$\boxed{L22}$ Last time we defined and studied the minimal polynomial $m_\alpha(x)$ of a linear map $\alpha: V \to V$ (V is f.d.)

Summary of tips to find $m_\alpha(x)$

(0) Choose any basis and compute $m_A(x)$ where A matrix representing α, to find $m_\alpha(x)$

(1) Compute $p_A(x) = \det(xI_n - A)$ $A \in M_n(\mathbb{K})$ and factorise it as much as possible. [Can depend on $\mathbb{K}$!]

(2) $m_A(x)$ divides $p_A(x)$ and must vanish on the same λ (by Theorem 5.18) so $x-\lambda$ is a factor of $p_A(x)$ then it is also a factor of $m_A(x)$. Write down all potential candidates for $m_A(x)$ consistent with these two facts.

(3) Find the smallest degree one of these that vanishes on $x = A$ so that $m_A(A) = 0$.

[(4) A is diagonalizable iff $m_A(x)$ is a product of distinct linear factors.]

<u>Example</u> If A is 3×3 with real entries.

$\Rightarrow$ $P_A(x)$ is cubic with real coefficient.

$\Rightarrow$ if λ is a root, then so is $\bar{\lambda}$,

so complex roots appear in pairs if we work over $\mathbb{C}$.

$\Rightarrow$ at least one of the roots is real.

$$P_A(x) = (x-\mu)\,(x^2+bx+c) \qquad (*)$$

real

$$= (x-\mu)(x-\lambda_+)(x-\lambda_-)$$

$$\lambda_\pm = \frac{-b \pm \sqrt{b^2-4c}}{2}$$

If $b^2-4c \geq 0$, $\lambda_\pm \in \mathbb{R}$

if $x < 0$, $\lambda_\pm$ are strictly complex

<u>Example</u> $A = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix}$, $P_A(x) = \begin{vmatrix} x-1 & 0 & -2 \\ 2 & x-1 & 0 \\ -4 & -2 & x-1 \end{vmatrix}$

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

$$= (x-1)^3 - 8(x-2)$$

$$= x^3 - 3x^2 + 3x - 1 - 8x + 16$$

$$= x^3 - 3x^2 - 5x + 15$$

$$= (x-3)(x+\sqrt{5})(x-\sqrt{5})$$

over $\mathbb{R}$.

1. guess by inspection
2. draw a graph
3. match to the formula (*) to find μ, b, c.

$m_A(x)$ contains each of these as factors and divide $P_A(x)$ $\Rightarrow$ $m_A(x) = P_A(x)$

$\therefore$ $m_A(x)$ is a product of distinct linear factors

$\therefore$ A diagonalizable over $\mathbb{R}$

If A is regarded over $\mathbb{Q}$ then

$$p_A(x) = (x-3)(x^2-5)$$

$$m_A(x) = (x-3) \quad \text{or} \quad (x-3)(x^2-5) = p_A$$

but $A \neq 3I_3$ so $m_A(x) \neq x-3$

$\therefore$ $m_A = p_A$ not a product of distinct linear factors $\therefore$ not diagonalizable over $\mathbb{Q}$

If A is regarded over $\mathbb{F}_5$ then

$$p_A(x) = (x-3)x^2$$ options for $m_A(x)$

are $(x-3)x$, $(x-3)x^2$

but $(A-3I_3)A \neq 0$ (work over $\mathbb{F}_5$)

so $m_A(x) = (x-3)x^2 = p_A(x)$

not a product of <u>distinct</u> linear factors

$\therefore$ not diagonalizable.

5.6 Jordan Form

<u>Definition 5.22</u> (a) A jordan block $J(n, \lambda)$ is an $n \times n$ matrix of the form

$$\begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

(b) A matrix is in <u>jordan form</u> if it can be written in block form with jordan blocks on the diagonal and zero elsewhere:

$$\begin{bmatrix} J_1 & & & & 0 \\ & J_2 & & & \\ & & J_3 & & \\ & & & \ddots & \\ 0 & & & & J_r \end{bmatrix}$$

(by definition $J(1,\lambda) = [\lambda]$ as a 1×1 matrix)

e.g. if a matrix is diagonal then its of jordan form with all the blocks of type $J(1,\lambda)$ for various λ.

