Analogous Results of Operators on Complex and Real Vector Spaces

Win Sandar[1], Win Yu Yu Maw[22]

Abstract

In this paper, some analogous results on complex and real vector spaces are described. The structures of operator on complex and real vector spaces are analyzed. Cayley-Hamilton Theorems for operators on complex and real vector spaces are described. The major structure theorems about operators on complex and real vector spaces are expressed.

Keywords: Basis, Upper-triangular Matrix, Real Vector Space, Complex Vector Space, Characteristic Polynomial, Eigenvalues.

Introduction

In this paper, we show that the results on real vector spaces are more complex than analogous results on complex vector spaces. Therefore, most of the results on complex vector spaces are proved first. The analogous results on real vector spaces are then proved. We define the characteristic polynomial of an operator on complex and real vector spaces. Suppose that V is a complex vector space and $L \in \mathcal{L}(V)$, the set of operators on V. We know that V has a basis with respect to which L has an upper-triangular matrix. Thus if L has dim V distinct eigenvalues, then each eigenvalues must appear exactly once on the diagonal of any upper-triangular matrix of L.

We prove that the characteristic polynomial of operator on compelx and real vector spaces must equal to zero. We describe that the proof uses the same idea as the proof of the analogous result in analyzing the structure of an operator on complex and real vector spaces. In analyzing the structure of an operator, the number of times an eigenvalue is repeated on the diagonal of an upper-triangulars matrix of L is independent of which particular basis we choose for a complex vector space. We find that the number of times a particular characteristic polynomial appears is independent of the choice of basis for a real vector space. These results will be our key tools in analyzing the structure of an operator on complex and real vector spaces.

We also show that the major structure theorem about operators on complex vector space and the corresponding result on real vector space.

Operators on Complex Vector Spaces

Definition 1

Suppose V is a complex vector space and $L \in \mathcal{L}(V)$. Any basis of V with respect to which L has an upper-triangular matrix of the form

$$\mathcal{M}(L) = \begin{bmatrix} \lambda_1 & & * \\ & . & \\ 0 & & \lambda_n \end{bmatrix}.$$

(1)

[1] Lecturer, Dr, Department of Mathematics, University of Mandalay
[2] Lecturer, Department of Mathematics, University of Mandalay

Then the **characteristic polynomial** of L is given by $(z-\lambda_1)...(z-\lambda_n)$, where $\lambda_1,...,\lambda_n$ denote the distinct eigenvalues of L.

Proposition 1

If a linear map $L \in \mathcal{L}(V)$, the set of operators on V and n is a nonnegative integer such that null $L^n = \text{null}\, L^{n+1}$, then $\text{null}\, L^0 \subset \text{null}\, L^1 \subset ... \subset \text{null}\, L^n = \text{null}\, L^{n+1} = \text{null}\, L^{n+2} ...$.

Proof: See Dr Win Sandar, (2019). □

Proposition 2

If $L \in \mathcal{L}(V)$, then $\text{null } L^{\dim V} = \text{null } L^{\dim V+1} = \text{null } L^{\dim V+2} = ...$.

Proof: See [Win Sandar, 2019]. □

Corollary 1

Suppose $L \in \mathcal{L}(V)$ and λ is an eigenvalue of L. Then the set of generalized eigenvectors of L corresponding to λ equals $\text{null}\,(L-\lambda I)^{\dim V}$.

Proof: See [Win Sandar, 2019]. □

Theorem 1

Let $L \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Then for every basis of V with respect to which L has an upper-triangular matrix, λ appears on the diagonal of the matrix of L precisely $\dim \text{null}\,(L-\lambda I)^{\dim V}$ times.

Proof: See [Win Sandar, 2019].
□

Proposition 3

If V is a complex vector space and $L \in \mathcal{L}(V)$, then the sum of the multiplicities of all the eigenvalues of L equals dim V.

Proof: See [Win Sandar, 2019]. □

Theorem 2 (Cayley-Hamilton theorem on a complex vector space)

Suppose that V is a complex vector space and $L \in \mathcal{L}(V)$. Let q be the characteristic polynomial of L. Then $q(L) = 0$.

Proof:

Suppose that $(v_1,...,v_n)$ is a basis of V with respect to which the matrix of L has an upper-triangular form (1).

