1

WEEK 1 AND 2 LECTURES

HUSNEY PARVEZ SARWAR

Abstract. Here we will briefly recall what we have taught in MATH-II.

Contents

1. Lecture - 1–5

MATHEMATICS – II

consists of the following four topics:

(1) Linear Algebra
(2) Numerical Analysis
(3) Integral Calculus
(4) Vector Calculus

We will begin with

(1) Linear Algebra

is about Matrices and Vector spaces.

Syllabus: **Matrix addition, multiplications, elementary row operations, row echelon form, Gauss elimination method to solve system of linear equations (homogenous and non-homogenous), rank of a matrix and its properties, solution of system of equations using rank concept, vector spaces, subspace, linearly indpendent vectors, linearly dependent vectors, basis, dimension, linear map, matrix representations of a linear map with respect to the given ordered basis, rank-nullity theorem, eigen values, eigen vectors, Cayley–Hamilton theorem, diagonalisation, Eigen values and eigen vectors of Hermitian (symmetric), skew-Hermitian (skew-symmetric), unitary (orthogonal) matrices.**

Here “scalars” means either real numbers or complex numbers.

Recall: Matrix addition and matrix multiplication. Multiplication of a matrix by a scalar.

Let A be an $m \times n$ matrix, x a column vector with n unknowns, b an another column vector with m entries.

$$Ax = b$$

yields m equations in n unknowns.

Problem: We want to solve this system of equations for x.

[1]If you find any errors or typos, please let me know AT parvez@maths.iitkgp.ac.in

Recall Gauss elimination method to solve it.

Elementary row operations:

1. Interchange of rows
2. Scalar multiple of one row adding to the another row
3. Multiply any row by a non-zero scalar.

In the language of system of equations, it means:

1. Interchange of equations (solutions do not change!)
2. Scalar multiple of one equation adding to the another equation (solutions do not change!)
3. Multiplying any equation by a scalar. (solutions do not change!)

Caution! These operations are for rows NOT for columns.

Definition 1.1. We call a linear system S_1 row equivalent to a linear system S_2 if S_1 can be obtained from S_2 by finitely many row operations.

Theorem 1.2. *Row equivalent linear systems have the same set of solutions.*

Given a matrix $A = [a_{ij}]$, how to make row echelon form.

Step - 1: By row interchange operations, keep all the zero rows in the below. Let us say that there are n-rows. We kept 1st m-non zero rows in the begining and $n - m$ zero rows at the bottom.

Step - 2: Look for 1st non-zero entry in the first row. Let us say that it is found in c_1-th column. Using that non-zero entry in c_1-column make all other entries below to that entry zero.

Step - 3: Look for 1st non-zero entry in the second row. Let us say that it is found in c_2-th column. If $c_2 > c_1$, then using that non-zero entry in c_2-column make all other entries below to that entry zero.

If $c_1 > c_2$, then interchange the rows. Now using that non-zero entry in c_2-column make all other entries below to that entry zero.

Step - 4: Go to the next rows and comtinue the same process.

Final form: If the non-zero entry of i-th row occurs in c_i-th column, then we have $c_1 < c_2 < \cdots < c_m$ and $a_{jc_i} = 0$ for $j < i$.

Definition 1.3. (Rank of a matrix) The rank of a matrix A is the maximum number of linearly independent vectors row vectors of A. It is denoted by $\operatorname{rank}(A)$.

Another equivalent definition: The highest order of the square sub-matrix of the given matrix whose minor is non-zero.

Remark 1.4. $\operatorname{rank}(A) = 0$ if and only if $A = [0]$- zero matrix.

We call a matrix A_1 row-equivalent to a matrix A_2 if A_1 can be obtained from A_2 by some elementary row operations.

Theorem 1.5. *Row-equivalent matrices have the same rank.*

Method to find the rank of a matrix A:

Theorem 1.5 tells us that we can determine the rank of a matrix by reducing the matrix to row-echelon form.

When the matrix is in the row-echelon form, we count the number of non-zero rows which is precisely the rank of the matrix.

Example 1.6. Let $A = \begin{bmatrix} 1 & 0 & 3 \\ 4 & -1 & 5 \\ 2 & 0 & 6 \end{bmatrix}$. Calculate the rank of A. Answer: $\operatorname{rank}(A) = 2$.

