# 2201 Algebra 3: Further Linear Algebra Notes

## Based on the 2016 autumn lectures by Dr I Strouthos

03-10-16 **Further Linear Algebra**

Isidoros Strouthos
i.strouthos@ucl.ac.uk
Room 501
Tues 2.15 - 4pm
Thurs 5.15 - 6pm
Fri 3.15 - 5pm

Course material: on Moodle
(lecture notes & exercises)
First deadline: 21-10-2016

Possibly useful textbooks:
- Cohn { Elements of Linear Algebra
  Algebra (Vol 1)
- Curtis - Abstract Linear Algebra
- Kaye & Wilson - Linear Algebra
- Lang - Linear Algebra
- Lipschutz & Lipson - Linear Algebra
- Lipschutz - 3000 solved problems in Linear Algebra.

Overview of Course
- Polynomial Rings:
  Studying polynomials, and try to examine how they are similar to the natural numbers or integers: there is a Euclidian algorithm.
- Linear maps and the Jordan normal form.
  Some matrices can be diagonalised, e.g. consider $\begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$:
  $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ is an eigenvector corresponding to the eigenvalue 2.
  $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ is an eigenvector corresponding to the eigenvalue 3.
  Not every square matrix can be diagonalised, e.g. consider $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. This has complex eigenvalues.
  But even with complex numbers not all matrices can be diagonalised as we might not have enough linearly independent eigenvectors.
  Jordan showed that any square matrix (over $\mathbb{C}$) can be transformed to a simpler form.

- Linear and bilinear forms

  Linear form: a linear map that takes as input a single <u>vector</u> and returns as output a single <u>number</u>.

  - correspond to "row matrices"

  Bilinear form: input = two vectors,

  output = a single vector

  e.g. $(x_1 \ y_1)\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$

  $= (x_1 \ y_1)\begin{pmatrix} x_2 + 2y_2 \\ 3x_2 + 4y_2 \end{pmatrix}$

  $= x_1x_2 + 2x_1y_2 + 3x_2y_1 + 4y_1y_2$

- Inner product spaces

  Special case of a bilinear form

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

  $(x \ y \ z \ t)\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = x^2 + y^2 + z^2 - t^2$

MATH 2201

04-10-16

## Some notation

We will use the "usual" symbols $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, to denote, respectively, the sets of integers, rational numbers, real numbers, complex numbers.

Note that, for us, $\mathbb{N}$ will denote the set of positive integers, and $\mathbb{N}_0$ will denote the set of non-negative integers.

$\mathbb{N} = \{1, 2, 3, \dots\}$

$\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$

# Chapter 1 - Polynomial rings

## Some basic objects in abstract algebra

We start with the notation of a group, a basic object involving a single operation.

Def: A group consists of a set, $G$, together with an operation, $*$, such that the following conditions are satisfied:

1). If $a, b \in G$, then $a * b \in G$ ← closure
2). If $a, b, c \in G$, then $a*(b*c) = (a*b)*c$ ← associativity
3). There exists an identity element, $e$, 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, $b$, such that: $a*b = e$ and $b*a = e$

Note: • A structure satisfying conditions (1) and (2) is a semigroup
• A structure satisfying conditions (1), (2) and (3) is a monoid.

Examples

- If we take $\mathbb{N}$ with addition as the operation, we have a semigroup.
- If we take $\mathbb{N}$ with multiplication as the operation, we have a monoid.
- Consider $\mathbb{N}_0$ with addition: it forms a monoid.
- " " multiplication: " " monoid also.
- Consider $\mathbb{Z}$ with addition $\rightarrow$ group
- " " multiplication $\rightarrow$ monoid
- Consider $\mathbb{R}$ with addition $\rightarrow$ group
- " " multiplication $\rightarrow$ monoid.
- However $\mathbb{R} \setminus \{0\}$ with multiplication $\rightarrow$ group

Note:

A semigroup, monoid or group that also satisfies: $\forall a, b \in G$, $a * b = b * a$ $\leftarrow$ commutativity is abelian.

Let's now consider a structure that involves two operations.

Def: A ring consists of a set $R$ together with two operations, addition, denoted by '$+$', and multiplication, denoted by '$\cdot$' (we often write '$a \cdot b$' as '$ab$') such that:

R forms an additive abelian group.

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

MATH 2201

04-10-16

R forms a multiplicative monoid
(6). If $a, b \in R$ : $ab \in R$
(7). If $a, b, c \in R$ : $a(bc) = (ab)c$
(8). There exists an identity element for multiplication, $1$, in $R$, s.t. $\forall a \in R$ : $1 \cdot a = a$ and $a \cdot 1 = a$

distributivity
(9). $\forall a, b, c \in R$ : $a(b+c) = ab + ac$
(10). $\forall a, b, c \in R$ : $(b+c)a = ba + ca$

Examples:
In the following, addition is "usual addition", multiplication is "usual multiplication", 0 denotes the number zero and 1 denotes the number one.
Then:
- $\mathbb{N}$ and $\mathbb{N}_0$ do not form rings
  conditions 3, 4 fail → condition 4 fails.
- $\mathbb{Z}$ forms a ring
- $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ form rings
  In fact, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ form fields
  A field is a ring which is commutative and where every non-zero element has a multiplicative inverse.

In general, if a ring also satisfies
11). $\forall a, b \in R$ : $ab = ba$
we have a commutative ring.

Any commutative ring that also satisfies
12). $\forall a \in R \setminus \{0\}$, $\exists$ an element $b$ s.t. $ab = 1$ and $ba = 1$
is a field.

(A ring that satisfies condition (12), but not necessarily (11) is often known as a division ring)

Most rings we will study in this course are commutative.

Example of non-commutative ring

Consider the set of $2\times 2$ matrices with real number entries

$$M_2(\mathbb{R}) = \left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} : a_{11}, a_{12}, a_{21}, a_{22} \in \mathbb{R} \right\}$$

with '+' being addition of matrices (identity element, 0, is $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$)

'$\cdot$' being multiplication of matrices (identity element, 1, is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$).

Then $M_2(\mathbb{R})$ forms a ring, but multiplication is not commutative

e.g. $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

$$\begin{pmatrix} 4 & 6 \\ 3 & 4 \end{pmatrix} \neq \begin{pmatrix} 1 & 3 \\ 3 & 7 \end{pmatrix}$$

07-10-16

From last time:
— definitions of group, semigroup, monoid, ring, field.
Key example of a ring (in this chapter / course):
The ring of polynomials over another ring.

Let $R$ be a ring. Then a polynomial over $R$ is a formal expression of the form
$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots$$
where $x$ is an indeterminate or a variable,
and $a_0, a_1, a_2, \dots$ are coefficients,
and where only finitely many of the coefficients are nonzero.

For example, if we take $R$ to be $\mathbb{R}$, the following are polynomials over $\mathbb{R}$:

$0 = 0 + 0x + 0x^2 + \dots$

$1 = 1 + 0x + 0x^2 + \dots$

$3$

$x + 1$

$x - 2$

$2x^2 - 2$

$x^2 + 5x + 6$

$x^2 + 1$

$x^2 - 2$

$-x^3 + 1$

However
$1 + x + x^2 + \dots + x^n + \dots$ (which goes on forever)
is not a polynomial over $\mathbb{R}$.

The zero polynomial is the polynomial where ever coefficient is zero: $a_0 = 0$, $a_1 = 0, \dots$
$\Rightarrow 0 + 0x + 0x^2 + \dots$

Two polynomials,
$$f(x) = a_0 + a_1 x + \ldots = \sum_i a_i x^i$$
and
$$g(x) = b_0 + b_1 x + \ldots = \sum_i b_i x^i,$$
are equal if all corresponding coefficients are equal.

i.e. $a_0 = b_0$, $a_1 = b_1$, ...

e.g. $1 + x + 2x^2 = 1 + x + 2x^2$

$1 + x \neq 1 + x^2$

A constant polynomial is one of the form $f(x) = a_0$.

Goal for remainder of chapter:

To show that polynomials over a ring are similar to $\mathbb{Z}$:

they form a ring and there is a Euclidian algorithm on them.

The set of all polynomials over a ring $R$ is denoted by $R[x]$.

Key notation to find Euclidean algorithm:

we need to compare the sizes of polynomials using the notion of degree.

Definition:

Given a polynomial $f \in R[x]$, for a ring $R$, the degree of $f$ is defined as follows:

- If $f$ is not zero, the degree of $f$, $\deg(f)$, is the largest non-negative integer for which $a_n \neq 0$, where $f = a_0 + a_1 x + a_2 x^2 + \ldots$

• If $f = 0$, then $\deg(f) = -\infty$ (:)

For any ring $R$, $R[x]$ becomes a ring under the following operations:

Addition: Suppose $f(x) = \sum_i a_i x^i = a_0 + a_1 x + \ldots$
and $g(x) = \sum_i b_i x^i = b_0 + b_1 x + \ldots$
then the sum of $f$ and $g$, $f+g$, is defined as follows:

$$(f+g)(x) = f(x) + g(x) = \sum_i (a_i + b_i) x^i$$
$$= (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \ldots$$

How is the degree affected by the sum?

examples:
- $f(x) = x+1$, $g(x) = x^2 + 5x + 6$, $(f+g)(x) = x^2 + 6x + 7$
  $\deg(f) = 1$, $\deg(g) = 2$, $\deg(f+g) = 2$
- $f(x) = 0$, $g(x) = x^2 + 3x + 7$, $(f+g)(x) = x^2 + 3x + 7$
  $\deg(f) = -\infty$, $\deg(g) = 2$, $\deg(f+g) = 2$
- $f(x) = x^2 + 1$, $g(x) = -x^2 + x + 1$, $(f+g)(x) = x + 2$
  $\deg(f) = 2$, $\deg(g) = 2$, $\deg(f+g) = 1$

In general: $\deg(f+g) \leq \max\{\deg(f), \deg(g)\}$

Multiplication

If $f(x) = a_0 + a_1 x + \ldots + a_n x^n$, $\deg(f) = n$, $a_n \neq 0$
and $g(x) = b_0 + b_1 x + \ldots + b_m x^m$, $\deg(g) = m$, $b_m \neq 0$
Then the product is defined as follows:

$$(fg)(x) = f(x)g(x)$$
$$= a_0 b_0 + (a_0 b_1 + a_1 b_0)x + (a_0 b_2 + a_1 b_1 + a_2 b_0)x^2 + \ldots + a_n b_m x^{n+m}$$
$$= \sum_i^{n+m} \left( \sum_{j=0}^{i} a_j b_{i-j} \right) x^i$$

eg. $f(x) = x + 1$, $g(x) = x^2 + 2$
then $(fg)(x) = (x+1)(x^2+2)$
$= x^3 + x^2 + 2x + 2$

If we take the coefficients to be in a field, then if $a \neq 0$ and $b \neq 0$ : $ab \neq 0$.
(Not true for modular arithmetic etc...)

So, if $R$ is a field (we will only consider $\mathbb{R}$ or $\mathbb{C}$) then we can say that $\deg(fg) = n+m = \deg(f) + \deg(g)$

This also works if $f = 0$ or $g = 0$.
Suppose $f = 0$. Then, for any $g$ : $fg = 0$.

$$\deg(fg) = \deg(f) + \deg(g) = -\infty + m = -\infty.$$

So, we can say that for any $f, g \in k(x)$, where $k$ is a field,

$$\deg(fg) = \deg(f) + \deg(g).$$

Under these operations, any set of polynomials over a field forms a ring.
Suppose that $f$ is a non-zero polynomial, $f = \sum_i a_i x^i$
If $n$ is the largest non-negative integer for which $a_n \neq 0$ (i.e. $f$ has degree $n$) then $a_n$ is the leading coefficient of $f$.

examples: $f(x) = x^3 + 3 \Rightarrow$ leading coefficient is 1.
$f(x) = -3x^2 + 7 \Rightarrow$ leading coefficient is $-3$.

A polynomial with leading coefficient equal to 1 is a <u>monic</u> polynomial.
e.g. $f(x) = x^3 + 3$ is monic
$f(x) = -3x^2 + 7$ is not monic.

Note if $f(x) = a_0$ and $a_0 \neq 0$, then $\deg(f) = 0$

MATH2201

07-10-16

# Divison in rings and the Euclidian algorithm

Def:

Suppose $a, b$ are elements in a ring $R$. Then $a$ divides $b$, or $a$ is a divisor of $b$, if there exists an element $c \in R$ such that
$$ac = b.$$
If so, we wride $a \mid b$.
If $a$ does note divide $b$, we write $a \nmid b$

Example

If $R = \mathbb{Z}$, then $\underset{2\times 3=6}{2 \mid 6}$, $\underset{-2\times -3=6}{-2 \mid 6}$, $\underset{1\times 3=3}{1 \mid 3}$, $\underset{3\times 1=3}{3 \mid 3}$, but $3 \nmid 7$.

Note: everything divides 0, but 0 only divides itself.
For any $n \in \mathbb{Z}$: $n \cdot 0 = 0$ so $n \mid 0$, but $0 \cdot c = n$ has no solution if $n \neq 0 \Rightarrow 0 \nmid n$.

If $R = \mathbb{R}[x]$ then $x+2$ divides $x^2+5x+6$, but $x+2$ does not divide $x^2+1$.

Again, the zero polynomial divides only itself, but every polynomial divides zero.

From the point of view of division, there are three types of elements in the rings we will study: for a ring $R$

1). an element $u \in R$ is a unit if there exists a multiplicative inverse of $u$ in $R$, i.e. an element $u' \in R$ such that $uu' = 1$ and $u'u = 1$.
The set of all units in $R$ is denoted by $U(R)$.
Examples: $U(\mathbb{Z}) = \{+1, -1\}$ $(-1)(-1) = 1$
What about $\mathbb{R}[x]$: Any polynomial of degree 1 or

or greater is not a unit

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

all non-zero, constant polynomials

e.g. $x+1$ is not a unit: there is no $f \in \mathbb{R}[x]$ such that $(x+1)\cdot f = 1$.

We can show this using

$$\deg(fg) = \deg(f)\deg(g).$$

Suppose that $f$ is a unit in $\mathbb{R}[x]$

Then, there is some $g \in \mathbb{R}[x]$ such that $fg = 1$.

So $\deg(f) + \deg(g) = \deg(\overbrace{fg}^{=1})$

(Possibilities: $\deg(f) = -\infty, 0, 1, 2, \ldots$)

So $\deg(f) + \deg(g) = 0$

ie. $\deg(f) = 0$ (and $\deg(g) = 0$.

So $f$ is a nonzero constant.

If $f$ is a non-zero constant ie.
$f = a \in \mathbb{R} \setminus \{0\}$ then $f$ is a unit: $a \cdot a^{-1} = 1$,
each non-zero constant in a field, like $\mathbb{R}$, is invertible.

Similarly $U(\mathbb{C}[x]) = \mathbb{C} \setminus \{0\}$ (non-zero, constant polynomials)

2). an element $a \in R$ is <u>irreducible</u> if $a$ is not a unit and if $a = bc$ for $b, c \in R$, then $b$ or $c$ is a unit.

Examples:

In $\mathbb{Z}$ the irreducible elements are essentially the primes: $\{\pm 2, \pm 3, \pm 5, \ldots\}$.

6 is not irreducible as $6 = 2 \times 3$.

In $\mathbb{C}[x]$, the irreducible elements are ....?

If $\deg(f) = 1$ then $f$ is irreducible.

We can show this, again, using

MATH 2201

07-10-16

$\deg(fg) = \deg(f) + \deg(g)$.
Suppose that $\deg(f) = 1$, and that $f = gh$.
Then $\deg(g) + \deg(h) = 1$,
i.e. $\deg(g) = 0$ or $\deg(h) = 0 \Rightarrow g$ or $h$ is a unit.
So $f$ is irreducible.
This also works in $\mathbb{R}[x]$: if $\deg(f) = 1$ then $f$ is irreducible.
If we work in $\mathbb{C}[x]$, then any polynomial of degree greater than 1 is <u>reducible</u>.

3). an element $a \in R$ is reducible if it is not a unit and not irreducible.
Every polynomial with complex coefficients can be factorised into linear factors (over $\mathbb{C}$):
e.g. $x^2 + 5x + 6 = (x+2)(x+3)$
$x^2 + 1 = (x+i)(x-i)$

Fundamental Thm of Algebra

For a proof, see MATH 2101.
So, in $\mathbb{C}[x]$ all polynomials of degree 2, 3, 4, ... are reducible.
Note: the zero polynomial is reducible: $0 = 0 \times 0$
i.e. 0 can be written as a product without using any products.

In $\mathbb{R}[x]$, all polynomials of degree 1 are irreducible, but there are also irreducible polynomials of greater degree.
e.g. $f(x) = x^2 + 1$ is irreducible in $\mathbb{R}[x]$,
we can only factorise it using non-zero constants:
$$x^2 + 1 = 1 \cdot (x^2+1) = 5\left(\tfrac{1}{5}(x^2+1)\right)$$

MATH 2201

10-10-16

Previously on Algebra 3:

- notion of degree in $\underbrace{k[x]}_{\text{ring of polynomials with coefficients in a field } k}$
- some results involving degrees:

For $f, g \in k[x]$ $\boxed{\deg(0) = -\infty}$

$\deg(f+g) \leq \max\{\deg f, \deg g\}$

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

Consequence: $\deg(-f) = \deg(-1 \cdot f)$
$= \deg(-1) + \deg(f)$
$= 0 + \deg(f)$

$\Rightarrow \deg(-f) = \deg(f)$

Motto: "There's no need for anger,
there's no need for blame,
there's nothing to prove here,
everything's still the same!"

We will try to "translate" the Euclidean algorithm and associated results from $\mathbb{Z}$ to $k[x]$.

In $\mathbb{Z}$, given $a, b \in \mathbb{Z}$ and $b \neq 0$, then there exists unique $q, r \in \mathbb{Z}$ such that

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

try to make remainder a number

e.g. $-7 = 0.3 - 7$ $\quad$ $r$ s.t. $0 \leq r < 3$.

$-7 = -1.3 + (3-7)$ $\qquad$ $-7 = -1.3 + (-4)$

$-7 = -2.3 + (6-7)$ $\qquad$ $-7 = -2.3 + (-1)$

$-7 = -3.3 + (9-7)$ $\qquad$ $-7 = \underbrace{-3}_{q} \cdot 3 + \underbrace{(2)}_{r}$

To show that $q, r$ exist in the above statement, just carry out the process of changing the

remainder step-by-step until it satisfies
$0 \le r < |b|$.

To show that $q, r$ are unique, suppose:

① — $a = q_1 b + r_1$ , $0 \le r_1 < |b|$

② — $a = q_2 b + r_2$ , $0 \le r_2 < |b|$

Then, subtracting ① from ②:

$$0 = (q_2 - q_1)b + (r_2 - r_1)$$

so $b(q_1 - q_2) = (r_2 - r_1)$

so $|b(q_1 - q_2)| = |r_2 - r_1|$

$\Rightarrow |b||q_1 - q_2| = |r_2 - r_1|$

(But note that $0 \le r_1, r_2 < |b|$
so $0 \le |r_2 - r_1| < |b|$ )

Hence $0 \le |b||q_1 - q_2| < |b|$

i.e. $0 \le |q_1 - q_2| < 0$

so $q_1 = q_2$ note $q_1, q_2 \in \mathbb{Z}$

Substituting $q_1 = q_2$ back into ①, ② we obtain
$r_1 = r_2$, so $q$ and $r$ are uniqely determined.

Let's now consider $k[x]$.
Here, for $a, b \in k[x]$, where $b \neq 0$, we can find unique $q, r \in k[x]$ such that
$a = qb + r$ where $\deg(r) < \deg(b)$.

MATH 2201

10-10-16

**Example**

Take $\mathbb{R}[x]$.

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

Then $\underbrace{(x^2+3x+2)}_{a} = \underbrace{(x+2)}_{q}\underbrace{(x+1)}_{b} + \underbrace{0}_{r}$

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

$-\infty < 1$.

whereas if $a = x^2+3x+3$, we obtain

$\underbrace{(x^2+3x+3)}_{a} = \underbrace{(x+2)}_{q}\underbrace{(x+1)}_{b} + \underbrace{1}_{r}$

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

$0 < 1$

**Detailed example** (how to find $q, r$):

Let $a = x^2+4x+7$, $b = 2x+4$

Start with any $q, r$ that work eg. $q = 0$

$x^2+4x+7 = \underbrace{0}_{q} \cdot (2x+4) + \underbrace{x^2+4x+7}_{r}$

$\deg(r) \not< \deg(b)$

as $2 \geq 1$ so we have to make $r$ a smaller degree.

Notice: $x^2+4x+7 = \frac{1}{2}x(2x+4) + (x^2+4x+7) - (x^2+2x)$

$= \frac{1}{2}x(2x+4) + (2x+7)$

here $\deg(r) = \deg(b)$ so continue the process to get $\deg(r) < \deg(b)$.

Notice: $\underbrace{x^2+4x+7}_{a} = \frac{1}{2}x(2x+4) + 1\cdot(2x+4) + \big((2x+7)-(2x+4)\big)$

$= \underbrace{(\tfrac{1}{2}x+1)}_{q}\underbrace{(2x+4)}_{b} + \underbrace{3}_{r}$

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

$0 < 1$

So, we have found $q, r$:

$q = \frac{1}{2}x + 1$, $r = 3$

This can be summarised in terms of a long division of polynomials.

$$
\begin{array}{r}
\tfrac{1}{2}x \quad + 1 \phantom{+7} \\
(2x+4)\,\overline{\big)\, x^2 + 4x + 7} \\
-(x^2 + 2x + 0) \\
\hline
2x + 7 \\
-(2x + 4) \\
\hline
3
\end{array}
$$

so $x^2 + 4x + 7 = (\tfrac{1}{2}x + 1)(2x+4) + 3$

Similarly for $a = 3x^3 + 5$, $b = 2x + 4$

$$
\begin{array}{r}
\tfrac{3}{2}x^2 + 3x \quad + 6 \phantom{+5} \\
(2x+4)\,\overline{\big)\, 3x^3 + 0x^2 + 0x + 5} \\
-(3x^3 + 6x^2 + 0x + 0) \\
\hline
6x^2 + 0x + 5 \\
-(6x^2 + 12x + 0) \\
\hline
12x + 5 \\
-(12x + 24) \\
\hline
-19
\end{array}
$$

So $a = qb + r$ for $\deg(r) < \deg(b)$
if $q = \frac{3}{2}x^2 - 3x + 6$, $r = -19$

Let's now prove the statement that, for any $a, b \in k[x]$, where $b \neq 0$, there exists unique $q, r \in k[x]$ such that $a = qb + r$ and $\deg(r) < \deg(b)$.

<u>Proof of uniqueness:</u>

Suppose that $a = q_1 b + r_1$ $\quad \deg(r_1) < \deg(b)$
$\qquad a = q_2 b + r_2$ $\quad \deg(r_2) < \deg(b)$

Then, as in the case of $\mathbb{Z}$:

$$b(q_1 - q_2) = r_2 - r_1$$

MATH 2201

10-10-16

Taking degrees:

$$\deg(b(q_1 - q_2)) = \deg(r_2 - r_1)$$

So $\deg(b) + \deg(q_1 - q_2) = \deg(r_2 - r_1)$

$$= \deg(r_2 + (-r_1))$$

Then $\deg(b) + \deg(q_1 - q_2) \leq \max\{\deg(r_2), \deg(r_1)\}$

$$< \deg(b)$$

---

11-10-16 Let's complete the proof from yesterday:

Proposition:

Suppose that $k$ is a field and that $a, b \in k[x]$, where $b \neq 0$.
Then, there exists unique $q, r \in k[x]$ such that $a = qb + r$ where $\deg(r) < \deg(b)$.

Proof:

Let's first show that such $q, r$ exist.
Start with any $q', r'$ that satisfy

$$a = q'b + r'$$

(e.g. start with $q' = 0$, $r' = a$: $a = 0 \cdot b + a$)
(At any stage) if $\deg(r') \geq \deg(b)$, then $r', b$ have the form:

$$r' = r_n x^n + \dots + r_0 \quad \text{where } r_n \neq 0$$

$$b = b_m x^m + \dots + b_0 \quad \text{where } b_m \neq 0 \qquad (b = b(x))$$

(note: $b \neq 0$ by assumption, so this is okay)
where $n \geq m$ (since $\deg(r') \geq \deg(b)$)

Then consider:

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

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

Then $r' - r_n b_m^{-1} x^{n-m} b$

$$= (r_n x^n + \dots + r_0) - (r_n x^n + \dots + r_n b_m^{-1} b_0 x^{n-m})$$

so $\deg(r' - r_n b_m^{-1} x^{n-m} b) < n = \deg(r')$

Then, we can find $q''$, $r''$ such that

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

namely since $a = q'b + r'$

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

Applying this always gives a remainder, $r''$, of degree strictly smaller than the degree of the previous remainder: $\deg(r'') < \deg(r')$.

As long as $\deg(r) \geq \deg(b)$, keep applying this until $\deg(r) < \deg(b)$.
This process will terminate after a finite number of steps.

At this point we will have $q, r \in k[x]$ such that $a = qb + r$ and $\deg(r) < \deg(b)$.

Let's now show that such $q, r$ are <u>uniquely</u> determined.

Suppose that $a = q_1 b + r_1$ ① , $\deg(r_1) < \deg(b)$
$a = q_2 b + r_2$ ② , $\deg(r_2) < \deg(b)$

Subtracting ① from ② gives:

$$0 = (q_2 - q_1) b + (r_2 - r_1)$$

and hence: $b(q_1 - q_2) = (r_2 - r_1)$

Taking degrees:
$$\deg(b(q_1-q_2)) = \deg(r_2-r_1)$$

Then $\deg(b) + \deg(q_1-q_2) = \deg(r_2-r_1)$

(using $\deg(fg) = \deg(f) + \deg(g)$ for $f, g \in k[x]$)

So $\deg(b) + \deg(q_1-q_2) \leq \max\{\deg(r_2), \deg(-r_1)$

(using $\deg(f+g) \leq \max\{\deg(f), \deg(g)\}$ for $f, g \in k[x]$)

i.e. $\deg(b) + \deg(q_1-q_2) \leq \max\{\deg(r_2), \deg(r_1)$

(using $\deg(f) = \deg(-f)$ for $f \in k[x]$)

But $\deg(r_1) < \deg(b)$ and $\deg(r_2) < \deg(b)$
so $\max\{\deg(r_1), \deg(r_2)\} < \deg(b)$

Therefore: $\deg(b) + \deg(q_1-q_2) < \deg(b)$

i.e. $\deg(q_1-q_2) < 0$
so $\deg(q_1-q_2) = -\infty$ (possible degrees: $-\infty, 0, 1, 2, \ldots$)
i.e. $q_1-q_2 = 0$ (using $\deg(f) = -\infty \Leftrightarrow f = 0$)

Thus $q_1 = q_2$ and substituting back into ①, ②: $r_1 = r_2$.
So $q_1 = q_2$, $r_1 = r_2$, this shows that $q$ and $r$ are uniquely determined.

$\square$

Using this process of division, we can prove the following:

<u>Remainder Theorem:</u>

Suppose $f(x) \in k[x]$ for some field $k$, and $a \in k$. Then $f(a) = 0$ if and only if $x - a$ divides $f$.

<u>Proof:</u>

Suppose first that $x-a$ divides $f$.
Then, for some $g(x) \in k[x]$: $f(x) = (x-a)g(x)$

Then $f(a) = (a-a)g(a)$
$= 0 \cdot g(a)$ i.e. $f(a) = 0$ as required.

Suppose that $f(a) = 0$
By the previous proposition, there exists $q(x), r(x)$ such that

$$f(x) = q(x)(x-a) + r(x) \quad \text{— ①}$$

where $\deg(r(x)) < \deg(x-a)$
i.e. $\deg(r(x)) < 1$
So $r(x)$ is a constant, $r(x) = c$ for some $c \in k$.

Substituting $x = a$ into ①:
$f(a) = q(a)(a-a) + r(a)$
so $f(a) = r(a)$, so $r(a) = 0$ (since $f(a) = 0$)
So $c = 0$ and $r(x) = 0$
Then $f(x) = q(x)(x-a)$, so $(x-a)$ divides $f(x)$, as required. $\square$

MATH 2201

11-10-16

Goal: show how we can use Euclidean division on $k[x]$ to obtain complete factorizations of polynomials in $k[x]$

First extend the notion of a greatest common divisor from $\mathbb{N}$ to $\mathbb{Z}$ and $k[x]$.

Suppose $a, b \in \mathbb{N}$. Then, $c \in \mathbb{N}$ is a common divisor of $a$ and $b$ if $c$ divides $a$ and $c$ divides $b$.
We can say the same in $\mathbb{Z}$ and $k[x]$.
In fact, for any ring $R$, $c \in R$ is a common divisor of $a$ and $b$ if $c \mid a$ and $c \mid b$.

The greatest common divisor of $a, b \in \mathbb{N}$ is a natural number $d$ such that:
- $d \mid a$ and $d \mid b$
- if, for some $c \in \mathbb{N}$: $c \mid a$ and $c \mid b$, then $c \leq d$ (or equivalently, $c \mid d$).

14-10-16

Greatest common divisors in $\mathbb{N}$:
Suppose $a, b \in \mathbb{N}$ where at least one of $a$ and $b$ is non zero.
A number $d \in \mathbb{N}$ is a greatest common divisor of $a$ and $b$ if the following conditions hold:

- $d | a$ and $d | b$
- If, for some $c \in \mathbb{N}$, $c | a$ and $c | b$, then $c | d$.

(note: this is equivalent to "if $c | a$ and $c | b$, then $c \leq d$."

In $\mathbb{N}$, a greatest common divisor of $a, b$ is unique.

If we take the same definition to $\mathbb{Z}$ or $k[x]$ (for a field $k$) this is not the case.

In $\mathbb{Z}$: if $a = 6$, $b = 9$, then both 3 and $-3$ would satisfy the corresponding definition.

