Course: Advanced Topics in Abstract Algebra-I

Unit-IV: (Linear Algebra, Canonical Forms)

Programme: M. Sc. Mathematics I[st] Semester

Written By:

Dr. Mukhtar Ahmad Khanday
Department of Mathematics
University of Kashmir, Srinagar-190006

4.1: Introduction

In linear algebra, Jordan normal form (often called Jordan canonical form) of a linear operator on a finite dimensional vector space is an upper triangular matrix of a particular form called Jordan matrix, representing the operator on some basis. The form is characterized by the condition that any non-diagonal entries that are non-zero must be equal to 1, be immediately above the main diagonal (on the super-diagonal), and have identical diagonal entries to the left and below them. If the vector space is over a field K, then a basis on which the matrix has the required form exists if and only if all eigenvalues of M lie in K, or equivalently if the characteristic polynomial of the operator splits into linear factors over K. This condition is always satisfied if K is the field of complex numbers. The diagonal entries of the normal form are the eigenvalues of the operator, with the number of times each one occurs being given by its algebraic multiplicity.

If the operator is originally given by a square matrix M, then its Jordan normal form is also called the Jordan normal form of M. Any square matrix has a Jordan normal form if the field of coefficients is extended to one containing all the eigenvalues of the matrix. In spite of its name, the normal form for a given M is not entirely unique, as it is a block diagonal matrix formed of Jordan blocks, the order of which is not fixed; it is conventional to group blocks for the same eigenvalue together, but no ordering is imposed among the eigenvalues, nor among the blocks for a given eigenvalue, although the latter could for instance be ordered by weakly decreasing size. The Jordan–Chevalley decomposition is particularly simple on a basis on which the operator takes its Jordan normal form. The diagonal form for diagonalizable matrices, for instance normal matrices, is a special case of the

Jordan normal form. The Jordan normal form is named after Camille Jordan. This chapter contains the results related to triangular form, Jordan canonical form, bilinear form, quadratic forms etc for the graduate students in Mathematics.

4.2: Similarity of Matrices

Definition: Let A and B are two square matrices. Then the matrix B is said similar to A if there exists an invertible matrix P such that

$$P^{-1}AP = B$$

Theorem : Similarity of matrices is an equivalence relation

Similarity of Linear transformation

Definition : Let V be an n-dimensional vector space over a field F. Let A(V) be the set of all linear transformations from V to V. Then two linear transformations $S, T \in A(V)$ are said to be similar if there exists an invertible linear transformation $C \in A(V)$ such that

$$C^{-1}SC = T$$

The relation on A(V) defined by similarity is an equivalence relation, thus, A(V) decomposes into equivalence classes, each is called similarity class. The existence of linear transformation in each similarity class whose matrix representation in some bases of V is of special form, such matrices are known as Canonical forms.

Now in order to check the two linear transformations are similar, we have to compute a particular canonical form for each and check if these are the same.

There are many canonical forms, but we shall discuss the following forms:

(i) Triangular form

(ii) Jordan form

(iii) Rational Canonical form

Theorem : $T:U\to V$ be a linear transformation and rank T is equal to r, then there exist bases of U and that of V such that the matrix representation of T has the form

$$A=\begin{pmatrix} I & O \\ O & O \end{pmatrix}$$

where I is the r-square identity matrix.

[The matrix A is known as normal or canonical form]

Proof: Let the dim U = m and dim V = n. Let W be the kernel of T and Im (T) the image of T.

Since the rank of T is r, therefore the dimension of the kernel space of T is m-r. Let $\{\alpha_1,\alpha_2,\cdots,\alpha_{m-r}\}$ be a basis of W. So it can be extended to form a basis of U. Let this extension be

$$\{v_1,v_2,\cdots,v_n,\alpha_1,\alpha_2,\cdots,\alpha_{m-r}\}$$

Now setting $u_1=T(v_1),u_2=T(v_2),\cdots,u_r=T(v_r)$

Observe that

$$T(v_1)=u_1=1.u_1+0.u_2+\cdots+0.u_r+0.u_{r+1}+\cdots+0.u_n$$

$$T(v_2)=u_2=0.u_1+1.u_2+\cdots+0.u_r+0.u_{r+1}+\cdots+0.u_n$$

$$T(v_3)=u_3=0.u_1+0.u_2+1.u_3+\cdots+0.u_r+0.u_{r+1}+\cdots+0.u_n$$

$$\cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots$$

$$T(v_r)=u_r=0.u_1+0.u_2+\cdots+1.u_r+0.u_{r+1}+\cdots+0.u_n$$

$$T(\alpha_1)=0=0.u_1+0.u_2+\cdots+0.u_r+0.u_{r+1}+\cdots+0.u_n$$

$$T(\alpha_2)=0=0.u_1+0.u_2+\cdots+0.u_r+0.u_{r+1}+\cdots+0.u_n$$

$$\cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots$$

$$T(\alpha_{m-r})=0=0.u_1+0.u_2+\cdots+0.u_r+0.u_{r+1}+\cdots+0.u_n$$

Thus the matrix representation of T is

$$A=\begin{pmatrix}1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0\\ 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0\\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0\end{pmatrix}_{n\times n}$$

or $$A=\begin{pmatrix}I & O\\ O & O\end{pmatrix}$$

4.3. Canonical Forms

Let T be a linear operator on a finite dimensional vector space, we know that T may not have a diagonal matrix representation ¥. The canonical form aims to simplify the matrix representation of T by means of primary decomposition theorem, Triangular, Jordan and rational canonical forms.

We note that that triangular and Jordan canonical forms exist for T if and only if the characteristic polynomial $\Delta(\lambda)$ of T has all its roots in the base field K. this is always true if K is the complex field C but may not be true if K is the real field R.

¥: The matrix $A=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$ is not diagonalizable, since the characteristic polynomial of A is $\Delta(\lambda)=(\lambda-1)^2$; hence 1 is the only eigen-value of A. We find a basis of the eigen-space of the eigen-value 1. Substitute $\lambda=1$ into the matrix $\lambda I-A$ to obtain

$$\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} \text{ or } \begin{cases}-y=0\\ 0=0\end{cases} \text{or } y=0$$

The system has only one independent solution, e.g., $x=1, y=0$. Hence $u=(1,0)$ forms a basis of the eigen-space of A. Since A has at most one independent eigen-value, A can not be diagonalizable.

4.4: Invariance

Let $T:V \to V$ be linear. A subspace W of a V is said to be invariant if T maps W into itself, i.e., if $v \in W$ implies $T(v) \in W$. In this case T restricted to W defines a linear operator on W; that is, T induces a linear operator $\hat{T}:W \to W$ defined by $\hat{T}(w)=T(w)$ for every $w \in W$.

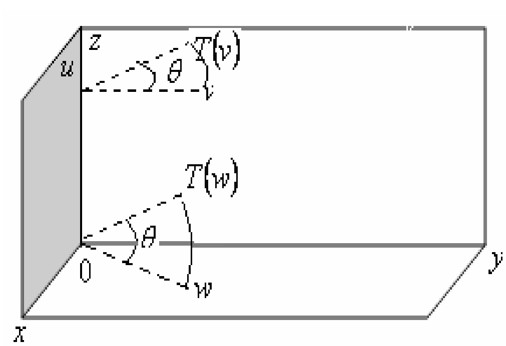
Invariant Subspaces

Definition: Let $T:V \to V$ be a linear transformation. Then a subspace W of V is invariant under T if $T(W) \subset V$ i.e., if $\alpha \in W$, then $T(\alpha) \in W$.

***Example** 1:* Let $T:R^3 \to R^3$ be the linear operator which rotates each vector about the z-axis by an angle θ:

$$T(x,y,z)=(x\cos\theta - y\sin\theta,\ x\sin\theta + y\cos\theta,\ z)$$

Observe that each vector $w=(a,b,0)$ in the xy plane W remains in W under the mapping T, i.e., W is T-invariant. Observe also that the Z-axis U is invariant under T. Furthermore, the restriction of T to W rotates each vector about the origin O, and the restriction of T to U is the identity mapping of U.



***Example** 2:* Non-zero eigen vectors of a linear operator $T:V\to V$ may be characterized as generators of T-invariant 1-dimensional subspaces. For suppose $T(v)=\lambda v$, $v\neq 0$, then $W=\{ku\ ;\ k\in K\}$, the 1-dimensional subspace generated by v, is invariant under T because

$$T(kv)=kT(v)=k(\lambda v)=k\lambda v\in W$$

Conversely, suppose that dim V = 1 and $u\neq 0$ generates U, and U is invariant under T. Then $T(u)\in U$ and so $T(u)$ is a multiple of u, i.e., $T(u)=\mu u$, Hence u is an eigen vector of T.

***Exercise** 1:* Suppose $T:V\to V$ is a linear operator, show that each of the following is invariant under T.

(i) {0} (ii) V (iii) Kernel of T (iv) Image of T

***Sol.**:* We have $T:V\to V$ a linear map.

(i) Clearly $T(0)=0\in\{0\}$

Hence $\{0\}$ is invariant under T.

(ii) For every $v\in V$, $T(v)\in V$,

Hence V is invariant under T.

(iii) Let $u\in\ker T$, then $T(u)=0\in\ker T$. Since the $\ker T$ is a subspace of V. Thus $\ker T$ is invariant under T.

(iv) Since $T(v)\in \mathrm{Im}(T)$, for every $v\in V$, it is certainly true if $v\in\mathrm{Im}(T)$. Hence the $\mathrm{Im}(T)$ is invariant under T.

***Exercise** 2:* Suppose $\{W_i\}$ is a collection of T-invariant subspaces of a vector space V. Show that the intersection $W=\bigcap W_i$ is also T-invariant.

***Sol.**:* Suppose $v\in W$; then $v\in W_i$ for every i. Since W_i is T-invariant, $T(v)\in W_i$ for every i. Thus $T(v)\in W=\bigcap W_i$, so W is T-invariant.

***Theorem** 2:* Let $T:V \to V$ be any linear operator and let $f(t)$ be any polynomial. Then the kernel of $f(T)$ is invariant under T.

***Proof**:* We have given that $f(t)$ is a polynomial and $T:V \to V$ is a linear map.

Now $\ker f(T) = \{u \in V : (f(T))(u) = 0\}$

Now suppose $v \in \ker f(T)$

i.e., $f(T)(v) = 0$

we need to show that $T(v)$ also belongs to the kernel of $f(T)$

i.e., $f(T)(T(v)) = 0$

since $f(t)t = tf(t)$

we have $f(T)T = Tf(T)$

Thus, $f(T)T(v) = Tf(T)(v) = T(0) = 0$

***Exercise** 3:* Find all invariant subspaces of $A = \begin{pmatrix} 2 & -5 \\ 1 & -2 \end{pmatrix}$ viewed as an operator on R^2.

***Sol.**:* First of all, we have that R^2 and $\{0\}$ are invariant under A. Now if A has any other invariant subspaces, then it must be 1-dimensional. However the characteristic polynomial of A is

$$\Delta(\lambda) = |A - \lambda I| = \begin{vmatrix} \lambda - 2 & 5 \\ -1 & \lambda + 2 \end{vmatrix} = \lambda^2 + 1$$

Hence A has no eigen values (in R^2) and so A has no eigen vectors. But the 1-dimensional invariant subspaces correspond to the eigen vectors; Thus R^2 and $\{0\}$ are the only subspaces invariant under A.

Theorem: If W is a subspace invariant under $T \in A(V)$, then T induces a linear transformation $\bar{T}$ on V/W defined by $\bar{T}(\alpha + W) = T(\alpha) + W$. Moreover if T satisfies the polynomial $q(x) \in F[x]$, then so is $\bar{T}$. Thus the minimal polynomial of $\bar{T}$ divides the minimal polynomial of T.

Proof: First we show that $\bar{T}$ is well defined. Let $\alpha + W$ and $\beta + W$ be any element of V/W.

If $\alpha + W = \beta + W$, then $\alpha - \beta \in W$. Since W is T-invariant, then

$$T(\alpha - \beta) = T(\alpha) - T(\beta) \in W$$

So, accordingly, $T(\alpha) + W = T(\beta) + W$

$\Rightarrow$ $\bar{T}(\alpha + W) = \bar{T}(\beta + W)$

Thus $\bar{T}$ is well defined.

We now show that $\bar{T}$ is linear. For which,

$$\bar{T}\{(\alpha + W) + (\beta + W)\} = \bar{T}(\alpha + \beta + W)$$
$$= T(\alpha + \beta) + W$$
$$= T(\alpha) + W + T(\beta) + W$$
$$= \bar{T}(\alpha + W) + \bar{T}(\beta + W)$$

Also $\bar{T}\{c(\alpha + W)\} = \bar{T}(c\alpha + W)$

$$= T(c\alpha) + W$$
$$= cT(\alpha) + W$$
$$= c(T(\alpha) + W)$$
$$= c\bar{T}(\alpha + W)$$

Thus, T is linear.

If $\alpha + W \in V/W$,

Then $\bar{T}^2(\alpha + W) = T^2(\alpha) + W = T(T(\alpha)) + W$

$$= \bar{T}(T(\alpha) + W) = \bar{T}(\bar{T}(\alpha + W))$$
$$= \bar{T}^2(\alpha + W)$$

Therefore, $(\bar{T}^2) = (\bar{T})^2$

Thus, we can easily show that

$$\left(\overline{T^n}\right)=\left(\overline{T}\right)^n \text{ for any } n \geq 0.$$

Now for any polynomial $q(x) \in F[x]$, given by

$$q(x)=a_n x^n + a_{n-1}x^{n-1}+\cdots+a_0$$

$$q(\overline{T})(\alpha+W)=q(T)(\alpha)+W = a_nT^n(\alpha)+a_{n-1}T^{n-1}(\alpha)+\cdots+a_0I(\alpha)+W$$

$$=\sum a_iT^i(\alpha)+W=\sum a_i\left(T^i(\alpha)+W\right)$$

$$=\sum a_i\overline{T}^i(\alpha+W)=\sum a_i\left(\overline{T}\right)^i(\alpha+W)$$

$$=\overline{q(T)}(\alpha+W)$$

Therefore, $q(\overline{T})=\overline{q(T)}$. Accordingly if T is a root of $q(x)=0$, then $q\left(\overline{\overline{T}}\right)=\overline{0}=W=q\left(\overline{T}\right)$. Thus $\overline{T}$ is also a root of $q(x)=0$.

Let $p_1(x)$ be the minimal polynomial over F satisfied by $\overline{T}$.

If $q(\overline{T})=0$ for $q(x)\in F[x]$, then $p_1(x) \mid q(x)$. If $p(x)$ be the minimal polynomial of T over F, then $p(T)=0$ implies $p(\overline{T})=0$. Hence $p_1(x) \mid p(x)$.

