{n} \langle u, e_i\rangle e_i, \sum_{i=1}^{n} \langle v, e_i\rangle e_i \right\rangle = \sum_{i=1}^{n} \langle u, e_i\rangle \overline{\langle v, e_i\rangle}$$ □

Theorem (Gram-Schmidt orthogonalisation process)**.** V inner product space let I countable (finite) est and $(v_i)_{i \in I}$ linearly independent. Then there exists a sequence $(e_i)_{i \in I}$ of *orthonormal* vectors such that

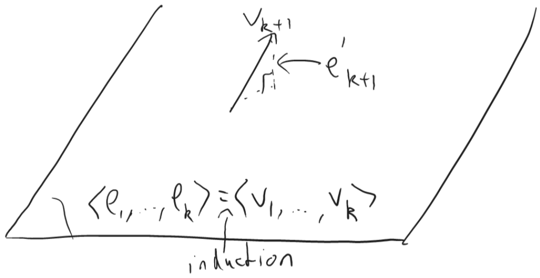
$$\operatorname{span}\langle v_1, \ldots, v_k\rangle = \operatorname{span}\langle e_1, \ldots, e_k\rangle$$

$\forall k \geq 1$.

$\rightarrow$ if $\dim V < \infty$, then we have existence of an *orthonormal* basis.

Proof. We construct the $(e_i)_{i \in I}$ family by induction on k.

- $k = 1$, $v_1 \neq 0 \implies e_1 = \frac{v_1}{\|v_1\|}$.
- Say we found $(e_1, \ldots, e_k)$, we look for e_{k+1}.



We define:

$$e'_{k+1} = v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, e_i \rangle e_i$$

- $e'_{k+1} \neq 0$. Indeed, otherwise,

$$v_{k+1} \in \langle e_1, \dots, e_k \rangle = \langle v_1, \dots, v_k \rangle$$

  which would contradict the fact that $(v_i)_{i \in I}$ is free.

- Pick $1 \leq j \leq k$:

$$\begin{aligned} \langle e'_{k+1}, e_j \rangle &= \left\langle v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, e_i \rangle e_i, e_j \right\rangle \\ &= \langle v_{k+1}, e_j \rangle - \langle v_{k+1}, e_j \rangle \\ &= 0 \end{aligned}$$

- $\langle v_1, \dots, v_{k+1} \rangle = \langle e_1, \dots, e_k, e'_{k+1} \rangle$.

- We take $e_{k+1} = \frac{e'_{k+1}}{\|e'_{k+!}\|}$ □

$\implies$ Gram Schmidt designs an algorithm to compute e_k for all k.

Corollary. V *finite dimensional* inner product space. Then any orthonormal 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, so we can extend $(e_1, \dots, e_k, v_{k+1}, \dots, v_n)$ basis of V. Apply Gram-Schmidt to this set noticing that there is no need to modify the first k vectors.

$$\implies (e_1, \dots, e_k, e_{k+1}, \dots, e_n)$$

orthonormal basis of V. □

Note. $A \in \mathcal{M}_n(\mathbb{R})$, then A has *orthonormal* column vectors if and only if

$$A^\top A = \text{id} \qquad (\mathbb{R})$$

$$A^\top \overline{A} = \text{id} \qquad (\mathbb{C})$$

Definition. (i) $A \in \mathcal{M}_n(\mathbb{R})$ is *orthogonal* if:

$$A^\top A = \text{id} \qquad (\iff A^{-1} = A^\top)$$

(ii) $A \in \mathcal{M}_n(\mathbb{C})$ is unitary if:

$$A^\top \overline{A} = \text{id} \qquad (\iff A^{-1} = \overline{A^\top})$$

Proposition. $A \in \mathcal{M}_n(\mathbb{R})$ (respectively $\mathcal{M}_n(\mathbb{C})$), then A can be written $A = RT$ where:

- T is upper triangular
- R is orthogonal (respectively unitary)

Proof. Exercise: apply Gram Schmidt to the $(c_1, \dots, c_n)$ column vectors of A. □

Orthogonal complement and projection

Definition.
- V inner product space
- $V_1, V_2 \leq V$.