Divisors of 6: $\pm 1, \pm 2, \pm 3, \pm 6$
" 9: $\pm 1, \pm 3, \pm 9$

Common divisors: $\pm 1, \pm 3$, each of these divides 3 and $-3$. Note there are two units ($\pm 1$) in $\mathbb{Z}$.

Similarly, we do not have uniqueness in $k[x]$

e.g. consider $\mathbb{R}[x]$, and let $a = x^2 - 1$, $b = x^2 + 3x + 2$
Then $a = (x+1)(x-1)$, divisors are: $1, x+1, x-1, x^2-1$ and any non-zero constant multiple.
and $b = (x+1)(x+2)$, divisors are: $1, x+1, x+2, x^2+3x+2$ and any non-zero constant multiple.
Common divisors: $1, x+1$ and all constant (non zero) multiples.

In $\mathbb{Z}$, we obtain uniqueness by insisting that the greatest common divisor is non-negative.

In $k[x]$, we insist that greatest common divisors are monic (leading coefficient 1).

<u>Definition</u>

Suppose $f(x), g(x) \in k[x]$ for a field $k$, and $f(x), g(x)$ are not both zero. A monic polynomial $d(x) \in k[x]$ is a greatest common divisor of $f(x)$ and $g(x)$ if:

- $d(x) \mid f(x)$ and $d(x) \mid g(x)$
- if, for some $c(x) \in k[x]$: $c(x) \mid f(x)$ and $c(x) \mid g(x)$, then $c(x) \mid d(x)$.

Let's show that, given $f(x), g(x) \in k[x]$, there is a <u>unique</u> greatest common divisor

<u>Helpful results:</u>

- If $a(x)$ and $b(x)$ are monic polynomials and $a(x) = u\, b(x)$ for $u \in k \setminus \{0\}$ (i.e. $\deg(u) = 0$) then $a(x) = b(x)$.

  <u>Proof</u>: $\deg(a) = \deg(u) + \deg(b)$
  $\Rightarrow \deg(a) = \deg(b)$
  So, for some $n \in \mathbb{N}_0$:
  $$\begin{cases} a = x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0 \\ b = x^n + b_{n-1}x^{n-1} + \dots + b_1 x + b_0 \end{cases} \quad (a, b \text{ monic}).$$
  Since $a(x) = u\, b(x)$:
  $$x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0 = ux^n + ub_{n-1}x^{n-1} + \dots + ub_1 x + ub_0$$
  Comparing coefficients: $u = 1$, $a_i = b_i$ $\forall\, 0 \leq i \leq n-1$.
  So $a(x) = b(x)$.

- If $d_1(x), d_2(x)$ are monic polynomials in $k[x]$ and if $d_1 \mid d_2$ and $d_2 \mid d_1$, then $d_1 = d_2$

14-10-16

Proof: Since $d_1 | d_2$ there is $c \in k[x]$ such that $d_2(x) = c(x) d_1(x)$. Also as $d_2 | d_2$ there is $c' \in k[x]$ such that $d_1(x) = c'(x) d_2(x)$.
Substituting $d_2(x) = c(x) d_1(x)$ into $d_1(x) = c'(x) d_2(x)$, we obtain $d_1(x) = c(x) c'(x) d_2(x)$
Then $\deg(c') + \deg(c) + \deg(d_1) = \deg(d_1)$
Since $d_1$ is monic, $d_1 \neq 0$, $\deg(d_1) \geq 0$
So $\deg(c') + \deg(c) = 0$
$\Rightarrow \deg(c') = \deg(c) = 0$.
So $d_2(x) = c(x) d_1(x)$, but $c(x)$ is a nonzero constant, so, by the previous helpful result:
$d_1(x) = d_2(x)$.

Now suppose that for $f(x), g(x) \in k[x]$ (not both zero), $d_1(x)$ and $d_2(x)$ are greatest common divisors.
$d_2 | f$ and $d_2 | g$, so $d_2 | d_1$ (since $d_1$ is a greatest common divisor).
Suppose also that $d_1 | f$ and $d_1 | g$, so $d_1 | d_2$ (as $d_2$ is a greatest common divisor).
So $d_1, d_2$ are monic, $d_1 | d_2$ and $d_2 | d_1$, hence, by the previous result

$$d_1(x) = d_2(x).$$

Hence, for $f(x), g(x) \in k[x]$ (not both zero), there is a <u>unique</u> greatest common divisor, which we denote by $\gcd(f, g)$.

<u>Note:</u>

1). An equivalent definition of $\gcd(f,g)$ involves the following: If $c | f$ and $c | g$, then $\deg(c) \leq \deg(d)$.
2). Note that any polynomial divides $0$, so for a non-zero $f \in k[x]$ $\gcd(f, 0)$ is the monic "version" of $f$. eg. $\gcd(2x+4, 0) = x + 2$.

How may we determine greatest common divisors in practice?
Use the Euclidean algorithm:

- Apply Euclidean division repeatedly, starting from two polynomials $f(x)$ and $g(x)$, where $g \neq 0$, until we get a zero remainder:

$$\begin{aligned} f &= q_1 g + r_1 \\ g &= q_2 r_1 + r_2 \\ &\vdots \\ &\quad + r_n \\ &\quad + 0 \end{aligned}$$

The final non-zero remainder in this process leads to $\gcd(f, g)$.
$\gcd(f,g)$ is the monic "version" of that remainder.
Also, by back substitution in the algorithm, we can determine polynomials $a(x)$, $b(x)$ such that
$\gcd(f,g) = a(x) f(x) + b(x) g(x)$.

Example

Take $f(x) = x^3 - 10x + 3$, $g(x) = x^2 - 9$.

Step 1: divide $f$ by $g$:

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

Step 2: divide $x^2 - 9$ by $-x + 3$

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

So we have

$$x^3 - 10x + 3 = x(x^2 - 9) + (-x + 3) \quad -①$$
$$x^2 - 9 = (-x - 3)(-x + 3) + 0 \quad -②$$

The final non-zero remainder is $-x + 3$.
Make it monic (divide through by $-1$): $x - 3$
So $\gcd(x^3 - 10x + 3, x^2 - 9) = x - 3$
Rearranging ①: $-x + 3 = 1 \cdot (x^3 - 10x + 3) - x(x^2 - 9)$

so $$x - 3 = \underbrace{-1}_{a(x)} \cdot \underbrace{(x^3 - 10x + 3)}_{f(x)} + \underbrace{x}_{b(x)} \underbrace{(x^2 - 9)}_{g(x)}$$

14-10-16

Example

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

Applying the algorithm:

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

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

$$x + 2 = \left(\tfrac{1}{15}(x+2)\right)\cdot 15 + 0 \quad -③$$

The final non-zero remainder is 15 so $gcf(f,g) = 1$.

($f$ and $g$ are coprime. Note for any $f \in k[x]$ $gcf(f,1) = 1$ so $f$ and 1 are coprime.)

Rearranging ① and ② (to make the remainders the subjects):

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

$$15 = 1\cdot(3x^2 - x + 1) - (3x-7)(x+2) \quad -④$$

Substitute ③ into ④:

$$15 = 1\cdot(3x^2 - x + 1) - (3x-7)\left((9x^3 - 3x^2 + 4x - 2) - 3x(3x^2 - x + 1)\right)$$

$$15 = (1 + 3x(3x-7))(3x^2 - x + 1) + (-3x+7)(9x^3 - 3x^2 + 4x - 2)$$

$$\Rightarrow 1 = \left(\frac{-3}{15}x + \frac{7}{15}\right)f(x) + \left(\frac{9}{15}x^2 - \frac{21}{15}x + \frac{1}{15}\right)g(x).$$

Example

Given:
- A. $f$ has degree 3
- B. $x^2+1$ divides $f(x)$
- C. if we divide $f(x)$ by $x-1$ the remainder is 16.
- D. if " " $f(x)$ " $x+2$ " " " $-5$

Find $f$.

$x^2+1$ divides $f(x)$: $f(x) = g(x)(x^2+1)$

Also $\deg(f) = 3$, so $\deg(g) = 1$, i.e. $g = a_1 x + a_0$ $(a_1 \neq 0)$

So $f(x)$ has the form $f(x) = (a_1 x + a_0)(x^2+1)$

From © $f(x) = q(x)(x-1) + 16$ for some $q(x)$

Substituting $x = 1$: $f(1) = 16$

From before $(a_1 x + a_0)(x^2+1) = f(x)$

Set $x = 1$ : $2(a_1 + a_0) = 16$

Similarly from ① : $f(-2) = -5$

So $(-2a_1 + a_0)(5) = -5$

So $a_1 + a_0 = 8$ , $-2a_1 + a_0 = -1$

MATH 2201

17-10-16

Let's now show that the Euclidian algorithm **does** lead to the greatest common divisor, in the way described last time.

Helpful result from earlier:
If $d_1, d_2$ are monic, and if $d_1 | d_2$ and $d_2 | d_1$, then $d_1 = d_2$.

We use this to show the following key result.

**Proposition**
Suppose that $f(x), g(x), q(x), r(x) \in k[x]$, for a field $k$, and suppose $g \neq 0$. If $f(x) = q(x)g(x) + r(x)$, then $\gcd(f,g) = \gcd(g,r)$

**Proof:**
Suppose that $d_1 = \gcd(f,g)$ and $d_2 = \gcd(g,r)$
Then $\underbrace{d_1 | f \text{ and } d_1 | g}$ , $\underbrace{d_2 | g \text{ and } d_2 | r}$

| | |
|---|---|
| Since: $d_1 \mid g$ , $d_1 \mid qg$ | Since: $d_2 \mid g$ , $d_2 \mid qg$ |
| Also: $d_1 \mid f$ | Also: $d_2 \mid r$ |
| So: $d_1 \mid f - qg$ | So: $d_2 \mid qg + r$ |
| i.e. $d_1 \mid r$ (and $d_1 \mid g$) | i.e. $d_2 \mid f$ (and $d_2 \mid g$) |
| But $d_2 = \gcd(g,r)$ | But $d_1 = \gcd(f,g)$ |

So, by definition of gcd: $d_1 | d_2$ and $d_2 | d_1$. Furthermore, $d_1$ and $d_2$ are monic, so using an earlier result: $d_1 = d_2$ as required. $\square$

Then viewing this in the context of the Euclidean algorithm, leads to the greatest common divisor.

Applying the algorithm to $f(x), g(x)$ $(g(x) \neq 0)$

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

Then
$$
\begin{aligned}
\gcd(f,g) &= \gcd(g, r_1) \\
\gcd(g, r_1) &= \gcd(r_1, r_2) \\
&\vdots \\
\gcd(r_{n-2}, r_{n-1}) &= \gcd(r_{n-1}, r_n) \\
\gcd(r_{n-1}, r_n) &= \gcd(r_n, 0)
\end{aligned}
$$

Overall: $\gcd(f,g) = \gcd(r_n, 0)$

Also: for any $f \in k[x]$, $\gcd(f, 0)$ is the monic "version" of $f$, eg. $\gcd(2x+3, 0) = x + \frac{3}{2}$

So $\gcd(f,g)$ is, as claimed before, $\gcd(r_n, 0)$ i.e. it is the monic version of $r_n$, the final non-zero remainder.

Note:

If the first remainder is already 0, i.e. if $f = qg (+0)$ then $\gcd(f,g)$ is the monic 'version' of $g$. eg. if $f(x) = x^2 - 1$, $g(x) = x+1$ then $x^2 - 1 = (x-1)(x+1)$, $\gcd(f,g) = x+1$.

Consider irreducible elements in $k[x]$ [in $\mathbb{R}[x]$ and $\mathbb{C}[x]$ really]

Reminder

In a ring $R$:

- $u$ is a unit if $\exists u' \in R$ such that $u \cdot u' = 1$ and $u' \cdot u = 1$
- $a$ is irreducible if it is not a unit and if, whenever $a = bc$, then $b$ or $c$ is a unit.
- $a$ is reducible if it is neither a unit nor irreducible, ie. if it is possible to find non-units $b, c$ such that $a = bc$.

Examples

In $\mathbb{Z}$: units are $\pm 1$

irreducible elements are $\pm 2, \pm 3, \pm 5, \pm 7, \dots$

reducible elements are $\pm 4, \pm 6, \pm 8, \pm 9, \dots$

What about irreducibles / units in $k[x]$?

From before:
- $f(x) \in k[x]$ unit iff $\deg(f) = 0$

  " " " iff $f$ is a non-zero constant.
- If $f(x) \in k[x]$ and $\deg(f) = 1$, then $f(x)$ is irreducible.

  But it is not always true that every irreducible element has degree 1.

  eg. in $\mathbb{R}[x]$: $x^2+1$ is irreducible.

The next result shows we can "test" degree 2 polynomials for irreducibility:

Proposition:
Suppose that $f(x) \in k[x]$ and $\deg(f) = 2$.
Then $f$ is reducible if and only if $\exists a \in k$ s.t. $f(a) = 0$. [i.e. there exists a root of $f$ in $k$]

Equivalently: $f$ is irreducible iff $f$ has no root in $k$.

[ Recall: Remainder theorem:
For $f(x) \in k[x]$, $x-a$ divides $f$ if and only if $f(a) = 0$. ]

Proof of proposition:
Suppose that $f = gh$.
Considering degrees: $2 = \deg f = \deg(g) + \deg(h)$.
So, we have the following possibilities:

$$\begin{cases} \deg(g) = 0, & \deg(h) = 2 \\ \deg(g) = 2, & \deg(h) = 0 \\ \deg(g) = 1, & \deg(h) = 1 \end{cases}$$

So, either $\deg(g) = 0$ or $\deg(h) = 0$,
or $\deg(g) = 1$ or $\deg(h) = 1$

i.e. either $g$ or $h$ is a unit,
or neither is a unit.

Equivalently: either $f$ is irreducible, or
$f$ is reducible.

$f$ reducible $\Rightarrow f(x) = g(x)h(x)$ where
$\deg(g) = \deg(h) = 1$
i.e. $f(x) = (a_1 x + a_0) h(x)$ for $a_0, a_1 \in k$,
$a_1 \neq 0$.

MATH 2201

17-10-16

ie $f\left(\frac{-a_0}{a_1}\right) = 0$

examples:

$$x^2 + 3x + 2 = (x+1)(x+2)$$
$$x^2 + 1 = 2\cdot\left(\tfrac{1}{2}(x^2+1)\right)$$

18-10-16

From last time

Proposition:
Suppose $f(x) \in k[x]$, for a field, where $\deg(f) = 2$. Then $f(x)$ is reducible iff $\exists a \in k$ st. $f(a) = 0$ (ie. $f$ has a root in the field $k$).

Proof:
Suppose, first, that $f(a) = 0$ for some $a \in k$. Then, using the Remainder Theorem, $x - a$ divides $f(x)$, so that, for some $g(x) \in k[x]$:
$$f(x) = (x-a)g(x).$$
Considering degrees:
$$\underbrace{\deg(f)}_{2} = \underbrace{1}_{\deg(x-a)} + \deg(g) \quad \text{so } \deg(g) = 1$$
So, this is a factorisation of $f(x)$ that does not involve units (each unit in $k[x]$ has $\deg = 0$)
So $f(x)$ is reducible, as required.

Now, suppose that $f(x)$ is reducible, ie. that for some $g(x), h(x) \in k[x]$,
$f(x) = g(x)h(x)$, where neither of $g(x), h(x)$ is a unit (ie. $\deg(g) \neq 0$, $\deg(h) \neq 0$).
By considering degrees:

$$\deg(f) = \deg(g) + \deg(h)$$
i.e. $\deg(g) + \deg(h) = 2$.
Since $\deg(g) \neq 0$, $\deg(h) \neq 0$, the only possibility is $\deg(g) = 1$ and $\deg(h) = 1$.
i.e. $g(x) = a_1 x + a_0$ for $a_0, a_1 \in k$, $a_1 \neq 0$
Then $f(x) = (a_1 x + a_0) h(x)$.

Now, set $a = \dfrac{-a_0}{a_1} \in k$:

$$f(a) = \left(a_1\left(\frac{-a_0}{a_1}\right) + a_0\right) h(x) = 0$$

So there exists $a \in k$ s.t. $f(a) = 0$, as required $\square$

Similarly we can show:

<u>Proposition</u>:
Suppose $f(x) \in k[x]$, for a field $k$, where $\deg(f) = 3$.
Then: $f(x)$ is reducible in $k[x]$ iff $\exists a \in k$ s.t. $f(a) = 0$

<u>Examples</u>
- The polynomial $x^2 - 2$ is irreducible in $\mathbb{Q}[x]$, since there is no $q \in \mathbb{Q}$ such that $q^2 - 2 = 0$, while $x^2 - 2$ <u>is</u> reducible in $\mathbb{R}[x]$ and $\mathbb{C}[x]$
  $x^2 - 2 = (x + \sqrt{2})(x - \sqrt{2})$.

- The polynomial $x^2 + 2$ is irreducible in $\mathbb{Q}[x]$ and $\mathbb{R}[x]$ since there is no real (or rational) number $r$ such that $r^2 + 2 = 0$
  i.e. such that $r^2 = -2$.
  However, $x^2 + 2$ is reducible in $\mathbb{C}[x]$:

MATH
2201

18-10-16

$$x^2 + 2 = (x + i\sqrt{2})(x - i\sqrt{2}) \quad \text{where } i^2 = 1$$

Note:

1). In $\mathbb{C}[x]$: if $\deg(f) \geq 2$ then $f(x)$ is reducible.
This follows from the Fundamental Theorem of Algebra.
Any polynomial in $\mathbb{C}[x]$ can be factorised into linear factors (of degree 1), ie. can be written as $c(x-a_1)(x-a_2)\dots(x-a_n)$ for $a_1, \dots, a_n \in \mathbb{C}$, $c \in \mathbb{C}$, $c \neq 0$.

2). In $\mathbb{R}[x]$, there exist reducible and irreducible polynomials of degree 2:
e.g. $x^2+2$ is irreducible, but $x^2-2$ is reducible.
In $\mathbb{Q}[x]$, we also have both reducible and irreducible polynomials of degree 2:
e.g. $x^2-2$ is irreducible, but $x^2-4$ is reducible.

3). It turns out that in $\mathbb{R}[x]$ every polynomial of degree greater than or equal to 3 is reducible.
In $\mathbb{Q}[x]$, things get more interesting.

We now try to show that every polynomial in $k[x]$ can be factorised in a unique way, in terms of irreducible elements.
(Just like every natural number can be factorised uniquely in terms of primes).

Consider a monic irreducible element $p(x)$ in $k[x]$.
Since $p(x)$ is irreducible, if $p(x) = f(x)g(x)$ then $f(x)$ or $g(x)$ is a unit, a non-zero constant.

So, any factorisation of $p(x)$ has the form $p(x) = u\left(\frac{1}{u}p(x)\right)$ where $u \in k$, $u \neq 0$.

So the only monic devisors of $p(x)$ are: $1$ and $p(x)$.

Hence, for any $f(x)$ in $k[x]$:
either $\underbrace{\gcd(f,p) = p(x)}_{p(x) \mid f(x)}$ or $\underbrace{\gcd(f,p) = 1}_{f(x),\, p(x) \text{ coprime}}$

Recall:
By simply using "back substitution" in the Euclidean algorithm, we obtain $a(x), b(x) \in k[x]$ s.t.
$$\gcd(f,g) = a(x)f(x) + b(x)g(x)$$
for every $f(x), g(x) \in k[x]$, $g(x) \neq 0$
This is known as Bezout's Lemma

MATH 2201

18-10-16

Key step to unique factorisation:

Proposition:
Suppose $p(x)$ is a monic irreducible polynomial in $k[x]$.
For any $f(x), g(x) \in k[x]$:
If $p(x) \mid f(x)g(x)$, then $p(x) \mid f(x)$ or $p(x) \mid g(x)$.

Proof:
We assume that $p(x) \mid f(x)g(x)$.
If $p(x) \mid f(x)$, then we are done.
Suppose now that $p(x) \nmid f(x)$. Let's try to show that $p(x) \mid g(x)$.
Since $p(x)$ does not divide $f(x)$: $\gcd(f, p) = 1$
So, using Bezout's Lemma, there exist $a(x), b(x) \in k[x]$ s.t. $a(x)f(x) + b(x)p(x) = 1$.

Multiplying through by $g(x)$:
$$a(x)f(x)g(x) + b(x)p(x)g(x) = g(x)$$

Note: · since $p(x) \mid f(x)g(x)$ : $p(x) \mid a(x)f(x)g(x)$
· since $p(x) \mid p(x)$ : $p(x) \mid b(x)p(x)g(x)$
$\Rightarrow$ So $p(x) \mid \underbrace{a(x)f(x)g(x) + b(x)p(x)g(x)}_{g(x)}$

Hence $p(x)$ divides $g(x)$ as required. $\square$

MATH 2201

21-10-16

From last time:

Proposition

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

For every $f(x), g(x)$ in $k[x]$: if $p(x) \mid f(x)g(x)$, then $p(x) \mid f(x)$ or $p(x) \mid g(x)$.

This can be extended to the following:

Proposition:

Suppose $p(x)$ is a monic irreducible polynomial in $k[x]$.

If, for $f_1(x), \ldots, f_n(x) \in k[x]$, $p(x) \mid f_1(x) \ldots f_n(x)$, then $p(x) \mid f_i(x)$ for some $i$ $(1 \leq i \leq n)$.

($p(x)$ divides at least one of the $f_i(x)$).

Proof By induction on $n$.

For the case $n=1$, the statement becomes:

If $p(x) \mid f_1(x)$, then $p(x) \mid f_1(x)$ (nothing to prove)

Now, assume that the result holds up to and including $n$:

i.e. if $p(x) \mid g_1(x) \ldots g_m(x)$, then $p \mid g_i(x)$ for $1 \leq i \leq m$, whenever $m \leq n$.

Next, suppose that $p(x) \mid (f_1(x) \ldots f_n(x)) f_{n+1}(x)$, so, by the previous proposition:

$$p(x) \mid f_1(x) \ldots f_n(x) \quad \text{or} \quad p(x) \mid f_{n+1}(x).$$

So, using the inductive assumption:

$$p(x) \mid f_i(x) \text{ for some } 1 \leq i \leq n \text{ or } p(x) \mid f_{n+1}(x).$$

as $p(x) \mid f_1(x) \dots f_n(x)$.

Overall, $p(x)$ divides $f_i(x)$ for some $i$, $1 \le i \le n+1$.

$\square$

We can use this to prove a theorem about unique factorisation in $k[x]$.

Two other results we will use are:

1). If $p(x), q(x)$ are monic irreducible and $p(x) \mid q(x)$, then $p(x) = q(x)$.

To see this, note, from earlier, that if $p(x)$ is a monic irreducible polynomial in $k[x]$, then, for any $f(x) \in k[x]$:

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

($f, p$ coprime) $\quad$ ($p(x) \mid f(x)$)

So since $p(x)$ is a monic irreducible polynomial:

$\gcd(p, q) = p(x)$ (we cannot have $\gcd(p, q) = 1$, since $p \mid q$)

Similarly, since $q(x)$ is monic and irreducible:

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

$\therefore p(x) = q(x)$. $\square$

2). The following holds for $a(x), b(x), c(x) \in k[x]$:

if $a(x) \neq 0$ and $a(x)b(x) = a(x)c(x)$, then $b(x) = c(x)$

If $a(x)b(x) = a(x)c(x)$, then $a(x)(b(x) - c(x)) = 0$.

So, $a(x) = 0$ or $b(x) - c(x) = 0$.

Since $a(x) \neq 0$, $b(x) = c(x)$ by assumption.

Theorem

Suppose $f(x)$ is a monic polynomial in $k[x]$ and $\deg(f) \geq 1$ (so $f$ is not a constant).
Then, there exist monic irreducible polynomials $p_1(x), \dots, p_r(x)$ such that

$$f(x) = p_1(x) \dots p_r(x).$$

Furthermore, the factorisation is unique (up to reordering):
If we also have

$$f(x) = q_1(x) \dots q_s(x) \text{ for } q_1(x), \dots, q_s(x) \in k[x]$$

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

Proof:

We prove this by induction on the degree of $f(x)$.
Let's first show that a suitable factorisation of $f(x)$ exists (by induction).
If $\deg(f) = 1$, then $f$ is irreducible, so the result holds (chose $r=1$ and $p_1(x) = f(x)$ : $f(x) = f(x)$ is a factorisation).

Now, suppose that the result holds for degree smaller than or equal to $n$, i.e. that, for each $g(x) \in k[x]$ such that $\deg(g) \leq n$, there exists a factorisation into irreducible elements.

Consider $f(x) \in k[x]$ such that $\deg(f)$ is either reducible or irreducible ($\deg(f) \geq 1$, so $f$ cannot be a unit)
If $f(x)$ is irreducible, then the result we are trying to prove holds (just as in the $n=1$ case)

If $f$ is reducible, then, for some $g(x), h(x) \in k[x]$ such that $f(x) = g(x) h(x)$ and where neither of $g(x)$, $h(x)$ is a unit (so $\deg(g) \geq 1$ and $\deg(h) \geq 1$).

Then $\underbrace{\deg(f)}_{n+1} = \deg(g) + \deg(h)$.

It follows that $\deg(g) \leq n$ and $\deg(h) \leq n$.

Hence, we can use our inductive assumption; we can factorise $g(x)$ and $h(x)$:
$g(x) = p_1(x) \dots p_t(x)$ for monic irreducible $p_1(x), \dots, p_t(x) \in k[x]$
$h(x) = q_1(x) \dots q_t(x)$ for " " $q_1(x), \dots, p_t(x) \in k[x]$

So, there (also) exists a factorisation of $f(x)$ into monic, irreducible polynomials: $(f(x) = g(x)h(x))$

$$f(x) = p_1(x) \dots p_t(x)\, q_1(x) \dots q_t(x).$$

Let's now show the uniqueness of a factorisation for $f(x)$, up to reordering (using induction).

- If $\deg(f) = 1$ then $f$ is irreducible, so one factorisation is $f(x) = f(x)$ (i.e. choose $r = 1$ and $p_1(x) = f(x)$).

  Suppose $f(x) = q_1(x) \dots q_s(x)$ ($q_i$: monic, irreducible)
  Then $\deg(f) = \deg(q_1) + \dots + \deg(q_s)$
  So $\deg(q_1) + \dots + \deg(q_s) = 1$
  Also $\deg(q_i) \geq 1$ for each $1 \leq i \leq s$ ($q_i$ irreducible)
  It follows that $s = 1$: $f(x) = q_1(x)$.
  So $r = s = 1$, and $p_1(x) = q_1(x)$ $(= f(x))$ as required.

- Now, suppose that the result holds for all polynomials of degree smaller that or equal to $n$: If $\deg(g) \leq n$ then the factorisation of $g(x)$ into irreducible elements is unique as described in the statement of the theorem.

  Consider $f(x) \in k[x]$ with $\deg(f) = n+1$,

MATH
2201

21-10-16

and suppose that
$f(x) = p_1(x) \dots p_r(x)$ for monic irreducible $p_1(x), \dots, p_r(x)$ in $k[x]$
and $f(x) = q_1(x) \dots q_s(x)$ " " " $q_1(x), \dots, q_s(x)$ " ".

Then, since $p_1(x) \mid f(x)$ : $p_1 \mid q_1(x) \dots q_s(x)$.
So, using an earlier result: $p_1 \mid q_j(x)$ for som $1 \le j \le s$.
Thus, using one of the earlier results: $p_1(x) = q_j(x)$
(both monic irreducible)
Therefore, we can "cancel out" $p_1(x), q_j(x)$ from:
$\cancel{p_1(x)} p_2(x) \dots p_r(x) = q_1(x) \dots q_{j-1}(x) \cancel{q_j(x)} q_{j+1}(x) \dots q_s(x)$.
So $p_2(x) \dots p_r(x) = q_1(x) \dots q_{j-1}(x) q_{j+1}(x) \dots q_s(x)$.
Since $\deg(p_1) = \deg(q_j) \geq 1$,
$\deg(p_2(x) \dots p_r(x)) < \deg(f) = n+1$. (*)

Hence: $\deg(p_2(x) \dots p_r(x)) \le n$, so we can use our
inductive assumption:

$$\underbrace{r-1}_{\text{no. of terms on left}} = \underbrace{s-1}_{\text{no. of terms on right}} \quad \text{of } *$$

and, for each $2 \le i \le r$ : $p_i(x) = q_t$ for some $t \neq j$
Also $p_1(x) = q_j(x)$.
So, overall, we obtain: $r = s$
and $p_i(x) = q_t(x)$ for some $1 \le t \le s$, for each $1 \le i \le r$.
Thus, the factorisation of $f(x)$ into irreducible
elements, is unique.

$\square$

## Chapter 2 - Linear maps and the Jordan normal form.

We start with the spaces on which linear maps are defined.

Definition:

A vector space consists of a set, $V$, together with an operation of addition denoted by $+$, and, for some field $k$, a scalar multiplication, denoted by $\cdot$, such that the following conditions are satisfied:

1). If $a, b \in V$, then $a+b \in V$

2). For all $a, b, c \in V$: $a+(b+c) = (a+b)+c$

3). There exists an identity element, $0$, in $V$ such that, for each $a \in V$: $a+0 = a$ and $0+a = a$.

4). For each $a \in V$, there exists an additive inverse, $-a$, in $V$, such that: $a+(-a) = 0$ and $(-a)+a = 0$.

5). For all $a, b \in V$: $a+b = b+a$

6). For each $\lambda \in k$, $a \in V$: $\lambda \cdot a \in V$

7). If $1$ is the multiplicative identity in $k$, then $1 \cdot v = v$ for each $v \in V$

8). For each $\lambda_1, \lambda_2$ in $k$, and $v$ in $V$: $\lambda_1 \cdot (\lambda_2 \cdot v) = (\lambda_1 \cdot \lambda_2) \cdot v$

9). For each $\lambda \in k$ and $a, b \in V$: $\lambda \cdot (a+b) = \lambda \cdot a + \lambda \cdot b$

10). For each $\lambda_1, \lambda_2$ in $k$, and $a \in V$: $(\lambda_1 + \lambda_2) \cdot a = (\lambda_1 \cdot a) + (\lambda_2 \cdot a)$

Notes:

1). The elements of $V$ are known as vectors,
" " " $k$ " " " scalars

2). We often "drop" the symbol for scalar multiplication: we write $\lambda v$ instead of $\lambda \cdot v$.