4.5: Triangular Forms

Definition: If $T:V\to V$ is a linear transformation of V over F, then the matrix of T in the basis $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ of V is triangular if

$$T(\alpha_1)=a_{11}\alpha_1$$

$$T(\alpha_2)=a_{21}\alpha_1+a_{22}\alpha_2$$

$$T(\alpha_3)=a_{31}\alpha_1+a_{32}\alpha_2+a_{33}\alpha_3$$

....

$$T(\alpha_n)=a_{n1}\alpha_1+a_{n2}\alpha_2+\cdots+a_{nn}\alpha_n$$

In other words, Let T be a linear operator on an n-dimensional vector space V. suppose T can be represented by the triangular matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \cdots & \cdots \\ & & & a_{nn} \end{pmatrix}$$

Then the characteristic polynomial of T

$$\Delta(\lambda) = |A - \lambda I| = (\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn})$$

is a product of linear factors and conversely.

Theorem: If $T \in A(V)$ has all its characteristic roots in F, then there is a basis of V in which the matrix of T is triangular.

Proof: We prove the theorem by induction on the dimension of V. If dim V =1, then every matrix representation of T is a matrix of order 1 x 1, which is trivially triangular.

Now suppose that the theorem is true for all vector spaces over F of dimension n -1. Let dim V = n >1. Since T has all its characteristic roots of F. Let $\lambda_1 \in F$ be a characteristic root of T. Then there exists a non-zero eigen vector α_1 corresponding to λ_1 such that $T(\alpha_1) = a_{11}\alpha_1$. Let W be the one dimensional subspace of V spanned by α_1, and is T-invariant. Set $\overline{V} = V/W$, then

$$\dim \overline{V} = \dim V - \dim W = n - 1$$

Thus by the above theorem T induces a linear transformation $\overline{T}$ on $\overline{V}$ whose minimal polynomial divides the minimal polynomial of T. Therefore, all the roots of the minimal polynomial of $\overline{T}$, being roots of the minimal polynomial of T must lie in F. Thus $\overline{V}$ and $\overline{T}$ satisfy the hypothesis of the theorem.

Since $\dim \overline{V} = n - 1$, therefore by induction hypothesis, there is a basis $\{\overline{\alpha}_1, \overline{\alpha}_2, \cdots, \overline{\alpha}_n\}$ of $\overline{V}$ such that

$$\bar{T}(\bar{\alpha}_2) = a_{22}\bar{\alpha}_2$$

$$\bar{T}(\bar{\alpha}_3) = a_{32}\bar{\alpha}_2 + a_{33}\bar{\alpha}_3$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\bar{T}(\bar{\alpha}_n) = a_{n2}\bar{\alpha}_2 + a_{n3}\bar{\alpha}_3 + \cdots + a_{nn}\bar{\alpha}_n$$

Now let $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ be the elements of V which belong to the cosets $\bar{\alpha}_1, \bar{\alpha}_2, \cdots, \bar{\alpha}_n$ respectively i.e., $\bar{\alpha}_i = \alpha_i + W$. Then $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ is a basis of V.

Since $\quad \bar{T}(\bar{\alpha}_2) = a_{22}\bar{\alpha}_2$

$\Rightarrow \quad \bar{T}(\alpha_2 + W) = a_{22}(\alpha_2 + W)$

$\Rightarrow \quad T(\alpha_2) + W = a_{22}(\alpha_2) + W$

$\Rightarrow \quad T(\alpha_2) - a_{22}(\alpha_2) \in W$

But W is spanned by α_1, so

$$T(\alpha_2) - a_{22}(\alpha_2) = a_{21}\alpha_1$$

$\Rightarrow \quad T(\alpha_2) = a_{21}\alpha_1 + a_{22}(\alpha_2)$

Similarly for $\bar{\alpha}_3, \bar{\alpha}_4, \cdots, \bar{\alpha}_n$, we have

$$T(\alpha_i) = a_{i1}\alpha_1 + a_{i2}\alpha_2 + \cdots + a_{in}\alpha_n$$

Thus, we have

$$T(\alpha_1) = a_{11}\alpha_1$$

$$T(\alpha_2) = a_{21}\alpha_1 + a_{22}\alpha_2$$

$$T(\alpha_3) = a_{31}\alpha_1 + a_{32}\alpha_2 + a_{33}\alpha_3$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$T(\alpha_n) = a_{n1}\alpha_1 + a_{n2}\alpha_2 + \cdots + a_{nn}\alpha_n$$

Hence the matrix of T in the basis $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ is triangular.

Alternative form of the above theorem: If a square matrix A has all its characteristic roots in F, then A is similar to a triangular matrix i.e., there exists an invertible matrix P such that $P^{-1}AP$ is triangular.

Remark: Let T be the representation of triangular matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{in} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

Then the characteristic polynomial of T is given by

$$\Delta(x) = | xI - A | = (x - a_{11})(x - a_{22})\cdots(x - a_{nn})$$

which is a product of linear factors.

Theorem: If dim V = n and if $T \in A(V)$ has all its roots in F, then T satisfies a polynomial of degree n and $\lambda_1, \lambda_2, \cdots, \lambda_n$ be the characteristic roots of F.

Proof: Since T has all its roots in F so there is a basis $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ of V such that

$$T(\alpha_1) = \lambda_1 \alpha_1$$
$$T(\alpha_2) = a_{21}\alpha_1 + \lambda_2 \alpha_2$$
$$T(\alpha_3) = a_{31}\alpha_1 + a_{32}\alpha_2 + \lambda_3 \alpha_3$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$T(\alpha_n) = a_{n1}\alpha_1 + a_{n2}\alpha_2 + \cdots + \lambda_n \alpha_n$$

Above equations are equivalent to

$$(T - \lambda_1 I)(\alpha_1) = 0$$
$$(T - \lambda_2 I)(\alpha_2) = a_{21}\alpha_1$$
$$(T - \lambda_3 I)(\alpha_3) = a_{31}\alpha_1 + a_{32}\alpha_2$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$(T - \lambda_n I)(\alpha_n) = a_{n1}\alpha_1 + a_{n2}\alpha_2 + \cdots + a_{n(n-1)}\alpha_{n-1}$$

Now $(T - \lambda_2 I)(T - \lambda_1 I)(\alpha_1) = (T - \lambda_1 I)(T - \lambda_2 I)(\alpha_2)$

$$[\because \ (T - \lambda_2 I)(T - \lambda_1 I) = (T - \lambda_1 I)(T - \lambda_2 I)]$$

$$= (T - \lambda_2 I)(a_{21}\alpha_1)$$

$$= a_{21}(T-\lambda_1 I)(\alpha_1) = 0$$

and $(T-\lambda_3 I)(T-\lambda_2 I)(T-\lambda_1 I)(\alpha_3) = (T-\lambda_2 I)(T-\lambda_1 I)(T-\lambda_3 I)(\alpha_3)$

$$= (T-\lambda_2 I)(T-\lambda_1 I)(a_{31}\alpha_1 + a_{32}\alpha_2)$$

$$= (T-\lambda_2 I)(T-\lambda_1 I)(a_{31}\alpha_1) + (T-\lambda_2 I)(T-\lambda_1 I)(a_{32}\alpha_2)$$

$$= a_{31}(T-\lambda_2 I)(T-\lambda_1 I)(\alpha_1) + a_{32}(T-\lambda_2 I)(T-\lambda_1 I)(\alpha_2)$$

$$= 0 + 0 = 0$$

Continuing in this way, we get

$$(T-\lambda_n I)(T-\lambda_{n-1} I)(T-\lambda_{n-2} I)\cdots(T-\lambda_1 I)(\alpha_n) = 0$$

Let $S = (T-\lambda_n I)(T-\lambda_{n-1} I)(T-\lambda_{n-2} I)\cdots(T-\lambda_1 I)$, it satisfies $S(\alpha_1) = 0$, $S(\alpha_2) = 0$, . . ., $S(\alpha_n) = 0$.

Thus annihilates a basis of V, thus S annihilates all of V. Therefore S = 0 which implies that

$$(T-\lambda_n I)(T-\lambda_{n-1} I)(T-\lambda_{n-2} I)\cdots(T-\lambda_1 I) = 0$$

Hence T satisfies a polynomial

$q(x) = (x-\lambda_n)(x-\lambda_{n-1})(x-\lambda_{n-2})\cdots(x-\lambda_1)$ in $F[x]$ of degree n.

***Theorem** 1:* Let $T: V \to V$ be a linear operator whose characteristic polynomial factors into linear polynomials. Then there exists a basis of V in which T is represented by a triangular matrix.

Alternative form: Let A be a square matrix whose characteristic polynomial factors into linear polynomials. Then A is similar to a triangular matrix, i.e., there exists an invertible matrix P such that $P^{-1}AP$ is triangular.

***Example** 1:* Let A be a square matrix over the complex field C. Suppose λ is an eigen value of A^2, Show that $\sqrt{\lambda}$ or $-\sqrt{\lambda}$ is an eigen-value of A.

We know by the above theorem that

$$B = \begin{pmatrix} \mu_1 & * & \cdots & * \\ & \mu_2 & \cdots & * \\ & & \cdots & \cdots \\ & & & \mu_n \end{pmatrix}$$

Hence A^2 is similar to the matrix

$$B^2 = \begin{pmatrix} \mu_1^2 & * & \cdots & * \\ & \mu_2^2 & \cdots & * \\ & & \cdots & \cdots \\ & & & \mu_n^2 \end{pmatrix}$$

Since similar matrices have the same eigen-values, $\lambda = \mu_i^2$ for some i. Hence $\mu_i = \pm\sqrt{\lambda}$, i.e., $\sqrt{\lambda}$ or $-\sqrt{\lambda}$ is an eigen-value of A.

***Theorem** 3:* Suppose W is invariant subspaces of $T : V \to V$. Then T has a block diagonal matrix representation $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$, where A is a matrix representation of the restriction $\hat{T}$ of T to W.

***Proof**:* We choose a basis $\{w_1, w_2, \cdots, w_r\}$ and extend it to a basis $\{w_1, w_2, \cdots, w_r, v_1, v_2, \cdots, v_s\}$ of V .

We have

$$\hat{T}(w_1) = T(w_1) = a_{11}w_1 + a_{12}w_2 + \cdots + a_{1r}w_r$$

$$\hat{T}(w_2) = T(w_2) = a_{21}w_1 + a_{22}w_2 + \cdots + a_{2r}w_r$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$T(v_1) = b_{11}w_1 + b_{12}w_2 + \cdots + b_{1r}w_r + c_{11}v_1 + \cdots + c_{1s}v_s$$

$$T(v_2) = b_{21}w_1 + b_{22}w_2 + \cdots + b_{2r}w_r + c_{21}v_1 + \cdots + c_{2s}v_s$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$T(v_s) = b_{s1}w_1 + b_{s2}w_2 + \cdots + b_{sr}w_r + c_{s1}v_1 + \cdots + c_{ss}v_s$$

But the matrix of T in this basis is the transpose of the matrix of coefficients of the above system of equations. Therefore, it has the form $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$, where A is the transpose of the matrix of coefficients for the obvious subsystem. By the same argument A is the matrix of $\hat{T}$ relative to the basis $\{w_i\}$ of W.

***Exercise** 4:* Let $\hat{T}$ denote the restriction of an operator T to an invariant subspace W, i.e., $\hat{T}(W) = T(W)$ for every $w \in W$. Prove that

(i) For any polynomial $f(t)$, $f(\hat{T})(w) = f(T)(w)$

(ii) The minimal polynomial of $\hat{T}$ divides the minimum polynomial of T.

***Sol.**:* (i) If $f(t) = 0$ or $f(t)$ is constant, i.e., of degree one, then the result clearly holds. Assume $\deg f = n > 1$ and the result holds for polynomials of degree less than n. Suppose that

$$f(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$$

Then

$$\begin{aligned} f(\hat{T})(w) &= \left(a_n \hat{T}^n + a_{n-1}\hat{T}^{n-1} + \cdots + a_1 \hat{T} + a_0 I\right)(w) \\ &= \left(a_n \hat{T}^{n-1}\right)\left(\hat{T}(w)\right) + \left(a_{n-1}\hat{T}^{n-1} + \cdots + a_1 \hat{T} + a_0 I\right)(w) \\ &= \left(a_n T^{n-1}\right)\left(T(w)\right) + \left(a_{n-1} T^{n-1} + \cdots + a_1 T + a_0 I\right)(w) \\ &= f(T)(w) \end{aligned}$$

(ii) Let $m(t)$ denote the minimum polynomial of T. Then by (i)

$$m(\hat{T})(w) = m(T)(w) = 0(w) = 0 \text{ for every } w \in W.$$

i.e., $\hat{T}$ is a zero of the polynomial $m(t)$. Hence the minimum polynomial of $\hat{T}$ divides $m(t)$.

4.6: Invariant Direct Sum Decomposition

A vector space V is termed as the direct sum of its subspaces $W_1, W_2, \cdots, W_r$ written as

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$$

If for every $v \in V$ can be written uniquely in the form

$$v = w_1 + w_2 + \cdots + w_r \text{ with } w_i \in W_i.$$

***Theorem**: If* $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$ where each subspace W_i is of dimensions n_i and is invariant under $T \in A(V)$, then a basis of V can be found so that the matrix of T in this basis is of the form

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & A_r \end{pmatrix}$$

where each A_i is an $n_i \times n_i$ matrix and is the matrix of the linear transformation induced by T on W_r.

Proof: Let $\{\alpha_1^{(1)}, \alpha_2^{(1)}, \cdots, \alpha_{n1}^{(1)}\}$, $\{\alpha_1^{(2)}, \alpha_2^{(2)}, \cdots, \alpha_{n2}^{(2)}\}$, . . . , $\{\alpha_1^{(r)}, \alpha_2^{(r)}, \cdots, \alpha_{nr}^{(r)}\}$ be the basis of $W_1, W_2, \cdots, W_r$ respectively.

Since $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$, therefore

$$\{\alpha_1^{(1)}, \alpha_2^{(1)}, \cdots, \alpha_{n1}^{(1)}, \alpha_1^{(2)}, \alpha_2^{(2)}, \cdots, \alpha_{n2}^{(2)}, \alpha_1^{(r)}, \alpha_2^{(r)}, \cdots, \alpha_{nr}^{(r)}\}$$

form a basis of V. Also each W_i is T-invariant, so that $T(\alpha_j^{(i)}) \in W_i$ and it is linear combination of $\{\alpha_1^{(i)}, \alpha_2^{(i)}, \cdots, \alpha_{ni}^{(i)}\}$, and of only these, that is,

$$T(\alpha_j^{(i)}) = a_1^{(i)}\alpha_1^{(i)} + a_2^{(i)}\alpha_2^{(i)} + \cdots + a_{ni}^{(i)}\alpha_{ni}^{(i)} \qquad \ldots(1)$$

for $j = 1,2,\cdots,n_1$, $j = 1,2,\cdots,n_2$, so on and $i = 1,2,\cdots,r$

Thus the matrix representation of T in a basis of V is obtained by (1) which is

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & A_r \end{pmatrix}$$

where A_i is the matrix of T_i induced on W_i by T.

