

Dispositions for
AdVec

Malthe Munk Karbo '14

November 6, 2017

CONTENTS

1	Basis and dimension	2
2	Dual space, dual basis	6
3	Tensor products	8
4	Alternating forms	11
5	Invariance, reducibility and projections	15
6	Combination of projections	17
7	Range, null-space, rank and nullity	19
8	Triangular form	22
9	Nilpotence	25
10	Jordan Normal Form	28
11	Perpendicular projections	31
12	Spectral theorem	33
13	Polar decomposition	35
14	Norms of linear transformations	37

BASIS AND DIMENSION

Agenda

1. Basic definitions: vector space, linear dependence, dimension and basis
2. Lemma about basis and spanning sets
3. Characterization of spanning sets, linear independent sets and bases of vector spaces.

Content

Definition 1.1

A **vector space** V of a field \mathcal{F} is an algebraic object consisting of a set V and a binary map $+: V \times V \rightarrow V$ turning V into an abelian group (with 0) and a map $F \times V \rightarrow V$ of scalar multiplication satisfying basic axioms:

- $\alpha(\beta v) = (\alpha\beta)v$ for all $v \in V$ and $\alpha, \beta \in \mathcal{F}$.
- $1_{\mathcal{F}}v = v$ for all $v \in V$.
- $\alpha(v + v') = \alpha v + \alpha v'$ for all $\alpha \in \mathcal{F}$ and $v, v' \in V$.
- $(\alpha + \beta)v = \alpha v + \beta v$ for all $\alpha, \beta \in \mathcal{F}$ and $v \in V$.

From now on V will denote a vector space over a field \mathcal{F} .

Definition 1.2

Given a set of vectors $v_1, \dots, v_n \in V$, a **linear combination** of these vectors is an element of the form

$$x = \sum_{i=1}^n \alpha_i v_i, \text{ for } \alpha_1, \dots, \alpha_n \in \mathcal{F}$$

Definition 1.3

Given a subset $S \subseteq V$, we define the **span** of S to be the set

$$\text{Span}_{\mathcal{F}}(S) = \left\{ \sum_{i \in F} \alpha_i x_i \mid F \subseteq \mathbb{N}_0 \text{ finite}, \alpha_i \in \mathcal{F}, x_i \in S \right\}$$

Definition 1.4

We say that a subset $S \subseteq V$ **spans** V if $\text{Span}(S) = V$.

Definition 1.5

If there is a finite subset $S \subseteq V$ spanning V , we say that V is **finite dimensional** (and infinite otherwise).

Definition 1.6

For finite dimensional V , we say that **the dimension** of V is the minimal number of vectors needed to span all of V , abbreviated $\dim(V)$. That is

$$\dim(V) = \min \{n \in \mathbb{N} \mid \text{there is } x_1, \dots, x_n \in V \text{ such that } \text{Span}_{\mathcal{F}}(\{x_1, \dots, x_n\}) = V\}$$

Definition 1.7

A subset $S \subseteq V$ is **linearly dependent** if there is $n > 0, x_1, \dots, x_n \in S$ and $\alpha_1, \dots, \alpha_n \in \mathcal{F}$ with at least one $1 \leq i \leq n$ such that $\alpha_i \neq 0$ such that

$$\sum_{1 \leq i \leq n} \alpha_i x_i = 0.$$

And we say that S is linearly independent otherwise.

Linear independence passes to subsets, and no set containing 0 can be linearly independent.

Theorem 1.8 (Do not show)

If $x_1, \dots, x_n \in V$, then the following are equivalent:

1. $\{x_1, \dots, x_n\}$ is lin. dependent
2. There is an integer $1 \leq k \leq n$ such that $x_k \in \text{Span}_{\mathcal{F}}\{x_1, \dots, x_{k-1}\}$.

Proof.

(2) \implies (1): If $x_k \in \text{Span}_{\mathcal{F}}\{x_1, \dots, x_{k-1}\}$, then there is $\alpha_1, \dots, \alpha_{k-1} \in \mathcal{F}$ such that

$$x_k = \sum_{i=1}^{k-1} \alpha_i x_i \implies \sum_{i=1}^{k-1} \alpha_i x_i - x_k = 0$$

so the set $\{x_1, \dots, x_{k-1}, x_k\}$ is linearly dependent, and since linear independence passes to subsets, we conclude that $\{x_1, \dots, x_n\}$ can't be linearly independent.

(1) \implies (2): Pick the smallest $1 \leq k \leq n$ such that $\{x_1, \dots, x_k\}$ is linearly dependent (by using the assumption of (1)). Then $\{x_1, \dots, x_{k-1}\}$ is linearly independent. Let $\alpha_1, \dots, \alpha_{k-1}, \beta \in \mathcal{F}$ be such that

$$\sum_{i=1}^{k-1} \alpha_i x_i + \beta x_k = 0 \implies \beta \neq 0$$

Since else $\{x_1, \dots, x_{k-1}\}$ would be linearly dependent. Then

$$x = \sum_{i=1}^{k-1} \left(-\frac{\alpha_i}{\beta}\right) x_i \in \text{Span}_{\mathcal{F}}\{x_1, \dots, x_{k-1}\}$$

□

Definition 1.9

A set of vectors $\{x_i\}$ is a basis for V if

1. It spans V
2. It is linearly independent.

Lemma 1.10 (do not show)

A set $S = \{x_i\}$ is a basis for V if and only if every $x \in V$ is represented in a unique way as a linear combination of x_i 's

Proof.

Proof is trivial: S spans $V \iff$ every element is represented by S is clear, and S is independent \iff every such representation is unique. □

Theorem 1.11 (do not show)

If $X = \{x_1, \dots, x_n\}, Y = \{y_1, \dots, y_m\}$ are finite subsets of V with X spanning V and Y linearly independent, then there is a subset $X' = \{x_{i_1}, \dots, x_{i_k}\} \subseteq X$ such that $Y \cup X'$ forms a basis for V .

Proof.

Form the set $S = \{y_1, \dots, y_m, x_1, \dots, x_n\} = Y \cup X$. Since $X \subseteq S$ we have S . If S is linearly independent, we are done. If not: First note that we may write

$$S = \{z_1, z_2, \dots, z_{n+m}\}, \text{ where } z_i = y_i \text{ for } 1 \leq i \leq m \text{ and } z_i = x_{i-m} \text{ for } m \leq i \leq m+n$$

By theorem above, we obtain $1 \leq i \leq n+m$ such that $\{z_1, \dots, z_{i-1}\}$ is linearly dependent. Since y_1, \dots, y_m is linearly independent, $m+1 < i \leq n+m$. Remove this element from the list S to form

$$S' = \{z_1, z_2, \dots, z_m, z_{m+1}, z_{m+3}, \dots, z_{n+m}\} = \{y_1, \dots, y_m, x_1, x_3, \dots, x_n\}$$

If S' is linearly independent, if not continue this process. Since S is a finite list, the process terminates at some point, finishing the proof. □

The above theorem implies that any finite dimensional space V has a basis. (X finite spanning, $Y = \emptyset$).

Lemma 1.12

If $X = \{x_1, \dots, x_n\}$ spans V and $Y = \{y_1, \dots, y_m\}$ is linearly independent, then $m \leq n$

Proof.

Assume towards a contradiction that $m > n$. Then X spans $V \implies S_1 := Y \cup X$ spans V . Use theorem §6 to pick $1 < i \leq n$ such that

$$x_i \in \text{Span}_{\mathcal{F}} \{y_1, x_1, \dots, x_{i-1}\}.$$

Set $X_1 := \{y_1, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$ and $Y_1 := \{y_2, \dots, y_m\}$. Remove x_i does not change the span, so X_1 spans all of V and Y_1 is linearly independent since $Y_1 \subseteq Y$. This process terminates after n steps, and we see that $X_n = \{y_1, \dots, y_n\}$ and $Y_n = \{y_{n+1}, \dots, y_m\}$ are non-empty sets ($m > n$) with X_n spanning V . Hence $y_{n+1} \in \text{Span}_{\mathcal{F}} X_n$ contradicting the fact that $\{y_1, \dots, y_n, y_{n+1}\}$ is linearly independent (linear independence passes to subsets) \square

This lemma has some nice implications:

Theorem 1.13

If V is a vector space with a basis of n elements ($n \in \mathbb{N}_0$). Then

1. Every lin. independent subset of V has $\leq n$ elements.
2. Every spanning set in V has $\geq n$ elements
3. Every basis has n elements

Proof.

The proof of (1) and (2) follows directly from the lemma. (3) follows from (1) and (2) \square

DUAL SPACE, DUAL BASIS

Agenda

1. Basic definitions:
2. Lemma about basis and spanning sets
3. Characterization of spanning sets, linear independent sets and bases of vector spaces.

Content

In the following we let V be a vector space over a field \mathcal{F} .

Definition 2.1

A linear function $y: V \rightarrow \mathcal{F}$ is a **linear functional**. The space of linear functionals is denoted by V' .

Definition 2.2

If $n = \dim(V) < \infty$, define the n 'th coordinate functional $y_j: V \rightarrow \mathcal{F}$ by

$$y_j(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n) = \alpha_j.$$

where x_1, \dots, x_n is a basis for V

For finite dimensional V , we then see that all $x \in V$ we have

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

and moreover

$$y_j(x_i) = \delta_{ij}, \text{ for } 1 \leq i, j \leq n$$

So that

$$y(x_j) = \sum_{i=1}^n y(x_i) \delta_{ij} = \sum_{i=1}^n y(x_i) y_i(x_j)$$

which implies that

Lemma 2.3

Given $y \in V'$, it holds that $y = \sum_{i=1}^n y(x_i)y_i$.

So that we have $\text{Span}\{y_1, \dots, y_n\} = V'$. In fact the set $\{y_1, \dots, y_n\}$ is a basis:

Theorem 2.4

$\dim(V') = n$ and $\{y_1, \dots, y_n\}$ is a basis for V' .

Proof.

