Project report on structure of Linear Maps

Jishu Das
Indian Institute of Science Education and Research (IISER), Kolkata
E-mail Id- jd13ms109@iiserkol.ac.in

January 7, 2015

Abstract

This is an project report about structure of linear maps which I studied under Dr. Shameek Paul of Centre for Excellence in Basic Sciences(UM-DAE CBS), Mumbai as a guide during the period of time from 14 December 2014 to 09 January 2014. I would like to thank Dr. Shameek Paul by giving his valuable time to guide me.

Signature of Guide
Dr. Shameek Paul

Signature of Student
Jishu Das

Definition 1:- Consider a linear map $T : V \to W$. Let $B_V = \{v_1, v_2, ..., v_n\}$ and $B_W = \{w_1, w_2, ..., w_m\}$ form a basis for V and W respectively. For a fixed vector $v_j \in B_V$ we can always write $T(v_j)$ as a unique linear combination of basis vectors for W. i.e. $T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$ for some scalars $a_{ij} \in R$. We call this (a_{ij}) as the coordinates of $T(v_j)$ w.r.t. basis B_W. Note that this (a_{ij}) is a collection of m no of scalars which can be seen as a column matrix in usual sense. Similarly we can write the coordinates of $T(v_j)$ for all $1 \leq j \leq n$. On putting these obtained coordinates (to be precise n no of coordinates of n vectors in W) together in a particular order we get a matrix of order $m \times n$ (as there are n column matrices with each matrix containing m no of scalars). We denote this obtained matrix by $[M(T)]_{B_W}^{B_V}$ whose jth column is the coordinates of $T(v_j)$ w.r.t. basis B_W.
Hence our $[M(T)]_{B_W}^{B_V}$ looks like

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

Let us see an application of this matrix $[M(T)]_{B_W}^{B_V}$.
We can uniquely write $v \in V$ as a linear combination of vectors in B_V. i.e. $v = \sum_{j=1}^{n} a_j v_j$. v w.r.t. B_V looks like

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

We denote it by $[v]_{B_V}$. Similarly We can also write $T(v)$ uniquely as $T(v) = \sum_{i=1}^{m} b_i w_i$. Hence $T(v)$ w.r.t. B_W looks like

$$\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

We denote it by $[T(v)]_{B_W}$. We claim that $[M(T)]_{B_W}^{B_V}[v]_{B_V} = [T(v)]_{B_W}$
Proof:- Let $T(v_j) = \sum_{i=1}^{m} c_{ij} w_i$ such that c_{ij} is the ith row and jth column of matrix $[M(T)]_{B_W}^{B_V}$.
$v = \sum_{j=1}^{n} a_j v_j$
$\Rightarrow T(v) = T(\sum_{j=1}^{n} a_j v_j)$ (applying T on both sides)
$\Rightarrow T(v) = \sum_{j=1}^{n} a_j T(v_j)$ (as T is a linear map)
$\Rightarrow \sum_{i=1}^{m} b_i w_i = \sum_{j=1}^{n} a_j (\sum_{i=1}^{m} c_{ij} w_i)$
$\Rightarrow \sum_{i=1}^{m} b_i w_i = \sum_{j=1}^{n} (\sum_{i=1}^{m} a_j c_{ij} w_i)$ (W is distributive over a scalar among vectors)
$\Rightarrow \sum_{i=1}^{m} b_i w_i = \sum_{j=1}^{n} (\sum_{i=1}^{m} a_j c_{ij}) w_i$ (W is distributive over a vector among scalars)
$\Rightarrow \sum_{i=1}^{m} b_i w_i = \sum_{i=1}^{m} (\sum_{j=1}^{n} a_j c_{ij}) w_i$ (W is associative over addition)
$\Rightarrow \sum_{i=1}^{m} b_i w_i = \sum_{i=1}^{m} (\sum_{j=1}^{n} c_{ij} a_j) w_i$ (R is commutative over addition)
$\Rightarrow \sum_{i=1}^{m} (b_i - \sum_{j=1}^{n} c_{ij} a_j) w_i = \theta_W$
$\Rightarrow b_i = \sum_{j=1}^{n} c_{ij} a_j$ for $1 \leq i \leq m$ (as B_W is linearly independent in W)

$\Rightarrow b_{i1} = \sum_{j=1}^{n} c_{ij} a_{j1}$. This proves our claim.
Let us look at an example.
Consider a linear map $R^3 \to R^2$ defined by $T(x,y,z) = (x+y, y)$. $B_V = \{(1,0,0),(0,1,0),(0,0,1)\}$ be basis for R^3 and $B_W = \{(1,0),(0,1)\}$ be basis for R^2. Take $v = (1,2,3) \in V$ on direct computation we get $T(1,2,3) = (3,2)$. Let us try the other way round. $(1,2,3) = 1(1,0,0)+2(0,1,0)+3(0,0,1)$. Hence
$[v]_{B_V} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$
$T(1,0,0) = (1,0) = 1(1,0)+0(0,1)$
$T(0,1,0) = (1,1) = 1(1,0)+1(0,1)$
$T(0,0,1) = (0,0) = 0(1,0)+0(0,1)$
$[M(T)]_{B_W}^{B_V} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$
$[M(T)]_{B_W}^{B_V}[v]_{B_V} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} = [T(v)]_{B_W}$
Hence $T(v) = 3(1,0)+2(0,1) = (3,2)$ which clearly matches with direct computation.