Home Assignments

1. Suppose W is invariant under $S: V \to V$ and $T: V \to V$. Show that W is also invariant under S + T and ST.
2. Suppose $T: V \to V$ is linear and $T = T_1 \oplus T_2$, with respect to a T-invariant direct sum decomposition $V = V_1 \oplus V_2$. Show that
   (i) $m(x)$ is the least common multiple of $m_1(x)$ and $m_2(x)$, where $m(x)$, $m_1(x)$ and $m_2(x)$ are minimal polynomials of T, T_1 and T_2 respectively.
   (ii) $\Delta(x) = \Delta_1(x)\Delta_2(x)$, where $\Delta(x)$, $\Delta_1(x)$ and $\Delta_2(x)$ are the characteristic polynomials of T, T_1 and T_2 respectively.
3. Let $T: V \to V$ be linear and let W be the eigenspace belonging to an eigenvalue λ of T. Show that W is T-invariant.
4. Prove that similar matrices have the same eigenvalues.

***Theorem** 4:* Suppose $W_1, W_2, \cdots, W_r$ are subspaces of V, and let $\{w_{11}, w_{12}, \cdots, w_{1n1}\}$, . . . , $\{w_{r1}, w_{r2}, \cdots, w_{rn1}\}$ are bases of $W_1, W_2, \cdots, W_r$ respectively, then V is the direct sum of the W_i ; $i = 1,2,\cdots,r$ if and only if the union $B = \{w_{11}, w_{12}, \cdots, w_{1n1}, \cdots, w_{r1}, w_{r2}, \cdots, w_{rnr}\}$ is a basis of V.

***Proof**:* Suppose $B = \{w_{11}, w_{12}, \cdots, w_{1n1}, \cdots, w_{r1}, w_{r2}, \cdots, w_{rnr}\}$ is a basis of V, then for every $v \in V$.

$$v = a_{11}w_{11} + a_{12}w_{12} + \cdots + a_{1n1}w_{1n1} + \cdots + a_{r1}w_{r1} + a_{r2}w_{r2} + \cdots + a_{rnr}w_{rnr}$$

$$= w_1 + w_2 + \cdots + w_r$$

where $w_i = a_{i1}w_{i1} + a_{i2}w_{i2} + \cdots + a_{ini}w_{ini} \in W_i$.

We next show that such a sum is unique.

Suppose $v = w_1' + w_2' + \cdots + w_r'$ where $w_i' \in W_i'$

Since $\{w_{i1}, w_{i2}, \cdots, w_{ini}\}$ is a basis of W_i,

we have $$w_i' = b_{i1}w_{i1} + b_{i2}w_{i2} + \cdots + b_{ini}w_{ini}$$

and so $$v = b_{11}w_{11} + b_{12}w_{12} + \cdots + b_{1n1}w_{1n1} + \cdots + b_{r1}w_{r1} + \cdots + b_{rnr}w_{rnr}$$

Since B is a basis of V, $a_{ij} = b_{ij}$ for each i and each j. Hence $w_i = w_i'$ and so the sum of v is unique. Accordingly V is the direct sum of W_i.

Conversely, suppose V is the direct sum of W_i. Then for any $v \in V$, $v = w_1 + w_2 + \cdots + w_r$ where $w_i \in W_i$. Since $\{w_{iji}\}$ is a basis of W_i, each w_i is a linear combination of the elements w_{iji} and so v is the linear combination of the elements of B. Thus B spans V. We now show that the elements in B are linearly independent.

Suppose $a_{11}w_{11} + a_{12}w_{12} + \cdots + a_{1n1}w_{1n1} + \cdots + a_{r1}w_{r1} + a_{r2}w_{r2} + \cdots + a_{rnr}w_{rnr} = 0$

Note that $a_{i1}w_{i1} + a_{i2}w_{i2} + \cdots + a_{ini}w_{ini} \in W_i$, we also have $0 = 0 + 0 + \cdots + 0$, where $0 \in W_i$. Since such a sum for 0 is unique.

Therefore, $a_{i1}w_{i1} + a_{i2}w_{i2} + \cdots + a_{ini}w_{ini} = 0$ for $i = 1, 2, \cdots, r$

The independence of the basis $\{w_{iji}\}$ imply that all $a's$ are zero. Thus B is linearly independent and hence is a basis of V.

Remark 1: Suppose that $T: V \to V$ is linear and V is the direct sum of (non-zero) T-invariant subspaces $W_1, W_2, \cdots, W_r$:

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_r \text{ and } T(W_i) \subset W_i;\ i = 1, 2, \ldots, r$$

Let T_i denote the restriction of T to W_i. Then T is said to be decomposed into the operators T_i or T is said to be direct sum of the T_i, written $T = T_1 \oplus T_2 \oplus \cdots \oplus T_r$. Also the subspaces $W_1, W_2, \cdots, W_r$ are said to reduce T or to form a T-invariant direct sum decomposition of V.

Remark 2: Consider the special case where two subspaces U and W reduce an operator $T : V \to V$; say dim U = 2 and dim W = 3 and suppose $\{u_1, u_2\}$ and $\{v_1, v_2, v_3\}$are basis of U and W respectively. If T_1 and T_2 denote respectively the restrictions of T to U and W, then

$$T_1(u_1) = a_{11}u_1 + a_{12}u_2$$

$$T_1(u_2) = a_{21}u_1 + a_{22}u_2$$

$$T_2(w_1) = b_{11}w_1 + b_{12}w_2 + b_{13}w_3$$

$$T_2(w_2) = b_{21}w_1 + b_{22}w_2 + b_{23}w_3$$

$$T_2(w_3) = b_{31}w_1 + b_{32}w_2 + b_{33}w_3$$

Hence

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

are matrix representations of T_1 and T_2 respectively. By the above theorem $\{u_1, u_2, v_1, v_2, v_3\}$ is a basis of V. Since $T(u_i) = T_1(u_i)$ and $T(w_j) = T_2(W_j)$. The matrix in the basis is the block diagonal matrix $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$.

Remark 3: Suppose $T : V \to V$ is linear and V is the direct sum of T-invariant subspaces $W_1, W_2, \cdots, W_r$. If A_i is the matrix representation of the restrictions of T to W_i, then T can be represented by the block diagonal matrix

$$M = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & A_r \end{pmatrix}$$

The block diagonal matrix M with diagonal entries $A_1, A_2, \cdots, A_r$ is sometimes called the direct sum of the matrices $A_1, A_2, \cdots, A_r$ and is denoted by $M = A_1 \oplus A_2 \oplus \cdots \oplus A_r$.

Theorem 7: Suppose $T : V \to V$ is linear and for $f(t) = g(t)h(t)$ are polynomials such that $f(T) = 0$ and $g(t)$ and $h(t)$ are relatively prime. Then V is the direct sum of the T-invariant subspaces U and W, where $U = \ker g(T)$ and $W = \ker h(T)$.

Proof: Note first that U and W are T-invariants. Now, since $g(t)$ and $h(t)$ are relatively prime, there exists polynomials $r(t)$ and $s(t)$ such that

$$r(t)g(t) + s(t)h(t) = 1$$

Hence for the operator T,

$$r(T)g(T) + s(T)h(T) = I \qquad \ldots(1)$$

Let $v \in V$, then by (1), we have

$$r(T)g(T)v + s(T)h(T)v = v$$

But the first term in this sum belongs to $W = \ker h(T)$. Since

$$h(T)r(T)g(T)v = r(T)g(T)h(T)v = r(T)f(T)v$$
$$= r(T).0 = 0$$

Similarly, the second term in this sum belongs to $U = \ker g(T)$.

Hence, V is the direct sum of U and W.

To prove $V = U \oplus W$, we must show that the representation $v = u + w$ is unique for v, u and w respectively, the elements of V, U and W.

Applying the operator $r(T)g(T)$ to $v = u + v$ using $g(T)u = 0,$ we obtain

$$r(T)g(T)v = r(T)g(T)u + r(T)g(T)w = r(T)g(T)w$$

Again applying (1) to w alone and using $h(T)w = 0$, we obtain

$$w = r(T)g(T)w + s(T)h(T)w = r(T)g(T)w$$

Both of the above formulae gives us $w = r(T)g(T)v$ and w is uniquely determined by v. Similarly u is uniquely determined by v. Hence $V = U \oplus W$, as required.

Corollary: If f(t) is the minimal polynomial of T [g(t) and h(t) are monic], then g(t) and h(t) are the minimal polynomials of the restrictions of T to U and W respectively.

4.7: Primary Decomposition Theorem

Statement: Let $T : V \to V$ be a linear operator with minimal polynomial

$$m(t) = f_1(t)^{n_1} f_2(t)^{n_2} \cdots f_r(t)^{n_r}$$

where $f_i(t)$ are the distinct monic irreducible polynomials. Then V is the direct sum of T-invariant subspaces $W_1, W_2, \cdots, W_r$, where W_i is the kernel of $f_i(T)^{n_i}$. Moreover $f_i(t)^{n_i}$ is the minimal polynomial of the restriction of T to W_i.

Proof: To prove this result, we use induction on r.

Clearly for r = 1, the result is trivial.

Assume that the result is true for all values up to r-1. By using above theorem, we can write v as the direct sum of T-invariant subspaces W_1 and V_1, where W_1 is the kernel of $f_1(T)^{n_1}$ and V_1 is the kernel of $f_2(T)^{n_2} \cdots f_r(T)^{n_r}$. Also by using above corollary, the minimal polynomial of the restrictions of T to W_1 and V_1 are respectively $f_1(T)^{n_1}$ and $f_2(T)^{n_2} \cdots f_r(T)^{n_r}$.

Denote the restrictions of T to V_1 by T_1. By the induction hypothesis V_1 is the direct sum of subspaces $W_2,\cdots,W_r$ such that W_i is the kernel of $f_i(T)^{n_i}$ and such that $f_i(t)^{n_i}$ is the minimal polynomial for the restriction of T to W_i. But the kernel of $f_i(T)^{n_i}$ for $i=1,2,\cdots,r$ is necessarily contained in V_1 since $f_i(t)^{n_i}$ divides $f_2(t)^{n_2}\cdots f_r(t)^{n_r}$. Thus the kernel of $f_i(T)^{n_i}$ is the same as the $f_i(T_1)^{n_i}$, which is W_i. Also the restriction of T to W_i is the same as the restriction of T_1 to $W_i; i=1,2,\cdots,r$. Hence $f_i(t)^{n_i}$ is also the minimal polynomial for the restriction of T to W_i, thus $V=W_1\oplus W_2\oplus\cdots\oplus W_r$ is the desired decomposition of T.

Theorem: A linear operator $T:V\to V$ has a diagonal matrix representation if and only if its minimal polynomial m(t) is a product of distinct linear polynomials.

Proof: Suppose m(t) is a product of distinct linear polynomials, say

$$m(t)=(t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_r)$$

where the λ_i are distinct scalars. By the primary decomposition theorem, V is the direct sum of subspaces $W_1,W_2,\cdots,W_r$, where $W_i=\ker(T-\lambda_i I)$, thus if $v\in W_i$, then $(T-\lambda_i I)v=0$ or $T(v)=\lambda_i v$. In other words, every vector in W_i is an eigenvector belonging to the eigenvalue λ_i. But we know that the union of bases for $W_1,W_2,\cdots,W_r$ is a bases of V. This basis consists of eigenvectors and so T is diagonalizable.

Conversely, suppose T is diagonalizable, i.e. V has a basis consisting of eigenvectors of T. Let $\lambda_1,\lambda_2,\cdots,\lambda_s$ be the distinct eigenvalues of T, then the operator

$$f(T)=(T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_s I)$$

maps each basis vector into zero.

Thus f(T) = 0 and hence the minimum polynomial m(t) of T divides the polynomial $f(t)=(t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_s)$

Accordingly m(t) is the product of distinct linear polynomials.

Example: Suppose $A \neq I$ is a square matrix for which $A^3 = I$. Determine whether or not A is similar to a diagonal matrix if A is a matrix over (i) the field of reals R, (ii) the complex field C.

Since $A^3 = I$, A is a zero of the polynomial $f(t)=t^3-1=(t-1)(t^2+t+1)$

The minimal polynomial m(t) of A can not be $(t-1)$, since $A \neq I$

Hence $m(t)=(t^2+t+1)$ or $m(t)=t^3-1$

Since neither polynomial is a product of linear polynomials over R, A is not diagonalizable over R. On the other hand each of the polynomial is a product of distinct linear polynomials over C. Hence A is diagonalizable over C.

4.8: Nilpotent Operators

A linear operator $T:V\to V$ is termed nilpotent if $T^n = 0$, for some positive integer n; we call k the index of nilpotency of T if $T^k = 0$ but $T^{k-1}\neq 0$. Analogously, a square matrix A is termed nilpotent if $A^n = 0$ for some positive integer n; and of index k if $A^k = 0$ but $A^{k-1}\neq 0$.

Clearly, the minimum polynomial of a nilpotent operator (matrix) of index k is $m(t)=t^k$, hence 0 is its only eigenvalue.

Theorem: Let $T:V\to V$ be linear and for $v\in V$, $T^k(v)=0$ but $T^{k-1}(v)\neq 0$. Prove that:

a) The set $S=\{v, T(v), \cdots, T^{k-1}(v)\}$ is linearly independent.
b) The subspace W generated by S is T-invariant.
c) The restriction $\hat{T}$ of T to W is nilpotent of index k.

d) Relative to the basis $\{T^{k-1}(v),\cdots,T(v),v\}$ of W, the matrix of T is of the form

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

Hence the above k-square matrix is nilpotent of index k.

Proof: (a) Suppose $a_0 v + a_1 T(v) + a_2 T^2(v) + \cdots + a_{k-1} T^{k-1}(v) = 0$ …(1)

Apply T^{k-1} to (1) and using $T^k(v)=0$, we obtain $aT^{k-1}(v)=0$; since $T^{k-1}(v)\neq 0,\ a=0$.

Now applying T^{k-2} to (1) and using $T^k(v)=0$ and $a=0$, we find $a_1 T^{k-1}(v)=0$; hence $a_1 = 0$.

Next applying T^{k-3} to (1) and using $T^k(v)=0$ and $a = a_1 = 0$, we find $a_2 T^{k-1}(v)=0$; hence $a_2 = 0$.