Suppose that $\alpha_1, \dots, \alpha_n \in \mathcal{F}$ such that

$$\sum_{i=1}^n \alpha_i y_i = 0$$

Then, for all $j = 1, \dots, n$ we have

$$0 = \sum_{i=1}^n \alpha_i y_i(x_j) = \sum_{i=1}^n \alpha_i \delta_{ij} = \alpha_j.$$

□

Definition 2.5

The basis $\{y_1, \dots, y_n\}$ is called the **dual basis**.

There is a natural correspondance between V and it's bidual $V'' := (V')'$.

Definition 2.6

The map $T: V \rightarrow V''$ given by $T(x)(y) = y(x)$ is called the **natural correspondance**. It is a linear map in x , clearly.

Applying the above theorem twice we see that for finite dimensional V we have $n = \dim(V) = \dim(V') = \dim(V'')$, and in fact the map T is a linear isomorphism:

Theorem 2.7

The natural correspondance $T: V \rightarrow V''$ is a linear isomorphism for $\dim(V) < \infty$.

Proof.

Choose a basis $\{x_1, \dots, x_n\}$ for V . We then claim that $\{T(x_1), \dots, T(x_n)\}$ is a basis for V'' , and it will follow from showing that it is a linear independent set. Linear independence comes from the fact that T is injective, for if $u \neq v$, then they differ in one of their coordinates, i.e.

$$T(u)(y_j) = y_j(u) \neq y_j(v) = T(v)(y_j)$$

for some $1 \leq j \leq n$. So T is injective. And injective linear functions carries linearly independent sets to linearly independent sets, finishing the proof. □

CHAPTER 3
—
TENSOR PRODUCTS

Agenda

1. Bilinear forms definition and space of bilinear forms.
2. Bilinear forms basis and characterisation as matrices
3. Tensor product as linear functionals on bilinear forms
4. True and proper characterization of tensor products

Content

Let U, V be vector spaces over \mathcal{F} .

Definition 3.1

A **bilinear form** on $U \times V$ is a map $w: U \times V \rightarrow \mathcal{F}$ such that

- The map $x \mapsto w(x, y)$ is linear for all $y \in V$
- The map $y \mapsto w(x, y)$ is linear for all $x \in U$.

(This is **not** a regular linear functional on $U \times V$)

Examples include $w(x, y) = axy$ for some $a \in \mathbb{R}$ as a bilinear form on \mathbb{R}^2 .

Definition 3.2

We denote by $\text{Bil}(U, V)$ the set of bilinear forms $U \times V \rightarrow \mathcal{F}$. This is a vector space over \mathcal{F} .

For finite dimensional U, V we achieve a nice characterization of $\text{Bil}(U, V)$:

Theorem 3.3

If $\dim(U) = n$ and $\dim(V) = m$ have bases $\{x_1, \dots, x_n\}$ and $\{y_1, \dots, y_m\}$, then the map

$$w \mapsto (w(x_i, y_j))_{i,j}$$

is a linear isomorphism $\text{Bil}(U, V) \rightarrow M_{n,m}(\mathcal{F})$.

Proof.

Linearity of the map is easy to convince oneself of. So we show injectivity and surjectivity instead:

Injectivity: If $\{u_1, \dots, u_n\}$ and $\{v_1, \dots, v_m\}$ are dual bases for U' and V' , then for $w \in \text{Bil}(U, V)$ we see that

$$w(x, y) = w\left(\sum_{i=1}^n u_i(x)x_i, \sum_{j=1}^m v_j(y)u_j\right) = \sum_{i,j} u_i(x)v_j(y)w(x_i, y_j)$$

for all $x \in U$ and $y \in V$. Setting $(a_{i,j}) := (w(x_i, y_j))$ we see that w is uniquely determined by this matrix, so if two bilinear forms have the same matrix then they are equal.

Surjectivity: Given $(a_{i,j}) \in M_{n,m}(\mathcal{F})$, define the bilinear form $w \in \text{Bil}(U, V)$ by

$$w(x, y) = \sum_{i,j} u_i(x)v_j(y)a_{i,j}$$

This is clearly a bilinear map, and moreover $w(x_p, y_q) = \sum_{i,j} \delta_{i,p}\delta_{j,q}a_{i,j} = a_{p,q}$, so w is mapped to $(w(x_i, y_j)) = (a_{i,j})$, so it is surjective. \square

Theorem 3.4

Let $w_{i,j}(x, y) := u_i(x)v_j(y)$, then $w_{i,j} \in \text{Bil}(U, V)$ and the set of $w_{i,j}$ form a basis for $\text{Bil}(U, V)$.

Proof.

We see that $w_{i,j}(x_p, y_q) = u_i(x_p)v_j(y_q) = \delta_{i,p}\delta_{j,q}$, so $w_{i,j}$ is mapped to the ij 'th matrix unit in $M_{n,m}(\mathcal{F})$. The matrix units form a basis, and since the correspondence is an isomorphism, the set $\{w_{i,j}\}$ form a basis of $\text{Bil}(U, V)$. \square

Definition 3.5

The **tensor product** of U, V is defined as $U \otimes V := \text{Bil}(U, V)'$, and the tensor product of x and y is defined on $w \in \text{Bil}(U, V)$ by the map $w \mapsto w(x, y)$

Theorem 3.6

The basis of U, V induces the basis of $U \otimes V$, i.e., the set $\{x_i \otimes y_j\}$ forms a basis for $U \otimes V$.

Proof.

It is not hard to see, for it is just the dual basis of $\{w_{p,q}\}$:

$$x_i \otimes y_j(w_{p,q}) = w_{p,q}(x_i, y_j) = \delta_{i,p}\delta_{j,q}$$

\square

One can generalize this notion:

Definition 3.7

A tensor product of any vector spaces U, V is a vector space T and a bilinear maps $B: U \times V \rightarrow T$ such that for all vector spaces S and bilinear maps $A: U \times V \rightarrow S$ we have the following diagram

$$\begin{array}{ccc} U \times V & \xrightarrow{B} & T \\ \downarrow A & \swarrow \exists \varphi & \\ S & & \end{array}$$

Theorem 3.8

There exists a tensor product and it is unique up to isomorphism.

Proof.

For finite dimensional case, the existence is shown above: If the setting is the same as above, then the map φ making the diagram commute is simply the map $\varphi(x_i \otimes y_j) = A(x_i, y_j)$.

For uniqueness, assume that (T_1, B_1) and (T_2, B_2) are two tensor products of U, V . Then we get

$$\begin{array}{ccc} U \times V & \xrightarrow{B_1} & T_1 \\ \downarrow B_2 & \swarrow \exists \varphi & \\ T_2 & & \end{array}$$

and

$$\begin{array}{ccc} U \times V & \xrightarrow{B_2} & T_2 \\ \downarrow B_1 & \swarrow \exists \psi & \\ T_1 & & \end{array}$$

Giving us the commutative diagram

$$\begin{array}{ccc} U \times V & \xrightarrow{B_1} & T_1 \\ \downarrow B_1 & \swarrow \psi \circ \varphi \circ B_1 & \\ T_1 & & \end{array}$$

But also the identity makes the diagram

$$\begin{array}{ccc} U \times V & \xrightarrow{B_1} & T_1 \\ \downarrow B_1 & \swarrow \text{id}_{T_1} & \\ T_1 & & \end{array}$$

So uniqueness ensures that $\psi \circ \varphi = \text{id}_{T_1}$ and similarly $\varphi \circ \psi$ is the identity on T_2 , so they are inverses hence isomorphisms. \square

ALTERNATING FORMS

Agenda

1. Basic definitions: Multilinear forms and the vector space of k -linear forms.
2. S_k action on $\mathcal{L}^k(V)$
3. Skew-symmetric k -forms, alternating k -forms definitions. Alt \implies Skew.
char(F) $\neq 2$: Skew \implies alt and lin dept set \implies Alt forms are 0
4. Wedge product definition and wedgeproduct is alternating
5. Define I_n^k , y_I and x_J . $y_I(x_J) = \delta_{I,J}$.
6. Basis of $A^k(V)$ is $\{y_I\}$ (Right transversal mode)
7. $\dim(A^n(V)) = 1$. Basis
8. Determinant

Content

In the following, we let V be a vector space over \mathcal{F} .

Definition 4.1

The set of multilinear k -forms is the set $\mathcal{L}^k(V)$ of maps $\varphi: V^k \rightarrow \mathcal{F}$ which is linear in each variable separately.

For $k = 2$ it is the set $\mathcal{L}^2(V) = \text{Bil}(V, V)$. There is a special subspace of multilinear k -forms:

Definition 4.2

A k -multilinear form $w \in \mathcal{L}^k(V)$ is said to be alternating if for all (x_1, \dots, x_n) with $x_i = x_j$ for some $i \neq j$ it holds that

$$w(x_1, \dots, x_n) = 0.$$

The set of alternating k -forms is a vector space over \mathcal{F} , which we denote by $A^k(V)$

There is an action $S_k \curvearrowright \mathcal{L}^k(V)$, given by

$$\sigma w(v_1, \dots, v_k) = w(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \text{ for all } (v_1, \dots, v_k) \in V^k, w \in \mathcal{L}^k(V).$$

Definition 4.3

A multilinear k -form w is called **skew** if $\sigma w = \text{sgn}(\sigma)w$ for all $\sigma \in S_k$

If w is alternating, then w is skew:

$$\begin{aligned} 0 = w(x_1 + x_2, x_1 + x_2, x_3, \dots, x_k) &= \underbrace{w(x_1, x_1, \dots)}_{=0} + w(x_2, x_1, \dots) + w(x_1, x_2, \dots) + \underbrace{w(x_2, x_2, \dots)}_{=0} \\ &= w(x_1, x_2, \dots) = w(x_2, x_1, \dots), \end{aligned}$$

here we just used the transposition (12), but it works for any transposition hence any permutation. If $\text{char}(\mathcal{F}) \neq 2$ then skew implies alternating:

$$w(x_1, x_1, x_3, \dots, x_k) = -w(x_1, x_1, x_3, \dots, x_k) \implies w(x_1, x_1, x_3, \dots, x_k) = 0.$$

Theorem 4.4

If $w \in A^k(V)$ and (x_1, \dots, x_k) is linearly dependent, then $w(x_1, \dots, x_k) = 0$.

