Invariant Subspaces And Direct Sums :-

It is useful to decompose a finite dimensional vector space V into a direct sum of as many T- invariant subspaces as possible because the behaviour of T on V can be inferred from its behaviour on the direct summands. For example, T is diagonalizable if and only if V can be decomposed into a direct sum of one-dimensional T- invariant subspaces.

Theorem 5.24. Let T be a linear operator on a finite-dimensional vector space V, and suppose that $V = W_1 \oplus W_2 \oplus \cdots \oplus W_K$, where W_i is a T- invariant subspace of V for each i $(1 \le i \le k)$. Suppose that $f_i(t)$ is the characteristic polynomial of T_{W_i} $(1 \le i \le K)$. Then $f_1(t) \cdot f_2(t) \cdots f_K(t)$ is the characteristic polynomial of T.

Proof. The proof is by mathematical induction on k. Let $f(t)$ denotes the characteristic polynomial of T. Suppose first that $k=2$. Let β_1 be an ordered basis for W_1, β_2 an ordered basis for W_2, and $\beta = \beta_1 \cup \beta_2$. Then β is an ordered basis for V by Theorem 5.10 (d). Let $A = [T]_\beta$, and $B_1 = [T_{W_1}]_{\beta_1}$ and $B_2 = [T_{W_2}]_{\beta_2}$. It follows that

$$\Rightarrow A = \begin{bmatrix} B_1 & O \\ O' & B_2 \end{bmatrix}$$

where O and O' are zero matrices of the appropriate sizes.

Then $f(t) = \det(A - tI)$

$$= \det\begin{pmatrix} B_1 - tI_1 & O \\ O' & B_2 - tI_2 \end{pmatrix}$$

$$= \det(B_1 - tI_1) \cdot \det(B_2 - tI_2)$$

$$\Rightarrow f(t) = f_1(t) \cdot f_2(t)$$

Here I_1 and I_2 are identity matrices of appropriate sizes. Since $f(t) = f_1(t) \cdot f_2(t)$. This proves the result for $K = 2$.

Now assume that the theorem is valid for $K-1$ summands, where $K-1 \geq 2$ and suppose that V is a direct sum of K subspaces, say

$$V = W_1 \oplus W_2 \oplus \dots \oplus W_K$$

Let $W = W_1 \oplus W_2 \oplus \dots \oplus W_{K-1}$. Since $W_1, W_2, \dots, W_{K-1}$ are T-invariant subspaces, W is also a T-invariant subspace of V and $V = W \oplus W_K$. So by the case for $K = 2$, $f(t) = g(t) \cdot f_K(t)$, where $g(t)$ is the characteristic polynomial of T_W. Clearly $W = W_1 \oplus W_2 \oplus \dots \oplus W_{K-1}$, and therefore $g(t) = f_1(t) \cdot f_2(t) \dots f_{K-1}(t)$ by the induction hypothesis. We conclude

$$f(t) = g(t) \cdot f_K(t)$$
$$= f_1(t) \cdot f_2(t) \dots, f_{K-1}(t) \cdot f_K(t).$$

As an illustration of this result, suppose that T is a diagonalizable linear operator on a finite-dimensional vector space V with distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_K$. By Theorem 5.11, V is a direct sum of the eigenspaces of T. Since each eigenspace is T-invariant, we can use Theorem 5.24. For each eigenvalue λ_i, the restriction of T to E_{λ_i} has characteristic polynomial $(\lambda_i - t)^{m_i}$, where m_i is the dimension of E_{λ_i}. By Theorem 5.24, the characteristic polynomial of T is the product

$$f(t) = (\lambda_1 - t)^{m_1} (\lambda_2 - t)^{m_2} \dots (\lambda_K - t)^{m_K}$$

It follows the multiplicity of each eigenvalue is equal to the dimension of the corresponding eigenspace.

Example 8. Let T be the linear operator on R^4 defined by

$$T(a,b,c,d) = (2a-b,\ a+b,\ c-d,\ c+d)$$

and let $W_1 = \{(s,t,0,0) : s,t \in R\}$ and

$$W_2 = \{(0,0,s,t) : s,t \in R\}.$$

for any $(a,b,0,0) \in W_1 \Rightarrow T(a,b,0,0) \in W_1$ since

$$T(a,b,0,0) = (2a-b,\ a+b,\ 0,\ 0)$$

$\Rightarrow W_1$ is T-invariant.