We need only show that $q(L)v_i = 0$ for $i = 1, ..., n$.

To do this, it sufficies to show that

$$(L-\lambda_1 I)\ \ldots\ (L-\lambda_i I)v_i = 0 \text{ for } i = 1, \ldots, n.$$
(6)

By induction on i, suppose that $i = 1$.

We have $L_1 v_1 = \lambda_1 v_1$, giving by (6).

Now suppose that $1 < i \le n$ and that

$$\begin{aligned} 0 &= (L-\lambda_1 I)v_1 \\ &= (L-\lambda_1 I)(L-\lambda_2 I)v_2 \\ &\ \vdots \\ &= (L-\lambda_1 I)\ \ldots\ (L-\lambda_{i-1} I)v_{i-1}. \end{aligned}$$

Because $\mathcal{M}(L,\ (v_1, \ldots, v_n))$ is given by (1), we see that $(L-\lambda_i I)v_i \in \operatorname{span}(v_1, \ldots, v_{i-1})$.

Thus, by induction hypothesis, $(L-\lambda_1 I)\ \ldots\ (L-\lambda_{i-1} I)$ applied to $(L-\lambda_i I)v_i$ gives 0. In other words, (6) holds, completing the proof. □

Proposition 4

If $L \in \mathcal{L}(V)$ and the polynomial $p \in \mathcal{P}(\mathbf{F})$, the set of all polynomials with coefficients in **F**, then null p(L) is invariant under L.

Proof. See [Win Sandar, 2019]. □

Theorem 3 (Major structure theorem on a complex vector space)

Suppose that V is a complex vector space and $L \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_n$ be the distinct eigenvalues of L, and let $U_1, \ldots, U_n$ be the corresponding subspaces of generalized eigenvectors. Then

(a) $V = U_1 \oplus \ldots \oplus U_n$;

(b) each U_i is invariant under L;

(c) each $\left(L-\lambda_i I\right)\big|U_i$ is nilpotent.

Proof: See [Win Sandar, 2019].
□

Operators on Real Vector Spaces

For operators on real vector spaces, we need to define the characteristic polynomial of 1-by-1 and 2-by-2 matrices with real entries.

Definition 2

The characteristic polynomial of a 2-by-2 matrix $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$ is $(x-a)(x-d)-bc$ for a real vector space.

Proposition 5

Suppose $L \in \mathcal{L}(V)$ and B is a matrix of L with respect to some basis of V. Then the eigenvalues of L are the same eigenvalues of B.

Proof: See [Axler, S., 1997].

□

Definition 3

A **block upper-triangular matrix** is a square matrix of the form

$$\begin{bmatrix} B_1 & & * \\ & \cdot & \\ 0 & & B_n \end{bmatrix},$$

(7)

where $B_1,...,B_n$ are square matrices lying along the diagonal, all entries below $B_1,...,B_n$ equal 0 and the * denotes arbitrary entries.

Theorem 4

Suppose V is a real vector space and $L \in \mathcal{L}(V)$. Then there is a basis of V with respect to which L has a block upper triangular matrix

$$\begin{bmatrix} B_1 & & * \\ & \cdot & \\ 0 & & B_n \end{bmatrix},$$

where each B_i is a 1-by-1 matrix or 2-by-2 matrix with no eigenvalues.

Proof:

If dim V =1, the result holds. Consider dim V = 2.

If L has an eigenvalue λ, then let $v_1 \in V$ be nonzero eigenvector.

Extend (v_1) to a basis (v_1, v_2) of V. Then L has an upper-triangular matrix with respect to this basis of the form

$$\begin{bmatrix} \lambda & a \\ 0 & b \end{bmatrix}.$$

If L has no eigenvalues, then choose any basis (v_1, v_2) of V. Then the matrix of L with respect to this basis has no eigenvalues. Thus we have the desired conclusion when dim V = 2.

Suppose now that dim V > 2 and the desired result holds for all real vector spaces with smaller dimension. If L has an eigenvalue, let U be a one-dimensional subspace of V that is invariant under L.

Choose any basis of U and let B_1 denote the matrix of L|U with respect to this basis. If B_1 is a 2-by-2 matrix, then L has no eigenvalues and thus L|U has no eigenvalues.