Proof.

Independence implies that one coordinate can be written as a sum of the others. Alternating forms are linear in one coefficient, so by definition it will be 0 on linear dependent sets. \square

This implies that $A^k(V) = \{0\}$ if $k > \dim(V)$.

Definition 4.5

If $u_1, \dots, u_k \in V'$ are any linear functionals, we define their **wedge-product** to be the k -linear form

$$u_1 \wedge \dots \wedge u_k := \sum_{\sigma \in S_k} \text{sgn}(\sigma) \sigma(u_1 \dots u_k)$$

Theorem 4.6

The wedge product is alternating, i.e., $u_1 \wedge \dots \wedge u_k \in A^k(V)$

Proof.

Let $w = u_1 \wedge \dots \wedge u_k$ and let $z_1, \dots, z_k \in V$ with $z_i = z_j$ for some $i < j$. The goal is to show that $u_1 \wedge \dots \wedge u_k(z_1, \dots, z_k) = 0$. Let $H = \{1, (ij)\} \leq S_k$ be the subgroup generated by the transposition (ij) . Pick a right transversal R of S_k/H , i.e. a set containing one representant for each right coset. Then we see that

$$G = \bigcup_{\tau \in R} [\tau]_{S_k/H}$$

so we see that

$$u_1 \wedge \dots \wedge u_k = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \sigma(u_1 \dots u_k) = \sum_{\tau \in R} \sum_{\sigma \in [\tau]} \text{sgn}(\sigma) \sigma(u_1 \dots u_k).$$

Since $\sigma \in [\tau]$ is of the form $\tau, \tau(ij)$, we see that the inner sum is

$$\begin{aligned} \sum_{\sigma \in \{\tau, \tau(ij)\}} \text{sgn}(\sigma) \sigma(u_1 \cdots u_k) &= \text{sgn}(\tau) \tau(u_1 \cdots u_k) + \underbrace{\text{sgn}(\tau(ij))}_{=-\text{sgn}(\tau)} (\tau \circ (ij)) (u_1 \cdots u_k) \\ &= \text{sgn}(\tau) (\tau(u_1 \cdots u_i \cdots u_j \cdots u_k) - \tau(u_1 \cdots u_j \cdots u_i \cdots u_k)) \end{aligned}$$

which applied to (z_1, \dots, z_k) gives

$$\begin{aligned} \sum_{\sigma \in \{\tau, \tau(ij)\}} \text{sgn}(\sigma) \sigma(u_1 \cdots u_k)(z_1, \dots, z_k) &= u_1(z_{\tau(1)}) \cdots u_i(z_{\tau(i)}) \cdots u_j(z_{\tau(j)}) \cdots u_k(z_{\tau(k)}) \\ &\quad - u_1(z_{\tau(1)}) \cdots u_i(z_{\tau(j)}) \cdots u_j(z_{\tau(i)}) \cdots u_k(z_{\tau(k)}) \\ &= 0 \end{aligned}$$

since $z_i = z_j$. □

Let $\{x_1, \dots, x_n\}$ be a basis for V and let $\{y_1, \dots, y_n\}$ be the associated dual basis. For $1 \leq k \leq n$ we define

$$I_n^k := \{I \in \{1, \dots, n\}^k : 1 \leq j \leq j' \leq k \implies I(j) < I(j')\}.$$

Then define $y_I := y_{i_1} \wedge \dots \wedge y_{i_k}$ and $x_I = (x_{i_1}, \dots, x_{i_k}) \in V^k$ for all $I = \{i_1, \dots, i_k\} \in I_n^k$.

Lemma 4.7

For $I, J \in I_n^k$ we have $y_I(x_J) = \delta_{I,J}$

Proof.

By definition

$$\begin{aligned} y_I(x_J) &= y_{i_1} \wedge \cdots \wedge y_{i_k}(x_{j_1}, \dots, x_{j_k}) = \sum_{\sigma \in S_k} \text{sgn}(\sigma) y_{i_1}(x_{j_{\sigma(1)}}) \cdots y_{i_k}(x_{j_{\sigma(k)}}) \\ &= \sum_{\sigma \in S_k} \text{sgn}(\sigma) \delta_{i_1, j_{\sigma(1)}} \cdots \delta_{i_k, j_{\sigma(k)}} \end{aligned}$$

which is non-zero only if $I = J$ and $\sigma = \text{id}$ in which case it is 1. □

Theorem 4.8

The set of $\{y_I\}_{I \in I_n^k}$ forms a basis for $A^k(V)$.

Proof.

We first prove linear independence: Assume $\sum_{I \in I_n^k} \alpha_I y_I = 0$ for some scalars α_I .

Then, for all $J \in I_n^k$ we have

$$0 = \left(\sum_{I \in I_n^k} \alpha_I y_I \right) (x_J) = \sum_{I \in I_n^k} \alpha_I \delta_{I,J} = \alpha_J$$

so it is linearly independent.

Now we prove that it spans: Let $w \in A^k(V)$, $I \in I_n^k$ and set $\alpha_I := w(x_I) \in \mathcal{F}$. Then we claim that $w = \sum_{I \in I_n^k} \alpha_I y_I$. To see this, we check on the k -tuples of basis vectors x_J , since both sides are multilinear:

Let $J \in \{1, \dots, n\}^k$. Since both w and $\sum_{I \in \{1, \dots, n\}^k} \alpha_I y_I$ are alternating, we may assume that x_J have distinct members: $j \neq j' \implies x_j \neq x_{j'}$, and since alternating forms are skew, we may assume that the order of J is increasing, i.e., that $J \in I_n^k$. Then

$$w(x_J) = \sum_{I \in I_n^k} \alpha_I y_I(x_J) = \alpha_J$$

as wanted. □

Since I_n^k has cardinality n -choose- k , we see that $\dim(A^k(V)) = \binom{n}{k}$. In particular:

Theorem 4.9

if $\dim(V) = n$ then $\dim(A^n(V)) = 1$ and it has basis $\{y_1 \wedge \dots \wedge y_n\}$.

Proof.

follows directly from the previous theorem. In particular, $w \in A^n(V) \implies w = cy_1 \wedge \dots \wedge y_n$ for some $c \in \mathcal{F}$. □

Which implies the following theorem

Theorem 4.10

If $w \in A^n(V)$ with $w \neq 0$ then $w(x_1, \dots, x_n) \neq 0$.

Proof.

By the above, $w = cy_1 \wedge \dots \wedge y_n$ for some $c \in \mathcal{F}^\times$, so

$$w(x_1, \dots, x_n) = cy_1 \wedge \dots \wedge y_n(x_1, \dots, x_n) = cy_1(x_1) \dots y_n(x_n) = c \neq 0.$$

□

This allows us to define the determinant of $A \in \text{End}(V)$ by letting $\bar{A} \in \text{End}(A^n(V))$ to be the map $\bar{A}(w)(x_1, \dots, x_n) = w(Ax_1, \dots, Ax_n)$. Since $\dim(A^n(V)) = 1$, there is $\delta \in \mathcal{F}$ such that $\bar{A}w = \delta w$ for all $w \in A^n(V)$.

Definition 4.11

The scalar δ is called the determinant of A

INVARIANCE, REDUCIBILITY AND PROJECTIONS

Agenda

Fix agenda

1. Basic definitions:
2. Lemma about basis and spanning sets
3. Characterization of spanning sets, linear independent sets and bases of vector spaces.

Content

In the following V is a vector space over \mathcal{F} . We use M, N to denote subspaces of V .

Definition 5.1

If $V = M \oplus N$, then a projection on M along N is a linear map $E: V \rightarrow V$ given by $Ez = x$ for $z = x + y$ with $x \in M$ and $y \in N$.

If E is a projection on M along N , then $1 - E$ is the projection on N along M . This is clear.

Theorem 5.2

There exists a direct sum $V = M \oplus N$ if and only if there is $E \in \text{End}(V)$ such that $E^2 = E$.

Proof.

If $V = M \oplus N$, then the projection on M along N satisfies $E(x + y) = x$ and $E^2(x + y) = Ex = x$ so $E^2 = E$. Conversely, suppose that $E^2 = E \in \text{End}(V)$. Then $M := \text{Range}(E)$ and $N := \ker(E)$ satisfy $M \cap N = \{0\}$, for if $Ez = z$ and $Ez = 0$ then $z = 0$. Also $M + N = V$, for if $z \in V$, then $z = Ez + (1 - E)z$, and $Ez \in M$ since $Ez = E(Ez) \in \text{Range}(E)$. And $E(1 - E)z = Ez - Ez = 0$ so $(1 - E)z \in \ker(E)$. \square

Definition 5.3

A subspace $M \subseteq V$ is invariant under $A \in \text{End}(V)$ if $A(M) \subseteq M$. If $V = M \oplus N$ with M and N invariant under A , then we say that A is reduced by M and N .

It is worth to note the following: If $A \in \text{End}(V)$ and $M \subseteq V$ is invariant under A , and $\{x_1, \dots, x_m, \dots, x_n\}$ is a basis for V with $\{x_1, \dots, x_m\}$ as a basis for M , then the matrix of A with respect to this basis is of the form

$$[A] = \begin{pmatrix} \bullet & \bullet \\ 0 & \bullet \end{pmatrix}$$

If moreover A is reduced by M, N with $V = M \oplus N$ and $\{x_{m+1}, \dots, x_n\}$ is a basis for N , then $[A]$ is of the form

$$[A] = \begin{pmatrix} \bullet_M & 0 \\ 0 & \bullet_N \end{pmatrix}$$

where $\bullet_m = [A|_M]$ and $\bullet_N = [A|_N]$, i.e. $[A] = [A|_M] \oplus [A|_N]$

Theorem 5.4

Let $V = M \oplus N$ with E the projection on M along N . For $A \in \text{End}(V)$ the following are equivalent:

1. A is reduced by N and M (both M, N are invariant under A)
2. $EA = AE$

Proof.

