MATH0014 Algebra 3: Further Linear Algebra Notes (Part 1 of 2)

Based on the 2019 autumn lectures by Dr I Strouthos

ALGEBRA 3

Bárbara Nieto Aguirre

MATH 0014 FURTHER LINEAR ALGEBRA

Contact email: i.strouthos@ucl.ac.uk

office hours for term 1: Room 501, 25 Gordon Street (weekly except reading week)

Tuesday: 3.30pm - 5:30pm

Fridays: 3.30pm - 5:30pm

First coursework submission deadline - 25/10/19

Some possible useful textbooks:

Cohn: Elements of linear Algebra.

Cohn: Algebra (volume 1)

Curtis: Abstract linear Algebra.

Kayle Wilson: Linear Algebra.

Lang: Linear Algebra.

Lipschutz and Lipson: Linear Algebra.

Lipschutz: solved problems in linear Algebra. (3000)

CHAPTER 1: POLYNOMIAL RINGS

We define sets of polynomials and study ways in which they behave like (more familiar) the sets of natural numbers and integers, and in particular, the presence of a Euclidan Algorithm.

CHAPTER 2: LINEAR MAPS AND THE JORDAN NORMAL FORM

Linear maps represented by matrices, concentrate on square matrices.

→ Some matrices over $\mathbb{R}$ (i.e with real number entries)

eg: $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ or $\begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$ eigenvalues 2 3

eigenvectors (corresponding) $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$

$\Rightarrow \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 2\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ✓

We may diagonalise $\begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$ if $P = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}$

$\Rightarrow P^{-1}\begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}P = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$

for other matrices, we may need to work over $\mathbb{C}$ to diagonalise

eg. $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has eigenvalues $+i, -i$

But, even over $\mathbb{C}$, some matrices cannot be diagonalised

(there might not be enough linearly independent eigenvectors)

Every square matrix can be reduced to a simpler (not necessarily diagonal) form of the Jordan normal form.

CHAPTER 3 LINEAR AND BILINEAR FORMS AND INNER PRODUCT

SPACES

INPUT — OUTPUT

Linear form: one vector — one number

$$f\begin{pmatrix} x \\ y \end{pmatrix} = x + 2y$$

$$(1\ 2)\begin{pmatrix} x \\ y \end{pmatrix} = x + 2y$$

Bilinear form: two vectors — one number

$$(x_1, y_1)\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = x_1x_2 + 2x_1y_2 + 3x_2y_1 + 4y_1y_2$$

Inner product: take same input twice

eg: $(x\ y)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2$

CHAPTER 1: POLYNOMIAL RINGS

Some notation:

We use $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ to denote the sets of integers, rational numbers, real numbers, complex numbers respectively, while, in our course

$\mathbb{N} = \{1, 2, 3, \dots\}$ natural numbers / positive integers (not zero)

$\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$ non-negative integers

1.1. Some basic objects and constructions

Start with (some) basic objects involving one (binary) operation

Definition A **group** consists of a **set**, G, together with an **operation**, denoted by $*$, such that

(1) If $a \in G, b \in G \Rightarrow a * b \in G$

(2) For all $a, b, c \in G$: $(a * b) * c = a * (b * c)$

(3) There exists an identity element, e, in G, such that for every $a \in G$: $a * e = a$ and $e * a = a$.

(4) For each a in G, there exists an inverse element, a^{-1}, in G, such that $a * a^{-1} = e$ and $a^{-1} * a = e$.

We refer to objects satisfying conditions (1), (2), above as **semi-groups**.

We refer to objects satisfying conditions (1), (2), (3) above as **monoids**

Examples:

- If we take G to be $\mathbb{Z}$, and $*$ to be **addition**, we have a **group** (where $e = 0$).
- If we take G to be $\mathbb{Z}$, and $*$ to be **multiplication**, we have a **monoid** (where $e = 1$)

(since for example 2 doesn't have an inverse $\in \mathbb{Z}$).

- If we take G to be $\mathbb{N}$, and $*$ to be addition, we have a semi-group (ex: 0 doesn't belong to $\mathbb{N}$).
- If we take G to be $\mathbb{N}$, and $*$ to be multiplication, we have a monoid (No inverse $\in \mathbb{N}$).
- If we take G to be $\mathbb{N}_0$, and $*$ to be addition, we would have a monoid
- If we take G to be $\mathbb{N}_0$, and $*$ to be multiplication, monoid.
- If we take G to be $\mathbb{R}$, and $*$ to be addition $\Rightarrow$ group ($e=0$)
- If we take G to be $\mathbb{R}$, and $*$ to be multiplication $\Rightarrow$ monoid ($e=1$) (0 doesn't have an inverse).
- If we take G to be $\mathbb{R}\setminus\{0\}$, and $*$ to be multiplication $\Rightarrow$ group ($e=1$).

All of the above are instances of abelian structures.

A semi-group / monoid / group G is abelian if $\forall a,b \in G \Rightarrow a*b = b*a$.

Definition: A ring consists of a set, R, together with two operations denoted by '+' (addition) and '.' (multiplication) s.t

(+)
(1) If $a \in R \Rightarrow a+b \in R$.
(2) $\forall a,b \in R \Rightarrow (a+b)+c = a+(b+c)$.
(3) $\exists$ an identity element, 0, in R s.t $\forall a \in R \Rightarrow a+0=a$ and $0+a=a$
(4) For each a in R, $\exists$ an (additive) inverse, $-a$, in R s.t $a+(-a)=0$ and $(-a)+a=0$.
(5) $\forall a,b,c \in R \Rightarrow a+b=b+a$

(1) - (5) $\rightarrow$ R is an additive abelian group. (from 1 to 5)

(.)
(6) If $a,b \in R \Rightarrow a \cdot b \in R$ closure
(7) $\forall a,b,c \in R \Rightarrow (a \cdot b) \cdot c = a \cdot (b \cdot c)$ associativity
(8) $\exists$ an identity element (for multiplication), 1 in R, s.t $\forall a \in R \Rightarrow a \cdot 1 = a$ and $1 \cdot a = a$ identity

(6) - (8) $\rightarrow$ R is a multiplicative monoid

(9) $\forall a,b,c \in R \Rightarrow a \cdot (b+c) = a \cdot b + a \cdot c$
(10) $\forall a,b,c \in R \Rightarrow (b+c) \cdot a = b \cdot a + c \cdot a$

} Distributivity of multiplication over addition

Note: If it also holds that, $\forall a,b \in R \Rightarrow a \cdot b = b \cdot a$
$\Rightarrow$ R is said to be a commutative ring

If, in a commutative ring R, it also holds that every non-zero element in R has a multiplicative inverse (i.e. if $a \in R \setminus \{0\}$, $\Rightarrow$ there is an a^{-1} in R s.t $a \cdot a^{-1} = 1$ (and $a^{-1} \cdot a = 1$))
$\Rightarrow$ R is a field.

Examples:

- Suppose that '+', '.' denote "usual" addition and multiplication, and that '0', '1' denote the number 'zero' and 'one' respectively. Then $\mathbb{Z}$ forms a ring (a commutative ring in fact).
  $\mathbb{N}$ does not form a ring (conditions (3), (4) fail to hold).
  $\mathbb{N}_0$ does not form a ring (conditions (4) fails to hold)
  $\mathbb{R}$ forms a ring (in fact, $\mathbb{R}$ forms a field).
  All the structures we have seen so far, in examples, are commutative.
- Consider the set of 2×2 matrices with real entries, using "usual" addition and multiplication for matrices, where $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ is the identity for addition and $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is the identity for multiplication.
  This forms a ring but not a commutative ring.

eg: $\underset{A}{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}} \underset{B}{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

$AB \neq BA$

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

Friday October 4th 2019

Definition: given a ring R, the set of all polynomials over R (in one variable indeterminate, x say) is denoted by $R[x]$ and is defined via:

$$R[x] = \left\{ \sum_{i=0}^{\infty} a_i x^i = a_0 + a_1 x + a_2 x^2 + \dots \right.$$

$a_i \in R$ for $i = 0, 1, 2, \dots$, and (only) finitely many of the a_i are non-zero $\Big\}$

Notes:

- For a polynomial as above, the element a_i from R is the coefficient of x^i (a_0 is the constant coefficient)
- For a polynomial $a(x) = a_0 + a_1 x + a_2 x^2 + \dots$, we often simply use '$a$' to refer to '$a(x)$'.
- A polynomial of the form $a(x)$ for which $a_1 = 0, a_2 = 0, \dots$ i.e. a polynomial of the form $a(x) = a_0$ is a constant polynomial

The zero polynomial is $a(x) = 0$ i.e. $a(x) = 0 + 0 \cdot x + 0 \cdot x^2 + \dots$

In this way, we have a "copy" of the ring R (when R is $\mathbb{R}$?) within $R[x]$

↓

(set of constant polynomials).

- The condition that only finitely many of the a_i are non-zero rules out expressions such as: $1 + x + x^2 + x^3 + \dots$ not a polynomial.

As such, each polynomial "terminates": it has the form

$$a(x) = a_0 + a_1 x + \dots + a_n x^n.$$

Examples of polynomials in $\mathbb{R}[x]$.

$0, 1, \frac{1}{3}, -5, x, x+2, x^2+1, -x^2+1, x^3+x-1$

Equality of polynomials:

Polynomials $a(x) = a_0 + a_1 x + \dots$ and $b(x) = b_0 + b_1 x + \dots$

are equal if $a_i = b_i$ for every $i = 0, 1, 2, \dots$

eg $2x + 3 = 2x + 3$

$a_0 + a_1 x + a_2 x^2 + a_3 x^3 = 2x^3 - x + 5$.

precisely if $a_0 = 5, a_1 = -1, a_2 = 0, a_3 = 2$.

The set of polynomials $R[x]$ becomes a ring when we define addition and multiplication 'as usual'.

Addition: If $a(x) = a_0 + a_1 x + a_2 x^2 + \dots$ and $b(x) = b_0 + b_1 x + b_2 x^2 + \dots$ the sum is:

$$(a+b)(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \dots$$

Multiplication: If $a(x) = a_0 + a_1 x + \dots + a_m x^m$ and $b(x) = b_0 + b_1 x + \dots + b_n x^n$

the product is

$$(a \cdot b)(x) = (a_0 b_0) + (a_0 b_1 + a_1 b_0)x + (a_0 b_2 + a_1 b_1 + a_2 b_0)x^2 + \dots + (a_m b_n)x^{m+n}$$

For instance, in $\mathbb{R}[x]$

$(x^2 + 2x + 1) + (3x - 6) = x^2 + 5x - 5$

$(x-1)(x^2 + 2x + 3) = x^3 + x^2 + x - 3$.

Key notation associated to polynomials:

Degree of polynomial: For a ring R and a polynomial $a(x)$ in $R[x]$, the degree of $a(x)$ denoted by $\deg(a)$, is defined as follows:

- If $a(x) = 0$, the zero polynomial, then $\deg(a) = -\infty$ $\deg(0) = -\infty$
- If $a(x) \neq 0$, then it has the form $a(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n$ for $a_n \neq 0$.

In this case $\deg(a) = n$, i.e. it is the largest non-negative integer i for which $a_i \neq 0$

Examples in $\mathbb{R}[x]$:

$\deg(0) = -\infty$, $\deg(1) = 0$, $\deg(-\frac{1}{5}) = 0$, $\deg(x+1) = 1$, $\deg(2x-8) = 1$

$\deg(-x^2+1) = 2$, $\deg(x^4-3) = 4$.

NOTES:

- For a general polynomial $a(x)$, the possible values of $\deg(a)$ are $-\infty, 0, 1, 2, 3 \dots$
- $\deg(a) = -\infty$ if and only if $a(x) = 0$ (zero polynomial).

$\deg(a) = 0$ if and only if $a(x) = a_0$, $a_0 \neq 0$ non-zero constant polynomial.

Given a non-zero polynomial $a(x)$:

$a(x) = a_0 + a_1 x + \dots + a_n x^n$ for $a_n \neq 0$

The leading coefficient of $a(x)$ is a_n. If the leading coefficient of $a(x)$ is $= 1$, i.e. if $a(x) = x^n + \dots$ $\Rightarrow$ $a(x)$ is monic

The monic "version" of a non-zero polynomial $a(x) = a_n x^n + \dots + a_1 x + a_0$ $(a_n \neq 0)$ is the polynomial

$$x^n + \frac{a_{n-1}}{a_n} x^{n-1} + \dots + \frac{a_1}{a_n} x + \frac{a_0}{a_n}$$

(estoy multiplicando por la inverse de a_n)

⚠ (I can always achieve this if a_n has an inverse in R, eg. always true in a field since $a_n \neq 0$)

porque para ser field una condición es tener inversa

corregirlo en la pag 13 de las notes.

Como son polynomial rings no puedo multiplicar por números que no $\in$ al ring: eg en $2x+3 \in \mathbb{Z}$, no puedo hacer $x + \frac{3}{2}$ since $\frac{1}{2} \notin \mathbb{Z}$.

How does notion of degree interact with sums / products?

SUMS

Let $a(x) = a_0 + \dots + a_m x^m$ $(a_m \neq 0)$

$b(x) = b_0 + \dots + b_n x^n$ $(b_n \neq 0)$

Then $\deg(a) = m$, $\deg(b) = n$.

- Suppose $n > m$: $(a+b)(x) = (a_0 + b_0) + \dots + b_n x^n$

$\deg(a+b) = n$.

eg: $(2x^3+1) + (5x^2 - 4) = 2x^3 + 5x^2 - 3$

degrees: 3 2 3.

- Similarly if $m > n$: $\deg(a+b) = m$
- Suppose $m = n$: $(a+b)(x) = (a_0 + b_0) + \dots + (a_n + b_n) x^n$

Then $\deg(a+b) \leq n$.

Could have $\deg(a+b) < n$; if $a_n + b_n = 0$.

eg: $(x^3 - x + 1) + (-x^3 + 3x + 4) = 2x + 5$

degrees: 3 3 1

In general: $\deg(a+b) \leq \max(\deg(a), \deg(b))$

$\downarrow$

maximum of $\deg(a)$, $\deg(b)$.

Note: This also works if $a(x) = 0$ and/or $b(x) = 0$

eg: if $a(x) = 0$

$(a(x) + b(x)) = 0 + b(x) = b(x)$

$\deg(a) = -\infty$, $\deg(a+b) = \deg(b)$

$\max(\deg(a), \deg(b)) = \max(-\infty, \deg(b)) = \deg(b)$.

So $\deg(a+b) = \max(\deg(a), \deg(b))$

More generally, if $a(x)$, $b(x)$ are in $R[x]$ for any ring R:

$\deg(a+b) \leq \max(\deg(a), \deg(b))$

In fact, for $f_1(x), \dots, f_k(x)$ in $R[x]$

$\deg(f_1 + \dots + f_k) \leq \max(\deg(f_1), \dots, \deg(f_k))$

PRODUCTS / MULTIPLICATION

Let $a(x) = a_0 + a_1x + \dots + a_mx^m \quad (a_m \neq 0)$

$b(x) = b_0 + b_1x + \dots + b_nx^n \quad (b_n \neq 0)$

Then $(a \cdot b)(x) = (a_0 \cdot b_0) + (a_0b_1 + a_1b_0)x + \dots + (a_mb_n)x^{m+n}$.

Note: To ensure that if $a_m \neq 0$ and $b_n \neq 0 \Rightarrow a_mb_n \neq 0$,

from now on, we assume that the coefficients are in a field

In general, for any field k, and $a, b \in k$:

if $a \neq 0, b \neq 0 \Rightarrow ab \neq 0$

i.e. if $a \cdot b = 0 \Rightarrow a = 0$ or $b = 0$

Then, returning to our product: $a_mb_n \neq 0$ so $\deg(ab) = m+n$

This leads to: $\deg(ab) = \deg(a) + \deg(b)$

This also holds if $a(x) = 0$ and or $b(x) = 0$.

For instance, if $a(x) = 0$: $a(x)b(x) = 0 \cdot b(x) = 0$ $\quad \deg(ab) = -\infty$

$\deg(a) + \deg(b) = "-\infty + \deg(b) = -\infty"$.

So $\deg(ab) = \deg(a) + \deg(b)$ as required.

More generally, if $f_1(x), \dots, f_e(x)$ in $k[x]$, for some field k.

$$\deg(f_1f_2 \cdots f_e) = \deg(f_1) + \deg(f_2) + \dots + \deg(f_e).$$

1.2. Division and the Euclidean algorithm

Consider a, b elements in some ring R. Then b divides a, denoted by $b|a$ if there exists some q in R s.t. $a = q \cdot b$

Otherwise, b does not divide a, denoted by $b \nmid a$.

eg: If R is $\mathbb{Z}$:

$3|3$, $3|12$, $-3|3$, $-3|12$

but $12 \nmid 3$, $2 \nmid 5$, $5 \nmid 2$

while if R is $\mathbb{R}[x]$:

$x+1 \mid 2x+2$, $x+1 \mid x^2-1$ since $(x+1)(x-1) = x^2-1$

$3x+3 \mid x^2-1$ since $(3x+3)(\frac{1}{3}x - \frac{1}{3}) = x^2-1$

But $x^2-1 \nmid x+1$

Note: "$\frac{1}{x-1}$" is not a polynomial and $x+1 \nmid x+2$.

Some basic properties of division

for any ring R:

(1) If $a|b \Rightarrow a|rb \quad \forall r \in R$

(2) If $a|b$ and $a|c \Rightarrow a|b+c$

From (1), (2) we obtain:

(3) If $a|b$ and $a|c \Rightarrow a|r_1b + r_2c \quad \forall r_1, r_2 \in R$

(4) If $a|b$ and $b|c \Rightarrow a|c$

Notes: ⚠

- for every $a \in R$: $a \cdot 0 = 0$, so $a|0$
- For every $a \in R$: $1 \cdot a = a$, so $1|a$.

Try to classify elements in rings from the point of view of division:

First, consider invertible elements or units.

In a ring R, an element u is a unit if $\exists$ an element u' in R s.t. $u \cdot u' = 1$ and $u' \cdot u = 1$.

The set of all units in R is denoted by $U(R)$.

- In a field k, every non-zero element is invertible

$U(k) = k \setminus \{0\}$.

→ $U(\mathbb{R}) = \mathbb{R} \setminus \{0\}$, $U(\mathbb{C}) = \mathbb{C} \setminus \{0\}$.

→ In $\mathbb{Z}$ $U(\mathbb{Z}) = \{\pm 1\} = \{+1, -1\}$.

$(+1)(+1) = +1$

$(-1)(-1) = +1$,

To show this, could use absolute value / modulus.

Note: For $a, b \in \mathbb{Z}$: $|ab| = |a| \cdot |b|$.

Suppose u is a unit in $\mathbb{Z}$, i.e. for some $u' \in \mathbb{Z}$, $u \cdot u' = 1$.

Then $|uu'| = |1|$ and $|u||u'| = 1$.

Then, we must have $|u|=1$ (and $|u'|=1$)

i.e. $u=+1$ or $u=-1$ (only two options)

October 7th

Lets now see how we may use the notion of degree to determine all units in $K[x]$ (→ all polynomials in a field K), for a field K.

Proposition Let K be a field.

An element $u(x)$ in $K[x]$ is a unit in $K[x]$ iff $u(x)$ is a non-zero constant polynomial (i.e. iff $\deg(u)=0$)

So $\mathcal{U}(K[x]) = K\setminus\{0\}$.

Proof ($\Leftarrow$) Suppose, first, that $u(x)=u$ is a non-zero constant.

Then, since K is a field and $u\neq 0$, there must exist u^{-1} in K, s.t. $u\cdot u^{-1}=1$ and $u^{-1}\cdot u=1$. So, every non-zero constant (polynomial) is a unit in $K[x]$, as required.

($\Rightarrow$) Next, suppose that $u(x)$ is a unit in $K[x]$.

Then, for some $u'(x)$ in $K[x]$ is s.t. $u(x)\,u'(x)=1$ [and $u'(x)\cdot u(x)=1$].

Applying degrees, $\deg(u\cdot u')=\deg(1)$

Then, using an earlier result:

$$\deg(u)+\deg(u')=\deg(1)$$

i.e. $\deg(u)+\deg(u')=0$

Note: possible values of degree are $-\infty, 0, 1, 2, 3, \ldots$

The only possibility is that $\deg(u)=0$

Since $\deg(u)=0$, the polynomial $u(x)$ must be a non-zero constant polynomial, as required. This completes the proof.

For example, some units in $\mathbb{R}[x]$ are: $1, -1, \frac{1}{2}, \sqrt{2}, -3, \pi, 0, \ldots$

$$\mathcal{U}(\mathbb{R}[x]) = \mathbb{R}\setminus\{0\}.$$

Now, study elements in rings that are similar to primes in $\mathbb{N}$: 2,3,5,7,...

p primes in $\mathbb{N}$ → if $p=bc \Rightarrow b=1$ or $c=1$.

- In general, an element a, in a ring R, is irreducible if a is not a unit and the following holds if $a=bc$ for some $b,c\in R \Rightarrow$ b or c is a unit.

(la única manera de descomponer el nº requiere utilizar un unit)

Examples:

In $\mathbb{Z}$, the units are $+1, -1$: the irreducible elements are

$\pm 2, \pm 3, \pm 5, \pm 7 \ldots$ → cuando hago su descomposición en factores primos es el número por 1 o (-1) al ser las unit de $\mathbb{Z}$ $\frac{\pm 1}{y}$ $\frac{\pm 1}{2} = 2\,1$

In $K[x]$, the degree can help us identify some irreducible polynomials.

Proposition Let K be a field and $a(x)$ be in $K[x]$.

Then, if $\deg(a)=1$, $a(x)$ is irreducible in $K[x]$. (Puede tb tener otro degree y ser irreducible)

Proof First note that, since $\deg(a)=1$, $a(x)$ can not be a unit (as shown above that units in $K[x]$ have degree 0).

Now, suppose that $a(x)=b(x)\cdot c(x)$.

This leads to $\deg(a)=\deg(bc)$

i.e. $\deg(a)=\deg(b)+\deg(c)$

i.e. $1=\deg(b)+\deg(c)$.

Given some possible values for degree, must have:

$\deg(b)=1$ and $\deg(c)=0$

or

$\deg(b)=0$ and $\deg(c)=1$

} Overall, $\deg(b)=0$ or $\deg(c)=0$.

So using previous result, b is a unit or c is a unit.

If $a(x)=b(x)\cdot c(x)$, $b(x)$ is a unit or $c(x)$ is a unit

This concludes the proof. $a(x)$ is irreducible

for instance, in $\mathbb{R}[x]$.

$x, x+1, x-1, 2x+5, \ldots$ are irreducible.

Note: A polynomial such as x^2+1 is irreducible in $\mathbb{R}[x]$ but can be reduced in $\mathbb{C}$ $x^2+1=(x+i)(x-i)$.

- Final type of element in a ring is:

An element in a ring R is reducible if it is neither a unit nor irreducible,

i.e. it is possible to express a as $a=bc$ for b, c in R and where b,c are not units.

Example: in every ring R, 0 is reducible

$0 = 0\cdot 0$

neither is a unit.

- In $\mathbb{Z}$, the reducible elements are $0, \pm 4, \pm 6, \pm 8, \ldots$
- In $\mathbb{R}[x]$, the following are reducible x^2-1, x^2+3x+2

$$x^2-1=(x+1)(x-1) \quad , \quad x^2+3x+2=(x+2)(x+1) \; .$$

Now, study division via **Euclidean division**:

⊙ In $\mathbb{N}$: For $m, n \in \mathbb{N}$ $\exists$ unique $q \in \mathbb{N}$, $r \in \mathbb{N}_0$ s.t $n = qm + r$ and $0 \leq r < m$.

⊙ Equivalent statement in $\mathbb{Z}$: For $a, b \in \mathbb{Z}$, where $b \neq 0$, $\exists$ unique $q, r \in \mathbb{Z}$ such that $a = qb + r$ and $0 \leq r < |b|$.

Example: Divide -7 by 3:

$$-7 = (-2)(3) + (-1)$$
$$-7 = (-3)(3) + 2 \; .$$

(satisfies conditions of statement above.)

October 8th 2019

$(-7) = (-3)(3) + 2$

$a = q \cdot b + r$

How can we arrive at this in a systematic way?

(Start from) $-7 = (0)(3) + (-7)$

$(-7) = (-1)(3) + (-4)$

$(-7) = (-2)(3) + (-1)$

$(-7) = (-3)(3) + (2)$

$0 \leq r < |b|$

$0 \leq 2 < |3|$.

How can we show uniqueness of q, r (in $\mathbb{Z}$)?

Suppose $a = qb + r$ where $0 \leq r < |b|$

and $a = \bar{q}b + \bar{r}$ (same across, numbers distinct) $0 \leq \bar{r} < |b|$

Then $qb + r = \bar{q}b + \bar{r}$, i.e. $b(q - \bar{q}) = \bar{r} - r$

so $|b(q-\bar{q})| = |\bar{r} - r|$, i.e. $|b||q-\bar{q}| = |\bar{r} - r|$

Note that, since $0 \leq r < |b|$, $0 \leq \bar{r} < |b|$.

we must have $-b < \bar{r} - r < b$, so $0 \leq |\bar{r} - r| < |b|$

Substituting into (*) $|b||q-\bar{q}| < |b|$

Since $b \neq 0$, $|b| \neq 0$ so we obtain $|q - \bar{q}| < 1$

But, $q, \bar{q}$ are integers, so we must have $|q - \bar{q}| = 0$.

So $q = \bar{q}$ it follows that $r = \bar{r}$.

So q, r are uniquely determined. ▨

In $k[x]$, instead of $\mathbb{Z}$, use degree instead of modulus (k is a field)

⊙ Euclidean division in $k[x]$:

$a(x) = q(x)\,b(x) + r(x)$ where $\deg(r) < \deg(b)$.

Example:

Start with $a(x) = x^2+3x+3$, $b(x) = x+1$

$$x^2+3x+3 = 0 \cdot (x+1) + (x^2+3x+3)$$
$$x^2+3x+3 = x(x+1) + (-x(x+1) + x^2+3x+3)$$
$$x^2+3x+3 = x(x+1) + (2x+3)$$
$$x^2+3x+3 = (x+2)(x+1) + 1$$

$a \quad q \quad b \quad r$

$\deg(r) < \deg(b)$

$0 < 1$.

we may express this whole process using a 'long division'.

$$\begin{array}{r|l} & x+2 \\ \hline x+1 & x^2+3x+3 \\ & -(x^2+x) \\ \hline & 2x+3 \\ & -(2x+2) \\ \hline & 1 \end{array}$$

$x^2+3x+3 = (x+2)(x+1)+1$.

Similarly:

$$\begin{array}{r|l} & \frac{1}{2}x+1 \\ \hline 2x+4 & x^2+4x+7 \\ & -(x^2+2x) \\ \hline & 2x+7 \\ & -(2x+4) \\ \hline & 3 \end{array}$$

$x^2+4x+7 = \left(\frac{1}{2}x+1\right)(2x+4)+3$.

a(x) b(x)

Lets also divide $3x^3+5$ by x^2+x-2

$$\begin{array}{r|l} & 3x-3 \\ x^2+x-2 & \overline{3x^3+0x^2+0x+5} \\ & -(3x^3+3x^2-6x) \\ \hline & -3x^2+6x+5 \\ & -(-3x^2-3x+6) \\ \hline & 9x-1 \end{array}$$

$3x^3+5 = (3x-3)(x^2+x-2) + (9x-1)$

$a \quad q \quad b \quad r$

$\deg(r) < \deg(b)$

$1 < 2$

Note: The remainder could also be 0:

eg: $x^2+3x+2 = (x+2)(x+1) + 0$.

Let's now show that Euclidean division works in general in $K[x]$:

Proposition Let K be a field, and $a(x), b(x)$ be in $K[x]$ where $b(x) \neq 0$. Then, there exists unique $q(x), r(x)$ in $K[x]$ s.t

$$a(x) = q(x)\,b(x) + r(x) \quad \text{and} \quad \deg(r) < \deg(b)$$

Proof:

First, show that such polynomials $q(x)$ and $r(x)$ exist.

Start the process by choosing any $q'(x), r'(x)$

satisfying $a(x) = q'(x)\,b(x) + r'(x)$

eg: set $q'(x) = 0$, $r'(x) = a(x)$: $a(x) = 0\,b(x) + a(x)$.

- If $\deg(r') < \deg(b)$, then we may end the process imediatly

  set $q(x) = q'(x)$, $r(x) = r'(x)$.

- Suppose instead that $\deg(r') \geq \deg(b)$

Then apply the following (perhaps repeatedly, if necessary):

let $r'(x) = r_n x^n + \dots + r_1 x + r_0 \quad r_n \neq 0$.

$b(x) = b_m x^m + \dots + b_1 x + b_0 \quad b_m \neq 0$.

Note: Such a $b(x)$ exists with $b_m \neq 0$, since $b(x) \neq 0$ by assumption.

Then, since $\deg(r') \geq \deg(b)$, $b(x) \neq 0$, also ensures that $r'(x) \neq 0$ (so can find $r_n \neq 0$, as above).

Also, since $\deg(r') \geq \deg(b)$, we have $n \geq m$.

Consider, $r_n b_m^{-1} x^{n-m}$ (note: b_m^{-1} exists since $b_m \neq 0$ and b_m is in a field K). Como estoy demostrando que existe $q'(x)$ y $r'(x)$, lo que hago es reducir el $\deg(r')$ hasta que cumpla $\deg r'(x) < \deg(b)$

Then, $r_n b_m^{-1} x^{n-m}\, b(x) = r_n b_m^{-1} x^{n-m} (b_m x^m + b_{m-1} x^{m-1} + \dots + b_0) =$

$$= r_n x^n + \frac{r_n b_{m-1}}{b_m} x^{n-1} + \dots + \frac{r_n b_0}{b_m} x^{n-m}$$

Then, we can use this to "decrease the degree of $r'(x)$".

$$r'(x) - r_n b_m^{-1} x^{n-m}\, b(x) = (r_n x^n + r_{n-1} x^{n-1} + \dots) - \left(r_n x^n + \frac{r_n b_{m-1}}{b_m} x^{n-1} + \dots\right)$$

$$r'(x) - r_n b_m^{-1} x^{n-m}\, b(x) = \left(r_{n-1} - \frac{r_n b_{m-1}}{b_m}\right) x^{n-1} + \dots$$

This leads to

$$a(x) = q'(x)\,b(x) + r'(x) =$$

$$= q'(x)\,b(x) + r_n b_m^{-1} x^{n-m}\, b(x) - r_n b_m^{-1} x^{n-m}\, b(x) + r'(x)$$

$$a(x) = \left(q'(x) + r_n b_m^{-1} x^{n-m}\right) b(x) + \left(r'(x) - r_n b_m^{-1} x^{n-m}\, b(x)\right)$$

Set $q''(x) = q'(x) + r_n b_m^{-1} x^{n-m}$, $r''(x) = r'(x) - r_n b_m^{-1} x^{n-m}\, b(x)$

we obtain

$$a(x) = q''(x)\,b(x) + r''(x)$$

and as shown above, $\deg(r'') \leq n-1$ i.e. $\deg(r'') < \deg(r')$ ($n-1 < n$)

We may repeat this process to obtain 'remainder' of strictly smaller degree each time.

After a finite number of steps, we will be able to find a remainder $r(x)$ satisfying $\deg(r) < \deg(b)$

October 11th 2019

Let's now show the uniqueness of $q(x), r(x)$.