<u>Theorem 5.23</u> Over $\mathbb{C}$, <u>any</u> square matrix is similar to a matrix of jordan form.

i.e. every linear map on a f.d. vector space over $\mathbb{C}$ can be represented by a jordan form w.r.t. some basis. Moreover, the jordan form is unique up to reordering of the blocks

<u>proof</u> – omitted.

<u>Example</u> $n=3$ for 3×3 matrices, could have

$$\begin{bmatrix} \lambda & & 0 \\ & \mu & \\ 0 & & \nu \end{bmatrix}, \quad \left[\begin{array}{cc|c} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ \hline 0 & 0 & \boxed{\mu} \end{array}\right], \quad \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

all 1-blocks (diagonal matrix) | 2-block and 1-block | 3-block these are all the possible jordan forms up to reordering of the blocks, for some $\lambda, \mu, \nu \in \mathbb{C}$.

The theorem says that every 3×3 matrix over $\mathbb{C}$ is similar to one of these.

5.7 Trace Recall that if $A = (a_{ij})$ an $n\times n$ matrix then $\mathrm{Tr}(A) := \sum_{i=1}^{n} a_{ii}$ (sum of the diagonal entries)

prop 5.7 (a) if $A, B \in M_n(\mathbb{C}/\mathbb{K})$ then $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$

(b) If $B = PAP^{-1}$ then $\mathrm{Tr}(B) = \mathrm{Tr}(A)$

proof (a) $\mathrm{Tr}(AB) = \sum_i (AB)_{ii} = \sum_{i,j} a_{ij} b_{ji}$

$(A = (a_{ij}),\ B = (b_{ij}))$

$$= \sum_{i,j} b_{ji} a_{ij} = \sum_j (BA)_{jj}$$

$$= \mathrm{Tr}(BA)$$

(b) By part (a):

$$\mathrm{Tr}(PAP^{-1}) = \mathrm{Tr}((PA)P^{-1}) = \mathrm{Tr}(P^{-1}(PA))$$

$$= \mathrm{Tr}((P^{-1}P)A) = \mathrm{Tr}(A).$$

QED.

As discussed before, property (b) $\Rightarrow$ there is a notion of $\mathrm{Tr}(\alpha)$ and $\alpha : V \to V$, V f.d.

defined as $\mathrm{Tr}(A)$ for any representative matrix A of α w.r.t. any basis.

(Similarly for $\det(\alpha)$, $P_\alpha(x)$)

~~Proof~~ <u>Prop 5.28</u> If $\alpha : V \longrightarrow V$, $\dim(V) = n$

(a) The coefficient of x^{n-1} in $P_\alpha(x)$ is $-\mathrm{Tr}(\alpha)$ (ie $P_\alpha(x) = x^n - \mathrm{Tr}(\alpha)x^{n-1} + \cdots$)

and coeff of $x^0 = 1$ in $P_\alpha(x)$ is $(-1)^n \det(\alpha)$ (so $P_\alpha(x) = x^n - \mathrm{Tr}(\alpha)x^{n-1} + \cdots + (-1)^n \det(\alpha)$)

$\boxed{\text{L23}}$ <u>proof</u> (a) $P_A(x) = \det(xI_n - A) = \begin{vmatrix} x-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & x-a_{22} & \cdots & -a_{2n} \\ \vdots & & \ddots & \vdots \\ -a_{n1} & \cdots & \cdots & x-a_{nn} \end{vmatrix}$

we'll use the Laplace expansion along the first row. The only terms containing x^{n-1} come from the first term $(x-a_{11}) \cdots \cdots (x-a_{nn})$ since the terms will cross out two linear factors

e.g. $\begin{vmatrix} \cancel{x-a_{11}} & \cancel{-a_{12}} & \cancel{\cdots} & \cancel{-a_{1n}} \\ & x-a_{22} & & \\ & & \ddots & \\ & & & x-a_{nn} \end{vmatrix}$ the remaining subdeterminant can't contain x^{n-1}.