1 \implies 2: First we show that $EAz = AEz$ for $z \in M$ and then for $z \in N$:

- If $z \in M$, then as $Az \in M$ and $Ez = z$ we have $EAz = Az = AEz$.
- If $z \in N$, then $Ez = 0$ and $Az \in N$, so $AEz = A0 = 0$ and $EAz = 0$ so $EAz = AEz$.

2 \implies 1:

- If $z \in M$ then $Ez = z$ and M is the image of E so $EAz = AEz = Az \in M$.
- If $z \in N$ then $Ez = 0$ and N is the kernel of E so $EAz = AEz = A0 = 0 \in N$

□

COMBINATION OF PROJECTIONS

Agenda

Fix agenda

1. Basic definitions:
2. Lemma about basis and spanning sets
3. Characterization of spanning sets, linear independent sets and bases of vector spaces.

Content

In the following, V denotes a vector space over a field \mathcal{F} , and $M, N \subseteq V$ will denote subspaces.

Definition 6.1

If $V = M \oplus N$, then a projection on M along N is a linear map $E: V \rightarrow V$ given by $Ez = x$ for $z = x + y$ with $x \in M$ and $y \in N$.

If E is a projection on M along N , then $1 - E$ is the projection on N along M . This is clear.

Theorem 6.2

Assume that E_i is the projection onto M_i along N_i for $i = 1, 2$ so $V = M_1 \oplus N_1 = M_2 \oplus N_2$. If $E_1 E_2 = E_2 E_1$ then $V = (M_1 \cap M_2) \oplus (N_1 + N_2)$ and $E := E_1 E_2$ is the projection onto $M_1 \cap M_2$ along $N_1 + N_2$.

Proof.

Clearly E is idempotent: $E^2 = (E_1 E_2)^2 = E_1 E_2 E_1 E_2 = E_1 E_1 E_2 E_2 = E_1 E_2 = E$. We need to show that $M_1 \cap M_2$ is the range of E and $N_1 + N_2$ is the kernel.

- If $Ez = z$, then $E_1 z = E_1 Ez = E_1 E_1 E_2 z = E_1 E_2 z = Ez$, so $Ez \in M_1$, and similarly $Ez \in M_2$. If $z \in M_1 \cap M_2$, then both $E_1 z = z = E_2 z$ so $Ez = E_1 E_2 z = z$ hence z is in the image of E .
- if $z \in N_i$ then $E_i z = 0$ so $EE_i z = Ez = 0$ for $i = 1, 2$. By linearity, if $z \in N_1 + N_2$ then $Ez = 0$. If $Ez = 0$, then $E_1(E_2 z) = 0$ so $E_2 z \in N_1$, but $(1 - E_2)z \in N_2$, so $z = E_2 z + (1 - E_2)z \in N_1 + N_2$, as wanted.

and whenever we have an idempotent we may decompose our domain into range and nullspace. \square

Theorem 6.3

If $E_1E_2 = E_2E_1 = 0$, then $V = M_1 \oplus M_2 \oplus (N_1 \cap N_2)$ and $E_1 + E_2$ is the projection onto $M_1 \oplus M_2$ along $N_1 \cap N_2$

Before proving this, we will need the following lemma, which we state without proving. The idea is to consider the projections E_1E_2 , $E_1(1 - E_2)$, $(1 - E_1)E_2$ and $(1 - E_1)(1 - E_2)$.

Lemma 6.4

If $E_1E_2 = E_2E_1$ then $V = (M_1 \cap M_2) \oplus (M_1 \cap N_2) \oplus (N_1 \cap M_2) \oplus (N_1 \cap N_2)$

Proof of Theorem 6.3.

Since $E_2E_1 = 0$ we see that $M_1 \subseteq N_2$, since $M_1 = \text{Rg}(E_1)$. So $M_1 \cap N_2 = M_1$. Similarly we see that $M_2 \cap N_1 = M_2$, and by theorem 6.1 $M_1 \cap M_2$ is the image of E_1E_2 , which is 0 by assumption, i.e. $M_1 \cap M_2 = 0$. So By the lemma above, we have

$$\begin{aligned} V &= \underbrace{(M_1 \cap M_2)}_{=0} \oplus \underbrace{(M_1 \cap N_2)}_{=M_1} \oplus \underbrace{(N_1 \cap M_2)}_{=M_2} \oplus (N_1 \cap N_2) \\ &= M_1 \oplus M_2 \oplus (N_1 \cap N_2). \end{aligned}$$

Since $E_1E_2 = E_2E_1 = 0$, we see that $E_1 + E_2$ is a projection:

$$(E_1 + E_2)^2 = E_1^2 + E_2^2 + E_1E_2 + E_2E_1 = E_1 + E_2$$

And it has range in $M_1 + M_2$ and kernel $N_1 \cap N_2$, so it is indeed the projection onto $M_1 \oplus M_2$ along $N_1 \cap N_2$ \square

Proof of converse for $\text{char}(\mathcal{F}) \neq 2$ of the above.

If $E_1^2 = E_1, E_2^2 = E_2$ and $(E_1 + E_2)^2 = E_1 + E_2$, then

$$E_1 + E_2 = (E_1 + E_2)^2 = E_1 + E_2 + E_1E_2 + E_2E_1 \implies E_1E_2 = -E_2E_1,$$

And then we see that

$$E_1E_2 = E_1(E_1E_2) = E_1(-E_2E_1) = (-E_1E_2)E_1 = E_2E_1E_1 = E_2E_1 \implies 2E_1E_2 = 0$$

and since $\text{Char}(\mathcal{F}) \neq 2$, this implies $E_1E_2 = 0$. \square

RANGE, NULL-SPACE, RANK AND NULLITY

Agenda

1.

Content

In the following we let U, V be vector spaces over the same field \mathcal{F} . For $A \in \text{Hom}(U, V)$ We let $\mathcal{R}(A) = \{Ax \mid x \in U\}$ and $\mathcal{N}(A) = \{x \in U \mid Ax = 0\}$ denote the range- and nullspace of A . In particular, we note that $\mathcal{R}(A) \subseteq V$ and $\mathcal{N}(A) \subseteq U$.

Definition 7.1

Let $A \in \text{Hom}(U, V)$ and let $K \subseteq U$ and $L \subseteq V$. Suppose that A maps K into L , i.e. that $\mathcal{R}(A) \subseteq L$. Then the operator $\bar{A}: U/K \rightarrow V/L$ given by

$$\bar{A}(x + K) = Ax + L$$

is called the **quotient map** of A

Lemma 7.2

The quotient map is well-defined and linear.

Proof.

As vector spaces in particular are abelian groups, the map is clearly well-defined, i.e. if $x + K = y + K$ then $x - y \in K$ so $Ax - Ay = A(x - y) \in L$ so $Ax + L = Ay + L$. Linearity follows from linearity of A and $(a + b) + L = (a + L) + (b + L)$. \square

In particular, if $K = \mathcal{N}(A)$, then $AK = \{0\}$, so if $L = \{0\}$ in the above, then as $V/\{0\} \cong V$ we achieve a map $\bar{A}: U/\mathcal{N}(A) \rightarrow V$ given by $x + \mathcal{N}(A) \mapsto Ax$, i.e., we actually have $\bar{A}: U/\mathcal{N}(A) \rightarrow \mathcal{R}(A)$.

Theorem 7.3

For $A \in \text{Hom}(U, V)$, the map $\bar{A}: U/\mathcal{N}(A) \rightarrow \mathcal{R}(A)$ above is an isomorphism.

Proof.

\bar{A} is linear by the above lemma. We see that \bar{A} is injective, for

$$0 = \bar{A}(x + \mathcal{N}(A)) = Ax \implies x \in \mathcal{N}(A) \iff x = 0 \in U/\mathcal{N}(A).$$

Also we see that $\bar{A}(U/\mathcal{N}(A)) = A(U) = \mathcal{R}(A)$, so \bar{A} is surjective. \square

We assign to an operator $A \in \text{Hom}(U, V)$ a few more things, namely two parameters which measure how well it behaves

Definition 7.4

Let $A \in \text{Hom}(U, V)$, then the **Rank** of A is $\rho(A) := \dim(\mathcal{R}(A))$ and the **nullity** of A is $\nu(A) := \dim(\mathcal{N}(A))$

And if $\dim(U) < \infty$ then we can calculate either if we know something about the other and the domain/codomain:

Corollary 7.5

If $\dim(U) = n < \infty$ and $A \in \text{Hom}(U, V)$, then $\rho(A) = n - \nu(A)$.

Proof.

By the theorem above, $\mathcal{R}(A) \cong U/\mathcal{N}(A)$, hence $\rho(A) = \dim(\mathcal{R}(A)) = \dim(U/\mathcal{N}(A)) = \dim(U) - \nu(A) = n - \nu(A)$. \square

With this in mind we can prove the following

Theorem 7.6

Let $A \in \text{Hom}(V, W)$ and $B \in \text{Hom}(U, V)$, i.e. $A: V \rightarrow W$ and $B: U \rightarrow V$ so $AB \in \text{Hom}(U, W)$. Then

1. $\rho(AB) \leq \min \{\rho(A), \rho(B)\}$
2. $\nu(A) \leq \nu(A) + \nu(B)$

Proof.