Suppose $\exists\, q(x), \bar{q}(x), r(x), \bar{r}(x)$ in $K[x]$ s.t

$$a(x) = q(x)\,b(x) + r(x), \quad \deg(r) < \deg(b) \quad (1)$$

$$a(x) = \bar{q}(x)\,b(x) + \bar{r}(x), \quad \deg(\bar{r}) < \deg(b) \quad (2)$$

Let's show that we must have

$$\bar{q}(x) = q(x) \quad \text{and} \quad \bar{r}(x) = r(x)$$

Note: $q(x)\,b(x) + r(x) = \bar{q}(x)\,b(x) + \bar{r}(x)$

Then, $b(x)\,(q(x) - \bar{q}(x)) = \bar{r}(x) - r(x)$

Consider degrees: $\deg\left(b'(q-\bar{q})\right) = \deg(\bar{r}-r)$

so $\deg(b) + \deg(q-\bar{q}) = \deg(\bar{r}-r) = \deg(\bar{r}+(-r))$

So: $\deg(b) + \deg(q-\bar{q}) \leq \max(\deg(\bar{r}), \deg(-r))$

[Note: $\deg(-r) = \deg((-1)\cdot r) = \deg(-1) + \deg(r) = 0 + \deg(r)$

so $\deg(-r) = \deg(r)$]

Then, $\deg(b) + \deg(q-\bar{q}) \leq \max(\deg(\bar{r}), \deg(r))$

Using $\deg(\bar{r}) \leq \deg(b)$ and $\deg(\bar{r}) < \deg(b)$ we obtain

$$\deg(b) + \deg(q-\bar{q}) < \deg(b)$$

Then, since $b(x) \neq 0$, $\deg(b) \geq 0$, so:

$$\deg(q-\bar{q}) < 0$$

The only option, given definition of degree, is $\deg(q-\bar{q}) = -\infty$

i.e. $q(x) - \bar{q}(x) = 0$.

Hence, $\bar{q}(x) = q(x)$, and by substituting into $q(x)b(x) + r(x) = \bar{q}(x)b(x) + \bar{r}(x)$

we obtain $\bar{r}(x) = r(x)$.

Overall: $\bar{q}(x) = q(x)$, $\bar{r}(x) = r(x)$ as required.

This completes the proof

We next study greatest common divisor in $k[x]$:

→ Starting from $\mathbb{N}$:

For $a, b \in \mathbb{N}$, the natural number d is a greatest common divisor of a and b if:

* $d \mid a$ and $d \mid b$, and
* if, for some $c \in \mathbb{N} \Rightarrow c \mid a$ and $c \mid b \Rightarrow c \mid d$.

Within $\mathbb{N}$, greatest common divisors are unique.

Ex: In $\mathbb{N}$, divisors of 6: 1, 2, 3, 6.
9: 1, 3, 9

common divisors: 1, 3.

greatest common divisors is 3.

$\gcd_{\mathbb{N}}(6,9) = 3$.

→ If we "transfer" the definition to $\mathbb{Z}$ or $k[x]$ we no longer have uniqueness

ex: in $\mathbb{Z}$: division of 6: $\pm 1, \pm 2, \pm 3, \pm 6$. (factorizo en factores primos y añado el 1 y el nº que estoy factorizando)
9: $\pm 1, \pm 3, \pm 9$.

Both 3 and (-3) satisfy the conditions given earlier (for $\mathbb{N}$).

→ In $\mathbb{R}[x]$: consider x^2-1 and x^2+3x+2

• divisors of x^2-1 are: $1, (x+1), (x-1), (x^2-1)$ and all non-zero multiples of these.

eg: $2(x+1)$ divides x^2-1: $x^2-1 = \left(\frac{1}{2}x - \frac{1}{2}\right)(2(x+1))$,

• divisors of x^2+3x+2: $1, x+1, x+2, x^2+3x+2$ (factorizo en factores primos añado 1 y el polinomio que estoy factorizando y las non-zero multiples)
and all non-zero multiples of these.

For instance: $x+1, 2(x+1), \frac{1}{3}(x+1), -(x+1)$ all divide x^2-1 and x^2+3x+2.

In $\mathbb{Z}$, can ensure uniqueness by insisting that greatest common divisors are positive, in $k[x]$, we insist that they are monic

Definition: Suppose $a(x), b(x)$, in $k[x]$ (for a field k), where at least one of $a(x), b(x)$ is non-zero

Then, a polynomial $d(x)$ in $k[x]$ is a greatest common divisor of $a(x)$ and $b(x)$ if $d(x)$ is monic and:

→ $d(x) \mid a(x)$ and $d(x) \mid b(x)$, and

→ If $c(x) \mid a(x)$ and $c(x) \mid b(x)$, for some $c(x) \in k[x]$
$\Rightarrow c(x) \mid d(x)$.

Let's first check that greatest common (factor) divisors are unique in $k[x]$:

Two helpful results, for $k[x]$, k a field:

(#1) Consider $a(x), b(x)$ s.t $a(x) = u\,b(x)$ for $u \in k \setminus \{0\}$ (u non-zero constant).

and $a(x), b(x)$ both monic $\Rightarrow a(x) = b(x)$ and $u = 1$

Proof: Since $a(x) = ub(x)$

$$\deg(a) = \deg(ub) = \deg(u) + \deg(b) = 0 + \deg(b)$$

Since u is a non-zero constant.

so $\deg(a) = \deg(b)$, say n:

Then, for $a_i, b_j \in k$ $(0 \le i \le n)$:

$a(x) = x^n + \ldots + a_0$. where $a_n \neq 0$

$b(x) = x^n + \ldots + b_0$ $\quad b_n \neq 0$

Then,

$ub(x) = u\,x^n + \ldots + ub_0$.

Since $a(x) = ub(x)$, by comparing coefficients of x^n, we obtain $u = 1$, i.e. $a(x) = b(x)$ as required.

- Suppose $a(x), b(x)$ are monic and where $a(x) \mid b(x)$ and $b(x) \mid a(x)$. Then $a(x) = b(x)$

Proof: Since $a(x) \mid b(x)$, $b(x) \mid a(x)$, for some $c(x), c'(x)$:

$a(x)\,c(x) = b(x)$ and $b(x)\,c'(x) = a(x)$

respectively.

Then $b(x) = c(x)\,a(x) = c(x)\,(c'(x)\,b(x))$

i.e $b(x) = c(x)\,c'(x)\,b(x)$

Then $\deg(b) = \deg(cc'b) = \deg(c) + \deg(c') + \deg(b)$

i.e. $\boxed{\deg(c) + \deg(c') = 0}$.

(note: $b(x)$ is monic, so $b(x) \neq 0$ and $\deg(b) \geq 0$)

Given definition of degree, must have:

$\deg(c) = 0$ (and $\deg(c') = 0$).

So $c(x)$ is a non-zero constant, say $c(x) = c$

Then $b(x) = c(x)\,a(x)$

$b(x) = c\,a(x)$

└ non-zero constant.

Hence, using previous result (#1)

$a(x) = b(x)$ as required

Finally, can show uniqueness:

Proposition: Suppose $a(x), b(x) \in k[x]$ (for a field k), such that at least one of $a(x), b(x)$ is non-zero

Suppose also that $d(x)$ and $d'(x)$ are both greatest common divisors of $a(x), b(x)$. Then $\boxed{d'(x) = d(x)}$

ie. greatest common divisors are unique (in $k[x]$)

Proof:

By definition of greatest common divisor: $d(x) \mid a(x)$ and $d'(x) \mid b(x)$

Since $d(x) \mid a(x)$, $d(x) \mid b(x)$ and $d'(x)$ is a greatest common divisor of $a(x)$ and $b(x)$ it follows that $d(x) \mid d'(x)$.

Similarly, since $d'(x) \mid a(x)$ and $d'(x) \mid b(x)$ and $d(x)$ is a gcd (of $a(x), b(x)$) $d'(x) \mid d(x)$

Then, $d(x) \mid d'(x)$, $d'(x) \mid d(x)$ and also, by definition of greatest common divisor, $d(x)$ and $d'(x)$ are both monic, we obtain (using the previous result): $d(x) = d'(x)$ as required

We may denote the unique gcd of suitable $a(x), b(x)$ by $\gcd(a, b)$, or equivalently: $\gcd(b, a)$.

Examples:

In $\mathbb{R}[x]$: $\gcd(x^2 - 1, x^2 + 3x + 2) = x + 1$

$\gcd(6, 9) = \underbrace{1}_{\text{in } \mathbb{R}[x], \text{ gcd has to be monic}}$

└ It is in polynomials !!

whereas, $\gcd(6, 9)$ in $\mathbb{N}$ $= 3$.

In general $\gcd(a,b)=1$ in $k[x]$ for any non-zero constants a,b.

Some other special cases:

Suppose $a(x)$ is in $k[x]$ and $a(x)\neq 0$

$a(x)=a_n x^n + a_{n-1}x^{n-1} + \dots + a_0$. for $a_0,\dots,a_n \in k$, $a_n \neq 0$.

Then, $\gcd(a(x),0) = x^n + \frac{a_{n-1}}{a_n}x^{n-1} + \dots + \frac{a_0}{a_n}$ "monic version" of $a(x)$.

- Suppose $a(x)$ in $k[x]$ and u to be non-zero constant $\gcd(a(x),u) = 1$

eg. $\gcd(x+1,2)=1$

$\gcd(x^2-1,3)=1$

In particular $\gcd(a(x),1)=1$ for any $a(x)$ in $k[x]$.

To compute $\gcd(a(x),b(x))$ in general, we use Euclidean algorithm.

Example: Consider $f(x)=x^3-10x+3$, $g(x)=x^2-9$

Start by dividing $f(x)$ by $g(x)$.

$$\begin{array}{r|l} & x \\ x^2-9 & x^3+0x^2-10x+3 \\ & -x^3 \quad -9x \\ \hline & -x+3 \end{array}$$

So we obtain $x^3-10x+3 = (x)(x^2-9)+(-x+3)$

Now, divide $g(x)=x^2-9$ by $(-x+3)$: (to r_1)

$x^2-9 = (-x-3)(-x+3)+0$.

Overall, the algorithm leads to the following:

$x^3-10x+3 = x(x^2-9)+(-x+3)$ (*)

$x^2-9 = (-x-3)(-x+3)+0$

(The algorithm stops when we arrive at a remainder of zero)

The final non-zero remainder then leads to $\gcd(f(x),g(x))$

Here, that remainder is $-x+3$:

We require $\gcd(f(x),g(x))$ to be monic, so divide through by leading coefficient of $(-x+3)$, namely -1, thus obtaining

$\gcd(\underbrace{x^3-10x+3}_{f(x)}, \underbrace{x^2-9}_{g(x)}) = (x-3)$ ↳ "monic version" of $(-x+3)$

From this process, can also find $a(x)$, $b(x)$ s.t $\gcd(f,g) = a(x)f(x)+b(x)g(x)$.

(using "back substitution" through the algorithm)

From (*):

$-x+3 = 1\cdot(x^3-10x+3) - x(x^2-9)$

So $\underbrace{x-3}_{\gcd(f,g)} = (-1)\underbrace{(x^3-10x+3)}_{f(x)} + \underbrace{(x^2-9)}_{g(x)}\cdot x$

October. 14th 2019

We now try to justify the validity of the method used in the previous example.

First, we indicate a connection between gcf of pairs of polynomials appearing in the same Euclidean division:

Proposition: Suppose $f(x), g(x), q(x), r(x)$ are in $k[x]$, for a field k, and that $f(x) = q(x)g(x)+r(x)$, where $g(x)$ is non-zero.

Then $\gcd(f,g) = \gcd(g,r)$.

Proof: First, note that, since $g(x)\neq 0$, both $\gcd(f,g)$ and $\gcd(g,r)$ are defined.

Let $d(x)=\gcf(f,g)$ and $d'(x)=\gcd(g,r)$

Consider $d'(x)=\gcd(g,r)$. Then by definition

$d'(x)\mid g(x)$ and $d'(x)\mid r(x)$.

Since $d'(x)\mid g(x)$, we obtain $d'(x)\mid q(x)g(x)$ and so $d'(x)\mid \underbrace{q(x)g(x)+r(x)}_{f(x)}$

(using $d'(x)\mid r(x)$)

So $d'(x) \mid f(x)$. Also $d'(x) \mid g(x)$, so, given that $d(x) = \gcd(f,g)$, we must have $\boxed{d'(x) \mid d(x)}$.

Consider $d(x) = \gcd(f,g)$. Then $d(x) \mid f(x)$ and $d(x) \mid g(x)$

Since $d(x) \mid g(x)$, $d(x) \mid -q(x)g(x)$, also $d(x) \mid f(x)$

overall $d(x) \mid f(x) - q(x)g(x)$ i.e $d(x) \mid r(x)$.

Then, given that $d(x) \mid g(x)$ and $d(x) \mid r(x)$, we deduce that $\boxed{d(x) \mid d'(x)}$ (since $d'(x) = \gcd(g,r)$)

By definition of gcd, $d(x)$ and $d'(x)$ are also monic (and $d(x) \mid d'(x)$, $d'(x) \mid d(x)$).

So, using an earlier result: $d(x) = d'(x)$

i.e $\gcd(f,g) = \gcd(g,r)$ as required

Let's now consider the steps of a general application of the Euclidean algorithm:

$$f(x) = q_1(x)\,g(x) + r_1(x)$$
$$g(x) = q_2(x)\,r_1(x) + r_2(x)$$
$$r_1(x) = q_3(x)\,r_2(x) + r_3(x)$$
$$\vdots$$
$$r_{n-2}(x) = q_n(x)\,r_{n-1}(x) + r_n(x)$$
$$r_{n-1}(x) = q_{n+1}(x)\,r_n(x) + \boxed{0}$$

$\deg(r_1) < \deg(g)$
$\deg(r_2) < \deg(r_1)$
$\vdots$
$\deg(r_n) < \deg(r_{n-1})$

⚠ NOTE: The degree of the remainder becomes strictly smaller with each application of Euclidean division, so, eventually, after a finite number of steps, the process will terminate, with a zero remainder

$\ldots < \deg(r_n) < \deg(r_{n-1}) < \ldots$

Using previous result:

$$\gcd(f,g) = \gcd(g,r_1)$$
$$\gcd(g,r_1) = \gcd(r_1,r_2)$$
$$\gcd(r_1,r_2) = \gcd(r_2,r_3)$$
$$\vdots$$
$$\gcd(r_{n-2},r_{n-1}) = \gcd(r_{n-1},r_n)$$
$$\gcd(r_{n-1},r_n) = \gcd(r_n,0)$$

$\gcd(f,g) = \gcd(g,r)$
$= \gcd(r_1,r_2)$
$\vdots$
$= \gcd(r_n,0)$

i.e. $\gcd(f,g) = \gcd(r_n,0)$

But, as seen earlier $\gcd(r_n,0)$ is the monic "version" of r_n

So, as in the previous example, $\gcd(f,g)$ is the monic "version" of the final non-zero remainder.

Further, we may rearrange the equations involved:

$$r_1(x) = f(x) - q_1(x)\,g(x)$$
$$r_2(x) = g(x) - q_2(x)\,r_1(x)$$
$$\vdots$$
$$r_{n-1}(x) = r_{n-3}(x) - q_{n-1}(x)\,r_{n-2}(x)$$
$$r_n(x) = r_{n-2}(x) - q_n(x)\,r_{n-1}(x)$$

and use "back substitution" to express $r_n(x)$, (and so, $\gcd(f,g)$) as a "combination of $f(x), g(x)$".

This is known as BEZOUT'S LEMMA

Proposition: Given $f(x), g(x)$ in $K[x]$ (for a field K), such that $f(x), g(x)$ are not both zero, $\exists\, a(x), b(x)$ in $K[x]$ s.t $\gcd(f,g) = a(x)f(x) + b(x)g(x)$

⚠ NOTE: If, in the algorithm $r_1(x) = 0$
i.e $f(x) = q(x)\,g(x) + 0$
we regard the monic "version" of $g(x)$ as $\gcd(f,g)$

eg: $\underbrace{x^2-1}_{f(x)} = (x-1)\underbrace{(x+1)}_{g(x)} + 0$ so $\gcd(x^2-1, x+1) = x+1$

In this sense, we may treat $g(x)$ as $r_0(x)$.

ex. $\gcd(12,3)$
como $12 = 4 \cdot 3 + 0$
↳ $3 = \gcd$.

Example: Consider $f(x) = 9x^3 - 3x^2 + 4x + 2$, $g(x) = 3x^2 - x + 1$

Applying the Euclidean algorithm leads to:

$$9x^3 - 3x^2 + 4x + 2 = (3x)(3x^2 - x + 1) + (x + 2)$$

$$3x^2 - x + 1 = (3x - 7)(x + 2) + 15$$

$$x + 2 = \left(\frac{1}{15}x + \frac{2}{15}\right)(15) + 0.$$

The final non-zero remainder is 15 but $\gcd(f,g)$ must be monic, so $\gcd(9x^3 - 3x^2 + 4x + 2, 3x^2 - x + 1) = 1 \quad \left(= \frac{1}{15} \cdot 15\right)$.

Also, $15 = (3x^2 - x + 1) - (3x - 7)(x + 2) =$

$$= (3x^2 - x + 1) - (3x - 7)\left((9x^3 - 3x^2 + 4x + 2) - (3x)(3x^2 - x + 1)\right) =$$

$$= g(x) - (3x - 7)\left(f(x) - 3x\,g(x)\right).$$

so $15 = (-3x + 7) f(x) + \left(1 + (3x - 7)(3x)\right) g(x)$

$$15 = (-3x + 7) f(x) + (9x^2 - 21x + 1) g(x)$$

Then, $1 = \frac{1}{15}(-3x + 7) f(x) + \frac{1}{15}(9x^2 - 2x + 1) g(x)$.

↑ $\gcd(f,g)$

Here: $\gcd(f,g) = a(x) f(x) + b(x) g(x)$

where $a(x) = \frac{-3}{15}x + \frac{7}{15}$, $b(x) = \frac{9}{15}x^2 - \frac{2}{15}x + \frac{1}{15}$.

October 15th 2019

We now try to study (the connection between the Euclidean algorithm and) factorizations of polynomials in $k[x]$.

1.3. Factorizations into irreducible polynomials

We first note some results related to (ir)reducibility from previously in the course.

If a polynomial $f(x)$ in $k[x]$ has degree 1 $\Rightarrow$ it is irreducible (for every field $k[x]$).

2 ESPECIAL CASES:

- If $k = \mathbb{C}$

Here, the following holds:

$\forall$ polynomial $f(x)$ in $\mathbb{C}[x]$ s.t $\deg(f) \geq 1$, $\exists$ a factorization

$f(x) = f_1(x) \cdots f_n(x)$ where, for $1 \leq i \leq n$ $f_i(x) \in \mathbb{C}[x]$ and $\deg(f) = 1$

As a consequence, every polynomial of degree greater than or equal to 2, in $\mathbb{C}[x]$, can be factorised, i.e. it is reducible overall, in $\mathbb{C}[x]$:

> A polynomial in $\mathbb{C}[x]$ is irreducible iff it has degree 1

For instance, the following are irreducible in $\mathbb{C}[x]$:

$x+1$, $x-i$, $2x+3$, $x+i$, $x+\sqrt{5}$, $x-i\sqrt{3}$.

whereas the following are all reducible:

$(x^2-1) = (x+1)(x-1)$

$(x^2+1) = (x+i)(x-i)$

$x^3 + x^2 + x + 1 \ldots$

$x^4 + ix + 3$

$x^3 - 17x^3 + 18x^2 + 19x - 20$

- If $k = \mathbb{R}$

Here, there (also) exist irreducible polynomials of degree 2 (but of no greater degree).

Any factorization of such a polynomial in $\mathbb{C}[x]$, may involve pairs of complex conjugate roots, which can be combined to give a degree 2 polynomial in $\mathbb{R}[x]$.

eg: over $\mathbb{R}$ $\underbrace{x^2+1}_{\text{irreducible over } \mathbb{R}} = \underbrace{(x+i)(x-i)}_{\text{reducible over } \mathbb{C}}$.

eg: consider x^3+x^2+x+1.

→ over $\mathbb{C}$: $x^3+x^2+x+1 = \underbrace{(x+1)(x+i)\cdot(x-i)}$.
all irreducible over $\mathbb{C}$.

→ over $\mathbb{R}$: $x^3+x^2+x+1 = \underbrace{(x+1)(x^2+1)}$.
all irreducible over $\mathbb{R}$.

In general, for $\mathbb{R}[x]$:

For every polynomial $f(x)$ in $\mathbb{R}[x]$ s.t $\deg(f) \geq 1$,
$\exists$ a factorisation $f(x) = f_1(x) \cdots f_m(x)$ where, for $1 \leq i \leq m$:
$f_i(x) \in \mathbb{R}[x]$ and
$\deg(f_i) = 1$ or
$\deg(f_i) = 2$.

As a consequence, every polynomial of degree greater than or equal to 3, in $\mathbb{R}[x]$, can be factorised, i.e. it is reducible.

NOTE: In $\mathbb{R}$, there are both reducible and irreducible polynomials of degree 2:

eg: $\underbrace{x^2-1}_{(x+1)(x-1)}$, $\underbrace{x^2+3x+2}_{(x+1)(x+2)}$ are reducible

but x^2+1, x^2+17 are irreducible

How do we determine which degree 2 polynomials in $\mathbb{R}[x]$ are reducible?

We may use the following (general) result:

Remainder Theorem: Let $f(x)$ be in $k[x]$, for a field k, and let α be an element of k.

Then: $x-\alpha$ divides $f(x)$ iff α is a root of $f(x)$ in k (i.e. $f(\alpha)=0$)

Proof:

($\Rightarrow$) First, suppose that $x-\alpha$ divides $f(x)$:

i.e. $f(x) = (x-\alpha)\,g(x)$ for some $g(x)$ in $k[x]$.

Set $x=\alpha$ (evaluate both sides at $x=\alpha$):

$f(\alpha) = (\alpha-\alpha)\,g(\alpha) = 0 \cdot g(\alpha)$ i.e $f(\alpha)=0$ as required.

($\Leftarrow$) Suppose that $f(\alpha)=0$.

Apply Euclidean division to $f(x)$ and $x-\alpha$:

$\exists\ q(x), r(x)$ in $k[x]$ s.t:

$f(x) = q(x)(x-\alpha) + r(x)$ where $\deg(r) < \deg(x-\alpha)$
i.e $\deg(r) < 1$
i.e. $r(x)$ is a constant
$r(x) = r \in k$ say.

Evaluate both sides at $x=\alpha$:

$f(\alpha) = q(\alpha)\cdot(\alpha-\alpha) + r(\alpha) = q(\alpha)\cdot 0 + r$

i.e. $f(\alpha) = r$.

By assumption $f(\alpha)=0 \Rightarrow r=0$ (remainder is zero).

Hence, $f(x) = q(x)(x-\alpha) + 0 = q(x)(x-\alpha)$ i.e $(x-\alpha)$ divides f.
as required.

This completes the proof ▨

The idea of this proof/result is used in the following:

Suppose $f(x)$ is in $\mathbb{R}[x]$, and:

- $\deg(f) = 3$.
- $x^2+1 \mid f(x)$
- After dividing $f(x)$ by $x-1$, the remainder is 16.
  $x+2$, the remainder is -5.

Since $x^2+1 \mid f(x)$

$f(x) = (x^2+1)\,g(x)$ where $\deg(f) = \deg(x^2+1) + \deg(g)$

$3 \qquad 2$

so $\deg(g) = 1$.

i.e. $f(x) = (x^2+1)(a_1x + a_0)$ where $a_0, a_1 \in \mathbb{R}$, $a_1 \neq 0$ (*)

- Divide $f(x)$ by $(x-1)$ (the remainder is 16)

$f(x) = g(x)(x-1) + r$ where $r = 16$

$f(x) = g(x)(x-1) + 16$.

Then $f(1) = g(1)\cdot(1-1) = 16$ so $f(1) = 16$.

Similarly, $f(-2) = -5$.

In (*).

- Using $f(1) = 16$: $f(1) = (1^2+1)(a_1 1 + a_0) \Rightarrow 2(a_0 + a_1) = 16$

i.e. $\boxed{a_0 + a_1 = 8}$

- Using $f(-2) = -5$: $f(-2) = ((-2)^2+1)(a_1\cdot(-2) + a_0) \Rightarrow 5\cdot(-2a_1 + a_0) = -5$

i.e. $\boxed{-a_0 + 2a_1 = 1}$

By solving simultaneously; $a_0 = 5$, $a_1 = 3$

So, from (*) $F(x) = (x^2+1)(3x+5)$ $\Rightarrow$ i.e. $\boxed{f(x) = 3x^3 + 5x^2 + 3x + 5}$

October 17th 2019.

To complete the classification of irreducible polynomials in $\mathbb{R}[x]$, we use the following:

Proposition: Let $f(x)$ be in $K[x]$ (for a field K), and suppose $\deg(f) = 2 \Rightarrow f$ is reducible

$\Rightarrow f$ is reducible iff there exists α in K s.t $f(\alpha) = 0$ (i.e. α is a root of $f(x)$ in K).

Proof: (⇐) First suppose there is an α in K s.t $f(\alpha) = 0$.

$\Rightarrow$ by the Remainder Theorem $(x-\alpha) \mid f(x)$ i.e. $f(x) = g(x)\cdot(x-\alpha)$

Note: $\deg(f) = \deg(g) + \deg(x-\alpha)$

$2 \qquad\qquad 1$

It follows $\deg(g) = 1$.

$\Rightarrow$ the $\deg(x-\alpha) = 1$ and $\deg(g(x)) = 1$

Neither of $g(x)$, $x-\alpha$ is a unit (all units in $K[x]$ have degree 0).

So, the factorization $f(x) = g(x)(x-\alpha)$ does not involve units.

Hence, $f(x)$ is reducible, by definition

(⇒) Now, suppose $f(x)$ is reducible, i.e. that $f(x) = g(x)h(x)$ where $g(x), h(x)$ are not units.

(so that $\deg(g) \neq 0$ and $\deg(h) \neq 0$).

Consider degrees:

$\deg(f) = \deg(g) + \deg(h)$

Since $\deg(f) = 2$ and $\deg(g) \neq 0$, $\deg(h) \neq 0$

the only possibility is that $\deg(g) = 1$ (and $\deg(h) = 1$).

Then $g(x) = a_1x + a_0$ for $a_0, a_1,$ in K, $a_1 \neq 0$.

Consider $\alpha = \dfrac{-a_0}{a_1} = -a_0a_1^{-1}$ (a_1^{-1} exists, since $a_1 \neq 0$ and K is a field)

Then $g(\alpha) = a_1\cdot\left(\dfrac{-a_0}{a_1}\right) + a_0 = 0$. and

$f(\alpha) = g(\alpha)\,h(\alpha) = 0\cdot h(\alpha)$

So, there ~~doesn't~~ exists α in K s.t. $f(\alpha) = 0$.

as required

We may prove a similar result for degree 3:

If $f(x) \in K[x]$ and $\deg(f) = 3$

$f(x)$ is reducible $\iff$ there is an α in K s.t $f(\alpha) = 0$

Using these results, as well as results from last time, for $f(x)$ in $\mathbb{R}[x]$:

$f(x)$ is irreducible iff:

- $\deg(f) = 1$, or
- $\deg(f) = 2$ and there is no $\alpha \in \mathbb{R}$ s.t $f(\alpha) = 0$.

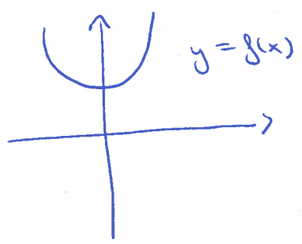
Note: In $\mathbb{Q}[x]$, $\exists$ irreducible polynomials for all positive degrees

e.g: $x^n - 2$ is irreducible in $\mathbb{Q}[x]$ for any n.

Consider $f(x) = x^4 + 1$

This can be factorised in terms of degree 1, polynomials in $\mathbb{C}[x]$:

$$x^4+1 = \left(x - \frac{\sqrt{2}}{2}(1+i)\right)\left(x - \frac{\sqrt{2}}{2}(1-i)\right)\left(x - \frac{\sqrt{2}}{2}(-1+i)\right)\left(x - \frac{\sqrt{2}}{2}(-1-i)\right)$$



In $\mathbb{R}[x]$, $f(x)$ factorises as:

$$x^4+1 = (x^2 - \sqrt{2}x + 1)(x^2 + \sqrt{2}x + 1)$$

$$\left(x - \frac{\sqrt{2}}{2}(1+i)\right)\left(x - \frac{\sqrt{2}}{2}(1-i)\right)\left(x - \frac{\sqrt{2}}{2}(-1+i)\right)\left(x - \frac{\sqrt{2}}{2}(-1-i)\right)$$

Note: In $\mathbb{Q}[x]$, $f(x)$ is in fact irreducible.

Let's now try to show that, in $K[x]$, every non-constant polynomial can be factorised uniquely in terms of irreducible polynomials.

Some preliminary results:

Proposition Suppose that $p(x)$ is a monic irreducible polynomial in $K[x]$.

Then, $\gcd(f, p) = 1$ or $\gcd(f, p) = p(x)$

Proof: Since $p(x)$ is irreducible, by definition:

if $p(x) = a(x)b(x)$ then $a(x)$ or $b(x)$ is a unit.

So, every factorization of $p(x)$ has the form

$$p(x) = u \cdot \left(\frac{1}{u} p(x)\right)$$ for u a unit in $K[x]$.
i.e. for u a non-zero constant.

The only divisors for $p(x)$ have one of the forms:

u or $\frac{1}{u} p(x)$ for u in K, $u \neq 0$.

So, the only monic divisors of $p(x)$ are 1 and $p(x)$.

Since $\gcd(f, p)$ must be monic polynomial and must also divide $p(x)$, we obtain:

$\gcd(f, p) = 1$ or $\gcd(f, p) = p(x)$ as required.

Proposition

Suppose that $p(x)$, $q(x)$ are monic irreducible polynomials in $K[x]$.

If $p(x) \mid q(x) \Rightarrow p(x) = q(x)$.

Proof: Applying the previous result.
Since p is monic, irreducible (set $f(x) = q(x)$ above)

$\gcd(q, p) = 1$ or $\gcd(q, p) = p(x)$.

But $p(x) \mid q(x)$, so we cannot have $\gcd(q, p) = 1$. It must be that $\gcd(q, p) = p(x)$.

Also, q is monic, irreducible. (set $f(x) = p(x)$ above):

$$\gcd(p, q) = 1 \quad \text{or} \quad \gcd(p, q) = q(x).$$

Since also, $\gcd(p, q) = \gcd(q, p) = p(x)$

Since $p(x) \neq 1$ ($p(x)$ is irreducible : $\deg(p) \geq 1$)
↳ since $p(x)$ is irreducible, whereas 1 is a non-zero constant (1 is a unit, not irreducible)

we must have $p(x) = q(x)$

We shall also make use of the following

Proposition:

Let $p(x)$ be a monic, irreducible polynomial in $K[x]$, and let $f(x), g(x)$ be in $K[x]$.

If $p(x) \mid f(x)\, g(x)$,

then, $p(x) \mid f(x)$ or $p(x) \mid g(x)$

Proof:

We assume $p(x) \mid f(x)\, g(x)$ ✱

We try to show that $p(x) \mid f(x)$ or $p(x) \mid g(x)$

If $p(x) \mid f(x)$, we are done

Otherwise, let's show that $p(x) \mid g(x)$

So suppose, $p(x)$ does not divide $f(x)$.

From previous result:

$\gcd(f, p) = 1$ or $\gcd(f, p) = p(x)$

But $p(x) \nmid f(x)$ so cannot have $\gcd(f, p) = p(x)$.

Hence, $\gcd(f, p) = 1$ (i.e. $f(x)$ and $p(x)$ are coprime).

By Bezout's Lemma:

$$a(x) f(x) + b(x) p(x) = 1$$
$$\text{↳ } \gcd(f, p)$$

for some $a(x), b(x)$ in $K[x]$.

Multiply through by $g(x)$:

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

Note: $p(x) \mid a(x) f(x) g(x)$ (since $p(x) \mid f(x) g(x)$ by assumption) ✱

Also, $p(x) \mid b(x) p(x) g(x) = b(x) g(x) p(x)$.