Hence if B_1 is a 2-by-2 matrix, then B_1 has no eigenvalues.

Let W be any subspace of V such that

$$V = U \oplus W.$$

Since W has dimension less than the dimension of V, We will proof the induction hypothesis to L|W. However, W might not be invariant under L, i.e., L | W might not be an operator on W.

Define $S \in \mathcal{L}(W)$ by $Sw = P_{W,U}(Lw)$ for $w \in W$.

Note that

$$\begin{aligned} Lw &= P_{U,W}(Lw) + P_{W,U}(Lw) \\ &= P_{U,W}(Lw) + Sw \end{aligned}$$

for every $w \in W$.

By induction hypothesis, there is a basis of W with respect to which S has a block upper-triangular matrix of the form (7), where each is a 1-by-1 matrix of a 2-by-2 matrix with no eigenvalues.

Adjoin this basis of W to the basis of U chosen above, getting a basis of V. By using this result matrix of L with respect to this basis is a block upper-triangular matrix of the form (7), completing the proof. □

Proposition 6

Suppose that V is a real vector space with dimension 2 and $L \in \mathcal{L}(V)$ has no eigenvalues. Let $p \in \mathcal{P}(\mathbf{R})$, be a monic polynomial with degree 2. Suppose A is the matrix of L with respect to some basis of V.

(a) If p equals the characteristic polynomial of A, then $p(L) = 0$.

(b) If p does not equal the characteristic polynomial of A, then $p(L)$ is invertible.

Proof:

(a) Suppose that Vis a real vector space with dimension 2 and $L \in \mathcal{L}(V)$.

Suppose that $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is the matrix of T with respect to some basis (v_1, v_2) of V.

If $b = 0$, then the matrix above is upper triangular. We know that L has characteristic polynomial $(x-a)(x-d)$.

When applied to L, the polynomial $(x-a)(x-d)$ gives 0 even when $b \neq 0$.

We have

$$\begin{aligned} (L-aI)(L-dI)v_1 &= (L-dI)(L-aI)v_1 \\ &= (L-dI)\,bv_2 = bcv_1 \end{aligned}$$

and $(L-aI)(L-dI)v_2 = (L-aI)cv_1 = bcv_2$. Thus $(L-aI)(L-dI)$ not equal to 0 unless $bc = 0$.

However, the above equations show that $(L-aI)(L-dI) - bcI = 0$.

Thus if $p(x) = (x-a)(x-d) - bc$, then $p(L) = 0$.

(b) Let q denote the characteristic polynomial of A and suppose $p \neq q$.

We can write $p(x) = x^2 + \alpha_1 x + \beta_1$ and $q(x) = x^2 + \alpha_2 x + \beta_2$ for some $\alpha_1, \beta_1, \alpha_2, \beta_2 \in \mathbf{R}$.

$$p(L) = p(L) - q(L) = (\alpha_1 - \alpha_2)L + (\beta_1 - \beta_2)I.$$

If $\alpha_1 = \alpha_2$, then $\beta_1 \neq \beta_2$.

Thus if $\alpha_1 = \alpha_2$, then $p(L)$ is a nonzero multiple of the identity and hence is invertible, as desired.

If $\alpha_1 \neq \alpha_2$, then $p(L) = (\alpha_1 - \alpha_2)\left(L - \frac{\beta_1 - \beta_2}{\alpha_1 - \alpha_2} I\right)$, which is an invertible operator because L has no eigenvalues. Thus, complete the proof. □

The following proof uses the same ideas as the proof of the analogous result on complex vector spaces, Theorem 1.

Theorem 5

Suppose that V is a real vector space and $L \in \mathcal{L}(V)$. Suppose that with respect to some basis of V, the matrix of L is

$$\begin{bmatrix} B_1 & & * \\ & \cdot & \\ 0 & & B_n \end{bmatrix},$$

where each B_i is a 1-by-1 matrix or a 2-by-2 matrix with no eigenvalues.

(a) If $\lambda \in \mathbf{R}$, then precisely dim null $(L - \lambda I)^{\dim V}$ of the matrices $B_1, ..., B_n$ equal the 1-by-1 matrix $[\lambda]$.