(1): First we note that $\mathcal{R}(AB) \subseteq \mathcal{R}(A)$, so $\rho(AB) \leq \rho(A)$. Also, we see that $\mathcal{R}(AB) = \mathcal{R}(A|_{\mathcal{R}(B)})$ so $\mathcal{N}(AB) = \mathcal{N}(A|_{\mathcal{R}(B)}) = \mathcal{N}(A) \cap \mathcal{R}(B)$. By the theorem above applied to $A|_{\mathcal{R}(B)}$ we see that

$$\mathcal{R}(AB) \cong \mathcal{R}(B)/(\mathcal{N}(A) \cap \mathcal{R}(B)) \implies \rho(AB) \leq \rho(B)$$

(2): Clearly B restricted to $\mathcal{N}(AB)$ maps onto $\mathcal{R}(B) \cap \mathcal{N}(A)$, for if $x \in \mathcal{N}(AB)$, then $A(Bx) = 0$ and $\mathcal{N}(B) \subseteq \mathcal{N}(AB)$ as well, hence by the above theorem we see that

$$\mathcal{R}(B) \cap \mathcal{N}(A) \cong \mathcal{N}(AB)/\mathcal{N}(B)$$

which implies that

$$\begin{aligned}
 \nu(AB) - \nu(B) &= \dim(\mathcal{N}(AB)) - \dim(\mathcal{N}(B)) \\
 &= \dim(\mathcal{N}(AB)/\mathcal{N}(B)) \\
 &= \dim(\mathcal{R}(B) \cap \mathcal{N}(A)) \\
 &\leq \dim(\mathcal{N}(A)) \\
 &= \nu(A) \\
 \implies \dim(\mathcal{N}(AB)) &\leq \nu(A) + \nu(B)
 \end{aligned}$$

as wanted

□

TRIANGULAR FORM

Agenda

1.

Content

In the following, V is a vector space with $\dim(V) = n < \infty$ over a field \mathcal{F} .

Definition 8.1

A **flag** in V is a chain of subspaces

$$0 = M_0 \subset M_1 \subset M_2 \subset \cdots \subset M_q = V.$$

A **full flag** is a flag $(M_i)_{i=0}^q$ satisfying $\dim(M_i) = i$, which in particular implies that $q = n$. For $A \in \text{End}(V)$, a flag is said to be *A-invariant* if each subspace is *A-invariant*.

Lemma 8.2

Every basis of V gives rise to a full flag, and every full flag gives rise to a basis of V .

Proof.

Given a basis $\{x_1, \dots, x_n\}$ of V , setting $M_j := \text{Span}\{x_1, \dots, x_j\}$ we obtain a full flag $0 = M_0 \subset M_1 \subset \cdots \subset M_n = V$. Conversely, assume that $0 = M_0 \subset M_1 \subset \cdots \subset M_n = V$, we can pick $x_i \in M_i$ consecutively such that $M_i = \text{Span}\{x_1, \dots, x_i\}$ and then $V = \text{Span}\{x_1, \dots, x_n\}$. \square

Lemma 8.3

Let $A \in \text{End}(V)$. Then the matrix $[A]$ with respect to a basis $\{x_1, \dots, x_n\}$ is **upper triangular** if and only if the corresponding full flag is invariant.

Proof.

Assume that $[A]$ is upper triangular. That means that $Ax_j \subseteq \text{Span}\{x_1, \dots, x_j\} = M_j$ so $AM_j \subset M_j$. If the full flag is invariant, then $AM_j \subset M_j$ so that $Ax_j \in \text{Span}\{x_1, \dots, x_j\}$ so $[A]$ is upper triangular w.r.t the basis \square

We recall that a field \mathcal{F} is algebraically closed if every polynomial in $\mathcal{F}[x]$ has all its roots in \mathcal{F} . And that for $A \in \text{End}(V)$ we define $A' \in \text{End}(V')$ by $A'(y) = y \circ A$ for all $y \in V'$.

Theorem 8.4

If \mathcal{F} is algebraically closed, then every $A \in \text{End}(V)$ admits a full invariant flag.

Proof.

Induction on the dimension $\dim(V) = n$: For $n = 0$ there is nothing to show, and the same is true for $n = 1$.

Assume that $n > 1$ and that it is true for $n - 1$. Since \mathcal{F} is algebraically closed, the characteristic polynomial $\chi_{A'}$ has a root in \mathcal{F} , i.e., an eigenvalue of A' and denote it by λ . Pick a non-zero eigenvector $y \in V'$ associated to the eigenvalue λ . Then $\dim(\mathcal{R}(y)) = 1$ since $y \neq 0$. By the fundamental theorem, $V/\mathcal{N}(y) \cong \mathcal{R}(y)$, so $\dim(\mathcal{N}(y)) = n - 1$. So let $M_{n-1} = \mathcal{N}(y)$. All we need to show is that M_{n-1} is A -invariant:

$$x \in M_{n-1} \iff y(x) = 0 \implies 0 = \lambda y(x) = A'y(x) = y(Ax) \implies Ax \in M_{n-1}.$$

Our induction hypothesis now ensures the existence of a full invariant flag

$$0 = M_0 \subset M_1 \subset \cdots \subset M_{n-1} \subset M_n = V$$

□

We can now show the following

Theorem 8.5

If \mathcal{F} is algebraically closed, then for every $A \in \text{End}(V)$ there is a basis for V such that $[A]$ is upper triangular.

Proof.

Given A , use the above theorem to obtain a full invariant flag. Use the lemma to obtain a basis for this full flag. Then $[A]$ is upper triangular with respect to this basis. □

We recall the definition of the characteristic polynomial

Definition 8.6

For $A \in \text{End}(V)$, the characteristic polynomial $\chi_A(x) := \det([A] - x)$ is a polynomial in \mathcal{F} , and since \det is independent of basis choice, so is the characteristic polynomial.

Theorem 8.7

When $[A]$ is upper triangular, the diagonal elements $\alpha_{11}, \dots, \alpha_{nn}$ are the eigenvalues of A , each occurring according to its algebraic multiplicity.

Proof.

If $[A]$ is triangular, then the determinant $\det([A])$ is the product of the diagonal elements of $[A]$. If $[A]$ is triangular, then $[A] - x$ is triangular, and so

$$\chi_A(x) = (\alpha_{11} - x) \cdots (\alpha_{nn} - x)$$

hence the roots of χ is the diagonal values of $[A]$. □

CHAPTER 9

NILPOTENCE

Agenda

fix agenda

1.

Content

In the following, we let V be a vector space with $\dim(V) = n < \infty$ over a field \mathcal{F} .

Definition 9.1

An endomorphism $A \in \text{End}(V)$ is nilpotent if there is $q \geq 0$ such that $A^q = 0$. The index is the smallest integer q such that $A^q = 0$ but $A^{q-1} \neq 0$.

We wish to characterise all nilpotent maps in the following. We will do this by the following theorem:

Theorem 9.2

Let $\dim(V) < \infty$ and $A \in \text{End}(V)$ nilpotent. Then there is $r \geq 0$ and integers $q_1 \geq q_2 \geq \dots \geq q_r > 0$ and $x_1, \dots, x_r \in V$ such that:

1. The vectors $A^j x_i$ where $\{0 \leq j < q_i\}$ for $1 \leq i \leq r$ gives a basis of V . In particular we note that this implies that $q_1 + q_2 + \dots + q_r = n$
2. $A^{q_i} x_i = 0$ for $1 \leq i \leq r$

to prove this theorem, we need the following:

Theorem 9.3

Assume $\dim(V) < \infty$ and $A \in \text{End}(V)$ of index q . Let $x_0 \in V$ satisfy $A^{q-1} x_0 \neq 0$ (which assume by definition of q , i.e. $A^{q-1} \neq 0$) Then

1. The vectors $x_0, Ax_0, \dots, A^{q-1} x_0$ are linearly independent.
2. The space $H = \text{Span}\{x_0, Ax_0, \dots, A^{q-1} x_0\}$ is A invariant.
3. There exists an A -invariant subspace K such that $V = H \oplus K$

It is not hard to see that the above theorem will imply the proof of the other theorem:

Proof of theorem 9.2.

First let $q := q_1$ and write $V = H_1 \oplus K$. Then clearly $A|_K$ is nilpotent of index $q_2 \leq q_1$. Apply the theorem again and again. It will terminate since $\dim(V) < \infty$. \square

We now prove the needed theorem:

Proof of theorem 9.3.

Recall that A have nilpotence index q and x_0 witnessed this: $A^{q-1}x_0 \neq 0$.

(1): Assume that the set $\{x_0, Ax_0, \dots, A^{q-1}x_0\}$ is linearly dependent. This means that there is $k \leq q_1$ (if $k = q - 1$ then the span is 0) such that

$$A^{q-1} \in \text{Span}\{A^{k+1}, \dots, A^{q-1}x_0\}$$

Then $A^k x_0 = \alpha_{k+1} A^{k+1} x_0 + \dots + \alpha_{q-1} A^{q-1} x_0$. Applying A^{q-1-k} to this we get $A^{q-1} x_0 = 0$, contradicting our assumption.

(2): This is clear, as $A(\text{Span}\{x_0, \dots, A^{q-1}x_0\}) = \text{Span}\{Ax_0, \dots, A^{q-1}x_0, 0\} = \text{Span}\{Ax_0, \dots, A^{q-1}x_0\}$.

(3): Now $H := \text{Span}\{x_0, \dots, A^{q-1}x_0\}$. We wish to find $K \subseteq V$ such that $A(K) \subseteq K$ and $V = H \oplus K$. We do so by induction on q : for $q = 1$ we have $A = 0$ so every subspace is invariant.

Assume it holds for all nilpotent of index $< q$. Set $R := \mathcal{R}(A)$. Then R is A -invariant and $A|_R$ is nilpotent of index $q - 1$, so we may apply the induction hypothesis to $y_0 = Ax_0 \in R$ to obtain decomposition $R = H_0 \oplus K_0$ where $H_0 = \text{Span}\{y_0, Ay_0, \dots, A^{q-2}y_0\} = A(H)$, since $y_0 = Ax_0$ and $K_0 \subset R$ is some invariant subspace.