3). Many other rules that hold in vector spaces follow from the above, e.g. $-1 \cdot v = -v$, $0 \cdot v = 0$, $\underset{\in k}{0} \cdot \underset{\in V}{0} = 0$

MATH 2201

24-10-16

Previously on MATH 2201....
definition of vector space $V$ over a field $k$.
Note: In this module, we will not in general, be underlining vectors.

Examples of vector spaces:

1). For a field $k$, consider $k^n = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} : a_1, \dots, a_n \in k \right\}$

If we define addition and scalar multiplication as follows:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}, \quad \lambda \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix} \text{ for } \lambda \in k$$

then $k^n$ is a vector space over the field $k$.
e.g. $\mathbb{R}^3$ is a vector space over $\mathbb{R}$.

2). For a field $k$, let $V = \{ (a_1, \dots, a_n) : a_1, \dots, a_n \in k \}$ and define addition and scalar multiplication as follows:

$(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, a_n + b_n)$,
$\lambda(a_1, \dots, a_n) = (\lambda a_1, \dots, \lambda a_n)$ for $\lambda \in k$.
Then $V$ is a vector space over $k$.

Each element of $V$ may be thought of as a function from $k^n$ to $k$.
e.g. $(1, 2, 3)$ corresponds to a function from $\mathbb{R}^3$ to $\mathbb{R}$ defined as follows:

$$(1, 2, 3) \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{\in \mathbb{R}^3} = x_1 + 2x_2 + 3x_3$$

3). For a field $k$, consider the set of all $2\times 2$ matrices with entries in $k$:

$$M_2(k) = \left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} : a_{11}, a_{12}, a_{21}, a_{22} \in k \right\}$$

This forms a vector space over $k$, using the following rules:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{pmatrix},$$

$$\lambda \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{pmatrix} \quad \text{for } \lambda \in k$$

As seen earlier, $M_2(k)$ also forms a ring (with multiplication defined as "usual" matrix multiplication) but here we "forget" that operation (not required to use it in order to satisfy the definition of a vector space).

In general $M_n(k)$ forms a vector space over $k$. (All $n\times m$ matrices with entries form a field over $k$, in fact).

4). For any field $k$, the ring of polynomials, $k[x]$, forms a vector space over $k$, using the rules:

$$\sum_i a_i x^i + \sum_i b_i x^i = \sum_i (a_i + b_i) x^i$$

$$\lambda(a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots) = \lambda a_0 + \lambda a_1 x + \dots + \lambda a_n x^n + \dots$$

Again this is in fact a ring, but we "forget" about multiplication of polynomials here.

MATH 2201

24-10-16

5). For any field $k$, consider $k[x]_n$, the set of all polynomials of degree up to and including $n$:

$$k[x]_n = \{a_0 + a_1 x + \dots + a_n x^n : a_0, \dots, a_n \in k\}$$

This forms a vector space over $k$, using rules as in example ④.

$$(a_0 + a_1 x + \dots + a_n x^n) + (b_0 + b_1 x + \dots + b_n x^n) = (a_0 + b_0)x + \dots + (a_n + b_n)x^n$$

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

Note: $k[x]_n$ is not a ring, e.g. in $\mathbb{R}[x]_2$ we have $1 + x^2,\ 2 + x \in \mathbb{R}[x]_2$, but

$$(1+x^2)(2+x) = 2 + x + 2x^2 + x^3 \notin \mathbb{R}[x]_2 \quad (\text{degree} > 2)$$

So if $a(x), b(x) \in \mathbb{R}[x]_2$, it does not follow necessarily that $a(x)b(x) \in \mathbb{R}[x]_2$

Subspaces are vector spaces within vector spaces.

Definition

Consider a subset $U$ of a vectorspace $V$ (over a field $k$). Then $U$ is a subspace of $V$ if the following are satisfied:

1). If $a, b \in U$ then $a + b \in U$

2). If $a \in U$ and $\lambda \in k$, then $\lambda a \in U$

3). The identity element of $V$, $0$, is in $U$: $0 \in U$.

Examples:

The subset of the vector space $\mathbb{R}^3$, defined as follows: $U = \left\{ \begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix} : x_1 \in \mathbb{R} \right\}$, is a subspace of $\mathbb{R}^3$.

e.g. to check condition ①; let $a, b \in U$, say $a = \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ 0 \\ 0 \end{pmatrix}$ for $a_1, b_1 \in \mathbb{R}$.

Then $a + b = \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} b_1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ 0 \\ 0 \end{pmatrix}$

So $a+b \in U$ ($x_1 = a_1 + b_1$ in def$^n$ of $U$).
For condition ②, consider $\begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix} \in U$ and $\lambda \in k$

Then $\lambda \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ 0 \\ 0 \end{pmatrix} \in U$

For condition ③, note that the identity in $\mathbb{R}^3$, for addition, is $0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, and this is in $U$ ($x_1 = 0$).

On the other hand

$$W = \left\{ \begin{pmatrix} 1 \\ x_2 \\ x_3 \end{pmatrix} : x_2, x_3 \in \mathbb{R} \right\}$$

is a subset of $\mathbb{R}^3$, but <u>not</u> a subspace of $\mathbb{R}^3$, as all conditions fail (only one failure needed).

MATH 2201

25-10-16

From last time:
Examples of vector spaces, subspaces.

Today:
Review of (other) notions related to vector spaces.

Suppose we are given vectors $v_1, \dots, v_n$ in a vector space $V$ over a field $k$.
A linear combination of $v_1, \dots, v_n$ over $k$ is an element of $V$ of the form $\lambda_1 v_1 + \dots + \lambda_n v_n$ where $\lambda_1, \dots, \lambda \in k$.
The span of $v_1, \dots, v_n$ over $k$ is the set of all linear combinations of $v_1, \dots, v_n$ over $k$; we denote this set by/as $\mathrm{span}_k\{v_1, \dots, v_n\}$

$$\mathrm{span}_k\{v_1, \dots, v_n\} = \{\lambda_1 v_1 + \dots + \lambda_n v_n : \lambda_1, \dots, \lambda_n \in k\}$$

We say that $v_1, \dots, v_n$ span a vector space $W$ (over $k$) if $v_1, \dots, v_n$ are vectors in $W$, and every vector in $W$ can be written as a linear combination of $v_1, \dots, v_n$ over $k$
ie. $\mathrm{span}_k\{v_1, \dots, v_n\} = W$.

Examples
1). Consider $\mathbb{R}^2 = \left\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1, x_2 \in \mathbb{R}\right\}$
Then $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ span $\mathbb{R}^2$ over $\mathbb{R}$.

For each $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $x_1, x_2 \in \mathbb{R}$.

Also $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ span $\mathbb{R}^2$: every vector in $\mathbb{R}^2$ is a linear combination of these vectors over $\mathbb{R}$.

2). Consider $k[x]_2$ over a field $k$.

$$k[x]_2 = \{a_0 + a_1 x + a_2 x^2 : a_0, a_1, a_2 \in k\}$$

Each element of $k[x]_2$ is a linear combination of $1, x, x^2$ ore $k$.

$$a_0 + a_1 x + a_2 x^2 = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2$$

So $1, x, x^2$ span $k[x]_2$ over $k$, or $\{1, x, x^2\}$ spans $k[x]_2$ over $k$.

The vectors $v_1, \dots, v_n$ are <u>linearly independent</u> over $k$ if the equation $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ has as its only solution $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$ $(\lambda_1, \dots, \lambda_n \in k)$.

Otherwise, if there exists a non-zero solution to $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ $(\lambda_1, \dots, \lambda_n \in k)$ we say that $v_1, \dots, v_n$ are <u>linearly dependent</u> over $k$.

<u>Examples:</u>

In $\mathbb{R}^2$, the vectors $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$ are linearly independent since $\lambda_1 \begin{pmatrix}1\\0\end{pmatrix} + \lambda_2 \begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \Rightarrow \lambda_1 = 0, \lambda_2 = 0$.

$$\begin{pmatrix}\lambda_1\\\lambda_2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \quad (\text{for } \lambda_1, \lambda_2 \in \mathbb{R})$$

<u>Note:</u>

Any non-zero vector in $\mathbb{R}^2$, on its own, is linearly independent.

e.g. $\lambda \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \Rightarrow \lambda = 0$

So $\{\begin{pmatrix}1\\0\end{pmatrix}$ is linearly independent$\}$

The zero vector is linearly dependent, there are many nonzero solutions to $\lambda \begin{pmatrix}0\\0\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$

MATH 2201

25-10-16

(could choose any $\lambda \in \mathbb{R}$).
So $\{\binom{0}{0}\}$ is linearly dependent.

In general, for a vector space $V$ over a field $k$ if $v \in V$ and $v \neq 0$, then
- $\{v\}$ is linearly independent, while
- $\{0\}$ is linearly dependent.

(In fact, any set containing the zero vector is linearly dependent).

Returing to $\mathbb{R}^2$, the vectors $\binom{1}{0}, \binom{0}{1}, \binom{1}{1}$ are linearly dependent since the equation
$\lambda_1 \binom{1}{0} + \lambda_2 \binom{0}{1} + \lambda_3 \binom{1}{1} = \binom{0}{0}$ has non-zero solutions
e.g. $\lambda_1 = 1$, $\lambda_2 = 1$, $\lambda_3 = -1$.

$\therefore$ In $k[x]_2$ the set $\{1, x, x^2\}$ is linearly independent
$\lambda_0 \cdot 1 + \lambda_1 \cdot x + \lambda_2 \cdot x^2 = 0 \Rightarrow \lambda_0 = 0, \lambda_1 = 0, \lambda_2 = 0.$
As is the set $\{1, x\}$: $\lambda_0 \cdot 1 + \lambda_1 \cdot x = 0 \Rightarrow \lambda_0 = 0, \lambda_1 = 0.$

"Every subset of a linearly independent set is linearly independent."

These notions lead to the idea of a basis:

The vectors $v_1, \dots, v_n$, in a vector space $V$ over a field $k$, form a <u>basis</u> for $V$ if:
- $v_1, \dots, v_n$ <u>span</u> $V$ over $k$: $V = \text{span}_k \{v_1, \dots, v_n\}$
- $v_1, \dots, v_n$ are <u>linearly independent</u> over $k$.

We also say that the set $\{v_1, \dots, v_n\}$ is a basis for $V$ over $k$.

## Examples

1). The set $\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$ is a basis for $\mathbb{R}^2$ over $\mathbb{R}$ as is the set $\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right\}$,

whereas $\underbrace{\left\{\begin{pmatrix}1\\0\end{pmatrix}\right\}}_{\text{does not span}}$ and $\underbrace{\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right\}}_{\text{not L.I.}}$ are not.

2). For the vector space $\{(x_1, x_2) : x_1, x_2 \in \mathbb{R}\}$ over the field $\mathbb{R}$, $\{(1,0), (0,1)\}$ is a basis.

$$\left[\begin{array}{l} \lambda_1(1,0) + \lambda_2(0,1) = 0 \Rightarrow \lambda_1 = \lambda_2 = 0 \\ \forall x_1, x_2 \in \mathbb{R} : (x_1, x_2) = x_1(1,0) + x_2(0,1) \end{array}\right]$$

3). Consider $M_2(\mathbb{C}) = \left\{\begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22}\end{pmatrix} : z_{11}, z_{12}, z_{21}, z_{22} \in \mathbb{C}\right\}$

Basis of $M_2(\mathbb{C})$ over $\mathbb{C}$: $\left\{\overset{\epsilon(1,1)}{\begin{pmatrix}1&0\\0&0\end{pmatrix}}, \overset{\epsilon(1,2)}{\begin{pmatrix}0&1\\0&0\end{pmatrix}}, \overset{\epsilon(2,1)}{\begin{pmatrix}0&0\\1&0\end{pmatrix}}, \overset{\epsilon(2,2)}{\begin{pmatrix}0&0\\0&1\end{pmatrix}}\right\}$

$$\forall z_{11}, z_{12}, z_{21}, z_{22} \in \mathbb{C} : \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22}\end{pmatrix} = z_{11}\begin{pmatrix}1&0\\0&0\end{pmatrix} + z_{12}\begin{pmatrix}0&1\\0&0\end{pmatrix} + z_{21}\begin{pmatrix}0&0\\1&0\end{pmatrix} + z_{22}\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

So $\operatorname{span}_{\mathbb{C}}\{\epsilon(1,1), \epsilon(1,2), \epsilon(2,1), \epsilon(2,2)\} = M_2(\mathbb{C})$

Also, these vectors are linearly independent:

$$\lambda_1\begin{pmatrix}1&0\\0&0\end{pmatrix} + \lambda_2\begin{pmatrix}0&1\\0&0\end{pmatrix} + \lambda_3\begin{pmatrix}0&0\\1&0\end{pmatrix} + \lambda_4\begin{pmatrix}0&0\\0&1\end{pmatrix} = \begin{pmatrix}0&0\\0&0\end{pmatrix}$$

$$\Rightarrow \begin{pmatrix}\lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4\end{pmatrix} = \begin{pmatrix}0&0\\0&0\end{pmatrix} \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$$

the same argument shows that we have linear independence over $\mathbb{R}$.

MATH 2201

28-10-16

From last time:

Examples of bases of vector spaces.

For $M_2(\mathbb{C}) = \left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} : a_{11}, a_{12}, a_{21}, a_{22} \in \mathbb{C} \right\}$

The following is a basis over $\mathbb{C}$:

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

Bases depend on the fields that we choose to work over:

e.g. $\mathbb{C}$ has the following basis over $\mathbb{C}$: $\{1\}$

Each $\underset{x, y \in \mathbb{R}}{z = x + iy}$ can be written as $x + iy = \underbrace{(x+iy)}_{\text{a const. in } \mathbb{C}} \cdot 1$

Over $\mathbb{R}$, however, we require two (real) numbers to express $x+iy$: $x + iy = \underbrace{x}_{\text{in } \mathbb{R}} \cdot 1 + \underbrace{y}_{\text{in } \mathbb{R}} \cdot i$

So a basis of $\mathbb{C}$ over $\mathbb{R}$ is $\{1, i\}$.

Returning to $M_2(\mathbb{C})$, a general element has the form $\begin{pmatrix} x_1 + iy_1 & x_2 + iy_2 \\ x_3 + iy_3 & x_4 + iy_4 \end{pmatrix}$

$$= \underbrace{(x_1 + iy_1)}_{\in \mathbb{C}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \underbrace{(x_2 + iy_2)}_{\in \mathbb{C}} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \underbrace{x_3 + iy_3}_{\in \mathbb{C}} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \underbrace{(x_4 + iy_4)}_{\in \mathbb{C}} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

In terms of real constants, we need eight matrices over $\mathbb{R}$:

$$\begin{pmatrix} x_1 + iy_1 & x_2 + iy_2 \\ x_3 + iy_3 & x_4 + iy_4 \end{pmatrix} = x_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + y_1 \begin{pmatrix} i & 0 \\ 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + iy_2 \begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix}$$

$$+ x_3 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + iy_3 \begin{pmatrix} 0 & 0 \\ i & 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + iy_4 \begin{pmatrix} 0 & 0 \\ 0 & i \end{pmatrix}$$

The matrices given here actually form a basis of $M_2(\mathbb{C})$ over $\mathbb{R}$:

$$\left\{\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}i&0\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&i\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&0\\i&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix},\begin{pmatrix}0&0\\0&i\end{pmatrix}\right\}$$

Examples

4). Consider the vector space $k[x]$ over a field $k$:

$$k[x] = \{a_0 + a_1x + a_2x^2 + \dots + a_nx^n + \dots \quad a_i \in k, \text{ only finitely many } a_i \text{ are non-zero}\}$$

Basis of $k[x]$ over $k$: $\{1, x, x^2, \dots, x^n, \dots\}$ ← infinite basis

5). Consider $k[x]_2 = \{a_0 + a_1x + a_2x^2 : a_0, a_1, a_2 \in k\}$

Basis of $k[x]_2$ over $k$: $\{1, x, x^2\}$

Some results concerning bases:

- Every vector space $V$ over a field $k$ has a basis over $k$. ($\exists$ basis for $V$)
  Also, any two bases of $V$ over $k$ contain the same number of elements. (Basis Theorem)
  The number of elements in a basis of $V$ over $k$ is called the dimension of $V$ over $k$, denoted by $\dim_k(V)$ (or $\dim(V)$ if $k$ is obvious).
  e.g. $\dim_{\mathbb{R}}(\mathbb{R}^3) = 3$, $\dim_{\mathbb{C}}(M_2(\mathbb{C})) = 4$, $\dim_{\mathbb{R}}(M_2(\mathbb{C})) = 8$, $\dim_{\mathbb{C}}(\mathbb{C}) = 1$, $\dim_{\mathbb{R}}(\mathbb{C}) = 2$, $\dim_k(k[x]_2) = 3$, $\dim_k(k[x]) = \infty$ (note we will study mostly finite-dimensional vector spaces).

MATH 2201

28-10-16

- Any linearly independent set of vectors is a subset of a basis (can be "extended" to a basis).
  e.g. $\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right\}$ is linearly independent, and can be extended to $\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$.

- Any spanning set has a basis as a subset (can be "reduced" to a basis).
  e.g. $\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\2\end{pmatrix}\right\}$ spans $\mathbb{R}^2$ over $\mathbb{R}$, and we can find a subset of this that is a basis of $\mathbb{R}^2$ over $\mathbb{R}$: $\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$.

- If $\dim_k(V) = n$ and $\{v_1, \dots, v_n\}$ is a linearly independent set in $V$ over $k$, then $\{v_1, \dots, v_n\}$ is a <u>basis</u> of $V$ over $k$. (i.e. the spanning property also holds).

- If $\dim_k(V) = n$ and $\{v_1, \dots, v_n\}$ that spans $V$ over $k$, then $\{v_1, \dots, v_n\}$ is a basis of $V$ over $k$. (i.e. the linear independence property also holds).

- Suppose that, over a field $k$, $U$ is a subspace of $V$. If $\dim(U) = \dim(V)$, then $U = V$. (For finite dimensional $U$ and $V$).

## An operation on (sub)spaces

Suppose that $U, W$ are subspaces of a vector space $V$. The sum of $U$ and $W$, $U+W$, is defined as $U+W = \{u+w : u \in U, w \in W\}$

e.g. if we take $V = \mathbb{R}^3$, and $U = \left\{\begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} : x \in \mathbb{R}\right\}$,

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

i.e. $U + W = \left\{\begin{pmatrix} x \\ 0 \\ z \end{pmatrix} : x, z \in \mathbb{R}\right\}$

The sum is <u>direct</u> if, in addition, $U \cap W = \{0\}$.
We write a direct sum as $U \oplus W$.
e.g. in the previous example, the sum is direct.
if $v \in U \cap W$, then $v \in U$, so $v = \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix}$ for $x \in \mathbb{R}$

and $v \in W$, so $v = \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$ for $z \in \mathbb{R}$

$\Rightarrow \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} \Rightarrow x = 0, \ z = 0$, so $v = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

i.e. $U \cap W = \{0\}$

Key property of direct sums:
If $U$ has a basis $\{u_1, \ldots, u_n\}$
and $W$ has a basis $\{w_1, \ldots, w_m\}$
and $U \cap W = \{0\}$, then the set
$\{u_1, \ldots, u_n, w_1, \ldots, w_m\}$ is a basis for the direct sum $U \oplus W$.

As a result $\boxed{\underbrace{\dim(U \oplus W)}_{n+m} = \underbrace{\dim(U)}_{n} + \underbrace{\dim(W)}_{m}}$

This is actually a special case of the more general result: $\dim(U+W) = \dim(U) + \dim(W) - \dim(U \cap W)$.

MATH 2201

28-10-16

For example

Consider $\mathbb{R}^3$ and subspaces

$U_1 = \left\{ \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}$, $U_2 = \left\{ \begin{pmatrix} 0 \\ y \\ z \end{pmatrix} : y, z \in \mathbb{R} \right\}$,

$U_3 = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\}$

$\dim(U_1) = 1$, $\dim(U_2) = 2$, $\dim(U_3) = 2$

$U_1 + U_2 = \mathbb{R}^3$, $U_1 \cap U_2 = \{0\}$ so $U_1 \oplus U_2 \cong \mathbb{R}^3$

$\dim(\mathbb{R}^3) = \dim(U_1) + \dim(U_2)$.

$U_2 \cap U_3 = \left\{ \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} : y \in \mathbb{R} \right\}$

$$\dim(U_2 + U_3) = \dim(U_2) + \dim(U_3) - \dim(U_2 \cap U_3)$$

$$3 = 2 + 2 - 1$$

We now define linear maps:

Suppose that $V, W$ are vector spaces over a field $k$. A function $T: V \mapsto W$ is a linear map if it satisfies the following conditions:

1). $\underset{\text{identity in } V,\ 0_V}{T(0)} = \underset{\text{identity in } W,\ 0_W}{0}$

2). For all $v_1, v_2 \in V$: $T(v_1 + v_2) = T(v_1) + T(v_2)$

3). For all $v \in V$, $\lambda \in k$ : $T(\lambda v) = \lambda T(v)$

Examples

Consider the function $T: \mathbb{R}^3 \to \mathbb{R}^2$ defined as follows $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 \\ 4x_3 \end{pmatrix}$

This is a linear map.

$T\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \cdot 0 - 0 \\ 4 \cdot 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$T\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 2a_1 - a_2 \\ 4a_3 \end{pmatrix}$, $T\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 2b_1 - b_2 \\ 4b_3 \end{pmatrix}$

$$T\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right) = T\begin{pmatrix} a_1+b_1 \\ a_2+b_2 \\ a_3+b_3 \end{pmatrix} = \begin{pmatrix} 2(a_1+b_1) - (a_2+b_2) \\ 4(a_3+b_3 \end{pmatrix}$$

$$= \begin{pmatrix} 2a_1 - a_2 \\ 4a_3 \end{pmatrix} + \begin{pmatrix} 2b_1 - b_2 \\ 4b_3 \end{pmatrix}$$

So $T\begin{pmatrix} a_1+b_1 \\ a_2+b_3 \\ a_3+b_3 \end{pmatrix} = T\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + T\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$

$$T\begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \lambda a_3 \end{pmatrix} = \begin{pmatrix} 2\lambda a_1 - \lambda a_2 \\ 4\lambda a_3 \end{pmatrix} = \begin{pmatrix} \lambda(2a_1 - a_2) \\ \lambda(4a_3) \end{pmatrix} = \lambda\begin{pmatrix} 2a_1 - a_2 \\ 4a_3 \end{pmatrix} = \lambda T\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

Linear maps can be represented by matrices.
e.g. $T$, from above, can be represented as follows

$$T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 \\ x_3 \end{pmatrix}$$

To any linear map, there are two important sets associated:

1). The kernel of a linear map $T: V \mapsto W$ is the set $\{v \in V : T(v) = 0\} = Ker(T)$.

2). The image of a linear map $T: V \mapsto W$ is the set $\{w \in W : w = T(v) \text{ for some } v \in V\} = Im(T)$.

Using matrices, we can employ row reduction to determine kernels and images of linear maps.

In notes:

- $R_2 \rightarrow R_2 + 5R_3$ is used to denote $\mathcal{E}(2,3;5)$
- $R_2 \leftrightarrow R_3$ is used to denote $\mathcal{P}(2,3)$
- $R_2 \rightarrow \frac{1}{5}R_2$ is used to denote $\mathcal{D}(2, 1/5)$

MATH 2201

28-10-16

By selecting bases, we can find (different) matrix representations of linear maps:

e.g. $D : \mathbb{R}[x]_2 \to \mathbb{R}[x]_2$

defined by $D(a_0 + a_1 x + a_2 x^2) = a_1 + 2a_2 x$ "differentiation"

Examples

- Show that $\{1, x-1, x^2-1\}$ is a basis for $\mathbb{R}[x]_2$

Spanning:

For any $a_0 + a_1 x + a_2 x^2$ in $\mathbb{R}[x]_2$ try to find $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ such that

$$a_0 + a_1 x + a_2 x^2 = \lambda_1 \cdot 1 + \lambda_2 (x-1) + \lambda_3 (x^2-1)$$
$$= (\lambda_1 - \lambda_2 - \lambda_3) + \lambda_2 x + \lambda_3 x^2$$

So
$$\left.\begin{aligned} \lambda_1 - \lambda_2 - \lambda_3 &= a_0 \\ \lambda_2 &= a_1 \\ \lambda_3 &= a_2 \end{aligned}\right\} \quad \begin{aligned} \lambda_1 &= a_0 + a_1 + a_2 \\ \lambda_2 &= a_1 \\ \lambda_3 &= a_2 \end{aligned}$$

So for any $a_0 + a_1 x + a_2 x^2$:

$$a_0 + a_1 x + a_2 x^2 = (a_0 + a_1 + a_2) \cdot 1 + a_1 \cdot (x-1) + a_2 \cdot (x^2-1)$$

Linear Independence

Solve $\lambda_1 + \lambda_2(x-1) + \lambda_3(x^2-1) = 0$ .... $\lambda_1 = 0, \lambda_2 = 0, \lambda_3 = 0.$

Express $D$ in terms of $A = \{1, x-1, x^2-1\}$

i.e. find $[D]_A^A$ $\quad M(D)_A^A$

1). Apply $D$ to the basis below, $A$

2). Express your answers in terms of the basis above, $A$.

3). The resulting numbers form (columns in) a matrix, $[D]_A^A$

$$\begin{aligned} D(1) &= 0 &&= 0 \cdot 1 + 0 \cdot (x-1) + 0 \cdot (x^2-1) \\ D(x-1) &= 1 &&= 1 \cdot 1 + 0 \cdot (x-1) + 0 \cdot (x^2-1) \\ D(x^2-1) &= 2x &&= 2 \cdot 1 + 2 \cdot (x-1) + 0 \cdot (x^2-1) \end{aligned}$$

1 $\qquad$ 2

$$\text{So } [D]_A^A = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

3

MATH 2201

31-10-16

Let's complete our review from last time.

For a linear map $T: V \mapsto W$ over a field $k$ (where $V$, $W$ are vector spaces over $k$)

- the rank of $T$ is the dimension of the image of $T$: $\text{rank}(T) = \dim_k(\text{Im}(T))$
- the nullity of $T$ is the dimension of the kernal of $T$: $\text{null}(T) = \dim_k(\text{Ker}(T))$.

The result: $\dim(\text{Ker}(T)) + \dim(\text{Im}(T)) = \dim(V)$ is known as the rank-nullity theorem (or the kernel-rank theorem)

Consider a linear map $T: \mathbb{C}^n \mapsto \mathbb{C}^n$ over $\mathbb{C}$. (in the remainder of this chapter, we will study linear maps of this form).

Given any basis $\mathcal{B}$ of $\mathbb{C}^n$, can find an $n \times n$ matrix, $[T]_{\mathcal{B}}^{\mathcal{B}}$ that represents the linear map in terms of $\mathcal{B}$.

Example

Consider the linear map $T: \mathbb{C}^2 \mapsto \mathbb{C}^2$ defined as follows

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ -2x_1 + 4x_2 \end{pmatrix}$$

Choose the standard basis $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ of $\mathbb{C}^2$ over $\mathbb{C}$.

Let's determine $[T]_{\mathcal{E}}^{\mathcal{E}}$

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

$$T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} = 1\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 4\begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \text{So } [T]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$$

Can we find a basis in terms of which the matrix is particularly simple?

eg. consider the basis $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$, and find $[T]_{\mathcal{B}}^{\mathcal{B}}$

$$T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$
$$T\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 0\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 3\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

So $[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$

Recall we can "connect" $[T]_{\mathcal{E}}^{\mathcal{E}}$ and $[T]_{\mathcal{B}}^{\mathcal{B}}$ through a change of basis matrix, obtained by using $[I]_{\mathcal{B}}^{\mathcal{E}}$ and $[I]_{\mathcal{E}}^{\mathcal{B}}$, where $I$ denotes the identity map: $I\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

Here: $I\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$I\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

So $[I]_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ this is the same as placing elements of $\mathcal{B}$, in order, as columns in a matrix.

While $I\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} - 1\begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$I\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -1\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 1\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ So $[I]_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$

Note $[I]_{\mathcal{E}}^{\mathcal{B}} = \left( [I]_{\mathcal{B}}^{\mathcal{E}} \right)^{-1}$

And we have:

$[T]_{\mathcal{B}}^{\mathcal{B}} = [I]_{\mathcal{E}}^{\mathcal{B}} [T]_{\mathcal{E}}^{\mathcal{E}} [I]_{\mathcal{B}}^{\mathcal{E}}$ (note we match diagonals!)

or $[T]_{\mathcal{B}}^{\mathcal{B}} = P^{-1} [T]_{\mathcal{E}}^{\mathcal{E}} P$ where $P = [I]_{\mathcal{B}}^{\mathcal{E}}$

MATH 2201

31-10-16

Here we obtain

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

The matrix $[T]_{\mathcal{B}}^{\mathcal{B}}$ is diagonal because the basis $\mathcal{B}$ consists of eigenvectors of $T$, and $2, 3$ are eigenvalues of $T$.

(Note: In general, for any linear map $T: \mathbb{C}^n \to \mathbb{C}^n$ over $\mathbb{C}$, and bases $A, \mathcal{B}$ of $\mathbb{C}^n$ over $\mathbb{C}$:

$$[T]_{\mathcal{B}}^{\mathcal{B}} = [I]_A^{\mathcal{B}} [T]_A^A [I]_{\mathcal{B}}^A$$

i.e. $[T]_{\mathcal{B}}^{\mathcal{B}} = P^{-1} [T]_A^A P$ for some invertible $n \times n$ matrix $P$.

Any two matrices representing $T$ may be related in this way.)