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. $V = V_1 \overset{\perp}{\oplus} V_2$ $(V = V_1 + V_2)$ TODO…

Remark. $v \in V_1 \cap V_2$, $\|v\|^2 = \langle v, v \rangle = 0 \implies v = 0$.

Definition (orthogonal)**.** V inner product space, $W \leq V$.

$$W^\perp = \{v \in V \mid \langle v, w \rangle = 0 \ \forall w \in W\} = \text{orthogonal of } W$$

Lemma. V inner product space, $\dim V < \infty$, $W \leq V$. Then

$$V = W \overset{\perp}{\oplus} W^\perp \tag{$*$}$$

Start of lecture 22

1.24 Orthogonal complement and adjoint map

Definition. Suppose $V = U \oplus W$ (U is *a* complement of W in V). We define $\pi : V \to W$, $v = u + w \mapsto w$. Then

- π is linear
- $\pi^2 = \pi$

We say that π is the *projector operator* onto W.

Remark. $\text{id}\pi \equiv$ projection onto $U \to V$ inner product space, W finite dimensional, then we can chose $U = W^\perp$ and π is explicit.

Lemma.
- Let V be an inner product space
- Let $W \leq V$, W finite dimensional.

Let $(e_1, \dots, e_k)$ be an orthonormal basis of W (given by Gram-Schmidt). Then

(i) $\pi(v) = \sum_{i=1}^k \langle v, e_i \rangle e_i \ \forall v \in V$ and $V = W \overset{\perp}{\oplus} W^\perp$.

(ii) $\forall v \in V$, $\forall w \in W$,

$$\|v - \pi(v)\| \leq \|v - w\|$$

with equality if and only if $w = \pi(v)$ (that is $\pi(v)$ is the point in W closest to v).

Remark. Infinite dimensional generalisation:

- V inner product space $\to$ V Hilbert space
- W finite dimensional $\to$ W closed (completeness)

$\to$ part II class "Linear Analysis".

Proof. (i) $W = \text{span}\langle e_1, \ldots, e_k \rangle$, $(e_i)_{1 \le i \le k}$ orthogonal. Let us define

$$\pi(v) = \sum_{i=1}^{k} \langle v, e_i \rangle e_i$$

Observation:

$$v = \underbrace{\pi(v)}_{\in W} + \underbrace{v - \pi(v)}_{\text{claim: } \in W^\perp}$$

Indeed

$$\begin{aligned} v - \pi(v) \in W^\perp &\iff \forall w \in W, \langle v - \pi(v) w \rangle = 0 \\ &\iff \forall 1 \le j \le k, \langle v - \pi(v), e_j \rangle = 0 \end{aligned}$$

We compute:

$$\begin{aligned} \langle v - \pi(v), e_j \rangle &= \left\langle v - \sum_{i=1}^{k} \langle v, e_i \rangle e_i, e_j \right\rangle \\ &= \langle v, e_j \rangle - \langle v, e_j \rangle \\ &= 0 \end{aligned}$$

We have shown $v - \pi(v) \in W^\perp$ Hence

$$v = \underbrace{\pi(v)}_{\in W} + \underbrace{(v - \pi(v))}_{\in W^\perp}$$

$$\implies V = W + W^\perp$$

And $v \in W \cap W^\perp$

$$\begin{aligned} &\implies \|v\|^2 = \langle \underbrace{v}_{\in W}, \underbrace{v}_{\in W^\perp} \rangle = 0 \\ &\implies v = 0 \end{aligned}$$