Set $K_1 := A^{-1}(K_0)$ to be the preimage of K_0 under A . Then $K_0 \subset K_1$, since if $x \in K_0$ then invariance implies $Ax_0 \in K_0$ so $x \in A^{-1}\{Ax_0\} \subseteq A^{-1}(K_0)$. We then have

- $H + K_1 = V$: For if $x \in V$ then $Ax \in R = H_0 \oplus K_0$ so $Ax = y + z$ for some $y \in H_0, z \in K_0$. Since $H_0 = A(H)$, there is $y_1 \in H$ so $y = Ay_1$. Then $A(x - y_1) = Ax - y = z \in K_0$ so $x - y_1 \in K_1$ so $x = y_1 + (x - y_1) \in H + K_1$.
- $H \cap K_1 = \text{Span}\{A^{q-1}x_0\}$: Let $x \in H \cap K_1$, then

$$Ax \in A(H \cap K_1) \subseteq A(H) \cap A(K_1) \subseteq H_0 \cap K_0 = \{0\}.$$

In particular, since $\{x_0, \dots, A^{q-1}x_0\}$ is linearly independent, this implies in particular that $x \in \text{Span}\{A^{q-1}x_0\}$, since $A(\sum \alpha_i A^i x_0) = 0$ implies that $\alpha_i = 0$ for all i not equal $q-1$ (Or equivalently, $H \cap \mathcal{N}(A) = \text{Span}\{A^{q-1}x_0\}$).

- $H \cap K_0 = \{0\}$: Since $K_0 \subset K_1$ and $A^{q-1}x_0 \notin K_0$, since $\mathcal{R}(A) = K_0 \oplus H_0$ and $A^{q-1}x_0 \in H_0$, so the statement holds.
- Let now K be any subspace of K_1 of codimension 1 containing K_0 but not $A^{q-1}x_0$. Then $H \oplus K = V$ and K is invariant under A :

$$A(K) \subset A(K_1) \subset K_0 \subset K$$

This finishes the proof. □

uniqueness

CHAPTER 10

JORDAN NORMAL FORM

Agenda

1.

Content

In the following, let V be a vector space over a field \mathcal{F} with $\dim(V) = n < \infty$. We wish to classify all $A \in \text{End}(V)$ for algebraically closed \mathcal{F} to be of the form (for a suitable base):

$$[A] = \begin{pmatrix} \bullet & & 0 \\ & \ddots & \\ 0 & & \bullet \end{pmatrix}$$

Where $\bullet = \begin{pmatrix} \lambda_i & & 0 \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix} + [\text{Nilpotent}]$ and λ_i is an eigenvalue, and each bullet correspond to one eigenvalue, and the size of the bullets is the algebraic multiplicity of the eigenvalue. To do so, we prove the following

Theorem 10.1

Let $A \in \text{End}(V)$ with spectrum $\sigma(A) = \{\lambda_1, \dots, \lambda_p\}$ (distinct). Then there exists a unique invariant subspace R such that

$$V = M_{\lambda_1} \oplus \dots \oplus M_{\lambda_p} \oplus R,$$

Where $M_{\lambda_i} = \bigcup_{k \geq 0} \mathcal{N}((A - \lambda_i)^k)$ are the generalized eigenspaces, and where $\chi_A|_R$ has no roots in \mathcal{F} .

The proof of this requires the following theorem

Theorem 10.2

For $A \in \text{End}(V)$, $\dim(V) < \infty$ there exists a unique reduction $V = R \oplus N$ such that $A|_R \in \text{End}(R)$ is invertible and $A|_N \in \text{End}(N)$ is nilpotent.

The proof follows nicely from the following two lemmas:

Lemma 10.3

Set $N := \{x \in V \mid \exists k > 0 : A^k x = 0\} = \bigcup_{k>0} \mathcal{N}(A^k)$. Then N is A -invariant and the restriction $A|_N \in \text{End}(N)$ is nilpotent, and every invariant subspace N' such that $A|_{N'} \in \text{End}(N')$ is nilpotent is contained in N

and

Lemma 10.4

Let $R = \{x \in V \mid \forall k > 0 : x \in \mathcal{R}(A^k)\} = \bigcap_{k>0} \mathcal{R}(A^k)$. Then R is A -invariant and the restriction $A|_R \in \text{End}(R)$ is invertible. Moreover, every invariant subspace R' such that $A|_{R'} \in \text{End}(R')$ is invertible is contained in R .

We prove both of them at the same time:

Proof of lemmas:

Set $N_k := \mathcal{N}(A^k)$ and $R_k := \mathcal{R}(A^k)$ for $k \in \mathbb{N}$. Then we obtain

$$\begin{aligned} 0 &= N_0 \subset N_1 \subset \dots \subset N_k \subset \dots \\ V &= R_0 \supset R_1 \supset \dots \supset R_k \supset \dots \end{aligned}$$

where $A(N_k) \subset N_{k-1} \subset N_k$ and $A(R_k) \subset R_{k+1} \subset R_k$ so they are invariant subspaces. Since $\dim(V)$ is finite, these stabilize. So there is $q > 0$ such that $N_k = N_q$ and $R_k = R_q$ for $k \geq q$. Then $N = N_q$ and $R = R_q$ are invariant subspaces.

Clearly $(A|_N)^q = 0$, so it is nilpotent. We also see that $A(R) = R_{q+1} = R_q$ so $A|_R \in \text{End}(R)$ is surjective and hence invertible.

uniqueness of N follows from if $A|_{N'}$ is nilpotent and N' invariant subspace, then $N' \subset N_k$ where k is the nilpotence index of $A|_{N'}$, so $N' \subseteq N$.

uniqueness of R follows from: If $A|_{R'} \in \text{End}(R')$ is invertible, then all powers of it is invertible, hence $R' \in R_k$ for all $k > 0$ so $R' \subset R$. \square

We can now prove Theorem 10.2

Proof of Theorem 10.2.

Let R and N as in the lemmas. Then invertibility of $A|_R$ implies that $A^k|_R$ is invertible for all k , so $A^k|_{R \cap N_k}$ is invertible for all k , but $A^k|_{N_k} = 0$ so $R \cap N_k = 0$ for all k so $R \cap N = 0$. Now, let q be such that $R = \mathcal{R}(A^q)$ and $N = \mathcal{N}(A^q)$. Then $R \cong V/N$ so $\dim(R) = \dim(V) - \dim(N) \iff \dim(R) + \dim(N) = \dim(V)$. Since their intersection is empty $R \oplus N \subset V$. Since their dimension is the same as V 's: $V \subset R \oplus N$. \square

We can now prove the main theorem:

Proof of Theorem 10.1.

Let $V = M_i \oplus R_i$ be reductions for $A - \lambda_i$. If $i \neq j$, then

$$\begin{aligned} \chi_{(A-\lambda_j)|_{M_i}}(x) &= \chi_{(A-\lambda_j+\lambda_j-\lambda_i)|_{M_i}}(x + \lambda_j - \lambda_i) \\ &= \chi_{(A-\lambda_i)|_{M_i}}(x + \lambda_j - \lambda_i) \\ &= (-x + \lambda_i - \lambda_j)^{m_i} \end{aligned}$$

where m_i is the algebraic multiplicity of λ_i . In particular, when evaluated at 0 we obtain the determinant of $\det((A - \lambda_j)_{M_i}) = (\lambda_i - \lambda_j)^{m_i} \neq 0$, so it is invertible. By the previous lemma, since M_i is invariant under A it is also invariant under $A - \lambda_j$, we see that $M_i \subseteq R_j$ (invertible and invariance).

Now, let E_i, E_j be the projections on M_i and M_j along R_i and R_j , respectively. As $M_j \subseteq R_i$ and $M_i \subseteq R_j$ we have

$$E_i E_j = 0 = E_j E_i$$

and $E_i + E_j$ is the projection on $M_i \oplus M_j$ along $R_i \cap R_j$ and $V = M_i \oplus M_j \oplus (R_i \cap R_j)$. This holds pairwise for all $i \neq j$, and the theorem of projections generalises, so we obtain $R := R_1 \cap \dots \cap R_p$ and

$$V = \bigoplus_{i=1}^p M_i \oplus R$$

and hence

$$\chi_A = \prod_{i=1}^p \left(\chi_{A|_{M_i}} \right) \chi_{A|_R} = \prod_{i=1}^p (\lambda_i - x)^{m_i} \chi_{A|_R}$$

as wanted. □

The jordan normal then follows since if \mathcal{F} is algebraically closed, then $V = M_{\lambda_1} \oplus \dots \oplus M_{\lambda_p}$ and

$$A|_{M_{\lambda_i}} = (A - \lambda_i + \lambda_i)|_{M_{\lambda_i}} = \lambda_i + \underbrace{(A - \lambda_i)|_{M_{\lambda_i}}}_{\text{nilpotent of index } m_i}$$

PERPENDICULAR PROJECTIONS

Agenda

1.

Content

In the following we let V be a vector space over $\mathcal{F} = \mathbb{R}$ or $\mathcal{F} = \mathbb{C}$. If $V = M \oplus N$ then the map $z = x + y \mapsto x \in M$ is the projection on M along N . This happens if and only if $E \in \text{End}(V)$ satisfy $E = E^2$ and $M = \mathcal{R}(E)$ and $N = \mathcal{N}(E)$. The map $1 - E$ is the projection on N along M in this case, and E is a projection if and only if $1 - E$ is a projection.

If $\dim(V) < \infty$, then every subspace M of V satisfy $V = M \oplus M^\perp$.

Definition 11.1

A **perpendicular projection** is a projection E such that $\mathcal{R}(E) \perp \mathcal{N}(E)$.

And we have a lemma

Lemma 11.2

If $V = M + N$ and $M \perp N$ then $V = M \oplus N$ and $N = M^\perp$

Proof.

From $M \perp N$ we see that $N \subset M^\perp$ by definition. If $x \perp x$ then $x = 0$, since $\langle x, x \rangle = 0 \iff x = 0$ so $M \cap M^\perp = 0$.

This implies that $M \cap N = 0$, so $V = M \oplus N$. We need $M^\perp \subset N$: If $z \in M^\perp$, then $z = x + y$ for $x \in M$ and $y \in N$, hence $x = y - z \in M \cap M^\perp = 0$ so $z = y \in N$ as wanted. \square

We can now completely classify when a projection is a perpendicular projection, by the following theorem:

Theorem 11.3

Let $V = M \oplus N$ and let E be the projection on M along N . Then the following are equivalent:

1. E is self-adjoint
2. E is a perpendicular projection ($\mathcal{R}(E) \perp \mathcal{N}(E)$).
3. $\|Ez\| \leq \|z\|$ for all $z \in V$

If these holds, then E is positive (self-adjoint and $\langle Ez, z \rangle \geq 0$ for all $z \in V$) and “=” holds in (3) only if $Ez = z$.