Consider now $T: \mathbb{C}^2 \to \mathbb{C}^2$, where $T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3x_1 + x_2 \\ -x_1 + x_2 \end{pmatrix}$

Then if $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ : $[T]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix} := M$

Let's determine the eigenvalues and eigenvectors of this matrix:

Eigenvalue: 2

Eigenvector: $\begin{pmatrix} x_1 \\ -x_1 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ for $x_1 \in \mathbb{C}$, $x_1 \neq 0$

Check: $T\begin{pmatrix} x_1 \\ -x_1 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ -2x_1 \end{pmatrix} = 2\begin{pmatrix} x_1 \\ -x_1 \end{pmatrix}$

Characteristic polynomial:

$$\det(t \cdot I_2 - M) = \det\begin{pmatrix} t-3 & -1 \\ 1 & t-1 \end{pmatrix}$$

$$= (t-2)^2$$

Algebraic multiplicity of the eigenvalue 2 is 2

Geometric " " " " 2 is 1.

These are not the same so we deduce that $M$ is not diagonalisable: cannot find an invertible matrix $P$ such that $P^{-1}MP$ is diagonal, ie. there is no basis $\mathcal{B}$ such that $[T]_{\mathcal{B}}^{\mathcal{B}}$ is diagonal.

Let's try to find a basis that includes the eigenvector $\binom{1}{-1}$ and "simplify" the matrix.

Try $\mathcal{B} = \left\{ \binom{1}{-1}, \binom{1}{0} \right\}$.

Let's compute $[T]_{\mathcal{B}}^{\mathcal{B}}$

$$T\binom{1}{-1} = \binom{2}{-2} = 2\binom{1}{-1} + 0\binom{1}{0}$$

$$T\binom{1}{0} = \binom{3}{-1} = 1\binom{1}{-1} + 2\binom{1}{0} \quad \text{So } [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

So this is "almost" diagonal, but has an entry equal to 1 "just above" the main diagonal.

MATH 2201

1-11-16

Let's see how some definitions / results concerning matrices extended to linear maps:

Suppose that $M$ is an $n \times n$ matrix over $\mathbb{C}$:

- The characteristic polynomial of $M$ is the polynomial $\det(tI_n - M)$, where det denotes determinant, and $I_n$ denotes the $n \times n$ identity matrix. It is also denoted by $ch_M(t) = \det(tI_n - M)$ and it is a monic polynomial, of degree $n$, in $\mathbb{C}[t]$.
- The characteristic equation of $M$ is $ch_M(t) = 0 \equiv \det(tI_n - M) = 0$.

An eigenvector, $x$, of $M$ is a vector in $\mathbb{C}^n$, that is non-zero, and such that $Mx = \lambda x$ for some $\lambda \in \mathbb{C}$.
Then, $\lambda$ is the eigenvalue of $M$ corresponding to $x$.

Note: $Mx = \lambda x \Leftrightarrow Mx = \lambda I_n x \Leftrightarrow Mx - \lambda I_n x = 0$
$\Leftrightarrow (M - \lambda I_n)x = 0$
or $(\lambda I_n - M)x = 0$

So, the set of eigenvectors corresponding to an eigenvalue $\lambda$, is: $\{x \in \mathbb{C}^n : (M - \lambda I_n)x = 0,\ x \neq 0\}$

If we include the zero vector, we obtain a subspace of $\mathbb{C}^n$, known as the eigenspace corresponding to the eigenvalue $\lambda$:

$$V_1(\lambda) = \{x \in \mathbb{C}^n : (M - \lambda I_n)x = 0\}$$

Note: $V_1(\lambda) = \operatorname{Ker}(M - \lambda I_n)$.

The eigenvalues of $M$ are the roots of the characteristic polynomial of $M$, i.e. for $\lambda \in \mathbb{C}$:
$\lambda$ is an eigenvalue of $M \Leftrightarrow \det(\lambda I_n - M) = 0$
(i.e. $\lambda$ solves the characteristic equation of $M$).

The algebraic multiplicity of an eigenvalue $\lambda$ is the number of times that $\lambda$ appears as a root of $\det(\lambda I_n - M)$.

The geometric multiplicity of an eigenvalue $\lambda$ is the dimension of $V_1(\lambda)$, i.e. it is $\dim_{\mathbb{C}}(V_1(\lambda))$

Given a linear map $T: \mathbb{C}^n \to \mathbb{C}^n$, we can associate eigenvalues independently of the matrix we choose to represent $T$.

Proposition:
Given bases $\mathcal{B}, \mathcal{C}$ for $\mathbb{C}^n$, $[T]_{\mathcal{B}}^{\mathcal{B}}$ and $[T]_{\mathcal{C}}^{\mathcal{C}}$ have the same characteristic polynomial.

Proof:
Note that there exists an invertible $n \times n$ matrix $P$ such that $[T]_{\mathcal{C}}^{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}^{\mathcal{B}} P$.
Consider $\det(t \cdot I_n - [T]_{\mathcal{C}}^{\mathcal{C}})$

$$\begin{aligned}
&= \det(tI_n - P^{-1}[T]_{\mathcal{B}}^{\mathcal{B}} P) \\
&= \det(tP^{-1}I_n P - P^{-1}[T]_{\mathcal{B}}^{\mathcal{B}} P) \\
&= \det(P^{-1}tI_n P - P^{-1}[T]_{\mathcal{B}}^{\mathcal{B}} P) \\
&= \det(P^{-1}(tI_n - [T]_{\mathcal{B}}^{\mathcal{B}})P) \\
&= \det(P^{-1})\det(tI_n - [T]_{\mathcal{B}}^{\mathcal{B}})\det(P) \\
&= \frac{1}{\det(P)}\det(tI_n - [T]_{\mathcal{B}}^{\mathcal{B}})\det(P)
\end{aligned}$$

MATH 2201

1-11-16

So $\det(tI_n - [T]_{\mathcal{C}}^{\mathcal{C}}) = \det(tI_n - [T]_{\mathcal{B}}^{\mathcal{B}})$ $\square$

Eigenvalues are associated to linear maps themselves, and not just to matrices that represent them.

This allows us to extend many results from matrices to linear maps.

Important example: Cayley-Hamilton theorem for matrices.
"An $n \times n$ matrix $M$ satisfies its own characteristic equation $ch_M(M) = 0$."
(non-trivial result from MATH 1202.)

Consequence:
Cayley-Hamilton Theorem for linear maps.

Theorem:
Given a linear map $T: \mathbb{C}^n \to \mathbb{C}^n$, and any bases $\mathcal{B}$, $\mathcal{C}$ of $\mathbb{C}^n$: $ch_{[T]_{\mathcal{B}}^{\mathcal{B}}}([T]_{\mathcal{B}}^{\mathcal{B}}) = ch_{[T]_{\mathcal{C}}^{\mathcal{C}}}([T]_{\mathcal{C}}^{\mathcal{C}})$
So, for any matrix $M$ representing $T$: $ch_M(M) = 0$

Proof
Similar to proof of previous proposition.

Examples:
① Let $T: \mathbb{C}^2 \mapsto \mathbb{C}^2$ be defined via $T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ -2x_1 + 4x_2 \end{pmatrix}$
with respect to the standard basis $\mathcal{E} = \{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$
$M = [T]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$
$ch_M(t) = \det(t \cdot I - M) = \begin{vmatrix} t-1 & -1 \\ 2 & t-4 \end{vmatrix} = t^2 - 5t + 6 = (t-2)(t-3)$

So $ch_M(t) = (t-2)(t-3)$.

Then, let's verify the Cayley-Hamilton theorem:
$ch_M(M) = (M - 2I)(M - 3I)$

So $ch_m(M) = \begin{pmatrix} -1 & 1 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} -2 & 1 \\ -2 & 1 \end{pmatrix}$

$= \begin{pmatrix} 2-2 & -1+1 \\ 4-4 & -2+2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ as required.

Monic divisors of $(t-2)(t-3)$ are: $\underset{f_1(t)}{1}$, $\underset{f_2(t)}{t-2}$, $\underset{f_3(t)}{t-3}$, $\underset{f_4(t)}{(t-2)(t-3)}$

"Substitute" $t = M$ into all of them, we obtain

$f_1(M) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$f_2(M) = M - 2I = \begin{pmatrix} 1 & 1 \\ -2 & 2 \end{pmatrix}$

$f_3(M) = M - 3I = \begin{pmatrix} -2 & 1 \\ -2 & 1 \end{pmatrix}$

$f_4(M) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

So the "smallest" polynomial that sends $M$ to $O$ is $(t-2)(t-3)$

MATH 2201

04-11-16

## From last time

Trying to find "smallest" polynomial that "sends" a matrix to zero.

## Another example

Consider the linear map $T: \mathbb{C}^2 \to \mathbb{C}^2$, where $T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_1 + 3x_2 \end{pmatrix}$

With respect to the standard basis $\mathcal{E} = \{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$ of $\mathbb{C}^2$, we obtain

$$[T]_\mathcal{E}^\mathcal{E} = \begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix}$$

Characteristic polynomial of $T$:

$$ch_T(t) = \det(tI - [T]_\mathcal{E}^\mathcal{E}) = \begin{vmatrix} t-1 & 1 \\ -1 & t-3 \end{vmatrix}$$

So
$$\begin{aligned} ch_T(t) &= (t-1)(t-3) + 1 \\ &= t^2 - 4t + 4 \\ &= (t-2)^2 \end{aligned}$$

By the Cayley-Hamilton Theorem: $(T - 2I)^2 = 0$, $ch_T(T) = 0$

i.e. $([T]_\mathcal{E}^\mathcal{E} - 2I)^2 = 0$

Let's see if a divisor of $(t-2)^2$ also sends $T$ to $0$.

Possible divisors: $1$, $t-2$, $(t-2)^2$

Apply $1$ to $T$: $I(T) = I([T]_\mathcal{E}^\mathcal{E}) = 1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Note: The polynomial $1$ represents $f(t) = 1$, a constant or $f(M) = I$ in terms of matrices.

This always gives as output the identity matrix; it is a constant map.

The identity map, represented by a poynomial, has the form $f(t) = t$ or $f(M) = M$ in terms of matrices.

Apply $t-2$: $[T]_\varepsilon^\varepsilon - 2I = \begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

So, in fact, $ch_T(t) = (t-2)^2$ is the "smallest" one that works: $(T-2I)^2 = 0$.

This is the minimal polynomial of $T$.

Definition:
Suppose that, for a field $k$, $T: V \mapsto V$ is a linear map over $k$, where $V$ is a finite dimensional vector space over $k$.
Then a minimal polynomial of $T$ is a monic polynomial $m(t)$ in $k[t]$ such that:
(1) $m(T) = 0$
(2) If, for some $f(t) \in k[t]$: $f(T) = 0$, then $\deg(f) \geq \deg(m)$ or $f = 0$ (the zero polynomial).

Note: for any $T$, there always exists at least one non-zero polynomial $f(t)$ such that $f(T) = 0$, e.g. the characteristic polynomial $ch_T(t)$:
$ch_T(T) = 0$ by the Cayley-Hamilton Theorem.

Let's now check that minimal polynomials are unique:
Suppose that, for a linear map $T$, $m_1(t)$ and $m_2(t)$ are minimal polynomials for $T$: $m_1(T) = 0$ and $m_2(T) = 0$.
Then $(m_1 - m_2)(T) = m_1(T) - m_2(T) = 0$
Since $m_1$ is minimal and $m_2(T) = 0$: $\deg(m_2) \geq \deg(m_1)$
(Note: $m_1, m_2$ are assumed to be monic, so $m_1 \neq 0$ and $m_2 \neq 0$).
Similarly, $m_2(t)$ is minimal and $m_1(T) = 0$, so $\deg(m_1) \geq \deg(m_2)$.
So $\deg(m_1) = \deg(m_2)$, and so $\deg(m_1 - m_2) < \deg(m_1)$

MATH 2201

04-11-16

(Say $m_1(t) = t^n + a_{n-1}t^{n-1} + \dots + a_1 t + a_0$

$m_2(t) = t^n + b_{n-1}t^{n-1} + \dots + b_1 t + b_0$

So $(m_1 - m_2)(t) = (a_{n-1} - b_{n-1})t^{n-1} + \dots + (a_1 - b_1)t + (a_0 - b_0)$)

So, using (2) in the definition of $m_1(t)$ as a minimal polynomial, we see that $(m_1 - m_2)(t) = 0$, i.e. $m_1 - m_2 = 0$, so that $m_1(t) = m_2(t)$.
So $(m_1 - m_2)(T) = 0$ and $\deg(m_1 - m_2) < \deg(m_1)$ so we must have $m_1 - m_2 = 0$.

So minimal polynomials are unique: $m_1(t) = m_2(t)$.

Let's now also show that, if $f(T) = 0$ for some non-zero $f(t) \in k[t]$, then the minimal polynomial $m(t)$ divides $f(t)$

To show this apply Euclidean division to $f(t)$ and $m(t)$:
$\exists\, q(t), r(t) \in k[t]$ such that:

$f(t) = q(t)m(t) + r(t)$ and where $\deg(r) < \deg(m)$

Apply both sides to $T$:

$$\underset{\overset{\parallel}{0}}{f(T)} = q(T)\underset{\overset{\parallel}{0}}{m(T)} + r(T)$$

So $0 = q(T) \cdot 0 + r(T)$ i.e. $r(T) = 0$
Then, $r(T) = 0$ and $\deg(r) < \deg(m)$, so, by (2) of definition of the minimal polynomial: $r(t) = 0$ i.e. $r$ is the zero polynomial.
Hence: $f(t) = q(t)m(t) + r(t) = q(t)m(t) + 0 = q(t)m(t)$
i.e. $m(t)$ divides $f(t)$ as required.

In particular, sinch $ch_T(T)=0$, for some linear map $T$: the minimal polynomial $m(t)$ is a monic divisor of the characteristic polynomial $ch_T(t)$.

So we can "search" for the minimal polynomial of $T$ by starting from $ch_T(t)$.

Example

Suppose, for some linear map $T$: $ch_T(t) = (t-2)(t-3)^2$.
Possible choices for the minimal polynomial, $m_T(t)$:
$1,\ t-2,\ t-3,\ (t-3)^2,\ (t-2)(t-3),\ (t-2)(t-3)^2$
Apply $T$ to each of these ~~and~~ find the smallest degree polynomial that "takes" $T$ to $0$:
Compute $T-2I,\ T-3I,\ (T-3I)^2,\ (T-2I)(T-3I),\ (T-2I)(T-3I)^2$ and check.

Let's now try to show that $m_T(t)$ must have every eigenvalue of $T$ as a root: if $\lambda$ is an eigenvalue of $T$, then $m_T(\lambda)=0$.

For instance in this example, we have eigenvalues 2 and 3:
For some non-zero vectors $v, w$: $Tv = 2v$, $Tw = 3w$
i.e. $(T-2I)v = 0$ and $(T-3I)w = 0$
and $t-2$ and $t-3$ will be factors of $m_T$.
Why? Let's prove it in general:
If $\lambda$ is an eigenvalue of a linear map $T$, then $Tv = \lambda v$ for some vector $v$, $v \neq 0$.
Note: If $T(v) = \lambda v$, then, for any $r \in \mathbb{N}$: $T^r(v) = \lambda^r v$
(eg. $T^2(v) = T(T(v)) = T(\lambda v) = \lambda T(v) = \lambda(\lambda v) = \lambda^2 v$ etc.)

Consider the minimal polynomial of $T$, $m_T(T)$ applied to $v$.

MATH 2201

04-11-16

Suppose $m_T(t) = t^n + a_{n-1}t^{n-1} + \dots + a_1 t + a_0$

then $m_T(T)(v) = (T^n + a_{n-1}T^{n-1} + \dots + a_1 T + a_0)(v)$

$= T^n(v) + a_{n-1}T^{n-1}(v) + \dots + a_1 T(v) + a_0 v$

$= \lambda^n v + a_{n-1}\lambda^{n-1} v + \dots + a_1 \lambda v + a_0 v$

$= (\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0) v$

So $m_T(T)(v) = m_T(\lambda) \cdot v$

By definition of $m_T$, $m_T(T) = 0$, so $(m_T(T))v = 0$

Hence: $m_T(\lambda) \cdot v = 0$

Then, since $v \neq 0$, must have $m_T(\lambda) = 0$, as required.
So, every eigenvalue of $T$ is a root of $m_T(t)$ (and of $ch_T(t)$).

So to find $m_T(t)$ for a general linear map $T: V \mapsto V$

- first compute $ch_T(t) = \det(tI - T)$
- consider all monic divisors of $ch_T(t)$ which include every root of $ch_T(t)$
- substitute "$t = T$" into these to find $m_T(t)$.

Examples

Consider the linear map $T: \mathbb{C}^3 \to \mathbb{C}^3$ such that

$$[T]_\varepsilon^\varepsilon = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{for} \quad \varepsilon = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

ie. $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 3x_2 \\ 3x_3 \end{pmatrix}$

Compute $ch_T(t) = \begin{vmatrix} t-2 & 0 & 0 \\ 0 & t-3 & 0 \\ 0 & 0 & t-3 \end{vmatrix} = (t-2)(t-3)^2$

Possibilities for $m_T(t)$: $\underbrace{(t-2)(t-3)}_{f_1(t)}, \underbrace{(t-2)(t-3)^2}_{f_2(t)}$

Apply $T$ and check:

$$f_1(T) = (T-2I)(T-3I) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

So $m_T(t) = (t-2)(t-3)$ ← we have found the minimal polynomial.

(Key result from later: because $m_T(t)$ contains precisely one copy of each eigenvalue, $T$ is diagonalisable.)

Suppose now, that, for a linear map $T : \mathbb{C}^3 \mapsto \mathbb{C}^3$,

$$[T]_{\varepsilon}^{\varepsilon} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 7 \\ 0 & 0 & 3 \end{pmatrix}, \quad ch_T(t) = \begin{vmatrix} t-2 & 0 & 0 \\ 0 & t-3 & -7 \\ 0 & 0 & t-3 \end{vmatrix} = (t-2)(t-3)^2$$

Possible choices for $m_T(t)$: $(t-2)(t-3)$, $(t-2)(t-3)^2$.

Apply $T$ and check: $m_T(T) = (T-2I)(T-3I)$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

So $m_T(t) \neq (t-2)(t-3)$.

It must be that $m_T(t) = (t-2)(t-3)^2$

Check: $(T-2I)(T-3I)^2 = \big((T-2I)(T-3I)\big)(T-3I)$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 7 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

As we will see: the presence of $(t-3)^2$, i.e. two "copies" of the eigenvalue 3 in $m$

MATH 2201

14-11-16

If we *cannot* diagonalise a complex square $n \times n$ matrix $M$, it means that, for some eigenvalue, $\lambda$, say, of $M$, we cannot find "enough" eigenvectors, i.e. the eigenspace

$$V_1(\lambda) = \{x \in \mathbb{C}^n : (M - \lambda I_n)(x) = 0\} = \ker(M - \lambda I_n)$$

is not big enough.

In such cases, we look at *generalised eigenspaces*.

Suppose that, for some $n \times n$ complex matrix $M$, $\lambda$ is an eigenvalue of $M$.
Then the $r^{th}$ generalised eigenspace corresponding to $\lambda$ is

$$V_r(\lambda) = \{x \in \mathbb{C}^n : (M - \lambda I_n)^r(x) = 0\}$$

(When $r = 1$, we obtain the "usual" eigenspace.)

$$V_r(\lambda) = \ker((M - \lambda I_n)^r)$$

Every non-zero element of $V_r(\lambda)$ is a *generalised eigenvector*.

Two key features of generalised eigenspaces:

- For any $n \times n$ matrix $A$, note that if $Ax = 0$, then $A^2 x = A(Ax) = A \cdot 0 = 0$
  In general, if $A^r x = 0$ and $m \geq r$, then $A^m x = A^{m-r}(A^r x) = A^{m-r}(0) = 0$
  So, if $x \in \ker(A^r)$, then $x \in \ker(A^m)$ $\forall m \geq r$.
  Take $A = M - \lambda I_n$ : if $x \in \ker((M - \lambda I_n)^r)$ then $x \in \ker((M - \lambda I_n)^m)$.
  So, for $m \geq r$ : if $x \in V_r(\lambda)$, then $x \in V_m(\lambda)$,
  so $\boxed{V_r(\lambda) \subseteq V_m(\lambda)}$

So, for a given eigenvalue $\lambda$:
$V_1(\lambda) \subseteq V_2(\lambda) \subseteq V_3(\lambda) \subseteq \ldots$

- For any $n \times n$ matrix $A$, if $A^2 x = 0$, then $A(Ax) = 0$.
  So if $x \in \ker(A^2)$ then $Ax \in \ker(A)$.
  In general: if $\underbrace{x \in \ker(A^{r+1})}_{A^{r+1}x=0}$, then $\underbrace{Ax \in \ker(A^r)}_{A^r(Ax)=0}$
  Take $A = M - \lambda I_n$:
  if $x \in \ker((M-\lambda I_n)^{r+1})$, then $(M-\lambda I_n)x \in \ker((M-\lambda I_n)^r)$
  so if $x \in V_{r+1}(\lambda)$, then $(M-\lambda I_n)x \in V_r(\lambda)$.

As we shall prove later, in order to find enough generalised eigenvectors to "reach" the algebraic multiplicity of an eigenvalue $\lambda$, we need to get to $V_b(\lambda)$ where $b$ is the exponent / power of $(t-\lambda)^b$ in the minimal polynomial.

For example, consider $A = \begin{pmatrix} 5 & 2 & 4 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$

Characteristic polynomial of $A$: $\det(tI - A)$
$$= \begin{vmatrix} t-5 & -2 & -4 \\ 0 & t-5 & 0 \\ 0 & 0 & t-5 \end{vmatrix}$$

So $ch_A(t) = (t-5)^3$

Possibilities for the minimal polynomial $m_A(t)$:
$t-5, (t-5)^2, (t-5)^3$

MATH 2201

14-11-16

$$A - 5I = \begin{pmatrix} 0 & 2 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0 \quad \text{so } m_A(t) \neq t-5$$

$$(A-5I)^2 = \begin{pmatrix} 0 & 2 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 2 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

so $m_A(t) = (t-5)^2$

The algebraic multiplicity of the eigenvalue 5 is 3, and $m_A(t) = (t-5)^2$ suggests that we need to go to $V_2(5)$ to find 3 linearly independent vectors.

Let's determine (bases for) $V_1(5)$, $V_2(5)$:

for $V_1(5)$, solve: $(A-5I)^1 x = 0$

i.e. $$\begin{pmatrix} 0 & 2 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

We obtain $2x_2 + 4x_3 = 0$ i.e. $x_2 = -2x_3$

So general solution is $\begin{pmatrix} x_1 \\ -2x_3 \\ x_3 \end{pmatrix}$ for $x_1, x_3 \in \mathbb{C}$

$$V_1(5) = \left\{ \begin{pmatrix} x_1 \\ -2x_3 \\ x_3 \end{pmatrix} = x_1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_3\begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} : x_1, x_3 \in \mathbb{C} \right\}$$

Basis for $V_1(5)$ over $\mathbb{C}$: $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right\}$

For $V_2(5)$, solve $(A-5I)^2 x = 0$ i.e. $$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

A general solution is $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ for $x_1, x_2, x_3 \in \mathbb{C}$.

$$V_2(5) = \mathbb{C}^3 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Basis for $V_2(5) = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

$V_2$ has dimension 3, confirming what we said earlier.

Basis for $V_1(5) = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right\}$ and $V_1(5) \subset V_2(5)$.

We know that $V_1(5) \subset V_2(5)$ and we wish to obtain a basis for $V_2(5)$ which includes a basis for $V_1(5)$.
Such a basis is called a pre-Jordan basis.

To find such a basis, try to replace vectors in the $V_2(5)$ basis by the vectors in the $V_1(5)$ basis, while keeping it a basis.

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

So replace $\begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$ by one of the vetors it "relates to".

No need to replace $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$: it's already in the basis for $V_2(5)$.

We obtain the pre-Jordan basis: $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

MATH 2201

15-11-16

From yesterday:

$$A = \begin{pmatrix} 5 & 2 & 4 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

Note: 'A' represents a linear map $T: \mathbb{C}^3 \mapsto \mathbb{C}^3$ where $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5x_1 + 2x_2 + 4x_3 \\ 5x_2 \\ 5x_3 \end{pmatrix}$

with respect to the standard basis

$\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ $\quad A = [T]_{\mathcal{E}}^{\mathcal{E}}$

$m_T(t) = (t-5)^2$

Basis for $V_1(5) = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right\}$

Basis for $V_2(5) = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

Try to replace vectors in the basis for $V_2(5)$ by vectors in the basis of $V_1(5)$ in a suitable way, so that we still have a basis of $V_2(5)$ at the end.

From the basis of $V_1(5)$: $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ already appears in the basis for $V_2(5)$, so no need to take further action.

$\begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$ is inside $V_2(5)$ (since $V_1(5) \subset V_2(5)$):

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

To obtain a basis of $V_2(5)$ containing $\begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$, we replace one of the vectors "related" to $\begin{pmatrix}0\\-2\\1\end{pmatrix}$, i.e. one of $\begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}$, by $\begin{pmatrix}0\\-2\\1\end{pmatrix}$.

e.g. if we replace $\begin{pmatrix}0\\-2\\1\end{pmatrix}$ by $\begin{pmatrix}0\\1\\0\end{pmatrix}$ say, to obtain the following set in $V_2(S)$: $\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\-2\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$.

This is still a basis of $V_2(S)$, e.g. we can retrieve $\begin{pmatrix}0\\1\\0\end{pmatrix}$ as a linear combination of these two vectors by rearranging (*):

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

The basis $\underbrace{\underbrace{\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\-2\\1\end{pmatrix}\right.}_{\text{basis of } V_1(S)}, \left.\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}}_{\text{basis of } V_2(S)}$ is a pre-Jordan basis.

General definition: Suppose that $\lambda$ is an eigenvalue of a linear map $T: \mathbb{C}^n \mapsto \mathbb{C}^n$, and that $\lambda$ has multiplicity $b$ in the minimal polynomial ($m_T(t)$) of $T$, i.e. the term $(t-\lambda)^b$ appears in $m_T(t)$.

Then $V_1(\lambda) \subseteq V_2(\lambda) \subseteq \dots \subseteq V_b(\lambda)$

A pre-Jordan basis for $\lambda$ is a basis of $V_b(\lambda)$ of the form $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_b$ such that:

$\mathcal{B}_1$ is a basis for $V_1(\lambda)$

$\mathcal{B}_1 \cup \mathcal{B}_2$ " " " $V_2(\lambda)$

$\mathcal{B}_1 \cup \dots \cup \mathcal{B}_b$ " " " " $V_b(\lambda)$

MATH 2201

15-11-16

For instance, in our example $\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\-2\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$ is a pre-Jordan basis for the eigenvalue 5, where $\mathcal{B}_1 = \left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\-2\\1\end{pmatrix}\right\}$, $\mathcal{B}_2 = \left\{\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$.

We next create "links" between vectors in $\mathcal{B}_2$ and $\mathcal{B}_1$ using an earlier result:

If $x \in V_{r+1}(\lambda)$, then $(A-\lambda I)(x) \in V_r(\lambda)$.

Take "each" vector in $\mathcal{B}_2$ and use this to find a linked vector in $V_1(5)$:

If $x \in \mathcal{B}_2$, then $x \in V_2(5)$.

Apply $A-5I$ to $\begin{pmatrix}0\\0\\1\end{pmatrix}$: $\begin{pmatrix}0&2&4\\0&0&0\\0&0&0\end{pmatrix}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}4\\0\\0\end{pmatrix}$

I would like $\begin{pmatrix}4\\0\\0\end{pmatrix}$ to be in $\mathcal{B}_1$: $\begin{pmatrix}4\\0\\0\end{pmatrix} = 4\begin{pmatrix}1\\0\\0\end{pmatrix} + 0\cdot\begin{pmatrix}0\\-2\\1\end{pmatrix}$

To do this and still have a basis, we replace $\begin{pmatrix}4\\0\\0\end{pmatrix}$ by the "related" vector $\begin{pmatrix}1\\0\\0\end{pmatrix}$ in $\mathcal{B}_1$

to obtain: $\mathcal{B}_1' = \left\{\begin{pmatrix}4\\0\\0\end{pmatrix}, \begin{pmatrix}0\\-2\\1\end{pmatrix}\right\}$, $\mathcal{B}_2' = \left\{\begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$ no change to $\mathcal{B}_2$.

Why do we want to have both $\begin{pmatrix}0\\0\\1\end{pmatrix}$ and the "linked" vector $\begin{pmatrix}4\\0\\0\end{pmatrix}$?

Since $(A-5I)\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}4\\0\\0\end{pmatrix}$, we obtain $A\begin{pmatrix}0\\0\\1\end{pmatrix} - 5\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}4\\0\\0\end{pmatrix}$.

Rearranging this: $A\begin{pmatrix}0\\0\\1\end{pmatrix} = 1\cdot\begin{pmatrix}4\\0\\0\end{pmatrix} + 5\cdot\begin{pmatrix}0\\0\\1\end{pmatrix}$ "nice" formula that leads to a "nice" matrix.

We can picture this as

link

$$\begin{array}{lcc} \mathcal{B}_2' & * & \\ \mathcal{B}_1' & * & * \end{array} \qquad\qquad \begin{array}{lcc} \mathcal{B}_2' & \begin{pmatrix}0\\0\\1\end{pmatrix} & \\ \mathcal{B}_1' & \begin{pmatrix}4\\0\\0\end{pmatrix} & \begin{pmatrix}0\\-2\\1\end{pmatrix} \end{array}$$

We now arrive at the final, Jordan, basis, by rearranging the three vectors, so that "linked" vectors appear next to each other in order (from $\mathcal{B}_1$ to $\mathcal{B}_2$).