Example 1.7. Let $A = \begin{bmatrix} \sin\theta & \sin\theta & \sin\theta \\ \sin\theta & \sin\theta & \sin\theta \\ \sin\theta & \sin\theta & \sin\theta \end{bmatrix}$, where θ is a real number. Calculate the rank of A.

Real Vector space: A non-empty set V of elements a, b ... is called a real vector space, and these elements are called vectors if, in V, there are defined two algebraic operations (called vector addition and scalar multiplication) as follows:

I. Vector addition associates with every pair of vectors a and b of V a unique vector of V, called the sum of a and b and denoted by $a+b$, such that the following axioms are satisfied:

I.1: For every vectors a and b of V, we have $a+b=b+a$ (commutativity).

I.2 For any three vectors a, b and c of V, we have $(a+b)+c = a+(b+c)$ (written $a+b+c$) (Associativity).

I.3 There is a unique vector in V, called the zero vector and denoted by 0, such that for every a in V, $a+0=0+a=a$.

I.4 For every a in V, there is a unique vector in V that is denoted by $-a$ and is such that $a+(-a)=0$. (Inverse).

II. Scalar multiplication. The real numbers are called scalars. Scalar multiplication associates with every a in V and every scalar r a unique vector of V, called the product of c and r denoted by ra or (ar) such that the following axioms are satisfied:

II.1 For every scalar r, and vectors a, b in V, $c(a+b) = ra+rb$ holds.

II.2 For scalars r and s, and every a in V, $(r+s)a = rs+sa$ holds.

II.1 and **II.2** are called distributive property.

II.3 (Associativity). For all scalars r and s, and for every a in V, $r(sa)=(rs)a$ holds.

II.4 For every a in V, $1a=a$.

Examples 1.8.

1. Check that $(\mathbb{R}^n, +)$ is a vector space over $\mathbb{R}$, where $+$ is coordinate wise addition i.e. if $a=(a_1,\ldots,a_n)$ and $b=(b_1,\ldots,b_n)$ in $\mathbb{R}^n$, then $a+b = (a_1+b_1,\ldots,a_n+b_n)$. The scalar multiplication is $c(a_1,\ldots,a_n) := (ca_1,\ldots,ca_n)$, $(a_1,\ldots,a_n)\in\mathbb{R}^n$, $c\in\mathbb{R}$ is scalar.

   Always keep the picture for $n=1,2,3$.
2. Let $M_{m\times n}(\mathbb{R})$ be the set of all $m\times n$ matrices over $\mathbb{R}$. Check that $M_{m\times n}(\mathbb{R})$ forms a vector space under usual matrix addition and scalar multiplication.

   Think about 2×2 and 3×3 matrices!
3. Let $P[x]_{\leq n} := \{a_0+a_1x+\cdots+a_nx^n : a_i\in\mathbb{R}, 1\leq i\leq n\}$ (set of all polynomials of degree $\leq n$). Let $p(x)=a_0+a_1x+\cdots+a_nx^n$, $q(x)=b_0+b_1x+\cdots+b_nx^n \in P[x]_{\leq n}$. Define

$$p(x)+q(x) := (a_0+b_0)+(a_1+b_1)x+\cdots+(a_n+b_n)x^n$$
$$cp(x) := ca_0+ca_1x+\cdots+ca_nx^n.$$

   Check that $P[x]_{\leq n}$ is a vector space over $\mathbb{R}$ with the above operations.

Definition: Complex Vector space. If, in the definition of real vector space, we take complex as scalars instead of real numbers, we obtain the axiomatic definition of complex vector space.

Examples 1.9. In above examples, if we change scalars from real number ($\mathbb{R}$) to complex numbers ($\mathbb{C}$), then all become complex vector space.

Fundamental Idea of linear algebra is to take "linear combinations"!

"linear combinations" means linear combination of vectors with scalar coefficients!

Question: What is the set $L\{(1,2)\}$ (in language linear span of the vector $(1,2)$) in $\mathbb{R}^2$? Find ten 10 different elemnts from $L\{(1,2)\}$. What about the set $L\{(1,2,0)\}$ in $\mathbb{R}^3$? What is $L\{(1,2,0),(1,-1,0)\}$?

Question: Let S be a finite subset of V. Is $L(S)$ a subspace of V?
Answer: Yes. Check it! (How to check?)