Proof.

1 \implies 2: Let $z \in V$. Then $\langle Ez, (1 - E)z \rangle = \langle z, E(1 - E)z \rangle = \langle z, 0 \rangle = 0$ so $\mathcal{R}(E) \perp \mathcal{N}(E)$ as wanted.

2 \implies 1. Let $z, w \in V = \mathcal{R}(E) \oplus \mathcal{N}(E)$. Then $z = Ez + (1 - E)z$ and $w = Ew + (1 - E)w$ and by assumption $(1 - E)z \perp Ew$ and $Ez \perp (1 - E)w$, so

$$\langle Ez, w \rangle \langle Ez, Ew \rangle = \langle z, Ew \rangle \implies E^* = E.$$

2 \implies 3: By pythagoras and the assumption $Ez \perp (1 - E)z$:

$$\|z\|^2 = \|Ez + (1 - E)z\|^2 = \|Ez\|^2 + \|(1 - E)z\|^2 \geq \|Ez\|^2.$$

3 \implies 2 Let $x \in M$ and $y \in N$. For all $t \in \mathbb{R}$ we get

$$\|x + ty\|^2 = \|x\|^2 + 2t\operatorname{Re}\langle x, y \rangle + t^2\|y\|^2 \geq \|x\|^2$$

which implies

$$f(t) = t^2\|y\|^2 - 2t\operatorname{Re}\langle x, y \rangle \geq 0$$

It has two roots, $t = 0$ and $t = -2\operatorname{Re}\langle x, y \rangle / \|y\|^2$ and between these roots it take negative values if $\operatorname{Re}\langle x, y \rangle \neq 0$, contradicting $f(t) \geq 0$. Hence $\operatorname{Re}\langle x, y \rangle = 0$. If $\mathcal{F} = \mathbb{R}$ we are done. If $\mathcal{F} = \mathbb{C}$ then consider iy in the above. Then we get

$$\begin{aligned} \langle x + ity, x + ity \rangle &= \|x\|^2 + \langle x, ity \rangle + \langle ity, x \rangle + t^2\|y\|^2 \\ &= \|x\|^2 + i(-t\langle x, iy \rangle + t\langle iy, x \rangle) + t^2\|x\|^2 \\ &= \|x\|^2 + 2t\operatorname{Im}(\langle x, y \rangle) + t^2\|y\|^2 \\ &\geq \|x\|^2 \end{aligned}$$

again implying that also $\operatorname{Im}\langle x, y \rangle = 0$. So $\mathcal{R}(E) \perp \mathcal{N}(E)$. □

We end with an example: If

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

then A is idempotent, but $A \neq A^*$. The range of A is the diagonal $\{(x, x)\}$ and the kernel is $\{(0, x)\}$ and these are not perpendicular.

SPECTRAL THEOREM

In the following V is a vector space over $\mathcal{F} = \mathbb{R}$ or $\mathcal{F} = \mathbb{C}$ of dimension $\dim(V) = n < \infty$ and $A \in \text{End}(V)$. For each $\lambda \in \sigma(A)$ we denote by V_λ the associated eigenspace, i.e., $V_\lambda = \mathcal{N}(A - \lambda)$ and we denote by E_λ the orthogonal projection hereon. Then we have the following theorem

Theorem 12.1 (Spectral theorem for self-adjoint operators)

Assume $A = A^*$ and $\dim(V) < \infty$. Then the following holds:

1. $\sigma(A) \subseteq \mathbb{R}$
2. $E_\lambda E_\mu = 0$ for $\lambda \neq \mu \in \sigma(A)$.
3. $\sum_{\lambda \in \sigma(A)} E_\lambda = 1$
4. $A = \sum_{\lambda \in \sigma(A)} \lambda E_\lambda$ (spectral resolution)
5. If p is a polynomial then $p(A) = \sum_{\lambda \in \sigma(A)} p(\lambda) E_\lambda$
6. For each $\lambda \in \sigma(A)$ there exists a polynomial p_λ such that $E_\lambda = p_\lambda(A)$
7. If $B \in \text{End}(V)$ then $BA = AB$ if and only if for all $\lambda \in \sigma(A)$ we have $BE_\lambda = E_\lambda B$.

Proof.

(1): If $\mathcal{F} = \mathbb{R}$ nothing to prove. If $\mathcal{F} = \mathbb{C}$ then let $\lambda \in \sigma(A)$ and let x be any eigenvector associated to λ . Without loss of generality assume $\|x\| = 1$. Then $\langle Ax, x \rangle = \langle \lambda x, x \rangle = \lambda$ and so

$$\lambda = \langle Ax, x \rangle = \langle x, A^* x \rangle = \langle x, Ax \rangle = \overline{\langle Ax, x \rangle} = \bar{\lambda}$$

so $\lambda \in \mathbb{R}$.

(2): If $\lambda \neq \mu \in \sigma(A)$ and $x \in V_\lambda$ and $y \in V_\mu$ then

$$\lambda \langle x, y \rangle = \langle Ax, y \rangle = \langle x, Ay \rangle = \bar{\mu} \langle x, y \rangle = \mu \langle x, y \rangle$$

and since $\mu \neq \lambda$ this implies $\langle x, y \rangle = 0$ so $V_\lambda \perp V_\mu$.

(3): Let $E := \sum_{\lambda \in \sigma(A)} E_\lambda$. This is an orthogonal projection since $E_\lambda \perp E_\mu$ for $\lambda \neq \mu \in \sigma(A)$ by (2). It is the projection onto $M := \bigoplus_{\lambda \in \sigma(A)} \mathcal{R}(E_\lambda)$ along $N = M^\perp$, and M is A invariant so N is A^* invariant, hence A invariant since $A^* = A$. All

eigenvectors of A are in M , and $A|_N$ is self-adjoint so every eigenvalue is real. If N is not 0, this would contradict the fact that $A^* = A \implies \sigma(A) \subseteq \mathbb{R}$ for both $\mathcal{F} = \mathbb{C}$ or $\mathcal{F} = \mathbb{R}$ which would give an eigenvector of A not contained in M . Hence $N = 0$ and $E = 1$.

(4): Apply A to E from both sides to see the desired result: $AE_\lambda x = \lambda E_\lambda x$ for all $\lambda \in \sigma(A)$ and $x \in V$, and $\sum_{\lambda \in \sigma(A)} E_\lambda = 1$ finishes the proof.

(5): We note that for any power $A^n = \sum_{\lambda \in \sigma(A)} \lambda^n E_\lambda$, as seen for $n = 2$ since $A^2 = \sum_{\lambda, \mu \in \sigma(A)} \lambda \mu E_\lambda E_\mu = \sum_{\lambda \in \sigma(A)} \lambda^2 E_\lambda$ since each eigenprojection is perpendicular to the others.

(6): Note that the lagrangian interpolation polynomial exists: given data points $(x_1, y_1), \dots, (x_n, y_n)$ there is a polynomial $F(x)$ such that $F(x_i) = y_i$ and it is the sum of polynomials $\ell_{y_i}(x)$ satisfying $\ell_{y_i}(x_j) = \delta_{ij}$, so setting $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$ the lagrangian polynomials for $(\lambda_1, \lambda_n), \dots, (\lambda_n, \lambda_n)$ satisfy $\ell_{\lambda_i}(A) = \ell_{\lambda_i}(\lambda_i) E_{\lambda_i} = E_{\lambda_i}$, the result follows from (5).

(7): by (4) we see that $BE_\lambda = E_\lambda B$ implies $BA = AB$. By 5 and 6, we see that, with p_λ being the polynomial in (6) that

$$Bp_\lambda(A) = B \sum_{\lambda \in \sigma(A)} p(\lambda) E_\lambda = BE_\lambda$$

but clearly $AB = BA$ implies $Bp_\lambda(A) = p_\lambda(A)B$. The claim follows now. \square

POLAR DECOMPOSITION

Agenda

1.

Content

We begin by noting the following: Given any set of values $(x_1, y_1), \dots, (x_n, y_n) \in \mathbb{C}^2$ the polynomial

$$f_k(t) := y_k \prod_{\substack{1 \leq j \leq n \\ j \neq k}} \frac{t - x_j}{x_k - x_j}$$

satisfy $f_k(x_j) = y_k \delta_{k,j}$. In the following V is a complex vector space (i.e. over the field $\mathcal{F} = \mathbb{C}$) of dimension $\dim(V) = n < \infty$. Any operator $A \in \text{End}(V)$ satisfies

$$A = \frac{A + A^*}{2} + i \frac{A - A^*}{2i}$$

and each part is self-adjoint. The spectral theorem for self-adjoint operators then ensure, together with the fact, that for normal operators A that $\text{Re}(A)$ and $\text{Im}(A)$ commute, so the spectral theorem for self-adjoint operators extend to the class of normal operators.

Then, any function which is defined on the spectrum $\sigma(A)$ for normal operators A will give rise to a well-defined element $f(A)$, since we may just use the interpolation polynomial above on the pairs $(\lambda, f(\lambda))$ for $\lambda \in \sigma(A)$. In particular, for any operator A , the element A^*A is positive and self-adjoint, so all its spectrum is contained in \mathbb{R}_+ , so we may define $\sqrt{A^*A}$ by the spectral theorem.

Theorem 13.1

If $A \in \text{End}(V)$ where V is a finite dimensional inner-product space, then there is a unique positive P and there is an isometry U such that $A = UP$. If A is invertible then U is also uniquely determined by A .

Proof.