Continuing this process, we find that all the a's are 0; hence S is linearly independent.

(b) Let $v\in W$. Then $v = bv + b_1 T(v) + b_2 T^2(v) + \cdots + b_{k-1} T^{k-1}(v)$

Using $T^k(v)=0$, we have that

$$T(v) = bT(v) + b_1 T^2(v) + b_2 T^3(v) + \cdots + b_{k-2} T^{k-1}(v) \in W$$

Thus W is T- invariant.

(c) By hypothesis $T^k(v)=0$, hence for $i = 0, 1, 2, \cdots, k-1$

$$\hat{T}(T^i(v)) = T^{k+i}(v) = 0$$

i.e., applying $\hat{T}^k$ to each generator of W, we obtain 0; hence $\hat{T}^k = 0$ and so $\hat{T}$ is nilpotent of index at most k. On the other hand $\hat{T}^{k-1}(v) = T^{k-1}(v) = 0$.

Hence T is nilpotent of index k.

(d) For the basis $\{T^{k-1}(v),\cdots,T(v),v\}$ of W,

$$\hat{T}(T^{k-1}(v)) = T^k(v) = 0$$

$$\hat{T}(T^{k-2}(v)) = \quad T^{k-1}(v)$$

$$\hat{T}(T^{k-3}(v)) = \quad T^{k-2}(v)$$

…………………………………

$$\hat{T}(T(v)) = \quad T^2(v)$$

$$\hat{T}(v) = \quad T(v)$$

Hence the matrix in this basis is

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

Theorem: Let $T:V\to V$ be linear and $U=\ker T^i$ and $W=\ker T^{i+1}$ Show that (i) $U\subset W$ but $T^{k-1}(v)\neq 0$ (ii) $T(W)\subset W$.

Proof: (i) Suppose $u\in U=\ker T^i$, then $T^i(u)=0$ and so $T^{i+1}(u)=T(T^i(u))=T(0)=0$. Thus $u\in \ker T^{i+1}=W$

But this is true for every $u\in U$. Hence $U\subset W$

(ii) Similarly if $w\in W=\ker T^{i+1}$ then $T^{i+1}(w)=0$

Thus $T^{i+1}(w)=T^i(T(w))=T^i(0)=0$

Therefore, $T(W)\subset W$

Theorem: Let $T:V\to V$ be a nilpotent operator of index k. Then T has a block diagonal matrix representation whose diagonal entries are of the form

$$N = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

There is at least one N of order k and all other N are of order less or equal to k. The number of N of each possible order is uniquely determined by T. Moreover, the total number of N of all orders is the nullity of T.

Proof: Suppose dimV = n and $W_1 = \ker T$, $W_2 = \ker T^2$, . . . , $W_k = \ker T^k$.

Set $m_i = \dim W_i$ for $i = 1, 2, \cdots, k$. Since T is of index k, $W_k = V$ and $W_{k-1} \neq V$, and so $m_{k-1} < m_k = n$. We know that $W_1 \subset W_2 \subset \ldots \subset W_k = V$.

Thus by induction, we can choose a basis $\{u_1, u_2, \cdots, u_n\}$ of V such that $\{u_1, u_2, \cdots, u_{m_i}\}$ is a basis of W_i.

We now choose a new basis for V with respect to which T has the desired form. It will be convenient to label the members of this new basis by pairs of indices. We begin by setting

$v(1,k) = u_{m_{k-1}+1}$, $v(2,k) = u_{m_{k-1}+2}$, . . . , $v(m_k - m_{k-1}, k) = u_{m_k}$ and setting $v(1,k-1) = Tv(1,k)$, $v(2,k-1) = Tv(2,k)$, . . . , $v(m_k - m_{k-1}, k-1) = Tv(m_k - m_{k-1}, k)$

We also know that $S_1 = \{u_1, \cdots, u_{m_{k-2}}, v(1,k-1), \cdots, v(m_k - m_{k-1}, k-1)\}$ is linearly independent subset of W_{k-1}.

We extend S_1 to a basis of W_{k-1} by adjoining new elements (if necessary) which can be done $v(m_k - m_{k-1} + 1, k-1), v(m_k - m_{k-1} + 2, k-1)$, . . . , $v(m_k - m_{k-2}, k-1)$

Next we set $v(1,k-2) = Tv(1,k-1)$, $v(2,k-2) = Tv(2,k-1)$, . . . , $v(m_{k-1} - m_{k-2}, k-2) = Tv(m_{k-1} - m_{k-2}, k-1)$

Again, we have $S_2 = \{u_1, \cdots, u_{m_{k-3}}, v(1,k-2), \cdots, v(m_{k-1} - m_{k-2}, k-2)\}$ is linearly independent subset of W_{k-2} which we can extend to a basis of W_{k-2} by adjoining elements

$v(m_{k-1} - m_{k-2} + 1, k-2)$, $v(m_{k-1} - m_{k-2} + 2, k-2)$, . . . , $v(m_{k-2} - m_{k-3}, k-2)$

Continuing in this manner, we get a new basis for V, which for convenient reference we arrange as follows

$$v(1,k), \cdots, v(m_k - m_{k-1}, k)$$

$$v(1,k-1), \cdots, v(m_k - m_{k-1}, k-1), \cdots, v(m_{k-1} - m_{k-2}, k-1)$$

……………………………………………………..

$$v(1,2), \cdots, v(m_k - m_{k-1}, 2), \cdots, v(m_{k-1} - m_{k-2}, 2), \cdots, v(m_2 - m_1, 2)$$

$$v(1,1), \cdots, v(m_k - m_{k-1}, 1), \cdots, v(m_{k-1} - m_{k-2}, 1), \cdots, v(m_2 - m_1, 1), \cdots, v(m_1, 1)$$

The bottom row forms a basis of W_1, the bottom two rows form a basis of W_2, etc. But what is important for us is that T maps each vector immediately below it in the table or into 0 if the vector is in the bottom is in the bottom row. That is

$$Tv(i,j) = \begin{cases} v(i, j-1) & for \quad j > 1 \\ 0 & for \quad j = 1 \end{cases}$$

Now it is clear from the above theorem-(iv) that T will have the desired form if the v(i, j) are ordered lexicographically: beginning with v(1, 1) and moving up the first column to v(1, k), then jumping to v(2, 1) and moving up the second column as far as possible, etc.

Moreover, there will be exactly

$m_k - m_{k-1}$ diagonal entries of order k

$(m_{k-1} - m_{k-2}) - (m_k - m_{k-1}) = 2m_{k-1} - m_k - m_{k-2}$ diagonal entries of order k-1

………………………………………………………………………………

$2m_2 - m_1 - m_3$ diagonal entries of order 2

$2m_1 - m_2$ diagonal entries of order 1.

as can be directly read off from the table.

In particular since the numbers $m_1, m_2, \cdots, m_k$ are uniquely determined by T.

Finally the identity

$$m_1 = (m_k - m_{k-1}) + (2m_{k-1} - m_k - m_{k-2}) + \cdots + (2m_2 - m_1 - m_3) + (2m_1 - m_2)$$

shows that the nullity m_1 of T is the total number of diagonal entries of T.

Theorem: Let $T : V \to V$ be linear and $X = \ker T^{i-2}$, $Y = \ker T^{i-1}$ and $Z = \ker T^i$, then $X \subset Y \subset Z$. Suppose $\{u_1, u_2, \cdots, u_r\}$, $\{u_1, u_2, \cdots, u_r, v_1, v_2, \cdots, v_s\}$ and $\{u_1, u_2, \cdots, u_r, v_1, v_2, \cdots, v_s, w_1, w_2, \cdots, w_t\}$ are bases of X, Y and Z respectively. Then show that $S = \{u_1, u_2, \cdots, u_r, T(w_1), T(w_2), \cdots, T(w_t)\}$ is contained in Yand is linearly independent.

Proof: From the above theorem, we can easily write $T(Z) \subset Y$ and hence $S \subset Y$.

Now suppose S is linearly independent, then there exists a relation

$$a_1u_1 + a_2u_2 + \cdots + a_ru_r + b_1T(w_1) + b_2T(w_2) + \cdots + b_tT(w_t) = 0$$

where at least one coefficient is non-zero, Further, since $\{u_i\}$ is linearly independent, at least one of the b_k must be non-zero. Transposing, we find

$$b_1T(w_1) + b_2T(w_2) + \cdots + b_tT(w_t) = -a_1u_1 - a_2u_2 - \cdots - a_ru_r \in X = \ker T^{i-2}$$

Hence $T^{i-2}(b_1T(w_1) + b_2T(w_2) + \cdots + b_tT(w_t)) = 0$

Thus, $T^{i-1}(b_1w_1 + b_2w_2 + \cdots + b_tw_t) = 0$ and so

$$b_1w_1 + b_2w_2 + \cdots + b_tw_t \in Y = \ker T^{i-1}$$

Since $\{u_i, v_j\}$ generate Y, we obtain a relation among the u_i, v_j and w_k where one of the coefficients i.e., b_k, is not zero. This contradicts the fact that $\{u_i, v_j, w_k\}$ is independent. Hence S must also be linearly independent.

4.9: Jordan Canonical Form

An operator T can be put into Jordan canonical form if its characteristic and minimal polynomial factor into linear polynomials. This is always true if K is the complex field C. In any case, we can always extend the base field K to a field in which the characteristic and minimum polynomial do factor into linear factors; thus in a broad sense every operator has a Jordan Canonical Form. Analogously, every matrix is similar to a matrix in Jordan Canonical form.

Exercise: Let

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \text{ then } A^2 = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \text{ and } A^3 = 0$$

Hence A is nilpotent of index 2.

Find the nilpotent matrix M in canonical form which is similar to A.

Solution: Since A is nilpotent of index 2, M contains a diagonal block of index 2 and none greater than 2. Note that rank A = 2; hence nullity of A = 5-2 = 3. Thus M contains 3 diagonal blocks. Accordingly M must contain 2 diagonal blocks of order 2 and 1 of order 1; that is

$$M = \begin{pmatrix} 0 & 1 & \vdots & 0 & 0 & & 0 \\ 0 & 0 & \vdots & 0 & 0 & & 0 \\ \cdots & \cdots & \vdots & \cdots & \cdots & \cdots & \\ 0 & 0 & \vdots & 0 & 1 & \vdots & 0 \\ 0 & 0 & \vdots & 0 & 0 & \vdots & 0 \\ 0 & 0 & \vdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \vdots & 0 \end{pmatrix}$$

Theorem: Let $T : V \to V$ be a linear operator whose characteristic and minimal polynomials respectively are

$\Delta(t)=(t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2}\cdots(t-\lambda_r)^{n_r}$ and $m(t)=(t-\lambda_1)^{m_1}(t-\lambda_2)^{m_2}\cdots(t-\lambda_r)^{m_r}$ where the $\lambda_i's$ are distinct scalars. The T has a block diagonal matrix representation J whose diagonal entries are of the form

$$J_{ij}=\begin{pmatrix}\lambda_i & 1 & 0 & \cdots & 0 & 0\\ 0 & \lambda_i & 1 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & \lambda_i & 1\\ 0 & 0 & 0 & \cdots & 0 & \lambda_i\end{pmatrix}$$

For each λ_i, the corresponding blocks J_{ij} have the following properties:

i) There is at least one J_{ij} of order m_i; all other J_{ij} are of the order $\leq m_i$.

ii) The sum of the orders of the J_{ij} is n_i.

iii) The number of J_{ij} equals the geometric multiplicity of λ_i.

iv) The number of J_{ij} of each possible order is uniquely determined by T.

Proof: By the Primary Decomposition theorem, T is decomposable into operators $T_1,T_2,\cdots,T_r$, i.e., $T=T_1\oplus T_2\oplus\cdots\oplus T_r$; where $(t-\lambda_i)^{m_i}$ is the minimal polynomial of T_i. Thus in particular,

$$(T_1-\lambda_i I)^{m_i}=0,\cdots,(T_r-\lambda_r I)^{m_r}=0$$

Set $N_i=T_i-\lambda_i I$; then for $i=1,2,\cdots,r$

$$T_i=N_i+\lambda_i I \quad \text{where } N_i^{m_i}=0$$

That is T_i is the sum of the scalar operator $\lambda_i I$ and a nilpotent operator N_i, which is of index m_i, Since $(t-\lambda_i)^{m_i}$ is the minimal polynomial of T_i.

Now by above theorem on nilpotent operators, we can choose a basis so that N_i is in canonical form. In this basis, $T_i=N_i+\lambda_i I$ is represented by a block diagonal matrix M_i whose diagonal entries are the matrices J_{ij}. The direct sum J of the matrices M_i is in Jordan canonical form and by remark-3, is a matrix representation of T.

Lastly, we must show that the blocks J_{ij} satisfy the required properties. Property (1) follows from the fact that N_i is of index m_i. Property (ii) is true since T and J have the same characteristic polynomial. Property (iii) is true since the nullity of $N_i = T_i - \lambda_i I$ is equal to the geometric multiplicity of the eigenvalue λ_i. Property (iv) follows from the fact that the T_i and hence the N_i are uniquely determined by T.

Remark: The matrix J appears in the above theorem is called the Jordan Canonical form of the operator T. A diagonal block J_{ij} is called a Jordan Block belonging to the eigenvalue λ_i.

Observe that

$$\begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix} = \begin{pmatrix} \lambda_i & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 0 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

i.e., $$J_{ij} = \lambda_i I + N$$

where N is the nilpotent block appearing in the previous theorem. In fact we have proved the same in the above theorem by showing that T can be decomposed into operators, each the sum of a scalar and a nilpotent operator.

Example: Suppose the characteristic and minimum polynomial of an operator T are respectively

$$\Delta(t) = (t-2)^4 (t-3)^3 \text{ and } m(t) = (t-2)^2 (t-3)^2$$

Then the Jordan canonical form of T is one of the following matrices

$$\begin{pmatrix} 2 & 1 & \vdots & & & & & & \\ 0 & 2 & \vdots & & & & & & \\ \cdots & \cdots & \vdots & \cdots & \cdots & \cdots & & & \\ & & \vdots & 2 & 1 & \vdots & & & \\ & & \vdots & 0 & 2 & \vdots & & & \\ & & \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ & & & & & \vdots & 3 & 1 & \vdots \\ & & & & & \vdots & 0 & 3 & \vdots \\ & & & & & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & & & & & \vdots & 3 \end{pmatrix} \text{ or } \begin{pmatrix} 2 & 1 & \vdots & & & & & & \\ 0 & 2 & \vdots & & & & & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & & & & \\ & & \vdots & 2 & \vdots & & & & \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & & \\ & & & & \vdots & 2 & \vdots & & \\ & & & & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & & & \vdots & 3 & 1 & \vdots \\ & & & & & & \vdots & 0 & 3 & \vdots \\ & & & & & & \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & & & & & & & \vdots & 3 \end{pmatrix}$$

The first matrix occurs if T has two independent eigenvectors belonging to its eigenvalue 2; and the second matrix occurs if T has three independent vectors belonging to 2.