So $p(x) \mid a(x) f(x) g(x) + b(x) p(x) g(x)$

i.e. $p(x) \mid g(x)$ as required ▨.

By repeatedly applying this result, we obtain:

Proposition: Let $p(x)$ be a monic, irreducible polynomial in $K[x]$ and $f_1(x), \dots, f_n(x)$ be in $K[x]$.

If $p(x) \mid f_1(x) \cdots f_n(x) \Rightarrow p(x) \mid f_i(x)$ for some i, $1 \leq i \leq n$
(p divides at least one of the $f_i(x)$).

Proposition:

Suppose $a(x), b(x), c(x)$ are in $K[x]$, and $a(x) \cdot b(x) = a(x)\, c(x)$ where $a(x) \neq 0$ $\Rightarrow b(x) = c(x)$

Finally

Theorem. Suppose that $f(x)$ is a non-constant monic polynomial in $K[x]$ (i.e. $\deg(f) \geq 1$). Then $f(x)$ can be factorised as a product of irreducible polynomials in $K[x]$.

Further, the factorisation is unique (up to re-ordering)

In particular, if $f(x) = p_1(x) \cdots p_r(x)$ for monic, irreducible $p_1(x), \dots, p_r(x)$ and $f(x) = q_1(x) \cdots q_s(x)$ for monic irreducible $q_1(x), \dots, q_s(x)$

then $s = r$ and for each i, $1 \leq i \leq r$: $p_i(x) = q_j(x)$ for some j, $1 \leq j \leq s$.

Proof

By induction on degree of $f(x)$ (note $\deg(f) \geq 1$)

First, let's show the existence of a suitable factorization

INDUCTION

- Suppose $\deg(f) = 1$

Then, since $\deg(f) = 1$, $f(x)$ is irreducible itself

A suitable factorisation is $f(x) = f(x)$ or $f(x) = p_1(x)$ where $r = 1$

$p_1(x) = f(x)$ above.

- Now, suppose result holds for each $g(x)$ in $K[x]$ with $\deg(g) \leq n$

(assume, for each such $g(x)$, a suitable factorisation exists).

- Consider $f(x)$ s.t $\deg(f) = n+1$

  - If $f(x)$ is irreducible, then, as in the case of degree 1, a suitable factorisation is $f(x) = f(x)$.

  - Otherwise, if $f(x)$ is reducible

    [$f(x)$ cannot be a unit, $f(x)$ is not a non-zero constant]

    then $f(x) = g(x)h(x)$

    where neither of $g(x)$ or $h(x)$ is a unit.

    (Also, $f(x) \neq 0$, so $g(x) \neq 0$ and $h(x) \neq 0$)

    It follows that $\deg(g) \geq 1$, $\deg(h) \geq 1$.

Note: $\deg(f) = \deg(g) + \deg(h)$

$\deg(f) = n+1$, $\deg(g) \geq 1$, $\deg(h) \geq 1$

We may deduce that $\deg(g) \leq n$, $\deg(h) \leq n$.

So by inductive hypothesis:

$g(x) = p_1(x) \cdots p_t(x)$ for monic irreducible $p_1(x), \ldots, p_t(x)$.

$h(x) = q_1(x) \cdots q_l(x)$ for monic irreducible $q_1(x), \ldots, q_l(x)$

Then, a suitable factorisation of $f(x)$ is:

$f(x) = p_1(x) \cdots p_t(x) q_1(x) \cdots q_l(x)$

This completes proof for existence

October 21st 2019

Let us now show uniqueness of such a factorisation, as described in the statement of the theorem.

We prove this by induction on the degree of f, $\deg(f)$

- First, suppose that $\deg(f) = 1$. Then $f(x)$ is irreducible, so a suitable factorisation is $f(x) = p_1(x)$ (set $r = 1$, $p_1(x) = f(x)$)

Suppose also that

$f(x) = q_1(x) \cdots q_s(x)$ for monic irreducible $q_1(x), \ldots, q_s(x)$ in $K[x]$.

Since $q_j(x)$ is irreducible, for $1 \leq j \leq s$ $\deg(q_j) \geq 1$

Then:

$\deg(f) = \deg(q_1) + \cdots + \deg(q_s)$

$= 1 \qquad \geq 1 \qquad \geq 1$.

So, we must have $s = 1$ $f(x) = q_1(x)$

Then, $f(x) = p_1(x)$, $f(x) = q_1(x)$ so $q_1(x) = p_1(x)$

So, there is a unique factorisation if $\deg(f) = 1$ $f(x) = f(x)$.

- Now, suppose result holds for every polynomial $g(x)$ s.t $\deg(g) \leq n$. (any such $g(x)$ has a unique factorisation).

- Consider a monic $f(x)$ s.t $\deg(f) = n+1$, and suppose

$f(x) = p_1(x) \cdots p_r(x)$ for monic irreducible $p_1(x), \ldots, p_r(x)$

$f(x) = q_1(x) \cdots q_s(x)$ " " $q_1(x), \ldots, q_s(x)$

Note: that $p_1(x)$ divides $f(x) = p_1(x) \cdots p_r(x)$

So, $p_1(x)$ also divides $q_1(x) \cdots q_s(x)$

Then, from a previous result, $p_1(x)$ must divide $q_m(x)$ for some m, $1 \leq m \leq s$

Since $p_1(x)$ is monic and irreducible,

and $q_m(x)$ is monic and irreducible

and $p_1(x) \mid q_m(x)$

it follows, from another previous result, that $p_1(x) = q_m(x)$.

Then, from $p_1(x) \cdots p_r(x) = q_1(x) \cdots q_{m-1}(x)\, q_m(x)\, q_{m+1}(x) \cdots q_s(x)$

$= q_1(x) \cdots q_{m+1}(x) \cdot p_1(x) \cdot q_{m+1}(x) \cdots q_s(x)$

we may "cancel out" $p_1(x)$ to obtain

$$p_2(x) \cdots p_r(x) = q_1(x) \cdots q_{m-1}(x) \cdot q_{m+1}(x) \cdots q_s(x)$$

Note: $\deg(f) = \deg(p_1 \cdots p_r) = \deg(p_1) + \deg(p_2 \cdots p_r)$ (with $n+1$, ≥ 1, $\leq n$)

and $\deg(f) = n+1$, $\deg(p_1) \geq 1$ since $p_1(x)$ is irreducible

so, $\deg(p_2 \cdots p_r) \leq n$.

Then, the inductive hypothesis holds for $p_2(x) \cdots p_r(x)$

i.e. inductively from (*) we deduce that,

$r-1 = s-1$ and, for each i, $2 \leq i \leq r$: $p_i(x) = q_j(x)$ for some j, $1 \leq j \leq s$, $j \neq m$

Combining with $p_1(x) = q_m(x)$, we obtain $r = s$ and, for each i, $1 \leq i \leq r$: $p_i(x) = q_j(x)$ for some j: $1 \leq j \leq s$.

Hence, the factorisation of $f(x)$ is unique in the required manner.

This completes the proof ▨

There is a similar, more general, result for not necessarily monic polynomials

Theorem:

Suppose $f(x)$ is a polynomial in $K[x]$ s.t $\deg(f) \geq 1$.

Then, $\exists$ a factorisation of $f(x)$ of the form

$f(x) = c \cdot p_1(x) \cdots p_r(x)$ for monic irreducible $p_1(x), \ldots, p_r(x)$ and a non-zero constant c.

Such a factorisation is essentially unique (up to reordering of terms).

Examples of factorisations, in $\mathbb{Q}[x]$, $\mathbb{R}[x]$, $\mathbb{C}[x]$

- $x^2 - 1 = (x+1)(x-1)$ $(x+1), (x-1)$ irreducible in each of $\mathbb{Q}[x], \mathbb{R}[x], \mathbb{C}[x]$ (they have degree 1).
- $x^2 + 1$ roots are $i, -i$.

  so x^2+1 has no real or rational roots, so x^2+1 is irreducible in $\mathbb{C}[x]$: $x^2+1 = (x+i)(x-i)$ is a factorisation using irreducible polynomials.
- $f(x) = x^3 + 2x^2 + 2x + 4$

  By "inspection" $f(-2) = 0$ so $x+2$ divides $f(x)$.

  $f(x) = (x+2)(x^2+2)$

  Factorisation into irreducible polynomials:

  In $\mathbb{Q}[x]$: $f(x) = (x+2)(x^2+2)$

  $\mathbb{R}[x]$: $f(x) = (x+2)(x^2+2)$

  $\mathbb{C}[x]$: $f(x) = (x+2)(x+i\sqrt{2})(x-i\sqrt{2})$

  una raíz no es $\mathbb{Q}$ porque para ser $\mathbb{Q}$ tiene que ser un decimal que se pueda escribir en fracción, ej $\frac{1}{2} = 0{,}5$ es decir un decimal no infinito
- $g(x) = x^4 - 2x^2 + 8$

  Treating $g(x)$ as "a quadratic in x^2", we obtain

  $g(x) = (x^2 - 4\ \ )(x^2+2)$

Factorisation into irreducible polynomials:

In $\mathbb{Q}[x]$: $(x+2)(x-2)(x^2+2)$

In $\mathbb{R}[x]$ $(x+2)(x-2)(x^2+2)$

In $\mathbb{C}[x]$ $(x+2)(x-2)(x+i\sqrt{2})(x-i\sqrt{2})$.

CHAPTER 2 : Linear maps and the Jordan normal form

We begin by reviewing some notions related to:

- vector spaces
  - subspaces
  - span
  - linear independance
  - basis
  - dimension
  - Exchange lemma.
- Linear maps.
  - kernel
  - Image
  - Change of basis / different matrix representations
  - Eigenvalues / eigenvectors
  - Diagonalisability

October 22nd 2019

We begin with a review of some material.

(note: a more detailed review appears in the online notes, on the moodle page).

2.1. Review of vector spaces and linear maps.

Let us start with some (general) examples of vector spaces, over fields, with specified 'operations' of addition and scalar multiplication

(1) The set $k^n = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \; a_1, \dots, a_n \in k \right\}$ is a vector space over the field k, where:

- $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1+b_1 \\ \vdots \\ a_n+b_n \end{pmatrix}$ addition in k^n.
- $\lambda \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix}$ scalar multiplication (for $\lambda \in k$).

(2) The set $V = \{(a_1, \dots, a_n) \; a_1, \dots, a_n \in k\}$ is (also) a vector space over the field k, where:

- $(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1+b_1, \dots, a_n+b_n)$
- $\lambda(a_1, \dots, a_n) = (\lambda a_1, \dots, \lambda a_n)$.

(3) For any field k, with additive identity 0, the set $\{0\}$ or

Indeed any set composed of a 'zero vector', $\{0\}$, is a vector space over k.

This is a trivial vector space, where

$$0+0 = 0$$
$$\lambda \cdot 0 = 0 \quad \forall \lambda \text{ in } k.$$

(4) The set $k[x]$, of all polynomials over k, is a vector space over k, where:

- $(a_0+a_1x+\dots) + (b_0+b_1x+\dots) = (a_0+b_0)+(a_1+b_1)x+\dots$
- $\lambda(a_0+a_1x+\dots+a_nx^n) = (\lambda a_0)+(\lambda a_1)x+(\lambda a_2)x^2+\dots+(\lambda a_n)x^n$.

(5) The following subset of $k[x]$.

$k[x]_n = \{a_0+a_1x+\dots+a_nx^n : a_0, \dots, a_n \in k\}$.

is a vector space over k ('operations' defined as in $k[x]$)

eg: $\mathbb{R}[x]_2 = \{a_0+a_1x+a_2x^2 : a_0, a_1, a_2 \in \mathbb{R}\}$.

but it is not a ring eg: $x, x^2 \in \mathbb{R}[x]_2$ but $x \cdot x^2 = x^3 \notin \mathbb{R}[x]_2$

A subspace of a vector space V (over a field K) is a subset $\mathcal{U}$ of V s.t.

(1) $0 \in \mathcal{U}$ (0 is a zero vector from V)

(2) If $a, b \in \mathcal{U} \Rightarrow (a+b) \in \mathcal{U}$

(3) If $a \in \mathcal{U}$ and $\lambda \in K \Rightarrow (\lambda a) \in \mathcal{U}$

Examples

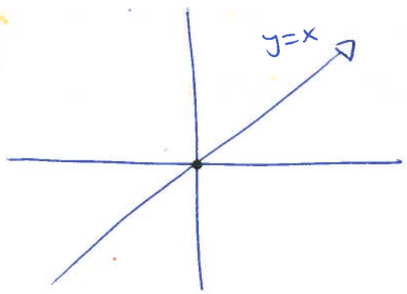
For any vector space V (over a field K), $\{0\}$ is a subspace of V over K, and the whole "space" V is also a subspace of V over K.

- Consider the vector space $\mathbb{R}^2$ over $\mathbb{R}$.

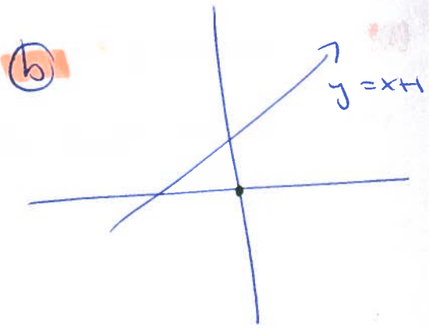
  Then, $\mathcal{U} = \left\{ \begin{pmatrix} x \\ x \end{pmatrix} : x \in \mathbb{R} \right\}$ is a subspace of $\mathbb{R}^2$ over $\mathbb{R}$. (a)

  whereas $\mathcal{U} = \left\{ \begin{pmatrix} x \\ x+1 \end{pmatrix} : x \in \mathbb{R} \right\}$ is not a subspace of $\mathbb{R}^2$ over $\mathbb{R}$ (b)

(a)



(b)



In a vector space / subspace V over K, a linear combination of vectors $v_1, \dots, v_m$ in V is an element (of V) of the form

$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_m v_m$ for $\lambda_1, \dots, \lambda_m$ in K.

eg: in $\mathbb{R}^2$ $\begin{pmatrix} 5 \\ 4 \end{pmatrix}$ is a linear combination of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$$\begin{pmatrix} 5 \\ 4 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 4 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In fact, for $\mathbb{R}^2$, every vector is a linear combination of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$ for some x_1, x_2 in $\mathbb{R}$.

We say that $\mathbb{R}^2 = \text{span}_{\mathbb{R}} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ or that $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ span $\mathbb{R}^2$ over $\mathbb{R}$.

In general,

For vectors $v_1, \dots, v_m$ in V (a vector space over K), the span of $v_1, \dots, v_m$ over K, $\text{span}_K \{v_1, \dots, v_m\}$, is the set of all linear combinations of $v_1, \dots, v_m$ over K:

$$\text{span}_K \{v_1, \dots, v_m\} = \{\lambda_1 v_1 + \dots + \lambda_m v_m : \lambda_1, \dots, \lambda_m \in K\}.$$

Note: For any such $v_1, \dots, v_m$ $\text{span}_K \{v_1, \dots, v_m\}$ is a subspace of V over K.

Given a subspace $\mathcal{U}$ of V, over K, we say that vectors $v_1, \dots, v_m$ span $\mathcal{U}$ over K if:

- Vectors $v_1, \dots, v_m \in \mathcal{U}$.
- Every vector in $\mathcal{U}$ is a linear combination of $v_1, \dots, v_m$

Then $\mathcal{U} = \text{span}_K \{v_1, \dots, v_m\}$.

eg every vector in $\mathcal{U} = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}$ (subspace of $\mathbb{R}^2$ over $\mathbb{R}$) is a linear combination of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$: $\begin{pmatrix} x \\ 0 \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

but $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \notin \mathcal{U}$: $\mathcal{U} \neq \text{span}_{\mathbb{R}} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$

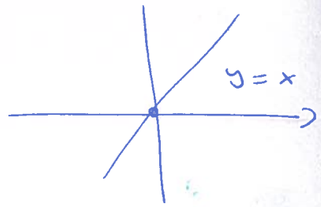
In fact: $\mathcal{U} = \text{span}_{\mathbb{R}} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$. $\begin{pmatrix} x \\ 0 \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

whereas, as indicated earlier $\text{span}_{\mathbb{R}} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} = \mathbb{R}^2$.

Also, $U = \text{span}_{\mathbb{R}}\left\{\begin{pmatrix}2\\0\end{pmatrix}\right\}$ or $U = \text{span}_{\mathbb{R}}\left\{\begin{pmatrix}x\\0\end{pmatrix}\right\}$ for $x \in \mathbb{R}$ $x \neq 0$.

Aim: to use the number of vectors in $\text{span}_K\{v_1, \dots, v_m\}$ to indicate the dimension of the (sub)space.

Example $U = \left\{\begin{pmatrix}x_1\\x_1\end{pmatrix} : x_1 \in \mathbb{R}\right\}$ is 1-dimensional



and $U = \text{span}_{\mathbb{R}}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$.

↳ 1 vector needed

Similarly: $\mathbb{R}^2 = \text{span}_{\mathbb{R}}\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$ $\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1\begin{pmatrix}1\\0\end{pmatrix} + x_2\begin{pmatrix}0\\1\end{pmatrix}$

2 dimensional — 2 vectors

but $\mathbb{R}^2 = \text{span}_{\mathbb{R}}\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right\}$ $\begin{pmatrix}x_1\\x_2\end{pmatrix} = (x_1-2)\begin{pmatrix}1\\0\end{pmatrix} + (x_2-2)\begin{pmatrix}0\\1\end{pmatrix} + 2\begin{pmatrix}1\\1\end{pmatrix}$.

October 24th 2019

Note: $\begin{pmatrix}1\\1\end{pmatrix}$ is "unnecessary"

$\begin{pmatrix}1\\1\end{pmatrix} = 1\begin{pmatrix}1\\0\end{pmatrix} + 1\begin{pmatrix}0\\1\end{pmatrix}$ $\Rightarrow \begin{pmatrix}1\\1\end{pmatrix}$ can be obtained from $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$

Equivalently: $1\begin{pmatrix}1\\0\end{pmatrix} + 1\begin{pmatrix}0\\1\end{pmatrix} - 1\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$

a "proper" combination of the three vectors goes to zero

vectors are dependent on each other.

Given a vector space V over a field K, vectors $v_1, \dots, v_m$ in V are linearly independent if the only solution to $\lambda_1 v_1 + \dots + \lambda_m v_m = 0$ (for $\lambda_1, \dots, \lambda_m$ in K) is the zero solution $\lambda_1 = 0, \dots, \lambda_m = 0$.

otherwise, if a non-zero solution exists to the given equation, $v_1, \dots, v_m$ are linearly dependent.

Notation $\{\underline{v}_1, \dots, \underline{v}_m\}$ linearly independent set iff $\underline{v}_1, \dots, \underline{v}_m$ are linearly independent.

Examples

- $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$ are linearly independent over $\mathbb{R}$: $\lambda_1\begin{pmatrix}1\\0\end{pmatrix} + \lambda_2\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$

$\iff \begin{pmatrix}\lambda_1\\\lambda_2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \iff \lambda_1 = 0, \lambda_2 = 0$.

- $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}$ are linearly dependent on $\mathbb{R}$

$1\begin{pmatrix}1\\0\end{pmatrix} + 1\begin{pmatrix}0\\1\end{pmatrix} - 1\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$

$\lambda_1\begin{pmatrix}1\\0\end{pmatrix} + \lambda_2\begin{pmatrix}0\\1\end{pmatrix} + \lambda_3\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$

has more solutions than $\lambda_1 = 0, \lambda_2 = 0, \lambda_3 = 0$

(e.g. $\lambda_1 = 1, \lambda_2 = 1, \lambda_3 = -1$).

Note for a single vector v:

$\{\underline{v}\}$ is linearly independent $\iff v \neq 0$

v is not the zero vector.

$\{\underline{0}\}$ is a linearly dependent set:

$\lambda \cdot 0 = 0$ for every λ (not just $\lambda = 0$).

More generally, any set of vectors that includes the zero vector is linearly dependent.

Equivalently, any linearly independent set does not contain a zero vector.

By combining "span" and "independence" we obtain:

Suppose $v_1, \dots, v_m$ are vectors in a vector space V over a field k.

Then, $\{\underline{v}_1, \dots, \underline{v}_m\}$ is a basis of (for) V, over k, if:

→ $v_1, \dots, v_m$ span V over k

→ $v_1, \dots, v_m$ are L.I over k.

The following important result holds:

Basis Theorem: Let $\mathcal{U}$ be a subspace of a vector space V, over a field k.

Then, there exists a basis of $\mathcal{U}$ over k, and any (two) such (finite) bases contain the same number of elements.

This allows us to define dimension "algebraically".

Given a subspace $\mathcal{U}$ of a vector space V, over k, the dimension of $\mathcal{U}$ over k, denoted by $\dim_k(\mathcal{U})$, is the number of vectors in a(ny) basis for $\mathcal{U}$ over k.

Examples:

- The trivial subspace, $\{0\}$, of any vector space, over any field k, is a space of dimension 0.

$$\dim_k(\{0\}) = 0.$$

We regard a basis as the empty set, $\emptyset$.

- The vector space $\mathbb{R}^2$ over $\mathbb{R}$ has a basis

$$\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\} \qquad \dim_{\mathbb{R}}(\mathbb{R}^2) = 2.$$

In general, for any field k and any $n \in \mathbb{N}$:

the vector space k^n over k has a "standard basis" over k:

$$\left\{\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}, \dots, \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}\right\} \qquad \dim_k(k^n) = n.$$

- The vector space $V = \{(a_1, \dots, a_n) : a_1, \dots, a_n \in k\}$ over k also satisfies $\dim_k(V) = n$; a "standard basis" here is:

$$\{(1,0,\dots,0), (0,1,0,\dots,0), \dots, (0,\dots,0,1)\}.$$

- The subspace $\mathcal{U} = \left\{\begin{pmatrix}x\\x\end{pmatrix} : x \in \mathbb{R}\right\}$ of $\mathbb{R}^2$ has basis $\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$ or

$\left\{\begin{pmatrix}2\\2\end{pmatrix}\right\}$ or, ..., over $\mathbb{R}$. so $\dim_{\mathbb{R}}(\mathcal{U}) = 1$.

- Consider $\mathbb{R}[x]_2 = \{a_0 + a_1x + a_2x^2 : a_0, a_1, a_2 \in \mathbb{R}\}$.

Then, we can check that $\{1, x, x^2\}$ is a basis for $\mathbb{R}[x]_2$ over $\mathbb{R}$.

$a_0 + a_1x + a_2x^2 = a_0 \cdot 1 + a_1 \cdot x + a_2x^2$ (span).

$\lambda_1 + \lambda_2x + \lambda_3x^2 = 0 \Rightarrow \lambda_1 = 0, \lambda_2 = 0, \lambda_3 = 0$ (L.I)

$$\dim_{\mathbb{R}}(\mathbb{R}[x]_2) = 3.$$

~~(- The vector space $\mathbb{R}^2$ over $\mathbb{R}$)~~

Some results involving notions related to basis:

Exchange Lemma

Let V be a vector over a field k. Suppose $\{u_1, u_2, \dots, u_m\}$ is linearly independent (over k), and $\{w_1, \dots, w_n\}$ spans V over k.

Then, ∃ a set that spans V, over k, of the form: $\{u_1, \dots, u_m, w'_{m+1}, \dots, w'_n\}$ where $w'_i \in \{w_1, \dots, w_n\}$ for each $m+1 \le i \le n$.

Now, suppose $\dim_K(V) = n$.

$\Rightarrow$ the following holds:

- If $\{\underline{v}_1, \dots, \underline{v}_n\}$ in V span V over K, then they are also linearly independent

  i.e. $\{\underline{v}_1, \dots, \underline{v}_n\}$ is a basis for V over K.

- If vectors $v_1, \dots, v_n$ in V are linearly independent over K, they also span V over K,

  i.e. $\{\underline{v}_1, \dots, \underline{v}_n\}$ is a basis for V over K.

- If U is a subspace of V over K, and $\dim_K(U) = n$ (also), then $U = V$.

Note on dimensions: The dimension depends on the field K, eg for

$\mathbb{C} = \{x + iy : x, y \in \mathbb{R}\} = \{z : z \in \mathbb{C}\}$.

$\dim_{\mathbb{R}}(\mathbb{C}) = 2$. with a basis over $\mathbb{R}$: $\{1, i\}$.

$\dim_{\mathbb{C}}(\mathbb{C}) = 1$ with a basis over $\mathbb{C}$: $\{1\}$.

Some operations on (sub) spaces:

Consider subspaces U, W of a vector space V over K.

- The intersection of U and W, denoted by $U \cap W$ is:

$$U \cap W = \{v \in V : v \in U \text{ and } v \in W\}.$$

- The sum of U and W, $U + W$, is:

$$U + W = \{v \in V : v = u + w \text{ for some } u \in U, w \in W\}.$$

Crucially: If U, W are subspaces of a vector space V over K, then:

- $U \cap W$ is also a subspace of V over K.
- $U + W$ is also a subspace of V over K.

Example: (in the setting of vector space $\mathbb{R}^3$ over $\mathbb{R}$)

$$U_1 = \left\{\begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix} : x_1 \in \mathbb{R}\right\} \qquad U_2 = \left\{\begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} : x_1, x_2 \in \mathbb{R}\right\}.$$

$$U_3 = \left\{\begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix} : x_3 \in \mathbb{R}\right\} \qquad U_4 = \left\{\begin{pmatrix} 0 \\ x_2 \\ x_3 \end{pmatrix} : x_2, x_3 \in \mathbb{R}\right\}.$$

Some corresponding bases:

For $U_1 = \left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right\}$ $\dim_{\mathbb{R}}(U_1) = 1$.

For $U_2 = \left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right\}$ $\dim_{\mathbb{R}}(U_2) = 2$.

For $U_3 = \left\{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ $\dim_{\mathbb{R}}(U_3) = 1$.

For $U_4 = \left\{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ $\dim_{\mathbb{R}}(U_4) = 2$.

Then $U_1 \cap U_2 = \left\{\begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix} : x_1 \in \mathbb{R}\right\} = U_1$

$U_1 + U_2 = U_2$.

Note:

$$\dim(U_1 + U_2) = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2)$$
$$2 = 1 + 2 - 1$$

Then $U_1 \cap U_3 = \left\{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right\} = \{0\}$.

$U_1 + U_3 = \left\{\begin{pmatrix} x_1 \\ 0 \\ x_3 \end{pmatrix} : x_1, x_3 \in \mathbb{R}\right\}$.

Note:

$$\dim(U_1 + U_3) = \dim(U_1) + \dim(U_3) - \dim(U_1 \cap U_3)$$
$$2 = 1 + 1 - 0.$$

Also, $U_2 \cap U_3 = \{0\}$

$$U_2 + U_3 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1, x_2, x_3 \in \mathbb{R} \right\} = \mathbb{R}^3$$

and

$$\dim(U_2 + U_3) = \dim(U_2) + \dim(U_3) - \dim(U_2 \cap U_3).$$
$$3 \qquad 2 \qquad 1 \qquad 0$$

$$U_2 \cap U_4 = \left\{ \begin{pmatrix} 0 \\ x_2 \\ 0 \end{pmatrix} : x_2 \in \mathbb{R} \right\}.$$

$$U_2 + U_4 = \mathbb{R}^3.$$

Note:

$$\dim(U_2 + U_4) = \dim(U_2) + \dim(U_4) - \dim(U_2 \cap U_4).$$

It holds in general that, for U, W subspaces of V over K:

$$\dim_K(U+W) = \dim_K(U) + \dim_K(W) - \dim_K(U \cap W).$$

Special case:

The subspace S of a vector space V over K is an internal direct sum of subspaces U, W (of V over K) if:

* $S = U + W$, and
* $U \cap W = \{0\}$

From our earlier example, two of the (four) sums are direct:

$U_1 + U_3$, $U_2 + U_3$

Notation: If S is a direct sum of U and W, we write

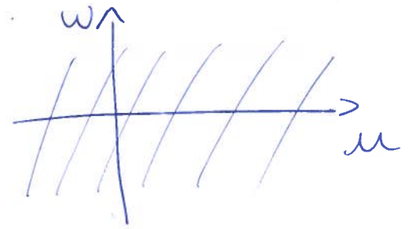
$$S = U \oplus W.$$

In the case of direct sums, the previous result on dimensions becomes:

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

$$U = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}, \quad W = \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix} : y \in \mathbb{R} \right\}.$$

$$U + W = \mathbb{R}^2.$$



At the "deeper" level of bases:

- If $U + W$ is direct ($U \cap W = \{0\}$), then, given any bases for U, W, say:

  $\{u_1, \dots, u_m\}$ for U over K $\quad \dim_K(U) = m$

  $\{w_1, \dots, w_n\}$ for W over K $\quad \dim_K(W) = n$.

  the set $\{u_1, \dots, u_m, w_1, \dots, w_n\}$ is a basis for $U \oplus W$ over K, $(\dim_K(U \oplus W) = m+n)$.

October 28th 2019

Let us now review material on maps "acting on vector spaces":

Let V, W be vector spaces over a field K. A linear map from V to W, over K, is a function $T: V \to W$ s.t.:

(1) T maps the zero vector from V to the zero vector in W

$$T(0_V) = 0_W \quad \text{or} \quad T(0) = 0$$

(2) for all $a, b \in V$, $T(a+b) = T(a) + T(b)$.

(3) for all $a \in V$, $\lambda \in K \Rightarrow T(\lambda a) = \lambda T(a)$.

Examples

- For any such V, W, over every field K, consider the zero map:
  $T: V \to W$, $T(v) = 0$ for each v in V. This is a linear map.

- for any vector space V over any field k, the identity map

$T: V \to V$, $T(v) = v$ for each v in V, is linear

- Consider $D: \mathbb{R}[x]_2 \to \mathbb{R}[x]_2$, defined via

$$D(a_0 + a_1 x + a_2 x^2) = a_1 + 2a_2 x \qquad D \iff \text{derivative}$$

Then D is a linear map.

Consider $T: \mathbb{R}^3 \to \mathbb{R}^2$ where $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - 4x_3 \\ 5x_2 \end{pmatrix}$

This is linear

(1) $T\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 - 4(0) \\ 5(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ so $T(0) = 0$ as required.

(2) For $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \in \mathbb{R}^3$ $\quad a+b = \begin{pmatrix} a_1+b_1 \\ a_2+b_2 \\ a_3+b_3 \end{pmatrix}$

$$T\begin{pmatrix} a_1+b_1 \\ a_2+b_2 \\ a_3+b_3 \end{pmatrix} = T(a+b) = \begin{pmatrix} (a_1+b_1) - 4(a_3+b_3) \\ 5(a_2+b_2) \end{pmatrix} = \begin{pmatrix} a_1 b_1 - 4a_3 - 4b_3 \\ 5a_2 + 5b_2 \end{pmatrix} =$$

$$= \begin{pmatrix} a_1 - 4a_3 \\ 5a_2 \end{pmatrix} + \begin{pmatrix} b_1 - 4b_3 \\ 5b_2 \end{pmatrix} = T(a) + T(b).$$

(3) For $a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \in \mathbb{R}^3$, $\lambda \in \mathbb{R}$: $\lambda a = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \lambda a_3 \end{pmatrix}$

$$T(\lambda a) = T\begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \lambda a_3 \end{pmatrix} = \begin{pmatrix} \lambda(a_1) - 4(\lambda a_3) \\ 5(\lambda a_2) \end{pmatrix} = \lambda \begin{pmatrix} a_1 - 4a_3 \\ 5a_2 \end{pmatrix}$$

so $T(\lambda a) = \lambda T(a)$ as required.

Note: $T: \mathbb{R}^2 \to \mathbb{R}^2$, $T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 1 \\ x_2 \end{pmatrix}$ is not linear

e.g. $T\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ condition (1) fails to hold.