So

$$V = W \overset{\perp}{\oplus} W^\perp$$

(ii) Indeed, let $w \in W$, then

$$\begin{aligned}\|v-w\|^2 &= \| \underbrace{v-\pi(v)}_{\in W^\perp} + \underbrace{\pi(v)-w}_{\in W} \|^2 \\ &= \langle v-\pi(v)+\pi(v)-w, v-\pi(v)+\pi(v)-w\rangle \\ &= \|v-\pi(v)\|^2 + \|\pi(v)-w\|^2 \\ &\geq \|v-\pi(v)\|^2\end{aligned}$$

With equality if and only if $w = \pi(w)$. PYTHAGORAS. □

Adjoint map

Definition. Let V, W be finite dimensional inner product spaces, let $\alpha \in \mathcal{L}(V, W)$. Then there exists a *unique* linear map

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

such that $\forall (v, w) \in V \times W$,

$$\langle \alpha(v), w\rangle = \langle v, \alpha^*(w)\rangle$$

Moreover, if $\mathcal{B}$ is an orthonormal basis of V and $\mathcal{C}$ is an orthonormal basis of W then

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

Proof. Computation: $\mathcal{B} = (v_1, \dots, v_n)$, $\mathcal{C} = (w_1, \dots, w_m)$, $A = [\alpha]_{\mathcal{B},\mathcal{C}} = (a_{ij})$. Existence

$$[\alpha^*]_{\mathcal{C},\mathcal{B}} = \overline{A}^\top = C = (c_{ij})$$

$c_{ij} = \overline{a_{ji}}$. We compute:

$$\begin{aligned}\left\langle \alpha\left(\sum_i \lambda_i v_i\right), \sum_j \mu_j w_j \right\rangle &= \left\langle \sum_{i,k} \lambda_i a_{ki} w_k, \sum_j \mu_j w_j \right\rangle \\ &= \sum_{i,j} \lambda_i a_{ji} \overline{\mu_j} \qquad \text{(orthonormal)}\end{aligned}$$

Then

$$\begin{aligned}\left\langle \sum_i \lambda_i v_i, \alpha^*\left(\sum_j \mu_j w_j\right)\right\rangle &= \left\langle \sum_i \lambda_i v_i, \sum_{j,k} \mu_j c_{kj} v_k \right\rangle \\ &= \sum_{i,j} \lambda_i \overline{c_{ij} \mu_j}\end{aligned}$$

So the expressions are equal because $\overline{c_{ij}} = a_{ji}$. So this proves existence. Uniqueness follows by computing $\alpha^*(w_j) \to$ exercise. □

Remark. We are using the same notation α^* for the adjoint of α and the dual of α. V, W are real product spaces, $\alpha \in \mathcal{L}(V, W)$,

$$\psi_{R,V} : V \xrightarrow{\simeq} V^*$$
$$v \mapsto \langle \bullet, v \rangle$$
$$\psi_{R,W} : W \xrightarrow{\simeq} W^*$$
$$w \mapsto \langle \bullet, w \rangle$$

Then the adjoint map of α is given by:

$$W \underset{\psi_{R,W}}{\longrightarrow} W^* \underset{\text{dual of } \alpha}{\longrightarrow} V^* \underset{\psi_{R,V}^{-1}}{\longrightarrow} V$$

Self adjoint maps and isometries

Definition. V inner product space finite dimensional $\alpha \in \mathcal{L}(V)$, $\alpha^* \in \mathcal{L}(V)$ the adjoint map. Then:

- $\langle \alpha v, w \rangle = \langle v, \alpha w \rangle \ \forall (v, w) \in V \times V \iff \alpha = \alpha^*$. We call such a map self adjoint. ($\mathbb{R}$ α symmetric, $\mathbb{C}$ α Hermitian).
- $\langle \alpha v, \alpha w \rangle = \langle v, w \rangle \ \forall (v, w) \in V \times V \iff \alpha^* = \alpha^{-1}$ we call an isometry. ($\mathbb{R}$ α orthogonal, $\mathbb{C}$ α unitary).

