Diagonalizability of Linear Operators

Definition: Let $F[x]$ be the set of polynomials with coefficients from the field F. $H \subset F[x]$ is called an **ideal** if

- H is closed under addition.
- If $f \in F[x]$ and $h \in H$ then $hf \in H$.

Definition: If V is a finite dimemsional vector space over the field F and $T \in L(V, V)$, let $H_T = \{f \in F[x] : f(T)(\alpha) = \vec{0} \text{ for all } \alpha \in V\}$. H_T is called the **annihilator** of T.

Definition: A polynomial is said to be monic if its leading coefficient is 1.

Theorem: If H_T is the annihilator of T there is a unique monic polynomial m such that for each $h \in H_t$ there is some $q \in F[x]$ such that $h = mq$. m is the non-zero polynomial of least degree such that $m(T)$ is the zero linear operator.

Definition: The polynomial in the preceding theorem is called the **minimal polynomial** of T and will be denoted by m_T.

Cayley-Hamilton Theorem: Let f_T be the characteristic polynomial of T. Then $f_T \in H_T$.

Theorem: $f_T(c) = 0$ if and only if $m_T(c) = 0$. Therefore the roots of m_T are exactly the eigenvalues of T.

Theorem: Suppose that

$$m_T(x) = \prod_{j=1}^{k}(x - c_j)$$

for k distinct scalars c_j. Then T is diagonalizable.

Proof: If $k = 1$ then $T = c_1 I$ and we are done. Suppose now that $k \geq 2$. For each $i \in \{1, \ldots, k\}$ put

$$\begin{aligned} q_i(x) &= \prod_{j \neq i}(x - c_j) \\ p_i(x) &= q_i(x)/q_i(c_i) \end{aligned}$$

The polynomials $p_1, \ldots p_k$ are linearly independent and span the vector space of polynomial of degree $k-1$ or less. Their dual basis are the evaluation functionals telling one to evaluate polynomials at c_1, c_2, $\ldots$, c_k. Therefore

$$\begin{aligned} 1 &= \sum_{j=1}^{k} p_j(x) \\ x &= \sum_{j=1}^{k} c_j p_j(x) \end{aligned}$$

Put $E_j = p_j(T)$ for $j \in \{1, \ldots, k\}$. Then

$$\begin{aligned} I &= \sum_{j=1}^{k} E_j \\ T &= \sum_{j=1}^{k} c_j E_j \end{aligned}$$

Since $p_j p_i$ is divisible by m_T for $j \neq i$ we have $E_i \circ E_j$ is the zero operator for $i \neq j$. On the other hand, for any $\alpha \in V$,

$$E_j(\alpha) = E_j(I(\alpha)) = E_j(E_1(\alpha)) + \cdots + E_j(E_k(\alpha)) = E_j^2(\alpha).$$

Therefore, if $\alpha \in V$ then

$$T(E_j(\alpha)) = c_1 E_1(E_j(\alpha)) + \cdots + c_k E_k(E_j(\alpha)) = c_j E_j(\alpha)$$

so either $E_j(\alpha) = \vec{0}$ or $E_j(\alpha)$ is an eigenvector of T. Therefore the range of E_j is contained in the eigenspace of T corresponding to c_j: $W_j = \{\alpha \in V : T(\alpha) = c_j \alpha\}$.

If we let R_j denote the range of E_j then

$$V = \bigoplus_{j=1}^{k} R_j \subset \bigoplus_{j=1}^{k} W_j \subset V$$

so

$$\bigoplus_{j=1}^{k} W_j = V$$

so T is diagonalizable.