Similarly, $T: \mathbb{R}^2 \to \mathbb{R}$ $\quad T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + x_2^2$, is not linear

e.g. $T\left(3\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = T\begin{pmatrix} 3 \\ 0 \end{pmatrix} = 3^2 + 0^2 = 9$ but $3T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 3 \cdot 1 = 3$

so $T\left(3\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) \neq 3T\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ condition (3) fails to hold.

Two important sets:

for a linear map $T: V \to W$ over k:

- The **kernel of T**, $\ker(T)$, is the set: $\{v \in V : T(v) = 0\}$.
- The **image of T**, $\mathrm{Im}(T)$, is the set: $\{w \in W : T(v) = w \text{ for } v \in V\}$.

Crucially, for a linear map $T: V \to W$ over k:

- The **kernel of T is a subspace of V over k**.
- The **image of T is a subspace of W over k**.

- The **dimension of $\ker(T)$** over k is the **nullity of T**.

$$\text{nullity}(T) = \dim_k(\ker(T))$$

- The **dimension of $\mathrm{Im}(T)$** over k is the **rank of T**.

$$\text{rank}(T) = \dim_k(\mathrm{Im}(T))$$

Example: $T: \mathbb{R}^3 \to \mathbb{R}^3$, $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \\ x_3 \end{pmatrix}$

Then, $\mathrm{Im}(T) = \left\{ \begin{pmatrix} x_1 \\ 0 \\ x_3 \end{pmatrix} : x_1, x_3 \in \mathbb{R} \right\}$ Basis over $\mathbb{R}$: $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

$\text{rank}_{\mathbb{R}}(T) = 2$.

Note: $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \\ x_3 \end{pmatrix}$ so $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \iff x_1 = 0, x_3 = 0$.

So $\ker(T) = \left\{ \begin{pmatrix} 0 \\ x_2 \\ 0 \end{pmatrix} : x_2 \in \mathbb{R} \right\}$ Basis over $\mathbb{R}$ $\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$.

$$\text{nullity}_{\mathbb{R}}(T) = 1 .$$

Note: $\text{null}_{\mathbb{R}}(T) + \text{rank}_{\mathbb{R}}(T) = 1 + 2 = 3$ ← dimension of the domain.

In general, for a linear map $T: V \to W$ over K:

$$\text{nullity}_K(T) + \text{rank}_K(T) = \dim(V).$$

This is the rank-nullity theorem

Note on notation when row reducing

We shall use:

- $R_i \longrightarrow R_i + \lambda R_j$ to denote this row operation corresponding to $E(i, j; \lambda)$.
- $R_i \longrightarrow \lambda R_i$: to denote operation corresponding to $D(i; \lambda)$
- $R_i \longleftrightarrow R_j$ to denote operation corresponding to $P(i; j)$.

For a given linear map $T: V \to W$, over K, we can determine a matrix expressing T once we choose bases for V and W over K.

Given bases $\{v_1, \dots, v_s\}$ for V over K.

$\{w_1, \dots, w_t\}$ for W over K.

we may find a suitable matrix using the following.

Apply T to each vector in $\{v_1, \dots, v_s\}$, obtaining $T(v_1), \dots, T(v_s)$.

Express each resulting answer in terms of $\{w_1, \dots, w_t\}$.

$$T(v_1) = m_{11} w_1 + m_{21} w_2 + \dots + m_{t1} w_t$$
$$T(v_2) = m_{12} w_1 + m_{22} w_2 + \dots + m_{t2} w_t$$
$$\vdots$$
$$T(v_s) = m_{1s} w_1 + m_{2s} w_2 + \dots + m_{ts} w_t .$$

Then, we form the $t \times s$ matrix $M = (m_{ij})$

Note: each row of coefficients becomes a column of M:

$$M = \begin{pmatrix} m_{11} & m_{12} & \dots & m_{1s} \\ & & & \\ m_{t1} & & \dots & m_{ts} \end{pmatrix}$$

Then, M is the required matrix $M = [T]^{\mathcal{B}_2}_{\mathcal{B}_1}$.

October 28th 2019

For a linear map $T: V \to W$ over a field K, given bases

$\mathcal{B}_1 = \{v_1, \dots, v_s\}$ for V over K.

$\mathcal{B}_2 = \{w_1, \dots, w_t\}$ for W over K.

there exists a $t \times s$ matrix $[T]^{\mathcal{B}_2}_{\mathcal{B}_1}$ representing T.

In our case, we assume all bases are ordered sets, in which case, given (ordered) bases $\mathcal{B}_1, \mathcal{B}_2$, as above, we obtain a specific (uniquely defined) matrix $[T]^{\mathcal{B}_2}_{\mathcal{B}_1}$

[Our bases are ordered sets, so that, for instance, we do not regard $A = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ and $B = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$ as being the same basis (for some vector space over some field)]

Some examples:

Consider an identity map $Id: \mathbb{R}^2 \to \mathbb{R}^2$, $Id \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

and bases $\mathcal{B} = \left\{ \underbrace{\begin{pmatrix} 1 \\ 3 \end{pmatrix}}_{b_1}, \underbrace{\begin{pmatrix} 0 \\ -2 \end{pmatrix}}_{b_2} \right\}$, $\Sigma = \left\{ \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{e_1}, \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{e_2} \right\}$

Let's find $[Id]^{\Sigma}_{\mathcal{B}}$ $Id \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$$Id \begin{pmatrix} 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (-2) \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

So $[Id]_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & 0 \\ 3 & -2 \end{pmatrix}$ has vectors of $\mathcal{B}$ as columns (in order).

In general, given an identity map, $Id : k^n \to k^n$ for some field k, then given a basis $\mathcal{B} = \{b_1, \dots, b_n\}$ for k^n over k and the standard basis $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}$ the

matrix $[Id]_{\mathcal{B}}^{\mathcal{E}}$ consists of the vectors of $\mathcal{B}$ (in order) placed as columns in a matrix

Consider $Id : k^n \to k^n$ again, and any basis $\mathcal{B} = \{b_1, \dots, b_n\}$ for k^n over k.

$$Id(b_1) = b_1 = 1 \cdot b_1 + 0 \cdot b_2 + \dots + 0 \cdot b_n$$
$$Id(b_2) = b_2 = 0 \cdot b_1 + 1 \cdot b_2 + \dots + 0 \cdot b_n$$
$$\vdots$$
$$Id(b_n) = b_n = 0 \cdot b_1 + \dots + 1 b_n$$

So $[Id]_{\mathcal{B}}^{\mathcal{B}} = I_n$.

Consider $T : \mathbb{R}^2 \to \mathbb{R}^2$, where $T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ 2x_2 \end{pmatrix}$

and let $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$, $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \end{pmatrix} \right\}$ bases for $\mathbb{R}^2$ over $\mathbb{R}$

Let's find $[T]_{\mathcal{E}}^{\mathcal{E}}$ $\quad T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$$T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$[T]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}.$$

Let's also find $[T]_{\mathcal{B}}^{\mathcal{E}}$

$$T\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix} = 4 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 6 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

$$T\begin{pmatrix} 0 \\ -2 \end{pmatrix} = \begin{pmatrix} -2 \\ -4 \end{pmatrix} = -2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 4 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$[T]_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 4 & -2 \\ 6 & -4 \end{pmatrix}.$$

Note: $[T]_{\mathcal{E}}^{\mathcal{E}} [Id]_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 3 & -2 \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ 6 & -4 \end{pmatrix} = [T]_{\mathcal{B}}^{\mathcal{E}}$.

More generally, for linear maps $T : U \to V$, $S : V \to W$ (over a field k), with bases $\mathcal{A}, \mathcal{B}, \mathcal{C}$ for U, V, W respectively, then:

$$[S]_{\mathcal{B}}^{\mathcal{C}} [T]_{\mathcal{A}}^{\mathcal{B}} = [S \circ T]_{\mathcal{A}}^{\mathcal{C}}$$

where $S \circ T : U \to W$ defined via $(S \circ T)(u) = S(T(u))$ for every u in U.

Some related results (for $U = V = W = k^n$ for some field k)

- If S, T are identity maps and $\mathcal{A} = \mathcal{C}$, then:

$$[Id]_{\mathcal{B}}^{\mathcal{A}} [Id]_{\mathcal{A}}^{\mathcal{B}} = \underbrace{[Id]_{\mathcal{A}}^{\mathcal{A}}}_{I_n} \quad \text{(and } [Id]_{\mathcal{A}}^{\mathcal{B}} [Id]_{\mathcal{B}}^{\mathcal{A}} = I_n \text{ also)}$$

So $[Id]_{\mathcal{A}}^{\mathcal{B}}$, $[Id]_{\mathcal{B}}^{\mathcal{A}}$ are invertible and $\left([Id]_{\mathcal{B}}^{\mathcal{A}}\right)^{-1} = [Id]_{\mathcal{A}}^{\mathcal{B}}$

$$\left([Id]_{\mathcal{A}}^{\mathcal{B}}\right)^{-1} = [Id]_{\mathcal{B}}^{\mathcal{A}}.$$

Special case:

As noted earlier, $[Id]_{\mathcal{B}}^{\mathcal{E}}$ consists of vectors in $\mathcal{B}$ as columns

This is invertible.

More generally A set of vectors $\mathcal{B} = \{b_1, \dots, b_n\}$ in k^n over k, is a basis for k^n over k, iff the matrix we obtain by placing vectors in $\mathcal{B}$ as columns is an invertible matrix $\left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \end{pmatrix} \right\}$ is a basis for $\mathbb{R}^2$ over $\mathbb{R}$. $\langle = \rangle$

$\Leftrightarrow \begin{pmatrix} 1 & 0 \\ 3 & -2 \end{pmatrix}$ is invertible

By twice applying the general rule above, given a linear map $T: k^n \to k^n$, we obtain:

$$[Id]^B_A \cdot [T]^A_A \cdot [Id]^A_B = [Id]^B_A \cdot \left([T]^A_A \cdot [Id]^A_B \right) =$$

$$= [Id]^B_A \cdot [T]^A_B$$

$$[Id]^B_A \, [I]^A_A \, [Id]^A_B = [T]^B_B$$

Note: $[Id]^A_B$ and $[Id]^B_A$ are inverses for each other.

In general, given a linear map $T: k^n \to k^n$ over a field k, and bases A, B for k^n over k, $\exists$ an invertible $r \times n$ matrix P s.t.

$$P^{-1} [T]^A_A P = [T]^B_B \quad (\text{set } P = [Id]^A_B \text{ above})$$

2.2. EIGENVALUES, EIGENVECTORS AND MINIMAL POLYNOMIALS

Consider a linear map $T: V \to V$ for a V a vector space over $\mathbb{C}$.

- An eigenvector of T is a non-zero vector v in V s.t.

  $T(v) = \lambda v$ for some complex number λ.

  The number λ is the eigenvalue corresponding to v.

Note: A zero vector is not defined as an eigenvector (but $\lambda = 0$ is a possible eigenvalue).

$$T(v) = 0 \cdot v = 0.$$

If we do include the zero vector in a set with all eigenvectors for a given eigenvalue, we obtain

$$V_1(\lambda) = \{ v \in V : T(v) = \lambda v \}$$ ← including $v = 0$

This is the eigenspace corresponding to λ.

Crucially, $V_1(\lambda)$ forms a subspace of the vector space V.

November 1ST 2019

From last time:

eigenvalues, eigenvectors, eigenspaces (in context of linear maps over $\mathbb{C}$)

These notations carry over to the setting of (individual) matrices

for an $n \times n$ matrix M, over $\mathbb{C}$:

An eigenvector of M is a non-zero vector v in $\mathbb{C}^n$ s.t. $Mv = \lambda v$ for some $\lambda \in \mathbb{C}$.

↑ the associated eigenvalue.

Related eigenspace: $V_1(\lambda) = \{ v \in \mathbb{C}^n : Mv = \lambda v \}$.

Note: $V_1(\lambda) = \{ v \in \mathbb{C}^n : Mv = \lambda v \} = \{ v \in \mathbb{C}^n : Mv = \lambda I_n v \} =$

$$= \{ v \in \mathbb{C}^n : (M - \lambda I_n) v = 0 \}$$

i.e. $V_1(\lambda) = \ker(M - \lambda I_n)$.

$V_1(\lambda)$ is a subspace of $\mathbb{C}^n$.

Example: If $M = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$, then $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector, for eigenvalue 2.

$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is an eigenvector for eigenvalue 3.

$$M \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \checkmark$$

$$M \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 2 \end{pmatrix} \checkmark$$

Note: λ is an eigenvalue of $M \iff \exists\, v \in \mathbb{C}^n$ s.t. $v \neq 0$ and $Mv = \lambda v$

$\iff \exists\, v \in \mathbb{C}^n$ s.t. $v \neq 0$ and $(M - \lambda I_n) v = 0$

$\iff (M - \lambda I_n) x = 0$ does not have a unique solution (eg $x = 0$, $x = v$ both work)

$\iff (M - \lambda I_n)$ is not invertible

$\iff \det(M - \lambda I_n) = 0$.

So, the solutions to the characteristic equation

$\det(M - tI_n) = 0$ are precisely the eigenvalues of M.

$\det(M - tI_n)$: characteristic polynomial

May also use such an equation to find eigenvalues of a linear map.

$T: \mathbb{C}^n \to \mathbb{C}^n$.

Choose any basis $\mathcal{A}$ for $\mathbb{C}^n$ over $\mathbb{C}$ and find eigenvalues of matrix $[T]_{\mathcal{A}}^{\mathcal{A}}$ (more on this later)

In our course, we assume some basic methods/results involving determinants:

- $\det \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = A_{11}A_{22} - A_{12}A_{21}$
- may use cofactor expansion method.
- For $n \times n$ matrices A, B:

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

Example

If $M = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ then $tI_2 - M = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} = \begin{pmatrix} t-1 & -1 \\ 2 & t-4 \end{pmatrix}$

So $\det(tI_2 - M) = \begin{vmatrix} t-1 & -1 \\ 2 & t-4 \end{vmatrix} = (t-1)(t-4) - (-1)\cdot 2$

i.e. $\det(tI_2 - M) = t^2 - 5t + 4 + 2 = t^2 - 5t + 6$.

So, characteristic equation of M is:

$t^2 - 5t + 6 = 0$ i.e. $(t-2)(t-3) = 0$

Hence, M has eigenvalues 2 and 3 (as indicated earlier)

Consider $A = \begin{pmatrix} -4 & 1 & 2 \\ 8 & -2 & -4 \\ 0 & 0 & 0 \end{pmatrix}$

$$\det(tI_3 - A) = \begin{vmatrix} t+4 & -1 & -2 \\ -8 & t+2 & 4 \\ 0 & 0 & t \end{vmatrix} =$$

(use cofactor expansion along row 3)

$$= t \cdot \begin{vmatrix} t+4 & -1 \\ -8 & t+2 \end{vmatrix} = t \cdot \left((t+4)(t+2) - 8\right) =$$

$$= t(t^2 + 6t).$$

So $\det(tI_3 - A) = t^2(t+6)$

It follows that A has eigenvalues 0 and -6.

Consider $M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ $\det(tI - M) = \begin{vmatrix} t & 1 \\ -1 & t \end{vmatrix}$

i.e. $\det(tI - M) = t^2 + 1 = 0$.

$\Rightarrow$ eigenvalues are $+i, -i$. (OK: here, we are working over $\mathbb{C}$)

Helpful result (over $\mathbb{C}$)

For a given $n \times n$ matrix M or linear map $T: \mathbb{C}^n \to \mathbb{C}^n$

If v_1, v_2 are eigenvectors corresponding to distinct eigenvalues λ_1, λ_2 respectively, then v_1, v_2 are L.I.

This helps when trying to diagonalise.

For a linear map $T: \mathbb{C}^n \to \mathbb{C}^n$, we say T is diagonalisable if $\exists$ a basis $\mathcal{A}$ of $\mathbb{C}^n$ over $\mathbb{C}$ s.t. the matrix $[T]^{\mathcal{A}}_{\mathcal{A}}$ is diagonal

Note: in such case, if $\mathcal{B}$ is another basis for $\mathbb{C}^n$ over $\mathbb{C}$:

$$[Id]^{\mathcal{A}}_{\mathcal{B}} [T]^{\mathcal{B}}_{\mathcal{B}} [Id]^{\mathcal{B}}_{\mathcal{A}} = [T]^{\mathcal{A}}_{\mathcal{A}}$$

($[T]^{\mathcal{A}}_{\mathcal{A}}$ is diagonal)

P^{-1} M P

In general: an $n \times n$ matrix M is diagonalisable if $\exists$ an invertible $n \times n$ matrix P s.t. $P^{-1}MP$ is diagonal matrix.

We can use any matrix of the form $[T]^{\mathcal{E}}_{\mathcal{E}}$ (for any suitable basis $\mathcal{E}$) to find the eigenvalues of T itself.

This follows from

Proposition: Suppose $T: V \to V$ is a linear map over $\mathbb{C}$, and let $\mathcal{A}, \mathcal{B}$ be bases for V over $\mathbb{C}$.

Then, $\det(tI - [T]^{\mathcal{A}}_{\mathcal{A}}) = \det(tI - [T]^{\mathcal{B}}_{\mathcal{B}})$

Equivalently, $ch_{[T]^{\mathcal{A}}_{\mathcal{A}}}(t) = ch_{[T]^{\mathcal{B}}_{\mathcal{B}}}(t)$

Notation for characteristic polynomials of matrix $[T]^{\mathcal{A}}_{\mathcal{A}}$ and $[T]^{\mathcal{B}}_{\mathcal{B}}$

Proof:

$$ch_{[T]^{\mathcal{B}}_{\mathcal{B}}}(t) = \det(tI - [T]^{\mathcal{B}}_{\mathcal{B}})$$

$$= \det(tI - P^{-1}[T]^{\mathcal{A}}_{\mathcal{A}} P)$$

($\exists$ such an invertible matrix P)

$$= \det(tP^{-1}IP - P^{-1}[T]^{\mathcal{A}}_{\mathcal{A}} P)$$

using $P^{-1}IP = I$.

$$= \det(P^{-1}(tI)P - P^{-1}[T]^{\mathcal{A}}_{\mathcal{A}} P)$$

$$= \det(P^{-1}(tI - [T]^{\mathcal{A}}_{\mathcal{A}})P)$$

Using $\det(AB) = \det A \cdot \det B$:

$$= \det(P^{-1}) \det(tI - [T]^{\mathcal{A}}_{\mathcal{A}}) \det P.$$

$\det(P^{-1})\det(P) = \det(P^{-1}P) = \det(I) = 1$.

$$= \det(tI - [T]^{\mathcal{A}}_{\mathcal{A}}) \cdot \det(P^{-1}) \det(P)$$

$$= \det(tI - [T]^{\mathcal{A}}_{\mathcal{A}})$$

so $ch_{[T]^{\mathcal{B}}_{\mathcal{B}}}(t) = ch_{[T]^{\mathcal{A}}_{\mathcal{A}}}(t)$ as required

Example:

Consider $T: \mathbb{C}^2 \to \mathbb{C}^2$

where $T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ -2x_1 + 4x_2 \end{pmatrix}$

In terms of $\mathcal{E} = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$ $[T]^{\mathcal{E}}_{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$

From earlier $ch_{[T]^{\mathcal{E}}_{\mathcal{E}}}(t) = (t-2)(t-3)$

Eigenvalues: 2, 3

Corresponding eigenvectors: $x\begin{pmatrix}1\\1\end{pmatrix}, x\begin{pmatrix}1\\2\end{pmatrix}$

For $x \in \mathbb{C}, x \neq 0$, $y \in \mathbb{C}, y \neq 0$.

Consider basis: $\mathcal{C} = \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\2\end{pmatrix}\right\}$

$\updownarrow$ eigenvectors of $[T]^{\mathcal{E}}_{\mathcal{E}}$

Let's find $[T]^{\mathcal{C}}_{\mathcal{C}}$.

$$T\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}2\\2\end{pmatrix} = 2 \cdot \begin{pmatrix}1\\1\end{pmatrix} + 0 \cdot \begin{pmatrix}1\\2\end{pmatrix}$$

$$T\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}3\\6\end{pmatrix} = 0 \cdot \begin{pmatrix}1\\1\end{pmatrix} + 3 \cdot \begin{pmatrix}1\\2\end{pmatrix}$$

So $[T]_{\mathcal{C}}^{\mathcal{C}} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$

$$ch_{[T]_{\mathcal{C}}^{\mathcal{C}}}(t) = \begin{vmatrix} t-2 & 0 \\ 0 & t-3 \end{vmatrix} = (t-2)(t-3) = ch_{[T]_{\mathcal{E}}^{\mathcal{E}}}(t)$$

Note $[Id]_{\mathcal{E}}^{\mathcal{C}} [T]_{\mathcal{E}}^{\mathcal{E}} [Id]_{\mathcal{C}}^{\mathcal{E}} = [T]_{\mathcal{C}}^{\mathcal{C}}$

i.e. here,

$$\underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}^{-1}}_{P^{-1}} \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}}_{P} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

P comprised of the (eigen) vectors in basis $\mathcal{C}$ placed as columns in a matrix.

A result involving matrices:

For any $n \times n$ matrix M over $\mathbb{C}$, M satisfies its own characteristic polynomial. $ch_M(M) = 0$.

CAYLEY-HAMILTON THEOREM for matrices

In the context of linear maps:

If $T: \mathbb{C}^n \to \mathbb{C}^n$ is a linear map

$ch_T(T) = 0$, i.e. for any basis $\mathcal{A}$ for $\mathbb{C}^n$ over $\mathbb{C}$:

$$ch_{[T]_{\mathcal{A}}^{\mathcal{A}}}\left([T]_{\mathcal{A}}^{\mathcal{A}}\right) = 0.$$

—

Examples from earlier:

If $M = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ $ch_M(t) = (t-2)(t-3)$

Then $(M-2I)\cdot(M-3I) =$

i.e. $= \begin{pmatrix} -1 & 1 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} -2 & 1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

If $M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $\Rightarrow ch_M(t) = (t-1)^2$.

$\Rightarrow (M-I)^2 = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ as required

Note for $M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $M - I = 0$.

There is a "smaller" polynomial (than $ch_M(t) = (t-1)^2$) that "sends M to zero".

In general:

For a linear map $T: V \to V$ over $\mathbb{C}$, $m_T(t)$ is a minimal polynomial for T if $m_T(t)$ is a monic polynomial in $\mathbb{C}[t]$ s.t.:

- $m_T(T) = 0$
- If, for some $f(t) \in \mathbb{C}[t]$, $f(T) = 0$
  $\Rightarrow f(t) = 0$ or $\deg(m_T) \leq \deg(f)$

Some properties of minimal polynomials.

Given a linear map $T: \mathbb{C}^n \to \mathbb{C}^n$, there is a unique minimal polynomial for T.

Proof Suppose $m_1(t)$, $m_2(t)$ are minimal polynomials for T.

Then $m_1(t)$, $m_2(t)$ are monic, so are non-zero polynomials.

Also, $m_1(T) = 0$

Since $m_2(t)$ is a minimal polynomial and $m_1(T) = 0$, where $m_1(t) \neq 0$, where $m_1(t) \neq 0$, we obtain, by definition:

$\deg(m_2) \leq \deg(m_1)$

(substitute $m_T = m_2$, $f = m_1$ in definition)

Similarly, $m_2(T) = 0$ and $m_2(t) \neq 0$, so

$\deg(m_1) \leq \deg(m_2)$

Overall $\deg(m_1) = \deg(m_2)$

We may then write:

$m_1(t) = t^n + a_{n-1}t^{n-1} + \dots + a_1 t + a_0$.

$m_2(t) = t^n + b_{n-1}t^{n-1} + \dots + b_1 t + b_0$.

Then, $\underbrace{m_1(t) - m_2(t)}_{(m_1-m_2)(t)} = (a_{n-1} - b_{n-1})t^{n-1} + \dots + (a_0 - b_0)$

Then $(m_1 - m_2)(T) = m_1(T) - m_2(T) = 0 - 0 = 0$.

By definition of minimal polynomial:

either $(m_1 - m_2)(t) = 0$ (the zero polynomial)

or

$\deg(m_1) \leq \deg(m_1 - m_2)$

Here, since $\deg(m_1) = n$

and $\deg(m_1 - m_2) \leq n-1$.

we cannot have $\deg(m_1) \leq \deg(m_1 - m_2)$

So, it must be the case that $(m_1 - m_2)(t) = 0$

i.e $m_1(t) - m_2(t) = 0$.

Hence, $m_1(t) = m_2(t)$

The minimal polynomial of T is unique, as required

November 11th 2019

Another helpful result:

If, for some $f(t)$ in $\mathbb{C}[t]$: $f(T) = 0$, then the minimal polynomial $m_T(t)$ say, divides $f(t)$: $m(t) \mid f(t)$

Proof: Apply Euclidean division to $f(t)$ and $m(t)$

(note: $m(t) \neq 0$ by definition; $m(t)$ is monic).

There exists (unique) $q(t), r(t)$ in $\mathbb{C}[t]$ s.t $f(t) = q(t)m(t) + r(t)$

where $\deg(r) < \deg(m)$.

$\Rightarrow f(T) = q(T)m(T) + r(T)$

so $0 = q(T) \cdot 0 + r(T)$

($f(T) = 0$ by assumption; $m(t)$ is minimal polynomial since $\deg(r) < \deg(m)$)

Hence, $r(T) = 0$.

Then, by definition of minimal polynomial, we must have $r(t) = 0$, the zero polynomial

(The def says that if $R(T) = 0 \Rightarrow$ either $r(t) = 0$ or $\deg(m_T) < \deg(r_T)$)

$\Rightarrow f(t) = q(t)m(t) + 0 = q(t)m(t)$

So $m(t)$ divides $f(t)$ as required

Finally,

- If λ is an eigenvalue of T, then λ is a root of the minimal polynomial of T, i.e $m_T(\lambda) = 0$.

Proof: Suppose v is an eigenvector corresponding to λ. Then $v \neq 0$ and $T(v) = \lambda v$.

for any $k \in \mathbb{N}$: $T^k(v) = T^{k-1}(T(v)) =$

$= T^{k-1}(\lambda v) =$

$= \lambda T^{k-1}(v) =$

$= \lambda T^{k-2}(T(v)) =$

$= \lambda T^{k-2}(\lambda v) =$

$= \lambda^2 T^{k-2}(v)$.

$\Rightarrow T^k(v) = \lambda^k v$ for $k = 1, 2, 3, \dots$ $(*)$

Consider the minimal polynomial:

$$m(t) = t^m + a_{m-1}t^{m-1} + \dots + a_1 t + a_0 \cdot 1$$

$$\Rightarrow m(T)(v) = \left(T^m + a_{m-1}T^{m-1} + \dots + a_1 T + a_0 \cdot Id\right)(v)$$

$$= T^m(v) + a_{m-1}T^{m-1}(v) + \dots + a_1 T(v) + a_0 Id(v)$$

So, using (*):

$$m(T)(v) = \lambda^m v + a_{m-1}\lambda^{m-1} v + \dots + a_1 \lambda v + a_0 v$$

$$= (\lambda^m + a_{m-1}\lambda^{m-1} + \dots + a_1 \lambda + a_0)\, v$$

$$m(T)(v) = m_T(\lambda)\, v.$$

Since $m(T) = 0$, $m_T(\lambda)\, v = 0$. By assumption, $v \neq 0$, so must have $m_T(\lambda) = 0$, i.e. λ is a root of $m_T(t)$ as required

Let's now consider some examples:

(using matrices)

$A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ $\quad ch_A(t) = (t-2)(t-3)$ from earlier.

Using previous results, the options for $m_A(t)$ (minimal polynomial) are: $(t-2)(t-3)$, ~~$t-2$~~, ~~$(t-2)^2$~~, ~~$(t-2)^3$~~

So, $m_A(t) = (t-2)(t-3)$

are not the minimal polynomial because they don't include all the possible eigenvalues, roots.

ex: $t-2$ has root 2 but not 3.

Consider $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 0 \\ 7 & 1 \end{pmatrix}$

$ch_B(t) = (t-1)^2$ $\quad ch_C(t) = (t-1)^2$

For either B or C, the minimal polynomial is one of

- $f_1(t) = t-1$ ← Always try the smallest degree 1st and if it works you're done, if not keep trying
- $f_2(t) = (t-1)^2$

$f_1(B) = B - I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ Hence $m_B(t) = t-1$

$f_1(C) = C - I = \begin{pmatrix} 1 & 0 \\ 7 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 7 & 0 \end{pmatrix} \neq 0$.

Since $f_1(C) \neq 0$, $f_2(t)$ must be the minimal polynomial of C.

Check: $f_2(C) = (C-I)^2 = \begin{pmatrix} 0 & 0 \\ 7 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 7 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ as required.

$M = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$ $\quad ch_M(t) = (t-5)(t-3)^2$.

options from $m_M(t) = (t-5)(t-3), (t-5)(t-3)^2$
$\qquad\qquad\qquad\qquad\quad f_1(t) \qquad\quad f_2(t)$

$f_1(M) = (M-5I)(M-3I) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{pmatrix}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} \neq 0$.

$f_1(M) \neq 0$, so $m_M(t) = f_2(t) = (t-5)(t-3)^2$.

$N = \begin{pmatrix} -4 & 1 & 2 \\ 8 & -2 & -4 \\ 0 & 0 & 0 \end{pmatrix}$ $\quad ch_N(t) = t^2(t+6)$.

Choices for $m_N(t) = t\cdot(t+6),\ t^2(t+6)$.
$\qquad\qquad\qquad\quad f_1(t) \qquad\quad f_2(t)$

$f_1(N) = N(N+6I) = \begin{pmatrix} -4 & 1 & 2 \\ 8 & -2 & -4 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 2 & 1 & 2 \\ 8 & 4 & 4 \\ 0 & 0 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

So $f_1(t)$ is the minimal polynomial $m_N(t) = t\cdot(t+6)$.

November 12th 2019

Suppose a matrix / linear map has characteristic polynomial $(t+1)^3(t-9)^2$.

Then, we may determine the relevant minimal polynomial by listing the possible choices, in some order of non-decreasing degree.

e.g:

① $(t+1)(t-9)$ → degree 2.

② $(t+1)^2(t-9)$ } degree 3.

③ $(t+1)(t-9)^2$ } degree 3.

④ $(t+1)^3(t-9)$ → degree 4.

⑤ $(t+1)^2(t-9)^2$ → degree 4.

⑥ $(t+1)^3(t-9)^2$ → degree 5.

and then proceeding to go through the list, in order, until we find the first polynomial that "sends the matrix / linear map to zero" (that should be the minimal polynomial).

2.3. The Jordan normal form

Earlier, we showed that for a linear map $T: V \to V$, over $\mathbb{C}$, and given any bases $\mathcal{A}, \mathcal{B}$ for V over $\mathbb{C}$.

$$ch\left([T]^{\mathcal{A}}_{\mathcal{A}}\right) = ch\left([T]^{\mathcal{B}}_{\mathcal{B}}\right)$$

It follows that: $ch\left([T]^{\mathcal{A}}_{\mathcal{A}}\right) = 0 \iff ch\left([T]^{\mathcal{B}}_{\mathcal{B}}\right) = 0$.

In fact, more generally, for any polynomial $f(t)$ in $\mathbb{C}[t]$:

$$f\left([T]^{\mathcal{A}}_{\mathcal{A}}\right) = 0 \iff f\left([T]^{\mathcal{B}}_{\mathcal{B}}\right) = 0.$$

(Proof: given in notes NFE)

So, in particular, we can associate a minimal polynomial to a linear map and not just a specific matrix.

Goal for rest of the chapter:

Given such a linear map $T: V \to V$

- If T is diagonalisable, find a basis $\mathcal{B}$ s.t $[T]^{\mathcal{B}}_{\mathcal{B}}$ is diagonal
- If T is not diagonalisable, find a basis $\mathcal{C}$ s.t $[T]^{\mathcal{C}}_{\mathcal{C}}$ is "almost" diagonal.