For instance, a possible Jordan basis here is:

$$\left\{ \overset{\mathcal{B}_1}{\begin{pmatrix}4\\0\\0\end{pmatrix}}, \overset{\mathcal{B}_2}{\begin{pmatrix}0\\0\\1\end{pmatrix}}, \overset{\mathcal{B}_1}{\begin{pmatrix}0\\-2\\1\end{pmatrix}} \right\} = \mathcal{E}_1$$

linked

Another Jordan basis would be $\left\{ \begin{pmatrix}0\\-2\\1\end{pmatrix}, \begin{pmatrix}4\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix} \right\} = \mathcal{E}_2$

Let's now express $T: \mathbb{C}^3 \to \mathbb{C}^3$, where $T\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}5x_1 + 2x_2 + 4x_3\\5x_2\\5x_3\end{pmatrix}$

- In terms of $\mathcal{E}_1$:

$$T\begin{pmatrix}4\\0\\0\end{pmatrix} = \begin{pmatrix}20\\0\\0\end{pmatrix} = 5\begin{pmatrix}4\\0\\0\end{pmatrix} + 0\begin{pmatrix}0\\0\\1\end{pmatrix} + 0\begin{pmatrix}0\\-2\\1\end{pmatrix}$$
$$T\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}4\\0\\5\end{pmatrix} = 1\begin{pmatrix}4\\0\\0\end{pmatrix} + 5\begin{pmatrix}0\\0\\1\end{pmatrix} + 0\begin{pmatrix}0\\-2\\1\end{pmatrix}$$
$$T\begin{pmatrix}0\\-2\\1\end{pmatrix} = \begin{pmatrix}0\\-10\\5\end{pmatrix} = 0\begin{pmatrix}4\\0\\0\end{pmatrix} + 0\begin{pmatrix}0\\0\\1\end{pmatrix} + 5\begin{pmatrix}0\\-2\\1\end{pmatrix}$$

So $[T]_{\mathcal{E}_1}^{\mathcal{E}_1} = \begin{pmatrix}5 & 1 & 0\\0 & 5 & 0\\0 & 0 & \boxed{5}\end{pmatrix}$

This is a Jordan normal form for $T$.
It is a matrix that consists of two Jordan blocks "placed diagonally".
For $\lambda \in \mathbb{C}$, the Jordan block $J_m(\lambda)$ is the $m \times m$ matrix defined as follows: $J_{i,i} = \lambda$ , $i = 1, \dots, m$

$J_{i,i+1} = 1$ , $i = 1, \dots, m-1$

$J_{i,j} = 0$ otherwise

MATH
2201

15-11-16

e.g. $[T]_{\mathcal{E}_1}^{\mathcal{E}_1}$ consists of $J_2(5)$ and a $J_1(5)$ "placed diagonally" to form a $3\times3$ matrix

MATH 2201

18-11-16

From last time:

Linear map $T: \mathbb{C}^3 \mapsto \mathbb{C}^3$ such that $[T]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} 5 & 2 & 4 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$

for standard basis $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Found a basis (a Jordan basis) $\mathcal{E}_1 = \left\{ \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right\}$

such that $[T]_{\mathcal{E}_1}^{\mathcal{E}_1} = \begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$.

We found a Jordan normal form for $T$.

A Jordan block, $J_r(\lambda)$, for some $r \in \mathbb{N}$ and $\lambda \in \mathbb{C}$, is a complex $r \times r$ matrix defined as follows:

$(J_r(\lambda))_{i,i} = \lambda$ for $i = 1, \ldots, r$

$(J_r(\lambda))_{i,i+1} = 1$ for $i = 1, \ldots, r$

$(J_r(\lambda))_{i,j} = 0$ otherwise.

So $J_r(\lambda) = \begin{pmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & 0 & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}$

A square complex matrix is in Jordan normal form if it is obtained by placing Jordan blocks diagonally (across the main diagonal), i.e. it is a matrix of the form

$$\begin{pmatrix} J_{r_1}(\lambda_1) & & & 0 \\ & J_{r_2}(\lambda_2) & & \\ 0 & & \ddots & \\ & & & J_{r_n}(\lambda_n) \end{pmatrix}$$

for $r_1, \ldots, r_n \in \mathbb{N}$, $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$

could have

$r_i = r_j$ for some $i, j$

$\lambda_i = \lambda_j$ for some $i, j$

Examples of matrices in Jordan normal form:

$$\begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix} : \begin{pmatrix} J_2(5) & 0 \\ 0 & J_1(5) \end{pmatrix}$$

$$\begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{pmatrix} : \begin{pmatrix} J_1(5) & 0 \\ 0 & J_2(5) \end{pmatrix}$$

$$\begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{pmatrix} : J_3(5)$$

$$\begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix} : \begin{pmatrix} J_1(5) & 0 & 0 \\ 0 & J_1(5) & 0 \\ 0 & 0 & J_1(5) \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} : \begin{pmatrix} J_2(2) & 0 & 0 \\ 0 & J_1(2) & 0 \\ 0 & 0 & J_2(3) \end{pmatrix}$$

Nonexamples of Jordan normal form:

$$\begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \leftarrow !! \quad , \quad \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} \leftarrow !!$$

MATH 2201
18-11-16

Any square complex matrix can be transformed to one in Jordan normal form.
In general, for any linear map, $T: \mathbb{C}^n \mapsto \mathbb{C}^n$, there exists a basis $\mathcal{B}$ such that $[T]_{\mathcal{B}}^{\mathcal{B}}$ is in Jordan normal form.

We can find such a basis, $\mathcal{B}$, using the following algorithm:

- Determine $m_T(t)$, the minimal polynomial of $T$ (perhaps by finding, $ch_T(t)$, the characteristic polynomial of $T$ first).

  $$m_T(t) = (t-\lambda_1)^{b_1} \dots (t-\lambda_r)^{b_r}$$

  where $\lambda_1, \dots, \lambda_r$ are the distinct eigenvalues of $T$.

- For each eigenvalue $\lambda_i$, $1 \leq i \leq r$:
  - determine $V_1(\lambda_i), V_2(\lambda_i), \dots, V_{b_i}(\lambda_i)$
  - find bases for $V_1(\lambda_i), \dots, V_{b_i}(\lambda_i)$
  - by perhaps suitably exchanging vectors, find a pre-Jordan basis for $V_{b_i}(\lambda_i)$, ie. a basis

    $$\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_{b_i}$$

    such that $\mathcal{B}_1$ is a basis for $V_1(\lambda_i)$, $\mathcal{B}_1 \cup \mathcal{B}_2$ is a basis for $V_2(\lambda_i)$, ..., $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_{b_i}$ is a basis for $V_{b_i}(\lambda_i)$
  - For each vector $v$ in the basis $\mathcal{B}_{b_i}$ and, after possibly exchanging suitably:

    make $(T-\lambda_i I)(v)$ appear in $\mathcal{B}_{b_i - 1}$, then

    make $(T-\lambda_i I)^2(v)$ " " $\mathcal{B}_{b_i - 2}$,

    ⋮

    make $(T-\lambda_i I)^{b_i - 1}(v)$ " " $\mathcal{B}_1$.

    Then, $v, (T-\lambda_i I)(v), (T-\lambda_i I)^2(v), \dots, (T-\lambda_i I)^{b_i - 1}(v)$ is a set of "linked" vectors
  - Then, rearrange the new basis, so that "linked" vectors appear next to each other

$$\underset{\in V_1(\lambda_i)}{(T-\lambda_i I)^{b_i-1}(v)}, \ldots, \underset{\in V_{b_i-1}(\lambda_i)}{(T-\lambda_i I)(v)}, \underset{\in V_{b_i}(\lambda_i)}{v}$$

This gives a Jordan basis for the eigenvalue $\lambda_i$.

- Place together Jordan basis for all eigenvalues $(\lambda_1, \ldots, \lambda_r)$. This will give a Jordan basis for $T$, i.e. a basis $\mathcal{B}$ such that $[T]_{\mathcal{B}}^{\mathcal{B}}$ is in Jordan normal form.

To show that this works in general, need to know:
1. the "powers" in the minimal polynomial will lead to generalised eigenspaces, whose bases will be "large enough" to find a basis in the end.
2. when we "exchange" vectors in this process, we are allowed to do so, and still have a basis at the end.

Let's start with 2:

We can exchange vectors using the Exchange Lemma:
Suppose that $\{v_1, \ldots, v_r\}$ is a linearly independent set in $V$, and $\{w_1, \ldots, w_m\}$ a spanning set for $V$ (where $V$ is a subspace of $\mathbb{C}^n$),
then $r \leq m$ and there exists a spanning set of $V$ of the form $\{v_1, \ldots, v_r, w'_{r+1}, \ldots, w'_m\}$ where each $w'_i \in \{w_1, \ldots, w_m\}$ for $r+1 \leq i \leq m$.

MATH 2201

18-11-16

We use this result twice in our algorithm:

(i) to obtain a pre-Jordan basis starting from a basis of $V_{b_i}(\lambda_i)$ : here, we already have elements in a basis, so they are linearly independent and we can apply the result.

(ii) while finding a Jordan basis, when, for each $v \in \mathcal{B}_{b_i}$, we try to "make" $v, (T-\lambda I)(v), \dots, (T-\lambda I)^{b_i-1}(v)$ appear in the new basis. To do this using the result, we must check that $v, (T-\lambda I)v, \dots, (T-\lambda I)^{b_i-1}(v)$ are linearly independent.

Note that in the pre-Jordan basis:

$\mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_{b_i-1} \cup \mathcal{B}_{b_i}$,

a vector $v$ in $\mathcal{B}_{b_i}$ satisfies:

$$v \in V_{b_i}(\lambda_i) \text{ but } v \notin V_{b_{i-1}}(\lambda_i)$$

Reminder:

$$V_r(\lambda_i) = \ker((T-\lambda_i I)^r) = \{x \in \mathbb{C}^n : (T-\lambda_i I)^r(v) = 0\}$$

So $(T-\lambda I)^{b_i}(v) = 0$ but $(T-\lambda I)^{b_i-1}(v) \neq 0$.

Similarly, for $v \in \mathcal{B}_{r+1}$: $(T-\lambda I)^{r+1}(v) = 0$, $v \in V_{r+1}(\lambda)$

but $(T-\lambda I)^r(v) \neq 0$, $v \notin V_r(\lambda)$.

<u>Using this we can prove</u>

<u>Proposition</u>

Suppose that, for some eigenvalue $\lambda$, of a linear map $T: \mathbb{C}^n \mapsto \mathbb{C}^n$, $v \in \mathcal{B}_b$ (where $b$ is the "power" of $(t-\lambda)$ in $m_T(t)$). Then, the vectors

$$v, (T-\lambda I)(v), (T-\lambda I)^2(v), \dots, (T-\lambda I)^{b-1}(v)$$

are linearly independent.

Proof:

Suppose that, for $\lambda_0, \dots, \lambda_{b-1} \in \mathbb{C}$:

$\lambda_0 v + \lambda_1 (T - \lambda I)(v) + \dots + \lambda_{b-1}(T - \lambda I)^{b-1}(v) = 0$

Apply $(T-\lambda I)^{b-1}$ to both sides:

$(T - \lambda I)^{b-1}(\lambda_0 v + \dots + \lambda_{b-1}(T-\lambda I)^{b-1}(v)) = (T - \lambda I)^{b-1}(0)$

So $\lambda_0 (T-\lambda I)^{b-1}(v) + \lambda_1 (T-\lambda I)^{b}(v) + \dots + \lambda_{b-1}(T-\lambda I)^{2b-2}(v) = 0$ by linearity.

Note $v \in B_b$ so $(T-\lambda I)^b(v) = 0$, so $(T-\lambda I)^r(v) = 0$ for each $r \geq b$.

Hence, the equation reduces to $\lambda_0 (T-\lambda I)^{b-1}(v) = 0$

Also note that $v \in B_b$, so $(T-\lambda I)^b(v) = 0$ but $(T-\lambda I)^{b-1}(v) \neq 0$.

So, since $\lambda_0 (T-\lambda I)^{b-1}(v) = 0$ but $(T-\lambda I)^{b-1}(v) \neq 0$, we must have $\lambda_0 = 0$.

Now $\lambda_1 (T-\lambda I)(v) + \dots + \lambda_{b-1}(T-\lambda I)^{b-1}(v) = 0$.

Proceed similarly: apply $(T-\lambda I)^{b-2}$ to this, and use $(T-\lambda I)^r(v) = 0$ for $r \geq b$, to obtain

$\lambda_1 (T-\lambda I)^{b-1}(v) = 0$

Since $(T-\lambda I)^{b-1}(v) \neq 0$, we obtain $\lambda_1 = 0$.

Using a similar process: $\lambda_2 = 0, \lambda_3 = 0, \dots, \lambda_{b-1} = 0$.

So $\lambda_i = 0$ for each $0 \leq i \leq b-1$, i.e. the given vectors are, indeed, linearly independent. $\square$

So ② is OK.

Let's now try to study ①:

We wish to show that if $m_T(t) = (t-\lambda_1)^{b_1} \dots (t-\lambda_r)^{b_r}$

then $\underset{\substack{\uparrow \\ \text{"whole space"}}}{\mathbb{C}^n} \cong V_{b_1}(\lambda_1) \oplus \dots \oplus V_{b_r}(\lambda_r)$

This will show ① is OK.

MATH 2201

18-11-16

To show this, we first prove:

Proposition

Suppose that the polynomials $f(t), g(t)$ in the ring of polynomials $k[t]$, for a field $k$, are coprime, i.e. that $\gcd(f,g) = 1$

Then: $\ker(\underbrace{fg}_{\text{product } f(x)\cdot g(x)}) = \ker(f) \oplus \ker(g)$

Proof:

Let's first show that $\ker(fg) = \ker(f) + \ker(g)$ and then show that the sum $\ker(f) + \ker(g)$ is direct.

$\ker(fg) = \ker(f) + \ker(g)$ ?

Suppose that $v \in \ker(f) + \ker(g)$

Then, for $v = v_1 + v_2$ where $v_1 \in \ker(f)$ i.e. $f(t))(v_1) = 0$

and $v_2 \in \ker(g)$ i.e. $(g(t))(v_1) = 0$

Apply $fg$ to $v$: $(f(t)g(t))(v) = (f(t)g(t))(v_1 + v_2)$

$= f(t)g(t)(v_1) + f(t)g(t)(v_2)$

$= g(t)f(t)(v_1) + g(t)f(t)(v_2)$

So $(f(t)g(t))(v) = g(t)(f(t)(v_1)) + f(t)(g(t)(v_2))$

$= g(t)(0) + f(t)(0)$

$= 0 + 0$

So $(f(t)g(t))(v) = 0$ so $v \in \ker(fg)$

So $\ker(f) + \ker(g) \subseteq \ker(fg)$

As we will see:

the presence of $(t-3)^2$, i.e. two "copies" of the eigenvalue 3 in $m_T(t)$ means that we will not be able to find enough eigenvectors for 3 to diagonalise $T$.

To find eigenvectors for 3, solve $([T]_\varepsilon^\varepsilon - 3I)x = 0$

i.e. $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 7 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

So $-x_1 = 0$, $0 = 0$, $7x_3 = 0$

So the general eigenvector for $\lambda = 3$ is $\begin{pmatrix} 0 \\ x_2 \\ 0 \end{pmatrix}$, $x_2 \in \mathbb{C}$, $x_2 \neq 0$.

The eigenspace $V_1(3) = \left\{ \begin{pmatrix} 0 \\ x_2 \\ 0 \end{pmatrix} : x_2 \in \mathbb{C} \right\}$ with basis $\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$

$\left.\begin{array}{l} \text{Algebraic multiplicity of 3 is 2} \\ \text{Geometric " " 3 " 1} \end{array}\right\}$ So $T$ not diagonalisable.

The presence of $(t-3)^2$ in $m_T(t)$ reveals that, to find enough vetors to "almost" diagonalise $T$, we need to consider the "generalised" eigenspace

$V_2(3) = \{ x \in \mathbb{C}^3 : ([T]_\varepsilon^\varepsilon - 3I)^2 x = 0 \}$

Note $V_1(3) = \{ x \in \mathbb{C}^3 : ([T]_\varepsilon^\varepsilon - 3I)^1 x = 0 \}$ - Eigenspace

MATH 2201

21-11-16

If $T: \mathbb{C}^n \to \mathbb{C}^n$ is a linear map and $f(t)$ is in $\mathbb{C}[t]$ then $f(T)$ is also a linear map $f(T): \mathbb{C}^n \mapsto \mathbb{C}^n$

Proposition

Suppose that $T: \mathbb{C}^n \mapsto \mathbb{C}^n$ is a complex linear map, and that the polynomials $f(t), g(t)$, in $\mathbb{C}[T]$, are coprime, so that $\gcd(f, g) = 1$.
Then $\ker(f(T)g(T)) = \ker(f(T)) \oplus \ker(g(T))$

Proof

Let's first show that $\ker(f(T)g(T)) = \ker(f(T)) + \ker(g(T))$ (and, then, that the sum is direct).
Suppose $v \in \ker(f(T)) + \ker(g(T))$
Then $v = v_1 + v_2$ where $v_1 \in \ker(f(T))$, i.e. $f(T)(v_1) = 0$
$v_2 \in \ker(g(T))$, $g(T)(v_2) = 0$

Then,
$$\begin{aligned}(f(T)g(T))(v) &= (f(T)g(T))(v_1+v_2)\\ &= (f(T)g(T))(v_1) + (f(T)g(T))(v_2)\\ &\quad \text{by linearity of } f(T)g(T)\\ &= g(T)\underbrace{f(T)(v_1)}_{0} + f(T)\underbrace{g(T)(v_2)}_{0}\\ &= g(T)(0) + f(T)(0)\\ &= 0 \text{ by linearity of } f(T), g(T).\end{aligned}$$

So $v \in \ker(f(T)g(T))$
Hence $\ker(f(T)) + \ker(g(T)) \subseteq \ker(f(T)g(T))$

Now suppose $v \in \ker(f(T)g(T))$
Since $f(t), g(t)$ are coprime, by Bezout's Lemma, there exist $a(t), b(t)$ in $\mathbb{C}[t]$ such that
$$a(t)f(t) + b(t)g(t) = 1 \qquad [1 = \gcd(f, g)]$$

i.e. $a(T)f(T) + b(T)g(T) = Id$ (identity map) i.e. $(a(T)f(T) + b(T)g(T))(v) = v$

So $(a(T)f(T))(v) + (b(T)g(T))(v) = v$

Let $v_1 = (b(T)g(T))(v)$ and $v_2 = (a(T)f(T))(v)$

Then $v = v_1 + v_2$

and $f(T)(v_1) = (f(T)b(T)g(T))(v) = b(T)(f(T)g(T)(v))$
$= b(T)(0)$ since $v \in \ker(f(T)g(T))$

$\Rightarrow f(T)(v_1) = 0$

So $v_1 \in \ker(f(T))$

Similarly: $g(T)(v_2) = (g(T)a(T)f(T))(v)$
$= a(T)((f(T)g(T))(v))$
$= a(T)(0)$

$g(T)(v_2) = 0$

So $v_2 \in \ker(g(T))$

Hence $v = v_1 + v_2$ where $v_1 \in \ker(f(T))$ and $v_2 \in \ker(g(T))$

So $v \in \ker(f(T)) + \ker(g(T))$

Thus $\ker(f(T)g(T)) \subseteq \ker(f(T)) + \ker(g(T))$

Overall, we obtain
$$\ker(f(T)g(T)) = \ker(f(T)) + \ker(g(T))$$

Let's now show the sum is direct.

Suppose $v \in \ker(f(T)) \cap \ker(g(T))$

So $f(T)(v) = 0$ and $g(T)(v) = 0$

Also, from above $v = v_1 + v_2$ where

$v_1 = b(T)g(T)(v) = b(T)(0) = 0$ since $g(T)(v) = 0$

$v_2 = a(T)f(T)(v) = a(T)(0) = 0$ since $f(T)(v) = 0$

Hence $v = v_1 + v_2 = 0 + 0$.

i.e. $\ker(f(T)) \cap \ker(g(T)) = \{0\}$

Consequently, the sum $\ker(f(T)) + \ker(g(T))$ is direct and

$$\ker(f(T)g(T)) = \ker(f(T)) \oplus \ker(g(T)) \quad \text{as required.}$$

$\square$

Let's now show the link between minimal polynomials and finding enough vectors for a basis:

Primary Decomposition Theorem:

Suppose that, for a linear map $T: \mathbb{C}^n \mapsto \mathbb{C}^n$, the minimal polynomial $m_T(t)$ is one of the form

$$m_T(t) = (t - \lambda_1)^{b_1} \dots (t - \lambda_r)^{b_r}$$

where $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ are the distinct eigenvalues of $T$. Then $\mathbb{C}^n = V_{b_1}(\lambda_1) \oplus \dots \oplus V_{b_r}(\lambda_r)$

Proof

Note that since $m_T(t)$ is the minimal polynomial for $T$, $m_T(T) = 0$, so $\ker(m_T(T)) = \ker(0) = \mathbb{C}^n$

Also $\ker(m_T(T)) = \ker((T-\lambda_1 I)^{b_1}(T-\lambda_2 I)^{b_2} \dots (T-\lambda_r I)^{b_r})$

$$= \ker((T-\lambda_1 I)^{b_1}) \oplus \ker((T-\lambda_2 I)^{b_2} \dots (T-\lambda_r I)^{b_r})$$

[using earlier proposition, $(T-\lambda_1 I)^{b_1}$ and $(T-\lambda_2 I)^{b_2} \dots (T-\lambda_r I)^{b_r}$ coprime]

$$= \ker((T-\lambda_1 I)^{b_1}) \oplus \ker((T-\lambda_2 I)^{b_2}) \oplus \dots \oplus \ker((T-\lambda_r I)^{r})$$

So $\mathbb{C}^n = V_{b_1}(\lambda_1) \oplus V_{b_2}(\lambda_2) \oplus \dots \oplus V_{b_r}(\lambda_r)$

$\square$

It follows that, by "placing together" bases for $V_{b_1}(\lambda_1), \dots, V_{b_r}(\lambda_r)$ we will have a basis for $\mathbb{C}^n$, so, as required, we can find "enough" vectors for a basis of $\mathbb{C}^n$ by looking at

$$V_{b_1}(\lambda_1), \dots, V_{b_r}(\lambda_r).$$

powers arising in minimal polynomial

This is a result of the following:

Proposition

Suppose that, for a field $k$, and finite dimensional vector spaces $U, V, W$ over $k$:

$$V = U \oplus W$$

If $\{u_1, \dots, u_l\}$ is a basis for $U$,
and $\{w_1, \dots, w_m\}$ is a basis for $W$,
then $\{u_1, \dots, u_l, w_1, \dots, w_m\}$ is a basis for $V$.

MATH 2201

22-11-16

From yesterday

Primary Decomposition Theorem
For a linear map $T: \mathbb{C}^n \mapsto \mathbb{C}^n$ : if $m_T(t) = (t-\lambda_1)^{b_1} \cdots (t-\lambda_r)^{b_r}$
where $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ are distinct eigenvalues, then
$\mathbb{C}^n = V_{b_1}(\lambda_1) \oplus \cdots \oplus V_{b_r}(\lambda_r)$

---

Consequence of this related to diagonalisability:
Proposition
Let $T: \mathbb{C}^n \mapsto \mathbb{C}^n$ be a linear map:
The map $T$ is diagonalisable if and only if
$m_T(t) = (t-\lambda_1) \cdots (t-\lambda_r)$ for distinct eigenvalues
$\lambda_1, \ldots, \lambda_r \in \mathbb{C}$

Proof
Suppose that $m_T(t) = (t-\lambda_1) \cdots (t-\lambda_r) = (t-\lambda_1)^1 \cdots (t-\lambda_r)^1$.
Then by the Primary Decomposition Theorem:
$\mathbb{C}^n = V_1(\lambda_1) \oplus \cdots \oplus V_1(\lambda_r)$
So we can form a basis for $\mathbb{C}^n$ consisting of vectors that form bases for $V_1(\lambda_1), \ldots, V_1(\lambda_r)$.

But $V_1(\lambda_1), \ldots, V_1(\lambda_r)$ are eigenspaces consisting of eigenvectors (as opposed to generalised eigenvectors).
So, there exists a basis for $\mathbb{C}^n$ consisting of eigenvectors, i.e. the linear map $T$ is diagonalisable (in terms of such a basis of eigenvectors).

Now suppose that $T$ is diagonalisable.
So, there exists a basis $\{v_1, \ldots, v_n\}$ consisting of eigenvectors of $T$.
Let $f(t) = (t-\lambda_1) \cdots (t-\lambda_r)$
We wish to show that $f(t) = m_T(t)$.

Let's show that $f(T) = 0$, the zero map.

[Note: for any eigenvalue $\lambda_i$, $1 \le i \le r$:
$$f(\lambda_i) = (\lambda_i - \lambda_1) \cdots \underbrace{(\lambda_i - \lambda_i)}_{=0} \cdots (\lambda_i - \lambda_r) = 0$$ ]

Take any vector $w$ in the basis $\{v_1, \dots, v_n\}$; then, $w$ is an eigenvector of $T$ for some eigenvalue, $\lambda_i$ say.

Compute $f(T)(w)$:
$f(T)(w) = f(\lambda_i)w$ (since $w$ is an eigenvector corresponding to $\lambda_i$).

[eg. if $f(t) = t^r + a_{r-1}t^{r-1} + \dots + a_1 t + a_0$
then
$$\begin{aligned} f(T)(w) &= (T^r + a_{r-1}T^{r-1} + \dots + a_1 T + a_0 I)(w) \\ &= (T^r(w) + a_{r-1}T^{r-1}(w) + \dots + a_1 T(w) + a_0 w) && \text{by linearity} \\ &= \lambda_i^r w + a_{r-1}\lambda_i^{r-1} w + \dots + a_1 \lambda_i w + a_0 w && \text{since } w \text{ is an eigenvector for } \lambda_i \\ &= (\lambda_i^r + a_{r-1}\lambda_i^{r-1} + \dots + a_1 \lambda_i + a_0) w \\ &= f(\lambda_i) w = 0 \cdot w = 0 \end{aligned}$$ ]

This works for any $w$ in the basis $\{v_1, \dots, v_n\}$:
$$f(T)v_1 = 0, \dots, f(T)v_n = 0$$

Let's now consider a general element, $v$ say, of $\mathbb{C}^n$.
Since $\{v_1, \dots, v_n\}$ is a basis for $\mathbb{C}^n$: $v = c_1 v_1 + \dots + c_n v_n$ for $c_1, \dots, c_n \in \mathbb{C}$
Then
$$\begin{aligned} f(T)(v) &= f(T)(c_1 v_1 + \dots + c_n v_n) \\ &= c_1 f(T)(v_1) + \dots + c_n f(T)(v_n) && \text{by linearity} \\ &= c_1 \cdot 0 + \dots + c_n \cdot 0 \\ &= 0 \end{aligned}$$

So $f(T)(v) = 0$ for each $v \in \mathbb{C}^n$. Hence $f(T) = 0$, the zero map.

MATH 2201

22-11-16

So, using the definition of the minimal polynomial $m_T(t)$, we deduce that $m_T(t) \mid f(t)$ (since $f(T)=0$).

But, also, $m_T(t)$ has each eigenvalue as a root, i.e.
$m_T(t) = (t-\lambda_1)^{b_1} \dots (t-\lambda_r)^{b_r}$ for $b_1 \geq 1, \dots, b_r \geq 1$
So $f(t) \mid m_T(t)$ also.

Since $f(t) \mid m_T(t)$ and $m_T(t) \mid f(t)$, and $m_T(t)$ and $f(t)$ are both monic.
Then, using an earlier result: $m_T(t) = f(t) = (t-\lambda_1)\cdots(t-\lambda_r)$ as required. $\square$

Examples

- Let $A = \begin{pmatrix} 2 & 5 \\ 0 & 1 \end{pmatrix}$, then $ch_A(t) = (t-1)(t-2)$.

  So, only one choice for minimal polynomial, $m_A(t)$.
  $m_A(t) = (t-1)(t-2$
  Since $m_A(t)$ contains only linear factors, we conclude that $A$ is diagonalisable.

- Let $B = \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix}$, $ch_A(t) = (t-2)^2$

  Choices for $m_B(t) = t-2, (t-2)^2$
  $(B-2I) = \begin{pmatrix} 0 & 5 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$
  $(B-2I)^2 = \begin{pmatrix} 0 & 5 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 5 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ So $m_B(t) = (t-2)^2$
  Since there is a multiplicity of 2 in $m_B(t)$, $B$ is not diagnalisable.

• Let $C = \begin{pmatrix} 5 & 2 & 4 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$, as shown earlier $m_C(t) = (t-5)^2$

So $C$ is not diagonalisable, but it does have a Jordan normal form however, e.g.
$\begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ consisting of two Jordan blocks: $J_2(5)$, $J_1(5)$.

• Let $D = \begin{pmatrix} 5 & 2 & 4 \\ 0 & 5 & 3 \\ 0 & 0 & 5 \end{pmatrix}$, $ch_D(t) = (t-5)^3$

Choices for $m_D(t)$: $(t-5)$, $(t-5)^2$, $(t-5)^3$

$$(D-5I) = \begin{pmatrix} 0 & 2 & 4 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$(D-5I)^2 = \begin{pmatrix} 0 & 2 & 4 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 2 & 4 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$(D-5)^3 = \begin{pmatrix} 0 & 2 & 4 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$\Rightarrow m_D(t) = (t-5)^3$, so $D$ is not diagonalisable, but we can find the Jordan normal form.
In fact, we can already determine this form without explicity finding a Jordan basis.
To find a Jordan basis, since $m_T(t) = (t-5)^3$, we would have to go up to $V_3(5)$. We would obtain a pre-Jordan basis $B_1 \cup B_2 \cup B_3$ of $V_3(5)$ where

$B_1$ is a basis for $V_1(5)$
$B_2$ " " " " $V_2(5)$
$B_3$ " " " " $V_3(5)$

Then for the Jordan basis, we would create a "link" from $B_3$ to $B_2$ to $B_1$. So the Jordan normal form will consist of three '5's all linked: $\begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{pmatrix}$.