(b) If $\alpha, \beta \in \mathbf{R}$, satisfy $\alpha^2 < 4\beta$, then precisely

$$\frac{\dim \text{ null}\,(L^2 + \alpha L + \beta I)^{\dim V}}{2}$$

of the matrices $B_1, ..., B_n$ have characteristic polynomial equal to $x^2 + \alpha x + \beta$.

Proof:

We construct one proof that can be used to prove both (a) and (b)

Let $\lambda, \alpha, \beta \in \mathbf{R}$ with $\alpha^2 < 4\beta$.

Define $p \in \mathcal{P}(\mathbf{R})$ by

$$p(x) = \begin{cases} x - \lambda, & \text{if we are trying to prove (a)} \\ x^2 + \alpha x + \beta, & \text{if we are trying to prove (b).} \end{cases}$$

Let d be the degree of p.

Thus $d = 1$ if we are trying to prove (a) and $d = 2$ if we are trying to prove (b).

We will prove this theorem by induction on n.

If $n = 1$, then $\dim V = 1$ or $\dim V = 2$ of which implies that the desired result holds.

Assume that $n > 1$ and that the desired result holds when n is replaced with $n - 1$.

Let $\dim V = m$.

Consider a basis of V with respect to which L has the block upper-triangular matrix (7).

Let U_i be the span of the basis vectors corresponding to A_i.

Thus $\dim U_i = 1$, if A_i is a 1-by-1 matrix and $\dim U_i = 2$, if A_i is a 2-by-2 matrix.

Let $U = U_1 + ... + U_{n-1}$.

Clearly U is invariant under L and the matrix of $L|U$ with respect to the obvious basis is

$$\begin{bmatrix} A_1 & & * \\ & . & \\ 0 & & A_{n-1} \end{bmatrix}.$$

Thus, by induction hypothesis,

precisely $\left(\frac{1}{d}\right)$dim null $p(L|U)^m$ of the matrices $A_1,...,A_{n-1}$ have characteristic polynomial p.
(8)

The induction hypothesis gives (8) with exponent dim U instead of n, but we can replace dim U with n to get the statement above.

Suppose $u_n \in U_n$.

Let $S \in \mathcal{L}(U_n)$ be the operator whose matrix with respect to the basis corresponding to U_n equals A_n.

In particular, $Su_n = P_{U_n,U} Lu_n$.

Now $Lu_n = P_{U,U_n} Lu_n + P_{U_n,U} Lu_n$

$= *_U + Su_n$,

where $*_U$ denotes a vector in U and $Su_n \in U_n$.

Thus applying to both sides of the equation above gives

$$L^2 u_n = *_U + S^2 u_n.$$

The last two equations show that

$$p(L)u_n = *_U + p(S)u_n$$

for some $*_U \in U$.

Thus iterating the last equation gives

$$p(L)^m u_n = *_U + p(S)^m u_n \tag{9}$$

for some $*_U \in U$ and $p(S)u_n \in U_n$.

The proof of theorem breaks into two cases.

First consider the case where the characteristic polynomial of B_n does not equal p.

We will show that in this case

$$\text{null } p(L)^m \subset U. \tag{10}$$

We know that null $p(L)^m = \text{null}\, p(L \mid U)^m$, and hence (8) will tell that precisely $(\frac{1}{d})$ dim null $p(L)^m$ of the matrices $B_1, \ldots, B_n$ have the characteristic polynomial p, completing the proof in the case where the characteristic polynomial of B_n does not equal p.

Suppose that $v \in \text{null } p(L)^m$. We can write $v = u + u_n$ where $u \in U$ and $u_n \in U_n$.

Using (9), we have $0 = p(L)^m v = p(L)^m u + p(L)^m u_n = p(L)^m u + *_U + p(S)^m u_n$, for some $*_U \in U$.

Since the vectors $p(L)^m u$ and $*_U$ are in U and $p(S)^m u_n \in U_n$, this implies that $p(S)^m u_n = 0$.

However, $p(S)$ is invertible, so $u_n = 0$.

Thus $v = u \in U$ completes the proof of (10).

Now we consider the case where the characteristic polynomial to B_n equals p.

We will show that $\dim \text{ null } p(L)^m = \dim \text{ null } p(L|_U)^m + d$,

(11)

which along with (8) complete the proof.

