Polynomial functions of endomorphisms

Aim lecture: We look at polynomial functions of endomorphisms which not only arise naturally, but are useful for understanding the endomorphisms themselves. In particular, we give the Cayley-Hamilton theorem.

Defn

Let $p(x) = \sum_{i=0}^{d} p_i x^i \in \mathbb{F}[x]$ be a polynomial & $T : V \longrightarrow V$ be linear. We define $p(T) = \sum_{i=0}^{d} p_i T^i$ where $T^0 = \text{id}$ & $T^n = T \circ \ldots \circ T$ is the composite of T with itself n times. This is a linear map from $V \longrightarrow V$.

E.g. If $p(x) = x^2 - 3x + 2$, then

$$p(\frac{d}{dx}) =$$

More examples of poly fns of endomorphisms

E.g. 1 If $R : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is reflection about a line $\mathbb{R}\mathbf{v}$ then we may simplify

$$4R^4 + 3R^3 + 5R =$$

E.g. 2 Let $N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

$$I + 2N + 3N^2 + N^3 =$$

Polynomial functions & co-ordinate change

Prop

Let $p(x) = \sum_{i=0}^{d} p_i x^i \in \mathbb{F}[x]$ & $C : W \longrightarrow V$ be an isomorphism of $\mathbb{F}$-spaces. For any linear $T : V \longrightarrow V$, we have

$$p(C^{-1} \circ T \circ C) = C^{-1} \circ p(T) \circ C.$$

In particular, if C is a co-ord system & A the matrix representing T wrt C, then the matrix representing $p(T)$ is $p(A)$.

Proof. We first verify the case $p(x) = x^n$.

$$(C^{-1} \circ T \circ C)^n = \qquad\qquad\qquad\qquad .$$

Hence, in general we have

$$p(C^{-1} \circ T \circ C) = \sum_{i=0}^{d} p_i (C^{-1} \circ T \circ C)^i = \sum_{i=0}^{d} p_i (C^{-1} \circ T^i \circ C) = C^{-1} \circ p(T) \circ C.$$

Cayley-Hamilton theorem

Theorem

Let $V =$ fin dim $\mathbb{F}$-space & $T : V \longrightarrow V$ be a linear map. Then $cp_T(T) = 0$.

Proof in case $\mathbb{F} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ only. (The same proof applies in general given the existence of algebraic closure). Suppose $C : \mathbb{F}^n \longrightarrow V$ is a co-ord system & $A = C^{-1} \circ T \circ C$ the matrix representing T. If we know the thm holds for the matrix A, we are done since then

$$C^{-1} \circ \mathrm{cp}_T(T) \circ C = \mathrm{cp}_T(A) = \mathrm{cp}_A(A) = 0$$

so we must have $\mathrm{cp}_T(T) = 0$.

We may thus assume $T = A$. Since any real or rational matrix is also a complex matrix, we may also assume $\mathbb{F} = \mathbb{C}$ & further, by the above argument, replace A with any similar matrix. By triangularisation, we may assume $A = (a_{ij})$ is upper triangular.

proof completed

Now in this case we see $\mathrm{cp}_A(\lambda) = (a_{11} - \lambda)(a_{22} - \lambda)\dots(a_{nn} - \lambda)$ so it suffices to show that

$$0 = (a_{11}I_n - A)(a_{22}I_n - A)\dots(a_{nn}I_n - A) \qquad (*)$$

Note that each factor $a_{ii}I_n - A$ is upper triangular with a 0 in the (i,i)-th entry. It suffices now to complete the following easy exercise in induction

Lemma

The bottom $n - i + 1$ rows of the matrix $(a_{ii}I_n - A)(a_{i+1,i+1}I_n - A)\dots(a_{nn}I_n - A)$ are zero.

Why? (Draw picture)

Example using the Cayley-Hamilton theorem

E.g. Consider the matrix

$$A = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}.$$

Write A^4 as a linear combn of A & I.

Criterion for diagonalisability via algebraic multiplicity

Prop

Let $A \in M_{nn}(\mathbb{F})$. Suppose the e-values are $\lambda_1, \ldots, \lambda_r$ with geometric multiplicities $n_1, \ldots, n_r$ & algebraic multiplicities $a_1, \ldots, a_r$.

1. For all i we have $n_i \leq a_i$.
2. A is diagonalisable over $\mathbb{F}$ iff $\text{cp}_A(\lambda)$ factors into linears over $\mathbb{F}$ & $n_i = a_i$ for all i.

Proof. Note 2) ($\Longrightarrow$) is easy whilst 2)($\Longleftarrow$) follows from 1) & our old criterion for diagonalisability & the fact that $\text{cp}_A(\lambda)$ factors into linears over $\mathbb{F} \Longrightarrow \sum_i a_i = n$.

Proof completed

We prove 1). Since geometric & algebraic multiplicities are similarity invariants, we may replace A with a similar matrix. Let $E \leq \mathbb{F}^n$ be the sum of the e-spaces & E' be a vector space complement to E in $\mathbb{F}^n$. Changing to a co-ord system adapted to $\mathbb{F}^n = E \oplus E'$ we may assume $A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$ is block upper triangular with $A_{11} = \lambda_1 I_{n_1} \oplus \ldots \oplus \lambda_r I_{n_r}$.

Now

$$\mathrm{cp}_A(\lambda) = \det \begin{pmatrix} A_{11} - \lambda I & A_{12} \\ 0 & A_{22} - \lambda I \end{pmatrix} = \det(A_{11} - \lambda I) \det(A_{22} - \lambda I)$$

(ex or by triangularising A_{22}). But clearly $(\lambda - \lambda_i)^{n_i} \mid \det(A_{11} - \lambda I)$ so $n_i \leq a_i$.