MATH 2201

25-11-16

Let's now see two examples involving Jordan normal forms:

Example:

Consider the (complex) linear map $T: \mathbb{C}^3 \mapsto \mathbb{C}^3$, where $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5x_1 + 2x_2 + 4x_3 \\ 5x_2 + 3x_3 \\ 5x_3 \end{pmatrix}$

Find a Jordan basis for $T$, and the corresponding Jordan normal form.

—

Consider the standard basis $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

Then $[T]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} 5 & 2 & 4 \\ 0 & 5 & 3 \\ 0 & 0 & 5 \end{pmatrix}$

Then we can use $[T]_{\mathcal{E}}^{\mathcal{E}}$ to find the characteristic and minimal polynomials of $T$:

$ch_T(t) = \det(tI_3 - [T]_{\mathcal{E}}^{\mathcal{E}}) = (t-5)^3$

Possible choices for $m_T(t)$: $(t-5)$, $(t-5)^2$, $(t-5)^3$.

$[T]_{\mathcal{E}}^{\mathcal{E}} - 5I = \begin{pmatrix} 0 & 2 & 4 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix} \neq 0$

$([T]_{\mathcal{E}}^{\mathcal{E}} - 5I)^2 = \begin{pmatrix} 0 & 0 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0$

$([T]_{\mathcal{E}}^{\mathcal{E}} - 5I)^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow m_T(t) = (t-5)^3$

(If it's the only option left, no need to calculate)

Since $m_T(t) = (t-5)^{\boxed{3}}$, we must go "up to" $V_3(5)$ to find enough vectors for a pre-Jordan or Jordan basis.

Let's determine $V_1(5)$, $V_2(5)$, $V_3(5)$:

$V_1(5) = \ker(T - 5I)$

To find this, solve $([T]_\varepsilon^\varepsilon - 5I)x = 0$, i.e. $\begin{pmatrix} 0 & 2 & 4 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

We obtain $2x_2 + 4x_3 = 0$, $x_3 = 0$, $0 = 0$,

so $x_3 = 0 \Rightarrow x_2 = 0$ ($x_1$ is a free variable)

So $V_1(5) = \left\{ \begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix} : x_1 \in \mathbb{C} \right\}$ Basis: $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$

For $V_2(5)$, solve $([T]_\varepsilon^\varepsilon - 5I)^2 x = 0$, i.e. $\begin{pmatrix} 0 & 0 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

so we obtain $6x_3 = 0$, $0 = 0$, $0 = 0$, i.e. $x_3 = 0$

($x_1$, $x_2$ are free)

So $V_2(5) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} : x_1, x_2 \in \mathbb{C} \right\}$, Basis: $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$

For a pre-Jordan basis we require a basis $\mathcal{B}_1 \cup \mathcal{B}_2$ such that $\mathcal{B}_1$ is a basis for $V_1(5)$,

$\mathcal{B}_1 \cup \mathcal{B}_2$ " " " " $V_2(5)$.

Here $\underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{\text{basis of } V_1(5)}$ already appears in $\underbrace{\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}}_{\text{basis of } V_2(5)}$, so no need to "exchange".

We may simply set $\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$, $\mathcal{B}_2 = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$

For $V_3(5)$, solve $([T]_\varepsilon^\varepsilon - 5I)^3 x = 0$, i.e. $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

Here, $x_1, x_2, x_3$ are all free:

$V_3(5) = \mathbb{C}^3 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1, x_2, x_3 \in \mathbb{C} \right\}$ Basis $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

A pre-Jordan basis is a basis $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3$ for $V_3(5)$ such that $\mathcal{B}_1$ is a basis for $V_1(5)$, $\mathcal{B}_1 \cup \mathcal{B}_2$ is a basis

MATH 2201

25-11-16

for $V_2(5)$ and $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3$ is a basis for $V_3(5)$.

The elements in the $\mathcal{B}_1$, $\mathcal{B}_2$ found above already appear in our basis for $V_3(5)$, so simply choose $\mathcal{B}_3 = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

Pre-Jordan basis: $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3$

with $\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$, $\mathcal{B}_2 = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$, $\mathcal{B}_3 = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

Let's now find a Jordan basis by creating a link going from the "top level", $\mathcal{B}_3$ dow to the "bottom level", $\mathcal{B}_1$, for each vector in the "top level".

Here there is only one vector in $\mathcal{B}_3$: $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$$\left([T]_\varepsilon^\varepsilon - 5I\right)\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 4 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix}$$ want this to appear in a basis for $V_2(5)$

Existing basis for $V_2(5)$ $\mathcal{B}_1 \cup \mathcal{B}_2$: $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$

and $\begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix} = 4\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

So $\begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix}$ is "related" to both $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

At this stage of the link, we want $\left([T]_\varepsilon^\varepsilon - 5I\right)\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix}$ to appear in the "new" $\mathcal{B}_2$, so we exchange $\begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix}$ for one of the "related" vectors in $\mathcal{B}_2 = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$.

So, exchange $\begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix}$ for $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

This leads to
$\mathcal{B}_3$ $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ linked
$\mathcal{B}_2$ $\begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix}$ ?
$\mathcal{B}_1$ $\begin{pmatrix} \vdots \end{pmatrix}$

To find a linked vector in the "new" $\mathcal{B}_1$, compute $([T]_\varepsilon^\varepsilon - 5I)\begin{pmatrix}4\\3\\0\end{pmatrix} = \begin{pmatrix}0&2&4\\0&0&3\\0&0&0\end{pmatrix}\begin{pmatrix}4\\3\\0\end{pmatrix} = \begin{pmatrix}6\\0\\0\end{pmatrix}$

In terms of $\mathcal{B}_1 = \left\{\begin{pmatrix}1\\0\\0\end{pmatrix}\right\}$ : $\begin{pmatrix}6\\0\\0\end{pmatrix} = 6\begin{pmatrix}1\\0\\0\end{pmatrix}$

So we replace $\begin{pmatrix}1\\0\\0\end{pmatrix}$ in $\mathcal{B}_1$ by $\begin{pmatrix}6\\0\\0\end{pmatrix}$

Now we have

$$\begin{array}{ll} \mathcal{B}_3 & \begin{pmatrix}0\\0\\1\end{pmatrix} \\ & \quad \text{linked} \\ \mathcal{B}_2 & \begin{pmatrix}4\\3\\0\end{pmatrix} \\ & \quad \text{linked} \\ \mathcal{B}_1 & \begin{pmatrix}6\\0\\0\end{pmatrix}. \end{array}$$

No more vectors to "link" so can now write down the Jordan basis, by writing the "linked" vectors in order, from $\mathcal{B}_1$ to $\mathcal{B}_3$.

This leads to the Jordan basis $\left\{\begin{pmatrix}6\\0\\0\end{pmatrix}, \begin{pmatrix}4\\3\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\} := \mathcal{B}$, say.

Then, $[T]_\mathcal{B}^\mathcal{B}$ is in Jordan normal form:

$$T\begin{pmatrix}6\\0\\0\end{pmatrix} = \begin{pmatrix}30\\0\\0\end{pmatrix} = 5\begin{pmatrix}6\\0\\0\end{pmatrix} + 0\begin{pmatrix}4\\3\\0\end{pmatrix} + 0\begin{pmatrix}0\\0\\1\end{pmatrix}$$

$$T\begin{pmatrix}4\\3\\0\end{pmatrix} = \begin{pmatrix}26\\15\\0\end{pmatrix} = 1\begin{pmatrix}6\\0\\0\end{pmatrix} + 5\begin{pmatrix}4\\3\\0\end{pmatrix} + 0\begin{pmatrix}0\\0\\1\end{pmatrix}$$

$$T\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}4\\3\\5\end{pmatrix} = 0\begin{pmatrix}6\\0\\0\end{pmatrix} + 1\begin{pmatrix}4\\3\\0\end{pmatrix} + 5\begin{pmatrix}0\\0\\1\end{pmatrix}$$

So $[T]_\mathcal{B}^\mathcal{B} = \begin{pmatrix}5&1&0\\0&5&1\\0&0&5\end{pmatrix}$

in Jordan normal form, as required.

MATH 2201

25-11-16

Example

Consider $C = \begin{pmatrix} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

Find a Jordan basis for $C$, and, hence, an invertible matrix $P$, and a matrix $J$ in Jordan normal form, such that $P^{-1}CP = J$.

Note:

This particular matrix appears in exercise 2 of 'Sample Exercises 3'; we will use some of the computations from the relevant solution to save time.

From that exercise: $ch_C(t) = (t-1)^2(t-2)^3$

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

Also: $(C-I) = \begin{pmatrix} 1 & 5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ $(C-2I) = \begin{pmatrix} 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}$

$(C-2I)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

To find enough vectors for a Jordan basis, need to go "up to" $V_1(1)$ and $V_2(2)$:

$V_1(1) = \ker(C-I)$, solve $\begin{pmatrix} 1 & 5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

We obtain $x_1 + 5x_2 = 0$, $x_2 = 0$, $x_4 = 0$

So $V_1(1) = \left\{ \begin{pmatrix} 0 \\ 0 \\ x_3 \\ 0 \\ x_5 \end{pmatrix} : x_3, x_5 \in \mathbb{C} \right\}$ Basis: $\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

This is already a pre-Jordan and Jordan basis for the eigenvalue 1. There is no need to go to $V_2(1)$, $V_3(1)$, ..., and then construct "linked" sets of vectors.

Let's now consider the eigenvalue 2, and find a Jordan basis:

$V_1(2) = \ker(C-2I)$ ; solve $\begin{pmatrix} 0&5&0&0&0\\0&0&0&0&0\\0&0&-1&0&0\\0&0&0&0&0\\0&0&0&0&-1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}$

from which we obtain $x_2 = 0$, $x_3 = 0$, $x_5 = 0$.

So $V_1(2) = \left\{\begin{pmatrix}x_1\\0\\0\\x_4\\0\end{pmatrix} : x_1, x_4 \in \mathbb{C}\right\}$ Basis for $V_1(2)$, $\mathcal{B}_1 = \left\{\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix}\right\}$

$V_2(2) = \ker((C-2I)^2)$ ; solve $\begin{pmatrix} 0&0&0&0&0\\0&0&0&0&0\\0&0&1&0&0\\0&0&0&0&0\\0&0&0&0&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}$

Obtain $x_3 = 0$, $x_5 = 0$.

$V_2(2) = \left\{\begin{pmatrix}x_1\\x_2\\0\\x_4\\0\end{pmatrix} : x_1, x_2, x_4 \in \mathbb{C}\right\}$ Basis: $\left\{\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix}\right\} = \mathcal{B}_2$

The basis $\mathcal{B}_1$ is already a subset of this basis, so for a pre-Jordan basis, simply set

$\mathcal{B}_1 = \left\{\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix}\right\}$ and $\mathcal{B}_2 = \left\{\begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix}\right\}$

For a Jordan basis for eigenvalue 2, link $\begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix}$, from $\mathcal{B}_2$, to a vector in the "new" $\mathcal{B}_1$.

$(C-2I)\begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix} = \begin{pmatrix} 0&5&0&0&0\\0&0&0&0&0\\0&0&-1&0&0\\0&0&0&0&0\\0&0&0&0&-1\end{pmatrix}\begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix} = \begin{pmatrix}5\\0\\0\\0\\0\end{pmatrix}$

In terms of $\mathcal{B}_1$: $\begin{pmatrix}5\\0\\0\\0\\0\end{pmatrix} = 5\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix} + 0\begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix}$

So $\begin{pmatrix}5\\0\\0\\0\\0\end{pmatrix}$ is "related" only to $\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}$, so we replace $\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}$ by $\begin{pmatrix}5\\0\\0\\0\\0\end{pmatrix}$

MATH 2201

25-11-16

Our "new" $\mathcal{B}_1$ is $\left\{\begin{pmatrix}5\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix}\right\}$, $\mathcal{B}_2 = \left\{\begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix}\right\}$

linked

A possible Jordan basis is $\left\{\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}, \begin{pmatrix}5\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix}\right\}$

Then, if $P = \begin{pmatrix}0&0&5&0&0\\0&0&0&1&0\\1&0&0&0&0\\0&0&0&0&1\\0&1&0&0&0\end{pmatrix}$,

then $\underbrace{P^{-1}CP}_{J} = \begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&2&1&0\\0&0&0&2&0\\0&0&0&0&2\end{pmatrix}$, a matrix in Jordan normal form.

We can often determine the Jordan normal form of a linear map using some key data:
Suppose, for a linear map $T$: $ch_T(t) = (t-5)^4$
Then the Jordan normal form is $\begin{pmatrix}5&?&0&0\\0&5&?&0\\0&0&5&?\\0&0&0&5\end{pmatrix}$

where '?' is 0 or 1 in each case.
If $m_T(t) = (t-5)^1$, all 4 vectors are in $\mathcal{B}_1$ and the matrix is diagonalisable.

If $m_T(t) = (t-5)^4$, the 4 vectors appear in $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \mathcal{B}_4$ then, we have a Jordan block of size $4\times 4$:
$\begin{pmatrix}5&1&0&0\\0&5&1&0\\0&0&5&1\\0&0&0&5\end{pmatrix}$.

In general, the multiplicity of an eigenvalue $\lambda$ in the <u>minimal polynomial</u> is the size of the largest block (there may be more than one such block).

If $m_T(t) = (t-5)^3$, then the 4 vectors appear in $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3$, we obtain $\begin{pmatrix}5&1&0&0\\0&5&1&0\\0&0&5&0\\0&0&0&5\end{pmatrix}$ for example.

Finally, if $m_T(t) = (t-5)^2$, then the 4 vectors are in $\mathcal{B}_1$, $\mathcal{B}_2$:

$$\begin{array}{ll} \mathcal{B}_2 & * \\ \mathcal{B}_1 & * \quad * \quad * \end{array} \qquad \begin{array}{ll} \mathcal{B}_2 & * \quad * \\ \mathcal{B}_1 & * \quad * \end{array}$$

leading to $\begin{pmatrix} 5 & 1 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$, $\begin{pmatrix} 5 & 1 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix}$ for example.

$\boxed{\dim(V_1(5)) = 3}$ $\qquad$ $\boxed{\dim(V_1(5)) = 2}$

MATH 2201

28-11-16

# Chapter 3 – Linear and bilinear forms

A *form* is a function on one or more vector spaces over a field $k$ that "returns" values in $k$ itself.

Let $k$ be a field and $V$ a (finite dimensional) vector space over $k$. A linear form over $k$ is a linear map $f: V \mapsto k$.
So $f(0) = 0$, $\forall v_1, v_2 \in V: f(v_1 + v_2) = f(v_1) + f(v_2)$,
$\forall v \in V$ and $\lambda \in k : f(\lambda v) = \lambda f(v)$.

Note:
If $V$ is a vector space over $k$, of dimension $n$ say, then $V$ is isomorphic to $k^n$.

This leads to the standard (and, in fact, only) examples of linear forms:

Example:
Let $k = \mathbb{R}$, and $V = \mathbb{R}^2$, a linear form is a linear map $f: \mathbb{R}^2 \mapsto \mathbb{R}$.
e.g. $f: \mathbb{R}^2 \mapsto \mathbb{R}$ where $f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3x_1 - 5x_2$

We can choose a basis of $\mathbb{R}^2$, in order to find a $(1 \times 2)$ matrix representing $f$:
e.g. $f\begin{pmatrix} 3 \\ 1 \end{pmatrix} = 3 \cdot 3 - 5 \cdot 1 = 4$

In terms of the standard basis $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$, the vector $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ satisfies $\begin{pmatrix} 3 \\ 1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, so $\left[\begin{pmatrix} 3 \\ 1 \end{pmatrix}\right]_{\mathcal{E}} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

Definition:

Suppose $f : k^n \to k$ is a linear form and $\mathcal{B} = \{b_1, \dots, b_n\}$ is a basis for $k^n$. Then the matrix (of size $1 \times n$) representing $f$ with respect to $\mathcal{B}$ is $[f]_{\mathcal{B}} = (f(b_1), f(b_2), \dots, f(b_n))$

(Just as for a vector $v$ in $k^n$, $[v]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$

where $v = \lambda_1 b_1 + \dots + \lambda_n b_n$ )

In our example $f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3x_1 - 5x_2$.

So $f\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 3$, $f\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -5$

and $[f]_{\mathcal{E}} = (3, -5)$

Then, in terms of $\mathcal{E}$; we can find $f(v)$, where $v = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, using $[f]_{\mathcal{E}} [v]_{\mathcal{E}} = (3, -5)\begin{pmatrix} 3 \\ 1 \end{pmatrix} = 4$. ✓

Note:

For a given linear form $f$, and vector $v$, the value of $f(v)$ is the same independently of the basis we choose.

e.g. consider the basis $\mathcal{B} = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$

Then $v = \begin{pmatrix} 3 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

So $[v]_{\mathcal{B}} = \left[\begin{pmatrix} 3 \\ 1 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

Also $f\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 3$ and $f\begin{pmatrix} 1 \\ 1 \end{pmatrix} = -2$. So $[f]_{\mathcal{B}} = (3, -2)$, and: $[f]_{\mathcal{B}} [v]_{\mathcal{B}} = (3, -2)\begin{pmatrix} 2 \\ 1 \end{pmatrix} = 4$. ✓

MATH 2201

28-11-16

Let's now consider bilinear forms:

- Given a field $k$ and a vector space $V$ over $k$, a bilinear form is a function $f: V \times V \longrightarrow k$ satisfying:

  - $f(0, v) = 0$ and $f(v, 0)$ for all $v$ in $V$.

  - For all $v_1, v_2, w$ in $V$: $f(v_1 + v_2, w) = f(v_1, w) + f(v_2, w)$, also $f(w, v_1 + v_2) = f(w, v_1) + f(w, v_2)$.

  - For all $v, w \in V$ and any $\lambda \in k$: $f(\lambda v, w) = \lambda f(v, w)$ and $f(v, \lambda w) = \lambda f(v, w)$.

e.g. consider $f: \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$, where
$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 + 3x_1 y_2 - 2x_2 y_1 + 7x_2 y_2.$$

We may verify that this is a bilinear form.

But, we can also represent $f$ using a matrix:
$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = (x_1, x_2)\begin{pmatrix} 1 & 3 \\ -2 & 7 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

So $f(x, y) = x^T \begin{pmatrix} 1 & 3 \\ -2 & 7 \end{pmatrix} y$

transpose of $x$

Check: $$(x_1, x_2)\begin{pmatrix} 1 & 3 \\ -2 & 7 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = (x_1, x_2)\begin{pmatrix} y_1 + 3y_2 \\ -2y_1 + 7y_2 \end{pmatrix}$$
$$= x_1 y_1 + 3x_1 y_2 - 2x_2 y_1 + 7x_2 y_2 \checkmark$$

Special type of bilinear form

- A symmetric bilinear form is a bilinear form $f: V \times V \longrightarrow k$, such that for all $x, y$ in $V$, $f(x, y) = f(y, x)$.

Symmetric bilinear form corresponds to symmetric matrices:

e.g. if $f(x,y) = f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = (x_1, x_2)\begin{pmatrix} 1 & 2 \\ 2 & 7 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

$$\left.\begin{aligned} f(x,y) &= x_1 y_1 + 2x_1 y_2 + 2x_2 y_1 + 7x_2 y_2 \\ f(y,x) &= y_1 x_1 + 2y_1 x_2 + 2y_2 x_1 + 7y_2 x_2 \end{aligned}\right\} \text{equal}$$

So $f$ is a symmetric bilinear form.

To row reduce symmetric matrices, while still maintaining their relevance in a bilinear form, we apply pairs of row and (corresponding) column operations:

e.g. let's reduce $\begin{pmatrix} 1 & 2 \\ 2 & 7 \end{pmatrix}$ in such a way.

$$\begin{pmatrix} 1 & 2 \\ 2 & 7 \end{pmatrix} \xrightarrow{R_2 \to R_2 - R_1} \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \xrightarrow{C_2 \to C_2 - 2C_1} \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \xrightarrow{R_2 \to \frac{1}{\sqrt{3}} R_2} \begin{pmatrix} 1 & 0 \\ 0 & \frac{3}{\sqrt{3}} \end{pmatrix} \xrightarrow{C_2 \to \frac{1}{\sqrt{3}} C_2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \checkmark$$

But we cannot reduce $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ (say) over the real numbers $(i^2 = -1)$.

MATH 2201

29-11-16

## Duality on linear forms

Suppose that $V$ is a vectorspace over a field $k$. Then, the set of all associated linear forms is
$$V^* = \{f : V \mapsto k \,;\, f \text{ is linear}\}.$$

Note $V^*$ is also a vector space over $k$, where:
- for $f, g \in V^*$: $(f+g)(v) = f(v) + g(v) \quad \forall v \in V$
- for $f \in V^*$, $\lambda \in k$: $(\lambda f)(v) = \lambda(f(v)) \quad \forall v \in V$.

eg. • $$\underbrace{\begin{pmatrix}1 & 2\end{pmatrix}\begin{pmatrix}x_1 \\ x_2\end{pmatrix}}_{f(v)} \underbrace{+}_{+} \underbrace{\begin{pmatrix}1 & -5\end{pmatrix}\begin{pmatrix}x_1 \\ x_2\end{pmatrix}}_{g(v)} \underbrace{=}_{=} \underbrace{\begin{pmatrix}2 & -3\end{pmatrix}\begin{pmatrix}x_1 \\ x_2\end{pmatrix}}_{(f+g)(v)}$$

where $(1 \;\; 2) + (1 \;\; -5) = (2 \;\; -3)$

• $$\underbrace{3\left(\begin{pmatrix}1 & 2\end{pmatrix}\begin{pmatrix}x_1 \\ x_2\end{pmatrix}\right)}_{\lambda \times f(v)} = \underbrace{\begin{pmatrix}3 & 6\end{pmatrix}\begin{pmatrix}x_1 \\ x_2\end{pmatrix}}_{(\lambda f)(v)}$$

We try to indicate a correspondence between $V$ and $V^*$ (here, we assume that $V$ has finite dimension, $n$ say):
Let $\{e_1, \dots, e_n\}$ be a basis for $V$.
Consider the set $\{f_1, \dots, f_n\}$ in $V^*$, where $f_1, \dots, f_n$ are defined as follows

$$f_1(e_1) = 1, \; f_1(e_2) = 0, \; f_1(e_3) = 0, \; \dots, \; f_1(e_n) = 0$$
$$f_2(e_1) = 0, \; f_2(e_2) = 1, \; f_2(e_3) = 0, \; \dots, \; f_2(e_n) = 0$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$
$$f_n(e_1) = 0, \; \dots \quad \dots, \; f_n(e_{n-1}) = 0, \; f_n(e_n) = 1$$

So $$f_i(e_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$
$$= \delta_{ij}$$

Example

Let $k = \mathbb{R}$, $V = \mathbb{R}^2$

Basis for $V$: $\left\{ \underset{e_1}{\begin{pmatrix}1\\0\end{pmatrix}}, \underset{e_2}{\begin{pmatrix}0\\1\end{pmatrix}} \right\}$

Then we should have

$$f_1\begin{pmatrix}1\\0\end{pmatrix} = 1,\ f_1\begin{pmatrix}0\\1\end{pmatrix} = 0 \Rightarrow f_1 \text{ is } (1\ 0) = e_1^T$$

$$f_2\begin{pmatrix}1\\0\end{pmatrix} = 0,\ f_2\begin{pmatrix}0\\1\end{pmatrix} = 1 \Rightarrow f_2 \text{ is } (0\ 1) = e_2^T$$

The set $\{f_1, \dots, f_n\}$ is a dual set, in $V^*$, to the set $\{e_1, \dots, e_n\}$ in $V$.

To show that $V^*$ is isomorphic to $V$, we show that, just as $\{e_1, \dots, e_n\}$ is a basis for $V$, $\{f_1, \dots, f_n\}$ is a basis for $V^*$.

$\{f_1, \dots, f_n\}$ spans $V^*$: Let's show this

Consider a general linear form, $f: V \mapsto k$, and a general vector $v$ in $V$. Then, since $\{e_1, \dots, e_n\}$ is a basis for $V$: $v = \lambda_1 e_1 + \dots + \lambda_n e_n$ for $\lambda_1, \dots, \lambda_n \in k$.

Then $f(v) = f(\lambda_1 e_1 + \dots + \lambda_n e_n)$
$= \lambda_1 f(e_1) + \dots + \lambda_n f(e_n)$ by linearity of $f$.

Set $a_1 = f(e_1), \dots, a_n = f(e_n)$ so that $f(v) = \lambda_1 a_1 + \dots + \lambda_n a_n$.

We want to show that $f = a_1 f_1 + \dots + a_n f_n$:

$(a_1 f_1 + \dots + a_n f_n)(v) = a_1 f_1(v) + \dots + a_n f_n(v)$ by linearity

$= a_1 f_1(\lambda_1 e_1 + \dots + \lambda_n e_n) + \dots + a_n f_n(\lambda_1 e_1 + \dots + \lambda_n e_n)$

$= a_1(\lambda_1 f_1(e_1) + \cancel{\dots} + \cancel{\lambda_n f_1(e_n)}) + \dots + a_n(\cancel{\lambda_1 f_n(e_1)} + \cancel{\dots} + \lambda_n f_n(e_n))$

$f_i(e_j) = \delta_{ij}$

$= a_1 \lambda_1 f_1(e_1) + a_2 \lambda_2 f_2(e_2) + \dots + a_n \lambda_n f_n(e_n)$

$= \lambda_1 a_1 + \dots + \lambda_n a_n = F(v)$

MATH 2201

29-11-16

So, for all $v \in V$: $f(v) = (a_1 f_1 + \dots + a_n f_n)(v)$
i.e. $f = a_1 f_1 + \dots + a_n f_n$.
So $\{f_1, \dots, f_n\}$ spans $V^*$.

<u>$\{f_1, \dots, f_n\}$ is linearly independent:</u> Let's show this
Suppose that, for $a_1, \dots, a_n \in k$: $a_1 f_1 + \dots + a_n f_n = 0$.
i.e. $(a_1 f_1 + \dots + a_n f_n)(v) = 0$ for each $v$ in $V$.

Set $v = e_1$: $(a_1 f_1 + \dots + a_n f_n)(e_1) = 0$
So $a_1 f_1(e_1) + \dots + a_n f_n(e_1) = 0$
i.e. $a_1 \cdot 1 + a_2 \cdot 0 + \dots + a_n \cdot 0 = 0$
$\Rightarrow a_1 = 0$.

In general, set $v = e_i$: $(a_1 f_1 + \dots + a_n f_n)(e_i) = 0$
Then $a_1 f_1(e_i) + \dots + a_i f_i(e_i) + \dots + a_n f_n(e_i) = 0$
Then $a_i = 0$. This works for each $i = 1, \dots, n$.

So $a_1 f_1 + \dots + a_n f_n = 0 \Rightarrow a_1 = 0, \dots, a_n = 0$
Thus $\{f_1, \dots, f_n\}$ is a linearly independent set.

---

This leads to the result that, if $V$ is a finite dimensional vector space over a field $k$, then $V \cong V^*$.

<u>Back to bilinear forms</u>
How do we represent bilinear forms using matrices? :
Given a vector space $V$ over a field, with a basis $\mathcal{B} = \{b_1, \dots, b_n\}$ for $V$, the matrix representing a bilinear form $f: V \times V \longrightarrow k$ with respect to $\mathcal{B}$ is

$$[f]_{\mathcal{B}}^{\mathcal{B}} = [f]_{\mathcal{B}} = \begin{pmatrix} f(b_1, b_1) & f(b_1, b_2) & \cdots & f(b_1, b_n) \\ \vdots & \vdots & & \vdots \\ f(b_n, b_1) & f(b_n, b_2) & \cdots & f(b_n, b_n) \end{pmatrix}.$$

Suppose $k = \mathbb{R}$, $V = \mathbb{R}^2$, and let $f : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ be defined by

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 - x_1 y_2 + 2 x_2 y_1 + 5 y_1 y_2$$

Let's find $[f]_{\mathcal{E}}$, the matrix of $f$ with respect to the basis $\mathcal{E}$, where $\mathcal{E} = \left\{ \underset{e_1}{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}, \underset{e_2}{\begin{pmatrix} 0 \\ 1 \end{pmatrix}} \right\}$.

Then $f(e_1, e_1) = f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = 1$

Similarly $f(e_1, e_2) = -1$, $f(e_2, e_1) = 2$, $f(e_2, e_2) = 5$

So $[f]_{\mathcal{E}} = \begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} f(e_1, e_1) & f(e_1, e_2) \\ f(e_2, e_1) & f(e_2, e_2) \end{pmatrix}$

MATH 2201

02-12-16

From last time:

Bilinear form $f: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 - x_1 y_2 + 2x_2 y_1 + 5x_2 y_1$$

With respect to standard basis $\mathcal{E} = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$

$$[f]_{\mathcal{E}}^{\mathcal{E}} = [f]_{\mathcal{E}} = \begin{pmatrix} f(e_1, e_1) & f(e_1, e_2) \\ f(e_2, e_1) & f(e_2, e_2) \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix}.$$

Let $v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, with respect to $\mathcal{E}$: $\begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1\begin{pmatrix} 0 \\ 1 \end{pmatrix}$,

so $[v]_{\mathcal{E}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Similarly, if $w = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, then $[w]_{\mathcal{E}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

In this case $f(v, w) = f\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = (2)(0) - (2)(1) + 2(1)(0) + 5(1)(3)$
$= 3$

transpose

Also $([v]_{\mathcal{E}})^T [f]_{\mathcal{E}}^{\mathcal{E}} [w]_{\mathcal{E}} = (2, 1)\begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = (4, 3)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 3$

So $f(v, w) = ([v]_{\mathcal{E}})^T [f]_{\mathcal{E}}^{\mathcal{E}} [w]_{\mathcal{E}}$

In fact, this holds for any suitable basis.

Consider the basis $\mathcal{B} = \left\{\underset{b_1}{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}, \underset{b_2}{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}\right\}$

Find "change of basis matrices" between $\mathcal{E}$ and $\mathcal{B}$.