(T can be not diagonalisable when there are not enough linearly independent eigenvectors).

In the latter case, we find extra more "generalised eigenvectors".

For an $n \times n$ complex matrix M, the r^{th} generalised eigenspace corresponding to an eigenvalue λ of M is

$$V_r(\lambda) = \ker\left((M-\lambda I_n)^r\right) = \{x \in \mathbb{C}^n : (M-\lambda I_n)^r x = 0\}$$

$(V_1(\lambda) = \ker(M-\lambda I_n))$

Note: If we set $r=1$, we obtain the "usual" eigenspace $V_1(\lambda)$

Note also that $V_r(\lambda)$ is a kernel and so it forms a subspace of $\mathbb{C}^n$.

Two basic properties of generalised eigenspaces:

- For any $k, r \in \mathbb{N}$, $V_r(\lambda) \leq V_{r+k}(\lambda)$

Proof: Suppose $x \in V_r(\lambda)$: $(M-\lambda I_n)^r x = 0$ (*).

Then $(M-\lambda I_n)^{r+k} x = \left((M-\lambda I_n)^k (M-\lambda I_n)^r\right) x$

$= (M-\lambda I_n)^k \left((M-\lambda I_n)^r \cdot x\right)$

$= (M-\lambda I_n)^k \cdot 0$ using (*).

So $(M-\lambda I_n)^{r+k} x = 0$ i.e. $x \in V_{r+k}(\lambda)$ as required. ▨

- Suppose, for $r \in \mathbb{N}$, $x \in V_{r+1}(\lambda)$. Then $(M-\lambda I_n)x \in V_r(\lambda)$

Proof: Suppose $x \in V_{r+1}(\lambda) \Rightarrow (M-\lambda I_n)^{r+1} x = 0$

Then consider $(M-\lambda I_n)x$

$(M-\lambda I_n)^r\left((M-\lambda I_n)\cdot x\right) = \left((M-\lambda I_n)^r (M-\lambda I_n)\right) \cdot x$

$= (M-\lambda I_n)^{r+1} x$

So $(M-\lambda I_n)^r((M-\lambda I_n)x) = 0$ i.e. $(M-\lambda I_n)\cdot x \in V_r(\lambda)$ as required.

Example: Consides $T: \mathbb{C}^2 \to \mathbb{C}^2$ where $T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4x_1 - x_2 \\ x_1 + 2x_2 \end{pmatrix}$

Then, in terms of $\mathcal{E} = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$ $[T]^{\mathcal{E}}_{\mathcal{E}} = \begin{pmatrix} 4 & -1 \\ 1 & 2 \end{pmatrix}$

characteristic polynomial: $ch_T(t) = (t-3)^2$.

Choices for $m_T(t)$: $(t-3), (t-3)^2$.

Note: $[T]^{\mathcal{E}}_{\mathcal{E}} - 3I = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \neq 0$ so $m(T) \neq t-3$.

but $m_T(t) = (t-3)^2$

$$\left(\left([T]^{\mathcal{E}}_{\mathcal{E}}\right) - 3I\right)^2 = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \checkmark$$

Then, to find $V_1(3)$, solve $\left([T]^{\mathcal{E}}_{\mathcal{E}} - 3I\right)x = 0$

i.e. $\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ This leads to $x_1 - x_2 = 0$. i.e. $x_1 = x_2$.

General vector in $V_1(3)$ is $\begin{pmatrix} x_2 \\ x_2 \end{pmatrix} = x_2\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ for $x_2 \in \mathbb{C}$.

$V_1(3)$ has (possible) basis $\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$.

Next, to find $V_2(3)$, solve $\left([T]^{\mathcal{E}}_{\mathcal{E}} - 3I\right)^2 x = 0$

ie. $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ This is solved by any $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ in $\mathbb{C}^2$.

So $V_2(3) = \mathbb{C}^2 = \left\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} x_1, x_2 \in \mathbb{C}\right\}$

Possible basis for $V_2(3) = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$.

Now, let's choose a basis to $V_2(3)$, which includes the vector from the basis of $V_1(3)$, namely $\begin{pmatrix}1\\1\end{pmatrix}$.

Note: $\begin{pmatrix}1\\1\end{pmatrix} = 1\begin{pmatrix}1\\0\end{pmatrix} + 1\begin{pmatrix}0\\1\end{pmatrix}$

$\begin{pmatrix}1\\1\end{pmatrix}$ is "related" to each of $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$

and may exchange for either $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$ to get a "new basis" for $V_2(3)$.

Let's exchange for $\begin{pmatrix}1\\0\end{pmatrix}$: $\begin{pmatrix}1\\1\end{pmatrix} \longleftrightarrow \begin{pmatrix}1\\0\end{pmatrix}$

New basis for $V_2(3) = \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$

a basis for $V_2(3)$ that "includes" a basis for $V_1(3)$, an example of a pre-Jordan basis).

Let's find $[T]^{\beta}_{\beta}$

$$T\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}3\\3\end{pmatrix} = 3\begin{pmatrix}1\\1\end{pmatrix} + 0\begin{pmatrix}0\\1\end{pmatrix}$$

$$T\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-1\\2\end{pmatrix} = -1\begin{pmatrix}1\\1\end{pmatrix} + 3\begin{pmatrix}0\\1\end{pmatrix}$$

i.e. $[T]^{\beta}_{\beta} = \begin{pmatrix} 3 & -1 \\ 0 & 3 \end{pmatrix}$

↑ "almost diagonal".

November 15th 2019

So found a pre-Jordan basis for M:

$$\beta = \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$$

where we may set $\beta_1(3) = \left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$

$\beta_2(3) = \left\{\begin{pmatrix}0\\1\end{pmatrix}\right\}$

Then, $\beta_1(3)$ is a basis for $V_1(3)$ and $\beta = \beta_1(3) \cup \beta_2(3)$ is a basis for $V_2(3)$

for the given linear map, T, $m[T]^{\mathcal{E}}_{\mathcal{E}}$ for $\mathcal{E} = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$

while $[T]^{\beta}_{\beta} = \begin{pmatrix} 3 & -1 \\ 0 & 3 \end{pmatrix}$

"almost diagonal"

In general, for a given $n \times n$ complex matrix M, with minimal polynomial:

$$m(t) = (t-\lambda_1)^{b_1} \cdots (t-\lambda_r)^{b_r}$$

(where $b_i \geq 1$ for each i, $1 \leq i \leq r$)

↳ $b_i = 2$ in a $m(t) = (t-\lambda)^2$; $b_i = 3$ in $m(t) = (t-\lambda)^3$.

⇒ A pre-Jordan basis for the eigenvalue λ_i $(1 \leq i \leq r)$ is a basis for $V_{b_i}(\lambda_i)$ of the form $\beta_1(\lambda_i) \cup \beta_2(\lambda_i) \cup \cdots \cup \beta_{b_i}(\lambda_i)$

s.t. the $\beta_j(\lambda_i)$ (todas las bases) are mutually disjoint sets and:

$\beta_1(\lambda_i)$ is a basis for $V_1(\lambda_i)$

$\beta_1(\lambda_i) \cup \beta_2(\lambda_i)$ is a basis for $V_2(\lambda_i)$

$\vdots$

$\beta_1(\lambda_i) \cup \cdots \cup \beta_{b_i}(\lambda_i)$ is a basis for $V_{b_i}(\lambda_i)$

⇒ A pre-Jordan Basis for λ_i is a pre-Jordan basis s.t.

(in addition) for each v in $\beta_{r+1}(\lambda_i)$ the vector $(M - \lambda_i I_n)v$ is in $\beta_r(\lambda_i)$ (for $r = 1, \ldots, b_i - 1$)

A Jordan basis for M is a union of Jordan bases for all eigenvalues of M (for $\lambda_1, \ldots, \lambda_r$), s.t. for each v in $\beta_{r+1}(\lambda_i)$, the vector $(M - \lambda_i I_n)v_i$ (from $\beta_r(\lambda_i)$) appears immediately to the left of v in the basis.

Returning to earlier example.

$[T]^{\mathcal{E}}_{\mathcal{E}} = M = \begin{pmatrix} 4 & -1 \\ 1 & 2 \end{pmatrix}$ $\quad \beta = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ pre-Jordan basis.

where $\beta_1(3) = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$, $\beta_2(3) = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$.

$[T]^{\beta}_{\beta} = \begin{pmatrix} 3 & -1 \\ 0 & 3 \end{pmatrix}$.

Let's transform β to a Jordan basis:

Consider $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \beta_2(3)$.

Then, want $(M-3I)\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to be in $\beta_1(3)$:

$$(M-3I)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

linked vectors.

Note: $\begin{pmatrix} -1 \\ -1 \end{pmatrix} = -1\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

May exchange $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ for $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$ in $\beta_1(3)$ and obtain the Jordan-basis:

$$\mathcal{C} = \left\{ \underbrace{\begin{pmatrix} -1 \\ -1 \end{pmatrix}}_{(M-3I)v}, \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{v} \right\}$$

Note: $\begin{pmatrix} -1 \\ -1 \end{pmatrix} \in V_1(3)$ is an eigenvector, so

$$M\begin{pmatrix} -1 \\ -1 \end{pmatrix} = 3\begin{pmatrix} -1 \\ -1 \end{pmatrix} = 3\begin{pmatrix} -1 \\ -1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and, since $(M-3I)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$

$$M\begin{pmatrix} 0 \\ 1 \end{pmatrix} - 3\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

$$\boxed{M\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} -1 \\ -1 \end{pmatrix} + 3\begin{pmatrix} 0 \\ 1 \end{pmatrix}}$$

Equivalently: $\mathcal{C} = \left\{ \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$.

$$T\begin{pmatrix} -1 \\ -1 \end{pmatrix} = 3\begin{pmatrix} -1 \\ -1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1\begin{pmatrix} -1 \\ -1 \end{pmatrix} + 3\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$\Rightarrow [T]^{\mathcal{C}}_{\mathcal{C}} = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$

even nicer than $[T]^{\beta}_{\beta}$: has a '+1' in row 1, column 2.

In general,

A Jordan block $J_r(\lambda)$ is an $r \times r$ matrix s.t.

$(J_r(\lambda))_{i,i} = \lambda$ for $i = 1, \dots, r$

$(J_r(\lambda))_{i,i+1} = 1$ for $i = 1, \dots, r-1$

$(J_r(\lambda))_{i,j} = 0$ for otherwise.

Examples:

$J_1(3) = (3)$, $J_2(3) = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$

$J_3(17) = \begin{pmatrix} 17 & 1 & 0 \\ 0 & 17 & 1 \\ 0 & 0 & 17 \end{pmatrix}$

$J_4(-2) = \begin{pmatrix} -2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix}$

A square matrix is in Jordan normal form if it has the form:

$$\begin{pmatrix} J_{r_1}(\lambda_1) & & & \\ & J_{r_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{r_k}(\lambda_k) \end{pmatrix}$$

For example, the following are in Jordan normal form:

$$\begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} \quad J_4(3)$$

$$\begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} \quad J_2(3), J_2(3).$$

$$\begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix} \quad J_2(3), J_2(5)$$

$$\begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \quad J_3(3), J_1(3)$$

$$\begin{pmatrix} -5 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & -5 \end{pmatrix} \quad J_1(-5), J_1(5), J_1(-5).$$

whereas the following are not in Jordan normal form:

$$\begin{pmatrix} 8 & 2 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 17 & 0 \\ 0 & 0 & 0 & 17 \end{pmatrix}, \quad \begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 1 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

$$\begin{pmatrix} 8 & 0 & 1 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}, \quad \begin{pmatrix} 8 & 0 & 0 & 0 \\ 1 & 8 & 0 & 0 \\ 0 & 0 & -17 & 1 \\ 0 & 0 & 0 & -17 \end{pmatrix}$$

Let's now try to determine a Jordan basis (and associated Jordan normal form) for a 3×3 matrix / a linear map $T: \mathbb{C}^3 \to \mathbb{C}^3$.

Consider $T: \mathbb{C}^3 \to \mathbb{C}^3$

where $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 + 7x_3 \\ 2x_2 + 3x_3 \\ 2x_3 \end{pmatrix}$

In terms of $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

$M := [T]^{\mathcal{E}}_{\mathcal{E}} = \begin{pmatrix} 2 & -1 & 7 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix}$

$$ch_M(t) = \det(tI_3 - M) = \begin{vmatrix} t-2 & 1 & -7 \\ 0 & t-2 & 3 \\ 0 & 0 & t-2 \end{vmatrix}$$ upper triangular

$ch_M(t) = (t-2)^3$.

Possibilities for minimal polynomial $m(t)$:

$\underbrace{t-2}_{f_1(t)}, \underbrace{(t-2)^2}_{f_2(t)}, \underbrace{(t-2)^3}_{f_3(t)}$

$$f_1(M) = M - 2I = \begin{pmatrix} 0 & -1 & 7 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix} \neq 0$$

$$f_2(M) = (M-2I)^2 = \begin{pmatrix} 0 & -1 & 7 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 & 7 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0.$$

So, $m(t)$ must be $f_3(t)$: $m(t) = (t-2)^3$

$f_3(M) = (M-2I)^3 = (M-2I) = 0$.

So to find a pre-Jordan basis, start by determining $V_1(2), V_2(2), V_3(2)$.

Start with $V_1(2) = \ker(M-2I)$

Solve $(M-2I)x = 0$ i.e. $\begin{pmatrix} 0 & -1 & 7 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

This leads to $-x_2 + 7x_3 = 0$, $3x_3 = 0$, $0 = 0$

i.e. to $x_2 = 0$, $x_3 = 0$

So a general element of $V_1(2)$ has the form

$\begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ for $x_1 \in \mathbb{C}$.

Possible basis for $V_1(2) = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$

For $V_2(2) = \ker((M-2I)^2)$

$$(M-2I)^2 x = 0 \Rightarrow \begin{pmatrix} 0 & 0 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

This leads to $-3x_3 = 0$, i.e. $x_3 = 0$.

i.e. $V_2(2) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} : x_1, x_2 \in \mathbb{C} \right\}$

Possible basis over $\mathbb{C}$: $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$

Finally, for $V_3(2)$, solve $(M-2I)^3 x = 0$

i.e. $0 \cdot x = 0$.

This holds for every $x \in \mathbb{C}^3$, so:

$V_3(2) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1, x_2, x_3 \in \mathbb{C} \right\} = \mathbb{C}^3$.

Possible basis over $\mathbb{C}$: $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

For a pre-Jordan basis, require a basis $\beta_1(2) \cup \beta_2(2) \cup \beta_3(2)$ for $V_3(2)$ s.t. $\beta_1(2)$ is basis for $V_1(2)$

$\beta_1(2) \cup \beta_2(2)$ is basis for $V_2(2)$

Here, from our earlier computations, may simply choose:

$\beta_1(2) = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$, $\beta_2(2) = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$, $\beta_3(2) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

For a Jordan basis, require, for each v in $\beta_{r+1}(2)$, $(M-2I)v$ to be in $\beta_r(2)$

Start from $\beta_3(2) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

$$(M-2I)\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 7 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix}$$

Note: $\begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix} = 7 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

$\begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix}$ is "related" to each of $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, and in particular

$\begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix}$ is "related" to $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

So, we may exchange $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ for $\begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix}$ to obtain a "new" β_2:

$\beta_2(2) = \left\{ \begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix}, \right.$ $\left. \right\}$

Consider a new $\beta_2(2) = \left\{ \begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix} \right\}$

$$(M-2I)\begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 7 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix}$$

$$\left[(M-2I)^2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right]$$

$\begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix} = -3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ where $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \in \beta_1(2)$.

May exchange $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ for $\begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix}$ in $\beta_1(2)$, to obtain a new $\beta_1(2)$: $\beta_1(2) = \left\{ \begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix} \right\}$

Overall, we now have:

$$\beta_1(2) = \left\{ \begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix} \right\}, \quad \beta_2(2) = \left\{ \begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix} \right\}, \quad \beta_3(2) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

"linked" "linked"

$$\begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix} = (M-2I)\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix} = (M-2I)\begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix} = (M-2I)^2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

This leads to a Jordan basis:

$$\beta = \left\{ \begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

$(M-2I)^2 v$ $(M-2I)v$ v

Let's find a matrix representing T in terms of the Jordan basis:

$$T\begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -6 \\ 0 \\ 0 \end{pmatrix} = 2\begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$T\begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 11 \\ 6 \\ 0 \end{pmatrix} = 1\begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 3 \\ 2 \end{pmatrix} = 0\begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 7 \\ 3 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

So $[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ ✓ a matrix in Jordan normal form.

where $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 + 7x_3 \\ 2x_2 + 3x_3 \\ 2x_3 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 7 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

November 18th 2019

Outline of general algorithm to find Jordan basis of complex linear map $T: \mathbb{C}^n \to \mathbb{C}^n$ (or $n \times n$ complex matrix M)

- Find $(ch(t))$ and $m(t) = (t-\lambda)^{b} \cdots (t-\lambda_r)^{b_r}$
- For each eigenvalue λ_i $(1 \le i \le r)$:
  - → Determine (bases) for $V_1(\lambda_i), \ldots, V_{b_i}(\lambda_i)$
  - → By suitably exchanging vectors, & where necessary, obtain (pre-Jordan and) Jordan basis for eigenvalue λ_i.
- Take union of Jordan basis for all eigenvalues. The resulting set is due to be a Jordan basis for the "whole" linear map / matrix.

In order for this algorithm to work in general, we must ensure that:

① When we combine Jordan bases for all eigenvalues at final step, we obtain a Jordan basis for the whole space, i.e. a suitable basis for $\mathbb{C}^n$.

② When we try to exchange vectors at various parts, (to obtain a pre-Jordan and Jordan basis), we may validly do so

Start with ①

Let's show that whenever we combine bases for $V_{b_1}(\lambda_1), V_{b_2}(\lambda_2), \ldots, V_{b_r}(\lambda_r)$ we obtain a (Jordan) basis for the whole of $\mathbb{C}^n$.

We show this by showing that (means that $V_{b_1}(\lambda_1) \cap V_{b_2}(\lambda_2) = 0$)

$$\mathbb{C}^n = V_{b_1}(\lambda_1) \oplus V_{b_2}(\lambda_2) \oplus \cdots \oplus V_{b_r}(\lambda_r).$$

eg: for an $n \times n$ matrix M. $m_1(t) = (t-3)(t-5)^2$

$$\Rightarrow \mathbb{C}^n = V_1(3) \oplus V_2(5) =$$

$$\mathbb{C}^n = \ker(M-3I) \oplus \ker\left((M-5I)^2\right)$$

Note: The 2 "parts" come from coprime polynomials $(t-3), (t-5)^2$

A key result to show $\mathbb{C}^n = V_{b_1}(\lambda_1) \oplus \cdots \oplus V_{b_r}(\lambda_r)$, is:

Proposition: Suppose $f(t)$ and $g(t)$ are coprime polynomials in $\mathbb{C}[t]$. Then

$$\ker\left(f(T)g(T)\right) = \ker\left(f(T)\right) \oplus \ker\left(g(T)\right)$$

for a linear map $T: \mathbb{C}^n \to \mathbb{C}^n$ over $\mathbb{C}$.

Proof: Note, $f(t), g(t)$ are coprime, so the $\gcd(f,g) = 1$

Then, by Bezout's lemma, $\exists\, a(t), b(t)$ in $\mathbb{C}[t]$ s.t. $a(t)f(t) + b(t)\cdot g(t) = 1$.

"Applying" both sides to T:

$$a(T)f(T) + b(T)g(T) = Id \quad (*)$$

To show $\ker\left(f(T)g(T)\right) = \ker\left(f(T)\right) \oplus \ker\left(g(T)\right)$

we will show that:

A $\ker\left(f(T)g(T)\right) = \ker\left(f(T)\right) + \ker\left(g(T)\right)$

B The sum is direct: $\ker\left(f(T)\right) \cap \ker\left(g(T)\right) = \{0\}$.

In turn, to show Ⓐ, we show:

A1 $\ker\left(f(T)\right) + \ker\left(g(T)\right) \subseteq \ker\left(f(T)g(T)\right)$

A2 $\ker\left(f(T)g(T)\right) \subseteq \ker\left(f(T)\right) + \ker\left(g(T)\right)$

Start with (A1) Let $v \in \ker(f(T)) + \ker(g(T))$

Then $v = v_1 + v_2$ where $v_1 \in \ker(f(T))$, $v_2 \in \ker(g(T))$

i.e. where $f(T)(v_1) = 0$ $\quad g(T)(v_2) = 0$

Consider $(f(T)g(T))(v)$

$$\begin{aligned}(f(T)g(T))(v) &= (f(T)g(T))(v_1 + v_2)\\ &= (f(T)g(T))(v_1) + (f(T)g(T))(v_2) \quad \text{by linearity}\\ &= (g(T)f(T))(v_1) + (f(T)g(T))(v_2) \quad \text{since } f(T)g(T) = g(T)f(T)\\ &= g(T)(f(T)(v_1)) + f(T)(g(T)v_2)\\ &= g(T)(0) + f(T)(0) \quad \text{using } f(T)(v_1) = 0,\ g(T)(v_2) = 0\\ &= 0 + 0.\end{aligned}$$

In this case they commute.

So $(f(T)g(T))(v) = 0$ i.e. $v \in \ker(f(T)g(T))$

So, if $v \in \ker(f(T)) + \ker(g(T))$, then $v \in \ker(f(T)g(T))$

i.e. $\ker(f(T)) + \ker(g(T)) \subseteq \ker(f(T)g(T))$

Now, consider (A2) Let $v \in \ker(f(T)g(T))$, so that $(f(T)g(T))(v) = 0$

From (*) (use Bezout's Lemma)

$a(T)f(T) + b(T)g(T) = Id$. Apply both sides to v.

$(a(T)f(T) + b(T)g(T))(v) = Id(v)$

i.e. $(a(T)f(T))(v) + (b(T)g(T))(v) = (v)$.

Set $v_1 = b(T)g(T)(v)$, $v_2 = a(T)f(T)(v)$

$\Rightarrow$ $v = v_1 + v_2$ and

$$\begin{aligned}f(T)(v_1) &= f(T)(b(T)g(T))(v)\\ &= f(T)b(T)g(T)(v)\end{aligned}$$

i.e. $$\begin{aligned}f(T)(v_1) &= b(T)(f(T)g(T))(v)\\ &= b(T)\cdot(0) \quad \text{since } f(T)g(T)(v) = 0.\end{aligned}$$

so $f(T)(v_1) = 0$ i.e. $v_1 \in \ker(f(T))$

Similarly: $v_2 \in \ker(g(T)) \Rightarrow g(T)(v_2) = 0$.

18th November 2019

$$\begin{aligned}g(T)(v_2) &= g(T)(a(T)f(T)(v))\\ &= g(T)a(T)f(T)(v) \quad \text{so } g(T)(v_2) = 0 \text{ i.e. } v_2 \in \ker(g(T))\\ &= a(T)f(T)g(T)(v)\\ &= a(T)((f(T)g(T))(v))\\ &= a(T)(0)\end{aligned}$$

so $v = v_1 + v_2$, where $v_1 \in \ker(f(T))$, $v_2 \in \ker(g(T))$

i.e. $v \in \ker(f(T)) + \ker(g(T))$

This shows $\ker(f(T)g(T)) \subseteq \ker(f(T)) + \ker(g(T))$

Overall, (A1) and (A2) together lead to:

(A) $\ker(f(T)g(T)) = \ker(f(T)) + \ker(g(T))$

Now, show the sum is direct, (B) $\ker(f(T)) \cap \ker(g(T)) = \{0\}$.

Let $v \in \ker(f(T)) \cap \ker(g(T))$

Then $v \in \ker(f(T))$ and $v \in \ker(g(T))$ i.e. $f(T)(v) = 0$, $g(T)(v) = 0$

Also, using (*) $v = a(T)f(T)(v) + b(T)g(T)(v)$

$$\begin{aligned}&= a(T)(f(T)(v)) + b(T)\cdot(g(T)(v))\\ &= a(T)(0) + b(T)(0) \quad \text{using } f(T)(v) = 0,\ g(T)(v) = 0.\\ &= 0 + 0.\end{aligned}$$

Then, it must be that $v = 0$ i.e. $\ker(f(T)) \cap \ker(g(T)) = \{0\}$ as required

Then, using (A) and (B), we finally obtain, as required

$\ker(f(T)g(T)) = \ker(f(T)) \oplus \ker(g(T))$.

This leads to a result known as the Primary Decomposition Theorem.

Theorem: Suppose that $T: \mathbb{C}^n \to \mathbb{C}^n$ is a linear map over $\mathbb{C}$, with minimal polynomial

$m_T(t) = (t-\lambda_1)^{b_1} \cdots (t-\lambda_r)^{b_r}$ for distinct $\lambda_1, \ldots, \lambda_r$ in $\mathbb{C}$.

Then, $\mathbb{C}^n = V_{b_1}(\lambda_1) \oplus \cdots \oplus V_{b_r}(\lambda_r)$.

Proof: Note that, since $m_T(t)$ is a minimal polynomial of T

$m_T(T) = 0$, the zero map.

Then, $\ker(m_T(T)) = \ker(0) = \{v \in \mathbb{C}^n : 0 \cdot v = 0\} = \mathbb{C}^n$.

Also, $\ker(m_T(T)) = \ker\left((T-\lambda_1 Id)^{b_1} \cdot (T-\lambda_2 Id)^{b_2} \cdots (T-\lambda_r Id)^{b_r}\right)$

$= \ker\left((T-\lambda_1 Id)^{b_1}\right) \oplus \ker\left((T-\lambda_2 Id)^{b_2} \cdots (T-\lambda_r Id)^{b_r}\right)$

(using previous result).

$= \ker\left((T_1-\lambda_1 Id)^{b_1}\right) \oplus \ker\left((T-\lambda_2 Id)^{b_2}\right) \oplus \ker\left((T-\lambda_3 Id)^{b_3} \cdots (T-\lambda_r Id)^{b_r}\right) =$

$= \ker\left((T_1-\lambda_1 Id)^{b_1}\right) \oplus \ker\left(T-\lambda_2 Id)^{b_2}\right) \oplus \cdots \oplus \ker\left(T-\lambda_r Id\right)^{b_r}\right)$

(by, perhaps repeatedly, applying the previous result).

By definition: $V_{b_i}(\lambda_i) = \ker\left((T-\lambda_i Id)^{b_i}\right)$, so overall, we obtain

$$\mathbb{C}^n = V_{b_1}(\lambda_1) \oplus V_{b_2}(\lambda_2) \oplus \cdots \oplus V_{b_r}(\lambda_r)$$

Using the Primary Decomposition Theorem and a result reviewed earlies in the course, we can obtain a basis for the "whole" direct sum, $\mathbb{C}^n$ by simply taking a union of basis for $V_{b_1}(\lambda_1), \ldots, V_{b_r}(\lambda_r)$.

So, the final part of "Jordan basis algorithm", where we combine Jordan bases for the individual eigenvalues, to obtain a Jordan basis for the "whole space", does work.

November 22nd 2019

Let's now consider a special case of the minimal polynomial $m_T(t)$.

Proposition: for a linear map $T: V \to V$ over $\mathbb{C}$, T is diagonalisable if and only if the minimal polynomial of T has the form $m_T(t) = (t-\lambda_1)\cdot(t-\lambda_2)\cdots(t-\lambda_r)$ for distinct (eigen)values $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$.

Proof ($\Rightarrow$) Suppose, first, that the minimal polynomial of T has the form

$m(T) = (t-\lambda_1)\cdots(t-\lambda_r) =$

$= (t-\lambda_1)^1 \cdots (t-\lambda_r)^1$

Then, by the Primary decomposition theorem:

$V = V_1(\lambda_1) \oplus \cdots \oplus V_1(\lambda_r)$.

So, we can form a basis for V by taking the union of bases for $V_1(\lambda_1), \ldots, V_1(\lambda_r)$.

But $V_1(\lambda_i)$, for $1 \le i \le r$ is the "usual" eigenspace corresponding to the eigenvalue of λ_i. So a basis for $V_1(\lambda_i)$ consists of eigenvectors of T.

Hence, we can find a basis for V consisting of eigenvectors of T.

Therefore, T is diagonalisable, as required.

($\Leftarrow$) Now, suppose that T is diagonalisable. So, $\exists$ basis for V consisting of eigenvectors for T, say $\{v_1, \ldots, v_n\}$.

Consider a vector v from $\{v_1, \ldots, v_n\}$.

Suppose v corresponds to eigenvalue λ_i (all eigenvalues are $\lambda_1, \ldots, \lambda_r$).

$\Rightarrow T(v) = \lambda_i v$

So $T^m(v) = T^{m-1}(T(v)) = T^{m-1}(\lambda v) = \lambda(T^{m-1}(v))$

$= \lambda\left(T^{m-2}(T(v))\right) =$

$= \lambda\left(T^{m-2}(\lambda v)\right) = \lambda^2\left(T^{m-2}(v)\right) =$

$\vdots$

$T^m(v) = \lambda^m v$ for $m = 1, 2, \ldots$

Then, for any polynomial

$$g(t) = a_m t^m + \dots + a_1 t + a_0$$

we obtain:

$$g(T)(v) = \left(a_m T^m + \dots + a_1 T + a_0 \mathrm{Id}\right)(v)$$

$$= a_m T^m(v) + \dots + a_1 T(v) + a_0(v)$$

$$= a_m \lambda_i^m (v) + \dots + a_1 \lambda_1 (v) + a_0 (v)$$

$$= \left(a_m \lambda_i^m + \dots + a_1 \lambda_i + a_0\right) \cdot v$$

So $g(T)(v) = g(\lambda_i)(v)$.

Consider $f(t) = (t-\lambda_1) \cdots (t-\lambda_r)$

$\Rightarrow f(t)(v) = f(\lambda_i)(v)$

But $f(\lambda_i) = 0$ since λ_i is a root of $f(t)$

So $f(\lambda_i) v = 0$ i.e. $f(T)(v) = 0$

This holds for every v in $\{v_1, \dots, v_n\}$

But $\{v_1, \dots, v_n\}$ is a basis for V, so for each $w \in V$: $w = c_1 v_1 + \dots + c_n v_n$ for some $c_1, \dots, c_n \in \mathbb{C}$.

Then: $f(T)(w) = f(T)(c_1 v_1 + \dots + c_n v_n)$

$$= c_1 f(T)(v_1) + \dots + c_n f(T)(v_n)$$

$$= c_1 \cdot 0 + \dots + c_n \cdot 0$$

(since $f(T)(v_1) = 0, \dots, f(T)(v_n) = 0$).

So $f(T)(w) = 0$ for each w in V.

Hence, $f(T) = 0$, the zero linear map.

Then, using a previous result: $m(t) \mid f(t)$

- Also, every eigenvalue of T is a root of $m(t)$ i.e. $m(t)$ has the form: $m(t) = (t-\lambda_1)^{b_1} \cdots (t-\lambda_r)^{b_r}$ for $b_1 \geq 1, \dots, b_r \geq 1$.

It follows from this that: $f(t) \mid m(t)$

$$\| $$

$$(t-\lambda_1) \cdots (t-\lambda_r)$$

Overall, $m(t)$ and $f(t)$ are monic polynomials such that $m(t) \mid f(t)$ and $f(t) \mid m(t)$.

Hence, $m(t) = f(t)$ (using an earlier result)

i.e. $m(t) = (t-\lambda_1) \cdots (t-\lambda_r)$ as required.

This completes the proof ▨

Example: of Primary Decomposition theorem

If, for $T: \mathbb{C}^n \to \mathbb{C}^n$, with

$$m_T(t) = (t-5)^3 (t-7)^2$$

$\Rightarrow m_T(T) = 0$, so $\ker(m_T(T)) = \mathbb{C}^n$

and $\mathbb{C}^n = \ker(m_T(T))$

$$= \ker\left((T-5\mathrm{Id})^3 (T-7\mathrm{Id})^2\right)$$

$$= \ker\left((T-5\mathrm{Id})^3\right) \oplus \ker\left((T-7\mathrm{Id})^2\right)$$

So $\mathbb{C}^n = V_3(5) \oplus V_2(7)$.