Result 1:-
Consider a vector space V over R. Let $B_1 = \{v_1, v_2, ..., v_n\}$ and $B_2 = \{w_1, w_2, ..., w_n\}$ are two basis for V. Say we know the coordinates of a vector $v \in V$ w.r.t. basis B_1. Question that comes obvious to mind that what will be the coordinates of v w.r.t. basis B_2. We intuitively consider Identity linear map from V to itself in a manner that we express coordinates of v w.r.t. basis B_1 and express its image (which is same as v) coordinates w.r.t. basis B_2 and try to find out how both of them are related.
$I(v_j) = v_j$
$\Rightarrow \sum_{i=1}^{n} b_{ij} w_i = v_j$
$v = \sum_{j=1}^{n} a_j v_j$
$\Rightarrow I(v) = I(\sum_{j=1}^{n} a_j v_j)$
$\Rightarrow \sum_{i=1}^{n} c_i w_i = \sum_{j=1}^{n} a_j v_j$
$\Rightarrow \sum_{i=1}^{n} c_i w_i = \sum_{j=1}^{n} a_j (\sum_{i=1}^{n} b_{ij} w_i)$
$\Rightarrow \sum_{i=1}^{n} (c_i) w_i = \sum_{i=1}^{n} (\sum_{j=1}^{n} b_{ij} a_j) w_i$ (on rearranging)
$\Rightarrow c_i = \sum_{j=1}^{n} b_{ij} a_j$
i.e. $[M(I)]_{B_2}^{B_1}[v]_{B_1} = [v]_{B_2}$

Result 2:-
Let U, V and W be three vector spaces. Consider three linear maps $T_1 : W \to V$, $T_2 : U \to W$ and $T_3 : U \to V$ such that $T_1 \circ T_2 = T_3$. Let $B_U = \{u_1, ..., u_r\}$ be a basis for U, $B_V = \{v_1, ..., v_n\}$ be a basis for V and $B_W = \{w_1, ..., w_m\}$ be a basis for W. We wish to find a relation between $[M(T_1)]_{B_V}^{B_W}$, $[M(T_2)]_{B_W}^{B_U}$ and $[M(T_3)]_{B_V}^{B_U}$
Claim:- $[M(T_2)]_{B_W}^{B_U}[M(T_1)]_{B_V}^{B_W} = [M(T_3)]_{B_V}^{B_U}$
Proof:- For $w_k \in B_W$, $T_1(w_k) = \sum_{i=1}^{n} a_{ik} v_i$, for $u_j \in B_U$, $T_2(u_j) = \sum_{k=1}^{m} b_{kj} w_k$ and for $u_j \in B_U$, $T_3(u_j) = \sum_{i=1}^{n} c_{ij} v_i$
$\Rightarrow \sum_{i=1}^{n} c_{ij} v_i = T_3(u_j) = T_1 \circ T_2(u_j)$

$= T_1(\sum_{k=1}^{m} b_{kj} w_k)$
$= \sum_{k=1}^{m} b_{kj} T_1(w_k)$
$= \sum_{k=1}^{m} b_{kj} (\sum_{i=1}^{n} a_{ik} v_i)$
$= \sum_{i=1}^{n} (\sum_{k=1}^{m} a_{ik} b_{kj}) v_i$ (on rearranging)
$\Rightarrow c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}$ (as B_V form a basis for V)

Definition 2:- Let V be a vector space over F. A map $T : V \to V$ is said to be diagonalisable if there exists a basis B_V for V such that $[M(T)]_{B_V}$ is a diagonal matrix.

Definition 3:- Let V be a vector space over F. $T : V \to V$ be a linear map. $\lambda \in F$ is said to be an eigenvalue of T if there exists $x \in V \setminus \{\theta_V\}$ such that $T(x) = \lambda x$. We call this $x \in V \setminus \{\theta_V\}$ eigenvector corresponding to eigenvalue λ.
Consider $E_\lambda (\subset V) = \{x \in V : T(x) = \lambda x\}$. Clearly $\theta_V \in E_\lambda$. Also it follows from definition that $E_\lambda \neq \{\theta_V\}$ iff λ is an eigenvalue of T. We claim that E_λ is a subspace of V. Take $x_1, x_2 \in E_\lambda$
$\Rightarrow T(x_1) = \lambda x_1$ and $T(x_2) = \lambda x_2$
$\Rightarrow T(x_1 + x_2) = T(x_1) + T(x_2) = \lambda x_1 + \lambda x_2 = \lambda(x_1 + x_2)$ (Using linearity of T and definition)
$\Rightarrow x_1 + x_2 \in E_\lambda$
Take $x \in E_\lambda$ and $\alpha \in F \Rightarrow T(x) = \lambda x$
$T(\alpha x) = \alpha T(x) = \alpha(\lambda x) = \lambda(\alpha x)$
$\alpha x \in E_\lambda$. Hence E_λ is a subspace.

Definition 4:- If λ is an eigenvalue of T, then E_λ is called the corresponding eigenspace of V.

Result 3:- Let λ and μ be two eigenvalues of T. $x(\neq \theta_V) \in E_\lambda \cap E_\mu$ iff $\lambda = \mu$
Proof:- Let $x(\neq \theta_V) \in E_\lambda \cap E_\mu$. $x \in E_\lambda \Rightarrow T(x) = \lambda x = \mu x$ (as $x \in E_\mu$ also)
$\Rightarrow (\lambda - \mu)x = \theta_V$
$\Rightarrow \lambda - \mu = \theta_V$ (as $x \neq \theta_V$)
$\Rightarrow \lambda = \mu$.
Let $\lambda = \mu$ as λ is an eigenvalue of T there exists $x(\neq \theta_V) \in E_\lambda (= E_\mu)$.