Also for any $(0,0,c,d) \in W_2 \Rightarrow T(0,0,c,d) \in W_2$ since

$$T(0,0,c,d) = (0,0,\ c-d,\ c+d)$$

$\Rightarrow W_2$ is also T-invariant.

Now, any $(a,b,c,d) \in R^4$ can be written as

$$(a,b,c,d) = (a,b,0,0) + (0,0,c,d)$$

where $(a,b,0,0) \in W_1$ and $(0,0,c,d) \in W_2$

$\Rightarrow R^4 = W_1 + W_2$

since $W_1 \cap W_2 = \{0,0,0,0\}$

$\Rightarrow \underline{R^4 = W_1 \oplus W_2}$.

Let $\beta_1 = \{e_1, e_2\}$, $\beta_2 = \{e_3, e_4\}$, and $\beta = \beta_1 \cup \beta_2 = \{e_1, e_2, e_3, e_4\}$

Then β_1 is an ordered basis for W_1, β_2 is an ordered basis for W_2, and β is an ordered basis for R^4. Let $\underline{A} = [T]_\beta$, $\underline{B_1} = [T_{W_1}]_{\beta_1}$, and $\underline{B_2} = [T_{W_2}]_{\beta_2}$. Then

$$T(e_1) = T(1,0,0,0) = (2,1,0,0) = 2e_1 + e_2$$

$$T(e_2) = (-1,1,0,0) = -e_1 + e_2$$

$$T(e_3) = (0,0,1,1) = e_3 + e_4$$

$$T(e_4) = (0,0,-1,1) = -e_3 + e_4$$

$$\Rightarrow B_1 = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

and

$$A = \begin{pmatrix} B_1 & O \\ O & B_2 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Let $f(t)$, $f_1(t)$, and $f_2(t)$ denote the characteristic polynomial of T, T_{W_1}, and T_{W_2}, respectively. Then

$$f(t) = \det(A - tI)$$

$$= \det(B_1 - tI) \cdot \det(B_2 - tI)$$

$$\Rightarrow f(t) = f_1(t) \cdot f_2(t).$$

Definition. Let $B_1 \in M_{m \times m}(F)$ and let $B_2 \in M_{n \times n}(F)$. We define the direct sum of B_1 and B_2, denoted by $B_1 \oplus B_2$, as the $(m+n) \times (m+n)$ matrix A such that

$$A_{ij} = \begin{cases} (B_1) & \text{for } 1 \le i, j \le m \\ (B_2)_{(i-m),(j-m)} & \text{for } m+1 \le i, j \le n+m \\ 0 & \text{otherwise} \end{cases}$$

If $B_1, B_2, \ldots, B_K$ are square matrices with entries from F, then we define the direct sum of $B_1, B_2, \ldots, B_K$ recursively by

$$B_1 \oplus B_2 \oplus \cdots \oplus B_K = (B_1 \oplus B_2 \oplus \cdots \oplus B_{K-1}) \oplus B_K.$$

If $A = B_1 \oplus B_2 \oplus \cdots \oplus B_K$, then we often write

$$A = \begin{bmatrix} B_1 & O & \cdots & O \\ O & B_2 & \cdots & O \\ \vdots & \vdots & & \vdots \\ O & O & \cdots & B_n \end{bmatrix}$$

Recursion → An expression such that each term is generated by repeating a particular mathematical operation.

Example 9. Let

$$B_1 = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}, \quad B_2 = (3), \quad \text{and} \quad B_3 = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}$$

Then

$$B_1 \oplus B_2 \oplus B_3 = \left(\begin{array}{cc|c|ccc} 1 & 2 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{array}\right)$$

Theorem 5.25. Let T be a linear operator on a finite-dimensional vector space V, and let $W_1, W_2, \ldots, W_K$ be T-invariant subspaces of V such that $V = W_1 \oplus W_2 \oplus \cdots \oplus W_K$. For each i, let β_i be an ordered basis for W_i, and let $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_K$. Let $A = [T]_\beta$ and $B_i = [T_{W_i}]_{\beta_i}$ for $i = 1, 2, \ldots, k$. Then $A = B_1 \oplus B_2 \oplus \cdots \oplus B_K$.

Matrix Polynomial :→

An expression of the form

$$F(x) = A_0 + A_1 x + A_2 x^2 + \dots + A_m x^m$$

where $A_0, A_1, \dots A_m$ are matrices each of order $n \times n$ over a field F, is called a matrix polynomial of degree m, provided $A_m \neq 0$.