Exercise: Determine all possible Jordan canonical forms for a linear operator $T : V \to V$ whose characteristic polynomial is $\Delta(t) = (t-2)^3 (t-5)^2$.

Solution: Since $t-2$ has exponent 3 in $\Delta(t)$ must appear three times on the main diagonal. Similarly 5 must appear twice. Thus, the possible Jordan canonical forms are:

(i) $\begin{pmatrix} 2 & 2 & & & \\ & 2 & 1 & & \\ & & 2 & & \\ & & & 5 & 1 \\ & & & & 5 \end{pmatrix}$ (ii) $\begin{pmatrix} 2 & 2 & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 5 & 1 \\ & & & & 5 \end{pmatrix}$ (iii) $\begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 5 & 1 \\ & & & & 5 \end{pmatrix}$

(iv) $\begin{pmatrix} 2 & 1 & & & \\ & 2 & 1 & & \\ & & 2 & & \\ & & & 5 & \\ & & & & 5 \end{pmatrix}$ (v) $\begin{pmatrix} 2 & 2 & & & \\ & 2 & 1 & & \\ & & 2 & & \\ & & & 5 & \\ & & & & 5 \end{pmatrix}$ (vi) $\begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 5 & \\ & & & & 5 \end{pmatrix}$

Exercise: Determine all possible Jordan canonical forms J for a matrix of order 5 whose minimal polynomial is $m(t) = (t-2)^2$

Solution: Clearly J must have one Jordan block of order 2 and the others must be of order 2 and 1. Thus, there are only two possibilities

$$J = \begin{pmatrix} 2 & 1 & & & \\ & 2 & & & \\ & & 2 & 1 & \\ & & & 2 & \\ & & & & 2 \end{pmatrix} \quad \text{or} \quad J = \begin{pmatrix} 2 & 1 & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 5 & \\ & & & & 5 \end{pmatrix}$$

Note that all the diagonal entries must be 2. Since 2 is the only eigenvalue.

Cyclic Subspaces

Let T be a linear operator on a vector space V of finite dimension over K. Suppose $v(\neq 0) \in V$, the set of all vectors of the form $f(T)(v)$, where $f(t)$ ranges over all polynomials over K, is a T-invariant subspace of V called the T-cyclic subspace of V generated by v; we denote it by $Z(v,T)$ and denote the restriction of T to $Z(v,T)$ by T_v. We could equivalently define $Z(v,T)$ as the intersection of all T-invariant subspaces of V containing v.

Remark 5: Consider the sequence

$$v, \ T(v), \ T^2(v), \ T^3(v), \ \cdots$$

of powers of T acting on v. Let k be the lowest integer such that $T^k(v)$ is linear combination of those vectors which precede it in the sequence; say

$$T^k(v) = -a_{k-1}T^{k-1}(v) - \cdots - a_1 T(v) - a_0 v$$

Then $$m_v(t) = t^k + a_{k-1}t^{k-1} + \cdots + a_1 t + a_0$$

is the unique monic polynomial of lowest degree for which $m_v(T)(v) = 0$, we call $m_v(t)$ the T-annihilator of v and $Z(v,T)$.

Remark 6: Suppose $Z(v,T)$, T_v and $m_v(t)$ be defined as above, then

i) The set $\{v,\ T(v),\ \cdots, T^{k-1}(v)\}$ is a basis of $Z(v,T)$; hence $\dim Z(v,T)$ = k

ii) The minimal polynomial of T_v is $m_v(t)$.

iii) The matrix representation of T_v in the above basis is

$$C = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & -a_{k-2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{k-1} \end{pmatrix}$$

The matrix C is called the companion matrix of the polynomial $m_v(t)$

Proof: (i) By definition of $m_v(t)$, $T_k(v)$ is the first vector in the sequence $v,\ T(v),\ T^2(v),\ T^3(v),\ \cdots$ which is a linear combination of those vectors which precede it in the sequence; hence the set $B = \{v,\ T(v),\ \cdots, T^{k-1}(v)\}$ is linearly independent. We now only have to show that $Z(v,T) = L(B)$, the linear span of B. But we have $T^k(v) \in L(B)$. We prove by induction that $T^n(v) \in L(B)$ for every n. Suppose $n > k$ and $T^{n-1}(v) \in L(B)$, i.e., $T^{n-1}(v)$ is a linear combination of $v,\ T(v),\ \cdots, T^{k-1}(v)$. Then $T^n(v) = T(T^{n-1}(v))$ is a linear combination of $T(v),\ \cdots, T^k(v)$, but $T^k(v) \in L(B)$; hence $T^n(v) \in L(B)$ for every n. Consequently $f(T)(v) \in L(B)$ for any polynomial $f(t)$.

Thus $Z(v,T) = L(B)$ and so B is a basis as claimed

(ii) Suppose $m(t) = t^s + b_{s-1}t^{s-1} + \cdots + b_1 t + b_0$ is a minimal polynomial of T_v. Then since $v \in Z(v,T)$

$$0 = m(T_v)(v) = m(T)(v) = T^s(v) + b_{s-1}T^{s-1}(v) + \cdots + b_1 T + b_0 v$$

Thus, $T^s(v)$ is a linear combination of $v,\ T(v), \cdots, T^{s-1}(v)$ and therefore $k \le s$. However, $m_v(T)=0$ and so $m_v(T_v)=0$. Then $m(t)$ divides $m_v(t)$ and so $s \le k$. Accordingly $s = k$ and hence $m(t) = m_v(t)$.

(iii) we have

$$T_v(v) = \qquad T(v)$$

$$T_v(T(v)) = \qquad T^2(v)$$

……………………………………………………….

$$T_v\left(T^{k-2}(v)\right) = \qquad T^{k-1}(v)$$

$$T^k(v) = T_v\left(T^{k-1}(v)\right) = -a_0 v - a_1 T(v) - a_2 T^2(v) - \ \cdots \ - a_{k-1} T^{k-1}(v)$$

By definition, the matrix T_v in this basis is the transpose of the matrix of coefficients of the above system of equations; hence it is C, as required.

4.10: Rational Canonical Form

In this section, we present the rational canonical form for a linear operator $T: V \to V$. We emphasize that this form exists even when the minimal polynomial can not be factorized into linear polynomials [Recall this is not the case in Jordan canonical form].

Lemma: Let $T: V \to V$ be a linear operator whose minimal polynomial is $f(t)^n$ where $f(t)$ is a monic irreducible polynomial. Then V is the direct sum

$$V = Z(v_2, T) \oplus Z(v_3, T) \oplus \cdots \oplus Z(v_r, T)$$

of T-cyclic subspaces $Z(v_i, T)$ with corresponding T-annihilators $f(t)^{n_1}, f(t)^{n_2}, \cdots, f(t)^{n_r}, \quad n = n_1 \ge n_2 \ge \cdots \ge n_r$

Any other decomposition of V into T-cyclic subspaces has the same number of components and the same set of T-annihilators.

We emphasize that the above lemma does not say that the vectors v_i or the T-cyclic subspaces $Z(v_i, T)$ are uniquely determined by T; but it does not say that the set of T-annihilators are uniquely determined by T. Thus T has a unique matrix representation

$$\begin{pmatrix} c_1 & & & \\ & c_2 & & \\ & & \cdots & \\ & & & c_3 \end{pmatrix}$$

where c_i are companion matrices. In fact, are the companion matrices to the polynomial $f(t)^{n_i}$

Theorem: Let $T : V \to V$ be linear. Let W be a T-invariant subspace of V and $\bar{T}$ the induced operator on V/W. Prove (i) The T-annihilator of $v \in V$ divides the minimal polynomial of T. (ii) The $\bar{T}$-annihilator of $\bar{v} \in V/W$ divides the minimal polynomial of T.

Proof: (i) The T-annihilator of $v \in V$ is the minimal polynomial of the restriction of T to $Z(v, T)$ and therefore as we know, it divides the minimal polynomial of T.

(ii) The $\bar{T}$-annihilator of $\bar{v} \in V/W$ divides the minimal polynomial of $\bar{T}$, which divides the minimal polynomial of T.

Note: In case the minimal polynomial of T $f(t)^n$ where $f(t)$ is a monic irreducible polynomial, then the T-annihilator of $v \in V$ and $\bar{T}$-annihilator of $\bar{v} \in V/W$ are of the form $f(t)^m$; where $m \le n$.

Remark: Let $T : V \to V$ be a linear operator with minimal polynomial $m(t) = f_1(t)^{m_1} f_2(t)^{m_2} \cdots f_r(t)^{m_s}$; where $f_i(t)$ are distinct monic irreducible polynomials. Then T has a unique block diagonal matrix representation

$$\begin{pmatrix} c_{11} & & & & & & \\ & \cdots & & & & & \\ & & c_{1r_1} & & & & \\ & & & \cdots & & & \\ & & & & c_{s1} & & \\ & & & & & \cdots & \\ & & & & & & c_{sr_s} \end{pmatrix}$$

where c_{ij} are companion matrices. In particular, the c_{ij} are the companion matrices of the polynomials $f_i(t)^{n_{ij}}$,

where $\quad m = n_{11} \geq n_{12} \geq \cdots \geq n_{1r1} \geq \cdots, m_s = n_{s1} \geq n_{s2} \geq \cdots \geq n_{srs}$

The above matrix representation of T is called its rational canonical form. The polynomials $f_i(t)^{n_{ij}}$ are called the elementary divisors of T.

Example: Let V be a vector space of dimension 6 over R, and let T be a linear operator whose minimal polynomial is $m(t) = (t^2 - t + 3)(t-2)^2$. Then the rational canonical form of T is one of the following direct sum of companion matrices

(i) $C(t^2 - t + 3) \oplus C(t^2 - t + 3) \oplus C(t-2)^2$

(ii) $C(t^2 - t + 3) \oplus C(t-2)^2 \oplus C(t-2)^2$

(iii) $C(t^2 - t + 3) \oplus C(t-2)^2 \oplus C(t-2) \oplus C(t-2)$

where $C(f(t))$ is the companion matrix of $f(t)$; that is

$$\left(\begin{array}{cc|cc|cc} 0 & -3 & & & & \\ 1 & 1 & & & & \\ \hline & & 0 & -3 & & \\ & & 1 & 1 & & \\ \hline & & & & 0 & -4 \\ & & & & 1 & 4 \end{array}\right) \left(\begin{array}{cc|cc|cc} 0 & -3 & & & & \\ 1 & 1 & & & & \\ \hline & & 0 & -4 & & \\ & & 1 & 4 & & \\ \hline & & & & 0 & -4 \\ & & & & 1 & 4 \end{array}\right) \left(\begin{array}{cc|cc|c|c} 0 & -3 & & & & \\ 1 & 1 & & & & \\ \hline & & 0 & -4 & & \\ & & 1 & 4 & & \\ \hline & & & & 2 & \\ \hline & & & & & 2 \end{array}\right)$$

(i) (ii) (iii)

Exercise: Let V be a vector space of dimension 7 over R, and let $T: V \to V$ be a linear operator whose minimal polynomial is $m(t) = (t^2+2)(t+3)^3$. Find the all possible rational canonical forms for T.

Solution: The sum of the degrees of the companion matrices must add up to 7. Also, one companion matrix must be (t^2+2) and one must be $(t+3)^3$. Thus the rational canonical form of T is exactly one of the following direct sum of companion matrices:

(i) $C(t^2+2) \oplus C(t^2+2) \oplus C(t-3)^3$

(ii) $C(t^2+2) \oplus C(t+3)^3 \oplus C(t+3)^2$

(iii) $C(t^2+2) \oplus C(t+3)^3 \oplus C(t+3) \oplus C(t+3)$

i.e.,

$$\begin{pmatrix} 0 & -2 & & & & & \\ 1 & 0 & & & & & \\ & & 0 & -2 & & & \\ & & 1 & 0 & & & \\ & & & & 0 & 0 & -27 \\ & & & & 1 & 0 & -27 \\ & & & & 0 & 1 & -9 \end{pmatrix} \begin{pmatrix} 0 & -2 & & & & & \\ 1 & 0 & & & & & \\ & & 0 & 0 & -27 & & \\ & & 1 & 0 & -27 & & \\ & & 0 & 1 & -9 & & \\ & & & & & 0 & -9 \\ & & & & 0 & 1 & -6 \end{pmatrix} \begin{pmatrix} 0 & -2 & & & & & \\ 1 & 0 & & & & & \\ & & 0 & 0 & -27 & & \\ & & 1 & 0 & -27 & & \\ & & 0 & 1 & -9 & & \\ & & & & & -3 & \\ & & & & 0 & & -3 \end{pmatrix}$$

(i) (ii) (iii)

Exercise 1: Find all possible rational canonical forms for:

i) Matrices of order 6 with minimal polynomial $(t^2+3)(t+1)^2$

ii) Matrices of order 6 with minimal polynomial $(t+1)^3$

iii) Matrices of order 8 with minimal polynomial $(t^2+2)^2\,(t+3)^2$

Exercise 2: Find the rational canonical form of the Jordan block

$$\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

4.12: Quotient Spaces:

Let V be a vector space over a field K and W be a subspace of V. If v is any vector in V, we write v + W for the set of sums v + w with $w \in W$ and $v \in V$

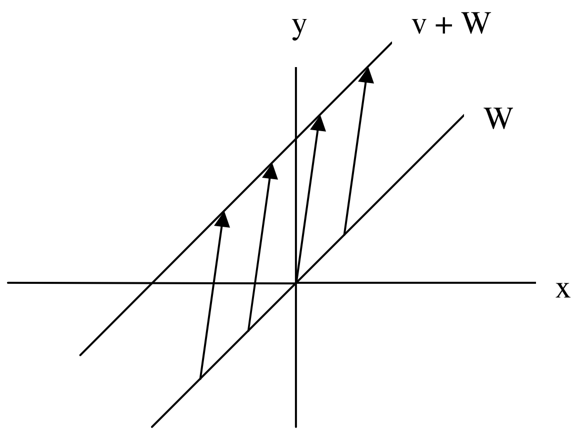
$$v + W = \{v + W : \quad w \in W\}$$

These sets are called the cosets of W in V.

Example: Let W be the subspace of R^2 defined by

$$W = \{(a, b) : a = b\}$$

i.e., W is the line given by the equation x – y = 0. We can view $v + W$ as a translation of the line, obtained by adding the vector v to each point in W. $v + W$ is also a line and is parallel to W. Thus the cosets of W in R^2 are precisely all the lines parallel to line.