Result 4:- Let x and y be two eigenvectors of T corresponding to two distinct eigenvalues λ and μ respectively. x and y form a linearly independent set in V.
Proof:- Let x and y are linearly dependent. i.e. there exists atleast one non zero coefficient in linear combination of these two vectors. Without loss of generality let $\alpha \neq 0$ in $\alpha x + \beta y = \theta_V \Rightarrow x = \frac{-\beta}{\alpha} y$
$\Rightarrow x \in E_\mu$ (as E_μ is a subspace of V)
$\Rightarrow x(\neq \theta_V)$ (as x is an eigenvector) $\in E_\lambda \cap E_\mu$
$\Rightarrow \lambda = \mu$ (contradiction)

Definition 5:- Let V be a vector space over F. $W \leq V$ ($\leq$ stands for subspace). Let a map $T : V \to V$ be linear. W is said to be $T-$invariant if $T(W) \leq W$. i.e. $\forall w \in W$, $T(w) \in W$.
Examples:- (i) $\{\theta\}$ and V are $T-$invariant.
(ii) ker T and Im T are $T-$invariant.
Proof:- Take $x \in \ker T \Rightarrow T(x) = \theta_V \in \ker T$.

Take $x \in \text{Im } T \Rightarrow T(x) \in \text{Im } T$ (as $T(x) \leq V$)
(iii)E_λ is $T-$invariant.
Proof:- $x \in E_\lambda \Rightarrow T(x) = \lambda x \in E_\lambda$ (as $E_\lambda \leq V$)

Definition 6:- Let V be a vector space over F. Let $W \leq V$ and $U \leq V$. We say that V is direct sum of W and U and write $W \oplus U$ if $\forall v \in V$ there exists unique $w \in W$ and $u \in U$ such that $v = u + w$.

Definition 7:-We say that $T : V \to V$ has an eigen basis if there exists a basis such that each vector in the basis is an eigenvector.
Consider a result involving definition 6 and definition 1. A map $T : V \to V$ is diagonalizable iff there exists an eigen basis. Let us try to prove it. Let $T : V \to V$ be diagonalizable over F and let dim $V = n$. There exists a basis B_V such that $[M(T)]_{B_V}$ is a diagonal matrix. We claim that this particular basis is an eigen basis. Let $B_V = \{v_1, ..., v_n\}$ be that basis.
j column of the diagonal matrix looks like $\begin{pmatrix} 0 \\ \vdots \\ \alpha_j \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. Hence $T(v_j) = 0.v_1 + ...\alpha_j v_j + 0.v_{j+1} + ... + 0.v_n$ or $T(v_j) = \alpha_j v_j$. This proves our claim for forward implication. Similarly we can prove reverse implication.

Result 5:- $V = U \oplus W$ iff $V = U + W$ and θ_V is a unique sum of vectors in U and W. $u + w = \theta_V \Rightarrow u = \theta_V = w$
Proof:- Forward implication is established as we take v in particular θ_V. For reverse implication let there exists $u_1, u_2 \in U$ and $w_1, w_2 \in W$ such that $v = u_1 + w_1 = u_2 + w_2 \Rightarrow (u_1 - u_2) + (w_1 - w_2) = \theta_V$ (as $U \leq V$ and $W \leq V \Rightarrow u_1 - u_2 \in U$ and $w_1 - w_2 \in W$)
$\Rightarrow u_1 = u_2$ and $w_1 = w_2$. Hence $V = U \oplus W$

Result 6:- Let V be vector space over F of dimension 2. $T : V \to V$ be linear. Let λ and μ be two distinct eigenvalues of T, then T has an eigen basis. More over $V = E_\lambda \oplus E_\mu$.
Proof:- Take $x \in E_\lambda$ and $y \in E_\mu$. $B_V = \{x, y\}$ form a linearly independent set in V by result 4. As dimension of V is 2, B_V is in fact a basis for V. Hence T has an eigen basis. For $v \in V$, $v = \alpha x + \beta y$ where $\alpha x \in E_\lambda$ and $\beta y \in E_\mu$. Hence $V = E_\lambda + E\mu$. Take $u \in E_\lambda$ and $w \in E_\mu$ such that $u + w = \theta_V$.
$\Rightarrow u = -w \Rightarrow u \in E_\mu$ (as E_μ is an subspace)
$\Rightarrow u \in E_\lambda \cap E_\mu$
$\Rightarrow u = \theta_V$ (as λ and μ are distinct from Result-3)
$\Rightarrow w = -u = \theta_V$
Hence $V = E_\lambda \oplus E\mu$

Result 7:- Let $V = U \oplus W$
(i) iff $V = U + W$ and $U \cap W = \{\theta_V\}$
(ii) dim V =dim U+dim W