Express $\mathcal{B}$ in terms of $\mathcal{E}$, using matrix $[Id]_{\mathcal{B}}^{\mathcal{E}}$.

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

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

So $[Id]_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

Also, express $\mathcal{E}$ in terms of $\mathcal{B}$:

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

So $[Id]_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix}1 & -1\\0 & 1\end{pmatrix}$.

Note: $[A]_{\mathcal{E}}^{\mathcal{B}} = \left([A]_{\mathcal{B}}^{\mathcal{E}}\right)^{-1}$

$v = \begin{pmatrix}2\\1\end{pmatrix} = 1\begin{pmatrix}1\\0\end{pmatrix} + 1\begin{pmatrix}1\\1\end{pmatrix}$ So $[v]_{\mathcal{B}} = \begin{pmatrix}1\\1\end{pmatrix}$

$w = \begin{pmatrix}0\\1\end{pmatrix} = -1\begin{pmatrix}1\\0\end{pmatrix} + 1\begin{pmatrix}1\\1\end{pmatrix}$ So $[w]_{\mathcal{B}} = \begin{pmatrix}-1\\1\end{pmatrix}$

Note:

$$[v]_{\mathcal{B}} = [Id]_{\mathcal{E}}^{\mathcal{B}}[v]_{\mathcal{E}} = \begin{pmatrix}1 & -1\\0 & 1\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}$$
$$[w]_{\mathcal{B}} = [Id]_{\mathcal{E}}^{\mathcal{B}}[w]_{\mathcal{E}} = \begin{pmatrix}1 & -1\\0 & 1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-1\\1\end{pmatrix}$$

Also $[f]_{\mathcal{B}}^{\mathcal{B}} = [f]_{\mathcal{B}} = \begin{pmatrix}f(b_1, b_1) & f(b_1, b_2)\\ f(b_2, b_1) & f(b_2, b_2)\end{pmatrix}$

$$= \begin{pmatrix}1 & 0\\3 & 7\end{pmatrix}$$

computed using the definition of $f$, e.g:

$$f(b_1, b_1) = f\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right) = (1)(1) - (1)(0) + 2(0,1) + 5(0,0) = 1$$

Then $([v]_{\mathcal{B}})^T[f]_{\mathcal{B}}^{\mathcal{B}}[w]_{\mathcal{B}} = (1, 1)\begin{pmatrix}1 & 0\\3 & 7\end{pmatrix}\begin{pmatrix}-1\\1\end{pmatrix} = (4, 7)\begin{pmatrix}-1\\1\end{pmatrix} = 3$

So $f(v,w) = ([v]_{\mathcal{E}})^T[f]_{\mathcal{E}}^{\mathcal{E}}[w]_{\mathcal{E}} = ([v]_{\mathcal{B}})^T[f]_{\mathcal{B}}^{\mathcal{B}}[w]_{\mathcal{B}} (= 3)$.

MATH 2201

02-11-16

In general:

Proposition

Let $k$ be a field, and $V$ a vector space of dimension $n$ over $k$, with a basis $\mathcal{B} = \{b_1, \dots, b_n\}$. Then, a bilinear form $f: V \times V \to k$ is uniquely determined by $[f]_{\mathcal{B}}^{\mathcal{B}}$; for all vectors $v, w$ in $V$:

$$f(v,w) = ([v]_{\mathcal{B}})^T [f]_{\mathcal{B}}^{\mathcal{B}} [w]_{\mathcal{B}}$$

Proof:

Suppose that, in terms of $\mathcal{B}$:

$v = x_1 b_1 + \dots + x_n b_n$ for $x_1, \dots, x_n \in k$

$w = y_1 b_1 + \dots + y_n b_n$ for $y_1, \dots, y_n \in k$

Then $[v]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $[w]_{\mathcal{B}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, $[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} f(b_1, b_1) & \cdots & f(b_1, b_n) \\ \vdots & & \vdots \\ f(b_n, b_1) & \cdots & f(b_n, b_n) \end{pmatrix}$

So $$([v]_{\mathcal{B}})^T [f]_{\mathcal{B}}^{\mathcal{B}} [w]_{\mathcal{B}} = (x_1, \dots, x_n) \begin{pmatrix} f(b_1, b_1) & \cdots & f(b_1, b_n) \\ \vdots & & \vdots \\ f(b_n, b_1) & \cdots & f(b_n, b_n) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= \left(x_1 f(b_1, b_1) + \dots + x_n f(b_n, b_1), \dots, x_1 f(b_1, b_n) + \dots + x_n f(b_n, b_1)\right) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= x_1 y_1 f(b_1, b_1) + \dots + x_n y_1 f(b_n, b_1)$$
$$+ \dots$$
$$+ x_1 y_n f(b_1, b_n) + \dots + x_n y_n f(b_n, b_n)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j f(b_i, b_j)$$

Also

$$\left.\begin{aligned} f(v,w) &= f(x_1 b_1 + \dots + x_n b_n, w) \\ &= f(x_1 b_1, w) + \dots + f(x_n b_n, w) \\ &= x_1 f(b_1, w) + \dots + x_n f(b_n, w) \end{aligned}\right\} \text{using definition of a bilinear map.}$$

So $f(v,w) = x_1 f(b_1, y_1 b_1 + \dots + y_n b_n) + \dots + x_n f(b_n, y_1 b_1 + \dots + y_n b_n)$

$$= x_1 f(b_1, y_1 b_1) + \dots + x_1 f(b_1, y_n b_n)$$
$$+ \dots$$
$$+ x_n f(b_n, y_1 b_1) + \dots + x_n f(b_n, y_n b_n)$$
$$= x_1 y_1 f(b_1, b_1) + \dots + x_1 y_n f(b_1, b_n) + \dots$$
$$+ x_n y_1 f(b_n, b_1) + \dots + x_n y_n f(b_n, b_n)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j f(b_i, b_j)$$

So $f(v,w) = ([v]_{\mathcal{B}})^T [f]_{\mathcal{B}}^{\mathcal{B}} [w]_{\mathcal{B}}$ $\square$

<u>Consequence of this:</u>

For any two bases, $\mathcal{B}$ and $\mathcal{C}$, of $V$:

$$f(v,w) = ([v]_{\mathcal{B}})^T [f]_{\mathcal{B}}^{\mathcal{B}} [w]_{\mathcal{B}}$$

$$f(v,w) = ([v]_{\mathcal{C}})^T [f]_{\mathcal{C}}^{\mathcal{C}} [w]_{\mathcal{C}}$$

So, as in the earlier example:

$$([v]_{\mathcal{B}})^T [f]_{\mathcal{B}}^{\mathcal{B}} [w]_{\mathcal{B}} = ([v]_{\mathcal{C}})^T [f]_{\mathcal{C}}^{\mathcal{C}} [w]_{\mathcal{C}} \quad (*)$$

This allows us to find a formula linking $[f]_{\mathcal{C}}^{\mathcal{C}}$ and $[f]_{\mathcal{B}}^{\mathcal{B}}$.

$$[w]_{\mathcal{B}} = [Id]_{\mathcal{C}}^{\mathcal{B}} [w]_{\mathcal{C}} \quad , \quad [v]_{\mathcal{B}} = [Id]_{\mathcal{C}}^{\mathcal{B}} [v]_{\mathcal{C}}$$

$$([v]_{\mathcal{C}})^T [f]_{\mathcal{C}}^{\mathcal{C}} [w]_{\mathcal{C}} = ([v]_{\mathcal{B}})^T [f]_{\mathcal{B}}^{\mathcal{B}} [w]_{\mathcal{B}}$$
$$= ([Id]_{\mathcal{C}}^{\mathcal{B}} [v]_{\mathcal{C}})^T [f]_{\mathcal{B}}^{\mathcal{B}} [Id]_{\mathcal{C}}^{\mathcal{B}} [w]_{\mathcal{C}}$$

$$= ([v]_{\mathcal{C}})^T ([Id]_{\mathcal{C}}^{\mathcal{B}})^T [f]_{\mathcal{B}}^{\mathcal{B}} [Id]_{\mathcal{C}}^{\mathcal{B}} [w]_{\mathcal{C}}$$

$$= ([v]_{\mathcal{C}})^T [f]_{\mathcal{C}}^{\mathcal{C}} [w]_{\mathcal{C}}$$

So $[f]_{\mathcal{C}}^{\mathcal{C}} = M^T [f]_{\mathcal{B}}^{\mathcal{B}} M$ where $M = [Id]_{\mathcal{C}}^{\mathcal{B}}$ is an invertible matrix.

MATH 2201

02-12-16

We say that two matrices $A$, $B$ representing bilinear forms, are equivalent if there exists an invertible matrix $M$ such that

$$B = M^T A M.$$

Example

Returning to the bilinear form from earlier:

$$[f]_\mathcal{E}^\mathcal{E} = \begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix}, \quad [f]_\mathcal{B}^\mathcal{B} = \begin{pmatrix} 1 & 0 \\ 3 & 7 \end{pmatrix}$$

$$[Id]_\mathcal{E}^\mathcal{B} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Then $$([Id]_\mathcal{E}^\mathcal{B})^T [f]_\mathcal{B}^\mathcal{B} [Id]_\mathcal{E}^\mathcal{B} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 3 & 7 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix} = [f]_\mathcal{E}^\mathcal{E}$$

---

Let's now concentrate on symmetric bilinear forms: bilinear form satisfying $f(v, w) = f(w, v) \quad \forall v, w.$

---

eg. the bilinear form $f: \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ given by

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 + 2x_1 y_2 + 2x_2 y_1 + 5x_2 y_2$$

is symmetric:

$$f\left(\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = y_1 x_1 + 2y_1 x_2 + 2y_2 x_1 + 5y_2 x_2$$
$$= f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right)$$

Note:

Symmetric bilinear forms correspond to symmetric matrices e.g. here $[f]_\mathcal{E}^\mathcal{E} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$ for basis $\mathcal{E} = \left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$

a symmetric matrix: $([f]_\mathcal{E}^\mathcal{E})^T = [f]_\mathcal{E}^\mathcal{E}$

or $([f]_\mathcal{E}^\mathcal{E})_{ji} = ([f]_\mathcal{E}^\mathcal{E})_{ij}$ for all $i, j$

To each symmetric bilinear form $f: V \times V \mapsto k$ we can associate a quadratic form $q: V \to k$ where for $v$ in $V$: $q(v) = f(v,v)$.

For example: to the bilinear form $f: \mathbb{R}^2 \times \mathbb{R}^2 \mapsto \mathbb{R}$, from above we associate the following quadratic form:
$q: \mathbb{R}^2 \to \mathbb{R}$
where $q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = x_1^2 + 2x_1x_2 + 2x_2x_1 + 5x_2^2$

So $q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 4x_1x_2 + 5x_2^2$

or $q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1 \; x_2)\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

Quadratic forms correspond to ways of measuring distances, e.g. a quadratic form $q: \mathbb{R}^2 \mapsto \mathbb{R}$ corresponding to the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is defined by:
$q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1 \; x_2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + x_2^2$
while bilinear forms correspond to ways of defining "inner products" (dot products):
$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = (x_1, x_2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x_1y_1 + x_2y_2$$

<u>Note</u>:
The same quadratic form can be obtained from more than one bilinear form:
e.g. $(x_1 \; x_2)\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 4x_1x_2 + 5x_2^2$ from above

but also $(x_1 \; x_2)\begin{pmatrix} 1 & 7 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 4x_1x_2 + 5x_2^2$

but it corresponds to a unique <u>symmetric</u> bilinear form.

MATH 2201

02-12-16

Given a bilinear form $f: V \times V \to k$, can define a quadratic form $q: V \to k$ on all values of $V$ using:
$$q(v) = f(v,v).$$

But, given a quadratic form $q: V \to k$, does this, on its own, completely define a particular symmetric bilinear form $f: V \times V \to k$?

YES! see below ↓

For all $v, w \in V$: try to define $f(v,w)$ in terms of $q$.

Consider
$$\begin{aligned} q(v+w) &= f(v+w, v+w) \\ &= f(v, v+w) + f(w, v+w) \\ &= f(v,v) + f(v,w) + f(w,v) + f(w,w) \\ &= f(v,v) + 2f(v,w) + f(w,w) \\ &\qquad \text{(since } f \text{ is assumed to be symmetric)} \\ &= q(v) + 2f(v,w) + q(w) \end{aligned}$$

So $f(v,w) = \frac{1}{2}\left(q(v+w) - q(v) - q(w)\right)$

This defines $f(v,w)$ $(\forall v, w)$ in terms of $q$.

This works in any field $k$, as long as $2$ is invertible (so we can write '$\frac{1}{2}$').
i.e. it works in any field where $2 \neq 0$.

Consider $\mathbb{F}_2 = \{0, 1\}$ (field with 2 elements, where $1+1=0$.
Then the two matrices $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ lead to the same quadratic form:
$$(x_1 \ \ x_2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + x_2^2$$
and
$$(x_1 \ \ x_2)\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + (1+1)x_1x_2 + x_2^2$$
$$= x_1^2 + x_2^2$$

Aim:
Given a symmetric bilinear form $f: V \times V \mapsto k$ identify a basis $\mathcal{B}$ of $V$ such that the matrix $[f]_{\mathcal{B}}$ is a diagonal matrix.

Suppose that $f: \mathbb{R}^2 \times \mathbb{R}^2 \mapsto \mathbb{R}$, where $[f]_{\mathcal{E}} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$.
where $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$.

Now consider the basis $\mathcal{B} = \left\{ \underset{b_1}{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}, \underset{b_2}{\begin{pmatrix} -4 \\ 2 \end{pmatrix}} \right\}$

Then $[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} f(b_1, b_1) & f(b_1, b_2) \\ f(b_2, b_1) & f(b_2, b_2) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$

where e.g. $f(b_1, b_2) = b_1^T [f]_{\mathcal{E}}^{\mathcal{E}} b_2 = (1, 0) \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} -4 \\ 2 \end{pmatrix} = (1, 2) \begin{pmatrix} -4 \\ 2 \end{pmatrix} = 0$

So with respect to $\mathcal{B}$ : $[f]_{\mathcal{B}}^{\mathcal{B}}$ is a diagonal matrix.

Here $f(b_1, b_1) = 1$, $f(b_1, b_2) = 0$, $f(b_2, b_1) = 0$, $f(b_2, b_2) = 4$.
This is an example of an orthogonal basis:

- Vectors $v, w$ are orthogonal with respect to a symmetric bilinear form $f$ if $f(v, w) = 0$.

- A set $\{v_1, \dots, v_n\}$ is orthogonal with respect to $f$ if $f(v_i, v_j) = 0$ for all $i, j$, $i \neq j$.

MATH 2201

05-12-16

Previously on MATH 2201

Considered symmetric bilinear form $f : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ given by $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_2 + 2x_1 y_2 + 2x_2 y_1 + 5x_2 y_2$

w.r.t. $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$, $[f]_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$

w.r.t $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ 2 \end{pmatrix} \right\}$, $[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$

Note that w.r.t. $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\} = \mathcal{C}$, $[f]_{\mathcal{C}}^{\mathcal{C}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Today: Sequel to row reduction from MATH 2201.
"This time, it's bilinear"

We will use double operations: row / column.

Notation:

- $E_r(i,j;\lambda)$ is the matrix corresponding to Row $i \to$ Row $i + \lambda$ Row $j$

$E_r(i,j;\lambda)$ is the same as the identity matrix, but with an entry of $\lambda$ in row $i$, column $j$

eg. on a $2 \times 2$ matrix: $E_r(1,2;2) = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$

Multiplying on the left by $E_r(1,2;2)$ "does" the corresponding operation:

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 3 & 4 \end{pmatrix}$$

To obtain the corresponding column operation, we multiply by the transpose matrix on the right:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 11 & 4 \end{pmatrix}$$

- Multiplying by $E_c(i,j;\lambda) = E_r(i,j;\lambda)^T$ on the right performs the corresponding column operation.

- $D_r(i;\lambda)$ corresponds to Row $i \to \lambda$ Row $i$, $\lambda \neq 0$
- $D_c(i,\lambda) = D_r(i,\lambda)^T$ corresponds to Column $i \to \lambda$ Column $i$

- $P_r(i,j)$ corresponds to Row $i \leftrightarrow$ Row $j$
- $P_c(i,j) = P_r(i,j)^T$ corresponds to Column $i \leftrightarrow$ Column $j$

Let's see a reduction of $\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} = [f]_{\mathcal{E}}^{\mathcal{E}}$ in light of this:

$$\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \xrightarrow{\mathcal{E}_r(2,1;-2)} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \xrightarrow{\mathcal{E}_c(2,1;-2)} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$E_r(2,1;-2) = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}, \quad E_c(2,1;-2) = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}.$$

This reduction may be expressed as the following matrix equation:

$$\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \underbrace{\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}}_{[f]_{\mathcal{E}}^{\mathcal{E}}} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{[f]_{\mathcal{C}}^{\mathcal{C}}}$$

The columns of the final matrix on the right of $[f]_{\mathcal{E}}^{\mathcal{E}}$ "reveal" the basis $\mathcal{C}$ : $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$

This agrees with

$$\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\left[ [Id]_{\mathcal{C}}^{\mathcal{E}} \right]^T [f]_{\mathcal{E}}^{\mathcal{E}} \underbrace{[Id]_{\mathcal{C}}^{\mathcal{E}}}_{} = [f]_{\mathcal{C}}^{\mathcal{C}}$$

colums of this form basis $\mathcal{C}$

In general, we can always reduce a symmetric matrix in this way

MATH 2201

05-12-16

Examples

1). $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \xrightarrow{\mathcal{E}_r(2,1;-2)} \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \xrightarrow{\mathcal{E}_c(2,1;-2)} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

most reduced form

2). $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \xrightarrow{\mathcal{E}_r(2,1;-2)} \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \xrightarrow{\mathcal{E}_c(2,1;-2)} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

most reduced form over $\mathbb{R}$

But over $\mathbb{C}$ we can do the following:

$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \xrightarrow{\mathcal{D}_r(2;i)} \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix} \xrightarrow{\mathcal{D}_c(2;i)} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

which is the most reduced form over $\mathbb{C}$.

3). $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

Try permutations:

$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \xrightarrow{\rho_r(1,2)} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \xrightarrow{\rho_c(1,2)} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ !!

Instead use 'E' type operations

$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \xrightarrow{\mathcal{E}_r(1,2;\frac{1}{2})} \begin{pmatrix} \frac{1}{2} & 1 \\ 1 & 0 \end{pmatrix} \xrightarrow{\mathcal{E}_c(1,2;\frac{1}{2})} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$

MATH 2201

06-12-16

From yesterday:

Examples of reducing matrices corresponding to symmetric bilinear forms or quadratic forms.

In general, we say that two matrices, A, B say, are congruent if there exists an invertible matrix M such that $B = M^T A M$.

For example, any two matrices representing the same bilinear form are congruent.

$$[f]_B = \left([Id]_B^A\right)^T [f]_A \, [Id]_B^A$$
$$B = M^T \quad A \quad M$$

Two quadratic forms are equivalent if we can obtain one from the other using a change of basis.

e.g. as we have seen

$\begin{pmatrix}1&2\\2&3\end{pmatrix}$ reduces to $\begin{pmatrix}1&0\\0&-1\end{pmatrix}$

so the bilinear form given by $f\left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}\right) = x_1y_2 + 2x_1y_2 + 2x_2y_1 + 3x_2y_2$

is equivalent to $g\left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}\right) = x_1y_1 - x_2y_2$.

OR, equivalently, the quadratic form given by $q\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1^2 + 4x_1x_2 + 3x^2$ is equivalent to the form given by $q'\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1^2 - x_2^2$

Similarly: $\begin{pmatrix}1&2\\2&4\end{pmatrix}$ reduces to $\begin{pmatrix}1&0\\0&0\end{pmatrix}$

So $q : \mathbb{R}^2 \mapsto \mathbb{R}$, $q\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1^2 + 4x_1x_2 + 4x_2^2$ is equivalent to $q' : \mathbb{R}^2 \mapsto \mathbb{R}$, $q'\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1^2$

And $\begin{pmatrix}1&2\\2&5\end{pmatrix}$ reduces to $\begin{pmatrix}1&0\\0&1\end{pmatrix}$

So $q: \mathbb{R}^2 \mapsto \mathbb{R}$ , $q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 4x_1x_2 + 5x_2^2$
is equivalent to $q': \mathbb{R}^2 \mapsto \mathbb{R}$ , $q'\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + x_2^2$

In general, given a quadratic form $q: \mathbb{R}^2 \mapsto \mathbb{R}$
there is a basis $\mathcal{B}$ of $\mathbb{R}^n$ such that

$$[q]_{\mathcal{B}}^{\mathcal{B}} = [q]_{\mathcal{B}} = \begin{pmatrix} I_r & & 0 \\ & -I_s & \\ 0 & & 0_{n-r-s} \end{pmatrix}$$ (can be obtained through reduction, for example)

where
$I_r = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & \ddots & 1 \end{pmatrix}$ size $= r \times r$

$-I_s = \begin{pmatrix} -1 & & 0 \\ & -1 & \\ 0 & \ddots & -1 \end{pmatrix}$ size $= s \times s$

$0_{n-r-s} = \begin{pmatrix} 0 & & 0 \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$ size $= (n-r-s) \times (n-r-s)$

This is the <u>real canonical form</u> of $q$.
The number of nonzero rows is the <u>rank</u> of $q$.
Here $\operatorname{rank}(q) = r+s$.

Suppose $q: \mathbb{C}^n \mapsto \mathbb{C}$ is a complex quadratic form.
Then there is a basis $\mathcal{B}$ of $\mathbb{C}^n$ such that

$$[q]_{\mathcal{B}}^{\mathcal{B}} = [q]_{\mathcal{B}} = \begin{pmatrix} I_a & \\ & 0_{n-a} \end{pmatrix}$$ where $\operatorname{rank}(q) = a$

This is the <u>complex canonical form</u> of $q$

MATH 2201

06-12-16

Examples of determining canonical forms over $\mathbb{R}$ and $\mathbb{C}$, and corresponding bases:

Consider the bilinear form $f: \mathbb{R}^2 \times \mathbb{R}^2 \mapsto \mathbb{R}$ given by $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_2 + x_2 y_1$

Then $[f]_\mathcal{E} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ , $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$

Associated quadratic form: $q: \mathbb{R}^2 \mapsto \mathbb{R}$, $q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2x_1 x_2$
(we also consider corresponding forms over $\mathbb{C}$).

Reduction of $[f]_\mathcal{E} = [f]_\mathcal{E}^\mathcal{E}$ using "double operations"

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \xrightarrow{\mathcal{E}_r(1,2;\frac{1}{2})} \begin{pmatrix} \frac{1}{2} & 1 \\ 1 & 0 \end{pmatrix} \xrightarrow{\mathcal{E}_c(1,2;\frac{1}{2})} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\xrightarrow{\mathcal{E}_r(2,1;-1)} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \xrightarrow{\mathcal{E}_c(2,1;-1)} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{final form over } \mathbb{R}$$

$$\xrightarrow{\mathcal{D}_r(2;i)} \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix} \xrightarrow{\mathcal{D}_c(2,i)} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Canonical form of $q$ over $\mathbb{R}$ is: $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
To find a basis $\mathcal{B}$ of $\mathbb{R}^2$ such that $[q]_\mathcal{B} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ express the whole reduction in terms of elementary matrices:

$$\underset{E_r(2,1;-1)}{\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}} \underset{E_r(1,2;\frac{1}{2})}{\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \underbrace{\underset{E_c(1,2;\frac{1}{2})}{\begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{pmatrix}} \underset{E_c(2,1;-1)}{\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}}}_{\text{columns of this product will form basis } \mathcal{B} \text{ of } \mathbb{R}^2} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & -1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}}_{M}$$

Then $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = M^T \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} M$

$[q]_\mathcal{B} = \left([Id]_\mathcal{B}^\mathcal{E}\right)^T [q]_\mathcal{E} [Id]_\mathcal{B}^\mathcal{E}$

Columns of $M = [Id]^{\mathcal{E}}_{\mathcal{B}}$ form basis $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} -1 \\ \frac{1}{2} \end{pmatrix} \right\}$

Check: If $b_1 = \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}$, $b_2 = \begin{pmatrix} -1 \\ \frac{1}{2} \end{pmatrix}$ then $[f]_{\mathcal{B}} = \begin{pmatrix} f(b_1, b_1) & f(b_1, b_2) \\ f(b_2, b_1) & f(b_2, b_2) \end{pmatrix}$
$= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

e.g. $f(b_2, b_2) = \underbrace{(-1, \tfrac{1}{2})}_{b_2^T} \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{[f]_{\mathcal{E}}} \underbrace{\begin{pmatrix} -1 \\ \frac{1}{2} \end{pmatrix}}_{b_2} = (\tfrac{1}{2}, -1) \begin{pmatrix} -1 \\ \frac{1}{2} \end{pmatrix} = -1$

Over $\mathbb{C}$, the canonical form is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

To find an associated basis $\mathcal{C}$, we again express the reduction in terms of elementary matrices:

$$\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\underbrace{\begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}}_{\text{columns of this form } \mathcal{C}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

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

So $\mathcal{C} = \left\{ \begin{pmatrix} 1 \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} -i \\ \frac{i}{2} \end{pmatrix} \right\}$

Then $[f]_{\mathcal{C}} = \begin{pmatrix} f(c_1, c_1) & f(c_1, c_2) \\ f(c_2, c_1) & f(c_2, c_2) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Quick example for useful permutations:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{P_r(1,3)} \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{P_c(1,3)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

MATH 2201

09-12-16

## Some proofs related to canonical forms

We first prove that we can diagonalise any symmetric bilinear form.

Suppose that, for a field $k$, and a vector space $V$ over $k$, where $\dim_k(V) = n$, we have a symmetric bilinear form $f: V \times V \mapsto k$, associated to a quadratic form $q: V \mapsto k$.

Then, if there exists a basis $\mathcal{B}$ of $V$ such that $[f]_{\mathcal{B}}^{\mathcal{B}}$ is diagonal, i.e. $[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} f(b_1, b_1) & \cdots & f(b_1, b_n) \\ \vdots & & \vdots \\ f(b_n, b_1) & \cdots & f(b_n, b_n) \end{pmatrix}$

$$= \begin{pmatrix} f(b_1, b_1) & 0 & & 0 \\ 0 & f(b_2, b_2) & & \\ & 0 & \ddots & \\ & & & f(b_n, b_n) \end{pmatrix}$$

then $f(b_i, b_j) = 0$ if $i \neq j$.
The basis $\mathcal{B}$ is orthogonal with respect to $f$.

So diagonalising $f \longleftrightarrow$ finding an orthogonal basis with respect to $f$.

Let's now show that any such bilinear form $f$ can be diagonalised (at least over $\mathbb{R}$ or $\mathbb{C}$).

Key notion:
Given a subset $S$ of $V$, the orthogonal compliment of $S$, denoted by $S^\perp$ with respect to $f$ is

$$S^\perp = \{ v \in V : f(v, w) = 0 \ \forall w \in S \}$$

Then; for any set $S$ in $V$, $S^\perp$ is a subspace of $V$.

Check:

1). Consider $0 \in V$.
Then $f(0,w) = 0 \quad \forall w \in S$ (in fact, $f(0,v)\ \forall v \in V$, by definition of a bilinear form).
So $0 \in S^\perp$.

2). Suppose $a, b \in S^\perp$ i.e. $\forall w \in S$, $f(a,w) = 0$ and $f(b,w) = 0$.
Then $\forall w \in S$: $f(a+b, w) = f(a,w) + f(b,w) = 0 + 0 = 0$.
So $a+b \in S^\perp$.

3). Suppose that $a \in S^\perp$: $\forall w \in S$, $f(a,w) = 0$.
then, $\forall w \in S$ and a constant $\lambda \in k$: $f(\lambda a, w) = \lambda f(a,w) = \lambda \cdot 0 = 0$.
So $\lambda a \in S^\perp$.

So $S^\perp$ is a subspace of $V$. //

Now the key proposition that leads to diagonalisability

Proposition

For any vector $v \in V$, if $f(v,v) \neq 0$, then
$$V = \operatorname{span}\{v\} \oplus \{v\}^\perp \qquad \text{set } S = \{v\}$$

Proof

Suppose that $w$ is a vector in $V$.

Set $w_1 = \dfrac{f(w,v)}{f(v,v)} v$, $w_2 = w - \dfrac{f(w,v)}{f(v,v)} v$

Then $w = w_1 + w_2$
and $w_1$ is a multiple of $v$,
so $w_1 \in \operatorname{span}\{v\}$
$(\operatorname{span}\{v\} = \{\lambda v : \lambda \in k\})$
Let's also check that $w_2 \in \{v\}^\perp$
$$f(w_2, v) = f\left(w - \frac{f(w,v)}{f(v,v)} v\right)$$

MATH
2201

09-12-16

So $f(w_2, v) = f(w, v) + f\left(-\frac{f(w,v)}{f(v,v)} v, v\right)$

$= f(w,v) - \frac{f(w,v)}{f(v,v)} f(v,v)$

$= f(w,v) - f(w,v) = 0$

So $w_1 \in \text{span}\{v\}$, $w_2 \in \{v\}^\perp$

So $V = \text{span}\{v\} + \{v\}^\perp$.

Let's now check that the sum is direct:
Suppose that $w \in \text{span}\{v\}$ and $w \in \{v\}^\perp$
Since $w \in \text{span}\{v\}$: $w = \lambda v$ for some $\lambda \in k$.
Since $w \in \{v\}^\perp$: $f(v, w) = 0$.
i.e. $f(v, \lambda v) = 0$, i.e. $\lambda f(v,v) = 0$.
Then, by assumption, $\lambda = 0$ and $w = \lambda v = 0$.
So $w = 0$, and $\text{span}\{v\} \cap \{v\}^\perp = \{0\}$
Hence, the sum is direct:

$V = \text{span}\{v\} \oplus \{v\}^\perp$

$\square$.