2. Lectures:6–8

Rank-Nullity Theorem.
Let V and W be two real (or complex) vectors spaces. Let $T : V \to W$ be a linear map.

Let $\mathrm{Ker}(T)$ denotes the kernel of the map T, i.e. $\{v \in V : T(v) = 0\}$ and $\mathrm{Im}\,(T)$ denotes the image of the linear map T, i.e. $\{w \in W : \exists v \in V$ such that $T(v) = w\}$.

Proposition 2.1. (1) $\mathrm{Ker}(T)$ *is a subspace of* V *and* (2) $\mathrm{Im}\,(T)$ *is a subspace of* W*.*

Proof. (1) Let $u, v \in \mathrm{Ker}(T)$. This means $T(u) = 0$ and $T(v) = 0$. Let c be any scalar. We have to show: $u + v \in \mathrm{Ker}(T)$ and $cu \in \mathrm{Ker}(T)$.

For this, $T(u + v) =^1 T(u) + T(v) = 0 + 0 = 0$. This implies that $u + v \in \mathrm{Ker}(T)$. Again $T(cu) =^2 cT(u) = c0 = 0$. Hence $cu \in \mathrm{Ker}(T)$. (Note that $=^1$ and $=^2$ hold because T is linear.) Therefore $\mathrm{Ker}(T)$ is a subspace of V.

(2) Let $w_1, w_2 \in \mathrm{Im}\,(T)$. This means $\exists v_1, v_2$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$. Let c be ascalar.

We have to show: $w_1 + w_2 \in \mathrm{Im}\,(T)$ and $cw_1 \in \mathrm{Im}\,(T)$. This means that we have to find v and v' in V such that $T(v) = w_1 + w_2$ and $T(v') = cw_1$.

Now observe that we can take $v = v_1 + v_2$ and $v' = cv_1$. Then $T(v_1 + v_2) = T(v_1) + T(v_2) = w_1 + w_2$ and $T(cv_1) = cT(v_1) = cw_1$. Therefore $\mathrm{Im}\,(T)$ is a subspace of W. □

Definition 2.2. $\mathrm{Nullity}\,(T) = \dim(\mathrm{Ker}(T))$ and $\mathrm{rank}(T) := \dim(\mathrm{Im}(T))$.

Theorem 2.3. *(Rank–Nullity Theorem) Let* V *and* W *be two vectors spaces. Assume* V *is finite dimensional. Let* $T : V \to W$ *be a linear map. Then* $\mathrm{Nullity}\,(T) + \mathrm{rank}(T) = \dim(V)$*.*

Corollary 2.4. *(Rank–Nullity Theorem– Real case.) Let* A *be an* $m \times n$ *real matrix. Then* $A : \mathbb{R}^n \to \mathbb{R}^m$ *is a linear map. Then* $\mathrm{Nullity}\,(A) + \mathrm{rank}(A) = n$*.*

Corollary 2.5. *(Rank–Nullity Theorem– Complex case.) Let* A *be an* $m \times n$ *complex matrix. Then* $A : \mathbb{C}^n \to \mathbb{C}^m$ *is a linear map between the complex vector spaces. Then* $\mathrm{Nullity}\,(A) + \mathrm{rank}(A) = n$*.*

Remark 2.6. Note that $\mathrm{Ker}(A)$ is a solution space (observe the promotion from "Solution Set" to "Solution Space"). Hence Nullity of A is the dimension of the solution space.

Example 2.7. **Verification of the Rank–Nullity Theorem map via an example.**
Define a linear map $T : \mathbb{R}^2 \to \mathbb{R}^2$ as $T(x, y) = (x, 0)$. To verify the rank–nullity theorem, we have to calculate $dim(\mathrm{Ker}(T))$ and $dim(\mathrm{Im}\,(T))$. Then we need to show that

$$\dim(\mathrm{Ker}(T)) + \dim(\mathrm{Im}\,(T) = 2.$$

For this, $\mathrm{Ker}(T) = \{(x, y) \in \mathbb{R}^2 : T(x, y) = (0, 0)\} = \{(x, y) \in \mathbb{R}^2 : (x, 0) = (0, 0)\} = \{(x, y) \in \mathbb{R}^2 : x = 0\} = \{(0, y) \in \mathbb{R}^2 : y \in \mathbb{R}\}$. Claim:

$$\mathrm{Ker}(T) = L(\{(0, 1)\}).$$

For this, note $(0,1) \in \operatorname{Ker}(T)$ as $T(0,1) = (0,0)$. Hence $L(\{(0,1)\}) \subset \operatorname{Ker}(T)$ since $\operatorname{Ker}(T)$ is a subspace (i.e. a vector space). Now let $(0,y) \in \operatorname{Ker}(T)$. Then $(0,y) = y(0,1)$, where $y \in \mathbb{R}$. This shows that $(0,y)$ is scalar multiple of $(0,1)$. Hence $(0,y) \in L(\{(0,1)\})$, whence $\operatorname{Ker}(T) \subset L(\{(0,1)\})$. Therefore $\operatorname{Ker}(T) = L(\{(0,1)\})$.

Since the set $\{(0,1)\}$ is linearly independent, we have $\dim(\operatorname{Ker}(T)) = 1$.

Similarly prove that $\dim(\operatorname{Im}(T)) = 1$. Hence we have verified the rank–nullity theorem.

Recall transpose of a matrix:

Let A be a $m \times n$ matrix. We define $A^t :=$ the transpose of the matrix A as $A^t = [b_{ij}]_{n\times m}$ where $b_{ij} = a_{ji}$, $1 \le i \le n$ and $1 \le j \le m$.

Verify the following properties: $(A+B)^t = A^t + B^t$ and $(AB)^t = B^t A^t$.

"dot" product in the vector space $\mathbb{C}^n$:

Let $u = \begin{bmatrix} z_1 \\ z_2 \\ \cdot \\ \cdot \\ \cdot \\ z_n \end{bmatrix} \in \mathbb{C}^n$ and $v = \begin{bmatrix} z'_1 \\ z'_2 \\ \cdot \\ \cdot \\ \cdot \\ z'_n \end{bmatrix} \in \mathbb{C}^n$. We define

$$u \bullet v = \overline{u}^t v = \begin{bmatrix} \overline{z}_1 & \overline{z}_2 & . & . & . & \overline{z}_n \end{bmatrix} \begin{bmatrix} z'_1 \\ z'_2 \\ \cdot \\ \cdot \\ \cdot \\ z'_n \end{bmatrix}$$

$$= \overline{z}_1 z'_1 + \overline{z}_2 z'_2 + \cdots + \overline{z}_n z'_n$$

Same definition works for the real vector space $\mathbb{R}^n$. If $u, v \in \mathbb{R}^n$, then in this case we know that $\overline{u} = u$. So $u \bullet v = z_1 z'_1 + z_2 z'_2 + \cdots + z_n z'_n$.

Definition 2.8. The length of a vector $u \in \mathbb{C}^n$ is a real number which is denoted by $||u||$ and is defined as $\sqrt{u \bullet u}$. So if $u = \begin{bmatrix} z_1 \\ z_2 \\ \cdot \\ \cdot \\ \cdot \\ z_n \end{bmatrix} \in \mathbb{C}^n$, then

$$||u|| = \sqrt{|z_1|^2 + |z_2|^2 + \ldots + |z_n|^2}.$$

Remark 2.9. Note that $||u|| = 0$ if and only if $u = 0$ i.e. $z_1 = 0, \ldots, z_n = 0$.

Example 2.10. (1) Let $u = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $v = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. Then $u \bullet v = \overline{u}^t v = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 3 + 8 = 11$.

$||u|| = \sqrt{1^2 + 2^2} = \sqrt{5}$ and $||v|| = \sqrt{3^2 + 4^2} = \sqrt{9+16} = 5$.

(2) Let $u = \begin{bmatrix} i \\ 2 \end{bmatrix}$ and $v = \begin{bmatrix} 1 \\ 4i \end{bmatrix}$. Then $u \bullet v = \overline{u}^t v = \begin{bmatrix} -i & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 4i \end{bmatrix} = -i + 8i = 7i$. In this case, $||u|| = \sqrt{|i|^2 + 2^2} = \sqrt{5}$ and $||v|| = \sqrt{1^2 + |4i|^2} = \sqrt{1+16} = \sqrt{17}$.

(3) Let $u = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $v = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Then $u \bullet v = \overline{u}^t v = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1 - 1 = 0$.