Proof:- (i)Let $V = U \oplus W$, by definition $V = U + W$. $x \in U \cap W$, $\theta_V = x - x$
$\Rightarrow x = \theta_V$ (using Result 5)
For reverse implication let $V = U + W$ and $U \cap W = \{\theta_V\}$. Let $x + y = \theta_V$ where ($x \in U$ and $y \in W$) $\Rightarrow x = -y \in W$
$\Rightarrow x \in U \cap W$
$\Rightarrow x = \theta_V$ and $y = -x = \theta_V$. Form Result 5 it follows that $V = U \oplus W$.
(ii) Let $B_U = \{u_1, ..., u_m\}$ and $B_W = \{w_1, ..., w_p\}$ be basis for U and W respectively. Our claim is that $B_U \cup B_W$ is a basis for V. Take $v \in V$. $v = u + w$ (for some $w \in W$ and $u \in U$)
$\Rightarrow v = \sum_{i=1}^{m} \alpha_i u_i + \sum_{j=1}^{p} \beta_j w_j$ for some scalars $\alpha_i, \beta_j \in F$ (as B_U and B_W are basis for U and W respectively)
$\Rightarrow v \in$ span $(B_U \cup B_W)$
Consider $\sum_{i=1}^{m} \alpha_i u_i + \sum_{j=1}^{p} \beta_j w_j = \theta_V$
$\sum_{i=1}^{m} \alpha_i u_i = -(\sum_{j=1}^{p} \beta_j w_j) \in U \cap W$
$\Rightarrow \sum_{j=1}^{p} \beta_j w_j = \theta_V \Rightarrow \beta_j = 0\ \forall j$ (as B_W is a basis for W)
$\Rightarrow \sum_{i=1}^{m} \alpha_i u_i = \theta_V \Rightarrow \alpha_i = 0\ \forall j$ (as B_V is a basis for V)
$B_U \cup B_W$ is a linearly independent set in V. Hence $B_U \cup B_W$ is a basis for V.
dim U+dim W=dim V (as B_U and B_W are disjoint sets. This follows from the fact that basis $B_U \cup B_W$ contains distinct vectors in V)

Result 8:- Let $T : V \to V$ and $V = U \oplus W$ such that both U and W are $T-$invariant. Let $B_U = \{u_1, ..., u_m\}$ and $B_W = \{w_1, ..., w_p\}$ be basis for U and W respectively. $B_V = B_U \cup B_W$ is a basis for V. We wish to see how the matrix $[M(T)]_{B_V}$ looks like. $T(u_i) \in U$ for $1 \leq i \leq m$, $T(u_i) = a_{1i}u_1 + a_{2i}u_2 + ... + a_{mi}u_m + 0.w_1 + ... + 0.w_p$ i.e. $[M(T)]_{B_V}$ will have all the terms zero after m th row upto $m + p$ th row for $1 \leq t \leq m$ (where t stands for t th column of the matrix. Similarly $T(w_i) \in W$ for $1 \leq i \leq p$, $T(w_i) = 0.u_1 + 0.u_2 + ... + 0.u_m + b_{1i}w_1 + b_{2i}w_2 + ... + b_{pi}w_p$ i.e. $[M(T)]_{B_V}$ will have all the terms zero starting from 1 st row upto m th row for $m \leq t \leq m + p$.

So matrix $[M(T)]_{B_V} = \begin{pmatrix} a_{11} & \cdots & a_{1m} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{p1} & \cdots & b_{pp} \end{pmatrix}$

or $[M(T)]_{B_V} = \begin{pmatrix} A & O \\ O & B \end{pmatrix}$ where A(size $m \times m$) , B(size $p \times p$) and O(one of size $p \times p$ and another is of size $m \times m$) are block matrices.
Similarly if $V = W_1 \oplus ... \oplus W_k$ where W_is are $T-$invariant subspaces of V with basis B_{w_i} respectively. Then we can say there exists a basis $B_V = \cup_{i=1}^{k} B_{W_i}$

such that $[M(T)]_{B_V} = \begin{pmatrix} [M(T)]_{B_{W_1}} & & & & \\ & [M(T)]_{B_{W_2}} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & [M(T)]_{B_{W_k}} \end{pmatrix}$

where $[M(T)]_{B_{W_i}}$ is the matrix corresponding to the same map restricted to W_i which is $T : W_i \to W_i$. Other blocks are O of corresponding size.

If $\lambda_1, ..., \lambda_k$ are k distinct eigenvalues with $E_{\lambda_1} + ... + E_{\lambda_k} = V$ then $V = E_{\lambda_1} \oplus ... \oplus E_{\lambda_k}$. Note that this follows from Result 3.

Result 9:- There exist a polynomial $f(X) \in R[X] \setminus \{0\}$ with degree less than equal to n^2 such that each matrix of size $n \times n$ satisfies.
Proof:- Let $V = M_n(R)$, where $M_n(R)$ is set of all $n \times n$ matrices. It is clear that dim $V = n^2$. Consider $A \in V$ and a set $\{I, A, A^2, ..., A^{n^2}\} \subset V$. It contains more than n^2 vectors in V. Hence it is a linearly dependent set in V i.e. there exists a_is not all zero such that $a_0 + a_1 A + ... + a_{n^2} A^{n^2} = O$ $(O \in V)$
$\Rightarrow \exists f(X) \in R[X] \setminus \{0\}$ with degree $f(X) \leq n^2$ such that $f(A) = O$

Definition 8:- Let A be a square matrix. The polynomial det $(A - XI)$ in X is called the characteristics polynomial of A. It is denoted by $p_A(X)$. Let $p_A(X) = (X - \lambda_1)^{n_1}...(X - \lambda_k)^{n_k}$. Then λ_i $(1 \leq i \leq k)$ are the eigenvalues of A.

Definition 9:- Lt A be a square matrix. We say that $m_A(X)$ is the minimal polynomial of A if $m_A(A) = O$ and it is of least degree.

Consider an example. Let $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Characteristics polynomial for A is $(X - I)^3$. But A also satisfies $(X - I)^2$. Hence $(X - 2)^2$ is the minimal polynomial of A. i.e. $p_A(X) = (X - I)^3$ and $m_A(X) = (X - 1)^2$.