Can find a Jordan basis for T by "combining" Jordan bases for the eigenvalues $5, 7$.

Let's now study an example involving 2 eigenvalues:

Consider $A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 3 & 5 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$

Let's assume that it is given that

$$ch_A(t) = (t-1)^2 (t-2)^3$$

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

According to the primary decomposition Theorem, can find a (Jordan) basis for $\mathbb{C}^5$ by "combining" (Jordan) bases for $V_2(1)$ and $V_2(2)$

Then, for any polynomial

$$g(t) = a_m t^m + \dots + a_1 t + a_0$$

we obtain:

$$g(T)(v) = \left(a_m T^m + \dots + a_1 T + a_0 Id\right)(v)$$

$$= a_m T^m(v) + \dots + a_1 T(v) + a_0(v)$$

$$= a_m \lambda_i^m (v) + \dots + a_1 \lambda_1(v) + a_0(v)$$

$$= \left(a_m \lambda_i^m + \dots + a_1 \lambda_i + a_0\right) \cdot v$$

So $g(T)(v) = g(\lambda_i)(v)$.

Consider $f(t) = (t-\lambda_1) \dots (t-\lambda_r)$

$\Rightarrow f(t)(v) = f(\lambda_i)(v)$

But $f(\lambda_i) = 0$ since λ_i is a root of $f(t)$

So $f(\lambda_i)v = 0$ i.e. $f(T)(v) = 0$

This holds for every v in $\{v_1, \dots, v_n\}$

But $\{v_1, \dots, v_n\}$ is a basis for V, so for each $w \in V$: $w = c_1 v_1 + \dots + c_n v_n$ for some $c_1, \dots, c_n \in \mathbb{C}$.

Then: $f(T)(w) = f(T)(c_1 v_1 + \dots + c_n v_n)$

$$= c_1 f(T)(v_1) + \dots + c_n f(T)(v_n)$$

$$= c_1 \cdot 0 + \dots + c_n \cdot 0$$

(since $f(T)(v_1) = 0, \dots, f(T)(v_n) = 0$).

So $f(T)(w) = 0$ for each w in V.

Hence, $f(T) = 0$, the zero linear map.

Then, using a previous result: $m(t) \mid f(t)$

- Also, every eigenvalue of T is a root of $m(t)$ i.e. $m(t)$ has the form: $m(t) = (t-\lambda_1)^{b_1} \dots (t-\lambda_r)^{b_r}$ for $b_1 \geq 1, \dots, b_r \geq 1$.

It follows from this that: $f(t) \mid m(t)$

$$\| $$

$$(t-\lambda_1) \dots (t-\lambda_r)$$

Overall, $m(t)$ and $f(t)$ are monic polynomials such that $m(t) \mid f(t)$ and $f(t) \mid m(t)$.

Hence, $m(t) = f(t)$ (using an earlier result)

i.e. $m(t) = (t-\lambda_1) \dots (t-\lambda_r)$ as required.

This completes the proof ▨

Example: of Primary Decomposition theorem

If, for $T: \mathbb{C}^n \to \mathbb{C}^n$, with

$$m_T(t) = (t-5)^3 (t-7)^2$$

$\Rightarrow m_T(T) = 0$, so $\ker(m_T(T)) = \mathbb{C}^n$

and $\mathbb{C}^n = \ker(m_T(T))$

$$= \ker\left((T-5Id)^3 (T-7Id)^2\right)$$

$$= \ker\left((T-5Id)^3\right) \oplus \ker\left((T-7Id)^2\right)$$

So $\mathbb{C}^n = V_3(5) \oplus V_2(7)$.

Can find a Jordan basis for T by "combining" Jordan bases for the eigenvalues $5, 7$.

Let's now study an example involving 2 eigenvalues:

Consider $A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 3 & 5 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$

Let's assume that it is given that

$$ch_A(t) = (t-1)^2 (t-2)^3$$

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

According to the primary decomposition Theorem, can find a (Jordan) basis for $\mathbb{C}^5$ by "combining" (Jordan) basis for $V_2(1)$ and $V_2(2)$

1st Let's start with the eigenvalue 1.

Find a Jordan basis for 1.

Note: $A - 1\cdot I = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & 5 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

First, find $V_1(1)$

Solve $(A-I)x = 0$.

i.e. $\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & 5 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

This leads to $V_1(1) = \left\{ \begin{pmatrix} 0 \\ x_2 \\ 0 \\ 0 \\ 0 \end{pmatrix} : x_2 \in \mathbb{C} \right\}$

A Possible basis for $V_1(1) = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$

B Similarly, for $V_2(1)$, solve $(A-I)^2 x = 0$

This leads to a possible basis for $V_2(1)$ of $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$

The vector from basis of $V_1(1)$ already appears in basis for $V_2(1)$, so may set

$$\beta_1(1) = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}, \quad \beta_2(1) = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

This gives a pre-Jordan basis for the eigenvalue 1.

For a Jordan basis, consider $v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ from $\beta_2(1)$

$$\Rightarrow (A-I)(v_1) = (A-I)\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$\begin{pmatrix} 0 \\ 3 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 3 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ so $\begin{pmatrix} 0 \\ 3 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ is "related" to $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

and may exchange $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ for $\begin{pmatrix} 0 \\ 3 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ in $\beta_1(1)$

to obtain:

$$\beta_1(1) = \underbrace{\left\{ \begin{pmatrix} 0 \\ 3 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}}_{(A-I)v_1}, \quad \beta_2(1) = \underbrace{\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}}_{v_1}$$

This leads to a Jordan basis for the eigenvalue 1.

$$\left\{ \underbrace{\begin{pmatrix} 0 \\ 3 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{(A-I)v_1}, \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{v_1} \right\}$$

2nd Next, consider eigenvalue 2.

Here, use $A - 2I = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 3 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

A Then, for $V_1(2)$, solve $(A-2I)x = 0$ to obtain

$$\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ -5/3 \\ 1 \end{pmatrix} \right\}$$ as a possible basis for $V_1(2)$.

B Similarly, for $V_2(2)$, solve $(A-2I)^2 x = 0$

to obtain a possible basis for $V_2(2)$, of:

$$\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

for a pre-Sordan basis, try to ensure that each of

$\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\-5/3\\1\end{pmatrix}$ appears in a basis for $V_2(2)$.

- $\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}$ already appears in chose basis for $V_2(2)$
- $\begin{pmatrix}0\\0\\0\\-5/3\\1\end{pmatrix} = -5/3 \begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix} + 1 \cdot \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}$

May exchange for either $\begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}$

Here, let's replace $\begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix}$ by $\begin{pmatrix}0\\0\\0\\-5/3\\1\end{pmatrix}$

This leads to a new basis for $V_2(2)$:

$$\left\{\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\-5/3\\1\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}\right\}$$

a pre-Jordan basis where:

$$\beta_1(2) = \left\{\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\-5/3\\1\end{pmatrix}\right\}, \quad \beta_2(2) = \left\{\begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}\right\}$$

for a Jordan basis (for 2), consider

$v_2 = \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix} \in \beta_2(2)$

Then, $(A-2I)\, v_2 = (A-2I) \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix} = \begin{pmatrix}0\\0\\5\\0\\0\end{pmatrix}$

Note: would like $(A-2I)\, v_2$ to appear in $\beta_1(2)$

Note: $\begin{pmatrix}0\\0\\5\\0\\0\end{pmatrix} = 5 \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix} + 0 \cdot \begin{pmatrix}0\\0\\0\\-5/3\\1\end{pmatrix} = 5 \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}$

So, we replace using the "related" vector $\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}$ to

obtain a Jordan basis for eigenvalue 2.

This leads to a new basis for $V_2(2)$:

$$\left\{\begin{pmatrix}0\\0\\5\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\0\\-5/3\\1\end{pmatrix}\right\}$$

$(A-2I)\, v_2 \qquad v_2$

where $\beta_1(2) = \left\{\begin{pmatrix}0\\0\\5\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\-5/3\\1\end{pmatrix}\right\}$ and $\beta_2(2) = \left\{\begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}\right\}$

Overall, we have Jordan bases for eigenvalues:

$$\left\{\begin{pmatrix}0\\3\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}\right\} \qquad \left\{\begin{pmatrix}0\\0\\5\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\0\\-5/3\\1\end{pmatrix}\right\}$$

3rd Then, a possible Jordan basis corresponding to the matrix A is:

$$\left\{\begin{pmatrix}0\\3\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\5\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\0\\-5/3\\1\end{pmatrix}\right\}$$

$\lambda=1 \quad \lambda=1 \quad \lambda=2 \quad \lambda=2 \quad \lambda=2$

Then, the corresponding Jordan normal form must be:

$$\begin{pmatrix}1 & * & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 2 & * & 0\\ 0 & 0 & 0 & 2 & *\\ 0 & 0 & 0 & 0 & 2\end{pmatrix} \longrightarrow \begin{pmatrix}1 & 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 2 & 1 & 0\\ 0 & 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 0 & 2\end{pmatrix}$$

Also, if we set $P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & -5/3 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$

Then, it holds that $P^{-1}AP = J$.

Quicker way of finding the final Jordan normal form

after finding $V_1(1), V_2(1)$ and $V_1(2), V_2(2)$

Here: $\dim(V_1(1)) = 1$, $\dim(V_2(1)) = 2$

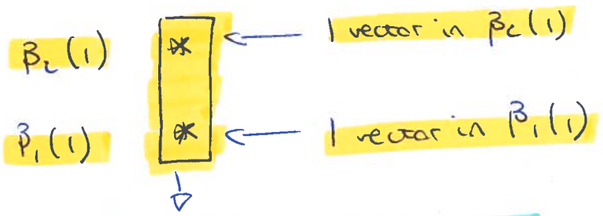
So, in any Jordan basis $\beta_1(1)$ consists of 1 vector.

$\beta_1(1) \cup \beta_2(1)$ consists of 2 vectors.

So $|\beta_1(1)| = 1$, $|\beta_2(1)| = 1$

size of the set

1st Visually

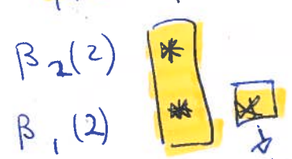


2nd Similarly

$\dim(V_1(2)) = 2$ $\dim(V_2(2)) = 3$.

So, in a Jordan basis,

$|\beta_1(2)| = 2$ $|\beta_2(2)| = 1$



November 25th 2019

Let's now return to the general algorithm for finding a Jordan basis and potential issue (2), as described in a previous lecture.

We would like to ensure that, when exchanging vectors to obtain a (pre-) Jordan basis, we may validly do so, using the Exchange Lemma, to do so, we verify, that relevant vectors are linearly independent.

Let's first verify this when trying to exchange all vectors in a given '$\beta_b(\lambda)$', in some pre-Jordan basis.

Proposition: Let M be an $n \times n$ complete matrix, and let λ be an eigenvalue of M, such that, in a relevant pre-Jordan basis for λ:

$\beta_1(\lambda) \cup \beta_2(\lambda) \cup \dots$

we have $\beta_b(\lambda) = \{v_1, \dots, v_m\}$ for some b, $b \geq 2$

Then $(M - \lambda I)v_1, \dots, (M - \lambda I)v_m$ are L.I.

(here I denote an $n \times n$ identity matrix).

Proof: Note that, by definition:

$\beta_1(\lambda) \cup \dots \cup \beta_{b-1}(\lambda) \cup \beta_b(\lambda)$ is a basis for $V_b(\lambda)$

$\beta_1(\lambda) \cup \dots \cup \beta_{b-1}(\lambda)$ is a basis for $V_{b-1}(\lambda)$.

Suppose $\beta_1(\lambda) \cup \dots \cup \beta_{b-1}(\lambda) = \{u_1, \dots, u_k\}$

So $\{u_1, \dots, u_k\}$ is a basis for $V_{b-1}(\lambda)$.

$\{u_1, \dots, u_k, v_1, \dots, v_m\}$ is a basis for $V_b(\lambda)$.

Also, note: $V_1(\lambda) \leq V_2(\lambda) \leq \dots \leq V_{b-1}(\lambda)$

so $\{u_1, \dots, u_k\}$; the basis for $V_{b-1}(\lambda)$, spans $V_1(\lambda)$, $V_2(\lambda)$, $V_{b-1}(\lambda)$

Suppose $\alpha_1(M - \lambda I)v_1 + \dots + \alpha_m(M - \lambda I)v_m = 0$

$\Rightarrow (M - \lambda I)(\alpha_1 v_1 + \dots + \alpha_m v_m) = 0$ by linearity

so $\alpha_1 v_1 + \dots + \alpha_m v_m \in \ker(M - \lambda I)$

i.e. $\alpha_1 v_1 + \dots + \alpha_m v_m \in V_1(\lambda)$

Since $\{u_1, \dots u_k\}$ spans $V_1(\lambda)$, there must exist $b_1, \dots, b_k$ in $\mathbb{C}$ s.t. $\alpha_1 v_1 + \dots + \alpha_m v_m = b_1 u_1 + \dots + b_k u_k$

$\Rightarrow \alpha_1 v_1 + \dots + \alpha_m v_m - b_1 u_1 - \dots - b_k u_k = 0$

Since $\{u_1, \dots, u_k, v_1, \dots, v_m\}$ is a basis for $V_b(\lambda)$.

$\{u_1, \dots, u_k, v_1, \dots, v_m\}$ is L.I.

So we must have $\alpha_1 = 0, \dots, \alpha_m = 0, -b_1 = 0, \dots, -b_k = 0$.

In particular $\alpha_1 = 0, \dots, \alpha_m = 0$

This shows that $(M-\lambda I)v_1, \dots, (M-\lambda I)v_m$ are L.I as required ///

Then, in this context, may apply Exchange Lemma to exchange other vectors in $\beta_1(\lambda) \cup \dots \cup \beta_{b-1}(\lambda)$ by (the linearly independent) $(M-\lambda I)v_1, \dots, (M-\lambda I)v_m$

In fact, can always exchange for vectors in $\beta_{b-1}(\lambda)$

To see this, suppose $v \in \beta_b(\lambda)$

Then, by definition of a pre-Jordan basis: $v \in V_b(\lambda)$ but $v \notin V_{b-1}(\lambda)$

i.e. $(M-\lambda I)v = 0$ but $(M-\lambda I)^{b-1} v \neq 0$.

So $(M-\lambda I)^{b-1}\big((M-\lambda I)v\big) = 0$, $(M-\lambda I)^{b-2}\big((M-\lambda I)v\big) \neq 0$.

Hence, $(M-\lambda I)v \in V_{b-1}(\lambda)$, but $(M-\lambda I)v \notin V_{b-2}(\lambda)$

So, it will be possible to exchange $(M-\lambda I)v$ for something in $\beta_{b-1}(\lambda)$

In general, can also exchange each of $(M-\lambda I)v_1, \dots (M-\lambda I)v_m$ for vectors in $\beta_{b-1}(\lambda)$.

A consequence of this is $|\beta_{b-1}(\lambda)| \geq m$. i.e. $|\beta_{b-1}(\lambda)| \geq |\beta_b(\lambda)|$

↑ size of $\beta_{b-1}(\lambda)$

More generally: $|\beta_1(\lambda)| \geq |\beta_2(\lambda)| \geq |\beta_3(\lambda)| \geq \dots$

(useful property of (pre-) Jordan basis).

Can often use this to determine Jordan normal forms.

Suppose that, for an $n \times n$ matrix M.

$ch_M(t) = (t-2)^{(4)}$ ← in final Jordan normal form, there will be 4 entries of the eigenvalue 2 on the diagonal.

Let's consider the different possibilities for the minimal polynomial, $m(t)$

- $m(t) = t-2 = (t-2)^{(1)}$ ← we can find enough vectors in $V_1(2)$

Visually: leads to

$\beta_1(2)$ ⊛ ⊛ ⊛ ⊛ — 4 (1×1) blocks.

So, a Jordan normal form must consist of 4 copies of $\underbrace{J_1(2)}_{1\times 1 \text{ block}}$.

eg. $\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

- $m(t) = (t-2)^{(2)}$ ← we can find enough (i.e. 4) vectors in $V_2(2)$, also $|\beta_1(2)| \geq |\beta_2(2)|$

while $|\beta_1(2)| + |\beta_2(2)| = 4$

There are 2 possibilities

Visually

$\beta_2(2)$ ⊛ ⊛

$\beta_1(2)$ ⊛ ⊛ ⊛

$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

or

$\beta_2(2)$ ⊛ ⊛

$\beta_1(2)$ ⊛ ⊛

$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ $\dim(V_1(2)) = 2$

November 26th 2019

Using this and some other data, we can often determine the (possible) forms of a) Jordan normal form in a given situation.

In general, given a matrix or linear map for which

$ch(t) = (t-\lambda_1)^{a_1} \cdots (t-\lambda_r)^{a_r}$, for distinct eigenvalues $\lambda_1, \dots, \lambda_r$

$m(t) = (t-\lambda_1)^{b_1} \cdots (t-\lambda_r)^{b_r}$ where, for each i: $1 \le b_i \le a_i$ $(1 \le i \le r)$

the following hold(s), for each eigenvalue λ_i $(1 \le i \le r)$

The number b_i, from the minimal polynomial $m(t)$, indicates that a Jordan basis for λ, is of the form $\beta_1(\lambda_i) \cup \dots \cup \beta_{b_i}(\lambda_i)$, where, by definition:

$\beta_1(\lambda_i), \dots, \beta_{b_i}(\lambda_i)$ are disjoint (non-empty) sets s.t.:

$\beta_1(\lambda_i)$ is a basis for $V_1(\lambda_i)$,

$\beta_1(\lambda_i) \cup \beta_2(\lambda_i)$ " " $V_2(\lambda_i)$

$\vdots$

$\beta_1(\lambda_i) \cup \beta_2(\lambda_i) \cup \dots \cup \beta_{b_i}(\lambda_i)$ " " $V_{b_i}(\lambda_i)$.

Using (all of) the above, we may obtain (and use):

→ The multiplicity of λ, in $ch(t)$, namely a_i, indicates the total number of occurrences of λ_i in the Jordan normal form.

In fact, $a_i = \dim(V_{b_i}(\lambda_i))$

So $a_i = |\beta_1(\lambda_i) \cup \dots \cup \beta_{b_i}(\lambda_i)| = |\beta_1(\lambda_i)| + |\beta_2(\lambda_i)| + \dots + |\beta_{b_i}(\lambda_i)|$

since $\beta_1(\lambda_i), \dots, \beta_{b_i}(\lambda_i)$ are disjoint.

→ The multiplicity of λ_i, in $m(t)$, namely b_i, indicates that a Jordan basis is of the form $\beta_1(\lambda_i) \cup \dots \cup \beta_{b_i}(\lambda_i)$

In a Jordan normal form, the largest possible Jordan block (for λ_i) is of the size $b_i \times b_i$ (and there is at least one such block).

→ By definition of a Jordan basis:

$\dim(V_1(\lambda_i)) = |\beta_1(\lambda_i)|$

$\dim(V_2(\lambda_i)) = |\beta_1(\lambda_i) \cup \beta_2(\lambda_i)| = |\beta_1(\lambda_i)| + |\beta_2(\lambda_i)|$

$\vdots$

$\dim(V_{b_i}(\lambda_i)) = \dots = |\beta_1(\lambda_i)| + |\beta_2(\lambda_i)| + \dots + |\beta_{b_i}(\lambda_i)|$

→ As stated earlier: $|\beta_1(\lambda_i)| \ge |\beta_2(\lambda_i)| \ge \dots \ge |\beta_{b_i}(\lambda_i)|$

Example: Suppose that, for some matrix / linear map

$ch(t) = (t-2)^4$.

So a Jordan normal consists of a matrix with 4 entries of the eigenvalue 2 on the main diagonal

$$\begin{pmatrix} 2 & \cdot & 0 & 0 \\ 0 & 2 & \cdot & 0 \\ 0 & 0 & 2 & \cdot \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Consider different choices for $m(t)$

• $m(t) = t-2 = (t-2)^1$.

The largest Jordan block for eigenvalue 2 is of size 1×1,

i.e. is a $J_1(2)$ block.

Only possibility: 4 copies of $J_1(2)$

So a Jordan normal form: $\begin{pmatrix} \boxed{2} & 0 & 0 & 0 \\ 0 & \boxed{2} & 0 & 0 \\ 0 & 0 & \boxed{2} & 0 \\ 0 & 0 & 0 & \boxed{2} \end{pmatrix}$ $\beta_1(2)$ $\beta_1(2)$ $\beta_1(2)$ $\beta_1(2)$

- $m(t) = (t-2)^2$

The largest relevant Jordan block is of size 2×2

There are 2 possibilities for the Jordan normal form

(A) $\begin{pmatrix} \boxed{2\ 1} & 0 & 0 \\ 0\ 2 & 0 & 0 \\ 0\ 0 & \boxed{2} & 0 \\ 0\ 0 & 0 & \boxed{2} \end{pmatrix}$ one copy of $J_2(2)$, 2 copies of $J_1(2)$

$\dim(V_1(2)) = 3$ here

$\beta_2(2)$ $\beta_1(2)$ $\beta_1(2)$

$J_2(2)$ $J_1(2)$ $J_1(2)$

(B) $\begin{pmatrix} \boxed{2\ 1} & 0 & 0 \\ 0\ 2 & 0 & 0 \\ 0\ 0 & \boxed{2\ 1} & 0 \\ 0\ 0 & 0\ 2 & \boxed{2} \end{pmatrix}$ 2 copies of $J_2(2)$

$\dim(V_1(2)) = 2$ here

$\beta_2(2)$ $\beta_1(2)$

$J_2(2)$ $J_2(2)$

Note: $\dim(V_1(2))$ denotes $|\beta_1(2)|$, this corresponds to the total number of Jordan blocks (for eigenvalue 2).

- $m(t) = (t-2)^3$

The largest relevant Jordan block is of size (3×3).

Only option: one copy of $J_3(2)$ one copy of $J_1(2)$.

$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ $\beta_3(2)$ $\beta_2(2)$ $\beta_1(2)$ $\beta_1(2)$

- $m(t) = (t-2)^4$.

The largest relevant Jordan block is of size 4×4.

Only possibility: one copy of $J_4(2)$

$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ $\beta_4(2)$ $\beta_3(2)$ $\beta_2(2)$ $\beta_1(2)$

Final result concerning linear independence of vectors within a "Jordan link"

Proposition: Suppose, for an $n \times n$ complex matrix M, over $\mathbb{C}$ with λ an eigenvalue of M: $v \in V_b(\lambda)$ but $v \notin V_{b-1}(\lambda)$.

Then, the following are LI:

$$v, (M-\lambda I)v, \dots, (M-\lambda I)^{b-1}v$$

Proof: Since $v \in V_b(\lambda)$: $(M-\lambda I)^b v = 0$.

So for any $m \geq b$: $(M-\lambda I)^m v = (M-\lambda I)^{m-b}\left((M-\lambda I)^b v\right)$

$= (M-\lambda I)^{m-b} \cdot 0$

i.e. $(M-\lambda I)^m v = 0$ $\forall$ $m = b, b+1, b+2, \dots$

Also: $v \notin V_{b-1}(\lambda)$: $(M-\lambda I)^{b-1} v \neq 0$.

It follows that $(M-\lambda I)v + \dots + \alpha_{b-1}(M-\lambda I)^{b-1}v = 0$ $(*)$

Multiply through by $(M-\lambda I)^{b-1}$

$$(M-\lambda I)^{b-1}\left(\alpha_0 v + \alpha_1 (M-\lambda I)v + \dots + \alpha_{b-1}(M-\lambda I)^{b-1}v\right) = (M-\lambda I)^{b-1} \cdot 0$$

i.e.

$$\alpha_0 (M-\lambda I)^{b-1}v + \alpha_1 \underbrace{(M-\lambda I)^b v}_{\to 0} + \alpha_2 \underbrace{(M-\lambda I)^{b+1} v}_{\to 0} + \dots + \alpha_{b-1} \underbrace{(M-\lambda I)^{2b-2} v}_{\to 0} = 0$$

Then, using (1), we obtain: $\alpha_0 (M-\lambda I)^{b-1} v = 0$.

Also $(M-\lambda I)^{b-1} v \neq 0 \Rightarrow \alpha_0 = 0$.

Then, $(*)$ becomes:

$$\alpha_1 (M-\lambda I)v + \dots + \alpha_{b-1}(M-\lambda I)^{b-1}v = 0$$

Multiply through by $(M-\lambda I)^{b-2}$, and simplify as before

$\alpha_1 (M-I\lambda)^{b-1} v = 0$ but $(M-\lambda I)^{b-1} v \neq 0$ so $\alpha_1 = 0$.

Similarly, can obtain $\alpha_2 = 0, \dots, \alpha_{n-1} = 0$

Overall; $\alpha_0 = 0, \alpha_1 = 0, \dots, \alpha_{b-1} = 0$, so the given vectors are LI

∎

November 29th 2019

CHAPTER 3 : LINEAR AND BILINEAR FORMS AND INNER PRODUCT SPACES :

- Mathematical "form" : a function that has "single numbers" as outputs.

3.1. LINEAR FORMS

- Linear forms : involve single vectors as "input values"
- For a vector space V over a field K, a linear form on V over K, is a function $f : V \to K$ s.t. $\forall\, a, b \in V$ every λ in K :
  - $f(0) = 0$
  - $f(a+b) = f(a) + f(b)$
  - $f(\lambda a) = \lambda f(a)$.

Example : Set $K = \mathbb{C}$, $V = \mathbb{C}^2$ $\Rightarrow$ Consider $f : \mathbb{C}^2 \to \mathbb{C}$, where

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3x_1 - 5x_2$$

Then f is a linear form on $\mathbb{C}^2$ (over $\mathbb{C}$).

eg. $f\begin{pmatrix} 17 \\ 11 \end{pmatrix} = -4$

In general, given a basis $\beta = \{b_1, \dots, b_n\}$ for V over K :

- for any vector v in V, there exists unique $\lambda_1, \dots, \lambda_n$ in K s.t. $v = \lambda_1 b_1 + \dots + \lambda_n b_n$.

$\Rightarrow$ we define :

$$[v]_\beta = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$ vector in K representing v in terms of basis β.

- Given a linear form $f : V \to K$, the matrix representing f in terms of β, is the $1 \times n$ matrix $[f]^\beta_\beta$, or $([f]_\beta)_i = f(b_i)$ for $i = 1, \dots, n$.

ie. $[f]_\beta = (f(b_1), f(b_2), \dots, f(b_n))$

eg. consider $f : \mathbb{C}^2 \to \mathbb{C}$ where $f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3x_1 - 5x_2$

Consider also the following bases for $\mathbb{C}^2$ over $\mathbb{C}$:

$$\mathcal{E} = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}, \quad \mathcal{B} = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$$

Set $v = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$

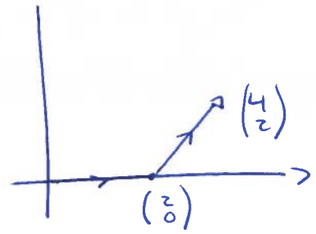
$\Rightarrow f(v) = f\begin{pmatrix} 4 \\ 2 \end{pmatrix} = 3(4) - 5(2) = 2$.

Let's express v in terms of $\mathcal{E}$ and $\mathcal{B}$:

$v = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 4\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ so $[v]_\mathcal{E} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$

↳ we usually call our vectors in terms of standard basis.

$v = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ so $[v]_\mathcal{B} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$



Let's now express f in terms of $\mathcal{E}$ and $\mathcal{B}$:

$f\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 3(1) - 5(0) = 3$, $f\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 3(0) - 5(1) = -5$.

so $[f]^\mathcal{E}_\mathcal{E} = (3, -5)$

$f\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 3$, $f\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3(1) - 5(1) = -2$

so $[f]^\mathcal{B}_\mathcal{B} = (3, -2)$.

Then, $[f]^\mathcal{E}_\mathcal{E}[v]_\mathcal{E} = (3, -5)\begin{pmatrix} 4 \\ 2 \end{pmatrix} = 12 - 10 = 2$.

i.e. $[f]^\mathcal{E}_\mathcal{E}[v]_\mathcal{E} = f(v)$.

while $[f]^\mathcal{B}_\mathcal{B}[v]_\mathcal{B} = (3, -2)\begin{pmatrix} 2 \\ 2 \end{pmatrix} = 6 - 4 = 2$.

i.e., $[f]^\mathcal{B}_\mathcal{B}[v]_\mathcal{B} = f(v)$.

In general:

for a linear form $f: V \to K$

if $\beta = \{b_1, \dots, b_n\}$ is a basis for V over K, and v is any vector in V,

then $f(v) = [f]_\beta^\beta \, [v]_\beta$

Proof: For a given v in V, there exists $\lambda_1, \dots, \lambda_n$ s.t. $v = \lambda_1 b_1 + \dots + \lambda_n b_n$

$$\Rightarrow [v]_\beta = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

while $[f]_\beta^\beta = \big(f(b_1), \dots, f(b_n)\big)$.

Then,

$$[f]_\beta^\beta \, [v]_\beta = \big(f(b_1), \dots, f(b_n)\big) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} =$$

$$= f(b_1)\cdot\lambda_1 + \dots + f(b_n)\,\lambda_n$$

i.e. $\boxed{[f]_\beta^\beta \, [v]_\beta = \lambda_1 f(b_1) + \dots + \lambda_n f(b_n)}$.

Let's now compute $f(v)$ directly:

$$f(v) = f(\lambda_1 b_1 + \dots + \lambda_n b_n)$$

↘ Linearity

$$= f(\lambda_1 b_1) + \dots + f(\lambda_n b_n)$$

↘ linearity

$$= \lambda_1 f(b_1) + \dots + \lambda_n f(b_n)$$

So $f(v) = [f]_\beta^\beta \, [v]_\beta$, as required ☑

This holds for any basis β.

As a consequence, if β, C are bases for V over K, then, for every v in V

$$[f]_\beta^\beta \, [v]_\beta = f(v) = [f]_C^C \, [v]_C$$

i.e. for every vector v in V: $[f]_\beta^\beta \, [v]_\beta = [f]_C^e \, [v]_C$.

Next, consider the set of all linear forms on V over K:

$V^* = \{f: V \to K \;:\; f \text{ is linear}\}$ dual set of/to V.

We know that V^* is (also) a vector space over K (just as V is), if we use the following operations, for all $f, g \in V^*$, $\lambda \in K$, and every v in V:

- $(f+g)(v) = f(v) + g(v)$ (addition)
- $(\lambda f)(v) = \lambda f(v)$ (scalar multiplication)

Example, Consider $\mathbb{R}^3$ over $\mathbb{R}$. Then, the dual vector space is

$$(\mathbb{R}^3)^* = \{f: \mathbb{R}^3 \to \mathbb{R} \;:\; f \text{ is linear}\}$$

If $\beta = \{b_1, b_2, b_3\}$ is a basis for $\mathbb{R}^3$ over $\mathbb{R}$, then we may represent a general $f: \mathbb{R}^3 \to \mathbb{R}$, from $(\mathbb{R}^3)^*$, via:

$$[f]_B = \big(f(b_1), f(b_2), f(b_3)\big)$$

So, we may contrast $\mathbb{R}^3$ and $(\mathbb{R}^3)^*$ as follows:

$$\mathbb{R}^3 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \;:\; x_1, x_2, x_3 \in \mathbb{R} \right\} \quad \text{column vectors}$$

$$(\mathbb{R}^3)^* = \{(a_1, a_2, a_3) \;:\; a_1, a_2, a_3 \in \mathbb{R}\} \quad \text{row vectors.}$$

$[f]_\beta = \big(f(b_1), f(b_2), f(b_3)\big)$

Consider a basis for $\mathbb{R}^3$:

$$\mathcal{E} = \left\{ \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{e_1}, \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{e_2}, \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}_{e_3} \right\}$$

Consider also a basis for $(\mathbb{R}^3)^*$

$$\mathcal{E}^* = \{\underbrace{(1,0,0)}_{e_1^*}, \underbrace{(0,1,0)}_{e_2^*}, \underbrace{(0,0,1)}_{e_3^*}\}$$

Note: $e_1^* \cdot (e_1) = (1,0,0) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1$

$e_1^* \cdot (e_2) = (1,0,0) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0$, $e_1^* \cdot (e_3) = (1,0,0) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$.