Definition 2.11. Two vectors u and v are said to be orthogonal to each other if $u \bullet v = 0$. A set of vectors $\{u_1, \ldots, u_m\}$ is said to be orthogonal if they are pairwise orthogonal, i.e. $u_i \bullet u_j = 0$ for all i, j, $i \neq j$. Geometrically, orthogonal vectors are perpendicular to each other.

Proposition 2.12. *Let $S = \{u_1, \ldots, u_n\}$ be a set of non-zero orthogonal vectors. Then S is a linearly independent set.*

Proof. Let

$$c_1 u_1 + \cdots + c_n u_n = 0. \tag{2.1}$$

Now we have to show that $c_1 = 0, \ldots, c_n = 0$.

Taking dot product with u_i on the both sides of the equation 2.1, we get $(c_1 u_1 + \cdots + c_n u_n) \bullet u_i = 0 \bullet u_i$. Since $u_j \bullet u_i = 0$ for $j \neq i$, we get $c_i u_i \bullet u_i = 0$ for all $i = 1, \ldots, n$. Since $u_i \neq 0$, $||u_i|| \neq 0$. Hence $||u_i||^2 \neq 0$ i.e. $u_i \bullet u_i \neq 0$. Therefore $c_i = 0$ for $i = 1, \ldots, n$. This shows that S is a linearly independent set. ☐

Recall how to calculate determinant:

Given a square matrix A, $\det(A)$ is a scalar. (Practice some examples.)

Properties of determinant: (i) $\det(AB) = \det(A)\det(B)$. and (ii) $\det(cA) = c^n \det(A)$, where c is a scalar and n is the size of the matrix.

3. Lecture–8

Let A be an $n \times n$ matrix. We have learnt that if A is a real matrix, then A gives rise to a linear map $A : \mathbb{R}^n \to \mathbb{R}^n$.

Let $A =$

$$\begin{bmatrix} 6 & 3 \\ 4 & 7 \end{bmatrix}.$$

Note that: $A \begin{bmatrix} 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 33 \\ 27 \end{bmatrix}$ and $A \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 30 \\ 40 \end{bmatrix} = 10 \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

We observe that A takes some vector to a scalar multiple of itself. Goemetrically, after applying A some vectors remain in the same line passing through the origin.

For a matrix A, we are interested in those vectors and scalars. Formally, we are interested in the following vector equation:

$$Ax = \lambda x \tag{3.1}$$

The problem of finding non-zero vector x's and scalars λ's that satisfy the equation (3.1) is called an eigenvalue problem.

Remark 3.1. We want non-zero x because if $x = 0$, L.H.S.=0=R.H.S. for any λ. This case is not interesting.

Definition 3.2. A value of λ for which 3.1 has a solution $x \neq 0$, is called an eigen value of the matrix A. The correspoding solution $x (\neq 0)$ of 3.1 are called eigenvectors of A corresponding to the eigenvalue λ.

How to find eigen values and eigen vectors:

Observation: $Ax = \lambda x \Rightarrow Ax = \lambda x = 0 \Rightarrow (A - \lambda I)x = 0$. Check that to have non-zero solution x, we have $det(A - \lambda I) = 0$.

Note that $det(A - \lambda I)$ is a polynomial $p(\lambda)$ of degree n (recall that n is the size of the matrix.) This polynomial is called the characteristic polynomial of the matrix A.

To get eigen values: Solve the polynomial equation $p(\lambda) = 0$.

To get eigen vectors: Let the eigen values be $\lambda_1, \lambda_2, \dots, \lambda_n$. Then put the value of λ_i in 3.1. Then solve for non-zero x.

Example 3.3. Let

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 1 \end{bmatrix}, \text{ then } A - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 5 & 1-\lambda \end{bmatrix}.$$

Therefore $det(A - \lambda I) = \lambda^2 - 2\lambda - 9$. The solutions are $1 \pm \sqrt{10}$. Hence eigen values are $1 \pm \sqrt{10}$.

Now check that $\begin{bmatrix} \sqrt{10}/5 \\ 1 \end{bmatrix}$ is an eigen vector corresponding to the eigen value $1 + \sqrt{10}$. Calculate the other one!