Proof. Check the equivalence that preserving the scalar product

$$(\langle \alpha v, \alpha w \rangle = \langle v, w \rangle \qquad \forall (v, w) \in V \times V)$$

is equivalent to (α invertible and $\alpha^* = \alpha^{-1}$)

$\Rightarrow$ $\langle \alpha v, \alpha w \rangle = \langle v, w \rangle \ \forall (v, w) \in V \times V$. Use $v = w$:

$$\|\alpha v\|^2 = \langle \alpha v, \alpha v \rangle = \langle v, v \rangle = \|v\|^2$$

(α preserves the norm: isometry) So $\ker \alpha = \{0\}$, so α bijective, α^{-1} well defined. (since finite dimensional). $\alpha \in \mathcal{L}(V)$, Then $\forall (v, w) \in V \times V$,

$$TODO$$

$$\implies \forall v \in V, \langle v, \alpha^* w \rangle = \langle v, \alpha^{-1} w \rangle$$

$$\implies \forall v \langle v, \alpha^* w - \alpha^{-1} w \rangle = 0$$

I take $v = \alpha^* w - \alpha^{-1} \omega$

$$\implies \alpha^* w = \alpha^{-1} w \ \forall w \in V$$

$$\implies \alpha^* = \alpha^{-1}$$

1. $\alpha \in \mathcal{L}(V)$, $\alpha^* = \alpha^{-1}$, then

$$\langle \alpha v, \alpha w \rangle = \langle v, \alpha^* \alpha w \rangle = \langle v, w \rangle$$

□

TODO
α isometry ($\alpha = \alpha^{-1}$)

$$\iff \forall (v, w) \in V \times V \langle \alpha v, \alpha w \rangle = \langle v, w \rangle$$

$$\iff \forall v \in V, \|\alpha(v)\| = \|v\|$$

(preservation of scalar product $\iff$ preservation of the norm)

Lemma. V finite dimensional real (complex) inner product space. Then $\alpha \in \mathcal{L}(V)$ is:

(i) Self adjoint if and only if in *any* orthonormal basis of V, $[\alpha]_{\mathcal{B}}$ is symmetric (Hermitian).

(ii) An isometry if and only if in *any* orthonormal basis of V, $[\alpha]_{\mathcal{B}}$ is orthogonal (unitary).

Proof. $\mathcal{B}$ orthonormal basis,

$$[\alpha^*]_{\mathcal{B}} = \overline{[\alpha]_{\mathcal{B}}^\top}$$

- Self adjoint $\overline{[\alpha^*]_{\mathcal{B}}^\top} = [\alpha]_{\mathcal{B}}$
- Isometry $\overline{[\alpha]_{\mathcal{B}}^\top} = [\alpha]_{\mathcal{B}}^{-1}$.

□

Definition. V finite dimensional inner product space.

- $F = \mathbb{R}$,

$$\theta(V) = \{\alpha \in \mathcal{L}(V), \alpha \text{ isometry}\} \equiv \text{orthogonal group of } V$$

- $F = \mathbb{C}$,

$$U(V) = \{\alpha \in L(V), \alpha \text{isometry}\} \equiv \text{unitary group of } V$$

Remark. V finite dimensional, $\{e_1, \dots, e_n\}$ orthonormal basis.

- $F = \mathbb{R}$, $\theta(V) \leftrightarrow \{\text{orthonormal basis of } V\}$, $\alpha \mapsto (\alpha(e_1, \dots, \alpha(e_n))$.
- $F = \mathbb{C}$, $U(V) \leftrightarrow \{\text{orthonormal basis of } V\}$, $\alpha \mapsto (\alpha(e_1, \dots, \alpha(e_n))$.