Exercise: Let W be a subspace of a vector space V. Show that the following are equivalent:

i) $u \in v + W$ ii) $u - v \in W$ iii) $v \in u + W$

Solution: Suppose $u \in v + W$, then there exists $w_0 \in W$ such that $u = v + w_0$ hence $u - v = w_0 \in W$

Conversely, suppose $u - v \in W$, then $u - v = w_0$, where $w_0 \in W$.

Hence $u = v + w_0 \in v + W$. Thus (i) and (ii) are equivalent.

We also have $u-v\in W$ if and only if $-(u-v)=v-u\in W$ if and only if $v\in u+W$

Thus (ii) and (iii) are also equivalent.

Exercise: Prove that the cosets of W in V partition V into mutually disjoint sets, i.e,

(i) Any two cosets u + W and v + W are either identical or disjoint; and

(ii) Each $v\in V$ belongs to a coset; in fact $v\in v+W$.

Furthermore, $u+W=v+W$ if and only if $u-v\in W$, and so $(u+w)+W=v+W$ for any $w\in W$.

Proof: Let $v\in V$, we have $v-v+0\in v+W$ as $0\in W$ proving (ii).

Now suppose the coset $u+W$ and $v+W$ are not disjoint, say the vector x belongs to both. Clearly, $u-x\in W$ and $x-v\in W$. For any $w_0\in W$, let $u+w_0$ be any element in $u+W$.

Clearly, $(u-w_0)-v=(u-x)+(x-v)+w_0\in W$

Thereby, it follows that $u+w_0\in v+W$ and hence the coset $u+W$ is contained in $v+W$. Similarly $v+W\subseteq u+W$ and thus $v+W=u+W$

The last statement follows from the fact that $v+W=u+W$ if and only if $u\in v+W$ that is equivalent to $u-v\in W$.

Home Assignments

Exercise 1: Let W be the solution space of homogeneous equation $2x+3y+4z=0$. Describe the cosets of W in R^3

Exercise 2: Given a subspace W of a vector space V, show that the natural map $\eta:V\to V/W$ defined by $\eta(v)=v+W$ is linear.

Exercise: Let W be a subspace of a vector space V. Suppose $\{w_1,w_2,\cdots,w_r\}$ is a basis of W and the set of cosets $\{\bar{v}_1,\bar{v}_2,\cdots,\bar{v}_s\}$ where $\bar{v}_j=v_j+W$ a basis of the quotient space is. Show that $B=\{v_1,v_2,\cdots,v_s,w_1,w_2,\cdots,w_r\}$ is a basis of V. Thus dimV = dimW + dimV/W

Solution: Suppose $u\in V$, since $\{\bar{v}_j\}$ is a basis of V/W.

$$\bar{u}=u+W=a_1\bar{v}_1+a_2\bar{v}_2+\cdots+a_s\bar{v}_s$$

Hence $u=a_1v_1+a_2v_2+\cdots+a_sv_s+w$, where $w\in W$

Since $\{w_i\}$ is a basis of W

$$u=a_1v_1+a_2v_2+\cdots+a_sv_s+b_1w_1+b_2w_2+\cdots+b_rw_r$$

Accordingly B generates V.

We now show that B is linearly independent.

Suppose $c_1v_1+a_2v_2+\cdots+c_sv_s+d_1w_1+d_2w_2+\cdots+d_rw_r=0$

Then $c_1\bar{v}_1+c_2\bar{v}_2+\cdots+c_s\bar{v}_s=\bar{0}=W$

Since $\{\bar{v}_j\}$ is independent, then c's are all zero. Therefore, we get $d_1w_1+d_2w_2+\cdots+d_rw_r=0$. Also $\{w_i\}$ is independent. Therefore all d's are zero and hence B is a basis of V.

Theorem: Suppose W is a subspace invariant under a linear operator $T:V\to V$. Then T induces a linear operator $\bar{T}$ on V/W defined by $\bar{T}(v+W)=T(v)+W$. Moreover, if T is a zero of any polynomial, then so is $\bar{T}$, thus, the minimal polynomial of $\bar{T}$ divides the minimal polynomial of T.

Proof: We first show that $\bar{T}$ is well defined, i.e., if $u+W=v+W$, then $\bar{T}(u+W)=\bar{T}(v+W)$, If $u+W=v+W$, then $u-v\in W$ and since W is T-invariant, $T(u-v)=T(u)+T(v)\in W$

Accordingly, $\bar{T}(u+W)=T(u)+W=T(v)+W=\bar{T}(v+W)$ as required

We next show that $\bar{T}$ is linear. we have

$$\bar{T}(u+W)+(v+W)=\bar{T}(U+v+W)=T(u+v)+W=T(u)+T(v)+W$$
$$=T(u)+W+T(v)+W=\bar{T}(u+W)+\bar{T}(v+W)$$

and similarly, we can show that

$$\overline{T}(ku+W)=\overline{T}(ku+W)=T(ku)+W=kT(u)+W$$
$$=k(T(u)+W)=k\overline{T}(u+W)$$

Thus $\overline{T}$ is linear.

Now for any coset u + W in V/W,

$$\overline{T}^2(u+W)=T^2(u)+W=T(T(u))+W=\overline{T}(T(u)+W)=\overline{T}(\overline{T}(u+W))=\overline{T}^2(u+W)$$

Hence $\overline{T}^2=\overline{T^2}$. Similarly, $\overline{T}^n=\overline{T^n}$ for any n.

Thus for any polynomial

$$f(t)=a_nt^n+\cdots+a_0=\sum a_it^i$$

$$\overline{f(T)}(u+W)=f(T)(u)+W=\sum a_iT^i(u)+W=\sum a_i\left(T^i(u)+W\right)$$
$$=\sum a_i\overline{T}^i(u+W)=\sum a_i\overline{T^i}(u+W)$$
$$=\left(\sum a_i\overline{T^i}\right)(u+W)=f(\overline{T})(u+W)$$

and so $f\left(\overline{T}\right)=f(\overline{T})$. Accordingly if T is a root of $f(t)$ then

$\overline{f(T)}=\overline{0}=W=f(\overline{T})$, i.e., $\overline{T}$ is also a root of $f(t)$.

Theorem: Let $T:V\to V$ be a linear operator whose characteristic polynomial factors into linear polynomials. Then V has a basis in which T is represented by a triangular matrix.

Proof: To prove this result, we use induction on the dimension of V. If dimV = 1, then every matrix representation of T is of order 1 which is clearly triangular.

Now suppose dimV = n > 1 and that the theorem holds for spaces of dimension less than n. Since the characteristic polynomial of T factors into linear polynomials, T has at least one eigen value and so at least one non-zero eigenvalue v, say $T(v) = a_{11}v$. Let W be the 1-dimensional subspace spanned by v. Set $\overline{V} = V/W$, we have $\dim \overline{V} = \dim V - \dim W = n-1$. Note also that W is invariant under T. Theorem by previous theorem T induces a linear operator $\overline{T}$ on $\overline{V}$ whose minimal polynomial divides the minimal polynomial of T. since the characteristic polynomial of T is a product of linear polynomials, so is its minimum polynomial; hence so are the minimum and characteristic polynomials of $\overline{T}$. Thus $\overline{V}$ and $\overline{T}$ satisfy the hypothesis of the theorem. Hence by induction, there exists a basis $\{\overline{v}_2, \overline{v}_3, \cdots, \overline{v}_n\}$ of $\overline{V}$ such that

$$\overline{T}(\overline{v}_2) = a_{22}\overline{v}_2$$

$$\overline{T}(\overline{v}_3) = a_{32}\overline{v}_2 + a_{33}\overline{v}_3$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\overline{T}(\overline{v}_n) = a_{n2}\overline{v}_2 + a_{n3}\overline{v}_3 + \cdots + a_{nn}\overline{v}_n$$

Let $v_2, v_3, \cdots, v_n$ be elements of V which belong to the cosets $\overline{v}_2, \overline{v}_3, \cdots, \overline{v}_n$ respectively. Then $\{v, v_2, \cdots, v_n\}$ is a basis of V. Since $\overline{T}(\overline{v}_2) = a_{22}\overline{v}_2$, we have

$$\overline{T}(\overline{v}_2) - a_{22}\overline{v}_2 = 0 \text{ and so } T(v_2) - a_{22}v_2 \in W$$

But W is spanned by v, hence $T(v_2) - a_{22}v_2$ is a multiple of v, say

$$T(v_2) - a_{22}v_2 = a_{21}v \text{ and so } T(v_2) = a_{21}v + a_{22}v_2$$

Similarly, for $i = 3, 4, \cdots, n$

$$T(v_i) = a_{i1}v + a_{i2}v_2 + \cdots + a_{ii}v_i$$

Thus

$$T(v) = a_{11}v$$

$$T(v_2) = a_{21}v + a_{22}v_2$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$T(v_n) = a_{n1}v + a_{n2}v_2 + \cdots + a_{nn}v_n$$

and hence the matrix of T in this basis is triangular.

4.13: Bilinear Forms

Let V be a vector space of finite dimensions over a field K. A bilinear form on V is a map $f: V\times V \to K$ which satisfies

i) $f(au_1 + bu_2, v) = af(u_1, v) + bf(u_2, v)$

ii) $f(u, av_1 + bv_2) = af(u, v_1) + bf(u, v_2)$

for all $a, b \in K$ and all $u_i, v_i \in V$. We express condition (i) by satisfying f is linear in the first variable, and condition (ii) by satisfying f is linear in the second variable.

Example: Let ϕ and σ are arbitrary linear functionals on V. Let $f: V\times V \to K$ be defined by $f(u, v) = \phi(u)\sigma(v)$. The f is bilinear because ϕ and σ are both linear. (such a bilinear form f turns out to be the tensor product of ϕ and σ and so is sometimes written as $f = \phi \otimes \sigma$).

Example 2: Let f be the dot product on R^n; i.e,

$f(u, v) = u.v = ab_1 + a_2b_2 + \cdots + a_nb_n$ where $u = (a_i)$, $v = (b_i)$.

Example 3: Let $A = (a_{ij})$ be any matrix of order n over a field K. Then A may be viewed as a bilinear form f on K^n by defining

$$f(X, Y) = X^t AY = (x_1, x_2, \cdots, x_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$= \sum_{i,j=1}^{n} a_{ij} x_i y_j = a_{11}x_1y_1 + a_{12}x_1y_2 + \cdots + a_{nn}x_ny_n$$

The above formal expression in variables x_i , y_i is termed the bilinear polynomial corresponding to the matrix A.

Remark: Let B(V) denote the set of bilinear forms on V. A vector structure is placed on B(V) by defining f + g and kf by:

$(f+g)(u,v)=f(u,v)+g(u,v)$

$(kf)=kf(u,v)$ for any $f,g\in B(V)$ and $k\in K$

Theorem: Let V be a vector space of dimension n over K. Let $\{\phi_1,\phi_2,\cdots,\phi_n\}$ be a basis of the dual space V^*. Then $\{f_{ij};\ i,\ j=1,2,\cdots,n\}$ is a basis of $B(V)$ where f_{ij} is defined by $f_{ij}(u,v)=\phi_i(u)\phi_j(v)$. Thus in particular, $\dim B(V)=n^2$.

Proof: Let $\{e_1,e_2,\cdots,e_n\}$ be a basis of V dual to $\{\phi_i\}$, we first show that $\{f_{ij}\}$ spans B(V). Let $f\in B(V)$ and suppose $f(e_i,e_j)=a_{ij}$, where $f=\sum a_{ij}f_{ij}$. It suffices to show that $f(e_s,e_t)=\sum a_{ij}f_{ij}(e_s,e_t)$ for $s,\ t=1,2,\cdots,n$. We have

$$\left(\sum a_{ij}f_{ij}\right)(e_s,e_t)=\sum a_{ij}f_{ij}(e_s,e_t)=\sum a_{ij}\phi_i(e_s)\phi_j(e_t)$$

$$=\sum a_{ij}\delta_{is}\delta_{jt}=a_{st}=f(e_s,e_t)$$

as required. Hence $\{f_{ij}\}$ spans B(V).

It remains to show that $\{f_{ij}\}$ is linearly independent. Suppose $\sum a_{ij}f_{ij}=0$ for $s,\ t=1,2,\cdots,n$.

$$0=0(e_s,e_t)=\left(\sum a_{ij}f_{ij}\right)(e_s,e_t)=a_{rs}$$

The last step follows as above. Thus $\{f_{ij}\}$ is independent and hence is a basis of B(V).

Exercise 1: Given $f(u,v)=3x_1y_1-2x_1y_2+5x_2y_1+7x_2y_2-8x_2y_3+4x_3y_2-x_3y_3$ where $u=(x_1,x_2,x_3)$ and $v=(y_1,y_2,y_3)$. Express f in matrix notation.

Exercise 2: Let $A=(a_{ij})$ be any matrix of order n over a field K. Show that the following map f is bilinear form on K^n ; $f(X,Y)=X^tAY$

Bilinear form and Matrices

Let f be a bilinear form on V, and let $\{e_1, e_2, \cdots, e_n\}$ be a basis of V. Suppose $u, v \in V$ such that $u = a_1e_1 + a_2e_2 + \cdots + a_ne_n$ and $v = b_1e_1 + b_2e_2 + \cdots + b_ne_n$

Then $f(u,v) = f(a_1e_1 + a_2e_2 + \cdots + a_ne_n, b_1e_1 + b_2e_2 + \cdots + b_ne_n) = \sum_{i,j}^{n} a_ib_j f(e_i, e_j)$

Thus f is completely determined by the n^2 values $f(e_i, e_j)$. Thus matrix $A = (a_{ij})$, where $a_{ij} = f(e_i, e_j)$ is called the matrix representation of f relative to basis $\{e_i\}$ or simply the matrix of f in $\{e_i\}$. It represents f in the sense that

$$f(u,v) = \sum_{i,j}^{n} a_ib_j f(e_i, e_j) = (a_1, a_2, \cdots, a_n)A\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = [u]_e^t A[v]_e$$

for all $u, v \in V$.

Definition: A matrix B is said to be convergent to a matrix A if there exists an invertible (or non-singular) matrix P such that $B = P^tAP$.

The rank of the bilinear form f on V is defined to be the rank of any matrix representation. We say that f is degenerate or non-degenerate according as to whether rank(f) < dimV or rank(f) = dimV.