Remark 3.4. Corresponding to the real matrices, eigen values need not be real. Hence eigen vector may not exist. For example,

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \text{ then } A - \lambda I = \begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix}.$$

Therefore $det(A - \lambda I) = \lambda^2 + 1$. The solutions are $\pm i$, where $i = \sqrt{-1}$. So the eigen values are $\pm i$.

For eigen vectors, we have $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = i \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. This implies $ix_1 - x_2 = 0$ and $-x_1 - ix_2 = 0$. Note that 2nd equation $= i \times$ 1st equation. So solution set $= \{c(1, i) : c \in \mathbb{C}\}$. Hence $\begin{bmatrix} 1 \\ i \end{bmatrix}$ is an eigen vector corresponding to the eigen value i.

Remark 3.5. Every real or complex matrix have complex eigen values and complex eigen vectors.

4. Lecture - 9, 10, 11

Recall. Let A be an $n \times n$ matrix. Then the degree n polynomial $p(\lambda) = \det(A - \lambda I)$ is called the characteristic polynomial. The equation $p(\lambda) = 0$ is called the characteristic equation.

Proposition 4.1. *If u, v are eigen vectors corresponding to the eigen value λ. Then $u + v (u \neq -v)$ and $cv (c \neq 0)$ are also eigen vectors corresponding to the eigen value λ.*

Proof. We have $Au = \lambda u$ and $Av = \lambda v$. Now $A(u + v) = Au + Av = \lambda u + \lambda v = \lambda(u + v)$. This shows that $u + v$ is an eigen vector corresponding to the eigen value λ. Note that $u \neq v$ is required to say that $u + v \neq 0$. Similarly, $A(cv) = cAv = c\lambda v = \lambda cv$. This shows that cv is an eigen vector corresponding to the eigen value λ. $\square$

Remark 4.2. The above proposition proves that the set of all eigen vectors corresponding to a eigen value λ forms a vector subspace along with the zero vector. This is called the eigen space corresponding to the eigen value λ. This is $\ker(A - \lambda I)$.

Algebraic Multiplicity and Geometric Multiplicity of an eigen value λ:

The algebraic multiplicity of an eigenvalue λ is the order of the multiplicity of the eigen value λ in the characteristic equation.

The geometric multiplicity of an eigenvalue λ is the dimension of the eigen space, i.e. $\dim(\ker(A - \lambda I))$.

Some examples:

Example 4.3. (1) Let $A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}$. The characteristic polynomial $\det(A-\lambda I) = -(\lambda-1)(\lambda-2)^2$. So eigenvalues are $1, 2, 2$. The algebraic multiplicity of the eigen value $\lambda = 1$ is one. The algebraic multiplicity of the eigen value $\lambda = 2$ is two.

First, let us find the eigen vectors corresponding to the eigen value $\lambda = 1$. We want to find a non-zero vector v such that $Av = \lambda v = v$, i.e. $(A - I)v = 0$ Now $A - I = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix}$. Let $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$. Then $(A - I)v = 0$ yields $2v_1 + v_2 - v_3 = 0$ and $2v_1 + 2v_2 - v_3 = 0$. So we want to find the following solution space $\{(v_1, v_2, v_3) : 2v_1 + v_2 - v_3 = 0 \text{ and } 2v_1 + 2v_2 - v_3 = 0\}$. This gives $v_2 = 0$ and $v_3 = 2v_1$. So we have the following solution space $\{(v_1, v_2, v_3); v_2 = 0; v_3 = 2v_1\} = \{(v_1, 0, 2v_1); v_1 \in \mathbb{R}\} = \{v_1(1, 0, 2) : v_1 \in \mathbb{R}\}$. So the eigen vector is $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$ corresponding to the eigen value 1. Also, notice $\dim(\ker(A - I)) = 1$ which is the geometric multiplicity of the eigen value 1.

For the eigen vectors corresponding to the eigen value 2. So we want to look for non-zero solution for the homogenous equation $(A-2I)v = 0$. Now $A-2I = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{bmatrix}$. So we want to find the following solution space $\{(v_1, v_2, v_3) : v_1 + v_2 - v_3 = 0 \text{ and } 2v_1 - v_3 = 0\} = \{(v_1, v_2, v_3) : v_2 = v_!; v_3 = 2v_1\} = \{(v_1, v_1, 2v_1) : v_1 \in \mathbb{R}\} = \{v_1(1, 1, 2) : v_1 \in \mathbb{R}\}$.