Start of lecture 23

1.25 Spectral theory for self adjoint maps

- **Spectral theory** $\equiv$ study of the spectrum of operators
  - $\rightarrow$ mathematics
  - $\rightarrow$ physics (QUANTUM MECHANICS)
  - $\Rightarrow$ INFINITE DIMENSIONAL. Finite dimension $\rightarrow$ infinite dimension. Linear maps $\rightarrow$ Hilbert space / compact operator.
- Adjoint operator: V, W finite dimensional inner product spaces, $\alpha \in \mathcal{L}(V, W)$, then the adjoint $\alpha^* \in \mathcal{L}(W, V)$ such that $\forall (v, w) \in V \times W$,

$$\langle \alpha(v), w \rangle_W = \langle v, \alpha^*(w) \rangle_V$$

  We defined:
  - Self adjoint maps, $V = W$, $\alpha = \alpha^*$,

$$\iff \forall (v, w) \in V \times V, \quad \langle \alpha v, w \rangle = \langle v, \alpha w \rangle$$

  - isometries $V = W$, $\alpha^* = \alpha^{-1}$

$$\iff \forall (v, w) \in V \times V, \quad \langle \alpha v, \alpha w \rangle = \langle v, w \rangle$$

  - $\mathbb{R}$: orthogonal group
  - $\mathbb{C}$: unitary group.

Spectral theory for self adjoint operators

Lemma. Let V be a finite dimensional inner product space. Let $\alpha \in \mathcal{L}(V)$ be self adjoint: ($\alpha = \alpha^*$). Then:

(i) α has real eigenvalues

(ii) eigenvectors of α with respect to *different* eigenvalues are orthogonal.

Proof. (i) $v \in V \setminus \{0\}$, $\lambda \in \mathbb{C}$ such that $\alpha v = \lambda v$. Then

$$\begin{aligned} \lambda \|v\|^2 &= \langle \lambda v, v \rangle \\ &= \langle \alpha v, v \rangle \\ &= \langle v, \alpha^* v \rangle \\ &= \langle v, \alpha v \rangle \\ &= \langle v, \lambda v \rangle \\ &= \overline{\lambda} \|v\|^2 \end{aligned}$$

So $(\lambda - \overline{\lambda}) \|v\|^2 = 0$. But $\|v\|^2 \neq 0$ since $v \neq 0$ so $\lambda = \overline{\lambda}$ so $\lambda \in \mathbb{R}$.

(ii) $\alpha v = \lambda v$, $\lambda \in \mathbb{R}$, $v \neq 0$. $\alpha w = \mu w$, $\mu \in \mathbb{R}$, $w \neq 0$. Also $\lambda \neq \mu$. Then

$$\begin{aligned}\lambda\langle v, w\rangle &= \langle \lambda v, w\rangle \\ &= \langle \alpha v, w\rangle \\ &= \langle v, \alpha^* w\rangle \\ &= \langle v, \alpha w\rangle \\ &= \langle v, \mu w\rangle \\ &= \overline{\mu}\langle v, w\rangle \\ &= \mu\langle v, w\rangle\end{aligned}$$

So $(\lambda - \mu)\langle v, w\rangle = 0$. But $\lambda \neq \mu$ so $\langle v, w\rangle = 0$.

□

Theorem. Let V be a finite dimensional inner product space. Let $\alpha \in \mathcal{L}(V)$ be self adjoint ($\alpha = \alpha^*$). Then V has an *orthonormal* basis made of eigenvectors of α.

→ We also say: α can be diagonalised in an orthonormal basis for V.

Proof. $F = \mathbb{R}$ or $\mathbb{C}$. We argue by induction on the dimension of V, $\dim_F V = n$.