Using the formula for the dimension of the sum of two subspaces, we have

$$\dim \text{ null } p(L)^m = \dim (U \cap \text{ null } p(L)^m) + \dim (U + \text{ null } p(L)^m) - \dim U$$

$$= \dim \text{ null } p(L|_U)^m + \dim (U + \text{null } p(L)^m) - (n-d).$$

If $U + \text{null } p(L)^m = V$, then $\dim (U + \text{null } p(L)^m) = m$, which when combined with the formula above for dim null $p(L)^m$ would give (11), as desired.

To prove that $U + \text{ null } p(L)^m = V$, suppose $u_n \in U_n$.

Because the characteristic polynomial of the matrix of S equals p, we have $p(S) = 0$. Thus $p(L)u_n \in U$.

Now $$p(L)^m u_n = p(L)^{m-1}(p(L)u_n) \in \text{range } p(L|_U)^{m-1}$$

$$= \text{range } p(L|_U)^m.$$

Thus we can choose $u \in U$ such that $p(L)^m u_n = p(L|_U)^m u$.

Now

$$p(L)^m (u_n - u) = p(L)^m u_n - p(L)^m u$$

$$= p(L)^m u_n - p(L|_U)^m u.$$

$$= 0.$$

Thus $u_n - u \in \text{null } p(L)^m$, and hence $u_n = u + (u_n - u)$ is in $U + \text{null } p(L)^m$.

In other words, $U_n \subset U + \text{null } p(L)^m$. Therefore $V = U + U_n \subset U + \text{null } p(L)^m$, and hence $U + \text{null } p(L)^m = V$, completing the proof. □

Definition 4

Suppose that V is a real vector space and $L \in \mathcal{L}(V)$. Suppose that with respect to some basis of V, L has a block upper-triangular matrix of the form

$$\begin{bmatrix} B_1 & & * \\ & \cdot & \\ 0 & & B_n \end{bmatrix},$$

where each B_i is a 1-by-1 matrix or 2-by-2 matrix with no eigenvalues. We define the **characteristic polynomial** of L to be the product of the characteristic polynomial of $B_1, ..., B_n$.

For each i, define $q_i \in \mathcal{P}(\mathbf{R})$ by

$$q_i(x) = \begin{cases} x - \lambda, & \text{if } B_i \text{ equals } [\lambda], \\ (x-a)(x-d) - bc, & \text{if } B_i \text{ equals } \begin{bmatrix} a & c \\ b & d \end{bmatrix}. \end{cases} \qquad (12)$$

Then the characteristic polynomial of L is

$$q_1(x) \ldots q_m(x).$$

Clearly the characteristic polynomial of L has degree dim V. □

Proposition 7

If V is a real vector space and $L \in \mathcal{L}(V)$, then the sum of the multiplicities of all the eigenvalues of L plus the sum of twice the multiplicities of all the eigenpairs of L equals dim V.

Proof:

Suppose that V is a real vector space and $L \in \mathcal{L}(V)$.

Then there is a basis of V with respect to which the matrix of L.

The multiplicity of an eigenvalue λ equals the number of times the 1-by-1 matrix $[\lambda]$ appears on the diagonal of this matrix.

The multiplicity of an eigenpair (α, β) equals the number of times $x^2 + \alpha x + \beta$ is the characteristic polynomial of a 2-by-2 matrix on the diagonal of this matrix.

Since the diagonal of this matrix has length dim V, the sum of the multiplicities of all the eigenvaues of L plus the sum of twice the multiplicities of all the eigenpairs of L must equal dim V. □

The following proof uses the same idea as the proof of the analogous result on complex vector spaces, Theorem 2.

Theorem 6

Suppose that V is a real vector space and $L \in \mathcal{L}(V)$. Let q denote the characteristic polynomial of L. Then $q(L) = 0$.

Proof:

Choose a basis of V with respect to which L has a block upper-triangular matrix of the form (7), where each B_i is a 1-by-1 matrix of a 2-by-2 matrix with no eigenvalues.

Suppose U_i is the one-or two-dimensional subspace spanned by the basis vectors corresponding to B_i. Define q_i as in (13).

To prove that $q(L) = 0$, we need only show that $q(L)|U_i = 0$ for i = 1, …, n. To do this, we will show that

$$q_1(L) \ldots q_i(L)|U_i = 0. \tag{14}$$

We will prove (14) by induction on i.