So the eigen vector is $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ corresponding to the eigen value 2. Also, notice $\dim(\ker(A - 2I)) = 2$ which is the geometric multiplicity of the eigen value 1.

Definition 4.4. (1) A $n \times n$ complex matrix A is called Hermitian if $\overline{A}^t = A$ and A is called skew-Hermitian if $\overline{A}^t = -A$. It is called unitary if $\overline{A}^t = A^{-1}$.

(2) If A is a real $n \times n$ matrix, then $A = \overline{A}$. In this case, Hermitian matrices are called symmetric, Skew-Hermitian matrices are called skew-symmetric and unitary matrices are called orthogonal.

Some observation: For Hermitian matrix $A = [a_{jk}]$, we have $a_{jj} = \overline{a_{jj}}$. This implies that a_{jj} are real. Hence the diagonal entries of a Hermitian matrix are real. Similarly, one checks that the diagonal entries of a skew-Hermitian matrix are purely imaginary or zero.

Theorem 4.5. (1) *The eigen values of a Hermitian matrix (and thus of a symmetric matrix) are real.*

(2) *The eigenvalues of a skew-Hermitian matrix (and thus of a skew-symmetric matrix) are pure imaginary or zero.*

(3) *The eigen values of a unitary matrix (and thus of an orthogonal matrix) have absolute value* 1.

Proof. Let λ be an eigen value of a matrix A and x an eigen vector correspoding to the eigen value λ. So we have $Ax = \lambda x$. Multiply this equation $\overline{x}^t$ from the left. Then we have

$\overline{x}^t Ax = \overline{x}^t \lambda x = \lambda \overline{x}^t x$. So $\lambda = \overline{x}^t Ax / \overline{x}^t x \cdots (i)$. Note that $\overline{x}^t x = |x_1|^2 + \cdots + |x_n|^2 > 0$ because $x \neq 0$.

Proof of (1). Since $\overline{x}^t x$ is real, it is enough to prove $\overline{x}^t Ax$ is real to prove λ is real. Since $\overline{x}^t Ax$ is scalar, we have $\overline{x}^t Ax = (\overline{x}^t Ax)^t = x^t A^t \overline{x} =^1 x^t \overline{A}\overline{x} = \overline{\overline{x}^t Ax}$. ($=^1$ holds because A is Hermitian, i.e. $A = \overline{A}^t$, i.e. $A^t = \overline{A}$.) Now we know that $z = \overline{z}$, then z is real. Hence $\overline{x}^t Ax$ is real.

Proof of (2). Let A be a skew-Hermitian matrix, i.e. $A = -\overline{A}^t$, i.e. $A^t = -\overline{A}$. Similarly as in (1), $\overline{x}^t Ax = (\overline{x}^t Ax)^t = x^t A^t \overline{x} =^1 -x^t \overline{A}\overline{x} = -\overline{\overline{x}^t Ax}$. ($=^1$ holds because A is skew-Hermitian.) Now we know that $z = -\overline{z}$, then z is pure imaginary or zero. Hence $\overline{x}^t Ax$ is pure imaginary or zero. From (i), we get that λ is pure imaginary or zero.

Proof of (3): Let A be a unitary matrix, i.e. $A^{-1} = \overline{A}^t$. From $Ax = \lambda x \cdots (i)$, we have $\overline{Ax}^t = \overline{\lambda x}^t \implies \overline{x}^t \overline{A}^t = \overline{\lambda}\overline{x}^t \cdots (ii)$. Using (i) and (ii), we get $\overline{x}^t \overline{A}^t Ax = \overline{\lambda}\overline{x}^t \lambda x$. But $\overline{A}^t A = I$. Hence we have $\overline{x}^t x = \overline{\lambda}\lambda \overline{x}^t x$. This implies $|\lambda|^2 = 1$. Hence $|\lambda| = 1$. $\square$

Definition 4.6. (Similar Matrices) An $n \times n$ matrix B is called similar to an $n \times n$ matrix A if $B = P^{-1}AP$ for some $n \times n$ non-singular matrix P.