Exercise: Let f be a bilinear form on R^2 defined by

$$f((x_1, x_2),(y_1, y_2)) = 2x_1y_1 - 3x_1y_2 + x_2y_2$$

i) Find the matrix A of f in the basis $\{u_1 = (1,0), u_2 = (1,1)\}$.

ii) Find the matrix B of f in the basis $\{v_1 = (2,1), v_2 = (1,-1)\}$

Solution: Set $A = (a_{ij})$ where $a_{ij} = f(u_i, u_j)$, we can easily calculate the values of all the entries of the matrix A.

Thus, $A = \begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix}$ is the matrix of f in the basis $\{u_i, u_j\}$.

(iii) Similarly, $B = \begin{pmatrix} 3 & 9 \\ 0 & 6 \end{pmatrix}$ is the matrix of f in the basis $\{v_i, v_j\}$

Definition: Let $\{e_1, e_2, \cdots, e_n\}$ be a basis of V and let $\{f_1, f_2, \cdots, f_n\}$ be another basis. Suppose

$$f_1 = a_{11}e_1 + a_{12}e_2 + \cdots + a_{1n}e_n$$

$$f_2 = a_{21}e_1 + a_{22}e_2 + \cdots + a_{2n}e_n$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$f_n = a_{n1}e_1 + a_{n2}e_2 + \cdots + a_{nn}e_n$$

The the transpose P of the above matrix of coefficients is termed the transition matrix from the old basis $\{e_i\}$ to the new basis $\{f_i\}$

$$P = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}$$

Since the vectors $\{f_i\}$ are linearly independent, the matrix P is invertible. In fact, its inverse P^{-1} is the transition matrix from the basis $\{f_i\}$ back to the basis $\{e_i\}$.

Example: Consider the following two basis of R^2

$$\{e_1 = (1,0),\ e_2 = (0,1)\} \text{ and } \{f_1 = (1,1),\ f_2 = (-1,0)\}$$

Then $f_1 = (1,1) = (1,0) + (0,1) = e_1 + e_2$

$$f_2 = (-1,0) = -(1,0) + 0(0,1) = -e_1 + 0.e_2$$

Hence the transition matrix P from the basis $\{e_i\}$ to the basis $\{f_i\}$ is

$$P = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$$

We also have

$$e_1 = (1,0) = 0(1,1) - (-1,0) = 0.f + (-1)f_2$$

$$e_2 = (0,1) = (1,1) + (-1,0) = f + f_2$$

Hence the transition matrix Q from the basis $\{f_i\}$ back to the basis $\{e_i\}$ is

$$Q = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$$

Observe that P and Q are invertible, i.e., PQ = I

Theorem: Let P be the transition matrix from a basis $\{e_i\}$ to a basis $\{f_i\}$ in a vector space V. Then for any $v \in V, P[v]_f = [v]_e$. Also $[v]_f = P^{-1}[v]_e$

Proof: Suppose for $i = 1,2,\cdots,,n$,

$$f_i = a_{i1}e_1 + a_{i2}e_2 + \cdots + a_{in}e_n = \sum_{j=1}^{n} a_{ij}e_j$$

Then P is the n-square matrix whose jth row is

$$(a_{1j} + a_{2j} + \cdots + a_{nj}) \qquad \text{….(1)}$$

Also suppose $v = k_1f_1 + k_2f_2 + \cdots + k_nf_n = \sum_{i=1}^{n} k_if_i$

Then writing a column vector as the transpose of a new vector,

$$[v]_f = (k_1, k_2 \cdots, k_n)^t \qquad \text{…(2)}$$

Similarly for f_i in the equation for v,

$$v = \sum_{i=1}^{n} k_if_i = \sum_{i=1}^{n} k_i\left(\sum_{j=1}^{n} a_{ij}e_j\right) = \sum_{j=1}^{n}\left(\sum_{j=1}^{n} a_{ij}k_i\right)e_j$$

$$= \sum_{j=1}^{n}(a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n)e_j$$

Accordingly, $[v]_e$ is the column vector whose jth entry is

$$a_{1j}k_1 + a_{2j}k_2 + \cdots + a_{nj}k_n \qquad \text{…(3)}$$

On the other hand, the jth entry of $P[v]_f$ is obtained by multiplying the jth row of P by $[v]_f$, i.e., (1) by (2). But the product of (1) and (2) is (3); hence $P[v]_f$ and $[v]_e$ have the same entries and thus $P[v]_f = [v]_e$

Furthermore, multiplying the above by P^{-1} gives $P^{-1}[v]_e = P^{-1}P[v]_f = [v]_f$.

Theorem: Let P be the transition matrix from one basis $\{e_i\}$ to a basis $\{e_i'\}$. If A is the matrix of f in the original basis $\{e_i\}$, then $B = P^t AP$ is the matrix of f in the new basis $\{e_i'\}$.

Proof: Let $u, v \in V$, since P is the transition matrix from $\{e_i\}$ to a basis $\{e_i'\}$, we have $P[u]_{e'} = [u]_e$ and $P[v]_{e'} = [v]_e$; hence $[u]_e^t = [u]_{e'}^t P^t$. Thus

$$f(u,v) = [u]_e^t A[v]_e = [u]_{e'}^t P^t AP[v]_{e'}$$

Since u and v are arbitrary elements of v. $P^t AP$ is the matrix of f in the basis $\{e_i'\}$.

Alternating Bilinear Forms

A bilinear form f on V is said to be alternating if

i) f(u,v) = 0 for any $v \in V$. If f is alternating, then

$0 = f(u+v, u+v) = f(u,u) + f(u,v) + f(v,u) + f(v,v)$ and so

ii) f(u,v) = - f(v,u) for every $u, v \in V$.

A bilinear form which satisfies (ii) is said to be skew-symmetric (or anti-symmetric). If $1+1 \neq 0$ in K, then condition (ii) implies $f(v,v) = -f(v,v)$ which implies condition (i). In other words, alternating and skew symmetric are equivalent when $1+1 \neq 0$.

Theorem: Let f be an alternating bilinear form on V. Then there exists a basis of V in which f is represented by a matrix of the form

$$\begin{pmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} & & & & & & & \\ & \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} & & & & & & \\ & & \ddots & & & & & \\ & & & \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} & & & & \\ & & & & [0] & & & \\ & & & & & [0] & & \\ & & & & & & \ddots & \\ & & & & & & & [0] \end{pmatrix}$$

Moreover, the number of $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is uniquely determined by f (because it is equal to $\frac{1}{2}[rank(f)]$.

Proof: If f = 0, then the theorem is obviously true, Also if dimV = 1, then

$$f(k_1u, k_2u) = k_1k_2 f(u,u) = 0 \text{ and so } f = 0.$$

Accordingly, we can assume that dimV >1 and $f \neq 0$

Since $f \neq 0$, there exists non-zero $u_2 = ku_1\, u_1, u_2 \in V$ such that $f(u_1, u_2) \neq 0$. In fact multiplying u_1 by an appropriate factor, we can assume $f(u_1, u_2) = 1$ and so $f(u_1, u_2) = -1$. Now u_1 and u_2 are linearly independent; because if say $u_2 = ku_1$, then

$f(u_1, u_2) = f(u_1, ku_1) = kf(u_1, u_1) = 0$. Let U be the subspace spanned by u_1 and u_2, i.e., $U = L(u_1, u_2)$.

Note: (i) the matrix representation of the restriction of f to U in the basis $\{u_1, u_2\}$ is $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

(ii) If $u \in U$ say $u = au_1 + bu_2$, then

$$f(u_1, u_2) = f(au_1 + bu_2, u_1) = -b$$

$$f(u_1,u_2)=f(au_1+bu_2,u_1)=a$$

Let W consists of those vectors $w\in W$ such that $f(w,u_1)=0$ and $f(w,u_2)=0$. Equivalently

$$W=\{w\in V:\ f(w,u)=0\ \ for\ \ every\ u\in U\}$$

We claim that $V=U\oplus W$ It is clear that $U\cap W=\{0\}$, and so it remains to show that $V=U+W$. Let $v\in V$.

Set $$u=f(v,u_2)u_1-f(v,u_1)u_2 \text{ and } w=v-u \qquad \ldots(1)$$

Since u is a linear combination of u_1 and u_2, $u\in U$. We show that $w\in W$. By (1) and (ii), $f(u,u_1)=f(v,v_1)$;

Hence $$f(w,u_1)=f(v-u,u_1)=f(v,u_1)-f(u,u_1)=0$$

Similarly, $f(u,u_2)=f(v,u_2)$ and so

$$f(w,u_2)=f(v-u,u_2)=f(v,u_2)-f(u,u_2)=0$$

Then $w\in W$ and so by (1), $v=u+w$, where $u\in U$ and $w\in W$.

The shows that $V=U+W$ and therefore $V=U\oplus W$.

Now the restriction of f to W is an alternating bilinear form on W. By induction, there exists a basis $\{u_3,u_4,\cdots,u_n\}$ of W in which the matrix representing f restricted to W has the desired form. Thus $\{u_1,u_2,\cdots,u_n\}$ is a basis of V in which matrix representing f has the desired form.

4.14: Symmetric bilinear form

A bilinear form f on V is said to be symmetric if

$$f(u,v)=f(v,u)$$

for every $u,v\in V$. If A is a matrix representation of f, we can write

$$f(X,Y)=X^tAY=(X^tAY)^t=Y^tA^tX$$

(we use the fact that $X^t AY$ is a scalar and therefore equals its transpose). Thus if f is symmetric,

$$Y^t A^t X = f(X,Y) = f(Y,X) = Y^t AX$$

and since this is true for all vectors X and Y, it follows that $A = A^t$ or A is symmetric. Conversely if A is symmetric, then f is symmetric.

Home Assignments

Exercise: Find the symmetric matrix which corresponds to each of the following quadratic polynomials

a) $q(x,y) = 4x^2 - 6xy - 7y^2$

b) $q(x,y) = xy + y^2$

c) $q(x,y,z) = 3x^2 + 4xy - y^2 + 8xz - 6yz + z^2$

d) $q(x,y) = x^2 - 2yz + xz$

Theorem: Let f be a symmetric bilinear form on V over K (in which $1+1 \neq 0$). Then V has a basis $\{v_1, v_2, \cdots, v_n\}$ in which f represented by a diagonal matrix, i.e., $f(v_i, v_j) = 0$ for $i \neq j$.

Proof: If f = 0 or if dimV = 1, then theorem clearly holds. Hence we can suppose $f \neq 0$ and dimV = n >1. If $q(v) = f(v,v) = 0$ for every $v \in V$, then the polar form of f: $f(u,v) = \frac{1}{2}(q(u+v) - q(u) + q(v))$ implies that f = 0. Hence we can assume there is a vector $v_1 \in V$ such that $f(v_1, v_1) \neq 0$. Let U be the subspace spanned by v_1 and let W consist of those vectors $v \in V$ for which $f(v_1, v) = 0$,

We claim that $V = U \oplus W$

(i) suppose $u \in U \cap W$, since $u \in U$, $u = kv_1$ for some scalar $k \in K$. Since $u \in W$, $0 = f(u,u) = f(kv_1, kv_1) = k^2 f(v_1, v_1)$ But $f(v_1, v_1) \neq 0$, hence $k = 0$ and therefore $u = kv_1$, thus $U \cap W = \{0\}$

(ii) To prove $V = U + W$. Let $v \in V$, set

$$w = v - \frac{f(v_1, v)}{f(v_1, v)} v_1 \qquad \ldots(1)$$

Then $f(v_1, w) = f(v_1, v) - \dfrac{f(v_1, v)}{f(v_1, v)} f(v_1, v_1) = 0$

Thus $w \in W$. By (1), v is the sum of an element of U and an element of W. Thus $V = U + W$. Therefore (i) and (ii) implies $V = U \oplus W$.

Now f restricted to W is a symmetric bilinear form on W. But dimW= n-1; hence by induction there is a basis $\{v_2, v_3, \cdots, v_n\}$ of W such that $f(v_i, v_j) = 0$ for $i \neq j$ and $2 \leq i,\ j \leq n$. But by the very definition of W, $f(v_1, v_j) = 0$ for $j = 2,3,\cdots,n$. Therefore the basis $\{v_1, v_2, \cdots, v_n\}$ of V has a required property that $f(v_i, v_j) = 0$ for $i \neq j$.

Example: Let $A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -4 \\ -3 & -4 & 8 \end{pmatrix}$, a symmetric matrix. It is convenient to form the block matrix (A, I):

We have

$$(A, I) = \left(\begin{array}{ccc|ccc} 1 & 2 & -3 & 1 & 0 & 0 \\ 2 & 5 & -4 & 0 & 1 & 0 \\ -3 & -4 & 8 & 0 & 0 & 1 \end{array}\right)$$

Applying operations $R_2 \to -2R_1 + R_2$ and $R_3 \to 3R_1 + R_3$ to (A, I), and then the corresponding operations $C_2 \to -2C_1 + C_2$ and $C_3 \to 3C_1 + C_3$ to A, to obtain

$$\begin{pmatrix} 1 & 2 & -3 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 2 & \vdots & -2 & 1 & 0 \\ 0 & 2 & -1 & \vdots & 3 & 0 & 1 \end{pmatrix} \text{ and then } \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 2 & \vdots & -2 & 1 & 0 \\ 0 & 2 & -1 & \vdots & 3 & 0 & 1 \end{pmatrix}$$

We next apply the operations $R_3 \to -2R_2 + R_3$ and then the corresponding operation $C_3 \to -2C_2 + C_3$ to obtain,

$$\begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 2 & \vdots & -2 & 1 & 0 \\ 0 & 0 & -5 & \vdots & 7 & -2 & 1 \end{pmatrix} \text{ and then } \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 1 & 0 & \vdots & -2 & 1 & 0 \\ 0 & 0 & -5 & \vdots & 7 & -2 & 1 \end{pmatrix}$$

Now A has been diagonalized. We set

$$P = \begin{pmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \text{ and then } P^t A P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{pmatrix}$$

Definition: A mapping $q : V \to K$ is called quadratic form if $q(v) = f(v, v)$ for some symmetric bilinear form on V.

We call q the quadratic form associated with the symmetric bilinear form f if $1+1 \neq 0$ in K, then f is obtainable from q according to the identity

$$f(u,v) = \frac{1}{2}\left(q(u+v) - q(u) - q(v)\right) \quad \text{(polar form of f)}$$

Remark: If f is represented by a symmetric matrix $A = (a_{ij})$, then q is represented in the form

$$q(X) = f(X, X) = X^T A X = (x_1, x_2, \cdots, x_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \sum_{i,j} a_{ij} x_i x_j = a_{11} x_1^{\,2} + a_{22} x_2^{\,2} + \cdots + a_{nn} x_n^{\,2} + 2\sum_{i<j} a_{ij} x_i x_j$$

The above formal expression in variable x_i is termed the quadratic polynomial corresponding to the symmetric matrix A. Observe that if the matrix A is diagonalizable, then q has the diagonal representation

$$q(X) = X^t AX = a_{11}x_1^{\ 2} + a_{22}x_2^{\ 2} + \cdots + a_{nn}x_n^{\ 2}$$

Example: Consider the following quadratic form on R^2

$$q(x, y) = 2x^2 - 12xy + 5y^2$$

Home Assignment

Exercise: Let q be the quadratic form associated with the symmetric bilinear form f. Verify

$$f(u,v) = \frac{1}{2}\left(q(u+v) - q(u) - q(v)\right) \quad \text{(Assume } 1+1 \neq 0\text{)}$$

4.15: Real Symmetric bilinear form

In this section, we treat symmetric bilinear forms and quadratic forms on vector spaces over the real field R.