Result 10:-Let $T : V \to V$ be linear. Suppose $m_T(X) = p(X)q(X)$ where $p(X)$ and $q(X)$ are relatively prime. $V =$ ker $p(T) \oplus$ ker $q(T)$.
Proof:- From Euclids algorithm we know that $\exists f(X), g(X) \in R[X]$ such that $p(X)f(X) + q(X)g(X) = 1$. In particular for T we have $p(T)f(T) + q(T)g(T) = I$.
$p(T)f(T) + q(T)g(T)(v) = I(v)$
$p(T)f(T)(v) + q(T)g(T)(v) = v$
Consider $q(T)(p(T)f(T)(v)) = f(T)(p(T)q(T)(v))$ (as $p(T), f(T), q(T)$ commute)
$= f(T)(\theta_V)$ (as $m_T(T) = P(T)q(T) = O$ where $O : V \to V$ is the zero map)
$= \theta_V$. Hence $(p(T)f(T))(v) \in \ker q(T)$. Similarly we can show that $(q(T)g(T))(v) \in$ ker $p(T)$. This shows that $V =$ ker $p(T)+$ ker $q(T)$
Consider $v \neq \theta_V \in$ ker $p(T) \cap$ ker $q(T)$. This implies $p(T)(v) = \theta_V$ and $q(T)(v) = \theta_V$
Consider $p(T)f(T) + q(T)g(T)(v)$
$= f(T)(p(T)(v)) + g(T)(q(T)(v))$
$= f(T)(\theta_V) + g(T)(\theta_V)$
$= \theta_V = I(v) = v$ (contradiction to $v \neq \theta_V$)
ker $p(T) \cap$ ker $q(T) = \{\theta_V\}$. Hence $V =$ ker $p(T) \oplus$ ker $q(T)$.

Result 11:-Let $T : V \to V$ be linear. Suppose $m_T(X) = (X - \lambda_1)^{m_1}...(X - \lambda_k)^{m_k}$ be minimal polynomial of T.
(i) If $p(T) = \theta_V$, $m_T(X)$ divides $p(X)$. In particular $m_T(X)$ divides $p_T(X)$ that follows from Cayley Hamilton Theorem.

(ii) If μ is an eigenvalue of T , Then $\mu = \lambda_i$ for some i
(iii) $V = \ker\, (T - \lambda_1)^{m_1} \oplus ... \oplus \ker\, (T - \lambda_k)^{m_k}$

Definition 10:- Let $V(\lambda) = \{v \in V : \exists k \in N \text{ such that } (T - \lambda I)^k v = \theta_V\}$ is called generalized eigenvalue corresponding to the eigenvalue λ.

Definition 10:- Let $T : V \to V$ be linear. $W \leq V$ is called cyclic if $\exists\ v \in V$, $\lambda \in C$, $k \in N$ such that $(T - \lambda I)^k = \theta_V$ and $\{v, (T - \lambda I)v, ..., (T - \lambda I)^{k-1}v\}$ is a basis for W.

Result 12:-(i) Let w be an eigenvector corresponding to eigenvalue μ , then $W = \text{span } (w)$ is cyclic.
(ii) Let W be $T-$ invariant and cyclic. Take $w_1 \neq \theta_V \in W \Rightarrow T(w_1) \in W$
$T(w_1) = T(w_1) - \lambda w_1 + \lambda w_1 = (T - \lambda I)w_1 + w_1$. Denote $(T - \lambda I)^i w_1 = w_{i+1}$. $T((T - \lambda I)w_1) = T(w_2) = (T - \lambda I)w_2 + w_2$. like wise $T(w_i) = \lambda w_i + w_{i+1}$ for $1 \leq i \leq k-1$ and $T(w_k) = \lambda w_k$. Consider $T|_W : W \to W$ with basis in reverse order i.e. $B_W = \{w_k, ..., w_1\}$. $[M(T)]_{B_W}$ looks like

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

of size $k \times k$. This matrix is called Jordan block of λ of size k. It is denoted by $J_k(\lambda)$.

Definition 12:- Let $T : V \to V$ be linear. T is said to be nilpotent of index k if there exists $k \in N$ such that $T^k = O$ but $T^{k-1} \neq O$.

Result 13:- Let T be nilpotent of index k.
Suppose μ is an eigenvalue of T. There exists $v \neq \theta_V$ such that $T(v) = \mu v$
$\theta_v = T^k(v) = \mu^k v$
$\Rightarrow \mu^k = 0 = \mu$
By definition there exists $w \in V$ such that $(v \neq \theta_V) = T^{k-1}w$
$\Rightarrow \exists v \neq \theta_V$ such that $T(v) = \theta_V = 0.v$
$\Rightarrow 0$ is an eigenvalue.

Result 14:- let $A : V \to V$ be a nilpotent map on a finite dimensional vector space over F. Assume that all the roots of the characteristics polynomial lies in F. Then there exists a Jordan basis of V such that $[M(T)]$ w.r.t. that basis looks like a Jordan block or a matrix whose diagonal consists of Jordan blocks.
Proof:- We prove this result by induction on dimension of V. If dim $V = 1$ and as T is nilpotent $T = O$. Any non zero vector is a Jordan basis of V. If $A = O$ and dim $V > 1$ any basis of V is a Jordan basis. In this induction we assume that the result is true for all non zero nilpotent maps on any finite dimensional vector space with dimension less than n where $n > 1$. Let dim $V = n$ and A be non zero and nilpotent. As $\ker A \neq \{\theta_V\}$ (as A is nilpotent) dim Im $A < n$. It is also invariant under A. Restriction of A to $W = \text{Im } A$ is again nilpotent on W. (as $W \leq V$). On applying induction hypothesis we get a Jordan basis for