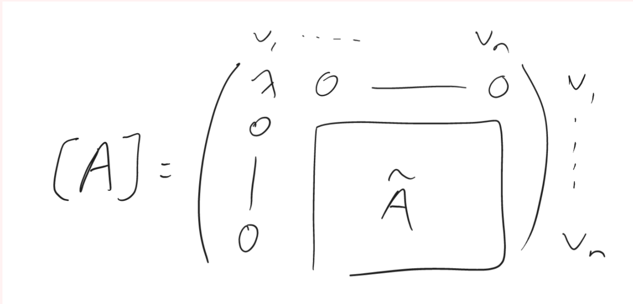
- $n = 1 \to$ trivial.
- $n - 1 \to n$. $\mathcal{B}$ any orthonormal basis of V say $A = [\alpha]_{\mathcal{B}}$. By the fundamental Theorem of Algebra, we know that $\chi_A(t)$ ($\equiv$ characteristic polynomial of A) has a *complex* root. This root is an eigenvalue of α and $\alpha = \alpha^* \implies$ this root is *real*. Let us call $\lambda \in \mathbb{R}$ this eigenvalue, pick an eigenvector $v_1 \in V \setminus \{0\}$ such that $\|v_1\| = 1$, $\alpha v_1 = \lambda v_1$. Let $U = \langle v_1\rangle^\perp \leq V$. Then KEY OBSERVATION: U stable by α, i.e. $\alpha(U) \leq U$. Indeed, let $u \in U$, then:

$$\begin{aligned}\langle \alpha u, v_1\rangle &= \langle u, \alpha v_1^*\rangle \\ &= \langle u, \alpha v_1\rangle \\ &= \langle u, \lambda v_1\rangle \\ &= \lambda\langle u, v_1\rangle \\ &= 0\end{aligned}$$

  So $\alpha(u) \in U$. This implies: we may consider $\alpha|_U \in \mathcal{L}(U)$ and self adjoint, and then $n = \dim V = \dim U + 1$, so $\dim U = n - 1$ so by induction hypothesis there exists $(v_2, \dots, v_n)$ orthonormal basis of eigenvectors for $\alpha|_U$. Then $V = \langle v_1\rangle \overset{\perp}{\oplus} U$ so $(v_1, \dots, v_n)$ orthonormal basis of V made of eigenvectors of α.

□

Remark. If you want to think in terms of matrices for the proof of (ii), then the choice of U means that $[A]$ is written as



Corollary. V finite dimensional inner product space. If $\alpha \in \mathcal{L}(V)$ is self adjoint, then V is the *orthogonal direct sum* of all the eigenspaces of α.

Spectral theory for unitary maps

Lemma. V be a *complex* inner product space (Hermitian sesquilinear structure). Let $\alpha \in \mathcal{L}(V)$ be unitary ($\alpha^* = \alpha^{-1}$). Then

(i) all the eigenvalues of α lie on the unit circle

(ii) eigenvectors corresponding to *distinct* eigenvalues are *orthogonal*.

Proof. (i) $\alpha v = \lambda v$, $v \neq 0$, $\lambda \in \mathbb{C}$.

- $\lambda \neq 0$: α unitary implies α invertible.
- $$\begin{aligned} \lambda\|v\|^2 &= \lambda\langle v, v\rangle \\ \langle \lambda v, v\rangle & \\ &= \langle \alpha v, v\rangle \\ &= \langle v, \alpha^* v\rangle \\ &= \langle v, \alpha^{-1} v\rangle \\ &= \left\langle v, \frac{1}{\lambda} v\right\rangle \\ &= \frac{1}{\overline{\lambda}}\|v\|^2 \end{aligned}$$

  So $\lambda\|v\|^2 = \frac{1}{\overline{\lambda}}\|v\|^2$ so since $v \neq 0$, $\lambda\overline{\lambda} = 1$, i.e. $|\lambda| = 1$.