In $(x-a_{11}) \cdots (x-a_{nn})$

the coeff. of x^{n-1} is from $-a_{ii}$ for one of the factors and x in the others. So this gives

$$-\sum_{i=1}^{n} a_{ii} = -\operatorname{Tr}(A).$$

The coefficient of x^0 in $P_A(x)$ is

$$P_A(0) = \det(0 I_n - A) = (-1)^n \det(A).$$

Prop 5.28 (ctd)

(b) If α is diagonalizable then the sum of its eigenvalues (counted with multiplicity) is $\operatorname{Tr}(\alpha)$ and the product (with multiplicity) is $\det(\alpha)$.

proof of (b) Choose a basis s.t. α is represented by

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \Rightarrow \lambda_i \text{ are the eigenvalues}$$

(not necessarily distinct)

$$\left.\begin{aligned} \operatorname{Tr}(\alpha) &= \operatorname{Tr}(A) = \sum_i \lambda_i \\ \det(\alpha) &= \det(A) = \prod_i \lambda_i \end{aligned}\right\} \text{ as stated.}$$

If an eigenvalue λ occurs p times in the list $\lambda_1, \dots, \lambda_n$ then it contributes $p\lambda$ in $\operatorname{Tr}(\alpha)$ and λ^p in $\det(\alpha)$

So you could do the sums and products over the distinct eigenvalues but with such multiplicities.

QED.

In the proof we didn't necessarily need α to be diagonalizable, only ~~at~~ that

$$P_\alpha(x) = (x-\lambda_1)\cdots(x-\lambda_n) \quad \text{for some } \lambda_i$$

(not necessarily distinct) – part (b) still applies, by part (a).

Look at Cwk 8

Q3 $A = \begin{bmatrix} 0 & -1 & -1 \\ 2 & 3 & 1 \\ 4 & 2 & 4 \end{bmatrix}$ has eigenvectors $v_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$, $v_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$

$\lambda_1 = 2$

$v_3 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$ $\lambda_3 = 3$

$\lambda_2 = 2.$

(a) Find $P, Q = P^{-1}$ s.t. QAP is the diagonal matrix

$$D = \begin{bmatrix} 2 & & 0 \\ & 2 & \\ 0 & & 3 \end{bmatrix}.$$

Soln $P = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ -2 & -1 & 2 \end{bmatrix}$, $\Rightarrow P^{-1} = \begin{bmatrix} 3 & 1 & 1 \\ -2 & 0 & -1 \\ 2 & 1 & 1 \end{bmatrix}$

(b) What is the limit of $3^{-t}A^t$ as $t \to \infty$?

Soln $A^t = (PDP^{-1})^t = PD^tP^{-1}$ so

$$\left(\frac{A}{3}\right)^t = P\left(\tfrac{1}{3}D\right)^t P^{-1} = P \begin{bmatrix} \left(\frac{2}{3}\right)^t & & 0 \\ & \left(\frac{2}{3}\right)^t & \\ 0 & & 1 \end{bmatrix} P^{-1}$$

$$\to P \begin{bmatrix} 0 & & 0 \\ & 0 & \\ 0 & & 1 \end{bmatrix} P^{-1} = \begin{bmatrix} -2 & -1 & -1 \\ 2 & 1 & 1 \\ 4 & 2 & 2 \end{bmatrix}$$

Q4 True or false?

(a) Let A be $n \times n$ with real entries

If A is diagonalizable viewed as a map on $\mathbb{C}^n$ then is it diagonalizable on $\mathbb{R}^n$?

yes ☐ no ☑

e.g. $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ eigenvalues $1 \pm i$

diagonal form over $\mathbb{C}$ is $\begin{bmatrix} 1+i & 0 \\ 0 & 1-i \end{bmatrix}$

$p_A(x) = x^2 - 2x + 2$ over $\mathbb{R}$,

$m_A(x) = p_A(x)$ not product of distinct linear factors $\Rightarrow$ not diagonalizable. But over $\mathbb{C}$, $m_A(x) = p_A(x)$ also, $= (x-1+i)(x-1-i)$

(b) If A is diagonalizable on $\mathbb{R}^n$ then is it diagonalizable on $\mathbb{C}^n$?

yes $\boxed{\checkmark}$ no $\square$

If $m_A(x)$ is a product of distinct linear factors $m_A(x) = (x-\lambda_1)\cdots(x-\lambda_r)$ over $\mathbb{R}$ so $\lambda_i \in \mathbb{R}$, this is still true over $\mathbb{C}$.