Example. Let

$$A = \begin{bmatrix} 1+2x+3x^2 & x^2 & 6-4x \\ 1+x^3 & 3-4x^2 & 1-2x+4x^3 \\ 2-3x+2x^3 & 5 & 6x^2+7x \end{bmatrix}$$

then, A can be written as

$$A = \begin{bmatrix} 1 & 0 & 6 \\ 1 & 3 & 1 \\ 2 & 5 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 0 & -4 \\ 0 & 0 & -2 \\ -3 & 0 & 7 \end{bmatrix} x + \begin{bmatrix} 3 & 1 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 6 \end{bmatrix} x^2 + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 4 \\ 2 & 0 & 0 \end{bmatrix} x^3$$

which is a matrix polynomial of degree 3.

Proof of Cayley-Hamilton Theorem For Matrices. Let A be an $n \times n$ matrix, and let $f(t)$ be the characteristic polynomial of A. Then $f(A) = 0$, the $n \times n$ zero matrix.

Let $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$

then the characteristic polynomial of A is

$$f(t) = \det(A - tI) = \det \begin{bmatrix} a_{11}-t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-t & \dots & a_{2n} \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} & \dots & a_{nn}-t \end{bmatrix}$$

Therefore, $f(t)$ can be written as

$$f(t) = (-1)^n \left[t^n + a_1 t^{n-1} + a_2 t^{n-2} + \dots + a_n \right]$$

where $a_1, a_2, \dots, a_n \in F$.

now to prove that $f(A) = O$, we must show that

$$\Rightarrow A^n + a_1 A^{n-1} + a_2 A^{n-2} + \dots + a_n I = O$$

Since all the elements of $A - tI$ are atmost of first degree in t, all the elements of $\text{adj}(A - tI)$ are polynomials in t of degree $(n-1)$ or less. (since elements of $\text{adj}(A - tI)$ are cofactors of the elements of $(A - tI)$).

$\therefore$ Hence, $\text{adj}(A - tI)$ can be written as a matrix polynomial in t of degree $(n-1)$

let $\text{adj}(A - tI) = B_0 t^{n-1} + B_1 t^{n-2} + \dots + B_{n-2} t + B_{n-1}$

where $B_0, B_1, B_2, \dots, B_{n-1}$ are square matrices of order $n \times n$.

Now we have

$$(A - tI) \, \text{adj}(A - tI) = \det(A - tI) \, I.$$

$$\left(\text{since } A \cdot \text{adj} A = \det(A) I \right)$$

$$\Rightarrow (A - tI)(B_0 t^{n-1} + B_1 t^{n-2} + \dots + B_{n-2} t + B_{n-1}) = \det(A - tI) \cdot I$$

$$\Rightarrow (A - tI)(B_0 t^{n-1} + B_1 t^{n-2} + \dots + B_{n-2} t + B_{n-1}) = (-1)^n \left[t^n + a_1 t^{n-1} + \dots + a_n \right] \cdot I$$

comparing coefficients of like powers of t, we obtain

$$-B_0 = (-1)^n I$$

$AB_0 - B_1 = (-1)^n a_1 I$

$AB_1 - B_2 = (-1)^n a_2 I$

$- - - - - - - - -$

$- - - - - - - - -$

$- AB_{n-1} = (-1)^n a_n I$

pre-multiplying the above equations ~~the~~ successively by $A^n, A^{n-1}, \ldots, I$ and adding, we obtain

$$O = (-1)^n [A^n + a_1 A^{n-1} + a_2 A^{n-2} + \ldots + a_n I]$$

$$\Rightarrow A^n + a_1 A^{n-1} + a_2 A^{n-2} + \ldots + a_n I = O \Rightarrow f(A) = O.$$

where 'O' is the zero matrix of order $n \times n$

Hence, A satisfies its characteristic equation.

Exercise 15. Use the Cayley-Hamilton theorem (Theorem 5.23) to prove its corollarly for matrices.

Theorem 5.23. Let T be a linear operator on a finite-dimensional vector space V, and let f(t) be the characteristic polynomial of T. Then ~~f(T)~~ $f(T) = T_0$, the zero transformation. That is, T "satisfies" its characteristic equation.

Solution. To prove this theorem for matrices, consider the linear transformation $L_A : F^n \to F^n$ such that $L_A(v) = Av$.