W, say ,$B_J = B_{J_1} \cup ... \cup B_{J_K}$ where each B_{J_i} is a Jordan string which acts as a basis for Jordan blocks inside the Jordan form obtained from basis B_J under A. Let $B_{J_1} = \{v_{i1}, ..., v_{in_i}\}$ with $Av_{i1} = \theta_V$ and $Av_{ij} = v_{ij-1}$ for $2 \leq n_i$, of course $n_1 + ... + n_k = \dim \text{ Im } A$.
B_J is an basis for Im A, the set $\{v_{i1} : 1 \leq i \leq k\}$ (first elements of the Jordan string J_i) is a linearly independent subset of V and also subset of ker A. We extend this set to a basis of ker A, say, $\{v_{11}, ..., v_{k1}, z_1, ..., z_r\}$. Each last element $v_{in_i} \in J_i$ lies in Im A hence we can find $v_{in_i+1} \in V$ such that $Av_{in_i+1} = v_{in_i}$. Let $B_i := B_{J_i} \cup \{v_{in_i+1}\}$ and $B := \cup_{i=1}^{k} B_i \cup \{z_1, ..., z_r\}$. We have removed k no of vectors from basis of Im A (v_{i1}s) and included k no of vectors in basis of Im $A(v_{in_i+1}$s). Using rank nullity theorem we get $|B| = n$. We claim that B is linearly independent.
Let $(a_{11}v_{11} + ... + a_{1n_1+1}v_{1n_1+1}) + ... + (a_{k1}v_{k1} + ... + a_{kn_k+1}v_{1n_k+1}) + b_1z_1 + ... + b_rv_r = \theta_V$ ....(i)
On applying A on both sides we have $A([a_{12}v_{12} + ... + a_{1n_1+1}v_{1n_1+1}] + ... + [a_{k2}v_{k2} + ... + a_{kn_k+1}v_{1n_k+1}]) = \theta_V$ (as $z_j, v_{i1} \in$ ker A for $1 \leq j \leq r$ and $1 \leq i \leq k$)
Since $Av_{ij} = v_{ij-1}$ for $1 \leq i \leq k$ and $2 \leq j \leq n_i + 1$, we get
$(a_{12}v_{11} + ... + a_{1n_1+1}v_{1n_1+1}) + ... + (a_{11}v_{11} + ... + a_{kn_k+1}v_{1n_k+1}) = \theta_V$ As v_{ij}'s appearing in this equation are linearly independent we deduce that $a_{ij} = 0$ for $1 \leq i \leq k$ and $2 \leq j \leq n_i + 1$
Thus (i) becomes $a_{11}v_{11} + ... + a_{k1}v_{k1} + b_1z_1 + ... + b_rz_r = \theta$. Vectors appearing in this equation are basis for ker A , so all the coefficients in the equation are zero. As a whole coefficient in equation (i) are zero. Hence B is linearly independent. Now we need to show that B is a Jordan basis. $B_i = \{v_{i1}, ..., v_{in_i+1}\}$ with $Av_{i1} = \theta_V$ and $Av_{ij} = v_{ij-1}$ for $2 \leq j \leq n_i + 1$ is a basis for i th Jordan block. $J_{k+m} = \{z_m\}$ for $1 \leq m \leq r$ is a basis as well as Jordan block. Hence we have total $k + r$ no of Jordan blocks in Jordan canonical form of A.

Result 15:- Let $T : (V = C^n) \to V$ be linear over field C and M be the matrix associated with the standard basis. $m_T(X)$ splits over C and let $(X - \lambda)$ be a linear factor of it. Clearly $T - \lambda I$ is nilpotent. This implies $J = C^{-1}(M - \lambda I)C$ is of Jordan form associated with the Jordan basis. i.e. $C^{-1}MC = J + \lambda I$ is Jordan form for M.

Result 16:-These are some important features of Jordan form of a linear map which are helpful in determining the Jordan canonical form.

(i) The sum of sizes of the blocks involving a fixed eigenvalue is equal to the algebraic multiplicity of that eigenvalue. Where algebraic multiplicity stands for the multiplicity of the eigenvalue as a root of the characteristics polynomial.
Proof:- Let J be the canonical form of M. M and J are similar. i.e. their exists an invertible matrix P such that $M = P^{-1}JP$. Consider $M - \lambda I = P^{-1}JP - P^{-1}\lambda IP = P^{-1}(J - \lambda I)P$. det $(M - XI) =$ det $(J - XI)$ as $M - \lambda I$ is similar to $J - \lambda I$. i.e. characteristics polynomials of A and J are same. The eigenvalues of a Jordan block $J_j(\lambda)$ is λ with algebraic multiplicity k.

(ii) The no of blocks involving an eigenvalue is equals to its geometric multiplicity, i.e. the dimension of the corresponding eigenspace.
Proof:- Eigenvalue λ of similar matrices have the same geometric multiplicity.

In order to show this consider $A, B \in L(V = C^n, C^n)$ and $[M(A)], [M(B)] \in M_n(C)$ (matrix w.r.t. standard basis) be similar. Observe that $E_\lambda(A) = \ker(A - \lambda I)$ and $E_\lambda(B) = \ker(B - \lambda I)$. As A is similar to B there exists an invertible map P such that $A = PBP^{-1}$. Take $x \in E_\lambda(A) \Rightarrow (A - \lambda I)x = \theta_V$
$\Rightarrow (PBP^{-1} - \lambda I)(x) = P((B - \lambda I)(P^{-1}x)) = \theta_V$
$\Rightarrow (B - \lambda I)(P^{-1}x) = \theta_V$ (As P is invertible, P is into i.e. $\ker P = \{\theta_V\}$)
$\Rightarrow P^{-1}(x) \in E_\lambda(B)$
Similarly taking $x \in E_\lambda(B)$ we can show that $P(x) \in E_\lambda(A)$ and $E_\lambda(A)$ and $E_\lambda(B)$ are in bijection. From this it follows that dim $E_\lambda(A) = E_\lambda(B)$.
Any Jordan block $J_k(\lambda)$ corresponding to λ has one dimensional eigen space. Hence no of blocks involving an eigenvalue is equals to dimension of E_λ as each block corresponds to one dimensional eigenspace and total no of blocks corresponds to dimension of $E_\lambda(J)$ which is equals to geometric multiplicity of λ corresponding to the original matrix.
(iii) The largest block involving an eigenvalue is equals the multiplicity of the eigenvalue as a root of minimal polynomial.
Proof :- First of all observe that map $T : (V = C^k) \to C^k$ defined as $T(c) = (J_k(\lambda) - \lambda I_{k\times k})(c)$ is nilpotent of index k. Let $B = \{e_1, ..., e_k\}$ be the standard basis for C^k.