(c) If A has n distinct real eigenvalues then it is diagonalizable viewed as a linear map on $\mathbb{R}^n$.

$\boxed{\checkmark}$ yes $\square$ no

$m_A(x) = (x-\lambda_1)\cdots(x-\lambda_n)$ as it must contain all these factors and already the maximum possible degree n. So A is diagonalizable

Q5 Suppose α linear map on $\mathbb{R}^2$. For each case (a)-(c) find A ~~st~~ representing

17 such that $m_\alpha(x)$ is:

(a) $m_\alpha(x) = x-1$

(b) $m_\alpha(x) = (x-1)^2$

(c) $m_\alpha(x) = x^2 - 2x + 2$

Soln (a) A ~~m~~ must obey $m_A(A) = 0$ so $A - I_2 = 0$, $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ will obey $m_A(A) = 0$

(c) here $m_\alpha(x) = P_\alpha(x)$ as m_α has degree 2. $m_\alpha(x) = (x-1)^2 + 1$ suggests $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ has this $P_\alpha(x)$.

L24

- Quiz 4 will go live at some point on thursday (due 11.59pm thursday next week)

- Week 10 quadratic forms
- week 11 inner product spaces.
- week 12 - revision lectures.

Chapter 6 Quadratic forms

Definition 6.1 Let V be a v.s. over $\mathbb{K}$.
A linear form on V is a linear map
$f: V \to \mathbb{K}$ ie $f(v_1+v_2) = f(v_1) + f(v_2)$
$f(cv) = cf(v)$
$\forall v, v_1, v_2 \in V, \quad c \in \mathbb{K}$
(Special case of $\alpha: V \to W$ with W 1-dimensional)

Remark if $\dim(V) = n$ and $v_1 \dots v_n$ a basis of V then $v \in V \longleftrightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n$

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

In the same way $f \longleftrightarrow [a_1 \dots a_n] \in \mathbb{K}^n$

$$f(v) = [a_1 \dots a_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_i a_i x_i$$

We can think of f as a linear function $f(x_1, \dots x_n) = \sum_{i=1}^{n} a_i x_i$
polynomials in n variables with all terms

(since a basis of degree 2 monomials is $\{ x_i x_j \mid i \leq j \}$)

6.1 Quadratic form functions

Our standing assumption for the rest of the chapter is $2 \neq 0$ in $\mathbb{K}$ ($\mathbb{K}$ is not "of characteristic 2") e.g. not $\mathbb{F}_2$.

Def 6.3 A quadratic form function in n variables $x_1, \ldots x_n$ over $\mathbb{K}$ is a polynomial of the form

$$q(x_1, \ldots x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

for some symmetric matrix $A = (a_{ij})$
(can also say $q_A(x_1, \ldots x_n)$ in this case)

Here $x_i x_j$ for $i \neq j$ enters twice in q with coefficients

$$c_{ij} = a_{ij} + a_{ji} = 2a_{ij} \quad i<j$$
$$c_{ii} = a_{ii}$$

of the basis elements $\{ x_i x_j \mid i \leq j \}$

w.l.o.g. we can assume that A is symmetric $A = A^t$. Conversely, for any desired coefficients c_{ij}, $i \leq j$, we can set

$$a_{ij} = \tfrac{1}{2} c_{ij} \quad i<j$$
$$a_{ii} = c_{ii}$$
$$a_{ij} = a_{ji} \quad i>j$$

as the symmetric matrix A.

of degree 1.