Let $\beta = \{e_1, e_2, \ldots, e_n\}$ ~~be the~~

be the standard basis for F^n. Then

$$\boxed{[L_A]_\beta = A}$$

Hence, A and L_A have the same characteristic polynomial $f(t)$. Let

$$f(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n \quad -(1)$$

Since L_A is a linear transformation, hence, from Theorem 5.23, it will satisfy its characteristic equations

$$\Rightarrow f(L_A) = (L_A)_0$$

$$\Rightarrow f(L_A)(v) = 0 \quad \text{for all } v \in F^n$$

$$\Rightarrow f(L_A)(v) = O(v).$$

$$\Rightarrow a_0 v + a_1 L_A(v) + a_2 L_A^2(v) + \cdots + a_n L_A^n(v) = O(v). -(2)$$

now $\quad L_A(v) = Av \Rightarrow L_A^2(v) = L_A(L_A(v)) = L_A(Av)$

$$= A \cdot Av = A^2 v$$

similarly $L_A^3(v) = L_A(L_A^2(v)) = L_A(A^2(v)) = A \cdot A^2 v = A^3 v$

and so on $\Rightarrow L_A^n(v) = A^n(v)$

substituting these values in equation (2), we get

$$\Rightarrow a_0 v + a_1 Av + a_2 A^2(v) + \cdots + a_n A^n v = O(v) = 0$$

$$\Rightarrow (a_0 + a_1 A + a_2 A^2 + \cdots + a_n A^n)(v) = 0$$

$$\Rightarrow f(A)(v) = 0 \quad \text{for all } v \in F^n$$

Thus $\quad f(A) = O$, the zero matrix of order $n \times n$.

Exercise 36. Let T be a linear operator on a finite-dimensional vector space V. Prove that T is diagonalizable if and only if V is the direct sum of one-dimensional T-invariant subspaces.

Solution. Let $\dim V = n$.

Suppose that V is the direct sum of one dimensional T-invariant subspaces, that is

$$V = W_1 \oplus W_2 \oplus \dots \oplus W_k$$

since W_i $(1 \le i \le k)$ is one-dimensional T-invariant subspace.

$$W_i = \text{span}\{v_i\} \quad \text{and} \quad T(W_i) \subseteq W_i$$

since $v_i \in W_i \Rightarrow T(v_i) \in W_i$

$\Rightarrow T(v_i) = \lambda v_i$ (linear combination of element of basis)

$\Rightarrow v_i$ is an eigenvector (for $1 \le i \le k$)

Thus $\beta_1 = \{v_1\}, \beta_2 = \{v_2\}, \dots, \beta_k = \{v_k\}$ are ordered bases of the subspaces $W_1, W_2, \dots, W_k$ respectively. Hence, from Theorem 5.10(d) $\beta_1 \cup \beta_2 \cup \dots \cup \beta_k = \{v_1, v_2, \dots, v_k\}$ is an ordered basis for V. But $\dim V = n$ $\Rightarrow$ Thus $k = n$

$\Rightarrow \beta = \{v_1, v_2, \dots, v_n\}$ is an ordered basis for V

where all v_i $(1 \le i \le n)$ are eigenvectors of T.

$\Rightarrow$ Thus there exists an ordered basis β for V consisting of eigenvectors of T. Hence, using Theorem 5.1, T is diagonalizable.

Now, suppose that T is diagonalizable, then there will exist an ordered basis for V consisting of eigenvectors of T.

Let $\beta = \{v_1, v_2, \dots, v_n\}$

be an ordered basis for V consisting of eigenvectors of T.

Now consider $W_i = \text{span}\{v_i\}$

$= \{\lambda v_i, \lambda \in F\}$

now for all $\lambda v_i \in W_i$

$T(\lambda v_i) = \lambda T(v_i)$

$= \lambda \cdot \lambda_i v_i$ (since v_i is an eigenvector

$= (\lambda \lambda_i) v_i \in W_i$ (since $\lambda \lambda_i \in F$)

$\Rightarrow T(\lambda v_i) \in W_i$ for all i ($1 \le i \le n$)

$\Rightarrow W_i$ is a T-invariant subspace of V and $\dim(W_i) = 1$

let $\beta_1 = \{v_1\}$, $\beta_2 = \{v_2\}$, $\dots$, $\beta_n = \{v_n\}$ are ordered bases

for the ~~subst~~ subspaces $W_1, W_2, \dots, W_n$ respectively and

$\beta_1 \cup \beta_2 \cup \dots \cup \beta_n = \{v_1, v_2, \dots, v_n\}$ is an ordered basis

for V. Then, from Theorem 5.10 (a) and (d), we

can say that

$V = W_1 \oplus W_2 \oplus \dots \oplus W_n$

$\Rightarrow$ V is the direct sum of one-dimensional T-invariant

subspaces of V.