We show first for A invertible: Since A^*A is positive, the element $P := \sqrt{A^*A}$ is the unique positive square root. Set $V = PA^{-1}$. Then $VA = P$, and we need to show

that V is an isometry, so $V^* = V^{-1}$ and setting $U := V^{-1}$ would satisfy $A = UP$. We see that

$$V^* = (A^{-1})^* P^* = (A^*)^{-1} P$$

so

$$V^* V = (A^*)^{-1} P P A^{-1} = (A^*)^{-1} \sqrt{A^* A} \sqrt{A^* A} A^{-1} = (A^*)^{-1} A^* A A^{-1} = 1$$

so V is an isometry. Assume now that $UP = U_0 P_0$. Then $PU^* = P_0 U_0^*$, so

$$P^2 = PU^* UP = P_0 U_0^* U_0 P_0 = P_0^2$$

And the transformation $P \mapsto P^2 = P_0^2$ has only one positive square root, $P = P_0$. If A is invertible, then P is invertible, since $P = U^{-1} A$, so $UP = U_0 P_0 = U_0 P \implies U = U_0$.

Now let $A \in \text{End}(V)$ be any operator not necessarily invertible. Set $P = \sqrt{A^* A}$ to be the positive square root. Letting $V_1 := \mathcal{N}(A)^\perp$ and $V_2 := \mathcal{R}(A)$, we see that $A|_{V_1} \in \text{Hom}(V_1, V_2)$ is surjective and hence bijective. We now have the following: $V = \mathcal{N}(A) \oplus \mathcal{N}(A)^\perp = \mathcal{R}(A)^\perp \oplus \mathcal{R}(A)$ and $\mathcal{N}(A)^\perp \cong \mathcal{R}(A)$ so $\mathcal{N}(A) \cong \mathcal{R}(A)^\perp$. By the above, there is $P_0 \in \text{End}(V_1)$ positive and isometry $U_0 \in \text{Hom}(V_1, V_2)$ such that $A|_{V_1} = U_0 P_0$. Pick now any isometry $U_2 \in \text{Hom}(V_1^\perp, V_2^\perp)$ then

$$U := U_1 \oplus U_2 \in \text{End}(V), \quad P := P_0 \oplus 0 \in \text{End}(V)$$

satisfy U is an isometry and P is positive and $A = UP$ for $Ax = Ax \oplus 0 = U_1 P_0 x \oplus 0 = UPx$. \square

We can by the above describe exactly when an operator is normal:

Theorem 13.2

Let $A \in \text{End}(V)$ with polar decomposition $A = UP$. Then A is normal if and only if $UP = PU$.

Proof.

$A^* A = A A^*$ if and only if $PU^* UP = U P P U^*$ and $PU^* UP = P^2$ if and only if $P^2 = U P^2 U^{-1}$ if and only if $P^2 U = U P^2$ if and only if $PU = UP$ by the spectral theorem for self-adjoint (e.g. positive as P) \square

NORMS OF LINEAR TRANSFORMATIONS

For a normed vector space V we say that $A \in \text{End}(V)$ is bounded if there is $K \geq 0$ such that for all $x \in V$ we have $\|Ax\| \leq K\|x\|$. In this case, we define

$$\|A\| := \sup_{\|x\|=1} \{\|Ax\|\} = \sup_{\|x\| \neq 0} \left\{ \frac{\|Ax\|}{\|x\|} \right\}.$$

The set of all bounded maps is denoted by $B(V)$.

Definition 14.1

The spectral radius of $A \in B(V)$ is the number $\rho := \sup \{|\lambda|, \lambda \in \sigma(A)\}$.

Theorem 14.2

Let $\dim(V) < \infty$ and let $\langle \cdot, \cdot \rangle$ be an inner-product on V . Then every $A \in B(V)$ normal satisfy

$$\rho = \sup \{|\langle Ax, x \rangle|, \|x\| = 1\} = \|A\|.$$

Proof.

Let $\gamma := \sup_{\|x\|=1} \{|\langle Ax, x \rangle|\}$. If $\|x\| = 1$ is any normalized eigenvector for an eigenvalue $\lambda \in \sigma(A)$ then $|\lambda| = |\lambda|\langle x, x \rangle| = |\langle Ax, x \rangle|$ so $\rho \leq \gamma$. We recall that for inner product spaces, the operator norm is also determined by the expression $\|A\| = \sup_{\|x\|=\|y\|=1} \{|\langle Ax, y \rangle|\}$, so in particular $\gamma \leq \|A\|$. We now have

$$\rho \leq \gamma \leq \|A\|.$$

Now, since A is normal, it is of the form $A = \sum_{\lambda \in \sigma(A)} \lambda E_\lambda$ where E_λ are the eigenspace projections on the eigenspaces of A . So

$$\|Ax\|^2 = \sum_{\lambda \in \sigma(A)} \|\lambda E_\lambda x\|^2 \leq \sum_{\lambda \in \sigma(A)} \rho^2 \|E_\lambda x\|^2 = \rho^2 \|x\|^2$$

so $\|A\| \leq \rho$, finishing the proof. □

We can now characterise all the eigenvalues of a normal operator by the following

Theorem 14.3

Let $A = A^* \in \text{End}(V)$ with $\dim(V) = n < \infty$. Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ be the eigenvalues of A (repetitions the number of alg. multiplicity of each eigenvalue). Define for each subspace $M \subseteq V$ the number $\mu(M) = \max_{\substack{\|x\|=1 \\ x \in M}} \{\langle Ax, x \rangle\}$. Then it holds that

$$\lambda_k = \mu_k := \inf \{\mu(M) : \dim(M) = n - k + 1\}.$$

Proof.

Since A is self-adjoint, we may by the spectral theorem pick a basis of orthonormal vectors x_1, \dots, x_n such that

$$Ax_i = \lambda_i x_i.$$

Set now $M_k := \text{Span}\{x_1, \dots, x_k\}$. We now show that $\lambda_k \leq \inf\{\mu(M) : \dim(M) = n - k + 1\}$. So let $M \subseteq V$ be a subspace of dimension $\dim(M) = n - k + 1$. Then $\dim(M_k \oplus M) = k + n - k + 1 = n + 1 > n = \dim(V)$ so $M \cap M_k \neq \{0\}$. By self-adjointness, the set $\{\langle Ax, x \rangle : \|x\| = 1\}$ is the convex hull of $\sigma(A)$, and since $A|_{M_k}$ is normal: $\{\langle Ax, x \rangle : x \in M_k\} = \text{conv}\left(\sigma\left(A|_{M_k}\right)\right) = [\lambda_k, \lambda_1]$ since $\sigma(A|_{M_k}) = \{\lambda_1, \dots, \lambda_k\}$ and $\lambda_k \leq \lambda_1$. In particular, we infer that

$$\langle Ax, x \rangle \geq \lambda_k \text{ for } x \in M_k \cap M, \|x\| = 1$$

And so the number $\mu(M) = \max\{\langle Ax, x \rangle : \|x\| = 1, x \in M\} \geq \lambda_k$, and since M was arbitrary this shows that the inf of such spaces satisfy this as well, i.e. $\lambda_k \leq \mu_k$.

To show the other implication, we note that if one space of dimension $n - k + 1$ satisfy $\mu(M) = \lambda_k$ then we are done, since the inf is greater than or equal to it. Let $M = \text{Span}\{x_k, \dots, x_n\}$. This is an A invariant space of dimension $n - k + 1$ and $\sigma(A|_M) = \{\lambda_k, \dots, \lambda_n\}$ so the set

$$\{\langle Ax, x \rangle : \|x\| = 1, x \in M\} \subseteq [\lambda_n, \lambda_k]$$

Moreover $\mu(M) = \lambda_k$, since $\langle Ax_k, x_k \rangle = \lambda_k$ and $x_k \in M$.

□

We now turn to discuss different properties of normed vector spaces and linear maps between them:

Definition 14.4

The norm topology on $B(V)$ is the topology generated by the operator norm.

Theorem 14.5

If V is a banach space, then so is $B(V)$: If any cauchy sequence converges in V then so will any cauchy sequence in $B(V)$ since $\|Ax\| \leq \|A\|\|x\|$.

Proof.

Set $Ax := \lim_n A_n x$ for each $x \in V$. Since A_n is cauchy, $\|A_n\| \leq K$ for some K . So $\|Ax\| = \lim \|A_n x\| \leq K\|x\|$. Clearly A is linear. And $A_n \rightarrow A$ for if $\varepsilon > 0$ then there is $N \geq 0$ such that $\|A_n - A_m\| < \varepsilon$ for $m, n \geq N$. Then

$$\|(A_n - A_m)x\| \leq \varepsilon\|x\|$$

for $m, n \geq N$ and hence

$$\|(A_n - A)x\| = \lim_m \|(A_n - A_m)x\| \leq \varepsilon \|x\|$$

for $n \geq N$. So $A_n \rightarrow A$. □

We equip $B(V)$ and V with different topologies.

Definition 14.6

The weak topology on V is the topology of point-wise convergence generated by V^ : $x_n \rightarrow x$ weakly if $f(x_n) \rightarrow f(x)$ for all $f \in V^*$. For finite dimensional vector spaces with inner product this is equivalent to $\langle x_n, y \rangle \rightarrow \langle x, y \rangle$ for all $y \in V$ by the riesz representation theorem.*

Definition 14.7

The weak operator topology is the topology generated by the family of functionals $F_{x,y} : T \mapsto \langle Tx, y \rangle$ for all $x, y \in V$. We say that $T_n \rightarrow T$ in the weak operator topology if and only if $\langle T_n x, y \rangle \rightarrow \langle Tx, y \rangle$ for all $x, y \in V$. It is the weakest such topology making the maps $F_{x,y}$ continuous for all x, y .

Definition 14.8

The strong operator topology is the topology generated by the family of semi-norms $F_x : T \mapsto \|Tx\|$. We say that $T_n \rightarrow T$ in the strong operator topology if and only if $\|T_n x - Tx\| \rightarrow 0$ for all $x \in V$. This is the weakest topology making the maps F_x continuous.

If $\dim(V) < \infty$ then the topologies agree, by an argument of hilbert-schmidt norms.