Definition 6.2 The set of all linear forms on V is a v.s., denoted V^*. If

$f, f_1, f_2 \in V^*$ $\quad (f_1+f_2)(v) = f_1(v)+f_2(v)$

$c \in \mathbb{K}$ $\quad (cf)(v) = c(f(v))$

$\forall v \in V$,

(similarly the set of linear maps $V \to W$ is a v.s.). V^* is called the "dual space" of V. By choosing a basis of V s.t. $V \cong \mathbb{K}^n$ column vectors, $V^* \cong \mathbb{K}^n$ row vectors

[Challenge question: how is $(V^*)^*$ related to V ?]

We will similarly study quadratic forms on a v.s. V both abstractly as a maps $q : V \longrightarrow \mathbb{K}$ which "depend quadratically on V", and as functions on x_i if we fix a basis of V.

$v \leftrightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $q(x_1 \ldots x_n)$ is a function of x_i which is a polynomial of degree 2 in each term,

$q(x_1, \ldots, x_n) = \sum_{i \leq j} c_{ij} x_i x_j$ for some coefficients c_{ij}

for a different matrix A'. We've proven:

<u>prop 6.4</u> A change of basis with transition matrix P replaces the symmetric matrix A representing the quadratic form q with respect to a basis to $P^T A P$.

Here $q: V \to \mathbb{K}$ is unchanged and exist independently of choice of basis, but the associated symmetric matrix representing q in the form $q_A(x_1, \ldots x_n)$ depends on the choice of basis. This is <u>different</u> from $\alpha: V \to V$ defining a matrix for any basis of V, hence behaves differently under a change of basis.

<u>Definition 6.5</u> Two (~~symmetric~~ symmetric) matrices A, A' over $\mathbb{K}$ are <u>congruent</u> if $A' = P^t A P$ for some invertible P.

<u>Prop 6.6</u> Two symmetric matrices are congruent iff they represent the same quadratic form with respect to possibly different bases (ie up to a linear change of variables)

<u>Remark</u> Congruence is an equivalence relation. (different from "similarity" relevant to $\alpha: V \to V$)

$\boxed{\text{So the data for a quadratic form function is just a symmetric matrix } A}$

In this case, if V is v.s. with basis $v_1 \dots v_n$ so $v \longleftrightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $v = \sum_i x_i v_i$, $x_i \in \mathbb{K}$

$q(x_1 \dots x_n)$ defines an abstract quadratic form $q : V \longrightarrow \mathbb{K}$,

$$\begin{aligned} q(v) &= q(x_1 \dots x_n) \\ &= \sum_{ij} a_{ij} x_i x_j \\ &= [x_1 \dots x_n] A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= \underbrace{\underline{x}^t}_{\text{row vector}} A \underbrace{\underline{x}}_{\text{column vector.}} \end{aligned}$$

If we change the basis of V with some transition matrix P then

$v \longleftrightarrow \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix}$ where $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = P \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix}$

replaces the x_i in terms of new x_i' as a linear change of variables.

Then
$$\begin{aligned} q(x_1 \dots x_n) = q_A(x_1 \dots x_n) &= [x_1 \dots x_n] A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= \left(P \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix}\right)^t A P \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix} \\ &= [x_1' \dots x_n'] \underbrace{P^T A P}_{A'} \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix} \\ &= q_{A'}(x_1' \dots x_n') \end{aligned}$$

Our next question is: can we put any quadratic form into a canonical shape ie any symmetric matrix A into "canonical form . for congruence"

6.2 Reduction of quadratic forms

<u>Theorem 6.7</u> Let q be a quadratic form function in n variables $x_1 \dots x_n$ over $\mathbb{K}$ (not of characteristic 2). By a suitable linear substitution to new variables $y_1 \dots y_n$ we can obtain $q = c_1 y_1^2 + \dots c_n y_n^2$ for some $c_i \in \mathbb{K}$.

Equivalently, if $q(x_1 \dots x_n) = q_A(x_1 \dots x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ then we can find P invertible such that $P^t A P = \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix}$ is diagonal.

i.e. every symmetric matrix is congruent to a diagonal one.