- $\alpha v = \lambda v$, $\alpha w = \mu w$, $\lambda, \mu \neq 0$, $\mu \neq \lambda$. Then

$$\begin{aligned}\lambda\langle v, w\rangle &= \langle \lambda v, w\rangle \\ &= \langle \alpha v, w\rangle \\ &= \langle v, \alpha^* w\rangle \\ &= \langle v, \alpha^{-1} w\rangle \\ &= \left\langle v, \frac{1}{\mu} w\right\rangle \\ &= \frac{1}{\overline{\mu}}\langle v, w\rangle \\ &= \mu\langle v, w\rangle\end{aligned}$$

(by (i)). So $(\lambda - \mu)\langle v, w\rangle = 0$. But $\lambda \neq \mu$ so $\langle v, w\rangle = 0$.

□

Theorem (Spectral theory for unitary maps)**.** Let V be a finite dimensional *complex* inner product space. Let $\alpha \in \mathcal{L}(V)$ be unitary ($\alpha^* = \alpha^{-1}$). Then V has an orthonormal basis made of eigenvectors of α.

→ Equivalently, α unitary on V Hermitian can be diagonalised in an orthonormal basis.

Proof. Pick $\mathcal{B}$ any orthonormal basis of V. $A = [\alpha]_{\mathcal{B}}$. Then $\chi_A(t)$ ($\equiv$ characteristic polynomial of A) has a *complex* root. So α has a *complex* eigenvalue. Fix $v_1 \in V \setminus \{0\}$ with $\|v_1\| \neq 0$, $\alpha v_1 = \lambda v_1$. Let $U = \langle v_1\rangle^\perp$. Then: KEY OBSERVATION: $\alpha(U) \le U$. Indeed: $u \in U$, then

$$\begin{aligned}\langle \alpha u, v_1\rangle &= \langle u, \alpha^* v_1\rangle \\ &= \langle u, \alpha^{-1} v_1\rangle \\ &= \left\langle u, \frac{1}{\lambda} v_1\right\rangle \\ &= \frac{1}{\overline{\lambda}}\langle u, v_1\rangle \\ &= 0\end{aligned}$$

$\implies$ $\alpha u \in U$, so $\alpha(U) \le U$. We argue by induction on $\dim_{\mathbb{C}} V = n$. We consider $\alpha|_U \in \mathcal{L}(U)$ which is unitary, and by the induction hypothesis, $\alpha|_U$ is diagonalisable in an orthonormal basis $(v_2, \dots, v_n)$ of U $\implies$ $(v_1, \dots, v_n)$ is an orthonormal basis of V, made of eigenvectors of α. □

Warning. We used the complex structure to make sure that there is an eigenvalue (which is a priori complex valued).
In general, a real valued orthonormal matrix ($AA^\top = \mathrm{id}$) *cannot* be diagonalised over $\mathbb{R}$.

Example (Rotation map in $\mathbb{R}^2$)**.**

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

$$\chi_A(\lambda) = (\cos\alpha - \lambda)^2 + \sin^2\alpha$$

Then the eigenvalues are $\lambda = e^{\pm i\alpha}$ ($\notin$ generally). (Hence diagonalisable in $\mathbb{C}$ but not $\mathbb{R}$).

Start of lecture 24

1.26 Application to bilinear forms

Diagonalisation of self adjoint / unitary operators.

Theorem 1. Let V be a finite dimensional inner product space (over $\mathbb{R}$ or $\mathbb{C}$). Let $\alpha \in \mathcal{L}(V)$ be *self adjoint* ($\alpha = \alpha^*$). Then there exists an ***orthonormal*** basis of V made of ***eigenvectors*** of α.

Theorem. Let V be a finite dimensional *complex* inner product space. Let $\alpha \in \mathcal{L}(V)$ be unitary ($\alpha^* = \alpha^{-1}$). Then there exists an ***orthonormal*** basis of V made of ***eigenvectors*** of α.

These theorems are so important we stated them twice!

- Translate these statements for bilinear forms.

Corollary. let $A \in \mathcal{M}_n(\mathbb{R})$ (respectively $\mathbb{C}$) be a symmetric (respectively Hermitian) matrix. Then there is an orthogonal (respectively unitary) matrix such that $P^\top AP$ (respectively $P^\dagger AP$) is diagonal with real valued entries.