Suppose that i = 1.

From Proposition 6, we have $q_1(L)|U_1 = 0$ if and by giving (14) when i = 1.

Now suppose that $1 < i \le m$ and that

$$\begin{aligned} 0 &= q_1(L)|U_1 \\ 0 &= q_1(L)q_2(L)|U_2 \\ &\vdots \\ 0 &= q_1(L)q_{i-1}(L)|U_{i-1}. \end{aligned}$$

If $v \in U_i$, then we see that $q_i(L)v = u + q_i(S)v$,

where $u \in U_1 + \ldots + U_{i-1}$ and $S \in \mathcal{L}(U_i)$ has characteristic polynomial q_i. Because $q_i(S) = 0$, the equation above shows that $q_i(L)v \in U_1 + \ldots + U_{i-1}$, where $v \in U_i$.

By induction hypothesis, $q_1(L) \ldots q_{i-1}(L)$ applied to $q_{i-1}(L)v = 0$ where $v \in U_i$.

Hence, complete the proof. □

The theorem below should be compared to Theorem 3, the corresponding result on complex vector spaces. The proof uses the same idea as the proof of the analogous result on complex vector spaces, Theorem 3.

Theorem 7 (Main structure theorem on a real vector space)

Suppose that V is a real vector space and $L \in \mathcal{L}(V)$. Let $\lambda_1, ..., \lambda_n$ be the distinct eigenvalues of L, with $U_1, ..., U_n$ the corresponding sets of generalized eigenvectors. Let $(\alpha_1, \beta_1), ..., (\alpha_N, \beta_N)$ be the distinct eigenpairs of L and let $V_i = \text{null}\,(L^2 + \alpha_i L + \beta_i I)^{\dim V}$.

Then (a) $V = U_1 \oplus ... \oplus U_n \oplus V_1 \oplus ... \oplus V_N$;

(b) each U_i and each V_i is invariant under L;

(c) each $(L - \lambda_i I) \,|\, U_i$ and each $(L^2 + \alpha_i L + \beta_i I) \,|\, V_i$ is nilpotent.

Proof:

From Proposition 7, we know that dim U_i equals the multiplicity of λ_i as an eigenvalue of L and dim V_i equals twice the multiplicity of (α_i, β_i) as an eigenpair of L. Thus

$$\dim V = \dim U_1 + ... + \dim U_n + \dim V_1 + ... + \dim V_N.$$

Let $U = U_1 + ... + U_n + V_1 + ... + V_N$.

Since U is invariant under L, we can define $S \in \mathcal{L}(U)$ by $S = L \,|\, U$.

S has the same eigenvalues, with the same multiplicities, as L because all the generalized eigenvectors of L are in U, the domains of S. Similarly, S has the same eigenpairs, with the same multiplicities, as L.

Thus applying Proposition 7, we get

$$\dim U = \dim U_1 + ... + \dim U_n + \dim V_1 + ... + \dim V_N.$$

This equation shows that $\dim V = \dim U$.

Because U is a subspace of V, this implies that $V = U$. By Proposition 6, we conclude that (a) holds.

From Proposition 4, we get the proof of (b). Clearly (c) follows from (b) and the definition of nilpotent, completing the proof. □

Conclusion

The main achievement of this paper is that important results of operators on real vector spaces are more complex than the analogous results on complex vector spaces.

Acknowledgements

The authors would like to express sincere gratitude to Dr, Thein Win, Rector of University of Mandalay for his kind consent to the submission of this research paper. The authors acknowledge to Professor Dr Khin Myo Aye, head of Mathematics Department and Professor Dr Khin Phyu Phyu Htoo for their encouragement on doing research. The authors are also thankful to their colleagues for their support while doing this research.

References

Axler, S. (1997) “Linear Algebra Done Right”, Second Edition, Springer-Verlag, New York.

Carrell, J. B. (2005), "Fundamentals of Linear Algebra", The University of British Columbia.

Dawkins, P. (2007), "Linear Algebra", Indiana University.

Win Sandar (2019), "*Some Structures of Operators on Complex Vector Spaces*", University of Mandalay Research Journal, Vol. 10, No.5, pp.38-47.