$$[M(T)]_B = J_k(\lambda) - \lambda I_{k\times k} = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$$

Clearly $T(e_1) = 0.e_1 + ... + 0.e_n = \theta_V$
$T(e_i) = 0.e_1 + ... + 1.e_{i-1} + 0.e_j + ... + 0.e_k = e_{j-1}$ for $2 \le i \le k$
$T^{k-1}(e_k) = T^{k-2}(T(e_k)) = T^{k-2}(e_{k-1}) = T^{k-3}(e_{k-2}) = ... = T(e_2) = e_1 \ne \theta_V$
hence $T^{k-1} \ne O$
$T^k(e_i) = T^{k-i+1}(e_1) = \theta_V$ for $k \le k - i + 1 \le 1$ or $1 \le i \le k$. $T^k = O$
If $J = \text{diag}\,(J_{n_1}(\lambda_1), ..., J_{n_k}(\lambda_k))$, then its minimal polynomial is the product of minimal polynomials of $J_n(\lambda_i)$.

Note 1:- $J_1(\lambda)$ looks like $\begin{pmatrix} \lambda \end{pmatrix}$. Let $T : V \to V$ be linear. Let M be the matrix associated with standard basis. Consider a Jordan canonical form J corresponding to matrix M consisting of Jordan block J_1 only. Clearly J is a diagonal matrix. i.e. Jordan basis itself is an eigen basis. As largest block size is 1 , multiplicity of eigenvalues is 1 i.e. minimal polynomial is a product of distinct linear factors. For a fixed eigenvalue sum of sizes of the blocks is equals to the no of blocks (as each and every block has size 1).In terms of geometric and algebraic multiplicity, they have to be the same for all possible eigenvalues. Now let M be any matrix associate with standard basis. If M has distinct eigenvalues, then characteristics polynomial will have distinct linear factors so also minimal polynomial have distinct linear factors. i.e. M is diagonalizable. It could be possible that M may not have distinct eigenvalue but M is diagonalizable. For that we need to have algebraic multiplicity same as geometric multiplicity or minimal polynomial is product of distinct linear factors. We can also say it like this. If minimal polynomial has atleast one repeated linear factor then M is not diagonalizable.

Problem 1:-

The characteristics polynomial of A is $(X-1)^3(X-2)^2$ and its minimal polynomial is $(X-1)^2(X-2)$. What is its Jordan form?
Solution :-Sum of sizes of blocks for eigenvalue 1 is 3 and maximum size of Jordan block for 1 is 2. For eigenvalue 1 there is one $J_2(1)$ and one $J_1(1)$. For eigenvalue 2 sum of sizes of the block is 2 with at most single size of the block. So for eigenvalue 2 there are two $J_1(2)$ block. Let J be Jordan form of A.

$$J = \begin{pmatrix} J_2(1) & & & \\ & J_1(1) & & \\ & & J_1(2) & \\ & & & J_1(2) \end{pmatrix} = \begin{pmatrix} 1 & 1 & & & \\ 0 & 1 & & & \\ & & 1 & & \\ & & & 2 & \\ & & & & 2 \end{pmatrix}$$

Problem 2:- Let $A \in M_n(C)$. Prove that A is similar to its transpose A^T.
Solution :- Consider $S : C^n \to C^n$ defined as $T(c) = Ac$. As $A \in M_n(C)$ characteristics polynomial splits over C. From result-15 it follows that there exists a Jordan basis B_J for C^n such that matrix associated with this is of Jordan canonical form. Say it J. As A is the matrix associated with the standard basis, A and J are similar. There exists $P \in M_n(C)$ which is invertible such that $A = P^{-1}JP$
$A^T = (P^{-1}JP)^T = P^T J^T (P^{-1})^T = (P^T)J^T(P^T)^{-1}$. Hence A^T is similar to J^T. Now if we show that J and J^T are similar we are done. Let there are k no of Jordan blocks inside J. We denote them as $J_1, ..., J_k$. Let B_{J_i} be Jordan basis corresponding to J_i. Clearly $B_J = B_{J_1} \cup .... \cup B_{J_k}$. Let $B_{J_i} = \{v_{i1}, ..., v_{in_i}\}$. Note that J is obtained when vectors in B_{J_i} are in the specific order as taken. Take $B'_{J_i} = \{v_{in_i}, ..., v_{i1}\}$ so that $B'_J = B'_{J_1} \cup .... \cup B'_{J_k}$. $J^T = [M(S)]_{B'_J}$. Hence there exists a basis for C^n such that $J^T = [M(S)]_{B'_J}$. J is similar to J^T.