Remark. $P^\dagger = \overline{P}^\top$

Proof. $F = \mathbb{R}$ ($\mathbb{C}$). Let $\langle , \rangle$ be the standard inner product over $\mathbb{R}^n$. Then $A \in \mathcal{L}(F^n)$ is self adjoint, hence we can find an orthonormal (for the standard inner product) basis of F^n such that A is diagonal in this basis, say $(v_1, \dots, v_n)$. Let $P = (v_1 \mid \cdots \mid v_n)$

$$\begin{aligned}(v_1, \dots, v_n) \text{ orthonormal basis} &\iff P \text{ orthogonal (unitary)} \\ &\iff P^\top P = \mathrm{id}(P^\dagger P = \mathrm{id})\end{aligned}$$

So $P^{-1}AP = P^\top AP = D$, and we know λ_i are real, they are the eigenvalues of a symmetric operator. □

Corollary. Let V be a finite dimensional real (complex) inner product space. Let $\varphi : V \times V \to F$ be a symmetric (Hermitian) bilinear form. Then there is an orthonormal basis of V such that φ in this basis is represented by a diagonal matrix.

Proof. Let $\mathcal{B} = \{v_1, \dots, v_n\}$ be any orthonormal basis of V. Let $A = [\varphi]_{\mathcal{B}}$. Then since φ is symmetric (Hermitian), $A^\top = A$ ($^\dagger = A$), hence there is an orthogonal (unitary) matrix P such that $P^\top AP$ ($P^\dagger AP$) is diagonal, say D. Let v_i be the i-th row of $P^\top$ ($P^\dagger$), then $\{v_1, \dots, v_n\}$ is an orthonormal basis say $\mathcal{B}'$ of V and $[\varphi]_{\mathcal{B}'} = D$. (We are using the change of basis for bilinear forms). □

Remark. Diagonal entries of $P^\top AP$ ($P^\dagger AP$) are exactly the eigenvalues of A. Moreover:

$$\Delta(\varphi) = \#(\text{positive eigenvalues of } A) - \#(\text{negative eigenvalues of } A)$$

(recall Δ is the signature of a bilinear form)

Important corollary

Corollary (Simultaneous diagonalisation of Hermitian forms)**.** Let V be a finite dimensional real (complex) vector space. Let:

$$\varphi, \psi : V \times V \to F$$

φ, ψ are bilinear symmetric (Hermitian) forms. And suppose φ is positive definite. Then there exists $(v_1, \dots, v_n)$ basis of V with respect to which *both* bilinear forms are represented by a diagonal matrix.

Proof. φ is positive definite so φ induces a scalar product on V, 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 which ψ is represented by a diagonal matrix. Observe that φ in this basis is represented by the Identity matrix (because the basis orthonormal for φ: $\mathcal{B} = (v_1, \ldots, v_n)$, $\langle v_i, v_j \rangle = \delta_{ij} = \varphi(v_i, v_j)$) So both matrices of φ and ψ in $\mathcal{B}$ are diagonal. $\square$

Corollary (Matrix reformulation of simultaneous diagonalisation)**.** Let $A, B \in \mathcal{M}_n(\mathbb{R})$ (respectively $\mathcal{M}_n(\mathbb{C})$), both symmetric (respectively Hermitian). Assume $\forall x \neq 0$, $\overline{x}^\top A x > 0$. Then there exists $Q \in \mathcal{M}_n(\mathbb{R})$ (respectively $\mathcal{M}_n(\mathbb{C})$) invertible such that *both* $Q^\top A Q$ (respectively $Q^\dagger A Q$) and $Q^\top B Q$ (respectively $Q^\dagger B Q$) are diagonal.

Proof. Direct consequence of the simultaneous diagonalisation Theorem. $\square$