Theorem: Let f be a symmetric bilinear form on V over R. Then there is a basis of V in which f is represented by a diagonal matrix, and every other diagonal representation of f has the same number of positive entries and the same number of negative entries.

Proof: We have by previous theorem, that there exists a basis $\{u_1, u_2, \cdots, u_n\}$ of V in which f is represented by a diagonal matrix, say with P positive and N negative entries. Now suppose $\{w_1, w_2, \cdots, w_n\}$is another basis of V in which f is represented by a diagonal matrix, say with P' positive and N' negative entries. We can assume without loss of generality that the positive entries in each matrix appear first. Since $rank(f) = P + N = P' + N'$, it suffices to prove that $P = P'$.

Let U be a linear span of $u_1, u_2, \cdots, u_P$ and let W be a linear span of $w_{P'+1}, w_{P'+2}, \cdots, w_n$. Then $f(v,v) > 0$ for every non-zero $v \in U$ and $f(v,v) \le 0$ for every non-zero $v \in W$. Hence $U \cap W = \{0\}$. Note that $\dim U = P$, $\dim W = n - P'$. Thus

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W) = P + (n - P') - 0 = P - P' + n$$

But $\dim(U+W) \le \dim V = n$. Hence $P - P' + n \le n$ or $P \le P'$.

Similarly $P \ge P'$ and therefore $P = P'$ as required

Remark: The above theorem and proof depend only on the concept of positivity, thus the theorem is true for any subfield K of the real field R.

Definition: A real symmetric bilinear form f is said to be non-negative semi-definite if $q(v) = f(v,v) \ge 0$ for every vector v and is said to be positive definite if $q(v) = f(v,v) > 0$ for every vector $v \ne 0$.

Remark: By the above theorem, the difference S = P – N is called the signature of f. Also

i) f is non-negative semi-definite if and only if $S = rank(f)$

ii) f is positive definite if and only $S = \dim(V)$

Example: Let f be the dot product on R^n, i.e.,

$$f(u,v) = u.v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

Where $u = (u_i)$ and $v = (v_i)$.

Note that f is symmetric, since

$$f(u,v) = u.v = v.u = f(v,u)$$

Furthermore, f is positive semi-definite because

$$f(u,u) = u.u = a_1^2 + a_2^2 + \cdots + a_n^2 > 0; \text{ where } u \ne 0$$

Corollary: Any real quadratic form q has a unique representation in the form

$$q(x_1, x_2, \cdots, x_n) = x_1^{\ 2} + x_2^{\ 2} + \cdots + x_s^{\ 2} - x_{S+1}^{\ 2} + x_{S+2}^{\ 2} + \cdots + x_r^{\ 2}$$

The above result for real quadratic forms is some times referred to as the Law of Inertia or Sylvester theorem.

Home Assignment

Exercise: For each of the following real symmetric matrices A, find a non-singular matrix P such that $P^t AP$ is diagonal and find its signature.

i) $A = \begin{pmatrix} 1 & -3 & 2 \\ -3 & 7 & -5 \\ 2 & -5 & 8 \end{pmatrix}$

ii) $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -2 & 2 \\ 1 & 2 & -1 \end{pmatrix}$

Exercise: Let $A = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}$, a diagonal matrix over K, Show that

i) For any non-zero scalar $k_1, k_2, \cdots, k_n \in K$, A is congruent to a diagonal matrix with diagonal entries $a_i k_i^{\ 2}$;

ii) If k is a complex field C, then A is congruent to a diagonal matrix with only 1's and 0's as diagonal entries;

iii) If k is the real field R, then A is congruent to a diagonal matrix with only 1's, -1's and 0's as diagonal entries.

Solution:

i) Let P be the diagonal matrix with diagonal entries k_i, then

$$P^t AP = \begin{pmatrix} k_1 & & & \\ & k_2 & & \\ & & \ddots & \\ & & & k_n \end{pmatrix} \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix} \begin{pmatrix} k_1 & & & \\ & k_2 & & \\ & & \ddots & \\ & & & k_n \end{pmatrix} = \begin{pmatrix} a_1 k_1^{\ 2} & & & \\ & a_2 k_2^{\ 2} & & \\ & & \ddots & \\ & & & a_n k_n^{\ 2} \end{pmatrix}$$

ii) Let P be the diagonal matrix with diagonal entries

$$b_i = \begin{cases} 1/\sqrt{a_i} & if \quad a_i \neq 0 \\ 1 & if \quad a_i = 0 \end{cases} \quad ;$$

Then $P^t AP$ has the required form.

iii) Let P be the diagonal matrix with diagonal entries

$$b_i = \begin{cases} 1/\sqrt{|a_i|} & if \quad a_i \neq 0 \\ 1 & if \quad a_i = 0 \end{cases} \quad ;$$

Then $P^t AP$ has the required form.

4.16: Hermitian Forms

Let V be a vector space of finite dimension over the complex field C. Let $f : V \times V \to C$ be such that

i) $f(au_1 + bu_2, v) = af(u_1, v) + bf(u_2, v)$

ii) $f(u,v) = \overline{f(v,u)}$

where $a, b \in C$ and $u_i, v \in V$. The f is called Hermitian form on V.

By (i) and (ii), we have

$$f(u, av_1 + bv_2) = \overline{f(av_1 + bv_2, u)} = \overline{af(v_1, u) + bf(v_2, u)}$$

$$= \bar{a}\ \overline{f(v_1, u)} + \bar{b}\ \overline{f(v_2, u)} = \bar{a}\ f\overline{(u, v_1)} + \bar{b}\ f\overline{(u, v_2)}$$

i.e., (iii) $f(u, av_1 + bv_2) = \bar{a}\ f\overline{(u, v_1)} + \bar{b}\ f\overline{(u, v_2)}$

Note that $f(u,v) = \overline{f(v,u)}$ and so $f(v,v)$ is real for any $v \in V$.

Exercise: Let A be a Hermitian matrix, Show that f is a Hermitian form on C^n; where f is defined by $f(X,Y) = X^t A\overline{Y}$.

Solution: For all $a, b \in C$ and $X_1, X_2, Y \in C^n$

$$f(aX_1 + bX_2, Y) = (aX_1 + bX_2)^t A\overline{Y} = (aX_1{}^t + bX_2{}^t)A\overline{Y}$$

Hence f is linear in the first variable. Also

$$\overline{f(X,Y)} = \overline{X^t A\overline{Y}} = \overline{(X^t A\overline{Y})^t} = \overline{\overline{Y}^t A^t Y} = Y^t A^* \overline{X} = Y^t A\overline{X} = f(Y,X)$$

Hence f is Hermitian form on C^n.

Exercise: Let f be a Hermitian form on V and H is the matrix of f in a basis $\{e_1, e_2, \cdots, e_n\}$of V. Show that

i) $f(u,v) = [u]_e^t H\,\overline{[v]}_e$ for all $u, v \in V$

ii) If P is the transition matrix from $\{e_i\}$ to a new basis $\{e_i'\}$ of V, then $B = P^t H\overline{P}$ (or $B = Q^* HQ$, where $Q = \overline{P}$) is the matrix of f in the new basis $\{e_i'\}$.

Proof: i) Let $u, v \in V$ and suppose

$u = a_1e_1 + a_2e_2 + \cdots + a_ne_n$ and $v = b_1e_1 + b_2e_2 + \cdots + b_ne_n$

Then $f(u,v) = f(a_1e_1 + a_2e_2 + \cdots + a_ne_n,\, b_1e_1 + b_2e_2 + \cdots + b_ne_n)$

$$\sum_{i,j} a_i\overline{b}_j f(e_i, e_j) = (a_1, a_2, \cdots, a_n)H\begin{pmatrix}\overline{b}_1\\ \overline{b}_2\\ \vdots\\ \overline{b}_n\end{pmatrix} = [u]_e^t H[\overline{v}]_e$$

iii) Since P is a transition matrix from $\{e_i\}$ to $\{e_i'\}$,

then $P[u]_{e'} = [u]_e$, $P[v]_{e'} = [v]_e$ and so,

$[u]_e^t = P[u]_{e'}^t P^t$, $\overline{[v]} = \overline{P}\,\overline{[v]}_{e'}$

Thus by (i), $f(u,v) = [u]_e^t H\,\overline{[v]}_e = [u]_e^t P^t H\overline{P}\,\overline{[v]}_{e'}$

But u and v are arbitrary elements of V, hence $P^t H\overline{P}$ is the matrix of f in the basis $\{e_i'\}$

Remark: The mapping $q: V \to R$ defined by $q(v) = f(v, v)$ is called Hermitian quadratic form or complex quadratic form associated with the Hermitian form f. We can obtain f from q according to the following identity called the polar form of f.

$$f(u,v) = \frac{1}{4}\{q(u+v) - q(u-v)\} + \frac{1}{4}\{q(u+iv) - q(u-iv)\}$$

Now suppose $\{e_1, e_2, \cdots, e_n\}$ is a basis of V, the matrix $H = (h_{ij})$, where $h_{ij} = f(e_i, e_j)$ is called the matrix representation of f in the basis $\{e_i\}$. Since $f(e_i, e_j) = f\overline{(e_j, e_i)}$; hence H is Hermitian and, in particular the diagonal entries of H are real.

Example: Let f be the dot product on C^n, that is

$$f(u,v) = u.v = z_1\overline{w}_1 + z_2\overline{w}_2 + \cdots + z_n\overline{w}_n \quad \text{where} \quad u = (z_i) \quad \text{and} \quad v = (w_i).$$

Then f is a Hermitian form on C^n. Moreover, f is positive definite since, for any $v \neq 0$

$$f(u,v) = z_1\overline{z}_1 + z_2\overline{z}_2 + \cdots + z_n\overline{z}_n = |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2 > 0$$

Exercise: Let H be a Hermitian matrix given below. Find the non-singular matrix P such that $P^t H\overline{P}$ is diagonal.

$$H = \begin{pmatrix} 1 & 1+i & 2i \\ 1-i & 4 & 2-3i \\ -2i & 2+3i & 7 \end{pmatrix}$$

Solution: First form the block matrix (H, I):

We have

$$(H,I)=\begin{pmatrix} 1 & 1+i & 2i & \vdots & 1 & 0 & 0 \\ 1-i & 4 & 2-3i & \vdots & 0 & 1 & 0 \\ -2i & 2+3i & 7 & \vdots & 0 & 0 & 1 \end{pmatrix}$$

Applying row operations $R_2 \to (-1+i)R_1 + R_2$ and $R_3 \to 2iR_1 + R_3$ to (H,I), and then the corresponding Hermitian column operations $C_2 \to (-1-i)C_1 + C_2$ and $C_3 \to -2iC_1 + C_3$ to H, to obtain

$$\begin{pmatrix} 1 & 1+i & 2i & \vdots & 1 & 0 & 0 \\ 0 & 2 & -5i & \vdots & -1+i & 1 & 0 \\ 0 & 5i & 3 & \vdots & 2i & 0 & 1 \end{pmatrix} \text{ and then } \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 2 & -5i & \vdots & -1+i & 1 & 0 \\ 0 & 5i & 3 & \vdots & 2i & 0 & 1 \end{pmatrix}$$

We next apply the operations $R_3 \to -5iR_2 + 2R_3$ and then the corresponding operation $C_3 \to 5iC_2 + 2C_3$ to obtain,

$$\begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 2 & -5i & \vdots & -1+i & 1 & 0 \\ 0 & 0 & -19 & \vdots & 5+9i & -5i & 2 \end{pmatrix} \text{ and then } \begin{pmatrix} 1 & 0 & 0 & \vdots & 1 & 0 & 0 \\ 0 & 2 & 0 & \vdots & -1+i & 1 & 0 \\ 0 & 0 & -38 & \vdots & 5+9i & -5i & 2 \end{pmatrix}$$

Now H has been diagonalized, We set

$$P=\begin{pmatrix} 1 & -1+i & 5+9i \\ 0 & 1 & -5i \\ 0 & 0 & 2 \end{pmatrix} \text{ and then } P^t A\overline{P}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -38 \end{pmatrix}$$

Observe that signature S of H is

S = 2-1=1

Theorem: Let f be a Hermitian form on V. Then there exists $\{e_1, e_2, \cdots, e_n\}$ of V in which f is represented by a diagonal matrix, i.e., $f(e_i, e_j) = 0$ for $i \neq j$. Moreover, every diagonal representation of f has the same number P of positive entries, and the same number N of negative entries. The difference S = P-N is called the signature of f. Analogously f is non-negative semi-definite if $q(v) = f(v,v) \geq 0$ for every $v \in V$ and if $q(v) = f(v,v) > 0$ for every $v \neq 0$, then f is positive definite.

Proof: Let f be a Hermitian form on V. Then there exists a basis $\{e_1, e_2, \cdots, e_n\}$ of V in which f is represented by a diagonal matrix, i.e., $f(e_i, e_j) = 0$ for $i \neq j$. Moreover, every diagonal representation of f has the same number P of positive entries, and the same number N of negative entries.

Note that the second part of the theorem does not hold for complex symmetric bilinear forms (as seen by part (ii) in exercise before Hermitian forms). However the proof of the previous theorem (Real symmetric bilinear form) does carry over the Hermitian cases.

Exercise: Show that any bilinear form f on V is the sum of a symmetric bilinear form and a skew-symmetric bilinear form.

Solution: Set $g(u,v) = \frac{1}{2}\{f(u,v) + f(v,u)\}$ and $h(u,v) = \frac{1}{2}\{f(u,v) - f(v,u)\}$

Clearly,

$$g(u,v) = \frac{1}{2}\{f(u,v) + f(v,u)\} = \frac{1}{2}\{f(v,u) + f(u,v)\} = g(v,u)$$

$\Rightarrow$ g is symmetric, and

$$h(u,v) = \frac{1}{2}\{f(u,v) - f(v,u)\} = -\frac{1}{2}\{f(v,u) - f(u,v)\} = -h(v,u)$$

$\Rightarrow$ h is skew-symmetric.

Furthermore

$$f = g + h$$