Theorem 4.7. *If B is similar to A, then B has same eigen values as A. Furthermore, if x is an eigen vector of A, then $y = P^{-1}x$ is an eigen vector of A.*

Proof. If we can prove that A and B have the same characteristic polynomial, then we will be done. The characterictic polynomial of B

$$
\begin{aligned}
&= det(B - \lambda I) = det(P^{-1}AP - \lambda I)\\
&= det(P^{-1}AP - \lambda P^{-1}P) = det(P^{-1}(A - \lambda I)P)\\
&= det(P^{-1})det(A - \lambda I)det(P) = det(A - \lambda I)det(P^{-1})det(P)\\
&= det(A - \lambda I)det(I) = det(A - \lambda I)\\
&= \text{the characteristic polynomial of } A.
\end{aligned}
$$

Let λ be an eigen value of a matrix A and x an eigen vector correspoding to the eigen value λ. So we have $Ax = \lambda x \implies P^{-1}Ax = P^{-1}\lambda x \implies P^{-1}APP^{-1}x = P^{-1}\lambda x$. Hence $B(P^{-1}x) = \lambda(P^{-1}x)$. Note that $P^{-1}x \neq 0$ since $x \neq 0$. Therefore $P^{-1}x$ is an eigen vector of B corresponding to the eigen value λ. $\square$

Definition 4.8. A square matrix is said to be diagonalizable if there exists a diagonal matrix D such that D is similar to A, i.e. $D = P^{-1}AP$ for some non-singular matrix P.

How to check a matrix A is diagonalizable and then how to find D and P:

Method.

Step 1. First find eigenvalues of the $n \times n$ matrix A. If it has n distinct eigenvalues, then A is diagonalizable. Let the eigenvalues be $\lambda_1, \ldots, \lambda_n$ and let $x_1, \ldots x_n$ be the eigenvectors corresponding to $\lambda_1, \ldots, \lambda_n$. Then $D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$ and $P = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$, where x_i's are column vector.

Step 2. If A does not have n distinct eigen values, then calculate algebraic multiplicity (AM) and geometric multiplicity (GM) for each eigenvalue. If there exists one eigenvalue λ such that $AM(\lambda) \neq GM(\lambda)$, then A is not diagonalizable.

Step 3. Otherwise, i.e. $AM(\lambda) = GM(\lambda)$ for all eigenvalue λ. Then A is diagonalizable. In this case $D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$ but some of λ_i's might be the same. And we will get n linearly independent eigenvectors $x_1, \ldots, x_n$ and $P = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$.

Remark 4.9. Hermitian, skew-Hermitian, Unitary matrices are diagonalizable.

Theorem 4.10. *(***Cayley–Hamilton Theorem***) Every square matrix satisfies its own characterictic equation.*

Example 4.11. (Tutorial–3; Problem–4) Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$. Prove that $A^n = A^{n-2} + A^2 - I$ for $n \geq 3$. Calculate A^{50}?

Solution: The characteristic polynomial $\det(A - \lambda I) = -(\lambda - 1)^2(\lambda + 1)$. The characterictic equation is $\lambda^3 - \lambda^2 - \lambda + 1 = 0$. Now Cayley–Hamilton Theorem says:

$$A^3 - A^2 - A + I = 0 \cdots (i).$$

We prove this by induction on n. For $n = 3$, this holds because of (i). Assume that this holds for $n = m$. Now we have to prove it for $n = m+1$. This holds for $n = m$ means $A^m = A^{m-2} + A^2 - I$. Multiply both sides by A. $A^{m+1} = A^{m-1} + A^3 - A \cdots (ii)$. But from (i), we have $A^3 - A = A^2 - I$. Putting back to (ii), we get $A^{m+1} = A^{m-1} + A^2 - I$. This is what we wanted to prove.

How to calculate A^{50}? Using equation (i), we get that $A^{50} = A^2 + 24(A^2 - I)$ (check!). Now $A^2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. and $A^2 - I = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$. Hence $A^{50} = \begin{bmatrix} 1 & 0 & 0 \\ 25 & 1 & 0 \\ 25 & 0 & 1 \end{bmatrix}$.

References

[1] E. Kreyszig, *Advanced engineering mathematics*, 10th edition. John Wiley & Sons, Inc., New York, 1999. xvi+1156+A97+I20 pp. ISBN: 0-471-15496-2

Office N354, Department of Mathematics, Indian Institute of Technology Kharagpur, Kharagpur 721302, West Bengal, India

E-mail address: `parvez@maths.iitkgp.ac.in`

E-mail address: `mathparvez@gmail.com`