So $e_1^*(e_1) = 1$

$e_1^*(e_2) = 0$

$e_1^*(e_3) = 0$.

Similarly:

$$e_2^*(e_1)=0\ ,\ e_2^*(e_2)=1,\ e_2^*(e_3)=0$$
$$e_3^*(e_1)=0\ ,\ e_3^*(e_2)=0\ ,\ e_3^*(e_3)=1$$

In this sense, the set $\{e_1^*, e_2^*, e_3^*\}$ is dual to $\{e_1, e_2, e_3\}$

In general, for a vector space V over K, with basis $B=\{b_1,\dots,b_n\}$ the dual set to B is the set $\{f_1,\dots,f_n\}$ in V^* defined via:

$$f_i(b_j)=\begin{cases}1 & \text{if } i=j\\ 0 & \text{if } i\neq j\end{cases}$$

i.e. $f_1(b_1)=1,\ f_1(b_2)=0,\dots,f_1(b_n)=0$.

Proposition: Suppose V is a vector space over a field K, and that $B=\{b_1,\dots,b_n\}$ is a basis for V over K.

$\Rightarrow$ The dual set to B in V^* $\{f_1,\dots,f_n\}$ say,

is a basis for V^* over K.

Proof: Suppose $\{f_1,\dots,f_n\}$ is dual to $\{b_1,\dots,b_n\}$.

i.e. for each i, $i=1,\dots,n$ f_i is a function from V to K,

and $$f_i(b_j)=\begin{cases}1 & \text{if } i=j\\ 0 & \text{if } i\neq j\end{cases}$$

1ST First, let's show that $\{f_1,\dots,f_n\}$ spans V^* over K.

Consider a function $f: V\to K$ in V^* and let v be a general vector in V.

Then, for some $\lambda_1,\dots,\lambda_n$: $v=\lambda_1 b_1+\dots+\lambda_n b_n$

Applying f to v gives

$$f(v)=f(\lambda_1 b_1+\dots+\lambda_n b_n)$$ ← By linearity of f.

$$f(v)=\lambda_1 f(b_1)+\dots+\lambda_n f(b_n)$$

Set $a_1=f(b_1),\dots,a_n=f(b_n)$

Consider: $a_1f_1+\dots+a_nf_n$ in V^*

Then, $(a_1f_1+\dots+a_nf_n)(v)=a_1f_1(v)+\dots+a_nf_n(v)$

by definition of V^* as a vector space.

$$=a_1f_1(\lambda_1b_1+\dots+\lambda_nb_n)+\dots+a_nf_n(\lambda_1b_1+\dots+\lambda_nb_n)$$

Then, by linearity:

$$(a_1f_1+\dots+a_nf_n)(v)=a_1\lambda_1f_1(b_1)+a_1\lambda_2f_1(b_2)+\dots+a_1\lambda_nf_1(b_n)$$
$$+\dots$$
$$+a_n\lambda_1f_n(b_1)+a_n\lambda_2f_n(b_2)+\dots+a_n\lambda_nf_n(b_n)$$

Now, use $f_i(b_j)=\begin{cases}1 & \text{if } i=j\\ 0 & \text{if } i\neq j\end{cases}$

So $(a_1f_1+\dots+a_nf_n)(v)=a_1\lambda_1\underbrace{f_1(b_1)}_{=1}+\dots+a_n\lambda_n\underbrace{f_n(b_n)}_{=1}$.

i.e. $$(a_1f_1+\dots+a_nf_n)(v)=a_1\lambda_1+\dots+a_n\lambda_n$$
$$=\lambda_1a_1+\dots+\lambda_na_n$$
$$=\lambda_1f(b_1)+\dots+\lambda_nf(b_n)$$
$$=f(v)$$

So, for each $v\in V$

$$f(v)=(a_1f_1+\dots+a_nf_n)(v)$$

Since this holds for each v:

$$f=a_1f_1+\dots+a_nf_n.$$

This holds for an arbitrary f in V^*

so $\{f_1,\dots,f_n\}$ spans V^* as required.

2nd Let's now show that $\{f_1, \dots, f_n\}$ is LI.

Suppose $a_1 f_1 + \dots + a_n f_n = 0$

↳ the zero map in V^*

Then, for every v in V:

$(a_1 f_1 + \dots + a_n f_n)(v) = 0$ (*)

Set $v = b_1$: $(a_1 f_1 + \dots + a_n f_n)(b_1) = 0$

i.e. $a_1 f_1(b_1) + a_2 \cancel{f_2(b_1)}^{\to 0} + \dots + a_n \cancel{f_n(b_1)}^{\to 0} = 0$

i.e. $a_1 f_1(b_1) = 0$.

Then, since $f_1(b_1) = 1$, we obtain $a_1 = 0$.

In general, if we set $v = b_i$ in (*)

we obtain $a_i f_i(b_i) = 0$ i.e. $a_i = 0$ (for $i = 1, \dots, n$).

So, overall, we obtain: $a_1 = 0, \dots, a_n = 0$

Hence, $\{f_1, \dots, f_n\}$ is LI.

This concludes the proof: $\{f_1, \dots, f_n\}$ is a basis for V^* over K ∎

It follows from this correspondence of bases:

$\beta = \{b_1, \dots, b_n\}$ for V over K, $\beta^* = \{f_1, \dots, f_n\}$ for V^* over K

that there is a bijection from V to V^*

$\varphi : V \to V^*$, where

$\varphi(\lambda_1 b_1 + \dots + \lambda_n b_n) = \lambda_1 f_1 + \dots + \lambda_n f_n$

$\varphi(b_1) = f_1, \dots, \varphi(b_n) = f_n$

This leads to an isomorphism of vector spaces $V \cong V^*$

3.2 Bilinear forms

Inputs: pairs of vectors.

- Suppose V is a vector space over K:

Then, a bilinear form on V is a function $f : V \times V \to K$ s.t.,

$\forall a, b, c$ in V and $\forall \lambda$ in K:

- $f(a, 0) = 0$ and $f(0, a) = 0$
- $f(a+b, c) = f(a, c) + f(b, c)$

  and

  $f(c, a+b) = f(c, a) + f(c, b)$
- $f(\lambda a, b) = \lambda f(a, b)$

  and

  $f(a, \lambda b) = \lambda f(a, b)$.

[Note: $f(\lambda a, \lambda b) = \lambda f(a, \lambda b) = \lambda \cdot \lambda f(a, b) = \lambda^2 f(a, b)$]

Given such a bilinear form f, and a basis $\beta = \{b_1, \dots, b_n\}$ for V over K, the matrix representing f with respect to β is the $n \times n$ matrix $[f]^\beta_\beta$, or $[f]_\beta$, defined via: $\left([f]^\beta_\beta\right)_{ij} = f(b_i, b_j)$

for $1 \le i \le n$, $1 \le j \le n$.

i.e. $[f]^\beta_\beta = \begin{pmatrix} f(b_1, b_1) & f(b_1, b_2) & \dots & f(b_1, b_n) \\ f(b_2, b_1) & & & \\ \vdots & & & \\ f(b_n, b_1) & & & f(b_n, b_n) \end{pmatrix}$

Examples:

Consider $K = \mathbb{R}$, $V = \mathbb{R}^2$, and let $f : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ defined via:

$$\underbrace{f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right)}_{f(x, y)} = x_1 y_1 + x_2 y_2$$

Then f is a bilinear form.

Consider $K = \mathbb{C}$, $V = \mathbb{C}^2$, and let $f : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ defined via

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 - x_1 y_2 + 2 x_2 y_1 + 5 x_2 y_2$$

This is also a bilinear form (on $\mathbb{C}^2$ over $\mathbb{C}$ here).

Let's consider $\mathcal{E} = \left\{ \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{e_1}, \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{e_2} \right\}$ and find $[f]_{\mathcal{E}}^{\mathcal{E}}$

$$f(e_1, e_1) = f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = 1 \cdot 1 - 1 \cdot 0 + 2 \cdot 0 \cdot 1 + 5 \cdot 0 \cdot 0$$

i.e. $f(e_1, e_1) = 1$

Similarly $f(e_1, e_2) = -1$, $f(e_2, e_1) = 2$, $f(e_2, e_2) = 5$.

Then, $[f]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} f(e_1, e_1) & f(e_1, e_2) \\ f(e_2, e_1) & f(e_2, e_2) \end{pmatrix}$

i.e. $[f]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix}$

How can we use this to compute f in general?

We may compute

$$\underset{\text{transpose}}{x^T} [f]_{\mathcal{E}}^{\mathcal{E}} y = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} =$$

$$= (x_1 \; x_2) \begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} =$$

$$= x_1 (y_1 - y_2) + x_2 (2 y_1 + 5 y_2)$$

So $x^T [f]_{\mathcal{E}}^{\mathcal{E}} y = x_1 y_1 - x_1 y_2 + 2 x_2 y_1 + 5 x_2 y_2$

i.e. $f(x, y) = x^T [f]_{\mathcal{E}}^{\mathcal{E}} y$

$[x]_{\mathcal{E}}^T [f]_{\mathcal{E}}^{\mathcal{E}} [y]_{\mathcal{E}}$.

December 1st 2019

Consider a particular case, where $x = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

- Using definition: $f\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = (2)(0) - (2)(1) + 2(1)(0) + 5(1)(1) = 3$

How can we compute this using $[f]_{\mathcal{E}}^{\mathcal{E}}$?

$x = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ i.e. $\left[\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right]_{\mathcal{E}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

$y = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ i.e. $\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]_{\mathcal{E}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Then $([x]_{\mathcal{E}})^T [f]_{\mathcal{E}}^{\mathcal{E}} [y]_{\mathcal{E}} = (2, 1) \begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = (2, 1) \begin{pmatrix} -1 \\ 5 \end{pmatrix} = 3$

Now, consider basis $\mathcal{B} = \left\{ \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{b_1}, \underbrace{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{b_2} \right\}$ for $\mathbb{C}^2$ over $\mathbb{C}$.

Then, $[f]_{\mathcal{B}} = \begin{pmatrix} f(b_1, b_1) & f(b_1, b_2) \\ f(b_2, b_1) & f(b_2, b_2) \end{pmatrix}$ here $[f]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 3 & 7 \end{pmatrix}$

where e.g. $f(b_1, b_2) = f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = (1)(1) - (1)(1) + 2(1)(0) + 5(0)(1)$ from definition of f.

i.e. $f(b_1, b_2) = 0$.

Similarly; $f(b_1, b_1) = 1$, $f(b_2, b_1) = 3$, $f(b_2, b_2) = 7$.

Also, find $[x]_{\mathcal{B}}$, $[y]_{\mathcal{B}}$:

$x = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ i.e. $[x]_{\mathcal{B}} = \left[\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

$y = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ i.e. $[y]_{\mathcal{B}} = \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

Then, $([x]_{\mathcal{B}})^T [f]_{\mathcal{B}}^{\mathcal{B}} [y]_{\mathcal{B}} = (1, 1) \begin{pmatrix} 1 & 0 \\ 3 & 7 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} =$

$= (1, 1) \begin{pmatrix} -1 \\ 4 \end{pmatrix} = -1 + 4 = 3 = f(x, y)$ ✓

In general, consides a bilinear form $f: V \times V \to K$ with

$\beta = \{b_1, \dots, b_n\}$ a basis for V over K

Then, for all vectors u, w in V:

$$f(u,w) = \left([u]_\beta\right)^T [f]_\beta^\beta \, [w]_\beta$$

Proof: Since β is a basis for V over K, can express u, w in terms of β.

$u = x_1 b_1 + \dots + x_n b_n$ for some $x_1, \dots, x_n$ in K.

$w = y_1 b_1 + \dots + y_n b_n$ for some $y_1, \dots, y_n$ in K.

$$\Rightarrow [u]_\beta = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad [w]_\beta = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$\Rightarrow$ Using the fact that f is bilinear:

$$f(u,w) = f(x_1 b_1 + \dots + x_n b_n, y_1 b_1 + \dots + y_n b_n)$$

$$= f(x_1 b_1, y_1 b_1 + \dots + y_n b_n) + \dots + f(x_n b_n, y_1 b_1 + \dots + y_n b_n)$$

$$= f(x_1 b_1, y_1 b_1) + \dots + f(x_1 b_1, y_n b_n) + \dots$$

$$\dots + f(x_n b_n, y_1 b_1) + \dots + f(x_n b_n, y_n b_n) =$$

$$= x_1 \cdot f(b_1, y_1 b_1) + \dots + x_1 f(b_1, x_n b_n) + \dots + x_n f(b_n, y_1 b_1) +$$

$$\dots + x_n f(b_n, y_n b_n)$$

$$= x_1 y_1 f(b_1, b_1) + \dots + x_1 y_n f(b_1, b_n) + \dots + x_n y_1 f(b_n, b_1) +$$

$$\dots + x_n y_n f(b_n, b_n)$$

So $$f(u,w) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j f(b_i, b_j)$$.

Similarly,

$$\left([u]_\beta\right)^T [f]_\beta^\beta \, [w]_\beta = (x_1, \dots, x_n) \begin{pmatrix} f(b_1,b_1) & \dots & f(b_1,b_n) \\ \vdots & & \\ f(b_n,b_1) & \dots & f(b_n,b_n) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} =$$

$$= (x_1, \dots, x_n) \begin{pmatrix} f(b_1,b_1) y_1 + \dots + f(b_1,b_n) y_n \\ \vdots \\ f(b_n,b_1) y_1 + \dots + f(b_n,b_n) y_n \end{pmatrix}$$

$$= x_1 \left(f(b_1,b_1) y_1 + \dots + f(b_1,b_n) y_n \right) + \dots + y_n \left(f(b_n,b_1) y_1 + \dots + f(b_n,b_n) y_n \right)$$

$$= x_1 y_1 f(b_1,b_1) + \dots + x_1 y_n f(b_1,b_n) + \dots + x_n y_1 f(b_n,b_1) + \dots + x_n y_n f(b_n,b_n)$$

So, $$\left([u]_\beta\right)^T [f]_\beta^\beta \, [w]_\beta = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j f(b_i, b_j)$$

Overall $f(u,w) = \left([u]_\beta\right)^T [f]_\beta^\beta \, [w]_\beta$ as required

This works for any basis, i.e. if $\mathcal{C}$ is another basis for V over K, we obtain, for all u, w in V

$$\left([u]_\beta\right)^T [f]_\beta^\beta \, [w]_\beta = f(u,w) = \left([u]_\mathcal{C}\right)^T [f]_\mathcal{C}^\mathcal{C} \, [u]_\mathcal{C}$$

Can use this to find a change of basis formula between $[f]_\beta^\beta$ and $[f]_\mathcal{C}^\mathcal{C}$

Returning to our previous example:

$$[f]_\mathcal{E}^\mathcal{E} = \begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix}, \quad [f]_\beta^\beta = \begin{pmatrix} 1 & 0 \\ 3 & 7 \end{pmatrix}$$

for $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ $\beta = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$.

$$[v]_\mathcal{E} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad [v]_\beta = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$[w]_\mathcal{E} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad [w]_\beta = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Let's find $[Id]_\beta^\mathcal{E}$

$$Id \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$Id \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\Rightarrow [Id]_\beta^\mathcal{E} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Note: $[Id]_\beta^\varepsilon [v]_\beta = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = [v]_\varepsilon$

In general (*) $\boxed{[Id]_\beta^e [v]_\beta = [v]_e}$ for suitable bases β, e and every vector v.

Let's now use this to find a formula relating $[f]_\beta^\beta$ and $[f]_e^e$

for all u, w in V:

$$([u]_e)^T [f]_e^e [w]_e = ([u]_\beta)^T [f]_\beta^\beta [w]_\beta$$

$$= \left([Id]_e^\beta [u]_e\right)^T [f]_\beta^\beta \left([Id]_e^\beta [w]_e\right)$$

[Using $[u]_\beta = [Id]_e^\beta [u]_e$, $[w]_\beta = [Id]_e^\beta [w]_e$ from (*)]

$$= [u]_e^T \left([Id]_e^\beta\right)^T [f]_\beta^\beta [Id]_e^\beta [w]_e$$

This holds for all u, w in V.

$$[f]_e^e = \left([Id]_e^\beta\right)^T [f]_\beta^\beta [Id]_e^\beta$$

If we set $M := [Id]_e^\beta$, $\Rightarrow$ This shows that there is an invertible matrix M s.t.:

$$[f]_e^e = M^T [f]_\beta^\beta M, \text{ namely } M = [Id]_e^\beta.$$

Returning to example from earlier $[f]_\varepsilon^\varepsilon = \begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix}$

$[f]_\beta^\beta = \begin{pmatrix} 1 & 0 \\ 3 & 7 \end{pmatrix}$ and $[Id]_\beta^\varepsilon = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Verify formula

$$\left([Id]_\beta^\varepsilon\right)^T [f]_\varepsilon^\varepsilon [Id]_\beta^\varepsilon = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^T \begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 3 & 7 \end{pmatrix} = [f]_\beta^\beta.$$

So $\left([Id]_\beta^\varepsilon\right)^T [f]_\varepsilon^\varepsilon [Id]_\beta^\varepsilon = [f]_\beta^\beta$ as required ▨

December 4th 2019

We now concentrate on a special kind of linear form:

Symmetric

A linear form $f: V \times V \to K$ is symmetric if $f(a,b) = f(b,a)$ $\forall a, b$ in V.

Note: for such a symmetric linear form, consider, $\forall$ basis $\beta = \{b_1, \dots, b_n\}$ of V over K, the matrix $[f]_\beta^\beta$.

$$[f]_\beta^\beta = \begin{pmatrix} f(b_1,b_1) & \cdots & f(b_1,b_n) \\ \vdots & & \\ f(b_n,b_1) & \cdots & f(b_n,b_n) \end{pmatrix}$$

Then, by symmetry of f: $f(b_1,b_n) = f(b_n,b_1)$

and, in general: $f(b_i,b_j) = f(b_j,b_i)$

i.e. $\left([f]_\beta^\beta\right)_{ij} = \left([f]_\beta^\beta\right)_{ji}$ $\forall i,j$

$\Rightarrow$ $\forall$ basis β, $[f]_\beta^\beta$ is a symmetric matrix.

(matrix H is symmetric if $H_{ij} = H_{ji}$ $\forall i,j$ or equivalently, $H^T = H$)

Examples:

① The linear form $f: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ where

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 + x_2 y_2$$

is symmetric, e.g. note that $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = (x_1, x_2)\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\text{symmetric}}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

$[f]_{\mathcal{E}}^{\mathcal{E}} \to \mathcal{E} = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$

② The bilinear form $f: \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$

where $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 + 2x_1 y_2 + 2x_2 y_1 + 5x_2 y_2 =$

$$= (x_1, y_2)\underbrace{\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}}_{\substack{f(e_1,e_1) \; f(e_1,e_2) \\ [f]_{\mathcal{E}}^{\mathcal{E}}}}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$[f]_{\mathcal{E}}^{\mathcal{E}} \Rightarrow$ symmetric matrix

Not every bilinear form is symmetric ⚠

e.g. $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 - x_1 y_2 + 2x_2 y_1 + 5x_2 y_2$

$$= (x_1\ x_2)\begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

is not symmetric

↳ $f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = -1$ whereas $f\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = 2$

(e_1, e_2 and e_2, e_1)

so, for $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ $\quad f(e_1, e_2) = -1$ and $f(e_2, e_1) = 2$

$f(e_1, e_2) \neq f(e_2, e_1)$.

Definition To any (not necessary symmetric) bilinear form $f: V \times V \to k$, we can associate a quadratic form q.

This quadratic form is a function $q: V \to k$ defined via

$$q(v) = f(v, v) \quad \forall v \in V$$

Examples

① For bilinear form $f: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$

where $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 + x_2 y_2$

The associated quadratic form is a function $q: \mathbb{R}^2 \to \mathbb{R}$

where $q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = x_1 x_1 + x_2 x_2$

i.e. $\boxed{q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + x_2^2}$

② For bilinear form $f: \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$

where $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 + 2x_1 y_2 + 2x_2 y_1 + x_2 y_2$

The associated quadratic form $q: \mathbb{C}^2 \to \mathbb{C}$ is defined by

$q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = x_1^2 + 2x_1 x_2 + 2x_2 x_1 + 5x^2$.

i.e. $\boxed{q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 4x_1 x_2 + 5x_2^2}$

Note that we may obtain the same quadratic form $q: \mathbb{C}^2 \to \mathbb{C}$ by starting from the (non-symmetric) bilinear form:

$f: \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ defined via

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 + 17 x_1 y_2 - 13 x_2 y_1 + 5 x_2 y_2 = (x_1, x_2)\begin{pmatrix} 1 & 17 \\ -13 & 5 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

Then, for the associated quadratic form $q: \mathbb{C}^2 \to \mathbb{C}$

$q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = x_1^2 + 17x_1 x_2 - 13 x_1 x_2 + 5x_2^2 =$

$= x_1^2 + 4x_1 x_2 + 5x_2^2$ ✓

(i.e. it's the same).

Let's show that, in many cases, a quadratic form corresponds to a unique symmetric bilinear form.

① To show this, first note that given any (symmetric) bilinear form $f: V \times V \to k$, we can define an associated quadratic form $q: V \to k$ by $\boxed{q(v) = f(v,v) \quad \forall\, v \text{ in } V}$

↑ defines q, given f.

② Let's now also show that we can define a symmetric bilinear form f, given a quadratic form q.

(For this to work, we must assume that we can divide by 2 in k, i.e. that $2 \neq 0$ in k)

↳ if k mod 2 $\Rightarrow 2 = 0$.

Proposition: Suppose that q is a quadratic form associated to a symmetric bilinear form $f: V \times V \to k$, in a field k where $2 \neq 0$.

Then, $f(u,w) = \frac{1}{2}\left(q(u+w) - q(u) - q(w)\right) \quad \forall u, w \text{ in } V$.

Proof: Consider, $\forall u, w$ in V: $q(u+w)$

$q(u+w) = f(u+w, w+u)$

$= f(u,u) + f(u,w) + f(w,u) + f(w,w)$ ← since f is bilinear

$= f(u,u) + f(u,w) + f(u,w) + f(w,w)$ ← f is symmetric so $f(w,w) = f(w,u)$.

$= f(u,u) + 2f(u,w) + f(w,w)$

$= q(u) + 2f(u,w) + q(w)$

So: $q(u+w) = q(u) + 2f(u,w) + q(w)$

Rearranging gives $2f(u,w) = q(u+w) - q(u) - q(w)$

finally, since $2 \neq 0$ and k is a field (any number has a multiplicative inverse), may divide through by 2 (in k, 2 has a multiplicative inverse), to obtain:

$$f(u,w) = \frac{1}{2}\left(q(u+w) - q(u) - q(w)\right) \quad \forall u, w \text{ in } V \quad \blacksquare$$

Note,

- Consider a field in which $2=0$, e.g. let $k = \mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z} = \{0,1\}$

(where $2 = 1+1 = 0$)

Then, for a given quadratic form, may find more than one symmetric bilinear form

eg consider $q: \mathbb{F}_2^2 \to \mathbb{F}_2$ where $q(x_1,x_2) = q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + x_2^2$

↑ Alternative way of finding $q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$\Rightarrow$ We can obtain q from the symmetric bilinear form

$f: \mathbb{F}_2^2 \times \mathbb{F}_2^2 \to \mathbb{F}_2$ where $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1y_1 + x_2y_2 =$

$= (x_1\ x_2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

$\Rightarrow f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1^2 + x_2^2 = q(x_1,x_2)$

$\Rightarrow$ But also, we can obtain q from the symmetric bilinear form

$f: \mathbb{F}_2^2 \times \mathbb{F}_2^2 \to \mathbb{F}_2$ where,

$f'\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2 = (x_1,x_2)\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

$\Rightarrow f'\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = x_1^2 + x_1x_2 + x_2x_1 + x_2^2 = x_1^2 + 2x_1x_2 + x_2^2 =$

$= x_1^2 + 0 + x_2^2 = q(x_1,x_2)$

↑ (since $2=0$ in $\mathbb{F}_2$).

December 5TH 2019

For a (symmetric) bilinear form $f: V \times V \to K$, and given bases $\mathcal{B}, \mathcal{C}$ for V over K, there is an invertible matrix M s.t

$$[f]^{\mathcal{C}}_{\mathcal{C}} = M^T [f]^{\mathcal{B}}_{\mathcal{B}} M$$

Aim: to show that, given a symmetric bilinear form f, or the associated quadratic form g, then, there is a basis, $\mathcal{B}$ say, s.t $[f]^{\mathcal{B}}_{\mathcal{B}}$, or, equivalently, $[q]^{\mathcal{B}}_{\mathcal{B}}$, is a diagonal matrix

(Set $\mathcal{B} = \{b_1, \dots, b_n\}$)

ie. such that

$$[f]^{\mathcal{B}}_{\mathcal{B}} = \begin{pmatrix} f(b_1,b_1) & f(b_1,b_2) & \dots & f(b_1,b_n) \\ \vdots & & & \\ f(b_n,b_1) & f(b_n,b_2) & \dots & f(b_n,b_n) \end{pmatrix}$$

is equal to

$$[f]^{\mathcal{B}}_{\mathcal{B}} = \begin{pmatrix} f(b_1,b_1) & & 0 \\ & \ddots & \\ 0 & & f(b_n,b_n) \end{pmatrix}$$

This is equivalent to insisting that $f(b_i, b_j) = 0$ for all i,j satisfying $i \neq j$.

Such a basis is known as an orthogonal basis with respect to f.

In fact, as we shall see, we may ensure that we obtain diagonal matrices of particularly simple form.

- An $n \times n$ matrix is in real canonical form if it has the form

$$\begin{pmatrix} I_r & & 0 \\ & -I_s & \\ 0 & & 0_{n-r-s} \end{pmatrix}$$

where I_r is an identity matrix of size $r \times r$

I_s is an identity matrix of size $s \times s$

0_{n-r-s} is a zero matrix of size $(n-r-s) \times (n-r-s)$.

eg: the following are in real canonical form:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$n=2$, $r=2$; $n=2$, $r=0$, $s=2$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$ $n=3$, $r=2$, $s=1$; $n=3$, $r=1$, $s=1$

whereas the following are not

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

- Similarly, an $n \times n$ matrix is in complex canonical form if it has the form:

$$\begin{pmatrix} I_m & 0 \\ 0 & 0_{n-m} \end{pmatrix}$$

where I_m is an identity matrix of size $m \times m$.

0_{n-m} is a zero matrix of size $(n-m) \times (n-m)$

eg. the following are in complex canonical form:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

but the following aren't:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Example: Consider the symmetric bilinear form f, where

$$f\left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}\right) = x_1y_1 + 2x_1y_2 + 2x_2y_1 + 5x_2y_2$$

In terms of $\mathcal{E} = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$

$$[f]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix}1 & 2\\ 2 & 5\end{pmatrix}.$$

This has associated quadratic form q, defined via:

$$q\begin{pmatrix}x_1\\x_2\end{pmatrix} = f\left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}x_1\\x_2\end{pmatrix}\right)$$

i.e. $$q\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1^2 + 4x_1x_2 + 5x_2^2$$

and in terms of $\mathcal{E}$:

$$[q]_{\mathcal{E}}^{\mathcal{E}} = [f]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix}1 & 2\\ 2 & 5\end{pmatrix}$$

Consider the basis $\beta = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}-4\\2\end{pmatrix}\right\}$

$b_1 = \beta_1$ $\quad$ $B_2 = b_2$

Then,

$$f(b_1,b_2) = f\left(\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}-4\\2\end{pmatrix}\right) = (1)(-4) + 2(1)(2) + 2(0)(-4) + 5(0)(2).$$

i.e. $f(b_1,b_2) = 0$.

Similarly: $f(b_1,b_2) = 1$, $f(b_2,b_1) = 0$, $f(b_2,b_2) = 4$.

S.o. $$[f]_{\beta}^{\beta} = \begin{pmatrix}1 & 0\\ 0 & 4\end{pmatrix}.$$

How can we make the entry of 4 into an entry of 1?

Note: $f(b_2,b_2) = 4$.

So, $f\left(\frac{1}{\sqrt{4}}b_2, \frac{1}{\sqrt{4}}b_2\right) = f\left(\frac{1}{2}b_2, \frac{1}{2}b_2\right)$

assim al final tenho $\frac{1}{4}$

$$= \frac{1}{2} f\left(b_2, \tfrac{1}{2}b_2\right)$$
$$= \frac{1}{2}\cdot\frac{1}{2} f(b_2,b_2)$$
$$= \frac{1}{4} f(b_2,b_2)$$
$$= \frac{1}{4}\cdot 4 = \boxed{1}$$

Replace b_2 by $\frac{1}{2}b_2$ in $\beta = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}-4\\2\end{pmatrix}\right\}$

b_1 $\quad$ b_2

$\begin{pmatrix}-4\\+2\end{pmatrix}$ $\quad$ $\begin{pmatrix}-2\\+1\end{pmatrix}$

to obtain $$\mathcal{C} = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}-2\\1\end{pmatrix}\right\}$$

c_1 $\quad$ c_2

In terms of $\mathcal{C}$:

$$f(c_1,c_2) = f\left(\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}-2\\1\end{pmatrix}\right) = (1)(-2) + 2(1)(1) + 2(0)(-2) + 5(0)(1) = 0.$$

Similarly: $f(c_1,c_1) = 1$, $f(c_2,c_1) = 0$.

$f(c_2,c_2) = 1$

so $$[f]_{\mathcal{C}}^{\mathcal{C}} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$ ← The real and complex, canonical form of f.

In terms of matrix product

$$[f]_{\mathcal{C}}^{\mathcal{C}} = M^T [f]_{\mathcal{E}}^{\mathcal{E}} M$$

where $M = [Id]_{\mathcal{C}}^{\mathcal{E}}$ here $M = \begin{pmatrix}1 & -2\\ 0 & 1\end{pmatrix}$.

Let's verify.

$$M^T [f]^{\mathcal{E}}_{\mathcal{E}} M = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$$

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

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = [f]^{\mathcal{C}}_{\mathcal{C}}$$

So, with respect to f, $\mathcal{B}$ and $\mathcal{C}$ are both orthogonal basis

$$[f]^{\mathcal{B}}_{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}, \quad [f]^{\mathcal{C}}_{\mathcal{C}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

but $\mathcal{C}$ also leads to canonical form

In general, may determine such bases, and corresponding canonical forms, by using paired elementary row/column operations, or "double" operations

Some terminology

- The elementary row operation $R_i \to R_i + \lambda R_j$ corresponds to multiplying on the left by the elementary matrix $E(i,j,\lambda)$
  We shall write $E(i,j,\lambda)$ as $E_r(i,j,\lambda)$.
- For corresponding column operation $C_i \to C_i + \lambda C_j$ corresponds to multiplying on the right by the transpose of $E(i,j,\lambda)$
  We shall write $(E(i,j,\lambda))^T$ as $E^T(i,j,\lambda)$ or $E_c(i,j,\lambda)$.

eg: Consider $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $E(1,2;3) = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$.

