Matrices and polynomials

Motivation (from numerical methods)

Take $\sigma \in L(V,V)$ $\quad (A = [\sigma]_{\mathcal{A}})$

composition $\sigma \circ \sigma$

What is the dimension of $\mathrm{Span}\{\iota, \sigma, \sigma^2, \sigma^3, \ldots\} \subseteq L(V,V)$

(what is $\dim \mathrm{Span}\{I, A, A^2, A^3, \ldots\} \subseteq M_{\dim V, \dim V}$)

Note: if $\sigma^n = a_0\iota + a_1\sigma + a_2\sigma^2 + \cdots + a_{n-1}\sigma^{n-1}$ $\quad ☆$

$\in \mathrm{Span}\{\iota, \sigma, \sigma^2, \ldots, \sigma^{n-1}\}$

apply σ to both sides $\Rightarrow$ $\sigma^{n+1} = a_0\sigma + a_1\sigma^2 + \cdots + a_{n-1}\sigma^n$

$\in \mathrm{Span}\{\iota, \sigma, \ldots, \sigma^n\} = \mathrm{Span}\{\iota, \sigma, \ldots, \sigma^{n-1}\}$

$\because \sigma^n \in \mathrm{Span}\{\iota, \sigma, \ldots, \sigma^{n-1}\}$ (5.3.11)

similarly, $\sigma^{n+k} \in \mathrm{Span}\{\iota, \sigma, \ldots, \sigma^{n-1}\}$ $\forall k \geq 0$.

In ☆, move σ^n to the other side:

$$0 = a_0\iota + a_1\sigma + \cdots + a_{n-1}\sigma^{n-1} - \sigma^n$$

i.e. σ satisfies

$$a_0 + a_1x + \cdots + a_{n-1}x^{n-1} - x^n$$

So we are interested in:

what polynomials f does σ satisfy, i.e. $f(\sigma) = \vec{0}$?

zero function ($\vec{0} \in L(V,V)$)

(or $f(A) = 0$)

zero matrix

Th.5.2.4: Cayley-Hamilton theorem:

if $\sigma \in L(V,V)$ and $\dim V < \infty$

then σ satisfies χ_σ

i.e. $A \in M_{n,n}(\mathbb{F})$ satisfies χ_A.

What this does NOT mean:

$$\chi_A(x) = \det(A - xI)$$

$$\underbrace{\chi_A(A)}_{\text{a matrix}} \neq \det(A - AI)$$

$$= \underbrace{\det(\text{zero matrix}) = 0}_{\text{a number}}$$

Ex: if $A = \begin{pmatrix} 3 & 1 \\ 4 & 5 \end{pmatrix}$, then $\chi_A(x) = \begin{vmatrix} 3-x & 1 \\ 4 & 5-x \end{vmatrix}$

can't substitute in $x=A$ before taking determinant

can also substitute $x=A$ here $\longrightarrow$ $= (3-x)(5-x) - 4$

$= x^2 - 8x + 11$

So $\chi_A(A) = A^2 - 8A + 11 I_2$

$= \begin{pmatrix} 3 & 1 \\ 4 & 5 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 4 & 5 \end{pmatrix} - 8\begin{pmatrix} 3 & 1 \\ 4 & 5 \end{pmatrix} + 11\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$= \begin{pmatrix} 13 & 8 \\ 32 & 29 \end{pmatrix} - \begin{pmatrix} 24 & 8 \\ 32 & 40 \end{pmatrix} + \begin{pmatrix} 11 & 0 \\ 0 & 11 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

Sketch proof (in $\mathbb{R}$ or $\mathbb{C}$)

First assume A is diagonalisable.

i.e. $A = PDP^{-1}$

$$P^{-1}AP = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

$$\chi_A(x) = \chi_D(x) = \begin{vmatrix} \lambda_1 - x & & 0 \\ & \ddots & \\ 0 & & \lambda_n - x \end{vmatrix}$$

$\because A, D$ similar

$$= (\lambda_1 - x)\cdots(\lambda_n - x).$$

For clarity, let $n=3$

so $\chi_A(x) = \chi_D(x) = (\lambda_1 - x)(\lambda_2 - x)(\lambda_3 - x)$.

$$\chi_A(D) = (\lambda_1 I - D)(\lambda_2 I - D)(\lambda_3 I - D)$$

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

$$= \begin{pmatrix} \boxed{0}(\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1) & & \\ & (\lambda_1 - \lambda_2)\boxed{0}(\lambda_3 - \lambda_2) & \\ & & (\lambda_1 - \lambda_3)(\lambda_2 - \lambda_3)\boxed{0} \end{pmatrix}$$

= zero matrix

Why $\chi_A(A) = 0$ also:

let $\chi_A = a_0 + a_1 x + \cdots + a_n x^n$

then $\chi_A(A) = a_0 I + a_1 A + \cdots + a_n A^n$

$$= a_0 I + a_1 (PDP^{-1}) + a_2 (PDP^{-1})^2 + \cdots + a_n (PDP^{-1})^n$$

$$\underbrace{(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})}_{n \text{ times}}$$

$$= a_0 PP^{-1} + a_1 PDP^{-1} + a_2 PD^2P^{-1} + \cdots + a_n PD^nP^{-1}$$

$$= P(a_0 I + a_1 D + a_2 D^2 + \cdots + a_n D^n)P^{-1}$$

$$= P\chi_A(D)P^{-1}$$

$$= P0P^{-1} = 0.$$

If A is not diagonalisable, then "approximate A by diagonalisable matrices (over $\mathbb{C}$) and use continuity".