This proposition leads to the required result:
Diagonalisation Theorem:
Suppose that $k$ is a field in which $1+1 \neq 0$, and that $V$ is a vector space over $k$, of finite dimension.
Any symmetric bilinear form $f: V \times V \to k$ can be diagonalised, i.e. there exists an orthogonal basis $\mathcal{B}$ of $V$ with respect to $f$.

Proof
(By induction on the dimension of $V$, $\dim(V)$.)

If $\dim(V)=1$, then with respect to any basis $\mathcal{C}$: $[f]_{\mathcal{C}}$ is a $1\times1$ matrix. Every $1\times1$ matrix is diagonal, so any basis $\mathcal{C}$ is orthogonal with respect to $f$.

Suppose that the result holds for all vector spaces of dimension smaller than or equal to $n$, and let $\dim(V)=n+1$.
If, for all $v$ in $V$, $f(v,v)=0$.
Then, using an earlier result:

$$\text{for } u,w \text{ in } V:\ f(u,w) = \tfrac{1}{2}\big(f(u+w,u+w) - f(u,u) - f(w,w)\big)$$
$$= \tfrac{1}{2}(0-0-0)$$

i.e. $f(u,w)=0 \ \forall u,w$.
So $f$ must be the zero bilinear form,
i.e. $[f]_{\mathcal{C}}$ is the zero matrix for any basis $\mathcal{C}$ which is diagonal.

Let's now consider the case where there exists at least one $v$ in $V$ such that $f(v,v)\neq 0$
Then, using the previous proposition:

$$V = \operatorname{span}\{v\} \oplus \{v\}^{\perp}$$

Since the sum is direct: $\dim(V) = \dim(\operatorname{span}\{v\}) + \dim(\{v\}^{\perp})$

$$\text{i.e.}\quad n+1 = 1 + \dim(\{v\}^{\perp})$$

Hence, $\dim(\{v\}^{\perp})=n$ and, by the inductive assumption, there is an orthogonal basis, $\mathcal{B}=\{b_1,\dots,b_n\}$ of $\{v\}^{\perp}$.
Thus $f(b_i,b_j)=0$ for $i,j=1,\dots,n$, $i\neq j$. (*)

Then, since $\{v\}$ is a basis for $\operatorname{span}\{v\}$ and the sum is direct:

$$\{b_1,\dots,b_n\} \text{ is a basis for } V.$$

MATH 2201

09-12-16

But $b_1, \dots, b_n \in \{v\}^\perp$, so, by definition $f(b_i, v) = 0$ for $i = 1, \dots, n$. (**)

Combining (*) and (**), we see that $\{b_1, \dots, b_n, v\}$ is an orthogonal basis for $V$.

$\square$.

So every symmetric bilinear form / quadratic form over $\mathbb{R}$ or $\mathbb{C}$ can be diagonalised.

Now, let's explain that, from a diagonal matrix, we can obtain a matrix in canonical form.

Example

Suppose that, for some orthogonal basis $\mathcal{B} = \{b_1, b_2\}$ such that $[f]_{\mathcal{B}} = \begin{pmatrix} 2 & 0 \\ 0 & -9 \end{pmatrix}$

So $f(b_1, b_1) = 2$, $f(b_2, b_2) = -9$

Then $f\left(\frac{b_1}{\sqrt{2}}, \frac{b_1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}} f\left(b_1, \frac{b_1}{\sqrt{2}}\right) = \frac{1}{2} f(b_1, b_1) = 1$

Similarly $f\left(\frac{b_2}{3}, \frac{b_2}{3}\right) = -1$

Over $\mathbb{C}$: $f\left(\frac{b_2}{3i}, \frac{b_2}{3i}\right) = -\frac{1}{9} f(b_2, b_2) = 1$

Hence in terms of basis $\mathcal{C}$, over $\mathbb{R}$: $\mathcal{C} = \left\{\frac{1}{\sqrt{2}} b_1, \frac{1}{3} b_2\right\}$:

$$[f]_{\mathcal{C}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Similarly over $\mathbb{C}$: $\mathcal{D} = \left\{\frac{1}{\sqrt{2}} b_1, \frac{1}{3i} b_2\right\}$ : $[f]_{\mathcal{D}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Hence, we can obtain canonical forms for $f$, over $\mathbb{R}$ or $\mathbb{C}$.

Note:
Canonical forms, over $\mathbb{R}$ and $\mathbb{C}$ are unique.

Let's consider two forms over $\mathbb{C}$:
Suppose that, for bases $\mathcal{B}$, $\mathcal{B}'$ of $\mathbb{C}^n$:

$$[f]_{\mathcal{B}} = \begin{pmatrix} I_m & 0 \\ 0 & 0_{n-m} \end{pmatrix} = \begin{pmatrix} \underbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}_{m} & 0 \\ 0 & \begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix} \end{pmatrix}$$

$$[f]_{\mathcal{B}'} = \begin{pmatrix} I_{m'} & 0 \\ 0 & 0_{n-m'} \end{pmatrix} = \begin{pmatrix} \underbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}_{m'} & 0 \\ 0 & \begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix} \end{pmatrix}$$

Let $\mathcal{B} = \{b_1, \dots, b_m, b_{m+1}, \dots, b_n\}$, $\mathcal{B}' = \{b_1', \dots, b_{m'}', b_{m'+1}', \dots, b_n'\}$

$f(b_i, b_i) = 1$ for $1 \le i \le m$ $\qquad$ $f(b_i', b_i') = 1$ for $1 \le i \le m'$

$f(b_i, b_i) = 0$ for $m+1 \le i \le n$ $\qquad$ $f(b_i', b_i') = 0$ for $m'+1 \le i \le n$

Consider the set $\mathcal{C} = \{b_1, \dots, b_m, b_{m'+1}', \dots, b_n'\}$ in $\mathbb{C}^n$

Let $U = \operatorname{span}\{b_1, \dots, b_m\}$, $W = \operatorname{span}\{b_1', \dots, b_{m'}'\}$

$m + (n - m') \le n \implies m \le m'$.

$\dim(U \oplus W)$

MATH 2201

12-12-16

From last time:

Suppose, that, for a bilinear form $f$:

$$[f]_{\mathcal{B}} = \begin{pmatrix} I_m & O \\ O & O_{n-m} \end{pmatrix}, \quad [f]_{\mathcal{B}'} = \begin{pmatrix} I_{m'} & O \\ O & O_{n-m'} \end{pmatrix}$$

for bases $\mathcal{B} = \{b_1, \dots, b_m, b_{m+1}, \dots, b_n\}$

$\mathcal{B}' = \{b_1', \dots, b_{m'}', b_{m'+1}', \dots, b_n'\}$

Let's show that $m = m'$, i.e. that the complex canonical form of $f$ is unique.

Suppose that $U$ is a subspace of $\mathbb{C}^n$ with basis $\{b_1, \dots, b_m\}$
and $W$ is a subspace of $\mathbb{C}^n$ with basis $\{b_{m'+1}', \dots, b_n'\}$

General elements $\begin{cases} u \text{ in } U: & u = u_1 b_1 + \dots + u_m b_m \\ w \text{ in } W: & w = w_{m'+1} b_{m'+1}' + \dots + w_n b_n' \end{cases}$

Then $f(u,u) = f(u_1 b_1 + \dots + u_m b_m,\ u_1 b_1 + \dots + u_m b_m)$

$= u_1^2 f(b_1, b_1) + \dots + u_m^2 f(b_m, b_m)$

$= u_1^2 + \dots + u_m^2$

If we assume that $u_1, \dots, u_m \in \mathbb{R}$, then $f(u,u) \geq 0$
and $f(u,u) = 0$ if and only if $u_1 = 0, \dots, u_m = 0$
i.e. if and only if $u = 0$, the zero vector.

Similarly: $f(w,w) = f(w_{m'+1} b_{m'+1}' + \dots + w_n b_n',\ w_{m'+1} b_{m'+1}' + \dots + w_n b_n')$

$= w_{m'+1}^2 f(b_{m'+1}', b_{m'+1}') + \dots + w_n^2 f(b_n', b_n')$

$= w_{m'+1}^2 \cdot 0 + \dots + w_n^2 \cdot 0$

$= 0$

So if $v \in U \cap W$, then $f(v,v) \geq 0$ and $f(v,v) = 0$, i.e. $v = 0$
So $U \cap W = \{0\}$. Hence, the sum $U + W$ is direct: $U + W = U \oplus W$.

So $\dim(U \oplus W) = \dim(U) + \dim(W)$
$= m + (n - m')$

$U \oplus W$ is a subspace of $\mathbb{R}^n / \mathbb{C}^n$, so $\dim(U \oplus W) \leq n$
So $m + (n - m') \leq n$
and then $m \leq m'$.

Similarly, if we choose $U$ with basis $\{b_1', \dots, b_{m'}'\}$
and $W$ " " $\{b_{m+1}, \dots, b_n\}$
we will show that $m' \leq m$.
Overall, we obtain $m = m'$. $\square$

—

We now try to use quadratic / bilinear forms to measure lengths.

Key requirement: $q(v,v)$ a real number
and $q(v,v) \geq 0 \ \forall \ v \in V$.

Suppose $q$ is a real quadratic form corresponding to the bilinear form $f$.
Then, to ensure the above conditions hold, we require the canonical form of $f$ to be $I_n$.

Any such form is <u>positive definite</u>.
A bilinear form $f: \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ is <u>positive definite</u> if the real canonical form of $f$ is the identity matrix.
Equivalently: $f(v,v) \geq 0 \ \forall \ v \in \mathbb{R}^n$, and $f(v,v) = 0$ only if $v = 0$.

<u>Example</u>
$f: \mathbb{R}^2 \times \mathbb{R}^2 \mapsto \mathbb{R}$, where $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = x_1 y_1 + x_2 y_2$
$= (x_1 \ x_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

MATH 2201

12-12-16

Then $f\left(\begin{pmatrix}x_1\\x_2\end{pmatrix}, \begin{pmatrix}x_1\\x_2\end{pmatrix}\right) = x_1^2 + x_2^2$
which satisfies the definition.

Whereas if a bilinear form is represented by

$[f]_{\mathcal{B}} = \begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}$ e.g. $(0,0,0,1)\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = -1$

So $f\left(\begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\right) < 0$

This leads to the notion of a product space.

An <u>inner product space</u>, over $\mathbb{R}$, is a vector space $V$ over $\mathbb{R}$, together with a positive definite symmetric bilinear form $f: V \times V \to \mathbb{R}$.

So, $f$ satisfies, for all $u, v, w \in V$, $\lambda \in \mathbb{R}$:

$f(u+v, w) = f(u,w) + f(v,w)$

$f(u, v+w) = f(u,v) + f(u,w)$

$f(\lambda u, v) = \lambda f(u,v) = f(u, \lambda v)$

$f(u,v) = f(v,u)$

$f(u,u) \geq 0$, and $f(u,u) = 0$ only if $u = 0$.

<u>Example</u>:

Take $V$ to be $\mathbb{R}^n$ and $f$ to be any symmetric bilinear form $f: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ whose canonical form is $I_n$.

e.g. from earlier $V = \mathbb{R}^2$ and $f$ represented by $\begin{pmatrix}1&2\\2&5\end{pmatrix}$

$(x_1 \ x_2)\begin{pmatrix}1&2\\2&5\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1^2 + 4x_1x_2 + 5x_2 \geq 0 \quad \forall x_1, x_2 \in \mathbb{R}$

and $x_1^2 + 4x_1x_2 + 5x_2 = 0$ if $x_1 = 0$, $x_2 = 0$.

In an inner product space, the norm of a vector $v$ in $V$ is $\sqrt{f(v,v)}$.
This is denoted by $\|v\|$.

Often, in an inner product space, $f(u,v)$ is written as $\langle u,v \rangle$.

MATH 2201

13-12-16

Let's now try to find forms inner products that measure lengths of complex vectors.

To do so, we modify the way bilinear forms are defined.

Consider "standard form", represented by $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$:

In $\mathbb{R}^2$: $(x_1 \ x_2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + x_2^2$ ← square root is the length / norm of $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

In $\mathbb{C}^2$: $(z_1 \ z_2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = z_1^2 + z_2^2$ ← does not relate directly to length.

The length of $z \in \mathbb{C}$ is given by $\sqrt{z\bar{z}}$.

To obtain such a product define forms as follows:

$$\langle v, w \rangle = v^T \underset{\substack{\uparrow \\ \text{matrix}}}{A} \bar{w} \longleftarrow \text{conjugate of the second vector}$$

e.g. if $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $v = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$, $w = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$

then $(z_1 \ z_2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \bar{z}_1 \\ \bar{z}_2 \end{pmatrix} = \underbrace{z_1\bar{z}_1 + z_2\bar{z}_2}$

we want this expression to be real and nonnegative.

To ensure this, we need the matrix $A$ to satisfy $A^T = \bar{A}$ and to be positive definite.

e.g. Suppose $A$ is a $2 \times 2$ complex matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad a_{11}, a_{12}, a_{21}, a_{22} \in \mathbb{C}$$

Then if $A^T = \bar{A}$: $\begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} = \begin{pmatrix} \bar{a}_{11} & \bar{a}_{12} \\ \bar{a}_{21} & \bar{a}_{22} \end{pmatrix}$

So $a_{11} = \bar{a}_{11}$, $a_{12} = \bar{a}_{21}$, $a_{21} = \bar{a}_{12}$, $a_{22} = \bar{a}_{22}$

So $a_{11}, a_{22} \in \mathbb{R}$, $a_{12} = \bar{a}_{21}$.

Then $(z_1 \quad z_2)\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} \bar{z}_1 \\ \bar{z}_2 \end{pmatrix}$

$$= a_{11} z_1 \bar{z}_1 + a_{12} z_1 \bar{z}_2 + a_{21} z_2 \bar{z}_1 + a_{22} z_2 \bar{z}_2$$

$$= \underbrace{a_{11} |z_1|^2}_{\in \mathbb{R}} + \underbrace{a_{12} z_1 \bar{z}_2 + a_{21} z_2 \bar{z}_1}_{\in \mathbb{R}} + \underbrace{a_{22} |z_2|^2}_{\in \mathbb{R}}$$

e.g. consider $A = \begin{pmatrix} -3 & 1+i \\ 1-i & 2 \end{pmatrix}$

Then $(2 \ , 1-i)\begin{pmatrix} -3 & 1+i \\ 1-i & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 1+i \end{pmatrix} \in \mathbb{R}$

But this is not positive definite

e.g. $(1 \quad 0)\begin{pmatrix} -3 & 1+i \\ 1-i & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = -3$ ← want a positive number so can take root.

The general definition is the following:

A <u>Hermitian form</u> on a complex on a complex vector space, $\mathbb{C}^n$ say, is a form $\langle \cdot , \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \mapsto \mathbb{C}$ satisfying:

$$\langle a+b, c\rangle = \langle a, c\rangle + \langle b, c\rangle$$
$$\langle a, b+c\rangle = \langle a, b\rangle + \langle a, c\rangle$$
$$\langle \lambda a, b\rangle = \lambda \langle a, b\rangle$$
$$\langle a, \lambda b\rangle = \bar{\lambda} \langle a, b\rangle$$
$$\langle a, b\rangle = \overline{\langle b, a\rangle}.$$

$\langle a, b\rangle = (a_1 \quad a_2)\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}\begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix}$

Any Hermitian form can be represented by a Hermitian matrix: a square $n \times n$ complex matrix $A$ satisfying $A^T = \bar{A}$.

e.g. if $A = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$, then if we define

$\left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\rangle = (a_1, a_2)\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}\begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \end{pmatrix}$ all conditions in the definition will hold.

MATH 2201

13-12-16

Here:

$$\left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \right\rangle = (a_1, a_2)\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}\begin{pmatrix} \bar{a}_1 \\ \bar{a}_2 \end{pmatrix}$$

$$= a_1\bar{a}_1 + \underbrace{ia_1\bar{a}_2 - ia_2\bar{a}_1}_{\in \mathbb{R} \text{ since } ia_1\bar{a}_2 = \overline{-ia_2\bar{a}_1}} + a_2\bar{a}_2$$

A complex inner product space is a vector space $\mathbb{C}^n$ with a Hermitian form $\langle \cdot , \cdot \rangle$ that is also positive definite.

ie. $\langle v, v\rangle > 0$ for non-zero $v$, $\langle v, v \rangle = 0$ if $v = 0$.

Note:

A Hermitian matrix $A$ corresponds to a positive definite form if the canonical form of $A$ is the identity matrix, $I_n$.

[ie. row reduce to see if you get $I_n$]

Real and complex inner products satisfy some imortant results:

* Cauchy-Schwarz inequality:

Let $\langle \cdot , \cdot \rangle$ represent a real or complex inner product, then for any vectors $a, b$: $|\langle a, b\rangle| \leq \|a\| \cdot \|b\|$

(where $\|a\| = \sqrt{\langle a, a\rangle}$)

MATH 2201

16-12-16

## From last time

Cauchy - Schwarz inequality

Let $V$ be an inner product space.

For all $u, v \in V$ : $|\langle u,v\rangle| \leq \|u\|\,\|v\|$

Some rules:

$\langle u, v\rangle = \overline{\langle v, u\rangle}$

$\langle \lambda u, v\rangle = \lambda\langle u,v\rangle$

$\langle u, \lambda v\rangle = \bar{\lambda}\langle u,v\rangle$

$\|u\| = \sqrt{\langle u, v\rangle}$

It follows $\langle u,v\rangle\langle v,u\rangle = \langle u,v\rangle\overline{\langle u,v\rangle} = |\langle u,v\rangle|^2$

Proof

Suppose $v = 0$.

Then $\langle u,v\rangle = \langle u,0\rangle = 0$, so $|\langle u,v\rangle| = 0$

$\|u\|\|v\| = \|u\|\cdot\|0\| = \|u\|\cdot 0 = 0$ so $\|u\|\|v\| = 0$.

The (in)equality holds for $v = 0$.

Suppose $v \neq 0$.

Consider $\langle u-\lambda v, u-\lambda v\rangle$

This is of the form $\langle w,w\rangle$, so $\langle u-\lambda v, u-\lambda v\rangle \geq 0 \ \forall \lambda$.

Also $\langle u-\lambda v, u-\lambda v\rangle = \langle u,u\rangle - \langle \lambda v, u\rangle - \langle u, \lambda v\rangle + \langle \lambda v, \lambda v\rangle$

$= \langle u,u\rangle - \lambda\langle v,u\rangle - \bar{\lambda}\langle u,v\rangle + \lambda\bar{\lambda}\langle v,v\rangle$

Set $\lambda = \dfrac{\langle u,v\rangle}{\langle v,v\rangle}$ (note $v \neq 0$ by assumption, so $\langle v,v\rangle \neq 0$)

we obtain from $\langle u-\lambda v, u-\lambda v\rangle \geq 0$

$$\langle u,u\rangle - \frac{\langle u,v\rangle}{\langle v,v\rangle}\langle v,u\rangle - \frac{\overline{\langle u,v\rangle}}{\overline{\langle v,v\rangle}}\langle u,v\rangle + \frac{|\langle u,v\rangle|^2}{|\langle v,v\rangle|^2}\langle v,v\rangle \geq 0$$

Note $\langle v,v\rangle \in \mathbb{R}$, so $\overline{\langle v,v\rangle} = \langle v,v\rangle$; also $\langle v,v\rangle > 0$.

So we may multiply through by $\langle v,v\rangle$ to obtain

$\langle u,u\rangle\langle v,v\rangle - \langle u,v\rangle\langle v,u\rangle - \overline{\langle u,v\rangle}\langle u,v\rangle + |\langle u,v\rangle|^2 \geq 0$

$\|u\|^2\|v\|^2 - \cancel{|\langle u,v\rangle|^2} - |\langle u,v\rangle|^2 + |\langle u,v\rangle|^2 \geq 0$

So $\|u\|^2\|v\|^2 \geq |\langle u,v\rangle|^2$

All terms are non-negative, so may safely take square roots:

$\|u\|\|v\| \geq |\langle u,v\rangle|$ or $|\langle u,v\rangle| \leq \|u\|\|v\|$, as required. $\square$

This leads to

The Triangle Inequality

Let $V$ be an inner product space.

For all $a,b \in V$ : $\|a+b\| \leq \|a\| + \|b\|$

Consider
$$\begin{aligned}\|a+b\|^2 &= \langle a+b, a+b\rangle \\ &= \langle a,a\rangle + \langle a,b\rangle + \langle b,a\rangle + \langle b,b\rangle \\ &= \langle a,a\rangle + \langle a,b\rangle + \overline{\langle a,b\rangle} + \langle b,b\rangle \\ &= \|a\|^2 + 2\,\mathcal{R}e\langle a,b\rangle + \|b\|^2 \\ &\leq \|a\|^2 + 2|\langle a,b\rangle| + \|b\|^2 \\ &\leq \|a\|^2 + 2\|a\|\|b\| + \|b\|^2 \quad \text{by Cauchy-Schwarz inequality} \\ &= (\|a\| + \|b\|)^2\end{aligned}$$

So since $\|a+b\|^2 \leq (\|a\| + \|b\|)^2$

$\|a+b\| \leq \|a\| + \|b\|$ $\square$

All real inner products correspond to an identity matrix as a canonical form, there is a basis $\{d_1, \dots, d_n\}$ of $V$ such that the matrix representing the inner product is

$$\begin{pmatrix} 1 & & 0 \\ & 1 & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix} = \begin{pmatrix} \langle b_1, b_1\rangle & \cdots & \langle b_1, b_n\rangle \\ \langle b_2, b_2\rangle & & \langle b_2, b_n\rangle \\ \vdots & & \vdots \\ \langle b_n, b_1\rangle & \cdots & \langle b_n, b_n\rangle \end{pmatrix}$$

So $\langle b_i, b_j\rangle = 0$ for all $i,j$, if $i \neq j$

i.e. $\{b_1, \dots, b_n\}$ is orthogonal

and $\langle b_i, b_i\rangle = 1 \;\forall\, i = 1, \dots, n$.

MATH 2201

16-12-16

Overall, we have an orthonormal basis:

- $\langle b_i, b_j \rangle = 0$ if $i \neq j$
- $\|b_i\| = \sqrt{\langle b_i, b_i \rangle} = 1$

If $A$ is a complex symmetric matrix (i.e. if $A = A^T$), then can also reduce $A$ to canonical form using row / column operations.

But complex inner products involve Hermitian matrices instead (where $\overline{A} = A^T$). Can reduce by pairing row operations with their conjugate column operations.

$$\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \xrightarrow{R_2 \to R_2 + iR_1} \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} \xrightarrow{C_2 \to C_2 - iC_1} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Another way to obtain an orthonormal basis for any inner product $\langle \, , \rangle$:

Gram-Schmidt process

Start with any basis $\{a_1, \dots, a_n\}$ of the space $V$.

Compute the following vectors:

(note: for each vector $a_i$ in a basis, $a_i \neq 0$, so $\langle a_i, a_i \rangle \neq 0$)

$$b_1 = a_1$$

$$b_2 = a_2 - \frac{\langle a_2, b_1 \rangle}{\langle b_1, b_1 \rangle} b_1$$

$$b_3 = a_3 - \frac{\langle a_3, b_1 \rangle}{\langle b_1, b_1 \rangle} b_1 - \frac{\langle a_3, b_2 \rangle}{\langle b_2, b_2 \rangle} b_2$$

$$\vdots$$

$$b_n = a_n - \frac{\langle a_n, b_1 \rangle}{\langle b_1, b_1 \rangle} b_1 - \dots - \frac{\langle a_n, b_{n-1} \rangle}{\langle b_{n-1}, b_{n-1} \rangle} b_{n-1}$$

Then $\{b_1, \dots, b_n\}$ is an orthogonal basis for $V$.

Check

$$\langle b_1, b_2\rangle = \langle b_1, a_2 - \frac{\langle a_2, b_1\rangle}{\langle b_1, b_1\rangle} b_1\rangle$$

$$= \langle b_1, a_2\rangle - \langle b_1, \frac{\langle a_2, b_1\rangle}{\langle b_1, b_1\rangle} b_1\rangle \quad \text{(assume over } \mathbb{R}\text{)}$$

$$= \langle b_1, a_2\rangle - \frac{\langle a_2, b_1\rangle}{\langle b_1, b_1\rangle}\langle b_1, b_1\rangle$$

$$= \langle b_1, a_2\rangle - \langle a_2, b_1\rangle$$

$$= 0$$

To obtain an orthonormal basis

set $c_1 = \frac{b_1}{\|b_1\|}$, $c_2 = \frac{b_2}{\|b_2\|}$, ..., $c_n = \frac{b_n}{\|b_n\|}$

Then $\{c_1, \dots, c_n\}$ is an orthonormal basis for V.

Example

Consider the standard inner product on $\mathbb{R}^3$

$$\left\langle \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \right\rangle = (a_1\ a_2\ a_3)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1b_1 + a_2b_2 + a_3b_3$$

Apply the process to the following basis of $\mathbb{R}^3$:

$$\left\{ \underbrace{\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}}_{a_1}, \underbrace{\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}}_{a_2}, \underbrace{\begin{pmatrix} 3 \\ 7 \\ 5 \end{pmatrix}}_{a_3} \right\}$$

$$b_1 = a_1 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$$

$$b_2 = a_2 - \frac{\langle a_2, b_1\rangle}{\langle b_1, b_1\rangle} b_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right\rangle}{\left\langle \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right\rangle}\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}$$

$$b_3 = a_3 - \frac{\langle a_3, b_1\rangle}{\langle b_1, b_1\rangle} b_1 - \frac{\langle a_3, b_2\rangle}{\langle b_2, b_2\rangle} b_2$$

$$= \begin{pmatrix} 3 \\ 7 \\ 5 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 3 \\ 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right\rangle}{\left\langle \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right\rangle}\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 3 \\ 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} \right\rangle}{\left\langle \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} \right\rangle}\begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \\ 5 \end{pmatrix} - \frac{6}{4}\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} - \frac{24}{8}\begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}$$

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

MATH 2201

16-12-16

So $\left\{\begin{pmatrix}2\\0\\0\end{pmatrix}, \begin{pmatrix}0\\2\\2\end{pmatrix}, \begin{pmatrix}0\\1\\-1\end{pmatrix}\right\}$ is an orthogonal basis

$\left\|\begin{pmatrix}2\\0\\0\end{pmatrix}\right\| = 2$ so $c_1 = \frac{b_1}{2} = \begin{pmatrix}1\\0\\0\end{pmatrix}$

$\left\|\begin{pmatrix}0\\2\\2\end{pmatrix}\right\| = 2\sqrt{2}$ so $c_2 = \frac{b_2}{2\sqrt{2}} = \begin{pmatrix}0\\1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}$

$\left\|\begin{matrix}0\\1\\-1\end{matrix}\right\| = \sqrt{2}$ so $c_3 = \frac{b_3}{\sqrt{2}} = \begin{pmatrix}0\\1/\sqrt{2}\\-1/\sqrt{2}\end{pmatrix}$

and $\{c_1, c_2, c_3\}$ is an orthonormal basis.

Hermitian matrices are related to the notion of the adjoint of a matrix / linear map.

Consider an inner product $\langle , \rangle$ in terms of an orthonormal basis, so that for $u, v \in V$ ($\dim V = n$)

$$\langle u, v \rangle = u^T \begin{pmatrix}1 & & 0\\ & \ddots & \\ 0 & & 1\end{pmatrix} \bar{v} = u^T \bar{v}$$

With respect to this, the adjoint of an $n \times n$ matrix $A$ is a matrix $A^*$ such that, for all $u, v \in V$:

$$\langle Au, v \rangle = \langle u, A^* v \rangle$$

Let's try to determine $A^*$ in terms of $A$:

$$\left.\begin{aligned}\langle Au, v \rangle &= (Au)^T \bar{v} = u^T A^T \bar{v}\\ \langle u, A^* v \rangle &= u^T \overline{(A^* v)} = u^T \overline{A^*}\, \bar{v}\end{aligned}\right\} \text{ so } A^T = \overline{A^*}$$

So $A^* = \overline{A^T}$

A matrix $A$ is is self-adjoint if $A^* = A$, i.e. if $A^T = \bar{A}$, i.e. if $A$ is Hermitian.

## Key result

If $M$ is a Hermitian matrix, then all its eigenvalues are real, and $M$ can be diagonalised using an orthonormal basis of eigenvectors.

## Example

$M = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$

characteristic equation: $\begin{vmatrix} \lambda-1 & -i \\ +i & \lambda-1 \end{vmatrix} = (\lambda-1)^2 - (-i^2)$
$= \lambda^2 - 2\lambda$

eigenvalues : 0, 2.

Eigenvectors for $\lambda=0$: $(M - 0I)x = 0$

i.e. $\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

We obtain $\left.\begin{array}{l} x_1 + ix_2 = 0 \\ -ix_1 + x_2 = 0 \end{array}\right\} x_1 = -ix_2$

General eigenvector: $\begin{pmatrix} -ix_2 \\ x_2 \end{pmatrix} = x_2\begin{pmatrix} -i \\ 1 \end{pmatrix}$ $x_2 \in \mathbb{C}$, $x_2 \neq 0$.

Similarly, for $\lambda = 2$, solving $(M - 2I)x = 0$ we find a general eigenvector

$\begin{pmatrix} ix_2 \\ x_2 \end{pmatrix} = x_2\begin{pmatrix} i \\ 1 \end{pmatrix}$ $x_2 \in \mathbb{C}$, $x_2 \neq 0$.

Basis of eigenvectors for $\mathbb{C}^2$: $\left\{ \underset{\lambda=0}{\begin{pmatrix} -i \\ 1 \end{pmatrix}}, \underset{\lambda=2}{\begin{pmatrix} i \\ 1 \end{pmatrix}} \right\}$

If $P = \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}$ then $P^{-1}MP = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$

Check that $\begin{pmatrix} -i \\ 1 \end{pmatrix}, \begin{pmatrix} i \\ 1 \end{pmatrix}$ are orthogonal:

$\left\langle \begin{pmatrix} -i \\ 1 \end{pmatrix}, \begin{pmatrix} i \\ 1 \end{pmatrix} \right\rangle = (-i)\overline{(i)} + (1)(1) = 0$