$\Rightarrow E_r(1,2;3) = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, $E_c(1,2,3) = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}^T = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}$

$\Rightarrow A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \xrightarrow{R_1 \to R_1 + 3R_2} \begin{pmatrix} 10 & 14 \\ 3 & 4 \end{pmatrix}$

and $E_r(1,2;3)A = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 10 & 14 \\ 3 & 4 \end{pmatrix}$.

while

$\Rightarrow A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \xrightarrow{C_1 \to C_1 + 3C_2} \begin{pmatrix} 7 & 2 \\ 15 & 4 \end{pmatrix}$

and $AE_c(1,2,3) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 2 \\ 15 & 4 \end{pmatrix}$

Similarly,

- A row operation $R_i \to \lambda R_i$ corresponds to $D(i;\lambda)$, while a column operation $C_i \to \lambda C_i$ corresponds to $D^T(i,\lambda)$
- A row operation $R_i \leftrightarrow R_j$ corresponds to $P(i,j)$, while a column operation $C_i \leftrightarrow C_j$ corresponds to $P^T(i,j)$.

Using pairs of row/column operations, we can find (real and complex) canonical forms of given matrices

Example:

Consider f defined (over $\mathbb{R}$ or $\mathbb{C}$) via:

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 + 2x_1 y_2 + 2x_2 y_1 + 3x_2 y_2$$

with corresponding quadratic form

$$q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 4x_1x_2 + 3x_2^2$$

In terms of the 'standard' basis: $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$

$$[f]_{\mathcal{E}}^{\mathcal{E}} = [q]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}.$$

I try, but see that it doesn't work

$[f]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ $\xrightarrow{R_2 \to R_2 - 2R_1}$ $\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$ $\xrightarrow{C_2 \to C_2 - 2C_1}$ $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

si hago algo con rows lo hago tb con columns

is real canonical form.

$\xrightarrow{R_2 \to -R_2}$ $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $\xrightarrow{C_2 \to -C_2}$ $\begin{pmatrix} 1 & 0 \\ 0 & \textcircled{-1} \end{pmatrix}$

complex doesn't have -1.

not in complex canonical form

so instead

$\xrightarrow{R_2 \to iR_2}$ $\begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix}$ $\xrightarrow{C_2 \to iC_2}$ $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

in complex canonical form.

The real canonical form of f is $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

The complex canonical form of f is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Note that, in each of these cases, the matrix involves 2 non-zero entries in the diagonal.

The number of such non-zero entries, on the diagonal of a real or complex canonical form, is the rank of f and/or q.

So, here, f and/or q has rank 2.

Let's now interpret the reduction in terms of bases:

At the end of the first pair of operations, we obtain

$$E(2,1;-2)\ [f]_{\mathcal{E}}^{\mathcal{E}}\ E_c(2,1;-2)$$

check, $\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ ✓

So, $\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}^T \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

i.e. $M^T\ [f]_{\mathcal{E}}^{\mathcal{E}}\ M = [f]_{\beta}^{\beta}$

where $M = [Id]_{\beta}^{\mathcal{E}} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$

So, the basis $\beta = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$ is an orthogonal basis over $\mathbb{R}$:

$$[f]_{\beta}^{\beta} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Then, at the end of the second step, we obtained

$$D_r(2;i)\ E_r(2,1;-2)\ [f]_{\mathcal{E}}^{\mathcal{E}}\ E_c(2,1;-2)\ D_c(2;i)$$

(first step.)

Set, $N = E_c(2,1;-2)\ D_c(2;i) =$

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

Then, for $N = \begin{pmatrix} 1 & -2i \\ 0 & i \end{pmatrix}$

$$N^T\ [f]_{\mathcal{E}}^{\mathcal{E}}\ N = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Check: $\begin{pmatrix} 1 & 0 \\ -2i & i \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 & -2i \\ 0 & i \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -2i & i \end{pmatrix}\begin{pmatrix} 1 & \textcircled{0} \\ 2 & 3i-4i \end{pmatrix} =$

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

Then, in terms of the complex basis: $\beta = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2i \\ i \end{pmatrix} \right\}$

$$[f]_\beta^\beta = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$ complex canonical form.

Some other reductions of matrices to canonical form(s):

(1) $\begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix} \xrightarrow{R_2 \to \frac{1}{\sqrt{5}} R_2} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{5} \end{pmatrix} \xrightarrow{C_2 \to \frac{1}{\sqrt{5}} C_2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

(2) $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \xrightarrow{C_1 \leftrightarrow C_2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

(3) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_1 + \frac{1}{2} R_2} \begin{pmatrix} 1/2 & 1 \\ 1 & 0 \end{pmatrix} \xrightarrow{C_1 \to C_1 + \frac{1}{2} C_2} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$

$\xrightarrow{R_2 \to R_2 - R_1} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \xrightarrow{C_2 \to C_2 - C_1} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

December 9th 2019

We aim to show that every real/complex symmetric bilinear form can be diagonalised (i.e. that we can find a corresponding orthogonal basis).

To do so, we make use of the notion of an orthogonal complement.

Definition Consider a symmetric bilinear form $f: V \times V \to k$, where V is a vector space over the field k. Then, for any subset S of V, the orthogonal complement of S w.r.t. f, denoted by $S^\perp$, is the set:

$$S^\perp = \{ v \in V : f(v,s) = 0 \ \forall s \text{ in } S \}.$$

Equivalently, (since f is symmetric):

$$S^\perp = \{ v \in V \ \ f(s,v) = 0 \ \forall s \text{ in } S \}$$
$$\| \quad f(v,s)$$

Example Consider $f: \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ defined by:

$$f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right) = x_1 y_1 + x_2 y_2 + x_3 y_3$$ ⇒ standard or "dot" product.

and let $S = \left\{ \begin{pmatrix} 0 \\ 0 \\ s \end{pmatrix} \ s \in \mathbb{R} \right\}$

This is a subspace of $\mathbb{R}^3$, with possible basis $\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

Try to determine $S^\perp$

for a general vector $v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ to be in $S^\perp$, it must hold that

$$f\left(\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ s \end{pmatrix} \right) = 0 \quad \text{i.e. } v_1 \cdot 0 + v_2 \cdot 0 + v_3 \cdot s = 0 \quad \text{i.e. } v_3 \cdot s = 0$$ for each s in $\mathbb{R}$.

This leads to $v_3 = 0$; a general vector in $S^\perp$ has the form:

$\begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix}$ for $v_1, v_2 \in \mathbb{R}$.

i.e. $S^\perp = \left\{ \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix} = v_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + v_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} : v_1, v_2 \in \mathbb{R} \right\}$

This is also a subspace of $\mathbb{R}^3$ with possible basis $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$

Also, note, we can "retrieve" the whole of $\mathbb{R}^3$ by combining S and $S^\perp$: $\mathbb{R}^3 = S \oplus S^\perp$

bases: $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \quad \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \quad \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$

In general, given a symmetric bilinear form, $f: V \times V \to K$, and any subset S of V, it holds that:

- The orthogonal complement $S^\perp$ is a subspace of V over K.

Proof: Let's verify the 3 defining conditions of a subspaces:

(1) Consider $0 \in V$. Then $f(0,s) = 0 \quad \forall s$ in S.
(in fact $f(0,v) = 0 \quad \forall v$ in V)
So $0 \in S^\perp$ as required.

(2) Suppose $x, y \in S^\perp$, i.e. $\forall s$ in S: $f(x,s) = 0$ and $f(y,s) = 0$
Consider $f(x+y, s) = f(x,s) + f(y,s)$
$= 0 + 0$ since $x, y \in S^\perp$
ie. $f(x+y, s) = 0 \quad \forall s$ in S.
So $x+y \in S^\perp$ (also) as required

(3) Suppose $x \in S^\perp$, ie. $f(x,s) = 0 \quad \forall s$ in S.
Let $\lambda \in K$. Then, for all s in S:
$f(\lambda x, s) = \lambda f(x,s) = \lambda \cdot 0$ (ie. $f(\lambda x, s) = 0 \quad \forall s$ in S)
Hence, $\lambda x \in S^\perp$ (also) as required

This completes the proof: $S^\perp$ is a subspace of V ∎

Next, consider (the subspace) $S^\perp$ in the case where S is one-dimensional

Proposition: for a given symmetric bilinear form $f: V \times V \to K$, suppose v is a vector in V s.t. $f(v,v) \neq 0$.
$\Rightarrow V = \text{span}\{v\} \oplus \{v\}^\perp$
where $\text{span}\{v\} = \{\lambda v : \lambda \in K\}$.

Proof: Let's first show that $V = \text{span}\{v\} + \{v\}^\perp$
and then that the sum is direct
(intersection $= 0$)

1st: Consider a general vector w in V
Then, set $w_1 = \frac{f(w,v)}{f(v,v)} \cdot v$, $w_2 = w - \frac{f(w,v)}{f(v,v)} \cdot v$
(Note that $f(v,v) \neq 0$ by the assumption).

Also: $w = w_1 + w_2$.

Consider w_1: w_1 is of the form λv, where $\lambda = \frac{f(w,v)}{f(v,v)}$
so, $w_1 \in \text{span}\{v\}$

Consider w_2: try to show that $w_2 \in \{v\}^\perp$ ie. that $f(w_2, v) = 0$.

$$f(w_2, v) = f\left(w - \frac{f(w,v)}{f(v,v)} v, v\right)$$
$$= f(w,v) + f\left(-\frac{f(w,v)}{f(v,v)} v, v\right)$$
$$= f(w,v) - \frac{f(w,v)}{f(v,v)} f(v,v)$$
$$= f(w,v) - f(w,v)$$

So $f(w_2, v) = 0$ ie. $w_2 \in \{v\}^\perp$.

So, for any $w \in V$: $w = w_1 + w_2$ where $w_1 \in \text{span}\{v\}$, $w_2 \in \{v\}^\perp$.

ie. $V \subseteq \text{span}\{v\} + \{v\}^\perp$

Also, $\text{span}\{v\}, \{v\}^\perp$ were both subspaces of V, hence $\text{span}\{v\} + \{v\}^\perp \subseteq V$.

Overall: $V = \text{span}\{v\} + \{v\}^\perp$.

2nd: Now, show that the sum is direct
$\text{span}\{v\} \cap \{v\}^\perp = \{0\}$
Let $w \in \text{span}\{v\} \cap \{v\}^\perp$
Then, $w \in \text{span}\{v\}$, so $w = \lambda v$ for some λ in K
and $w \in \{v\}^\perp$, so $f(w,v) = 0$.

Then, $f(w,v) = f(\lambda v, v) = \lambda f(v,v)$

So, $f(w,v) = 0 \Rightarrow \lambda f(v,v) = 0$.

By assumption: $f(v,v) \neq 0$ so we must have $\lambda = 0$.

Then, $w = \lambda v = 0 \cdot v = 0$!

So $\text{span}\{v\} \cap \{v\}^{\perp} = \{0\}$ as required.

This completes the proof: $V = \text{span}\{v\} \oplus \{v\}^{\perp}$ ☑

December 10TH 2019

We make use of this proof on our proof of the following:

Diagonalisation Theorem for symmetric bilinear forms:

Let K be a field in which $2 \neq 0$, and let V be a vector space of finite, non-zero dimension over K.

Then, for every symmetric bilinear form $f: V \times V \to K$, f can be diagonalised, i.e. there exists an orthogonal basis for V, over K, with respect to f.

Proof. By induction on $\dim(V)$, the dimension of V over K.

- If $\dim(V) = 1 \Rightarrow \forall$ any basis C of V over K, $[f]_C^C$ is a 1×1 matrix, which is necessarily diagonal. i.e. every basis of V over K is orthogonal (this result holds).
- Next, assume the result holds for every vector space over K, with dimension equal to n (inductive assumption)
- Consider a vector space V, over K, for which $\dim(V) = n+1$
  - a). First, suppose that, for all vectors v in V, $f(v,v) = 0$ (*)
    (i.e. there is no vector v in V s.t $f(v,v) \neq 0$)
    Let's try to show that, in this case:
    $$f(u,w) = 0 \quad \forall u, w \text{ in } V.$$
    Note that, using a previous result, if $2 \neq 0$ in K
    $$\Rightarrow \forall u, w \text{ in } V \quad f(u,w) = \frac{1}{2}\left(f(u+w,u+w) - f(u,u) - f(w,w)\right)$$
    Then, using (*): $f(u,u) = 0$, $f(w,w) = 0$, $f(u+w,u+w) = 0$.

So, as required: $f(u,w) = 0$ for all u, w in V

i.e. f is the zero bilinear form.

Then, for any basis C, of V over K, $[f]_C^C$ is the zero matrix, which is diagonal, equivalently, every basis C is orthogonal.

  - b). Next, suppose that ~~there~~ exists a vector v in V s.t $f(v,v) \neq 0$

Using the final result from yesterday

$$V = \text{span}\{v\} \oplus \{v\}^{\perp}$$

So $\dim(V) = \dim(\text{span}\{v\} \oplus \{v\}^{\perp})$

Since $f(v,v) \neq 0$, v is not the zero vector ($v \neq 0$), so $\text{span}\{v\} = \{\lambda v : \lambda \in K\}$ has dimension 1 over K, with possible basis $\{v\}$, for $\text{span}\{v\}$ over K.

$$\Rightarrow \dim(\{v\}^{\perp}) = \dim(V) - \dim(\text{span}\{v\})$$
$$(n+1) \quad - \quad (1)$$

i.e. $\dim(\{v\}^{\perp}) = n$.

$\Rightarrow$ By the inductive assumption, there exists an orthogonal basis, $\{b_1, \ldots, b_n\}$ say, of $\{v\}^{\perp}$ over K.

Then $f(b_i, b_j) = 0 \quad \forall i, j$ satisfying $i \neq j$ (**)

Let's now "combine" bases for $\text{span}\{v\}$ and $\{v\}^{\perp}$ to obtain a basis $\{v, b_1, \ldots, b_n\}$ for V over K.

By definition of $\{v\}^{\perp}$: $f(v,s) = 0$ for every s in $\{v\}^{\perp}$

In particular: $f(v, b_i) = 0$ for $i = 1, \ldots, n$ (***)

Together, (**) and (***) show that $\beta = \{v, b_1, \ldots, b_n\}$ is an orthogonal basis for V over K, w.r.t. f.

Hence, f is diagonalisable as required, $[f]_{\beta}^{\beta}$ is diagonal matrix ☑

Having obtained a diagonal form for a real or complex symmetric bilinear form, we can proceed to obtain a matrix in real or complex canonical form respectively.

eg. suppose that $f: \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ is a symmetric bilinear form, such that, for a basis $\mathcal{A} = \{a_1, a_2, a_3\}$

$$[f]^{\mathcal{A}}_{\mathcal{A}} = \begin{pmatrix} f(a_1,a_1) & f(a_1,a_2) & f(a_1,a_3) \\ f(a_2,a_1) & f(a_2,a_2) & f(a_2,a_3) \\ f(a_3,a_1) & f(a_3,a_2) & f(a_3,a_3) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

i.e. $f(a_1,a_1) = 0$, $f(a_2,a_2) = 9$, $f(a_2,a_2) = 2$

$f(a_i,a_j) = 0$ whenever $i \neq j$.

By rearranging vectors in $\mathcal{A}$, we may obtain $\mathcal{A}' = \{a_2, a_3, a_1\}$

and $$[f]^{\mathcal{A}'}_{\mathcal{A}'} = \begin{pmatrix} f(a_2,a_2) & 0 & 0 \\ 0 & f(a_3,a_3) & 0 \\ 0 & 0 & f(a_1,a_1) \end{pmatrix} = \begin{pmatrix} 9 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Then, by "adjusting" vectors, we can obtain a basis leading to a real canonical form.

eg: $f(a_2,a_2) = 9 \Rightarrow f\left(\frac{1}{3}a_2, \frac{1}{3}a_2\right) = f(a_2,a_2) \cdot \frac{1}{3}\cdot\frac{1}{3} = \frac{1}{9}\cdot 9 = 1$

$f(a_3,a_3) = -2 \Rightarrow f\left(\frac{1}{\sqrt{2}}a_3, \frac{1}{\sqrt{2}}a_3\right) = -1$

so if we select the basis $\mathcal{B} = \left\{\frac{1}{3}a_2, \frac{1}{\sqrt{2}}a_3, a_1\right\}$

$$[f]^{\mathcal{B}}_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$ in real canonical form.

In general, for every symmetric bilinear form f, over $\mathbb{R}$ there exists a basis, $\mathcal{B}$ say, such that

$[f]^{\mathcal{B}}_{\mathcal{B}}$ is in real canonical form.

Similarly:

- For every symmetric bilinear form f, over $\mathbb{C}$, there exists a basis, $\mathcal{C}$ say, s.t. $[f]^{\mathcal{C}}_{\mathcal{C}}$ is in complex canonical form.

In addition, the following holds:

Sylvester's Law of Inertia

For every real or complex symmetric bilinear form, there is a corresponding unique real or complex canonical form, respectively.

Next, study real bilinear forms that share same key properties with the standard "dot product" bilinear form.

eg: consider $f: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, where:

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1y_1 + x_2y_2$$

Then, for any $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$: $f(x,x) = x_1^2 + x_2^2$

Note: $f(x,x) = x_1^2 + x_2^2 \geq 0$ for every x in $\mathbb{R}^2$.

$f(x,x) = x_1^2 + x_2^2 = 0 \iff x = 0$ (i.e. $x_1 = 0, x_2 = 0$).

December 13TH 2019

Then, we can define length of vector

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ as } \sqrt{f(x,x)} = \sqrt{x_1^2 + x_2^2}$$

We now study types of forms, known as inner products, that allow us to generalise notion of "length" to what is known as a norm, in the settings of real, and complex, vector spaces.

§ 3.3 Inner product spaces

A **real inner product space** consists of a **finite dimensional** vector space V, over $\mathbb{R}$, together with an **inner product** function, $\langle \cdot , \cdot \rangle : V \times V \longrightarrow \mathbb{R}$ satisfying for all $a, b, c \in V$ and every $\lambda \in \mathbb{R}$:

- $\langle a+b, c \rangle = \langle a, c \rangle + \langle b, c \rangle$ and
  $\langle c, a+b \rangle = \langle c, a \rangle + \langle c, b \rangle$
- $\langle \lambda a, b \rangle = \lambda \langle a, b \rangle$ and $\langle a, \lambda b \rangle = \lambda \langle a, b \rangle$
- $\langle a, b \rangle = \langle b, a \rangle$
- $\langle a, a \rangle \geq 0$.

while $\langle a, a \rangle = 0$ iff $a = 0$ (so $\langle a, a \rangle > 0$ whenever $a \neq 0$)

Eg Using the "standard" dot product

$f : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ from earlier, we obtain a real inner product

In general, we can use the real canonical form of a real symmetric bilinear form f to determine if, it forms a real inner product

> f defines an inner product
>
> iff
>
> the real canonical form of f is an identity matrix.

Suppose the basis $\beta = \{b_1, \dots, b_n\}$ of V over $\mathbb{R}$, leads to the real canonical form of f:

$$[f]^{\beta}_{\beta} = \begin{pmatrix} f(b_1, b_1) & & 0 \\ & \ddots & \\ 0 & & f(b_n, b_n) \end{pmatrix}$$

($\Rightarrow$) First, assume f is an inner product.

Then, $f(v, v) > 0$ whenever $v \neq 0$ (for $v \in V$)

In particular,

$f(b_i, b_i) > 0$ for each i, $1 \leq i \leq n$.

Choices for $f(b_i, b_i) = -1, 0, +1$.

(by definition of real canonical form).

Since $f(b_i, b_i) > 0$, must have

$f(b_i, b_i) = 1$ for $i = 1, \dots, n$

i.e. $$[f]^{\beta}_{\beta} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = I_n$$

an identity matrix as required.

($\Leftarrow$) Now, suppose $[f]^{\beta}_{\beta} = I_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$

Consider $v \in V$.

Since $\beta = \{b_1, \dots, b_n\}$ is a basis of V, $v = x_1 b_1 + \dots + x_n b_n$ for some $x_1, \dots, x_n \in \mathbb{R}$.

Then,

$$f(v, v) = [v]^T_{\beta} \, [f]^{\beta}_{\beta} \, [v]_{\beta}$$

Here, $[v]_{\beta} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ $[f]^{\beta}_{\beta} = I_n$

So $$f(v, v) = (x_1, \dots, x_n) \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$ $\therefore f(v, v) = x_1^2 + \dots + x_n^2$

Hence, $f(v, v) \geq 0$ for each $v \in V$, and $f(v, v) = 0 \Leftrightarrow x_1^2 + \dots + x_n^2 = 0$

$\Leftrightarrow x_1 = 0, \dots, x_n$

$\Leftrightarrow v = 0$ as required

Note If for some basis $\mathcal{B} = \{b_1, \dots, b_n\}$ and some form f:

$$[f]_{\mathcal{B}}^{\mathcal{B}} = I_n \quad \text{i.e.} \quad \left.\begin{array}{l} f(b_i, b_j) = 0 \\ \text{whenever } i \neq j \end{array}\right\} \text{ orthogonal}$$

and $f(b_i, b_i) = 1$ for each i,

then the basis $\mathcal{B} = \{b_1, \dots, b_n\}$ is orthonormal w.r.t. f.

In a real inner product space V, since $\langle v, v \rangle \geq 0$ for each v in V,

we can "safely" define the "length" or norm of a given vector $v \in V$

as $\|v\| = \sqrt{\langle v, v \rangle}$

Let's now consider complex vector spaces

for a real number $r \in \mathbb{R}$ we can define the length as $|r| = \sqrt{r \cdot r} = \sqrt{r^2}$

but for a complex number $z = a + ib$, we use, instead, a formula

such as $|z| = \sqrt{z \cdot \bar{z}}$ i.e. $|a+ib| = \sqrt{(a+ib)(a-ib)} = \sqrt{a^2+b^2}$.

($\bar{z}$: conjugate of z)

To define complex inner products more generally, we make use of this "conjugation":

- A complex inner product space is a finite dimensional vector space V, over $\mathbb{C}$, with a complex inner product function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$, s.t. $\forall a, b, c \in V$ and every $\lambda \in \mathbb{C}$

(1) $\langle a+b, c \rangle = \langle a, c \rangle + \langle b, c \rangle$

$\langle c, a+b \rangle = \langle c, a \rangle + \langle c, b \rangle$

(2) $\langle \lambda a, b \rangle = \lambda \langle a, b \rangle$

$\langle a, \lambda b \rangle = \bar{\lambda} \langle a, b \rangle$

(3) $\langle a, b \rangle = \overline{\langle b, a \rangle}$

(4) $\langle a, a \rangle \geq 0$

$\langle a, a \rangle = 0$ iff $a = 0$

Note Setting $b = a$ in (3)

$\langle a, a \rangle = \overline{\langle a, a \rangle}$ so $\langle a, a \rangle$ is a real number.

Note A form satisfying conditions (1), (2), (3) (but not necessarily (4))

is called Hermitian

Example

Consider $f : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$

given by $\langle u, w \rangle = u^T I_2 \bar{w} = u^T \bar{w}$

This defines a complex inner product space.

where, for instance

$$\left\langle \begin{pmatrix} 1+i \\ 1 \end{pmatrix}, \begin{pmatrix} 1+2i \\ 4 \end{pmatrix} \right\rangle = (1+i, 1)^T \overline{\begin{pmatrix} 1+2i \\ 4 \end{pmatrix}}$$

$$= (1+i)\overline{(1+2i)} + (1)(\bar{4})$$

$$= (1+i)(1-2i) + (1)(4)$$

$$= 1 + i - 2i + 2 + 4$$

$$= 7 - i$$

while,

$$\left\langle \begin{pmatrix} 1+2i \\ 2 \end{pmatrix}, \begin{pmatrix} 1+2i \\ 2 \end{pmatrix} \right\rangle = (1+2i)\overline{(1+2i)} + (2)(\bar{2})$$

$$= (1+2i)(1-2i) + (2)(2)$$

$$= 5 + 4$$

so, $\left\langle \begin{pmatrix} 1+2i \\ 2 \end{pmatrix}, \begin{pmatrix} 1+2i \\ 2 \end{pmatrix} \right\rangle = 9 \in \mathbb{R}$

We can define norms in the same way as before:

eg $\left\| \begin{pmatrix} 1+2i \\ 2 \end{pmatrix} \right\| = \sqrt{\left\langle \begin{pmatrix} 1+2i \\ 2 \end{pmatrix}, \begin{pmatrix} 1+2i \\ 2 \end{pmatrix} \right\rangle} = \sqrt{9} = 3$

In this example, the inner product we have used is the standard complex inner product: $\langle u, w \rangle = u^T \bar{w}$.

In general, what type of matrix A can we choose so that

$\langle u, w \rangle = u^T A \bar{w}$

satisfies conditions (1), (2), (3) in definition, i.e. is a Hermitian form?

For condition (3) in particular, require

$\langle u, w \rangle = \overline{\langle w, u \rangle} \quad \forall u, w \in V$.

Then: $\overline{\langle w, u \rangle} = \overline{w^T A \bar{u}}$

$= \bar{w}^T \bar{A} \bar{\bar{u}}$

i.e. $\langle \overline{w, u} \rangle = \bar{w}^T \bar{A} u$ (*)

$\langle u, w \rangle = u^T A \bar{w}$. $\quad [(3)^T = (3)]$

Since $\langle u, w \rangle \in \mathbb{C}$, $\langle u, w \rangle^T = \langle u, w \rangle$

So, $\langle u, w \rangle = (u^T A \bar{w})^T$

$= \bar{w}^T A^T (u^T)^T$ $\quad \} (MN)^T = N^T M^T$.

Then, $\langle u, w \rangle = \bar{w}^T A^T u$ (**)

To ensure that (*) and (**) give the same answer $\forall u, w \in V$

(i.e. to ensure $\langle u, w \rangle = \overline{\langle w, u \rangle} \quad \forall u, w \in V$)

we require $\boxed{\bar{A} = A^T}$

Any $n \times n$ complex matrix satisfying this is a Hermitian matrix.

Example

If $A = \begin{pmatrix} 17 & 2+i \\ 2-i & 5 \end{pmatrix}$ $\quad \} \Rightarrow A^T = \begin{pmatrix} 17 & 2-i \\ 2+i & 5 \end{pmatrix}$

so $\bar{A} = A^T$.

i.e. A is Hermitian

$\bar{A} = \begin{pmatrix} \overline{17} & \overline{2+i} \\ \overline{2-i} & \overline{5} \end{pmatrix} = \begin{pmatrix} 17 & 2-i \\ 2+i & 5 \end{pmatrix}$

Key result concerning Hermitian matrices: Spectral Theorem for Hermitian matrices:

Let M be an $n \times n$ Hermitian matrix,

Then:

(1) M is diagonalisable

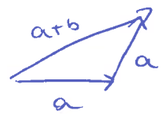
(2) Every eigenvalue of M is a real number.

(3) Eigenvectors of M corresponding to distinct eigenvalues are orthogonal for each other in terms of the standard complex inner product.

(4) There exists an orthonormal basis of eigenvectors for M.

---

2 important general results concerning complex (or real) inner product spaces:



Let V be a complex inner product space with inner product

$\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{C}$

The following holds:

$\forall a, b \in V \quad |\langle a, b \rangle| \leq \|a\| \, \|b\|$

(Cauchy-Schwarz inequality).

Proof: Consider $\langle a - \lambda b, a - \lambda b \rangle$ for $\lambda \in \mathbb{C}$.

Note: This is of the form $\langle v, v \rangle$, so $\langle a - \lambda b, a - \lambda b \rangle \geq 0$ (*)

Then, $\langle a - b\lambda, a - \lambda b \rangle = \langle a, a \rangle + \langle a, -\lambda b \rangle + \langle -\lambda b, a \rangle + \langle -\lambda b, -\lambda b \rangle =$

$= \langle a, a \rangle + -\bar{\lambda} \langle a, b \rangle - \lambda \langle b, a \rangle + (-\lambda)(-\bar{\lambda}) \langle b, b \rangle$

So, $\langle a - \lambda b, a - \lambda b \rangle = \|a\|^2 - \bar{\lambda} \langle a, b \rangle - \lambda \overline{\langle a, b \rangle} + (\lambda \bar{\lambda}) \|b\|^2 =$

$= \|a\|^2 - \bar{\lambda} \langle a, b \rangle - \lambda \overline{\langle a, b \rangle} + |\lambda|^2 \|b\|^2$

Set $\lambda = \dfrac{\langle a, b \rangle}{\langle b, b \rangle} = \dfrac{\langle a, b \rangle}{\|b\|^2}$ assuming $b \neq 0$.

We obtain

$$\langle a-\lambda b, a-\lambda b\rangle = \|a\|^2 - \frac{\overline{\langle a,b\rangle}}{\|b\|^2}\langle a,b\rangle + \frac{\langle a,b\rangle}{\|b\|^2}\overline{\langle a,b\rangle} + \left(\frac{|\langle a,b\rangle|}{\|b\|^2}\right)^2 \|b\|^2$$

Note: $\langle a,b\rangle\overline{\langle a,b\rangle} = |\langle a,b\rangle|^2$

$z\bar{z} = |z|^2$

Then, $\langle a-\lambda b, a-\lambda b\rangle = \|a\|^2 - \frac{|\langle a,b\rangle|^2}{\|b\|^2} - \frac{|\langle a,b\rangle|^2}{\|b\|^2} + \frac{|\langle a,b\rangle|^2}{\|b\|^2} -$

$$= \|a\|^2 - \frac{|\langle a,b\rangle|^2}{\|b\|^2}$$

Using (*):

$$\|a\|^2 - \frac{|\langle a,b\rangle|^2}{\|b\|^2} \geq 0$$

i.e. $\|a\|^2 \geq \frac{|\langle a,b\rangle|^2}{\|b\|^2}$ ↘ Since $\|b\|^2 \geq 0$.

Then, $\|a\|^2\|b\|^2 \geq |\langle a,b\rangle|^2$

finally, since $\|a\| \geq 0$, $\|b\| \geq 0$, $|\langle a,b\rangle| \geq 0$ we can "safely" take square roots to obtain, as required:

$$|\langle a,b\rangle| \leq \|a\|\,\|b\| \quad \text{for } b \neq 0.$$

for $b=0$, note $\|b\|=0$ so $\|a\|\,\|b\| = 0$

while $|\langle a,b\rangle| = |\langle a,0\rangle| = |0| = 0$

So $|\langle a,b\rangle| \leq \|a\|\,\|b\|$ holds here too $(0 \leq 0)$.

We can now prove the Triangle Inequality for such a complex inner product space V:

For all $a,b \in V$: $\|a+b\| \leq \|a\| + \|b\|$

Proof: Consider $\|a+b\|^2 = \langle a+b, a+b\rangle$

Try to show that $\|a+b\|^2 \leq (\|a\| + \|b\|)^2$

$$\|a+b\|^2 = \langle a+b, a+b\rangle$$
$$= \langle a,a\rangle + \langle a,b\rangle + \langle b,a\rangle + \langle b,b\rangle$$
$$= \|a\|^2 + \langle a,b\rangle + \overline{\langle a,b\rangle} + \|b\|^2$$

Note: for any $z = x+iy \in \mathbb{C}$

$z + \bar{z} = (x+iy)(x-iy) = 2x = 2\underbrace{R(z)}_{\text{real part of } z}$.

Also, $x \leq \sqrt{x^2+y^2}$ i.e. $R(z) \leq |z|$.

Then, $\|a+b\|^2 = \|a\|^2 + 2R(\langle a,b\rangle) + \|b\|^2 \leq \|a\|^2 + 2|\langle a,b\rangle| + \|b\|^2$

$\leq \|a\|^2 + 2\|a\|\,\|b\| + \|b\|^2$ (by The Cauchy-Schwarz inequality)

So $\|a+b\|^2 \leq (\|a\| + \|b\|)^2$

Note $\|a\| \geq 0$, $\|b\| \geq 0$, $\|a+b\| \geq 0$ so may safely take square roots to obtain $\|a+b\| \leq \|a\| + \|b\|$.