Problem 3:- Show that there is no $A \in M_3(R)$ whose minimal polynomial is X^2+1 but there is $B \in M_2(R)$ as well as $C \in M_3(C)$ whose minimal polynomial is X^2+1.
Solution:- $A \in M_3(R)$ its characteristics polynomial is a third degree polynomial. It has atleast one real root. X^2+1 does not have any real root. Hence X^2+1 can not be the minimal polynomial of A as it contradicts that minimal polynomial vanishes at roots of characteristics polynomial.

Take $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Clearly $A \in M_2(R)$ satisfies X^2+1 and there is no polynomial of degree less than 2 that A satisfies.

Take $A = \begin{pmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}$ Clearly $A \in M_3(C)$ and X^2+1 is its minimal polynomial as $i, -i$ are the only eigenvalues of A.

Problem 4:- Let $A^{k+1} = A$ for some $k \in N$. Show that A is diagonalizable.
Solution :- $A^{k+1} = A$
$\Rightarrow A(A^K - I) = O$
i.e. either $A = O$ or A satisfies $X^k - 1 = 0$ or k th root of unity. As k th root of unity are distinct eigenvalues are also distinct. Hence A is diagonalizable.

Problem 5:- Show that $\begin{pmatrix} cos\phi & -sin\phi \\ sin\phi & cos\phi \end{pmatrix}$ and $\begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix}$ are similar.

Solution:- On computing we get $e^{i\phi}$ and $e^{-i\phi}$ as two eigenvalues. When $\phi = 0$, $e^{i\phi} = e^{-i\phi} = 0$. The both matrix turns out to be the same. So they are similar when $\phi = 0$. When $\phi \neq 0$ eigenvalues are different. Minimal polynomial will be same as characteristics polynomial which is equals to $(X - e^{i\phi})(X - e^{-i\phi})$. Jordan block will be $\begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix}$. As J is similar to A, $\begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix}$ is similar to $A = \begin{pmatrix} cos\phi & -sin\phi \\ sin\phi & cos\phi \end{pmatrix}$

Result 17:-$card(A)^{card(B)+card(C)} = card(A)^{card(B)}card(A)^{card(C)}$
Proof:- We wish to define a bijection from $A^{B \sqcup C}$ to $A^B \times A^C$. Let ϕ be such a map that sends $f \in A^{B \sqcup C}$ to $\phi(f) \in A^B \times A^C$. f is function from $B \sqcup C \to A$ i.e. $f(1,b)$ and $f(2,c) \in A$ for some $(1,b),(2,c) \in B \sqcup C$. $\phi(f) = (\phi_f^B, \phi_f^C)$. $\phi_f^B \in A^B$ i.e. $\phi_f^B(b) \in A$. On setting $\phi_f^B(b) = f(1,b)$ will serve our purpose. Similarly for $c \in C$, $\phi_f^C(c) = f(1,c)$.
For bijection we define $\psi : A^B \times A^C \to A^{B \sqcup C}$ in the manner explained earlier as $\psi_f(1,b) = f^B(b)$ and $\psi_f(2,c) = f^C(c)$. Now we need to show that $\phi\psi = I$. Consider $\phi\psi(f)$ where $f = (f^B, f^C) \in A^B \times A^C$. Let $g = \psi(f)$, $\phi(g) \in A^B \times A^C$ so $\phi(g) = (\phi_g^B, \phi_g^C)$. Now $\phi_g^B(b) = g(1,b) \Rightarrow \phi_{\psi(f)}^B(b) = \psi(f(1,b)) = f^B(b)$
$\Rightarrow \phi\psi((f^B)(b)) = f^B(b)$
Similarly we can show that $\phi\psi((f^C)(c)) = f^C(c)$. Likewise we can show that $\psi\phi = I$

Result 18 :- $(card(A)^{card(B)})^{card(C)} = card(A)^{card(B)card(C)}$
Proof :- We wish to define a bijection from $(A^B)^C$ to $A^{B\times C}$. Let ϕ be a map that sends an element $f \in (A^B)^C$ to $\phi(f) \in A^{B\times C}$. So we have a given function $f \in (A^B)^C$ and we want to send it to $\phi(f) \in A^{B\times C}$. As $f \in (A^B)^C$ by definition of $(A^B)^C$, f is a function from $C \to A^B$. For an element $c \in C$, $f(c) \in A^B$. By definition , $f(c)$ is a function from $B \to A$ i.e. for an element $b \in B$, $f(c)(b) \in A$. Let us consider $\phi(f)$. $\phi(f)$ is a function from $B \times C \to A$. For $(b,c) \in B \times C$ we wish to send $\phi(f)(b,c)$ to an element in A. For a given f , $f(c)(b) \in A$. Hence by taking $\phi(f)(b,c) = f(c)(b)$ will serve our purpose for a given $f \in (A^B)^C$.
Now we need to show that $\phi : (A^B)^C \to A^{B\times C}$ is a bijection. In order to show that we wish to find a function $\psi : A^{B\times C} \to (A^B)^C$ such that $\psi\phi = \phi\psi = I$ where I is the identity map to the corresponding domain. Fix a $f \in A^{B\times C}$, $f : B \times C \to A$. For $(b,c) \in B \times C$, $f(b,c) \in A$. $\psi(f) \in (A^B)^C$ i.e. $\psi(f)(c) \in A^B$ or $\psi(f)(c)(b) \in A$. So taking $\psi(f)(c)(b) = f(b,c)$ will serve our purpose.
We need to check whether $\psi\phi = \phi\psi = I$ or not. First take $\psi(\phi(f))$ for a fixed $f \in (A^B)^C$. $\psi(\phi(f)(b,c)) = \psi(f(c)(b) = f(b,c))$ for any fixed $(b,c) \in B \times C$. Similarly we can prove $\phi\psi = I$